

# **Regioselective hydration of terpenoids using cofactor-independent hydratases**

Regioselektive Hydratisierung von Terpenoiden mittels  
Cofaktor-unabhängigen Hydratase

Von der Fakultät 3 (Chemie) der Universität Stuttgart  
zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.)  
genehmigte Abhandlung

vorgelegt von

**Jens Schmid**

aus Stuttgart

Hauptberichter: Prof. Dr. Bernhard Hauer  
Mitberichter: Prof. Dr. Sabine Laschat  
Vorsitzender: Prof. Dr. Albert Jeltsch

Tag der mündlichen Prüfung: 02.08.2019

Institut für Biochemie und Technische Biochemie der Universität Stuttgart  
Abteilung Technische Biochemie

2019



PhD Thesis

**Regioselective hydration of  
terpenoids using cofactor-independent  
hydratases**

Jens Schmid

Institute of Biochemistry and Technical Biochemistry  
Department of Technical Biochemistry  
University of Stuttgart

2019

The presented work was developed at the suggestion and under supervision of Prof. Dr. Bernhard Hauer from April 2015 to June 2018 at the Institute of Biochemistry and Technical Biochemistry, Department of Technical Biochemistry at the University of Stuttgart and was funded by Givaudan SA.

Within the scope of this work the following publications have been released:

Demming, R.\*, Fischer, M.-P.\*, **Schmid, J.\*** & Hauer, B. (2018). (De) hydratases- recent developments and future perspectives. *Current opinion in chemical biology*, 43, 43-50

**Schmid, J.\***, Steiner, L.\*, Fademrecht, S., Pleiss, J., Otte, K., & Hauer, B. (2016). Biocatalytic study of novel oleate hydratases. *Journal of Molecular Catalysis B: Enzymatic*, 133, S243-S249.

\*authors contributed equally to the work

## **Erklärung über die Eigenständigkeit der Dissertation**

Ich versichere, dass ich die vorliegende Arbeit mit dem Titel

„Regioselektive Hydratisierung von Terpenoiden mittels  
Cofaktor-unabhängigen Hydratasen“

selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe; aus fremden Quellen entnommene Passagen und Gedanken sind als solche kenntlich gemacht.

## **Declaration of Authorship**

I hereby certify that the dissertation entitled

“Regioselective hydration of terpenoids using  
cofactor-independent hydratases”

is entirely my own work except where otherwise indicated. Passages and ideas from other sources have been clearly indicated.

Name/Name: Jens Schmid

Unterschrift/Signed: \_\_\_\_\_

Datum/Date: 08.03.2019

## Danksagungen

Zunächst möchte ich mich bei meinem Doktorvater Prof. Dr. Bernhard Hauer für die Möglichkeit bedanken, meine Dissertation am Institut für Technische Biochemie anzufertigen. Zu besonderem Dank verpflichtet bin ich ihm für sein Vertrauen und seine Unterstützung auch in schwierigen Phasen des Projekts sowie die Zuversicht, dass ich es zu einem erfolgreichen Abschluss bringe.

Frau Prof. Dr. Sabine Laschat möchte ich ganz herzlich für die Übernahme des Zweitgutachtens danken. Herrn Prof. Dr. Albert Jeltsch gilt mein Dank für die Übernahme des Prüfungsvorsitzes.

In diesem Zusammenhang danke ich meinem Industriepartner, der Givaudan SA, für die Finanzierung und allen am Projekt Beteiligten für das Interesse an meinem Thema und die stets freundliche, offene und hilfreiche Diskussion meiner Ergebnisse. Besonders danke ich den Projektverantwortlichen Dr. Felix Flachsmann und Dr. Eric Eichhorn, die das Projekt neben ihrer täglichen Arbeit begleitet haben. Felix und seinem Team möchte ich außerdem für die Synthese von Substraten und Produktstandards sowie für die Analyse von Proben danken.

Prof. Dr. Per-Olof Syrén und seinen Kollegen von der KTH Stockholm danke ich für die Sequenzierung der gDNA des *Geobacillus*. Dr. Sandra Facey gilt mein Dank für die Hilfe bei der Isolierung der gDNA und bei Fragen zur Biologie. Dr. Silvia Fademrecht möchte ich für die Erstellung der HyED und die gelungene und angenehme Zusammenarbeit an unserem Paper danken. Bedanken möchte ich mich auch bei meinen Studenten Manuel Lauinger und Johanna Dollinger, die mich tatkräftig in der frühen Phase meiner Doktorarbeit unterstützt haben.

Dr. Bettina Nestl und Dr. Bernd Nebel gilt der Dank für das Zusammenhalten des Labors in allen Situationen, Bettina darüber hinaus für das Diskutieren aussichtloser Ideen und Korrekturlesen etlicher Manuskripte, Bernd für die Unterstützung bei der Etablierung der LC-MS Analytik. Lisa Kontny, Melanie Allgaier, Sven Richter und Christine Klug danke ich für ihre tolle Arbeit, die häufig im Hintergrund stattfindet. Besonders erwähnen möchte ich noch weitere Kollegen am ITB.

Dem Team (De)hydratase bestehend aus Rebecca Demming und Max-Philipp Fischer, zu Beginn auch Dr. Konrad Otte, danke ich für Material- und Wissensaustausch sowie die Diskussion während und außerhalb unserer „regelmäßiger“ Meetings.

Dr. Stephan Hammer, meinem Betreuer aus der Bachelorarbeit, danke ich dafür, dass er das lodernde Feuer der Biokatalyse in mir als Chemiker entfacht hat. Zugleich ist er für mich so etwas wie ein Mentor geworden, weil er sich immer Zeit für einen nimmt und Interesse am Thema und an den Problemen anderer findet. Meinen Kollegen, egal ob unsere gemeinsame Zeit am ITB nun eher lang oder kurz war, danke ich für die tollen Erfahrungen, lustigen Momente, fachliche Diskussionen, aber auch ernste Gespräche und das gute Zureden in schwierigen Zeiten. Besonders verbunden bleibe ich für immer dem Biokatalyse 1-Labor, das heute auf den Namen Rosalind Franklin hört und in dem mit der Bachelorarbeit meine wissenschaftliche Karriere begann. Danke also den „alten Hasen“ Sara, Maike und Julia aber auch den „jungen Wilden“ Niels, Benjamin, Sebastian, Ludwig und Pascal. Meinem langjährigen Nebensitzer Lars danke ich für all die tollen, lustigen und interessanten Gespräche, die Fachsimpeleien aus Biologie und Chemie und die vielen erteilten Ratschläge.

Sabrina und Svenja und vor allem Nico, meinem treuen Mitstreiter in Sachen „Biologie“ während meines gesamten Chemiestudiums, bin ich sehr dankbar für die Freundschaft und die gemeinsame Zeit mit Ihnen, sowohl im als auch außerhalb des Labors. Meinen Kommilitonen des AK Mensa, der auch die Zeit der Promotion überdauert hat, danke ich für all die Momente, die das Studium zu einer schönen Zeit und einem Lebensabschnitt gemacht haben, den ich nicht missen möchte. Johannes, meinem Laborpartner fast aller Praktika, möchte ich besonders danken. Allen, die ich während meiner Zeit beim JCF Stuttgart kennenlernen durfte, danke ich für Ihre Unterstützung bei unseren Kindergartenexperimenten und für die Zeit bei weiteren gemeinsamen Aktionen wie dem Eismachen, unseren Vorträgen und das gesellige Beisammensein bei unserem Stammtisch.

Lisa, Sabrina, Nico und insbesondere Stephan, möchte ich auch noch für die sehr gewissenhafte Korrektur meiner Doktorarbeit danken.

Zuletzt möchte ich meiner Familie und meiner Freundin für ihre Zuneigung, bedingungslose Liebe und die unendliche Unterstützung in allen Phasen meines Lebens danken.





*Für Lisa Steiner*

*Life is not the amount of breaths you take, it's the moments that take your breath away.*

aus Hitch, Columbia Pictures, 2005

## List of abbreviations

x g	gravitational acceleration
°C	degree Celsius (unit of temperature $T$ )
$\delta$	chemical shift (physical quantity [ppm])
$\tilde{\nu}$	wavenumber (physical quantity [ $\text{cm}^{-1}$ ])
$^1\text{H-NMR}$	proton-NMR
$^{13}\text{C-NMR}$	carbon-13-NMR
Å	angstrom (unit of length $l$ )
A	ampere (unit of electric current $I$ )
ACN	acetonitrile
ADH	alcohol dehydrogenase
ATCC	American Type Culture Collection
BCA	bicinchoninic acid
BLAST	Basic Local Alignment Search Tool
bp	base pair(s)
BSA	bovine serum albumin
$c$	molar concentration (physical quantity [ $\text{M}$ or $\text{mol L}^{-1}$ ])
CAST	Combinatorial Active-site Saturation Test
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
conc.	concentrated
COSY	Correlation Spectroscopy
CrtC	carotenoid hydratase
cww	cell wet weight
d	day(s)
Da	dalton (unit of mass $m$ )
DAD	diode array detector
ddH <sub>2</sub> O	double-demineralized water (ultrapure water, purified with EMD Millipore Milli-Q)
DMDDAC	5,9-dimethyldeca-4,8-dienoic acid
DMDDAL	5,9-dimethyldeca-4,8-dienal
DMDDIOX	2-(4,8-dimethylnona-3,7-dien-1-yl)-5,5-dimethyl-1,3-dioxane
DMDDOL	5,9-dimethyldeca-4,8-dienol
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DSM	catalogue number at DSMZ
DSMZ	<i>Deutsche Sammlung von Mikroorganismen und Zellkulturen</i>
DTT	dithiothreitol
DWP	deep well plate
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
EI	electron ionization
eq	equivalent(s)
ESI	electrospray ionization

Et <sub>2</sub> O	diethyl ether
EtOH	ethanol
eV	electronvolt (unit of energy <i>E</i> )
FAD	flavin adenine dinucleotide
FAH	fatty acid hydratase
FID	flame ionization detector
FT	Fourier-transform
fw	forward
g	gram (unit of mass <i>m</i> )
GC	gas chromatography
GC-FID	gas chromatography coupled with flame ionization detector
GC-MS	gas chromatography coupled with mass spectrometry
gDNA	genomic DNA
h	hour(s) (unit of time <i>t</i> )
HDMDAC	9-hydroxy-5,9-dimethyldec-4-enoic acid
HDMDAL	9-hydroxy-5,9-dimethyldec-4-enal (Mahonial)
HDMDIOX	9-(5,5-dimethyl-1,3-dioxan-2-yl)-2,6-dimethylnon-6-en-2-ol
HDMDOL	5,9-dimethyldec-4-ene-1,9-diol
HFam	homologous family
HMBC	Heteronuclear Multiple-bond Correlation
HPLC	high performance liquid chromatography
HSQC	Heteronuclear Single Quantum Coherence
HT	high throughput (screening)
HyED	Hydratase Engineering Database
Hz	hertz (unit of frequency <i>f</i> )
IPTG	isopropyl β-D-1-thiogalactopyranoside
IR	infrared
ISM	iterative saturation mutagenesis
ITS	internal transcribed spacer
<i>J</i>	coupling constant (physical quantity [Hz])
K	Kelvin (unit of temperature <i>T</i> )
Kan	kanamycin
kb	kilo basepairs
KHS	kievitone hydratase
KP <sub>i</sub>	potassium phosphate
L	liter (unit of volume <i>V</i> )
LAH	linoleic acid hydratase
LB	lysogeny broth (medium)
LC	liquid chromatography
LC-MS	liquid chromatography coupled with mass spectrometry
LIH	limonene hydratase
<i>m</i>	mass (physical quantity [g])
m	meter (unit of length <i>l</i> )
M	molar (mol L <sup>-1</sup> ; unit of molar concentration <i>c</i> )
MCRA	myosin cross-reactive antigen

MCS	multiple cloning site, in this case name for the empty (no hydratase inserted) pDHE plasmid
MD	molecular dynamics
MeOH	methanol
min	minute(s) (unit of time $t$ )
MOPS	3-morpholinopropane-1-sulfonic acid
MS	mass spectrometry
MTBE	methyl <i>tert</i> -butyl ether
$m/z$	mass-to-charge ratio
NAD <sup>+</sup> /NADH	nicotinamide adenine dinucleotide
NHase	nitrile hydratase
NMR	nuclear magnetic resonance
NOESY	Nuclear Overhauser Enhancement Spectroscopy
NTP	nucleoside triphosphate
OAH	oleate hydratase
OD <sub>600</sub>	optical density measured at 600 nm
ORF	open reading frame
ORI	origin of replication
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
pH	measure of acidity
ppm	parts per million (unit for chemical shift $\delta$ )
<i>i</i> PrOH	<i>iso</i> -propanol
QM/MM	quantum mechanics / molecular mechanics
RNA	ribonucleic acid
rpm	revolutions per minute
rv	reverse
s	seconds (unit of time $t$ )
SDP	specificity determining position
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SHC	squalene-hopene cyclase
TAE	TRIS-acetate-EDTA buffer
TB	terrific broth (medium)
Tfb	transformation buffer
THF	tetrahydrofuran
TLC	thin layer chromatography
TRIS	tris(hydroxymethyl)aminomethane
TSB	tryptone soy broth
U	(enzyme) unit (unit for catalytic activity)
V	volt (unit for voltage $U$ )
V	volume (physical quantity [L])

## **Bacteria strains as origin of hydratases**

### Oleate hydratase

<i>Cg</i>	<i>Chryseobacterium gleum</i>
<i>Db</i>	<i>Desulfomicrobium baculatum</i>
<i>Em</i>	<i>Elizabethkingia meningoseptica</i>
<i>Gm</i>	<i>Gemella morbillorum</i>
<i>Hf</i>	<i>Holdemania filiformis</i>
<i>La</i>	<i>Lactobacillus acidophilus</i>
<i>Lj</i>	<i>Lactobacillus johnsonii</i>
<i>Mp</i>	<i>Mucilaginibacter paludis</i>

### Carotenoid hydratase

<i>Rg</i>	<i>Rubrivivax gelatinosus</i>
<i>Tr</i>	<i>Thiocapsa roseopersicina</i>

### Limonene hydratase

<i>Gs</i>	<i>Geobacillus stearothermophilus</i>
-----------	---------------------------------------

Myraldene, Lyral and Mahonial are registered trademarks of the respective manufacturers.



# Table of content

<b>Zusammenfassung</b> .....	<b>1</b>
<b>Abstract</b> .....	<b>4</b>
<b>1 Introduction</b> .....	<b>6</b>
1.1 LILY-OF-THE-VALLEY IN FRAGRANCE INDUSTRY .....	6
1.1.1 <i>Lily-of-the-valley</i> .....	7
1.1.2 <i>Synthesis of lily-of-the-valley hydroxy aldehydes</i> .....	8
1.2 METHODS FOR SYNTHESIS OF TERTIARY ALCOHOLS .....	12
1.2.1 <i>Chemical methods</i> .....	12
1.2.2 <i>Examples from biology for the synthesis of tertiary terpenoid alcohols by hydration</i> .....	15
1.3 HYDRATASES .....	17
1.3.1 <i>Linalool dehydratase-isomerase</i> .....	18
1.3.2 <i>Kievietone hydratase</i> .....	19
1.3.3 <i>Carotenoid hydratases</i> .....	20
1.3.4 <i>Limonene hydratase</i> .....	23
1.3.5 <i>Oleate hydratases</i> .....	25
1.4 DEVELOPMENT OF NEW BIOCATALYSTS BY ENZYME ENGINEERING .....	32
<b>2 Motivation and aim of the project</b> .....	<b>34</b>
<b>3 Experimental</b> .....	<b>35</b>
3.1 MATERIALS .....	35
3.1.1 <i>Consumables</i> .....	35
3.1.2 <i>Chemicals</i> .....	36
3.1.3 <i>Marker &amp; dyes</i> .....	37
3.1.4 <i>Enzymes &amp; kits</i> .....	37
3.1.5 <i>Devices</i> .....	38
3.1.6 <i>Bacteria strains</i> .....	39
3.1.7 <i>Primers</i> .....	39
3.1.8 <i>Plasmids</i> .....	46
3.1.9 <i>Buffers, media &amp; antibiotics</i> .....	49
3.2 MOLECULAR BIOLOGICAL METHODS .....	52
3.2.1 <i>Isolation of genomic DNA</i> .....	52
3.2.2 <i>Polymerase chain reaction</i> .....	53
3.2.3 <i>Gibson Assembly</i> .....	54
3.2.4 <i>Site-directed and site-saturation mutagenesis</i> .....	59
3.2.5 <i>Agarose gel electrophoresis &amp; gel extraction</i> .....	61
3.2.6 <i>Plasmid preparation</i> .....	61
3.2.7 <i>Determination of DNA concentration (NanoDrop)</i> .....	61

3.2.8	<i>Sequencing</i> .....	62
3.2.9	<i>Preparation of chemo-competent cells</i> .....	62
3.2.10	<i>Transformation</i> .....	62
3.3	MICROBIOLOGICAL METHODS .....	63
3.3.1	<i>Cultivation of Geobacillus stearothermophilus</i> .....	63
3.3.2	<i>Cultivation of E. coli</i> .....	64
3.3.3	<i>Cultivation of E. coli in 96-DWP</i> .....	64
3.4	PROTEINBIOCHEMICAL METHODS .....	65
3.4.1	<i>Protein production</i> .....	65
3.4.2	<i>Cell disruption</i> .....	66
3.4.3	<i>SDS-PAGE</i> .....	66
3.4.4	<i>BCA assay</i> .....	67
3.5	BIOTRANSFORMATIONS .....	68
3.5.1	<i>Analytical biotransformations</i> .....	68
3.5.2	<i>Screening of EmOAH variants in 96-DWP</i> .....	69
3.5.3	<i>Preparative whole cell biotransformation of DMDDAC</i> .....	70
3.6	ANALYTICAL METHODS .....	72
3.6.1	<i>Gas chromatography (GC)</i> .....	72
3.6.2	<i>Liquid chromatography (LC)</i> .....	73
3.6.3	<i>Nuclear magnetic resonance spectroscopy (NMR)</i> .....	75
3.6.4	<i>IR spectroscopy</i> .....	75
3.7	COMPUTATIONAL METHODS .....	75
3.7.1	<i>Docking studies</i> .....	75
3.7.2	<i>Molecular dynamics (MD Refinement)</i> .....	76
3.8	CHEMICAL SYNTHESIS .....	77
3.8.1	<i>Synthesis of 10-hydroxy-10-methylundecanoic acid</i> .....	77
<b>4</b>	<b>Results</b> .....	<b>78</b>
4.1	GEOBACILLUS STEAROTHERMOPHILUS & LIMONENE HYDRATASE.....	78
4.1.1	<i>Recovery and identification</i> .....	78
4.1.2	<i>Cultivation</i> .....	80
4.1.3	<i>Isolation of gDNA and sequencing</i> .....	83
4.1.4	<i>Proof of limonene pathway and search for limonene hydratase</i> .....	84
4.1.5	<i>Biotransformations</i> .....	85
4.2	CAROTENOID HYDRATASE .....	87
4.2.1	<i>Cloning and expression of TrCrtC and RgCrtC in E. coli ITB 94</i> .....	88
4.2.2	<i>Biotransformations</i> .....	89
4.3	CONSTRUCTION OF EMPTY VECTOR PDHE-MCS .....	91
4.4	OLEATE HYDRATASE.....	92
4.4.1	<i>Preliminary experiments with EmOAH</i> .....	92
4.4.2	<i>Screening of oleate hydratase homologs</i> .....	93
4.4.3	<i>Cloning and expression of EmOAH in E. coli ITB 94</i> .....	95



4.4.4	<i>In silico docking studies</i> .....	96
4.4.5	<i>Identification of amino acid residues for saturation mutagenesis by comparison of docking results</i> .....	98
4.4.6	<i>Initial mutant library</i> .....	103
4.4.7	<i>Semi-rational site-directed mutagenesis (DWP-assay)</i> .....	106
4.4.8	<i>Screening of EmOAH saturation mutagenesis libraries</i> .....	106
4.4.9	<i>Biological replicates of selected EmOAH variants</i> .....	113
4.4.10	<i>Screening of reaction conditions</i> .....	114
4.4.11	<i>Preparative biotransformations</i> .....	114
4.4.12	<i>Substrate scope</i> .....	116
<b>5</b>	<b>Discussion</b> .....	<b>119</b>
5.1	SELECTION OF HYDRATASES .....	119
5.2	<i>GEOBACILLUS STEAROTHERMOPHILUS</i> .....	122
5.2.1	<i>Conditions for cultivation</i> .....	122
5.2.2	<i>Sequencing, identification and taxonomic classification</i> .....	123
5.2.3	<i>Putative limonene hydratase</i> .....	124
5.3	CAROTENOID HYDRATASES.....	127
5.4	OLEATE HYDRATASES .....	129
5.4.1	<i>Wild type OAH screening</i> .....	129
5.4.2	<i>Selection of hot spot amino acid residues</i> .....	131
5.4.3	<i>Iterative saturation mutagenesis</i> .....	134
5.4.4	<i>Validation of selected variants &amp; upscaling</i> .....	138
5.4.5	<i>Substrate scope</i> .....	139
5.5	BIOCATALYTIC HYDRATION IN THE CONTEXT OF ORGANIC SYNTHESIS AND THE SUPPLY CHAIN IN FRAGRANCE CHEMISTRY .....	140
<b>6</b>	<b>Conclusion &amp; outlook</b> .....	<b>142</b>
<b>7</b>	<b>Appendix</b> .....	<b>145</b>
7.1	DNA SEQUENCES OF WILD TYPE ENZYMES .....	145
7.2	VECTOR MAPS.....	154
7.3	NMR/GC/HPLC/MS SPECTRA .....	161
7.4	SUBSTRATE SCOPE STUDIES: CHEMICAL SYNTHESIS OF 10-HYDROXY-10-METHYLUNDECANOIC ACID .....	170
7.5	METHODS FOR THE SCREENING OF OAH MUTANT LIBRARIES .....	172
7.5.1	<i>Assessment of a photometric HT-assay</i> .....	172
7.5.2	<i>Chromatographic methods</i> .....	173
<b>8</b>	<b>Literature</b> .....	<b>175</b>



## Zusammenfassung

Biokatalytische Prozesse bieten eine Alternative zu chemischen Synthesen, vor allem wenn klassische Methoden nicht die gewünschte Selektivität für eine Reaktion bereitstellen können. Eine schwierige und komplexe Aufgabe stellt dabei das Finden eines geeigneten Enzyms für die gewünschte Reaktion dar.

Im Rahmen dieser Arbeit sollte ein biokatalytischer Weg gefunden werden, um Mahonial (HDMDAL), ein Molekül mit an Maiglöckchen erinnernden Geruch, aus 5,9-Dimethyldeca-4,8-dienal (DMDDAL) darzustellen.

Die Klasse der Maiglöckchendüfte stellt eine der bedeutendsten in der Parfümindustrie dar. Jedoch konnte bisher kein Öl aus Maiglöckchen gewonnen werden. Synthetische Moleküle, die den charakteristischen Duft am besten nachbilden, sind häufig Hydroxyaldehyde mit terpenoide Struktur. Die selektive Synthese solcher Hydroxyaldehyde wie das Mahonial ist jedoch eine Herausforderung. Besonders schwierig ist die Multifunktionalität des Moleküls, da sowohl die Aldehydgruppe als auch der tertiäre Alkohol empfindlich für harsche Reaktionsbedingungen sind. Eine späte Einführung der Hydroxyfunktion unter milden und regioselektiven Bedingungen durch Hydratisierung ist daher wünschenswert. Chemische Methoden, wie die Oxymercurierung, bieten sich für die Laborsynthese an, sind jedoch auf Grund der stöchiometrischen Mengen an giftigen Reagenzien für eine industrielle Anwendung ungeeignet.

Eine vielversprechende Alternative ist die enzymatische Hydratisierung der Prenyl-Doppelbindung mittels einer regioselektiven Hydratase. Hydratasen katalysieren die redox-neutrale Anlagerung von Wasser an eine Doppelbindung unter Bildung eines Alkohols. Hydratasen, die an isolierten Doppelbindungen wirken, sind von besonderem Interesse für die industrielle Anwendung.

Für die gewünschte Reaktion wurden als mögliche Enzyme die Limonenhydratase (LIH) aus *Geobacillus stearothermophilus*, die Carotinoidhydratasen (CrtCs) aus *Rubrivivax gelatinosus* und *Thiocapsa roseopersicina* sowie Oleathydratasen (OAHs) aus acht verschiedenen Organismen identifiziert. Im ersten Teil der Arbeit wurden die verschiedenen Hydratasen bereitgestellt, charakterisiert und auf Eignung für die gesuchte Reaktion getestet.

Entgegen der Angabe einer Literaturquelle konnte die Existenz einer LIH in *Geobacillus stearothermophilus* nicht nachgewiesen werden und wurde folglich nicht

weiter untersucht. Hingegen gelang es die CrtC in *E. coli* zu klonieren und heterolog zu exprimieren. Eine Untersuchung des Substratspektrums ergab, dass neben dem natürlichen Substrat Lycopin, ein C40 Terpen, unter den getesteten Bedingungen keine Hydratisierung kürzerer Substrate erfolgte. Im Gegensatz dazu konnte das breite Substratspektrum von OAHs bezüglich Fettsäuren verschiedener Kettenlängen (C11-C22) erfolgreich genutzt werden. Jedoch erwies sich die essentielle Säuregruppe im Substrat als limitierender Faktor. Aus diesem Grund wurde für weitere Experimente das Surrogat-Substrat 5,9-Dimethyldeca-4,8-diensäure (DMDDAC) verwendet. Mit diesem Substrat wurde eine Auswahl von acht homologen OAHs getestet. Durch eine umfangreiche Optimierung der Reaktionsbedingungen und die Etablierung einer hochsensitiven LC-MS Analytik konnte das Hydratisierungsprodukt von DMDDAC für fünf OAHs im Spurenbereich nachgewiesen werden.

Im zweiten Teil dieser Arbeit wurde die OAH aus *Elizabethkingia meningoseptica* (EmOAH) für weitere Mutagenesestudien ausgewählt, da zu diesem Enzym auch eine Kristallstruktur literaturbekannt ist. Für den gewählten Mutageneseansatz wurden zunächst *in silico* Bindungsstudien im Enzym durchgeführt. Dabei wird das Substrat flexibel in das aktive Zentrum des Enzyms modelliert. Anhand dieser Dockingstudien wurden elf Aminosäurereste als *Target* identifiziert und durch ortsgerichtete Mutagenese verändert. So wurde eine Mutantenbibliothek von 21 Varianten generiert und getestet. Das Screening der Bibliothek brachte die Varianten A248L, T549A und W389F hervor, welche eine ähnliche oder bessere Produktbildung als der Wildtyp zeigten. In weiteren Experimenten wurde durch iterative Sättigungsmutagenese an diesen Positionen die Variante A248I/T549S als die aktivste identifiziert. Die Verwendung dieses optimierten Biokatalysators ermöglichte erstmals die biokatalytische Darstellung des Säureanalogons von Mahonial, der 9-Hydroxy-5,9-dimethyldec-4-ensäure (HDMDAC). Dabei wurde mittels 2D-NMR Techniken die hohe Selektivität der Olehydratase nachgewiesen, da aus einem Isomerengemisch von *E/Z*-DMDDAC reine *E*-HDMDAC hergestellt werden konnte.

Zusammengefasst konnte am Beispiel der Synthese von HDMDAC demonstriert werden, dass die noch wenig erforschte Enzymklasse der Hydratase eine vielversprechende Alternative zu chemischen Methoden bietet. Es wurde gezeigt, wie durch einen Vergleich verschiedener Hydratase erstmals ein geeignetes Enzym mit einer Startaktivität für die gesuchte Reaktion gefunden werden konnte. Anschließend konnte diese durch *enzyme engineering* verbessert werden, wodurch das gewünschte

Produkt selektiv dargestellt und isoliert werden konnte. Damit zeigt diese Arbeit wie ein begrenztes Substratspektrum durch das Erstellen maßgeschneiderter Enzymvarianten überwunden werden kann.

## Abstract

Biocatalytic processes offer an alternative to chemical syntheses, especially when classical methods cannot provide the desired selectivity for a reaction. However, one of the biggest challenges and a complex task is finding a suitable enzyme for the desired reaction. In this work I aimed to explore a biocatalytic approach to prepare Mahonial (HDMDAL), a compound with a pleasant scent of lily-of-the-valley, from 5,9-dimethyldeca-4,8-dienal (DMDDAL).

The class of lily-of-the-valley fragrances represents one of the most important in the perfume industry, even though no essential oil could be obtained so far from lily-of-the-valley. Synthetic molecules that best replicate the characteristic scent are often hydroxy aldehydes with a terpenoid structure. The selective synthesis of such hydroxy aldehydes like Mahonial, however, is a challenge. Particularly difficult is the multifunctionality of the molecule, since both the aldehyde group and the tertiary alcohol are sensitive to harsh reaction conditions. Therefore, the late introduction of the hydroxy function under mild and regioselective conditions by hydration is desirable. Chemical methods, such as oxymercuration, are appropriate for laboratory scale synthesis, but are unsuitable for industrial application because of the need for stoichiometric amounts of toxic reagents.

A promising alternative is the enzymatic hydration of the isoprenyl double bond by means of a regioselective hydratase. Hydratases catalyze the redox-neutral addition of water to a double bond forming an alcohol. Hydratases which act on isolated double bonds are of particular interest for industrial use.

The limonene hydratase (LIH) from *Geobacillus stearothermophilus*, the carotenoid hydratases (CrtC) from *Rubrivivax gelatinosus* and *Thiocapsa roseopersicina*, and oleate hydratases (OAH) from eight different organisms were identified as potentially suitable enzymes. In the first part of the work, the various hydratases were provided, characterized and tested for the desired reaction.

In contrast to the indication of a literature source, the existence of a LIH in *Geobacillus stearothermophilus* could not be verified and consequently was not further investigated. On the other hand, CrtC was cloned and heterologously expressed in *E. coli*. Investigation of the substrate spectrum revealed that in addition to the natural substrate lycopene, a C40 terpene, no hydration of shorter substrates occurred under the tested conditions. On the contrary, the broad substrate scope of OAHs with respect to fatty

acids of different chain lengths (C11-C22) has been successfully exploited. However, the essential acid moiety in the substrate was a limiting factor. For this reason, the surrogate substrate 5,9-dimethyldeca-4,8-dienoic acid (DMDDAC) was used for further experiments. A selection of eight homologous OAHs was tested with this substrate. By extensive optimization of the reaction conditions and the establishment of a highly sensitive LC-MS analysis method, the hydration product of DMDDAC could be detected in traces for five OAHs.

In the second part of this work, the OAH from *Elizabethkingia meningoseptica* (EmOAH) was chosen for further mutagenesis studies, since a crystal structure of this enzyme is known in literature. Initially, *in silico* binding studies were performed for the chosen mutagenesis approach. In doing so, the substrate is flexibly modeled into the active center of the enzyme. Based on these docking studies, eleven amino acid residues were identified as target and altered by site-directed mutagenesis. Thus, a mutant library of 21 variants was generated and tested. Screening of the library revealed variants A248L, T549A and W389F, which showed similar or better product formation than the wild-type. In further iterative saturation mutagenesis experiments, the variant A248I/T549S was identified as the most active. The use of this optimized biocatalyst enabled for the first time the biocatalytic preparation of the acid analogue of Mahonial, 9-hydroxy-5,9-dimethyldec-4-enoic acid (HDMDAC). Thereby, the high selectivity of oleate hydratase was demonstrated by means of 2D-NMR techniques, since pure *E*-HDMDAC could be prepared from a mixture of isomers of *E/Z*-DMDDAC.

In summary, it was demonstrated by the example of the synthesis of HDMDAC that the yet poorly understood enzyme class of hydratases offers a promising alternative to chemical methods. Initially, a suitable enzyme with a starting activity for the desired reaction was found by comparing different hydratases. Subsequently, this activity could be improved by enzyme engineering in a way that the desired product could be selectively synthesized on preparative scale. Thus, this work shows how a limited substrate spectrum can be overcome by creating customized enzyme variants.

# 1 Introduction

In recent years, biocatalysis has steadily gained in importance and the number of applications, including industrial examples for the production of fine chemicals and even of bulk chemicals, has increased rapidly.<sup>[1–6]</sup> More and more classic chemical catalysts have been replaced by enzymes because they combine environmental benefits with cost-effectiveness and high selectivities.<sup>[2,7–9]</sup> In addition, the use of biocatalysts offers interesting opportunities for organic synthesis, as even abiological reactions can be performed.<sup>[9–12]</sup> Ecological benefits include, for example, biocompatibility, because enzymes are in general free of heavy metals and biodegradable.<sup>[2,9]</sup> Further, enzymes work under mild conditions (pH, temperature, pressure) and show high chemo-, regio-, and stereoselectivity.<sup>[3,4,7,13]</sup> This, together with a high activity (enzymes accelerate reactions typically by a factor of  $10^8$ - $10^9$ ), also allows access to products which cannot be synthesized, or only with difficulty, by classical organic methods.<sup>[1,13,14]</sup>

Despite advances in recent years, enzymes for many chemical reactions and processes are missing.<sup>[15–19]</sup> Therefore, an important task of biocatalysis is to identify or engineer these key enzymes.<sup>[4,16]</sup> A particular interesting example is the regioselective hydration of non-activated double bonds in terpenoids. The redox-neutral addition of water to alkenes by cofactor-independent hydratases is a remarkable example of a “dream reaction” with 100 % atom economy.<sup>[20–22]</sup> Moreover, a myriad of valuable fragrance compounds could be obtained from easily available starting materials using highly regioselective hydratases.

## 1.1 Lily-of-the-valley in fragrance industry

In 2018, the market of flavors and fragrances reached an estimated volume of US\$ 45.6 billion and will further grow over the next five years.<sup>[23]</sup> In most cases, these flavor and fragrance molecules are chemically synthesized or extracted from natural sources. Both approaches, however, exhibit significant disadvantages.<sup>[23,24]</sup> For extraction, main drawbacks are often low concentration of the target molecule, seasonal variation of the raw material, risk of plant diseases, compound instability and trade restrictions. All in all, this can lead to severe availability issues and high prices. Chemical synthesis is in general the cheapest method, but usually requires harsh conditions including toxic catalysts and solvents and often lacks adequate regio- and enantioselectivity. This can



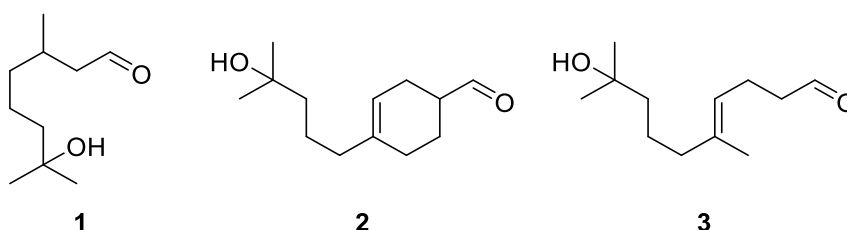
lead to undesired product mixtures which influence the odor of the fragrance.<sup>[25]</sup> Furthermore, the chemically synthesized flavors and fragrances have to be labeled as “artificial” or “nature identical” which decreases their economic value as this mode of production is less accepted by the consumer.<sup>[23,24]</sup> For this reason, biotechnological production methods like *de novo* microbial processes (fermentation) or bioconversions of natural precursors with microbial cells or enzymes (biocatalysis) have seen an increasing interest as the products can be classified as “natural”.<sup>[24,26,27]</sup>

### 1.1.1 Lily-of-the-valley

Currently, lily-of-the-valley fragrances are the most important odor family both in terms of market value and volume, accounting for 20 % of the total raw material market of the perfume industry.<sup>[25]</sup> However, the unique scent of lily-of-the-valley (“muguet”) cannot be obtained as essential oil neither by enfleurage (extraction with odorless fats) nor by distillation.<sup>[25,28,29]</sup> To this end, perfumers have always tried to imitate the scent as precisely as possible using mixtures of alternative essential oils.<sup>[28]</sup> As the chemical synthesis of fragrances has progressed, molecules have also become accessible that exhibit a scent reminiscent of lily-of-the-valley, though it cannot be mimicked by a single compound.<sup>[29,30]</sup> Today, hydroxycitronellal **1** (Figure 1), which was introduced in 1905 by Knoll & Co, is regarded as the prototype for muguet fragrances as its smell comes close to that of natural lily-of-the-valley.<sup>[25,29-31]</sup>

Significant effort has been made to establish a relationship between the molecule’s structure and its scent. The best known rule has been postulated by Pelzer *et al.* in 1992.<sup>[25]</sup> The rule describes the importance of two structural fragments (“Pelzer fragments”) and is based on a hydroxyl moiety (“hydroxy fragment”) and carbonyl function (“carbonyl fragment”), both in a specific chemical surrounding (sterical and electronical aspects are considered). Each of these structural motifs should cause a lily-of-the-valley character.<sup>[25]</sup> For this reason, the perfume industry is searching for hydroxy aldehydes, such as hydroxycitronellal **1**, which combine both substructures and therefore might possess a distinct lily-of-the-valley scent.<sup>[30,32]</sup> Meanwhile, hydroxycitronellal **1** and related compounds like the widely used Lyril **2** have been classified by the EU as skin-sensitizing or allergenic (Figure 1).<sup>[25,32-34]</sup> Nevertheless, these fragrances are still widely applied with an annual production volume of more than 1000 t in case of Lyril **2**.<sup>[25]</sup>

Due to the high demand on the one hand and regulatory pressure on the other hand, fragrance industry is forced to find new substances with equivalent or better perfume properties (with regard to odor characteristics, stability and skin tolerance). A substance with such properties is 9-hydroxy-5,9-dimethyldec-4-enal **3** (HDMDAL). It was introduced by Givaudan in 2014 as an alternative to Lyrals **2** and was named Mahonial **3** after *Mahonia japonica* for its diffusive muguet smell.<sup>[30]</sup> Structurally similar, Mahonial **3** only lacks the CH<sub>2</sub>-bridge which forms the cyclohexene ring of Lyrals **2** (Figure 1).



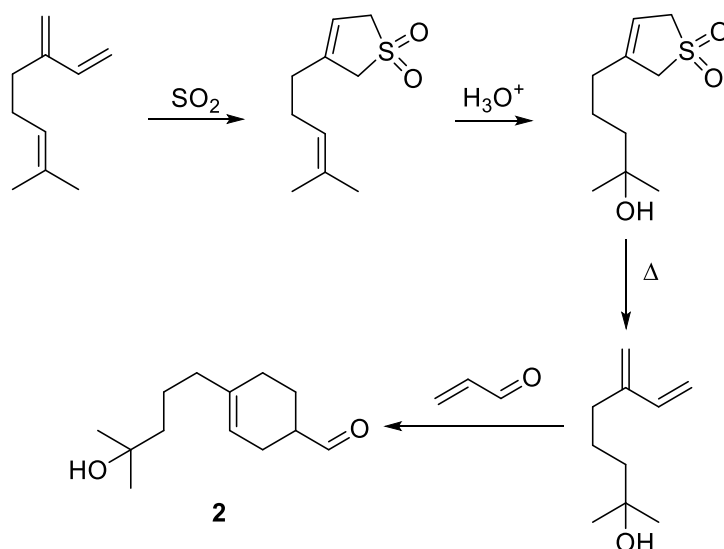
**Figure 1:** Hydroxycitronellal **1**, Lyrals **2** (4-[4-Hydroxy-4-methylpentyl]-1-cyclohex-3-enecarboxaldehyde) and Mahonial **3** are typical representatives from the group of muguet aldehydes.

An even wider application of lily-of-the valley scents is not only limited by regulatory issues but also due to complex syntheses of the hydroxy aldehydes. The development of short synthetic routes using highly selective reactions under mild conditions therefore remains an existing challenge for the organic chemist.

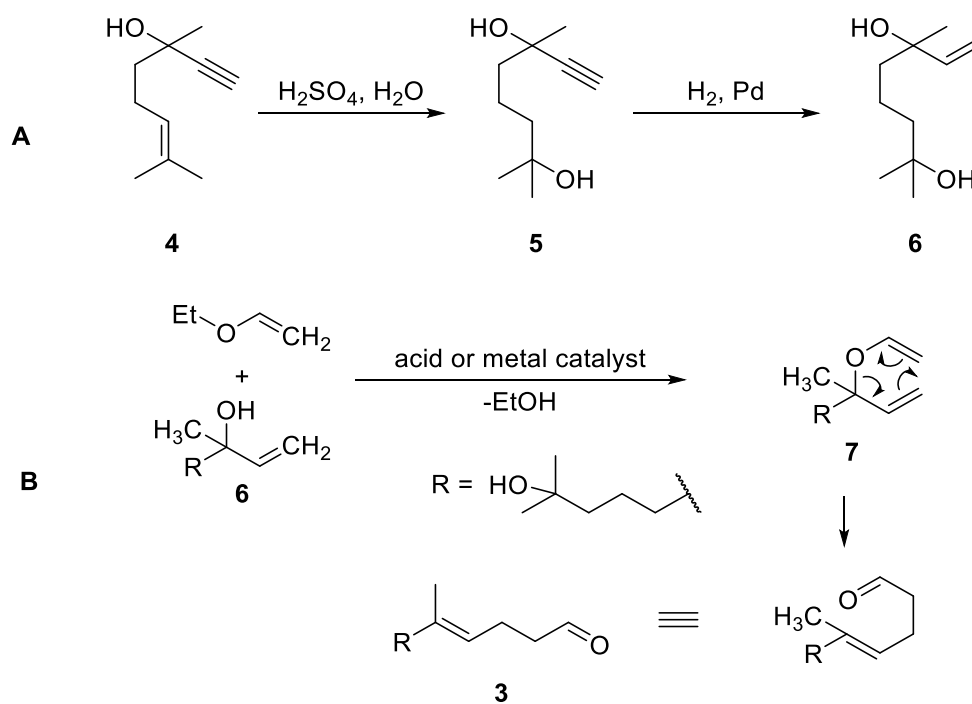
### 1.1.2 Synthesis of lily-of-the-valley hydroxy aldehydes

The chemical synthesis of terpenoid hydroxy aldehydes often begins with already hydroxylated starting material or the synthesis thereof followed by the construction of the target molecule.

For instance, the synthesis of Lyrals **2** starts with the hydration of myrcene to myrcene hydrate (or myrcenol), which can be either achieved directly<sup>[35]</sup>, or preferable by initial protection of the diene with sulfur dioxide to form the cyclic myrcene sulfone and subsequent hydration of the remote double bond with sulfuric acid before being decomposed by gas-phase pyrolysis. Finally, the addition of acrolein gives the desired product Lyrals **2** (Figure 2).<sup>[25,36,37]</sup>



**Figure 2:** Schematic illustration of a synthesis route to Lyral **2**. The figure is created after C. Sell (A fragrant introduction to terpenoid chemistry, page 47).<sup>[36]</sup>



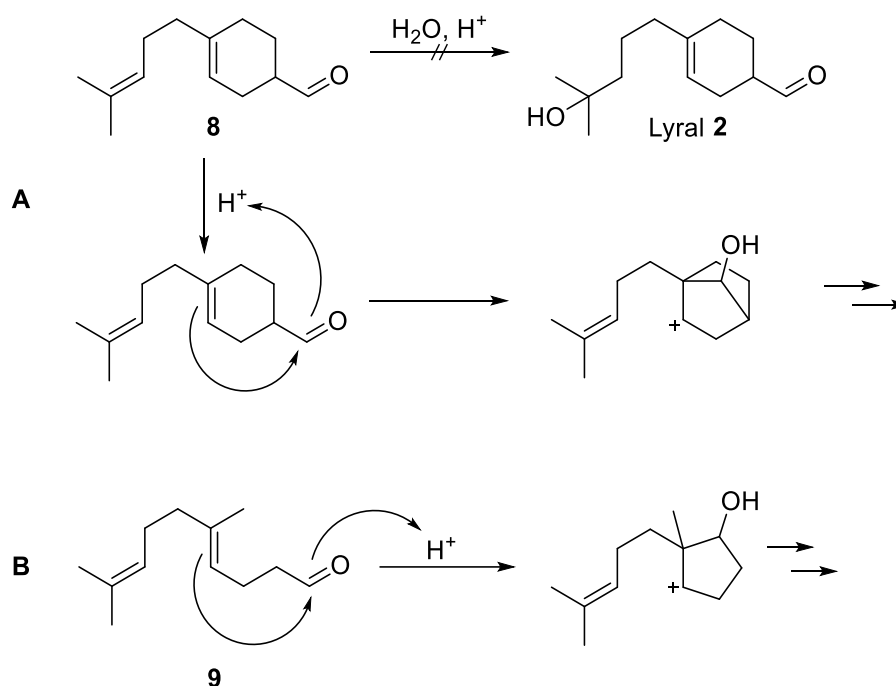
**Figure 3:** Synthesis of Mahonial **3**: A) General preparation of the commonly available starting material hydroxylinalool **6**. B) Proprietary synthesis route to Mahonial **3**. Hydroxylinalool **6** is reacted with ethyl vinyl ether to the allyl vinyl ether compound **7** which then undergoes a Claisen rearrangement to the desired product Mahonial **3**.<sup>[32]</sup>

A similar approach was chosen for Mahonial **3** (Figure 3). Its chemical synthesis and application was patented by Alchenberger *et al.* (Givaudan SA, Vernier, CH).<sup>[32]</sup> The synthesis starts from hydroxylinalool **6**, which is commonly available or can be produced from 3,7-dimethyloct-6-en-1-yn-3-ol **4**.<sup>[32]</sup> In a first step, **4** is hydrated in aqueous sulfuric acid to compound **5** followed by hydrogenation of the triple bond using a palladium catalyst to give hydroxylinalool **6** (Figure 3). Hydroxylinalool **6** is then reacted with ethyl vinyl ether using an acid or metal catalyst. Initially, an allyl vinyl ether **7** is formed by elimination of ethanol. Subsequently, a Claisen rearrangement yields the desired product **3** (Figure 3).<sup>[32]</sup>

The main challenge of both syntheses is the hydration step which requires harsh reaction conditions like, in particular, strong acids. As a consequence, only poor selectivities can be obtained or elaborate countermeasures have to be taken.

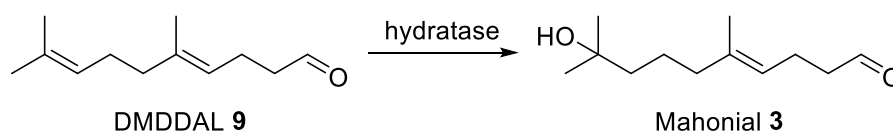
In the case of Lyral **2**, the diene of myrcene is protected as a cyclic sulfone before hydration can take place and has to be deprotected in heat prior to the Diels-Alder-type reaction (Figure 2). A more elegant way would be the hydration of Myraldene **8**, which can be easily obtained by addition of acrolein to myrcene, yielding Lyral **2**.<sup>[25,36]</sup> However, only a poor yield is achieved by direct hydration, as competing side-reaction occur.<sup>[36]</sup> Following a general principle in organic chemistry, the intramolecular reaction (in this case a Prins reaction) to form a 5- or 6-membered ring (Figure 4) is preferred over the intermolecular (hydration) reaction, which is why in this case the aldehyde function has to be protected (for example as an acetal).<sup>[36]</sup>

The same side-reactions are imaginable for the synthesis of Mahonial **3** by direct hydration using strong acids. Starting with the Claisen rearrangement instead of the hydration, linalool can be converted to DMDDAL **9**. A subsequent hydration does not lead to the desired Mahonial **3** but more likely the intramolecular Prins reaction takes place (Figure 4). In general, Mahonial **3** is susceptible to high temperatures and extreme pH values (pH < 4 and pH > 10), with rapid decomposition even at room temperature.<sup>[32]</sup> These problems are addressed in the patent of Alchenberger *et al.* by elaborate washing steps to eliminate residual acids or bases before the final distillation step or by the use of protecting groups during synthesis at the expense of additional reaction steps.<sup>[32]</sup> Especially the tertiary alcohol in the allyl alcohol hydroxylinalool **6** is prone to dehydration and has to be protected as acetate, silyl ether or mixed acetal.<sup>[32]</sup>



**Figure 4:** A): Side reaction of Myraldene **8** during acid-catalyzed hydration. Instead of the desired hydration to Lyrals **2**, rearrangements are observed. The figure is created after C. Sell (A fragrant introduction to terpenoid chemistry, pages 46+47).<sup>[36]</sup> B) Corresponding side reaction of DMDDAL **9** (instead of hydration to Mahonial **3**).

To avoid all the described problems, a sophisticated alternative would be the late introduction of the tertiary alcohol moiety by hydration of the respective alkene under the control of a highly selective catalyst. Hence, a biocatalytic synthesis route would be desirable as the use of enzymes allows mild reaction conditions combined with high selectivities (Figure 5).



**Figure 5:** Envisioned enzymatic synthesis of Mahonial **3** by hydration of the alkene DMDDAL **9** using a regioselective hydratase.

## 1.2 Methods for synthesis of tertiary alcohols

### 1.2.1 Chemical methods

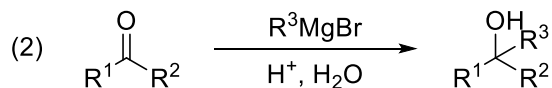
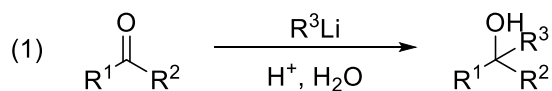
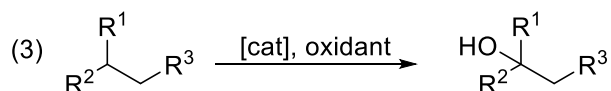
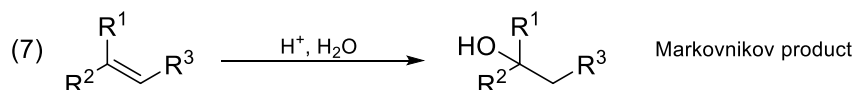
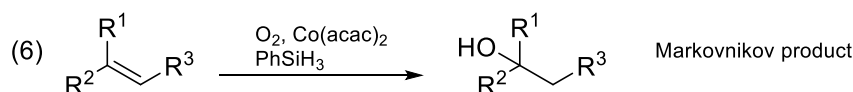
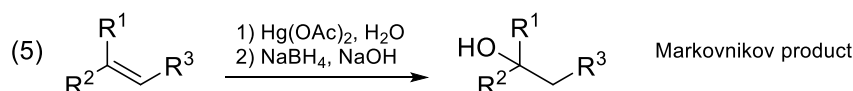
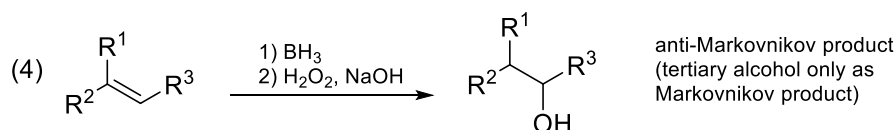
Several chemical methods, including some textbook examples, are described for the synthesis of tertiary alcohols (Figure 6).<sup>[38,39]</sup> Starting from appropriate carbonyl compounds or epoxides, nucleophilic addition of organolithium (RLi) or Grignard reagents (RMgBr) leads to the desired alcohol. Both methods are non-catalytic and further limited by the accessibility of the carbonyl compound. Other methods independent from carbonyls are the selective hydroxylation (C-H oxidation) of non-activated tertiary  $sp^3$  C-H-bonds<sup>[40–42]</sup> catalyzed by different species or the hydration of non-activated C-C double bonds. Hydration can be either achieved by stoichiometric two-step approaches (hydroboration-oxidation or oxymercuration-demercuration) or catalytically either by the radical Mukaiyama reaction or by electrophilic addition of water catalyzed by strong acids.

Hydroboration allows the regioselective functionalization of alkenes, as borane adds to C-C double bonds without catalytic activation.<sup>[39,43–45]</sup> In the case of differently substituted alkenes, the boron atom adds to the sterically less hindered carbon atom. However, the selectivity of hydroboration for similarly substituted alkenes is relatively low. This is overcome by the use of sterically hindered organoboranes like 9-borabicyclo[3.3.1]nonane (9-BBN). The oxidative workup with hydrogen peroxide and sodium hydroxide gives in general the favored anti-Markovnikov alcohol.<sup>[39,43–45]</sup> In comparison, oxymercuration-demercuration selectively provides access to alcohols from alkenes according to Markovnikov's rule. The alkene is reacted with mercury acetate (or other Hg(II) salts) in a mixture of a protic solvent such as THF and water to produce a mercury alcohol. The mercury compound is then reduced (demercuration) to the desired alcohol with sodium borohydride and sodium hydroxide solution.<sup>[39,43–46]</sup> Hydroboration and oxymercuration are useful methods for the regioselective lab-scale synthesis of alcohols from olefins, but are impractical for industrial use due to the stoichiometric amounts of toxic reagents. The issue of bad atom economy has been tackled by development of catalytic hydroboration-oxidation methods for specific cases of fluorinated olefins using an appropriate combination of rhodium catalysts and boranes.<sup>[47]</sup> In this rare example, the authors were also able to make tertiary alcohols (Markovnikov product).<sup>[47]</sup>

Mukaiyama hydration allows formal hydration (oxidation-reduction) of olefins with high chemo- and regioselectivity under mild conditions.<sup>[48]</sup> The method uses molecular oxygen, phenylsilane and bis(acetylacetonato)cobalt(II) for the formation of alcohols according to Markovnikov's rule.<sup>[48]</sup> The possibility to form tertiary alcohols from prenyl double bonds was exploited in the total synthesis of several natural products e.g. (±)-garsubellin<sup>[49]</sup> or stigmolone<sup>[50]</sup>.

Direct hydration of olefins by electrophilic addition of water is an alternative with outstanding atom efficiency.<sup>[38]</sup> The reaction is typically catalyzed by strong acids. However, the reverse dehydration can also take place (equilibrium reaction) and the formation of tertiary alcohols is only favored at low temperatures.<sup>[38]</sup> Additionally, acid-catalysis is unsuited for more complex substrates, as the lack of selectivity leads to a variety of by-products through isomerization, rearrangements and other acid-induced reactions and thereby to a low yield.<sup>[38,51–53]</sup> Demanding molecules like terpenoids are therefore not easy to obtain by simple acid-catalyzed hydration (1.1.2).<sup>[54]</sup> Hydration is rather achieved by using solid acids including zeolites, heteropoly acids or resins as alternative systems.<sup>[53–56]</sup> Examples are the hydration of dihydromyrcene to dihydromyrcenol by using a biphasic system with a zeolite catalyst<sup>[57]</sup> or trifluoromethanesulfonic acid in ionic liquids<sup>[55]</sup>. Similar studies on  $\alpha$ -pinene yielded only product mixtures.<sup>[53,58–60]</sup>

Despite advances of heterogeneous and homogeneous catalysis in the development of hydration methods for the synthesis of tertiary alcohols, there is still room for improvement due to the lack of regio- and stereocontrol and increasing demand for “greener” methods. Biocatalysis can provide ideal alternatives with highly regioselective and atom-economic enzymes working under mild reaction conditions. The identification of suitable enzymes in databases and literature is an important and challenging task.

**Nucleophilic addition****C-H Oxidation (Hydroxylation)****Hydration**

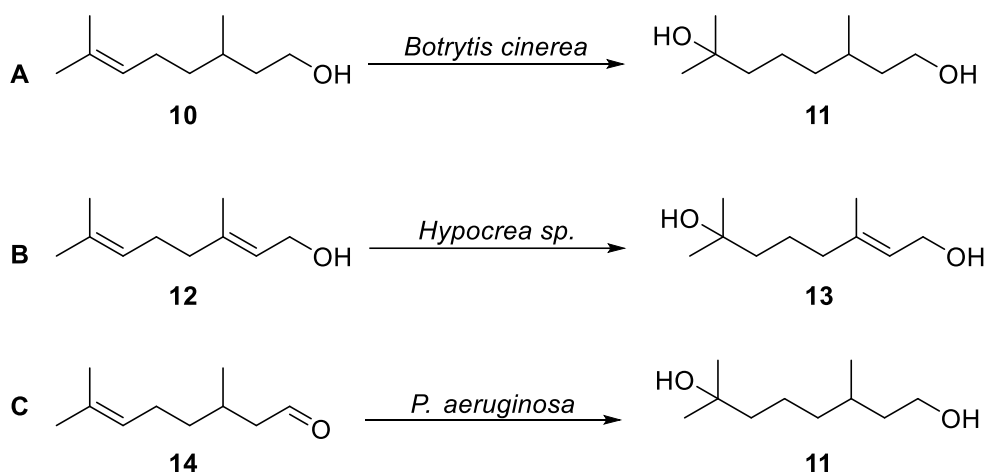
**Figure 6:** Text book examples for the synthesis of tertiary alcohols. Syntheses by nucleophilic addition (1, 2) of metalorganic compounds (organolithium or Grignard reagent) to carbonyls are limited by the accessibility of the carbonyl compounds. An alternative can be the selective hydroxylation (C-H oxidation) by activation of tertiary C-H bonds (3) or the hydration of non-activated C-C double bonds in alkenes. Hydration can be achieved in two-step approaches either by hydroboration-oxidation (4) or oxymercuration-demercuration (5). Mukaiyama hydration (6) leads to a formal hydration of water to an olefin. Direct hydration by electrophilic addition of water catalyzed by strong acids (7) is an alternative with outstanding atom efficiency. Tertiary alcohols are only obtained in hydration reactions yielding the Markovnikov product (5-7). Hydroboration (4) in general prefers the anti-Markovnikov product for trisubstituted alkenes. However, tertiary Markovnikov product is formed to a lesser extent or regioselective in specific examples.



### 1.2.2 Examples from biology for the synthesis of tertiary terpenoid alcohols by hydration

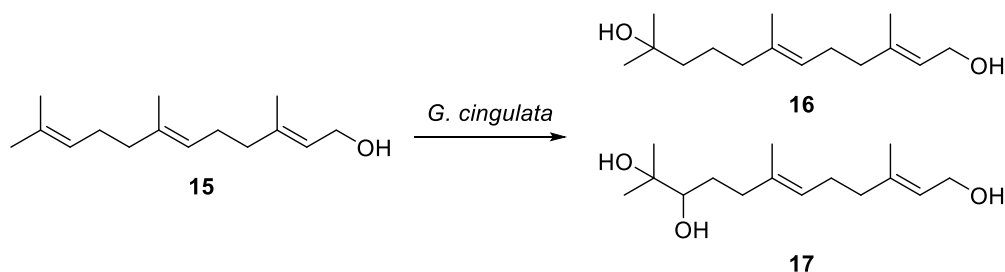
In literature, several vague examples for the formal hydration of terpenoids are described using a variety of fungi and bacteria. However, mainly fermentation or whole cell biotransformations are described and the isolation and characterization of responsible enzymes is missing in most cases.

Brunerie *et al.* used different *Botrytis cinerea* (“grey mold”) strains, which are responsible for noble rot in wine grapes, for the conversion of citronellol **10** to more than ten products (Figure 7).<sup>[61]</sup> Main products were 2,6-dimethyl-1,8-octandiol and (*E*)-2,6-dimethyl-2-octen-1,8-diol. The hydration product 3,7-dimethyloctane-1,7-diol **11** was formed in a range of only 5-15 % (0.3-0.6 mg L<sup>-1</sup>) but was also found in the negative controls without *B. cinerea* (0.2 mg L<sup>-1</sup>). The hydration product was only formed when grape must was used instead of a synthetic medium.<sup>[61]</sup> In comparison, studies of the same group using *B. cinerea* for the biotransformation of linalool showed no hydration product.<sup>[62]</sup> In 2009, Leutou *et al.* demonstrated the regioselective hydration of geraniol **12** by marine-derived fungus *Hypocrea* sp. to (*E*)-1,7-dihydroxy-3,7-dimethyl-oct-2-ene **13** (Figure 7). The product was isolated from crude extract by silica gel column chromatography (45 % isolated yield), whereby also substrate was recovered (28 %).<sup>[63]</sup> No indications on side-products were made. Joglekar and Dhavlikar investigated earlier in 1969 the metabolization of the terpenoid aldehydes citronellal **14** and citral by soil bacterium *Pseudomonas aeruginosa* (Figure 7).<sup>[64]</sup> The strain was able to use these terpenes as the sole source of carbon and energy. Mainly, oxidation took place and the corresponding acids citronellic acid and geranic acid were obtained. However, in the case of citronellal **14**, also crystals with pleasant odor were obtained and identified by the authors as the diol **11**. Further metabolites were citronellol, dihydrocitronellol and menthol. No hydration product was detected for citral.<sup>[64]</sup> While regioselectivity was not relevant for hydration of substrates **10** and **14** as they have only one double bond, the hydration of geraniol **12** to **13** is a remarkable example for a regiospecific enzyme.



**Figure 7:** A) Biochemical conversion of citronellol **10** to 3,7-dimethyloctane-1,7-diol **11** formed as a side product by *Botrytis cinerea*.<sup>[61]</sup> B) Biotransformation of geraniol **12** with the marine-derived fungus *Hypocrea sp.* to (*E*)-1,7-dihydroxy-3,7-dimethyl-oct-2-ene **13**.<sup>[63]</sup> C) Biotransformation of citronellal **14** using *P. aeruginosa*.<sup>[64]</sup>

Another hypothetical regioselective hydration at the remote double bond was also claimed by Miyazawa and coworkers when incubating the plant pathogenic fungus *Glomerella cingulata* with (*2E,6E*)-farnesol **15** as one of the main metabolites was the hydration product **16** next to triol **17** (Figure 8).<sup>[65]</sup> The authors claimed, that the two products are formed by hydration (in case of **16**) or epoxidation and subsequent hydrolysis (in case of **17**) before being further metabolized.<sup>[65]</sup> Earlier studies of the authors on the biotransformation of *cis*-nerolidol<sup>[66]</sup>, nerylacetone<sup>[66]</sup>, *trans*-nerolidol<sup>[67]</sup> and geranylacetone<sup>[67]</sup> suggested that an oxygenation is more likely than a redox-neutral hydration as in all cases the diol is formed and the selectivity differs depending on the configuration of the substrate (*cis* = diol, *trans* = “hydration” product is preferred). Nevertheless, in all cases, the putative hydratase was not identified and an unselective dioxygenase cannot be ruled out as responsible enzyme for both products.



**Figure 8:** Biotransformation of farnesol **15** with the plant pathogenic fungus *Glomerella cingulata*. Besides the hydration product (*2E,6E*)-3,7,11-trimethyldodeca-2,6-diene-1,11-diol **16**, also the (*2E,6E*)-3,7,11-trimethyldodeca-2,6-diene-1,10,11-triol **17** was formed from the double bond, which was speculated by Miyazawa *et al.* to be the product of epoxidation and subsequent hydrolysis.<sup>[65]</sup>

The regioselective hydroxyl functionalization of terpenoids has been described in diverse cases and even the formation of tertiary alcohols from prenyl double bonds was reported. Involved enzymes are likely mono- or dioxygenases or epoxide hydrolases. Hydratases can be a promising yet almost unexplored option to those enzymes. Main advantages of hydratases are their independency of cofactors, unmatched atom-economy (only water is needed) and high selectivities. Therefore, hydratases for the regioselective hydration of isolated non-activated double bonds are discussed in detail in the following chapter.

### 1.3 Hydratases

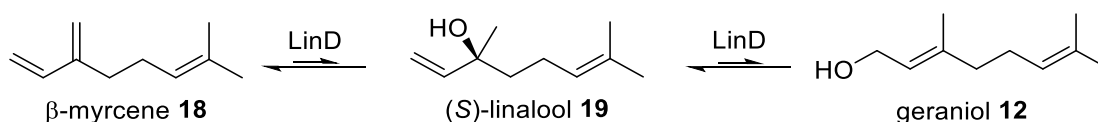
Hydratases are enzymes that catalyze the addition of water or the reverse elimination and belong to the class of C-O lyases (EC 4.2.1.X).<sup>[52,68]</sup> They provide a promising option to established enzymes for the synthesis of valuable alcohols as they show good activity for their natural substrate, are highly regio- and stereoselective, in general well expressed and reasonable stable.<sup>[69]</sup> This makes them also an alternative to chemical hydration methods (1.2), which are suffering from low selectivities caused by the use of strong acids and harsh reaction conditions.<sup>[52,53]</sup> Moreover, hydratases are interesting for industrial applications since their properties make the use of protective groups unnecessary. Thus, reaction steps can be saved and by-products can be avoided.<sup>[26,70]</sup>

Hydratases acting on C-C double bonds can be divided into two distinct groups: hydratases that act on  $\alpha,\beta$ -unsaturated carbonyl compounds (Michael systems) and hydratases for the addition of water to isolated C-C double bonds.<sup>[53]</sup> The reactions catalyzed by hydratases often play a central role in primary or secondary metabolism. The addition of water to a Michael system is for instance part of the citric acid cycle. An example is the reaction of fumaric acid to (*S*)-malate, which is catalyzed by fumarase (EC 4.2.1.2).<sup>[71]</sup> In contrast, the hydration of isolated double bonds is of central importance in the formation and degradation of terpenes and carotenoids.<sup>[72]</sup>

Highly regio- and stereoselectivity hydratases which catalyze the hydration of isolated C-C double bonds are cofactor-independent and particularly interesting for use in organic synthesis. Possible hydratases for the synthesis of Mahonial **3** by regioselective hydration of a prenyl moiety (Figure 5) are discussed in the following chapters.<sup>[69,73]</sup>

### 1.3.1 Linalool dehydratase-isomerase

Linalool dehydratase-isomerase (LinD or LDI, EC 4.2.1.127)<sup>[69,73]</sup> was first identified in 2010 in the  $\beta$ -proteobacterium *Castellaniella defragans* by Brodkorb *et al.*<sup>[74]</sup> The native enzyme has a molecular mass of 160 kDa and gel electrophoresis suggested that it consisted of four 40 kDa homotetramer units.<sup>[74]</sup> LinD reversibly catalyzes the enantioselective isomerization of the primary monoterpene alcohol geraniol **12** into the tertiary alcohol (*S*)-(+)-linalool **19** and the dehydration to  $\beta$ -myrcene **18** (Figure 9).<sup>[73]</sup> This equilibrium reaction favors the formation of the thermodynamically most stable  $\beta$ -myrcene.<sup>[75]</sup> Expression is likely to be periplasmic because the open reading frame features a precursor protein with leader sequence.<sup>[74,76,77]</sup>



**Figure 9:** Reaction of  $\beta$ -myrcene **18** to linalool **19** and isomerization to geraniol **12**, catalyzed by linalool dehydratase-isomerase (LinD).<sup>[73]</sup>

The crystal structure of LinD was independently solved by Nestl *et al.*<sup>[78]</sup> and Weidenweber *et al.*<sup>[75]</sup> in 2017. In contrast to what was previously assumed by Brodkorb and coworkers, both crystal structures revealed a homopentamer with the monomers consisting mainly of  $\alpha$ -helices and the active site was found to be located at the interface of two neighboring monomers.<sup>[73]</sup> Both groups also speculated on the mechanism of LinD based on the identified residues in the active site. While Weidenweber and colleagues only proposed an acid/base mechanism, Nestl *et al.* additionally suggested a covalent mechanism.<sup>[75,78]</sup> Recently, more insights on the importance of the protonation state of active site residues for the mechanism were gained by QM/MM studies on two models which have been constructed based on LinD-linalool and LinD-myrcene enzyme-substrate complexes.<sup>[79]</sup> The studies found the proposed mechanism of Nestl *et al.* to be unlikely.<sup>[79]</sup>

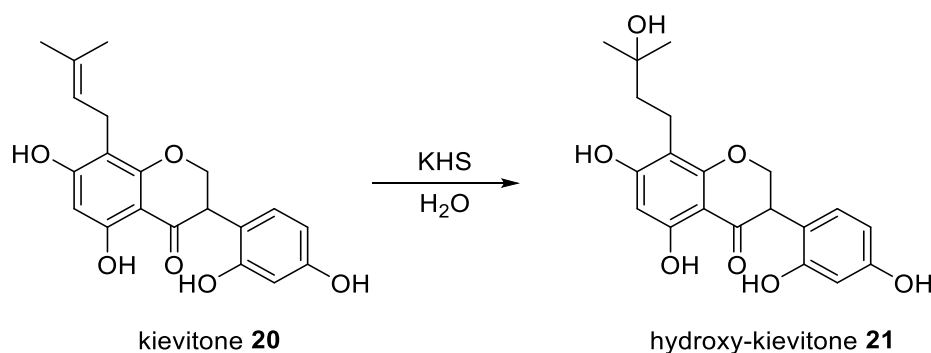
Nestl *et al.* further investigated the substrate scope of LinD and revealed that the enzyme has a broader substrate spectrum than previously assumed.<sup>[78]</sup> In addition to the monoterpenes (C10) myrcene, linalool and geraniol, a range of truncated and elongated aromatic and aliphatic tertiary alcohols were accepted. For instance, the sesquiterpenes (C15) farnesol and nerolidol as well as the C5 alcohol prenol (3-methyl-2-buten-1-ol)

could be converted. However, only those substrates were accepted, which have a tertiary  $\alpha$ -methylallyl alcohol as a motif, whereas the organic group may be variable. It should be noted that the substrate scope was tested with focus on the dehydration reaction.<sup>[73,78]</sup>

The potential application of this enzyme has already been applied for patent in several cases. For example, the production of butadiene, 1-propanol or 1,2-propanediol from a fermentable carbon source was claimed in 2015. In this pathway, a linalool dehydratase-isomerase was used to convert crotyl alcohol to butadiene.<sup>[80]</sup> The production of volatile dienes by enzymatic dehydration of short-chain alcohols by means of an alkenol dehydratase was also filed for a patent by Marliere mentioning LinD from *Castellaniella defragans* as an example.<sup>[81]</sup>

### 1.3.2 Kievitone hydratase

The addition of water to the plant isoflavonoid kievitone **20** is catalyzed by kievitone hydratase (KHS; EC 4.2.1.95).<sup>[69]</sup> Thereby, the prenyl moiety of kievitone **20** is transformed into a tertiary alcohol giving hydroxy-kievitone **21** (Figure 10).<sup>[82]</sup> Fungal pathogens use this reaction as a defense mechanism against phytoalexins from the host.



