

# Structure-function relationships in wax producing enzymes

## Dissertation

for the award of the degree

“Doctor rerum naturalium”

of the Georg-August-Universität Göttingen

within the doctoral program “GGNB Biomolecules: Structure-Function-Dynamics”

of the Georg-August University School of Science (GAUSS)

submitted by

**Steffen Justus Kawelke**

from Aachen

Göttingen, 2014



## **Members of the Thesis Committee**

### **Prof. Dr. Ivo Feußner**

Department for Plant Biochemistry, Albrecht-von-Haller-Institute for Plant Sciences, University of Göttingen

### **Prof. Dr. Jörg Stülke**

Department of General Microbiology, University of Göttingen

### **Dr. Karin Kühnel**

Department of Neurobiology, Max Planck Institute for Biophysical Chemistry, Göttingen

## **Members of the Examination Board**

### Reviewer

#### **Prof. Dr. Ivo Feußner**

Department for Plant Biochemistry, Albrecht-von-Haller-Institute for Plant Sciences, University of Göttingen

### Reviewer

#### **Prof. Dr. Jörg Stülke**

Department of General Microbiology, University of Göttingen

### **Dr. Karin Kühnel**

Department of Neurobiology, Max Planck Institute for Biophysical Chemistry, Göttingen

### **Prof. Dr. Marina Bennati**

Electron Paramagnetic Resonance Spectroscopy Group, Max Planck Institute for Biophysical Chemistry, Göttingen

### **PD Dr. Thomas Teichmann**

Department of Plant Cell Biology, Albrecht-von-Haller-Institute for Plant Sciences, University of Göttingen

### **Prof. Dr. Ralph Kehlenbach**

Department of Biochemistry I, Göttingen University Medical School

**Date of the oral examination: December 3, 2014**

## **Affidavit**

Herewith I declare that I wrote the PhD thesis on my own and without other sources and aids than quoted.

Steffen Justus Kawelke

Göttingen, October 2014

## INDEX

<b>1</b>	<b>INTRODUCTION .....</b>	<b>1</b>
1.1	Physical properties of wax esters (WEs) .....	2
1.2	Industrial application of WEs .....	3
1.3	Enzymatic basis of WE biosynthesis .....	4
1.4	Fatty acyl reductases (FARs).....	5
1.4.1	Phylogenetic distribution and subcellular localisation of FARs .....	6
1.4.2	Catalytic motifs in FARs .....	8
1.4.3	Substrate specificity of FARs .....	9
1.4.4	Biotechnological applications of FARs.....	11
1.4.5	Role of free fatty alcohols in nature.....	12
1.5	Wax synthases (WSs).....	13
1.5.1	Phylogenetic distribution of WSs .....	14
1.5.2	Protein structure and subcellular localisation of WSs .....	16
1.5.3	Catalytic motifs within WSs.....	18
1.5.4	Substrate specificities of WSs .....	20
1.5.4.1	Substrate specificities of WSD-type WS .....	20
1.5.4.2	Substrate specificities of DGAT2-type WS .....	21
1.5.4.3	Substrate specificities of DGAT1-type WS .....	22
1.5.5	Biotechnological applications of WSs .....	23
1.5.6	Role of WEs in nature.....	25
<b>2</b>	<b>OBJECTIVES OF THIS WORK.....</b>	<b>28</b>
<b>3</b>	<b>MATERIALS.....</b>	<b>29</b>
3.1	Software and web-based services.....	29
3.2	Machines and equipment .....	30
3.3	Chemicals .....	31
3.4	Kits, markers and consumables.....	31
3.5	Commercial crystallisation screens .....	32
3.6	Customised crystallisation grid-screens .....	32
3.7	Enzymes.....	34
3.8	Strains and organisms .....	35
3.9	Columns for chromatography .....	35
3.10	Oligonucleotides.....	36
3.11	DNA constructs.....	38
3.12	Buffers, media and antibiotics .....	40
<b>4</b>	<b>METHODS.....</b>	<b>42</b>
4.1	Statistical methods.....	42
4.1.1	Student's t-test.....	42
4.1.2	Tukey's test .....	42
4.2	Molecular biology methods .....	42
4.2.1	Preparation of competent <i>E. coli</i> cells .....	42
4.2.2	Heat shock transformation of <i>E. coli</i> .....	43
4.2.3	Preparation of competent <i>S. cerevisiae</i> cells .....	43
4.2.4	Transformation of <i>S. cerevisiae</i> .....	43
4.2.5	Standard PCR.....	44
4.2.6	Colony PCR .....	44
4.2.7	Overlap extension PCR .....	45
4.2.8	Quick-Change PCR .....	45
4.2.9	Restriction of DNA.....	45
4.2.10	Ligation of DNA .....	46
4.2.11	Plasmid preparation from <i>E. coli</i> .....	46
4.2.12	Separation of DNA by agarose gel-electrophoresis .....	46
4.2.13	Cultivation of <i>E. coli</i> Lemo21 (DE3) for protein expression screenings .....	46

## INDEX

4.2.14	Cultivation of <i>E. coli</i> expression cultures for protein purification .....	47
4.2.15	Cultivation of <i>E. coli</i> for production of 6xHis-TF-AbWSD1 selenoprotein.....	47
4.2.16	Cultivation of <i>S. cerevisiae</i> .....	48
4.3	Analytical methods.....	48
4.3.1	WE extraction from <i>S. cerevisiae</i> and <i>E. coli</i> .....	48
4.3.2	Fatty alcohol extraction from <i>S. cerevisiae</i> and <i>E. coli</i> .....	48
4.3.3	Thin layer chromatography analysis .....	48
4.3.4	GC-FID analysis .....	49
4.3.5	WE analysis via GC-MS .....	49
4.3.6	Wax ester molecular species profiling by nano-ESI-MS/MS .....	49
4.3.7	Fatty alcohol analysis via GC-MS.....	50
4.3.8	Extraction and analysis of acyl-CoA from <i>S. cerevisiae</i> .....	50
4.3.9	Thermal shift assay (Thermofluor).....	51
4.3.10	Multiangular light scattering (MALS) .....	51
4.4	Biochemical methods.....	52
4.4.1	Preparation of cell pellets for protein purification .....	52
4.4.2	Protein purification .....	52
4.4.2.1	Amylose resin affinity chromatography.....	52
4.4.2.2	Ni <sup>2+</sup> -based immobilised metal affinity chromatography.....	52
4.4.2.3	Affinity chromatography for separation of TF and fusion partner .....	53
4.4.2.4	Buffer compositions for purification of MaFAR1 from <i>M. aquaeolei</i> .....	54
4.4.2.5	Buffer compositions for purification of DmFAR1 from <i>D. melanogaster</i> .....	55
4.4.2.6	Buffer compositions for separation of 6xHis-TF and its fusion partner.....	56
4.4.2.7	Buffer compositions for purification of AtDGA3 from <i>A. thaliana</i> .....	57
4.4.3	Size exclusion chromatography (SEC) .....	57
4.4.4	SDS polyacrylamide gel electrophoresis (SDS-PAGE).....	58
4.4.5	Determination of protein concentrations according to Bradford.....	59
4.4.6	DTNB-based <i>in vitro</i> test of acyltransferases .....	59
4.4.7	NADPH-based <i>in vitro</i> test of acyltransferases .....	60
4.4.8	<i>In vitro</i> assays with [ <sup>14</sup> C] labelled substrates.....	60
4.5	Protein crystallisation.....	61
4.5.1	Grid screens for optimisation of crystallisation-conditions .....	61
<b>5</b>	<b>RESULTS .....</b>	<b>62</b>
5.1	Purification and characterisation of heterologously expressed FARs and WSs.....	62
5.1.1	Purification and analyses of FARs .....	64
5.1.1.1	DmFAR1 from <i>D. melanogaster</i> .....	66
5.1.1.1.1	DmFAR1-mediated fatty alcohol production in <i>S. cerevisiae</i> and <i>E. coli</i> .....	67
5.1.1.1.2	Recombinant expression of DmFAR1 in <i>E. coli</i> .....	71
5.1.1.2	TcFAR1 from <i>T. castaneum</i> .....	72
5.1.1.3	MmFAR1 from <i>M. musculus</i> .....	73
5.1.1.4	MaFAR1 from <i>M. aquaeolei</i> .....	74
5.1.1.4.1	Purification of heterologously produced MaFAR1.....	74
5.1.1.4.2	MaFAR1 is active <i>in vivo</i> and <i>in vitro</i> .....	78
5.1.2	Purification and analyses of acyltransferases .....	80
5.1.2.1	AtDGAT3 from <i>A. thaliana</i> .....	81
5.1.2.2	AbWSD1 from <i>A. baylyi</i> .....	84
5.1.2.2.1	Purification of a fusion protein from TF and AbWSD1 from <i>A. baylyi</i> .....	85
5.1.2.2.2	Separation of TF and AbWSD1 via chromatographic techniques .....	86
5.1.2.2.3	Structure prediction of AbWSD1.....	88
5.1.2.2.4	Purification of hexahistidine-tagged AbWSD1 from <i>A. baylyi</i> .....	89
5.2	Crystallisation of the 6xHis-TF-AbWSD1 fusion protein.....	90
5.3	Elucidation of substrate determinants in a murine WS.....	94

## INDEX

5.3.1	MmAWAT2 and MmDGAT2 show distinct acyl chain incorporation patterns in WE biosynthesis.....	94
5.3.2	Acyl-CoA pool composition in <i>S. cerevisiae</i> H1246 upon feeding of different alcohols	96
5.3.3	MmAWAT2 and MmDGAT2 show distinct specificities in respect to acyl chain incorporation into WEs .....	97
5.3.4	Construction of domain swap variants on the basis of MmAWAT2 and MmDGAT2 ....	98
5.3.5	Substrate specificities of domain swap variants .....	101
5.3.6	The predicted N-terminal transmembrane domain in MmAWAT2 is highly conserved among vertebrates.....	105
5.3.7	The substrate specificity in MmAWAT2 is influenced by different residues in the predicted transmembrane region .....	106
5.3.8	Activity of MmDGAT2 F80A and MmDGAT2 L91R .....	110
5.3.9	The putative transmembrane domains in MmAWAT2 might be dispensable for enzyme activity .....	111
5.3.10	Modelling of MmAWAT2 and MmDGAT2 .....	113
<b>6</b>	<b>DISCUSSION .....</b>	<b>115</b>
6.1	Purification of FARs and Ws .....	115
6.1.1	Enzymatic activity of DmFAR1.....	115
6.1.2	Purification of DmFAR1 .....	116
6.1.3	TcFAR1.....	117
6.1.4	Interaction of TF with MaFAR1 and DmFAR1.....	117
6.1.5	MaFAR1 .....	118
6.1.6	Models of FARs.....	119
6.1.7	AtDGAT3.....	121
6.1.8	Interaction of TF and AbWSD1.....	123
6.2	Substrate specificities of MmAWAT2.....	126
6.2.1	Altered WE/TAG ratios in MmAWAT2 variants.....	126
6.2.2	Incorporation of VLC acyl-CoA by MmAWAT2 variants .....	127
6.2.3	Role of a putative neutral lipid binding sequence in MmAWAT2.....	127
6.2.4	VLC WE production might be explained by altered substrate specificities of MmAWAT2 variants.....	128
6.2.5	VLC WE production might be explained by access of MmAWAT2 variants to new substrate pools.....	129
6.2.6	Truncation of MmAWAT2 .....	131
<b>7</b>	<b>SUMMARY .....</b>	<b>132</b>
<b>8</b>	<b>OUTLOOK AND PERSPECTIVE .....</b>	<b>133</b>
<b>9</b>	<b>LITERATURE.....</b>	<b>136</b>
<b>10</b>	<b>LIST OF ABBREVIATIONS .....</b>	<b>153</b>
<b>11</b>	<b>APPENDIX.....</b>	<b>155</b>
<b>12</b>	<b>ACKNOWLEDGEMENTS.....</b>	<b>171</b>



# 1 INTRODUCTION

Today, the global economy is mainly reliant on fossil carbon sources like oil, coal and gas as industrial feedstock and for energy production. Potential dangers accompanied with the use of fossil resources are illustrated by disasters like the Deepwater Horizon incident in 2010. As a long term effect of burning fossil resources, global warming has become an accepted scientific fact (Intergovernmental Panel on Climate Change, 1996). In addition, availability of fossil resources is not infinite. "Peak oil" describes the time point, on which the global oil production is highest and declines afterwards. Today it's assumed, that this event already happened in the last decade or is imminent (Alekklett et al., 2010). As a consequence, the public perception for clean and renewable resources is growing. To date, the biggest efforts towards the utilization of renewable resources are made in the energy and transportation sectors. In Germany, a total of 12,156 electric cars without conventional internal combustion engine were registered in January 2014 (Kraftfahrt-Bundesamt, 2014). This is an increase of approximately 100 % in comparison to the beginning of 2013, when 7114 electric cars were registered ( Kraftfahrt-Bundesamt, 2013). In the energy field, 40-45 % of the German national energy consumption in 2025 is planned to derive from renewable sources (Bundesministerium für Wirtschaft und Energie, 2014).

As in many technological fields, interests in contributing to a sustainable economy have also arisen in the biotechnological sector. Like in the general debate concerning renewable resources, public interests are mainly focused on biological solutions for the supply of energy and alternative car fuels. In this respect, crop plants like maize or sugar cane are grown in large scale in order to produce ethanol, biogas or biodiesel. A fundamental issue accompanied with this strategy is an ethical conflict termed as "food versus fuel" (BBC, 2007; Committee on world food security, 2013; International Centre for Trade and Sustainable Development, 2011). Taking into account that in some parts of the world sufficient nutrition is not ensured, burning crops for transportation purposes is highly questionable. To circumvent this conflict, algae-derived fuel could be an alternative (Chisti, 2013; Gimpel et al., 2013). Since algae don't need fertile soil, they can basically be grown anywhere. As a consequence, algae don't compete with crop plants. Moreover, most algae do not need fresh water since they grow in sea- or even wastewater. However, biotechnological efforts to optimise algal strains for production of high amounts of oil, starch or even hydrogen are in their infancy.

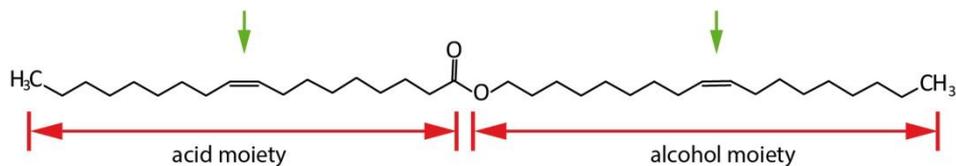
To date it seems rather unrealistic to be able to cover the entire world's demand of oil for energy production and transportation-purposes by biological sources. In Europe, around 2.87 million tons of fossil oil equivalent of biodiesel have been produced in 2005. This amount corresponds to a proportion of only 1.6 % of the total annual European diesel consumption (European Union, 2006). The worldwide annual production of vegetable oils is 142 million tons, which corresponds to only about 3 % of the annually consumed fossil oil. Of these 142 million tons, 80 % are processed into food, so that the actual amount which is available for other purposes is only about 0.6 % of the annually consumed fossil oil (Carlsson et al., 2011). Processing such a low amount of vegetable oil to a low price product like fuel appears to make little sense. In contrast, selling plant derived oils in high end application fields like chemical feedstock represents an economically attractive alternative. High end applications of plant oils include utilization as feedstock for chemical or pharmaceutical industry. As such, oils can be applied in cosmetics, as lubricants or as food coatings (Endlein and Peleikis, 2011). Wax esters (WE) are a specific type of plant derived oil and especially feasible for those applications (Carlsson, 2006; 2010).

## INTRODUCTION

### 1.1 Physical properties of wax esters (WEs)

WEs are esters from a fatty alcohol and a fatty acid. As such, WEs in general are linear, hydrophobic molecules (Figure 1). Depending on the nature of the esterified moieties, WEs can cover a broad range of physical properties. For instance, the melting temperature ( $T_m$ ) of WEs can reach from below zero degrees Celsius for oleyl-oleate to over 75 °C for tetracosanoyl-tetracosanoate (Patel et al., 2001). Thus, the overall chain length of a WE is determining its melting temperature, but also its mechanical stability. Short WEs are generally more stable and have a lower melting temperature, which is increasing with the chain length of WEs. In detail, the  $T_m$  of fully saturated WEs is increased by 1-2 °C with every additional single carbon atom (Patel et al., 2001). Interestingly, WEs of the same overall chain length can exhibit different melting temperatures, depending on where the ester bond is located within the molecule. For instance, WEs composed of two C16 moieties have a higher melting temperature in comparison to WEs composed of a C20 and a C12 moiety. Positioning of the ester bond in the middle of the molecule results in the highest melting temperature, while WEs with the same overall carbon number but moieties which are not of the same length possess lower melting temperatures.

Furthermore, the grade of saturation of a WE chain has an influence on its oxidation stability as well as on its melting temperature. As such, every double bond within a molecule increases the WE's susceptibility for oxidation (Hagemann and Rothfus, 1979). In addition, double bonds introduce kinks in the carbon chain, which lowers the molecule's melting temperature. Again, the position of a double bond within WEs is crucial for their  $T_m$ . Moreover, melting temperatures are also dependent on the general localisation of a double bond in either the alcohol or the acid moiety. For instance, octadecanoyl-octadecenotae (18:0-18:1) has a higher  $T_m$  than octadecenoyl-octadecanoate (18:1-18:0), although chain length and number of double bonds in both molecules are exactly the same (Patel et al., 2001).



**Figure 1: Chemical structure of WEs.** WEs can vary in chain length (red arrows) as well as in grade of desaturation and positions of double bonds (green arrows). Shown here is the structure of oleyl-oleate (18:1-18:1).

The majority of WEs found in nature is derived from primary alcohols and fatty acids. Apart from that, also WEs made from secondary alcohols and fatty acids have been described (Blomquist et al., 1972). Due to their fundamental different geometry, these WEs show different properties in comparison to their primary alcohol-derived isomers. Instead of being linear molecules, WEs from secondary alcohols adopt a T-shape. This T-shape induces severely altered physical properties in comparison to its linear isomer. For example, fully saturated WEs derived from secondary alcohols and comprising 40 carbon atoms melt at 5-10 °C, while linear species of the same number of carbon atoms melt at 65-70 °C (Patel et al., 2001).

## 1.2 Industrial application of WEs

WEs are being used for industrial purposes and in daily applications for a long time. In the 19<sup>th</sup> and early 20<sup>th</sup> century, sperm whale-derived waxes were intensely used as lamp oil and lubricants. In the USA, the peak of spermaceti oil import was reached in 1837, when approximately five million gallons were imported (Tower, 1907). With an average density of  $0.884 \text{ kg} \times \text{m}^{-3}$  (Dieterichs, 1916), this would correspond to 16,000 tons. With the rise of mineral oil, the market declined, but experienced a renaissance after the Second World War, when spermaceti oil was rediscovered as an additive in high pressure industrial lubricants. The newly emerged whaling industry peaked in 1964, when almost 30,000 animals were killed per year (Whitehead, 2009). Since these practices nearly led to extinction of the sperm whale, a global ban on whaling was spoken in the 1980s (<http://iwc.int/catches>). In the field of natural-derived lubricants, Jojoba oil was found to be an appropriate substitute for spermaceti oil. Jojoba-derived oil and spermaceti oil have similar properties, but exhibit important differences as well. Both oils have comparable viscosity indices of 190-230 and 180, respectively (American Oil Chemists' Society, 1988; Forbes, 1943). These arbitrary values indicate how the viscosity of a lubricant changes with temperature. Values over 110 represent the highest classification of the ranking, making both oils high class lubricants. The flash point of Jojoba oil is 295 °C (Wisniak, 1987), and thus a little higher than the respective value of spermaceti oil, which is 260 °C (Pohanish, 2004). The melting temperature of Jojoba oil is around 3.8-7 °C (Bassam, 1997) and thus lower than the one of spermaceti oil, which is around 30 °C (Morris, 1973). Hence, Jojoba-based lubricants are applicable in a wider temperature range than respective products based on spermaceti oil. This might be especially important for applications, in which the lubricant is temporarily exposed to cold temperatures, exemplary in car engines. Another important difference among both oils is their viscosity, which is 1.27 cm<sup>2</sup>/s for Jojoba oil (Bassam, 1997) and thus five times higher than the viscosity of spermaceti oil, which is 0.22 cm<sup>2</sup>/s (<http://www.engineeringtoolbox.com>, 2014).

However, today natural-derived WEs account for only a minor proportion of the global WE market. In 2010, the global WE consumption comprised three million tons (Wei, 2012). With a share of 85 %, mineral WEs are the major products among the global wax demand, while synthetic WEs comprise for 11 %. In contrast, animal and vegetable WEs accounted for only 4 %, corresponding to approximately 120,000 tons (Kline & Company, Inc., 2010). WEs are processed for a wide range of applications in industry. Interestingly, approximately 50 % of the global annual WE production in 2010 was used in a rather traditional way: for the production of candles (Kline & Company, Inc., 2010; Wei, 2012). Another major application of waxes is packaging. According to Wei, 2012, waxes for packaging purposes accounted for 21 % of the global demand in 2010, whereas Kline & Company account packaging for only 10 % of the demand. Packaging includes coatings of paper drinking cups, milk cartons or waxed papers. Further purposes are very widespread and include utilization in crayons, cosmetics, inks, foods, tires or lubricants (American Fuel & Petrochemical Manufacturers, 2014). Two of the main advantages of natural wax are its renewability as well as its biodegradability. The latter is particularly relevant in the field of lubricants, since it's estimated that approximately 50 % of the global production of lubricants end up in the environment (Mang, 1998). An estimated 95 % of these lubricants are mineral oil-based (Bart et al., 2013), often toxic for animals and plants and thus entail great dangers for the environment.

Nowadays, it is assumed that 90 % of all common industrial lubricant substances can be replaced by counterparts of plant origin with equivalent quality (Mang, 1998). Since lubricants make up a huge

global market with an annual demand of approximately 40 million tons, utilization of plant derived WEs in this field has a huge economical potential. Using plant derived substances in high end applications like lubricants is furthermore economically more reasonable than selling them in the low cost fuel field. Chemically, a high proportion of the plant-based lubricants used today are triglycerides (TAG), i.e. ester of a glycerol backbone and three long chain carboxylic acids (Figure 6). On the one hand, TAG-based lubricants are characterised by high flash points and low volatilities (Dyer et al., 2008). On the other hand, they also exhibit negative properties like high viscosity at low temperatures, high susceptibility to oxidation and low stability at high temperatures (Wagner et al., 2001). Because of their linear structure and a single ester linkage instead of three, waxes exhibit superior properties in comparison to TAG. In particular, WEs possess better oxidation as well as better shear stability (Carlsson, 2006). WEs furthermore act anti-corrosive, anti-foaming and minimise wear (Bisht et al., 1993; El Kinawy, 2004). In summary, utilization of natural WEs in the field of lubricants is economically useful, sustainable and thus highly promising. Apart from that, substitution of fossil oil derived substances will become inevitable as soon as the remaining supplies are depleted.

### 1.3 Enzymatic basis of WE biosynthesis

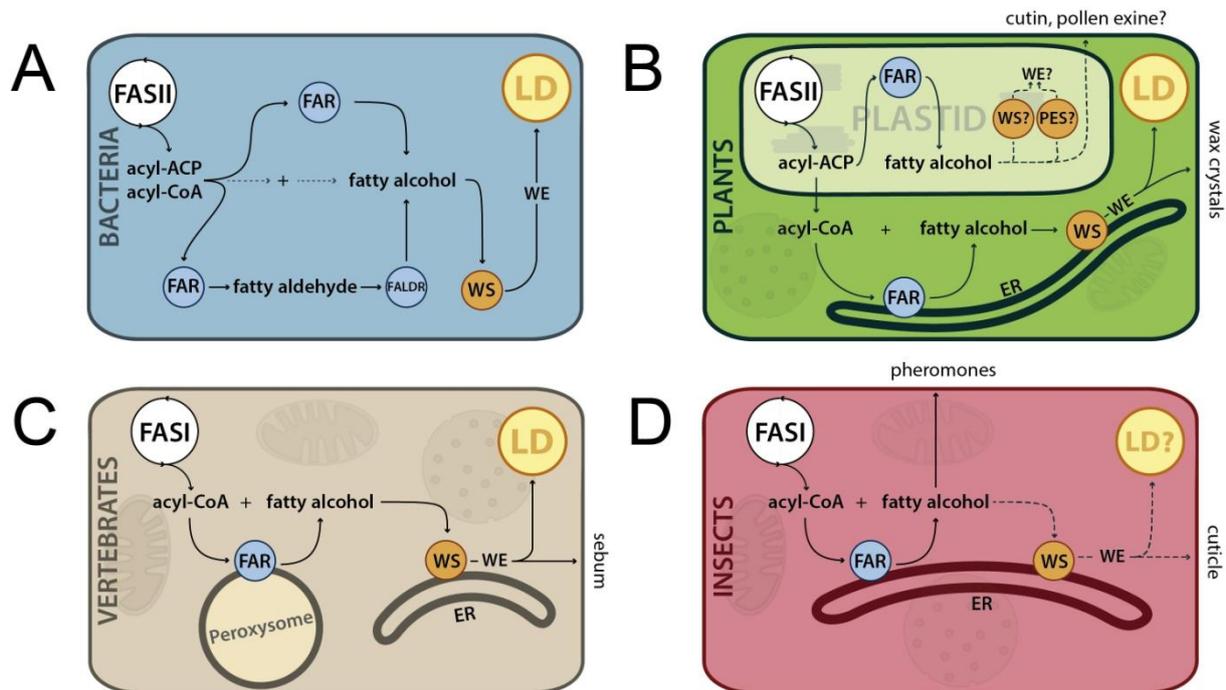
WEs are the product of a condensation reaction of a fatty alcohol and a long chain fatty acid. Long chain fatty acid biosynthesis is a process, which is highly conserved all over the phylogenetic tree with only few exceptions such as mycoplasmas (Rottem, 1980). Thus, the enzymatic steps of fatty acid biosynthesis in pro- and eukaryotes are highly similar. In both cases, a fatty acyl chain is built through rounds of successive condensation of acetyl units, each round being accompanied with two reduction and one dehydration reactions. In prokaryotes, plants and mitochondria, these reactions are catalysed by the fatty acid synthase (FAS) II (Figure 2 A&B). FAS II is a multimeric enzyme complex comprised of single enzymes for each reaction during fatty acid biosynthesis. In contrast, FAS I in mammals and insects is a single enzyme which catalyses all of these reactions in one chain and is located in the cytoplasm (White et al., 2005) (Figure 2 C&D). Since free fatty acids are toxic (Sikkema et al., 1995), the products of fatty acid biosynthesis are esterified to acyl-CoAs in the case of eukaryotes or ACPs in the case of prokaryotes, plastids and mitochondria. Acyl-ACP or acyl-CoAs represent one of the two substrates for WE biosynthesis (Röttig and Steinbüchel, 2013).

The second substrates are fatty alcohols. Unlike fatty acid biosynthesis, the ability to produce fatty alcohols is not as highly conserved in nature. In eukaryotes and most prokaryotes, fatty alcohols are synthesised by a single enzyme called fatty acyl reductase (FAR) (Figure 2/ Figure 3). FARs produce fatty alcohols by reducing fatty acyl-CoA or fatty acyl-ACP molecules. In contrast, fatty alcohol production is thought to be reliant on two separate enzymes in *Acinetobacter*. The first enzyme reduces acyl-CoA to the corresponding fatty aldehydes, while the second enzyme reduces these aldehydes to the corresponding fatty alcohols (Hofvander et al., 2011; Reiser and Somerville, 1997). To date, only the first enzyme producing fatty aldehydes has been identified, while the fatty aldehyde reductase (FALDR) has not been discovered yet (Hofvander et al., 2011). As FARs use acyl-ACP or acyl-CoA as the substrate, they are highly dependent on fatty acid biosynthesis and fatty acid editing.

Fatty alcohols and fatty acyl-CoAs or -ACPs can be esterified to each other by enzymes called acyltransferases. Acyltransferases which exclusively or mostly synthesise WEs are called wax synthases (WS). However, most WSs are also able to catalyse the transfer of a fatty acyl moiety to

## INTRODUCTION

acceptor molecules other than fatty alcohols, for example diacylglycerol (DAG) (Biester et al., 2012a; Cheng, 2004a; Röttig and Steinbüchel, 2013; Yen et al., 2005).

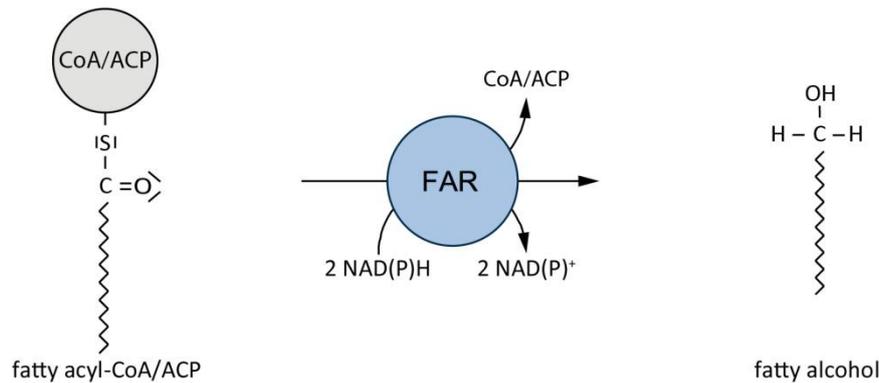


**Figure 2: WE synthesis in A) bacteria, B) plants, C) animals and D) insects.** Dashed lines indicate hypothetical routes for fatty alcohols. It is not fully understood what the fate of plastidial fatty alcohols is. Also, no WSs from insects are described to date. Therefore, the WE biosynthesis pathway in D) is indicated by a dashed line as well. However, insect-derived WEs were already described, making the existence of WSs in insects mandatory. ACP = acyl carrier protein, ER = endoplasmic reticulum, FALDR = fatty aldehyde reductase, FAR = fatty acyl reductase, FAS = fatty acid synthase, LD = lipid droplet, PES = phytol ester synthase, WE = wax ester, WS = wax synthase.

### 1.4 Fatty acyl reductases (FARs)

In the past, FAR was used as the abbreviation for “fatty acyl-CoA reductase”. However, since nowadays also enzymes are known which use acyl-ACPs as substrates (Doan et al., 2012; Hofvander et al., 2011), the general term “fatty acyl reductase” is more appropriate for this class of enzymes. FARs reduce either fatty acyl-CoA or fatty acyl-ACP molecules to the respective fatty alcohols with the use of two molecules of NADH or NADPH as cosubstrates (Riendeau and Meighen, 1985) (Figure 3). First reports on FAR-mediated fatty alcohol production were already published in the 1970s. At that time, respective studies were mainly carried out on the basis of the unicellular protist *Euglena gracilis* as well as on the already mentioned desert shrub *Simmondsia chinensis* or “Jojoba” (Khan and Kolattukudy, 1973; Pollard et al., 1979). It was shown that fatty alcohol synthesis is NADH-dependent in *E. gracilis*, whereas cell-free homogenates from *S. chinensis* were reliant on NADPH for fatty alcohol production. The corresponding enzyme as well as the cDNA from *S. chinensis* was purified and analysed for the first time in 2000 (Metz et al., 2000). Until today, a continuously growing number of FARs from bacteria, vertebrates, insects and plants are described in detail. Respective studies are mainly dedicated to elucidation of the substrate specificity of the enzymes. However, insights into structure-function relationships of FARs remain scarce.

## INTRODUCTION



**Figure 3: FARs catalyse the NAD(P)H-dependent reduction of either fatty acyl-CoA or fatty acyl-ACP to the corresponding fatty alcohol and free CoA or ACP, respectively.**

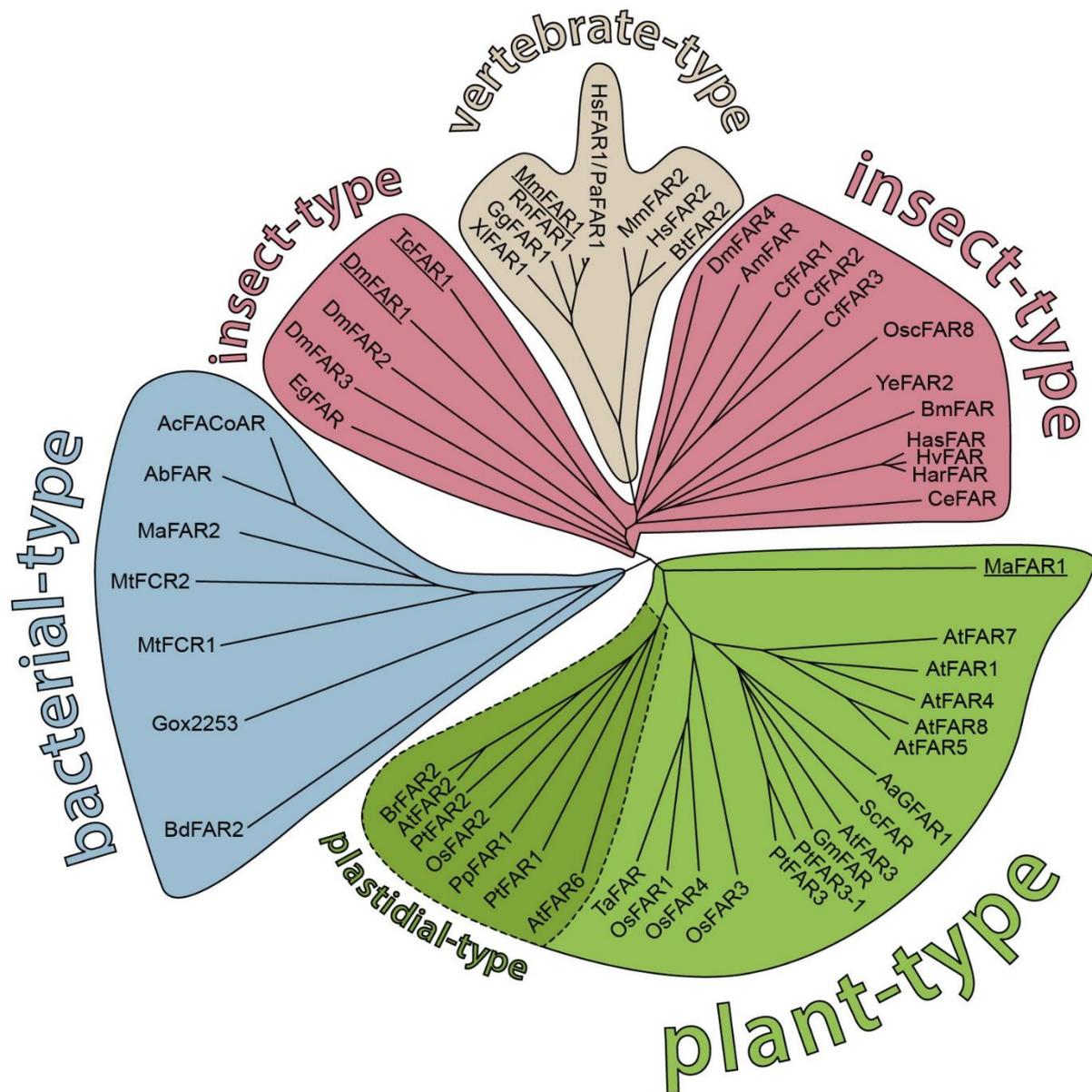
### 1.4.1 Phylogenetic distribution and subcellular localisation of FARs

Phylogenetic analyses of FAR amino acid sequences result in the formation of four main groups, which are predominantly comprised of either bacterial, plant, insect or vertebrate-derived enzymes. Interestingly, the insect-derived group consists of two subclades (Figure 4).

The model plant *Arabidopsis thaliana* possesses eight FAR sequences, which can be further divided into three localisation-based subgroups (Rowland and Domergue, 2012) (Figure 4). Enzymes of the first subgroup (AtFAR 1, 4, 5, 7, 8) localise to the endoplasmic reticulum (ER) (Rowland and Domergue, 2012), although most TM prediction services listed by the ARAMEMNON-service do not predict any transmembrane (TM) domains (Schwacke et al., 2003). The second group is plastid-localised (AtFAR2 & 6). As such, respective enzymes do not contain any predicted TM domains and are putatively soluble. The third group consists of only one enzyme, which is AtFAR3. It localises to the ER upon heterologous expression in yeast (Rowland et al., 2006). Most of the services listed by ARAMEMNON predict an N-terminal TM domain, which makes it different from AtFAR 1, 4, 5, 7 and 8. In contrast, the predictions for two C-terminal TM domains are similar to AtFAR 1, 4, 5, 7 and 8 (Rowland and Domergue, 2012; Schwacke et al., 2003).

The localisation of vertebrate-derived FAR enzymes within the cell has not been examined as intense as it has been for plants. After all, a peroxisomal membrane localisation was shown for murine FAR1 and FAR2 as well as for a FAR from guinea pig. In case of the murine enzymes, the peroxisomal localisation is mediated by two predicted C-terminal TM domains (Burdett et al., 1991; Cheng, 2004b; Heilmann et al., 2012). These domains can also be found in the two putative guinea pig FARs which are listed in the UniProt database (identifier H0V9X5 and H0UZ19).

INTRODUCTION



**Figure 4: Phylogram showing relationships among different types of fatty acyl CoA reductases.** UniProt IDs are listed behind respective enzyme abbreviations. FAR = fatty acyl reductase, Ab = *Acinetobacter baylyi*, Ac = *Acinetobacter calcoaceticus*, Am = *Apis mellifera*, At = *Arabidopsis thaliana*, Bm = *Bombyx mori*, Bt = *Bos taurus*, Ce = *Caenorhabditis elegans*, Cf = *Calanus finmarchicus*, Dm = *Drosophila melanogaster*, Gg = *Gallus gallus*, Har = *Helicoverpa armigera*, Has = *Helicoverpa assulta*, Hs = *Homo sapiens*, Hv = *Heliothis virescens*, Ma = *Marinobacter aqueolei*, Mm = *Mus musculus*, Mt = *Mycobacterium tuberculosis*, Osc = *Ostrinia scapulalis*, Pa = *Pongo abelii*, Rn = *Rattus norvegicus*, Sc = *Simmondsia chinensis*, Ta = *Triticum aestivum*, Tc = *Tribolium castaneum*, Xl = *Xenopus laevis*, Ye = *Yponomeuta evonymella*, AaGFAR = E9KL86, AbFAR = Q6F7B8, AcFACoAR = D0S4I2, AmFAR = D9MWM7, AtFAR1 = Q39152, AtFAR2 = Q08891, AtFAR3 = Q93ZB9, AtFAR4 = Q9LXN3, AtFAR5 = Q0WRB0, AtFAR6 = B9TSP7, AtFAR7 = Q9FMQ9, AtFAR8 = Q1PEI6, BdFAR2 = I1H9P9, BmFAR = Q7YTA9, BrFAR2 = A6MHV6, BtFAR2 = Q0P5J1, CeFAR = Q9TZL9, CffAR1 = G3KIJ8, CffAR2 = G3KIJ9, CffAR3 = G3KIK0, DmFAR1 = Q8MS59, DmFAR2 = Q960W6, DmFAR3 = A1ZAI3, DmFAR4 = A1ZAI5, EgFAR = D7PN08, GgFAR1 = Q5ZM72, GmFAR = I1M4E4, Gox2253 = Q5FNRO, HarFAR = I3PN86, HasFAR = I3PN85, HsFAR1 = Q8WVX9, HsFAR2 = Q96K12, HvFAR = D2SNU9, MaFAR1 = A1U2T0, MaFAR12 = A1U3L3, MmFAR1 = Q922J9, MmFAR2 = Q7TNT2, MtFCR1 = O50417, MtFCR2 = 6YAV6, OscFAR8 = B6SDC3, OsFAR1 = Q0IZI9, OsFAR2 = Q8S7T9, OsFAR3 = Q7XRZ6, OsFAR4 = Q6ZJ06, PaFAR1 = Q5R834, PpFAR1 = A9RVF6, PtFAR1 = B9IHM0, PtFAR2 = B9IID5, PtFAR3 = B9H1Z2, PtFAR3-1 = B9H1F3, RnFAR1 = Q66H50, ScFAR1 = Q9XGY7, TaFAR = Q8L4C3, TcFAR1 = D2A5A7, XIFAR1 = Q7ZXF5, YeFAR2 = D7P5E3. Sequence alignment and construction of phylogram was done with ClustalX using the default settings (Thompson, 1997). Underlined FARs were analysed in the course this work.

## INTRODUCTION

Unfortunately, the authors did not provide an identifier for the guinea pig FAR in their study (Burdett et al., 1991). However, it would be interesting to elucidate the role of the domains as a general peroxisomal targeting signal in FAR enzymes. At least one putative C-terminal TM domain is also present in avian enzymes, which consistently are found in the microsomal fraction upon expression in yeast (Hellenbrand et al., 2011). Other vertebrate-derived FARs were mostly found to be present in membrane fractions as well, underlining the membrane-associated nature of the proteins (Hellenbrand et al., 2011; Kolattukudy and Rogers, 1986; Riendeau and Meighen, 1985; Wang and Kolattukudy, 1995).

DmFAR1 from *Drosophila melanogaster* localised to the ER when expressed in *Drosophila* S2 cells (Jaspers et al., 2014). Other insect-derived enzymes have been found in the microsomal fraction upon expression in yeast, indicating a similar localisation (Moto et al., 2003; Teerawanichpan and Qiu, 2010). Finally, bacterial sequences do not contain any predicted TM domains and are thus most likely soluble enzymes. However, expression and purification of these enzymes in *Escherichia coli* was shown to be not as trivial as it could be expected for soluble enzymes, making it necessary to fuse the respective enzymes to solubility-promoting proteins like the maltose binding protein (MBP) (Hofvander et al., 2011; Wahlen et al., 2009; Willis et al., 2011).

### 1.4.2 Catalytic motifs in FARs

All FAR enzymes are members of the extended short chain dehydrogenase/reductase (SDR) enzyme family (Kallberg et al., 2010). This family is a vast group of NAD(P)H-dependent enzymes with highly diverse roles in the metabolism of lipids, carbohydrates, amino acids, hormones or within redox sensing (Kavanagh et al., 2008). In 2009, approximately 47,000 members of the SDR family were known. Most of them share little sequence identities of about 25 %, making it difficult to elucidate relationships and define subclades among them (Kallberg et al., 2010; Persson et al., 1991). In contrast, the respective enzyme structures are often highly similar.