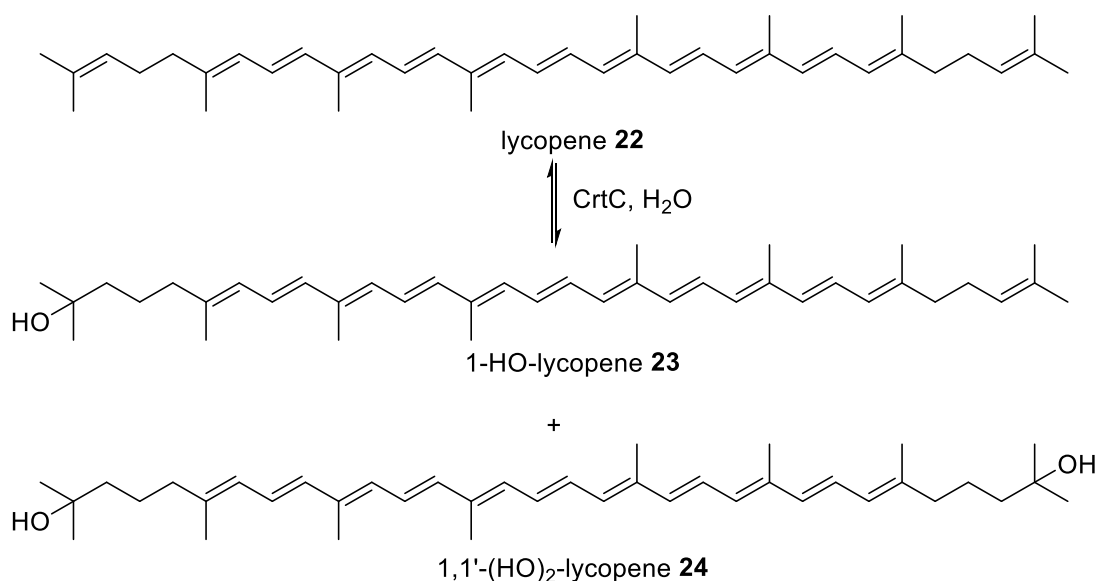
**Figure 10:** Hydration of plant isoflavonoid kievitone **20** to hydroxy-kievitone **21** by kievitone hydratase.

Kievitone hydratase was first described in 1979 by Kuhn and Smith in *Fusarium solani* f. sp. *phaseoli* (FsKHS).<sup>[82]</sup> Later, Cleveland and Smith (1983) performed first investigations using partial purified KHS while it was not until 1995 when the gene sequence of FsKHS was identified, also in the related teleomorph *Nectria haematococca* (NhKHS).<sup>[82,83]</sup> Based on the gene and cDNA sequences, the group of Schardl also found that the enzyme secretion is enabled through an N-terminal transit peptide of 19 amino acids, which is removed during maturation.<sup>[83]</sup>

Only recently, Engleder *et al.* searched for sequences similar to FsKHS in *N. haematococca* by BLAST suite of NCBI. The alignment revealed five significant sequences, each representing a hypothetical protein with predicted hydroxyneurosporene synthase (equivalent to carotenoid hydratase, CrtC, chapter 1.3.3) or peptidase activity without indication for kievitone hydratase activity.<sup>[84]</sup> The sequence with highest identity (58 %) to FsKHS was heterologously expressed using *Pichia pastoris* in secretory mode and proofed activity for kievitone **20**. Beside its eponymous substrate, KHS also accepts other flavonoids bearing a 3-methylbut-2-en-1-yl moiety, like xanthohumol, isoxanthohumol and 8-prenylnaringenin.<sup>[84]</sup> So far, neither structural nor mechanistic studies on KHS have been performed. However, conserved regions in different putative KHS were identified by amino acid sequence alignment suggesting an function in substrate binding or catalysis.<sup>[69,84]</sup>

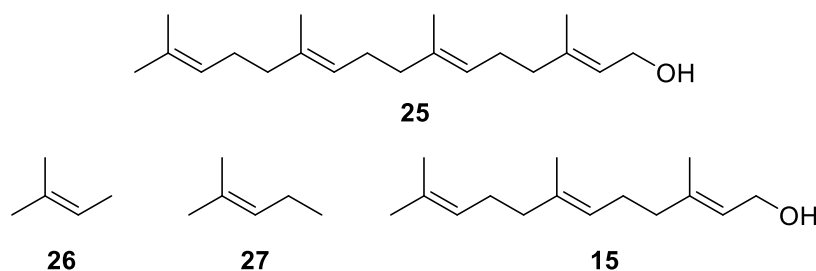
### 1.3.3 Carotenoid hydratases

Another class of enzymes from terpene metabolism capable of hydrating a prenyl motif are carotenoid hydratases (CrtC, EC 4.2.1.131). CrtCs catalyze the terminal hydroxyfunctionalization of carotenoids and are also (according to the nomenclature of carotenoids) referred to as carotenoid 1,2-hydratases since the addition of water to the C-C double bond takes place in 1,2-position.<sup>[51,85,86]</sup> The hydration mechanism was investigated and proofed by labeling experiments using D<sub>2</sub>O and <sup>18</sup>O-labeled water.<sup>[87,88]</sup> CrtCs are found in several photosynthetic and non-photosynthetic bacteria and a distinction is made between two carotenoid hydratases with respect to the function and the type of natural substrate: CrtC and CruF.<sup>[69,86,89-95]</sup> CruF catalyzes the hydration of  $\gamma$ -carotene and was found in *Deinococcus radiodurans* R1 and *Deinococcus geothermalis* DSM 11300.<sup>[86]</sup> By contrast, the CrtC from *Rubrivivax gelatinosus* is a membrane-bound enzyme that catalyzes the hydration of neurosporene into 1-hydroxyneurosporene and also of lycopene **22** to 1-hydroxylycopene **23** (rhodopin) and the dihydroxylated product 1,1'-dihydroxylycopene **24** (Figure 11).<sup>[91,92]</sup> However, only the monohydroxylated product was observed for the CrtC from *Rhodobacter capsulatus*, as 1-HO carotenoids are themselves poor substrates for this enzyme.<sup>[92]</sup>



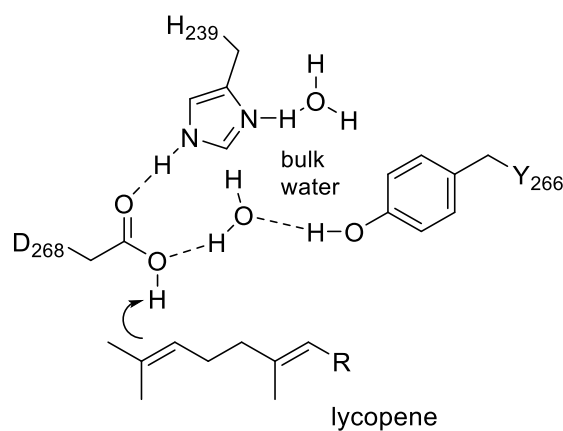
**Figure 11:** Reaction of lycopene **22** to 1-hydroxylycopene **23** (rhodopin) and the twice hydrated product 1,1'-dihydroxylycopene **24**.<sup>[51,96]</sup>

In 2011, Hiseni *et al.* reported the cloning and successful expression of CrtC from *Rubrivivax gelatinosus* (RgCrtC) and *Thiocapsa roseopersicina* (TrCrtC) in *E. coli*.<sup>[96]</sup> SDS-PAGE analysis showed for both proteins bands of 44 kDa with an around two times higher expression level for RgCrtC. However, for TrCrtC there was an additional weak band around 38 kDa indicating protein processing and loss of an N-terminal proline rich part. The same study revealed an optimum activity for the hydration of the natural substrate lycopene at 30 °C and pH 8. In addition, the substrate specificity was investigated with respect to shorter analogues, all of which have a terminal alkenyl function. For the hydration of geranylgeraniol **25** (C20), a peak, which was absent in the negative control experiment, was detected for both CrtCs (about 5 % conversion). However, the study lacks the proof with an authentic standard or GC-MS spectrum. The substrates 2-methyl-2-butene **26** (C5), 2-methyl-2-pentene **27** (C6) and farnesol **15** (C15) could not be converted (Figure 12). These results indicate that CrtCs are promising enzymes for the formation of tertiary alcohols, although no (industrial) application has been reported so far.<sup>[51,96]</sup>



**Figure 12:** Shorter substrates tested in the study by Hiseni *et al.*<sup>[96]</sup> The diterpenoid (C20) geranylgeraniol **25** was successfully hydrated while the even shorter substrates such as the sesquiterpenoid (C15) farnesol **15**, 2-methyl-2-pentene **27** (C6) and the hemiterpenoid 2-methyl-2-butene **26** (C5) were not accepted.

Hiseni and coworkers also reported in 2016 on catalytically relevant amino acid residues and proposed a putative mechanism (Figure 13) although no crystal structure is available.<sup>[73,97]</sup> The residues H239, W241, Y266 and D268 were identified (for RgCrtC, corresponding positions in TrCrtC) based on a homology model and a multi-sequence alignment of 100 CrtC-like sequences. Subsequently, an alanine scan at these positions was performed to confirm the importance of the conserved residues. The proposed mechanism starts with an initial protonation step by a highly acidic aspartic acid protonating the double bond and thereby generating an intermediate carbocation at C2 of the substrate, which is subsequently attacked by water to give the tertiary alcohol. The protonation step was compared by the authors to the mechanism of squalene-hopene cyclase from *Alicyclobacillus acidocaldarius*, which shares active site residues with conserved amino acids in CrtC.<sup>[69,73,97]</sup>

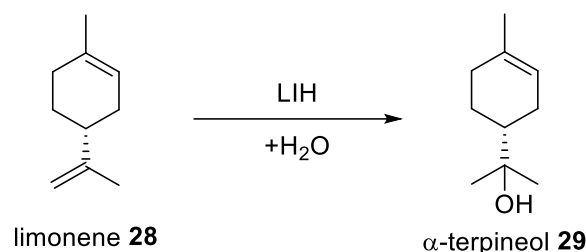


**Figure 13:** Catalytic mechanisms of CrtC. Lycopene hydration is initiated by protonation via a highly acidic aspartic acid.<sup>[73,97]</sup>



### 1.3.4 Limonene hydratase

The addition of water to the 8,9-double bond of (*R*)-(+)-limonene **28** leading to (*R*)-(+)- $\alpha$ -terpineol **29** is a rare example for the regio- and stereoselective enzymatic hydration of monoterpenes (Figure 14) catalyzed by putative limonene hydratase (LIH).<sup>[69,98]</sup>



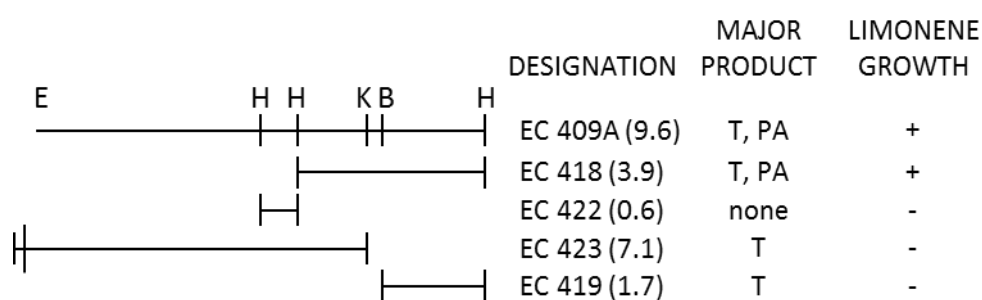
**Figure 14:** Reaction of limonene **28** to  $\alpha$ -terpineol **29** catalyzed by putative limonene hydratase.<sup>[99]</sup>

So far, the production of  $\alpha$ -terpineol from limonene has been reported mainly for various fungi (*Fusarium oxysporum*, *Pleurotus sapidus*, *Aspergillus niger* spp., *Penicillium* spp.) but also for some bacteria (*Bacillus stearothermophilus*, *Pseudomonas gladioli*, *Sphingobium* spp.).<sup>[51,53,69,99–101]</sup> However, the reported biotransformations were performed predominantly in whole cells, while the isolation and characterization of the responsible enzymes are missing in most cases.<sup>[69]</sup> This can make a revision of the proposed reaction path necessary, as in the case of formal hydration of limonene to  $\alpha$ -terpineol in *Penicillium digitatum*. The reaction was later found to take place via a two-step reaction comprising the epoxidation of the double bond and subsequent oxidative cleavage of the epoxide.<sup>[69]</sup>

An exception is the LIH from *Pseudomonas gladioli* isolated by Cadwallader *et al.*<sup>[102,103]</sup> The authors designated the enzyme as  $\alpha$ -terpineol dehydratase, to avoid the ambiguous name limonene hydratase, even though a dehydration reaction of  $\alpha$ -terpineol could not be detected even after 20 h. The hydratase could be solubilized partially from the membrane using 2 % *m V*<sup>-1</sup> Triton X-100. The molecular weight was determined to be 94.5 kDa for the monomer and 206.5 kDa for the dimer, which was probably due to an enzyme-detergent complex.<sup>[102,103]</sup>

The only work to date on a thermostable limonene hydratase heterologously expressed in *E. coli* is by the group of Oriol.<sup>[104–106]</sup> They isolated the thermophilic *Bacillus stearothermophilus* BR388 (classified as *Geobacillus stearothermophilus* since 2001) using a limonene enrichment culture from orange peel.<sup>[105]</sup> Subsequently, the pathway responsible for limonene degradation was cloned as a 9.6 kb chromosomal fragment in

*E. coli* and the successful transformation was confirmed by growth of *E. coli* EC409A on limonene as sole carbon source. The complete oxidation of limonene to CO<sub>2</sub> and H<sub>2</sub>O via perillyl alcohol, perillyl aldehyde and perillic acid was described as main degradation path. As a byway, the hydration of limonene to  $\alpha$ -terpineol by a hydratase has been proposed.<sup>[104]</sup> The production of  $\alpha$ -terpineol and perillyl alcohol with this strain has been reported and the process filed for patent by Savithiry *et al.* in 1998.<sup>[107]</sup> Further studies on the localization of the hydratase in the limonene pathway were carried out by digesting the 9.6 kb fragment by restriction digestion into smaller fragments of 3.9 kb, 0.6 kb, 7.1 kb and 1.7 kb, which were again cloned into *E. coli* (EC 418, EC 422, EC 423, EC 419) (Figure 15).



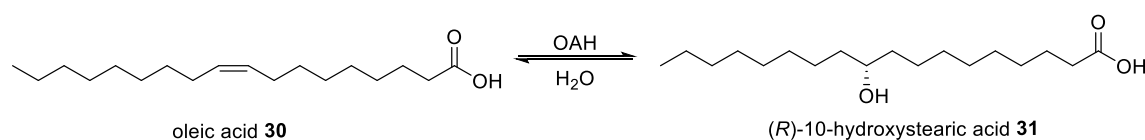
**Figure 15:** Subcloning of the 9.6 kb fragment by restriction digest. Restriction enzymes: E = EcoRI, H = HindIII, K = KpnI, B = BglII. Products: T =  $\alpha$ -terpineol, PA = perillyl alcohol.<sup>[107]</sup>

It could be shown that EC 418 has both methyloxidase and hydratase activity and grows on limonene. EC 419 showed only hydratase activity, with the only product being  $\alpha$ -terpineol. Remarkably, EC 423 also showed limonene hydratase activity, suggesting that two different hydratases in the pathway are responsible for the production of  $\alpha$ -terpineol. An initial characterization of the hydratase has been conducted as crude extract. Broad substrate specificity was found as besides the hydration of limonene also the nitrile group of cyanopyridine was converted to nicotinamide. The latter reaction was also used as assay and the product was detected by HPLC. Further investigations showed that the hydratase from EC 419 was thermally unstable (inactive above 40 °C), while the hydratase from EC 423 showed a temperature optimum of about 55 °C.<sup>[106,107]</sup>

The industrial attractive conversion of the cheap limonene to (*R*)-(+)- $\alpha$ -terpineol in whole cell processes has already been reported several times.<sup>[53,108–111]</sup> The highest yield so far (130 g L<sup>-1</sup> in 96 h) was achieved with resting cells from *Sphingobium* sp. in a biphasic system.<sup>[53,112]</sup>

### 1.3.5 Oleate hydratases

Fatty acid hydratases (FAH) are remarkable examples of hydratases that can add water to isolated C-C double bonds of free fatty acids. With more than 25 examples, FAHs are the far most studied group of cofactor-independent hydratases.<sup>[69]</sup> The first description is from 1962 by Wallen *et al.* when they observed hydration of oleic acid **31** to 10-hydroxystearic acid **32** (Figure 16) in *Pseudomonas* sp. 3266.<sup>[113]</sup> However, it was only in 2009 when Bevers *et al.* isolated the responsible enzyme, characterized it biochemically after heterologous expression and designated it as oleate hydratase (OAH, EC 4.2.1.53).<sup>[114]</sup>



**Figure 16:** Reaction of oleic acid **30** to (R)-10-hydroxystearic acid **31** catalyzed by oleate hydratase.<sup>[73]</sup>

Meanwhile, many FAHs were found in a broad range of Gram-negative and Gram-positive bacteria and identified as OAH (most cases), linoleic acid hydratase (LAH) or general FAH, named after the substrate with the highest activity.<sup>[69,73]</sup> Recently, our group established the *Hydratase Engineering Database* (HyED; URL: <https://hyed.biocatnet.de>) comprising 2046 sequences of known and putative hydratases, including mainly sequences of bacterial sources, but also 103 from fungi and 24 from archaea, which were allocated to eleven homologous families (HFam1-11).<sup>[115]</sup>

The natural function of OAH has not yet been clarified.<sup>[114]</sup> It is assumed that the hydration of unsaturated fatty acids with FAHs such as MCRA (myosin cross-reactive antigen) proteins and OAHs is a detoxification mechanism. Unsaturated fatty acids are toxic to many organisms as they inhibit enoyl-ACP reductases and, beyond, their angled structure interferes with the assembly of cell membranes by disturbing the lipid bilayer formation.<sup>[116-119]</sup>

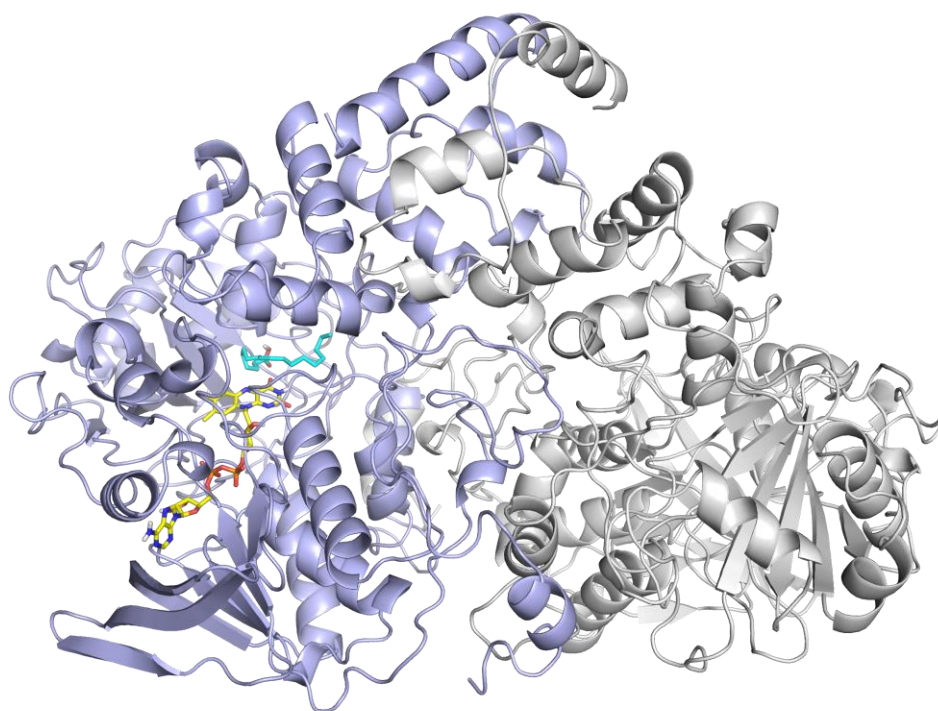
### 1.3.5.1 Structural aspects

In contrast to the fact that OAHs are described as cofactor-independent, this is only true for the actual hydration reaction. Indeed, all characterized FAHs share a conserved N-terminal nucleotide binding motif for non-covalent binding of an essential flavin adenine dinucleotide (FAD) cofactor, whose distinct function is unclear so far. It is therefore speculated that FAHs belong to the 10 % of flavoenzymes harboring a non-redox active cofactor, which might play a structural role or is relevant for charge stabilization.<sup>[69,73,117,119,120]</sup> Schmid *et al.* identified in total 80 highly conserved residues of structural and functional importance, which are present in > 90 % of the sequences from the HyED and mainly located in the areas responsible for FAD binding, substrate recognition and catalysis.<sup>[69,115]</sup> Elucidation of crystal structures revealed more insights on the location of these conserved residues. Meanwhile, the crystal structures of oleate hydratases from *Lactobacillus acidophilus*<sup>[117]</sup>, *Elizabethkingia meningoseptica*<sup>[120]</sup>, *Rhodococcus erythropolis*<sup>[121]</sup> and *Stenotrophomonas* sp. KCTC 12322<sup>[122]</sup> are available, belonging to the homologous families HFam2, HFam11, HFam3 and HFam11, respectively. Additionally, a homology model of a FAH from *L. plantarum* CFQ-100 (LPH) was created based on the crystal structure of *L. acidophilus*.<sup>[123]</sup>

The first crystal structure was reported in 2013 by Volkov *et al.* for the FAH from *Lactobacillus acidophilus* (LaOAH, nomenclature in the primary literature: LAH).<sup>[117]</sup> The authors obtained structures of apo-LAH (without substrate) and LA-LAH with bound linoleic acid at resolutions of 2.3 Å and 1.8 Å, respectively (PDB: 4IA5 & 4IA). As the co-crystallization of FAD was not successful, it was not possible to locate the active site. LAH was found to be a homodimer with each protomer consisting of four connected domains. The substrate- and putative FAD-binding sites are formed by three domains and exhibit structural similarity to several other FAD-dependent enzymes like amine oxidoreductases. The mainly  $\alpha$ -helical fourth domain is located at the C-terminus and covers the entrance to a hydrophobic substrate channel, which extends from the protein surface to the putative active site. This lid-domain changes its conformation upon binding of linoleic acid and thereby opens the substrate channel of the other protomer.<sup>[117]</sup>

Only in 2015, the groups of Pichler and Gruber succeeded in crystallizing the OAH from *Elizabethkingia meningoseptica* (EmOAH, nomenclature in the primary literature: OhyA) in the presence of FAD by incubating the protein with FAD prior to the last step of purification (PDB: 4UIR).<sup>[120]</sup> The structure was solved at 2.75 Å and revealed,

similar to LAH, a homodimer composed of four domains with a non-covalently bound FAD co-crystallized in chain A at the interface of a Rossmann-type domain and the substrate-binding domain. Both chains are structurally very similar with a notable difference in a loop region (R<sub>118</sub>GGREM<sub>123</sub>) covering residues in the FAD-binding pocket, which adopted a well-ordered conformation in the presence of the cofactor, while being less ordered in the absence of FAD. This suggests a structural role of cofactor binding. The authors identified a narrow, elongated and mainly hydrophobic cavity close to the FAD as the V-shaped active site. Moreover, they successfully docked the substrate oleic acid with the double bond at the bend of pocket (Figure 17).<sup>[120]</sup>



**Figure 17:** Active site of dimeric EmOAH (chain A: purple, chain B: grey) with bound FAD (yellow) and oleic acid (cyan) docked into the active site. Docking was performed using YASARA while the figure was compiled using PyMol (adapted from Engleder *et al.*).<sup>[120]</sup>

The structures of LA-LAH and OhyA displayed notable differences in the loop region L<sub>98</sub>-M<sub>123</sub>, which is covering the substrate channel, indicating conformational changes upon FAD binding and a possible gating function of this region.<sup>[69,120]</sup> Engleder *et al.* also found distinctly located binding sites and orientations of the substrate, which prompted them to guess that the crystal structure of LA-LAH rather reflects the initial substrate recognition mode at the surface, while the substrate binding in the actual active site is depicted in OhyA.<sup>[69,117,120,123]</sup>

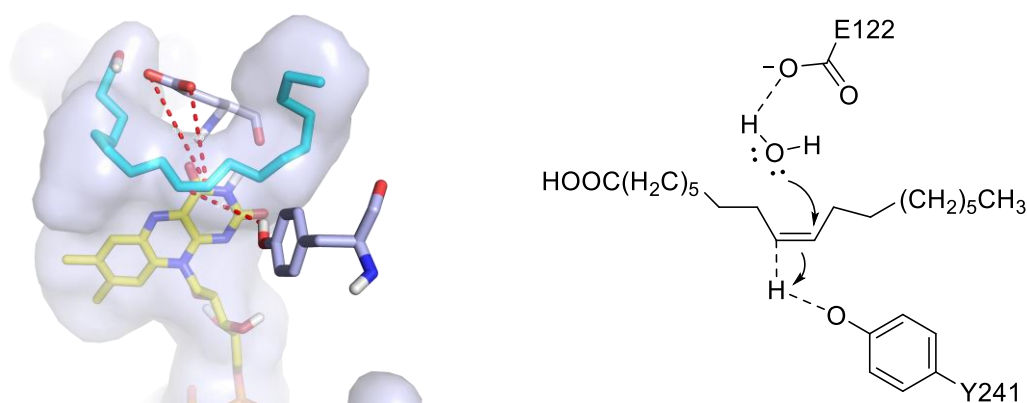
In 2018, Lorenzen *et al.* solved the crystal structure of OAH from the marine bacterium *Rhodococcus erythropolis* (OhyRe) at a resolution of 2.64 Å in the apo-state, that is, without co-crystallized FAD (PDB: 5ODO).<sup>[121]</sup> OhyRe has a sequence identity of only 35 % to LAH and 34 % to OhyA. Nevertheless, the monomeric OhyRe is made up by four domains (similar to the others) with the domains I-III being responsible for defining the FAD-binding site and the active site. Large differences were found in domain IV, as the domain in OhyRe is composed of merely one single  $\alpha$ -helix and truncated by 32 or 48 amino acids. It was speculated that the N- and C-terminal regions in OhyA and LAH are responsible for oligomerization and the lack of this regions in OhyRe prevents the formation of a dimer, as it might be the case for all members of HFam3.<sup>[121]</sup>

The crystal structure of the dimeric oleate hydratase from *Stenotrophomonas* sp. KCTC 12322 (OhySt) was recently solved by Park and coworkers at a resolution of 2.91 Å in its apo-state (PDB: 5Z70).<sup>[122]</sup> As OhySt also belongs to the HFam11, a direct comparison to the crystal structure of OhyA with FAD was possible. The comparison revealed a remarkable conformational change upon cofactor binding in the loop region surrounding the FAD acting thereby as a lid of the FAD binding pocket. This and also the observed significant change of the two key residues responsible for the hydration of oleic acid (E122 and Y241 in EmOAH) confirmed the structural role of FAD in active site arrangement enabling catalysis (1.3.5.2). However, a complete OAH structure with bound FAD and substrate is still required to unambiguously confirm the mode of substrate binding and the catalytic mechanism.<sup>[69]</sup>

### 1.3.5.2 Mechanism

Since the first report on OAHs, many more studies on kinetics, stereoselectivity and the mechanism of the reaction have been carried out.<sup>[114]</sup> Early isotope labeling experiments using deuterated and <sup>18</sup>O-labeled water confirmed the hydration mechanism, as <sup>18</sup>O was incorporated at the C10 position of 10-hydroxystearic acid.<sup>[124]</sup> The hydroxyl group is added in the (*R*)-configuration.<sup>[125]</sup> The equilibrium of this reversible reaction was also investigated and the concentration ratio was determined to be in the range of 85:15 on the side of the alcohol, while later studies for the truncated substrate (*Z*)-undec-9-enoic acid showed that it was 99:1 in favor of the hydrated product 10-hydroxyundecanoic acid.<sup>[114,126,127]</sup> A reaction via the epoxide and its subsequent reductive opening could be

ruled out as 9,10-epoxyoctadecanoic acid was not converted to 10-hydroxystearic acid. Further, the reaction with oleic acid was achieved under anaerobic conditions.<sup>[125,126]</sup> Speculations on detailed mechanisms arose together with the first crystal structures. Hence, Engleder *et al.* also proposed the first reaction mechanism for an OAH.<sup>[120]</sup> In a structure-based mutagenesis study, the active site residues E122 and Y241 were identified as essential for the activation of the water molecule and the initial protonation of the double bond, respectively. In detail, the authors suggested an acid-base mechanism for the *anti*-addition of water to oleic acid (Figure 18).<sup>[99,120]</sup> In a more or less concerted reaction, Y241 protonates the carbon-carbon double bond at the C9 atom of oleic acid leading to the formation of a carbocation intermediate at C10. The carbocation is subsequently attacked from the *re*-side by a water molecule which is activated by E122.<sup>[120]</sup>



**Figure 18:** Proposed hydration mechanism of EmOAH. *left:* Detailed view of the active site of EmOAH showing the docked substrate oleic acid (cyan) and the location of the bound FAD cofactor (yellow) at the bend of the substrate pocket. The amino acid residues E122 and Y241 were proposed to play an essential role in catalysis. *right:* Schematic depiction of the mechanism. In a more or less concerted reaction, Y241 protonates the double bond at the C9 atom while a water molecule, activated by E122, attacks the carbocation intermediate to give 10-hydroxystearic acid (adapted from Engleder *et al.*).<sup>[120]</sup>

Although the work of Engleder *et al.* gave important hints on the catalytically relevant residues and the positioning of the substrate in the active site, more details on the mechanism and the role of FAD are necessary to have clear evidence. This is even more important as the sequence alignment of OhyRe with LAH and OhyA revealed that a methionine and a valine in OhyRe replaced the catalytically relevant glutamate and a threonine with a proposed role in substrate binding indicating a distinct catalytic mechanism for oleate hydratases depending on the HFam, in this particular case for HFam3 and HFam1.<sup>[69,121]</sup> Mutagenesis studies with the variants M77E and V393T revealed a drastically reduced activity supporting this assumption.<sup>[121]</sup> Therefore, no

general statements can be made for the mechanism of oleate hydratases and detailed studies are crucial.

### 1.3.5.3 Substrate scope and selectivity

Due to their important function in the primary metabolism oleate hydratases are considered to have a high selectivity and narrow substrate specificity. However, a broad substrate scope would be desirable for use in chemical synthesis.<sup>[53]</sup> In the past years, oleate hydratases have seen a notable extension of their substrate scope. This can be tracked in the decrease of limitations discussed in the reviews by Hiseni *et al.*<sup>[128]</sup> (2015), Demming *et al.*<sup>[73]</sup> (2018) and Engleder *et al.*<sup>[69]</sup> (2018). Recently, Engleder *et al.* listed the current state of research by giving an overview about the substrate scope and the regioselectivity of so far investigated FAHs and summarized the following five essential requirements<sup>[69]</sup>:

- 1) a *cis*-carbon-carbon double bond
- 2) a carboxyl group (a free carboxylate of the fatty acid substrate)
- 3) a chain length of at least C11 of an unsaturated fatty acid<sup>[115]</sup>
- 4) a minimum of seven carbons between the carboxyl group and the to be hydrated double bond
- 5) addition of terminal OH-group is not possible<sup>[127]</sup>

However, inspired by a patent on the dehydration of short-chain alcohols by Marlière<sup>[129]</sup>, studies of our group have shown that these limitations can be overcome by using a short-chain saturated fatty acid as decoy molecule (“dummy substrate”).<sup>[127,130]</sup> Using this approach combined with rational enzyme engineering, Demming *et al.* demonstrated that oleate hydratases are capable of hydrating a broad range of nonactivated aliphatic alkenes with high regio- and stereoselectivity.<sup>[130]</sup> The study comprised overall 23 alkenes, among others, short-chain 1-alkenes (C5-C8 and C10), functionalized and internal alkenes. Besides the short-chain 1-alkenes, which were all hydrated to the corresponding 2-alcohol, 7-bromohept-1-ene, 7-octen-1-ol, 4-phenyl-1-butene, *trans*-2-octene, *cis*-2-octene were also accepted as substrates. Interestingly, while *cis*-2-octene was converted to the expected (*S*)-3-decanol, *trans*-2-octene yielded (*S*)-2-decanol. Additionally, the triple bond of 1-octyne was hydrated leading to 2-octanone as a result of keto-enol tautomerism.<sup>[130]</sup>



Concerning the hydration of substrates with multiple double bonds, like polyunsaturated fatty acids, FAHs are also strikingly regioselective. However, this is depending on the employed FAH (Table 1). Besides the large group of OAHs, which are more or less selective for the *cis*-9 double bond, several studies reported on FAHs with other selectivities. For instance, Joo *et al.* found that the OAH from *Macrococcus caseolyticus* is capable of hydrating both the *cis*-9 and the *cis*-12 carbon-carbon double bond of linoleic acid,  $\alpha$ -linoleic acid and  $\gamma$ -linoleic acid to the corresponding 10-hydroxy fatty acid or 10,13-dihydroxy fatty acids.<sup>[116]</sup> Similar results were obtained by Volkov *et al.* for the MCRA of *Streptococcus pyogenes* M49 for the hydration of linoleic acid.<sup>[119]</sup> Interestingly, several enzymes with distinct selectivities were found in *Lactobacillus acidophilus* spp. The group of Oh reported in 2015 two linoleic acid hydratases (LAH), with one selective for the hydration of C12-C13-double bonds and the other selective for hydration of the C9-C10-double bond.<sup>[131,132]</sup> In the same year, Hirata and coworkers also reported on two further FAH.<sup>[133]</sup> FA-HY2 was identified as a LAH selective for the formation of 10-hydroxy fatty acids. On the contrary, FA-HY1 converted preferable the *cis*-12 double bond when available, but was also able to hydrate the *cis*-9, *cis*-11, *cis*-13 and *cis*-14 double bond depending on the substrate. Thus, FA-HY1 with its relaxed substrate scope and broad regioselectivity is a unique exemption from the strictly selective FAHs.

**Table 1:** Regioselectivity of different FAHs.

enzyme	organism	regioselectivity	literature
OAH	<i>Macrococcus caseolyticus</i>	<i>cis</i> -9, <i>cis</i> -12	Joo <i>et al.</i> <sup>[116]</sup>
MCRA	<i>Streptococcus pyogenes</i>	<i>cis</i> -9, <i>cis</i> -12	Volkov <i>et al.</i> <sup>[119]</sup>
LHT-13	<i>Lactobacillus acidophilus</i>	<i>cis</i> -12	Oh <i>et al.</i> <sup>[131,132]</sup>
LHT-10	<i>Lactobacillus acidophilus</i>	<i>cis</i> -9	Oh <i>et al.</i> <sup>[131]</sup>
FA-HY1	<i>Lactobacillus acidophilus</i>	<i>cis</i> -12, <i>cis</i> -9, <i>cis</i> -11, <i>cis</i> -13, <i>cis</i> -14	Hirata <i>et al.</i> <sup>[133]</sup>
FA-HY2	<i>Lactobacillus acidophilus</i>	<i>cis</i> -9	Hirata <i>et al.</i> <sup>[133]</sup>

#### 1.3.5.4 Application

OAHs are promising enzymes for application in synthesis as they offer in general good stability and activity, easy expression in *E. coli* and high regio- and enantioselectivity without the need for stoichiometric amounts of cofactors.<sup>[53,69,134]</sup> Therefore, oleate hydratases are already widely used in the production of hydroxy fatty acids like 10-hydroxystearic acid<sup>[135–137]</sup> or 13-hydroxyoctadecanoic acid<sup>[132]</sup> with good volumetric productivity of  $16 \text{ g L}^{-1} \text{ h}^{-1}$ <sup>[135]</sup> and  $26.3 \text{ g L}^{-1} \text{ h}^{-1}$ <sup>[132]</sup>, respectively.<sup>[73]</sup> Additionally, they have huge potential when combined in enzyme cascades, e.g. for the synthesis of  $\omega$ -hydroxy fatty acids<sup>[138]</sup> and dicarboxylic acids<sup>[139]</sup> as monomer building blocks for plastics or even lactones ( $\gamma$ -dodecalactone<sup>[53,140]</sup>).<sup>[141]</sup>

## 1.4 Development of new biocatalysts by enzyme engineering

The identification of new enzymes can be achieved in various ways, for example by screening of microorganisms or screening of enzyme databases.<sup>[142,143]</sup> However, the enzyme found does not always meet the requirements of a given synthetic problem. Tailored biocatalysts derived by enzyme engineering from wild-type enzymes are used to circumvent this problem. Enzyme engineering can be achieved by two distinct strategies: directed evolution and rational design.<sup>[142–144]</sup>

The nobel prize 2018 in chemistry was awarded to Frances Arnold “for the directed evolution of enzymes”, which means to speed up evolution to develop new catalysts.<sup>[144,145]</sup> In her seminal work from 1993, she and her coworkers demonstrated the engineering of a subtilisin E variant, which was 256 times more active in 65 % DMF than the wild type, by successive rounds of random mutagenesis, screening and selection of improved variants.<sup>[2,21,146]</sup> Meanwhile, questions like the acceptance of non-natural substrates, the improvement of thermal or pH stability and the enantioselectivity were tackled by different research groups for the development of customized enzymes.<sup>[2,4]</sup> For the application of this approach, no information about the mechanism, the protein structure or sequence is required.<sup>[142]</sup> However, the method is limited by the large size of the created library.<sup>[142,144]</sup> Large numbers of variants have to be screened which makes an efficient (high-throughput) screening mandatory.<sup>[142,143]</sup> On the contrary, rational design requires the availability of the protein structure and detailed knowledge about the relationship between mechanism, structure and sequence.<sup>[21,142]</sup>

Small libraries are obtained by site-directed mutagenesis (introduction of point mutations) limited to a predetermined site (typically the active site) reducing the screening effort.<sup>[143]</sup> The increased use of rational design in the past years was supported by tremendous advances in DNA technologies (next-generation DNA sequencing and low-cost DNA synthesis), the development of novel tools in bioinformatics (multiple sequence alignment, homology modeling) and the significant rise of available structural data of proteins (PDB database: 147604 entries in 2018<sup>[147]</sup>).<sup>[4,142]</sup>

Nowadays, mainly a beneficial combination of both strategies called semi-rational design is applied.<sup>[143,148]</sup> Small focused libraries are created by site-saturation mutagenesis methods like ISM<sup>[149]</sup> (iterative site-saturation mutagenesis) or CAST<sup>[150]</sup> (combinatorial active-site saturation test).<sup>[143]</sup> ISM combines the rational selection of multiple targets sites with the subsequent randomization (exchange of the native amino acid against all 20 canonical amino acids) of each residue by saturation mutagenesis.<sup>[149]</sup> The residues are selected on the basis of structural and mechanistic data depending on the nature of the property to be improved. The libraries are screened and the best variant (“hit”) of each library is used as template for another round of saturation mutagenesis experiments at the respective other sites. The process is continued iteratively until sufficient improvement is achieved.<sup>[149]</sup> To further limit the screening effort in ISM to an extent that GC or HPLC analysis is still possible, smarter randomized libraries can be created by use of codons with reduced degeneracy.<sup>[151–153]</sup>

## 2 Motivation and aim of the project

In this thesis, the hydration of 5,9-dimethyldeca-4,8-dienal (DMDDAL) to Mahonial was selected for investigation as an industrial relevant case study (Figure 5). The selective hydration of such alkenes represents the key step in the synthesis of hydroxy aldehydes which are important fragrance compounds reminiscent of lily-of-the-valley. Chemical syntheses suffer from poor selectivities or can only be accomplished at the expense of additional steps for protection of other functional groups.

Hydratases exhibit outstanding regioselectivity combined with high atom economy in the hydration of C-C double bonds to form alcohols and can help to overcome these problems. Nevertheless, these enzymes are so far only rarely used in industrial biocatalysis due to their limited substrate scope and the lack of profound structural and biochemical information. No suitable enzyme for the hydration of DMDDAL or similar terpenoids with other functional groups was established at the beginning of this study.

The aim of this thesis was therefore to study the enzymatic hydration of prenyl double bonds. The discovery of a suitable regioselective hydratase and subsequent improvement of the biocatalyst by directed evolution constituted key milestones of the project.

## 3 Experimental

### 3.1 Materials

#### 3.1.1 Consumables

In the following (Table 2), the consumables which were used for this work and the suppliers thereof are listed. Pipette tips, reaction tubes and deep well plates were dry-autoclaved prior to use.

**Table 2:** Consumables and suppliers thereof.

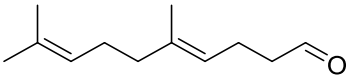
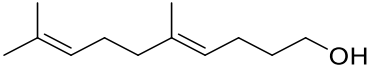
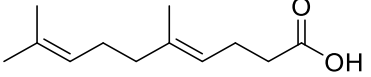
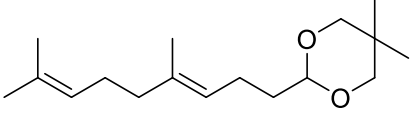
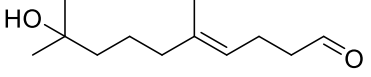
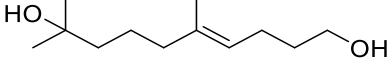
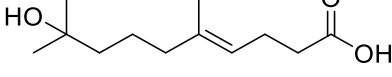
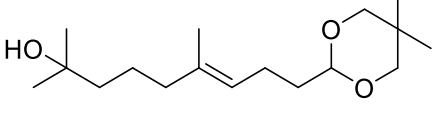
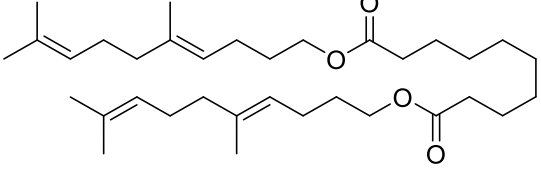
supplier	consumables
B. Braun (Melsungen, Germany)	Sterican single-use needles (different sizes)
Carl Roth (Karlsruhe, Germany)	Omnifix and Injekt single-use syringes (different sizes)
Eppendorf (Hamburg, Germany)	2.5 mL sample vial, flat bottom, PE stopper (0425.1)
GenScript (Piscataway, USA)	5000 µL pipette tips (0030.000.978)
Greiner Bio-One (Frickenhausen, Germany)	ExpressPlus PAGE gel, 10x8, 4-20 %, 12 wells (M42012)
Macherey-Nagel (Düren, Germany)	ExpressPlus PAGE gel, 10x8, 12 %, 12 wells (M01212)
Merck (Darmstadt, Germany)	Tris-MOPS-SDS running buffer powder (M00138)
Mettler Toledo - Rainin (Gießen, Germany)	Microplate 96-well, PS, F-bottom, clear (655101)
Neolab (Heidelberg, Germany)	TLC sheets ALUGRAM Xtra SIL G/UV <sub>254</sub> , 0.20 mm, 20x20cm (818333)
Ratiolab (Dreieich, Germany)	Silica 60 M, 0.04-0.063 mm for column chromatography (815381.1)
Ritter (Schwabmünchen, Germany)	MF-Millipore Membrane Filters, 13 mm, 0.025 µm (VSWP01300)
Sarstedt (Nümbrecht, Germany)	250 µL / 1200 µL multichannel pipette tips (17000506/17006324)
Sigma-Aldrich (St. Louis, USA)	Hirschmann glass inserts for 96-well plates (9240596)
VWR (Darmstadt, Germany)	Cuvets, semi-micro, PS (2712120)
WICOM (Heppenheim, Germany)	Riplate SW 96, 96 square wells, 2 mL, PP, (43001-0020)
	PCR tubes, 200 µL, single tube or strip of 8 (72.737.002/72.991.002)
	1.5 mL, 2 mL, 5 mL reactions tubes (72.690.001/72.691/72.701)
	15 mL / 50 mL centrifuge tubes (62.554.502/62.547.254)
	20 µL / 200 µL / 1000 µL pipette tips (70.1116/70.760.002/70.762)
	Breath easier sealing membrane for 96-well plates (Z763624-100EA)
	Aluma Seal II aluminum sealing foil for 96-well plates (A2350-100EA)
	Sterile syringe filters, 0.2 µm PES membrane (514-0073)
	2 mL autosampler crimp vial 12x32 mm (WIC 42000)
	300 µL micro insert (WIC 47000)
	11 mm alu crimp caps, butyl rubber/PTFE (WIC 44200)

### 3.1.2 Chemicals

Standard chemicals were purchased from Alfa-Aesar (Ward-Hill, USA), Carl Roth (Karlsruhe, Germany), Merck (Darmstadt, Germany), Peqlab (Erlangen, Germany), Sigma-Aldrich (St. Louis, USA) and VWR (Darmstadt, Germany).

Substrates for enzymatic hydration reactions and product standards (Table 3) were synthesized and provided by Givaudan (Dübendorf, Switzerland). The product standard 10-hydroxy-10-methylundecanoic acid was synthesized as described in chapter 3.8.1.

**Table 3:** Substrates and product standards synthesized by Givaudan.

structure	name or abbreviation	systematic name
	DMDDAL	5,9-dimethyldeca-4,8-dienal
	DMDDOL	5,9-dimethyldeca-4,8-dienol
	DMDDAC	5,9-dimethyldeca-4,8-dienoic acid
	DMDDIOX	2-(4,8-dimethylnona-3,7-dien-1-yl)-5,5-dimethyl-1,3-dioxane
	Mahonial HDMDAL	9-hydroxy-5,9-dimethyldec-4-enal
	HDMDOL	5,9-dimethyldec-4-ene-1,9-diol
	HDMDAC	9-hydroxy-5,9-dimethyldec-4-enoic acid
	HDMDIOX	9-(5,5-dimethyl-1,3-dioxan-2-yl)-2,6-dimethylnon-6-en-2-ol
		bis-(5,9-dimethyldeca-4,8-dien-1-yl)-decanedioate

### 3.1.3 Marker & dyes

Marker (GeneRuler 1 kb Plus DNA Ladder) and loading dye (6 x DNA loading dye) for agarose gels and marker for SDS-PAGE (PageRuler Prestained Protein Ladder) were purchased from Thermo Scientific (Waltham, USA). Midori Green Advance DNA stain for agarose gel staining was obtained from Nippon Genetics Europe (Dueren, Germany).

### 3.1.4 Enzymes & kits

In the following (Table 4), enzymes and appropriate buffers used for molecular biological methods and utilized kits are listed with the suppliers thereof.

**Table 4:** Enzymes and buffers used for molecular biological methods and utilized kits with suppliers thereof.

supplier	consumables
Agilent (Santa Clara, USA)	PfuUltra II Fusion HS DNA Polymerase
New England Biolabs (Ipswich, USA)	10 x PfuUltra buffer T5 exonuclease (10 U $\mu\text{L}^{-1}$ ) Phusion HF DNA polymerase (2 U $\mu\text{L}^{-1}$ ) Taq DNA ligase (40 U $\mu\text{L}^{-1}$ )
Pierce Biotechnology (Rockford, USA)	Pierce BCA Protein Assay Kit
Solis Biodyne (Tartu, Estonia)	5 x Hot FirePol Blend Master Mix (12.5 mM)
Thermo Scientific (Waltham, USA)	Pfu DNA polymerase (5 U $\mu\text{L}^{-1}$ ) DpnI 10 x Pfu buffer GeneJET Genomic DNA Purification Kit
Zymo Research (Irvine, USA)	DNA Clean & Concentrator Kit Zymoclean Gel DNA Recovery Kit Zyppy Plasmid Miniprep Kit

### 3.1.5 Devices

In the following (Table 5), the devices and the manufacturers thereof are listed.

**Table 5:** Devices and manufacturers thereof.

<b>manufacturer</b>	<b>device</b>
Amersham Biosciences (Amersham, UK)	Ultrospec 3100 pro
Beckman Coulter (Brea, USA)	Avanti J-26S XP
Bio-Rad (Hercules, USA)	Power Pac 300 BioRad Mini Protean Tetra Cell
BMG Labtech GmbH (Ortenberg, Germany)	PolarStar Omega 96
Branson (Danbury, USA)	Sonifier 250
Edmund Bühler GmbH (Bodelshausen, Germany)	TiMix 5 control TH30
Eppendorf (Hamburg, Germany)	Mastercycler eppgradient Thermomixer comfort Centrifuges 5415R / 5424 / 5810R
Infors HT (Basel, Switzerland)	Minitron Multitron Pro
Molecular Devices (San Jose, USA)	SpectraMax 340PC
Peqlab (Erlangen, Germany)	PerfectBlue Mini
Thermo Scientific (Waltham, USA)	NanoDrop 2000



### 3.1.6 Bacteria strains

Different *E. coli* strains were used in this work. The cells were obtained from stocks of the in-house strain collection of IBTB, department of Technical Biochemistry. DH5 $\alpha$  and XL1-blue cells were used for plasmid cloning, site-directed mutagenesis and storage. Protein production was conducted for pET-28a(+) vectors (3.1.8) in the expression strain *E. coli* BL21 (DE3). Additionally, the ADH deficient strain *E. coli* ITB 94 was used for pDHE vectors (3.1.8). This strain is based on the commercially available TG1 strain. Moreover, the strain is lacking L-rhamnose-isomerase which was replaced by tetracycline resistance gene as marker. Furthermore, two genes encoding for unspecific alcohol dehydrogenases were deleted ( $\Delta$ *yahK* and  $\Delta$ *yjgB*) while the two chaperone plasmids pHSG (chloramphenicol resistance) and pAgro4 (spectinomycin resistance) were added.

*Geobacillus stearothermophilus* BR388 was obtained as patent deposit from ATCC / LGC Standards, was used to isolate gDNA (3.2.1) and for test of biotransformations of limonene, DMDDAL and its derivatives.<sup>[154]</sup>

### 3.1.7 Primers

All primers (Table 6-Table 12) were purchased from Metabion (Planegg, Germany). The primers for Gibson Assembly were designed using the Gibson Assembly tool of SnapGene (Version 3.1.4., GSL Biotech, Chicago, USA). Primers for site-directed mutagenesis were designed following the guidelines of the QuikChange II Site-Directed Mutagenesis Kit (Agilent, Santa Clara, USA). For redundant codons, the most abundant codon in *E. coli* K12 was used. Exchanged codons are underlined. Melting temperatures of primers were calculated using SnapGene and the Oligocalc online tool (<http://biotools.nubic.northwestern.edu/OligoCalc.html>).

**Table 6:** List of primers used for DNA Sanger sequencing.

<b>name</b>	<b>sequence of primer (5'→3')</b>
FITS <sup>[155]</sup>	GGG GAA GCG CCG CGT TCG G
RITS <sup>[155]</sup>	GTG CAA GCA CCC TTG CAG GCG AAG A
F8 <sup>[156]</sup>	AGA GTT TGA TCC TGG CTC AG
1492R <sup>[156]</sup>	CGG TTA CCT TGT TAC GAC TT
JSC_EmOAH_Seq_mid	GCA ATT ATT CGC CTG CTG C
JSC_pDHE_Seq_1	AGT ATT CAA CAT TTC CGT GTC G
JSC_pDHE_Seq_2	TGG ATG AAC GAA ATA GAC AGA TCG
JSC_pDHE_Seq_3	CCC TGA TTC TGT GGA TAA CCG
JSC_pDHE_Seq_4	TTT GCA GCA GCA GTC GCT TCA C
JSC_pDHE_Seq_5	GAA GAT CAA CTG GAA AAC GTG
pGL3-1 (GATC in-house)	GAG CTG ACT GGG TTG AAG
pTI2-1-2 (GATC in-house)	GCG TTT CAC TTC TGA GTT CG
T7 (GATC in-house)	TAA TAC GAC TCA CTA TAG GG
pET-RP (GATC in-house)	CTA GTT ATT GCT CAG CGG

**Table 7:** List of primers used for Gibson Assembly for the construction of pDHE vectors containing the MCS (empty vector), CrtC and EmOAH.

<b>name</b>	<b>sequence of primer (5'→3')</b>
JSC_TrCrtC_fw	ATG CGT GCA GCA GGT ATT CTG ACA CCG GG
JSC_TrCrtC-His rev	TCA GTG GTG GTG GTG GTG CTC GAG TGC ACC ACC AAC ACG TGG CAT ACG
JSC_pDHE_TrCrtC_H_vf	ACC ACC ACC ACC ACC ACT GAG GGT ACC GGC TG
JSC_pDHE_TrCrtC_H_vr	AGA ATA CCT GCT GCA CGC ATA TGT ATA TCT CCT TCT TAA GAA TTG TTC ATT ACG ACC AGT CTA AAA AGC G
JSC_pDHE_TrCrtC_H_ff	CTT AAG AAG GAG ATA TAC ATA TGC GTG CAG CAG GTA TTC TGA CAC C
JSC_pDHE_TrCrtC_H_fr	CGC CAA AAC AGC CGG TAC CCT CAG TGG TGG TGG TGG TGG TGC T
JSC_pDHE_RgCrtC_vf	GTA TGC CTC GTC GTA GCT AAG GGT ACC GGC TGT TTT GGC G
JSC_pDHE_RgCrtC_vr	TGA TGA TGG CTG CTG CCC ATA TGT ATA TCT CCT TCT TAA GAA TTG TTC ATT ACG ACC AGT CTA AAA AGC G
JSC_pDHE_RgCrtC_ff	CTT AAG AAG GAG ATA TAC ATA TGG GCA GCA GCC ATC ATC ATC ATC ATC ACA GC
JSC_pDHE_RgCrtC_fr	CGC CAA AAC AGC CGG TAC CCT TAG CTA CGA CGA GGC ATA CGC C
JSC_pDHE_EmOAH_vf	CAC CAC CAC CAC CAC CAC TGA GGG TAC CCG ATT AG
JSC_pDHE_EmOAH_vr	CAA ATT TGC TGG TAA TCG GAT TCA TAT GTA TAT CTC CTT CTT AAG AAT TGT TCA TTA CGA CC
JSC_pDHE_EmOAH_ff	CTT AAG AAG GAG ATA TAC ATA TGA ATC CGA TTA CCA GCA AAT TTG ATA AAG TGC TGA ATG
JSC_pDHE_EmOAH_fr	GAA TCC CTA ATC GGG TAC CCT CAG TGG TGG TGG TGG TG
pDHE_MCS_vf	ACC ACC ACC ACC ACC ACT GAG GGT ACC GGC TGT TTT GG
pDHE_MCS_vr	TGA TGA TGG CTG CTG CCC ATA TGT ATA TCT CC
pDHE_MCS_ff	CTT AAG AAG GAG ATA TAC ATA TGG GCA GCA GCC ATC ATC A
pDHE_MCS_fr	CGC CAA AAC AGC CGG TAC CCT CAG TGG TGG TGG TGG TGG TGC

**Table 8:** List of primers for site-directed mutagenesis of EmOAH wild type. Several positions in the active site were mutated for the initial library.

mutation	sequence of	forward primer (5'→3') reverse primer (5'→3')
Y127S	GAA ATG GAT ATG ACC <u>AGC</u> GAA AAT CTG TGG GAC GTC CCA CAG ATT TTC <u>GCT</u> GGT CAT ATC CAT TTC	
Y127C	GAA ATG GAT ATG ACC <u>TGC</u> GAA AAT CTG TGG GAC GTC CCA CAG ATT TTC <u>GCA</u> GGT CAT ATC CAT TTC	
Y127A	GAA ATG GAT ATG ACC <u>GCG</u> GAA AAT CTG TGG GAC GTC CCA CAG ATT TTC <u>CGC</u> GGT CAT ATC CAT TTC	
Y127W	GAA ATG GAT ATG ACC <u>TGG</u> GAA AAT CTG TGG GAC GTC CCA CAG ATT TTC <u>CCA</u> GGT CAT ATC CAT TTC	
F227W <sup>[157]</sup>	CC TTT TGG CGT ACC ATG <u>TGG</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>CCA</u> CAT GGT ACG CCA AAA GG	
F227A	C TTT TGG CGT ACC ATG <u>GCG</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>CGC</u> CAT GGT ACG CCA AAA G	
A228F	GG CGT ACC ATG TTT <u>TTT</u> TTT GAA AAT TGG C G CCA ATT TTC AAA <u>AAA</u> AAA CAT GGT ACG CC	
A228W	GG CGT ACC ATG TTT <u>TGG</u> TTT GAA AAT TGG C G CCA ATT TTC AAA <u>CCA</u> AAA CAT GGT ACG CC	
Y241A	CTG GAA CTG AAA CTG <u>GCG</u> ATG CAT CGT TTT CTG CAG AAA ACG ATG CAT <u>CGC</u> CAG TTT CAG TTC CAG	
F260W	CTG AGC AGC CTG GTT <u>TGG</u> CCG AAA TAC AAC GTT GTA TTT CGG <u>CCA</u> AAC CAG GCT GCT CAG	
Y456A	CTG GTT CTG TGG GTT <u>GCG</u> GCA CTG TTT ATG CAT AAA CAG TGC <u>CGC</u> AAC CCA CAG AAC CAG	
F514A	CCG TAT ATT ACC TCA ATG <u>GCG</u> ATG CCT CGT GCC GGC ACG AGG CAT <u>CGC</u> CAT TGA GGT AAT ATA CGG	
F548A	C AAT AAT GAT GTG GTG <u>GCG</u> ACA ATG GAA AGC AG CT GCT TTC CAT TGT <u>CGC</u> CAC CAC ATC ATT ATT G	
F548W	C AAT AAT GAT GTG GTG <u>TGG</u> ACA ATG GAA AGC AG CT GCT TTC CAT TGT <u>CCA</u> CAC CAC ATC ATT ATT G	

**Table 9:** List of primers for site-directed saturation mutagenesis of EmOAH at position A248 (designed and provided by R. Demming).<sup>[157]</sup>

mutation	sequence of forward primer (5'→3') reverse primer (5'→3')
A248C	G CAT CGT TTT CTG CAT <u>TGC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GCA</u> ATG CAG AAA ACG ATG C
A248D	G CAT CGT TTT CTG CAT <u>GAC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GTC</u> ATG CAG AAA ACG ATG C
A248E	G CAT CGT TTT CTG CAT <u>GAA</u> ATT GAT GGT CTG AAT G C ATT CAG ACC ATC AAT <u>TTC</u> ATG CAG AAA ACG ATG C
A248F	G CAT CGT TTT CTG CAT <u>TTC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GAA</u> ATG CAG AAA ACG ATG C
A248G	G CAT CGT TTT CTG CAT <u>GGC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GCC</u> ATG CAG AAA ACG ATG C
A248H	G CAT CGT TTT CTG CAT <u>CAC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GTG</u> ATG CAG AAA ACG ATG C
A248I	G CAT CGT TTT CTG CAT <u>ATC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GAT</u> ATG CAG AAA ACG ATG C
A248K	G CAT CGT TTT CTG CAT <u>AAA</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>TTT</u> ATG CAG AAA ACG ATG C
A248L	G CAT CGT TTT CTG CAT <u>CTG</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>CAG</u> ATG CAG AAA ACG ATG C
A248M	G CAT CGT TTT CTG CAT <u>ATG</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>CAT</u> ATG CAG AAA ACG ATG C
A248N	G CAT CGT TTT CTG CAT <u>AAC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GTT</u> ATG CAG AAA ACG ATG C
A248P	G CAT CGT TTT CTG CAT <u>CCG</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>CGG</u> ATG CAG AAA ACG ATG C
A248Q	G CAT CGT TTT CTG CAT <u>CAG</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>CTG</u> ATG CAG AAA ACG ATG C
A248R	G CAT CGT TTT CTG CAT <u>CGC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GCG</u> ATG CAG AAA ACG ATG C
A248S	G CAT CGT TTT CTG CAT <u>TCC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GGA</u> ATG CAG AAA ACG ATG C
A248T	G CAT CGT TTT CTG CAT <u>ACC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GGT</u> ATG CAG AAA ACG ATG C
A248V	G CAT CGT TTT CTG CAT <u>GTG</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>CAC</u> ATG CAG AAA ACG ATG C
A248W	G CAT CGT TTT CTG CAT <u>TGG</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>CCA</u> ATG CAG AAA ACG ATG C
A248Y	G CAT CGT TTT CTG CAT <u>TAC</u> ATT GAT GGT CTG AAT GAT C G ATC ATT CAG ACC ATC AAT <u>GTA</u> ATG CAG AAA ACG ATG C

**Table 10:** List of primers for site-directed saturation mutagenesis of EmOAH at position T549.

mutation	sequence of	forward primer (5'→3')	reverse primer (5'→3')
T549A	GAT GTG GTG TTC	<u>GCG</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>CGC</u> GAA CAC CAC ATC
T549C	GAT GTG GTG TTC	<u>TGC</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>GCA</u> GAA CAC CAC ATC
T549D	GAT GTG GTG TTC	<u>GAT</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>ATC</u> GAA CAC CAC ATC
T549E	GAT GTG GTG TTC	<u>GAA</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>TTC</u> GAA CAC CAC ATC
T549F	GAT GTG GTG TTC	<u>TTT</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>AAA</u> GAA CAC CAC ATC
T549G	GAT GTG GTG TTC	<u>GGC</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>GCC</u> GAA CAC CAC ATC
T549H	GAT GTG GTG TTC	<u>CAT</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>ATG</u> GAA CAC CAC ATC
T549I	GAT GTG GTG TTC	<u>ATT</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>AAC</u> GAA CAC CAC ATC
T549K	GAT GTG GTG TTC	<u>AAA</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>TTT</u> GAA CAC CAC ATC
T549L	GAT GTG GTG TTC	<u>CTG</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>CAG</u> GAA CAC CAC ATC
T549M	GAT GTG GTG TTC	<u>ATG</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>CAT</u> GAA CAC CAC ATC
T549N	GAT GTG GTG TTC	<u>AAC</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>GTT</u> GAA CAC CAC ATC
T549P	GAT GTG GTG TTC	<u>CCG</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>CGG</u> GAA CAC CAC ATC
T549Q	GAT GTG GTG TTC	<u>CAG</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>CTG</u> GAA CAC CAC ATC
T549R	GAT GTG GTG TTC	<u>CGC</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>GCG</u> GAA CAC CAC ATC
T549S	GAT GTG GTG TTC	<u>AGC</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>GCT</u> GAA CAC CAC ATC
T549V	GAT GTG GTG TTC	<u>GTG</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>CAC</u> GAA CAC CAC ATC
T549W	GAT GTG GTG TTC	<u>TGG</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>CCA</u> GAA CAC CAC ATC
T549Y	GAT GTG GTG TTC	<u>TAT</u> ATG GAA AGC AGC GTG	CAC GCT GCT TTC CAT <u>ATA</u> GAA CAC CAC ATC

**Table 11:** List of primers for site-directed saturation mutagenesis of EmOAH at position W389.

mutation	sequence of	forward primer (5'→3')	reverse primer (5'→3')
W389A	GAA AAA AGC GCA	<u>GCG</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389C	GAA AAA AGC GCA	<u>TGC</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389D	GAA AAA AGC GCA	<u>GAT</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389E	GAA AAA AGC GCA	<u>GAA</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389F	GAA AAA AGC GCA	<u>TTT</u> GAA AGC GCA ACC CTG	CAG GGT TGC GCT TTC
W389G	GAA AAA AGC GCA	<u>GGC</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389H	GAA AAA AGC GCA	<u>CAT</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389I	GAA AAA AGC GCA	<u>ATT</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389K	GAA AAA AGC GCA	<u>AAA</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389L	GAA AAA AGC GCA	<u>CTG</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389M	GAA AAA AGC GCA	<u>ATG</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389N	GAA AAA AGC GCA	<u>AAC</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389P	GAA AAA AGC GCA	<u>CCG</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389Q	GAA AAA AGC GCA	<u>CAG</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389R	GAA AAA AGC GCA	<u>CGC</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389S	GAA AAA AGC GCA	<u>AGC</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389T	GAA AAA AGC GCA	<u>ACC</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389V	GAA AAA AGC GCA	<u>GTG</u> GAA AGC GCA ACC	GGT TGC GCT TTC
W389Y	GAA AAA AGC GCA	<u>TAT</u> GAA AGC GCA ACC	GGT TGC GCT TTC

**Table 12:** List of primers for semi-rational site-directed (saturation) mutagenesis of EmOAH at positions R121, M226/F227, F514 and F548/T549 in a CAST-like approach using degenerated codons. X stands for one of the other 19 amino acids that could replace the residue in the parent. The degenerated codons have the following meaning: NNK = all 20 amino acids; NDT = 12 amino acids: R, N, D, C, G, H, I, L, F, S, Y, V; VHG = 9 amino acids: A, Q, E, L, K, M, P, T, V; TGG = W; NYK = 9 amino acids: A, I, L, M, F, P, S, T, V; TAT = Y.

mutation	degenerated codons	sequence of forward primer (5'→3') reverse primer (5'→3')
R121X	NNK	C ATT CGT GGT GGT <u>NNK</u> GAA ATG GAT ATG ACC GGT CAT ATC CAT TTC <u>MNN</u> ACC ACC ACG AAT G
M226X_F227X	NDT, NDT	C TTT TGG CGT ACC <u>NDT</u> <u>NDT</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>AHN</u> <u>AHN</u> GGT ACG CCA AAA G
	NDT, VHG	C TTT TGG CGT ACC <u>NDT</u> <u>VHG</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>CDB</u> <u>AHN</u> GGT ACG CCA AAA G
	NDT, TGG	C TTT TGG CGT ACC <u>NDT</u> <u>TGG</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>CCA</u> <u>AHN</u> GGT ACG CCA AAA G
	VHG, NDT	C TTT TGG CGT ACC <u>VHG</u> <u>NDT</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>AHN</u> <u>CDB</u> GGT ACG CCA AAA G
	VHG, VHG	C TTT TGG CGT ACC <u>VHG</u> <u>VHG</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>CDB</u> <u>CDB</u> GGT ACG CCA AAA G
	VHG, TGG	C TTT TGG CGT ACC <u>VHG</u> <u>TGG</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>CCA</u> <u>CDB</u> GGT ACG CCA AAA G
	TGG, NDT	C TTT TGG CGT ACC <u>TGG</u> <u>NDT</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>AHN</u> <u>CCA</u> GGT ACG CCA AAA G
	TGG, VHG	C TTT TGG CGT ACC <u>TGG</u> <u>VHG</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>CDB</u> <u>CCA</u> GGT ACG CCA AAA G
	TGG, TGG	C TTT TGG CGT ACC <u>TGG</u> <u>TGG</u> GCC TTT GAA AAT TG CA ATT TTC AAA GGC <u>CCA</u> <u>CCA</u> GGT ACG CCA AAA G
F514X	NYK	CCG TAT ATT ACC TCA ATG <u>NYK</u> ATG CCT CGT GCC GGC ACG AGG CAT <u>MRN</u> CAT TGA GGT AAT ATA CGG
	TAT	CCG TAT ATT ACC TCA ATG <u>TAT</u> ATG CCT CGT GCC GGC ACG AGG CAT <u>ATA</u> CAT TGA GGT AAT ATA CGG
	TGG	CCG TAT ATT ACC TCA ATG <u>TGG</u> ATG CCT CGT GCC GGC ACG AGG CAT <u>CCA</u> CAT TGA GGT AAT ATA CGG
F548X_T549X	NYK, NYK	GAT GTG GTG <u>NYK</u> <u>NYK</u> ATG GAA AGC AGC GTG CAC GCT GCT TTC CAT <u>MRN</u> <u>MRN</u> CAC CAC ATC
	NYK, TAT	GAT GTG GTG <u>NYK</u> <u>TAT</u> ATG GAA AGC AGC GTG CAC GCT GCT TTC CAT <u>ATA</u> <u>MRN</u> CAC CAC ATC
	NYK, TGG	GAT GTG GTG <u>NYK</u> <u>TGG</u> ATG GAA AGC AGC GTG CAC GCT GCT TTC CAT <u>CCA</u> <u>MRN</u> CAC CAC ATC
	TAT, NYK	GAT GTG GTG <u>TAT</u> <u>NYK</u> ATG GAA AGC AGC GTG CAC GCT GCT TTC CAT <u>MRN</u> <u>ATA</u> CAC CAC ATC
	TAT, TAT	GAT GTG GTG <u>TAT</u> <u>TAT</u> ATG GAA AGC AGC GTG CAC GCT GCT TTC CAT <u>ATA</u> <u>ATA</u> CAC CAC ATC
	TAT, TGG	GAT GTG GTG <u>TAT</u> <u>TGG</u> ATG GAA AGC AGC GTG CAC GCT GCT TTC CAT <u>CCA</u> <u>ATA</u> CAC CAC ATC
	TGG, NYK	GAT GTG GTG <u>TGG</u> <u>NYK</u> ATG GAA AGC AGC GTG CAC GCT GCT TTC CAT <u>MRN</u> <u>CCA</u> CAC CAC ATC
	TGG, TAT	GAT GTG GTG <u>TGG</u> <u>TAT</u> ATG GAA AGC AGC GTG CAC GCT GCT TTC CAT <u>ATA</u> <u>CCA</u> CAC CAC ATC
	TGG, TGG	GAT GTG GTG <u>TGG</u> <u>TGG</u> ATG GAA AGC AGC GTG CAC GCT GCT TTC CAT <u>CCA</u> <u>CCA</u> CAC CAC ATC

### 3.1.8 Plasmids

In this work, both pDHE (Table 13) and pET28a(+) vectors (Table 14) were used. The vector maps of pDHE and pET28a(+) vectors containing genes of the wild type enzymes EmOAH, RgCrtC and TrCrtC, respectively, can be found in the appendix (7.2). pET28a(+) vectors feature a kanamycine (KanR) resistance for selection, a *lac* operon, a T7 promotor and gene expression can be induced by IPTG. pDHE vectors feature an ampicillin resistance (AmpR) for selection, a rhamnose promotor (*rhaB*), a basis of mobility (*bom*) region for horizontal gene transfer and a repressor of primer (*rop*) site to keep the copy number in *E. coli* low.<sup>[158]</sup> Plasmids were stored isolated at -20 °C and in *E. coli* cells as glycerol stocks at -80 °C.