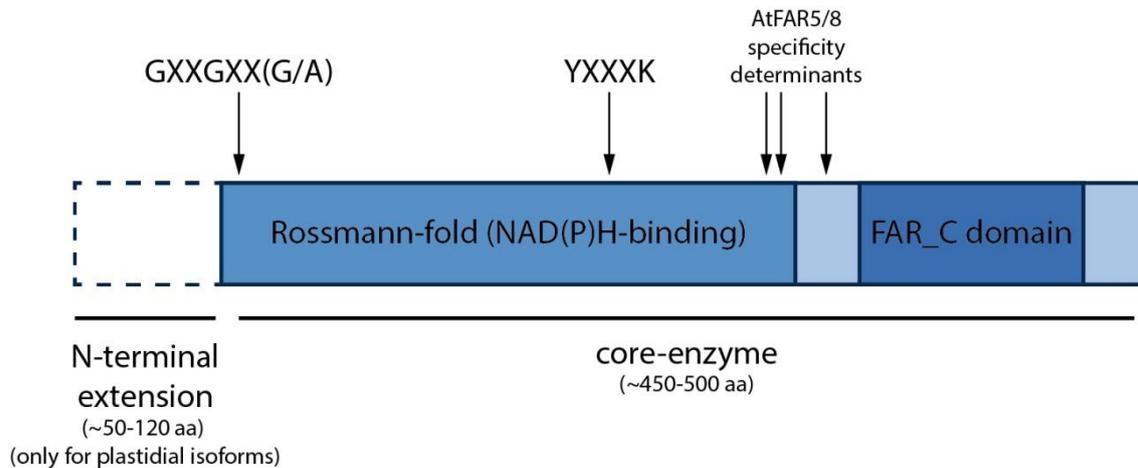
A conserved motif in all SDR and thus FAR enzymes is a Rossmann fold, which is encoded by the N-terminal part of the sequence (Kavanagh et al., 2008). Rossmann folds are responsible for dinucleotide cofactor-binding, i.e. for NADH or NADPH-binding, in SDR enzymes. The conserved folding pattern of Rossmann folds is built from alternating  $\beta$ -sheets and  $\alpha$ -helices. Together, they build a central array of six to seven parallel  $\beta$ -sheets, which is surrounded by three to four  $\alpha$ -helices on each side (Lesk, 1995) (Figure 13). NAD(P)H binding is mediated by a conserved, glycine-rich GXXGXXG motif in the loop between the first  $\beta$ -sheet and  $\alpha$ -helix of extended SDR (Brändén, 1980; Kavanagh et al., 2008; Lesk, 1995).

The active site of extended SDR enzymes is built by a conserved YXXXK motif within the Rossmann fold in the middle of the protein sequence (Figure 5). The indispensability of the tyrosine as well as of the lysine residue of this motif for enzyme activity was shown by site directed mutagenesis. Upon replacement of either the tyrosine by a phenylalanine or the lysine by an isoleucine in FAR5 from *A. thaliana*, the resulting variants did not show activity any more (Chacon 2013). Today, the generally accepted catalytic mechanism of SDR is based on a central role of the tyrosine residue acting as an acid/base catalyst, while the lysine residue is believed to be involved in modulation of the tyrosine hydroxyl's pKa as well as in NAD(P)H-binding (Kavanagh et al., 2008; Koumanov et al., 2003).

The N-terminal part of FAR sequences was used to be called "male sterility-domain", according to the effect caused by a knockout of the plastidial FAR2 of *A. thaliana* (Chen et al., 2011b). Disruption of the respective homologue in rice caused male sterility as well (Shi 2011). However, since male

## INTRODUCTION

sterility is not a general effect associated with disruption of FAR-encoding genes, the term “male sterility domain” is not used any longer. Instead, the FAR C-terminus was termed FAR\_C-domain in the meantime.



**Figure 5: FAR structural domains.** Approximate positions of the GXXGXX(G/A) NAD(P)H binding motif, the catalytic YXXXK-motif as well as the substrate specificity determinants identified in AtFAR5 and AtFAR8 from *A. thaliana* (Chacon et al., 2013) are indicated. The C-terminal FAR\_C domain was used to be called male sterility domain. Since not all FARs cause male sterility upon gene disruption, the respective domain was termed FAR\_C-domain in the meantime. Figure modified according to Rowland and Domergue (2012).

### 1.4.3 Substrate specificity of FARs

Substrate specificity determination of FARs is often conducted by expression of the respective enzyme in yeast, followed by analyses of the produced alcohols by GC or GC-MS measurements. However, the potential substrate range for FARs in *Saccharomyces cerevisiae* is limited. It does neither contain polyunsaturated fatty acids nor fatty acids longer than C26. Specificities of FARs for respective species thus stay concealed and cannot be elucidated without supplementation of the corresponding precursors. Furthermore, evaluation of the available reports on FARs from a variety of organisms must lead to the general conclusion that the alcohol blend composition, which is produced by a FAR, is often highly dependent on the underlying substrate pool composition. For instance, FARs from different moths were found to exhibit a clear preference for the production of 9Z-tetradecanol (9Z-14:1-OH) upon expression in yeast and feeding equal amounts of methyl-9Z-tetradecenoate, methyl-9Z-hexadecenoate and methyl-11Z-hexadecenoate. However, when the ratio of the fed methyl esters was not equal any more, the proteins in particular synthesised the fatty alcohol derived from the component, which was most abundant in the mixture (Hagström et al., 2012). Another example is AtFAR6 from *A. thaliana*. Upon expression in yeast, the enzyme produced almost exclusively hexadecanol (16:0-OH), whereas *in vitro* tests on the basis of the purified enzyme also showed a substantial activity for the synthesis of octadecanol (18:0-OH) (Doan et al., 2012). In their review about FARs in plants, Rowland et al. mention this substrate-pool dependency as well. Comparison of fatty alcohol blends produced by three different *Arabidopsis* FARs upon expression in either *E. coli* or *S. cerevisiae* illustrate a big discrepancy between the alcohol-blends of one enzyme expressed in different systems (Rowland and Domergue, 2012). Hence, enzymes which act upstream of the fatty alcohol synthesis play a key role for the FAR product composition *in vivo*, as they are involved in the composition of the FAR's substrate pool. Moreover, the results reflect a rather low

## INTRODUCTION

substrate specificity of many FARs known to date. *In vivo* substrate specificity determination carried out in heterologous systems other than the respective natural host of the enzyme should therefore be treated with caution.

Nevertheless, respective experiments might give a first hint on the basic substrate specificity of a FAR, for example in order to distinguish between rather broad or narrow specificities. For example, FAR5 and FAR8 from *Arabidopsis* predominantly produce 18:0-OH and 16:0-OH, respectively, upon expression in yeast, indicating a rather narrow and specific substrate range (Chacon et al., 2013). In contrast, HvFAR from *Heliothis virescens* was shown to produce a wide range of fatty alcohols from C8 to C16 upon expression in yeast, indicating less specificity (Hagström et al., 2012).

However, the only exact way to determine substrate specificities is to measure enzyme kinetics by the use of *in vitro* assays with the purified enzyme. While this was shown to be feasible for plastid-localised or prokaryotic FARs (Chen et al., 2011a; Doan et al., 2012; Hofvander et al., 2011; Willis et al., 2011), membrane-localised FARs have not been purified yet (Teerawanichpan et al., 2010).

Despite of the described difficulties, a first success in elucidation of substrate specificity determinants in FARs was achieved in 2013. *Arabidopsis* FAR5 and FAR8 are highly similar on the sequence level, sharing 85 % identity. Despite of that, both proteins exhibit distinct substrate specificities upon expression in yeast. FAR5 is highly specific for the production of 18:0-OH, while FAR8 produces 16:0-OH. In the course of a domain swap attempt between both enzymes followed by studies on single amino acid exchange variants, three residues with an influence on the FARs' substrate specificity were identified. By exchanging the respective residues of AtFAR5 for the corresponding residues of AtFAR8 and vice versa, the authors were able to completely exchange the substrate specificities of both enzymes for one another (Chacon et al., 2013). The crucial residues are located at the end of the Rossmann fold domain and in the intersection between Rossmann fold and FAR\_C domain, respectively (Figure 5). This is not surprising, since the C-terminal domain of SDRs is generally anticipated to be involved in substrate specificity of this class of enzymes (Kavanagh et al., 2008).

Most of the described FARs enzymes are strictly dependent on NADPH. However, there are also FARs which are reliant on NADH (Khan and Kolattukudy, 1973; Teerawanichpan and Qiu, 2010) or can use both (Chen et al., 2011a). Elucidation of the determinants for this cofactor specificity in FARs was not concerned until now. However, the Rossmann fold of FARs is structurally highly conserved and thus similar to Rossmann folds of other SDR enzymes. In the SDR aldehyde reductase Gox2253 from *Gluconobacter oxydans*, the strict NADPH dependency of the enzyme was recently converted to a bivalent cofactor specificity. By replacement of an arginine residue involved in the binding of NADPH by either a lysine or a tyrosine, the protein was able to efficiently use both, NADPH and NADH as reductants. The authors assume, that the molecular mechanism for the phenomenon was based on additional hydrogen bonds between the respective residues and the NADH, which could stabilise the NADH molecule upon binding to the protein (Yin et al., 2014).

#### 1.4.4 Biotechnological applications of FARs

To date, biotechnological applications of FARs are mostly connected with the heterologous production of WEs in different organisms. An outline concerning this topic can be found in section 1.5.5. In contrast, applications with the sole purpose of fatty alcohol production are scarcer, but have been reported before.

An established strategy for fatty alcohol production in *E. coli* relies on the expression of a thioesterase, an acyl-CoA ligase and a FAR (Liu et al., 2013; Steen et al., 2010; Zheng et al., 2012). Here, the underlying theory is to cleave acyl chains from fatty acid synthesis-derived acyl-ACP by the use of the thioesterase. The free fatty acids are then ligated to CoA via the acyl-CoA ligase and the resulting acyl-CoA molecules are reduced to the corresponding fatty alcohols by the FAR. It has been shown that the fatty alcohol blend can significantly be altered by expression of thioesterases with different acyl chain length specificities (Steen et al., 2010).

Apart from the thioesterase as a key component for the composition of the produced fatty alcohol blend, also the choice of the FAR was shown to contribute to both, amount and quality of the produced alcohols. The sole expression of the FAR from *S. chinensis* in *E. coli* was shown to result in mainly C18 alcohols, whereas a FAR from *Acinetobacter calcoaceticus* mainly produced C12 and C14 alcohols (Zheng et al., 2012). The choice of the expressed FAR was shown to also have influence on the quantity of the produced fatty alcohols. Steen et al. (2010) achieved a maximum alcohol titer of 60 mg/l by expressing a FAR from *A. calcoaceticus* together with a modified version of the *E. coli* thioesterase 'TesA, the acyl-CoA ligase FadD, and deletion of the acyl CoA dehydrogenase gene (*fadE*) in *E. coli*. In contrast, expression of MaFAR1 from *Marinobacter aquaeolei* (Hofvander et al., 2011; Wahlen et al., 2009) resulted in 31-fold higher fatty alcohol production in comparison to the FAR from *A. calcoaceticus* upon expression in *E. coli* (Liu et al., 2013).

The overall amount of *E. coli*-produced fatty alcohols was shown to also be increasable by optimisation of fermentation conditions. From an initial titer of 60 mg/l reported by Steen et al. (2010) upon expression of a FAR from *A. calcoaceticus*, a maximum yield of 598.6 mg/l was achieved with the same strain under optimised fermentation conditions. The highest volumetric fatty alcohol yield obtained by biotechnological means to date was achieved by Liu et al. (2013), who utilised an *E. coli fadE* knockout strain to express FadD and 'TesA as well as MaFAR1 from *M. aquaeolei*. Under fermentation conditions, fatty alcohol yields of 1.725 g/l have been achieved like this.

Growth of microorganisms like *E. coli* requires substances like carbon sources, amino acids, vitamins and likewise. Usually, these compounds are derived from plants and animals, making the production of chemicals in bacteria rather expensive. An economical alternative is the direct production of high value substances like fatty alcohols in photosynthetic organisms. A proof of concept regarding this idea was performed in cyanobacteria. Here, introduction of the FAR from *S. chinensis* yielded maximum fatty alcohol amounts of approximately 200 mg/l when the respective cultures were grown in a photoreactor for 18 days (Tan et al., 2011). This value is not as high as yields achieved in *E. coli*. However, it was achieved by expression of a single enzyme. Optimisation through expression of thioesterases or likewise could thus further increase the production of fatty alcohols in photosynthetic organisms.

Apart from fatty alcohol productions intended to serve as chemical feedstock, a more sophisticated application of photosynthetically produced fatty alcohol was presented recently. Insects, predominantly moths, use a blend of unsaturated fatty acid derivatives like fatty alcohols, aldehydes and acetylated fatty alcohols as sex pheromones. Ding et al. (2014) used a modified version of the pathway shown to be suitable for fatty alcohol production in *E. coli* in order to be able to produce

## INTRODUCTION

sex-pheromones *in planta*. A set of genes was transiently expressed in *Nicotiana benthamiana*, comprising different thioesterases, desaturases, FARs and an acetyltransferase. The thioesterases were shown to be able to produce C14 substrates, which are rather limited in the wild type plant. The introduced desaturases and FARs further created the desired desaturated fatty alcohols, which partly were further acetylated by the acetyltransferase (Ding et al., 2014). Although the authors finally extracted the fatty alcohols produced in the plant and acetylated them chemically, this approach is an interesting demonstration of how manifold fatty alcohol production in plants can be utilised.

### 1.4.5 Role of free fatty alcohols in nature

In nature, fatty alcohols do either occur esterified in the form of WEs (see 1.5.6) or as free substances. The free form of fatty alcohols is mainly known to function as signalling molecules or as components in cuticular wax layers of plants or insects. Signalling effects of fatty alcohols are predominantly described in moths (Antony et al., 2009; Ding et al., 2014; Hagström et al., 2012; Lienard et al., 2010; Moto et al., 2003). Here, they are produced in special glands of the animals and can be further oxidised or acetylated, yielding a blend of unsaturated fatty alcohols, fatty aldehydes and acetate esters (Roelofs and Wolf, 1988). The composition of this blend is unique for different species, assuring distinguishable mating specificities (Moto et al., 2003). Interestingly, there are several substances which are present in the pheromone blend of many species. As a consequence, production of rather few substances in genetically modified plants could serve as an attractant in biological traps for pest control (Ding et al., 2014).

In plants, free fatty alcohols occur as components of the cuticular wax layer, which mostly consists of fully saturated WEs, very long chain (VLC) alcohols, aldehydes, ketones and alkanes. The wax layer is present on environment-exposed parts of the plant and protects it from stresses like UV radiation, pathogen attack or extended water loss (Rowland and Domergue, 2012). In contrast to long chain free fatty alcohols acting as pheromones in insects, cuticular free fatty alcohols are derived from VLC fatty acids, ranging from 24 to 32 carbon atoms. Respective substances can constitute to the wax layer to an enormous extend. For instance, free 32:0-OH accounts for 63 % of the weight of seedling leaves of maize (Bianchi et al., 1984). The important role of fatty alcohols upon cuticular wax formation is illustrated by the *Arabidopsis cer-4* knockout phenotype. The corresponding gene encodes a FAR which is responsible for the formation of C24-C30 fatty alcohols (Rowland et al., 2006). Deletion of the enzyme's activity yields in an almost complete lack of C24-C28 fatty alcohols in the cuticular waxes of the plant's stem. As a consequence, the stems do not look whitish any longer, but have a glossy appearance. Notably, measurements of the total amount of the wax load on the stems did not reveal significant differences between the knockout mutants and the wild type. The reason for the glossy appearance was shown to be a thick film of wax instead of wax crystals on the stems' surface. The glossy appearance is thus caused by a differential reflection of light by the smooth wax film in comparison to the flaky wax crystals (Rowland et al., 2006). Interestingly, a knockout of the WE forming WSD1 in *Arabidopsis* did not yield a glossy phenotype (Li et al., 2008). Hence, not the absence of the esterified, but the free form of fatty alcohols seems to be crucial for wax crystal formation in that case.

Plastid-derived fatty alcohols in plants are involved in the synthesis of the anther cuticle and sporopollenin for pollen. The corresponding plastidial FARs from rice and *Arabidopsis* were characterized and have been found to mainly synthesise C16 primary alcohols (Chen et al., 2011a; Shi

## INTRODUCTION

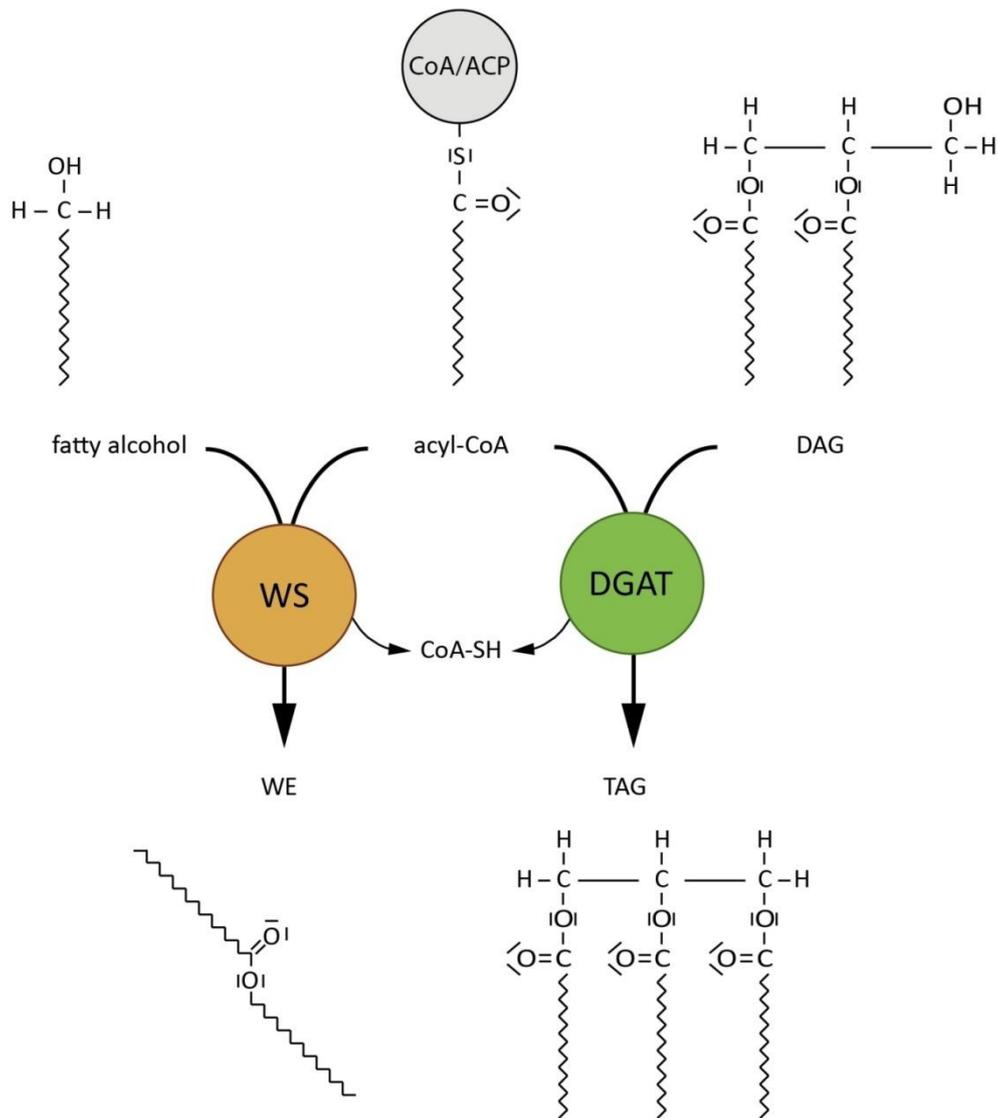
et al., 2011). Consistently, analyses of free fatty alcohols and free fatty acids in anthers of respective rice knockout mutants revealed decreased levels of C16 fatty alcohols and increased levels of total free fatty acids (Shi et al., 2011). Furthermore, pollen grains of the knockout mutants showed a thin pollen wall with a reduced exine layer, leading to male sterility (Aarts et al., 1997; Chen et al., 2011a). Apart from that, the amount of cutin in the anther cuticle of the mutants was significantly reduced. Analyses of the cutin monomers derived from mutant anther cutin polyester, e.g. hydroxyacids, epoxyacids and diacidmonomers showed decreased values throughout. These observations led the authors to the assumption, that plastid-derived fatty alcohols might be exported to the cytosol in order to be processed to cutin and sporopollenin precursors by ER-bound enzymes. Absence of these precursors might result in the observed phenotypes (Shi et al., 2011).

Hence, the fate of fatty alcohols synthesised by plastidial FARs is not fully deciphered yet. In theory, another possible sink for plastid-derived fatty alcohols might be the formation of WEs by plastidial phytyl ester synthases or plastidial WSs (Figure 2). The plastidial phytyl esters synthases AtPES1 and AtPES2 have recently been described in *Arabidopsis*. It has been shown, that the enzymes predominantly synthesise phytyl esters (Lippold et al., 2012). Upon plastidial coexpression with a FAR in *N. benthamiana*, AtPES2 was furthermore able to efficiently synthesise WEs (Aslan et al., 2014). However, to date there is no study about intrinsic, PES-mediated synthesis of WEs in plastids. Likewise, no plastidial WS is known, making a plastidial synthesis of WEs by WSs or PES a speculative sink for plastidial fatty alcohols.

### 1.5 Wax synthases (WSs)

FAR-derived fatty alcohols can be esterified to acyl chains by WSs, yielding WEs. All WSs are members of the enzyme family of acyltransferases. As such, they catalyse the esterification of an activated acyl chain with an acyl acceptor. In the last decade, knowledge about WSs has increased significantly. Accompanied with the discovery and analyses of numerous enzymes, it became clear that the vast majority of WSs are not only able to catalyse esterification of acyl-CoAs with fatty alcohols, but exhibit additional acyltransferase activities. Numerous described WSs also exhibit diacylglycerol O-acyltransferase (DGAT) activity and can catalyse esterification of acyl-CoAs to DAG, yielding TAG (Cheng, 2004a; Röttig and Steinbüchel, 2013; Stöveken et al., 2005). Simultaneously, several enzymes described to be a DGAT are capable of WE synthesis as well (Figure 6) (Cheng, 2004a; Du et al., 2014).

## INTRODUCTION

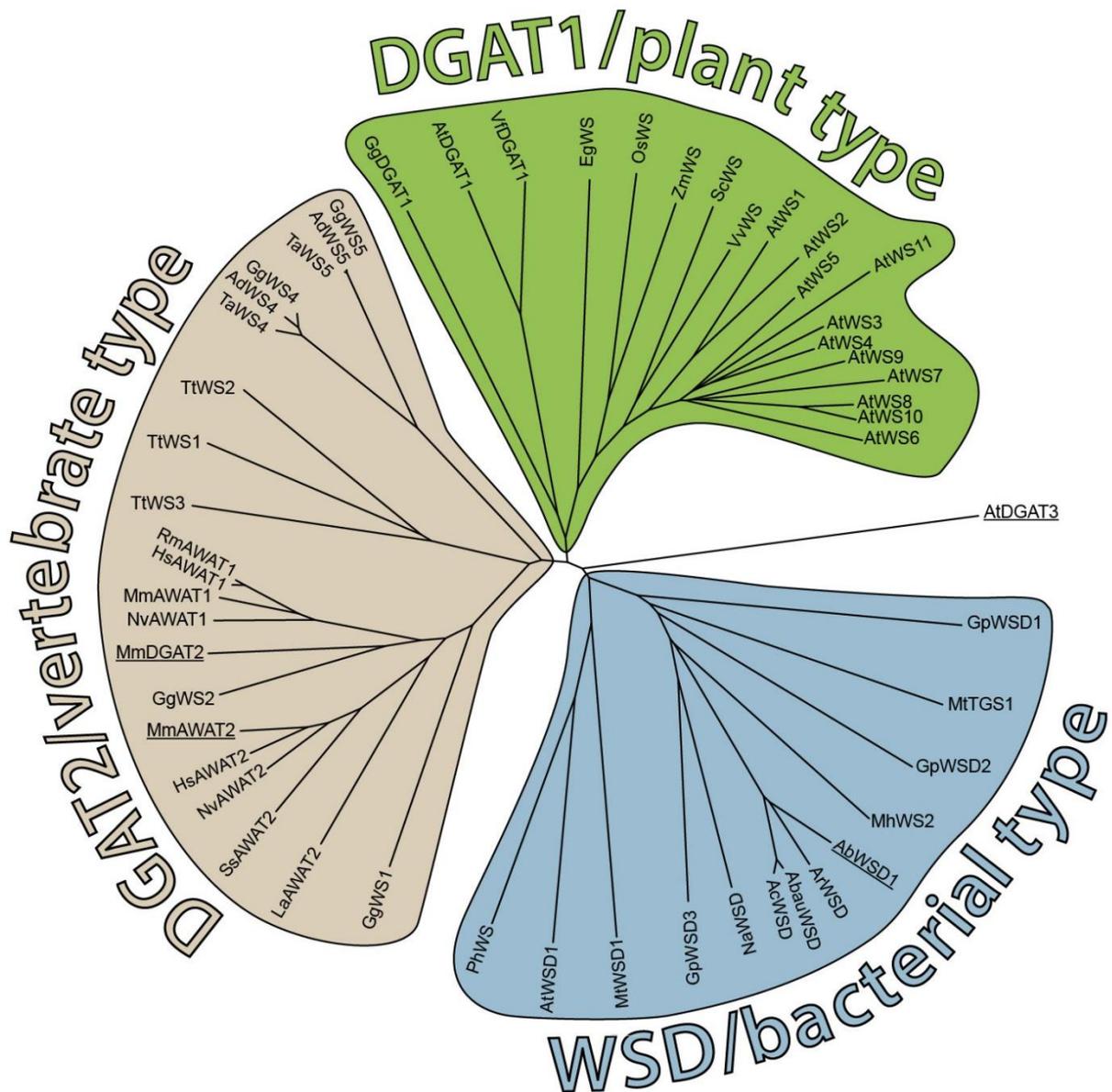


**Figure 6: Comparison of two related enzyme activities.** WSs primarily catalyse the condensation of a fatty alcohol with an acyl-CoA, thereby forming WEs and liberating a molecule of CoA-SH. In contrast to that, diacylglycerol O-acyltransferases (DGAT) primarily catalyse the condensation of an acyl-CoA molecule with diacylglycerol (DAG) to yield triacylglycerol (TAG), liberating one molecule of free CoA-SH as well. However, numerous WSs were also shown to exhibit DGAT-activity and vice versa.

### 1.5.1 Phylogenetic distribution of WSs

The ability to synthesise WEs is widely distributed over the phylogenetic tree, including the kingdoms of bacteria, plants and animals (Cheng, 2004a; Metz et al., 2000; Röttig and Steinbüchel, 2013). Phylogenetic analyses of WS sequences yield in the formation of three major clades, which harbour sequences mainly derived from either vertebrates, plants or bacteria (Figure 7). Vertebrate-type WSs cluster with DGAT2, while plant-like WSs cluster with DGAT1 sequences. In contrast, bacterial-type WSs do not cluster with a reference group. Since respective enzymes were first discovered in bacteria, bacterial sequences are eponymous for this type of WSs. However, nowadays also plant-derived sequences are known which cluster to this group (King et al., 2007).

INTRODUCTION



**Figure 7: Phylogram showing relationships among different types of Ws.** UniProt IDs are listed behind respective enzyme abbreviations. AWAT = acyl-CoA wax alcohol acyltransferase, WS = wax synthase, WSD = bifunctional wax synthase/diacylglycerol O-acyltransferase, Ab = *Acinetobacter baylyi*, Abau = *Acinetobacter baumannii*, Ac = *Acinetobacter calcoaceticus*, Ad = *Anser anser domesticus*, Ar = *Acinetobacter radioresistens*, At = *Arabidopsis thaliana*, Eg = *Euglena gracilis*, Gg = *Gallus gallus*, Gp = *Gordonia polyisoprenivorans*, Hs = *Homo sapiens*, La = *Loxodonta africana*, Mh = *Marinobacter hydrocarbonoclasticus*, Mm = *Mus musculus*, Mt = *Medicago truncatula*, MtTGS1 = *Mycobacterium tuberculosis*, MtWSD1 = *Medicago truncatula*, Na = *Nocardia asteroides*, Nv = *Neovison vison*, Os = *Oryza sativa*, Ph = *Petunia hybrida*, Rm = *Macaca mulatta*, Sc = *Simmondsia chinensis*, Ss = *Sus scrofa*, Ta = *Tyoto alba*, Tt = *Tetrahymena thermophila*, Vv = *Vitis vinifera*, Zm = *Zea mays*, AbauWSD = D0CDL4, AbWSD1 = Q8GGG1, AcWSD = N8N9S3, AdWS5 = H6W8E9, ArWSD = K6VXX4, AtDGAT1 = Q9SLD2, AtDGAT3 = Q9C5W0, AtWS1 = Q9FJ72, AtWS2 = Q9FJ73, AtWS3 = Q9FJ74, AtWS4 = Q9FJ75, AtWS5 = Q9FJ76, AtWS6 = 9FJ77, AtWS7 = Q9FJ78, AtWS8 = Q9LNL1, AtWS9 = Q4PT07, AtWS10 = Q3ED15, AtWS11 = , AtWSD1 = Q93ZR6, EgWS = D7PN09, GgDGAT1= E1BTG6, GgWS1 = UPI0002C86613, GgWS2 = H6W8E6, GgWS4 = H6W8E5, GgWS4 = UPI0000447B65, GgWS5 = Q5ZJD8, GpWSD1 = 6MTQ1, GpWSD2 = H6MS36, GpWSD3 = H6MYJ4, GpWSD3 = H6MYJ4, HsAWAT1 = Q58HT5, HsAWAT2 = Q6E213, LaAWAT2 = 3T8K5, MhWS2 = A3RE51, MmAWAT1 = A2ADU9, MmAWAT2 = Q6E1M8, MmDGAT2 = Q9DCV3, MtTGS1 = P9WK9, MtWSD1 = G7JTU6, NaWSD = 5E762, NvAWAT1 = U6DU75, NvAWAT2 = U6CZ66, OsWS = Q6K7A7, PhWS = A3QME3, RmAWAT1 = F6SLT8, ScWS = 9XGY6, SsAWAT2 = K7GQC2, TaWS4 = H6W8E7, TaWS5 = H6W8E8, TtWS1 = I7MN05, TtWS2 = Q24DK3, TtWS3 = Q22SB3, TtWS4 = 23ME5, VvDGAT1 = Q0QJ11, VvWS = Q84XY9, ZmWS = K7TU84. Sequence alignment and construction of phylogram was done with ClustalX using default settings (Thompson, 1997). Underlined enzymes were analysed in the course of this work. AtDGAT1 and VvDGAT1 were included in the phylogram to illustrate the homology to DGAT1-type Ws, but are not described to possess WS-activity on their own.

## INTRODUCTION

Insect-derived WSs have not been described to date. In contrast, several studies report insect-derived WEs (Blomquist et al., 1972; Buckner et al., 2009; Rottler et al., 2013). Thus, the existence of WSs in insects is more than likely. However, since no direct experimental evidence for respective enzymes is available, the WE biosynthesis pathway in Figure 2 D is indicated by a dashed line. Consequently, also the phylogram in Figure 7 does not contain insect-derived sequences.

AtDGAT3 is reported to belong to a class of soluble acyltransferases (Hernandez et al., 2012). While DGAT3-mediated WE formation was shown for an enzyme from peanut (Saha et al., 2006) upon expression in *S. cerevisiae*, AtDGAT3 was not tested for WS activity yet. As this type of enzymes does not share significant sequence similarity with other acyltransferases, its position in the acyltransferase-based phylogram is rather remote (Figure 7).

### 1.5.2 Protein structure and subcellular localisation of WSs

According to the three distinct clades WSs form upon phylogenetic analyses, they also differ in their predicted membrane topology as well as their subcellular localisation.

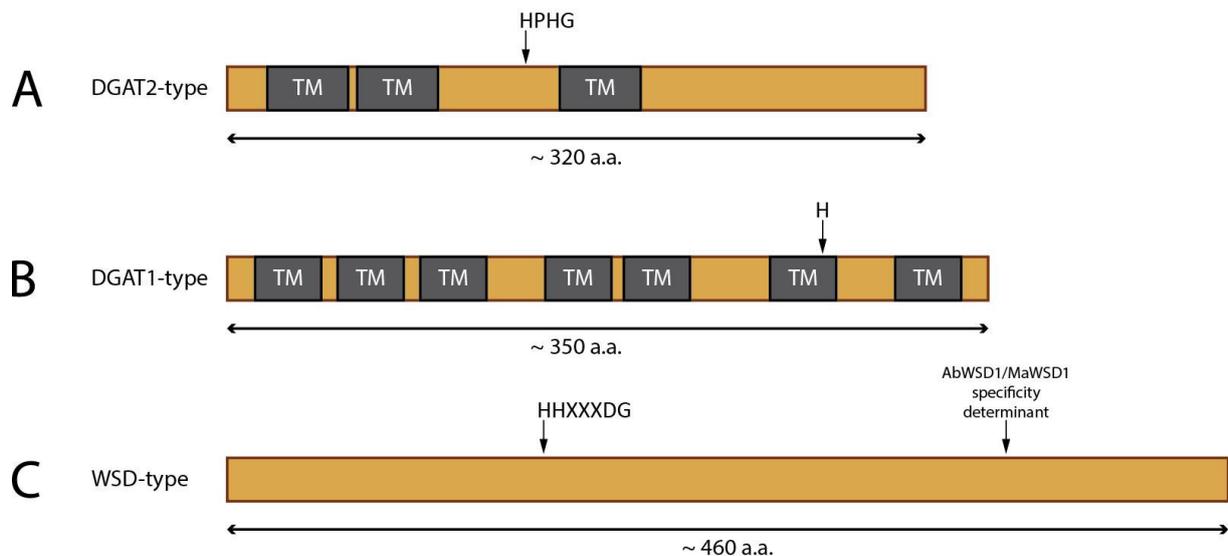
Bacterial WS sequences comprise of 450 to 500 amino acids on the average and generally do not contain any predicted TM domains. Despite of the fact that this holds also true for AbWSD1 from *A. baylyi*, the majority of its activity was reported to be present in the lipid inclusion fraction of its natural host. While considerable activity was also found to be associated with the total membrane fraction of the cell, only low activity was found in the cytosolic fraction (Stöveken et al., 2005). Thus, although the enzyme does not contain any obvious TM domains, it nevertheless seems to be highly membrane associated through currently unknown structures or mechanisms. Interestingly, the two bacterial type WSs from the plants *Petunia hybrida* (PhWS) and *A. thaliana* (AtWSD1) both contain a predicted TM domain in the middle of their sequence. In case of AtWSD1, a localisation of the enzyme at the ER was shown upon expression of a respective YFP fusion in *A. thaliana* (Li et al., 2008).

Sequences of DGAT2-like WSs comprise about 320 residues on the average. TM predictions of DGAT2 as well as DGAT2-like WSs indicate the presence of one to three TM domains within the sequence (Figure 8 A). Either one or two of these are predicted to be located near the N-terminus and are separated by a short stretch of approximately 4-5 amino acids (Figure 39). The third one is located in the middle of the sequence. However, in some sequences, this domain is predicted to only form a hydrophobic stretch rather than a TM domain. Experimental determination of the membrane topology of a DGAT2-like WS has not been done yet. However, in case of the murine DGAT2, respective experiments verified the presence of the two N-terminal TM domains. Apart from this, no other TM domains were identified within the sequence (McFie et al., 2014; Stone et al., 2006). The N-terminal TM domains in murine DGAT2 were shown to act as an ER targeting signal (McFie et al., 2011). Deletion of the respective part of the sequence resulted in a loss of the ER localisation of the enzyme upon expression in mammalian cells. Instead, the respective variant showed a more punctuated distribution and co-localised with the mitochondrial marker HSP-70 to mitochondria associated membranes (MAM), a special part of the ER. A small portion of the variant was moreover detected in the cytosolic fraction. However, this was also the case for the wild type version of MmDGAT2. Also, the TM domain deficient variant was still able to localise to lipid bodies upon oleate loading of the cells. Thus, the TM domains seem to determine the ER localisation of the enzyme, but are not the only structure for membrane binding of MmDGAT2 (McFie et al., 2011). While the mitochondrial localisation signal was found to be encoded by a conserved, positively charged RXKXXK

## INTRODUCTION

motif directly N-terminal to the TM regions (Stone et al., 2009), the part of the sequence responsible for lipid droplet-localisation was reported to be encoded in the part C-terminal of the enzyme, but was not specified in more detail yet (McFie et al., 2014). The N-terminal part of DGAT2-type Ws in front of the first predicted TM domain is usually shorter than in DGAT2 sequences, mostly consisting of about 10 to 20 residues. In contrast, the respective part in DGAT2 sequences is about 65 residues in length.

To date, knowledge about plant-type Ws is comparably limited. Respective sequences consist of approximately 350 residues, hence their length is similar to DGAT2 type Ws. The only members of this family, which were characterised in more detail are the Ws from *S. chinensis* (ScWS) and *E. gracilis* (EgWS) (Lardizabal et al., 2000; Teerawanichpan and Qiu, 2010). Plant type Ws cluster with sequences of DGAT1 and show a comparable predicted membrane topology. ScWS and EgWS are predicted to harbour seven TM domains. In case of the ScWS, TM domain predictions indicate the presence of three TM regions at the direct N-terminus of the sequence, while two are located in the middle and two are encoded by the C-terminus of the enzyme (Figure 8 B). Analyses of other plant type Ws result in the presence of generally six to eight TM domains, which are similar distributed as in ScWS. In detail, they cluster at the N-terminus, the C-terminus and the middle part of the sequence. One exception is AtWS11 (UniProt ID = Q9FGN1), which contains eleven predicted TM domains. In comparison, DGAT1 are predicted to contain eight to ten TM domains (Liu et al., 2012) with a comparable distribution to plant type WS: One cluster of TM domains at each the N- and C-terminus as well as a third cluster in the middle of the sequence.



**Figure 8: Different types of Ws.** **A)** DGAT2 or vertebrate-type Ws are ER-localised and contain 2–3 predicted transmembrane domains. The anticipated active site of the enzymes is a highly conserved HPHG motif between TM domain two and three **B)** DGAT1-type Ws are ER-localised as well and are mostly found in plants. Respective enzymes contain 6-8 predicted transmembrane domains. A possible active site residue is indicated in analogy to the anticipated active site histidine in DGAT1. **C)** Bifunctional wax synthase/diacylglycerol O-acyltransferase (WSD)-type enzymes are bifunctional Ws which also show DGAT activity. They are predicted to be soluble and are mostly found in bacteria. The approximate positions of the highly conserved HHXXXDG potential active site motif as well as a residue reported to be involved in substrate determination of AbWSD1 and MaWSD1 (Barney et al., 2012b) are indicated.

Despite of the high number of predicted TM domains, experimental elucidation of the membrane topology of the murine DGAT1 indicated the presence of only three TM domains. According to the model derived from these data, the majority of the enzyme including the C-terminus would be located on the luminal side of the ER, with only the N-terminus and a short cytosolic stretch between

## INTRODUCTION

TM domain two and three being exposed to the cytosol (McFie et al., 2010). In contrast, both N- and C-terminus of DGAT1 from *Vernicia fordii* (tung tree) were found to be located on the cytosolic site of the ER, indicating an even number of TM domains (Shockey, 2006). It would thus be interesting to map the TM topology of a plant-type WS in order to evaluate possible similarities or differences in comparison to DGAT1.

### 1.5.3 Catalytic motifs within WSs

Although the catalytic motif of each class of WSs is different, the crucial feature in all of them is anticipated to be a catalytic histidine residue. Multiple sequence alignments of bacterial bifunctional wax synthase/diacylglycerol O-acyltransferases (WSDs) revealed a highly conserved HHXXXDG motif, which was assumed to be the catalytic motif due to its conserved nature also in other acyltransferases (Kalscheuer, 2002; Wältermann et al., 2007). The motif is thus not exclusive to WS/DGAT enzymes, but also known from nonribosomal peptide synthases, chloramphenicol acetyltransferases or polyketide-associated proteins (Buglino et al., 2004; De Crécy-Lagard et al., 1995; Kleanthous et al., 1985; Marahiel et al., 1997). The second histidine within this motif was supposed to play a role in the catalytic mechanism of chloramphenicol acyltransferases already quite early (Shaw, 1983), and results supporting this hypothesis were published shortly thereafter.

Chloramphenicol acyltransferases catalyse the acetylation (and thereby the inactivation) of chloramphenicol, yielding acetylchloramphenicol. 3-(bromoacetyl)chloramphenicol can serve as a product analogue of acetylchloramphenicol. Upon incubation with chloramphenicol acyltransferase, it was found to covalently bind exclusively to the second histidine of the HHXXXDG motif, resulting in an inhibition of the enzyme (Kleanthous et al., 1985). In nonribosomal peptide synthases as well as in the Pap5 polyketide synthase from *M. tuberculosis*, the essential role of the second histidine was shown by studies on single amino acid exchange variants. Upon replacement of the histidine by valine or alanine, respectively, both enzymes were severely impaired in their activity (Onwueme et al., 2004; Stachelhaus et al., 1998). In contrast, replacement of the first histidine did only result in a moderate decreased enzyme activity of Pap5 (Onwueme et al., 2004).

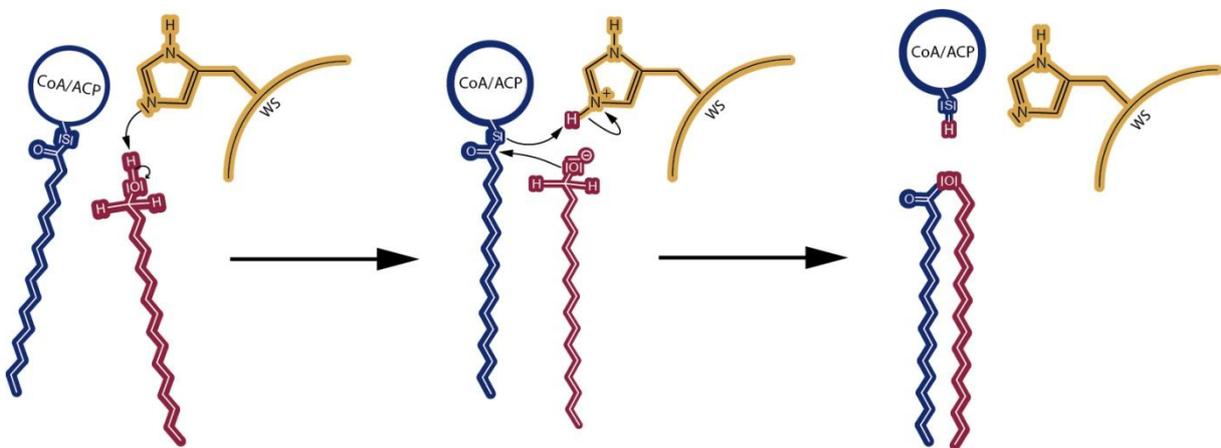
A structural role of the aspartate residue within the HHXXXDG motif was discovered in the crystal structure of the type III chloramphenicol acyltransferase from *E. coli*. Together with a conserved arginine near the N-terminus of the sequence, it was found to stabilise the catalytic histidine in an unusual confirmation, allowing it to interact with the chloramphenicol molecule and thereby abstract a proton of its primary hydroxyl group (Leslie et al., 1988). Consistently, replacement of the aspartate residue of the HHXXXDG motif resulted in a highly decreased activity of the corresponding variants of a nonribosomal peptide synthase from *Bacillus brevis* as well as in the Pap5 polyketide synthase from *M. tuberculosis* (Bergendahl et al., 2002; Buglino et al., 2004). Analogous to the findings in the chloramphenicol acyltransferase, the aspartate in nonribosomal peptide synthases and polyketide synthases is thought to play a structural role as well, rather than to be involved in the catalytic mechanism (Bergendahl et al., 2002; Buglino et al., 2004).

In bacterial WSD, the importance of the conserved HHXXXDG motif in mediating catalysis was also illustrated by the use of single amino acid exchange variants. However, not only exchange of the second histidine, but also exchange of the first histidine in WSD from *A. baylyi* and *M. hydrocarbonoclasticus* VT8 caused severely decreased enzymatic activities. Notably, exchange of the second histidine had a more pronounced effect than exchange of the first (Stöveken et al., 2009; Villa et al., 2013). Thus, it remains elusive if both histidines can act as a catalytic base, or if one of

## INTRODUCTION

them is more important for the structural integrity of the catalytic motif or the enzyme in general. The role of the conserved aspartate within the HHXXXDG motif seems to vary among different WSD. While exchange of the respective residue resulted in unaltered enzymatic activity in the WSD from *A. baylyi* (Stöveken et al., 2009), mutation in a WSD from *M. hydrocarbonoclasticus* resulted in a decrease of the enzymatic activity of approximately 90 % (Villa et al., 2013).

The generally accepted model for the catalytic mechanism of WSD-type Ws starts with the abstraction of a proton from the hydroxyl group of the fatty alcohol by the imidazole ring of the catalytic histidine residue (Figure 9). The resulting alcoholate anion then acts as a nucleophile and starts a nucleophilic attack towards the thioester bond of the acyl-CoA molecule. The thioester bond is thereby cleaved. A new oxoester bond is then formed between the acyl-CoA acyl chain and the alcoholate anion, resulting in formation of a wax ester. Finally, the free CoA gets protonated by the histidine residue, which is thereby regenerated to its initial state. The products of the reaction are a wax ester and a free CoA (Figure 9).



**Figure 9: Proposed catalytic mechanism of bacterial-type Ws.** A catalytic histidine residue of the HHXXXDG motif acts as a base and abstracts a proton from the hydroxyl group of the fatty alcohol. The resulting alcoholate anion acts as a nucleophile and attacks the thioester bond of the acyl-CoA molecule, resulting in a cleavage of the bond and the formation of a new oxoester bond between the acyl chain and the fatty alcohol anion. The cleaved CoA molecule is protonated by the catalytic histidine residue of the protein, thereby restoring the initial situation of the enzyme. Figure was modified according to Stöveken et al. (2009).

In DGAT2-type Ws, a highly conserved HPHG motif was identified to be crucial for the enzyme activity. While exchange of the first histidine within this motif in yeast DGAT2 almost completely abolished the enzymatic activity, the same exchange caused a 50 % decrease in enzyme activity in case of the murine DGAT2. In case of the second histidine, only 10 % residual activity was found for the respective variant of murine DGAT2, while the yeast enzyme completely lost its activity (Liu et al., 2011; Stone et al., 2006). Besides that, also exchange of the conserved proline within the HPHG motif of murine DGAT2 resulted in a 70 % decrease of enzymatic activity (Stone et al., 2006). Taken together, both histidine residues of the HPHG motif show a high influence on the catalytic pattern of the enzyme. Nevertheless, it remains elusive if they act as a catalytic base in a similar way like it's proposed for the respective residues in WSD.

Due to the low number of studies on plant-type Ws, no findings about residues which are critical for the catalytic process were published yet. Considering the phylogenetic relationship to sequences of DGAT1, one may speculate that both classes of enzymes might share a similar active site motif. In case of the murine DGAT1 it was shown that exchange of a highly conserved histidine within the C-

terminal part of the enzyme towards alanine abolished enzymatic activity (McFie et al., 2010). This histidine and the likewise conserved glutamate at the next position appear to be conserved in plant-type Ws as well and are located in the C-terminal part of the sequence (Figure 8 B). It would thus be interesting to see if a comparable mutation in plant Ws would have an impact on the enzymatic activity as well. Apart from this, no substrate specificity determining factors or other important residues within plant-type Ws are known to date.

#### 1.5.4 Substrate specificities of Ws

##### 1.5.4.1 Substrate specificities of WSD-type Ws

Bacterial-type Ws exhibit the broadest substrate range among the three classes of Ws and are thus rather unspecific. The ability to esterify an impressively wide range of acyl acceptors and acyl donors was shown on the example of AbWSD1 from *A. baylyi*. The enzyme can use linear alcohols ranging from chain lengths of 2 to 30 carbon atoms as well as branched, cyclic and aromatic alcohols. Moreover it can acylate monoacylglycerides (MAG) yielding DAG and diacylglycerides yielding TAG (Röttig and Steinbüchel, 2013). Apart from that, AbWSD1 was also shown to accept diols, thiols and dithiols to produce wax diester, wax thioester and wax dithioester, respectively (Kalscheuer et al., 2003; Uthoff et al., 2005). In contrast, the enzyme is unable to utilise polar substrates such as sugars, amino acids or organic acids (Stöveken et al., 2009). Moreover, it does not possess lysophosphatidyl acyltransferase activity and can thus not acylate lysophosphatidic acid. AbWSD1 shows its highest activity towards the synthesis of WEs, which is approximately ten times higher in comparison to its DGAT activity (Kalscheuer et al., 2003). In detail, the substrates yielding highest activity were found to be hexadecanoyl-CoA (16:0-CoA) in combination with octadecenol (18:1-OH), which was shortly followed by tetradecanol (14:0-OH), octadecanol (18:0-OH) and hexadecanol (16:0-OH) (Stöveken et al., 2005). AbWSD1 from *A. baylyi* is by far the best studied member of all WSD-type enzymes and became a model for the class of WSD-enzymes.

However, WSD from other prokaryotes have been studied as well. On the basis of *in vitro* experiments with purified enzymes produced in *E. coli*, WSDs from *A. baylyi*, *Psychrobacter cryohalolentis* K5, *Rhodococcus jostii* RHA1 as well as two enzymes from *M. aquaeolei* VT8 have been described (Barney et al., 2012a). Similar to AbWSD1, a broad substrate range was shown for all of the five enzymes. Like AbWSD1, all of the enzymes esterified a broad range of fatty alcohols and acyl-CoAs, ranging from eight to eighteen carbon atoms in length in case of fatty alcohols and eight to sixteen carbon atoms in case of acyl-CoAs. Moreover, all of the enzymes were capable of utilizing branched (isoamyl alcohol), medium chain (hexanol), and aromatic alcohols (phenylethanol) as substrates (Barney et al., 2012a). Despite of this broad range of accepted substrates, a certain preference for single species was nevertheless determined for each of the enzymes. In detail, all of them showed highest activity for 14:0-CoA as the acyl donor, while the most preferred acyl acceptors were 11:0-OH and 12:0-OH. Interestingly, the lipid compositions in the natural hosts significantly differed from the determined substrate specificities of the corresponding enzymes. These findings support an assumption also made by others (Röttig and Steinbüchel, 2013), according to which the lipid composition synthesised by WSD is likely to not only be determined by the WSD, but also by upstream factors which are involved in fatty acid synthesis and editing.

The activity of the WSD from *A. baylyi*, *R. jostii* RHA1 and *M. aquaeolei* VT8 were significantly negatively influenced, when the acyl-CoA substrate was added to *in vitro* experiments as the first

## INTRODUCTION

component. In contrast, the enzymatic activities of the WSD from *P. cryohalolentis* K5P and of Ma2 from *M. aquaeolei* VT8 were only slightly influenced. Based on these findings, the authors speculate that the effect may resemble an allosteric control of the enzyme. When the level of fatty alcohols in the cell rises, the enzyme gets activated in order to catalyse the sink reaction of fatty alcohols into WEs. Low levels of fatty alcohols on the other site might inactivate the enzyme (Barney et al., 2012a). Unfortunately, a possible DGAT activity of the enzymes was not elucidated in the study, although substantial amounts of TAG were found in extracts from *A. baylyi* and *R. jostii* (Barney et al., 2012a). To date, two plant-derived WSD-type WSs are described. PhWSD from *P. hydrida* was shown to possess a narrower substrate range in comparison to prokaryotic WSD enzymes. It was neither able to restore TAG synthesis in the neutral lipid deficient *S. cerevisiae* strain H1246, nor to acylate MAG or DAG in microsome-based *in vitro* experiments (King et al., 2007). Moreover, PhWSD did not accept secondary alcohols, primary alcohols longer than 14 carbon atoms or phenylethanol. Instead, it was shown to preferentially take fatty alcohols of 8 to 12 carbon atoms and long, fully saturated acyl-CoAs of 20 or 22 carbon atoms. In comparison to that, the substrate specificity of the second described plant-derived WSD, AtWSD1 from *A. thaliana*, was not described as detailed. *In vivo* experiments in yeast indicated, that the enzyme is capable of utilizing saturated fatty alcohols of 16, 24 and 28 carbon atoms in length. Furthermore, *in vitro* experiments on the basis of cell lysates from AtWSD1 expressing *E. coli* cultures showed the ability of AtWSD1 to acylate DAG, while AtWSD1-mediated restoration of TAG synthesis upon expression in yeast was not possible (Li et al., 2008). Thus, it still remains elusive if all members of the WSD-type WSs are bifunctional enzymes and thus capable of TAG synthesis.

Despite of the low specificity of WSD, an alteration of the preference for fatty alcohols of a WSD from *M. aquaeolei* (MaWSD1) and *A. baylyi* (AbWSD1) was achieved by the construction of single amino acid exchange variants. Respective residues were identified through structural alignments of the crystal structures of the polyketide-associated protein PapA5 and several carnitine acetyltransferases with bound substrates. PapA5, carnitine acyltransferases and WSDs are phylogenetically related enzymes and share the same catalytic motif (see section 1.5.3). By aligning the structures of substrate-bound carnitine acyltransferases with PapA5, residues involved in substrate binding in PapA5 were estimated. In a second step, the corresponding residues in the sequences of WSDs were identified by sequence alignment of AbWSD1 and PapA5 (Barney et al., 2012b). The identified residue is located in proximity to the active site motif of PapA5 in its three dimensional structure. Alignments of the sequences of PapA5 with AbWSD1 and MaWSD1 indicated a glycine to be the corresponding residue in MaWSD1 and AbWSD1. By exchanging this glycine towards an isoleucine, the substrate specificity of the respective variants was shifted towards shorter chain lengths in both cases (Barney et al., 2012b).

### 1.5.4.2 Substrate specificities of DGAT2-type WS

DGAT2-type WSs are predicted to be integral membrane enzymes which mostly harbour between one and three TM domains and are located at the ER (Biestler et al., 2012a, 2012b; Cheng, 2004a; Heilmann et al., 2012). Despite of the difficult handling of membrane-bound enzymes, several members of this class of enzymes were described in detail. The first described DGAT2-like WS was the murine acyl-CoA wax alcohol acyltransferase 2 (MmAWAT2). It was shown to be capable of WE, cholesteryl ester and TAG synthesis and to reside in the ER (Cheng, 2004a; Miklaszewska et al., 2013). While MmAWAT2 has a relatively pronounced substrate preference for 14:0-CoA and 18:1-CoA

## INTRODUCTION

(Miklaszewska et al., 2013), the selectivity for fatty alcohols with a chain length from 10 to 18 carbon atoms seems to be not as distinct. In contrast, the activity of MmAWAT2 for alcohols from 10 to 2 carbon atoms gradually decreases (Shi et al., 2012). Moreover, a pronounced difference between the preferences for saturated and monounsaturated fatty alcohols longer than 18 carbon atoms was described. Respective activities for monounsaturated species were measured to be significantly higher than for their saturated counterparts, which were only poorly utilised. Moreover, MmAWAT2 is not able to efficiently utilise ricinoyl-CoA, acyl-CoA carrying acyl chains longer than 18 carbon atoms or methylated acyl chains (Miklaszewska et al., 2013).

Other WSs of the DGAT2-type have been described from birds (Biester et al., 2012b) and the protozoan *Tetrahymena* (Biester et al., 2012a). Not all of the enzymes showed DGAT-activity in yeast microsome-based *in vitro* experiments, while some of them showed even higher DGAT than WS activities (Biester et al., 2012a, 2012b). Moreover, the acceptance for methylated acyl chains, diols and isoprenols like phytol or geranylgeraniol was highly varying among the different enzymes. In contrast, the selectivity of the enzymes for fatty alcohols was rather similar to each other, as they all showed the highest activity with fully saturated fatty alcohols of acyl chains lengths from 10 to 12 carbon atoms (Biester et al., 2012a, 2012b).

In summary, DGAT2-type WSs are generally more specific to certain substrates in comparison to WSD-type WSs as they exhibit a smaller substrate range which can significantly differ among distinct enzymes. To date, no studies concerning the elucidation of substrate specificity determinants in DGAT2 or DGAT2-type WSs have been published.

### 1.5.4.3 Substrate specificities of DGAT1-type WS

DGAT1-type WSs are the least explored group of WSs so far. In fact, only a WS from *E. gracilis* (EgWS) and a WS of *S. chinensis* (ScWS) are described in more detail to date (Lardizabal et al., 2000; Teerawanichpan and Qiu, 2010). The substrate specificity of EgWS has been tested via *in vitro* experiments upon expression in *S. cerevisiae* on the basis of yeast microsomes. Respective results indicate a rather narrow substrate range of the enzyme. It accepted both saturated fatty acids and fatty alcohols from 12 to 16 carbon atoms as well as 16:1-OH and 16:1-CoA, but no saturated or monounsaturated species longer than 16 or shorter than 12 carbon atoms. Furthermore, no TAG accumulation was observed *in vivo* (Teerawanichpan and Qiu, 2010). ScWS was characterised on the basis of membrane extracts from *S. chinensis*. Respective experiments showed a broader substrate range of ScWS in comparison to EgWS, as the enzyme accepted acyl-CoAs and fatty alcohols from 8 to 24 carbon atoms and also used unsaturated fatty alcohols and acyl-CoAs (Lardizabal et al., 2000). Interestingly, ScWS showed the highest activity towards 18:1-OH, although the most abundant alcohol moiety in WEs from *S. chinensis* seeds is 22:1-OH (Lardizabal et al., 2000). This observation once more illustrates the big influence of factors acting upstream of enzymes involved in WE synthesis on the actual WE compositions. Taken together, available findings on substrate specificity determination in DGAT1-type WSs are rather incomplete. *In vivo* experiments in case of EgWS could lead to falsifying effects caused by unknown conditions and substrate availability in the cells. In case of characterisation of ScWS, contaminating effects of unknown enzymes in membrane preparations of *S. chinensis* can't be ruled out. As for DGAT2-type WSs, no determinants of substrate specificity are known in DGAT1-type WSs to date.

### 1.5.5 Biotechnological applications of Ws

Due to the wide field of industrial applications of WEs, biological WE synthesis has become a popular topic in biotechnology and reports on successful WE production in heterologous systems are accumulating. In general, two main approaches for biological WE production can be distinguished. On the one side, esters of acyl-CoAs with short alcohols like ethanol or methanol are able to serve as biodiesel and can be used for fuel purposes. On the other side, long chain WEs of fatty acyl-CoA and fatty alcohols can be utilised as industrial lubricants.

Biodiesel is a generic term for esters of fatty acids of different chain lengths and either methanol (fatty acid methyl esters, FAME) or ethanol (fatty acid ethyl esters, FAEE). To date, the possibility of biological biodiesel production was exploited in both prokaryotic and eukaryotic systems. Since *S. cerevisiae* is capable of both, ethanol production and fatty acyl-CoA synthesis, it was anticipated to be an appropriate host for the production of biodiesel (Shi et al., 2012). Employment of the unspecific WSD from *M. hydrocarbonoclasticus* as the FAEE forming enzyme in *S. cerevisiae* resulted in yields of 6.3 mg/l FAEE in otherwise unaltered *S. cerevisiae* cells. In comparison, the murine DGAT2-type WS MmAWAT2 yielded 1.3 mg/l FAEE, which is in accordance with its low reported activity for short chain alcohols (Shi et al., 2012). The authors of the study were able to further increase the maximum FAEE yield of 6.3 mg/l to 8.2 mg/l by expressing an acetyl-CoA carboxylase. The enzyme is responsible for the formation of malonyl-CoA, which is the first substrate of fatty acid biosynthesis. Overexpression of the enzyme is thus intended to boost fatty acyl-CoA production (Tehlivets et al., 2007; Wattanachaisaereekul et al., 2008). An ensuing study of the same group describes the stable integration of multiple copies of the WSD gene from *M. hydrocarbonoclasticus* in the yeast genome. The approximate 30 mg/l of FAEE, which were produced by the respective strain, were further increased by the combined coexpression of an acyl-CoA binding protein (ACBP) to increase the amount of acyl precursors on the one side and a NADP<sup>+</sup> dependent glyceraldehyde-3-phosphate dehydrogenase (GAPN) in order to increase NADPH levels on the other side. The respective strain was able to produce FAEE up to nearly 50 mg/l (Shi et al., 2014). Another report on FAEE production in *S. cerevisiae* focused on increasing the endogenous ethanol production of the cell to boost FAEE synthesis. By expression of AbWSD1 from *A. baylyi* and genes responsible for glycerol utilization combined with a knockout of two genes involved in glycerol synthesis and export, the authors achieved a maximum yield of 520 mg/l FAEE. However, oleate was supplemented exogenously to the cultures in this study (Yu et al., 2012).

As for FAEE production in yeast, the preferred enzymes for production of FAEE in *E. coli* are unspecific bacterial WSD as well. Since the intrinsic ethanol production from *E. coli* was shown to be insufficient for FAEE synthesis upon expression of AbWSD1, additional genes encoding a pyruvate decarboxylase and an alcohol dehydrogenase from the strictly anaerobic ethanologenic gram-negative bacterium *Zymomonas mobilis* were expressed (Kalscheuer et al., 2006a). Under anaerobic conditions and supplementation of glucose and sodium-oleate, yields of up to 1.28 g/l of FAEE were realised after 72 h of growth. Interestingly, combination of all heterologous expressed genes on a single plasmid increased the yield by two-fold in comparison to expression of each gene in a single plasmid (Kalscheuer et al., 2006a). The potential of the plasmid was further maximised in fed-batch cultures under optimised conditions, which yielded a final amount of 19 g/l of FAEE (Elbahloul and Steinbüchel, 2010). However, a major drawback of both studies is the exogenous supplementation of fatty acids, which makes the approach expensive.

*De novo* FAEE production was shown to be achievable by additional engineering of the *E. coli* metabolism. Similar to the strategy used for fatty alcohol production, combined expression of a

## INTRODUCTION

thioesterase and an acyl-CoA ligase in combination with the knockout of the acyl coenzyme A dehydrogenase were shown to be suitable for improving the content of fatty acyl substrates in the cell. Appropriate amounts of ethanol were further produced by expressing a pyruvate decarboxylase as well as an alcohol dehydrogenase (Ingram et al., 1987). Combination of these strategies with the expression of two copies of AbWSD1 in *E. coli* resulted in a maximum yield of almost 700 mg/l of FAEE. To reach this yield, the culture was overlaid with dodecane in order to prevent FAEE evaporation (Steen et al., 2010). This strategy was further optimised by additional expression of an acetyl-CoA carboxylase in order to boost fatty acid and thereby acyl-CoA synthesis, as described before (Davis, 2000). In combination with optimised culture conditions, the approach yielded maximum amounts of 922 mg/L FAEE (Duan et al., 2011).

Besides WE synthesis for utilization as biofuels, the production of long chain WEs was explored in the recent years as well. In contrast to microbial-based FAEE production, the production of long chain WEs was also conducted in plants. Plants offer several advantages in comparison to microbial systems. For instance, neither a supplementation of carbon sources nor energy for mixing or heating of large scale fermenters is needed. Feasibility of WE production in plants was already demonstrated as part of the study on the initial discovery of the gene for the WS from *S. chinensis* (Lardizabal et al., 2000). Here, the FAR and WS from *S. chinensis* were expressed in combination with a  $\beta$ -ketoacyl-CoA synthase (KCS) from *Lunaria annua* as part of the fatty acid elongase system (Lassner et al., 1996) in *Arabidopsis* under the control of the napin promoter. In this approach, the KCS was intended to ensure supplementation of acyl chains of appropriate length. Seeds of respective transformants showed WE accumulation reaching a content of 50 % of all oil molecules as determined by NMR measurements (Lardizabal et al., 2000).

Another study reports the expression of a murine FAR and WS in *Arabidopsis*, yielding a mixture of different WEs mainly consisting of monounsaturated C18-C20 fatty alcohol moieties and polyunsaturated 18:2 and 18:3 fatty acyl moieties (Heilmann et al., 2012). Since the aim of this study was the production of oleyl-oleate, the genes were furthermore expressed in an *Arabidopsis* mutant line devoid of a fatty acid elongase and a fatty acid desaturase. As result, the respective plants accumulate oleate. Seeds of respective lines contained more than 60 % of oleyl-oleate in the WE fraction, while expression of the murine FAR and WS in wild type *Arabidopsis* yielded less than 5 % oleyl-oleate (Heilmann et al., 2012). The study is thus another example for the big influence of the FAR and WS substrate pools on the final WE composition, caused by an insufficient substrate specificity of WE synthesising enzymes. Moreover, the study describes an attempt of colocalisation a FAR and a WS within the cell. Fusing of the enzymes to oleosin, which is a lipid droplet (LD) located protein in *Arabidopsis* (Siloto, 2006), resulted in increased WE amounts upon expression in yeast. However, it did not yield a comparable effect upon expression in *Arabidopsis* (Heilmann et al., 2012). Other enzymes expressed in plants under the seed specific napin promoter include a combination of MaFAR1 from *M. aquaeolei* and ScWS from *S. chinensis* (Iven et al., 2013). In this case, both moieties of the WEs were found to be mainly monounsaturated chains of 18 to 22 carbon atoms in length.

Most approaches for lipid production in plants are targeted to seeds, as these are the main storage compartments for respective lipids in the plant. However, seeds only represent a small amount of a plant's biomass. Thus, WE production in plastids was considered to be an attractive alternative (Aslan et al., 2014). A respective approach was shown to be feasible by transient expression of a plastidial FAR, a prokaryotic FAR, a prokaryotic WS as well as a phytol ester synthase (PES) in different combinations in leaves of *N. benthamiana*. All enzymes were fused to a plastidial localisation signal from *Arabidopsis* to ensure a plastidial localisation. A maximum yield of WEs corresponding to

## INTRODUCTION

approximately 0.9 % of the leaves' dry weight was achieved by expression of MaFAR1 and AtPES2 (Aslan et al., 2014).

Besides biotechnological WE production for utilization in industrial processes, enhanced wax production on plant surfaces can serve as an improvement of the plant's drought tolerance. A proof of principle concerning this theory was demonstrated by expression of the *Arabidopsis* transcription factor MYB96 under the control of the CaMV35S promoter in *Camellina sativa*. Respective plants were found to possess a significantly increased drought resistance in comparison to the wild type due to elevated wax loads on its leaf surface. Transcriptomal analyses revealed an up-regulation of multiple wax biosynthetic genes (Lee et al., 2014).

To date, heterologous production of long chain WE production in microbes is not as well explored as it is in plants. Nevertheless, it has been shown to be possible by expression of the FAR from *S. chinensis* in combination with AbWSD1 from *A. baylyi* in *E. coli* in the presence of oleate. Respective cultures produced WEs up to 1 % of the cellular dry weight (Kalscheuer et al., 2006b).

### 1.5.6 Role of WEs in nature

In nature, WEs are incredibly widespread substances. They occur in all kind of organisms, e.g. bacteria (Barney et al., 2012a; Kalscheuer, 2002; Röttig and Steinbüchel, 2013; Santala et al., 2014), insects (Blomquist et al., 1972; Buckner et al., 2009; Rottler et al., 2013), vertebrates (Bagge et al., 2012; Cheng, 2004a; Miklaszewska et al., 2013; Rantamäki et al., 2013; Takagi and Itabashi, 1977), nematodes (Penkov et al., 2014) and plants (Bernard and Joubès, 2013; Duncan et al., 1974; Razeq et al., 2014; Yeats and Rose, 2013). In this respect, functions of WEs are highly diverse as well, ranging from acting as storage compounds, structural components or UV-barriers to functioning as repellents or buoyancy-regulating agents.

In humans, sebaceous glands-derived sebum consists of approximately 25 % WEs (Downing and Strauss, 1974). The actual role of sebum on our skin is not fully understood yet (Smith and Thiboutot, 2008), but it's anticipated to deliver antioxidants, pheromones and antibiotic lipids to the body surface (Kligman, 1963; Thiele et al., 1999). *In vivo* experiments furthermore indicate an efficient evaporation-retarding effect of WEs in the tear film of the eye (Craig and Tomlinson, 1997). After initial problems upon reproduction of these findings *in vitro* (Herok et al., 2009), it's known today that the evaporation-retarding effects of WEs in tear films seem to be closely associated with their specific melting temperature. WEs adopt a unique physical state at their melting temperature, which makes them able to efficiently cover a liquid's surface, inducing an anti-evaporative effect (Rantamäki et al., 2013).

A unique case of utilization of WEs in mammals can be found in the sperm whale. Unlike other mammals, sperm whales possess a specialised organ in their forehead which can harbour huge amount of a so called "spermaceti oil" (Perry et al., 1999). The biological function of the spermaceti organ is a matter of discussion until today. Theories span over influence of the spermaceti oil on the whale's buoyancy (Clarke, 1978) to a function as a weapon (Carrier et al., 2002) or in orientation through the animal's sonar (Zimmer et al., 2005). Spermaceti oil consists of 70 to 100 % of WEs, while the second major constituent is TAG. The WE fraction of spermaceti oil consists mainly of monoesters from 16:0-OH esterified to fatty acids ranging from 10 to 20 carbon atoms (Morris, 1973). However, data on the detailed fatty acid moiety composition in spermaceti WEs diverge. Wellendorf determined tetradecanoic (14:0) and hexadecanoic acid (16:0) as the main constituents (Wellendorf, 1963), whereas Morris claims dodecanoic acid (12:0) to be the most abundant fatty acid

## INTRODUCTION

moiety (Morris, 1973, 1975). In contrast, the Japanese lipid bank website asserts hexadecanoic acid to be the major fatty acid moiety with amounts of over 90 % (Japanese Conference on the Biochemistry of lipids (JCBL), 2014).

The best known insect-derived WEs are produced by honey bees like *Apis mellifera*. Bees use beeswax as a structural component in order to build honeycombs which can house larvae, honey or pollen. Chemically, beeswax is a blend of an estimated 300 substances. With 35 %, monoesters make up the major part of beeswax. Further constituents comprise hydrocarbons (14 %), diesters (14 %), free acids (12 %) and to a lower extent also triesters, hydroxymonoesters, hydroxypolyesters, acid polyesters and free alcohols (Tulloch, 1970, 1971). In comparison to that, the wax of bumble bees contains approximately twice the amount of hydrocarbons (Tulloch, 1970). Interestingly, bees seem to use their waxes not only for production of honeycombs, but also for orientation. It was shown, that members of different nests also differ in their cuticular wax composition. Traces of these cuticular wax compositions are left behind by the insects at their nests, allowing them to efficiently distinguish between their own and foreign nests (Buckner et al., 2009; Rottler et al., 2013). The use of WEs as structural components is not limited to bees. Larvae of the nematode *Pristionchus pacificus* uses a very long chain WE in the course of host finding. Coverage of their cuticle with the WEs make the larvae sticky, resulting in a congregation of up to one thousand individuals. The so formed structures are called “dauer towers” and enable the larvae to reach a potential host more easily. The respective WEs which are synthesised by the larvae derive from two C30 chains carrying six double bonds each. With 60 carbon atoms in length, these WEs are one of the longest WEs ever described in a living organism (Penkov et al., 2014).

As WEs consist of highly reduced carbon atoms, possess a high energy density and are osmotically inert, they are perfect storage compounds (Wältermann and Steinbüchel, 2005). Despite of the fact that most bacteria known today synthesise polyhydroxyalkanoate (PHA) or poly(3-hydroxybutyrate) (PHB) as a carbon storage compound, several species which synthesise TAGs or WEs have been described as well (Alvarez and Steinbüchel, 2002; Wältermann et al., 2007). WE synthesising bacteria are mostly derived from the gram-negative species of *Acinetobacter*, in which WE accumulation was first described 1971 (Gallagher, 1971). Today, wax production and accumulation has also been discovered in other species, such as *Alcanivorax*, *Marinobacter*, *Thalassolituus*, *Psychrobacter*, *Micrococcus*, *Moraxella*, *Corynebacterium* or *Nocardia* (Röttig and Steinbüchel, 2013). In all of those species, WEs are accumulated as a supply for times of food deficiency. In contrast, the gram-positive actinobacterium *Mycobacterium tuberculosis* is reliant on WEs in the process of dormancy induction for persistence of the organism in a host. This induction is impaired in knockout mutants devoid of two FARs responsible for the synthesis of the WEs' fatty alcohol moiety. As a consequence, the bacterium is not able to seal its cell wall with the help of WEs and prevent nutrient uptake. Decreased nutrient uptake is necessary for the inhibition of replication (Sirakova et al., 2012).

In contrast to the well described synthesis of WEs as storage compounds in *Acinetobacter* and several other bacterial species, there is only one known example for a comparable WE utilization in eukaryotes. *S. chinensis*, also called Jojoba, is a slowly growing desert scrub, which is believed to reach a live span of over one hundred years (Gentry, 1958). Its seeds contain approximately 50 % of fluid WEs instead of TAGs as storage lipids (Greene and Foster, 1933). Since Jojoba WEs are the only natural source of renewable WEs to date (Ash et al., 2005), Jojoba plants are being cultivated in specialised farms in order to harvest the WE containing seeds. However, the costs for Jojoba-derived WEs are rather high due to its slow growth and its shrubby habitus, making the harvest of the seeds more complicated than for normal crop plants. After spermaceti oil became unavailable for lubrication purposes in the 1980s, Jojoba-derived WEs were used in the industry as supplements for

## INTRODUCTION

lubricants. Other utilization of Jojoba oil in industry include cosmetics, in which Jojoba oil can be used in shampoos, lotions, crèmes etc. (Green, 2013; McCarthy, 2014). Interestingly, the wax also seems to be suitable for treatment of powdery mildew, a fungal disease affecting a wide range of plants (Hicks and Siemer, 2001). The molecular composition of Jojoba wax comprises mostly of dienoic WEs of 40 and 42 carbon atoms (27 % and 48 % of the total WE load, respectively) (Tada et al., 2005). The major fatty acid moiety within Jojoba wax is eicosenoic acid (20:1-COOH), whereas the most abundant alcohol moieties are eicosenol (20:1-OH) and docosenol (22:1-OH) (Duncan et al., 1974; Spencer et al., 1977). Jojoba WE molecules are thus different from spermaceti-derived WEs in respect to their higher chain lengths as well as in respect to their higher grade of desaturation.

## 2 OBJECTIVES OF THIS WORK

Current knowledge about WE biosynthesis indicates low substrate specificities for most described FARs and WSs. This circumstance results in problems concerning application of FARs and WSs for biotechnological approaches, since their low substrate specificities are often not able to realize distinct WE compositions. Instead, WE blends are often highly influenced by the substrate pool compositions of FARs and WSs. Thus, factors acting upstream of WE synthesising enzymes have a prominent role in determination of the final WE composition. As a consequence, the production of defined WE blends in genetically modified systems by the sole expression of FARs and WSs is difficult. One strategy to overcome this issue is to adjust respective substrate pools by metabolic engineering in order to obtain a desired WE blend. However, respective modifications complicate the approach, as they have to be done in addition to the expression of FARs and WSs. Moreover, they may interfere with the metabolism of the host, resulting in further difficulties. Also, the provision of two different acyl-CoA substrate pools within one cell is a problem, as respective modifications in the hosts metabolism might interfere with each other. Yet, different acyl-CoAs would be necessary in order to produce WEs consisting of two moieties differing in chain length, grade of desaturation or methyl branching.

In order to overcome these difficulties, modified enzymes with high substrate specificities might be an appropriate tool. If a certain specificity for both, FARs and WSs, could be assured regardless of the composition of the substrate pool, no additional metabolic engineering would be necessary in order to produce WEs of defined compositions. Moreover, modified FARs and WSs with high specificities could be used for the production of compounds which are only synthesised to an unsatisfactory extend by currently known wild type enzymes, for instance FAEE and FAME.

The construction of optimised FARs and WSs with distinct substrate specificities requires detailed knowledge about substrate specificity determining structures in respective enzymes. However, to date this knowledge is scarce. Neither structures of FARs nor of WSs are reported until now. Moreover, there is only one study describing a change of substrate specificity in FARs through protein engineering (Chacon et al., 2013). Likewise, only one report on changing substrate specificity in a WSD-type WS through protein engineering was published (Barney et al., 2012b). Although both studies are important contributions towards a detailed understanding of substrate specificity in FARs and WSs, they also illustrate the infancy of the whole field. Hence, significant efforts are necessary in order to be able to construct WE biosynthetic enzymes with distinct substrate specificities.

In the present study, the main aims were therefore to contribute to the elucidation of structure-function relationships in FARs and WSs. For this purpose, biochemical and biophysical properties of four FARs, two WSs as well as of a soluble DGAT3 have been studied in order to obtain structural insights into the respective enzyme classes.

### 3 MATERIALS

#### 3.1 Software and web-based services

Computational work and processing of data in this study was done with the respective programs and services listed in Table 1.

**Table 1: Software and web-based services used in this study**

Name	Purpose	Reference/Manufacturer
Caver	protein tunnel analysis software	CaverSoft s.r.o., Czech Republic
ChemStation	analysis of HPLC and GC-FID-derived data	Agilent, Germany
ClustalX	construction of phylogenetic trees	Thompson, 1997
Excel	statistics, data analyses, diagrams	Microsoft, USA
Graphical Codon Usage Analyser	graphical codon usage analyses	(Fuhrmann et al., 2004)
Illustrator	generation figures and schemes	Adobe, USA
ImageJ	quantification of graphical data	Schneider et al., 2012
NIST MS Search 2.0 library	identification of Mass spectra	<a href="http://www.nist.gov/">http://www.nist.gov/</a>
OriginPro	statistics, data analyses, diagrams	OriginLab, USA
Phobius	transmembrane prediction	Käll et al., 2007
Photoshop	figures and schemes	Adobe, USA
Phyre Server 2.0	protein-modelling (intense mode)	Kelley and Sternberg, 2009
PyMol	visualization of protein-models	Schrödinger, 2010
SOSUI	transmembrane prediction	Hirokawa et al., 1998
TMHMM	transmembrane prediction	Sonnhammer et al., 1998
TreeView	construction of phylogenetic trees	Page, 1996
Xcalibur	analysis of GC-MS-derived data	Thermo Electron Corp., Germany

## 3.2 Machines and equipment

**Table 2: Machines used in this study**

Item	Manufacturer
6890 Series GC System	Agilent, USA
AUTOMATIC TLC SAMPLER 4	CAMAG, Switzerland
ÄKTAprime™ plus	GE Healthcare, USA
BAS-MP 2040S IMAGING PLATE	FUJI PHOTO FILM CO., LT
Cartesian Microsys robot	Cartesian Technologies, USA
CARY 100 Bio UV-vis Spectrophotometer	Varian, Germany
Centrifuge 5417 R	Eppendorf, Germany
Centrifuge 5810 R	Eppendorf, Germany
CFX96 realtime PCR cycler	BioRad, Germany
CHROMATOGRAM IMMERSION DEVICE III	CAMAG, Switzerland
Corning® Spin-X® UF Concentrators	Life Sciences (Lowell, MA, USA)
Diana documentation system	Raytest, Germany
Fuji FLA-3000	Raytest, Germany
IDA gel documentation system	Raytest, Germany
Imaging plate	FUJIFILM, Japan
LC Agilent 1100 Serie	Agilent, USA
Mastercycler personal	Eppendorf (Hamburg, Germany)
Mastercycler gradient	Eppendorf (Hamburg, Germany)
miniDAWN	Wyatt, Germany
Mini-PROTEAN3 Electrophoresis System	Bio-Rad Laboratories GmbH, Germany
Polaris Q mass selective detector	Thermo Finnigan, USA
Rock Imager System	Formulatrix, USA
Sonifier® Cell Disruptor B15	Branson, Germany
Sterile bench Prettl® Telstar Bio II A	Telstar, Spain
TLC PLATE HEATER III	CAMAG, Switzerland
TLC SPRAY CABINET III	CAMAG, Switzerland
Trace gas chromatograph	Thermo Finnigan, USA
UV-table 312 nm	Raytest, Germany

## MATERIALS

### 3.3 Chemicals

Chemicals were obtained from Sigma-Aldrich (USA) or Carl Roth (Germany), unless stated otherwise. Oligonucleotides depicted in Table 9 were obtained from Sigma-Aldrich (USA).

**Table 3: Chemicals used in this study**

Chemical	Manufacturer
[ <sup>14</sup> C]-labelled substances	CPS, Germany
Acetonitrile	Baker, USA
Agar	Duchefa Biochemie, Netherlands
Agarose	Biozyme Scientific GmbH, Germany
Carbenicillin	Duchefa Biochemie, Netherlands
Hexane	Baker, USA
Kanamycin	Duchefa Biochemie, Netherlands
Methanol	Baker, USA
SYPRO Orange	Sigma, Germany

### 3.4 Kits, markers and consumables

All kits and markers in this study were used according to the respective manufacturer's instructions.

**Table 4: Kits, markers and consumables used in this study**

Item	Manufacturer
CloneJETPCR Cloning Kit	Thermo Scientific, USA
GeneRuler -50 bp DNA Ladder	Thermo Scientific, USA
GeneRuler 1 kb DA Ladder	Thermo Scientific, USA
Polystyrene cuvettes	SARSTEDT, Germany
NucleoSpin® Gel and PCR Clean-up	Macherey-Nagel, Germany
NucleoSpin™ Plasmid	Macherey-Nagel, Germany
Protino® Ni-NTA Agarose	Macherey-Nagel, Germany
Roti®-Mark STANDARD Protein-Molekulargewichtsmarker	Carl Roth, Germany
Spin-X® UF Concentrators	CORNING, USA
TLC Silica gel 60	Merck, Germany
Unstained Protein Molecular Weight Marker	Thermo Scientific, USA

## MATERIALS

### 3.5 Commercial crystallisation screens

For crystallisation of the 6xHis-TF-AbWSD1 fusion protein, the following commercially available screens were used:

**Table 5: Commercially available crystallisation screens used in this study**

Screen name	Manufacturer
Classics	Qiagen, Netherlands
JCSG +	Qiagen, Netherlands
Midas	Molecular Dimensions, United Kingdom
MPD	Qiagen, Netherlands
Nucleix	Qiagen, Netherlands
PACT	Qiagen, Netherlands
PEG II	Qiagen, Netherlands
Wizard 1 & 2	Rigaku Reagents, USA

### 3.6 Customised crystallisation grid-screens

Customised crystallisation screens were designed together with Dr. Karin Kühnel (Department of Neurobiology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany) and Felix Lambrecht (Department of Plant Biochemistry, University of Göttingen, Germany) according to the most promising results obtained in commercially available screens (listed in 3.5). A respective master-block was mixed by a Tecan robot-system. For crystallisation screens, either 1 ml of each condition was pipetted in 24-well hanging drop plates or 100  $\mu$ l of each condition was pipetted in 96-well plates to serve as the reservoir. Components depicted in the middle of the tables were constantly present in each condition.

#### Custom Screen 24-1

##### 30 % (v/v) Jeffamine ED-2001

	15 %	20 %	25 %	30 %	35 %	40 %
0.1 M HEPES pH 6.0						
pH 6.5						
pH 7.0						
pH 7.5						

MATERIALS

Custom Screen 24-2

PEG 4000

		20 %	24 %	26 %	32 %	36 %	40 %
0.1 M HEPES	pH 6.0						
	pH 6.5						
	pH 7.0						
	pH 7.5						

10 % (v/v) Isopropanol, 0.2 M MgCl<sub>2</sub>

Custom Screen 24-3

Ethanol

		10 %	15 %	20 %	25 %	30 %	40 %
0.1 M HEPES	pH 6.0						
	pH 6.5						
	pH 7.0						
	pH 7.5						

0.2 M MgCl<sub>2</sub>

Custom Screen 96-1

Isopropanol (% v/v)

		5.0	9.1	13.2	17.3	21.4	25.5	29.6	33.7	37.7	41.8	45.9	50.0
0.1 M HEPES pH 6.5													
0.1 M HEPES pH 7.5													
Tris pH 8.0													
Tris pH 9.0													

0.2 M MgCl<sub>2</sub>

tert-Butanol (% v/v)

		5.0	9.1	13.2	17.3	21.4	25.5	29.6	33.7	37.7	41.8	45.9	50.0
HEPES pH 6.5													
0.1 M HEPES pH 7.5													
Tris pH 8.0													
Tris pH 9.0													

0.2 M MgCl<sub>2</sub>

## MATERIALS

### Custom Screen 96-2

	1,6-Hexandiol (% v/v)											
	5.0	9.1	13.2	17.3	21.4	25.5	29.6	33.7	37.7	41.8	45.9	50.0
0.1 M HEPES pH 6.5												
0.1 M HEPES pH 7.5												
0.1 M Tris pH 8.0												
0.1 M Tris pH 9.0												

0.2 M MgCl<sub>2</sub>

	Ethanol (% v/v)											
	5.0	9.1	13.2	17.3	21.4	25.5	29.6	33.7	37.7	41.8	45.9	50.0
0.1 M HEPES pH 6.5												
0.1 M HEPES pH 7.5												
0.1 M Tris pH 8.0												
0.1 M Tris pH 9.0												

0.2 M MgCl<sub>2</sub>

### 3.7 Enzymes

All enzymes used in this study were handled according to the manufacturer's instructions.