**Table 13:** List of pDHE plasmids and their source.

ITB number	insert	variant	source
pITB5136	MCS of pET28a(+)	-	this work
pITB5137	RgCrtC	wild type	this work
pITB5138	TrCrtC	wild type	this work
pITB5139	EmOAH	wild type	this work
pITB5140	EmOAH	Y127A	this work
pITB5141	EmOAH	Y127C	this work
pITB5142	EmOAH	Y127S	this work
pITB5143	EmOAH	Y127W	this work
pITB5144	EmOAH	F227A	this work
pITB5145	EmOAH	F227W	this work
pITB5146	EmOAH	A228F	this work
pITB5147	EmOAH	A228W	this work
pITB5148	EmOAH	Y241A	this work
pITB5149	EmOAH	F260W	this work
pITB5150	EmOAH	W389A	this work
pITB5151	EmOAH	W389F	this work
pITB5152	EmOAH	W389L	this work
pITB5153	EmOAH	Y456A	this work
pITB5154	EmOAH	F514A	this work
pITB5155	EmOAH	F548A	this work
pITB5156	EmOAH	F548W	this work
pITB5157	EmOAH	A248C	this work
pITB5158	EmOAH	A248D	this work
pITB5159	EmOAH	A248E	this work
pITB5160	EmOAH	A248F	this work
pITB5161	EmOAH	A248G	this work
pITB5162	EmOAH	A248H	this work
pITB5163	EmOAH	A248I	this work
pITB5164	EmOAH	A248K	this work
pITB5165	EmOAH	A248L	this work
pITB5166	EmOAH	A248M	this work
pITB5167	EmOAH	A248N	this work
pITB5168	EmOAH	A248P	this work



---

pITB5169	EmOAH	A248Q	this work
pITB5170	EmOAH	A248R	this work
pITB5171	EmOAH	A248S	this work
pITB5172	EmOAH	A248T	this work
pITB5173	EmOAH	A248V	this work
pITB5174	EmOAH	A248W	this work
pITB5175	EmOAH	A248Y	this work
<hr/>			
pITB5176	EmOAH	T549A	this work
pITB5177	EmOAH	T549C	this work
pITB5178	EmOAH	T549D	this work
pITB5179	EmOAH	T549E	this work
pITB5180	EmOAH	T549F	this work
pITB5181	EmOAH	T549G	this work
pITB5182	EmOAH	T549H	this work
pITB5183	EmOAH	T549I	this work
pITB5184	EmOAH	T549K	this work
pITB5185	EmOAH	T549L	this work
pITB5186	EmOAH	T549M	this work
pITB5187	EmOAH	T549N	this work
pITB5188	EmOAH	T549P	this work
pITB5189	EmOAH	T549Q	this work
pITB5190	EmOAH	T549R	this work
pITB5191	EmOAH	T549S	this work
pITB5192	EmOAH	T549V	this work
pITB5193	EmOAH	T549W	this work
pITB5194	EmOAH	T549Y	this work
<hr/>			
pITB5195	EmOAH	A248L/T549A	this work
pITB5196	EmOAH	A248L/T549C	this work
pITB5197	EmOAH	A248L/T549D	this work
pITB5198	EmOAH	A248L/T549E	this work
pITB5199	EmOAH	A248L/T549F	this work
pITB5200	EmOAH	A248L/T549G	this work
pITB5201	EmOAH	A248L/T549H	this work
pITB5202	EmOAH	A248L/T549I	this work
pITB5203	EmOAH	A248L/T549K	this work
pITB5204	EmOAH	A248L/T549L	this work
pITB5205	EmOAH	A248L/T549M	this work
pITB5206	EmOAH	A248L/T549N	this work
pITB5207	EmOAH	A248L/T549P	this work
pITB5208	EmOAH	A248L/T549Q	this work
pITB5209	EmOAH	A248L/T549R	this work
pITB5210	EmOAH	A248L/T549S	this work
pITB5211	EmOAH	A248L/T549V	this work
pITB5212	EmOAH	A248L/T549W	this work
pITB5213	EmOAH	A248L/T549Y	this work
<hr/>			
pITB5214	EmOAH	A248C/T549S	this work
pITB5215	EmOAH	A248D/T549S	this work
pITB5216	EmOAH	A248E/T549S	this work
pITB5217	EmOAH	A248F/T549S	this work
pITB5218	EmOAH	A248G/T549S	this work
pITB5219	EmOAH	A248H/T549S	this work
pITB5220	EmOAH	A248I/T549S	this work
pITB5221	EmOAH	A248K/T549S	this work

---

pITB5222	EmOAH	A248M/T549S	this work
pITB5223	EmOAH	A248N/T549S	this work
pITB5224	EmOAH	A248P/T549S	this work
pITB5225	EmOAH	A248Q/T549S	this work
pITB5226	EmOAH	A248R/T549S	this work
pITB5227	EmOAH	A248S/T549S	this work
pITB5228	EmOAH	A248T/T549S	this work
pITB5229	EmOAH	A248V/T549S	this work
pITB5230	EmOAH	A248W/T549S	this work
pITB5231	EmOAH	A248Y/T549S	this work
pITB5232	EmOAH	A248I/T549S/W389A	this work
pITB5233	EmOAH	A248I/T549S/W389C	this work
pITB5234	EmOAH	A248I/T549S/W389D	this work
pITB5235	EmOAH	A248I/T549S/W389E	this work
pITB5236	EmOAH	A248I/T549S/W389F	this work
pITB5237	EmOAH	A248I/T549S/W389G	this work
pITB5238	EmOAH	A248I/T549S/W389H	this work
pITB5239	EmOAH	A248I/T549S/W389I	this work
pITB5240	EmOAH	A248I/T549S/W389K	this work
pITB5241	EmOAH	A248I/T549S/W389L	this work
pITB5242	EmOAH	A248I/T549S/W389M	this work
pITB5243	EmOAH	A248I/T549S/W389N	this work
pITB5244	EmOAH	A248I/T549S/W389P	this work
pITB5245	EmOAH	A248I/T549S/W389Q	this work
pITB5246	EmOAH	A248I/T549S/W389R	this work
pITB5247	EmOAH	A248I/T549S/W389S	this work
pITB5248	EmOAH	A248I/T549S/W389T	this work
pITB5249	EmOAH	A248I/T549S/W389V	this work
pITB5250	EmOAH	A248I/T549S/W389Y	this work

**Table 14:** List of pET28a(+) plasmids and their source.

ITB number	insert	variant	source
pITB1213	CgOAH	wild type	Schmid <i>et al.</i> <sup>[115]</sup>
pITB1217	DbOAH	wild type	Schmid <i>et al.</i> <sup>[115]</sup>
pITB1211	GmOAH	wild type	Schmid <i>et al.</i> <sup>[115]</sup>
pITB1216	HfOAH	wild type	L. Steiner <sup>[159]</sup>
pITB1210	LaOAH	wild type	Schmid <i>et al.</i> <sup>[115]</sup>
pITB1214	LjOAH	wild type	Schmid <i>et al.</i> <sup>[115]</sup>
pITB1215	MpOAH	wild type	Schmid <i>et al.</i> <sup>[115]</sup>
pITB868	EmOAH	wild type	Otte <i>et al.</i> <sup>[139]</sup>
	RgCrtC	wild type	M. Fischer (PhD thesis)
	TrCrtC	wild type	M. Fischer (PhD thesis)
pITB1890	EmOAH	A248L	R. Demming <sup>[157]</sup>
pITB1885	EmOAH	F227Y	R. Demming <sup>[157]</sup>
pITB1884	EmOAH	F227W	R. Demming <sup>[157]</sup>
pITB5251	EmOAH	A248I/T549S	this work
pITB5252	EmOAH	A248L/T549S	this work
pITB5253	EmOAH	A248M/T549S	this work
pITB5254	EmOAH	A248T/T549S	this work
pITB5255	EmOAH	A248Y/T549S	this work

### 3.1.9 Buffers, media & antibiotics

Buffers and media were autoclaved at 121 °C or filtered through 0.2 µm syringe filters prior to use for sterilization if necessary. The pH was adjusted with 1 M NaOH or 1 M HCl unless otherwise stated. Agar plates were prepared by adding 1.5 % agar-agar to the medium. Antibiotics were added to media as selection agent in different final concentrations (kanamycin: 30 µg mL<sup>-1</sup>, ampicillin: 100 µg mL<sup>-1</sup>, spectinomycin: 50 µg mL<sup>-1</sup>, chloramphenicol: 5 µg mL<sup>-1</sup> and tetracyclin: 15 µg mL<sup>-1</sup>). Stocks of antibiotics were prepared by dissolving the antibiotic in the following concentration using the appropriate solvent (kanamycin: 30 mg mL<sup>-1</sup> in ddH<sub>2</sub>O, ampicillin: 100 mg mL<sup>-1</sup> in ddH<sub>2</sub>O, spectinomycin: 50 mg mL<sup>-1</sup> in ddH<sub>2</sub>O, chloramphenicol: 5 mg mL<sup>-1</sup> in EtOH and tetracyclin: 5 mg mL<sup>-1</sup> in 70 % EtOH). Given percentages are specified in  $V V^{-1}$  for liquids and in  $m V^{-1}$  for solids.

#### Potassium phosphate buffer (50 mM)

KH <sub>2</sub> PO <sub>4</sub>	3.40 g
K <sub>2</sub> HPO <sub>4</sub>	4.36 g
ddH <sub>2</sub> O	ad 1 L

#### Citrate buffer (50 mM, pH 6.0)

citric acid	9.61 g
ddH <sub>2</sub> O	ad 1 L
adjust pH with conc. NaOH	

#### TfbI buffer (pH 5.8)

KOAc	0.59 g
RbCl	2.42 g
CaCl <sub>2</sub>	0.29 g
MnCl <sub>2</sub> · 4 H <sub>2</sub> O	2.0 g
glycerol	30 mL
ddH <sub>2</sub> O	ad 200 mL
adjust pH with acetic acid	

#### TfbII buffer (pH 6.5)

MOPS	0.21 g
RbCl	0.12 g
CaCl <sub>2</sub>	1.10 g
glycerol	15 mL
ddH <sub>2</sub> O	ad 100 mL

**TAE buffer (50 x, pH 8.3)**

TRIS	242 g
acetic acid	57 mL
EDTA	18.6 g
ddH <sub>2</sub> O	ad 1 L

**10 x TB buffer**

KH <sub>2</sub> PO <sub>4</sub>	23.1 g
K <sub>2</sub> HPO <sub>4</sub>	125.4 g
ddH <sub>2</sub> O	ad 1 L

**LB medium (pH 7.0)**

tryptone	10 g	1.0 %
yeast extract	5 g	0.5 %
NaCl	5 g	0.5 %
ddH <sub>2</sub> O	ad 1 L	

**TB medium (pH 7.2)**

tryptone	12 g	1.2 %
yeast extract	24 g	2.4 %
glycerol	4 mL	0.4 %
ddH <sub>2</sub> O	ad 900 mL	

adjust pH with 100 mL sterile 10 x TB buffer (0.17 M KH<sub>2</sub>PO<sub>4</sub>, 0.72 M K<sub>2</sub>HPO<sub>4</sub>) added after autoclaving to cold medium prior to use

**Nutrient broth**

peptone (from meat)	5 g	0.5 %
meat extract	3 g	0.3 %
ddH <sub>2</sub> O	ad 1 L	

**DP salt medium (pH 7.2)<sup>[160]</sup>**

NH <sub>4</sub> Cl	1 g	0.1 %
K <sub>2</sub> HPO <sub>4</sub>	0.5 g	0.05 %
MgSO <sub>4</sub> · 7 H <sub>2</sub> O	20 mg	0.002 %
ddH <sub>2</sub> O	ad 1 L	

MgSO<sub>4</sub> was added as 1000 x stock for easier handling. A vitamin complex (as 1000 x) can also be added to the minimal media (for 10 mL stock solution: 4 mg nicotinamide, 4 mg thiamine hydrochloride, 20 µg biotin).

**M9 salts (5 x)**

NH <sub>4</sub> Cl	6 g	0.6 %
K <sub>2</sub> HPO <sub>4</sub>	15 g	1.5 %
Na <sub>2</sub> HPO <sub>4</sub>	33.9 g	3.39 %
NaCl	2.5 g	0.25 %
ddH <sub>2</sub> O	ad 1 L	

For preparation of M9 medium, 800 mL ddH<sub>2</sub>O was added to 200 mL of the M9 salt solution. Additionally, 2 mL 1 M MgSO<sub>4</sub> and 100 µL 1 M CaCl<sub>2</sub> was supplemented.

**Tryptone soya broth (TSB) (pH 7.3)**

casein hydrolysate	17 g	1.7 %
peptone from soybean, enzymatic digest	3 g	0.3 %
NaCl	5 g	0.5 %
K <sub>2</sub> HPO <sub>4</sub>	2.5 g	0.25 %
ddH <sub>2</sub> O	ad 900 L	

TSB (DSMZ Medium 545) is a rich complex medium for fastidious microorganisms. pH was adjusted to 7.3 if necessary. D-(+)-glucose (2.5 g L<sup>-1</sup>) was added as 100 mL sterile 10 x stock solution after autoclaving.

## 3.2 Molecular biological methods

### 3.2.1 Isolation of genomic DNA

Genomic DNA (gDNA) was isolated from *Geobacillus stearothermophilus* BR388 after 24 h of cultivation at 55 °C without shaking in 200 mL TSB medium. The cells were harvested in three parallel batches by centrifugation (4 °C, 3220 x g, 30 min). Subsequently, the supernatant was discarded and the cells were centrifuged again for 15 min to ensure complete removal of medium.

The isolation of gDNA was conducted using the GeneJET Genomic DNA Purification Kit following manufacturer's purification protocol for Gram-positive bacteria likewise in three parallel batches. All resuspending and mixing steps were conducted by pipetting the sample carefully and incubation steps were performed in Thermomixer. In a first step, the cells were resuspended in 180 µL gram-positive bacteria lysis buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 1.2 % Triton X-100, add directly before use 20 mg mL<sup>-1</sup> lysozyme). The cell suspension was transferred to a 2 mL microreaction tube and incubated (37 °C, 30 min, 0 rpm). Then, 200 µL Lysis solution and 20 µL Proteinase K were added and everything was mixed until a uniform suspension was obtained. Subsequently, the cells were incubated (56 °C, 60 min, 300 rpm) while the sample was additionally mixed through manually every 5 min. After complete lysis of cells was reached, 20 µL RNase A solution was added and the sample was mixed before another incubation step of 10 min at room temperature followed. Finally, 400 µL of 50 % ethanol was added and the sample was mixed. The lysate was transferred to a purification column, centrifuged (6000 x g, 1 min) and the flow-through was discarded. The gDNA was washed by adding 500 µL of Wash buffer I and subsequent centrifugation (8000 x g, 1 min) before a second wash step was performed with 500 µL Wash Buffer II (centrifugation: 15000 x g, 3 min, discard flow-through, another 1 min) and the column was transferred to an 1.5 mL microreaction tube. The gDNA was eluted stepwise from the columns with 120 µL Elution buffer and incubation for 2 min following centrifugation (10000 x g, 1 min) whereby the flow-through of the preceding elution step was used for the next column. The quality and the concentration of the obtained gDNA were determined by agarose gel electrophoresis compared to defined amounts of λ DNA and additionally via absorbance measurement (3.2.7).

### 3.2.2 Polymerase chain reaction

Polymerase chain reaction (PCR) is a method to amplify DNA and generate millions of copies of a desired sequence. The reaction mixture was set up on ice in PCR tubes and comprises a thermostable polymerase, dNTPs, the DNA template and specifically designed primers (3.1.7). PCR was conducted in a thermocycler (Mastercycler ep gradient, Eppendorf, Hamburg, Germany).

The method was used in different contexts: a) to generate inserts and linearized vectors with overlapping sequences in Gibson Assembly (3.2.3), b) to introduce mutations for site-directed/site-saturation mutagenesis of EmOAH using the QuikChange method (3.2.4) and c) to identify *G. stearotherophilus* by species-specific PCR assay targeting the ITS 16S-23S rRNA region (Table 15 and Table 16) or by 16S rRNA sequencing (Table 17 and Table 18).<sup>[155,161]</sup> While the according set-ups of the PCR reaction and the methods for a) and b) are described in the respective chapters, information for c) is given in the following.

**Table 15:** Reaction mixture of PCR for the identification of *G. stearotherophilus* using the species-specific primers FITS and RITS targeting the ITS 16S-23S rRNA region.

substance	volume [ $\mu\text{L}$ ]	final concentration/amount
template gDNA	1	50 ng
FITS (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
RITS (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
dNTPs (10 mM each)	1.25	0.25 mM each dNTP
MgCl <sub>2</sub> (25 mM)	5	2.5 mM
10 x Taq buffer	5	1 x
Taq polymerase (1 U $\mu\text{L}^{-1}$ )	2.5	2.5 U
ddH <sub>2</sub> O	ad 50 $\mu\text{L}$	-

**Table 16:** Temperature program for the identification of *G. stearotherophilus* using the species-specific primers FITS and RITS targeting the ITS 16S-23S rRNA region. Lid temperature was set to 105 °C.

step	temperature [ $^{\circ}\text{C}$ ]	duration [s]	cycles
initial denaturation	94	180	1
denaturation	94	15	25
annealing/elongation	69	30	

A 302 bp fragment should be produced by PCR and the product should be identified by agarose gel electrophoresis (3.2.5) using a 2 % *m V*<sup>-1</sup> agarose gel and 7 V cm<sup>-1</sup> (60 V) for 2 h.

**Table 17:** Reaction mixture of PCR for the identification of *G. stearothermophilus* using the primers 8F and 1492R for 16S rRNA sequencing.

substance	volume [ $\mu\text{L}$ ]	final concentration/amount
template gDNA	1	50 ng
8F (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
1492R (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
dNTPs (10 mM each)	1.25	0.25 mM each dNTP
MgCl <sub>2</sub> (25 mM)	5	2.5 mM
10 x Pfu buffer	5	1 x
Pfu polymerase (2.5 U $\mu\text{L}^{-1}$ )	1	2.5 U
ddH <sub>2</sub> O	ad 50 $\mu\text{L}$	-

**Table 18:** Temperature program for the identification of *G. stearothermophilus* using the primers 8F and 1492R for 16S rRNA sequencing. Lid temperature was set to 105 °C.

step	temperature [ $^{\circ}\text{C}$ ]	duration [s]	cycles
initial denaturation	94	180	1
denaturation	94	60	
annealing	54	45	30
elongation	72	120	
final elongation	72	420	1
cooling	8	hold	1

The obtained PCR fragments were purified by DNA Clean & Concentrator Kit according to manufacturer's protocol and eluted in 20  $\mu\text{L}$  ddH<sub>2</sub>O. The concentration was measured by NanoDrop (3.2.7), purity of the DNA was determined by agarose gel electrophoresis (3.2.5) and the DNA was sent for sequencing using the same primers.

### 3.2.3 Gibson Assembly

Gibson Assembly<sup>[162]</sup> was used for cloning of hydratase genes from pET28a(+) in pDHE vectors and for construction of empty vector pDHE-MCS. In a first step, DNA fragments with overlapping ends are generated by two separate PCR reactions: one is for the linearization of the plasmid backbone, the other is for the amplification of the insert. Therefore, four primers (Table 7) are needed for each Gibson Assembly.



### Cloning of TrCrtC and RgCrtC

Cloning of TrCrtC was used as an occasion to transfer the His<sub>6</sub>-tag from the N- to the C-terminus. The PCR reaction set-up and used temperature program for this purpose can be found in Table 19 and Table 20. Four PCR reactions were conducted in parallel for actual Gibson Assembly of TrCrtC and RgCrtC. Table 21 shows which template DNA, forward and reverse primers were used and which annealing temperature was applied during PCR. The conditions for the PCR reaction are summarized in Table 22 and Table 23.

**Table 19:** Reaction mixture of PCR for transfer of His<sub>6</sub>-tag in TrCrtC.

substance	volume [ $\mu\text{L}$ ]	final concentration/amount
pET28a(+)-TrCrtC	1	100 ng
JSC_TrCrtC_fw (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
JSC_TrCrtC-His rev (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
dNTPs (10 mM each)	1	0.2 mM each dNTP
MgSO <sub>4</sub> (25 mM)	4	2 mM
10 x Pfu buffer	5	1 x
Pfu polymerase (5 U $\mu\text{L}^{-1}$ )	1	5 U
ddH <sub>2</sub> O	ad 50 $\mu\text{L}$	-

**Table 20:** Temperature program of PCR for transfer of His<sub>6</sub>-tag in TrCrtC. Lid temperature was set to 105 °C.

step	temperature [ $^{\circ}\text{C}$ ]	duration [s]	cycles
initial denaturation	95	120	1
denaturation	95	30	
annealing	65	30	30
elongation	72	180	
final elongation	72	600	1
cooling	8	hold	1

**Table 21:** List of used templates and primers for Gibson PCR of TrCrtC and RgCrtC.

	TrCrtC vector	TrCrtC fragment	RgCrtC vector	RgCrtC fragment
template DNA	pDHE_pelB_ LinD-His_noADH	TrCrtC-His	pDHE_pelB_ LinD-His_noADH	pET28-RgCrtC
fw primer (melting temp.)	JSC_pDHE_ TrCrtC_H_vf (69 °C)	JSC_pDHE_ TrCrtC_H_ff (65 °C)	JSC_pDHE_ RgCrtC_vf (59 °C)	JSC_pDHE_ RgCrtC_ff (68 °C)
rev primer (melting temp.)	JSC_pDHE_ TrCrtC_H_vr (65 °C)	JSC_pDHE_ TrCrtC_H_fr (64 °C)	JSC_pDHE_ RgCrtC_vr (65 °C)	JSC_pDHE_ RgCrtC_fr (64 °C)
PCR annealing temp.	64 °C	64 °C	60 °C	64 °C

**Table 22:** Reaction mixture of PCR for vector linearization or insert amplification of CrtCs.

substance	volume [ $\mu\text{L}$ ]	final concentration/amount
template DNA	1	
fw primer (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
rev primer (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
dNTPs (10 mM each)	1	0.2 mM each dNTP
MgSO <sub>4</sub> (25 mM)	4	2 mM
10 x Pfu buffer	5	1 x
Pfu polymerase (5 U $\mu\text{L}^{-1}$ )	1	5 U
ddH <sub>2</sub> O	ad 50 $\mu\text{L}$	-

**Table 23:** Temperature program for Gibson Assembly of CrtCs: PCR for vector linearization and insert amplification. Lid temperature was set to 105 °C.

step	temperature [ $^{\circ}\text{C}$ ]	duration [s]	cycles
initial denaturation	95	120	1
denaturation	95	30	
annealing	60/64	30	30
elongation	72	720	
final elongation	72	600	1
cooling	8	hold	1

### *Cloning of EmOAH*

The conditions for the PCR reactions for Gibson assembly of EmOAH are summarized in Table 24 to Table 26.

**Table 24:** List of used templates and primers for Gibson PCR of EmOAH.

	vector	fragment
template DNA	pDHE_pelB_LinD-His_noADH	pET28a(+)-EmOAH
fw primer (melting temp.)	JSC_pDHE_EmOAH_vf (70 °C)	JSC_pDHE_EmOAH_ff (65 °C)
rev primer (melting temp.)	JSC_pDHE_EmOAH_vr (66 °C)	JSC_pDHE_EmOAH_fr (70 °C)
PCR annealing temp.	66 °C	65 °C

**Table 25:** Reaction mixture of PCR for vector linearization or insert amplification of EmOAH.

substance	volume [ $\mu\text{L}$ ]	final concentration/amount
template DNA	1	
fw primer (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
rev primer (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
dNTPs (10 mM each)	1	0.2 mM each dNTP
MgSO <sub>4</sub> (25 mM)	4	2 mM
10 x Pfu buffer	5	1 x
Pfu polymerase (5 U $\mu\text{L}^{-1}$ )	1	5 U
ddH <sub>2</sub> O	ad 50 $\mu\text{L}$	-

**Table 26:** Temperature program for Gibson Assembly of EmOAH: PCR for vector linearization and insert amplification. Lid temperature was set to 105 °C.

step	temperature [°C]	duration [s]	cycles
<b>initial denaturation</b>	95	120	1
<b>denaturation</b>	95	30	
<b>annealing</b>	65/66	30	30
<b>elongation</b>	72	600	
<b>final elongation</b>	72	600	1
<b>cooling</b>	8	hold	1

### *Cloning of pDHE-MCS*

The conditions for the PCR reactions for Gibson assembly of pDHE-MCS are summarized in Table 27 to Table 29.

**Table 27:** List of used templates and primers for Gibson PCR of pDHE-MCS.

	vector	fragment
<b>template DNA</b>	pDHE_RgCritC	pET28a(+)
<b>fw primer</b> (melting temp.)	JSC_pDHE_MCS_vf (73 °C)	JSC_pDHE_MCS_ff (65 °C)
<b>rev primer</b> (melting temp.)	JSC_pDHE_MCS_vr (64 °C)	JSC_pDHE_MCS_fr (76 °C)
<b>PCR annealing temp.</b>	65 °C	65 °C

**Table 28:** Reaction mixture of PCR for vector linearization or insert amplification of pDHE-MCS.

substance	volume [µL]	final concentration / amount
<b>template DNA</b>	1	
<b>fw primer (10 µM)</b>	1	0.2 µM
<b>rev primer (10 µM)</b>	1	0.2 µM
<b>dNTPs (10 mM each)</b>	1	0.2 mM each dNTP
<b>10 x PfuUltra buffer</b>	5	1 x
<b>PfuUltra II polymerase (2.5 U µL<sup>-1</sup>)</b>	1	2.5 U
<b>ddH<sub>2</sub>O</b>	ad 50 µL	-

**Table 29:** Temperature program for Gibson Assembly of pDHE-MCS: PCR for vector linearization and insert amplification. Lid temperature was set to 105 °C.

step	temperature [°C]	duration [s]	cycles
<b>initial denaturation</b>	95	120	1
<b>denaturation</b>	95	30	
<b>annealing</b>	65	30	30
<b>elongation</b>	72	100	
<b>final elongation</b>	72	600	1
<b>cooling</b>	8	hold	1

The obtained PCR fragments are purified by DNA gelelectrophoresis (3.2.5) and consequently extracted from the agarose gel using Zymoclean Gel DNA Recovery Kit according to manufacturer's protocol. The DNA is eluted in 10  $\mu\text{L}$  ddH<sub>2</sub>O before the concentration is measured by NanoDrop (3.2.7).

The actual Gibson assembly is an isothermal reaction comprising three enzyme activities:

- 1) an exonuclease which generates single-stranded regions at the 3'-ends which can anneal to each other due the overlapping sequence (*sticky ends*)
- 2) a DNA polymerase filling the gaps by incorporating nucleotides
- 3) a DNA ligase for removal of nicks in the DNA

The composition of the Gibson Assembly master mix is listed in the following (Table 30 and Table 31).<sup>[163]</sup>

**Table 30:** Composition of the 5 x ISO reaction buffer for Gibson Assembly master mix. The components were mixed until all components were homogeneously dissolved and subsequently filter-sterilized.

<b>component</b>	<b>amount</b>	<b>final concentration</b>
PEG-8000	1.5 g	25 % $m V^{-1}$
1 M Tris-HCl (pH 7.5)	3000 $\mu\text{L}$	500 mM
2 M MgCl <sub>2</sub>	150 $\mu\text{L}$	50 mM
1 M DTT	300 $\mu\text{L}$	50 mM
dNTPs (100 mM each)	240 $\mu\text{L}$	1 mM each
100 mM NAD <sup>+</sup>	300 $\mu\text{L}$	5 mM
ddH <sub>2</sub> O	ad 6000 $\mu\text{L}$	

**Table 31:** Gibson Assembly master mix. Aliquots of 15  $\mu\text{L}$  were stored in PCR tubes at -20 °C until needed.

<b>component</b>	<b>volume [<math>\mu\text{L}</math>]</b>
5 x ISO reaction buffer	320
T5 exonuclease (10 U $\mu\text{L}^{-1}$ )	0.64
Phusion HF DNA polymerase (2 U $\mu\text{L}^{-1}$ )	20
Taq DNA ligase (40 U $\mu\text{L}^{-1}$ )	160
ddH <sub>2</sub> O	ad 1200

For Gibson Assembly, about 200 ng linearized vector DNA was mixed in a molar ratio of 1:5 with insert DNA and filled up to a volume of 5  $\mu\text{L}$  with ddH<sub>2</sub>O. The required amounts of DNA were calculated using the online tool NEBioCalculator (<https://nebiocalculator.neb.com/#!/ligation>). An aliquot of the Gibson Assembly master mix (15  $\mu\text{L}$ ) was thawed on ice and the DNA mixture (5  $\mu\text{L}$ ) was added. The isothermal assembly reaction was performed in a thermocycler (Mastercycler ep gradient, Eppendorf, Hamburg, Germany) at 50 °C for 1 h. The crude mixture was directly used after reaction to transform *E. coli* DH5 $\alpha$  or XL-1 blue competent cells and spread on

LB agar plates. Single colonies were used to inoculate a LB precultures from which plasmids were isolated. The correct construction of the plasmids was confirmed by Sanger sequencing.

### 3.2.4 Site-directed and site-saturation mutagenesis

QuikChange<sup>[164]</sup> is a method for *in vitro* site-directed or site-saturation mutagenesis. It is used for vector modifications by one or more mutations using mutagenic primers (in the case of site-saturation mutagenesis primers with different degenerated codons were used, 3.1.7) and PfuUltra II DNA polymerase in a PCR-like procedure (Table 32 and Table 33).

**Table 32:** Reaction mixture for QuikChange reaction.

substance	volume [ $\mu\text{L}$ ]	final concentration/amount
template plasmid	1	50-80 ng
fw primer (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
rev primer (10 $\mu\text{M}$ )	1	0.2 $\mu\text{M}$
dNTPs (10 mM each)	1	0.2 mM each dNTP
DMSO	1	2 %
10 x PfuUltra buffer	5	1 x
PfuUltra II polymerase (2.5 U $\mu\text{L}^{-1}$ )	0.75	1.875 U
ddH <sub>2</sub> O	ad 50 $\mu\text{L}$	-

For set-up of the reaction mixtures, if possible, a mastermix was prepared by adding the following components: water, buffer, DMSO, dNTPs and template. Primers and polymerase were added to aliquots of the mastermix before the reaction was started.

**Table 33:** Temperature program for QuikChange reaction. Lid temperature was set to 105 °C.

step	temperature [ $^{\circ}\text{C}$ ]	duration [s]	cycles
initial denaturation	95	120	1
denaturation	95	30	
annealing	55	45	20
elongation	72	150	
final elongation	72	300	1
cooling	8	hold	1

The mutagenic primers anneal to the vector and are elongated subsequently by the polymerase resulting in a mutated plasmid with staggered nicks. Afterwards, the crude mixture is treated with 1  $\mu\text{L}$  DpnI (10 U  $\mu\text{L}^{-1}$ ) overnight at 37 °C to digest the methylated template DNA. Copies of the plasmid are not methylated and therefore remain undigested in the PCR mixture. For purification, 20  $\mu\text{L}$  of the PCR mixture was

dialyzed for 20 min against water using a nitrocellulose membrane (MF-Membrane Filters, 0.025  $\mu\text{m}$  VSWP, Merck Millipore, Darmstadt, Germany). Finally, the purified DNA was directly used to transform 50  $\mu\text{L}$  competent *E. coli* cells (3.2.9). The cells repair the nicks in the mutated plasmid. Single colonies were picked from the agar plate and used to inoculate a LB preculture from which the plasmid was isolated and the sequence was confirmed by Sanger sequencing.

For simultaneous site-saturation mutagenesis in a CAST-like procedure using degenerated codons, special primer mixtures have been prepared according to Table 34.

**Table 34:** Equimolar primer mixtures [10  $\mu\text{M}$ ] for the construction of EmOAH libraries were prepared by mixing the given amounts of 100  $\mu\text{M}$  primer stocks. When the required volume was too low for pipetting, 100  $\mu\text{L}$  of a 1:10 dilution was made before finally added to the primer mix.

		amount of codons			for 10 $\mu\text{L}$	dilution	volume in mix [ $\mu\text{L}$ ]	ad 100 $\mu\text{L}$ H <sub>2</sub> O
<b>R121</b>	NNK	32	1	<b>32</b>	-	-	10	90.00
<b>M226+F227</b>	NDT, NDT	12	12	144	2.975		2.98	
	NDT, VHG	12	9	108	2.231		2.23	
	NDT, TGG	12	1	12	0.248	1:10	2.48	
	VHG, NDT	9	12	108	2.231		2.23	
	VHG, VHG	9	9	81	1.674		1.67	
	VHG, TGG	9	1	9	0.186	1:10	1.86	
	TGG, NDT	1	12	12	0.248	1:10	2.48	
	TGG, VHG	1	9	9	0.186	1:10	1.86	
	TGG, TGG	1	1	1	0.021	1:10	0.21	
	<b>total</b>			<b>484</b>	10.000		18.00	82.00
<b>F514</b>	NYK	16	1	16			8.89	
	TAT	1	1	1			0.56	
	TGG	1	1	1			0.56	
	<b>total</b>			<b>18</b>			10.00	90.00
<b>F548+T549</b>	NYK, NYK	16	16	256	7.901		7.90	
	NYK, TAT	16	1	16	0.494		0.49	
	NYK, TGG	16	1	16	0.494		0.49	
	TAT, NYK	1	16	16	0.494		0.49	
	TAT, TAT	1	1	1	0.031	1:100	3.09	
	TAT, TGG	1	1	1	0.031	1:100	3.09	
	TGG, NYK	1	16	16	0.494		0.49	
	TGG, TAT	1	1	1	0.031	1:100	3.09	
	TGG, TGG	1	1	1	0.031	1:100	3.09	
	<b>total</b>			<b>324</b>	10.000		22.22	77.78

### 3.2.5 Agarose gel electrophoresis & gel extraction

Agarose gel electrophoresis was used to analyze and separate DNA according to their size. Agarose gels (0.7 %, 1 % or 2 %  $m V^{-1}$ ) were prepared by dissolving agarose in 1 x TAE buffer in heat. Midori Green (0.1  $\mu L mL^{-1}$ ) was added to the still warm solution to visualize the DNA under UV light. After solidification, samples were mixed with a fifth of 6 x DNA loading dye (composition according to the supplier Thermo Scientific: 10 mM Tris-HCl pH 7.6, 0.03 % bromophenol blue, 0.03 % xylene cyanol FF, 60 % glycerol, 60 mM EDTA) and loaded into the wells. GeneRuler 1 kb plus (5  $\mu L$ ) was used as size standard. Electrophoresis was conducted in a horizontal gel system (PerfectBlue Mini, Peqlab, Erlangen, Germany) filled with 1 x TAE. Separation was achieved by application of, unless otherwise stated, 110 V (Power Pac 300, Bio-Rad, Hercules, USA) for about 60 min. After electrophoresis, the bands on the gel were visualized with UV and documented.

Agarose gel electrophoresis was also used to purify DNA after PCR for amplification of insert and linearization of vector in Gibson Assembly (3.2.3). Therefore DNA bands of desired size were cut out under UV light. The DNA was recovered from agarose gel by gel extraction using Zymoclean Gel DNA Recovery Kit according to manufacturer's protocol.

### 3.2.6 Plasmid preparation

Plasmids were isolated from 5 mL LB cultures using Zyppy Miniprep Kit according to manufacturer's protocol. Plasmid DNA was eluted in 30  $\mu l$  ddH<sub>2</sub>O and the concentration was measured using NanoDrop.

### 3.2.7 Determination of DNA concentration (NanoDrop)

The concentration of DNA was measured at 260 nm against water background with NanoDrop 2000 (Thermo Scientific, Waltham, USA) using 1  $\mu L$  of DNA sample.

### 3.2.8 Sequencing

DNA Sanger sequencing was performed by GATC Biotech AG (Konstanz, Germany) using sent in or in-house primers (20  $\mu\text{L}$ , 10 pmol  $\mu\text{L}^{-1}$ , Table 6). The concentrations of 20  $\mu\text{L}$  DNA samples were adjusted to 10-50 ng  $\mu\text{L}^{-1}$  for PCR fragments and 30-100 ng  $\mu\text{L}^{-1}$  for plasmids.

### 3.2.9 Preparation of chemo-competent cells

Chemo-competent cells were prepared according to the rubidium chloride method. A preculture was used to inoculate 100 mL LB medium to an  $\text{OD}_{600}$  of 0.05 followed by incubation (37 °C, 180 rpm) until an  $\text{OD}_{600}$  of 0.5 was reached. Cells were harvested after 5 min on ice by centrifugation (3220 x g, 4 °C, 10 min; Eppendorf Centrifuge 5810R, Hamburg, Germany). Subsequently, the pellet was resuspended in 40 mL TfbI buffer (3.1.9). After incubation for 15 min on ice, the cells were harvested by centrifugation (3220 x g, 4 °C, 10 min) and were resuspended in 5 mL TfbII buffer. The cells were incubated again for 15 min on ice, divided into 50  $\mu\text{L}$  aliquots and frozen with liquid nitrogen before storing at -80 °C.

### 3.2.10 Transformation

*E. coli* chemo-competent cells (3.2.9) were transformed with plasmids by the heat shock method. Therefore 1  $\mu\text{L}$  of the plasmid was added to 50  $\mu\text{L}$  thawed cells. After incubation on ice for 30 min, the mixture was treated for 45 s at 42 °C in a hot water bath. Subsequently, the cells were cooled on ice when 800  $\mu\text{L}$  cold LB medium was added. Following the incubation (37 °C, 1 h, 180 rpm), cells were spread on LB agar plates containing appropriate antibiotics as selection agent. Finally, the plates were incubated overnight at 37 °C. For transformation with plasmids obtained after QuikChange (3.2.4) or Gibson Assembly (3.2.3), higher amounts were used as specified in the respective chapters.



### 3.3 Microbiological methods

#### 3.3.1 Cultivation of *Geobacillus stearothermophilus*

##### *Revival of G. stearothermophilus*

*Geobacillus stearothermophilus* BR388 (ATCC 55596) was purchased from ATCC / LGC standards as frozen liquid culture and revived according to supplier's protocol. Therefore, the culture (~0.5 mL) was thawed in a water bath at 55 °C for 2 min until completely melted and the entire culture was used to inoculate 5 mL LB medium under strictly sterile conditions. From this culture, 0.5 mL was used to inoculate a secondary culture (5 mL LB medium) and 100 µL were spread on an LB agar plate. The cultures were incubated overnight (55 °C, 125 rpm).

##### *Cultivation of G. stearothermophilus on plates*

Cells were either spread on TSB agar plates (3.1.9) and incubated in an incubator overnight at 55 °C under a beaker to retain water and prevent dehydration of agar or on M9 minimal agar plates which were set into a desiccator equipped with a small amount (2 mL) of neat (*R*)-(+)-limonene. The plates were incubated in the closed desiccator for 4-6 d at 55 °C which leads to evaporation of the limonene and can thereby be used by the cells as carbon source.

##### *Cultivation of G. stearothermophilus in liquid culture*

In the following, the best method which was found by an evaluation of several cultivation conditions is described. Liquid cultures were grown in baffled flasks of five times the culture volume in TSB medium. It was important to use “peptone from soybean-enzymatic digest” (Sigma-Aldrich 87972) and “casein hydrolysate for microbiology” (Sigma-Aldrich 22090) for successful cultivation. The medium was inoculated either by cells from previous liquid cultures or from agar. The culture was incubated at 55 °C and 180 rpm overnight. Cell growth was followed through measurement of absorption at 600 nm (OD<sub>600</sub>) with a photometer (Ultrospec 3100 pro, Amersham Biosciences, Amersham, UK).

### 3.3.2 Cultivation of *E. coli*

#### *Agar plates*

*E. coli* cells were spread on LB agar plates containing appropriate antibiotics (3.1.9) as selection agent and incubated overnight at 37 °C.

#### *Precultures*

Appropriate antibiotics were added to 5 mL LB medium and the medium was inoculated from a LB agar plate. The preculture was incubated overnight at 37 °C and 180 rpm.

#### *Liquid cultures (for protein production)*

Liquid cultures were grown in Erlenmeyer flasks (with and without baffles) of five times the culture volume. Appropriate antibiotics were added to the TB medium. The culture was incubated, unless otherwise stated, at 37 °C and 180 rpm.

Cell growth was followed through measurement of absorption at 600 nm (OD<sub>600</sub>) with a photometer (Ultraspec 3100 pro, Amersham Biosciences, Amersham, UK)

#### *Glycerol stocks for long-term storage*

For the preparation of glycerol stocks, a part of the preculture was added to the same volume of sterile glycerol and resuspended. *E. coli* cell cultures were stored at -80 °C.

### 3.3.3 Cultivation of *E. coli* in 96-DWP

#### *Precultures*

LB medium (1 mL) with ampicillin was dispensed in 96-DWP and each well was inoculated by a distinct single colony from agar plates after transformation of *E. coli* ITB 94 with the respective plasmid library created by site-saturation mutagenesis (3.2.4). Empty wells and wells inoculated with wild type EmOAH served as control. The DWP was sealed with an air-permeable membrane and incubated at 37 °C at 350 rpm for 21 h. OD<sub>600</sub> was measured in a 1:10 dilution at a total volume of 200 µL per well in 96-well plates using a plate reader (PolarStar Omega 96, BMG Labtech GmbH, Ortenberg, Germany).

### *Glycerol stocks for long-term storage*

To a 96-DWP filled with 500  $\mu$ L sterile glycerol the same volume of the preculture plate was added well per well and mixed. The glycerol stock plate was sealed with an aluminium sealing foil and stored at  $-80\text{ }^{\circ}\text{C}$ .

## **3.4 Proteinbiochemical methods**

### **3.4.1 Protein production**

For protein production, *E. coli* BL21 (DE3) or *E. coli* ITB 94 cells were freshly transformed (3.2.10) and a preculture was inoculated (3.3.2).

#### *Procedure for E. coli BL21 (DE3)*

The preculture was used to inoculate TB medium containing kanamycin (3.1.9) to an  $\text{OD}_{600}$  of 0.05. The culture was incubated at  $37\text{ }^{\circ}\text{C}$  and 180 rpm until an  $\text{OD}_{600}$  of 0.5 was reached (about 2 h). A sample normalized to  $\text{OD}_{600} = 0.25$  was taken for SDS-PAGE (3.4.3), showing the protein composition before induction. Expression was induced by adding 1 M IPTG solution (final concentration: 0.5 mM) at  $20\text{ }^{\circ}\text{C}$ . The expression culture was allowed to shake overnight (20-21 h) at  $20\text{ }^{\circ}\text{C}$  and 180 rpm. Another sample normalized to  $\text{OD}_{600} = 0.25$  was taken for SDS-PAGE, showing the protein composition after induction. Cells were harvested by centrifugation ( $4\text{ }^{\circ}\text{C}$ ,  $7500 \times g$ , 20 min) and directly used in biotransformations or stored at  $-20\text{ }^{\circ}\text{C}$ .

#### *Procedure for E. coli ITB 94*

TB medium containing ampicillin (3.1.9) and rhamnose (final concentration:  $0.5\text{ g L}^{-1}$  added from a  $50\text{ g L}^{-1}$  stock) was inoculated with 1 %  $V V^{-1}$  preculture. The culture was incubated at  $30\text{ }^{\circ}\text{C}$  and 180 rpm for about 24 h. Samples normalized to  $\text{OD}_{600} = 0.25$  were taken for SDS-PAGE (3.4.3), showing the protein composition after incubation. Cells were harvested by centrifugation ( $4\text{ }^{\circ}\text{C}$ ,  $7500 \times g$ , 20 min) and directly used in biotransformations or stored at  $-20\text{ }^{\circ}\text{C}$ .

#### *Procedure for E. coli ITB 94 in 96-DWP*

After cultivation of a preculture 96-DWP (3.3.3), 50  $\mu\text{L}$  cell suspension was used to inoculate another 96-DWP with 950  $\mu\text{L}$  (per well) TB medium containing 100  $\mu\text{g mL}^{-1}$  ampicillin and 0.5  $\text{g L}^{-1}$  rhamnose for induction. The DWP was sealed with an air-permeable membrane and incubated at 37  $^{\circ}\text{C}$  at 350 rpm for about 22 h.  $\text{OD}_{600}$  was measured in a 1:10 dilution at a total volume of 200  $\mu\text{L}$  per well in 96-well plates using a plate reader (PolarStar Omega 96, BMG Labtech GmbH, Ortenberg, Germany). The cells were harvested by centrifugation (4  $^{\circ}\text{C}$ , 3220  $\times$  g, 15 min) and the supernatant was discarded. The 96-DWP with cells containing the over-expressed EmOAH variants were either directly used for whole cell biotransformations or sealed with an aluminium foil and stored at -80  $^{\circ}\text{C}$ .

#### **3.4.2 Cell disruption**

Cells were resuspended in buffer (5 or 10  $\text{mL g}^{-1}$  cell pellet) and subsequently disrupted on ice by sonication (Sonifier 250, Branson, Danbury, USA). Sonication was conducted depending on the buffer volume and the cell amount for 5-10 min (duty cycle = 35 %, output = 3-5) until the suspension turns from milky-opaque to a clearer solution. The crude lysate was centrifuged for smaller volumes in reaction tubes (4  $^{\circ}\text{C}$ , 16000  $\times$  g, 20 min) or for larger volumes in falcon tubes (4  $^{\circ}\text{C}$ , 7500  $\times$  g, 20 min). The supernatant was used as protein solution while the precipitate was discarded. The lysate was analyzed by SDS-PAGE (3.4.3) and the concentration was determined by BCA assay (3.4.4).

#### **3.4.3 SDS-PAGE**

SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) was used to analyze the size and purity of proteins. Commercially available, precast gels were used (GenScript ExpressPlus PAGE gel, 10x8, 4-20 % gradient or 12 %, 12 wells). Protein samples were prepared as follows: protein solutions were diluted to a concentration of 2  $\text{g L}^{-1}$  and mixed with the same volume of 2  $\times$  SDS sample buffer (Table 35). The mixture was heated to 95  $^{\circ}\text{C}$  for 10 min for protein denaturation. Samples of cell pellet taken during protein production (3.4.1) are resuspended in 50  $\mu\text{L}$  2  $\times$  SDS sample buffer (Table 35) and denatured at 95  $^{\circ}\text{C}$  for 10 min.

The electrophoresis chamber (Mini Protean Tetra Cell, Biorad, Hercules, USA) was filled with SDS running buffer (buffer was prepared by adding 1 L ddH<sub>2</sub>O to one bag of GenScript TRIS-MOPS-SDS running buffer powder, which is, according to the manufacturer website, composed of 6.06 g L<sup>-1</sup> TRIS base, 10.46 g L<sup>-1</sup> MOPS, 1 g L<sup>-1</sup> SDS, 0.3 g L<sup>-1</sup> EDTA). The denatured samples (10 µL for protein solutions, 15 µL for pellets) were then loaded along with 5 µL protein ladder (3.1.3) on the gel and the electric field was applied. Electrophoresis was conducted at 140 V (Power Pac 300, Bio-Rad, Hercules, USA) until the bromophenol blue band passed almost through the gel (about 50-60 min). Afterwards, the gels were stained overnight with Coomassie staining solution (60 % ddH<sub>2</sub>O, 30 % EtOH, 10 % pure acetic acid, 0.25 % *m V*<sup>-1</sup> Coomassie Brilliant Blue) before finally destained overnight (60 % ddH<sub>2</sub>O, 30 % EtOH, 10 % pure acetic acid).

**Table 35:** SDS sample buffer.

<b>substance</b>	<b>amount/volume</b>
<b>1 M TRIS-HCl pH 6.8</b>	2 mL
<b>MgCl<sub>2</sub></b>	190 mg
<b>glycerol</b>	1 mL
<b>SDS</b>	0.8 g
<b>bromophenol blue</b>	2 mg
<b>DTT</b>	310 mg
<b>ddH<sub>2</sub>O</b>	ad 20 mL

#### 3.4.4 BCA assay

The concentration of protein solutions was determined with the photometric BCA assay using Pierce BCA Protein Assay Kit according to manufacturer's microwell plate protocol. In principle, the assay combines the reduction of Cu<sup>2+</sup>-ions to Cu<sup>+</sup>-ions by proteins under alkaline conditions with the colorimetric detection of the cuprous cation (Cu<sup>+</sup>) by bicinchoninic acid (BCA). BCA forms a purple complex with the Cu<sup>+</sup>-ions which can be detected. The absorbance was measured using the implemented BCA assay tool of the microplate reader software (software: SoftMaxPro; microplate reader: SpectraMax 340PC, Molecular Devices, San Jose, USA) at 562 nm after incubation for 15 min at 37 °C and compared to BSA standard solutions of known concentration treated under the same conditions.

### 3.5 Biotransformations

In this chapter, the standard conditions for biotransformations which have been found by an evaluation of several conditions are described. Depending on the investigated enzyme and substrate, these conditions (e.g. buffer, temperature, reaction vessel) may vary. Furthermore, biotransformations were conducted as whole cell and lysate reactions. More details on differing conditions and specific negative controls are given in the respective results part. In general, unless otherwise stated, the reactions were conducted in technical triplicates, which means that the reactions were performed three times using cells from the same expression or the same lysate batch.

#### 3.5.1 Analytical biotransformations

##### *Lysate biotransformations with CrtC*

For the investigation of the substrate scope of CrtCs, biotransformations were conducted in 500  $\mu\text{L}$  scale in 2 mL glass vials using lysate of *E. coli* ITB 94 cells containing over-expressed CrtC (3.4.2). Experiments were set as follows: 25 g  $\text{L}^{-1}$  lysate, 20  $\mu\text{M}$  lycopene (from 1 mM stock in acetone), 0.5/1/2.5 % Triton X-100 as detergent (added in various amounts from a 10 % solution). Reactions were filled up with 50 mM  $\text{KPi}$  buffer pH 8.0. Incubation was performed in a Thermomixer (Eppendorf, Hamburg, Germany) at 28  $^{\circ}\text{C}$  and 800 rpm for different periods of time and protected from light by aluminium foil. Reactions were conducted in triplicates using the same lysate. Negative controls (without lysate / without substrate / cell lysate containing EmOAH) were performed as single experiments.

The reaction mixture was extracted twice with MTBE by adding 500  $\mu\text{L}$  MTBE, vortexing and centrifugation (room temperature, 7000 x g, 15 min) whereby 300  $\mu\text{L}$  extract was transferred into autosampler vials. The combined organic layer was analyzed by HPLC-DAD or LC-MS.

### *Whole cell biotransformations with OAH*

Biotransformations of DMDDAC or derivatives with OAHs were performed in 500  $\mu\text{L}$  scale in 2 mL glass vials (50 mM sodium citrate buffer pH 6.0, 50  $\text{g}_{\text{cww}} \text{L}^{-1}$  cells, 500  $\mu\text{M}$  substrate, 1 % DMSO, 100 mM glucose, 0.3 mM FAD, 3 mM NADH, 2 mM DTT) and either incubated in a shaking incubator (Infors HT Minitron, Basel, Switzerland) at 180 rpm and 30 °C for 4 d or in a Thermomixer (Eppendorf, Hamburg, Germany) at 600 rpm and 30 °C for 4 d. Negative controls were performed using cells containing an expressed empty vector instead of OAH. Subsequently, 900  $\mu\text{L}$  MTBE (occasionally 100  $\mu\text{M}$  octanoic acid was used as internal standard) was added and the mixture was vortexed for extraction. After centrifugation (room temperature, 7800 x g, 5 min) 600  $\mu\text{L}$  of the organic phase was transferred to autosampler vials and the solvent was removed. Finally, the residue was dissolved in the same amount acetonitrile (600  $\mu\text{L}$ ) and measured using LC-MS.

### **3.5.2 Screening of EmOAH variants in 96-DWP**

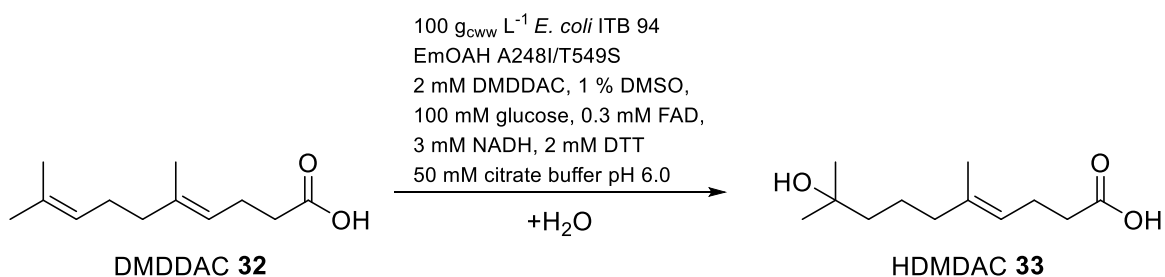
*E. coli* ITB 94 cells containing over-expressed variants of EmOAH were resuspended in 395  $\mu\text{L}$  50 mM citrate buffer pH 6.0 and were transferred from the 96-DWP used for protein production into another 96-DWP equipped with glass inserts for 96-DWP (Neolab, Heidelberg, Germany).

The cell suspension was supplemented with 100  $\mu\text{L}$  per well of an additive mastermix (25  $\mu\text{L}$  per additive, stock concentrations: 2 M glucose, 6 mM FAD, 60 mM NADH, 40 mM DTT, equivalent to a final concentration of 100 mM, 0.3 mM, 3 mM and 2 mM, respectively). Finally, the biotransformations were started by adding 5  $\mu\text{L}$  50 mM DMDDAC in DMSO (final concentrations: substrate 500  $\mu\text{M}$ , 1 % DMSO) and the plates were sealed with a cap mat. The plates were incubated for 4 d at 30 °C and 800 rpm using a microplate shaker in an incubator hood (TiMix 5 control and TH30, Edmund Bühler GmbH, Bodelshausen, Germany).

The reaction mixture was extracted by adding 500  $\mu\text{L}$  MTBE and subsequent mixing followed by centrifugation (room temperature, 3220 x g, 15 min). The organic layer (230  $\mu\text{L}$ ) was transferred to autosampler vials and analyzed by LC-MS.

### 3.5.3 Preparative whole cell biotransformation of DMDDAC

The preparative enzymatic synthesis of HDMDAC **32** from DMDDAC **31** was conducted in 800 x scale of analytical biotransformations. Therefore, *E. coli* ITB 94 whole cell catalyst containing EmOAH variant A248I/T549S was pooled from 7 x 400 mL expression cultures. The cells (40 g) were resuspended in 320 mL citrate buffer pH 6.0 to get a suspension of 125 g L<sup>-1</sup>. Stock solutions of additives were prepared by dissolving 100 mg FAD, 7920 mg glucose, 860 mg NADH and 124 mg DTT in citrate buffer (ad 20 mL), respectively.



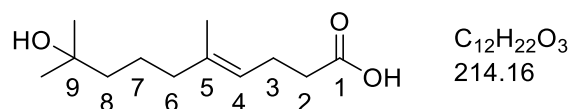
**Figure 19:** Preparative biotransformation of DMDDAC **32** to HDMDAC **33**.

In a 500 mL round-bottom flask, the cell suspension was supplemented with the additives to give a total volume of 400 mL with the following final concentrations: 100 g<sub>cww</sub> L<sup>-1</sup> cells, 100 mM glucose, 0.3 mM FAD, 3 mM NADH and 2 mM DTT. Finally, DMDDAC (174 μL, 157 mg, 0.8 mmol) was added as DMSO stock equivalent to a final concentration of 2 mM substrate and 1 % DMSO and the mixture was stirred for 8 d at 28 °C.

The reaction mixture was extracted with MTBE (3 x 400 mL), the combined organic layers were washed with brine (2 x 100 mL) before dried over MgSO<sub>4</sub> and concentrated *in vacuo* to give 412 mg of yellow-brown crude product. The crude product was purified by column chromatography (using cyclohexane / ethyl acetate 2:1 + 0.1 % formic acid) to yield 23 mg *E*-HDMDAC (0.11 mmol, 13 % based on *E/Z*-DMDDAC, 40 % based on 34 % *E*-DMDDAC and 98 % purity of substrate) as orange oil (LC-MS purity: 90 %).

TLC was used to monitor column chromatography. TLC plates were stained using a staining solution by Seebach *et al.*<sup>[165]</sup> (25 g phosphomolybdic acid, 10 g cerium (IV) sulfate, 60 mL conc. H<sub>2</sub>SO<sub>4</sub>, 940 mL ddH<sub>2</sub>O) and were developed in heat.





$^1H$ -NMR (500 MHz,  $CDCl_3$ ):  $\delta$  = 1.21 (s, 6H, 9-Me), 1.39-1.48 (m, 4H, 7-H, 8-H), 1.61 (s, 3H, 5-Me), 1.98 (t,  $J$  = 6.6 Hz, 2H, 6-H), 2.34 (m, 2H, 3-H), 2.37 (m, 2H, 2-H), 5.12 (t,  $J$  = 6.4 Hz, 1H, 4-H) ppm.

$^{13}C$ -NMR (125 MHz,  $CDCl_3$ ):  $\delta$  = 15.8 (Me-5), 22.3 (C-7), 23.4 (C-3), 29.1 (Me-9), 34.3 (C-2), 39.8 (C-6), 43.1 (C-8), 71.4 (C-9), 122.3 (C-4), 136.8 (C-5), 178.7 (C-1) ppm.

FT-IR (ATR):  $\tilde{\nu}$  = 3400 (w), 2968 (m), 2934 (s), 2872 (m), 2859 (m), 1711 (s), 1379 (m), 1261 (m), 1200 (m), 1148 (m), 911 (w)  $cm^{-1}$ .

LC-MS (ESI):  $m/z$  = 212.7  $[M-H]^-$ , 248.6  $[M+Cl]^-$ .

GC-MS (EI):  $m/z$  = 196.0  $[M-H_2O]^+$  (other main fragments: 181.0, 140.0, 95.0).

## 3.6 Analytical methods

### 3.6.1 Gas chromatography (GC)

GC methods for analysis of undec-9-enoic acid are described elsewhere.<sup>[127,157]</sup> The method for analysis of HDMDAC obtained from preparative biotransformation of DMDDAC is described in the following. The temperature program of the GC-MS method is shown in Table 36.

GC-MS parameters:

Agilent Technologies (Santa Clara, USA) 7890A GC system with 5975C MSD mass selective detector (electron impact, 70 eV), 7693 autosampler / autoinjector (injection volume: 1  $\mu$ L, split ratio: 1:5 or 1:20, injection temperature: 250  $^{\circ}$ C), carrier gas: helium (15.497 psi constant pressure), column: DB-5 polyphenylmethylsiloxane (Agilent J&W, Santa Clara, USA) (length: 30 m, inner diameter: 0.25 mm, film: 0.25  $\mu$ m), detector temperature FID: 320  $^{\circ}$ C; MS: ion source temperature 230  $^{\circ}$ C, transfer line temperature 250  $^{\circ}$ C, quad temperature: 150  $^{\circ}$ C, Scan parameters:  $m/z$  = 80-220, SIM:  $m/z$  = 74 or 140, solvent delay 2 min.

**Table 36:** GC-MS temperature program for analysis of DMDDAC and HDMDAC.

heating rate [ $\text{K min}^{-1}$ ]	temperature [ $^{\circ}\text{C}$ ]	hold time [min]
-	70	1
15	200	-
50	310	2

### 3.6.2 Liquid chromatography (LC)

Biotransformations were analyzed after extraction and sample workup (3.5) using high-performance liquid chromatography coupled with diode-array detection (HPLC-DAD) or mass spectrometry (LC-MS).

#### *HPLC-DAD analysis of lycopene biotransformations with CrtC*

An Agilent 1200 system equipped with a 1260  $\mu$ -degasser (G1379B), a quaternary pump (G1311A), an ALS autosampler (G1329A), a TCC thermostatted column compartment (G1316A) and a DAD diode array detector (G1315D) was used.

Method parameters: column: Prontosil 120-3-C30, 3  $\mu$ m, 100 x 4.6 mm (Bischoff Chromatography, Leonberg, Germany), oven temperature: 30 °C, flow: 1.2 mL min<sup>-1</sup>, stop time: 25 min, solvent: premixed MeOH:MTBE 60:40 + 0.1 % formic acid, DAD: 476 nm (spectrum 380-600 nm).

#### *LC-MS analysis of lycopene biotransformations with CrtC*

An Agilent 1260 Infinity system equipped with a 1260 HiP degasser (G4225A), a binary pump SL (G1312B), a 1260 HiP ALS autosampler (G1367E), a 1260 TCC thermostatted column compartment (G1316A) and 1260 DAD VL diode array detector (G1315D) was used. Mass analysis was conducted on a 6130 quadrupole LC/MS single quadrupole mass detector.

Method parameters: column: Prontosil 120-3-C30, 3  $\mu$ m, 100 x 4.6 mm (Bischoff Chromatography, Leonberg, Germany), oven temperature: 30 °C, flow: 1.2 mL min<sup>-1</sup>, stop time: 33 min, solvent: premixed MeOH:MTBE 60:40 + 0.1 % formic acid, DAD: 476 nm (spectrum 380-600 nm). MS (positive): SCAN: 300-650, SIM: 536, 554, 572.

MS

MM-APCI

Drying gas flow [L min<sup>-1</sup>]: 5.0

Drying gas temperature [°C]: 325

Nebulizer pressure [psig]: 40

Vaporizer temperature [°C]: 200

Capillary voltage (positive) [V]: 2500

Corona current [ $\mu$ A]: 6.0

Charging voltage [V]: 2000

*LC-MS analysis of biotransformations with EmOAH*

An Agilent 1260 Infinity system equipped with a 1260 HiP degasser (G4225A), a binary pump SL (G1312B), a 1260 HiP ALS autosampler (G1367E), a 1260 TCC thermostatted column compartment (G1316A) and 1260 DAD VL diode array detector (G1315D) was used. Mass analysis was conducted on a 6130 quadrupole LC/MS single quadrupole mass detector. General parameters applied for all methods and the applied solvent gradients (Table 37 and Table 38) are described below. MS settings for the individual methods for each substrate are specified in Table 39.

**General parameters**

Column: Supelco Supelcosil LC 18-T (ITB Nr. 31)  
 Oven temperature: 45 °C  
 Flow: 1 mL min<sup>-1</sup>  
 Solvent: A: ddH<sub>2</sub>O + 0.1 % formic acid  
 B: ACN

**Table 37:** Solvent gradient used in the methods for the substrates DMDDAC, DMDDOL, 10-methylundecenoic acid and 9-undecynoic acid.

time [min]	A [%]	B [%]	flow [mL min <sup>-1</sup> ]
0.00	65.0	35.0	1.000
2.00	65.0	35.0	1.000
8.50	40.0	60.0	1.000
10.00	40.0	60.0	1.000
10.01	65.0	35.0	1.000
15.00	65.0	35.0	1.000

**Table 38:** Solvent gradient used in the method for the substrate DMDDIOX.

time [min]	A [%]	B [%]	flow [mL min <sup>-1</sup> ]
0.00	65.0	35.0	1.000
15.00	10.0	90.0	1.000
16.00	10.0	90.0	1.000
16.01	65.0	35.0	1.000
20.00	65.0	35.0	1.000

DAD 198 nm ref. wavelength 300 nm, SCAN (190-400 nm, 2 nm)

MS

API-ES

Drying gas flow [L min<sup>-1</sup>]: 12.0

Nebulizer pressure [psig]: 40

Drying gas temperature [°C]: 350

Capillary voltage [V]: 4000 (positive), 3500 (negative)

**Table 39:** Method parameters for MS.

substrate	ionization mode	SCAN	SIM
DMDDAC	negative	120-250	143+213
DMDDOL	positive	120-250	183+201
DMDDIOX	positive	100-300	162+267+285
10-methylundecenoic acid	negative	120-250	215
9-undecynoic acid	negative	120-250	199

### 3.6.3 Nuclear magnetic resonance spectroscopy (NMR)

NMR spectra ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^1\text{H}$ - $^1\text{H}$ -COSY,  $^1\text{H}$ - $^{13}\text{C}$ -HMBC,  $^1\text{H}$ - $^{13}\text{C}$ -HSQC,  $^1\text{H}$ - $^1\text{H}$ -NOESY) were recorded on a Bruker Avance 500 spectrometer at room temperature using  $\text{CDCl}_3$  as solvent. The spectra were referenced internally using the residual protio solvent resonance relative to tetramethylsilane ( $\delta = 0$  ppm). The chemical shifts ( $\delta$ ) are quoted in ppm and coupling constants ( $J$ ) in Hertz (Hz). Multiplicities were abbreviated as follows: s (singlet), d (doublet), t (triplet), m (multiplet). Residual solvent signals were assigned by comparing the chemical shifts with literature.<sup>[166]</sup>

### 3.6.4 IR spectroscopy

IR spectra were recorded on a Bruker Vector 22 FT-IR-spectrometer with a MKII Golden Gate Single Reflection Diamond ATR-System. Absorption bands are given in whole wavenumbers  $\tilde{\nu}$  [ $\text{cm}^{-1}$ ]. The intensities are described with the following abbreviations: s (strong), m (medium), w (weak).