**Table 6: Enzymes used in this study.**

Enzyme	Manufacturer
Phusion High Fidelity DNA Polymerase	Finnzymes, Finland
GoTaq® DNA Polymerase	Promega, USA
T4 DNA Ligase	Fermentas, USA
Factor Xa	Prozyme, USA
Thrombin	Sigma-Aldrich, USA
Restriction Endonucleases	NEB, USA
Restriction Endonucleases	Fermentas, USA

## MATERIALS

### 3.8 Strains and organisms

**Table 7: Strains and organisms used in this study**

Strain	Features	Reference
<i>E. coli</i> XL1-blue	endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB+ lacIq Δ(lacZ)M15] hsdR17(rK- mK+)	Agilent Technologies, USA
<i>E. coli</i> BL21* (DE3)	fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHlo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5	NEB, USA
<i>E. coli</i> Lemo21 (DE3)	fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/ pLemo(CamR) λ DE3 = λ sBamHlo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5 pLemo = pACYC184-PrhaBAD-lysY	NEB, USA
<i>E. coli</i> ArcticExpress (DE3)	<i>E. coli</i> B F- ompT hsdS(rB- mB-) dcm+ Tetr gal λ(DE3) endA Hte [cpn10 cpn60 Gentr]	Agilent Technologies, USA
<i>S. cerevisiae</i> H1246	MATα are1-Δ::HIS3 are2-Δ::LEU2 dga1-Δ::KanMX4 Iro1-Δ::TRP1 ADE2	(Sandager, 2001)
<i>S. cerevisiae</i> W303	MATα leu2-3, 112 his3-11, 15 trp1-1 can1-100 ade2-1 ura3-1	(Sandager, 2001)

### 3.9 Columns for chromatography

**Table 8: Columns used in this study**

Item	Purpose	Manufacturer
HisTrap HP	IMAC	GE Healthcare, USA
MBP-Trap HP	affinity chromatography	GE Healthcare, USA
HiLoad 26/60 Superdex S200	SEC	GE Healthcare, USA
HiLoad 16/60 Superdex 200	SEC	GE Healthcare, USA
HiTrap PhenylHP 1 ml	hydrophilic interaction chromatography	GE Healthcare, USA
Mono Q	anion exchange chromatography	GE Healthcare, USA
Mono S	cation exchange chromatography	GE Healthcare, USA
Source 15 Phe	hydrophilic interaction chromatography	GE Healthcare, USA
DB-23 (30m x 0.25 mm; 0.25 mm coating thickness)	GC-FID / GC-MS	Agilent, USA
19091j-413 HP5 5 % Phenyl Methyl Siloxane	GC-FID	Agilent, USA
Restek Rxi™-5ms capillary	GC-MS	Restek, USA
LUNA 150 x 2,0 mm	HPLC, Acyl-CoA analyses	Phenomex, Germany

## MATERIALS

### 3.10 Oligonucleotides

**Table 9: Oligonucleotides used in this study**

Construct	Forward 5' → 3'	Reverse 5' → 3'
pET28a_AbWSD1	ACGGGATCCATGCGTCCGCTGCACCCGAT	ACGAAGCTTTTAGTTGGCTGTCTTAATGT
pET28a_AbWSD1 I358S I359S	GGCTTAAATAGCAGCAGCGGCATGATGCCGAAGC	CATGCCGCTGTGCTATTTAAGCCGGCAGGGCCG
pCold_AtDGAT3	CAGGAATTCATGGAGAAGGAGAAGAAGGCGT	CAGTCTAGAATATGAGACAGAACCGAGTC
pET28a_AtDGAT3	CAGGAATTCATGGAGAAGGAGAAGAAGGCGT	CAGCTCGAGATATGAGACAGAACCGAGTC
pYES2/NT_AtDGAT3	CAGGAATTCATGGAGAAGGAGAAGAAGGCGT	CAGTCTAGAATATGAGACAGAACCGAGTC
pET28_MmAWAT2 pYES2/NT_MmAWAT2	ACGGGATCCATGTTCTGGCCTACTAAG	ACGCTCGAGTTAGACGATAACCAGCTC
pYES2/NT_MmAWAT2ΔN pET28_MmAWAT2ΔN pCold_MmAWAT2ΔN	ACGGGATCCGAAAGAGGAGGTAGAAGG	ACGCTCGAGTTAGACGATAACCAGCTC
pYES2/NT_AtOLE3-MmAWAT2ΔN (AtOLE3-Part)	ACGGGTACCATGGCGGACCAAACAAGAACC	CGCCACTTTCTCACACATGTGAAAGAACTTGGG
pYES2/NT_AtOLE3-MmAWAT2ΔN (MmAWAT2-Part)	GTCCAACACCAACAAGTTTCTTTCACATGTGTGAG	ACGCTCGAGTTAGACGATAACCAGCTC
pYES2/NT_MmDGAT2	ACGGGATCCATGAAGACCCTCATCGC	ACGCTCGAGTCAGTTCACCTCCAGCA
pYES2/NT_MmDGAT2ΔN	CCTGCCACCTTTCTTGGGCGTGGATCCCGT	ACGCTCGAGTCAGTTCACCTCCAGCA
pET28_MmDGAT2ΔN	ACGGGATCCACGCCAAGAAAGGTGGCAGG	ACGCTCGAGTCAGTTCACCTCCAGCA
pYES2/NT_V1 (MmDGAT2-part)	ACGGGATCCATGAAGACCCTCATCGC	AAAGACCTCCATGGCCTGCAGCTGTTTTCCAC
pYES2/NT_V1 (MmAWAT2-part)	GAAAAACAGCTGCAGGCAATGGAGGTTTTCTGCT	ACGCTCGAGTTAGACGATAACCAGCTC
pYES2/NT_V2 (MmDGAT2-part)	ACGGGATCCATGTTCTGGCCTACTAAGAAGGACCTTAAGACTCAGCTGCAGGTCATCTCA	TCTACCTCCTTTTCGGGCGTGTTCAGTCAAAA
pYES2/NT_V2 (MmAWAT2-part)	GACTGGAACACGCCCGAAAGAGGAGGTAGAAAGG	ACGCTCGAGTTAGACGATAACCAGCTC
pYES2/NT_V3 (MmAWAT2-part)	ACGGGATCCATGTTCTGGCCTACTAAG	CCTTCTACCTCCTTTTTAGGTGTTTTCCAATCGAATGC
pYES2/NT_V3 (MmDGAT2-part)	GGAAAACACCTAAAAAAGGAGGTAGAAGGTCGCAGTGGG	ACGCTCGAGTCAGTTCACCTCCAGCA
pYES2/NT_V4 (MmAWAT2-part 1)	ACGGGATCCATGTTCTGGCCTACTAAG	CGCACCCACTGCGACCTTCTACCTCCTTTCA
pYES2/NT_V4 (MmDGAT2-part)	GAGGAGGTAGAAGGTCGAGTGGGTGCGAAACT	GATACCAGGGAAGGTCTTGCTGACTTCAGTAGC
pYES2/NT_V4 (MmAWAT2-part 2)	TACTGAAGTCAGCAAGACCTTCCCTGGTATCACAC	ACGCTCGAGTTAGACGATAACCAGCTC
pYES2/NT_V5 (MmDGAT2-part)	ACGGGATCCATGAAGACCCTCATCGC	CTCACACATGTGAATCTCCTGCCACCTTTCTTG
pYES2/NT_V5 (MmAWAT2-part)	GAAAGGTGGCAGGAGATTCACATGTGTGAGAAAG	ACGCTCGAGTTAGACGATAACCAGCTC
pYES2/NT_V6 (MmAWAT2-part)	ACGGGATCCATGTTCTGGCCTACTAAG	GGGCAGATGCCTCCTGACATGACGTAATCTCTC

## MATERIALS

pYES2/NT_V6 (MmDGAT2-part)	GATTACGTCATGTCAGGAGGCATCTGCCCTGT C	ACGCTCGAGTCAGTTCACCTCCAGCA
pYES2/NT_V7 (MmAWAT2-part)	ACGGGATCCATGTTCTGGCCTACTAAG	AAGTGGGAACCAGAGAAACCCCATGCTTTAGTG C
pYES2/NT_V7 (MmDGAT2-part)	GCATGGGGTTTCTCTGGTCCCACTTATTCCTT T	ACGCTCGAGTCAGTTCACCTCCAGCA
pYES2/NT_MmAWAT2 E14Q	GACTGCAATGCAGGTTTTCGCTCTTTTCCAATG GG	GAGCGAAAACCTGCATTGCAGTCTTAAGGTCC
pYES2/NT_MmAWAT2 A25F	TGGGCTTTGTCTTTCCTCGTTATCGTTACTACC GTC	AACGATAACGAGGAAAGACAAAGCCCATTGGA AAAGAGCG
pYES2/NT_MmAWAT2 T30A T31A	CGTTATCGTTGCTGCCGTATCATCGTGAACCT GTACC	CGATGATGACGGCAGCAACGATAACGAGTGCA GAC
pYES2/NT_MmAWAT2 N36A	GTCATCATCGTGGCCCTGTACCTTGTGTGTT ACTTC	CAAGGTACAGGGCCACGATGATGACGGTAGTA AC
pYES2/NT_MmAWAT2 N36K	CCGTATCATCGTGAACTGTACCTTGTGTGT TCAC	CACAACAAGGTACAGTTTCACGATGATGACGGT AGTAACG
pYES2/NT_MmAWAT2 N36L	CCGTATCATCGTGCTCCTGTACCTTGTGTGT TCAC	CAAGGTACAGGAGCACGATGATGACGGTAGTA ACG
pYES2/NT_MmAWAT2 N36R	CGTCATCATCGTGCGCCTGTACCTTGTGTGTT CAC	CAAGGTACAGGGCCACGATGATGACGGTAGTA ACG
pYES2/NT_MmAWAT2 N36W	CCGTATCATCGTGUGGCTGTACCTTGTGTGT TTCAC	CACAACAAGGTACAGCCACACGATGATGACGG TAGTAACG
pYES2/NT_MmAWAT2 L39-	CGTGAACCTGTACGTTGTGTTCACTTCATACTG GCC	GTGAACACAACGTACAGGTTACGATGATGACG G
pYES2/NT_MmAWAT2 F42C	CCTTGTGTGTGCACTTCATACTGGCCAGTTAC C	CAGTATGAAGTGACACAACAAGGTACAGGTT ACG
pYES2/NT_MmAWAT2 S44D	GTTGTGTTCACTGACTACTGGCCAGTTACCGTT TTGATGC	GGTAACTGGCCAGTAGTCAGTGAACACAACAA GGTACAGG
pYES2/NT_MmAWAT2 E64K	TGGAAAACACCTAAAAGAGGAGGTAGAAGGT TCACAT	TCTACCTCCTCTTTTAGGTGTTTTCCAATCGAAT GC
pYES2/NT_MmDGAT2 F80A	TGGGTCCTATCCGCCCTGGTGCTAGGAGTGCC CTGCAG	CTCCTAGCACCAGGGCCGGATAGGACCCATTGTA GTAC
pYES2/NT_MmDGAT2 L91R	GCAGTGTATCAGGATGTACACCTTCTGCACA GAC	GGTGTACATCCTGATGACACTGCAGGCCACTCC
pCold_DmFAR1 pYES2/NT_DmFAR1	ACGCTCGAGATGGATGATCCCAAATAATGAA CA	ACGTCTAGACTATAAGAATAGCTTGAGCAGGGC
pET24_DmFAR1	ACGTCTAGAATGGATGATCCCAAATAATGAA CA	AGTCTCGAGTAAGAATAGCTTGAGCAGGGCAT AG
pET28_DmFAR1	ATCGAATTCATGGATGATCCCAAATAATGAA CA	AGTCTCGAGCTATAAGAATAGCTTGAGCAGGGC ATAG
pMAL_DmFAR1	ACGTCTAGAATGGATGATCCCAAATAATGAA CA	ACGGTCGACCTATAAGAATAGCTTGAGCAGGG C
pCold_MaFAR1	CAGGGTACCATGGCAATCCAGCAGGTC	AGTGGATCCTCATGCCGCTTTTTACGTTGACG
pET24_MaFAR1	CAGGGATCCATGGCAATCCAGCAGGTC	AGTCTCGAGTGCCGCTTTTTACGTTGACGGGC
pMAL_MmFAR1	ACGTCTAGAATGGTTTCAATCCAGAATA	ACGGTCGACTCAGTATCTCATAGTGCT

## MATERIALS

### 3.11 DNA constructs

**Table 10: DNA-constructs generated in the course of this study**

Gene	Reference	Backbone	Restriction sites (5'/3')	Resistance (bacteria)/ Auxotrophy (yeast)
AbWSD1	Stöveken et al., 2005	pCold	BamHI/EcoRI	Ampicillin
		pET28a	BamHI/HindIII	Kanamycin
AtDGAT3	this study	pCold	EcoRI/XbaI	Ampicillin
		pET28a	EcoRI/XhoI	Kanamycin
		pYES2/NT	EcoRI/XbaI	Ampicillin/Uracil
MmAWAT2	Heilmann et al., 2012	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
		pET28a	BamHI/XhoI	Kanamycin
MmAWAT2ΔN	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
		pET28a	BamHI/XhoI	Kanamycin
		pCold	BamHI/XhoI	Ampicillin
AtOLE3- MmAWAT2ΔN	this study	pYES2/NTc	KpnI/XhoI	Ampicillin/Uracil
MmAWAT2 E14Q	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 A25F	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 A25F N36R	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 T30A T31A	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 T30A T31A N36R	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 N36A	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 N36K	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 N36L	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 N36R	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 N36R F42R S44D	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 N36W	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 L39-	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 S44D	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 S44D E64K	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 E64K	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 C72W	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmAWAT2 C106Y	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmDGAT2	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmDGAT2ΔN	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
		pET28a	BamHI/XhoI	Kanamycin

## MATERIALS

MmDGAT2 F80A	Stone et al., 2006	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
MmDGAT2 L91R	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
V1	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
V2	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
V3	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
V4	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
V5	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
V6	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
V7	this study	pYES2/NTc	BamHI/XhoI	Ampicillin/Uracil
DmFAR1	Jaspers et al., 2014	pMAL	XbaI/SalI	Ampicillin
		pCold	XhoI/XbaI	Ampicillin
		pET24a	XbaI/XhoI	Kanamycin
		pET28a	EcoRI/XhoI	Kanamycin
MaFAR1	Hofvander et al., 2011; Wahlen et al., 2009	pMAL	KpnI/BamHI	Ampicillin
		pCold	KpnI/BamHI	Ampicillin
		pET24a	BamHI/XhoI	Kanamycin
MmFAR1	Heilmann et al., 2012	pMAL	XbaI/SalI	Ampicillin
AbWSD1	Stöveken et al., 2005	pCold	BamHI/EcoRI	Ampicillin
		pET28a	BamHI/HindIII	Kanamycin
TcFAR1	Sequence kindly provided by Sabrina Lehmann, Georg-August-University Göttingen Johann-Friedrich-Blumenbach-Institute of Zoology and Anthropology Department of DEV BIOL (UniProt ID = D2A5A7)	pYES2/NTc	NotI/XbaI	Ampicillin/Uracil

## MATERIALS

### 3.12 Buffers, media and antibiotics

All media were prepared according to Ausubel et al. (1993), unless otherwise stated. Media were autoclaved at 120 °C for 20 min prior use.

#### Lysogeny broth (LB)-medium

amount	component
10 g/l	peptone
10 g/l	yeast extract
5 g/l	NaCl

#### 2xYT-medium

amount	component
30 g/l	peptone
20 g/l	yeast extract
5 g/l	NaCl

#### Yeast extract peptone dextrose (YPD)-medium

amount	component
10 g/l	yeast extract
10 g/l	peptone
20 g/l	glucose (added after autoclaving)

#### M9 medium for SeMet protein expression

5x M9stock solution	
30 g/l	Na <sub>2</sub> HPO <sub>4</sub>
15 g/l	KH <sub>2</sub> PO <sub>4</sub>
5 g/l	NH <sub>4</sub> Cl
2.5 g/l	NaCl

M9 medium	
200 ml/l	5x M9 stock solution
800 ml/l	H <sub>2</sub> O
1 ml/l	1 M MgSO <sub>4</sub>
100 ml	4 % (w/v) glucose in H <sub>2</sub> O
100 µl	0.5 % (w/v) thiamine
1 ml	4.2 g/l FeSO <sub>4</sub>

#### Single drop-out (SD)-medium

amount	component
6.7 g/l	Yeast Nitro Base
5 g/l	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>
20 g/l	glucose or galactose (added after autoclaving)
1 g/l	SC drop-out powder (added after autoclaving)

For SD-Agar, medium was prepared with additional 20 g/l of agar.

## MATERIALS

### Single drop-out powder

component	amount	component
2 g adenine hemisulfate	2 g	methionine
2 g arginine	3 g	phenylalanine
2 g histidine	6 g	homoserine
2 g isoleucine	3 g	tryptophan
4 g leucine	2 g	tyrosine
2 g lysine HCl	9 g	valine

### Antibiotics

antibiotic	stock-solution	final concentration
carbenicillin	100 mg/ml in H <sub>2</sub> O	100 µg/ml
kanamycin	50 mg/ml in H <sub>2</sub> O	25 µg/ml
gentamycin	20 mg/ml in H <sub>2</sub> O	20 µg/ml

### Tris buffered saline (TBS)

component	final concentration
NaCl	150 mM
Tris/HCl	50 mM

adjust to pH = 7.6 with HCL

## 4 METHODS

### 4.1 Statistical methods

#### 4.1.1 Student's t-test

Student's t-tests in this work were calculated with the Excel software (Microsoft, USA). All tests were two-sided type 3 tests.

#### 4.1.2 Tukey's test

Tukey's tests in this work were calculated with the OriginPro software (OriginLab, USA). For this, a one-way ANOVA test was conducted together with a Tukey's test at a significance level of 0.05.

### 4.2 Molecular biology methods

All molecular and microbiological methods in this chapter have been performed according to either Sambrook et al. (1989) or Ausubel et al. (1987), unless otherwise stated.

#### 4.2.1 Preparation of competent *E. coli* cells

Preparation of competent *E. coli* cells was done according to Inoue et al. (1990). A 5 ml overnight culture of the respective cells was inoculated directly from a respective agar plate. On the next day, 125 ml of LB-media was inoculated with the culture and incubated shaking at 180 rpm and 37 °C until the culture reached  $OD_{600} = 0.4-0.75$ . Culture was transferred to sterile 50 ml-falcon tubes and incubated on ice for 10 min. Afterwards, cells were centrifuged for 10 min at 4 °C and 1000 x g. Supernatant was removed and pellet was resuspended in 20 ml TFB-buffer per falcon. After 10 min of incubation on ice, falcons were centrifuged for 10 min at 4 °C and 1000 x g again. Supernatant was removed and the pellets were resuspended in 4 ml of TFB-buffer and combined in one falcon. After the addition of DMSO to a final concentration of 7 % (v/v), cells were further incubated on ice for 10 min. Finally, cells were divided in 200 µl-aliquots and flash-frozen in liquid nitrogen. Aliquots were stored at -80 °C.

#### TFB-buffer

amount	final concentration	component
0.605 g	10 mM	PIPES
0.441 g	15 mM	CaCl <sub>2</sub> x 2H <sub>2</sub> O
3.728 g	250 mM	KCl
	adjust pH to 6.7	
2.18 g	55 mM	MnCl <sub>2</sub> x 4H <sub>2</sub> O

## METHODS

### 4.2.2 Heat shock transformation of *E. coli*

Heat-shock transformation of *E. coli* was done according to Mandel and Higa (1970). Competent *E. coli* cells were mixed with DNA in a maximal ratio of 10:1 (v/v). Cell-DNA mixture was incubated on ice for 20 min. Afterwards, a heat-shock was performed at 42 °C for 45 sec. After further incubation on ice for two minutes, 600 µl of LB medium was added. Sample was incubated at 37 °C for 30 min in case of plasmids harbouring an ampicillin resistance and for 90 min in case of plasmids harbouring a kanamycin resistance. In case of retransformations, 100 µl of the sample was plated out on LB-plates containing the respective antibiotics. In all other cases, the complete sample was pelleted via centrifugation, supernatant was discarded and remaining pellet was plated out. Plates were incubated over night at 37 °C.

### 4.2.3 Preparation of competent *S. cerevisiae* cells

Preparation of competent *S. cerevisiae* cells was done according to Ito et al., a method which was further refined by Gietz and Schiestl (Gietz and Schiestl, 2007; Ito et al., 1983). For preparation of competent yeast cells, 50 ml of YPD medium was inoculated with a single colony of the respective strain and incubated over night at 30 °C and 180 rpm. Five millilitres of this culture were used to further inoculate 100 ml of YPD medium. The resulting culture was grown at 30 °C for approximately 4 h to an OD<sub>600</sub> of 0.6-0.7. Cells were pelleted via centrifugation at 4000 x g for 5 min and washed with 20 ml solution A. Afterwards, cells were resuspended in 4 ml solution A. Subsequently, 100 µl of a 10 mg/ml solution of heat-denaturated sperm-DNA and 100 µl of a 1 M histamine solution were added and mixed. Cells were divided in 200 µl aliquots and frozen at -70 °C.

#### Solution A

amount	component
1 M	sorbit
10 mM	tricine
3 % (w/v)	ethylene glycol

### 4.2.4 Transformation of *S. cerevisiae*

Transformation of *S. cerevisiae* was done according to Gietz and Schiestl (2007) and Ito et al. (1983). For transformations, 5 µg of plasmid-DNA was added to a 200 µl aliquot of frozen *S. cerevisiae* cells. Sample was incubated in a thermoshaker at medium speed and 37 °C for 5 min. Afterwards, cells were diluted in 1.2 ml solution B and further incubated at 30 °C for 1 h. Subsequently, cells were pelleted via centrifugation at 4000 x g for 1 min and washed three times with solution C. Cells were finally resuspended in 300 µl solution C and plated out on respective SD-agar-plates.

## METHODS

### Solution B

amount	component
40 %	PEG3350
200 mM	Tricine-NaOH

### Solution C

amount	component
0.15 mM	NaCl
10 mM	Tricine-NaOH

### 4.2.5 Standard PCR

Standard PCR was performed according to Mullis et al. (1986). Unless otherwise stated, standard PCRs were done using the Phusion-DNA polymerase (Finnzymes, Finland). Sample compositions and temperature gradients were done as follows:

amount	component	
0.5 µl	dNTP (each 10 µM)	98 °C 2 min
0.5 µl	Primer A (10 µM)	98 °C 1 min
0.5 µl	Primer B (10 µM)	25 x { 55 °C 30 sec.
~ 3 ng	DNA template	72 °C 30 sec / kb
1 U	DNA polymerase	72 °C 10 min.
ad 25 µl	H <sub>2</sub> O	

### 4.2.6 Colony PCR

Colony PCR was done according to Woodman (2008) in order to verify the successful ligation of a desired DNA-fragment in a respective vector. A PCR with primers specific for the respective vector was performed on the basis of whole cells as DNA-templates. A single *E. coli* colony was picked from an agar plate with a sterile toothpick. The cells were resuspended in a single PCR-sample and furthermore spread out on a LB-Agar plate supplemented with respective antibiotics for later use. The PCR sample was composed as described in 4.2.5, except for that GoTaq-DNA Polymerase (Promega, USA) was used. The temperature-gradient was as follows:

		94 °C 4 min
		94 °C 1 min
25 x	{	60 °C 30 sec.
		72 °C 1 min / kb
		72 °C 10 min.

## METHODS

### 4.2.7 Overlap extension PCR

Domain swap variants of MmAWAT2 and MmDGAT2, as described in 5.3.4, were created by overlap extension PCR according to Horton et al. (1989). Respective fragments were at first amplified from wild type sequences. All primers for overlap extension PCR were designed with a 20 bp-overhang corresponding to the adjacent fragment in case of forward-primers and a 15 bp overhang in case of reverse primers. Fragments were purified via gel-extraction and eluted in 30  $\mu$ l H<sub>2</sub>O. Of this eluate, 2  $\mu$ l of the respective shorter and 3  $\mu$ l of the respective longer fragment were used for overlap extension PCR using a standard PCR protocol and the Phusion DNA-Polymerase (Finnzymes, Finland).

### 4.2.8 Quick-Change PCR

All amino acid exchange variants used in this study were created by site directed mutagenesis (Hutchison et al., 1978). Respective primers listed in Table 9 were designed as follows: ~10 bases in direction of 5'-end of the mutation, ~20 bases in direction of 3'-end of the mutation, ~10 non-overlapping bases at the 3'-end of both primers. Phusion DNA polymerase (Finnzymes, Finland) was used to ensure correct amplification. Temperature gradient and sample composition was used as follows:

amount	component		
			94 °C 4 min
0.5 $\mu$ l	dNTP (each 10 $\mu$ M)		
0.5 $\mu$ l	Primer A (10 $\mu$ M)		94 °C 1 min
0.5 $\mu$ l	Primer B (10 $\mu$ M)	18 x {	55 °C 30 sec.
~ 3 ng	DNA template		72 °C 1 min / kb
1 U	DNA polymerase		
ad 50 $\mu$ l	H <sub>2</sub> O		72 °C 10 min.

After PCR amplification of the respective plasmid, sample was incubated with 10 U of DpnI at 37 °C over night in order to eliminate methylated template-DNA.

### 4.2.9 Restriction of DNA

DNA restriction was performed by utilization of DNA-restriction endonucleases, as reviewed by Boyer (1971) in the buffer recommended by the respective manufacturer and according to the manufacturer's protocol.

## METHODS

### 4.2.10 Ligation of DNA

Ligation of DNA was done by utilization of the T4-ligase (Weiss and Richardson, 1967) according to the instructions provided by the manufacturer.

### 4.2.11 Plasmid preparation from *E. coli*

For plasmid preparation, a respective 2 ml *E. coli* overnight-culture was inoculated with a single colony and supplemented with required antibiotics. On the next day, plasmid-DNA was prepared using the NucleoSpin™ Plasmid-Kit (Table 4) according to the instructions provided by the manufacturer. The procedure is based on a publication of Birnboim and Doly (1979).

### 4.2.12 Separation of DNA by agarose gel-electrophoresis

This method is based on a principle published by Aaij and Borst (1972). DNA-samples were mixed with an appropriate volume of 6x loading dye. Preparation of agarose-gels was done by boiling 1 % agarose (*w/v*) in TAE-buffer. Electrophoresis was performed for approximately 20 min, until the bromophenol-blue running front left the gel. Afterwards, gel was incubated in TAE supplemented with 2 µg/ml of ethidium bromide. DNA was visualised in UV-light.

TAE-buffer		6x loading dye	
amount	component	amount	component
40 mM	Tris/HCl, pH 7.0	40 mM	Tris/acetate pH 8.5
20 mM	acetic acid	100 mM	EDTA
1 mM	EDTA	0.1 % ( <i>w/v</i> )	SDS
		50 % ( <i>v/v</i> )	glycerol
		0.25 % ( <i>w/v</i> )	xylencyanol blue
		0.25 % ( <i>w/v</i> )	bromophenol blue

### 4.2.13 Cultivation of *E. coli* Lemo21 (DE3) for protein expression screenings

Lemo21 (DE3) cells were used to screen for optimal expression conditions for pMAL\_MmFAR1 and pGEX\_DmFAR. Lemo21 cells contain a plasmid which codes for the natural T7 RNA polymerase inhibitor T7 lysozyme, which is expressed from a titratable rhamnose promoter. It is thus possible to adjust the level of T7 RNA polymerase in order to control expression conditions.

For screening experiments, two times 10 ml expression cultures were supplemented with either 0, 100, 250, 500, 750, 1000 or 2000 µM of L-rhamnose, appropriate antibiotics and 200 µl of an overnight culture. Cultures were grown at 30 °C and 180 rpm to an OD<sub>600</sub> of 0.4-0.8. One of each culture was then induced with 400 µM IPTG and further incubated at 30 °C and 180 rpm for four

## METHODS

additional hours. Afterwards, OD<sub>600</sub> of the culture was measured. Samples were taken according to the formula  $[0.5/OD_{600} = \text{ml sample}]$ . Samples were pelleted, resuspended in 50  $\mu\text{l}$  1x Laemmli buffer and incubated at 95 °C for 5 min. 10  $\mu\text{l}$  of the samples were applied on a SDS-gel.

### 4.2.14 Cultivation of *E. coli* expression cultures for protein purification

For cultivation of *E. coli* expression cultures, 50 ml of 2xYT-medium supplemented with the appropriate antibiotics was inoculated with a single colony and incubated shaking at 200 rpm over night at 37 °C. On the next day, 1 l of 2xYT-medium, supplemented with appropriate antibiotics, was inoculated with 50 ml of an overnight culture. Culture was incubated shaking in 2 l-flasks at 200 rpm and 37 °C for 60-120 min, until the culture reached an OD<sub>600</sub> of 0.4-0.8. Protein expression was induced by the addition of 400  $\mu\text{M}$  IPTG. The culture was further incubated at 200 rpm and 16 °C for ~62 h, unless stated otherwise. Cultures were harvested by centrifugation at 3000 x g and resulting pellets were flash frozen in liquid nitrogen and stored at -20 °C unless stated otherwise.

### 4.2.15 Cultivation of *E. coli* for production of 6xHis-TF-AbWSD1 selenoprotein

Production of the 6xHis-TF-AbWSD1 selenoprotein was done according to Ausubel et al. (2001) and performed by Dr. Karin Kühnel with assistance of Felix Lambrecht. For the production of 6xHis-TF-AbWSD1 carrying selenomethionine instead of normal methionine, cultures were grown in a special M9 minimal medium supplemented with the appropriate antibiotics.

For expression, 250 ml of culture were incubated in a 1 l baffled flask in order to ensure good aeration of the culture. Flask and medium were equilibrated to 37 °C before inoculation. For inoculation of the medium, 1 ml of an overnight culture, grown in LB-medium, was pelleted, resuspended in M9-medium and used for inoculation of 250 ml M9 medium. Culture was incubated shaking at 180 rpm and 37 °C until it reached an OD<sub>600</sub> of 0.3. Solid amino-acid supplements were added to the cultures to the final concentrations:

amount	component
100mg/l	L-Lysine
100mg/l	L-Phenylalanine
100mg/l	L-Threonine
50mg/l	L-Isoleucine
50mg/l	L-Leucine
50mg/l	L-Valine
50mg/l	L-Selenomethionine

Protein expression was induced after 15 min of amino acid supplementation using 1 mM IPTG. Growth was continued for additional 6 h. Cells were harvested by centrifugation at 3000 x g.

#### 4.2.16 Cultivation of *S. cerevisiae*

Cultivation of *S. cerevisiae* expression cultures was done in accordance to Heilmann et al. (2012). In detail, 20 ml of SD-URA-media + 2 % galactose was inoculated to an  $OD_{600} = 0.05$  from a preculture. Precultures were grown in SD-URA + 2 % galactose at 30 °C and 180 rpm for 48 h. Expression cultures were incubated for five days at 30 °C and 180 rpm.

### 4.3 Analytical methods

#### 4.3.1 WE extraction from *S. cerevisiae* and *E. coli*

WE extraction from *S. cerevisiae* or *E. coli* was done similar as described by Heilmann et al. (2012). In detail, cells corresponding to 50  $OD_{600}$ -units were harvested by centrifugation at 1000 x g for 10 min. After centrifugation, 1 ml methanol was added to the cell-pellets together with 0.5 mm glass beads. The samples were vortexed for 15 min at room temperature (RT). Subsequently, 2 ml hexane was added, samples were vortexed for another 15 min at RT. The upper phase was removed, evaporated under  $N_2$  and resolved in 200  $\mu$ l of hexane. Sample was transferred to GC-vials and stored at -20 °C for further analyses.

#### 4.3.2 Fatty alcohol extraction from *S. cerevisiae* and *E. coli*

Fatty alcohol extraction from *S. cerevisiae* or *E. coli* was performed analogous to 4.3.1. The only difference was that the final extracts were not resolved in 200  $\mu$ l of hexane, but in 100  $\mu$ l of acetonitrile. From these extracts, 5  $\mu$ l were mixed with 5  $\mu$ l of BSTFA and analysed via GC-MS. Alternatively, alcohols were analysed non-derivatised.

#### 4.3.3 Thin layer chromatography analysis

Thin layer chromatography of WEs and fatty alcohols was done as described previously (Heilmann et al., 2012). For TLC analyses of lipid extracts, 25  $\mu$ l sample were applied on TLC plates with the help of an AUTOMATIC TLC SAMPLER 4 (CAMAG, Switzerland). The TLC plate was developed in a glass chamber containing hexane:diethyl ether:acetic acid (80:20:1, v/v/v) as a running solvent in case of analyses of WEs and hexane:diethyl ether:acetic acid (65:35:1, v/v/v) in case of analyses of fatty alcohols. Separated lipids were visualised by soaking the plates with  $CuSO_4$  and subsequent heating to 190 °C. Spots were assigned to different lipid classes according to their migration behaviour in comparison to respective standard substances. Standard-substances were applied at amounts of 50  $\mu$ g each. For WE, heptadecanoyl-heptadecanoate was used, while for steryl ester steryloleate, for fatty alcohols hexadecanoyl-alcohol, for fatty acids octadecanoic acid and for triacylglycerols olive oil were applied on respective TLC plates.

#### 4.3.4 GC-FID analysis

For relative quantification of WE species, yeast WE preparations were prepared according to section 4.3.1. Of these samples, 2  $\mu\text{l}$  were subjected to GC FID using a 6890 Series GC System equipped with an Agilent 19091j-413 HP5 5 % Phenyl Methyl Siloxane column (30 m x 320  $\mu\text{m}$  x 0.25  $\mu\text{m}$  film thickness; Agilent, Santa Clara, CA, USA). Helium was used as carrier gas (1.5 ml min<sup>-1</sup>). The split-ratio was 5:1. The temperature gradient was 2 min at 60 °C, 60 – 200 °C at 40 K min<sup>-1</sup>, 2 min at 200 °C, 200–325 °C at 3 K min<sup>-1</sup> and 325 °C for 16 min. WEs were identified according to their characteristic retention time as elucidated by GC-MS analyses. Detection was performed by flame ionization detection.

#### 4.3.5 WE analysis via GC-MS

WE analysis via GC-MS was done according to Heilmann et al. (2012). Lipid extracts from yeast were prepared according to section 4.3.1. Of these samples, 2  $\mu\text{l}$  were subjected to GC-MS analysis using a Polaris Q mass selective detector connected to a Trace gas chromatograph (Thermo Finnigan, Austin, TX, USA) equipped with a Restek Rxi™-5ms capillary column (15 m x 0.25 mm, 0.25  $\mu\text{m}$  film thickness; Restek, Bellefonte, PA, USA). Helium was used as carrier gas (1.5 ml min<sup>-1</sup>). The temperature gradient was 2 min at 60 °C, 60–200 °C at 40 K min<sup>-1</sup>, 2 min at 200 °C, 200 – 340 °C at 3 K min<sup>-1</sup> and 340 °C for 16 min. The WEs were detected by electron impact ionization (-70 eV, ion source 200 °C, Aux-line 350 °C) in a mass range of 50–730 amu.

#### 4.3.6 Wax ester molecular species profiling by nano-ESI-MS/MS

Wax ester molecular species profiling by nano-ESI-MS/MS was performed by Dr. Tim Iven (Department of Plant Biochemistry, University of Göttingen, Germany). For nano ESI-MS/MS wax ester molecular species profiling, the protocol described previously was followed (Iven et al., 2013). Briefly, wax esters were isolated from a crude lipid extract by preparative thin layer chromatography (TLC). For the measurement, wax esters were dissolved in methanol:chloroform (2:1, v/v) containing 5 mM ammonium acetate. The analysis was performed using an Applied Biosystems 3200 hybrid triple quadrupole/linear ion trap mass spectrometer (ABSciex, Darmstadt, Germany). Direct injection nano-ESI analysis was achieved using a chip ion source (TriVersa NanoMate; Advion BioSciences, Ithaca, NY, USA). 10  $\mu\text{l}$  of wax ester extract were subjected to nano-ESI in positive ionization mode with ionization voltage of 1.5 kV and backpressure of 0.4 psi. The mass analysers were adjusted to a resolution of 0.7 amu full width at half-height. Peak intensities of 785 MRM transitions, corresponding to even chain wax ester molecular species with acyl moieties of C16–C24 containing 0–3 double bonds and C26–C32 with 0–1 double bond, were collected with the Analyst 1.5.1 software (AB Sciex, Darmstadt, Germany). Signal intensities were extracted from the raw data file with the LipidView software (AB Sciex, Darmstadt, Germany) and a type I <sup>13</sup>C isotope correction of raw intensities was performed.

## METHODS

### 4.3.7 Fatty alcohol analysis via GC-MS

Identification of fatty alcohol species was performed by GC-MS as described by Heilmann et al. (2012). Lipid extracts from *S. cerevisiae* or *E. coli* were prepared according to section 4.3.2. Of these samples, 2  $\mu\text{l}$  were subjected to GC-MS analyses using an Agilent 5973 mass selective detector connected to an Agilent 6890 gas chromatograph equipped with a capillary DB-23 column (30m x 0.25 mm; 0.25 mm coating thickness; J&W Scientific, Agilent, Waldbronn, Germany). Alternatively, a Polaris Q mass selective detector connected to a Trace gas chromatograph (Thermo Finnigan, Austin, TX, USA) equipped with the same column was used. Helium was used as carrier gas at a flow rate of 1 ml min<sup>-1</sup>. The temperature gradient was 150 °C for 1 min, 150 – 200 °C at 8 K min<sup>-1</sup>, 200 – 250 °C at 25 K min<sup>-1</sup> and 250 °C for 6 min. For the MS analysis, an electron energy of 70 eV, an anion source temperature of 230 °C and a temperature of 260 °C for the transfer line were used. The ions were detected in scan mode in an m/z range from 50 to 650. Fatty alcohols were identified using the NIST MS Search 2.0 library. Accuracy of the NIST-hits was further verified based on published mass spectra of TMS-derivatised fatty alcohols as shown in the AOCS lipid library (<http://lipidlibrary.aocs.org/>).

### 4.3.8 Extraction and analysis of acyl-CoA from *S. cerevisiae*

Acyl-CoA extraction was done according to Rosendal and Knudsen (1992). Yeast cells corresponding to 80 OD<sub>600</sub> units were harvested and washed with 10 mM of perchloric acid. Afterwards, 10  $\mu\text{l}$  of 50  $\mu\text{M}$  heptadecanoyl-CoA (17:0-CoA) was added as an internal standard. After addition of 800  $\mu\text{l}$  H<sub>2</sub>O, cells were extracted by vortexing for 10 min with 3 ml chloroform/methanol mixture (2:1, v/v) and 0.5 mm glass beads. Afterwards, cells were further mixed at 4 °C on a shaker at full speed for 30 min. After adding of 1 ml chloroform and 1 ml H<sub>2</sub>O, vortexing for 30 sec and subsequent centrifugation, aqueous and organic phase were discarded, while the interphase was dried under a stream of nitrogen. After addition of 400  $\mu\text{l}$  extraction buffer, 10  $\mu\text{l}$  saturated (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and 1.2 ml methanol-chloroform mixture (2:1, v/v), samples were vortexed and incubated for 20 min at room temperature. After centrifugation to remove debris, supernatant was dried under nitrogen and subsequently derivatised with 200  $\mu\text{l}$  derivatization buffer at 85 °C for 20 min. Samples were then cooled down to 4 °C for 30 min, 15  $\mu\text{l}$  of the sample were further analysed by HPLC.

Extraction buffer		Derivatization buffer	
amount	component	amount	component
2 ml	isopropanol	0.5 M	chloroacetaldehyde
2 ml	KH <sub>2</sub> PO <sub>4</sub> , pH = 7.2	0.15 M	citrate buffer, pH = 4.0
50 $\mu\text{l}$	acetic acid	0.5 % (w/v)	SDS
80 $\mu\text{l}$	BSA (50 mg/ml in H <sub>2</sub> O, defatted)		

The samples were analysed on a LC Agilent 1100 Series-HPLC machine, equipped with a LUNA 150 x 2.0 mm column. Machine was operated by Sabine Freitag (Department of Plant Biochemistry, University of Göttingen, Germany). Elution of the synthesised acyl-etheno-CoA esters was done under basic conditions in the presence of triethylamine using a quaternary pump system. Detection

## METHODS

was done by excitement of the fluorescent esters at 230 nm, while the read out of emitted light was performed at 420 nm. Running solvents as well as gradients used for the analyses of acyl-CoA are listed in Table 11 and

Table 12, respectively. The acyl-CoA species were assigned to their molecular identity by coelution with known standard substances. In case of 24:1-CoA and 26:1-CoA, no standards were available and the molecular nature was concluded in analogy to the typical elution pattern of 16:1-CoA - 22:1-CoA (Figure 37 C).

**Table 11: Running solvents for acyl-CoA analyses.**

MeCN = acetonitrile, TEA = triethanolamine.

running solvent	composition (per volume)
A	100:1 H <sub>2</sub> O:CH <sub>3</sub> COOH
B	90:9:1 MeCN:H <sub>2</sub> O:CH <sub>3</sub> COOH
C	100:0,25 H <sub>2</sub> O:TEA
D	90:10 MeCN:H <sub>2</sub> O

**Table 12: Gradient for acyl-CoA analyses**

time (min)	% B	% C	% D	flow rate (ml/min)	pressure (bar)
0	10	0	0	0.4	300
5	80	0	0	0.4	300
5.1	0	80	0	0.4	300
7	0	97	3	0.4	300
10	0	95	5	0.4	300
10.1	0	95	5	0.4	300
50	0	55	45	0.4	300
51.1	0	0	100	0.4	300
52	0	0	100	0.4	300
62	0	0	100	0.4	300
62.1	10	0	0	0.4	300

### 4.3.9 Thermal shift assay (Thermofluor)

Thermofluor analyses were done according to Ericsson et al. (2006) at the Department of Molecular Structural Biology, University of Göttingen, Germany with the help of Dr. Achim Dickmanns (Department for Molecular structural Biology, University of Göttingen, Germany). In detail, 16 µl of 1 µM protein solution was mixed with 2 µl of 50 mM SYPRO Orange (Sigma, Germany) and 2 µl of different 0.1 M buffer solutions in a 96-well plate. After pipetting, 96-well plate was briefly centrifuged to ensure mixing of all components and analysed in a CFX96 realtime PCR cycler (BioRad, Germany). The heat gradient of the cycler was as follows: 20-95 °C in steps of 1 K, each step was hold for 30 sec each. Fluorescence intensities were recorded at 570 nm. Data was evaluated with the MaxFit-software (Department for Cryo Electron Microscopy, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany).

### 4.3.10 Multiangular light scattering (MALS)

MALS measurements were done at the Department for Molecular structural Biology, University of Göttingen, Germany by Dr. Thomas Monecke (Department for Molecular structural Biology,

## METHODS

University of Göttingen, Germany) and Felix Lambrecht. For analyses, a miniDAWN machine (Wyatt, Germany) was connected to an ÄKTA purifier equipped with a Hiload 16/60 superdex 200 SEC column (GE, USA). The SEC column was equilibrated with 20 mM glycine/NaOH pH = 10.0, 150 mM NaCl, 10 % glycerol until the scattering signal reached a stable baseline. Afterwards, protein sample was loaded onto the column with a flow rate of 0.5 ml min<sup>-1</sup>. Obtained scattering data were analysed using the Wyatt software (Wyatt, Germany).

### 4.4 Biochemical methods

#### 4.4.1 Preparation of cell pellets for protein purification

Frozen *E. coli* cell pellets were diluted with 50 ml of buffer A. Pellets were resuspended by vigorous vortexing. Small amounts of DNase and lysozyme as well as 1 mM PMSF were added. Pellet was incubated shaking at 4 °C for 30 min. Subsequently, suspension was applied to sonication for 15 x 1 min, with 1 min of break between the cycles. During sonication, cell-suspension was incubated on ice. Afterwards, sample was centrifuged for 20 min at 15000 x g in order to pellet cell debris. Supernatant was recovered and used for loading of a respective affinity-chromatography column. In case of purification of AbWSD1, cell disruption was done by applying the samples to a fluidizer for 3-5 cycles.

#### 4.4.2 Protein purification

For purification of heterologously produced protein from *E. coli*, the principle of affinity chromatography (AC) was used throughout the study. The coding sequences of respective proteins were cloned in to several *E. coli* expression vectors (Table 16), resulting in the production of the proteins fused to different tags. For actual purifications, either the 6xHis-tag or the MBP-tag was used. 6xHis-tagged proteins were immobilised at Ni<sup>2+</sup>-NTA-agarose resin, while MBP-tagged proteins were bound on Dextrin Sepharose High Performance resin (respective columns listed in Table 8). All protein purifications were either performed on an ÄKTAprime™ plus or an ÄKTAFPLC™ system (GE Healthcare, USA).

##### 4.4.2.1 Amylose resin affinity chromatography

MBPTrap-columns (GE Healthcare, USA) were equilibrated with buffer A until a stable baseline was formed. For detailed composition of buffer A see tables depicted on page 54-56. Cell-lysate was loaded onto the columns with a flowrate of 1 ml/min. Column was subsequently washed with buffer A containing 1 M of NaCl in order to elute loosely bound proteins. Afterwards, bound protein was eluted with buffer A containing 10 mM of maltose.

##### 4.4.2.2 Ni<sup>2+</sup>-based immobilised metal affinity chromatography

## METHODS

Ni-NTA-agarose columns were equilibrated with buffer A until a stable baseline was formed. For detailed composition of buffer A see tables depicted on page 54-56. Cell-lysate was loaded onto the column with a flowrate of 1 ml/min. Afterwards, unbound protein was removed from the column by washing with buffer A. Loosely bound protein was further removed by washing the column with 5 % of buffer B. In general, buffer B was composed as buffer A with the addition of 500 mM imidazole. Elution of the bound protein was performed with 50 % buffer B. To make sure that all proteins eluted, column was furthermore washed with 100 % of buffer B in the end. Buffer compositions for all proteins are summarised in the tables shown below.

### 4.4.2.3 Affinity chromatography for separation of TF and fusion partner

250 µg of 6xHis-TF-fusion partner were incubated with 10 U thrombin at 37 °C for 4 h. The sample was subsequently incubated with Ni-NTA agarose, which itself was equilibrated with one of the respective buffer depicted in Table 15. Sample was incubated spinning at 4 °C over night. On the next morning, resin with bound sample was poured into a gravity flow column, supernatant was recovered. Resin was washed with respective buffer A in order to elute non-bound fusion partner. Bound TF was eluted with buffer A containing 500 mM imidazole.

4.4.2.4 Buffer compositions for purification of MaFAR1 from *M. aquaeolei*

Table 13: Different conditions for the purification of MaFAR1 tested during this study.

construct	growing conditions	buffer affinity chromatography (IMAC, unless stated otherwise)	result affinity chromatography	buffer SEC	result SEC
pCold_MaFAR	2 d @ 16 °C	TBS, 10 % glycerol, 0.1 % Triton X-100	unpure eluate	TBS, 10 % glycerol, 0.1 % Triton X-100	aggregates
		TBS, 10 % glycerol, 0.3 % CHAPS	nearly pure eluate	TBS, 10 % glycerol, 0.3 % CHAPS	aggregates
		TBS, 10 % glycerol, 0.1 % Triton X-100	highly unpure eluate	500 mM HEPES pH = 7, 0.1 % Triton X-100	aggregates + highly unpure peak at ~ 250 kDa
pMAL_MaFAR	2 d @ 16 °C	50 mM Tris/HCl pH = 8, 300 mM NaCl, 10 % glycerol, 0.1 % Triton X-100	highly pure eluate	500 mM HEPES pH = 7, 0.1 % Triton X-100	aggregates
		50 mM Tris/HCl pH = 8, 300 mM NaCl, 10 % glycerol, 0.3 % CHAPS	highly pure eluate	50 mM HEPES pH = 7, 0.3 % CHAPS	aggregates
		20 mM Tris/HCl pH = 7, 50 mM NaCl, 1 mM EDTA (dextrin sepharose)	highly pure eluate	x	x
pET24_MaFAR	4 h @ 30 °C	100 mM PO <sub>4</sub> pH = 7, 40 mM imidazole	highly pure eluate	100 mM PO <sub>4</sub> pH = 7	aggregates
		50 mM Tris/HCl pH = 8, 300 mM NaCl, 10 % glycerol, 0.3 % CHAPS, 30 mM imidazole	no eluate	100 mM PO <sub>4</sub> pH = 7, 10 % glycerol, 0.3 % CHAPS	aggregates
		50 mM Tris/HCl pH = 8, 300 mM NaCl, 10 % glycerol, 0.3 % CHAPS, 30 mM imidazole	highly pure eluate	50 mM NaCitrate pH = 5.0, 10 % glycerol	aggregates
	2 d @ 16 °C	50 mM Tris/HCl pH = 8, 150 mM NaCl, 10 % glycerol, 0.3 % CHAPS, 30 mM imidazole	highly pure eluate	TBS, 10 % glycerol	aggregates
		50 mM Tris/HCl pH = 8, 150 mM NaCl, 10 % glycerol, 0.3 % CHAPS, 30 mM imidazole	highly pure eluate	TBS, 10 % glycerol	aggregates
		30 mM CHES pH = 9.5, 150 mM NaCl, 1 mM EDTA, 5 mM MgCl	unpure eluate <sup>1</sup>	CHES pH = 9, 150 mM NaCl, 10 % glycerol, 50 mM arginine, 50 mM glutamate	aggregates + 2 peaks at ~ 600 kDa, ~ 161 kDa and a shoulder at ~ 87 kDa
		50 mM Tris/HCl pH = 8, 500 mM NaCl, 10 % glycerol	highly pure eluate	30 mM Tris/HCl pH = 8, 500 mM NaCl,	aggregates + peak at ~ 600 kDa

<sup>1</sup> not eluted with imidazole, but 50 mM arginine and glutamine

METHODS

4.4.2.5 Buffer compositions for purification of DmFAR1 from *D. melanogaster*

Table 14: Buffer compositions for purification of heterologously produced DmFAR1. SEC = size exclusion chromatography.

construct	growing conditions	buffer affinity chromatography	result affinity chromatography	buffer SEC	result SEC
pET24_DmfAR	4 h @ 30 °C	50 mM Tris/HCl pH = 8	no eluate	x	x
		300 mM NaCl, 10 % glycerol 0.3 % CHAPS, 30 mM imidazole			
	2 d @ 16 °C	50 mM Tris/HCl pH = 8	smearly eluate	50 mM Tris/HCl pH = 8 150 mM NaCl, 10 % glycerol 0.3 % CHAPS	peaks corresponding to 445 kDa 170 kDa 114 kDa
		500 mM NaCl, 10 % glycerol 0.3 % CHAPS, 30 mM imidazole			
pET28_DmfAR	4 h @ 30 °C	50 mM Tris/HCl pH = 7.6	little eluate, smearly band underrepresented but present	x	x
		150 mM NaCl, 10 % glycerol 0.3 % CHAPS, 30 mM imidazole			
	2 d @ 16 °C	50 mM Tris/HCl pH = 8	no eluate	x	x
		300 mM NaCl, 10 % glycerol 0.3 % CHAPS, 30 mM imidazole			
pCold_DmfAR	2 d @ 16 °C	50 mM Tris/HCl pH = 8	unpure eluate	x	x
		500 mM NaCl, 10 % glycerol 0.3 % CHAPS, 30 mM imidazole			
	4 h @ 30 °C	50 mM Tris/HCl pH = 7.6	nearly pure eluate	50 mM HEPES pH = 7.0 500 mM NaCl	aggregates
		300 mM NaCl, 10 % glycerol 0.3 % CHAPS, 30 mM imidazole			
pMAL_DmfAR	4 h @ 30 °C	50 mM Tris/HCl pH = 7.6	little, unpure eluate	x	x
		200 mM NaCl, 10 % glycerol			
	2 d @ 16 °C	50 mM Tris/HCl pH = 7.6	nearly pure eluate	x	x
		200 mM NaCl, 10 % glycerol 0.05 % CHAPS			
4 h @ 30 °C	50 mM Tris/HCl pH = 7.6	nearly pure eluate	50 mM HEPES pH = 7.0 500 mM NaCl	aggregates	
	200 mM NaCl, 10 % glycerol 0.3 % CHAPS				

## METHODS

### 4.4.2.6 Buffer compositions for separation of 6xHis-TF and its fusion partner

**Table 15: Buffer compositions for separation of 6xHis-TF and its fusion partner AbWSD1.**

chromatography	column	buffer composition	elution with (length of gradient)	result
anion exchange chromatography	Mono S	50 mM Na-acetate pH = 5.0	0 - 1 M NaCl (80 ml)	insufficient binding on column
		50 mM Na-citrate pH = 5.0	0 - 1 M NaCl (50 ml)	no binding on column
		50 mM Tris/HCl pH = 7.0	0 - 1 M NaCl (50 ml)	co-elution
		50 mM Tris/HCl pH = 8.0	0 - 1 M NaCl (40 ml)	co-elution
		50 mM Bis-Tris pH = 6.0	0 - 1 M NaCl (30 ml)	no binding on column
		50 mM HEPES pH = 7.0	0 - 1 M KCl (25 ml)	co-elution
cation exchange chromatography	Mono Q	50 mM Tris/HCl pH = 8.0	step-gradient from 0 - 1 M NaCl	separation of TF and AbWSD1, but AbWSD1 fraction contaminated with other proteins
		50 mM Tris/HCl pH = 8.0	step-gradient from 0 - 1 M NaCl	co-elution
		100 mM Tris/HCl pH = 8.0, 0.1 % Triton X-100	0 - 1 M NaCl (50 ml)	co-elution
		50 mM Tris/HCl pH = 8.4	0 - 1 M NaCl (30 ml)	co-elution
		50 mM Tris/HCl pH = 7.0 1.5 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1.5 M - 0 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> (40 ml)	co-elution
		50 mM Tris/HCl pH = 7.5 150 mM NaCl 30 mM imidazole, 0.1 % triton X-100	+ 500 mM imidazole	co-elution
hydrophobic interaction chromatography	Source 15 Phe	50 mM Tris/HCl pH = 7.5 1 M NaCl 30 mM imidazole, 0.1 % triton X-100	+ 500 mM imidazole	co-elution
		50 mM Tris/HCl pH = 7.5 1 M NaCl 30 mM imidazole, 0.1 % triton X-100	+ 500 mM imidazole	co-elution
		50 mM Tris/HCl pH = 7.5 500 mM NaCl 30 mM imidazole, 0.1 % triton X-100	+ 500 mM imidazole	co-elution
		50 mM Tris/HCl pH = 7.5 700 mM NaCl 30 mM imidazole, 0.1 % triton X-100	+ 500 mM imidazole	co-elution
		50 mM Tris/HCl pH = 7.5 150 mM NaCl 30 mM imidazole, 0.1 % triton X-100	+ 500 mM imidazole	co-elution
		50 mM Tris/HCl pH = 7.5 150 mM NaCl 30 mM imidazole, 0.1 % triton X-100	+ 500 mM imidazole	co-elution
affinity chromatography	Ni-NTA Agarose, gravity-flow column	50 mM Tris/HCl pH = 7.5 150 mM NaCl 30 mM imidazole, 0.1 % triton X-100	+ 500 mM imidazole	co-elution
		50 mM Tris/HCl pH = 7.5 150 mM NaCl 30 mM imidazole, 0.1 % triton X-100	+ 500 mM imidazole	co-elution

## METHODS

For IMAC purification of 6xHis-TF-AbWSD1, freshly harvested, non-frozen cell pellets were resuspended in 50 mM Tris/HCl pH = 7.6, 150 mM NaCl, 30 mM imidazole, 10 % glycerol (buffer A). Cells were prepared according to 4.4.1 and disrupted by fluidizing. After protein loading and washing of the Ni-NTA column as described in 4.4.2.2, bound proteins were eluted with buffer A containing 500 mM imidazole (buffer B). Fractions containing 6xHis-TF-AbWSD1 were concentrated via spin-concentrators and applied to SEC. Initially, 50 mM Tris/HCl pH = 7.6, 150 mM NaCl, 10 % glycerol was used for SEC. After Thermofluor-measurements, the optimised buffer containing 20 mM glycine/NaOH pH = 10.0, 150 mM NaCl, 10 % glycerol was used throughout.

### 4.4.2.7 Buffer compositions for purification of AtDGA3 from *A. thaliana*

Initial IMAC purifications of 6xHis-TF-AtDGAT3 and 6xHis-AtDGAT3 was done in TBS, 10 % glycerol, 0.3 % CHAPS as buffer A. Buffer B was composed as buffer A plus 500 mM imidazole. Better results were achieved with 30 mM HEPES/HCl pH = 7.5, 500 mM NaCl, 10 % glycerol and 30 mM imidazole as IMAC buffer A. Therefore, the HEPES buffer was used in all AtDGAT3 IMAC purifications except for the initial ones. IMAC Buffer B was adapted depending on the following chromatography:

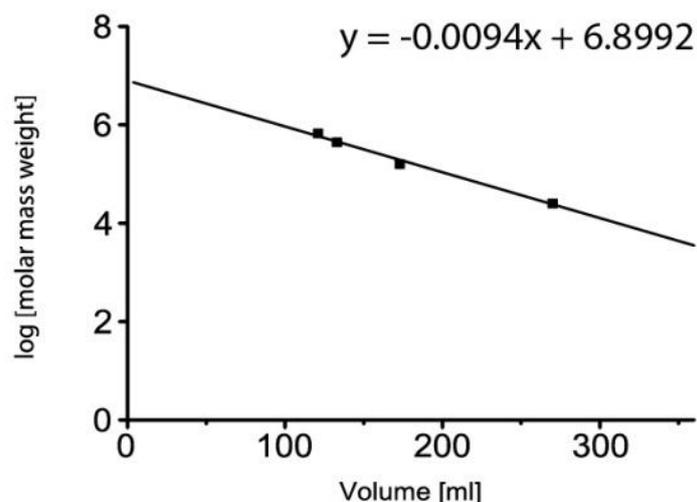
In case of further purification of the IMAC-derived eluates via hydrophobic interaction chromatography (HIC), IMAC buffer B contained 30 mM HEPES/HCl pH = 7.5, 1 M NaCl, 10 % glycerol and 500 mM imidazole. For HIC, IMAC-derived samples were further diluted 10 times in 30 mM HEPES/HCl pH = 7.5, 2 M NaCl, 10 % glycerol (HIC buffer A) and loaded on a on a 1 ml HiTrap PhenylHP column (GE Healthcare, USA). Bound protein was eluted with a continuous gradient of 20 ml length from HIC buffer A to HIC buffer B (30 mM HEPES/HCl pH = 7.5, 10 % glycerol).

In case of subsequent utilization of the protein in activity assays, IMAC buffer B consisted of 30 mM HEPES/HCl pH = 7.5, 500 mM NaCl, 10 % glycerol and 500 mM imidazole.

### 4.4.3 Size exclusion chromatography (SEC)

Determination of the enzymes' multimeric states as well as further purification of affinity chromatography-derived proteins was achieved by size exclusion chromatography (SEC). The SEC column was prepared by washing with one column volume of H<sub>2</sub>O followed by one column volume of SEC-buffer. The sample was loaded onto the column with a flowrate of 1 ml/min. Fractions were collected in a size of 5 ml, starting at an elution-volume of 90 ml after sample injection. Molecular weight of eluted protein was determined according to the equation shown in Figure 10.

## METHODS



**Figure 10: Standard curve for determination of molecular weights according with SEC.** Standard curve was determined by Dr. Florian Brodhun (Department of Plant Biochemistry, University of Göttingen, Germany) according to elution volumes of HiLoad 26/60 Superdex S200 SEC-column.

### 4.4.4 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was done according to Laemmli (1970). For SDS-PAGE, protein-samples were mixed with Laemmli buffer (Newie, Germany) and denaturated for 5 min at 95 °C. Of the denaturated samples, 10 µl per well were loaded onto the gels. Additionally, 5 µl of a molecular weight marker was loaded in one of the wells (Table 4). SDS-PAGE was run on Mini-PROTEAN3 electrophoresis Systems (Bio-Rad, Germany) at 40 mA and 300 V until the bromophenol-blue front migrated out of the gel. Proteins were stained in Coomassie-staining solution for 2 h at RT while shaking and unstained with destaining solution (composed as staining solution, but without Coomassie). Additionally, gels were supplemented with 2,2,2-trichloroethanol (TCE) according to Ladner et al. (2004) in order to achieve a first, fast visualization of the separated proteins in the gel. Respective gels were developed for 2 min under UV-radiation at 312 nm. Subsequently, gels were exposed to UV-light for another 20 sec while the image was recorded upon utilization of the SybrGreen-filter in a Diana machine (Raytest, Germany).

#### Stacking gels (4.8 %)

amount	component
5.95 ml	ddH <sub>2</sub> O
1.34 ml	acrylamide (30 %, w/v)
2.50 ml	0.5 M Tris/HCl, pH = 6.8
100 µl	SDS (10 %, w/v)
40 µl	APS (25 %, w/v)
10 µl	TEMED

#### Separation gels (10 %)

amount	component
6.35 ml	ddH <sub>2</sub> O
5.35 ml	acrylamide (30 %, w/v)
4.00 ml	0.5 M Tris/HCl, pH = 6.8
160 µl	SDS (10 %, w/v)
64 µl	APS (25 %, w/v)
16 µl	TEMED
120 µl	TCE

## METHODS

### Coomassie-staining solution

amount	component
50 % (v/v)	H <sub>2</sub> O
40 % (v/v)	methanol
10 % (v/v)	acetic acid
0.5 % (w/v)	Coomassie Brilliant Blue G250

### SDS-PAGE buffer

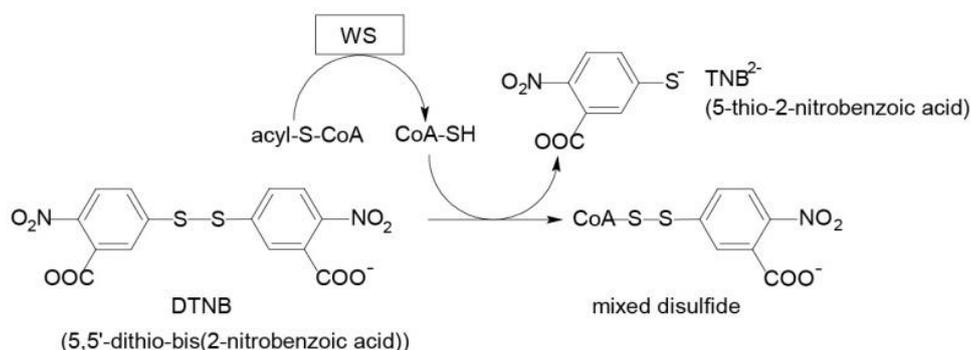
amount	component
25 mM	Tris/HCl, pH = 8.0
200 mM	glycine
0.1 % (w/v)	SDS

### 4.4.5 Determination of protein concentrations according to Bradford

Determination of protein concentrations was done according to Bradford (1976). A respective calibration curve was prepared with BSA as the standard in the linear range of 20 µg/ml to 1000 µg/ml. For preparation of the curve as well as for determination of protein concentration of unknown samples, 1 ml of Bradford-reagent (35 g Serva Blue G in 25 ml EtOH, 50 ml 85 % (v/v) phosphoric acid, H<sub>2</sub>O ad 500 ml) was mixed with 20 µl sample and briefly vortexed. Sample was incubated in the dark for 10 minutes and the absorption of the sample at 595 nm was measured in spectrophotometer against 1 ml of Bradford-reagent without sample as a blank-value.

### 4.4.6 DTNB-based *in vitro* test of acyltransferases

This method was modified according to Willis et al. (2011). DTNB-based *in vitro* assays were used to monitor enzyme-mediated cleavage of acyl-CoA-derived thioester-bonds. Enzyme mediated acyl-CoA cleavage results in the liberation of free CoA-SH molecules. Free CoA-SH reacts with the intramolecular disulphide bond of a DTNB molecule, which thereby is cleaved. The reaction results in a disulphide consisting of a CoA and a 5-thio-2-nitrobenzoic acid (TNB) on the one side and a free TNB<sup>2-</sup> ion on the other side. The TNB<sup>2-</sup> absorbs at 412 nm and appears yellow.



**Figure 11: Reaction scheme of the DTNB-reaction.** DTNB is cleaved, forming a mixed disulphide with free CoA molecules and a TNB<sup>2-</sup>-ion which absorbs at 412 nm.

The samples contained 0.2 mg DTNB/ml, 20 µM fatty alcohol, 2 µl of protein solution and 12.5 µM of acyl-CoA in a total volume of 1 ml TBS. Samples were pipetted in polystyrene cuvettes without acyl-

## METHODS

CoA and vortexed for 5 sec. Cuvettes were placed in Cary 100 Bio spectrophotometer (Varian, Germany). Absorbance was taken as a blank-value. After absorbance stabilised to a continuous baseline, acyl-CoA substrate was added, sample was vigorously mixed and increase off absorbance was recorded at 410 nm.

### 4.4.7 NADPH-based *in vitro* test of acyltransferases

NADPH-based *in vitro* tests were done to monitor FAR-mediated NADPH oxidation. Reactions were carried out according to 4.4.6 with the following differences: No DTNB was added to the samples. Absorbance level was not monitored and recorded at 410 nm, but at 340 nm. Also it should be noted, that the absorbance levels decrease in case of monitoring NADPH-oxidation.

### 4.4.8 *In vitro* assays with [<sup>14</sup>C] labelled substrates

#### FAR

*In vitro* activity tests for FARs on the basis of [<sup>14</sup>C]-labelled substrates were carried out according to Hofvander et al. (2011). Samples were composed as follows: 0.5 µg of purified protein, 25 µM [<sup>14</sup>C]-18:1-CoA, 3 mg/ml BSA (fatty acid free), 10 mM NADPH in a total reaction volume of 50 µl TBS. Samples were incubated for 30 min at 30 °C. Afterwards, reaction was quenched by the addition of 10 µl acetic acid and extracted with 375 µl methanol:chloroform 1:1 (v/v), thorough vortexing, and subsequent addition of 125 µl chloroform and 250 µl H<sub>2</sub>O. Sample was centrifuged at 6000 x g for 2 min in order to achieve phase separation. Subsequently, the chloroform phase was transferred to a fresh vial, dried under a stream of N<sub>2</sub> and resolved in 50 µl of chloroform. Sample was loaded on a TLC plate, which was developed in hexane:diethyl ether:acetic acid in a ratio of 55:45:0.5 (v/v/v) as a mobile phase. Afterwards, TLC plate was dried and placed on an image plate (FUJIFILM, Japan). Fatty alcohols were visualised using a Fuji FLA-3000 phosphor imager after 3 days of exposure.

#### AtDGAT3

*In vitro* tests of AtDGAT3 were done according to Rani et al. (2010) with either 30 µM, 60 µM or 200 µM of 1,2-dilinoleoylglycerol. For all samples, freshly prepared protein was used which still showed a brownish colour. In case of utilization of 30 µM and 60 µM 1,2-dilinoleoylglycerol, 5 µg of purified protein were incubated together with 10 µM [<sup>14</sup>C]-labelled linoleoyl-CoA (18:2-CoA) in a total volume of 100 µl of 30 mM HEPES/HCl pH = 7.5, 500 mM NaCl, 10 % glycerol, 30 mM imidazole. In case of utilization of 200 µM 1,2-dilinoleoylglycerol, 20 µM [<sup>14</sup>C]-labelled 18:2-CoA were incubated in combination with 50 µg of purified protein. All samples were incubated at 30 °C for 30 min. Afterwards, the complete sample was loaded on a TLC plate in case of the 200 µM DAG containing samples, while the other samples were extracted twice with 200 µl hexane. TLC were developed in hexane:diethyl ether:acetic acid 80:20:2 (v/v/v). TLC was wrapped in plastic foil and placed on an imaging plate (FUJIFILM, Japan) for 2 days. Imaging plate was analysed using a Fuji FLA-3000 phosphor imager.

## 4.5 Protein crystallisation

Protein crystallisation was done together with Dr. Karin Kühnel and Felix Lambrecht. Felix Lambrecht pipetted most of the crystallisation screens. SEC-purified protein, derived from fractions which eluted at a volume corresponding to a trimeric state of the fusion protein (Peak III, Figure 25 B), was used for crystallisation screens. As glycine-buffer with a pH of 10.0 was found to be the optimal buffer for the protein in the course of Thermofluor analyses, the protein was stored in 20 mM glycine/NaOH pH = 10.0 containing 150 mM NaCl in all cases. First protein crystallisation trails were performed with commercially available crystallisation screens in 96-well plates (Table 5). In the initial screens, sitting drops containing 100 nl of protein solution and 100 nl of precipitant as well as drops containing 200 nl of protein together with 100 nl of precipitant were used. Since precipitate in drops containing only 100 nl of protein solution did not form as often as in drops containing 200 nl of protein solution, later only drops containing 200 nl of protein were used to increase the protein concentration. For *in situ* proteolysis, either trypsin or proteinase K was added in a protease/protein ration of 1:2000 (*w/w*) or 1:500 (*w/w*), respectively. Drops were pipetted with a Cartesian Microsys robot (Cartesian Technologies, USA). After pipetting, plates were sealed with a transparent plastic foil in order to prevent the drops and reservoirs from drying out. Plates were stored in a Rock Imager system (Formulatrix, USA) at 4 °C in case of plates containing protease and at 20 °C in case of plates without protease. Plates were periodically imaged with the Formulatrix system. Protein crystals were flash frozen in liquid nitrogen with 30 % ethylene glycol, 17 % ethanol, 0.2 M MgCl<sub>2</sub>, 0.1 M HEPES pH = 6.5 as a cryoprotectant. X-ray diffraction data were measured at 100 K at the X10SA-beamline at the Swiss Light Source (Paul Scherrer Institute) in Villigen, Switzerland.

### 4.5.1 Grid screens for optimisation of crystallisation-conditions

Refinement of the crystallisation conditions was done with the help of custom-made grid screens (listed in 3.6). Screens were designed together with Dr. Karin Kühnel and Felix Lambrecht according to promising conditions found in commercially available crystallisation screens. Respective conditions were taken as a starting point, around which different parameters like pH or concentration of precipitant were systematically altered. Screens were set up as sitting drops in 96-well plates or as hanging drops in 24-well plates. Reservoir-volume was 75 µl in case of 96-well plates and 1 ml in case of 24-well plates. Drop composition in case of grid screens was 100 nl precipitant and 200 nl of protein solution throughout.

## 5 RESULTS

In order to elucidate structure-function relationships of FARs and WSs, one major goal of this study was to obtain protein crystals of respective enzymes for structure determination. The first part of the results section describes the outcome concerning expression, purification and characterisation of different FARs and WSs for crystallisation studies.

Some FARs and WSs are described to be integral membrane proteins, whose crystallisation is often difficult due to their hydrophobic nature. Alternatively, studies on chimera and single amino acid exchange variants of integral membrane proteins are suitable tools for the elucidation of specificity determining structures. In this study, according experiments have been performed with a murine DGAT2-type WS and the murine DGAT2. Respective findings are described in the second part of this section.

### 5.1 Purification and characterisation of heterologously expressed FARs and WSs

In order to gain detailed insights into the biochemical and biophysical properties of FARs and WSs, the possibility of heterologous expression and purification of respective enzymes from different phylogenetic groups in *E. coli* was investigated during this study. Starting with eight sequences, those which did not fulfil necessary criteria for successful purification and crystallisation of the respective protein were sorted out systematically (Figure 12). In the course of this process, all of the proteins were first assessed for *in vivo* activity in *E. coli* or *S. cerevisiae*, respectively. Subsequently, the possibility of purification of the proteins after cloning respective sequences into different *E. coli* expression vectors was elucidated (Table 16). Since several of the enzymes were predicted to contain one or more TM domains (Figure 14, Figure 22), the pMAL and pCold vector systems were used in order to increase the solubility of the enzymes. Upon cloning of the sequences in these vectors, resulting proteins are produced as fusion proteins with an N-terminal *E. coli* maltose binding protein (MBP) in case of pMAL (Sharff et al., 1993) or the *E. coli* derived Trigger Factor (TF) (Ferbitz et al., 2004; Stoller et al., 1995) in case of pCold. Both proteins are known to efficiently enhance the solubility of fusion partners (di Guan et al., 1988; Kim et al., 2012; Maina et al., 1988; Qing et al., 2004; Ray et al., 2012). Successfully purified proteins were furthermore checked for their oligomeric status to ensure a monodisperse, non-aggregated sample suitable for crystallisation screens. In the course of the project, the repertoire of sequences was continuously narrowed, finally focusing on a bifunctional WS/DGAT (WSD) from *A. baylyi* (AbWSD1). The enzyme was found to be active and easy to purify. Moreover, the enzyme was initially obtained in a non-aggregated state, while aggregation was the major problem for preparations of the other enzymes. In the following section, results upon purification and analyses of every enzyme as well as accompanied difficulties are described. Moreover, Figure 12 provides a comprehensive overview of the process of enzyme selection for crystallisation studies.

## RESULTS

**Table 16: Constructs used for activity tests and purification of FAR and WS enzymes.** UniProt IDs of the enzymes are given in brackets. AbWSD1 = *Acinetobacter baylyi* WSD1 (Q8GGG1), AtDGAT3 = *Arabidopsis thaliana* DGAT3 (Q9C5W0), DmFAR1 = *Drosophila melanogaster* FAR1 (Q8MS59), MaFAR1 = *Marinobacter aquaeolei* FAR1 (A1U2T0), MmAWAT2 = *Mus musculus* AWAT2 (Q6E1M8), MmAWAT2ΔN = *M. musculus* AWAT2, N-terminal truncated version, MmFAR1 = *M. musculus* FAR1 (Q922J9), TcFAR1 = *Tribolium castaneum* FAR1 (D2A5A7). The plasmids pCold, pET24, pET28, pGEX and pMAL are bacterial expression vectors and were used for expression in *E. coli*, while pYES2 and pYES2/NT were used for studies in *S. cerevisiae*.

		<u>sequence</u>							
		DmFAR1	MmFAR1	MaFAR1	TcFAR1	AbWSD1	AtDGAT3	MmAWAT2	MmAWAT2ΔN
<u>expression-plasmid</u>	pCold	○		○		○	○		○
	pET24	○		○					
	pET28	○				○	○	○	○
	pGEX		○						
	pMAL	○	○	○					
	pYES2		○						
	pYES2/NT	○			○		○	○	○

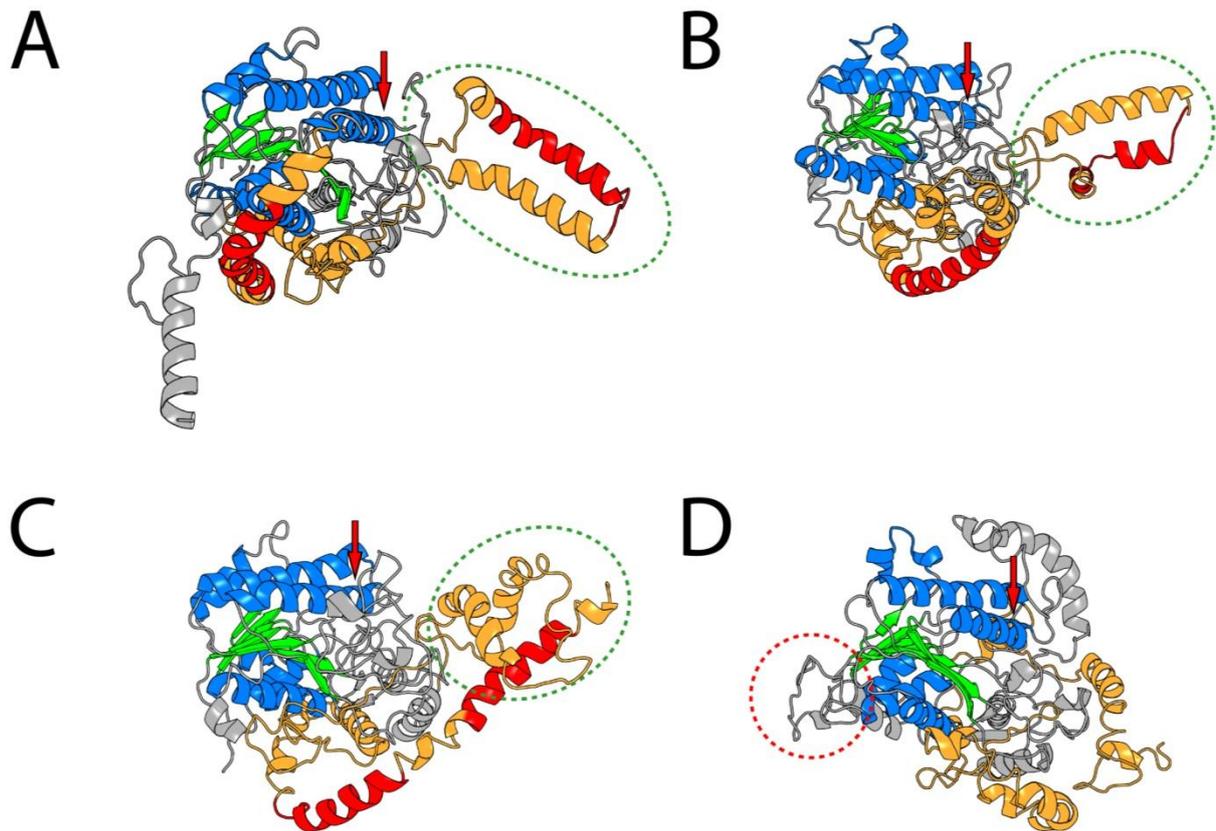
ACTIVITY	EXPRESSION & PURIFICATION	MULTIMERIC STATE	CRYSTALLIZATION
DmFAR1 ↓ MaFAR1 ↓ MmFAR1 ↓ TcFAR1 x AbWSD1 ↓ AthDGAT3 x MmAWAT2 ↓ MmAWAT2ΔN ↓	DmFAR1 ↓ MaFAR1 ↓ MmFAR1 x AbWSD1 ↓ MmAWAT2 x MmAWAT2ΔN x	DmFAR1 x MaFAR1 x AbWSD1 ↓	AbWSD1

**Figure 12: Selection of enzymes for crystallisation studies.** Each enzyme was tested for the criteria shown on the left side of the respective grey box. Whether an enzyme fulfilled the respective criterion or not is indicated by a colour code: green indicates positive results, red indicates negative results and yellow indicates unsatisfactory results with possibilities for improvement. Successfully tested enzymes were subsequently tested for the next criteria. Enzymes with negative results were excluded from further experiments. The "activity" criterion indicates if the respective enzyme activity was detected *in vivo* upon expression in *E. coli* or *S. cerevisiae*. "Expression & purification" points out, if purification of the respective enzyme in a pure form upon expression in *E. coli* was successful. "Multimeric state" refers to the proteins' multimeric states, as concluded from SEC-analyses. Green indicates a physiological realistic multimeric state (e.g. trimeric in case of AbWSD1). DmFAR1 was obtained in a dimeric/trimeric form once and the result still has to be verified. MaFAR1 was obtained in a quaternary complex corresponding to a decamer or in aggregated form.

### 5.1.1 Purification and analyses of FARs

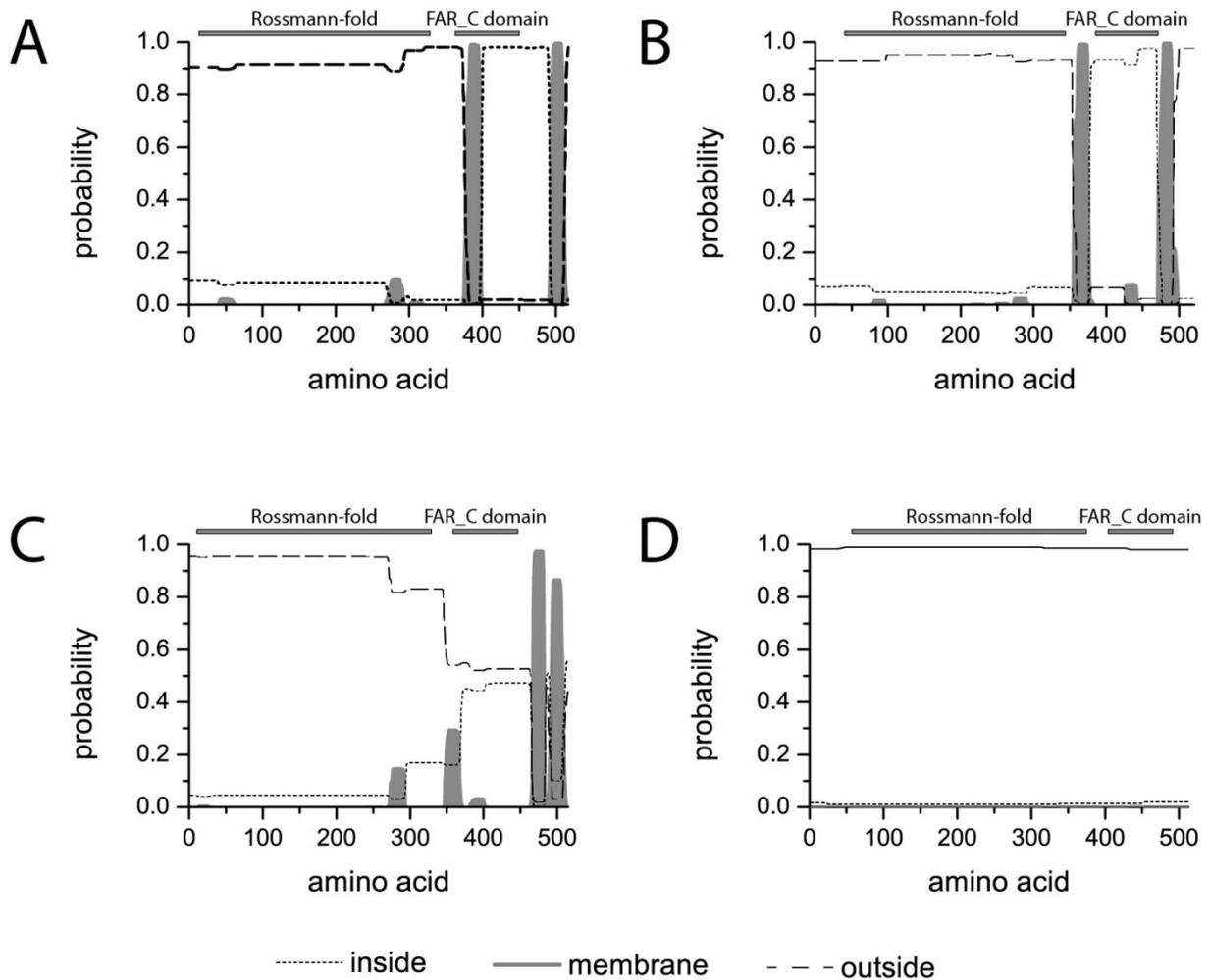
In the course of this study, four different FARs, derived from the mammal *Mus musculus* (MmFAR1), the insects *D. melanogaster* (DmFAR1) and *Tribolium castaneum* (TcFAR1) as well as from the marine bacterium *M. aquaeolei* (MaFAR1) were analysed. MmFAR1 and MaFAR1 have been described before (Cheng, 2004b; Heilmann et al., 2012; Hofvander et al., 2011; Wahlen et al., 2009). MmFAR1 was shown to be NADPH-dependent, to localise to peroxisomes via two predicted C-terminal TM domains and to produce both saturated and polyunsaturated fatty alcohols of 16 to 20 carbon atoms in chain length (Cheng, 2004b; Heilmann et al., 2012). MaFAR1 does not contain any predicted TM domains, is NADPH dependent as well and exhibits a broad substrate range from acyl-CoAs with ten to twenty carbon atoms in length (Hofvander et al., 2011; Wahlen et al., 2009). The sequences of TcFAR1 and DmFAR1 have been discovered by cooperation partners during this study and were kindly provided for further analyses. In case of DmFAR1, the present study contributed to the publication of Jaspers et al. (2014).

Phylogenetic analyses of the sequences indicated a low similarity compared to each other (Figure 4), which may reflect a structural diversity among the enzymes and might thus be beneficial in order to sort out an optimal candidate for purification and crystallisation. In detail, MmFAR1 clusters with FARs from other vertebrates, while TcFAR1 and DmFAR1 cluster with other insect-derived FAR sequences. Interestingly, MaFAR1 from *Marinobacter* clusters with plastidial, plant-type FARs rather than with bacterial sequences. In contrast, MaFAR2, which derives from the same organism, clusters in the bacterial-type FAR subclade. The four FAR sequences utilised in this study were also submitted to the Phyre server (Kelley and Sternberg, 2009) and the TMHMM online service (Sonnhammer et al., 1998). The Phyre server offers a structure modelling service (Figure 13), while TMHMM predicts possible TM domains within a sequence (Figure 14).