## 3.7 Computational methods

### 3.7.1 Docking studies

Molecular docking is a method to model a flexible substrate in the active site of a protein. It was used to simulate the substrate binding and thereby to visualize the positioning of different substrates in the active site of EmOAH to identify interesting amino acid residues for site-directed and site-saturation mutagenesis. Docking studies were performed in this work using YASARA (version: 16.12.29; YASARA Biosciences GmbH, Vienna, Austria).

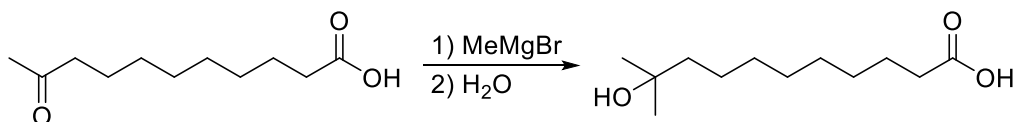
EmOAH was loaded as pdb file (4UIR) and the residual crystal water was deleted from the solved crystal structure. The enzyme was then energy-minimized. Therefore, a simulation cell was placed around the whole enzyme and the cell was entirely filled with water by the implemented energy minimization function of YASARA before the actual energy minimization was performed. A new simulation cell of 20 Å (oleic acid) or 10 Å (all other substrates) was defined around the N5 atom of the FAD cofactor (oleic acid) or the oxygen atom of E122 (for all other substrates) in the active site and the water from energy minimization was removed. PDB files of substrates were created using Chem3D Pro 14.0 and were energy-minimized separately using YASARA and loaded into the YASARA scene with the enzyme just before docking. Subsequently, the docking macro (dock\_run.mcr) was started. YASARA uses VINA and the AMBER03 force field for docking. The macro was used in default settings, which means that 25 docking runs are performed and the individual results are clustered afterwards by the algorithm to several different enzyme-substrate complexes. Docking results were visualized and processed by PyMOL (Schrödinger LLC, New York, USA).

### **3.7.2 Molecular dynamics (MD Refinement)**

After docking (3.7.1) of the substrates in the rigid environment of the protein, short molecular dynamics (MD) simulations were performed for the different enzyme-substrate complexes obtained from docking. Therefore, the macro for MD refinement (md\_refinement.mcr) which is included in YASARA was used. This tool allows the adjustment of amino acid residues after docking and thereby takes the dynamics of the active site upon substrate binding into account.<sup>[158]</sup> The MD refinement was performed for 0.5 ns and the whole protein was simulated under periodic boundary conditions in water (pH 7.4 and 298 K) using the YAMBER03 force field. Simulation snapshots were taken every 25 ps. Simulation results were visualized and processed using the MD trajectory player in YASARA (md\_play.mcr) and by PyMOL (Schrödinger LLC, New York, USA).

### 3.8 Chemical synthesis

#### 3.8.1 Synthesis of 10-hydroxy-10-methylundecanoic acid



**Figure 20:** Synthesis of 10-hydroxy-10-methylundecanoic acid from 10-oxoundecanoic acid.

In a dry Schlenk flask, 0.2 mL of a 3 M MeMgBr solution in Et<sub>2</sub>O (4.8 eq, 0.6 mmol) was added slowly on ice to 10-oxoundecanoic acid (1 eq, 0.125 mmol, 25 mg, HPLC purity: ca. 55 %) dissolved in Et<sub>2</sub>O (5 mL) before stirred at room temperature for 3 h. Subsequently, the reaction was quenched with 15 mL ice-cold water when a white precipitate was formed which was dissolved by adding 8 mL 1 M HCl. Then the solution was extracted with Et<sub>2</sub>O (3 x 10 mL) and the combined organic layer was washed with saturated NaHCO<sub>3</sub> solution and brine (10 mL each) before dried over MgSO<sub>4</sub> and finally concentrated *in vacuo* to yield the Grignard product 10-hydroxy-10-methylundecanoic acid (5 mg, HPLC purity: ca. 61 %) as viscous yellow oil which was used without further purification as standard in LC-MS analysis.

## 4 Results

Hydratases catalyzing the addition of water to non-activated isolated C-C double bonds are of special interest for organic synthesis and application in chemical industry. However, the use of hydratases is limited as suitable enzymes are missing for most relevant substrates. The identification of hydratases for desired reactions is difficult because only a few enzymes which are very specific for certain substrate classes are known so far. New enzyme activities can be found in various ways, e.g. by screening of microorganism or by using the flexible substrate scope of wild type enzymes (substrate promiscuity) including subsequent improvement by enzyme engineering.

In this work, three promising enzymes (limonene hydratase, carotenoid hydratase and oleate hydratase) of the relatively unexplored enzyme class of hydratases were identified as targets for detailed investigations and tested for their suitability in the regioselective hydration of DMDDAL to Mahonial (HDMDAL), an important fragrance compound reminiscent of lily-of-the-valley.

### 4.1 *Geobacillus stearothermophilus* & limonene hydratase

*Geobacillus stearothermophilus* BR388 is a thermophilic, spore-forming bacterium isolated from orange peel.<sup>[105]</sup> The unique ability of BR388 to grow on limonene as sole carbon source makes this strain interesting for the identification of new enzymatic activities. Chang and Oriel described the responsible limonene pathway, including a novel putative limonene hydratase (LIH). However, little information about the enzyme is known as it was not isolated and characterized and no DNA or protein sequence was reported. Therefore, the original strain was purchased for further investigations. (1.3.4)

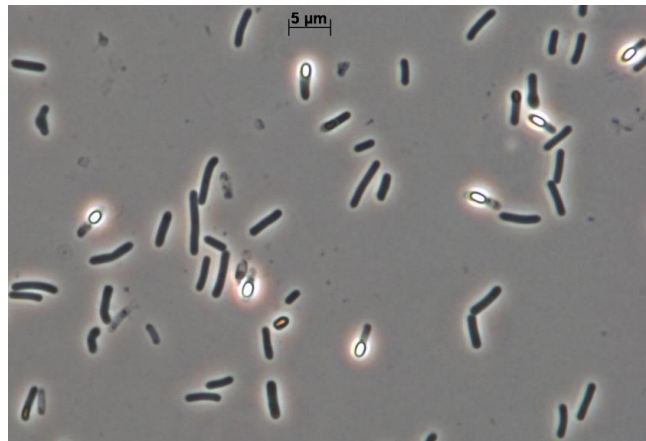
#### 4.1.1 Recovery and identification

*Geobacillus stearothermophilus* BR 388 was purchased from the ATCC collection (ATCC 55596) as liquid culture. The culture was revived according to the supplier's instructions and overnight cultures as well as an LB agar plate were inoculated with cells from the initial culture. After incubation at 55 °C, no growth was observed for the liquid cultures but some colonies were visible on the agar plate. The results are in accordance with the product specification sheet<sup>[154]</sup>, where it is stated that poor growth



in broth but heavy growth on agar is typical for *G. stearothermophilus*. Furthermore, colony variation was observed on agar as described in the product sheet. Hence, further experiments were conducted using cells from agar and the culture was preserved by continuous cultivation on LB-agar plates.

Early experiments focused on the clear identification of the *G. stearothermophilus* cells. *E. coli* and other gram-negative organisms could be ruled out by a KOH quick test which was carried out using a 3 % potassium hydroxide solution for cell lysis. While DNA strains could be extracted from *E. coli* cells, no DNA was visible in case of the putative *G. stearothermophilus* providing first evidence that the cells are gram-positive. *G. stearothermophilus* cells were studied under an optical microscope and likewise compared to *E. coli* cells. As expected, the sample showed rod-shaped bacteria forming strong refractive terminal endospores. The appearance was in accordance with that of *G. stearothermophilus* cells shown on a picture provided by R. Pukall from DSMZ (Figure 21).



**Figure 21:** Microscopic view of *G. stearothermophilus*. The figure shows the rod-shaped cells of the Gram-positive organism in sporulation. The spores are formed terminally and are clearly recognizable by the strong refraction of light (kindly provided by Dr. R. Pukall, DSMZ).

## 4.1.2 Cultivation

Further experiments including the isolation of gDNA for whole genome sequencing and whole cell biotransformations made larger quantities and consequently cultivation of *G. stearothermophilus* in liquid culture absolutely necessary. Thus, suitable growth conditions had to be found. Several literature-known conditions for cultivation of *Geobacillus* spp. have been studied.<sup>[104–106,155,167]</sup>

### 4.1.2.1 Evaluation of suitable media and conditions for liquid culture

In a first attempt, several media were tested for cultivation. Incubation (55 °C, ca. 24 h) was performed initially without shaking, as it was suspected that movement is one major difference between cultivation on agar and in liquid culture. *Nutrient broth* (5 g L<sup>-1</sup> peptone, 3 g L<sup>-1</sup> meat extract) which is recommended by ATCC and DSMZ for *G. stearothermophilus* was not appropriate as no growth was observed. The addition of 10 mg L<sup>-1</sup> MnSO<sub>4</sub> · H<sub>2</sub>O as sporulation agent led to limited growth after 24 h and a small pellet could be obtained after centrifugation. Besides complex media, the defined *DP salt* medium (1 g L<sup>-1</sup> NH<sub>4</sub>Cl, 0.5 g L<sup>-1</sup> K<sub>2</sub>HPO<sub>4</sub>, 20 mg L<sup>-1</sup> MgSO<sub>4</sub> · 7 H<sub>2</sub>O) was studied according to literature of Chang and Oriol.<sup>[105,160]</sup> The authors used this minimal medium and a Nephelo culture flask without baffles to provide a limonene saturated atmosphere. Limonene as sole carbon source, however, was not sufficient and no growth was achieved in the present study. Chang and Oriol additionally used a vitamin complex (0.4 mg L<sup>-1</sup> nicotinamide, 0.4 mg L<sup>-1</sup> thiamine and 2 µg L<sup>-1</sup> biotin) and 0.0125 % yeast extract for growth enhancement,<sup>[105]</sup> which also did not lead to growth in this case.

Taking this into account, it was further assumed that the size of the inoculum (or the amount of cells) is a critical parameter and could have been too low for the volume of the main culture.<sup>[168,169]</sup> Experiments with smaller cultures, which have been used again after incubation as inoculum, and likewise, filter paper which was added to enable the formation of a biofilm<sup>[170,171]</sup>, failed to generate a growth improvement. It was therefore concluded from the experiments that although *G. stearothermophilus* was growing, the chosen conditions are not yet optimal, especially as the cells seem to grow to a certain (albeit small) cell density, but then die off again, most likely due to a deficiency.

Because optimization of incubation conditions did not enable the cultivation of *G. stearothermophilus* BR388 in one of the tested media, the cultivation experiments were continued with tryptone soy broth (TSB, DSMZ medium 545). TSB is a medium which is used for particularly fastidious organisms. Eventually, it was essential for the prosperous cultivation to prepare the TSB medium by using “casein hydrolysate” instead of “peptone from casein” as stated in the formula of DSMZ.

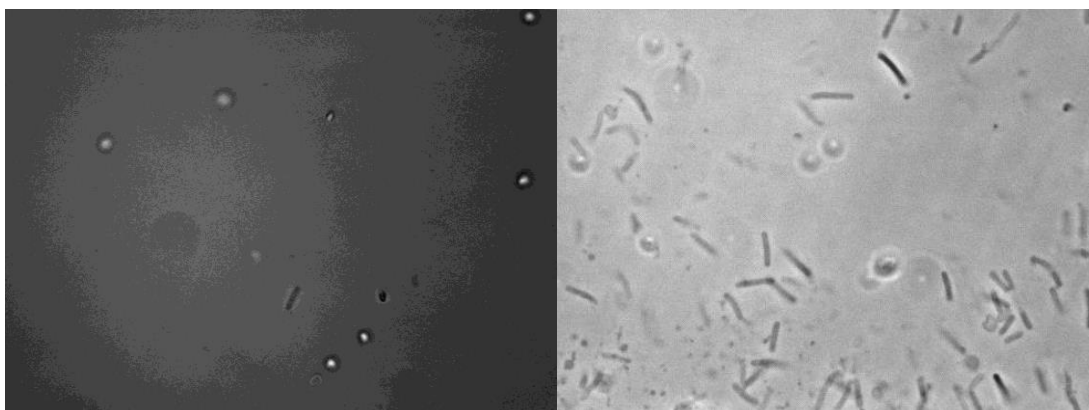
After the optimal composition of TSB was found, incubation conditions were again optimized. Despite the earlier assumption that shaking is disadvantageous, overall best results were achieved for the cultivation at 55 °C and 160 rpm in shaking flasks with big baffles as high optical densities and substantial cell mass was achieved. Cultivation under these conditions yielded sufficient cell mass for the isolation of gDNA (4.1.3) and whole cell biotransformations (4.1.5).

#### 4.1.2.2 Cultivation in TSB with different carbon sources

*Geobacillus* strains show different behavior in growth and metabolism when exposed to different sugars or glycerol as carbon source.<sup>[167]</sup> Therefore, cells were used to inoculate TSB medium containing glucose (2.5 g L<sup>-1</sup>), xylose (2.08 g L<sup>-1</sup>), galactose (2.5 g L<sup>-1</sup>) or glycerol (1 %). Growth after 20 h of incubation at 55 °C and 160 rpm was detected for all carbon sources: glucose OD<sub>600</sub> = 0.45, xylose OD<sub>600</sub> = 1.57, galactose OD<sub>600</sub> = 1.59 and glycerol OD<sub>600</sub> = 1.82. The lower OD for glucose might be due to formation of lumps and biofilm at the edge of the medium in the shaking flask. No distinction between different *Geobacillus* strains could be made based on this cultivation studies.

#### 4.1.2.3 M9 and TSB agar

The cultivation on M9 agar plates was challenging, as the volatile and cell toxic limonene had to be provided in an appropriate way so that it could be used by *G. stearothermophilus* as carbon source. The problem could be solved by using a desiccator. A small beaker with limonene was placed in the lower part of the desiccator while the agar plates were positioned above. The incubation at 55 °C provided enough limonene in the vapor phase to support growth. After two days, significant growth on M9 agar plates was visible. Negative controls cultivated outside the desiccator under otherwise identical conditions also showed some colonies but to a much lesser extent. Growth on the negative control is likely due to small amounts of exploitable carbon sources in the agar. In summary, the cultivation on M9 agar was successful and was used for identification of limonene metabolites (4.1.5.1). Cultivation on TSB agar was also successfully tested and good growth was achieved, which can be seen from the low number of spores and currently sporulating cells (Figure 22).



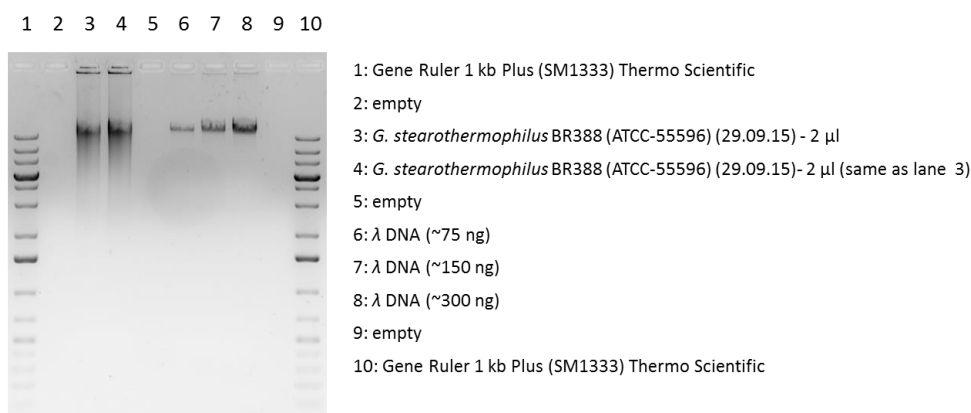
**Figure 22:** Microscopic view of *G. stearothermophilus*. *left:* Cultivation on M9 agar. The section shows mainly spores and one cell. *right:* Cultivation on TSB agar. The section shows only some spores, but mostly cells. Other areas of the preparation even had significantly more cells without spores.

### 4.1.3 Isolation of gDNA and sequencing

Genomic DNA (gDNA) was isolated with the GeneJET Genomic DNA Purification Kit (ThermoScientific) and was of sufficient quality and amount for sequencing. This was confirmed by agarose gel electrophoresis compared to defined amounts of  $\lambda$  DNA (Figure 23). Additionally, the concentration was determined via absorbance measurement (Nanodrop) to  $418 \text{ ng } \mu\text{L}^{-1}$ .

The gDNA was used to provide a clear identification of *G. stearotherophilus*. For this purpose, a highly specific PCR method according to Pennacchia *et al.* was applied.<sup>[155]</sup> This PCR targets the ITS-16S-23S rRNA (primers FITS and RITS) and produces a single PCR fragment of 302 bp in length. However, the method was unsuccessful as no PCR fragments could be obtained after several attempts, which is why classical 16S rRNA sequencing (primers F8 and 1492R) was used. This method gave fragments in the desired range of 1.5-1.6 kb which were sequenced. Subsequently, a BLAST analysis (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) showed 99 % identity with *G. stearotherophilus*, but the same probability for the closely related *G. thermoglucosidasius*.

Whole genome sequencing was performed in cooperation with Dr. Per-Olof Syrén from SciLifeLab at KTH Stockholm on a PacBio RSII system. Raw data, as well as assembled and annotated sequence data was obtained. Sequencing yielded 18 unitigs with a total number of 4'020'976 bases (GC content 43.9 %) and the N50 value was 719133. The draft genome sequence contained 4447 coding sequences (CDS), 28 rRNAs, 90 tRNAs and 72 miscRNA.

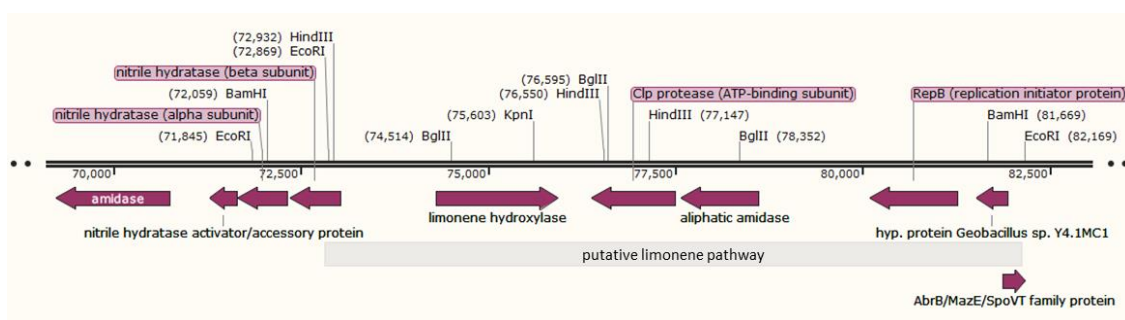


**Figure 23:** Agarose gel showing the isolated gDNA of *G. stearotherophilus* BR388 in comparison to defined amounts of  $\lambda$  DNA as a proof of purity and to estimate the concentration of the gDNA sample.

#### 4.1.4 Proof of limonene pathway and search for limonene hydratase

The existence of a putative limonene hydratase (LIH) and the limonene degrading pathway, which is responsible for the ability of *G. stearothermophilus* to grow on limonene as a sole carbon source was studied by analyzing the gDNA sequence data (this chapter) followed by biotransformations of limonene (4.1.5.1).

For analysis of gDNA, the sequence was compared with the information given in the literature<sup>[106]</sup>. The described limonene degrading pathway could be located on the genome on unitig\_6 and the *open reading frames* (ORFs) have been analyzed by BLAST (Figure 24). A limonene hydratase could not be identified. However, the limonene hydroxylase<sup>[172]</sup> and the aliphatic amidase<sup>[173]</sup>, which were described by the same authors, could be located within the putative limonene degrading pathway. In the vicinity of these genes, also a nitrile hydratase (separated in a activator / accessory protein and its  $\alpha$ - and  $\beta$ -subunits), a CLP protease (ATP-binding subunit), a RepB (replication initiator protein), an unspecified hypothetical protein from *G. stearothermophilus* Y4.1MC1 and a AbrB/MazE/SpoVT family protein are located.



**Figure 24:** Gene cluster of the putative limonene degrading pathway in *G. stearothermophilus* BR388. The pathway was found on unitig\_6 using information from the literature and is in accordance with the described restriction digest pattern.<sup>[106,172,173]</sup> The pathway contains the literature-known limonene hydroxylase and aliphatic amidase, which were described by the same authors as the entire pathway. However, no distinct limonene hydratase was identified. Investigated ORFs were assigned by BLAST as (from left to right): amidase, nitrile hydratase (activator / accessory protein, alpha-subunit, beta-subunit), limonene hydroxylase, Clp protease (ATP-binding subunit), aliphatic amidase, RepB (replication initiator protein), hypothetical protein found in *Geobacillus* sp. (protein predicted by sequencing without experimental evidence for *in vivo* expression whose function cannot be assigned yet) and AbrB/MazE/SpoVT family protein.

## 4.1.5 Biotransformations

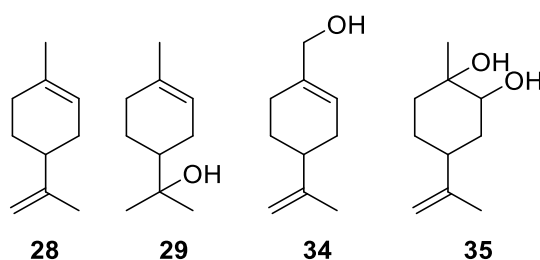
### 4.1.5.1 Analysis of limonene metabolites

After successful cultivation of *G. stearothermophilus*, different experiments on the transformation of limonene were performed.

In one experiment, cells were grown at 55 °C for 24 h followed by the addition of 0.1 %  $V V^{-1}$  limonene to the culture. Further incubation took place at room temperature due to the flash point (48 °C) and high volatility of limonene. After additional 24 h, the culture was extracted with ethyl acetate and cyclohexane. The analysis of the extract with GC-FID and GC-MS in comparison to authentic monoterpene standards proofed the presence of traces of limonene diol **35**.

In another experiment, *G. stearothermophilus* cells were grown on a M9 minimal agar plate in (*R*)-limonene atmosphere (in a desiccator, 55 °C, 4 d). For the analysis of the metabolites, the agar plate was first rinsed with 5 mL of water and the aqueous phase was extracted with 5 mL of MTBE. Subsequently, the agar plate was rinsed off with 5 mL of MTBE and 5 mL of water were added to the organic phase in order to obtain a clear phase separation during the extraction. Finally, the plate was rinsed again with 5 mL of water, but the grown cells were scraped off the agar surface with a pipette tip and rinsed with a little MTBE. Again, the water phase was extracted with MTBE.

From all three samples, 100  $\mu$ L of the organic phase was analyzed by GC-FID and GC-MS. The pure substances limonene **28**,  $\alpha$ -terpineol **29**, perillyl alcohol **34** and limonene diol **35** ([1*S*, 2*S*, 4*R*]-(+)-limonene diol) served as references. Based on comparison to the standards and the fragmentation pattern, high concentrations of limonene **28** could be detected. Furthermore, limonene diol **35** was found as a metabolite, while perillyl alcohol **34** and  $\alpha$ -terpinol **29** could not be detected. The chromatograms also showed many other substances that could not be assigned.



#### 4.1.5.2 Biotransformation of linear terpenoids

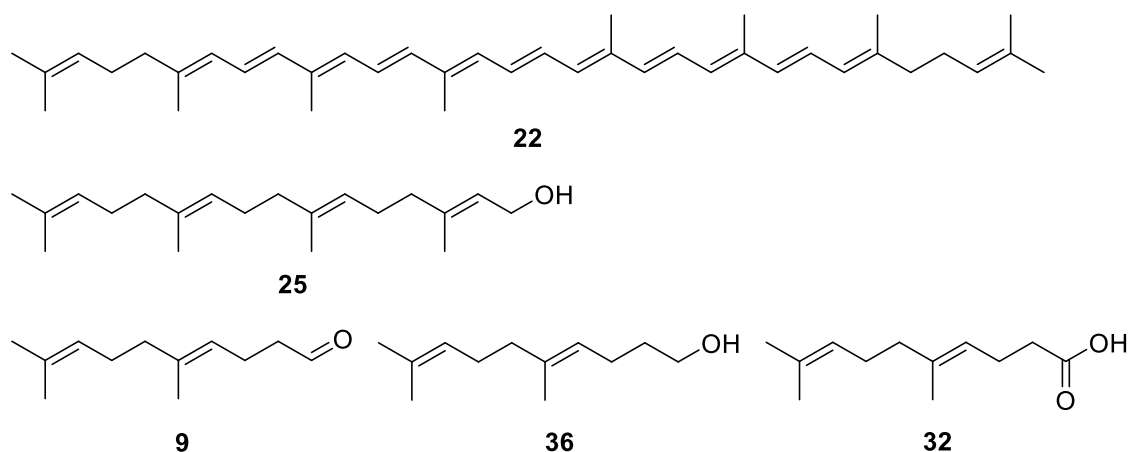
*G. stearothermophilus* cells which were grown in TSB medium were used for biotransformation of DMDDAL, DMDDAC, DMDDOL, citral and linalool, all potential substrates for alkene hydration (conditions: 500  $\mu\text{M}$  substrate and 50  $\text{g}_{\text{cww}} \text{L}^{-1}$  whole cell catalyst in a total volume of 500  $\mu\text{L}$  citrate buffer pH 6.0, 30  $^{\circ}\text{C}$ , 600 rpm, 1-2 days reaction time). After extraction with ethyl acetate, the samples were analyzed by GC-FID and GC-MS. No hydration products could be detected under the tested conditions.

As no LIH could be identified on the gDNA and no hydration was detected in biotransformations of limonene or linear terpenoids, we planned to clone the identified ORFs (4.1.4) in *E. coli* for heterologous protein production and further investigations. However, the project was not further pursued due to time limitations.



## 4.2 Carotenoid hydratase

Carotenoid-1,2-hydratases (CrtC) catalyze the regioselective hydration of the terminal prenyl double bonds of lycopene **22** and related compounds. This makes them promising candidates for the hydration of DMDDAL **9** and related compounds in Mahonial synthesis (Figure 25). (1.3.3)



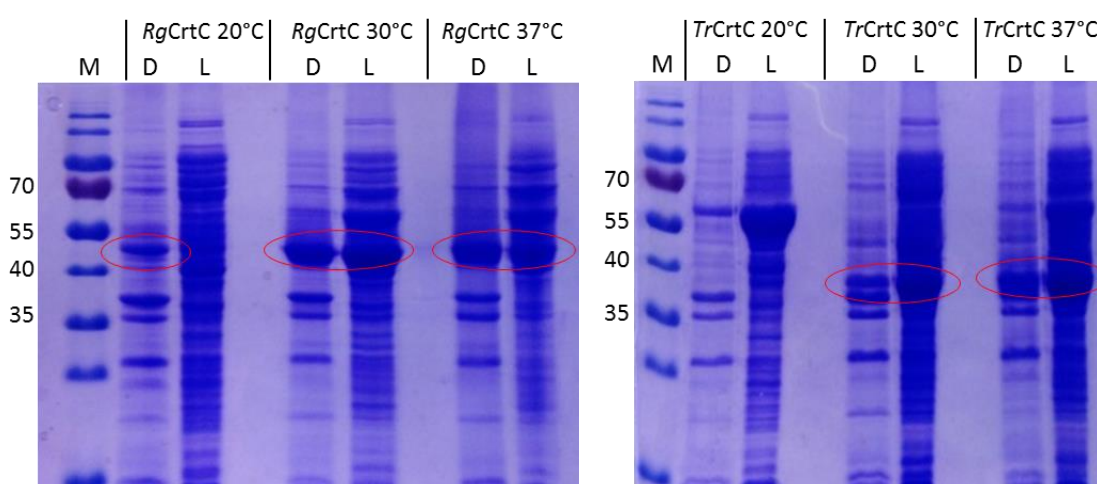
**Figure 25:** Potential substrates for carotenoid-1,2-hydratases (CrtCs). Lycopene **22**, the natural substrate, is hydrated at both terminal prenyl double bonds. The shorter diterpenoid geranylgeraniol **25** was also described as a substrate for CrtCs.<sup>[96]</sup> Therefore, DMDDAL **9**, DMDDOL **36** and DMDDAC **32** were assumed to be additional promising substrates.

CrtCs from *Rubrivivax gelatinosus* (RgCrtC) and *Thiocapsa roseopersicina* (TrCrtC) were chosen for investigations as they are known to be functionally expressed in *E. coli*.<sup>[96]</sup> Both enzymes were cloned, produced by heterologous expression (4.2.1) and assessed in biotransformations using lycopene **22** and shorter linear terpenoids (4.2.2).

#### 4.2.1 Cloning and expression of TrCrtC and RgCrtC in *E. coli* ITB 94

The hydratases TrCrtC and RgCrtC were provided by M. Fischer (ITB) with an N-terminal His<sub>6</sub>-tag in pET28a(+) vectors. However, Hiseni *et al.*<sup>[96]</sup> described that a 6 kDa proline-rich membrane anchor is split off from the N-terminus of TrCrtC (44 kDa) during purification and storage. Thus, the His<sub>6</sub>-tag of TrCrtC was moved to the C-terminus during cloning by Gibson Assembly and the successful construction of pDHE vectors carrying TrCrtC and RgCrtC was confirmed by Sanger sequencing. The switch from pET28a(+) to pDHE vector enabled heterologous expression using the ADH deficient *E. coli* ITB 94 strain. Besides the knockout of two ADHs, ITB 94 also possesses chaperones for enhanced functional expression and is suitable for induction with rhamnose, as the rhamnose degradation is switched off. The CrtC production was conducted in TB medium and induced by rhamnose (final concentration 0.5 g L<sup>-1</sup>). Three incubation temperatures were compared (20 °C, 30 °C or 37 °C) to find a suitable temperature with high protein yields (Figure 26).

The expression was successful for both CrtCs. For RgCrtC, the highest amount was produced at an incubation temperature of 30 °C and 37 °C. Significantly worse results were achieved at 20 °C. TrCrtC production could only be achieved at 30 °C and 37 °C in comparable amounts. Only very faint bands can be seen for the TrCrtC with membrane anchor (about 45 kDa), while strong bands of the mature protein are visible at about 38 kDa both in the cell debris and in the lysate.

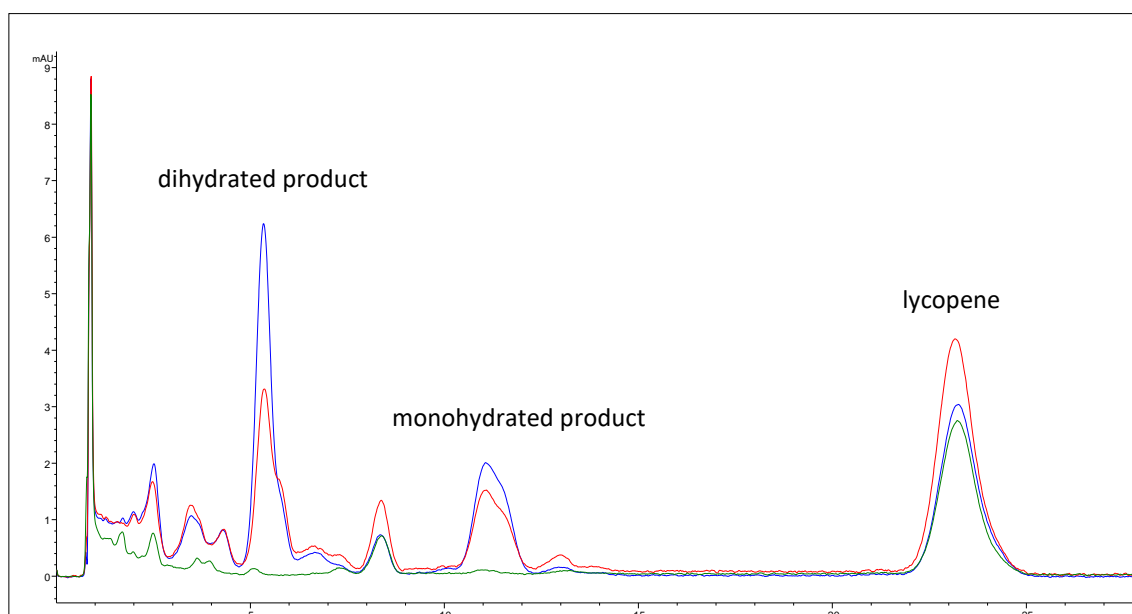


**Figure 26:** SDS-PAGE of heterologous expression of RgCrtC and TrCrtC. Cell debris (D) and lysates (L) are shown, compared to a size standard (M). The protein bands of the hydratases are marked with red circles. The observed protein sizes are in line with the expectations for RgCrtC (44.7 kDa, 47 kDa with His<sub>6</sub>-tag) and TrCrtC (44.1 kDa, 45.3 kDa with His<sub>6</sub>-tag). TrCrtC has been described (Hiseni *et al.*)<sup>[96]</sup> as active enzyme fragment with a truncated size of 38 kDa, which can be seen for expression at 30 °C and 37 °C.

## 4.2.2 Biotransformations

### 4.2.2.1 Lycopene

Biotransformations with the natural substrate lycopene were performed under different conditions with both carotenoid hydratases, TrCrtC and RgCrtC. Because detergents in biotransformations significantly increase enzyme activity, several detergents and detergent concentrations were tested. In contrast to literature<sup>[96]</sup> where phosphatidylcholine was used, successful hydration could only be observed by addition of Triton X-100. Biotransformations of 20  $\mu\text{M}$  lycopene with cell lysate (final total protein concentration: 25  $\text{g L}^{-1}$ ) and detergent concentrations of 0.5 % and 1 % led to a product formation of about 20 % monohydrated product (rhodopin) **23** and about 20-35 % for the dihydrated lycopene **24** (Table 40). The values are given as percentage of peak area in HPLC analysis. Higher conversions were observed for RgCrtC (about 50-60 %) compared to TrCrtC (about 40 %). Interestingly, the higher conversion of RgCrtC led mainly to an increased formation of the dihydration product, while the amount of the monohydrate remained almost the same (Figure 27). No obvious difference was observed for the biotransformations of lycopene with RgCrtC for 20 h and 24 h.



**Figure 27:** Exemplary HPLC chromatograms of lycopene biotransformations. Depicted are the spectra measured at a wavelength of 476 nm (VIS). Formation of the monohydrated product rhodopin **23** (retention time 11 min) and the dihydrated product **24** (5.5 min) was observed in biotransformations of lycopene (23 min) by RgCrtC (blue) and TrCrtC (red) while the peaks were absent in control experiments without lysate. The peaks were identified by mass using LC-MS.

**Table 40:** Results of biotransformations of 20  $\mu\text{M}$  lycopene (from 1 mM stock in acetone) using cell lysate (final concentration: 25  $\text{g L}^{-1}$ ) of *E. coli* ITB 94 cells containing overexpressed wild-type CrtCs after 20 h and 24 h. MTBE extracts of the reaction were analyzed by HPLC. Peak area values are given in percent for the substrate and the mono- and the dihydration products.

enzyme	time [h]	Triton X-100 [%]	lycopene	area [%]	
				monohydrate	dihydrate
RgCrtC	20	0.5	48 $\pm$ 2	21 $\pm$ 1	31 $\pm$ 1
		1	42 $\pm$ 1	22 $\pm$ 1	36 $\pm$ 0
	24	0.5	45 $\pm$ 5	23 $\pm$ 3	33 $\pm$ 2
		1	49 $\pm$ 3	20 $\pm$ 1	31 $\pm$ 2
TrCrtC	24	0.5	58 $\pm$ 2	22 $\pm$ 3	20 $\pm$ 1
		1	60 $\pm$ 0	18 $\pm$ 1	22 $\pm$ 1

Biotransformations with lycopene (20  $\mu\text{M}$ ) were repeated for cell lysate (20  $\text{g L}^{-1}$ ) of *E. coli* ITB 94 cells containing RgCrtC with only 4 h reaction time and compared to 24 h reactions. The conversions were incomplete in both cases as between 38 % and 66 % lycopene remained in the biotransformations (Table 41). However, the results show similar product formation already after 4 h. Especially for the monohydrate, which is formed in amounts of about 20-25 %, the results are in the same range considering the standard deviation. Thus, the hydration of lycopene could be an equilibrium reaction. Only the formation of the dihydration product seems to be increased after 24 h, particularly when using 1 % Triton X-100 as detergent because 39 % dihydrate is formed. This is a considerable increase compared to 14 % and 17 % (after 4 h using 0.5 % and 1 % Triton X-100, respectively) and 21 % (also after 24 h but with only 0.5 % detergent).

These experiments confirmed activity and functional expression of both CrtCs which were subsequently screened with non-natural substrates.

**Table 41:** Results of biotransformations of 20  $\mu\text{M}$  lycopene (from 1 mM stock in acetone) by lysate (final concentration: 20  $\text{g L}^{-1}$ ) of *E. coli* ITB 94 cells containing overexpressed RgCrtC after 4 h and 24 h. MTBE extracts of the reaction were analyzed by HPLC. Peak area values are given in percent for the substrate and the mono- and the dihydration products.

enzyme	time [h]	Triton X-100 [%]	lycopene	area [%]	
				monohydrate	dihydrate
RgCrtC	4	0.5	63 $\pm$ 5	23 $\pm$ 2	14 $\pm$ 3
		1	66 $\pm$ 7	17 $\pm$ 2	17 $\pm$ 5
	24	0.5	54 $\pm$ 7	25 $\pm$ 2	21 $\pm$ 5
		1	38 $\pm$ 2	24 $\pm$ 1	39 $\pm$ 1

#### 4.2.2.2 Shorter terpenoids

Biotransformations of geranylgeraniol **25** (Figure 25) with cell lysate of *E. coli* strain ITB 94 containing either RgCrtC or TrCrtC were conducted at different conditions. Analysis by GC-MS revealed no new peaks of potential products. Surprisingly, it was also not possible to recover the substrate geranylgeraniol **25** in some cases, as no peak with the retention time of the authentic geranylgeraniol **25** standard was observed in the biotransformation samples. Biotransformation of DMDDAL **9**, DMDDOL **36** and DMDDAC **32** (Figure 25) were also not successful under the tested conditions.

Despite the successful expression in *E. coli* and the hydration of the natural substrate lycopene, it was not possible to proof the capability of CrtCs to hydrate shorter substrates. Therefore, this subproject was stopped.

### 4.3 Construction of empty vector pDHE-MCS

A pDHE vector without a hydratase or any other enzyme under control of the rhamnose promotor was constructed for use as negative control. Such a vector was not available from the ITB in-house plasmid collection. Therefore, the hydratase gene was replaced by the *Multiple Cloning Site* of pET28a(+) by Gibson Assembly (3.2.3). Additionally, the possibility was retained to use both an N- and a C-terminal His<sub>6</sub>-tag which also allows the use of the vector for classical molecular biological work including restriction-ligation cloning methods. The successful cloning was confirmed by Sanger sequencing and the plasmid was named pDHE-MCS.

## 4.4 Oleate hydratase

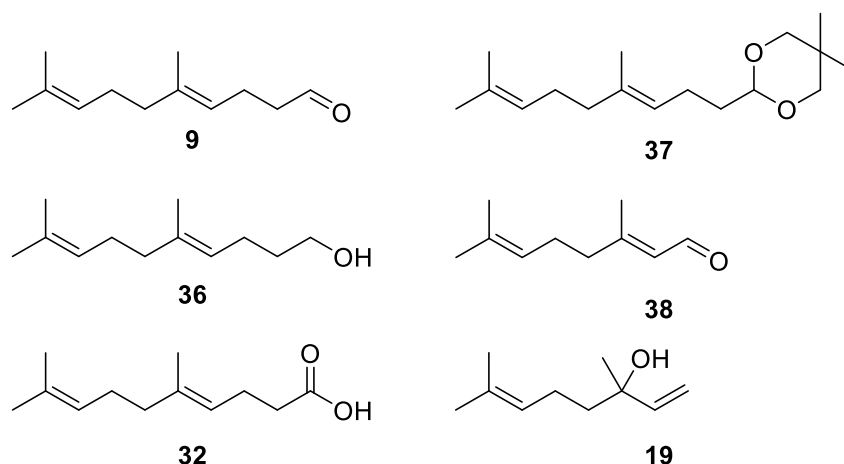
Oleate hydratases (OAHs) belong to the large group of fatty acid hydratases and are by far the most studied enzymes capable of adding water to isolated C-C double bonds. OAHs catalyze the hydration of *cis*-9-double bonds in unsaturated fatty acids leading to the corresponding 10-hydroxy fatty acid as, for instance, in the conversion of oleic acid to 10-hydroxystearic acid. In recent years, OAHs aroused interest in biocatalysis, as it was demonstrated that they also accept considerable shorter fatty acids.<sup>[115]</sup> Furthermore, it has even been demonstrated that OAHs can be applied in the asymmetric hydration of terminal and internal short-chain alkenes in the presence of a carboxylic acid decoy molecule.<sup>[130]</sup> (1.3.5)

Here, the enzyme family of OAHs was investigated for the regioselective hydration of highly substituted alkenes to generate tertiary alcohols. The particular focus was on the regioselective hydration of terpenoids such as DMDDAL and related compounds (4.4.1). As these experiments revealed initial activity for DMDDAC, this substrate was used to screen several homologous OAHs from our in-house collection (4.4.2). In further experiments, the OAH from *Elizabethkingia meningoseptica* (EmOAH) was engineered. *In silico* docking studies (4.4.4) revealed target sites for mutational studies (4.4.5) which was followed by directed evolution using ISM (4.4.8). A preparative biotransformation of DMDDAC with the engineered EmOAH variant enabled HDMDAC synthesis, the carboxylic acid analog of Mahonial, which was used for spectroscopic characterization of the product (4.4.11).

### 4.4.1 Preliminary experiments with EmOAH

Initial experiments were performed using EmOAH as biocatalyst. The enzyme was produced according to established protocols from our institute.<sup>[115,127]</sup> Biotransformations of DMDDAL **9**, DMDDOL **36**, DMDDAC **32**, DMDDIOX **37**, citral **38** and linalool **19** (Figure 28) were performed using different reaction conditions. While no activity was measurable for most of the studied substrates, traces of hydration product could be identified in the conversion of DMDDAC **32** using the conditions described by Demming *et al.* (50 g L<sup>-1</sup> *E. coli* whole cell catalyst containing overexpressed EmOAH, 100 mM glucose, 2 mM DTT, 3 mM NADH, 0.3 mM FAD; 500  $\mu$ M substrate)<sup>[127]</sup>. Side reactions caused by ADHs were observed for biotransformation of DMDDAL **9** with *E. coli* BL21 (DE3). Protein purification led to a

catalytically active EmOAH isolate for the hydration of oleic acid, but still no hydration of DMDDAL **9** could be detected.

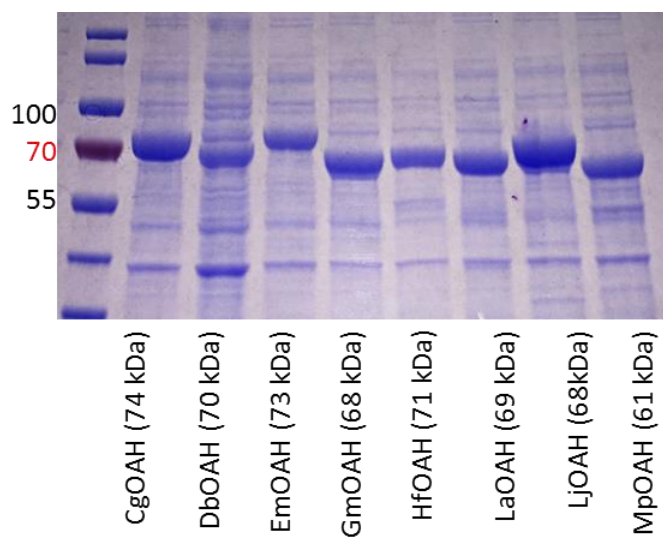


**Figure 28:** Substrates tested in preliminary experiments with the oleate hydratase from *Elizabethkingia meningoseptica* (EmOAH): DMDDAL **9**, DMDDOL **36**, DMDDAC **32**, DMDDIOX **37**, citral **38** and linalool **19**.

#### 4.4.2 Screening of oleate hydratase homologs

As EmOAH showed initial activity in the hydration of DMDDAC, related OAHs were screened to determine their substrate promiscuity. Wild type oleate hydratases from *Chryseobacterium gleum* (Cg), *Desulfomicrobium baculatum* (Db), *Elizabethkingia meningoseptica* (Em), *Gemella morbillorum* (Gm), *Holdemania filiformis* (Hf), *Lactobacillus acidophilus* (La), *Lactobacillus johnsonii* (Lj) and *Mucilaginibacter paludis* (Mp) were selected from our in-house collection and screened with DMDDAC as substrate. As no activity was found for the other substrates with EmOAH, we proposed that the carboxyl moiety as present in DMDDAC is essential for substrate acceptance of fatty acid hydratases (4.4.1).

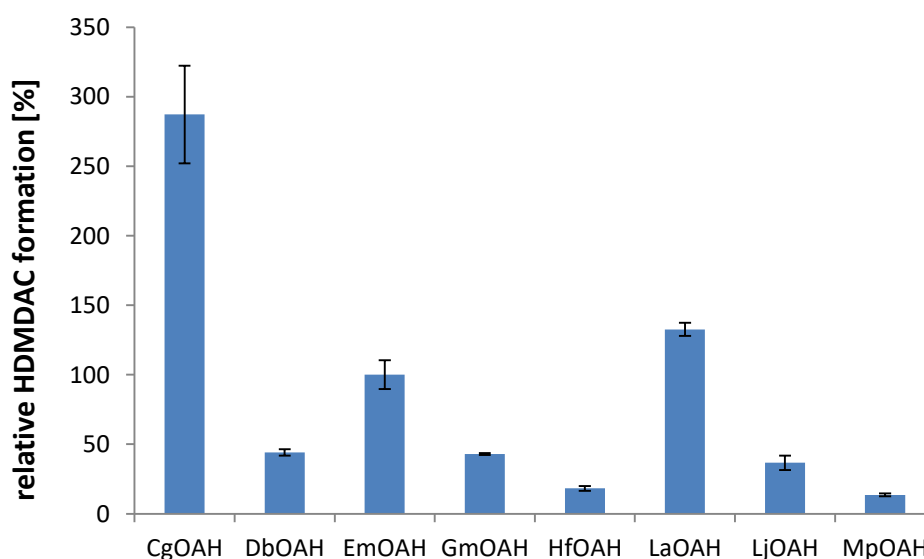
The enzymes were produced in *E. coli* BL21 (DE3) and the expression levels were verified by SDS-PAGE (Figure 29). All hydratases were produced in high amounts, whereas CgOAH and LjOAH had slightly higher expression levels.



**Figure 29:** SDS-PAGE of whole cell samples collected after expression of OAH WT enzymes just before harvesting. Despite all OAHs were expressed, this allows no conclusion of functional and soluble expression.

Biotransformations revealed that almost every tested OAH converted DMDDAC to the corresponding tertiary alcohol. The highest amount of HDMDAC was produced by cells containing CgOAH (287 % relative to EmOAH), followed by LaOAH (133 %) and EmOAH. DbOAH (44 %), GmOAH (43 %) and LjOAH (37 %) showed less conversion (Figure 30). HfOAH (18 %) and MpOAH (14 %) can be considered as almost inactive as the amount of formed HDMDAC was in the range of the standard deviation of the negative control without cells. Taken into account that these initial reactions are not normalized by enzyme concentration, the results should be interpreted with caution.





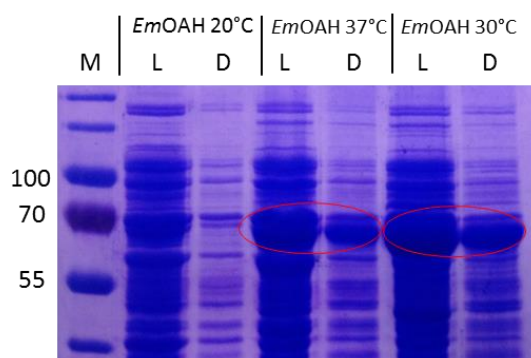
**Figure 30:** Results of the oleate hydratase wild type screening. Depicted is the formed amount of HDMDAC (analytical yield in LC-MS) relative to EmOAH as model OAH. Traces of the product are also detected in the negative control (buffer control). The area values of the negative control were subtracted from the area values of experiments.

While CgOAH showed the highest activity in the initial screening, EmOAH was selected for further enzyme engineering experiments as EmOAH is well studied and a crystal structure (Engleder *et al.*)<sup>[120]</sup> is available. Structural information is essential for a rational enzyme engineering approach which was considered to be more feasible than random mutagenesis for the envisioned directed evolution.

#### 4.4.3 Cloning and expression of EmOAH in *E. coli* ITB 94

While the initial experiments (4.4.1) and the screening of OAH homologs (4.4.2) were performed using *E. coli* BL21 (DE3) as host which reduces the aldehyde moiety of DMDDAL, more detailed studies were envisioned with an ADH-deficient *E. coli* to avoid these problems in future experiments with aldehydes. The use of the ADH-deficient *E. coli* strain ITB 94 made it necessary to clone the gene sequence coding for EmOAH from pET28a(+) into a pDHE vector (4.2.1). Cloning via Gibson Assembly (3.2.3) was successful and the correct sequence was confirmed by Sanger sequencing. EmOAH production in *E. coli* ITB 94 cells was conducted in TB medium and induced by rhamnose (final concentration 0.5 g L<sup>-1</sup>). Three incubation temperatures were compared (20 °C, 30 °C or 37 °C) to find a suitable temperature with high protein yields

(Figure 31). EmOAH (ca. 75 kDa) was produced at all tested temperatures. The highest amount of EmOAH was obtained at 30 °C and 37 °C. All following experiments were performed with *E. coli* ITB 94 cells with EmOAH expressed at 30 °C.

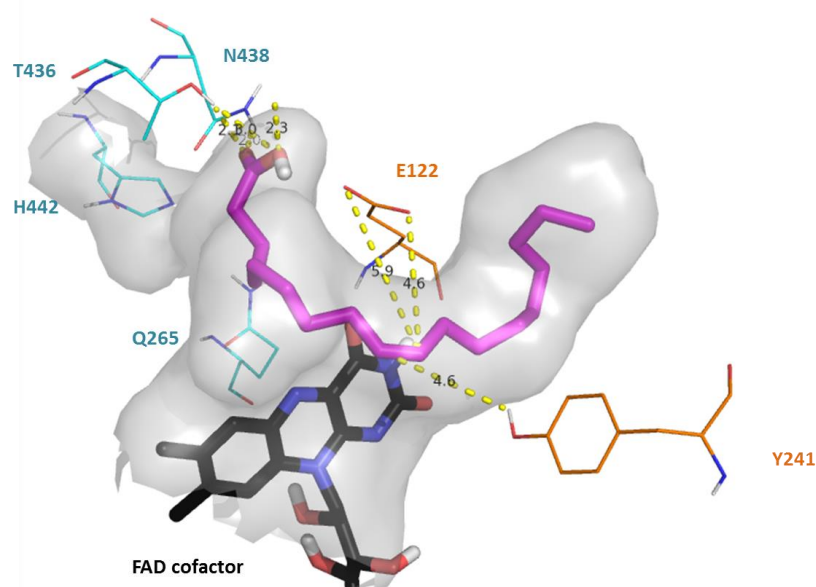


**Figure 31:** SDS-PAGE of cell debris (D) and lysate (L) of EmOAH production in *E. coli* ITB 94 cells at different temperatures compared to a size standard (M). EmOAH was found at the expected size of about 75 kDa (protein bands marked in red).

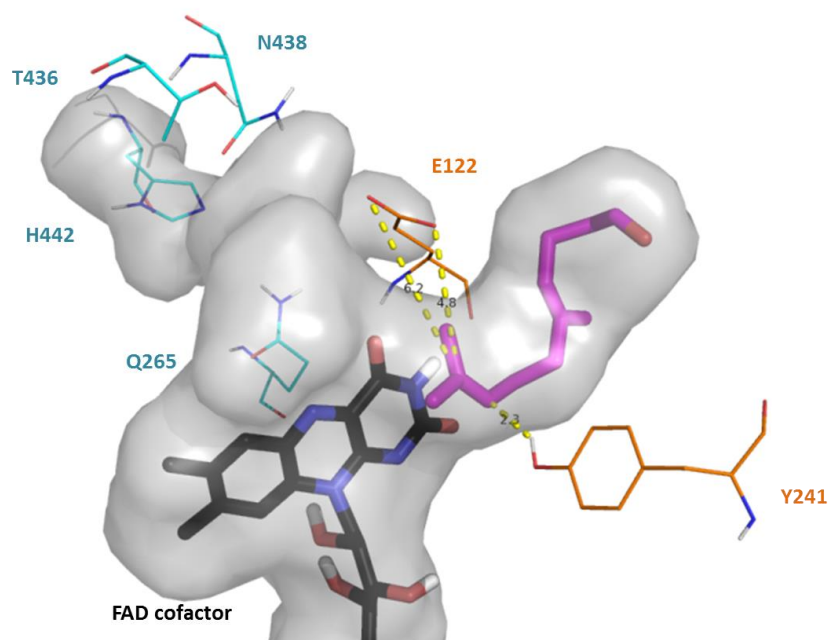
#### 4.4.4 *In silico* docking studies

Docking studies with EmOAH using YASARA were performed to find interesting amino acid residues for construction of a focused mutant library (3.7.1). YASARA allows docking of flexible substrates (ligands) into rigid enzymes (receptor). However, it should be noted that YASARA works with the crystal structure of the enzyme. This is a limitation, as it means that protein dynamics are not considered and only the substrate is kept flexible, although the substrates and the protein structure were energy-minimized prior to docking. YASARA calculates the binding energy and the dissociation constant and additionally reports amino acid residues which have contact to the docked substrate. In default settings, YASARA analyzes 25 possibilities (“runs”), which are then summarized based on similar energies to so-called clusters. To limit the range of possibilities (especially attachment to the protein surface) the simulation cell was chosen only around the active site of EmOAH (PDB: 4UIR).

Dockings were performed for *E*-DMDDAL, *E*-DMDDOL, *E*-DMDDAC, *E*-DMDDIOX, (*Z*)-undec-9-enoic acid and 10-methylundec-9-enoic acid. For comparison and as positive control, the natural substrate oleic acid was docked into EmOAH. The docking of oleic acid was successful as the first cluster was in accordance with the docking results published by Engleder *et al.*<sup>[120]</sup> (Figure 32).



**Figure 32:** Docking of the natural substrate oleic acid (purple) into the active site of EmOAH. Catalytically active amino acids E122 and Y241 (according to Engleder *et al.*)<sup>[120]</sup> are depicted in orange. The amino acid residues N438, T436, H442 and Q265 (cyan) have been described in literature as essential for the stabilization of the carboxyl group.<sup>[120]</sup> The FAD cofactor (black) presumably does not participate in the reaction and has only a structure-stabilizing function. Yellow dashed lines indicate the distances between the amino acid residues and the double bond or the carboxyl group, respectively.



**Figure 33:** Docking of *E*-DMDDAL (purple) into the active site of EmOAH. In the depicted binding mode (cluster 5 of 11, sorted by decreasing energy), *E*-DMDDAL lays in the active site in a reverse orientation compared to the most likely structure of oleic acid. However, this is just one of many options, which was issued by YASARA.

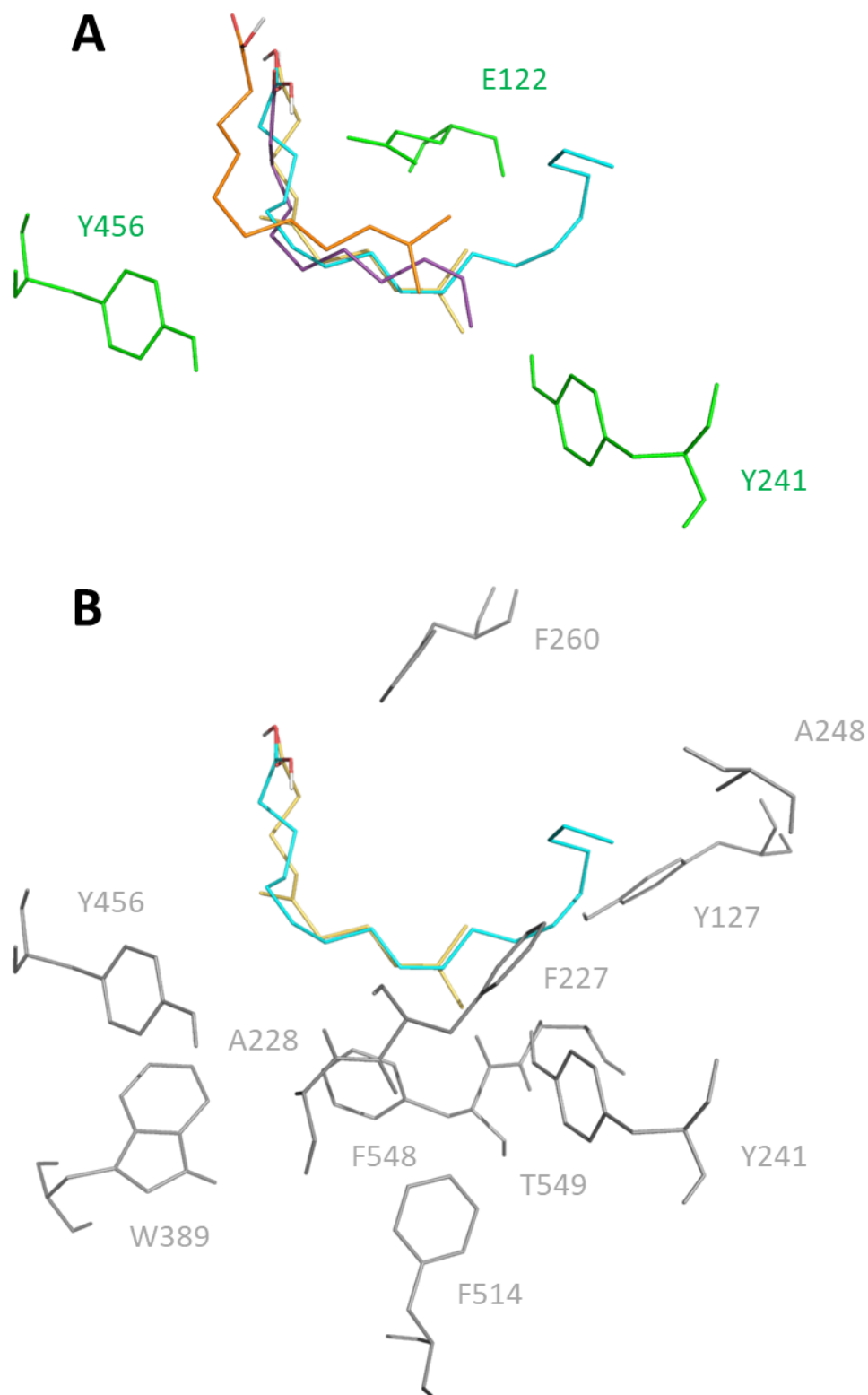
For all substrates except for *E*-DMDDIOX, YASARA found at least one or more cluster with the substrate located in the active site under the chosen simulation conditions. In general (independent of the docked substrate), the results showed many binding modes with the substrate fully or partially outside the active site. This can be a hint that shorter and branched (terpenoid) substrates are not as efficiently bound in the narrow substrate pocket of the enzyme as the natural substrate oleic acid. Further, even if the substrate was docked completely into the active site, several different binding modes and locations of the docked substrate were suggested by YASARA: at the far end where the alkyl part of oleic acid is bound, at the bend near to the catalytic amino acid residues and near the entrance with the carboxyl binding site. Interestingly, docking of *E*-DMDDAL gave also reversed binding modes, which means that the carbonyl function is bound inversely in the active site compared to the natural substrate oleic acid (Figure 33). However, it is not possible to make a statement how likely such a binding mode is only based on the determined energies. Therefore, the docking runs were considered individually (not clustered) and independently of the binding energies. Moreover, account was taken of the orientation of the substrate and the mandatory requirement of the double bond in the vicinity of the catalytically active amino acids. In this regard, plausible results were obtained for (*Z*)-undec-9-enoic acid, 10-methylundec-9-enoic acid and *E*-DMDDAC.

#### **4.4.5 Identification of amino acid residues for saturation mutagenesis by comparison of docking results**

In order to identify promising amino acid residues for mutagenesis studies, the docking results of oleic acid, (*Z*)-undec-9-enoic acid, 10-methylundec-9-enoic acid and *E*-DMDDAC were compared. Therefore, contacting amino acids (residues closer than 4 Å to the docked substrate) were listed and the residues were located in the active site. While the YASARA cluster with the highest binding energy (“yob 1”) could be used in case of the docking of the natural substrate oleic acid, single docking runs issued by YASARA were used for (*Z*)-undec-9-enoic acid (run 21), 10-methylundec-9-enoic acid (run 19) and *E*-DMDDAC (run 12). YASARA sorts the runs by decreasing binding energy. However, the runs were not selected for the binding energy, but for the reasonable orientation compared to the docking of oleic acid (Table 42 and Figure 34). Only runs with contact to the proposed catalytic amino acid Y241 were selected. The focus was on the comparison between the dockings of oleic acid and *E*-DMDDAC.

**Table 42:** Comparison of the docking results of oleic acid, (*Z*)-undec-9-enoic acid, 10-methylundec-9-enoic acid and *E*-DMDDAC. Listed are the contacting amino acids of the individual docking runs or cluster with binding energy and dissociation constant. A green field indicates that the amino acid residue had contact to the respective substrate; a red field indicates no contact to the substrate. Amino acid residues marked in bold were selected for mutagenesis and the column “mutants” shows the amino acids into which the residues were exchanged for the initial mutant library. E122 and Y241 have been described as catalytic amino acids while T436 and N438 are essential for binding of the carboxylate.<sup>[120]</sup>

docking	oleic acid	( <i>Z</i> )-undec-9-enoic acid	10-methylundec-9-enoic acid	<i>E</i> -DMDDAC	mutants
RUN/YOB	YOB 1	RUN 21	RUN 19	RUN 12	
binding energy [kcal mol <sup>-1</sup> ]	6,9780	3,8230	4,7330	5,5030	
dissoc. const. [pM]	7674643	1576535808	339363392	92522536	
GLY 119					
GLY 120					
ARG 121					
GLU 122		CATALYTIC AMINO ACID, water activation			
MET 123					
ASP 124					
<b>TYR 127</b>					S, C, A, W
THR 225					
MET 226					
<b>PHE 227</b>					A, W
<b>ALA 228</b>					F, W
<b>TYR 241</b>		CATALYTIC AMINO ACID, protonation			A
PHE 245					
<b>ALA 248</b>					L
LEU 252					
LEU 255					
LEU 258					
<b>PHE 260</b>					W
<b>TRP 389</b>					F, L, A
GLY 422					
ILE 423					
THR 436		CARBOXYLATE BINDING			
ASN 438		CARBOXYLATE BINDING			
TRP 454					
<b>TYR 456</b>					A
<b>PHE 514</b>					A
<b>PHE 548</b>					A, W
<b>THR 549</b>					A, Y, W
FAD 701					



**Figure 34:** A) Overlay of docking results for oleic acid (cyan), (*Z*)-undec-9-enoic acid (purple), 10-methylundec-9-enoic acid (orange) and *E*-DMDDAC (yellow). Polar amino acid residues E122, Y241 and Y456 which are close to the double bond at the bend of the active site, are depicted in green. E122 and Y241 are assumed to be essential for catalysis (according to Engleder *et al.*<sup>[120]</sup>). B) Overlay of the docking results for oleic acid (cyan) and *E*-DMDDAC (yellow) into the active site of EmOAH. The mutated amino acids which are targeted in the creation of an initial library are shown in grey.

In total 28 amino acid residues were found as relevant by comparison of contacting amino acids from the docking results. Interestingly, docking revealed additional amino acids, namely G119, L255, W389 and F514 which do not have contact to oleic acid. From the identified amino acids, several residues were excluded from the mutational study. G119 and L255 were not considered as they were not found in the docking of DMDDAC and due to their proximity to the carboxylate binding site. E122 was excluded as it is part of the proposed catalytic machinery.

The residues W389 and F514 were selected for mutagenesis, as both are located in areas where the terpenoid DMDDAC possesses additional methyl groups (compared to the natural substrate oleic acid). Smaller amino acids were proposed to be beneficial for substrate binding and therefore the mutants W389F, W389L and W389A were supposed to increase activity.

Other main differences between the dockings of oleic acid and the shorter substrates were mainly found for the residues in the area where the alkyl chain (more precisely the carbon atoms C11-C18) of oleic acid is located. The alkyl pocket is made up of two different loops with the oleic acid contacting residues M123, D124, Y127 being part of a loop region located under the plane of the double bond while F245, A248, L252 and L258 are part of a loop or  $\alpha$ -helix above the double bond. From both groups of residues, one position each was chosen for mutagenesis. Y127 was favored over M123 and D124 to prevent changes in the direct vicinity of E122. The polar tyrosine at position 127 was exchanged to other polar but smaller (Y127C, Y127S), smaller (Y127A) and bigger (Y127W) amino acids. Out of the second group, A248 was selected as it was found by Demming *et al.* to be a specificity determining position (SDP).<sup>[130]</sup> The size of residues introduced at position A248 has been shown to correlate with the size of accepted substrates. Therefore, A248 was also included in the initial mutagenesis study, with leucine (A248L) as a slightly larger residue for the initial library (4.4.6). Other target positions to generate an initial library were mainly located around the double bond and chosen to screen for their influence on substrate binding and catalysis. In general, the residues were exchanged to alanine (A) when smaller or tryptophan (W) and phenylalanine (F) when larger amino acids were predicted to be beneficial. Overall, the following mutants were created additionally to the ones mentioned above: F227A, F227W, A228F, A228W, F260W, F548A, F548W, T549A, T549Y and T549W.

In addition, Engleder *et al.* reported mutational studies on the three polar residues E122, Y241 and Y456 at the bend of the active site.<sup>[120]</sup> The variants E122A and Y241F were reported to be completely inactive, while the mutant Y456F was even slightly more active than the wild type. However, the authors mentioned that some activity was retained for Y241F when FAD was photoreduced. Therefore, we designed the mutant Y456A as potentially beneficial catalyst and generated Y241A as a knockout (negative control).

As a result of these docking studies, eleven positions were identified and a total of 20 beneficial variants plus one knockout variant were postulated by this rational enzyme design approach (Figure 34).

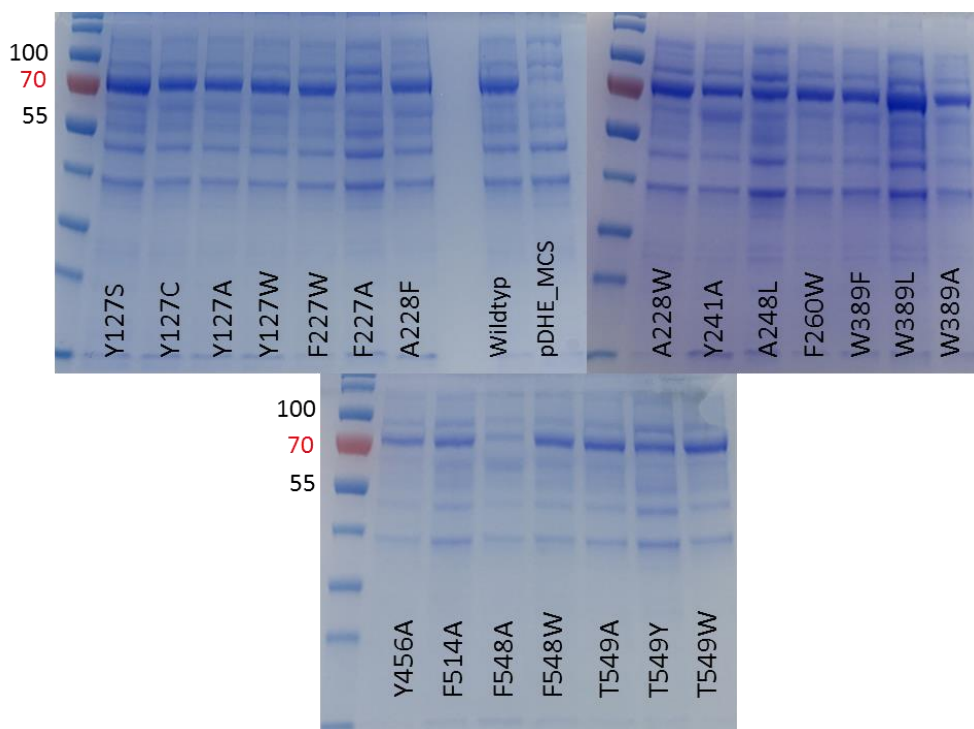


#### 4.4.6 Initial mutant library

The initial EmOAH mutant library consisted of 21 variants and was constructed using the *QuikChange* method (3.2.4). Mutagenic primers with the most frequently used codon in *E. coli* (“codon usage”) were chosen in case of redundancy. Successful mutagenesis was confirmed by sequencing the mutated region of the gene.

##### 4.4.6.1 Expression

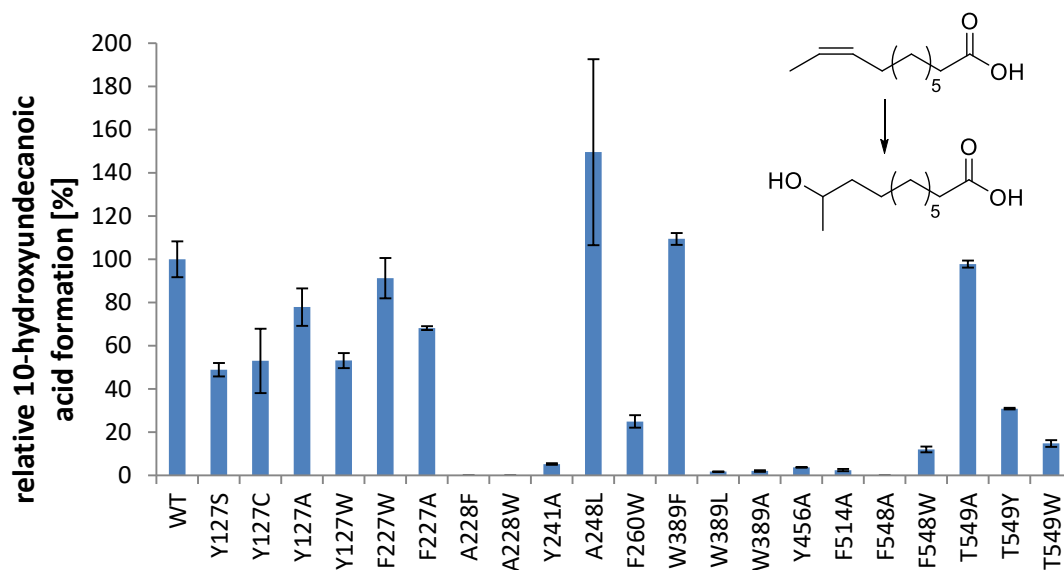
EmOAH variants were produced in the ADH-deficient strain *E. coli* ITB 94 and the expression level was monitored by SDS-PAGE (Figure 35). Expression was confirmed for all variants, however, the variants F227A and Y456A were less clearly and the variant F548A was hardly overexpressed. Cells transformed with the vector pDHE-MCS (an empty vector lacking the oleate hydratase gene) was used as control.



**Figure 35:** SDS-PAGE of EmOAH variants of the initial mutant library produced in *E. coli* ITB 94. Bands of the OAH are clearly visible at a size of about 75 kDa. Cells containing the empty vector pDHE-MCS were used as negative control.

4.4.6.2 Activity test with (*Z*)-undec-9-enoic acid

To test the variants of the initial library, activity towards hydration of (*Z*)-undec-9-enoic acid was analyzed in a first experiment. (*Z*)-undec-9-enoic acid is the shortest fatty acid which has been successfully hydrated by OAH wild type enzymes.<sup>[115]</sup> The reactions were conducted as described in 3.5.1 using a temperature of 30 °C and 800 rpm shaking speed for 40 h and analyzed by LC-MS (Figure 36).

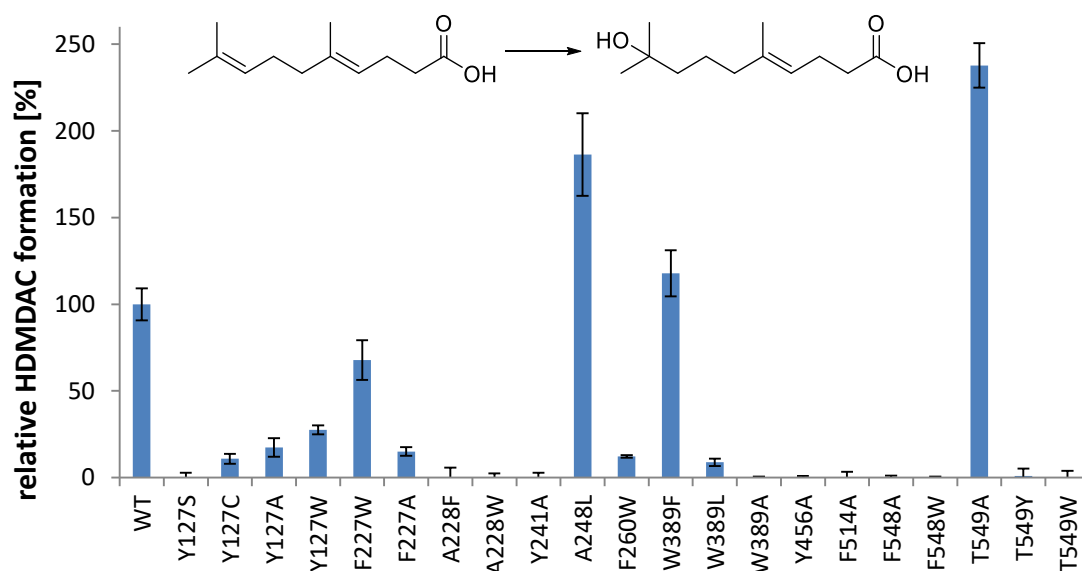


**Figure 36:** Results of the biotransformations using (*Z*)-undec-9-enoic acid as substrate. Depicted are the product formations of 10-hydroxyundecanoic acid in experiments using EmOAH variants of the initial library compared to the EmOAH wild type (WT). No background was observed in control experiments using cells containing an empty vector.

Most of the variants were active under the tested conditions, while seven can be considered as almost inactive (A228F, A228W, W389L, W389A, Y456A, F514A, F548A). In case of F548A, the inactivity can primarily be explained by the poor expression. Remarkable, the knockout variant Y241A showed some residual activity although the tyrosine at position 241 is described in the literature<sup>[120]</sup> as a catalytically active amino acid. Predominantly, variants showed decreased or wild type-like (Figure 36) product formation. Only the best variant A248L surpassed EmOAH wild type activity. However, it has to be considered that the EmOAH wild type and variant A248L already showed almost quantitative substrate conversion. Considering the standard deviation, both are in the same range and no statement can be made whether the mutation lead to further improvements. For reason of better comparison of the individual mutants, lower conversion would be preferred.

## 4.4.6.3 Screening with DMDDAC

The library was also screened in whole cell biotransformations using DMDDAC as substrate and the product formation was determined after 4 d of reaction by LC-MS (Figure 37).



**Figure 37:** Results of biotransformations using EmOAH variants of the initial mutant library, which was constructed based on the results of docking studies. Depicted are the product formations of HDMDAC in experiments using variants compared to the EmOAH wild type (WT). The results are adjusted by the background which was measured in control experiments using cells containing an empty vector.

The screening revealed 10 of 21 variants to be able to hydrate DMDDAC to HDMDAC under the tested conditions. Satisfyingly, three mutations thereof (A248L, T549A, W389F) showed significantly increased activity towards the target substrate up to a factor of 2.4. While many of the other mutants revealed reduced activity (Y127C, Y127A, Y127W, F227W, F227A, F260W, W389L), a big portion of the variants (Figure 37) were completely inactive including the knockout variant Y241A (negative control). The finding of a large number of inactive mutants is in contrast to the activity profile using (*Z*)-undec-9-enoic acid as substrate (4.4.6.2) where most of the mutations showed at least residual activity.

#### 4.4.7 Semi-rational site-directed mutagenesis (DWP-assay)

To increase the enzyme activity using an evolutionary approach, site-saturation libraries were screened independently or in combination with other amino acids. In more detail, R121, a residue in direct vicinity to the catalytically active E122 (see mechanism: 1.3.5.2), was randomized using the degenerate NNK codon (all 20 amino acids) and position 514 was diversified using the degenerate NYK, TAT and TGG codon (A, I, L, M, F, P, S, T, V, W and Y). Further, the two residues at M226 and F227 were mutated in a combinatorial approach using the 22c-trick<sup>[174]</sup> (NDT, VHG, TGG degeneracy, all 20 amino acids) and the residues F548 and T549 were combinatorial randomized using NYK, TAT and TGG degenerate codons (A, I, L, M, F, P, S, T, V, W and Y). All mutated residues are in 5 Å distance to the prenyl double bond. The four individual mutant libraries were generated successfully and the amino acid exchange was proven on DNA level by Sanger sequencing according to the Quick Quality Control protocol described by Reetz.<sup>[153]</sup> After construction of the libraries, the variants were produced in 96-deep well plates (DWP) and screened in biotransformations of DMDDAC.

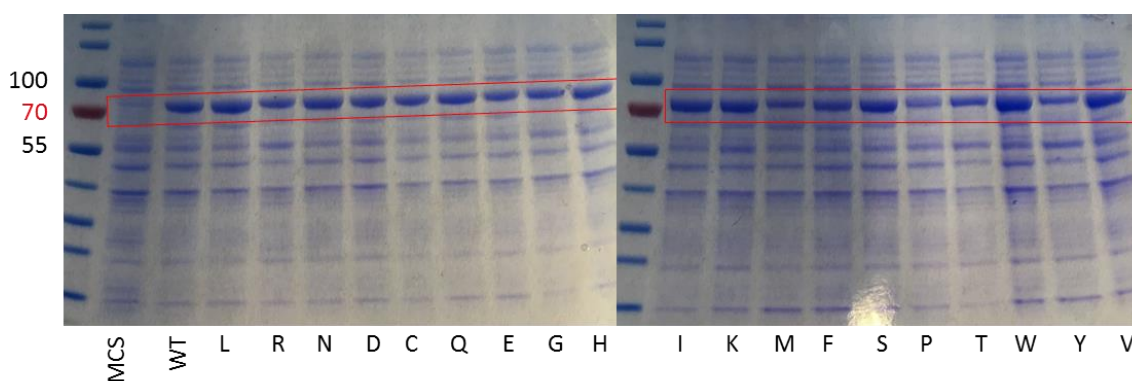
However, no clear results were obtained in LC-MS analysis as the product formation, if present at all, was in the range of the background. This might be due to the fact that the amount of active catalyst per well was below the threshold of activity detection. Therefore, the DWP screening was aborted after screening in total eleven plates (R121: 2 DWP, M226/F227: 8 DWP, F514: 1 DWP, F548/T549 no DWP screened) and the strategy was switched back to a more rational approach as applied for the initial library.

#### 4.4.8 Screening of EmOAH saturation mutagenesis libraries

Based on the results of the screening of the initial mutant library, saturation mutagenesis was performed for the position A248 and T549 and the mutants were screened for higher HDMDAC formation. In this screening, the variants A248L and T549S were found as the best from the respective library. A subsequent combinatorial saturation mutagenesis yielded the variant A248I/T549S showing further increased HDMDAC production. Using this double mutant as parent, a third round of saturation mutagenesis at W389 was performed. However, this did not lead to a triple variant with better activity. The detailed results of the saturation mutagenesis experiments can be found in the following subchapters.

## 4.4.8.1 A248X

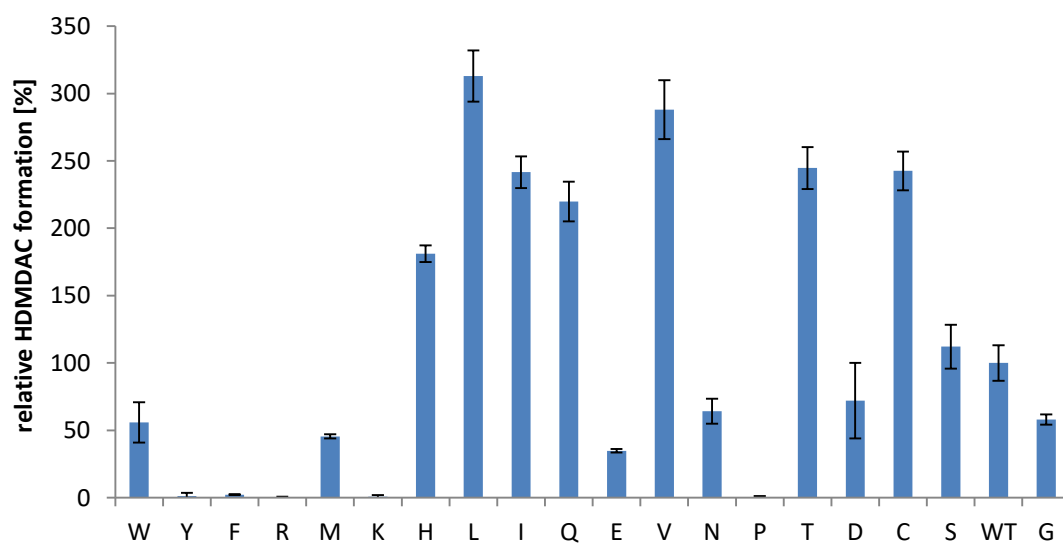
In a first round of saturation mutagenesis, alanine at position 248 was randomized. The variants were generated by the QuikChange method and produced in *E. coli* ITB 94 as described in the experimental section. The variants were produced in more or less similar amounts except for the variants A248P and A248Y which had apparently a lower expression level (Figure 38).



**Figure 38:** SDS-PAGE showing the expression of EmOAH A248 variants with a size of about 74.6 kDa. All variants were produced in more or less equal amounts except for the variants A248P and A248Y which apparently had a lower expression level.

Whole cell biotransformations of DMDDAC yielded 5 inactive variants (A248Y, F, R, K, P). In the case of A248P and A248Y, this might be due to the low expression level. Two variants with charged amino acids (A248D: 72 %, A248E: 35 %) showed less activity than the WT. An exception was found for the variant with a basic histidine (A248H: 181 %). In general, smaller (A248G: 58 %) or too large amino acid residues like tryptophan (A248W: 56 %) and methionine (A248M: 46 %) are not beneficial and exhibited only residual activity.

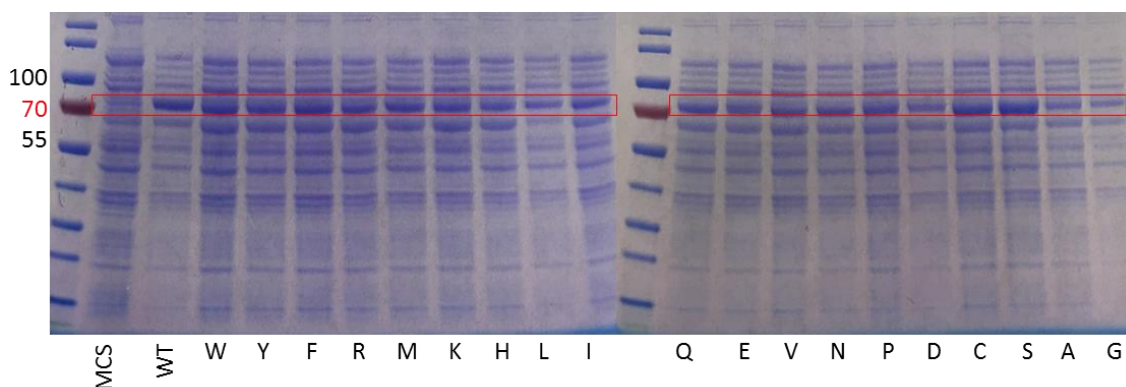
A significant improvement was obtained for medium-size amino acids. Variants with polar, neutral amino acids showed slightly increased (A248S: 112 %) or noticeably increased product formation (A248Q: 220 %, A248C: 243 %, A248T: 245 %) product formation. Best results were obtained for variants with nonpolar, neutral amino acids (A248L: 313 %, A248V: 288 %, A248I: 242 %).



**Figure 39:** Results of biotransformations using EmOAH variants from site-directed saturation mutagenesis at position A248. Depicted are the product formations of HDMDAC in experiments using variants compared to the EmOAH wild type (WT). The results are adjusted by the background which was measured in control experiments using cells containing an empty vector and ordered by decreasing volume of the amino acid residues (according to Esque *et al.*)<sup>[175]</sup>.

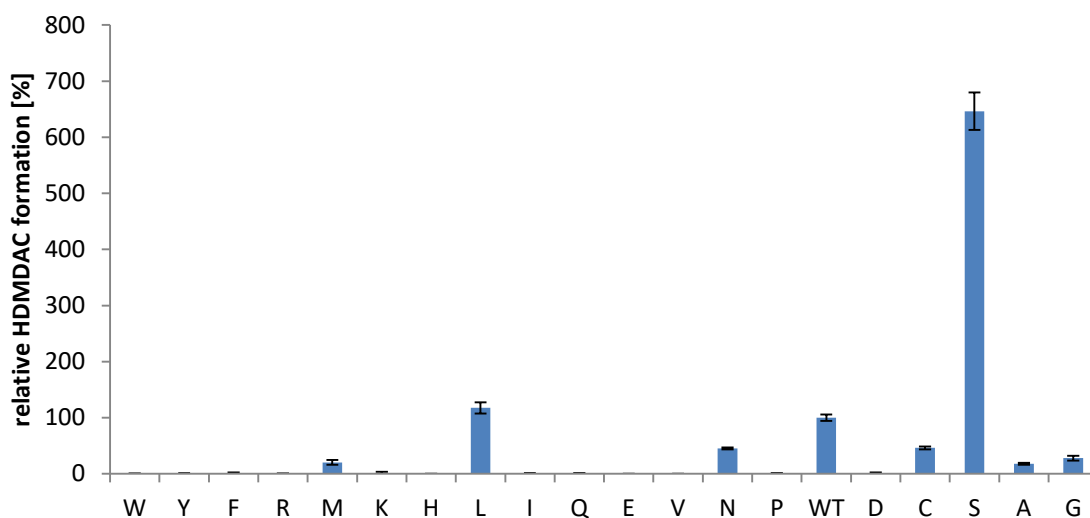
#### 4.4.8.2 T549X

Position T549 is of particular interest due to its proximity to the methyl group of the substrate's prenyl unit. The variants were generated by the QuikChange method and produced in *E. coli* ITB 94 (Figure 40) as described in the experimental section. Small variations in the expression level were observed. In general, the expression level seemed to be lower compared to the expression of A248 variants (4.4.8.1).



**Figure 40:** SDS-PAGE showing the expression of EmOAH wild type (WT) and T549 variants with a size of about 74.6 kDa compared to cell containing an empty vector (pDHE-MCS).

Whole cell biotransformations of DMDDAC revealed mainly (12 of 20) inactive variants (Figure 41). Especially variants with larger amino acid residues were inactive, while residual activity was detected for the medium-size nonpolar, neutral T549M (20 %) and WT-like activity for T549L (118 %). Interestingly, the related variant T549I, which has a similar size as T549L, was completely inactive. Some activity was also retained for small to medium-size amino acid residues with a polar side chain, as it is the case for T549N (45 %), T549C (46 %) and T549G (28 %). An exception is the variant T549A with a nonpolar alanine and 18 % relative product formation compared to the EmOAH wild type (WT). A distinct improvement of the activity was obtained by the introduction of a threonine. The variant T549S displayed a significant increase in product formation, as 6.5-times (646 %) the amount of HDMDAC was formed. Hence, the intention to make room for the additional methyl group and thereby taking account of the increased steric demand of the trisubstituted double bond was successful.



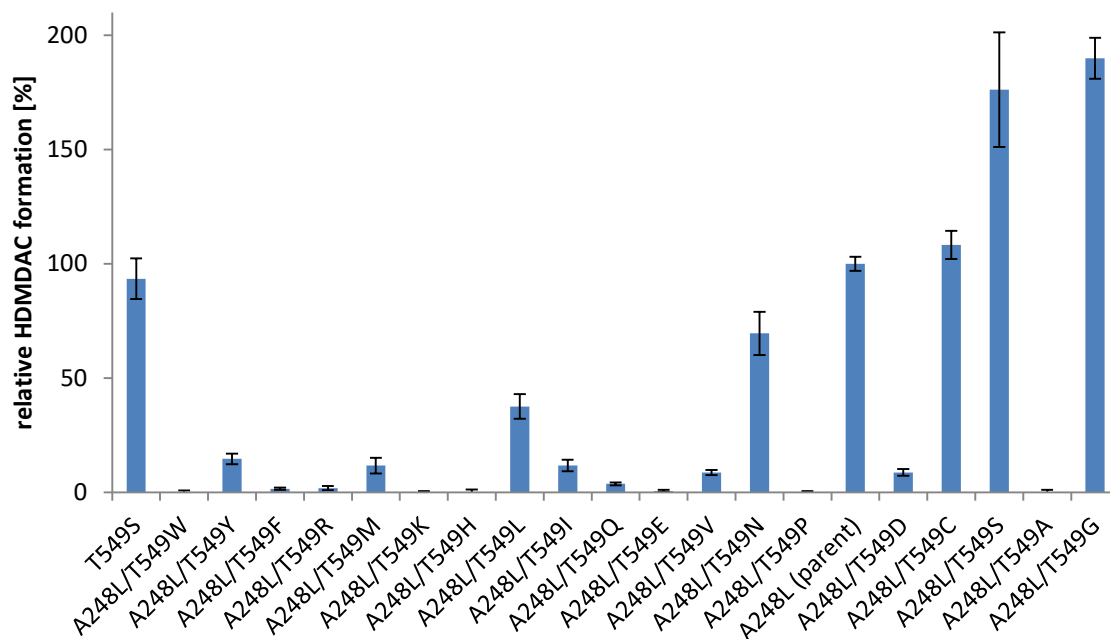
**Figure 41:** Results of biotransformations using EmOAH variants from site-directed saturation mutagenesis at position T549. Depicted are the product formations of HDMDAC in experiments using variants compared to the EmOAH wild type (WT). The results are adjusted by the background which was measured in control experiments using cells containing an empty vector and ordered by decreasing volume of the amino acid residues (according to Esque *et al.*)<sup>[175]</sup>.

#### 4.4.8.3 A248L/T549X and A248X/T549S

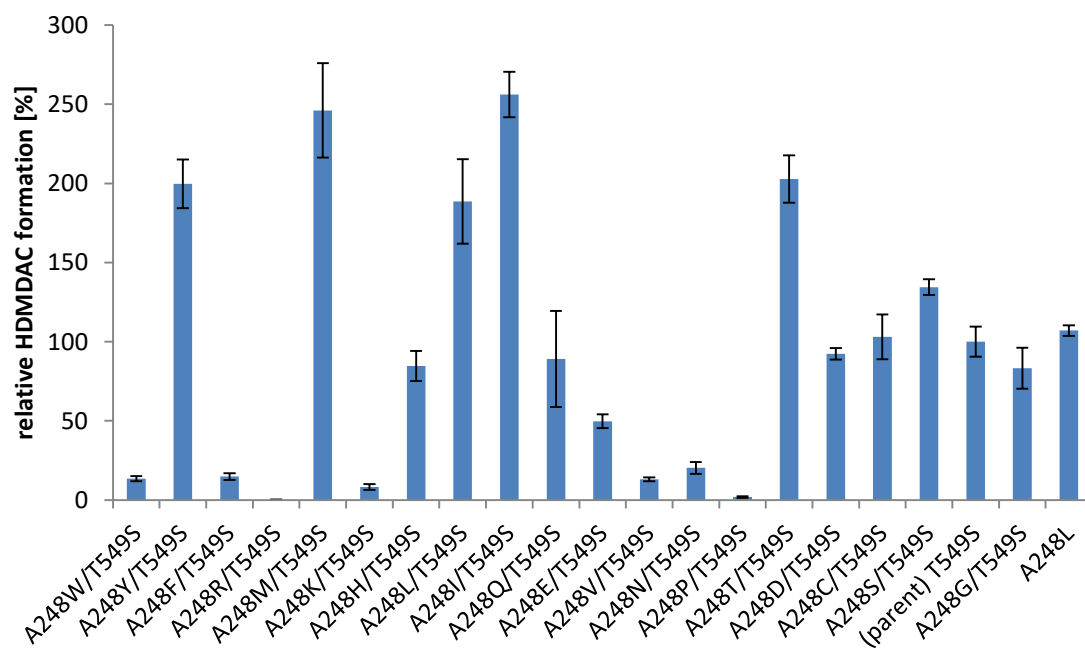
Motivated by the positive influence of mutations at positions A248 and T549, two additional libraries were generated in the context of the most beneficial variants A248L and T549S, respectively. Position T549 was saturated based on the parent A248L (Figure 42). Vice versa, position A248 was saturated using the EmOAH variant T549S as parent (Figure 43). The variants were generated by QuikChange, produced in *E. coli* ITB 94 and assessed in whole cell biotransformation as described in the experimental section.

The biotransformations revealed that the combination of mutations at position T549 and A248 has an influence on HDMDAC formation. Both parents produced a comparable amount of HDMDAC (A248L 107 % compared to T549S or T549S has 93 % compared to A248L). In general, variants based on T549S seemed to be more active, while many A248L/T549X variants have been less active than the parent or inactive (which is in accordance to the results of the single mutants). The variant A248L/T549G (190 % relative to A248L) along with the combinatorial variant A248L/T549S (176 % relative to A248L or 189 % relative to T549S, respectively) were rare exceptions. On the other hand, variants based on the T549S parent were more often equal or even more active. In total six variants exhibit higher HDMDAC formation compared to the parent T549S, namely the variants A248S/T549S (134 %), the combinatorial variant A248L/T549S (189 %), A248Y/T549S (200 %), A248T/T549S (203 %), A248M/T549S (246 %) and A248I/T549S, the overall best variant in this screening with 256 % relative product formation (239 % compared to A248L). Therefore, the EmOAH variant A248I/T549S was used for further experiments.





**Figure 42:** Results of biotransformations using EmOAH variants from site-directed saturation mutagenesis at position T549 using the single mutants A248L as parent. Depicted are the product formations of HDMDAC in experiments using variants compared to the EmOAH A248L parent. For reason of comparison, the variant T549S is also shown. The results are adjusted by the background which was measured in control experiments using cells containing an empty vector and ordered by decreasing volume of the amino acid residues (according to Esque *et al.*)<sup>[175]</sup>.

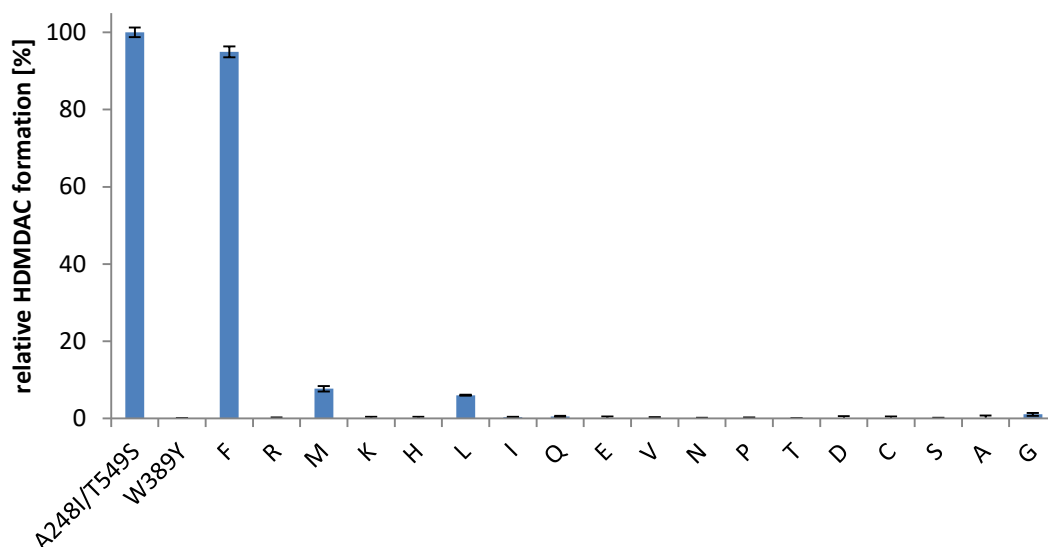


**Figure 43:** Results of biotransformations using EmOAH variants from site-directed saturation mutagenesis at position A248 using the single mutants T549S as parent. Depicted are the product formations of HDMDAC in experiments using variants compared to the EmOAH T549S parent. For reason of comparison, the variant A248L is also shown. The results are adjusted by the background which was measured in control experiments using cells containing an empty vector and ordered by decreasing volume of the amino acid residues (according to Esque *et al.*)<sup>[175]</sup>.

## 4.4.8.4 Saturation of position W389 based on EmOAH A248I/T549S

Iterative site-directed saturation mutagenesis was continued at position W389 based on the parent EmOAH A248I/T549S. The variants were generated by QuikChange, produced in *E. coli* ITB 94 and assessed in whole cell biotransformation (Figure 44) as described in the experimental section.

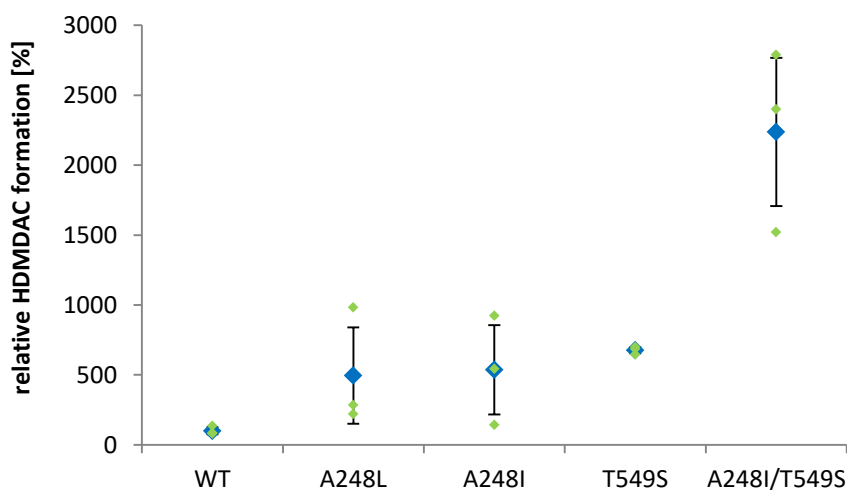
Formation of HDMDAC was only observed for the EmOAH variants A248I/T549S/W389F (95% compared to A248I/549S), A248I/T549S/W389M (8 %) and A248I/T549S/W389L (6 %). The results show that smaller amino acid residues at this position are not beneficial and reduce the activity of already evolved double mutant A248I/T549S in the hydration of DMDDAC. Further, hydrophobic (neutral, nonpolar) amino acids seem to be preferred in this position.



**Figure 44:** Results of biotransformations using EmOAH triple mutants derived from site-directed saturation mutagenesis at position W389 based on the A248I/T549S parent. Depicted are the product formations of HDMDAC in experiments using W389 triple variants compared to A248I/T549S. The results are adjusted by the background which was measured in control experiments using cells containing an empty vector. For better illustration, the results are ordered by decreasing volume of the amino acid residues (according to Esque *et al.*)<sup>[175]</sup>.

#### 4.4.9 Biological replicates of selected EmOAH variants

Biological replicates were conducted to confirm trends observed in the screening of mutants as high deviations are typical for whole cell biotransformation reactions. This is caused by the various physiological states cells used may be in: viable and growing; viable, but non growing; and non-viable. Therefore, control and reproducibility of bioconversions with whole cells are more difficult to accomplish.<sup>[176]</sup> The single point mutants A248L, A248I and T549S as well as the double mutant A248L/T549S were compared in biotransformations of 500  $\mu$ M DMDDAC to the results of the EmOAH wild type (WT).



**Figure 45:** Results for whole cell biotransformation of DMDDAC to HDMDAC with selected EmOAH variants conducted in biological triplicates. Depicted is the product formation of the variant compared to the EmOAH wild type (WT) in a swarm plot. The blue square shows the mean value of the three triplicates with standard deviation while the green squares show the single experiments. All values are corrected by the background product formation detected in control experiments using an empty vector and referred to the mean value of EmOAH WT.

In general, the screening results could be confirmed with three biological independent experiments (Figure 45). EmOAH WT showed only poor HDMDAC formation. A first improvement of activity was reached with single mutants (A248L:  $495 \pm 345$  %, A248I:  $536 \pm 319$  %, T549S:  $676 \pm 22$  %). While triplicates of WT and the variant T549S gave all similar results, higher deviations were observed for A248L and A248I. A significant enhancement is achieved by introducing a second mutation leading to the double mutant A248I/T549S ( $2237 \pm 530$  %). This means an average 22-fold improvement compared to the wild type.

#### 4.4.10 Screening of reaction conditions

Reaction conditions were screened using EmOAH variant A248I/T549S (only single reactions were performed) to find suitable parameters for preparative biotransformation of DMDDAC to HDMDAC using whole cell biocatalysts. Investigated conditions were inducer concentration during catalyst production ( $0.5 \text{ g L}^{-1}$  and  $1 \text{ g L}^{-1}$  rhamnose), reaction temperature ( $25 \text{ }^{\circ}\text{C}$ ,  $30 \text{ }^{\circ}\text{C}$  and  $37 \text{ }^{\circ}\text{C}$ ), substrate concentration ( $500 \text{ }\mu\text{M}$ ,  $1 \text{ mM}$  and  $2 \text{ mM}$ ) and (co-)solvent for dilution of substrate (DMSO or *i*PrOH) as well as different shaking methods ( $180 \text{ rpm}$  in an incubator or  $800 \text{ rpm}$  in an Eppendorf Thermomixer).

As a result, no distinct improvement was reached by altering reaction conditions. Concerning the inducer concentration, results for cells induced with  $0.5 \text{ g L}^{-1}$  rhamnose were comparable to those of cells induced with the doubled amount of rhamnose. A reaction temperature of  $30 \text{ }^{\circ}\text{C}$  was already the optimum, while  $25 \text{ }^{\circ}\text{C}$  gives comparable results and  $37 \text{ }^{\circ}\text{C}$  led to a decreased product formation. No obvious difference was also observed for co-solvent selection and shaking conditions. However, HDMDAC formation was improved by doubling the cell mass ( $100 \text{ g}_{\text{cww}} \text{ L}^{-1}$  instead of  $50 \text{ g}_{\text{cww}} \text{ L}^{-1}$ ). The use of higher substrate concentrations ( $2 \text{ mM}$  instead of  $500 \text{ }\mu\text{M}$ ) lead also to higher product formation.

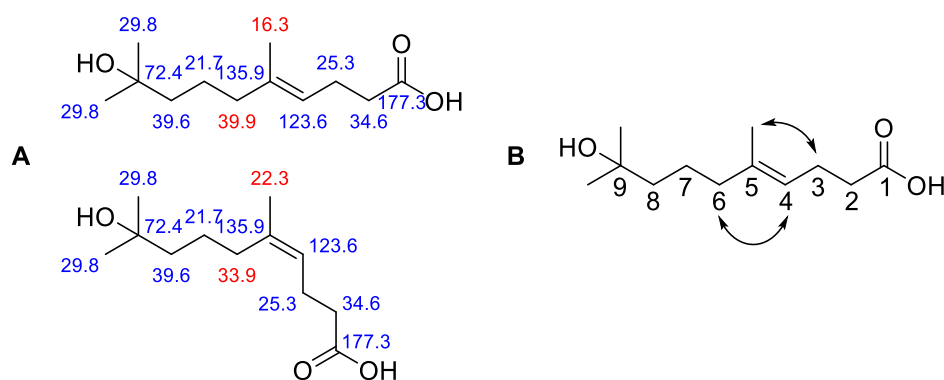
#### 4.4.11 Preparative biotransformations

Several attempts were made to isolate HDMDAC produced by hydration of DMDDAC using an *E. coli* whole cell catalyst containing EmOAH variant A248I/T549S. Improved conditions for these preparative scale biotransformations were explored in preliminary experiments (4.4.10). It was found that  $2 \text{ mM}$  substrate was a good trade-off for larger scale reactions, as it increases the total amount of formed product, while the yield was still in an acceptable range. Furthermore, the amount of cells was doubled to  $100 \text{ g}_{\text{cww}} \text{ L}^{-1}$ . In the first attempt ( $40 \text{ mL}$  scale), no product was obtained as the amount of crude product was too small for isolation via column chromatography. Therefore, the reaction volume was increased to  $400 \text{ mL}$ .

For the  $400 \text{ mL}$  scale, a total amount of  $157 \text{ mg}$  *E/Z*-DMDDAC ( $0.8 \text{ mmol}$ ) was used while  $23 \text{ mg}$  *E*-HDMDAC could be isolated by column chromatography from biotransformation (3.5.3). This corresponds to a yield of  $13 \%$  based on *E/Z*-DMDDAC but increases to  $40 \%$  when only the *E*-DMDDAC is considered as substrate. The

isolated amount of product was sufficient for detailed characterization and was analyzed by several analytical methods including GC-MS, LC-MS and NMR ( $^1\text{H-NMR}$ ,  $^{13}\text{C-NMR}$ , COSY, HSQC, HMBC and NOESY) techniques.

The two HDMDAC isomers can be clearly distinguished by their  $^{13}\text{C-NMR}$  spectra as the chemical shifts of C-6 and Me-5 are characteristic for the respective *cis* or *trans* configuration (Figure 46). The findings ( $\delta = 15.8$  ppm for Me-5 and  $\delta = 39.8$  ppm for C-6) were in accordance with the ChemDraw prediction for the *E*-isomer. The results were confirmed by NOESY (Figure 46) as the expected signals of a correlation between 4-H and 6-H and also between 5-Me and 3-H were present, while no signal was found for the potential *Z*-isomer were a 5-Me to 4-H correlation or a 3-H to 6-H correlation would be expected.

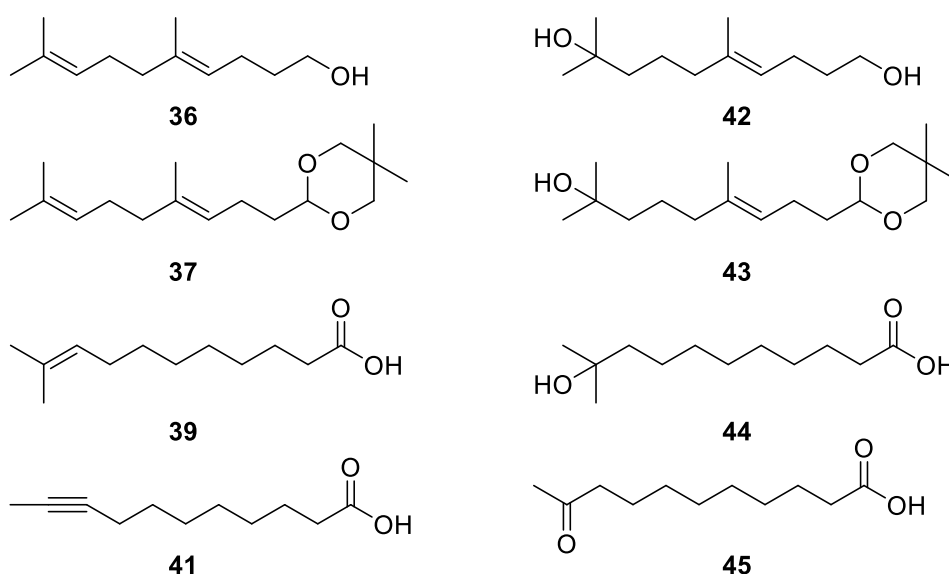


**Figure 46:** A) Chemical shifts predicted by ChemDraw for the  $^{13}\text{C-NMR}$  spectra of *E*-HDMDAC and *Z*-HDMDAC. A clear distinction between the two diastereomers can be made based on the chemical shifts of the C-6 and the methyl group adjacent to C5. B) Selected NOESY correlations in *E*-HDMDAC.

Therefore, the product was identified as the *E*-isomer of HDMDAC with traces of formic acid from the isolation process but without the *Z*-isomer (spectra can be found in the appendix 7.3).

#### 4.4.12 Substrate scope

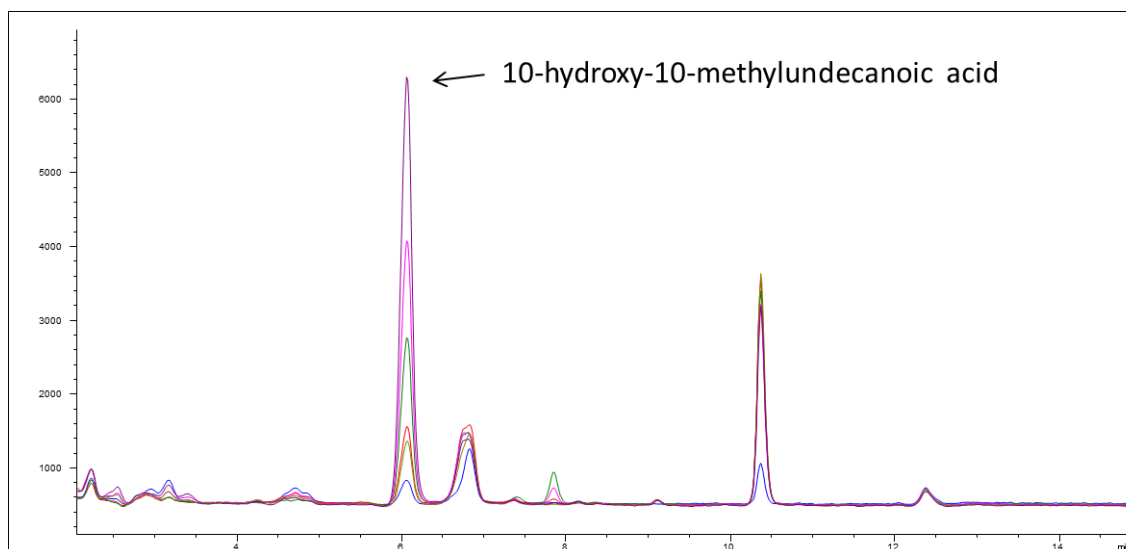
In order to investigate the substrate scope of the evolved hydratase and evolutionary intermediated, derivatives of DMDDAC **32** were studied (Figure 47). The corresponding alcohol DMDDOL **36** and the acetal DMDDIOX **37** were tested to explore the influence of the carboxylic acid binding motive. 10-Methylundec-9-enoic acid **39** was tested to study the influence of an additional methyl group adjacent to the double bond of (*Z*)-undec-9-enoic acid **40**, the so far smallest fatty acid substrate which was hydrated by EmOAH. Undec-9-ynoic acid **41** was included in the substrate screening to explore the activity of EmOAH in alkyne hydration. DMDDOL **36** and DMDDIOX **37** were also tested using heptanoic acid as decoy molecule.



**Figure 47:** Substrates DMDDOL **36**, DMDDIOX **37**, 10-methylundec-9-enoic acid **39** and undec-9-ynoic acid **41** for screening with EmOAH variants and the expected products (**42**, **43**, **44** and **45**) of the hydration reaction, respectively.

The biotransformations (96 h, 30 °C, 800 rpm in Eppendorf Thermomixer) were conducted as described in the experimental using 100 g<sub>cww</sub> L<sup>-1</sup> cells and 2 mM substrate. EmOAH variants for this experiment were A248L, A248I, T549S and A248I/T549S. Controls were performed using cells expressing empty vector (MCS) and EmOAH WT. As a positive control, the biotransformations of DMDDAC **32** were conducted under the same conditions. All biotransformations were performed only as single reactions.

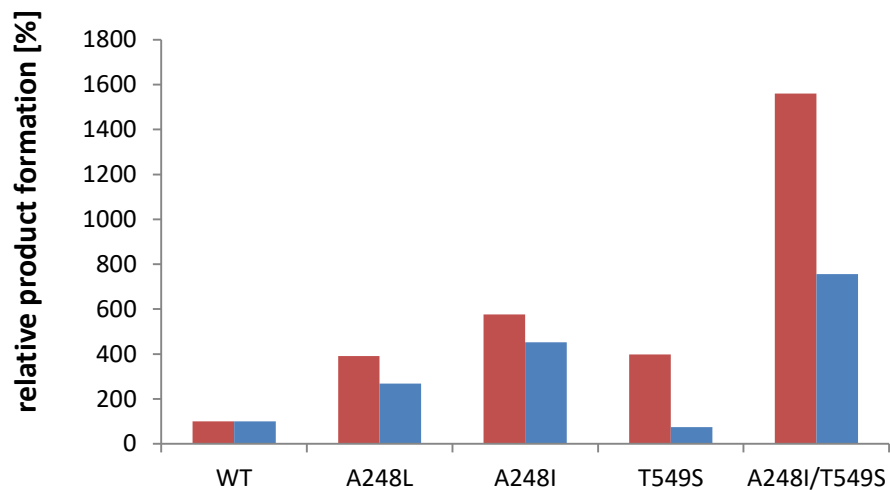
As observed in previous experiments, trace amounts of HDMDAC are also detectable in the empty vector control but the formed product increases significantly for wild type and mutants. This was also observed for the hydration of 10-methylundec-9-enoic acid **39** to 10-hydroxy-10-methylundecanoic acid **44** (Figure 48). Therefore, the background was subtracted from the other results.



**Figure 48:** LC-MS chromatogram: Depicted is the SIM (negative ionization) with mass to charge ratio of 215 of extracts from biotransformations of 10-methylundec-9-enoic acid with different EmOAH variants. The product 10-hydroxy-10-methylundecanoic acid has a retention time of 6 min. Other peaks are unknown impurities. Blue: empty vector, red: wild type, green: A248L, pink: A248I, light green: T549S and purple: A248I/T549S.

Compared to the EmOAH WT (100 %), the variant T549S showed only 75 % formation of 10-hydroxy-10-methylundecanoic acid, while the other single mutants have produced 2.7 times (A248L: 268 %) and 4.5 times (A248I: 452 %) the amount of product, respectively. The double mutant A248L/T549S showed with 755 % a significant higher product formation (Figure 49, blue). The results are comparable to those for the biotransformation of DMDDAC **32** (Figure 49, red). Therefore, it could be demonstrated that evolution of EmOAH for DMDDAC **32** also optimized the enzyme for the hydration of 10-methylundec-9-enoic acid. However, more detailed conclusions cannot be drawn as only single reactions were performed.

No distinct product formation was observed for the substrates DMDDOL **36** and DMDDIOX **37**. The use of heptanoic acid as dummy substrate and thereby activating the enzyme, did not lead to the hydration of DMDDOL **36** and DMDDIOX **37**, either. The hydration of the triple bond in undec-9-ynoic acid **41** was not possible.



**Figure 49:** Results of the biotransformations of DMDDAC (red) and 10-methylundec-9-enoic acid (blue). Depicted is the product formation compared to the EmOAH wild type (100 %).



## 5 Discussion

In this work, the biocatalytic alkene hydration was investigated as an alternative to chemical synthesis. Classical alkene hydration is limited as this chemistry mainly depends on strong acids with limited catalyst control. This leads to poor selectivities and low functional group tolerance. Enzymes commonly provide high selectivities at mild reaction conditions and can be therefore considered as “green”. Biocatalysis in fact meets ten of twelve principles of green chemistry, with the remaining two (design for safer products and for degradation) are related to the product, which is independent of the way of production.<sup>[7,21]</sup> Hydratases do not only combine all these advantages, they also optimally meet the first and second green chemistry principles. Through directly hydrating the substrate whereby only water is needed, hydratases prevent waste and show excellent (100 %) atom economy.<sup>[69]</sup> For the application in industry, cofactor-independent enzymes provide several advantages, as no costly cofactor has to be added and regenerated. Another advantage of hydratases is that only one catalyst is required in contrast to redox cascade approaches with combined oxidation and subsequent selective reduction.<sup>[177]</sup> Such redox cascades have already been explored in the past but are limited to primary and secondary alcohols. On the contrary, such an approach cannot be applied to generate tertiary alcohols. As tertiary alcohol products are of high interest in fragrance industry, the regioselective hydration of prenyl double bonds in terpenoids was investigated in this study.

### 5.1 Selection of hydratases

Out of the five hydratases (1.3) acting on non-activated isolated C-C double bonds known to date namely linalool dehydratase-isomerase, kievitone hydratase, carotenoid hydratase, fatty acid hydratases and limonene hydratase, none was known to hydrate DMDDAL or similar substrates.<sup>[69]</sup> At the beginning of the project, these potential enzymes were therefore evaluated concerning their literature-known substrate scope and other relevant properties (Table 43).

Linalool dehydratase-isomerase (LinD or LDI) was not considered in this study, as it in general favors dehydration reactions and is limited to an  $\alpha$ -methyl allyl alcohol motif.<sup>[75,78]</sup> Kievitone hydratase (KHS) was also excluded from further investigations, although it catalyzes the desired hydration of a prenyl moiety.<sup>[69,84]</sup> Firstly, the natural

substrate kievitone is not commercially available making activity tests impossible.<sup>[84]</sup> Secondly, the natural substrate differs widely from the target terpenoid substrates except for the prenyl moiety. We suspected that the substrate is recognized in the active site at several positions by polar interactions such as hydrogen bonds or interaction between the aromatic ring system and aromatic amino acid residues. These hypotheses have been meanwhile further supported by Engleder *et al.* who investigated the substrate scope of KHS. The authors speculated that the substrate binding is mediated by  $\pi$ - $\pi$ -stacking of the aromatic (prenylated) A-ring with aromatic amino acids and/or interaction of polar side chains with the hydroxyl-substituent at C7.<sup>[84]</sup>

On the contrary, carotenoid-1,2-hydratases (CrtCs) were considered for detailed investigations. CrtC regioselectively hydrates the terminal prenyl double bond of linear carotenoids. Moreover, the shorter geranylgeraniol (C20 terpenoid) was described to be accepted while the obtained product has never been characterized.<sup>[96,128,178]</sup> The authors proposed a minimum substrate size of C20 as shorter terpenoids (C5, C6, C15) were not converted.<sup>[96]</sup> This motivated us to shift the substrate scope by enzyme engineering towards shorter terpenoids like DMDDAL. Although a crystal structure was missing<sup>[97]</sup>, the availability of several biochemical characterization studies<sup>[91,92,96,97,128,178]</sup> let us consider a rational approach to be possible.

Out of the five enzyme candidates, oleate hydratases (OAHs) are by far the most broadly studied.<sup>[69,73]</sup> Since the first report of EmOAH in 1962, countless findings in different organisms and the characterization of more than 25 representatives followed.<sup>[69,113]</sup> Meanwhile, a hydratase engineering database (HyED) including sequences of 2046 putative fatty acid hydratases was established and several investigations on the substrate scope and optimization of the reaction conditions were performed, in part also by our research group. While first findings indicated that OAHs are limited to unsaturated *cis*-9-fatty acids with a chain length of at least C11, recent studies also demonstrated the hydration of shorter olefins using a carboxylic acid as decoy molecule.<sup>[127,130,157,179]</sup> OAHs were therefore included in this study to a later time point as a promising hydration catalyst.

The least studied enzyme of this group is the so-called limonene hydratase (LIH), which potentially hydrates (*R*)-(+)-limonene to  $\alpha$ -terpineol and represents a rare example for the enzymatic addition of water to monoterpenes.<sup>[69]</sup> Mainly,  $\alpha$ -terpineol production from limonene was reported in whole cells and the responsible enzymes have never been isolated or described in detail. A thermostable LIH was subcloned in *E. coli* as part

of a limonene degradation pathway of *Geobacillus stearothermophilus* BR388 obtained by restriction digest of genomic DNA and heterologously expressed.<sup>[106]</sup> Characterization showed not only hydratase activity for limonene, but surprisingly also for the nitrile group of cyanopyridine, which is a remarkably example of accidental catalytic promiscuity and suggesting a broad substrate scope.<sup>[106,180]</sup> In general, enzymes with a high catalytic versatility are such enzymes which recognize their substrate by hydrophobic interactions in the active site instead of very specific polar interaction like hydrogen bonds.<sup>[181]</sup> This was the motivation to further study *G. stearothermophilus* LIH in this project, although investigations had to be started from the original strain as no DNA or protein sequence of the investigated hydratase was available.

**Table 43:** Overview of the potential candidates for enzymatic hydration of highly substituted alkenes like the prenyl double bond in terpenoids. Advantages and disadvantages which are known from literature are compared.

Hydratase	Advantage	Disadvantage
LinD	<ul style="list-style-type: none"> <li>• Crystal structure available</li> <li>• Activity on monoterpenes confirmed</li> </ul>	<ul style="list-style-type: none"> <li>• Substrate scope: limited to <math>\alpha</math>-methyl allyl motif</li> <li>• Dehydration preferred</li> </ul>
KHS	<ul style="list-style-type: none"> <li>• Hydration of prenyl motif to tertiary alcohol known</li> </ul>	<ul style="list-style-type: none"> <li>• Natural substrate for activity test not commercially available</li> <li>• Activity on polar, cyclic substrates confirmed</li> <li>• Limited information available</li> </ul>
CrtC	<ul style="list-style-type: none"> <li>• Hydration of prenyl double bonds in linear terpenoids known</li> <li>• Mechanism has been proposed</li> </ul>	<ul style="list-style-type: none"> <li>• Substrate scope: limited to C20-C40 terpenoids</li> <li>• No crystal structure available</li> </ul>
OAH	<ul style="list-style-type: none"> <li>• Most studied alkene hydratase</li> <li>• Many characterized representatives known</li> <li>• Crystal structure available</li> <li>• Putative mechanism described</li> <li>• Broad substrate scope for fatty acids, terminal and internal alkenes</li> </ul>	<ul style="list-style-type: none"> <li>• Carboxylate essential for activation of OAH, either in the substrate or in a decoy molecule</li> </ul>
LIH	<ul style="list-style-type: none"> <li>• Hydration of small non-polar substrates (monoterpenes)</li> <li>• Broad substrate scope: acting also on C-N triple bonds</li> <li>• Thermostable enzyme</li> </ul>	<ul style="list-style-type: none"> <li>• Not isolated and purified</li> <li>• Not characterized</li> <li>• No sequence available</li> </ul>

## 5.2 *Geobacillus stearothermophilus*

### 5.2.1 Conditions for cultivation

For successful growth of *G. stearothermophilus*, it was necessary to find the appropriate cultivation conditions. The most important issue was to find the right composition of the medium (4.1.2). Tryptone soya broth (TSB) was finally used as it is a typical medium for fastidious microorganisms and is meanwhile described in several studies as suitable medium.<sup>[169,182,183]</sup> However, Javed *et al.* found that growth of *Geobacillus* spp. in rich media (agar) is most often inconsistent and non-reproducible, which is in accordance with the findings in this study.<sup>[184]</sup> It was stated, that *Geobacillus* spp. are particularly demanding in their vitamin and amino acid requirements, which are usually satisfied with high levels of undefined compounds of complex nutritional sources. This can be obstructive as for instance in the present case in the analysis of metabolites of limonene. Therefore, the authors developed a semi-defined minimal medium, using glycerol and pyruvate as carbon source supplemented with very low levels of yeast extract and tryptone which supported good growth with high and reproducible viable cell count.<sup>[184]</sup> However, no growth was achieved in the present study for cultivation of *G. stearothermophilus* in the similar DP salt medium supplemented with yeast extract, vitamins and limonene as carbon source as described by Chang and Oriel.<sup>[105]</sup>

Another reason why initially no consistent cultivation conditions could be found in this work may be the variability of spore heat resistance which was attributed by Wells-Bennik *et al.* to strain variability and the conditions of sporulation (particularly the sporulation medium and addition of ions like  $\text{Ca}^{2+}$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Mn}^{2+}$ ).<sup>[182]</sup> One more issue is the heterogeneity of the lag time of *G. stearothermophilus* single spores (15 h-60 h at 45 °C).<sup>[183]</sup> The mean lag time decreases from 37.9 h at 45 °C to an optimum of 10.9 h at 55 °C, which was also the temperature with the smallest standard deviation of 9.2 h.<sup>[183]</sup> Compared to a study of the same group in 2016, the temperature dependence of the mean lag time was different.<sup>[183,185]</sup> It can be assumed that the lag time and the growth in general may differ depending on the physiological state and the ratio between spores and vegetative cells.<sup>[183]</sup> This is also relevant for the composition of the inoculum. The influence of the inoculum size on spore germination was investigated in several studies.<sup>[168,169]</sup> These studies agreed that larger inoculum sizes (higher spore concentrations) lead to less variation and more rapid germination as a result of quorum sensing. Quorum sensing might also be the reason for *Geobacillus* spp. to grow in

biofilms on surfaces at the air-liquid interface rather than in submerged biofilms or preferable on agar.<sup>[170]</sup>

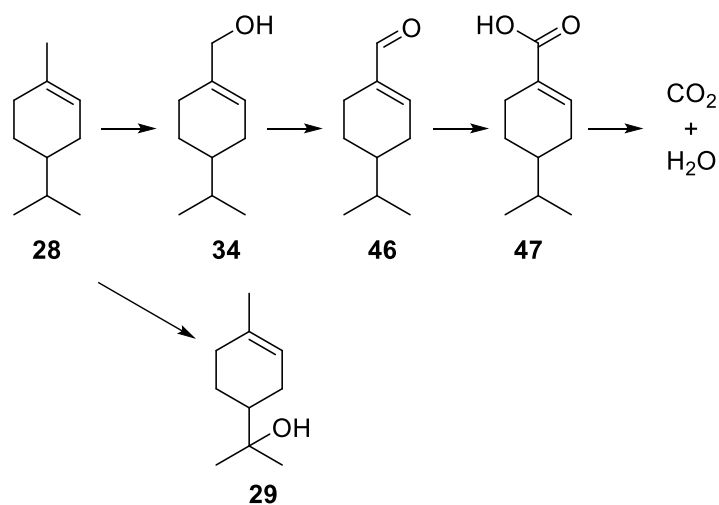
All things considered, cultivation of *Geobacillus* spp. is not trivial and the optimal conditions have to be found for the individual strain. In the present study, TSB as a rich medium enabled growth of *G. stearothermophilus* BR 388 in liquid culture and allowed us to generate enough cells for isolation and sequencing of genomic DNA. On the other hand, M9 minimal agar was used to analyze metabolism of limonene which was introduced as vapor.

### 5.2.2 Sequencing, identification and taxonomic classification

The strain BR388 was originally isolated in 1994 by Chang and Oriel from orange peel and tentatively assigned as *Bacillus stearothermophilus* due to its rod shaped cells of 3-4  $\mu\text{m}$  length, variable Gram staining, oval spores, and the ability to grow well under aerobic conditions at 65 °C.<sup>[105]</sup> *Bacillus stearothermophilus* was reassigned in 2001 to *Geobacillus stearothermophilus*.<sup>[167]</sup> In 2016, studies by Aliyu *et al.* based on whole genome sequences, however, suggested that two distinct clades are existing.<sup>[186]</sup> Clade I typically has a G+C content of 48.8-53.1 %, while clade II has 42.1-44.4 % G+C content. The species of *Geobacillus* residing in clade II, e.g. *G. thermoglucosidasius*, were suggested to be classified as a new genus called *Parageobacillus*.<sup>[186]</sup> Considering that the whole genome sequencing of putative *G. stearothermophilus* BR388 gave a GC-content of 43.9 % and that 16S rRNA sequencing gave ambiguous results, it is likely that it should be in fact assigned as *Parageobacillus thermoglucosidasius*. This is supported by the fact that the species-specific PCR method of Pennachia *et al.*<sup>[155]</sup> using a specific pair of primers for *G. stearothermophilus* gave no visible bands in agarose gel electrophoresis and that BLAST of unitigs of gDNA sequencing results are in accordance with *Parageobacillus thermoglucosidasius*. Further experimental evidence could not be provided by the present study, as cultivation experiments in TSB medium using different sugars and glycerol as additional carbon source were not clear. The cells grew with all carbon sources. This leads to the assumption that other components of the TSB medium provided enough C-sources making it impossible to distinguish between the different species based on specific nutritional demands.

### 5.2.3 Putative limonene hydratase

Hints on the existence of a limonene hydratase (LIH) in *G. stearothermophilus* BR388 were provided by Chang and Oriel in 1994 upon isolation of the microorganism.<sup>[105]</sup> They found  $\alpha$ -terpineol **29** as a side product during growth studies of *G. stearothermophilus* BR388 on DP minimal medium and introduction of limonene **28** vapor as carbon source.<sup>[105]</sup> Main metabolites were perillyl alcohol **34** and to a lesser extend the further oxidized perillyl aldehyde **46**.<sup>[105]</sup> In contrast, our experiments could not approve these findings. The cultivation (55 °C, 4 d) of *G. stearothermophilus* BR388 on M9 minimal agar in limonene **28** vapor gave only limonene diol as metabolite and high concentrations of limonene **28** were isolated. Neither  $\alpha$ -terpineol **29** nor perillyl alcohol **34** or perillyl aldehyde **46** could be detected. This can have several reasons. Deviating from the studies by Chang and Oriel<sup>[105]</sup>, we did not supplement yeast extract as additional nutrient. Therefore, limonene was the only carbon source and presumable metabolized with higher rates (Figure 50). As a consequence, small amounts of metabolites may be formed but instantly metabolized further. This makes the detection of perillyl alcohol **34** and perillyl aldehyde **46** difficult. The problem is enhanced by the fact that the metabolites are also exposed to the high temperature which is needed for cultivation and may be evaporating readily. Another reason may be the time point of analysis, as Chang and Oriel stated that the highest yield of perillyl alcohol **34** was reached in the exponential phase<sup>[105]</sup>, while the culture in the present study may have been already in the stationary phase.



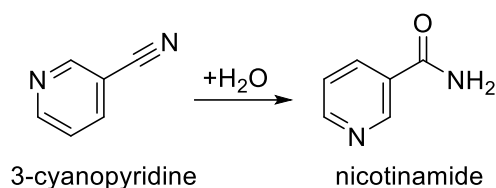
**Figure 50:** Limonene degradation pathway proposed by Chang *et al.* Limonene **28** is catabolized by oxidation via perillyl alcohol **34** and perillyl aldehyde **46** to perillic acid **47** and finally carbon dioxide and water. Limonene **28** is also converted in a byway to the dead end metabolite  $\alpha$ -terpineol **29**.<sup>[104]</sup>

Concerning the metabolites, the results were confirmed in biotransformations of limonene at room temperature using resting cells in buffer, as only the substrate limonene and limonene diol were detected. The cells for this experiment were grown in TSB complex medium and were used after harvesting. No further growth and metabolism is estimated under the applied conditions. Therefore mainly  $\alpha$ -terpineol was expected as product of biotransformation. One might speculate that the responsible enzyme, the putative LIH, was also not expressed during cultivation. However, the group of Oriol omitted to argue about the function of the dead end metabolite  $\alpha$ -terpineol or the reason for its production during limonene degradation, though the authors stated that it is, along with limonene diol, a typical metabolite found in microorganisms growing on limonene.<sup>[104,105]</sup> Limonene diol is presumable the product of a dioxygenase activity, which could be another detoxification mechanism. Oxygenation of hydrophilic compounds makes them more water soluble and thereby in general less toxic.

The potential of the responsible enzymes from the limonene degradation pathway for the biotechnological production of valuable monoterpenoids motivated the group of Oriol to clone the limonene degradation pathway into *E. coli*.<sup>[104,105]</sup> Therefore, the authors cloned DNA fragments larger than 5 kb which were obtained by partial restriction digestion of gDNA with EcoRI and found a 9.6 kb large fragment which conferred growth on limonene and assigned the clone EC409A.<sup>[104]</sup> Subcloning of fragments allowed the group of Oriol also the identification and characterization of an aliphatic amidase<sup>[173]</sup> and a limonene hydroxylase<sup>[172]</sup>. With the given sequence information for the two enzymes, we were able to locate the putative limonene degradation pathway on unitig\_6 of the sequenced gDNA. The size and the restriction pattern were more or less in accordance with the literature<sup>[106,107]</sup>. However, the exact size and location could not be followed up, as the given information is presumably based only on electrophoresis and therefore not very accurate. For instance, the two studies mention different sizes for the subfragments EC 419 (1.6 or 1.7 kb) and EC 423 (7.3 or 7.1 kb).

In the section of the limonene pathway, we found ORFs which were not investigated in detail by Chang and coworkers. A BLAST analysis suggested the presence of a CLP protease and RepB protein and a protein of the AbrB/MazE/SpoVT family. Another hypothetical protein known from sequence data of different *Geobacillus* spp. was also

found. Unfortunately, no LIH could be identified by BLAST as so far no sequence data has been deposited in databases. However, ORFs coding for the activator protein and the  $\alpha$ - and  $\beta$ -subunit of a nitrile hydratase (NHase) were found at the edge of the putative limonene pathway. This is interesting, as Savithiry *et al.* described a HPLC-based LIH assay, exploiting the broad substrate specificity of LIH for the hydration of the nitrile group in 3-cyanopyridine to nicotineamide (Figure 51).<sup>[106,107]</sup> In fact, this is a typical NHase activity which is even applied in industry.<sup>[187,188]</sup> We therefore suspected the opposite case, namely the hydration of limonene as a promiscuous activity of NHase. In principal, such an activity is covered by the catalytic mechanisms which are discussed in literature for NHases, though the exact mechanism still remains unsolved.<sup>[188,189]</sup> It could also be speculated that the NHase was evolved to LIH as *Geobacillus* spp. can rapidly adapt to environmental changes via genome diversification by horizontal gene transfer.<sup>[190]</sup> Additionally, *Geobacillus* spp. diversify their genome using inductive mutations and transposable elements. This leads to adaption to the stressor (in this case limonene) and is the reason why *Geobacillus* spp. efficiently generated mutants of thermostable enzyme variants under appropriate selection pressure.<sup>[190]</sup> Surprisingly, Savithiry *et al.* reported on two fragments with LIH activity indicating the presence of two distinct hydratases in the limonene degradation pathway.<sup>[106,107]</sup> Therefore, further research has to be done and the characterization of all identified ORFs is still inevitable to solve the mystery of LIH.



**Figure 51:** Hydration of 3-cyanopyridine is a typical activity for nitrile hydratases (NHase).<sup>[187,188]</sup> Promiscuous activity of NHase could therefore be the reason for limonene hydration in *G. stearothermophilus* BR388.