**Figure 13: Modelled FAR-structures.** Amino acid sequences of single enzymes were submitted to the Phyre server in the “intense” mode (Kelley and Sternberg, 2009). All structures were modelled against a non-ribosomal peptide synthetase from *M. tuberculosis* (PDB accession code 4DQV, (Chhabra et al., 2012), as automatically chosen by the Phyre service. Models are coloured according to the following properties:  $\beta$ -sheets of Rossmann folds are green, while the characteristic  $\alpha$ -helices of the Rossmann folds are blue. The C-terminal FAR\_C-domains, as annotated by BLAST search (Altschul et al., 1997), are coloured in orange, while predicted transmembrane domains according to TMHMM (Sonnhammer et al., 1998) are red. Moreover, a characteristic hairpin structure is circled in a green dashed line. The position of the catalytic motif is indicated by red arrows. **A)** Model of DmFAR1. The server modelled 73 % of the residues with > 90 % confidence. **B)** Model of TcFAR1. 73 % of all residues were modelled at > 90 % confidence. **C)** Model of MmFAR1. 73 % of all residues were modelled at > 90 % accuracy. Note: The red helix within the dashed green circle is, seen in perspective, behind the actual hairpin motif of the enzyme and does structurally contribute to it. The two predicted transmembrane helices depicted in red have been shown to mediate a peroxisomal localisation of MmFAR1 (Heilmann et al., 2012). **D)** Model of MaFAR1. 71 % of all residues were modelled at > 90 % accuracy. In case of MaFAR1, no characteristic hairpin structure was found in the model. The dashed red circle indicates the N-terminus of the protein, which is further discussed in section 6.1.5.

## RESULTS



**Figure 14: Predicted transmembrane structure of FARs analysed in this study.** Data for respective plots were generated with the TMHMM online service (Sonnhammer et al., 1998). The parts of the sequences which are annotated as Rossmann fold and FAR\_C-domain upon BLAST analyses (Altschul et al., 1997) are indicated by grey bars, respectively. TM prediction for **A)** DmFAR1, **B)** TcFAR1, **C)** MmFAR1 and **D)** MaFAR1.

### 5.1.1.1 DmFAR1 from *D. melanogaster*

The gene CG1443 from *D. melanogaster* was identified in the course of a RNAi-screening (Jaspers et al., 2014), aimed for the identification of genes which are involved in tracheal liquid clearing in *Drosophila* larvae. During this process, liquid in the tracheal tubes is replaced by gas shortly prior to hatching of the larvae (Elias and O’Brodivich, 2006; Förster and Woods, 2013). Upon knock-down of CG1443, tracheal clearing in *Drosophila* is impaired, suggesting a fundamental role of the gene product in the course of this process.

Since it contributes to the waterproof properties of tracheal tubes after liquid clearing, the gene was initially termed *wat1* as a short form of “waterproof”. However, BLAST analyses (Altschul et al., 1997, 2005) of the gene’s coding sequence revealed the typical conserved domain structure of a FAR (Appendix 1), consisting of a N-terminal Rossmann fold domain and a C-terminal FAR\_C domain (Figure 13 A, Figure 14 A). Hence, the corresponding protein is called “DmFAR1” throughout this study. In phylogenetic analyses, DmFAR1 does not cluster with the majority of the other insect-derived FAR sequences, but is part of sub-cluster together with two other *Drosophila*-derived FAR, TcFAR1 from *T. castaneum* and EgFAR from *E. gracilis* (Figure 4).

## RESULTS

The sequence is predicted to contain two TM regions. One of those domains is located at the very C-terminus of the enzyme, whereas the second domain is anticipated to be about 100 residues in front of the C-terminus (Figure 14 A). Phyre based modelling (Kelley and Sternberg, 2009) of the DmFAR1 structure resulted in 74 % of the residues modelled with over 90 % confidence. The template structure for the model was the reductase (R) domain of a non-ribosomal peptide synthetase from *M. tuberculosis* (PDB accession code 4DQV, (Chhabra et al., 2012)), as automatically chosen by the Phyre service. The confidence for real homology between both enzymes was 100 %. The modelled structure shows a typical Rossmann fold pattern, with seven  $\beta$ -sheets sandwiched by three  $\alpha$ -helices on each side (Figure 13 A, depicted in green and blue, respectively). The N-terminus of the enzyme is predicted to form an isolated  $\alpha$ -helix, which is connected to the enzyme with a flexible linker. The predicted TM domains are calculated to form surface exposed  $\alpha$ -helices (Figure 13 A, depicted in red), of which the non C-terminal one forms a hairpin like structure with a second  $\alpha$ -helix, which is not predicted to be a TM domain (Figure 13 A, circled in dashed green line).

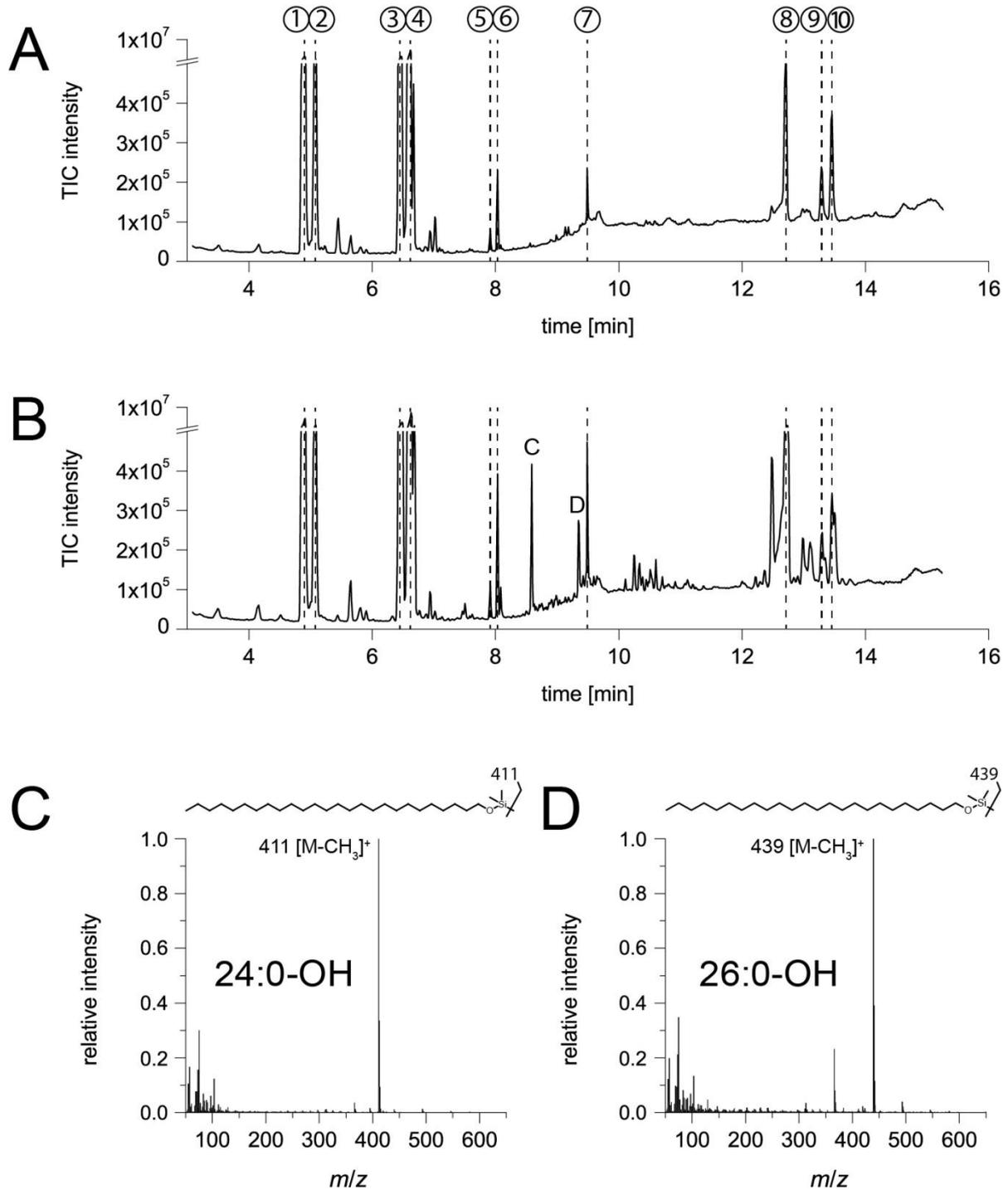
### 5.1.1.1.1 DmFAR1-mediated fatty alcohol production in *S. cerevisiae* and *E. coli*

In order to verify the annotated function of the enzyme, the sequence of DmFAR1 was cloned into the pYES2/NT vector and expressed in the *S. cerevisiae* strain H1246 (Sandager, 2001). The neutral lipid fraction of respective expression cultures was extracted and trimethylsilyl- (TMS) derivatised with N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA). GC-MS analyses showed that cultures transformed with pYES2/NT\_DmFAR1 and grown under inducing conditions contained the very long chain fatty alcohols tetracosanol (24:0-OH) (Figure 15 B&C) and hexacosanol (26:0-OH) (Figure 15 B&D). Mass spectra of respective total ion count (TIC) peaks showed the characteristic  $[M-CH_3]^+$ -fragment of TMS-derivatised fatty alcohols and match the published fragmentation patterns (<http://lipidlibrary.aocs.org>). Apart from these two very long chain fatty alcohols, no further fatty alcohols were detected. In contrast, cultures which grew under non-inducing conditions or expressing the pYES2/NT empty vector control did not contain detectable amounts of fatty alcohols (Figure 15 A).

In order to test a DmFAR1-mediated fatty alcohol production in *E. coli*, the sequence was furthermore cloned into the bacterial expression vectors pMAL and pET24 and lipid extracts of respective expression cultures were prepared. Since DmFAR1 was found to produce 24:0-OH and 26:0-OH upon expression in *S. cerevisiae*, *E. coli* expression cultures were supplemented with 1 mM hexacosanoic acid (26:0-COOH) in an additional experiment. Analogous to yeast extracts, the bacteria-derived samples were TMS-derivatised and analysed via GC-MS. Although the intensity of the signals was low, fatty alcohols were detected in samples from cultures expressing pMAL\_DmFAR, but not in samples from cultures expressing the pMAL empty vector control (Figure 16 A&B). Respective signals were assigned to tetradecanol (14:0-OH), hexadecanol (16:0-OH), hexadecenol (16:1-OH) and octadecenol (18:1-OH) (Figure 16). In case of cultures expressing pET24\_DmFAR, only very low signals for 14:0-OH were detected in two out of three independent samples (Figure 16 C). Analyses of cultures supplemented with 26:0-COOH yielded in results which were similar to those of the non-fed cultures. No 26:0-OH was detected in respective samples of cultures expressing either pMAL\_DmFAR1 or pET24\_DmFAR1 (data not shown).

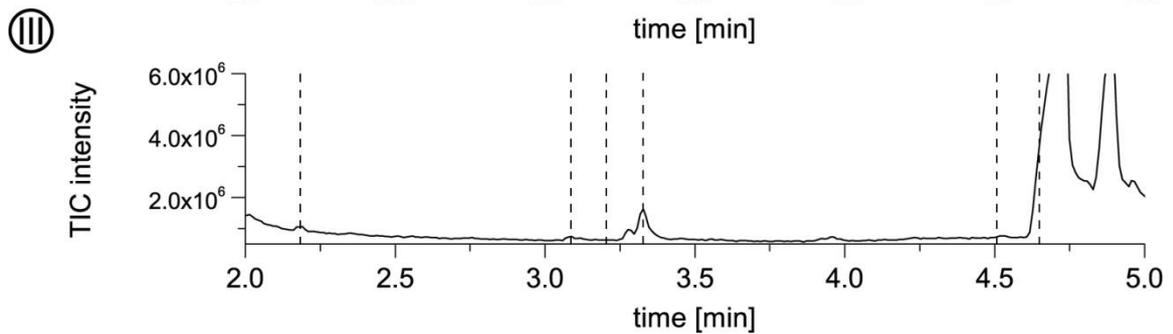
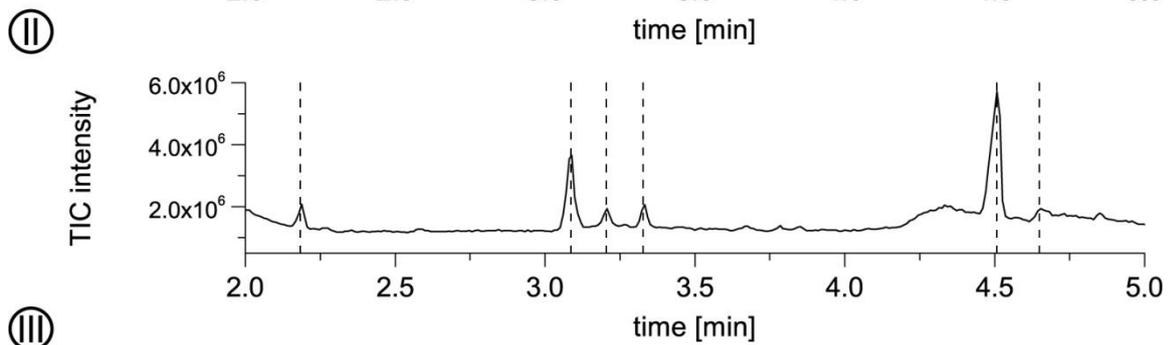
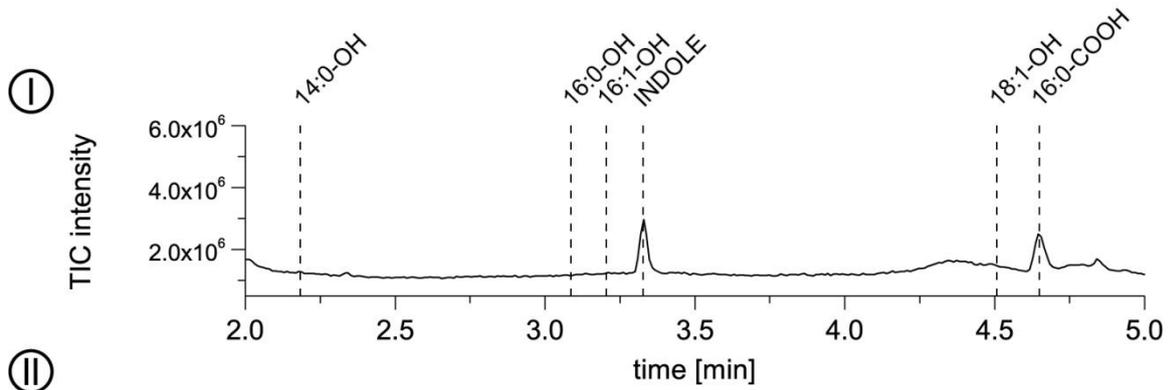
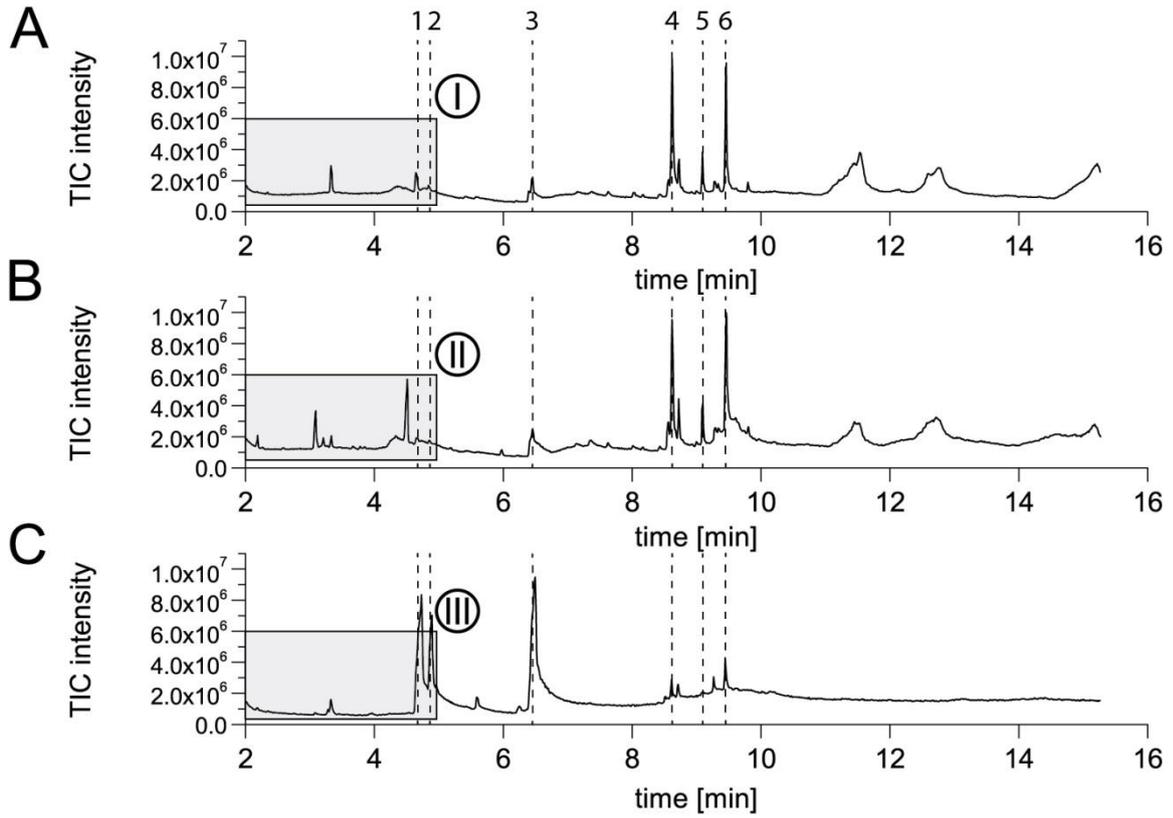
Taken together, it was shown that DmFAR1 is an active protein upon expression in either *S. cerevisiae* or *E. coli* and exhibits strongly different substrate specificities in both organisms.

## RESULTS

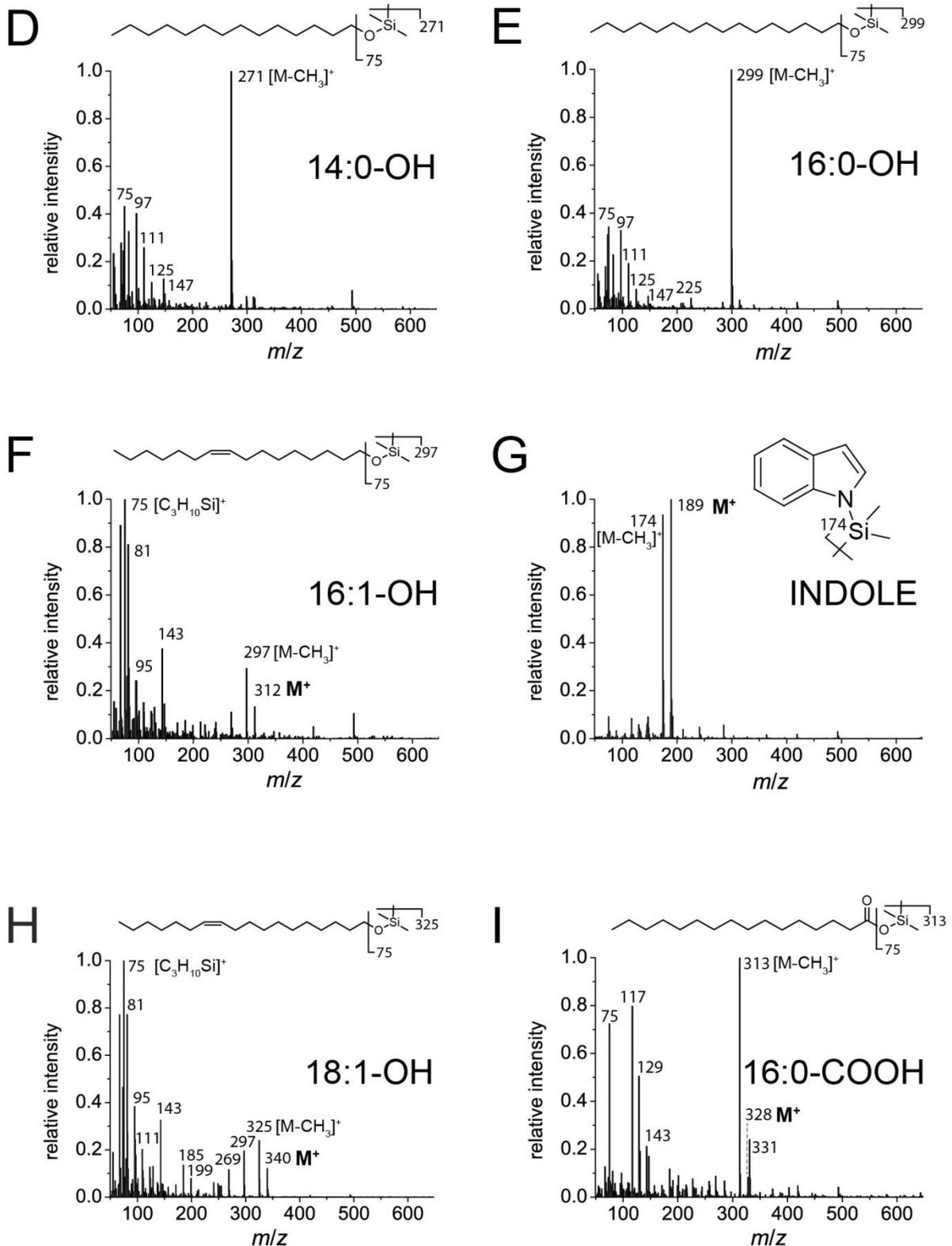


**Figure 15: DmFAR1-mediated production of fatty alcohols in *S. cerevisiae*.** **A)** GC-MS total ion count (TIC) of lipid extracts from *S. cerevisiae* H1246 expressing pYES2/NT as an empty vector control and **B)** pYES2/NT\_DmFAR1. Molecular identities of the numbered TIC peaks in A and B were identified by the NIST MS Search 2.0 library. Probabilities for respective hits are given in brackets. **1** = hexadecanoic acid, trimethylsilyl ester (96%), **2** = hexadecanoic acid, trimethylsilyl ester (56.6%), **3** = octadecanoic acid, trimethylsilyl ester (93.4%), **4** = octadecanoic acid, trimethylsilyl ester (22.5%), **5** = eicosanoic acid, trimethylsilyl ester (31.9%) **6** = eicosanoic acid, trimethylsilyl ester (42.6%), **7** = squalene (26.6%), **8** = ergosterol, trimethylsilyl ether (78.4%), **9** = cholesterol, trimethylsilyl ether (38.6%), **10** = stigmaterol, trimethylsilyl ether (58.0%) **C)** Mass spectrum of peak "C", showing the typical fragmentation pattern of the tetracosanol (24:0-OH) trimethylsilyl ether. The most abundant species with a mass of 411 represents the TMS-derivatised molecule of 24:0-OH lacking one of the TMS-derived methyl groups. **D)** MS-spectrum of peak "D", showing the typical fragmentation pattern of the hexacosanol (26:0-OH) trimethylsilyl ether. The most abundant species with a mass of 439 represents the TMS-derivatised molecule of 26:0-OH lacking one of the TMS-derived methyl groups. Fatty alcohols were detected in more than three samples derived from independent cultures.

RESULTS



RESULTS



**Figure 16: DmFAR1-mediated production of fatty alcohols in *E. coli*.** **A)** GC-MS total ion count (TIC) of lipid extracts from an *E. coli* BL21\* (DE3) culture expressing pMAL as an empty vector control. Data are also representative for extracts from cultures expressing the pET24 empty vector control. **B)** GC-MS TIC of lipid extracts from an *E. coli* BL21\* (DE3) culture expressing pMAL\_DmFAR1 and **C)** pET24\_DmFAR1. Molecular identities of the numbered TIC peaks in A-C were identified by the NIST MS Search 2.0 library. Probabilities for respective hits are given in brackets. 1 = hexadecanoic acid, trimethylsilyl ester (77.39 %), 2 = hexadecenoic acid, trimethylsilyl ester (58.67 %) 3 = octadecanoic acid, trimethylsilyl ester (20.42 %), 4 = hexadecanoic acid, 2,3-bis[(trimethylsilyl)oxy]propyl ester (74.42 %), 4 = 5-chloro-6beta-nitro-5alpha-cholestan-3-one (54.41 %), 6 = 1-monooleoylglycerol trimethylsilyl ether (35.50 %). Enlargement of the insets depicted by grey boxes are shown in I-III, respectively. Molecular identities of the TIC peaks, as deduced from respective mass spectra, are indicated on top of the trace. **D-I)** Mass spectra of respective peaks indicated in A-C. Mass spectra were obtained by subtraction of the

## RESULTS

pMAL empty vector control signal from the pMAL\_DmFAR1 signal by utilization of the “Subtract Background”-tool of the Xcalibur software suite under standard settings. **D)** Mass spectrum of the “14:0-OH”-peak, corresponding to the fragmentation pattern of the tetradecanol trimethylsilyl ether. The most abundant species with a mass of 271 represents the TMS-derivatised molecule of 14:0-OH lacking one of the TMS-derived methyl groups. **E)** Mass spectrum of the “16:0-OH”-peak, corresponding to the fragmentation pattern of the hexadecanol trimethylsilyl ether. Again, the most abundant species with a mass of 299 represents the TMS-derivatised molecule of 16:0-OH lacking one of the TMS-derived methyl groups. **F)** Mass spectrum of the “16:1-OH”-peak, corresponding to the fragmentation pattern of the hexadecenol trimethylsilyl ether. In desaturated fatty alcohols, the most abundant ion species represents the trimethylsilane group with a mass of 75. Moreover, a molecular ion can be detected, which has a mass of 312 in this case. The fragment with a mass of 297 represents the mass of the hexadecenol trimethylsilyl ether molecular ion lacking one TMS-derived methyl group **G)** Mass spectrum of the “indole”-peak. This peak was detected in both, empty vector control and DmFAR1-expressing cultures. The molecular identity of the peak was determined to be indole by the NIST MS Search 2.0 library. **H)** Mass spectrum of the “18:1-OH”-peak, corresponding to the fragmentation pattern of the octadecenol trimethylsilyl ether. The fragmentation pattern of 18:1-OH is analogous to 16:1-OH. **I)** Mass spectrum of the “16:0-COOH”-peak. The molecular identity of the peak was determined by the NIST MS Search 2.0 library. A molecular ion can be detected, which has a mass of 328. The most abundant ion has a mass of 313 and resembles the molecular ion lacking one of the TMS-derived methyl groups. The intensity of the “16:1-COOH” TIC-peak varied among the samples and was not always as minor as shown in B. Data are representative for three samples from independent cultures in each case.

### 5.1.1.1.2 Recombinant expression of DmFAR1 in *E. coli*

As the DmFAR1 was shown to be active upon expression in both, *S. cerevisiae* and *E. coli*, it was considered to be an interesting candidate for crystallisation attempts. In order to explore the possibility of purifying heterologously produced DmFAR1 for crystallisation studies, the corresponding sequence was cloned into different *E. coli* expression-vectors (Table 16).

It was previously shown, that the enzyme is natively located at the ER (Jaspers et al., 2014). Moreover, TM predictions suggest the presence of two C-terminal TM motifs in the DmFAR1 protein (Figure 14 B), making it unlikely that the enzyme could be expressed in a soluble form. To support the production of soluble enzyme upon heterologous expression in *E. coli*, the DmFAR1 coding sequence was initially cloned into the pMAL and the pCold expression vectors.

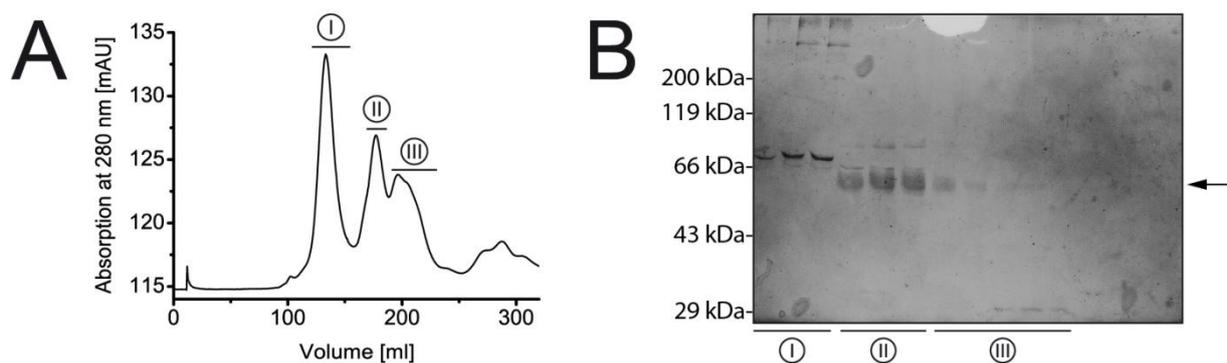
In both cases, expression of the respective construct led to a strong accumulation of protein with an estimated size of approximately 110 kDa, as judged by SDS-PAGE. This matched the calculated sizes of 105.6 kDa for MBP-DmFAR1 and 112.5 kDa for 6xHis-TF-DmFAR1. MBP-DmFAR1 and 6xHis-TF-DmFAR1 were both efficiently purified via affinity chromatography, using amylose resin in case of MBP-DmFAR1 and Ni<sup>2+</sup>-based IMAC in case of 6xHis-TF-DmFAR1. Both purification methods yielded an almost pure eluate (Appendix 2 & Appendix 3). However, SEC with these purified protein preparations resulted in elution of the protein in the void-volume of the column (representatively shown in Figure 19 A). The resolution of the used HiLoad 26/60 Superdex S200 column covers a range from 10-600 kDa and the void volume represents protein species with molecular weights  $\geq$  600 kDa. Thus, both 6xHis-TF-DmFAR1 and MBP-DmFAR1 eluted in oligomeric and most likely aggregated forms. Separation of TF and DmFAR1 via thrombin mediated cleavage and subsequent attempts to separate both proteins via IMAC resulted in constant co-elution of both proteins (data not shown, conditions analogous to “affinity chromatography” section in Table 15).

The *in vitro* activity of 6xHis-TF-DmFAR1 and MBP-DmFAR1 was tested in an assay utilizing [<sup>14</sup>C] labelled 18:1-CoA. Respective samples were extracted with chloroform:methanol (1:1, v/v). No signals corresponding to fatty alcohols were detected upon subsequent TLC analyses of the samples (Appendix 6). However, the assay was only done once under standard conditions. Due to time constraints, no attempts of further optimisation were performed.

The DmFAR1 coding sequence was furthermore cloned into the expression vectors pET24 and pET28. Both vectors encode the sequence of a hexahistidine-tag, which is fused either C-terminal (pET24) or

## RESULTS

N-terminal (pET28) to the insert. Only impure eluate was obtained upon an initial attempt of purification of 6xHis-DmFAR1 via IMAC. SDS-PAGE analyses of respective eluates showed several bands of equal intensity. Both bands migrated at positions which seemed to be too low for the calculated molecular weight of 6xHis-DmFAR1, which is 63 kDa (Appendix 4). In case of purification of DmFAR1-6xHis from cultures expressing pET24\_DmFAR, analyses of IMAC eluates via SDS-PAGE showed generally less bands than the respective gel of 6xHis-DmFAR1. Moreover, a smeary band was observed in the elution fractions. Its position might resemble the calculated molecular weight of DmFAR1-6xHis, which is 60.5 kDa (Appendix 5). Respective fractions were concentrated with Spin-X® UF concentrators (CORNING, USA) and analysed via SEC. The sample eluted in three peaks, corresponding to an approximate size of 445 kDa, 170 kDa and 110 kDa (Figure 17 A). The corresponding SDS-PAGE results showed a smeary band similar to the band which was found in the IMAC fractions. It was present in the peaks corresponding to 170 kDa and 114 kDa in almost pure form (Figure 17 B). The molecular weights of 170 kDa and 114 kDa may represent a dimeric and a trimeric form of DmFAR1-6xHis, respectively. However, SEC of DmFAR1-6xHis under the given conditions was done only once. An additional purification of the DmFAR1-6xHis with buffer containing 150 mM instead of 500 mM NaCl resulted in elution fractions of lower purity in comparison to the high salt buffer, but the smeary band was observed again (data not shown). All buffer compositions and conditions used for purifications of DmFAR1 are summarised in Table 14.



**Figure 17: SEC analysis of DmFAR1-6xHis. A)** Chromatogram of SEC analysis of DmFAR1-6xHis. Peak I corresponds to a molecular weight of about 445 kDa, while peak II corresponds to about 170 kDa and peak III to about 110 kDa. **B)** SDS-PAGE of samples derived from SEC shown in A. Circled numbers indicate lanes corresponding to the respective peaks in A. The potential DmFAR1 band is indicated by a black arrow. Data represent a single experiment.

### 5.1.1.2 TcFAR1 from *T. castaneum*

The sequence of a putative FAR from the flour beetle *T. castaneum* was kindly provided by Sabrina Lehmann (Department of DEV BIOL, University of Göttingen, Germany). Since no FAR from *T. castaneum* is described to date, the enzyme is called “TcFAR1” throughout this study. BLAST analyses (Altschul et al., 1997) of the respective amino acid sequence indicated that the enzyme might be a FAR, since the typical N-terminal Rossmann fold as well as a C-terminal FAR\_C domain were identified in the sequence (Appendix 7). Phylogenetic analyses revealed that TcFAR1 clusters with DmFAR1, DmFAR2, DmFAR3 and EgFAR. Among the sequences which were analysed during this study, TcFAR1 is closest related to DmFAR1 (Figure 4). Upon sequence based modelling of the enzyme, the Phyre server (Kelley and Sternberg, 2009) automatically chose the same template structure as for DmFAR1. Phyre was able to model 73 % of all residues at more than 90 % confidence.

## RESULTS

The confidence for real homology between DmFAR1 and the chosen template structure was 100 %. The model supports the predicted presence of a Rossmann fold in the protein, as it contains the typical seven  $\beta$ -sheets, which are sandwiched by three  $\alpha$ -helices on each side (Figure 13 B, depicted in green and blue, respectively). Bioinformatic analyses further predict the presence of two TM domains in the C-terminal part of the sequence (Figure 14 B). The positions of both TM domains within the sequence are highly similar to the two predicted TM domains in DmFAR1. Similar to the modelled structure of DmFAR1, respective parts consistently form surface exposed,  $\alpha$ -helical structures in the modelled TcFAR1 protein (Figure 13 B, depicted in red). Apart from that, the TcFAR1 model also contains the characteristic hairpin-like structure which was already observed in the DmFAR1 model. Again, the hairpin consists of the predicted, non-C-terminal TM domain and a second helix (Figure 13 B circled in dashed green line). To analyse a possible FAR activity of the enzyme, the corresponding coding sequence was cloned in the pYES2/NT vector for yeast expression. However, upon extraction of respective lipid extracts from yeast, TMS-derivatization and GC-MS analyses, no fatty alcohols were detected in the samples. Therefore, work on this sequence was stopped.

### 5.1.1.3 MmFAR1 from *M. musculus*

FAR1 from *M. musculus* was already shown to be active upon expression in yeast earlier in the work of Cheng (2004b) and Heilmann et al. (2012). Analyses of the amino acid sequence of the enzyme in terms of hydrophobicity via TMHMM (Sonnhammer et al., 1998) suggested the presence of two TM domains at the C-terminus of the enzyme (Figure 14 C). It was shown before, that this part of the enzyme is responsible for peroxisomal localisation of MmFAR1 and thus might build a TM anchor (Heilmann et al., 2012). Phyre assisted modelling of the enzyme (Kelley and Sternberg, 2009) was done with the same template as for DmFAR1 and TcFAR1, as chosen by Phyre. 73 % of all amino acid residues were threaded with a confidence of over 90 %. The confidence in real homology between both enzymes was 100 %. The model showed the typical Rossmann fold domain, which was also observed in the models of the other FARs (Figure 13 C, depicted in green and blue). The two predicted TM domains responsible for peroxisomal targeting of the enzyme are predicted to form a long, surface exposed and  $\alpha$ -helical structure (Figure 13 C, depicted in red). Furthermore, the MmFAR1 model also shows the hairpin structure, which was already observed in the models of DmFAR1 and TcFAR1. Similar to DmFAR1 and TcFAR1, the hairpin structure is formed by two helices, which are located 100 residues in front of the C-terminus (Figure 14). In contrast to DmFAR1 and TcFAR1, this part of the MmFAR1 sequence is not predicted to form a TM domain. TMHMM based predictions indicate the probability for TM domains in this part of the MmAWAT2 sequence to be only 30 % (Figure 14 C).

In order to increase the probability of the production of a soluble protein upon expression in *E. coli*, the MmFAR1 coding sequence was cloned in the pGEX and pMAL expression vectors. Since expression screenings of pMAL\_MmFAR1 and pGEX\_MmFAR1 transformed in *E. coli* Lemo21 (DE3) cells did not show any prominent bands in SDS-PAGE analyses of the respective *E. coli* cell lysates (data not shown), work on MmFAR1 was not continued.

## RESULTS

### 5.1.1.4 MaFAR1 from *M. aquaeolei*

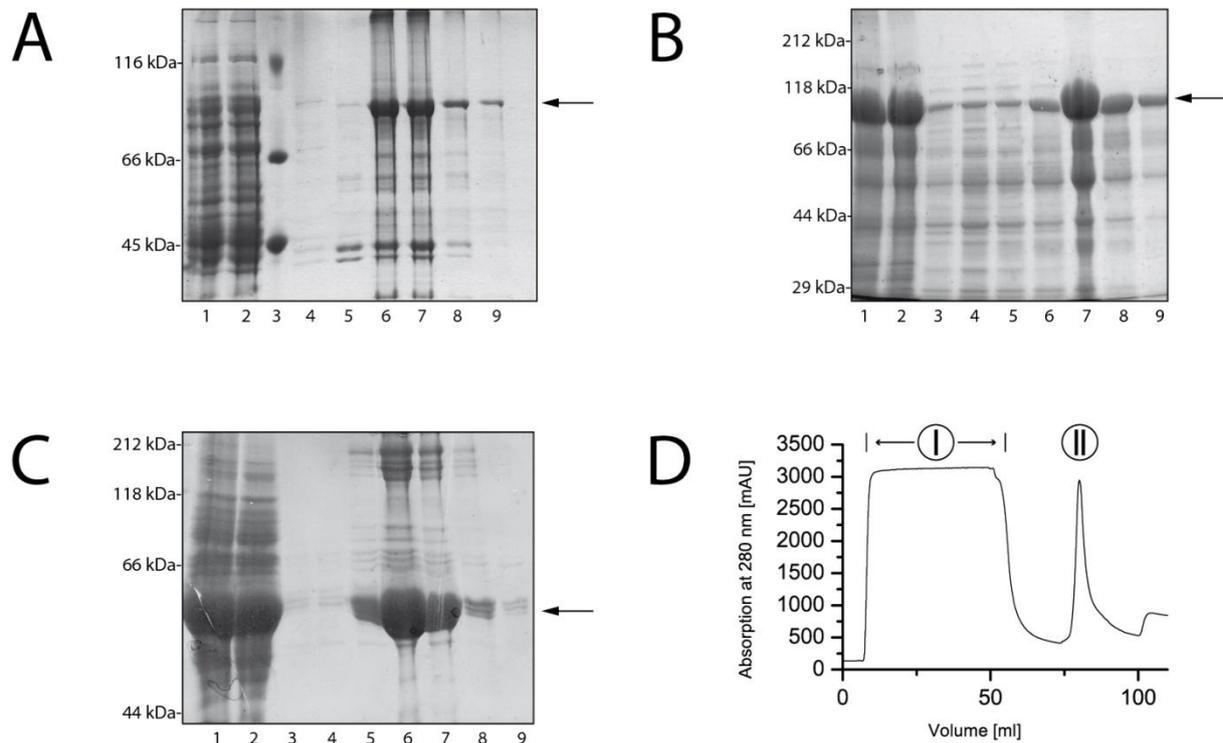
#### 5.1.1.4.1 Purification of heterologously produced MaFAR1

In contrast to the other three analysed FARs, MaFAR1 from *M. aquaeolei* is a prokaryotic enzyme. Analysis of the sequence's hydrophobicity using the TMHMM online service (Sonnhammer et al., 1998) did not reveal any putative TM domains or hydrophobic areas within the enzyme (Figure 14 D), supporting the anticipated soluble nature of MaFAR1 in its natural state. As for the other FARs, the enzyme structure of MaFAR1 was modelled with the Phyre server (Kelley and Sternberg, 2009). The service chose the same template as it did for the other FARs and thereby modelled 71 % of all residues at over 90 % confidence. The confidence in real homology was 100 % again. The model shows the typical Rossmann fold pattern, as already described for the other FARs (Figure 13 D, Rossmann fold depicted in green and blue). In contrast, the characteristic hairpin structure of the other FAR models is missing in the MaFAR1 model.

Despite of the anticipated soluble nature of the enzyme, the authors of the publication which first described the enzyme found an N-terminal hexahistidine-tagged version of the protein to be mainly present in the insoluble fraction upon expression in *E. coli*. To circumvent this issue, they used a fusion protein of the actual MaFAR1 and the maltose binding protein (Hofvander et al., 2011; Wahlen et al., 2009) for purification studies. The respective construct was kindly provided by Dr. Per Hofvander (Department of Plant Breeding, Uppsala, Swedish University of Agricultural Sciences, Sweden) for experiments in the present study. The sequence for MaFAR1 was furthermore cloned in additional *E. coli* expression vectors (Table 16) in order to perform activity tests and prepare heterologously produced enzyme with a purity which was suitable for crystallisation screens. Purification of cell lysates from *E. coli* expressing the 6xHis-MBP-MaFAR1 fusion protein via dextrin-sepharose affinity chromatography or IMAC resulted in a high accumulation of a protein which matched the calculated size of 6xHis-MBP-MaFAR1 of 100.3 kDa (Figure 18 A). Subsequent SEC analyses further increased the purity of the enzyme sample. However, the early time point of elution from the SEC in the void volume of the column indicated an aggregated form of the 6xHis-MBP-MaFAR1 fusion protein (representatively shown in Figure 19A). Since the authors of the initial study did not comment on the multimeric state of the 6xHis-MBP-MaFAR1 protein in their publication, it remains unknown whether they encountered the same problems with protein aggregation. To prevent aggregation of the protein, its stability in different buffer compositions was tested by the Thermofluor method. Respective measurements indicated the highest stability of 6xHis-MBP-MaFAR1 in a buffer containing 50 mM sodium-citrate at a pH of 5.0 and 10 % of glycerol (data not shown). However, the optimised buffer did not result in a change of the aggregation-state. Expression and purification of MaFAR1 fused to the TF (6xHis-TF-MaFAR1) resulted in similar observations. As shown in Figure 18B, IMAC elution fractions contained high proportions of a protein migrating little under the 118 kDa marker, matching the calculated size of 6xHis-TF-MaFAR1 of 109.3 kDa. Comparable to the 6xHis-MBP-tagged version of MaFAR1, also 6xHis-TF-MaFAR1 eluted in the void-volume of the SEC column, thus indicating a highly multimeric state and most likely an aggregated form of the protein (representatively shown in Figure 19 A). Addition of either 0.3 % (w/v) CHAPS or 0.1 % (v/v) Triton X-100 to the respective buffers did not result in a change of the quaternary structure as judged by SEC analyses. All conditions used for purification of 6xHis-TF-MaFAR1 are listed in

Table 13.