### 5.3 Carotenoid hydratases

Carotenoid-1,2-hydratases (CrtCs) are enzymes involved in the metabolism of carotenoids (C40 terpenoids) in purple photosynthetic bacteria. The membrane-bound enzymes were described to be able to hydrate lycopene and neurosporene as their natural substrate at the  $\psi$ -ends and in addition to that the shorter geranylgeraniol.<sup>[96]</sup>

In the present study, the carotenoid hydratases from *Thiocapsa roseopersicina* (TrCrtC) and *Rubrivivax gelatinosus* (RgCrtC) were selected for the investigation of the regioselective hydration of shorter linear terpenoids like DMDDAL. Both enzymes were heterologously expressed in *E. coli* strain ITB 94. The use of this strain overall improved expression compared to *E. coli* BL21 (data not shown) and successful expression was achieved for incubation at different temperatures (4.2.1). Apparently, higher expression levels were obtained for RgCrtC (Figure 26) which is in accordance with the findings of Hiseni *et al.* who described an around two times higher expression level for RgCrtC compared to TrCrtC.<sup>[96]</sup>

Lysate (cell free extracts) of *E. coli* cells containing the overexpressed CrtC was used for hydration reactions. Separation from the membrane environment during purification can lead to rapid loss of activity of carotenoid enzymes<sup>[90,191]</sup> and hydration of lycopene was reported in literature to be very slow using isolated CrtCs.<sup>[96]</sup> Therefore, Hiseni *et al.* determined enzyme kinetic parameters for lycopene hydration using lysate.<sup>[96]</sup> However, even with lysate, the enzyme is not very active for the putative natural substrate, as can be deduced from the determined parameters (RgCrtC:  $K_m$ : 24  $\mu\text{M}$ ,  $V_{max}$ : 0.31  $\text{nmol h}^{-1} \text{mg}^{-1}$ ,  $V_{max}/K_m$ : 1.3 ( $\times 10^2$ ),  $k_{cat}/K_m$ : 0.16  $\text{s}^{-1} \text{M}^{-1}$ ; TrCrtC:  $K_m$ : 9.5  $\mu\text{M}$ ,  $V_{max}$ : 0.15  $\text{nmol h}^{-1} \text{mg}^{-1}$ ,  $V_{max}/K_m$ : 1.6 ( $\times 10^2$ ),  $k_{cat}/K_m$ : 0.19  $\text{s}^{-1} \text{M}^{-1}$ ).<sup>[96]</sup> This is in accordance with comparable experiments by Steiger *et al.* for RgCrtC and the CrtC of *Rhodobacter capsulatus*.<sup>[92]</sup> Both studies used phosphatidylcholine as a detergent during the reaction which prompted us to start with these conditions and extended reaction times of 20-24 h compared to 4 h (Hiseni *et al.*<sup>[96]</sup>) or overnight reactions (Steiger *et al.*<sup>[92]</sup>). Unfortunately, we did not observe hydration of lycopene under the applied conditions. We further assumed that other detergents like SDS, CHAPS or Triton X-100 are more feasible as our research group has broad experience using these detergents for the membrane-bound squalene-hopene cyclase (SHC) which shares structural and mechanistic aspects with CrtC.<sup>[97,192-194]</sup> Experiments were conducted eventually using Triton X-100, the most frequently used detergent in SHC biotransformations<sup>[195]</sup>, enabling the mono- and dihydration of the natural substrate lycopene. Lycopene

conversion was in the range of 40 % for TrCrtC and up to 60 % for RgCrtC (Table 40). Hiseni *et al.* reported only conversions of 30 % for activity tests (reaction time presumable 4 h) of isolated RgCrtC and TrCrtC without stating the exact amount of used enzyme.<sup>[96]</sup> However, it has to be considered that for the results obtained in the present study, only the HPLC area values of substrate and products were analyzed and substrate loss (e.g. due to autoxidation, isomerization, ...) was not considered, leading in general to a bias in support of higher conversion. Interestingly, the product ratio (2:1) given by Hiseni and coworkers differed significantly from the present findings which are in the range of about 1:1 (TrCrtC) to 1:1.5 (RgCrtC). This can be explained with a more advanced reaction (due to a higher amount of enzyme or extended reaction time). More lycopene is consumed and reacted in a first step to the monohydrated product before this is again hydrated to the dihydrated product. Depending on the equilibrium of the reaction (or the different reaction rates), a certain constant amount of monohydrate could be necessary to be present in the reaction mixture while the dihydrate increases with decreasing substrate. This assumption is supported by the results we obtained in RgCrtC biotransformations of lycopene with shortened reaction times (4 h) and lower lysate concentrations (Table 41). When conversion is lower, products are formed in equal amounts or the monohydrate still predominates.

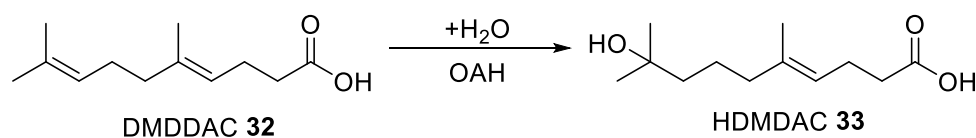
The successful conversion of lycopene using heterologously expressed CrtCs motivated us to test shorter terpenoids like the non-natural substrates geranylgeraniol or DMDDAL (Figure 25). However, we were not able to reproduce the hydration of geranylgeraniol (5 % conversion) reported by Hiseni and coworkers<sup>[96]</sup> or convert even shorter substrates with our reaction set-up. In general, shorter substrates might be hydrated but the concentration of the product could be too low for detection. As for SHCs, we suggest that CrtCs with their large active site are able to accommodate rather large compounds (like their natural carotenoid substrates) and exhibit decreasing activity with decreasing substrate size due to a poorer positioning and stabilization.<sup>[96,97,195]</sup> This subproject was stopped at that point, as the work with membrane-bound enzymes and the subsequent GC-analysis can be per se tedious and challenging due to the applied detergent and therefore too many parameters had to be considered for improvement. In further studies, the limited substrate scope of CrtCs could be overcome with increasing availability of structural and sequence data by the discovery of novel CrtCs with a distinct substrate scope for shorter substrates or by enzyme engineering towards truncated substrates.

## 5.4 Oleate hydratases

Oleate hydratases (OAHs) are promising enzymes for the regioselective hydration of a variety of substrates. Besides their natural substrate, oleic acid, OAHs also accept smaller *Z*-9-unsaturated fatty acids (C16-C11).<sup>[69,115]</sup> Meanwhile, even the hydration of non-activated terminal and internal alkenes was demonstrated using short carboxylic acids as decoy molecules and engineered OAHs.<sup>[130]</sup> In these conversions, OAHs exhibit excellent regioselectivity due to the exceptional shape of the substrate binding pocket. Only secondary alcohols were provided by hydration with OAHs so far. The setback with LIH and CrtCs motivated us to further extend the substrate scope of OAHs by branched alkenes (terpenoids) and to investigate the regioselective hydration of terminal prenyl bonds. A rational approach was envisioned, comprising the screening for initial hydration activity in OAH wild-type enzymes and subsequent improvement of this activity by enzyme engineering.

### 5.4.1 Wild type OAH screening

In preliminary experiments, EmOAH was found to be active in the regioselective hydration of DMDDAC **32** to HDMDAC **33** (Figure 52). The substrate was therefore chosen for detailed investigations of OAH-catalyzed hydration reactions. Initial activity delivered only trace amounts of HDMDAC **33**. Hence, a main goal was to find a biocatalyst with improved properties. In general, two approaches are established: the engineering of the wild-type enzyme or the screening of homologous enzymes from databases to find homologs with more suitable substrate specificity. Our strategy comprised the combination of both approaches starting with the screening of selected OAH wild-type enzymes and the subsequent enzyme engineering in a rational fashion.



**Figure 52:** Regioselective hydration of DMDDAC **32** to HDMDAC **33** catalyzed by OAHs. Initial activity was found for the wild-type oleate hydratase from *Elizabethkingia meningoseptica* (EmOAH).

We explored the sequence space of OAHs based on our in-house Hydratase Engineering Database (HyED).<sup>[115]</sup> Eight OAHs were selected and screened for initial activity in the hydration of DMDDAC. Compared to EmOAH, CgOAH (2.9 x) and LaOAH (1.3 x) even produced higher amounts of HDMDAC under screening conditions. The high activity of CgOAH is not surprising as it is closely related to EmOAH and belongs to the same homologous family (HFam11). The results are similar to what we previously observed in our study on the substrate scope of OAHs regarding smaller fatty acids.<sup>[115]</sup> In general, the OAHs from HFam2, which share only a sequence identity of about 40 % with EmOAH, are less active (Table 44).

**Table 44:** Results of the OAH wild type screening. Depicted is the formed amount of HDMDAC relative to EmOAH in regard to the sequence identity of chosen OAHs. The affiliation of the respective OAH to the homologous family (HFam) is according to the classification in the Hydratase Engineering Database (HyED).<sup>[115]</sup>

	HDMDAC formation relative to EmOAH	Sequence identity to EmOAH	Affiliation to HFam
Cg	287 %	82 %	11
Db	44 %	38 %	2
Em	-	-	11
Gm	43 %	41%	2
Hf	18 %	41%	2
La	133 %	39%	2
Lj	37 %	39%	2
Mp	14 %	35%	1

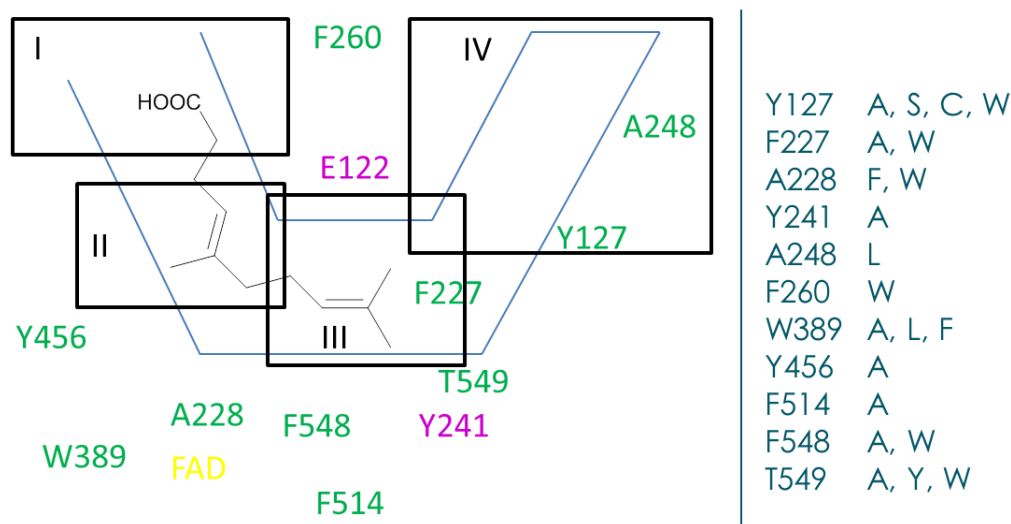
Despite showing only the third best initial activity in this screening for the hydration of DMDDAC, EmOAH was selected as template for enzyme engineering as it represents the model enzyme for the class of OAHs. Several studies about its biochemical properties<sup>[69,73,115,127,128,130,196,197]</sup> and a crystal structure<sup>[120]</sup> are available in literature, which is beneficial for rational enzyme design.<sup>[142]</sup>

#### 5.4.2 Selection of hot spot amino acid residues

EmOAH was selected in the present study as target for a rational protein engineering approach. So far, mutagenesis of EmOAH was limited mainly to positions which were studied with focus on mechanistic or biochemical aspects. For instance, Engleder *et al.* investigated positions G69 and G71 for their influence in FAD binding using alanine mutants and additionally created the variants E122A, Y241F and Y456F for the investigation of the contribution of these polar residues in catalysis. Previous work of our group by S. Kleehammer<sup>[198]</sup> and A. Hunold<sup>[199]</sup> within the framework of the PhD Thesis of R. Demming<sup>[157]</sup> investigated the role of several residues in catalysis or cofactor binding (E122, M123, Y241, T511) as well as substrate recognition (N438, H442) with selected variants. The variants were designed by conservation analysis using the HyED<sup>[115]</sup>. Random mutagenesis of EmOAH by epPCR was performed by A. Hiseni during her PhD work at TU Delft to demonstrate the applicability of a photometric assay (7.5.1).<sup>[178]</sup> The library was screened for hydration of oleic acid and revealed variants with improved activity. However, Hiseni did not characterize any of the generated mutants.

The first report on protein engineering of EmOAH to expand the substrate scope (towards smaller/shorter olefins and fatty acids) was by Demming *et al.*<sup>[130,157]</sup> The authors performed saturation mutagenesis at position A248 and in addition studied the variants F227W, F227Y and F227L. In the present study, we envisioned the first evolution of EmOAH. We targeted the regioselective hydration of terpenoids, an entire new substrate class for fatty acid hydratases, by iterative rounds of saturation mutagenesis (ISM)<sup>[149]</sup>. As the efficient screening with a high-throughput assay is not possible in the present case (7.5) the size of the library should be as limited as possible to reduce the screening effort and enable LC-MS screening. Our approach therefore comprised detailed *in silico* analysis by docking studies. The use of computational methods like docking can give guidance to identify mutational hot spots for the generation of small but smart libraries.<sup>[200]</sup> Docking of oleic acid (natural substrate) and the target substrate DMDDAC revealed the possibility for similar binding modes (Figure 34). The selected binding mode of DMDDAC was considered to be productive based on the orientation in the active site although it was not the energetically most favored (lower binding energy). We analyzed docking runs with the (prenyl) double bond in the proximity to the catalytic amino acid Y241 (Table 42).

Target amino acids were chosen based on comparison of these docking results and contacting residues covering almost the whole active site were selected (Figure 53). The residues were divided into different regions. Residues (e.g. N438, H442) in area I were not selected for mutagenesis, as they are essential for substrate recognition and binding of the carboxylate.<sup>[115,120,157]</sup> On the contrary, we decided to investigate the hydration of terpenoids using the acid substrate DMDDAC. The use of short fatty acid substrates as decoy molecules to enable hydration without the need for a carboxyl head group (as demonstrated by Demming *et al.*<sup>[130]</sup>) was excluded in this study because it entails the disadvantage of low atom economy or increases process complexity. Residues which play a role in FAD-binding were also avoided as OAHs are non-redox enzymes but may be FAD<sub>red</sub>-dependent (comparable to the lycopene cyclase CrtY).<sup>[115,120,201]</sup> Instead, the active site should be shortened by introducing larger residues in area IV to improve substrate binding, which is why residues A248 and F227 were selected. Area II took account for increased steric demand by the additional methyl group of the terpenoid substrate and residues Y456 and W389 were replaced by smaller, non-polar amino acids. However, the main focus for the selection of residues for an initial library was on area III. Here lays the main difference between *cis*-fatty acid or olefin substrates and the higher substituted terpenoid double bonds. Consequently, multiple residues in this area were picked for mutagenesis (Y127, F227, F514, F548, T549) and mainly replaced by smaller amino acids to make room for the prenyl double bond (Figure 53).



**Figure 53:** Schematic view of the active site of EmOAH with docked DMDDAC and position of FAD cofactor (yellow). The residues are projected onto the plane of the double bond (top view on double bond). The catalytic active amino acids are depicted in pink, while other residues selected for mutagenesis are shown in green. The active site was divided in four distinct areas, with different function accommodating the (non-natural) substrate. The list on the right summarizes the variants which were screened within the scope of an initial library.

The initial mutant library was designed as very small focused library to evaluate mutational hot spots which were identified by *in silico* docking studies and thereby to find suitable residues for ISM. Screening of this library was started with (*Z*)-undec-9-enoic acid (Figure 36), the shortest fatty acid accepted by OAHs, to identify residues which are essential for hydration activity and therefore should be excluded from mutagenesis. The results were compared to biotransformations of DMDDAC (Figure 37). We expected that mutation of conserved residues (Table 45) predominantly leads to inactive variants as active-site architecture is commonly at least partially conserved.<sup>[12]</sup> Conserved residues are often essential for the enzymatic function in one way or another (substrate binding, stabilization of transition state / reactive species, selectivity, ...).<sup>[202]</sup> However, redesign of enzyme functions has been achieved in several cases by substitution of (just a few) conserved active site residues leading to new activities.<sup>[12,203]</sup>

**Table 45:** Conservation analysis of OAHs based on the HyED (according to Schmid *et al.*)<sup>[115]</sup> Positions are given according to the EmOAH numbering.

position	conservation	position	conservation
127	not conserved	389	97 % W
227	98 % F	456	99 % Y
228	97 % A	514	98 % F
241	97 % Y	548	98 % F
248	not conserved	549	98 % T
260	not conserved		

Interestingly, while mutations at the conserved residues A228, Y241, Y456, F514 and F548 yielded nonfunctional variants, F227 variants retained activity so that no trend or statement about the influence of conserved residues could be derived from that data. Moreover, W389F and T549A retained or improved hydration ability, respectively, whereas the other mutants of conserved positions 389 and 549 were inactive. Non-conserved residues showed also no clear trend. In particular, the variants A248L, W389F and T549A aroused our interest as they exhibit enhanced hydration activity in the initial screening. Overall, the findings are in accordance with our assumptions. Beneficial mutations were mainly found at positions were the natural substrate and DMDDAC revealed major differences and steric hindrance in docking. Therefore, the positions 248, 389 and 549 were chosen as target residues for iterative site-saturation mutagenesis.

### 5.4.3 Iterative saturation mutagenesis

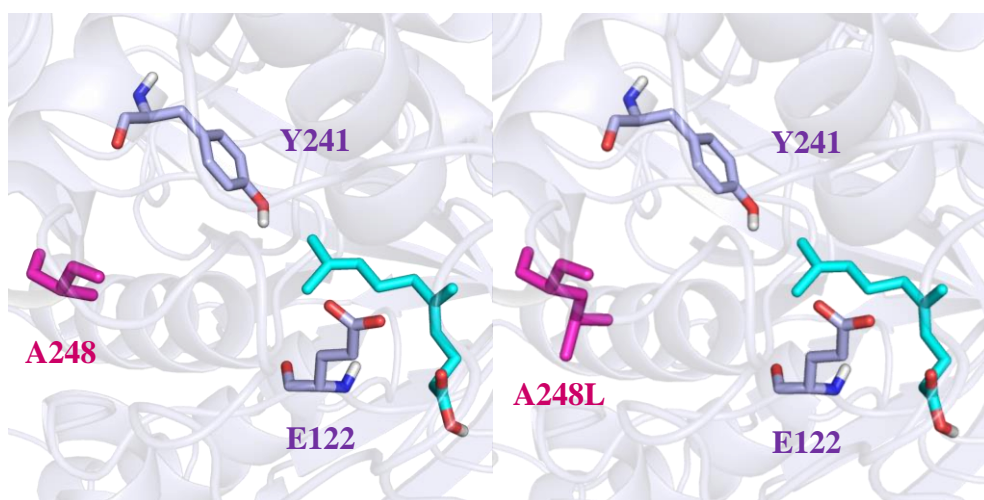
For the activity improvement of EmOAH in the regioselective hydration of DMDDAC to HDMDAC, the first iterative saturation mutagenesis for the evolution of this enzyme was performed. Hot spot amino acid residues were selected based on *in silico* docking studies and subsequent prescreening of a small focused library. As a result, the residues A248, T549 and W389 were selected.

#### *Saturation mutagenesis of position A248*

Saturation experiments were started at A248 as this residue is known to be a specificity-determining position (SDP).<sup>[130,157]</sup> Demming *et al.* investigated the hydration of 1-olefins by EmOAH wild-type using a short-chain carboxylic acid as decoy molecule and observed highly reduced conversion of shorter olefins like 1-octene compared to 1-decene.<sup>[130]</sup> The authors assumed that shorter substrates are not as efficiently bound into the large active site (comparable to what was observed for shorter fatty acids as substrate) and harnessed the diversity of fatty acid hydratases by identifying SDPs using the HyED.<sup>[115,130]</sup> SDPs are proposed to be conserved within functional subfamilies but differ between them.<sup>[130,204]</sup> EmOAH from HFam11 has an alanine (A) at position 248 while HFam1 possess a glutamate (E) and HFam2 a histidine or glutamine (H/Q), respectively. Larger amino acid residues at position A248 seem to be beneficial as in the present study the hydration activity was higher for LaOAH (HFam2) but not the only reason as also CgOAH (HFam11, same as EmOAH) showed higher hydration activity for DMDDAC.

Saturation mutagenesis of position A248 at the end of the alkyl binding pocket improved (C7+C8) or enabled (C5+C6) hydration of 1-alkenes.<sup>[130]</sup> Shorter substrates required larger amino acid residues, which was also the case for hydration of shorter fatty acid substrates without decoy molecules.<sup>[130,157]</sup> In the present study, the results could be confirmed for the hydration of short terpenoids. Mutagenesis at position A248 revealed a wide variety of active mutants (Figure 39). While mutants with large and also charged amino acid residues were less active, best activities were found for variants with mid-size, non-polar and neutral amino acids. The screening of saturation mutagenesis variants at position A248 revealed A248L as the overall best variant (Figure 54).



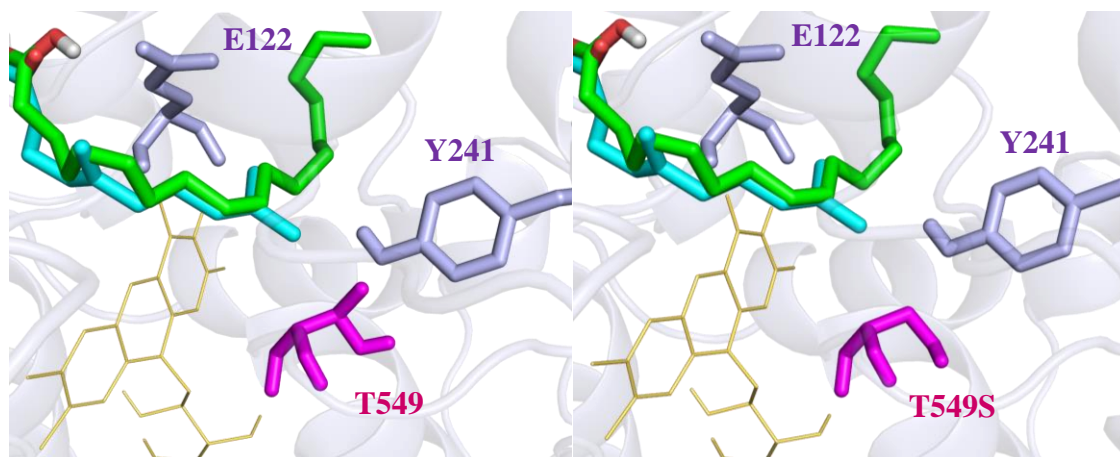


**Figure 54:** DMDDAC (cyan) docked into the crystal structure of the wild-type EmOAH (left). Depicted is the active site with the catalytically active residues E122 and Y241. The picture on the right shows the best variant A248L which was created by *in silico* mutagenesis using PyMOL.

#### *Saturation mutagenesis of position T549*

Mutagenesis at position 549 has not been described so far. The residue is highly conserved (98 % T)<sup>[115]</sup> among OAHs and is in close vicinity to the prenyl double bond of DMDDAC. We intended to generate space for the additional methyl group of branched terpenoid substrates. The screening of the initial mutant library revealed T549A as beneficial variant, whereas larger residues decreased HDMDAC formation (Figure 37). Saturation mutagenesis confirmed that changes at position T549 have a drastic influence on the hydration activity of EmOAH (Figure 41). However, the results of the preliminary screening could not be affirmed and variant T549A even showed less activity compared to the wild type. This was surprising as we expected smaller amino acid residues in general to be beneficial, especially as the variant T549A exhibited 2.5 fold relative activity (compared to the WT) in DMDDAC hydration but almost exactly the same activity in (*Z*)-undec-9-enoic acid hydration. This may be due to the almost full conversion of the fatty acid substrate while only small product concentrations were obtained for the terpenoid. Results for DMDDAC are therefore more prone to errors and even minor deviations in the activity of the WT and the variants result in significant overall differences. On the contrary, the variant T549S exhibits almost 7-fold HDMDAC formation relative to the WT. Remarkably, serine of variant T549S lacks a methyl group compared to the side chain of wild-type threonine compensating exactly the additional steric demand of the terpenoid substrate DMDDAC bearing a prenyl moiety with its additional methyl group compared to *cis*-9-fatty acids (Figure 55).

This highlights the importance and advantages of subtle steric changes introduced by mutagenesis for precise enzyme optimization, as observed previously for example in the development of a selective alkene anti-Markovnikov oxygenase.<sup>[205]</sup>



**Figure 55:** DMDDAC (cyan) and oleic acid (green) docked into the crystal structure of wild-type EmOAH (left) and the best variant T549S (right) which was created by *in silico* mutagenesis using PyMOL. Depicted is the active site with the catalytically active residues E122 and Y241 (purple). Clearly visible is the prenyl group of the terpenoids substrate DMDDAC with its additional methyl group (compared to oleic acid) pointing towards the catalytic Y241 and T549. The variant T549S offers more room to accommodate terpenoid substrates as serine exactly lacks a methyl group in the side chain compared to threonine in the wild type.

#### *Combinatorial mutagenesis of position A248 and T549*

The introduction of subtle amino acid changes at position A248 and T549 led to improved activity. Variants A248L (3-fold) and T549S (7-fold) have been identified as the so far best variants at the respective hot spots. Encouraged by these findings, we continued the evolution of EmOAH towards terpenoid substrates by combining these positions. Therefore, saturation mutagenesis was performed at position 549 based on EmOAH variant A248L and vice versa, T549S was saturated at position 248. In general, effects of amino acid changes (concerning the size and polarity of residues) which were discovered in the previous screenings of single mutations were confirmed. Fortunately, several improved variants were found during the screening of both libraries and additive or even synergistic effects could be observed.

The overall best variant A248I/T549S exhibits additional 2.5-fold activity compared to its parent T549S. In the screening of these two libraries the comparability to the EmOAH WT was not given, as only the direct parents were tested. Therefore, the results were later confirmed by comparison of WT and variants A248L, A248I, T549S and A248I/T549S in independent biological triplicates (Figure 45). The screening of the

double mutants revealed reverse activity for the parents A248L and T549S as the variant A248L was slightly more active in this case, differing thereby from the screening of single mutants and the confirmation in triplicates. Nevertheless, this is in the range of the standard deviations as indicated by the activities of the individual experiments for the single mutants (Figure 45). Interestingly, A248L/T549S obtained by combination of the best mutations at the respective position was not the overall best variant. Instead, A248I/T549S was 1.4-fold more active than the variant A248L/T549S indicating that improvement of the enzyme occurs by “wise-like” adaptation to the new substrate. This further demonstrates that mutations can have additive and non-additive effects as it has been shown in many other enzyme engineering experiments.<sup>[206]</sup>

#### *Saturation mutagenesis of position W389 based on the variant A248I/T549S*

Position 389 was selected as promising residue for evolution based on docking results and preliminary screening of the initial mutant library. Because distinct activity improvement was achieved with EmOAH variant A248I/T549S, saturation mutagenesis of W389 was based on this variant. Non-polar amino acids (large to mid-size) were expected to be beneficial as the preliminary screening using the wild-type enzyme as parent revealed slightly improved activity for W389F but almost total loss of activity for the variants W389L and W389A.

However, no additional activity enhancement was achieved by saturating position 389 in the presence of the A248I/T549S double mutation. While the variant A248I/T549S/W389F retained similar activity as the parent, all other variants were (almost) inactive. This is in agreement with the high conservation of the residue W389. Furthermore, W389 was only a contacting amino acid in the docking studies of (*Z*)-undec-9-enoic acid and 10-methylundec-9-enoic acid but not for oleic acid and *E*-DMDDAC (Table 42). It can be concluded that more space for the internal prenyl group is actually not necessary. Instead, overlay of the docking results rather suggests that the additional methyl group of DMDDAC perfectly fits into a small pocket of the active site in this area, which is filled in the case of oleic acid by two CH<sub>2</sub>-groups (Figure 34). This can also explain why the C12 terpenoid is accepted as substrate although it is one carbon atom shorter when considering the distance of the double bond to the carboxyl head group.

#### 5.4.4 Validation of selected variants & upscaling

Iterative saturation mutagenesis of EmOAH for the regioselective hydration of the terpenoid substrate DMDDAC revealed variant A248I/T549S with significantly improved properties. However, deviating activities were observed during the screening of the initial mutant library and the saturation libraries. This was explained by different physiological states of the whole cell catalyst and the fact that only technical triplicates (three reactions with one single cell preparation from the same expression) were performed to reduce experimental effort. We therefore aimed to confirm our results by performing biological triplicates derived from independent experiments (Figure 45). The preliminary observed trends were approved in this series of experiments and the evolution of EmOAH could be followed in a remarkable way. Starting from only elementary DMDDAC-hydration capability of the wild type over improved product formation for the single mutants an exceptional 22-fold increase was achieved with only two subtle amino acid exchanges.

The evolved EmOAH variant A248I/T549S was used to improve biotransformation conditions. This enabled upscaling of the reaction to a preparative scale to demonstrate the feasibility of enzymatic hydration using engineered OAH variants. Beside that, we were particularly interested in obtaining an authentic product standard and to determine its configuration. We assumed from the peak shape in liquid chromatography (see appendix for chromatograms) that only one diastereomer (*E* or *Z*) is formed as hydration product. We were able to isolate a sufficient product amount (23 mg) to indisputably determine it as *E*-HDMDAC. This is in accordance with our expectations, as *in silico* docking studies were performed from the beginning with the *E*-DMDDAC. In summary, hydration of an *E/Z*-mixture of DMDDAC led to *E*-HDMDAC with good yield and purity and thereby demonstrated the remarkable regio- and diastereospecificity of the engineered EmOAH.

### 5.4.5 Substrate scope

Oleate hydratases were investigated in this study in the light of the regioselective hydration of terpenoid substrates. The excellent selectivity featured by OAHs is said to be bought by narrow substrate specificity. On the contrary, the successful hydration of DMDDAC in the present study is a remarkable extension of the substrate scope of OAHs by branched substrates as earlier attempts to hydrate citronellene failed.<sup>[130]</sup> In view of our mutagenesis strategy towards terpenoids, we wanted to further evaluate the substrate scope of the evolved variants. Several aspects were examined in this regard.

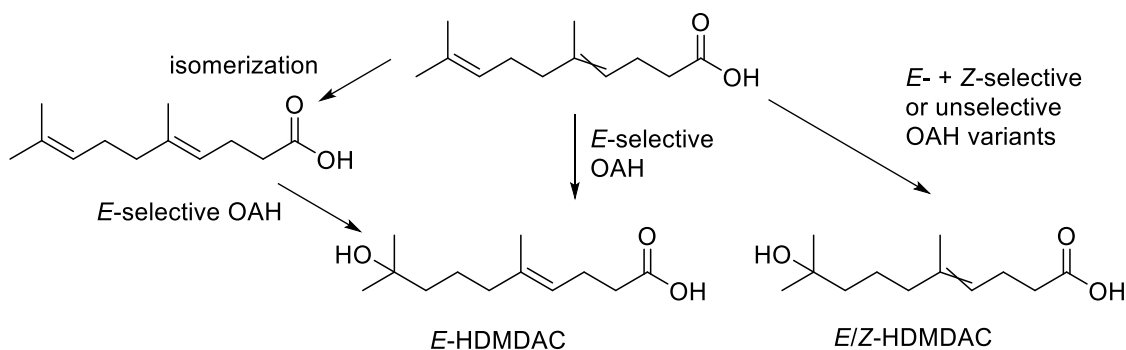
Substrate analogs of DMDDAC, namely the aldehyde DMDDAL, the alcohol DMDDOL and the acetal DMDDIOX were tested to challenge OAH with different functional groups. However, these substrates were not accepted. This is in accordance with previous studies demonstrating that OAHs are limited either to substrates with a carboxyl as head group or a carboxylic acid has to be used as decoy molecule for activation of the OAH.<sup>[115,127,130]</sup> However, first indications exist that this is not true for specific substrates like 1-decene.<sup>[130]</sup> The study by Demming *et al.* demonstrated the hydration of a variety of substrates including several alkenes but also the alkyne 1-octyne.<sup>[130]</sup> In the present study, we could not demonstrate the hydration of a C-C triple bond in the case of undec-9-ynoic acid.

Former studies of the substrate scope also showed that OAHs accept truncated fatty acids like (*Z*)-undec-9-enoic acid.<sup>[115,127]</sup> This substrate was one of the starting points for our rational engineering approach and used along with 10-methylundec-9-enoic acid in the docking studies. We therefore investigated the influence of introduced mutations with the synthetic 10-methylundec-9-enoic acid, which can be seen as a bridge to terpenoid substrates, as it differs only by one additional methyl group at the double bond compared to (*Z*)-undec-9-enoic acid. This can give insights into the importance of the different EmOAH variants for the hydration of branched substrates. We found similar results for the branched fatty acid compared to DMDDAC, as single mutations improved hydration activity slightly but variant A248I/T549S showed significantly (8-fold) increased activity for 10-methylundec-9-enoic acid. Interestingly, variant T549S was not beneficial in this case as it was even less active than the wildtype.

This shows that the introduced mutations optimized the enzyme in general for branched substrates but individual engineering has to be performed for every substrate to obtain a custom-made catalyst.

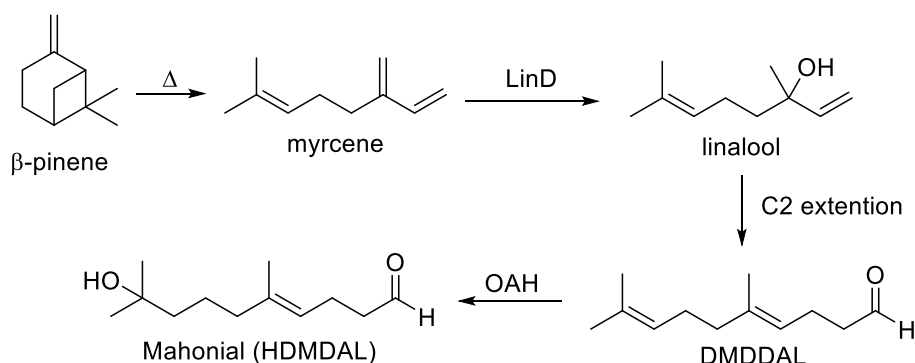
## 5.5 Biocatalytic hydration in the context of organic synthesis and the supply chain in fragrance chemistry

The direct enzymatic hydration of alkenes for the synthesis of alcohols offers considerable advantages over classical chemical methods including non-toxic catalysts and 100 % atom economy. Hydratases also allow for shortened synthesis routes as protecting steps are in general not necessary due to the mild reaction conditions and high selectivities. In this study, the regioselective hydration of prenyl double bonds has been demonstrated for terpenoids which are important in fragrance chemistry. The engineered EmOAH variant A2248I/T549S was found after only two rounds of iterative saturation mutagenesis and preparative synthesis of HDMDAC was demonstrated with this biocatalyst. However, some challenges remained open. This includes application-related aspects which are typical for biocatalysis, like long reaction times for non-natural substrates and low substrate loads leading in sum to poor space-time yields. On the contrary, further improvement can be achieved on the enzyme level by enzyme engineering as for instance by finding more active variants. Moreover, the yield could be increased when both diastereomers of *E/Z*-DMDDAC are used (Figure 56). This can be achieved either by development of an additional *Z*-selective EmOAH variant or by engineering an unselective OAH variant. Isomerization of the substrate would be a third option – albeit unlikely in the case of an isolated double bond. The downstream processing of HDMDAC to the more desired fragrance compound Mahonial is possible by chemical methods or, in future, also biocatalytic methods like the reduction using carboxylate reductases (CARs) are conceivable. Another option is to engineer OAHs for direct hydration of the aldehyde substrate DMDDAL, e.g. by targeting the residues H442, N438, T436 and Q265 which are supposed to play an essential role in recognition of carboxylic acid substrates.<sup>[120,157]</sup>



**Figure 56:** Increasing the yield of the *E/Z*-DMDDAC hydration reaction. Isomerization to *E*-DMDDAC allows in theory full conversion to pure *E*-HDMDAC. Other options are the development of an unselective OAH variant or of a complementary *Z*-selective OAH variant. Both ways lead to a product mixture of *E/Z*-HDMDAC. The strategy to choose depends on the desired product.

Up to now, Mahonial, like many important fragrance compounds, is synthesized starting from petroleum-based materials. Instead, future chemo-enzymatic routes provide alternative access starting from natural sources like  $\beta$ -pinene which can be easily obtained in large amounts from turpentine (Figure 57). The simple thermal conversion to myrcene is followed by biocatalytic hydration using linalool dehydratase-isomerase (LinD), which is also investigated at our institute. The subsequent C2-extension is not possible by biocatalytic methods and the standard acid-catalyzed Claisen rearrangement could be used (Figure 3). DMDDAL can then be hydrated by an engineered OAH to Mahonial. As this hydration step at a late stage of synthesis is the key step in the formation of many fragrance compounds, we believe that the biocatalytic access to such terpenoids is superior to classical methods as it offers efficient synthesis routes by avoiding extensive protecting group chemistry.



**Figure 57:** Envisioned alternative access to Mahonial (HDMDAL) combining biocatalysis and classical chemical methods starting from  $\beta$ -pinene. Pinene is thermally converted to myrcene and the subsequent hydration to linalool can be achieved by LinD. The C2 extension by Claisen rearrangement is realized by acid catalysis. Finally, Mahonial could be obtained by hydration with an engineered OAH.

## 6 Conclusion & outlook

Lily-of-the-valley-type scents are one of the most employed in fragrance industry. This important group of odorants is mainly made up of terpenoids, more precisely hydroxy aldehydes. However, the synthesis of these molecules can be tedious. In this work, the advantages of enzymatic procedures were investigated using hydratases for regioselective hydration of DMDDAL to Mahonial.

So far, no enzymatic reaction was described for this specific case, though several promising enzymes were identified in literature: the limonene hydratase from *Geobacillus stearothermophilus*, the carotenoid hydratase from *Rubrivivax gelatinosus* or *Thiocapsa roseopersicina* and the oleate hydratase from *Elizabethkingia meningoseptica*. In a first step, the enzymes had to be made available and characterized. While it was not possible to prove the existence of a limonene hydratase in *G. stearothermophilus*, the carotenoid hydratases could be heterologously expressed in *E. coli* and activity was detected for the natural reaction, the hydration of lycopene to the corresponding alcohol 1-HO-lycopene or diol 1,1'-(HO)<sub>2</sub>-lycopene.

Oleate hydratase from *Elizabethkingia meningoseptica* (EmOAH) on the other hand was known to be well expressible in *E. coli* and the substrate scope has been extended in several previous studies. However, only secondary alcohols were accessible so far, as only unsaturated *cis*-9-fatty acids and non-branched olefins were converted. In this study, it was demonstrated that EmOAH and homologous enzymes from other organisms were also able to hydrate the prenyl double bond of DMDDAC, a C<sub>12</sub>-terpenoid and the acid-derivative of DMDDAL, forming the corresponding hydroxy acid (HDMDAC). The reaction was not very efficient and product was merely detected in traces, not enough for further characterization and follow-up reactions.

For improvement of the reaction, only few and small amino acid changes were needed. Therefore, *in silico* docking experiments were performed for the natural substrate oleic acid and DMDDAC and the results were compared to identify possible target amino acids for a site-directed mutagenesis approach. Eleven amino acid residues in the active site were selected as promising targets and 21 variants (in certain cases more than one variant per residue) were constructed for an initial mutant library. This preliminary screening yielded variants A248L, T549A and W389F as advantageous. In further experiments iterative site-directed saturation mutagenesis was performed at these promising positions. It was found that the double mutant A248I/T549S was the best



variant tested for the hydration of DMDDAC to the desired product HDMDAC. The reaction was then performed in preparative scale and about 20 mg of *E*-HDMDAC could be isolated and fully characterized. The reaction was found to be strictly regioselective and diastereoselective for the *E*-isomer.

Beside this success, several issues remained open and improvements should be achieved by further investigations. Foremost, the iterative site-directed mutagenesis should be continued to find more active variants which enable the full conversion of DMDDAC in an even shorter period of time and at higher substrate concentrations. This approach could be expanded by random mutagenesis (provided that an efficient high-throughput screening is available) which gives also access to improved variants beyond the active site, for example in the substrate channel or other areas of the enzyme. Although the hydration of a mixture of *E/Z*-DMDDAC lead to only one of the possible two diastereomers and thereby shows remarkable selectivity, on the other hand more than half of the substrate remains in the reaction mixture. To increase the conversion, there are three options depending on whether the target is pure *E*-HDMDAC or a mixture of both:

- 1) finding an unselective OAH variant capable of hydrating both DMDDAC isomers,
- 2) finding an additional OAH variant selective for the hydration of *Z*-DMDDAC,
- 3) isomerization of *Z*-DMDDAC into *E*-DMDDAC and subsequent hydration using the present OAH variant (similar to a dynamic kinetic resolution)

These options can lead to full conversion either giving both product isomers (1 and 2) or only one (3).

Considering an application in perfume industry, HDMDAC is not of interest, as it is not volatile and is lacking the desired lily-of-the-valley odor. Therefore additional steps are necessary to reduce the acid to the wanted aldehyde Mahonial. This can be achieved in a classical chemical way or using the emerging class of carboxylate reductase (CAR) enzymes. However, CAR enzymes have so far not been used for the reduction of this kind of terpenoid. An even more sophisticated approach would be the design of OAH variants by enzyme engineering which are not dependent on the acid moiety and with that except other substrates with functional groups like alcohols or preferable aldehydes allowing the direct hydration of DMDDAL to Mahonial. This could be achieved for instance by targeting the amino acid residues which are responsible for the recognition of the carboxylic acid moiety. Other open questions for industrial application would be

the use of isolated and immobilized enzymes or lyophilized whole cell catalysts to improve storage and handling issues.

From an academic point of view, more investigations on the usage of cofactors are meaningful. This will also bring more mechanistically insights, as the role of FAD remains so far unclear.

Taken all together, considerable endeavors in the field of enzyme engineering of OAH are necessary making a medium- to high-throughput screening indispensable. Therefore, the deep well plate assay developed in this work should be improved, especially concerning the expression of the oleate hydratases in 96-deep well plates, and a faster HPLC method or colorimetric assay would be crucial. With advances in the field catalytic activity and substrate scope, OAH will be a promising enzyme for industrial as well as lab applications.

## 7 Appendix

### 7.1 DNA sequences of wild type enzymes

Oleate hydratase from *Chryseobacterium gleum* (CgOAH)

```
ATGAGTACGATCAATTCAAAATTCGATAAAGTTTTAAATGCCTCTGACCAGTTTGGAAA
TGTA AACACGAGCCGGATTCAAGCAAAGAAGTTCAGATTAACACTCCTGAAAAAACGA
TGCC TTTCTCCGATCAGATCGGAAACTATCAACGTAATAAAGGTATTCC TTTACAGTCT
TACG AAAACAGTAAAATCTATATTGTAGGAAGCGGAATTGCCGGCATGTCTGCCGCTTA
CTAT TTCATCCGCGATGGCCGTGTACCCGGCAAAAATATTATTTTCCTCGATCAGCTTA
ACGTTGAAGGCGGATCTCTTGATGGAGCCGGCAACGCAAAGACGGCTATATTATCCGG
GGAG GAAGGGAAATGGATATGACCTATGAAAATTTATGGGACATGTTCCAGGATATTCC
TGCT TTTGGAGCTGCC TGCACCTTACAGTGTATTGGATGAATACCGCCTGTTAATGACA
ACGAC CCGAATTATTCCAAAGCAAGACTCATT CACAACCAGGGACAGATCAAAGATTT C
AGTA AGTTCGGACTTGAGAAAAAAGATCAGCTGGCAATTGTA AAACTGTTGTTGAAGAA
AAA GAAGAGCTAGACGATCTCACGATTGAAGATTATTTTTCTGAATCTTTCCTCAACA
GTA ATTTCTGGTTCTTCTGGCGTTCTATGTTTGCCTTCGAAA ACTGGCACAGCTTACTG
GAA CTGAACTTTATATGCACAGGTTCC TTCATGCCATTGACGGAATGAAAGACTTCTC
ATGC CTGGTATTCCCGAAATACAACCAATATGACACCTATGTA ACTCCACTAAAAA ACT
TCCT GGTAGAAAAAGGTGTACAGATTCAGTTTAATACCCTTGTA AAAGATCTTGATATT
CAT ATTAACACGGAAGGAAAAACCGTAGAAGGAATTATTACTGAACAAAACGGAGAGGA
AGT AAAATACCAATCAGTAAAGAGGATTACGTTATTGTA ACTACCGGCTCTATGACCG
AAAG TACTTTCTACGGTGATAACAATACCGTTCC TGAAGTTACAATAGATAACAGCAGC
GCAG GACAAAGTGCCGGATGGAAACTCTGGAAAAATCTTGCCGCGAAATCTGAAGTATT
CGGG AAACCTGAAAAATTCTGTAGTCATATTGAAAAATCTTCCTGGGAATCTGCAACCT
TAAC ATGCCGTCCTTCTGCTTTTACAGAGAACTAAAAGA ACTGTGTGTAATGATCCT
TATT CAGGAAGA ACTGCTACAGGAGGAATTATTACCATTACAGACTCCAATTGGGTGAT
GAG TTTTACCTGCAACAGACAGCCGCACTTCCCTACTCAGCCGGATGATATCCTTG TAG
TTT GGGTGTATGCTTTACTGATGGATAAAGAAGGTAATTATATCAAAAAA ACTATGCCT
CAAT GTACCGGAAACGAGATTCTTGCCGAGTTATGCTACCACCTTGGTATAACAGACCA
GCT GGATAATGTTACTGAAAATACCATTGTACGTACAGCATT CATGCCTTATATCACAT
CTAT GTTTATGCCAAGAGCGATGGGAGACCGTCCGAGAGTGGTTCC TGAAGGATGTACC
AACT TAGGTCTTG TAGGACAGTTTGTAGAAACCAATAATGATGTTGTATT CACCATGGA
AAG CTCTGTAAGAACTGCAAGAATAGCTGTTTACAACCTGCTT AACCTTAATAAACAGG
TTCC GGCATCAATCCACTGCAGTATGATATCAGACATTTATTAAAAGCGACTCAGGCT
CTGA ATGACTATAAACCTTTCTTAGGAGAAGGAATTTTAAGAAAAATAT TAAAAGGAAC
TTAC TTTGAACACATTCTGGTAAACCGCCCGGAAGAAAAAGAAGAACACGAATCTTCT
TAAC CAGATTT CAGGAATGGGTAAAAGGAGTAAAAGATTGA
```

Oleate hydratase from *Desulfomicrobium baculatum* (DbOAH)

ATGACCTGGCAAAAAACCGACCGGCCGGTCATGTCCACCCCGTTAGGCAAACAAAAAGA  
GAGCGTTATGTATTACAGTAGTGGCAACTACGAAGCCTTCGCCCGCCCGCGCAAACCCA  
AGGGCGTAGAGAACAAGACGGCCTGGTTTTGTCGGCGCGGGCCTGGCGGCATTGGCGGGC  
GCGGCCTTCTGATCCGCGACGGCCAGATGCCGGGCAACCGGATCACCATCCTGGAGCA  
GCAACAGTTGCCCGGTGGCGCGCTGGACGGCATCAAGGAGCCGAACAAGGGCTTCGTGA  
TCCGCGGGCGGGCGGAGATGGAAGAGCACTTCGAGTGCTTGTGGGACCTGTACCGCTCG  
ATCCCGTTCGCTGGAGATCGAGGGCGCCAGCGTGCTCGACGAGTTCTATTGGCTCGACAA  
GGACGATCCGAACTCCTCTTTGCAGCGCGCCACTGTGAAGCGAGGCGAGGACGCGCATA  
CCGACGGCCTGTTACCCCTCAGTGAGCAGGCGCAGAAGGAGATCGTCAAGCTCTTCCTC  
GCCACCGGCGAGGAGATGGAGAACAAGCGCATAAGCGAGGTATTCGGCAAGGAATTCTT  
GGACAGCAATTTCTGGCTGTATTGGCGCACGATGTTTGCTTTTCGAGGAGTGGCATTTCGG  
CGCTGGAGATGAAGCGGTACCTGCACCGTTTCATCCACCAGATCAAAGGCATGCCGGAT  
TTCTCCACGCTGAAGTTCACCAAATCAACCAGTACGAATCGCTGGTGCTGCCGCTGGT  
GAAGTGGCTGCAGGAACATGGTGTGGTGTTCGAGTACGGTACCGAGGTCACCGATGTCCG  
ATTTTGACATTGCACCCGGACGTAACAGGGCGACCCGCATCCACTGGCGCAAGGACGCA  
GGCGAAGGCGGCGTTCGATCTCGGCCCGACGACCTGGTCTTCATGACCATCGGTTTCGCT  
GACCGAGAACTCCGACAACGGCGACCACCATAACGCCGGCCAAGCTCAATGAGGGGCCTG  
CGCCGGCCTGGGATCTGTGGCGTCGCATCGCCGCCAAGGATCCGGCGTTCGGACGCCCG  
GACGTGTTTCGGCGCGCACATCCCCGAGACCAAATGGGAGTCGGCCACCGTCACTACGCT  
CGACCAGCGCATCCCGAAGTACATCCAGAAGATCGCCAAGCGTGATCCGTTACAGCGGCA  
AGGTGGTGACAGGCGGCATCGTGACCGCGAAGGACTCATCCTGGCTGCTCAGTTGGACG  
GTGAACCGGCAGCCGCACTTCAAAGAGCAGCCCAAGGACCAGATCGTGGTTTTGGGTCTA  
CGCGCTGTTTCGTTCGAGCAACCCGGCGATTACGTGAAGAAGCCCATGCAGGACTGCAGCG  
GCGAGGAGATCACCCAGGAATGGCTCTACCACCTGGGCGTGCCGGTTCGAGGACATTCCC  
GAACTGGCGGCAACCGGTGCGATAACCGTGCCGGTGATGATGCCGTACGTGACCGCCTT  
CTTCATGCCCGGCCAGGCCGGTGATCGCCCGGACGTGGTGCCCGAGGGCGCTGTCAACT  
TCGCCTTCATTGGGCAGTTTCGCGGAATCCAGGGAGCGCGACTGCATCTTCACCACCGAG  
TATTCGTTACGAACGCCGATGGAGGCTGTCTACACACTGCTCGACGTCGAGCGCGGGCGT  
GCCGGAGGTATTCAACTCCACCTACGACATTCGTACGCTCCTGTCCGCGACCGGCCGCC  
TGCGCGACGGCAAGGAGATCGACATCCCGGGTCTGCGTTCGTGCGCAATCTGCTGATG  
AAGAAGCTCGACAAAACCCAGATCGGCGCACTGTTGCGCGAGTTCGGTATTGTCGGGGA  
CGATTGA

Oleate hydratase from *Elizabethkingia mengioseptica* (EmOAH)

ATGGGAATGAATCCGATTACCAGCAAATTTGATAAAGTGCTGAATGCCAGCAGCGAATA  
TGGTCATGTTAATCATGAACCGGATAGCAGCAAAGAACAGCAGCGTAATACACCGCAGA  
AAAGCATGCCGTTTAGCGATCAGATTGGTAATTATCAGCGCAATAAAGGTATTCCGGTG  
CAGAGCTATGATAACAGCAAAATTTACATTATCGGCAGCGGTATTGCAGGTATGAGCGC  
AGCATACTATTTTATCCGTGATGGCCATGTTCCGGCAAAAAACATTACCTTTCTGGAAC  
AGCTGCATATTGATGGTGGTAGCCTGGATGGTGCAGGTAATCCGACCGATGGTTATATC  
ATTCGTGGTGGTTCGTGAAATGGATATGACCTATGAAAATCTGTGGGACATGTTTCAGGA  
TATTCGGGCACTGGAAATGCCTGCACCGTATAGCGTTCTGGATGAATATCGTCTGATCA  
ATGATAACGACAGCAACTATAGCAAAGCCCCTGCTGATTAATAACAAAGGCGAGATCAA  
GATTTTAGCAAATTCGGCCTGAATAAAATGGATCAGCTGGCAATTATTCGCCTGCTGCT  
GAAAAACAAAGAAGAACTGGACGATCTGACCATCGAAGATTATTTAGCGAATCCTTCC  
TGAAAAGCAACTTTTGGACCTTTTGGCGTACCATGTTTGCCTTTGAAAATTGGCATAGC  
CTGCTGGAACGAACTGTATATGCATCGTTTTTCTGCATGCCATTGATGGTCTGAATGA  
TCTGAGCAGCCTGGTTTTTCCGAAATACAACCAGTATGACACCTTTGTTACACCGCTGC  
GCAAATTTCTGCAAGAAAAAGGCGTTAACATCCATCTGAATACCCTGGTTAAAGATCTG  
GACATTCACATTAACACCGAAGGTAAAGTGGTGGAAAGGCATTATTACCGAACAGGATGG  
TAAAGAAGTAAAAATCCGGTTGGCAAAAACGATTATGTGATTGTTACCACCGGTAGCA  
TGACCGAAGATACCTTTTATGGCAATAACAAAACCGCACCGATCATTGGTATCGATAAT  
AGCACCAGCGGTGAGAGCGCAGGTTGGAAACTGTGGAAAAATCTGGCAGCAAAAAGCGA  
AATTTTCGGCAAACCGGAAAAATTTCTGCTCCAACATTGAAAAAAGCGCATGGGAAAGCG  
CAACCCTGACCTGTAAACCGAGCGCACTGATTGATAAACTGAAAGAATATAGCGTGAAC  
GATCCGTATAGTGGTAAAACCGTTACCGGTGGCATTATCACCATTACCGATAGCAATTG  
GCTGATGAGCTTTACCTGTAATCGTCAGCCGCATTTTCCGGAACAGCCGGATGATGTTT  
TGGTTCTGTGGGTTTTATGCACTGTTTTATGGATAAAGAGGGCAATTATATCAAAAAAACC  
ATGCTGGAATGCACCGGTGATGAAATTTCTGGCAGAACTGTGTTATCATCTGGGTATTGA  
AGATCAACTGGAAAACGTGCAGAAAAATACCATTGTTTCGTACCGCATTCATGCCGTATA  
TTACCTCAATGTTTATGCCTCGTGCCAAAGGTGATCGTCCGCGTGTGTGCCGGAAGGT  
TGTAAAAACCTGGGTCTGGTTGGTTCAGTTTGTGGAAACCAATAATGATGTGGTGTTCAC  
AATGGAAAGCAGCGTGCCTACCGCACGTATTGCCGTGTATAAACTGCTGAACCTGAACA  
AACAGGTGCCGATATTAATCCGCTGCAGTATGATATTTCGGCATCTGCTGAAAGCAGCC  
AAAACCCTGAATGATGATAAACCGTTTGTGGTGAAGGTCTGCTGCGTAAAGTTCTGAA  
AGGCACCTATTTTGAACATGTGCTGCCTGCCGGTGCAGCCGAAGAAGAAGAACACGAAA  
GCTTTATTGCCGAACACGTGAACAAATTTTCGCGAATGGGTTAAAGGCATTTCGTGGCCTC  
GAGCACCACCACCACCACCCTGA

Oleate hydratase from *Gemella morbillorum* (GmOAH)

ATGTATTATTCAAACGGAAATTATGAAGCATTGCTAAACCAAAAAACCAGCAGGAGT  
TGATAAAAAATCGGCTTATTTAGTAGGTTCTGGACTAGCTTCTCTTGCGGCAGCAGCCT  
TTTTAGTTCGTGATGGACAAATGAAAGGAGAAAGAATCCATATTCTTGAAGAACTACCT  
ATCGCAGGAGGTAGCTTAGATGGTATTATGAATCCAACCTCGTGGATTATTATTTCGTGG  
TGGTCGTGAAATGGAAGATCACTTTGAATGTCTTTGGGATCTTTTCAGATCTATTTCCTT  
CATTAGAGGTAGAGGATGCATCTGTTTTAGATGAGTTCTACTGGTTAAATAAAGAAGAT  
CCTAACTTTTCAAAATGTCGTGTTATTGAAAACCGTGGACAACAAATTCCTAATGATGG  
ACTTTTTGCATTATCTGATAAATCAAGTCAGGAATTAGTAAAACCTATATTTGACTTCTG  
AATCAGAACTTCAAGGTATGAAAATAACAGATTTCTTTAGTAGTGAATTTTTTTGAGTCT  
AATTTTTGGACTTATTGGGCAACAATGTTTGCTTTTGAAAAATGGCATTTCAGTTGCTGA  
AATGCGTTCGATATATGTTGAGATTTATTCATCATATTAAGGACTTCCTGATTTATCAG  
CACTTAAATTCACTAAATATAACCAATATGAATCATTAGTTCTACCATTAGTTAAATAC  
TTAGAAAGTCATGGTGTAACTTTTGAGTACAATACTGTAGTTAAAGATATTAAGTTGA  
AAAACAAGGAAAAGATCTAGTAGCTAAGCATCTGATCTTAGAAGTAAACGGAGAACCTG  
TTGTTAGAGAATTAAGTGAAGATGACCTAGTATTTGTTACAAATGGTAGTATCACTGCA  
TCTACAACATATGGTAACAATGATACTTCAGCTCCAGTCAGCAAAGAACTTGGTGGAGC  
TTGGCAACTTTGGAAAACTTAGCTAAACAAGATGAACGTTTTTGGACGTCCAGAAGTAT  
TTTGTGAAAATCTTCCAGATCAAAGTTGGTTTGTTCGGCTACGACGACTGTAACAGAC  
AAACGAATCGCAGAAATATATAGAAAAAATTTGTAAACGTGATCCTTATGCAGGTAAAGT  
AGTAACTGGTGGTATAGTAACAGCAGGACTCAAATTTGGATGATGAGCTTTTACACTTA  
ACCGTCAACCTCACTTTAAATCTCAATCAAAGATGAATTAGTAGTTTGGATTTATGGA  
TTATATTCAAATATTTTCAGGTAACCTATATAAAAAACCAATAGAAGCTTGTACAGGTAT  
TGAAATCGCAGAAGAATGGCTTTACCATATGGGGTTCCCTGAACAATTTATTCACGAAT  
TTGCAAGCAAAGGTTGTAGCACAGTTCATGTTATATGCCTTATATAACAAGTTACTTTT  
ATGCCAAGACACGATGGTGTAGACCATTAGTTATCCCGAAGGATCTAAAACTTAGC  
ATTTATAGGTAACCTTCTCTGAAACACCTCGTGATACTGTCTTCACTACAGAATATTCTG  
TTCGTACAGCTATGGAAGCAGTATACACATTAAGTAGATGTTGATAGAGGTGTACCTGAA  
GTATTTAATTCTGCATATGATTTACGTGTCATGACTAAATCAACTAATCGCCTAATGGA  
TGGTAAAAAACTTGAAGAAGTTAAACTACCTTTCTTTGCCAAAGCAATTAGTAAACGTG  
TATTGAAAAAATCAAAGGAACATATATTGAAGAGATCCTAAAAGATGCAGGATTGATT  
TAG

Oleate hydratase from *Holdemania filiformis* (HfOAH)

ATGCTAAGAACGGTTAGAAAAAGAGAAACAGGAGGAATAAAAAATTATGTATTACAGCAA  
TGGCAACTACGAAGCGTTCGCGCATCCGCGCAAGCCGGAAGGCGTTGACAACAAGTCCG  
CATATCTGGTCTGGCTCCGGACTGGCGGCGCTGTCCGCGGCGTGTTCCTTAGTTCGTGAC  
GGACAGATGAAGGGCGAACACATCCATATCCTGGAGGAAGGGAAGCTGCCCGGCGGCGC  
CTGCGACGGCATTAAAGGATCCGCAGAAAGGTTTCATCATCCGCGGCGGCGGGAAATGG  
AAAATCACTTTGAATGTTTGTGGGATCTGTTCCGTTCGATTCCGTCTTTAGAAACGGAA  
GGCGTTTCCGTA CTGGACGAATACTACTGGCTGAACAAAGATGACCCGAACTATTCGCT  
GATGCGCGCGACGATCAACCGCGGCGAAGATGCGCATAACCGACGGCAAATTCGGTCTGA  
GCGACAAAGCGGCGATGGAGATCATGAAGCTGTTCTTTACGAAGGACGAGGATCTCTAT  
GATAAAAAGATTACCGACGTCTTCGATGACGAATTCTTCTCCAGCAATTTCTGGCTGTA  
CTGGCGCACGATGTTTCGCGTGTGAGGACTGGCATTTCAGCTCTGGAAATGAAGCTGTATA  
TCCAGCGCTTCATCCATCACGTCCGCGGCGCTGCCGGATTTTTCCGCGCTGAAATTCACG  
AAGTACAATCAGTACGAATCCCTGATTCTGCCGATGGTAAACTATCTGGAAGAGCATCA  
CGTGCAGTTTGAATATGATACGCGCGTGACCAATGTTGAATTTGAAATCACCGGAGAGA  
AGAAGGTGGCCCGCCGGATCGTCTGTGTCCGGGAAGGGCAGGAACTACGATTCCGCTG  
ACGGAACGATCTTGTCTTCGTGACCAACGGCAGCTGCACGGAAAATTCCTCGTTGGG  
CGACAACGATCACGCGCCGCAGTTTAAACGATCAAAATGGCGGATGCTGGCAGCTGTGGC  
GAAACATCGCGGCGCAGGATCCGGCGTGTGGCCATCCGGATAAATTCCTGCACGCAGACC  
GATAAGACTTATTGGGAATCCGCGACGGTCACCACGCTGGACGACCGCATTCCGCCGTA  
TATCCAGAAGATCTGCAAGCGTGATCCGTTTCAGCGGCAAGGTCGTGACCGGCGGCATCA  
TCACGGTCAAGGATTCCAACCTGGCTGATGAGCTACACGCTGAACCGCCAGCCGCATTTTC  
AAGGATCAGCCGAAGGATCAGCTGGTCGTCTGGGTTTACGGCCTGTTTCGGCGATGTGCC  
GGGCAACTACGTAAAAAAGCCGATGCGCGAGTGCACCGGTACGGAAATCACCGAGGAAT  
GGCTGTACCATCTGGGAGTGCCGGAGGATCAGATTTCATGACATGGCCGTCAATTCGCC  
CGCTGCATCCCCTGCATGATGCCGTATATCACCGCCTTCTTCATGCCGCGCACCGCGGG  
CGACCGGCCGAAGGTTGTACCGGAGGGCTGCGTGAACCTTCGCGTTCATTGGCCAGTTTG  
CCGATACTGTCCGCGACACGGTGTTCACGACGGAATATTCGCTCCGGACCGGCATGGAA  
GCCGTTTATACGCTCTTAGACGTCGACCGCGGCGTGCCGGAGGTCTTCAATTCGGTCTA  
TGACGTCCGTGTGCTGATGGATTCAACCTATAAGATGCTGGACGGGCGCAAGCTGGAAG  
ACATCAAGGTGCCGTGGCTGGTTCACTGGATTGAAAAGAAGGGCCTGAAAAAGATCCAG  
GGCACGGTGATTGAGGAACTGCTGCAAGAATATCATCTGATTTAA

Oleate hydratase from *Lactobacillus acidophilus* (LaOAH)

ATGTATTATTCCAATGGTAATTACGAAGCATTTCAGATCCTAAGAAACCTGCAGGCGT  
GGACAAGAAATCTGCTTATATTATTGGTAGTGGTCTAGCTGGTCTTTCTACAGCCGTAT  
TTTTAGTACGAGATGCACAAATGAAAGGTGAAAATATTCATATTTTAGAAGAATTACCA  
GTTGCTGGTGGTTCTCTTGATGGTGCAGATCGTCCCAATGCGGGCTTTGTTGTTTCGTGG  
TGGACGCGAAATGGAGAATCATTTTTGAATGTTTATGGGATATGTACCGCTCAATTCCAA  
GTTTGGAAAGTTCAGGCGCATCGTATCTTGATGAATATTATTGGCTTGACAAAGAAGAT  
CCCAATTCATCAAATGTTCGTTTAAATTTATAATCGTGGAGATCGTCTTCCAAGTGATGG  
ACAATATGGTTTAGGCAAATGTGCTAATGAAATTGTCAAGTTGATTATGACCCCTGAAA  
AGGAAATTGAAGGGCAAACCTATTGAAGAATTTTTTCAGTGATGAGTTCCTTTAAGACTAAT  
TTCTGGACATACTGGTCAACAATGTTTGCTTTTGAAAAATGGCATTTCATTAGCCGAAAT  
GCGCCGTTATGCAATGCGCTTTATTCATCATATTGATGGGTTGCCTGATTTTACTGCCT  
TAAAGTTTAAATAAGTATAATCAATATGAATCAATGGTTAAACCACTTCTTGCATATCTT  
AAAGATCATGGTGTGCAATTTGAATATGACTGCCATGTTAAGAATGTTCGAGGTAGATCA  
TGAAGGCGACAGTAAAATTGCCAAGAAGATTGTTATGACGCAGAATGGCAAAGATAAAG  
AAATTGATTTAACACATAATGACATCGTCTTTGTAACCTAACGGTTCAATTACTGAAAGT  
TCTACTTATGGTGATCAGAATACTCCAGCTCCAATTACTAATGCTAAAGGTGATTCATG  
GAAGTTATGGGAAAATTTGGCTAAGCAAGATCCAGCTTTTCGGTCATCCAGATGTATTCT  
GCGAAAACCTACCAGAACGTAGTTGGTTTTGTTTCCGCAACTGCTACATTGGAGAATAAG  
AAGCTTGCACCATATTTTGAGCGCTTAACCAAGCGCAGTTTGTATGATGGTAAGGTTAA  
CACAGGTGGTATTATTACTATCGTTGATTCTAACTGGGAACTTAGTTTTACTATTACC  
GTCAACCACATTTTAAGAGTCAAACCCAGACCAAATTGTTGTTTGGATTTATGCACTT  
TATTCAGATACTGAAGGTAACCTATATTTAAAAAGAGAATTGTTGATTGTACTGGCAAAGA  
AATTGCAGAAGAGTTGCTTTACCCTTAGGTGTTCCCGAAAGCCAAATTAGTGAATTGG  
CCAGCGAGGAAAATATGAATACCGTACCAGTTTATATGCCATATATTACTAGTTACTTC  
ATGCCTCGTCGTGATGGTGATCGTCCAGATGTTGTGCCAGAAGGATCAATAAACCTTGC  
CTTTATTGGTAATTTTGCAGAATCTCCAACAAGAGATACTGTGTTTACTACTGAGTATT  
CTGTCAGAACAGCTATGGAGGCTGTTTATACATTACTTAATGTTGATCGTGGCGTACCA  
GAAGTATTTGACTCAATTTATGATATTCGCCAGCTTTTACGTGCTATGTACTACATGTC  
AGACAAGAAGAAGCTAGCAGATCAAGATATGCCTCTTCTGAAAAGCTTGCCGTAAAGA  
CAGGAATGAGAAAGATTTAAAAAGACTTGGGTAGAAGAACTACTTAAAGAAGCAAATTTA  
GTCTAA



**Oleate hydratase from *Lactobacillus johnsonii* (LjOAH)**

ATGTA CTATTCAAAGGAAACTATGAAGCTTTTGC GGATACTAAAAAGCCGGCTGGTGT  
AGACAAGAAATCTGCTTACATAATCGGTAGTGGATTAGCGGGCCTTTCAGCAGCATTCT  
TTTTGGTTTCGAGATGCTAAGATGTCTGGTGATAAAAATTCATATTTTAGAAGAATTGGAT  
GTTGCGGGCGGCTCTCTAGATGGGACTAACCGACCTAATGCTGGTTTTGTAGTTCGAGG  
TGGTAGAGAAATGGAAGACCACCTTGAATGTATGTGGGACATGTATCGATCAGTTCCAT  
CATTGCGGATAACCAGGAGCTTCTTACTTAGATGAATATTACTGGCTTGATAAAGAGGAT  
CCTAACTCCTCGAATTGTCGATTGATTACATAAGCGTGGTAATGAATTGCCAACAGATGG  
ACTATATCAATTGGGTTTCGCATGCTAGTGAAATTGTTAAGTTAGTTTTAACACCAGAAT  
CTGAAATTC AAGGAAAAACAATTGAAGAATGGTTTTACCTGAATTTTTTGAACCTAAT  
TTTTGGACTTATTGGTCAACGATGTTTGCTTTTGAAAAATGGCATTCTCTTGCAGAGAT  
GAGACGCTATTGTATGCGTTTCATTCACCATATTGATGGTTTACCTGACTTTACTGCCT  
TGAAGTTTAACAAATATGATCAATATGATTCTATGGTGAAGCCACTACTTAAATATCTA  
AAAGACCATGGCGTTAAATTTGAATATGGTGTTC AAGTAGAAAATGTATTAGTTGACCA  
TGAAGGAAATGAAAAAGTTGCTAAAAAGATTGTGATGAAGAAGAACGGTAAGCAAGAAG  
ATATTGATTTAACTGAAGATGACATTGTCTTTGTA ACTAATGGATCAATTACTGAAAGC  
TCTACTTATGGGAACCAAACCTACGCCTGCGCCAATTACTCATGCTAAGGGTGGATCTTG  
GAAGTTATGGGAAACTTAGCTAAACAAGATCCGGCATT TGGTCATCCTGAGGTCTTTA  
GTGAAACTTACCTGAAAAATCATGGTTTTGTTTCTGCAACTACAACGTTAAAGAATAAA  
AAAGTAGCCCCTTACTTTGAAAGATTGACTAAACG TAGTTTGTATGATGGAAAAGTAAA  
TACTGGTGGAAATTATTACGGTTACAGATTCAAAT TGGGGTCTAAGTTTTACTATTCACC  
GTCAACCTCATTTTCCAACCTCAAAGCCAAATGAGATTGTAGTCTGGATCTATGCTTTG  
TATTCAGATACTGAGGGGAATTATATTAAGAAAAAGGTAGTCGACTGTACTGGACAAGA  
AATTGCTGAAGAAATGCTTTATCATTTGGGTGTGCCAGAAAGTGAATTAAGAATTGT  
CTTCAGAAGAAAATATGAACACCGTTCCGGTATATATGCCTTATATTACTAGTTACTTC  
ATGCCTCGTCATGATGGTGACCGGCCAGCTGTTGTACCAGAAGGATCAAAAAATTTAGC  
ATTTATTGGTAACTTTGCTGAAAGTCCTACTAGAGATACAGTGTTTACGACAGAATATT  
CAGTTAGAACAGCTATGGAATCTGTTTATACACTACTTGATGTAGATCGAGGAGTTCCA  
GAAGTTTGGAGTTCTGTTTATGATATCCGGGAATTACTTCGTGCTATGTATTACATGTC  
TGATAAAAAGAAATTAGCAGACCAAGAAATGCCATTACCAGAAAAGTTAGCTGTAAAAG  
CTGGTATGAAGAAGATTAAGGGAACATGGATTGAAGAATTGCTTGAAGAAGCAAATCTA  
ATTTAA

Oleate hydratase from *Mucilaginibacter paludis* (MpOAH)

ATGAATAATCTAAATAAAAAACGGGTTAACCCGACTCAACACCGGAAGGTCTATCTTAT  
CGGCGGCGGTATCGCATCTTTAGCAAGTGCTGCTTATTTTCATAAGAGATGGCCAAATTA  
ATCCTGATAATATCATAATTTACGAAGAGCTGAATGTTGCTGGCGGCAGTCTTGATAGT  
GCAGGTAACCCGGAAGATGGTTACGTGATGCGTGGCGGCAGGATGCTGAATTTTCAGCTA  
CCTGTGTACGTATGATCTTTTTTTCATTTATTCCATCGCTGTCCGATCCGGGCATCATGG  
TACTGGATGAAATTAAGGCCTTCAACCGGAACATTAAAACCCATTCACAGGCCCGGGTA  
GTGGAAAACGGCCATATATTAGATGTTTTCGTCAATGGGATTTAGCAATAAGGACAGACT  
TGATCTGATCGAGATGATGGCTGTAGGCGAAAACCATTTGGGCATCAAACGCATCAATG  
AATGGTTTTCGCGCCGGAATTTTTCAAGACCAACTTCTGGTTTATGTGGGATACCATGTTT  
GCCTTTCAGCCCTGGCACAGCGCTGTGGAGTTTAAACGTTATCTACATCGTTTTTATCCA  
TGAGTTTAAACGGATCAATACTTTAGCTGGCGTAGACAGGACGCCCTATAACCAATTTG  
ATTCTTTAGCCAAGCCGTTAATCAACTGGTTAAAACAGCAGGGCGTTCATTTTTCAAATA  
GGTGTACGCGTAACCGATCTGGATTTCTCATTAAACGGGCGAAAAGCTGACCGTGCAGCA  
AATTCATTTAATTGATGGCGGAAAAAAGAAATCTGTAAAGATTGGCGTAAACGATATGG  
TACTGGTAACCATTGGCTCAATGACTGCTGATTCAAGTTTAGGATCGATGCATGCTGCC  
CCCGAACTCATTACTGATAAACGTGACGGCAGCTGGAAATTTATGGGAAGCCATTGCAAA  
ACACAGTGACCAATTCGGTCATCCTTCGGTATTTGACAACCATGTTCGATGAATCAAAAT  
GGGAATCATTTACGGTAACTTGCCAGGGTACAGACTTTTTTGAACGGATGAAAACGTTT  
TCGGGCAATGACGCCGGAACCGGCGGCCTGGTAACTTTTAAGGACTCCAATTGGCTGAT  
GTCAATTGTTTTAGCCTACCAGCCTCATTTTCATCGGCCAGCCTGATGACGTGACCGTTT  
TTTGGGGGTATGGCTTGTTCCCCGATAATGAGGGAAACTTTGTAAAAAAGAAAATGGCC  
GACTGTACCGGTACCGAAATACTGACCGAATTATTTTCCCACCTTCAATTTGACACGTT  
AAACGACCGGCTACTGAAAACCGCTAACTGCATTCCCTGCATGTTGCCGTACATCACCA  
GTCAGTTTTTAATACGGGCACCGGGCGACAGGCCGCAGGTTATACCAGAAATATCGGAA  
AACATGGCCTTTATCGGGCAGTTCGCAGAAGTGCCCGACGACGTGGTATTCACCGTAGA  
ATATTCAGTAAGAACTGCGCAGACCGCAGTTTACGGATTATTGGAACCTCGATAAAAAAA  
CGATTCCAATGTATAAGGGCGATCATCATATTGATGTATTGTTTGGATGCAATGAAAAC  
ATGATTTTCATAA

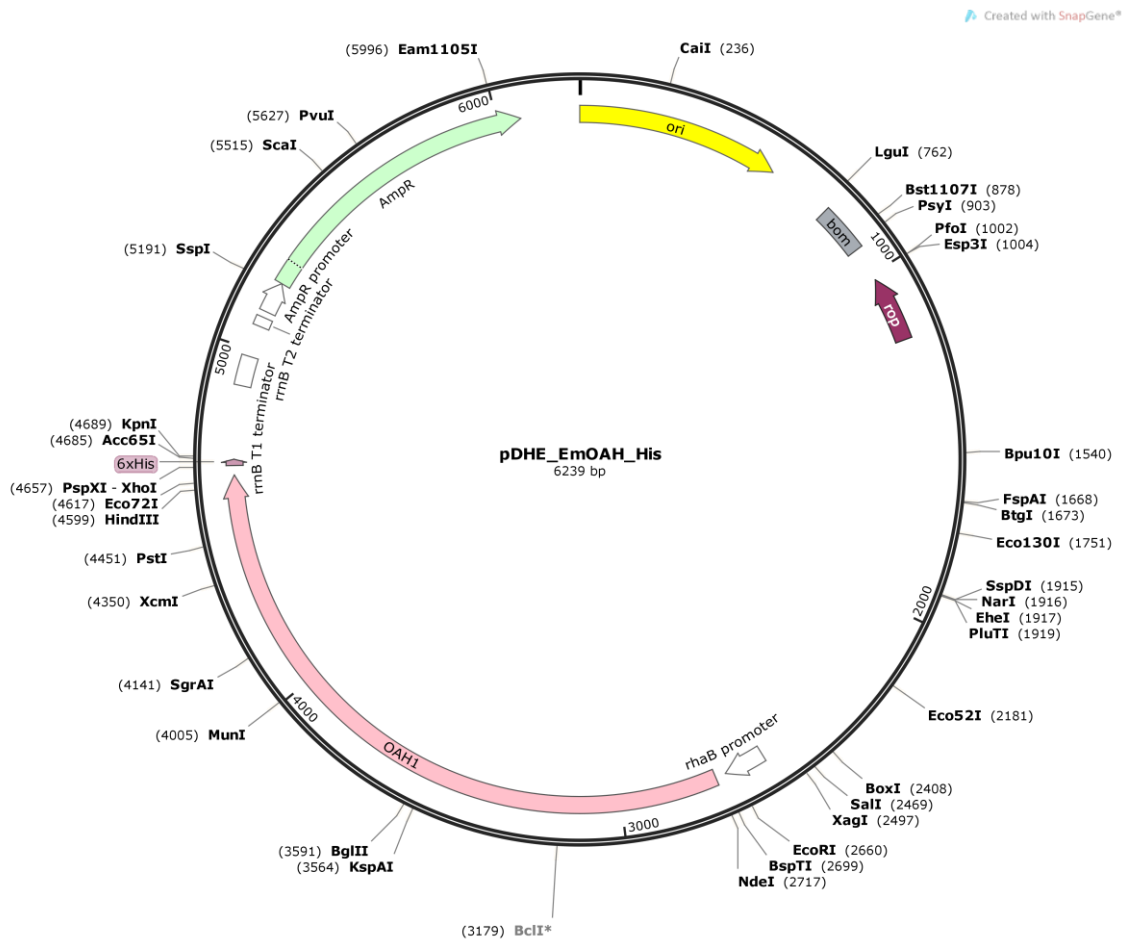
Carotenoid hydratase from *Rubrivivax gelatinosus* (RgCrtC) – codon-optimized

ATGCGTGCAGCAGAAAGCGGTGCAGATGCACGTGTTCTCGTCCGGATGATCGTCTGGAACC  
 GGCAGATGCCCTGCCGGTGATGCGGGTGAACCTGCGTGCACCGGTTCCGGGTGATGGTG  
 GTAGCGCAGTTCGTCCCTGGTGCACGTCTGGATGTTCTGGTTCGGCCTGGTCTGGTT  
 GATGAACCTGCAGCAGGTCCGGTTCCTGGTGGTCTGACGCTGCTCCGGGTGCAGGTGC  
 TGCAGATGGTGGTGCATGTGCGTCCGGCAGGCGGTCTGATGCCGATGGTGCACCGCGTT  
 TTGATCAGCCGTTCCCTCCGGGTGGTTATCTGTGGTGGTATGTTGATGCAGTTAGTGAT  
 GATGGTCTGCATGGTCTGACCTTTATTGCATTTGTTGGTAGCGTTTTTCAGCCCCTATTA  
 TGCATGGGCTGGCGGTCCGAATGCAGATCGTGCCGATCCGGAAAATCATTGTGCACTGA  
 ATATTGCACTGTATGGTGCATGCCGTAACCGTTGGACCATGACCGAACGTGGTCTGCTCGT  
 TGGATGCGTCTGATGCCGTGATGAATTTGTTATTGGTCCGAGCCGTCTGCATTGGGATGG  
 TGAAAGCCTGCTGGTTGAATTTGATGAAATGGTGTTCGGATTCCGCGTCTGTTAAAG  
 GTCGTGTTCTGTGTTTGGCCGAAAGCACTGTGTCTGTTTGTACCAGCCTGGATAGCGGT  
 GGTCGCCATCGTTGGGGTCCGATTGCACCGTGTAGCCGTATTGAAGTTGAACCTGGATAG  
 CCCGAATGCACGTTGGAGCGGTGCATGCCTATGTTGATAGCAATGAAGGTGATGAACCGA  
 TTGATCGTCCGTTTCTGTGAATGGGATTGGAGCCGTGCAACCATGGCAGATGGTTC AAC  
 GCAGTTATTTATGATGTTTCGCCAGAAACGTGATGGTGCATCGTGTATTATGCAGAACGTTT  
 TCTGCTGGATGGTAGCACCAGAAAGCTTTGAAGCACCGCCTCGTCAGCCGCTGCCGACCA  
 CACTGTGGCGTATTGATCGTACCATGCGTACCGAACCGGGTGTTCGGGCATTTGTGGAA  
 CAGACCCTGGAAGATACCCCGTTTTATGCACGTAGCATGGTTTCGTAGCGGTCTGCTGGG  
 TGAAGTTGTGACCAGCGTTCATGAAACCATGCTGCTGCCACGTGTTATTACTGCCGG  
 TGCGTCTGATGCTGCCGTGGCGTATGCCTCGTCTGATGCTAA

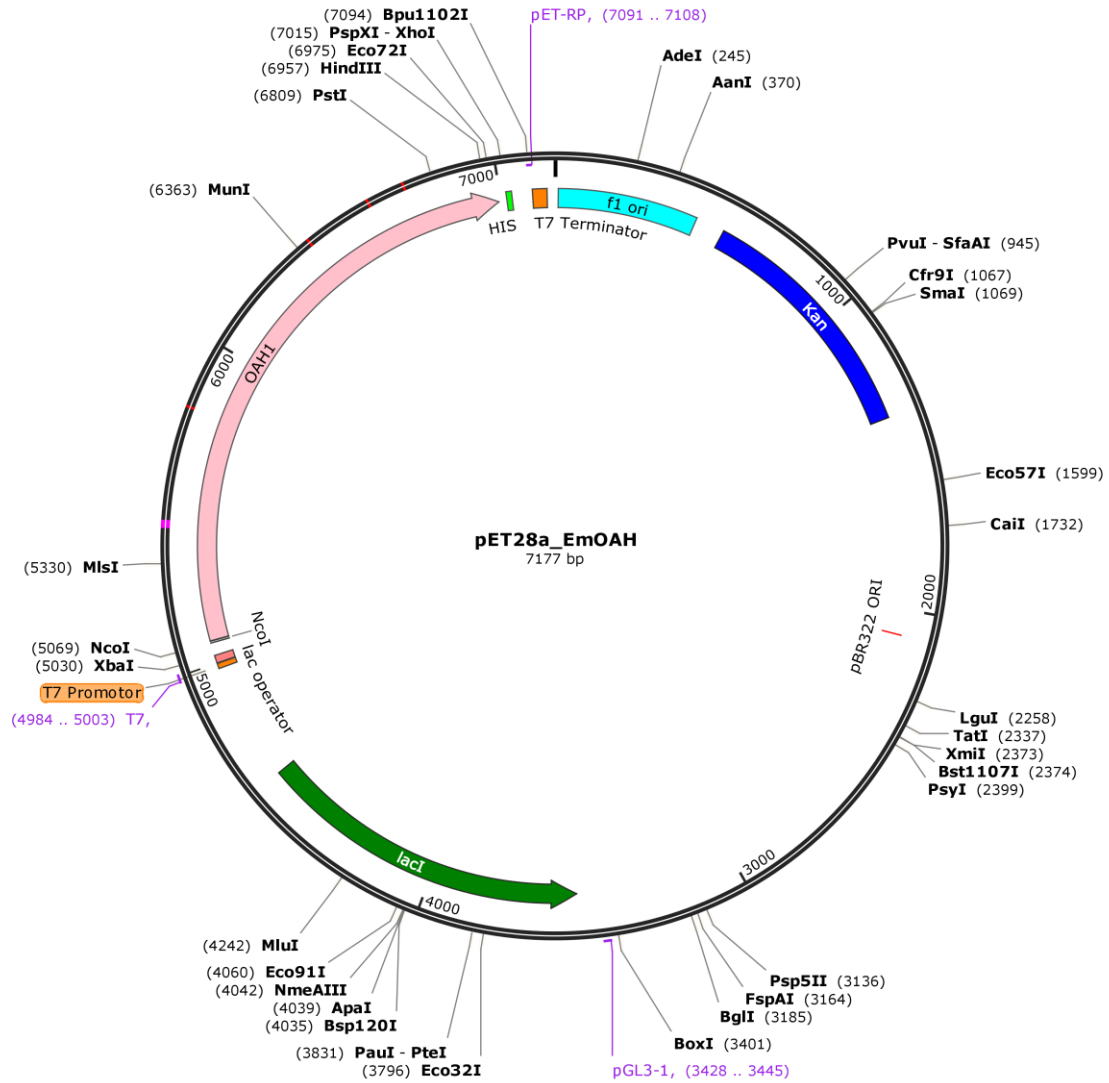
Carotenoid hydratase from *Thiocapsa roseopersicina* (TrCrtC) – codon-optimized

ATGCGTGCAGCAGGTATTCTGACACCGGGTGCACCTGTGGGCACCGGGTCCGAGCGATAC  
 CCGTGATGAACGTCGTCATGATGCAGGTCGTCTGCAGCCTGCACTGCCTGGTGATGGTC  
 GTGGTCCGCTGCGTCCGGGTGTTACCCGTATGGAAGGTCTGCTGCATGCAAGCTCAGTT  
 GCACAGCAGGGTGCCGGTGCACCTGAGCGGTCTGGTGAACGTGCAAGCGGTAGCCGTGG  
 CACCGATGGTGGTGCATGTTGGCACTCCGGGTGGTAGCGATCCGGCACGCGGTCCGCGTT  
 TTGATCTGCGTATTACCCCTGGTGGTTATCTGTGGTGGTATCTGGATGCACTGAGTGAT  
 GATGGTGCATGGTCTGACCATTATTGCAATGCTGGGTAGCGTTTTTTAGCCCGTATTA  
 TGCATGGGCACGTCGTCGTGGTAATCCTGATCCGCTGAATCATTGTGCACTGAATGTTG  
 CACTGTATGGTAAAGCAGGTAAACCGTTGGACCATGACCGAACGTGGTCTGTAAGCACTG  
 CGTCAGGCTCCGGGTGCTCTGGATATTGGTCCGAGCCATCTGACCTGGGATGGCACCGC  
 ACTGACCATTGATGTTAATGAAATTACCGCACCGATTCCGAGCCGTGTTCTGTGGTCCGA  
 TTCGTGTTATTCGGGCAGCAGTTAATGCACGTGAATTCACCCTGGACCCTGCAGAACGT  
 CATGTTTGGTGGCCGATTGCACCGATTAGCCGTGTTGAAGTTGATCTGGAAAAACCGGC  
 ACTGCGTTGGAGTGGTGCATGGCTATCTGGATAGCAATCGCGGTGAAGAACCCTGGAAG  
 ATGCATTTTCAGTGTGGGATTGGAGCCGTGCAAATACCCCGAGCGGCACCACCATGCTG  
 TATGATGTTACCGCACGTCATGGTACAGGTGCAAGCCTGGCCCTGCGTTTTAATGCCAG  
 CCGTGAAGTTGAAGAATTTCCGCCTCCGCTCGTGTTCGTCTGCCGACCACCGGTATTT  
 GGCGTATTAACGTGGCACCCAGTGTGAAGCAGGTGCATCAGGCACGTGTTGTTGAAACA  
 CTGGAAGATACCCCGTTTTATGCACGTAGCCTGGTTGAAACCCGTCTGGCAGGCGAAAC  
 CGCAACCTGTGTTTCATGAAAGCCTGAGCCTGGATCGTTTTTGCAAGTCCGGTTGTTT CAGC  
 TGATGCTGCCGTTTTCTGATGCCACGTGTTGGTGGTTAA

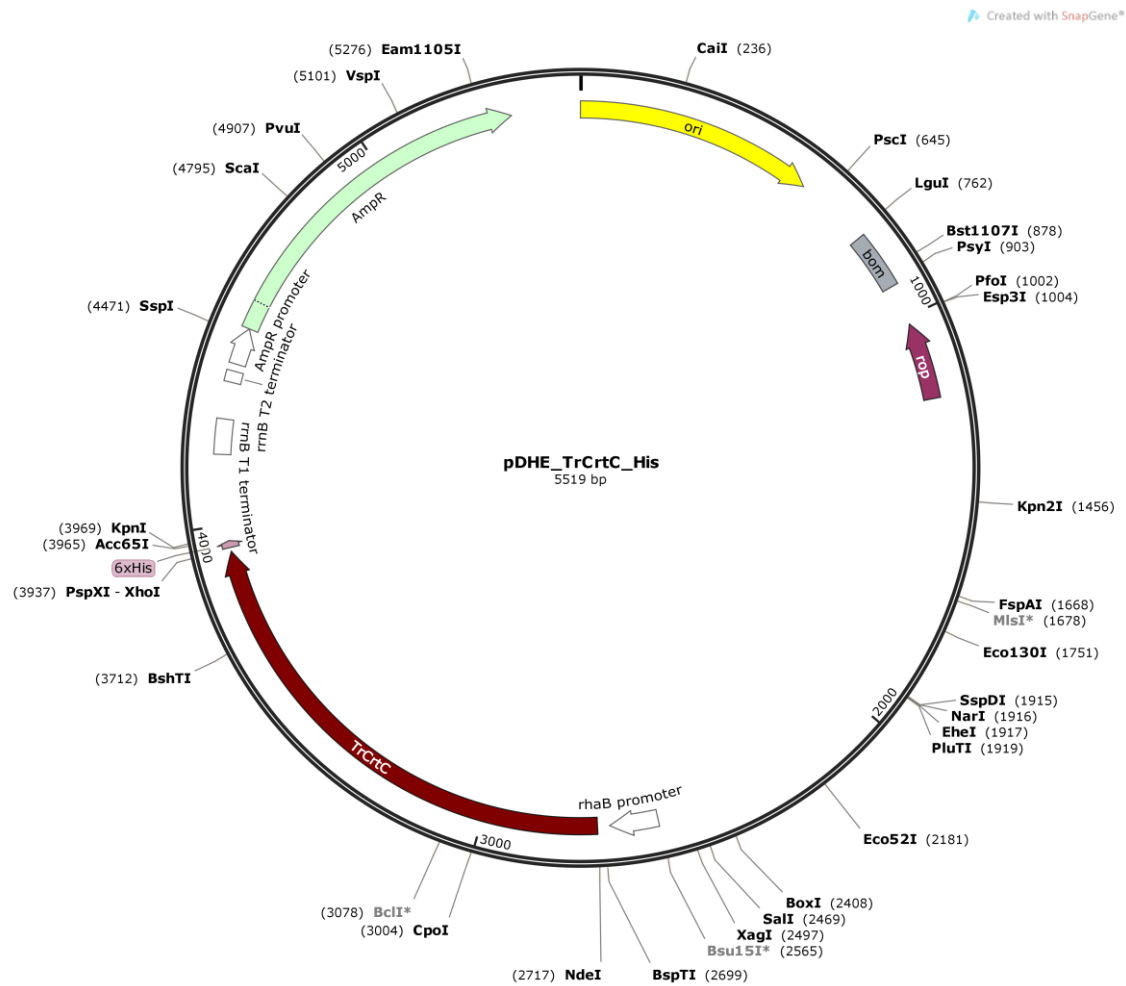
## 7.2 Vector maps



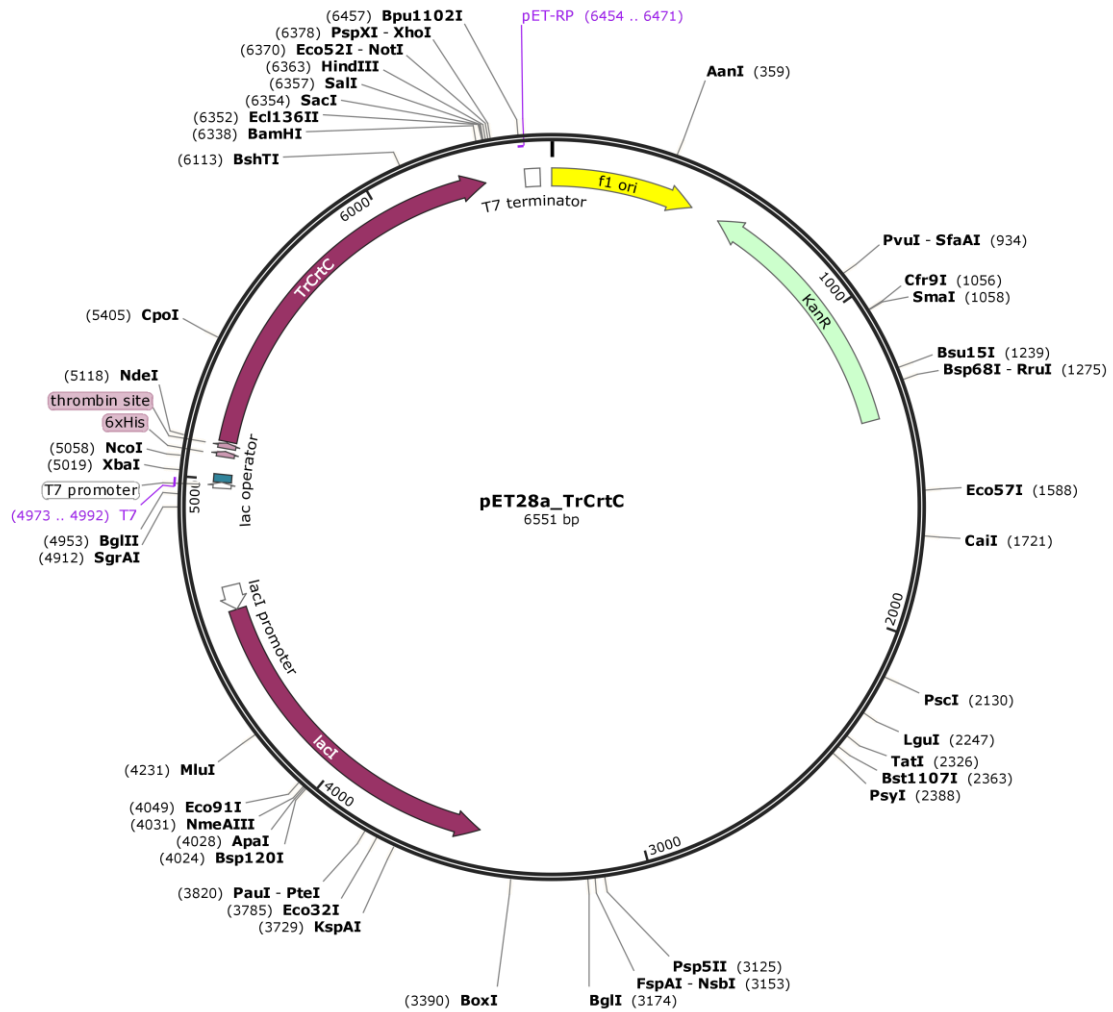
**Figure 58:** Vector map of pDHE-EmOAH wild type construct. The map shows oleate hydratase gene under the control of the rhamnose promoter. A C-terminal His<sub>6</sub>-tag was added in frame to the EmOAH gene (map compiled using SnapGene).



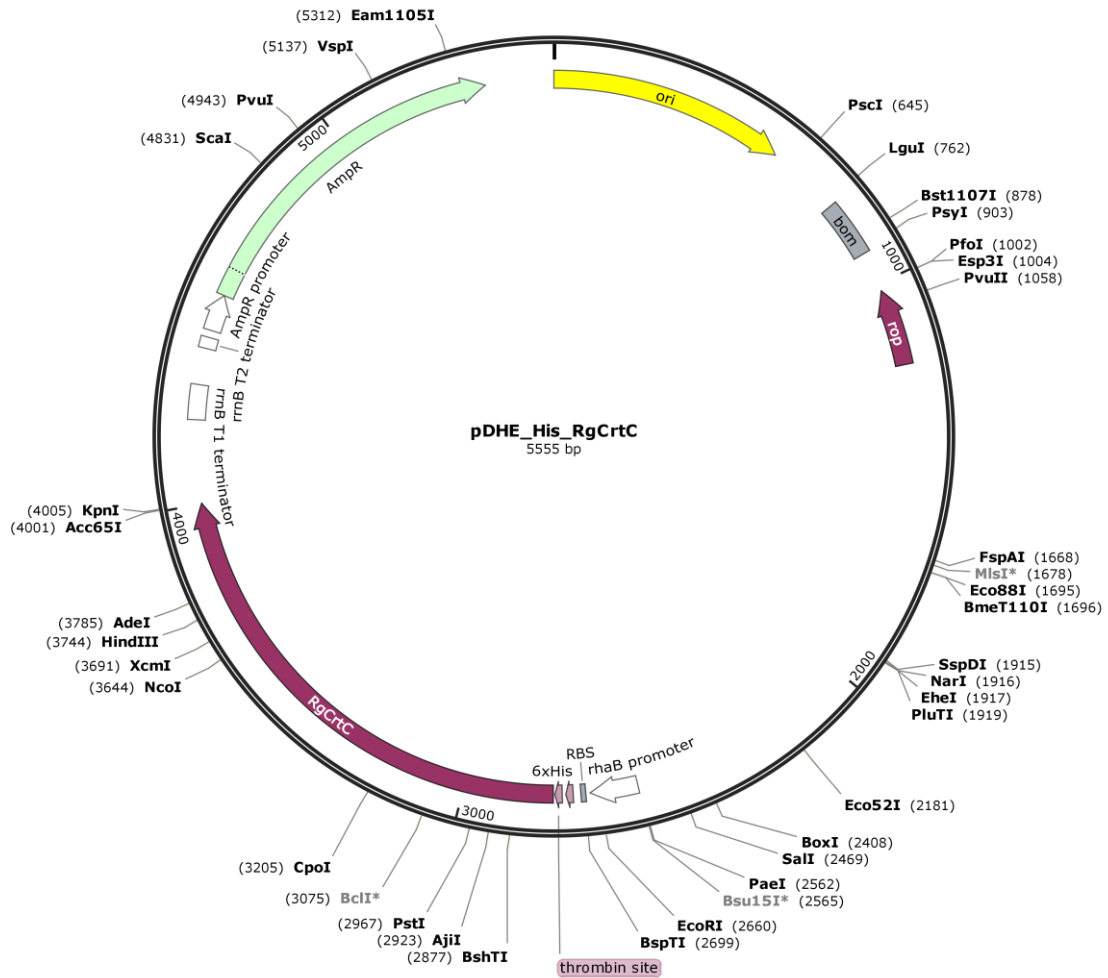
**Figure 59:** Vector map of pET28a(+)-EmOAH wild type construct. The map shows oleate hydratase gene under the control of the T7 promoter. A C-terminal His<sub>6</sub>-tag was added in frame to the EmOAH gene (map compiled using SnapGene).



**Figure 60:** Vector map of pDHE-TrCrtC wild type construct. The map shows carotenoid hydratase gene under the control of the rhamnose promoter. A C-terminal His<sub>6</sub>-tag was added in frame to the TrCrtC gene (in favor of the N-terminal one in pET28) during cloning procedure via Gibson Assembly as the N-terminal proline rich part of the protein is processed.<sup>[96]</sup> The map was compiled using SnapGene.

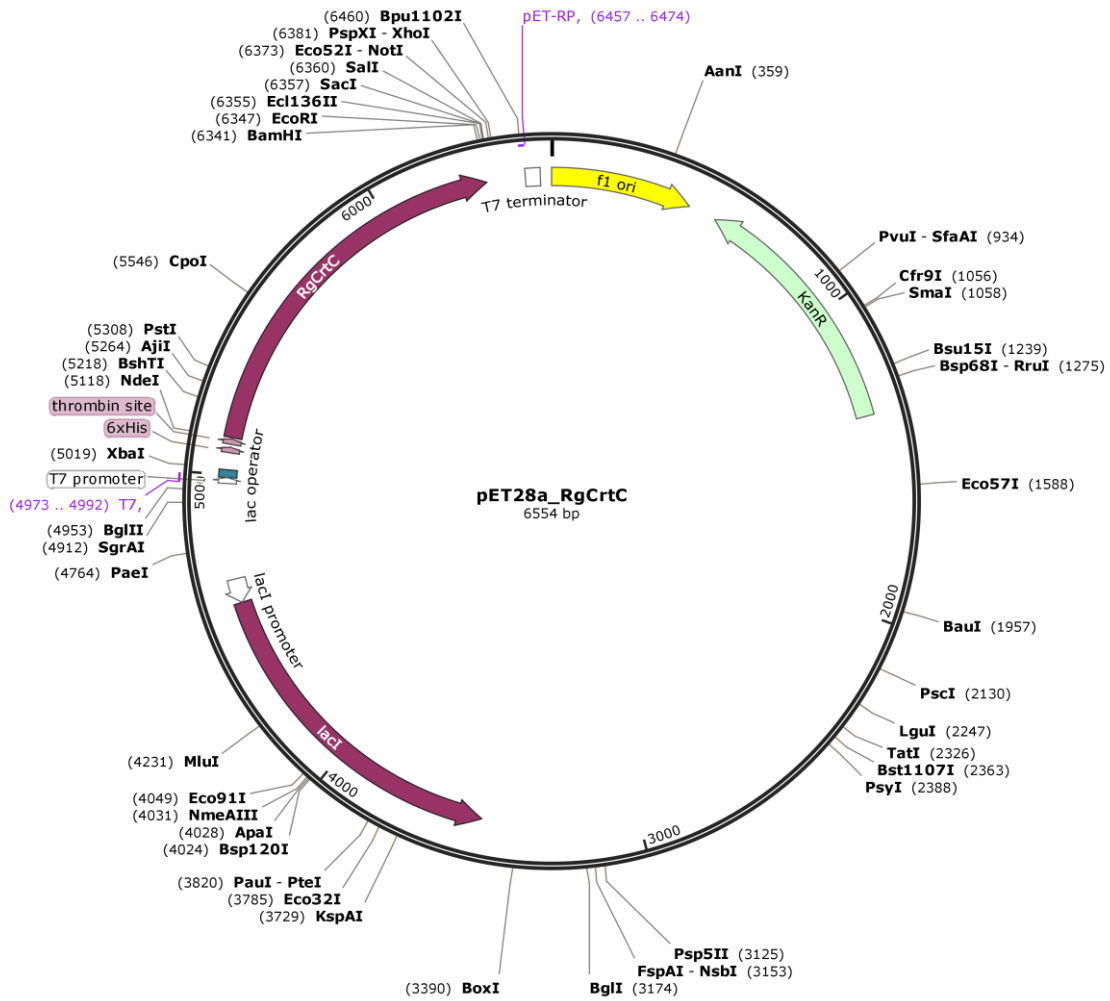


**Figure 61:** Vector map of pET28a(+)-TrCrtC wild type construct. The map shows carotenoid hydratase gene under the control of the T7 promoter. A N-terminal His<sub>6</sub>-tag was added in frame to the TrCrtC gene. The map was compiled using SnapGene.

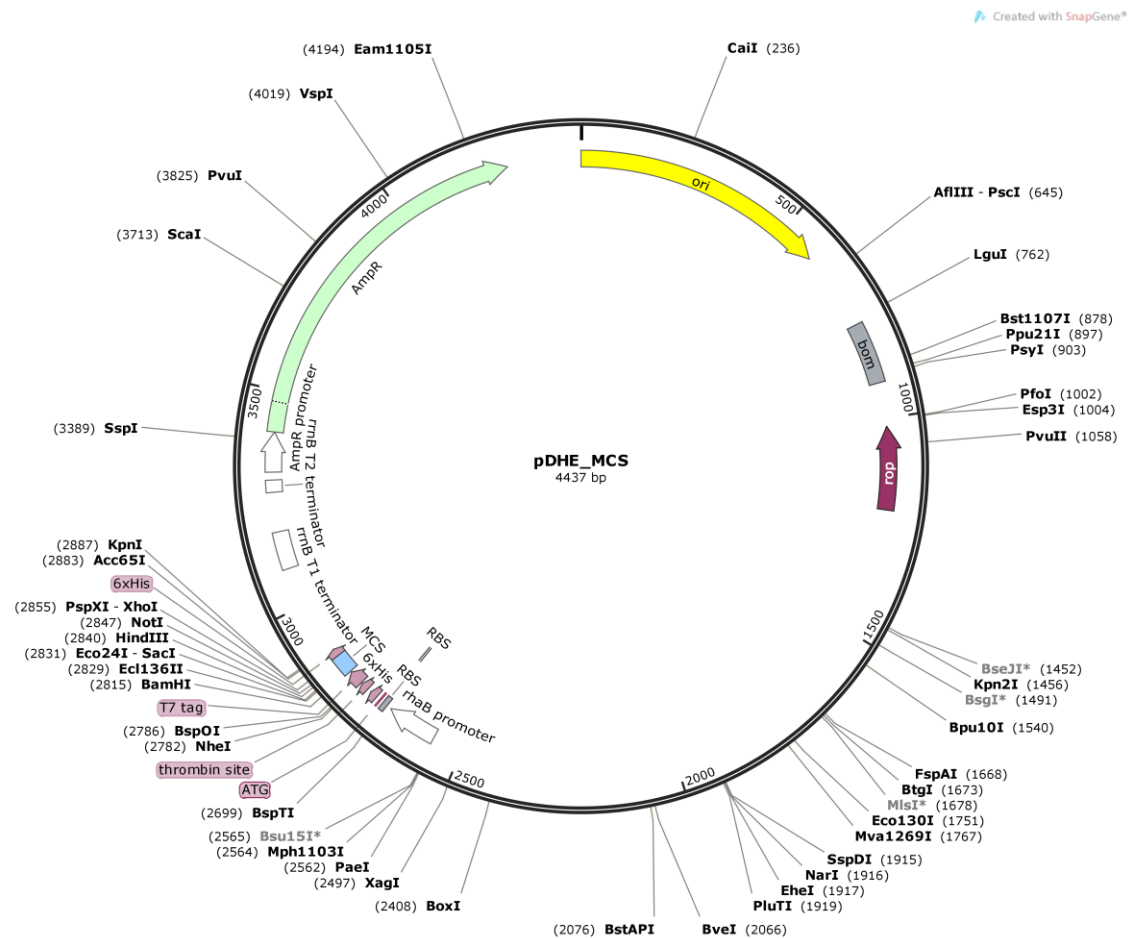


**Figure 62:** Vector map of pDHE-RgCrtC wild type construct. The map shows carotenoid hydratase gene under the control of the rhamnose promoter. An N-terminal His<sub>6</sub>-tag was added in frame to the RgCrtC gene. The map was compiled using SnapGene.





**Figure 63:** Vector map of pET28a(+)-RgCrtC wild type construct. The map shows carotenoid hydratase gene under the control of the T7 promoter. An N-terminal His<sub>6</sub>-tag was added in frame to the RgCrtC gene. The map was compiled using SnapGene.



**Figure 64:** Vector map of pDHE-MCS construct. The map shows the vector which was used as a negative control. For the construction of this empty pDHE vector, the RgCrtC gene was replaced by the multiple cloning site (“MCS”) of pET28a(+) by Gibson Assembly. The map was compiled using SnapGene.

### 7.3 NMR/GC/HPLC/MS spectra

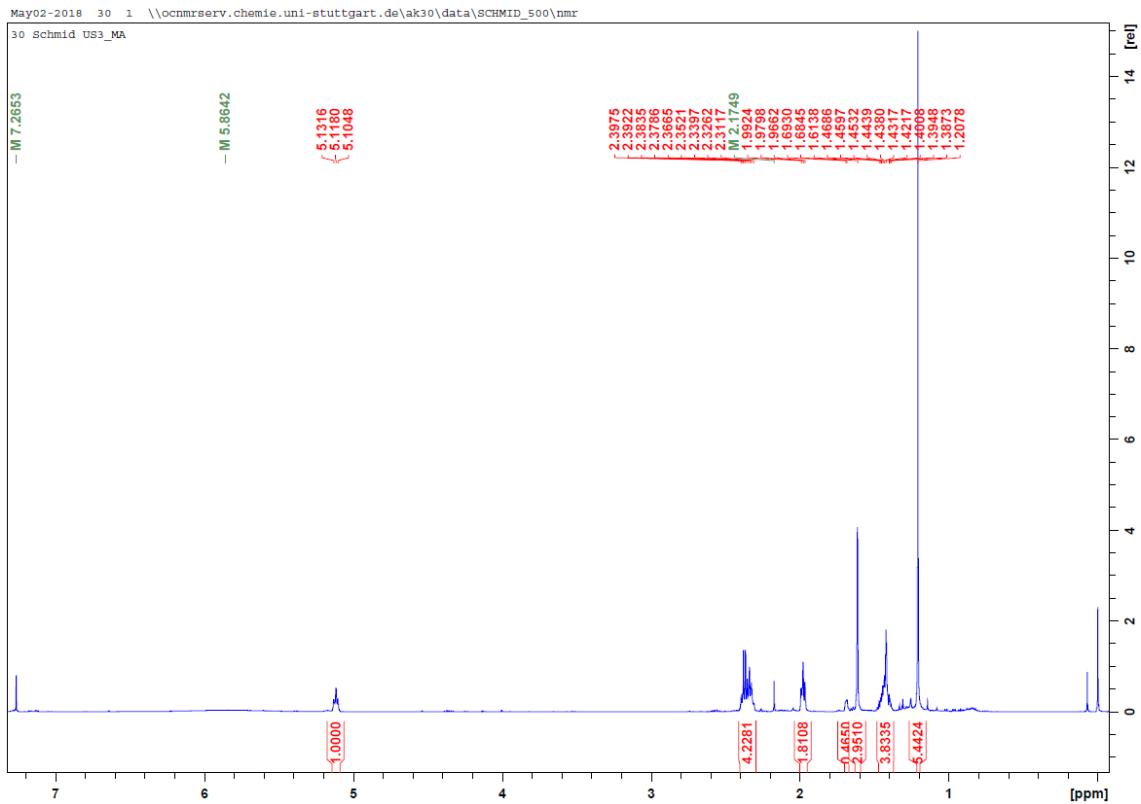


Figure 65:  $^1\text{H}$ -NMR of isolated HDMDAC.

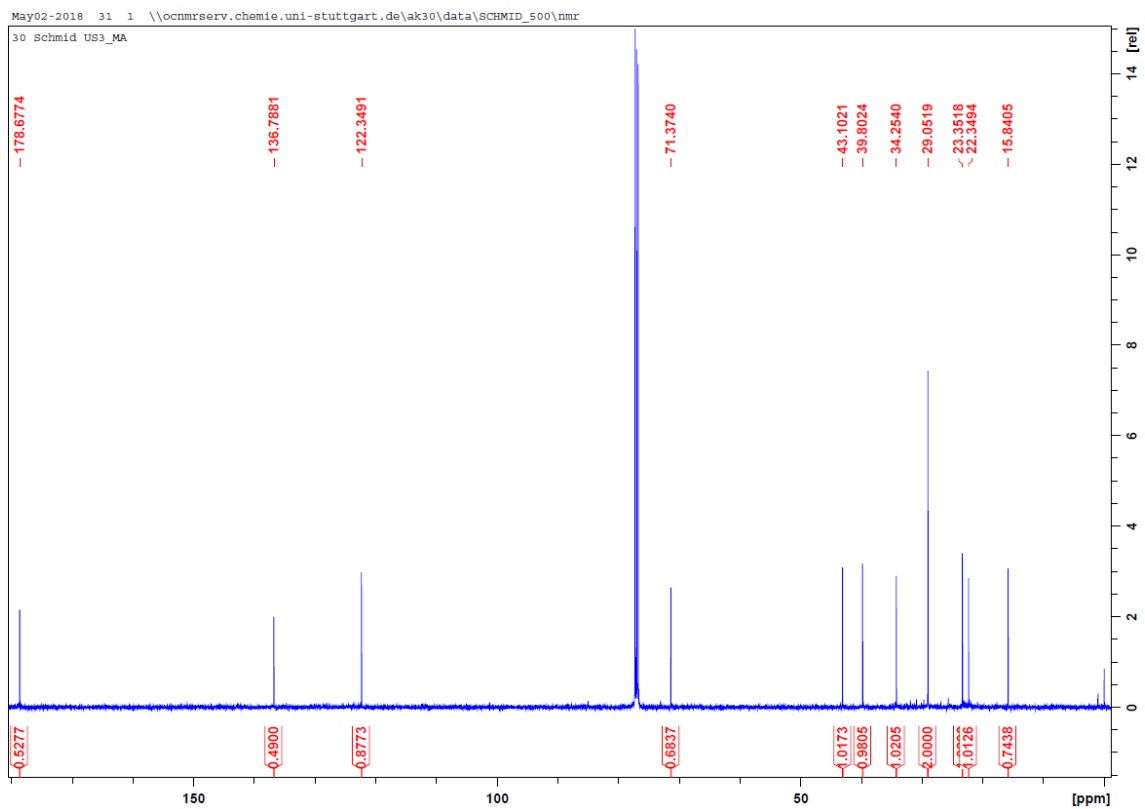
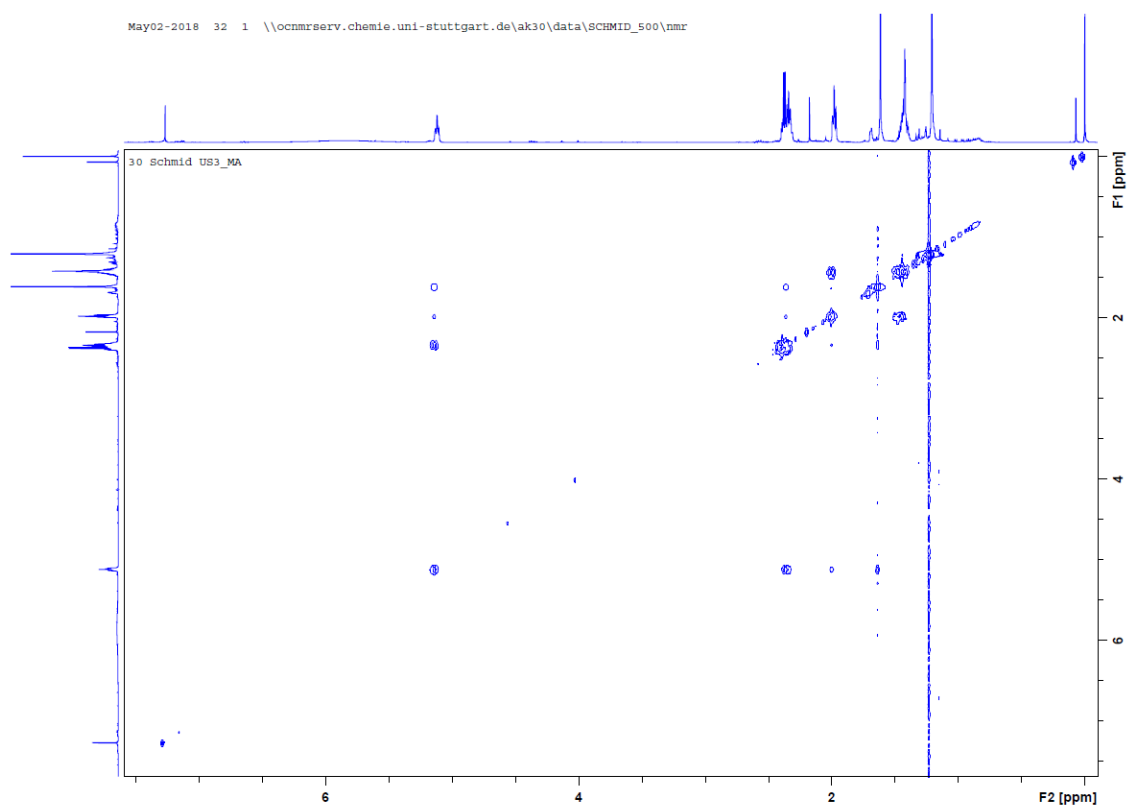
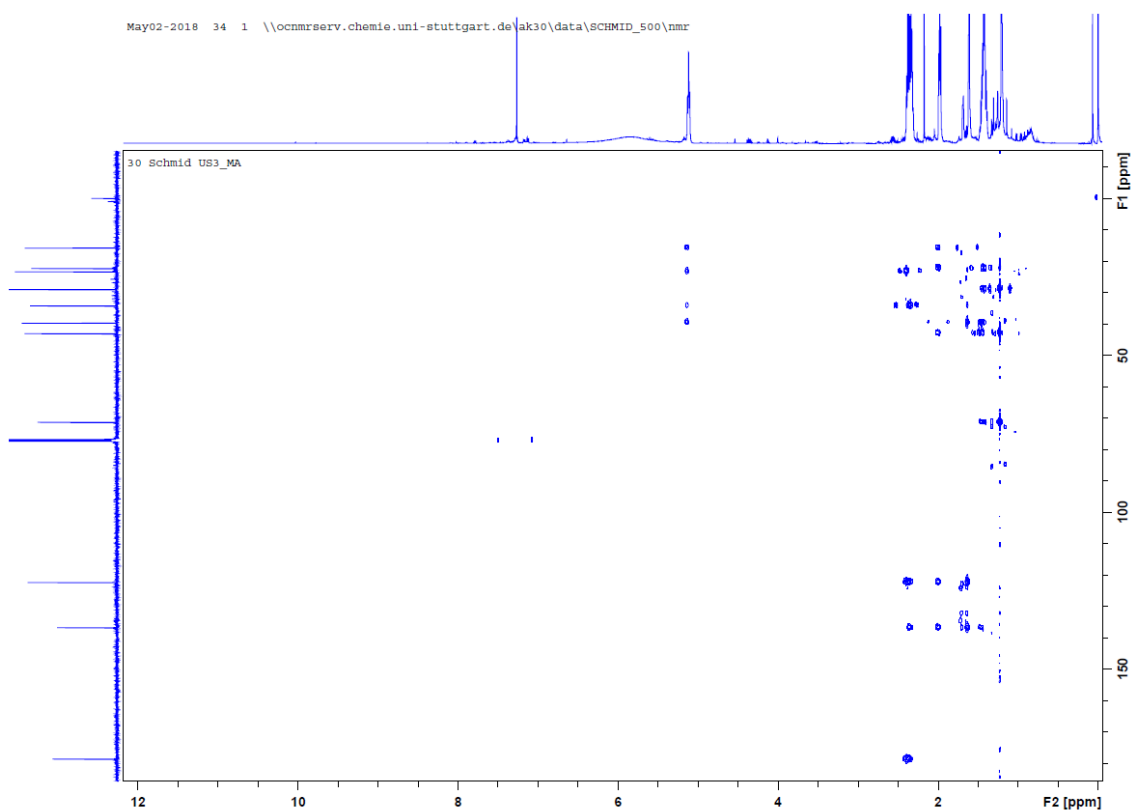


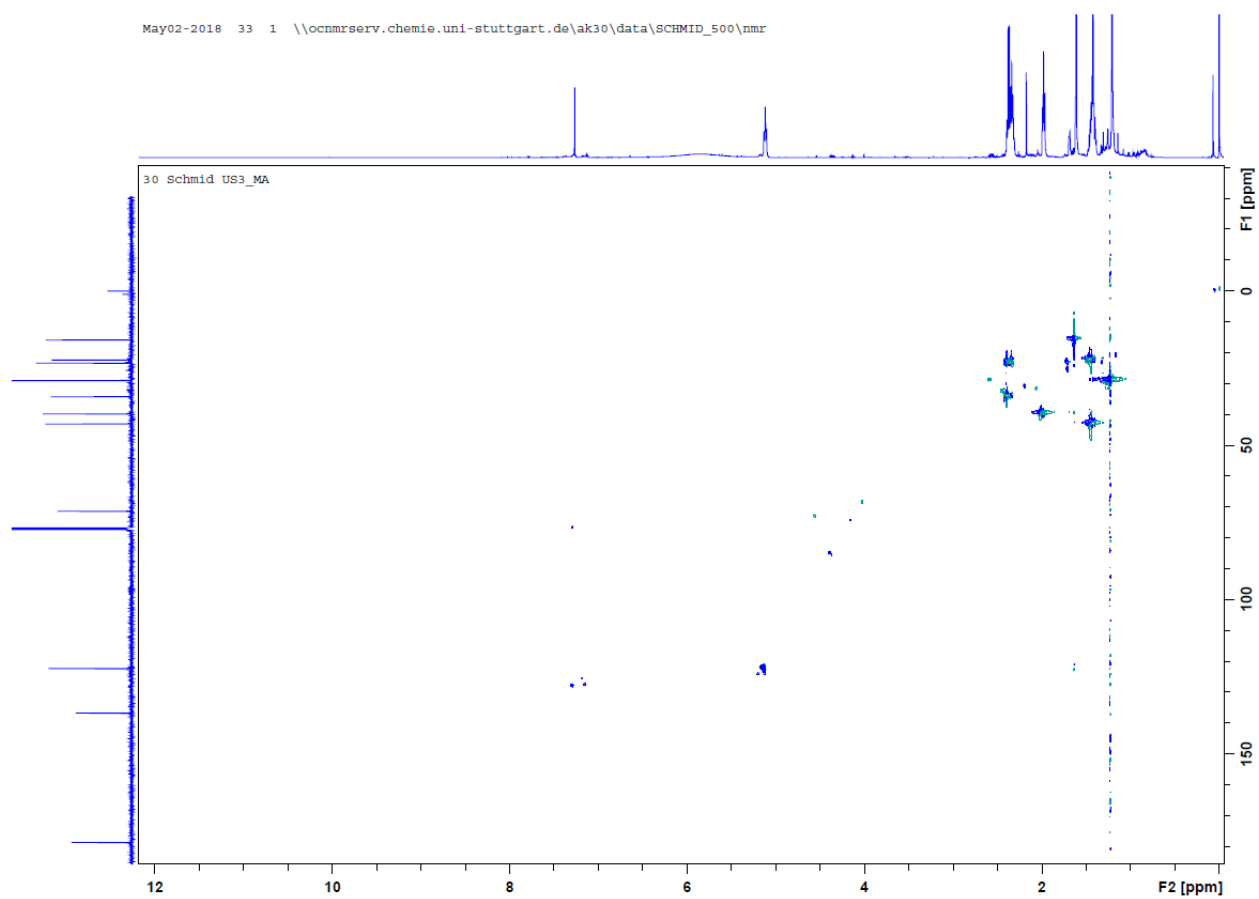
Figure 66:  $^{13}\text{C}$ -NMR of isolated HDMDAC.



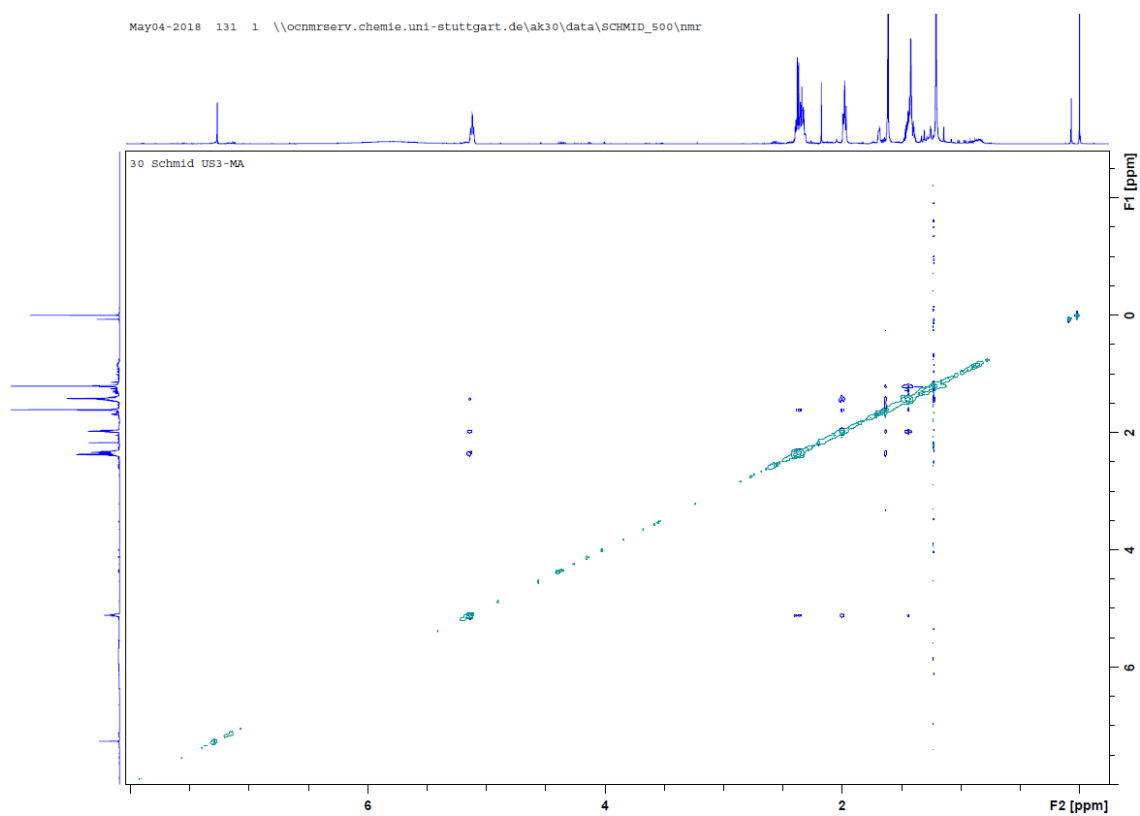
**Figure 67:** H-H-COSY of isolated HDMDAC.



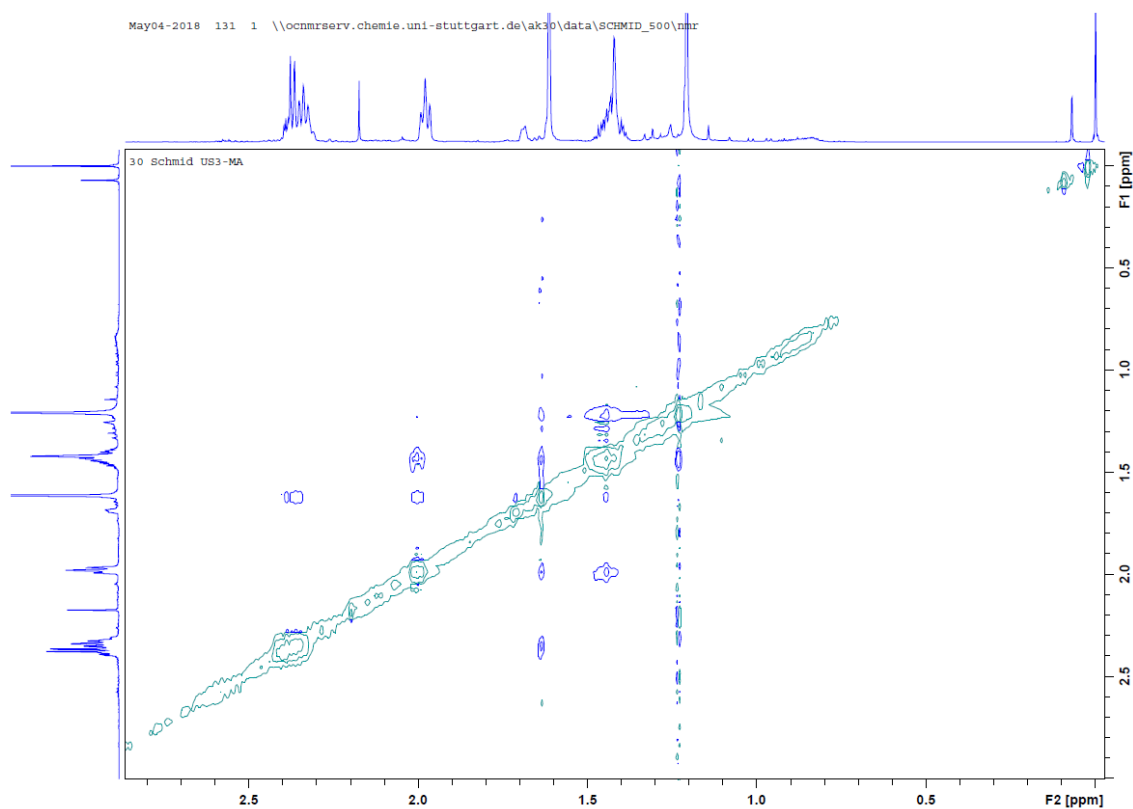
**Figure 68:** HMBC of isolated HDMDAC.



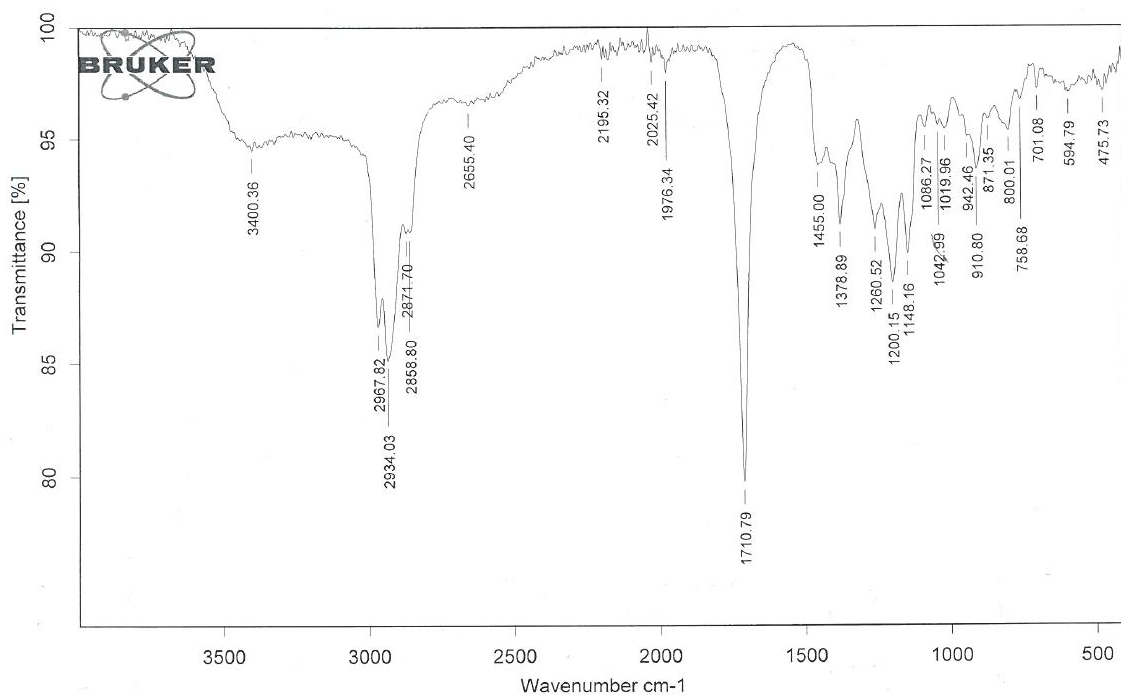
**Figure 69:** HSQC of isolated HDMDAC.



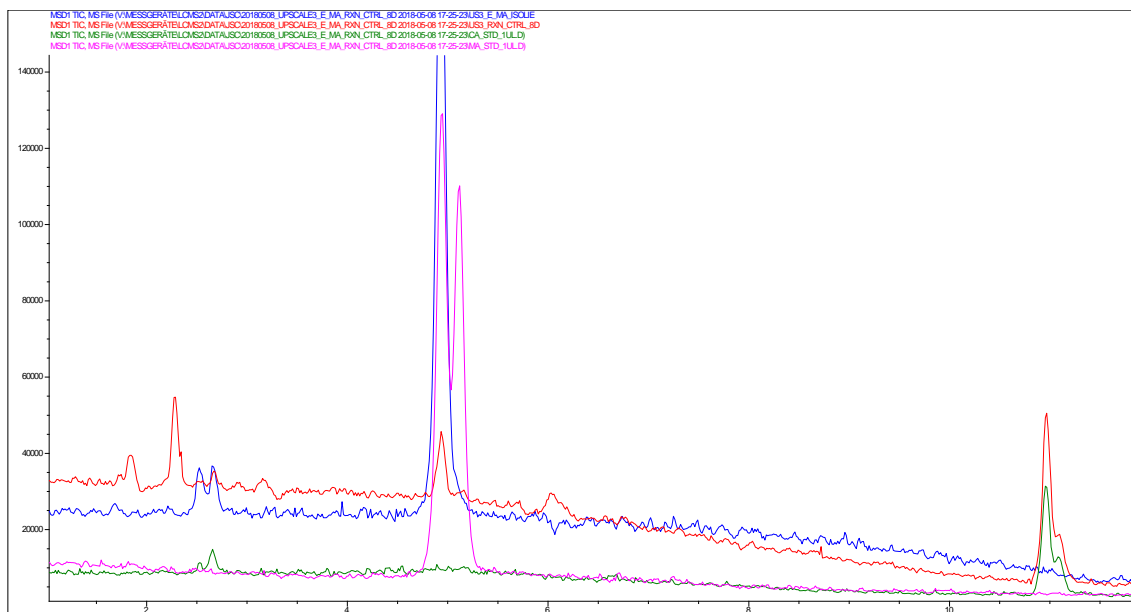
**Figure 70:** NOESY of isolated HDMDAC.



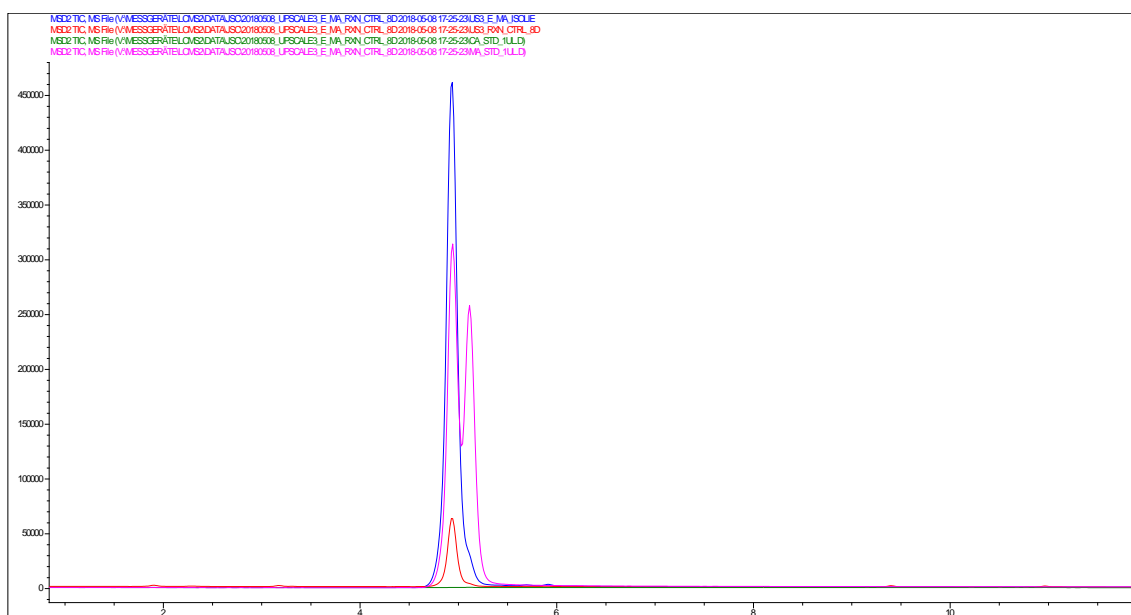
**Figure 71:** Detailed view of NOESY of isolated HDMDAC showing the area of 0-3 ppm.



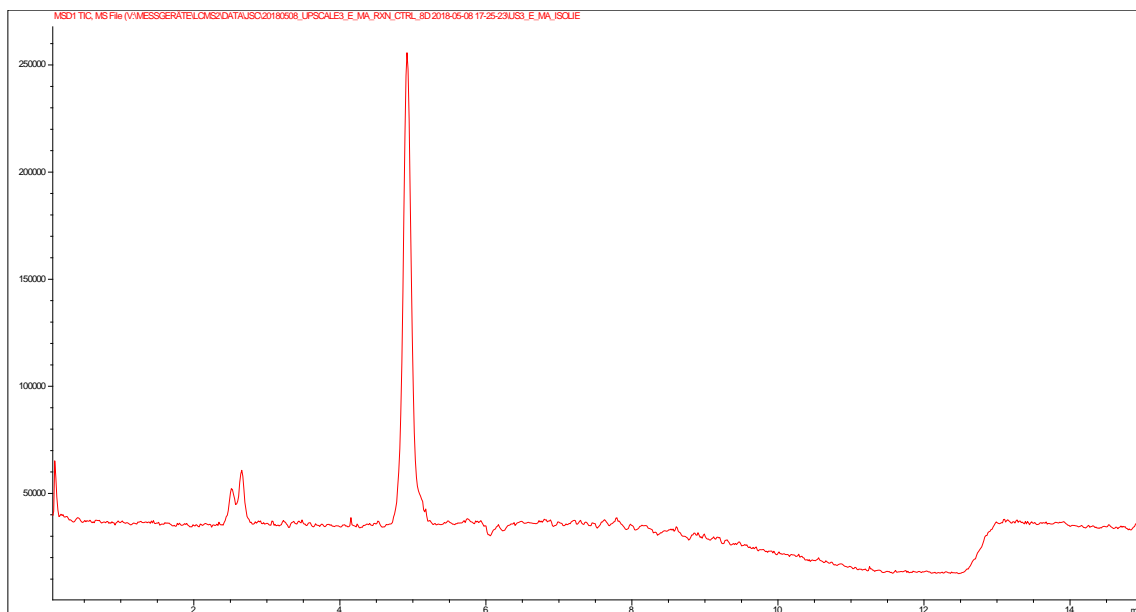
**Figure 72:** IR-spectrum of isolated HDMDAC.