## RESULTS

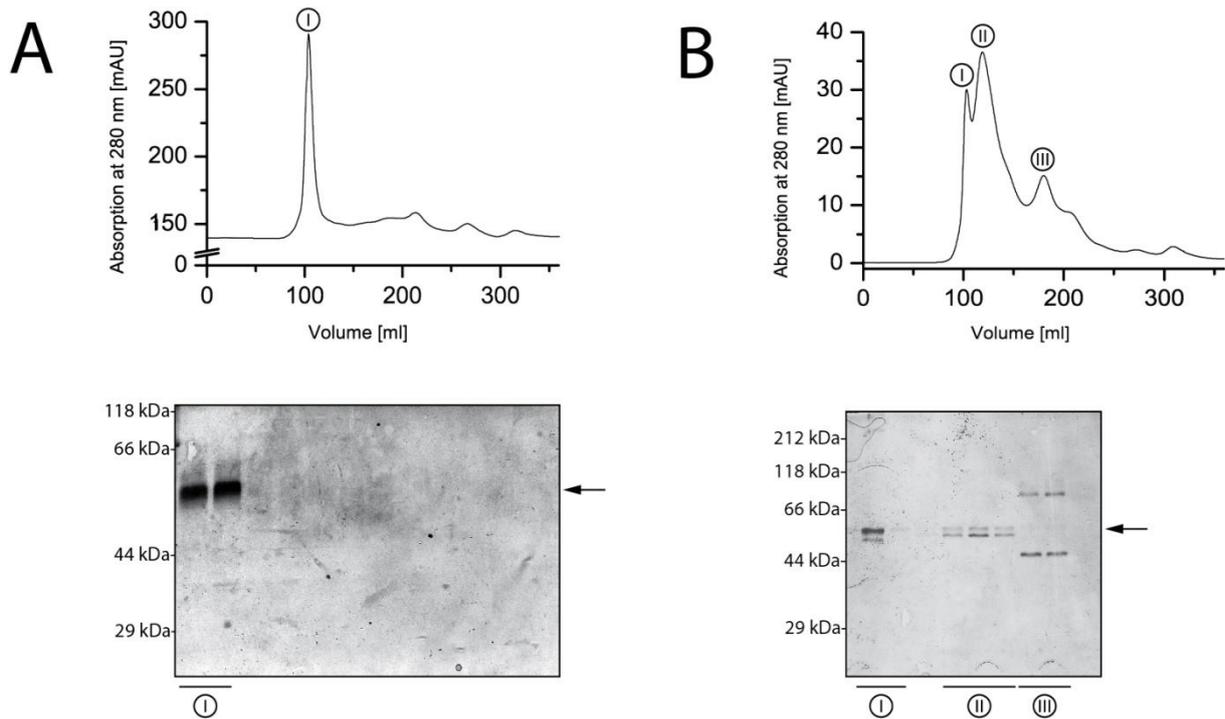


**Figure 18: IMAC purification of heterologous MaFAR1 from *E. coli*.** **A)** Purification of 6xHis-MBP-MaFAR1 fusion protein via dextrin sepharose resin. 1 = load, 2 = flow-through, 3 = molecular weight marker, 4-5 = wash, 6-9 = elution. The position of 6xHis-MBP-MaFAR1 is indicated by a black arrow. **B)** Purification of 6xHis-TF-MaFAR1 fusion protein via IMAC. 1 = load, 2 = flow-through, 3-5 = wash, 6-9 = elution. The position of 6xHis-TF-MaFAR1 is indicated by the black arrow. **C)** Purification of MaFAR1-6xHis via IMAC. 1 = load, 2 = flow-through, 3-9 = elution. The position of MaFAR1-6xHis is indicated by the black arrow. **D)** Representative example for an IMAC purification of MaFAR1. I) Protein loading II) Elution of bound protein from the column. All experiments have been performed at least three times.

Since purification of fusions of big protein tags and MaFAR1 yielded in an aggregated state of the respective protein, the MaFAR1 coding sequence was furthermore cloned into the pET24 vector, encoding for a hexahistidine-tag at the C-terminus of the insert. Cell lysates of respective *E. coli* expression cultures showed a strong accumulation of protein in the range of the expected size of MaFAR1-6xHis (58.2 kDa). The enzyme was efficiently purified via IMAC, yielding high amounts of almost pure protein (Figure 18 C). However, the purified protein precipitated shortly after elution from the Ni<sup>2+</sup>-sepharose column in the elution buffer, resulting in a cloudy, whitish solution. Precipitation of the IMAC-derived protein eluate was avoided by adding 300 mM NaCl and 1 mM betaine to respective expression cultures according to Oganessian et al. (2007). Under these conditions, betaine is thought to act as a compatible solute and promote correct protein folding (Oganessian et al., 2007). Subsequent SEC analyses with non-precipitated protein again showed the characteristic elution pattern of aggregated protein. This issue was partly circumvented by the use of a basic 50 mM CHES buffer with pH = 9 during SEC. Under this condition, the protein did not only elute in the void volume of the column, but also as a fraction with the calculated size of approximately 600 kDa (Figure 19 B). Given the fact that the calculated monomeric size of the MaFAR1-6xHis is 58.2 kDa, the 600 kDa-fraction would correspond to a decamer, which is not likely to be a physiological state. Moreover, SDS-PAGE analyses indicated the presence of two protein species in respective samples, migrating shortly in succession to one another (Figure 19 B). Apart from attempts to change the highly oligomeric state of the MaFAR1-6xHis protein by alterations in the buffer compositions and culture conditions, bypassing of aggregation of the protein was furthermore tried by on-column refolding. For this purpose, *E. coli* cell lysate derived protein was

## RESULTS

loaded on a Ni<sup>2+</sup>-IMAC column. Subsequently, bound protein was gradually exposed to a high concentration of guanidine hydrochloride in order to unfold the protein. Alternatively, protein was unfolded prior to column loading in buffer containing 6 M of guanidine hydrochloride. Prior to elution of the protein, the guanidine hydrochloride concentration was slowly decreased again in order to allow slow and controlled refolding of the protein. However, no protein was eluted from the column after refolding, indicating possible problems with 6xHis-mediated Ni<sup>2+</sup>-NTA sepharose binding in the presence of guanidine hydrochloride. An overview about all attempts to avoid aggregate formation of MaFAR1-6xHis is given in Table 17.



**Figure 19: SEC of 6xHis-tagged MaFAR1 produced in *E. coli*.** **A)** Representative example for SEC of purified MaFAR1 in non-optimised buffer. The peak eluting at about 100 ml represents the void volume of the column, indicating an aggregated form of heterologous produced MaFAR1-6xHis. Data are representative for more than three independent purifications. Position of MaFAR1-6xHis is indicated by black arrow. **B)** SEC of MaFAR1-6xHis from cultures grown and purified under optimised conditions. The SEC buffer consisted of 30 mM CHES pH = 9.5, 150 mM NaCl, 1 mM EDTA, 5 mM MgCl<sub>2</sub>. The protein elutes both as aggregates and in a highly multimeric state. Data are representative for two independent purifications. The position of MaFAR1-6xHis is indicated by a black arrow.

## RESULTS

**Table 17: Strategies to optimise the multimeric state of heterologously produced C-terminal hexahistidine-tagged MaFAR1.** Numbers of purifications under the respective conditions are indicated on the right side of the table.

strategy	parameters		result	
on-column refolding	denaturation prior column binding	6 M guanidine hydrochloride	no elution of protein	n = 1
	denaturation after column binding	6 M guanidine hydrochloride	no elution of protein	n = 1
Ion strength	high salt	500 mM NaCl	aggregates	n = 1
	low salt	150 mM NaCl	aggregates	n > 3
pH	high pH	CHES pH = 9.0 / 9.5	aggregates + highly multimeric state	n = 2
	neutral pH	Tris pH = 7.6	aggregates	n = 3
	low pH	MES pH = 5.5	aggregates	n = 1
additives	detergent	CHAPS	aggregates	n > 3
	compatible solute	300 mM NaCl + 1 mM betaine during culture growth	no obvious precipitation of IMAC eluate	n > 3

## RESULTS

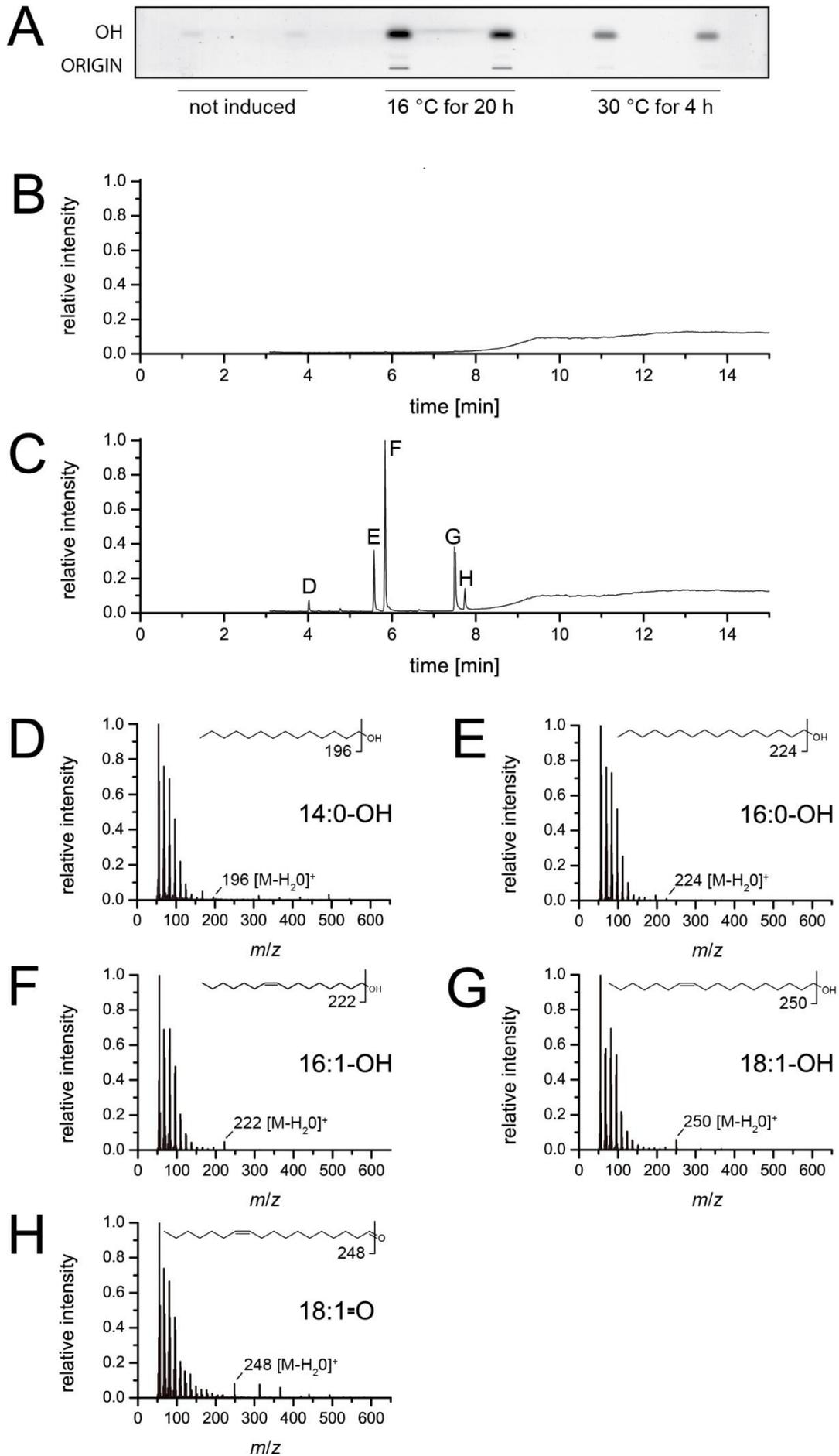
### 5.1.1.4.2 MaFAR1 is active *in vivo* and *in vitro*

Despite of the enzyme's aggregated state, heterologously produced MaFAR1 was active *in vivo* and *in vitro*. In the course of harvesting *E. coli* cells from MaFAR1 expressing cultures, an occurrence of whitish flakes, which floated on top of the medium after centrifugation at 4 °C, was noticed. Non-induced or empty vector expressing cultures did not show accumulation of these flakes. To elucidate the molecular nature of the flakes, apolar extracts of media samples from induced cultures expressing MaFAR1 and respective samples of not induced cultures were compared via TLC analyses. Additionally, apolar extracts of medium samples containing high amounts of respective flakes versus samples of the same medium containing no visible flakes were analysed via GC-MS.

TLC analyses showed a prominent band in samples derived from induced cultures, which co-migrated with a fatty alcohol standard. The band was only faintly visible in non-induced cultures (Figure 20 A). Moreover, GC-MS analyses of flakes-enriched medium showed several prominent peaks in the TIC, which were assigned as 14:0-OH, 16:0-OH, 16:1-OH, 18:1-OH and octadecenal (18:1=O) according to the characteristic fragmentation pattern of non-derivatised, free fatty alcohols and aldehydes (Figure 20 C-H). These fragmentation patterns show a characteristic species corresponding to [M-18]<sup>+</sup> (<http://lipidlibrary.aocs.org>). Respective fragments can be found in each of the peaks depicted in Figure 20. In contrast, the TIC of samples gained from media not containing whitish flakes or from cultures expressing empty vector controls did not show any prominent peaks (Figure 20 B). Thus, apart from its aggregated state upon purification, the MaFAR1-protein was able to synthesise fatty alcohols *in vivo*.

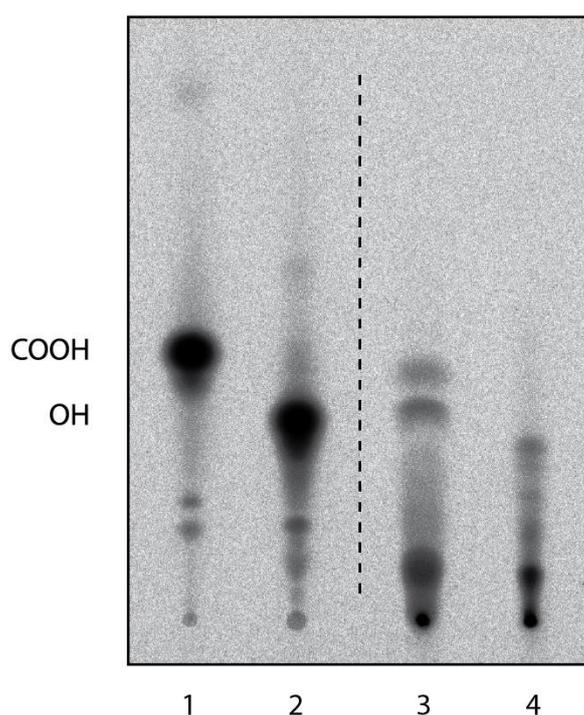
To elucidate the *in vitro* activity of purified MaFAR1 from *E. coli*, enzyme-mediated NADPH oxidizing activity was measured photometrically at 340 nm. No significant activity was observed in respective experiments (data not shown). Also DTNB-based activity assays at 410 nm, in which the liberation of free CoA upon acyl-CoA reduction is measured, did not yield any activity (data not shown). As an alternative to these kinetic methods, an end point activity assay with purified 6xHis-MBP-MaFAR1 and [<sup>14</sup>C] labelled fatty acyl-CoA substrates was performed. To visualise fatty alcohols produced by MaFAR1, the apolar phase of respective samples was extracted with a mixture from chloroform and methanol (1:1, v/v) and separated via TLC. Phosphoimager-derived data confirmed the presence of fatty alcohols in samples containing acyl-CoA, MaFAR1 and NADPH, but not in the negative control without enzyme (Figure 21). Notably, activity was only observed upon addition of at least 10 mM NADPH in the assays. Taken together, MaFAR1 was successfully expressed and purified from *E. coli* in an active form. However, considering the aggregation tendency of the purified enzyme showed in the tested conditions, further work was focused on enzymes which were more suitable for crystallisation trails.

## RESULTS



## RESULTS

**Figure 20: MaFAR1 is active *in vivo* upon expression in *E. coli*.** Upon cooling of MaFAR1 expressing *E. coli* cultures, fatty alcohols precipitate as whitish flakes which float on top of the medium surface. **A)** TLC analyses of apolar extracts of culture medium of independent *E. coli* BL21\* (DE3) cultures expressing pET24\_MaFAR1. In non-induced cultures, the signals representing fatty alcohols are only faint, while in induced cultures, significant amounts of fatty alcohols are accumulated. Data represent the result of a single experiment with samples derived from independent cultures each. **B)** GC-MS total ion count (TIC) of apolar extracts of medium from a pCold\_MaFAR1 expressing *E. coli* BL21\* (DE3) culture not containing whitish flakes. Data are also representative for apolar extracts from medium of cultures expressing empty vector controls. **C)** GC-MS TIC of apolar extracts from culture medium from a pCold\_MaFAR1 expressing *E. coli* BL21\* (DE3) culture containing high amounts of whitish flakes. Whitish flakes were observed during the majority of purifications of MaFAR1. Data shown here represent the results of a single experiment for pCold\_MaFAR1 and are furthermore representative for more than three samples derived from independent cultures expressing pET24\_MaFAR1. **D)** Mass spectrum of peak “D”, representing the fragmentation pattern of tetradecanol (14:0-OH). **E)** Mass spectrum of peak “E”, representing the fragmentation pattern of hexadecanol (16:0-OH). **F)** Mass spectrum of peak “F”, matching the fragmentation pattern of hexadecenol (16:1-OH). **G)** Mass spectrum of peak “G”, matching the fragmentation pattern of octadecenol (18:1-OH). **H)** Mass spectrum of peak “H”, matching the fragmentation pattern of octadecenal (18:1=O). The respective characteristic species corresponding to the [M-18]<sup>+</sup> fragment, as published at <http://lipidlibrary.aocs.org>, is indicated in each spectrum D) to H).



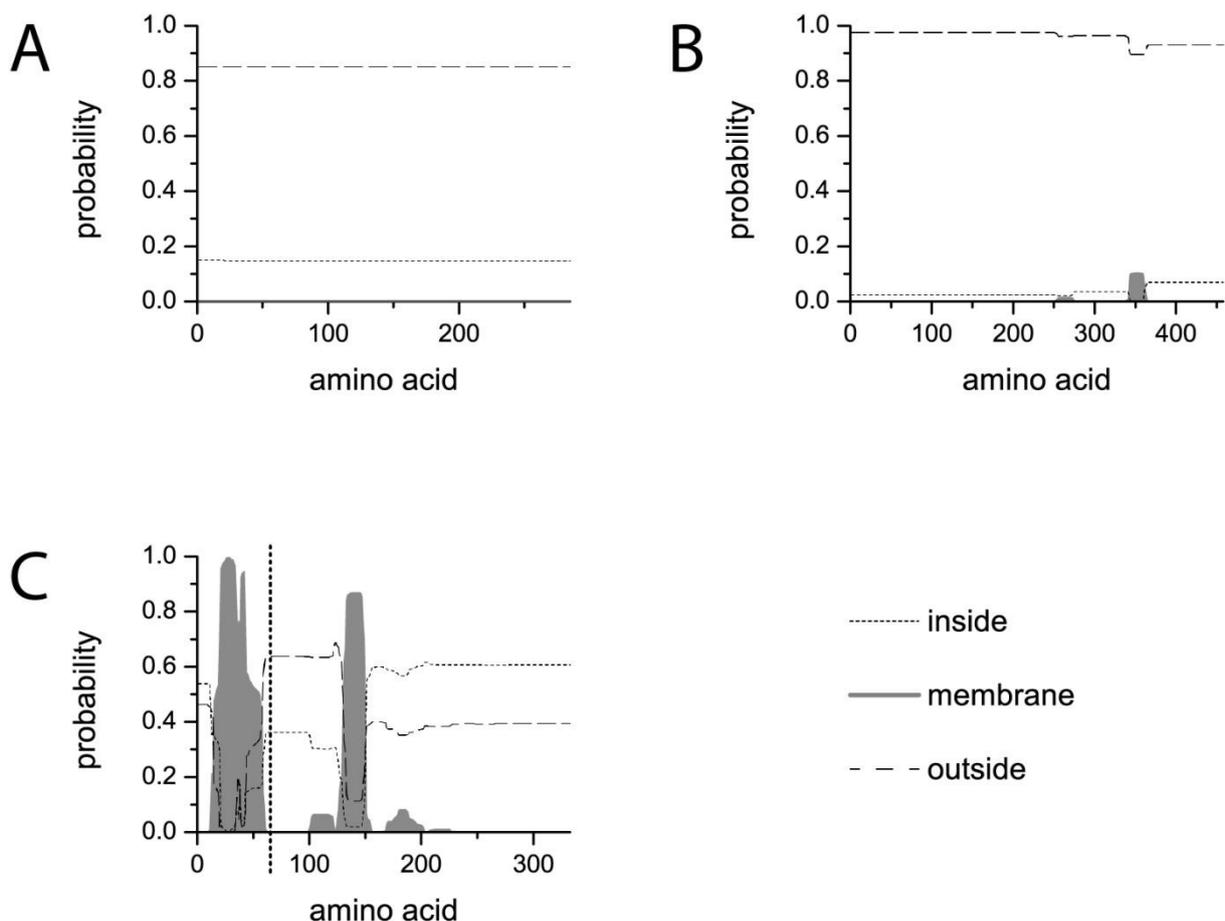
**Figure 21: Example of *in vitro* activity of purified MaFAR1.** **1** = [<sup>14</sup>C] octadecenoic acid standard, **2** = [<sup>14</sup>C] octadecenol standard, **3** = sample containing 5 µg of purified 6xHis-MBP-MaFAR1, 50 µM of [<sup>14</sup>C] octadecenyl-CoA, 10 mM NADPH and 3 mg/ml BSA in a total volume of 50 µl of 100 mM PO<sub>4</sub> buffer pH = 7. Samples were incubated for 30 min at 30 °C. **4** = negative control composed as 3 but without MaFAR1 enzyme. TLC was developed in hexane:diethyl ether:acetic acid (65:35:1, v/v/v), placed on an imaging plate (Fuji, Japan) for 2 days and analysed with a Fuji FLA-3000 (Raytest, Germany). All samples are derived from the same TLC, but unrelated samples between left and right side of dashed line have been removed by image processing. The image is representative for three independent experiments.

### 5.1.2 Purification and analyses of acyltransferases

The second step in WE biosynthesis is catalysed by WE synthases which belong to the group of acyltransferases. In the course of this project, three acyltransferases, originating from different organisms and classes of WSs, were analysed (Figure 7). Since DGAT1 like WSs are predicted to contain up to eight TM domains (Lardizabal et al., 2000; Liu et al., 2012), this group was excluded from purification attempts in this study. Instead, the two putatively soluble enzymes DGAT3 from *A. thaliana* and the bifunctional WS/DGAT from *A. baylyi* as well as the DGAT2-type WS AWAT2 from

## RESULTS

*M. musculus* were chosen for a detailed characterisation. For reasons of better understanding, results concerning purification of MmAWAT2 are shown later in section 5.3.9. Analogous to the bioinformatic analyses which were performed with the FAR sequences, the sequences of all three acyltransferases were submitted to the TMHMM online service (Sonnhammer et al., 1998) (Figure 22) and the Phyre server (Kelley and Sternberg, 2009) (Figure 47 A/data not shown). Moreover, a phylogenetic analysis with numerous WS sequences was conducted in order to generate a homology based classification of these enzymes. As expected, MmAWAT2 clusters in the group of vertebrate or DGAT2-like WSs, whereas AbWSD1 clusters in the bacterial-type WS subclade. In contrast, AtDGAT3 does not cluster in any of the known WS groups (Figure 7). This is not surprising, since until today only one study was published about the enzyme. Hence, it is not even known if the enzyme possesses acyltransferase activity at all (Hernandez et al., 2012).



**Figure 22: Predicted transmembrane structure of acyltransferases examined in this study.** Data for respective plots were generated with the TMHMM online service (Sonnhammer et al., 1998). **A)** AtDGAT3 **B)** AbWSD1 **C)** MmAWAT2. Dashed line in C) indicates the approximate site of truncation for expression of the C-terminal part of the enzyme (MmAWAT2ΔN) as described in 5.3.9.

### 5.1.2.1 AtDGAT3 from *A. thaliana*

Only a few members of the class of soluble DGAT enzymes have been described so far (Hernandez et al., 2012; Rani et al., 2010, 2012; Saha et al., 2006). One of them is the DGAT3 from *A. thaliana*, termed AtDGAT3. Since numerous DGATs are known to not only catalyse acylation of DAG, but also exhibit WS activity, a possible WS activity of the AtDGAT3 was elucidated in this study. The enzyme

## RESULTS

was described to be soluble (Hernandez et al., 2012), which is further supported by the result of a TM prediction based on its amino acid sequence. As shown in Figure 22 A, the probability for the presence of a TM domain within the protein is negligible throughout the whole sequence. Thus, the potential to efficiently purify the enzyme in a monodisperse state upon expression in *E. coli* is generally higher in comparison to an ER localised DGAT1 or DGAT2. Modelling of the enzyme structure using the Phyre server (Kelley and Sternberg, 2009) did only yield unsatisfactory results. The Phyre server threaded only 28 % of the residues at a confidence of over 90 % according to the structure of a thioredoxin-like [2FE-2S] ferredoxin from *Aquifex aeolicus* (PDB accession code = 1F37, Yeh et al., 2000). Moreover, the secondary structure prediction report indicated 40 % of the sequence to be disordered (data not shown).

In order to verify the DGAT activity of AtDGAT3, the enzyme coding sequence was amplified from cDNA of *A. thaliana* Col-0 and cloned into the pYES2/NT vector. The corresponding construct pYES2/NT\_AtDGAT3 was expressed in the *S. cerevisiae* strain H1246 under inducing conditions. The strain is devoid of all four genes responsible for neutral lipid synthesis (Sandager, 2001). As a consequence, all neutral lipids present in cultures expressing heterologous genes can directly be linked to the activity of respective gene products. However, expression of AtDGAT3 did neither yield detectable accumulation of TAGs nor WEs upon TLC analyses of lipid extracts from respective cultures (Appendix 12). Additionally, the enzyme was expressed in the corresponding wild type *S. cerevisiae* W303, which, in contrast to *S. cerevisiae* H1246, is capable of neutral lipid synthesis. Lipid extracts from respective cultures were compared with those of control cultures expressing the pYES2/NT empty vector. TLC analyses of extracts from three independent cultures each expressing either AtDGAT3 or the empty vector control showed an elevated TAG accumulation in the expression strain. However, the TAG signals for both, expression and empty vector control samples were rather low (Appendix 13). Immunodetection of the AtDGAT3 protein in yeast lysates was not performed. Hence, presence of the protein cannot be assured.

In addition to the experiments in yeast, the AtDGAT3 coding sequence was cloned into the bacterial expression vectors pET28 and pCold (Table 16). Expression of pCold\_AtDGAT3 in *E. coli* BL21\* (DE3) was done once and resulted in an accumulation of a prominent double band between 66 and 118 kDa (Appendix 14). This is in agreement with the theoretical mass of 6xHis-TF-AtDGAT3, which is 84 kDa. IMAC purification of the 6xHis-TF-AtDGAT3 yielded in accumulation of the band between 66 and 118 kDa in the elution fractions. Furthermore, a second prominent band between 43 kDa and 66 kDa was detected in those fractions upon SDS analyses.

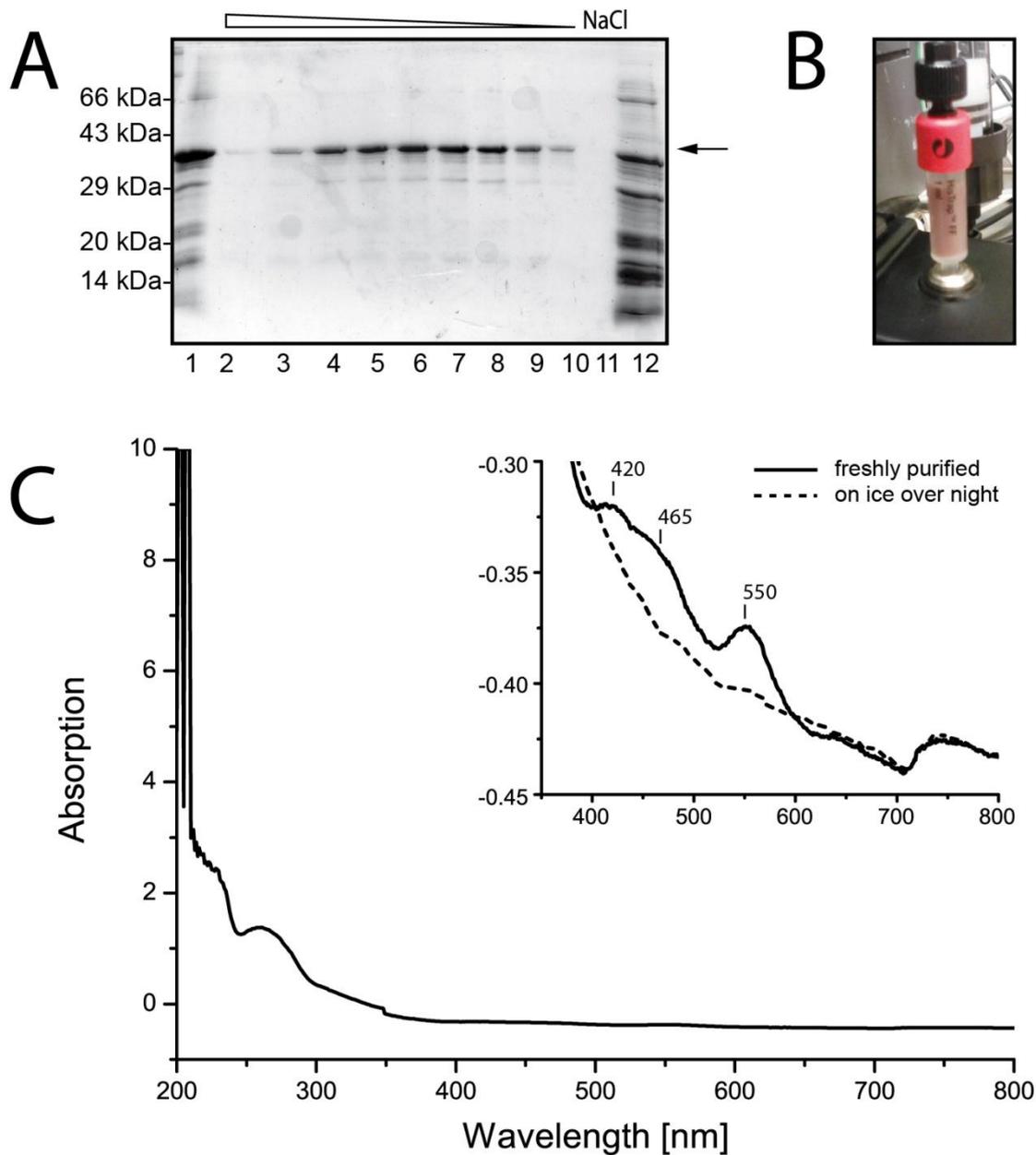
As judged from SDS-PAGE results, expression of pET28\_AtDGAT3 in *E. coli* resulted in a slight accumulation of a protein between 29 kDa and 40 kDa in respective cell lysates, which matched the calculated size of 31.8 kDa for 6xHis-AtDGAT3 (data not shown). The protein was purified via IMAC, resulting in elution fractions containing two major bands running little underneath the 29 kDa and between the 29 kDa and the 40 kDa band of the molecular weight marker, respectively (Appendix 15). The purity of the fractions was further enhanced by HIC, performed by Felix Lambrecht. The corresponding eluates contained almost pure AtDGAT3 protein (Figure 23 A). Interestingly, a brown colour accumulated on the Ni-sepharose column in case of purification of 6xHis-AtDGAT 3 (Figure 23 B). Consequently, also the 6xHis-AtDGAT3 elution fractions of the IMAC had a brown colour, while the 6xHis-TF-AtDGAT3 elution fractions were colourless. Upon photometric analyses of 6xHis-AtDGAT3 samples, a UV/VIS spectrum showing the spectral properties of a [2Fe2S]-cluster (Yuvaniyama et al., 2000) was obtained with shoulders at approximately 420 nm and 465 nm as well as a peak at 550 nm (Figure 23 C). However, after storage of the respective elution fractions on ice overnight, the colour disappeared, resulting in a completely colourless solution. Spectra of samples

## RESULTS

which lost their colour did not show any characteristic UV/VIS spectra (Figure 23 C, dashed line). BLAST analyses of the AtDGAT3 amino acid sequence revealed the presence of a possible [2Fe2S] cluster binding site (Appendix 10). Moreover, the only part of the protein which was modelled with high confidence is homologous to a thioredoxin-like [2Fe-2S] ferredoxin from *A. aeolicus* (PDB accession code 1F37), further supporting the presence of a [2Fe2S]-cluster.

Having the purified protein in hand, further investigations regarding a possible acyltransferase activity of AtDGAT3 were addressed by performing *in vitro* experiments. Most *in vitro* experiments were done together with Felix Lambrecht. First attempts to photometrically show a possible AtDGAT3-mediated cleavage of acyl-CoA substrates via the DTNB assays did not indicate any activity (data not shown). To circumvent possible issues caused by a low sensitivity of this assay, further experiments using [<sup>14</sup>C] labelled substrates were done. Since the initial publication on AtDGAT3 indicated a specificity towards linoleic (18:2) and linolenic (18:3) acid (Hernandez et al., 2012), 1,2-dilinoleoylglycerol and [<sup>14</sup>C] labelled linoleoyl-CoA (18:2-CoA) were used as substrates in the assays. The experiments were done with freshly prepared protein, which still showed a brownish colour. The assay was done three times with different amounts of DAG (30 µM, 60 µM and 200 µM), since an excess of DAG was shown to impair the enzymatic activity of a DGAT3 from peanut (Saha et al., 2006). After the reaction, samples were extracted with hexane twice and analysed via TLC in case of utilization of 30 µM and 60 µM DAG. In case of utilization of 200 µM DAG, the complete samples were applied on a TLC plate without extraction. No signals above the origin of migration were detected in any of the samples, regardless of the DAG concentration (Appendix 16). Therefore, the work on the enzyme was stopped in favour of proteins proven to be active.

## RESULTS



**Figure 23: Purification and spectral properties of AtDGAT3.** **A)** SDS-PAGE showing HIC purified 6xHis-AtDGAT3. 1 = Load, derived from purification of 6xHis-AtDGAT3 via IMAC. 3-11 = Elution fractions of 6xHis-AtDGAT3 from HIC. 12 = flow-through of HIC. Sample was loaded on a 1 ml HiTrap PhenylHP column (GE Healthcare, USA) in buffer containing 30 mM HEPES/NaOH pH = 7.5, 2 M NaCl, 10 % glycerol (HIC buffer A) and eluted with a gradient of 20 ml length from HIC buffer A to HIC buffer B (30 mM HEPES/NaOH pH = 7.5, 10 % glycerol). Grade of purity is representative for two separate HIC experiments, whereas the second experiment was conducted with 3 M NaCl in HIC buffer A. The position of 6xHis-AtDGAT3 is indicated by a black arrow. **B)** IMAC column with bound AtDGAT3 after washing of the column to remove unbound protein. Column shows the brownish colour of the AtDGAT3 protein. Image is representative for more than three independent purifications. **C)** UV/VIS spectra of purified AtDGAT3 before (solid line) and after incubation on ice overnight (dashed line). Spectra represent data obtained with protein from two independent purifications.

### 5.1.2.2 AbWSD1 from *A. baylyi*

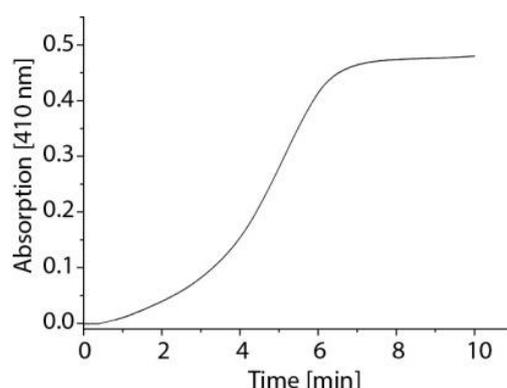
The AbWSD1 from *A. baylyi* developed to become a role model for bacterial-type, bifunctional WS/DGAT enzymes in the recent years. Due to the fact that bacterial WS/DGAT (WSD) are reported to be soluble enzymes, they may be efficiently purified from *E. coli* expression cultures (Stöveken et al., 2005). Thus, research on this enzyme is relatively advanced in comparison to WSs of other

## RESULTS

classes, which are typically integral membrane proteins. As expected, hydrophobicity analyses of the amino acid sequence of AbWSD1 did not indicate the presence of TM domains in the protein, but revealed a rather hydrophobic patch within the sequence only. The probability for a TM domain in this region, which spans from residue L344 to I359, is merely 10 % throughout (Figure 22 B). Being a soluble enzyme, AbWSD1 was considered to be a highly promising candidate for crystallisation studies.

### 5.1.2.2.1 Purification of a fusion protein from TF and AbWSD1 from *A. baylyi*

The coding sequence of AbWSD1, optimised for the codon usage of *E. coli*, was cloned in the pCold vector system and expressed in *E. coli*. The size of the fusion protein of 6xHis-TF and AbWSD1 (6xHis-TF-AbWSD1) was calculated to be 106 kDa. The enzyme was purified from respective *E. coli* BL21\* (DE3) expression cultures via IMAC. Upon elution of bound protein from the Ni-NTA agarose column, high amounts of almost pure protein were obtained, showing only minor contaminations. The protein ran at about 110 kDa and thus matched the calculated size of the 6xHis-TF-AbWSD1 fusion protein (representatively shown in Figure 26 A, lane 1). The *in vitro* activity of the purified protein was detectable via the DTNB assay (Figure 24) and was thus efficiently verified after purification.



**Figure 24:** *In vitro* activity of IMAC purified 6xHis-TF-AbWSD1 can be monitored by the DTNB assay. The assays contained 0.2 mg DTNB/ml, 20  $\mu$ M fatty alcohol, 2  $\mu$ l of purified protein solution and 12.5  $\mu$ M of acyl-CoA in a total volume of 1 ml TBS. Data are representative for more than three experiments with protein derived from independent purifications.

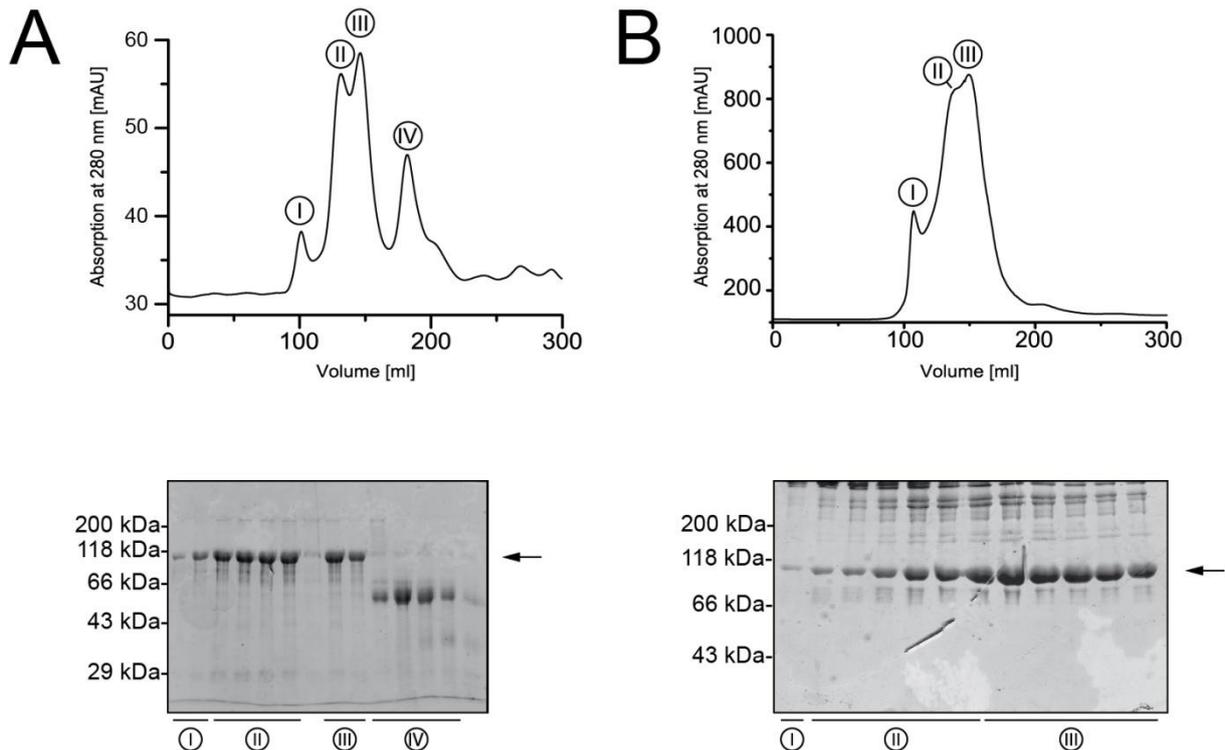
Initial SEC analyses of the 6xHis-TF-AbWSD1 IMAC eluates resulted in four peaks. The elution volumes of these peaks corresponded to aggregated protein eluting in the void volume of the column (peak I), at about 390 kDa (peak II), at about 315 kDa (peak III) and at about 130 kDa (peak IV) (Figure 25 A). Peak III (315 kDa) may correspond to a trimer of 6xHis-TF-AbWSD1 (3 x 106 kDa). Notably, SDS-PAGE analyses did not show any obvious differences among fractions of peaks I-III (Figure 25). According to SDS-PAGE analyses, peak IV contained protein with a molecular weight between 43 kDa and 66 kDa (Figure 25 A). The elution volume corresponding to 130 kDa may thus represent dimers of the sole AbWSD1 or the sole TF, respectively. Moreover, it may also represent the cleaved fusion protein of TF and AbWSD1. All three species would correspond to an elution volume of a molecular weight of approximately 130 kDa and migrate at approximately 60 kDa in SDS-PAGE analyses. However, the respective fractions did not show activity upon DTNB-based *in vitro* tests (data not shown) and thus, peak IV most likely represents a TF dimer.