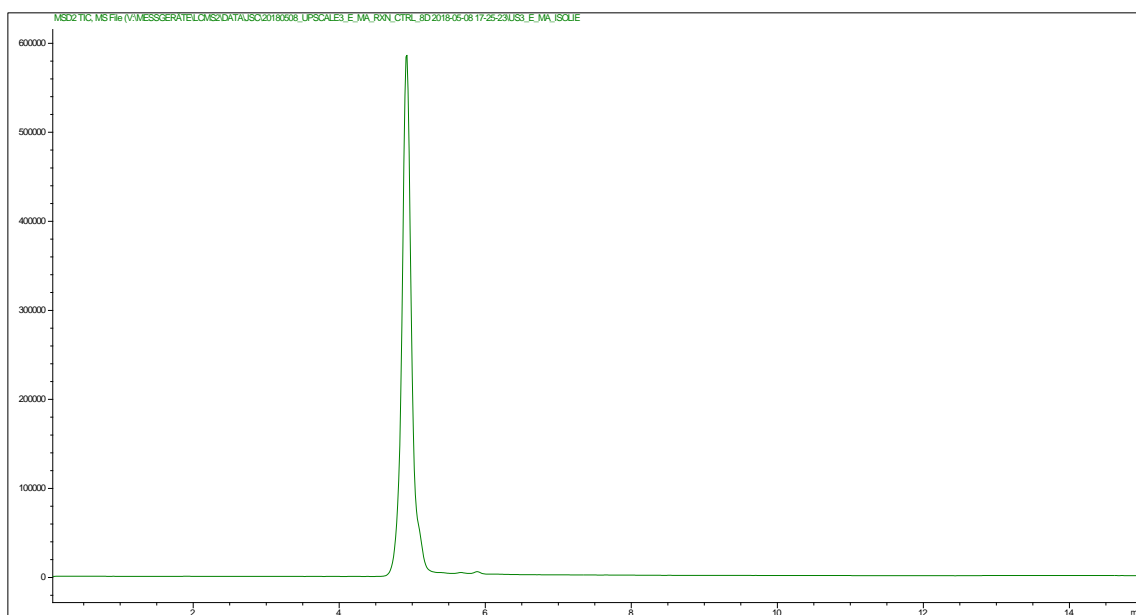
**Figure 73:** LC-MS chromatograms (SCAN, negative mode, 120-250) of isolated HDMDAC (blue), reaction control 7 d 19.25 h (red), DMDDAC reference (green), HDMDAC reference (pink).



**Figure 74:** LC-MS chromatograms (SIM, negative mode, 143+213) of isolated HDMDAC (blue), reaction control 7 d 19.25 h (red), DMDDAC reference (green), HDMDAC reference (pink).

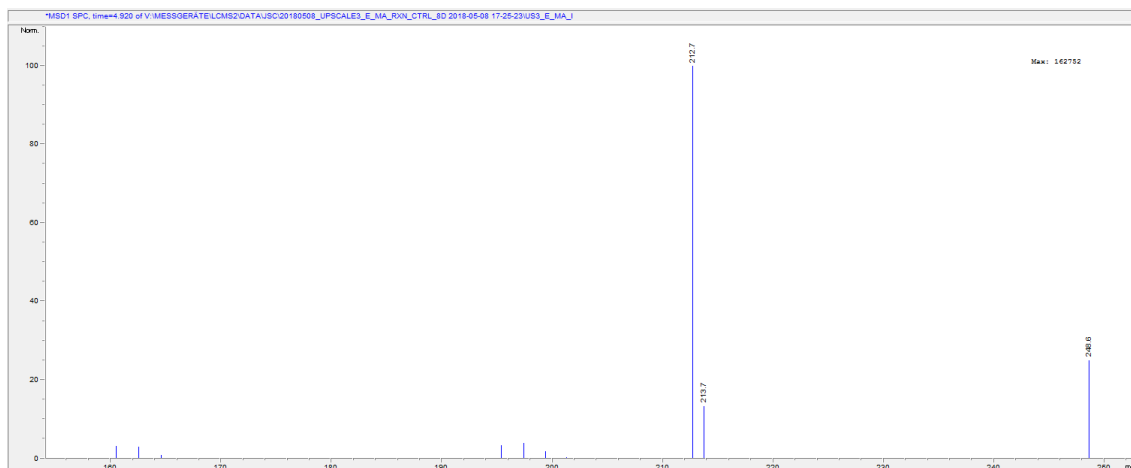


**Figure 75:** LC-MS chromatogram (SCAN, negative mode, 120-250) of isolated HDMDAC.

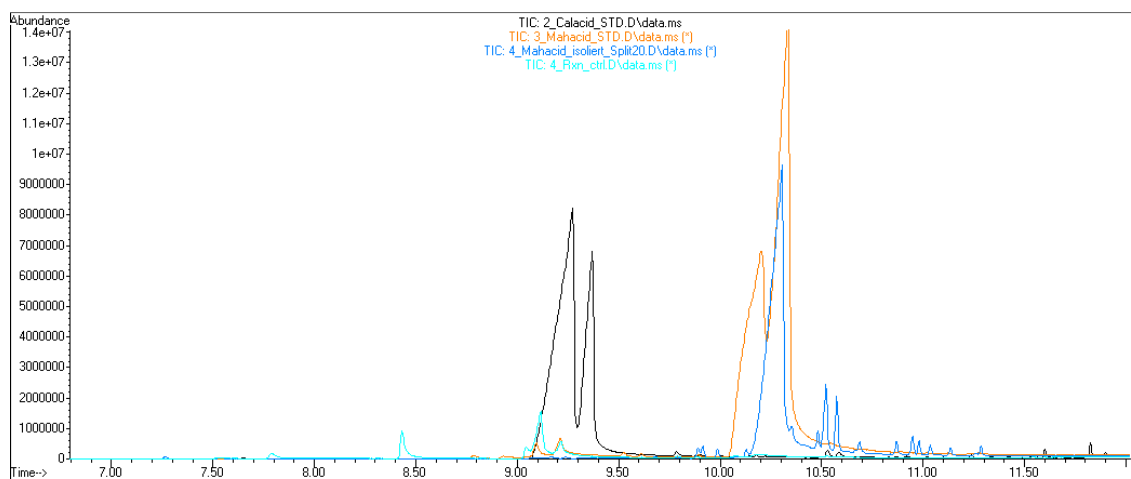


**Figure 76:** LC-MS chromatogram (SIM, negative mode, 143+213) of isolated HDMDAC.

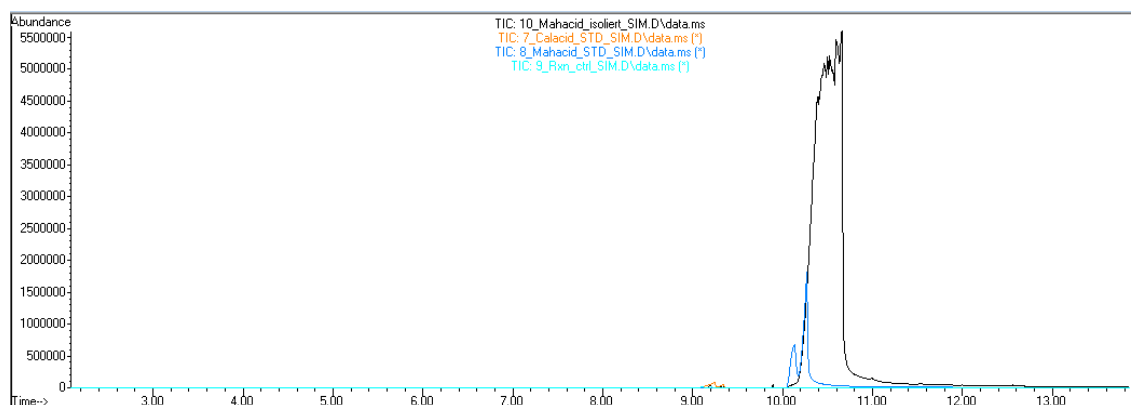




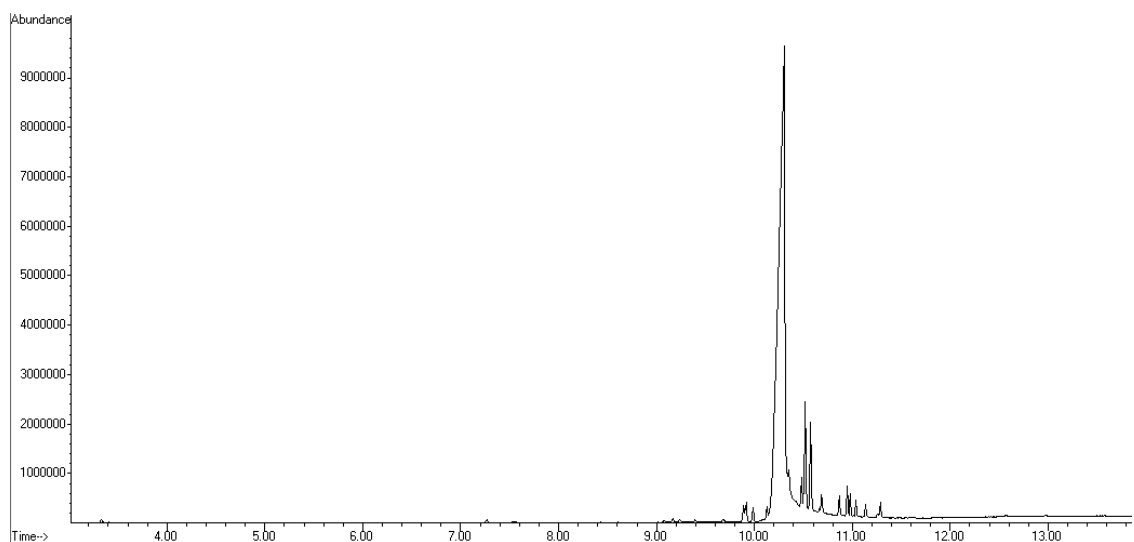
**Figure 77:** LC-MS chromatogram (SCAN, negative mode, 120-250) of isolated HDMDAC. Depicted is the mass spectrum of the product peak at 4.920 min. Peak at 212.7 corresponds to the  $[M-H]$  peak, 248.6 is the  $[M+Cl]$  peak typical for ESI negative ionization mode. Smaller peaks in the area of 195 correspond to the dehydration product  $[M-H_2O-H]$  under conditions of analysis.



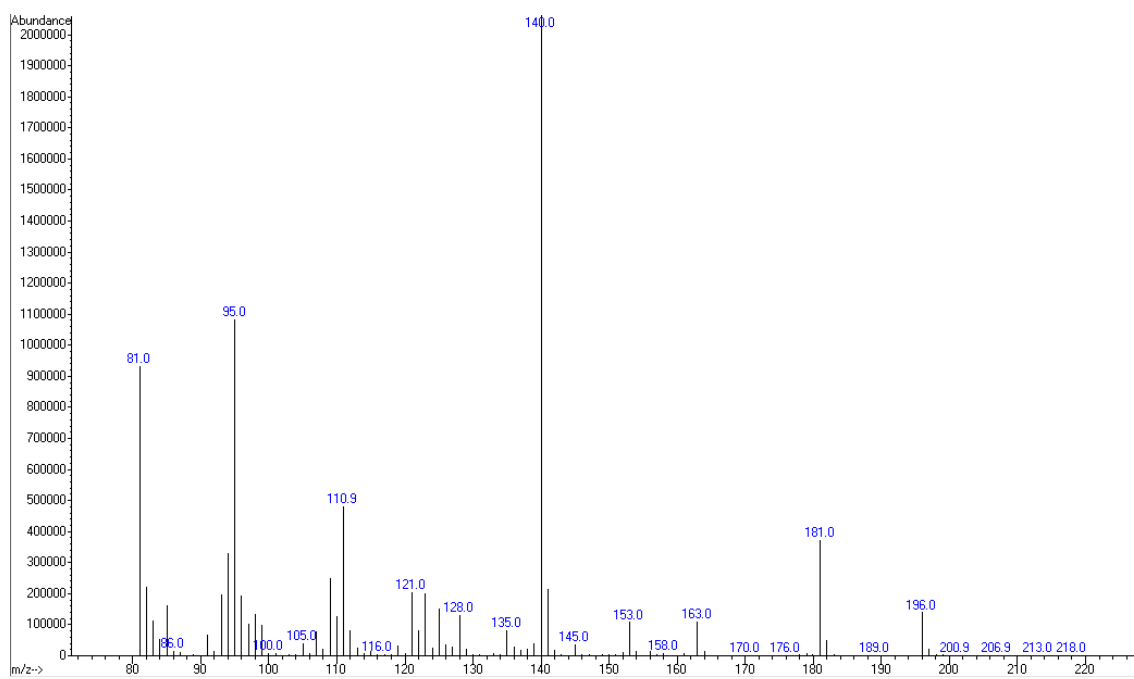
**Figure 78:** GC-MS chromatogram (SCAN, 80-220) of isolated HDMDAC (blue), reaction control 7 d 19.25 h (cyan), DMDDAC reference (black), HDMDAC reference (yellow).



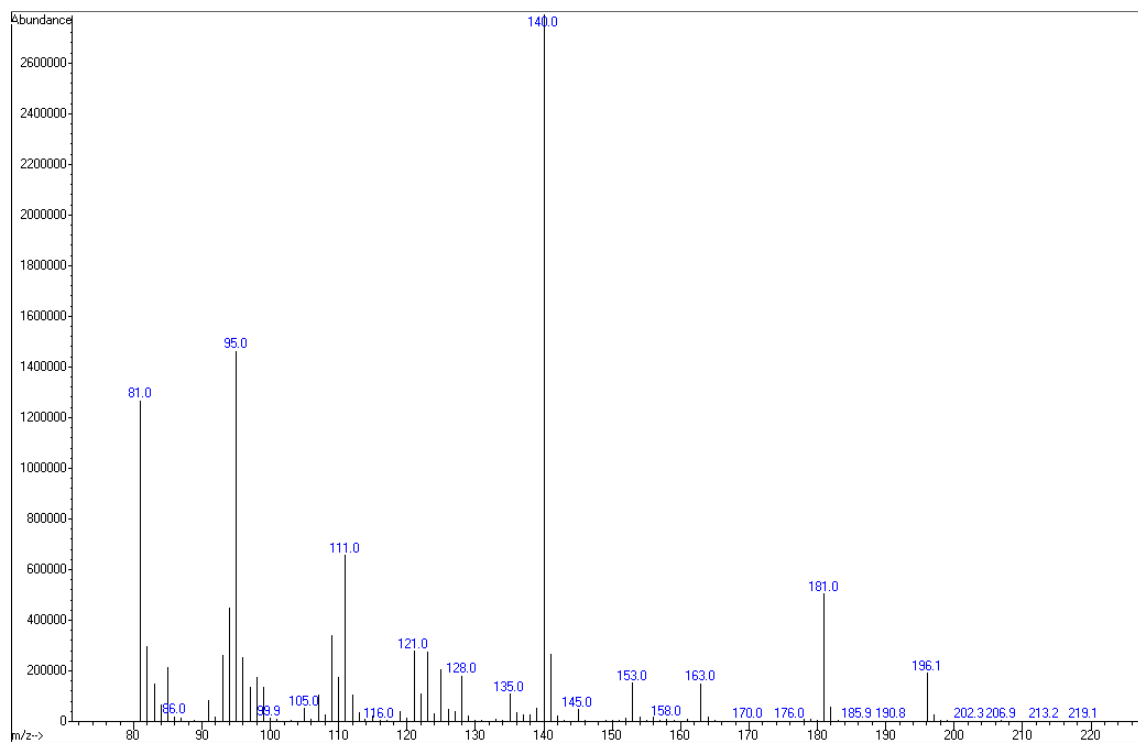
**Figure 79:** GC-MS chromatogram (SIM, 140 = main fragment) of isolated HDMDAC (black), reaction control 7 d 19.25 h (cyan), DMDDAC reference (yellow), HDMDAC reference (blue).



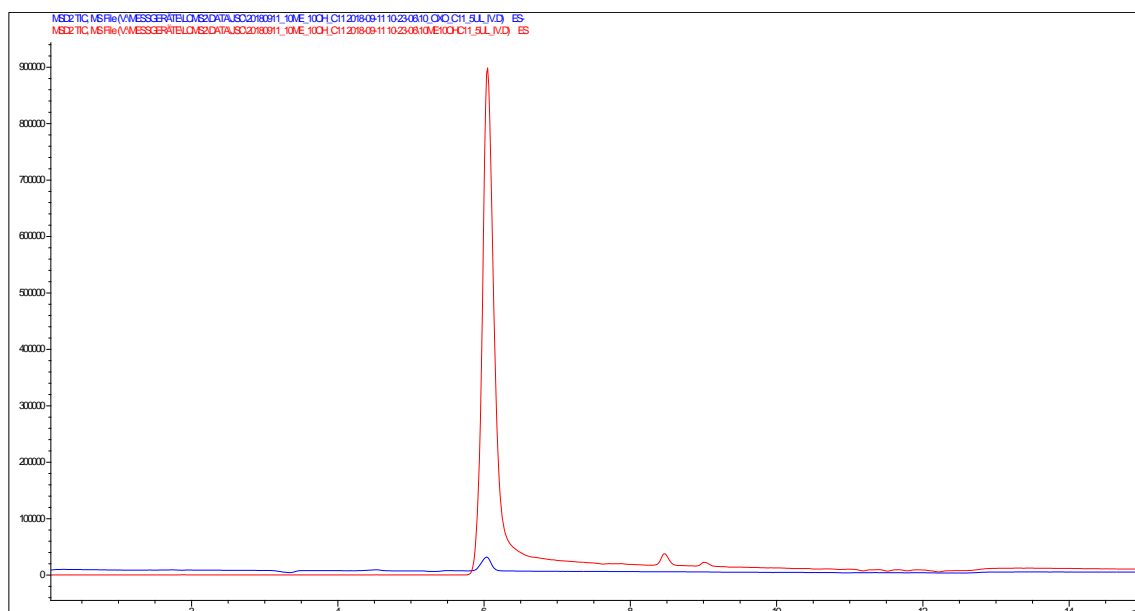
**Figure 80:** GC-MS chromatogram (SCAN, 80-220) of isolated HDMDAC.



**Figure 81:** Mass spectrum of GC-MS (SCAN, 80-220) at 10.300 min of isolated HDMDAC (*E*-isomer). The pattern is in accordance with the standard (Figure 82) and the MS data provided by Givaudan, main fragments are at 196 [ $M^+ - H_2O$ ], 181, 140, 95 and 81.



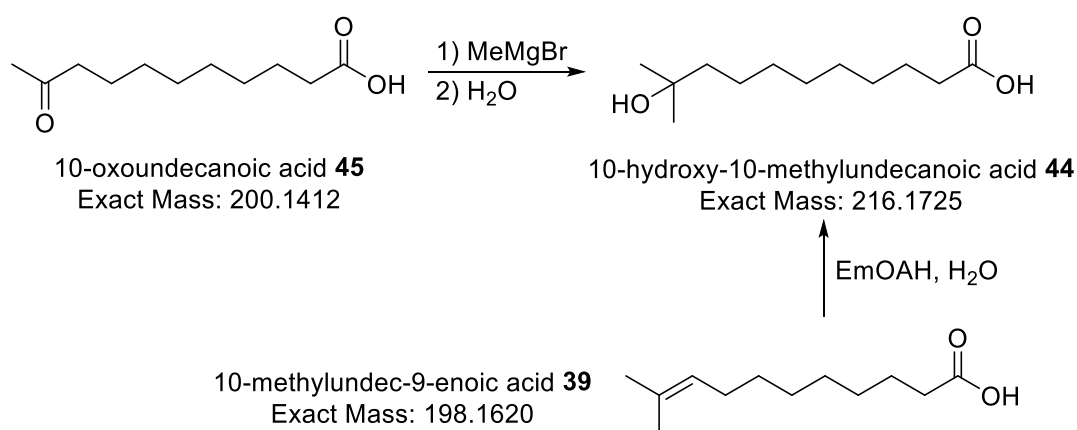
**Figure 82:** Mass spectrum of GC-MS (SCAN, 80-220) at 10.314 min (corresponds to the peak of the *E*-isomer) of HDMDAC for comparison with the spectrum of isolated HDMDAC. The spectra show the same fragment pattern and even ratios are apparently the same. (No difference was observed between the *E*- and the *Z*-isomer).



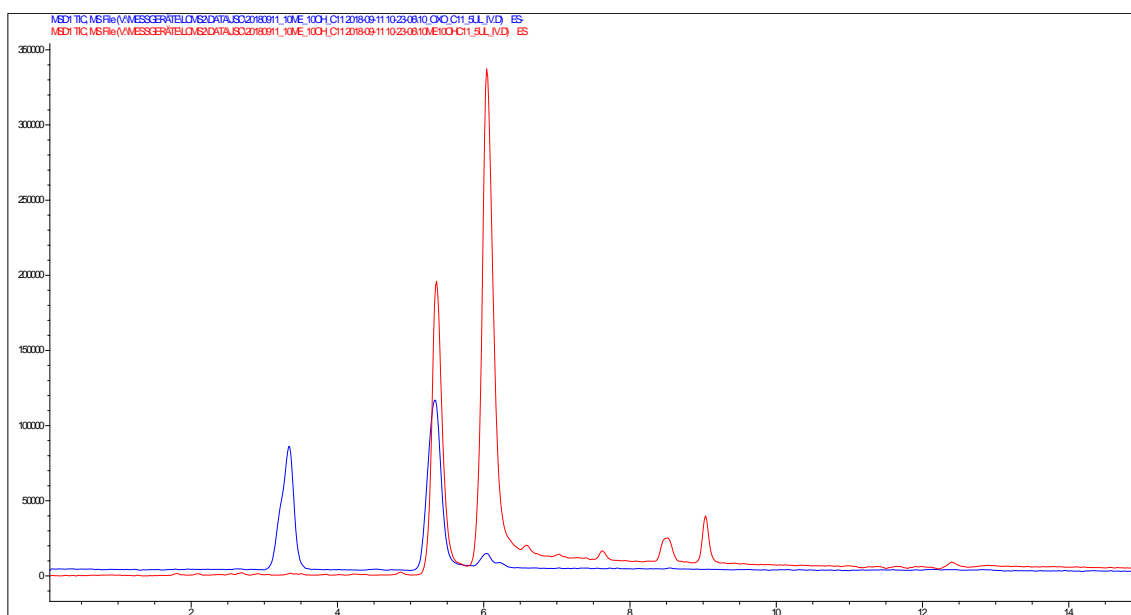
**Figure 83:** LC-MS chromatograms (SIM, negative mode, 215) of purchased 10-oxo-undecanoic acid (blue) and synthesized 10-hydroxy-10-methyl-undecanoic acid (red).

## 7.4 Substrate scope studies: Chemical synthesis of 10-hydroxy-10-methylundecanoic acid

10-Methylundec-9-enoic acid **39** was investigated for the influence of a methyl group adjacent to the double bond in the hydration of sterically more demanding trisubstituted alkenes (4.4.12). It has an additional methyl group compared to (*Z*)-undec-9-enoic acid (4.4.6.2). While the substrate **39** was provided by Givaudan, the hydration product 10-hydroxy-10-methylundecanoic acid **44** was synthesized (3.8.1) from 10-oxoundecanoic acid **45** by Grignard reaction with MeMgBr to obtain an authentic standard for the enzymatic hydration of 10-methylundec-9-enoic acid **39** with EmOAH (Figure 84).

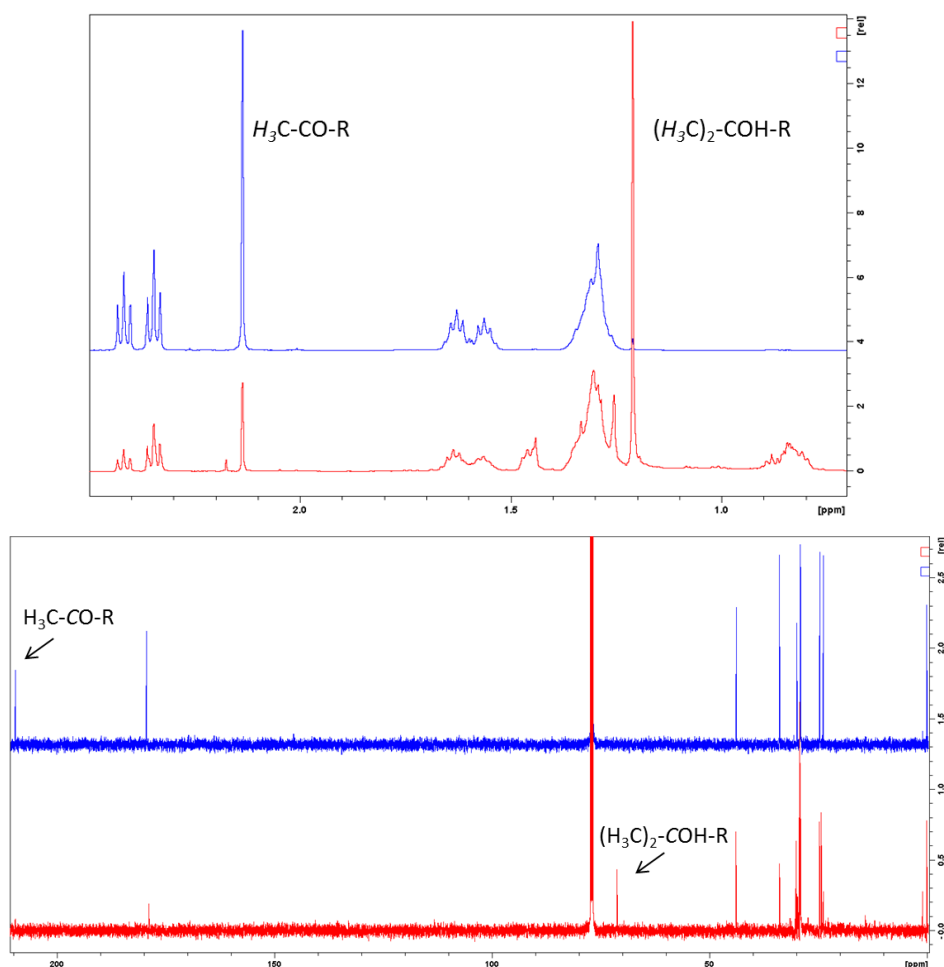


**Figure 84:** 10-Hydroxy-10-methylundecanoic acid **44** both as a product of Grignard reaction of 10-oxoundecanoic acid **45** with MeMgBr and H<sub>2</sub>O and as hydration product of 10-methylundec-9-enoic acid **39** in whole cell biotransformation with EmOAH.



**Figure 85:** LC-MS chromatograms (SCAN, negative mode, 120-250) of purchased 10-oxo-undecanoic acid **45** (blue) and synthesized 10-hydroxy-10-methylundecanoic acid **44** (red).

LC-MS analysis of the purchased 10-oxoundecanoic acid **45** (Figure 85, blue) showed only a poor purity (55 %) with even little impurities of the same retention time ( $t = 6.0$  min) as the product **44** (Figure 85, red). The peak of the main impurity ( $t = 3.3$  min) showed a mass signal of 201.2 while the peak at 5.3 min had a mass signal for  $[M-H] = 199.2$  which is in accordance with the desired 10-oxoundecanoic acid **45**. During Grignard reaction and subsequent work-up, the impurity vanishes and the educt peak decreases. The increasing peak was identified as 10-hydroxy-10-methylundecanoic acid **44** by its mass ( $[M-H] = 215.2$ ). Additionally, substrate **45** and product **44** were analyzed by  $^1\text{H}$ -,  $^{13}\text{C}$ - and  $^1\text{H}$ - $^1\text{H}$ -COSY-NMR. While the complete assignment of signals was restricted by the poor purity of both substances, the significant changes of the carbonyl group to the alcohol and the additional methyl group were clearly visible in  $^1\text{H}$ - and  $^{13}\text{C}$ -NMR (Figure 86). 10-Hydroxy-10-methylundecanoic acid **44** had a LC-MS purity of about ca. 61 % and was used without further purification as product standard.



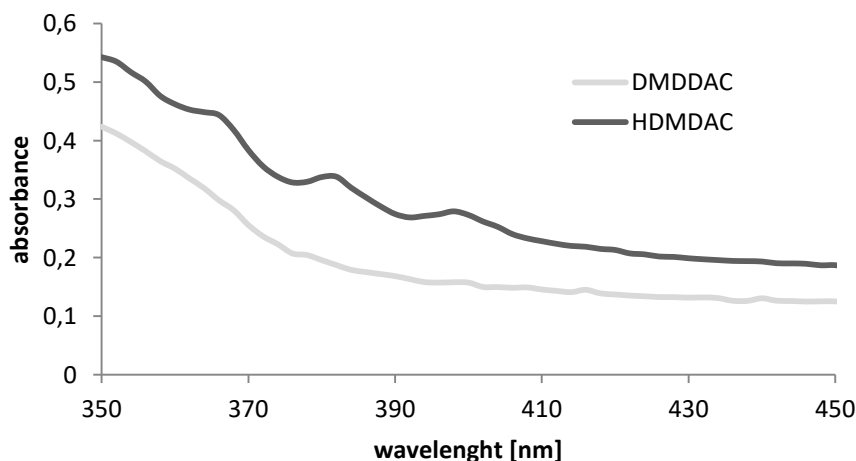
**Figure 86:** Comparison of educt (blue) and product (red) spectrum. a)  $^1\text{H}$ -NMR b)  $^{13}\text{C}$ -NMR

## 7.5 Methods for the screening of OAH mutant libraries

### 7.5.1 Assessment of a photometric HT-assay

The larger mutant libraries are the more necessary is a reliable and fast screening method. Screening of large libraries is often based on photometric assays, like tracing the NAD(P)H depletion or based on colorimetric reactions, sometimes coupled to another enzyme. However, the formation of a tertiary alcohol by cofactor-independent hydratases is not easy to monitor, as further oxidation e.g. with an enzyme is not possible and therefore no direct or indirect cofactor change is detectable. Furthermore, isolated double bonds are almost impossible to detect in UV-VIS.

An alternative high-throughput assay for the spectrophotometric detection of alcohols was developed by Hiseni *et al.*<sup>[197]</sup> OAHs were used as a model enzyme. The assay is based on the selective detection of alkyl nitrites formed by the reaction of the OH-group and nitrous acid. The assay even is able to distinguish between primary, secondary and tertiary alcohols, as tertiary alcohols have a maximum at 400 nm, which is absent for primary and secondary alcohols. However, Hiseni and coworkers reported lower detection limits in the range of 1.5 mM (for most compounds) to 3 mM (for 2-methyl-1-butanol and 2-methyl-2-butanol) with good  $Z'$ -factors. Lower  $Z'$ -factors were observed for enzymatic reaction using OAH depending on the enzyme preparation. Therefore, the authors concluded that this assay is not applicable to enzymatic reactions where low activities are to be discovered, especially for whole cells. For the hydration of oleic acid using whole cells, at least 40 mM of the hydration product 10-hydroxystearic acid must be produced.



**Figure 87:** Nitrite assay of 20 mM DMDDAC and its hydration product 20 mM HDMDAC (according to Hiseni *et al.*<sup>[197]</sup>).

The assay was tested for DMDDAC and its hydration product HDMDAC to determine the concentration needed for these specific substrates. A first test using 5 mM samples of substrate and product in buffer was not successful as the concentrations of the analytes were, as expected, too low. Higher concentrations (20 mM) revealed the typical pattern for tertiary alcohols (Figure 87) as described by Hiseni *et al.* However, the concentrations were out of the suitable range for the conducted biotransformations as the achieved product concentrations were in the micro molar range. Therefore, GC or HPLC based assays were considered to monitor the hydration reaction.

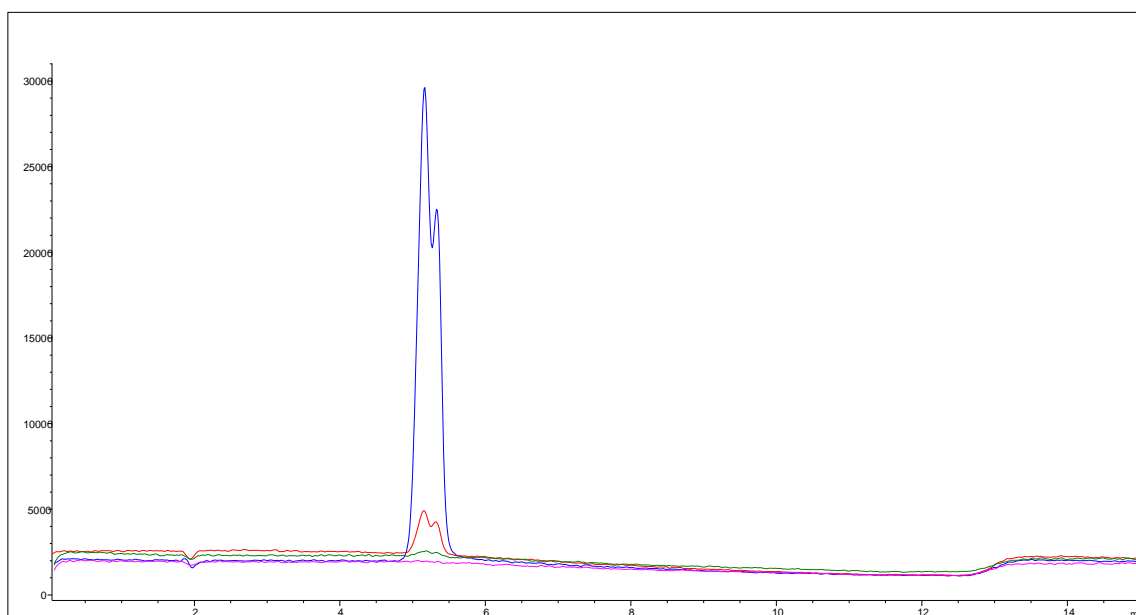
### 7.5.2 Chromatographic methods

Initially, GC methods were tested for the screening of small OAH mutant libraries as all authentic standards of the substrates and their hydration products could be detected by GC analysis. However, the proof of the desired tertiary alcohol products in the biotransformation mixtures was challenging. Two problems of GC analysis were identified:

- 1) Conditions of analysis lead to background hydration of the double bond or vice versa to dehydration of the tertiary alcohols (potentially caused in the liner during injection). This made quantification or a statement about the quality of the biocatalyst impossible as always traces of substrate and product were found in variable, non reproduceable amounts.
- 2) Limit of detection for HDMDAC was too low for screening of OAH-biotransformations of DMDDAC

GC analysis is in general possible when high product concentrations can be provided. With decreasing concentration of the tertiary alcohols, especially HDMDAC, the peaks “dilute” and get broad until they vanish completely. The limit of detection for HMDDAC was measured by Givaudan and determined to be in the range of 50-100 ppm ( $100 \mu\text{g mL}^{-1} = 467 \mu\text{M}$ ).

Finally, the switch to LC-MS analysis led to reliable results. HDMDAC formation could be monitored in biotransformation in adequately small amounts. Without the disturbing matrix of whole cell biotransformations, the authentic standard of *E/Z*-HDMDAC was still detectable in the range of 100 nM and quantification was possible for concentrations in the 1  $\mu\text{M}$  range (Figure 88).



**Figure 88:** LC-MS chromatogram (SIM, negative mode, 143+213) of *E/Z*-HDMDAC (58 % *E* isomer). Standard solutions of different concentrations have been analyzed using the same method parameters as for analysis of biotransformations. Blue: 10 μM, red: 1 μM, green: 100 nM, pink: 10 nM.



## 8 Literature

- [1] B. M. Nestl, S. C. Hammer, B. A. Nebel, B. Hauer, *Angew. Chemie Int. Ed.* **2014**, *53*, 3070–3095.
- [2] G. Hughes, J. C. Lewis, *Chem. Rev.* **2018**, *118*, 1–3.
- [3] C. Bernal, K. Rodríguez, R. Martínez, *Biotechnol. Adv.* **2018**, *36*, 1470–1480.
- [4] U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185–94.
- [5] A. J. J. Straathof, S. Panke, A. Schmid, *Curr. Opin. Biotechnol.* **2002**, *13*, 548–556.
- [6] N. J. Turner, *Nat. Chem. Biol.* **2009**, *5*, 567–573.
- [7] D. Hülsewede, L.-E. Meyer, J. von Langermann, *Chem. - A Eur. J.* **2019**, DOI 10.1002/chem.201804970.
- [8] M. D. Truppo, *ACS Med. Chem. Lett.* **2017**, *8*, 476–480.
- [9] R. A. Sheldon, P. C. Pereira, *Chem. Soc. Rev.* **2017**, *46*, 2678–2691.
- [10] U. T. Bornscheuer, R. J. Kazlauskas, *Angew. Chemie* **2004**, *116*, 6156–6165.
- [11] S. C. Hammer, A. M. Knight, F. H. Arnold, *Curr. Opin. Green Sustain. Chem.* **2017**, *7*, 23–30.
- [12] H. Renata, Z. J. Wang, F. H. Arnold, *Angew. Chemie - Int. Ed.* **2015**, *54*, 3351–3367.
- [13] K. Faber, *Biotransformations in Organic Chemistry: A Textbook*, Springer Verlag, Heidelberg, **2011**.
- [14] G. Torrelo, U. Hanefeld, F. Hollmann, *Catal. Letters* **2014**, *145*, 309–345.
- [15] B. M. Nestl, B. A. Nebel, B. Hauer, *Curr. Opin. Chem. Biol.* **2011**, *15*, 187–93.
- [16] R. A. Sheldon, D. Brady, *Chem. Commun.* **2018**, *54*, 6088–6104.
- [17] U. T. Bornscheuer, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2018**, *376*, 1–7.
- [18] J. L. Porter, R. A. Rusli, D. L. Ollis, *ChemBioChem* **2016**, *17*, 197–203.
- [19] A. S. Bommarius, *Annu. Rev. Chem. Biomol. Eng.* **2015**, *6*, 319–345.
- [20] M. Beller, W. Leitner, J. O. Metzger, P. Saling, *Nachrichten aus der Chemie* **2008**, *56*, 480–484.
- [21] R. A. Sheldon, J. M. Woodley, *Chem. Rev.* **2018**, *118*, 801–838.
- [22] B. List, J. L. Kennemur, *Synfacts* **2018**, *14*, 1300.
- [23] A. Braga, C. Guerreiro, I. Belo, *Food Bioprocess Technol.* **2018**.
- [24] E. J. Vandamme, *Fungal Divers.* **2003**, *13*, 153–166.
- [25] G. Ohloff, W. Pickenhagen, P. Kraft, *Scent and Chemistry - The Molecular World of Odors*, Wiley-VCH, Weinheim, **2012**.
- [26] J. Schrader, M. M. W. Etschmann, D. Sell, J.-M. Hilmer, J. Rabenhorst, *Biotechnol. Lett.* **2004**, *26*, 463–472.
- [27] C. Gupta, D. Prakash, S. Gupta, *J. Microbiol. Exp.* **2015**, *2*, 1–8.

- [28] R. Kaiser, P. Kraft, *Chemie unserer Zeit* **2001**, *35*, 8–23.
- [29] P. Kraft, J. a Bajgrowicz, C. Denis, G. Fráter, *Angew. Chemie* **2000**, *112*, 3106–3138.
- [30] A. Goeke, P. Kraft, D. Lelievre, A. Alchenberger, *Perfum. Flavorist* **2018**, *43*.
- [31] W. Skorianetz, H. Giger, G. Ohloff, *Helv. Chim. Acta* **1971**, *54*, 1797–1801.
- [32] A. Alchenberger, C. Berbez, C. Finn, D. Lelievre, M. A. Lovchik, R. Poignon-Martel, G. Romey, *Improvements in or Relating to Organic Compounds*, **2014**, WO2014/198709 A1, US14651719, US20150307432A1, US9469590B2, US15212864.
- [33] AccuStandard Europe, “Allergens in Cosmetics and Personal Care Products,” can be found under [http://www.vogel-gmbh.ch/accustandard/pdf/allergens/allergens\\_brochure.pdf](http://www.vogel-gmbh.ch/accustandard/pdf/allergens/allergens_brochure.pdf), **2004**.
- [34] P. J. Frosch, J. D. Johansen, T. Menne, S. C. Rastogi, M. Bruze, K. E. Andersen, J. P. Lepoittevin, E. Gimenez Arnau, C. Pirker, A. Goossens, et al., *Br. J. Dermatol.* **1999**, *141*, 1076–1083.
- [35] W. J. Houlihan, J. Levy, J. Mayer, *J. Am. Chem. Soc.* **1959**, *81*, 4692–4694.
- [36] C. S. Sell, *A Fragrant Introduction to Terpenoid Chemistry*, The Royal Society Of Chemistry, Cambridge, **2003**.
- [37] J. H. Blumenthal, *Myrcene Sulfone Hydrate and Process for Making Same*, **1965**, US 3,176,022.
- [38] K. P. C. Vollhardt, N. E. Schore, *Organische Chemie*, Weinheim, **2005**.
- [39] R. Brückner, *Reaktionmechanismen*, Elsevier, München, **2004**.
- [40] K. Chen, P. S. Baran, *Nature* **2009**, *459*, 824–828.
- [41] C. J. Pierce, M. K. Hilinski, *Org. Lett.* **2014**, *16*, 6504–6507.
- [42] E. McNeill, J. Du Bois, *J. Am. Chem. Soc.* **2010**, *132*, 10202–10204.
- [43] T. Laue, A. Plagens, *Namen- Und Schlagwort-Reaktionen Der Organischen Chemie*, Vieweg + Teubner, Wiesbaden, **2009**.
- [44] “Namensreaktionen - Hydroborierung, Portal für Organische Chemie,” can be found under <http://www.organische-chemie.ch/OC/Namen/Hydroborierung.htm>, **2015**.
- [45] K. A. Bichler, S. G. Van Ornum, M. C. Franz, A. M. Imhoff, C. M. May, *J. Chem. Educ.* **2015**, *92*, 1426–1428.
- [46] S. Kubik, “Römpf-Online - Hofmann-Sand-Reaktion,” can be found under <https://roempp.thieme.de/roempp4.0/do/data/RD-08-01579>
- [47] P. V. Ramachandran, M. P. Jennings, *Org. Lett.* **2001**, *3*, 3789–3790.
- [48] S. Isayama, T. Mukaiyama, *Chem. Lett.* **1989**, *18*, 1071–1074.
- [49] A. Kuramochi, H. Usuda, K. Yamatsugu, M. Kanai, M. Shibasaki, *J. Am. Chem. Soc.* **2005**, *127*, 14200–14201.
- [50] D. Enders, A. Ridder, *Synthesis* **2000**, *13*, 1848–1851.
- [51] V. Resch, U. Hanefeld, in *Sci. Synth. - Biocatal. Org. Synth. 2*, Georg Thieme Verlag, Stuttgart, **2015**, pp. 261–290.
- [52] J. Jin, U. Hanefeld, *Chem. Commun.* **2011**, *47*, 2502–10.

- [53] V. Resch, U. Hanefeld, *Catal. Sci. Technol.* **2015**, *5*, 1385–1399.
- [54] J. L. F. Monteiro, C. O. Veloso, *Top. Catal.* **2004**, *27*, 169–180.
- [55] P. N. Davey, M. J. Earle, J. T. Hamill, S. P. Katdare, D. W. Rooney, K. R. Seddon, *Green Chem.* **2010**, *12*, 628.
- [56] Y. Liu, Z. Zhou, G. Yang, Y. Wu, Z. Zhang, *Int. J. Chem. React. Eng.* **2010**, *8*, 1–15.
- [57] P. Botella, A. Corma, J. M. L. Nieto, S. Valencia, M. . Lucas, M. Sergio, *Appl. Catal. A Gen.* **2000**, *203*, 251–258.
- [58] N. Comelli, M. C. Avila, C. Volzone, M. Ponzi, *Cent. Eur. J. Chem.* **2013**, *11*, 689–697.
- [59] M. C. Ávila, N. A. Comelli, E. Rodríguez-Castellón, A. Jiménez-López, R. Carrizo Flores, E. N. Ponzi, M. I. Ponzi, *J. Mol. Catal. A Chem.* **2010**, *322*, 106–112.
- [60] T. Mochida, R. Ohnishi, N. Horita, Y. Kamiya, T. Okuhara, *Microporous Mesoporous Mater.* **2007**, *101*, 176–183.
- [61] P. Brunerie, I. Benda, G. Bock, P. Schreier, *Appl. Microbiol. Biotechnol.* **1987**, *27*, 6–10.
- [62] G. Bock, I. Benda, P. Schreier, *J. Food Sci.* **1986**, *51*, 659–662.
- [63] A. S. Leutou, G. Yang, V. N. Nenkep, X. N. Siwe, Z. Feng, T. T. Khong, H. D. Choi, J. S. Kang, B. W. Son, *J. Microbiol. Biotechnol.* **2009**, *19*, 1150–1152.
- [64] S. S. Joglekar, R. S. Dhavlikar, *Appl. Microbiol.* **1969**, *18*, 1084–1087.
- [65] M. Miyazawa, H. Nankai, H. Kameoka, *Phytochemistry* **1996**, *43*, 105–109.
- [66] M. Miyazawa, H. Nankai, H. Kameoka, *Phytochemistry* **1995**, *40*, 1133–1137.
- [67] M. Miyazawa, H. Nankai, H. Kameoka, *J. Agric. Food Chem.* **1996**, *44*, 1543–1547.
- [68] B. N. M. van Leeuwen, A. M. van der Wulp, I. Duijnste, A. J. a van Maris, A. J. J. Straathof, *Appl. Microbiol. Biotechnol.* **2012**, *93*, 1377–87.
- [69] M. Engleder, H. Pichler, *Appl. Microbiol. Biotechnol.* **2018**, *102*, 5841–5858.
- [70] U. Krings, R. G. Berger, *Appl. Microbiol. Biotechnol.* **1998**, *49*, 1–8.
- [71] J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemie*, Spektrum, Heidelberg, **2007**.
- [72] D. Umeno, A. V Tobias, F. H. Arnold, *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 51–78.
- [73] R. M. Demming, M.-P. Fischer, J. Schmid, B. Hauer, *Curr. Opin. Chem. Biol.* **2018**, *43*, 43–50.
- [74] D. Brodkorb, M. Gottschall, R. Marmulla, F. Lüddecke, J. Harder, *J. Biol. Chem.* **2010**, *285*, 30436–42.
- [75] S. Weidenweber, R. Marmulla, U. Ermler, J. Harder, *FEBS Lett.* **2016**, 1–9.
- [76] F. Lüddecke, J. Harder, *Z. Naturforsch. C.* **2011**, *66*, 409–12.
- [77] F. Lüddecke, Genetische Und Biochemische Charakterisierung von Enzymen Des Anaeroben Monoterpen-Abbaus in Castellaniella Defragrans, Universität Bremen, **2011**.
- [78] B. M. Nestl, C. Geinitz, S. Popa, S. Rizek, R. J. Haselbeck, R. Stephen, M. A.

- Noble, M.-P. Fischer, E. C. Ralph, H. T. Hau, et al., *Nat. Chem. Biol.* **2017**, *13*, 275–281.
- [79] B. Ling, X. Wang, H. Su, R. Liu, Y. Liu, *Phys. Chem. Chem. Phys.* **2018**, *20*, 17342–17352.
- [80] M. S. Garcez Lopes, A. M. Slovic, P. Luiz de Andrade Coutinho, *Modified Microorganisms and Methods of Co-Producing Butadiene with 1-Propanol and/or 1,2-Propanediol*, **2015**, US 2015/0152440 A1.
- [81] P. Marliere, *Production of Volatile Dienes by Enzymatic Dehydration of Light Alkenols*, **2014**, US 8,703,455 B2.
- [82] P. J. Kuhn, D. A. Smith, *Physiol. Plant Pathol.* **1979**, *14*, 179–190.
- [83] D. Li, K. R. Chung, D. A. Smith, C. L. Schardl, *Mol. Plant. Microbe. Interact.* **1995**, *8*, 388–397.
- [84] M. Engleder, M. Horvat, A. Emmerstorfer-Augustin, T. Wriessnegger, S. Gabriel, G. Strohmeier, H. Weber, M. Müller, I. Kaluzna, D. Mink, et al., *PLoS One* **2018**, *13*, DOI 10.1371/journal.pone.0192653.
- [85] A. R. Moise, S. Al-Babili, E. T. Wurtzel, *Chem. Rev.* **2014**, *114*, 164–93.
- [86] Z. Sun, S. Shen, C. Wang, H. Wang, Y. Hu, J. Jiao, T. Ma, B. Tian, Y. Hua, *Microbiology* **2009**, *155*, 2775–2783.
- [87] A. A. Yeliseev, S. Kaplan, *FEBS Lett.* **1997**, *403*, 10–14.
- [88] N. Patel, G. Britton, T. Goodwin, *Biochim. Biophys. Acta - Gen. Subj.* **1983**, *760*, 92–96.
- [89] P. A. Scolnik, M. A. Walker, B. L. Marrs, *J. Biol. Chem.* **1980**, *255*, 2427–2432.
- [90] G. A. Armstrong, M. Alberti, F. Leach, J. E. Hearst, *MGG Mol. Gen. Genet.* **1989**, *216*, 254–268.
- [91] S. Steiger, S. Takaichi, G. Sandmann, *J. Biotechnol.* **2002**, *97*, 51–58.
- [92] S. Steiger, A. Mazet, G. Sandmann, *Arch. Biochem. Biophys.* **2003**, *414*, 51–58.
- [93] A. T. Kovacs, G. Rakhely, K. L. Kovacs, *Appl. Environ. Microbiol.* **2003**, *69*, 3093–3102.
- [94] D. Umeno, A. V Tobias, F. H. Arnold, *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 51–78.
- [95] J. E. Graham, D. A. Bryant, *J. Bacteriol.* **2009**, *191*, 3292–3300.
- [96] A. Hiseni, I. W. C. E. Arends, L. G. Otten, *Appl. Microbiol. Biotechnol.* **2011**, *91*, 1029–1036.
- [97] A. Hiseni, L. G. Otten, I. W. C. E. Arends, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 1275–1284.
- [98] R. Marmulla, J. Harder, *Front. Microbiol.* **2014**, *5*, 1–14.
- [99] B.-S. Chen, L. G. Otten, U. Hanefeld, *Biotechnol. Adv.* **2015**, *33*, 526–546.
- [100] W. A. Duetz, H. Bouwmeester, J. B. Beilen, B. Witholt, *Appl. Microbiol. Biotechnol.* **2003**, *61*, 269–277.
- [101] M. R. Maróstica, G. M. Pastore, *Food Chem.* **2007**, *101*, 345–350.
- [102] K. R. Cadwallader, R. J. Braddock, M. E. Parish, D. P. Higgins, *J. Food Sci.* **1989**, *54*, 1241–1245.

- [103] K. R. Cadwallader, R. J. Braddock, M. E. Parish, *J. Food Sci.* **1992**, *57*, 241–244.
- [104] H. C. Chang, D. A. Gage, P. J. Oriel, *J. Food Sci.* **1995**, *60*, 551–553.
- [105] H. C. Chang, P. Oriel, *J. Food Sci.* **1994**, *59*, 660–662.
- [106] N. Savithiry, T. K. Cheong, P. Oriel, *Appl. Biochem. Biotechnol.* **1997**, *63–65*, 213–220.
- [107] N. Savithiry, P. J. Oriel, *Method for Production of Monoterperene Derivates of Limonene*, **1998**, US 5,763,237.
- [108] J. L. Bicas, F. F. C. Barros, R. Wagner, H. T. Godoy, G. M. Pastore, *J. Ind. Microbiol. Biotechnol.* **2008**, *35*, 1061–1070.
- [109] Q. Tan, D. F. Day, *Appl. Microbiol. Biotechnol.* **1998**, *49*, 96–101.
- [110] M. Pescheck, M. A. Mirata, B. Brauer, U. Krings, R. G. Berger, J. Schrader, *J. Ind. Microbiol. Biotechnol.* **2009**, *36*, 827–836.
- [111] G. A. Prieto, J. A. Perea V., C. C. Ortiz L., *Vitae, Rev. la Fac. Quim. Farm.* **2011**, *18*, 163–172.
- [112] J. L. Bicas, P. Fontanille, G. M. Pastore, C. Larroche, *Process Biochem.* **2010**, *45*, 481–486.
- [113] L. L. Wallen, R. G. Benedict, R. W. Jackson, *Arch. Biochem. Biophys.* **1962**, *99*, 249–253.
- [114] L. E. Bevers, M. W. H. Pinkse, P. D. E. M. Verhaert, W. R. Hagen, *J. Bacteriol.* **2009**, *191*, 5010–2.
- [115] J. Schmid, L. Steiner, S. Fademrecht, J. Pleiss, K. B. Otte, B. Hauer, *J. Mol. Catal. B Enzym.* **2016**, *133*, S243–S249.
- [116] Y.-C. Joo, K.-W. Jeong, S.-J. Yeom, Y.-S. Kim, Y. Kim, D.-K. Oh, *Biochimie* **2012**, *94*, 907–15.
- [117] A. Volkov, S. Khoshnevis, P. Neumann, C. Herrfurth, D. Wohlwend, R. Ficner, I. Feussner, *Acta Crystallogr. D. Biol. Crystallogr.* **2013**, *69*, 648–57.
- [118] C. J. Zheng, J.-S. Yoo, T.-G. Lee, H.-Y. Cho, Y.-H. Kim, W.-G. Kim, *FEBS Lett.* **2005**, *579*, 5157–62.
- [119] A. Volkov, A. Liavonchanka, O. Kamneva, T. Fiedler, C. Goebel, B. Kreikemeyer, I. Feussner, *J. Biol. Chem.* **2010**, *285*, 10353–10361.
- [120] M. Engleder, T. Pavkov-Keller, A. Emmerstorfer, A. Hromic, S. Schrempf, G. Steinkellner, T. Wriessnegger, E. Leitner, G. A. Strohmeier, I. Kaluzna, et al., *ChemBioChem* **2015**, *16*, 1730–1734.
- [121] J. Lorenzen, R. Driller, A. Waldow, F. Qoura, B. Loll, T. Brück, *ChemCatChem* **2018**, *10*, 407–414.
- [122] A. K. Park, G. H. Lee, D. W. Kim, E. H. Jang, H. T. Kwon, Y. M. Chi, *Biochem. Biophys. Res. Commun.* **2018**, 1–5.
- [123] J. Ortega-Anaya, A. Hernández-Santoyo, *Biochim. Biophys. Acta - Biomembr.* **2015**, *1848*, 3166–3174.
- [124] S. Koritala, C. T. Hou, C. W. Hesseltine, M. O. Bagby, *Appl. Microbiol. Biotechnol.* **1989**, *32*, 299–304.
- [125] W. G. Niehaus, A. Torkelson, A. Kisic, D. J. Bednarczyk, G. J. Schroepfer, *J. Biol. Chem.* **1970**, *245*, 3790–7.

- [126] W. G. Niehaus, G. J. Schroepfer, *Biochem. Biophys. Res. Commun.* **1965**, *21*, 271–275.
- [127] R. M. Demming, K. B. Otte, B. M. Nestl, B. Hauer, *ChemCatChem* **2017**, *9*, 758–766.
- [128] A. Hiseni, I. W. C. E. Arends, L. G. Otten, *ChemCatChem* **2015**, *7*, 29–37.
- [129] P. Marliere, *Method for Producing an Alkene Comprising the Step of Converting an Alcohol by an Enzymatic Dehydration Step*, **2011**, WO2011/076691A1 / EP2336341A1.
- [130] R. M. Demming, S. C. Hammer, B. M. Nestl, S. Gergel, S. Fademrecht, J. Pleiss, B. Hauer, *Angew. Chemie Int. Ed.* **2019**, *58*, 173–177.
- [131] K.-R. Kim, H.-J. Oh, C.-S. Park, S.-H. Hong, J.-Y. Park, D.-K. Oh, *Biotechnol. Bioeng.* **2015**, *112*, 2206–2213.
- [132] J.-Y. J.-B. Park, S.-H. Lee, K.-R. Kim, J.-Y. J.-B. Park, D.-K. Oh, *J. Biotechnol.* **2015**, *208*, 1–10.
- [133] A. Hirata, S. Kishino, S.-B. Park, M. Takeuchi, N. Kitamura, J. Ogawa, *J. Lipid Res.* **2015**, *56*, 1340–1350.
- [134] K. Zorn, I. Oroz-Guinea, H. Brundiek, U. T. Bornscheuer, *Prog. Lipid Res.* **2016**, *63*, 153–164.
- [135] B. N. Kim, Y. C. Joo, Y. S. Kim, K. R. Kim, D. K. Oh, *Appl. Microbiol. Biotechnol.* **2012**, *95*, 929–937.
- [136] Y.-C. Joo, E.-S. Seo, Y.-S. Kim, K.-R. Kim, J.-B. Park, D.-K. Oh, *J. Biotechnol.* **2012**, *158*, 17–23.
- [137] W.-R. Kang, M.-J. Seo, K.-C. Shin, J.-B. Park, D.-K. Oh, *Biotechnol. Bioeng.* **2017**, *114*, 74–82.
- [138] S. Kim, Y. Park, *Appl. Microbiol. Biotechnol.* **2018**, DOI 10.1007/s00253-018-9503-6.
- [139] K. B. Otte, E. Maurer, M. Kirtz, D. Grabs, E. Althoff, S. Bartsch, A. Vogel, B. M. Nestl, B. Hauer, *ChemCatChem* **2017**, *9*, 1378–1382.
- [140] S. Serra, D. De Simeis, *Catalysts* **2018**, *8*, 109.
- [141] H.-J. Oh, S.-U. Kim, J.-W. Song, J.-H. Lee, W.-R. Kang, Y.-S. Jo, K.-R. Kim, U. T. Bornscheuer, D.-K. Oh, J.-B. Park, *Adv. Synth. Catal.* **2015**, *357*, 408–416.
- [142] U. T. Bornscheuer, M. Pohl, *Curr. Opin. Chem. Biol.* **2001**, *5*, 137–43.
- [143] T. Davids, M. Schmidt, D. Böttcher, U. T. Bornscheuer, *Curr. Opin. Chem. Biol.* **2013**, *17*, 215–220.
- [144] F. H. Arnold, *Acc. Chem. Res.* **1998**, *31*, 125–131.
- [145] “The Nobel Prize in Chemistry 2018,” can be found under <https://www.nobelprize.org/prizes/chemistry/2018/press-release/>, **2018**.
- [146] K. Chen, F. H. Arnold, *Proc. Natl. Acad. Sci.* **1993**, *90*, 5618–5622.
- [147] “PDB Statistics,” can be found under <https://www.rcsb.org/stats/growth/overall>
- [148] R. A. Chica, N. Doucet, J. N. Pelletier, *Curr. Opin. Biotechnol.* **2005**, *16*, 378–384.
- [149] M. T. Reetz, J. D. Carballeira, *Nat. Protoc.* **2007**, *2*, 891–903.
- [150] M. T. Reetz, M. Bocola, J. D. Carballeira, D. Zha, A. Vogel, *Angew. Chemie -*

- Int. Ed.* **2005**, *44*, 4192–4196.
- [151] M. T. Reetz, D. Kahakeaw, R. Lohmer, *ChemBioChem* **2008**, *9*, 1797–1804.
- [152] M. T. Reetz, *J. Am. Chem. Soc.* **2013**, *135*, 12480–12496.
- [153] C. G. Acevedo-Rocha, M. T. Reetz, Y. Nov, *Sci. Rep.* **2015**, *5*, 1–12.
- [154] ATCC, Geobacillus Stearotherophilus BR388 - Product Sheet
- [155] C. Pennacchia, P. Breeuwer, R. Meyer, *Food Microbiol.* **2014**, *43*, 41–49.
- [156] O. Reznicek, S. J. Facey, B. Hauer, *Genome Announc* **2015**, *3*, 1999–2000.
- [157] R. M. Demming, Enzymatische Hydratisierung Kurzkettiger Fettsäuren Und Alkene, University of Stuttgart, **2018**.
- [158] N. Kreß, Development of a Chemoenzymatic (-)-Menthol Synthesis, University of Stuttgart, **2018**.
- [159] L. Steiner, Biokatalytische Dehydratisierung von Alkoholen Mittels Oleathydratasen, Universität Stuttgart, **2014**.
- [160] H. C. Chang, P. J. Oriel, *Preparation of Perillyl Compounds Using Bacillus Stearotherophilus*, **1996**, 5,487,988.
- [161] W. G. Weisburg, S. M. Barns, D. a Pelletie, D. J. Lane, D. A. Pelletier, D. J. Lane, *J. Bacteriol.* **1991**, *173*, 697–703.
- [162] D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, *Nat. Methods* **2009**, *6*, 343–345.
- [163] D. Gibson, *Protoc. Exch.* **2009**, DOI 10.1038/nprot.2009.77.
- [164] Agilent, *QuikChange Site-Directed Mutagenesis Kit*, La Jolla, **2010**.
- [165] D. Seebach, R. Imwinkelried, G. Stucky, *Helv. Chim. Acta* **1987**, *70*, 448–464.
- [166] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512–7515.
- [167] T. N. Nazina, T. P. Tourova, A. B. Poltarau, E. V. Novikova, A. A. Grigoryan, A. E. Ivanova, A. M. Lysenko, V. V. Petrunyaka, G. A. Osipov, S. S. Belyaev, et al., *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 433–446.
- [168] M. L. Caipo, S. Duffy, L. Zhao, D. W. Schaffner, *J. Appl. Microbiol.* **2002**, *92*, 879–884.
- [169] M. K. Llaudes, L. Zhao, S. Duffy, D. W. Schaffner, *Food Microbiol.* **2001**, *18*, 395–405.
- [170] Y. Zhao, M. P. M. Caspers, K. I. Metselaar, P. de Boer, G. Roeselers, R. Moezelaar, M. Nierop Groot, R. C. Montijn, T. Abee, R. Kort, *Appl. Environ. Microbiol.* **2013**, *79*, 5652–5660.
- [171] S. Flint, J. Palmer, K. Bloemen, J. Brooks, R. Crawford, *J. Appl. Microbiol.* **2001**, *90*, 151–157.
- [172] T. K. Cheong, P. J. Oriel, *Appl. Biochem. Biotechnol.* **2000**, *84–86*, 903–915.
- [173] T. K. Cheong, P. J. Oriel, *Enzyme Microb. Technol.* **2000**, *26*, 152–158.
- [174] S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z. G. Zhang, D. J. Opperman, M. T. Reetz, J. P. Acevedo, *ACS Synth. Biol.* **2013**, *2*, 83–92.
- [175] J. Esque, C. Oguey, A. G. de Brevern, *J. Chem. Inf. Model.* **2010**, *50*, 947–960.
- [176] C. C. C. R. de Carvalho, M. M. R. da Fonseca, *Biotechnol. Adv.* **2006**, *24*, 134–142.

- [177] I. Schnapperelle, W. Hummel, H. Gröger, *Chem. - A Eur. J.* **2012**, *18*, 1073–1076.
- [178] A. Hiseni, Study Towards Carotenoid 1,2-Hydratase and Oleate Hydratase as Novel Biocatalysts, Technische Universiteit Delft, **2014**.
- [179] M. Kadisch, M. K. Julsing, M. Schrewe, N. Jehmlich, B. Scheer, M. von Bergen, A. Schmid, B. Bühler, *Biotechnol. Bioeng.* **2016**, *9999*, 1–11.
- [180] K. Hult, P. Berglund, *Trends Biotechnol.* **2007**, *25*, 231–238.
- [181] A. Babbie, N. Tokuriki, F. Hollfelder, *Curr. Opin. Chem. Biol.* **2010**, *14*, 200–207.
- [182] M. H. J. Wells-Bennik, P. W. M. Janssen, V. Klaus, C. Yang, M. H. Zwietering, H. M. W. den Besten, *Int. J. Food Microbiol.* **2018**, DOI 10.1016/j.ijfoodmicro.2018.11.005.
- [183] M. Kakagianni, J. S. Aguirre, A. Lianou, K. P. Koutsoumanis, *Food Microbiol.* **2017**, *67*, 76–84.
- [184] M. Javed, N. Baghaei-Yazdi, W. Qin, S. Amartey, *J. Microbiol. Methods* **2017**, *132*, 116–118.
- [185] M. Kakagianni, M. Gougouli, K. P. Koutsoumanis, *Food Microbiol.* **2016**, *57*, 28–35.
- [186] H. Aliyu, P. Lebre, J. Blom, D. Cowan, P. De Maayer, *Syst. Appl. Microbiol.* **2016**, *39*, 527–533.
- [187] J. Eyal, M. Charles, *J. Ind. Microbiol.* **1990**, *6*, 185–190.
- [188] S. Prasad, T. C. Bhalla, *Biotechnol. Adv.* **2010**, *28*, 725–741.
- [189] C. A. MacDonald, R. J. Boyd, *Comput. Theor. Chem.* **2015**, *1070*, 48–54.
- [190] H. Suzuki, *Appl. Microbiol. Biotechnol.* **2018**, DOI 10.1007/s00253-018-9422-6.
- [191] P. Bramley, R. Taylor, *Biochim. Biophys. Acta - Gen. Subj.* **1985**, *839*, 155–160.
- [192] S. C. Hammer, A. Marjanovic, J. M. Dominicus, B. M. Nestl, B. Hauer, *Nat. Chem. Biol.* **2014**, *11*, 121–126.
- [193] M. Seitz, J. Klebensberger, S. Siebenhaller, M. Breuer, G. Siedenburg, D. Jendrosseck, B. Hauer, *J. Mol. Catal. B Enzym.* **2012**, *2–7*.
- [194] S. C. Hammer, J. M. Dominicus, P.-O. Syrén, B. M. Nestl, B. Hauer, *Tetrahedron* **2012**, *68*, 7624–7629.
- [195] E. Eichhorn, E. Locher, S. Guillemer, D. Wahler, L. Fourage, B. Schilling, *Adv. Synth. Catal.* **2018**, *4*, DOI 10.1002/adsc.201800132.
- [196] A. Todea, A. Hiseni, L. G. Otten, I. W. C. E. Arends, F. Peter, C. G. Boeriu, *J. Mol. Catal. B Enzym.* **2015**, *119*, 40–47.
- [197] A. Hiseni, R. Medici, I. W. C. E. Arends, L. G. Otten, *Biotechnol. J.* **2014**, *9*, 814–821.
- [198] S. Kleehammer, Engineering Der Oleat Hydratase Aus Elizabethkingia Meningoseptica, University of Stuttgart, **2016**.
- [199] A. Hunold, Active-site investigation of the oleate hydratase OAH1 from Elizabethkingia meningoseptica, Heinrich Heine University Düsseldorf, **2016**.
- [200] M. C. Ebert, J. N. Pelletier, *Curr. Opin. Chem. Biol.* **2017**, *37*, 89–96.
- [201] Q. Yu, P. Schaub, S. Ghisla, S. Al-Babili, A. Krieger-Liszkay, P. Beyer, *J. Biol.*



- Chem.* **2010**, 285, 12109–12120.
- [202] M. S. Humble, P. Berglund, *European J. Org. Chem.* **2011**, 2011, 3391–3401.
- [203] J. A. Gerlt, P. C. Babbitt, *Curr. Opin. Chem. Biol.* **2009**, 13, 10–18.
- [204] D. Suplatov, E. Kirilin, V. Takhaveev, V. Švedas, *J. Biomol. Struct. Dyn.* **2014**, 32, 1752–1758.
- [205] S. C. Hammer, G. Kubik, E. Watkins, S. Huang, H. Minges, F. H. Arnold, *Science* **2017**, 358, 215–218.
- [206] M. T. Reetz, *Angew. Chemie Int. Ed.* **2013**, 52, 2658–2666.