To increase the protein stability, optimal buffer conditions were determined via a ThermoFluor approach. Analyses of the measured data indicated highest stability of 6xHis-TF-AbWSD1 in 50 mM

## RESULTS

glycine/NaOH pH = 10.0. The purification conditions of 6xHis-TF-AbWSD1 were further optimised by Felix Lambrecht. He noticed, that upon purification of 6xHis-TF-AbWSD1 from freshly harvested, non-frozen cultures, the 130 kDa peak IV was absent in SEC analyses (Figure 25 B).

To further verify the molecular weights of proteins eluted during SEC, multiangular light scattering (MALS) was performed with SEC eluates directly upon elution from the column. Respective measurements were done with the help of Dr. Achim Dickmanns and resulted in the detection of four major protein species corresponding to 1695 kDa, 297 kDa, 75 kDa and 9 kDa, respectively (Appendix 18). Thus, peak III in Figure 25 may indeed correspond to a trimeric state of the 6xHis-TF-AbWSD1 protein, while peak I represents highly oligomeric protein aggregates.



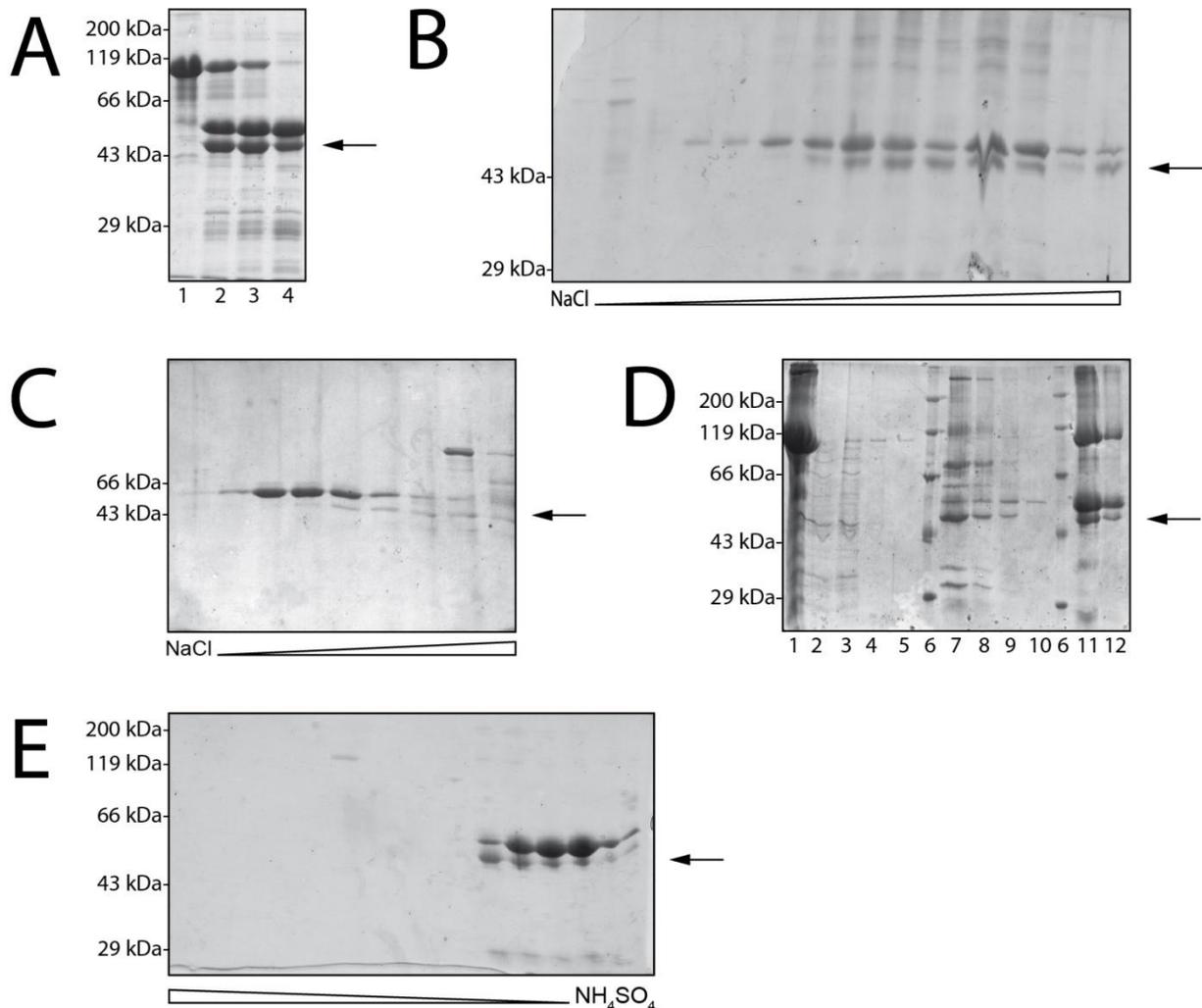
**Figure 25: SEC of the 6xHis-TF-AbWSD1 fusion protein. A)** SEC of 6xHis-TF-AbWSD1 fusion protein under non-optimised conditions. Peak I refers to the void volume of the SEC column and thus represents the aggregate peak. Peak II refers to a calculated molecular weight of about 390 kDa, while peak III corresponds to about 315 kDa and peak IV to about 130 kDa. Position of 6xHis-TF-AbWSD1 is indicated by a black arrow. **B)** SEC of 6xHis-TF-AbWSD1, purified under optimised conditions. Position of 6xHis-TF-AbWSD1 is indicated by a black arrow. Data are representative for more than three purifications.

### 5.1.2.2.2 Separation of TF and AbWSD1 via chromatographic techniques

The probability for the formation of protein crystals is higher, the less intramolecular flexibility the respective proteins possess. A protein composed of a single domain is thus in general more likely to form crystals than a protein containing multiple domains. In the present case, the fusion of the *E. coli* derived TF with the bacterial AbWSD1 may potentially hinder successful crystallisation, since both proteins are connected by an unstructured, flexible linker region. Moreover, the TF itself consists of multiple domains (Ferbitz et al., 2004) (Figure 27 B), which may likewise interfere with crystallisation of the fusion protein due to an increased flexibility.

## RESULTS

Having these possible issues in mind, attempts to separate both parts of the fusion protein after cleaving in order to obtain pure AbWSD1 protein for crystallisation studies were conducted (Figure 26).



**Figure 26: Attempts of separation of 6xHis-TF and AbWSD1 by different chromatographies.** The position of AbWSD1 in the respective pictures is indicated by black arrows. **A)** Cleavage of 100 µg 6xHis-TF-AbWSD1 fusion protein with 1 µg Factor Xa at 37 °C. 1 = uncleaved protein, 2 = after 1 h of incubation, 3 = after 2 h of incubation, 4 = after 16 h of incubation. Data are representative for more than three experiments. Treatment of the protein with thrombin at a concentration of 2 U thrombin / mg fusion protein at 16 °C for 16 h resulted in a similar cleavage of the fusion protein. **B)** Cation exchange chromatography with cleaved 6xHis-TF-AbWSD1 fusion protein using a 40 ml gradient from 50 mM Bis-Tris/HCl pH = 6.0, 10 % glycerol to 50 mM Bis-Tris/HCl pH = 6.0, 10 % glycerol, 1 M NaCl. **C)** Anion exchange chromatography with cleaved 6xHis-TF-AbWSD1 fusion protein using a 50 ml gradient from 50 mM Tris/HCl pH = 8.0, 10 % glycerol to 50 mM Tris/HCl pH = 8.0, 10 % glycerol, 1 M NaCl. **D)** Representative example for on-column cleavage of 250 µg of 6xHis-TF-AbWSD1 bound to Ni<sup>2+</sup>-sepharose. Uncleaved protein (lane 1) was incubated with Ni<sup>2+</sup>-sepharose resin, gently shaking at 4 °C over night. The resin was then washed with buffer (50 mM Tris/HCl pH = 7.5, 500 mM NaCl, 0.1 % Triton X-100) (lanes 2-5). 10 U of thrombin were applied to the sample and incubated for 6 h at 37 °C. The resin was washed with buffer (lanes 7-10) to recover cleaved AbWSD1. Bound TF and uncut fusion protein were eluted using buffer containing 500 mM imidazole (lanes 11-12). 6 = molecular weight marker **E)** HIC with cleaved 6xHis-TF-AbWSD1 fusion protein using a 30 ml gradient from 50 mM Tris/HCl pH = 7.0, 1.5 M NH<sub>4</sub>SO<sub>4</sub> to 50 mM Tris/HCl pH = 7.0 on a Source15PHE column (GE Healthcare, USA). Data shown in B-E are representative for single experiments with the specific conditions described in the text. All conditions for the respective chromatographic methods are listed in Table 15.

The linker between the TF and the actual AbWSD1 protein in the 6xHis-TF-AbWSD1 fusion protein contains recognition sites for thrombin, the HRV 3C protease as well as for the factor Xa protease. Incubation of factor Xa with the fusion protein in a ratio of 1:100 (w/w) resulted in almost complete

## RESULTS

cleavage of the protein upon incubation at 37 °C for 16 h (Figure 26 A). SDS-PAGE analyses of respective samples clearly showed the presence of two major bands in the height of the calculated size of the single AbWSD1 protein (54.08 kDa) and the single TF (51.9 kDa). Treatment of the protein with thrombin at a concentration of 2 U thrombin/mg fusion protein at 16 °C for 16 h resulted in a similar cleavage of the fusion protein (data not shown). In both cases, the two bands for AbWSD1 and TF migrated in close proximity upon SDS-PAGE analyses.

Since both, 6xHis-TF and AbWSD1, are almost equal in their molecular weights, separation of both constituents via SEC is not feasible. Instead, separation of the two proteins was tried by anion and cation exchange chromatography, HIC as well as IMAC. Despite of the alteration of numerous parameters for each type of chromatographic method, it was not possible to separate TF and AbWSD1 satisfactorily. As representatively shown for each chromatographic method in Figure 26 B-E, both proteins coeluted from the respective column in every case. A complete list of conditions and chromatographic steps which were carried out in order to separate TF and AbWSD1 is given in Table 15.

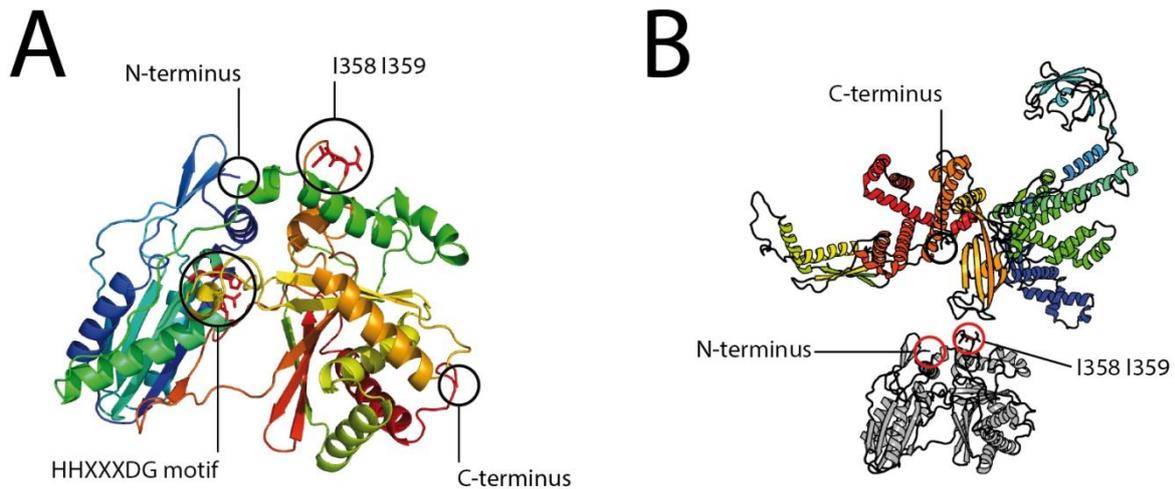
Identification of the band corresponding to AbWSD1 in SDS-PAGE gels was done by DTNB-based activity tests. Fractions, which only contained the upper one of the two major bands after protease cleavage, did not show activity. In contrast, fractions containing proteins of both bands were active (data not shown). Hence, AbWSD1 in Figure 26 is represented by the lower one of the two major bands, as indicated by arrows.

### 5.1.2.2.3 Structure prediction of AbWSD1

As it was not possible to separate 6xHis-TF and AbWSD1 by chromatographic techniques, possible reasons for the obvious affinity of both enzymes towards each other were considered.

Today, it is already known that TF forms a cradle-like shape, in which newly emerging peptides, coming from the ribosome, can fold in an enclosed environment (Ferbitz et al., 2004). It was furthermore shown, that these nascent peptides bind to numerous defined, TF-derived residues in the course of entering the cradle-like structure. Consequently, it is reasonable to assume that hydrophobic patches of AbWSD1 bind tightly to TF, making a chromatographic separation difficult.

In order to reveal possible hydrophobic patches on the surface of the AbWSD1 protein, a structure prediction of the enzyme was carried out using the Phyre server (Kelley and Sternberg, 2009). The system determined the best matching structure for threading of AbWSD1 to be surfactin synthetase subunit 3 (PDB accession code 2VSQ) (Tanovic et al., 2008). The confidence for true homology between the AbWSD1 sequence and the surfactin synthetase subunit 3 was 100 %. Phyre moreover modelled 96 % of the AbWSD1 residues with over 90 % confidence. The generated model shows a two domain architecture of the AbWSD1 protein, which divides the sequence in an N-terminal and a C-terminal domain. Both domains are arranged to each other in a shape resembling a capital lambda letter ( $\Lambda$ ) (Figure 27 A), with the “top” of the two domains connected to each other and a gap between both domains at the “bottom”. Both parts are connected via an unstructured loop. Interestingly, the designated HHXXXDG active motif of the protein is located in the N-terminal domain, although it is encoded on a C-terminal part of the amino acid sequence (Figure 27 A). It resides in the gap between the two domains, which in turn may form the binding pocket for substrates.



**Figure 27: Structure prediction of AbWSD1. A)** Model of AbWSD1 generated with the Phyre server (Kelley and Sternberg, 2009) according to a surfactin synthetase subunit 3 (PDB accession code 2VSQ) (Tanovic et al., 2008). N-terminus, C-terminus and the catalytic HHXXXDG motif are indicated. Isoleucines at positions 392 and 393 form a hydrophobic patch which may have an impact on the aggregation state of heterologously produced protein from *E. coli*. **B)** TF (upper structure, coloured, PDB accession code 1W26, (Ferbitz et al., 2004)) and the predicted structure of AbWSD1 (lower structure, grey) would allow the TF to shield both isoleucines at position 392 and 393 in AbWSD1 upon expression as a fusion protein. By shielding both residues, TF might prevent interaction of the residues with other proteins and thus help to avoid aggregate formation.

Since hydrophobicity predictions indicated the presence of a rather hydrophobic spot within the AbWSD1 (Figure 22 A), the localisation of this patch within the structure was analysed by Felix Lambrecht. The respective part of the sequence is located in a surface exposed, unstructured loop according to the Phyre model of AbWSD1. The most hydrophobic residues in this patch are two isoleucine residues at position 358 and 359 as well as the two valin residues at position 349 and 350. All four residues are predicted to be highly exposed to the solvent (Appendix 17, Figure 27). Upon *in silico* exchange of both isoleucine residues at positions 358 and 359 to serine, TMHMM-based hydrophobicity prediction (Sonnhammer et al., 1998) did not show any hydrophobic patches throughout the sequence any more. The probability for the presence of a TM domain in the respective area of the sequence for the V349S V350S variant decreased to approximately 1.1 % (data not shown). In order to elucidate possible effects of the mutations on the structure, the sequence of AbWSD1 I358S I359S was submitted to the Phyre service as well. The resulting model of AbWSD1 I385S I359S was highly similar to the wild type version of AbWSD1. According to the model, S358 and S359 were still exposed to the solvent in the variant. The only difference among both models was an  $\alpha$ -helical structure in the wild type, which was missing in the variant and predicted to be an unordered loop instead (Appendix 17).

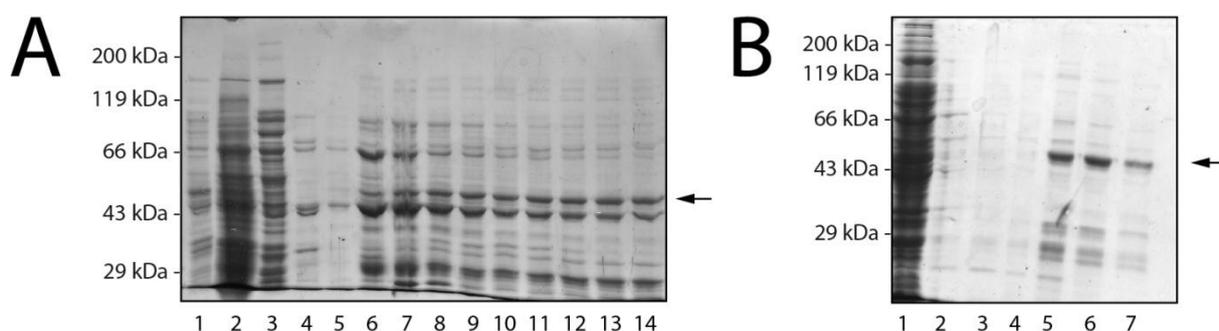
#### 5.1.2.2.4 Purification of hexahistidine-tagged AbWSD1 from *A. baylyi*

In order to avoid problems upon crystallisation of the 6xHis-TF-AbWSD1 fusion protein due to a probably high flexibility caused by the unstructured linker between both proteins, the AbWSD1-coding sequence was cloned into the pET28-vector. Expression of the construct results in an N-terminal hexahistidine-tagged version of the protein. Initial attempts to purify the protein via IMAC, done by Felix Lambrecht, resulted in impure elution fractions, in which the band estimated to represent the AbWSD1 enzyme was of the same intensity as two other bands migrating little

## RESULTS

underneath the AbWSD1 band and at approximately 30 kDa, respectively (Figure 28 A). Interestingly, elution fractions instantly precipitated upon elution from the IMAC column, resulting in a whitish, milky solution.

Since *in silico* analyses indicated positions I358 and I359 within the AbWSD1 enzyme to be highly exposed to the surface and thus represent a possible oligomerisation domain, the AbWSD1 I358S I359S variant was constructed. In a first IMAC purification experiment of a respective *E. coli* cell lysate, the corresponding elution fractions did not precipitate. Additionally, SDS-PAGE analyses of the elution fractions indicated a higher purity in comparison to purification of the wild type AbWSD1, since except for the major AbWSD1-band, only less intense bands around 30 kDa were detected as impurities (Figure 28 B). Taken together, these results strongly indicate a major role of positions I358 and I359 in the observed aggregation pattern of wild type AbWSD1. Work on the AbWSD1 I358S I359S variant has begun in a late phase of this study, thus the results shown here are incomplete. Nevertheless, the this variant is a highly promising candidate for crystallisation studies on the basis of the separate AbWSD1 protein, since possible issues accompanied with crystallisation of fusion proteins would be avoided.



**Figure 28: Purification of hexahistidine-tagged AbWSD1. A)** IMAC purification of 6xHis-AbWSD1 in 50 mM Tris/HCl pH = 7.5, 150 mM NaCl, 30 mM imidazole. Elution of bound protein was done in the same buffer without NaCl and 500 mM imidazole. 1 = whole cell lysate of *E. coli* culture expressing 6xHis-AbWSD1, 2 = insoluble fraction of cell lysate, 3 = soluble fraction of cell lysate, 4 = protein eluted upon washing with 30 mM imidazole, 5 = protein eluted upon washing with 5 mM ATP, 5 mM MgCl<sub>2</sub>, 6-14 = protein eluted with 500 mM imidazole. Data are representative for more than three independent purifications. **B)** IMAC purification of the 6xHis-AbWSD1 I358S I359S variant in 50 mM Tris/HCl pH = 7.5, 300 mM NaCl, 30 mM imidazole. Elution of bound protein was done in the same buffer with 500 mM imidazole. 1 = whole cell lysate of *E. coli* culture expressing pET28a\_AbWSD1 I358S I359S, 2-7 = protein eluted with 500 mM imidazole. The position of 6xHis-AbWSD1 I358S I359S is indicated by a black arrow. Data are representative for a single purification.

### 5.2 Crystallisation of the 6xHis-TF-AbWSD1 fusion protein

Full length 6xHis-TF-AbWSD1 fusion protein was used for crystallisation because it was not possible to separate 6xHis-TF and AbWSD1 after protease-mediated cleavage. TF is a member of the group of ribosome associated chaperones (Kramer et al., 2002) and has a physiological role in assisting nascent peptides in correct folding (Deuerling et al., 1999). Thus, its presence during crystallisation might have a beneficial effect on crystallisation, which may even compensate disadvantages associated with the flexible domain architecture. Hydrophobic patches of AbWSD1 could be shielded from interaction with hydrophobic parts of other proteins by the TF, preventing the formation of aggregates and possibly promoting crystal formation (Figure 27 B).

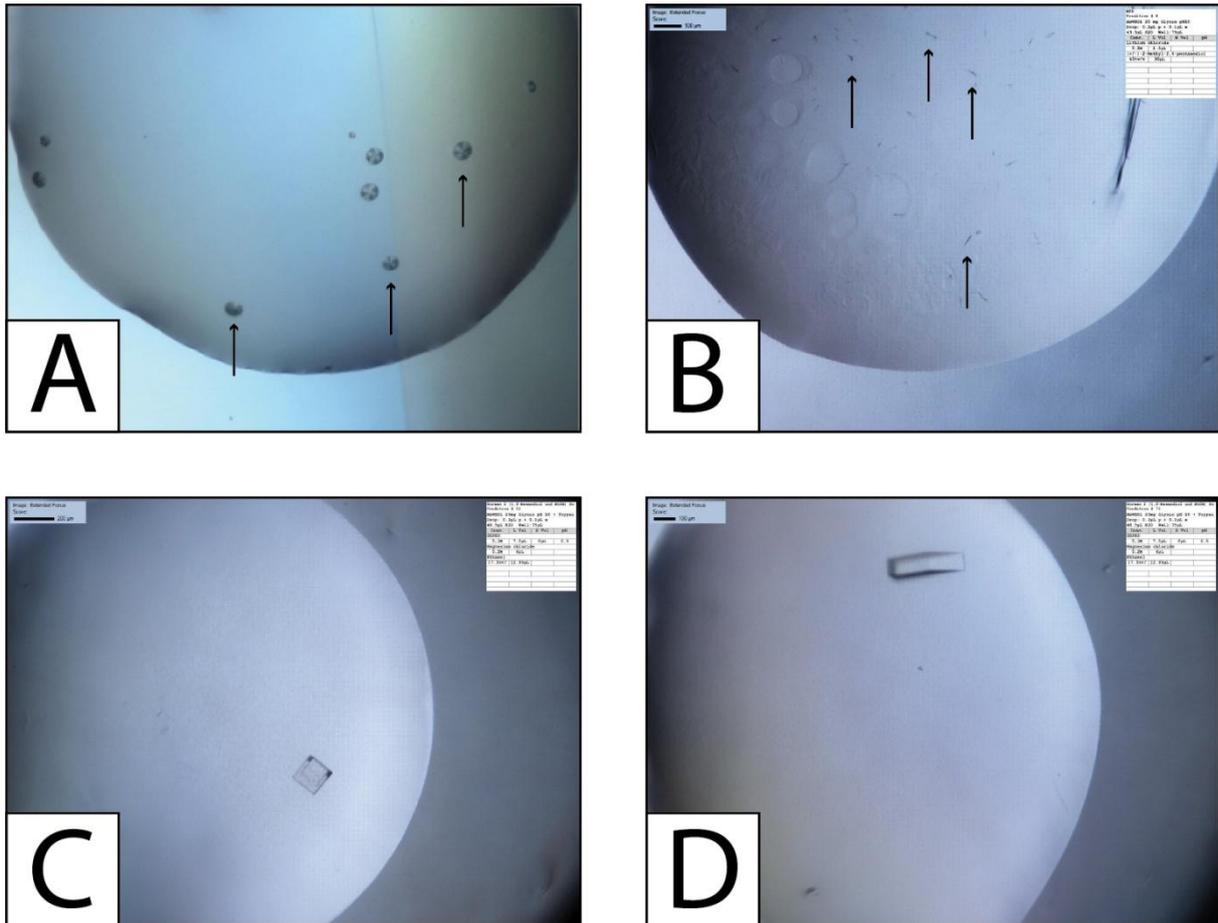
All crystallisation experiments were done with the assistance of Dr. Karin Kühnel. Most of the crystallisation screens were pipetted by Felix Lambrecht. The screening process was started with commercially available screens in a 96-well format (Table 5) using a Cartesian pipetting robot. In all

## RESULTS

conditions containing  $\text{Ca}^{2+}$ -ions, spherulites were formed (Figure 29 A). Whether those spherulites were protein or derived from precipitation of calcium glycinate, originating from the glycine buffer, remains elusive. In addition, needle or snowflake-like structures were formed in numerous conditions after an average of approximately one week (representatively shown in Figure 29 B). Almost all of these conditions contained 0.2 M  $\text{MgCl}_2$ . Moreover, isopropanol, 1,6-hexandiol, ethanol, *tert*-butanol, 2-methyl-2,4-pentadiole (MPD) as well as different Jeffamines were promising precipitants in many cases. Except for the Jeffamines, which are polyetheramines, all of these compounds are alcohols. Since alcohols are substrates of AbWSD1, it is conceivable that the enzyme could be stabilised by these precipitants. The crystalline structures found in the commercial screens did not grow bigger and were too small for further analyses. Hence, grid screens (listed in 3.6) were designed to improve crystal formation by systematically varying pH values and the precipitant concentrations. In the case of ethanol, two crystals grew, which both had a size of at least 100  $\mu\text{m}$  in each dimension (Figure 29 C&D). Both crystals were formed in conditions containing 17.3 % (v/v) of ethanol, 0.2 M  $\text{MgCl}_2$  as well as trypsin in a protease/protein ratio of 1:200 (w/w) in a sitting drop 96-well plate. Both drops were pipetted with 200 nl protein solution (20 mg/ml in 20 mM glycine/NaOH pH = 10, 150 mM NaCl) and 100 nl of precipitant. The buffers in the two conditions were 0.1 M HEPES pH = 6.5 (Figure 29 C) and 0.1 M Tris/HCl pH = 8.0 (Figure 29 D), respectively. Crystals were first detected in images acquired after 56 days of incubation at 4 °C and were reproducibly obtained under the mentioned conditions thereafter. Crystals were transferred into a cryoprotectant composed of precipitant supplemented with 30 % ethylene glycol to prevent ice formation and flash-cooled in liquid nitrogen. Diffraction data were collected by Dr. Karin Kühnel at beamline X10SA at the Swiss Light Source in Villigen, Switzerland. The crystal grown in 0.1 M Tris/HCl pH = 8.0, 0.2 M  $\text{MgCl}_2$ , 17.3 % EtOH, trypsin, 1:200 (w/w) (Figure 29 D) diffracted to a resolution of up to 2.1 Å (Figure 30). The crystal belonged to the centred tetragonal space group I442, with unit cell dimensions of  $a = b = 117.1$  Å,  $c = 141.2$  Å and  $\alpha = \beta = \gamma = 90$  °. Statistics of the collected dataset are shown in Table 18. The crystal which grew in a similar condition with 0.1 M HEPES pH = 6.5 did not diffract, probably due to damage in the freezing process.

The unit cell content can be analysed by calculation of the Matthews Coefficient ( $V_M$ ). This gives an estimation of the solvent content and the number of molecules present in the asymmetric unit. The asymmetric unit is the smallest fraction of the unit cell and by applying of the symmetry operations, the entire unit cell is assembled from an asymmetric unit. In space group I442, the unit cell is composed of 16 asymmetric units.  $V_M$  is the ratio of the unit cell volume divided through the molecular weight of the protein multiplied with the number of asymmetric units and the number of molecules in the asymmetric unit. Based on the calculated Matthews Coefficient ( $V_M$ ), it was concluded that a crystallisation of the complete 6xHis-TF-AbWSD1 fusion protein in the diffracting crystal was very unlikely. Under the measured conditions,  $V_M$  for a single molecule with a molecular weight of 106 kDa per asymmetric unit is 1.14. This  $V_M$  corresponds to a solvent content of merely 7.74 %, a value too low to represent a realistic solvent content of a protein crystal. Typically, protein crystals have a solvent content in the range of 25-80 %.

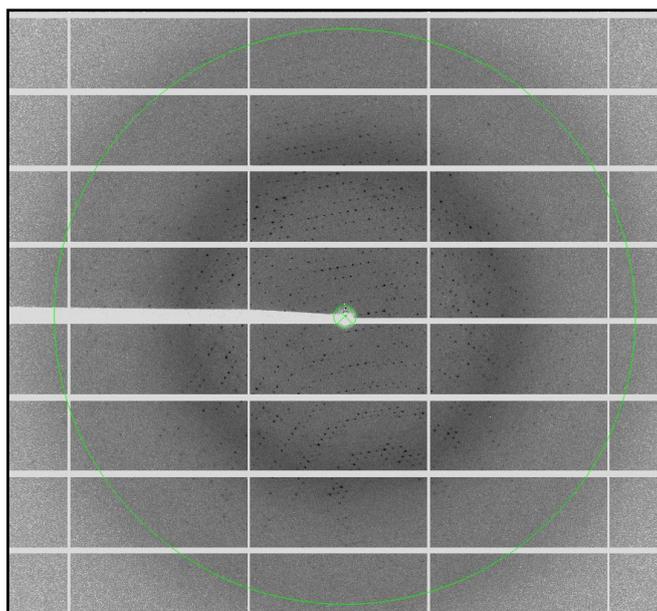
## RESULTS



**Figure 29: Crystallisation of the 6xHis-TF-AbWSD1 fusion protein.** **A)** In all conditions containing  $\text{Ca}^{2+}$ , spherulites were formed (indicated by black arrows). **B)** Example of small, spider-like crystals (indicated by black arrows). Crystals shown in **C)** (0.1 M HEPES pH = 6.5, 0.2 M  $\text{MgCl}_2$ , 17.3 % EtOH, trypsin, 1:200 (w/w)) and **D)** (0.1 M Tris/HCl pH = 8.0, 0.2 M  $\text{MgCl}_2$ , 17.3 % EtOH, trypsin, 1:200 (w/w)) were tested for diffraction at the Paul Scherrer Institute in Villigen, Switzerland. The crystal shown in D) diffracted to a resolution of 2.1 Å.

Under the given parameters,  $V_M$  for a protein of 52000 Da is 2.33, which corresponds to a calculated solvent content of 47.2 %. Hence, it is very likely that the asymmetric unit of the crystal contains either a single AbWSD1 molecule or a single TF molecule, since both proteins are similar in size. In order to determine the structure of the crystallised protein, molecular replacement using the *Vibrio cholera* TF structure (PDB accession code 1T11) (Ludlam et al., 2004) and the *E. coli* TF structure (PDB accession code 1W26) (Ferbitz et al., 2004) was tried, but attempts were not successful. Moreover, molecular replacement was tried with the surfactin synthetase subunit 3 (PDB accession code 2V5Q) (Tanovic et al., 2008), the structure which was chosen by Phyre (Kelley and Sternberg, 2009) for modelling of the AbWSD1 sequence. However, molecular replacement based on this structure was not successful either. Soaking of the crystals with iodide for single wavelength anomalous diffraction (SAD) phasing resulted in dissolving of the crystals. The goal was then to determine structure through SAD phasing with selenomethionine labelled protein. The protein was expressed in minimal medium supplemented with selenomethionine. Purification of 6xHis-TF-AbWSD1 carrying selenomethionine under the same conditions as described for the unlabelled 6xHis-TF-AbWSD1 produced sufficient amounts of protein for further crystallisation screenings (data not shown). However, respective crystallisation screenings were ongoing at the time of writing this thesis, but hadn't been finished.

## RESULTS



**Figure 30: X-ray diffraction pattern of 6xHis-TF-AbWSD1.** The respective protein crystal of 6xHis-TF-AbWSD1 was grown in 0.1 M Tris/HCl pH = 8.0, 0.2 M MgCl<sub>2</sub>, 17.3 % EtOH, trypsin, 1:200 (w/w).

**Table 18: Statistics of the X-ray diffraction data set which was collected for the 6xHis-TF-AbWSD1 protein crystal.**

parameter	value
Beam line	SLS X10SA
Detector distance (mm)	270
$\varphi_{\text{start}}$ / oscillation $\Delta \varphi$ (°)	259° / 0.5°
Exposure time (s)	0,5 s
Beam intensity	40 %
$\lambda$ (Å)	0.979 Å
No. of frames	180
Space group	I422
<b>Cell dimensions</b>	
a, b, c (Å)	117.1, 117.1, 141.2
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90
Resolution (highest res. shell) (Å)	50-2.10 (2.22-2.10)
R <sub>merge</sub> (%)	7.7 (60.2)
no. of observed reflections / unique reflections	381836/29155 (58900/4565)
I/ $\sigma$ (I)	26.6 (4.3)
Completeness (%)	99.6 (98.6)
Wilson B factor (Å <sup>2</sup> )	38.4

### 5.3 Elucidation of substrate determinants in a murine WS

The majority of WSs are integral membrane proteins. Heterologous expression in bacteria and subsequent purification of such enzymes are often accompanied with difficulties due to their hydrophobic nature. In order to get insights in structure-function relationships of integral membrane WSs anyhow, *in vivo* analyses of corresponding enzyme variants are an attractive alternative to experiments with purified enzymes. In this study, comparative and mutational studies on a murine DGAT2-type WS and a murine DGAT2 have been performed in order to identify determinants of substrate specificity within the murine WS. The obtained results are presented in the following section.

#### 5.3.1 MmAWAT2 and MmDGAT2 show distinct acyl chain incorporation patterns in WE biosynthesis

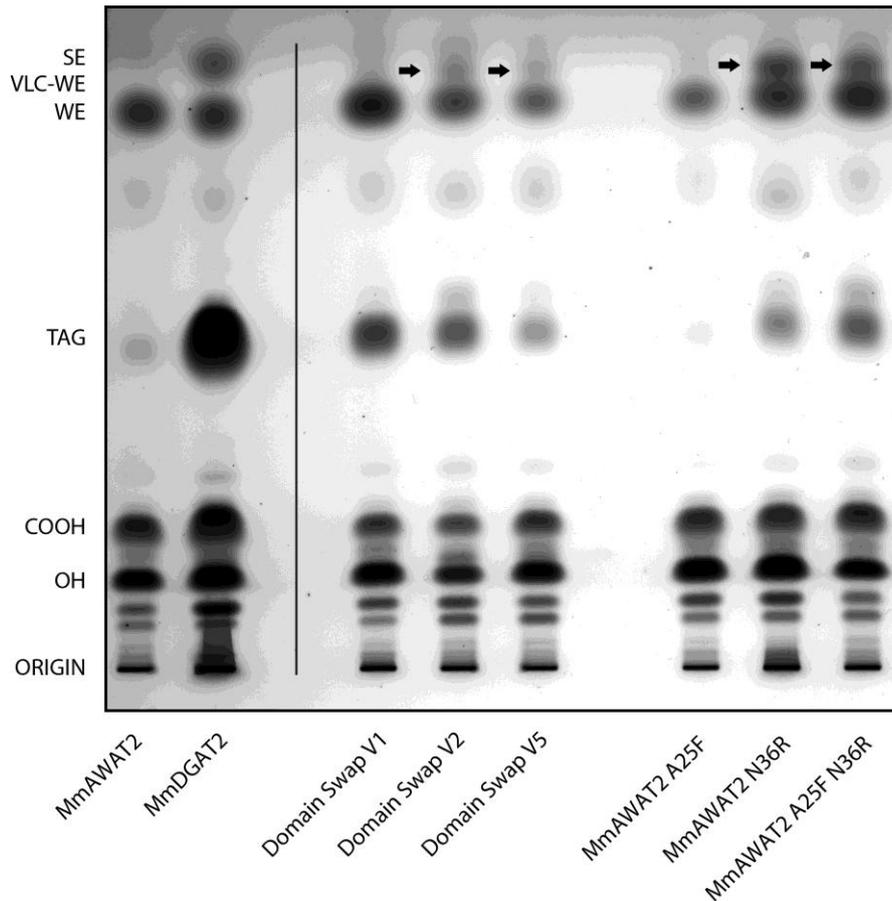
Murine acyl-CoA wax alcohol acyltransferase 2 (MmAWAT2) and murine DGAT2 (MmDGAT2) are both acyltransferases of the DGAT2-type. Despite of the fact that they share a similarity of 76 % on the sequence level (Appendix 19), the main products of the two enzymes are different. While MmAWAT2 predominantly produces WEs, the main products of MmDGAT2 are TAGs. However, both enzymes are not restricted to the synthesis of a single class of neutral lipids. It was shown before that MmAWAT2 is also capable of TAG formation, while MmDGAT2 can also catalyse the production of WEs (Cheng, 2004a). In order to further quantify the ratios of WEs and TAGs which are produced by the respective enzymes, both enzymes were expressed in the *S. cerevisiae* quadruple knockout strain H1246 (Figure 31, left lanes). The strain is devoid of all four genes responsible for neutral lipid synthesis (Sandager, 2001). As a consequence, all neutral lipids present in cultures expressing heterologous genes can directly be linked to the activity of respective gene products. Using the ImageJ software (Schneider et al., 2012), the ratio of WE/TAG production of MmAWAT2 was quantified densitometrically after TLC separation.

When feeding the cultures with the monounsaturated fatty alcohol 18:1-OH, MmAWAT2 produced approximately six times more WEs than TAGs (ratio of WE/TAG = 6.35:1 ( $\pm$  1.36)). In case of feeding the fully saturated fatty alcohol 16:0-OH, MmAWAT2 produced 10 times more WEs than TAGs (ratio of WE/TAG = 10.10:1 ( $\pm$  0.84)).

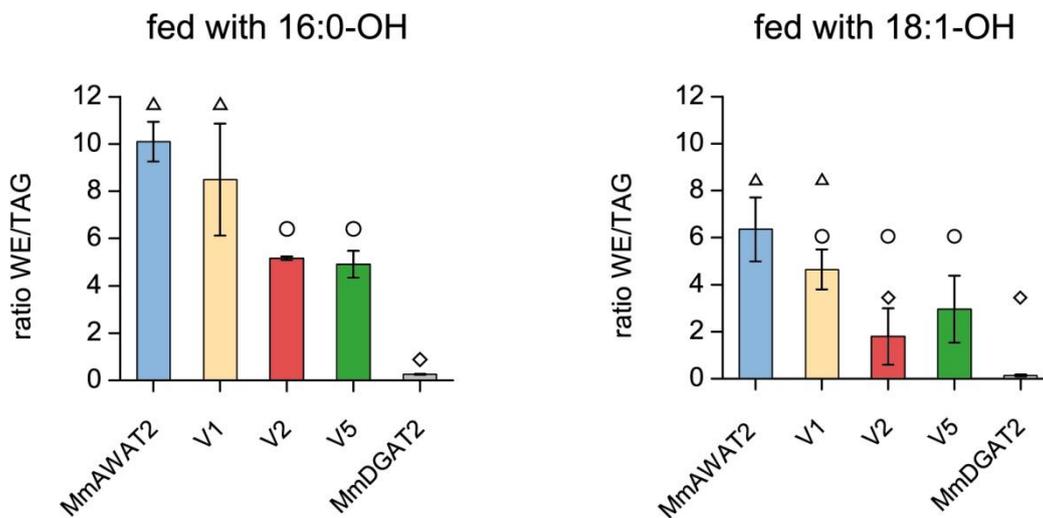
In contrast to that, MmDGAT2 produced approximately seven times more TAGs than WEs upon feeding 18:1-OH (ratio of WE/TAG = 0.14:1 ( $\pm$  0.04)), while feeding of 16:0-OH resulted in four times more TAGs than WEs (ratio of WE/TAG = 0.25:1 ( $\pm$  0.02)) (Figure 32). All values represent data derived from samples of at least three independent cultures.

In summary, MmAWAT2 prefers fatty alcohol as the acyl acceptor, whereas MmDGAT2 prefers DAG. Apart from that, feeding of 18:1-OH resulted in more TAG accumulation in comparison to feeding of 16:0-OH for both enzymes, MmAWAT2 and MmDGAT2 (Figure 32).

## RESULTS



**Figure 31: TLC analyses of WS and DGAT activity of MmAWAT2, MmDGAT2 and respective variants.** Separation of neutral lipids from yeast strains expressing wild type versions of MmAWAT2 and MmDGAT2 (left side), domain swap variants derived from those two enzymes and amino acid exchange variants of MmAWAT2 (right side). All constructs were expressed in *S. cerevisiae* H1246 and fed with 1 mM 18:1-OH. Very long chain (VLC) acyl chain containing WEs are indicated by arrows. The TLC is also representative for cultures fed with 16:0-OH in respect to the additional VLC WE spot. Data are representative for samples derived from at least three independent cultures for each construct. SE = sterol esters, VLC-WE = very long chain acyl chain containing wax esters, WE = wax esters, TAG = triacylglycerols, COOH = free fatty acids, OH = free fatty alcohols, ORIGIN = start of migration.



**Figure 32: Ratios of WE/TAG produced by MmAWAT2, domain swap variants V1, V2 and V5 as well as MmDGAT2.** Ratios were determined densitometrically after separation of respective lipid extracts via TLC. For quantification of the ratios, the ImageJ software was used with default settings (Schneider et al., 2012). Symbols indicate significant differences between the values at a level of  $p < 0.05$ , as derived from a Tukey's test. Data represent mean and standard deviation of samples derived from at least three independent cultures in every case.

## RESULTS

Since both enzymes are capable of WE synthesis, the detailed composition of the WE blends of both enzymes was analysed in order to elucidate potential differences in respect to specificities for certain acyl chains. *S. cerevisiae* H1246 does not possess an intrinsic FAR activity, so fatty alcohols for WE biosynthesis had to be fed to the cultures. Upon feeding a single alcohol species, all synthesised WEs carry the respective alcohol moiety and only differ in the acyl chain moiety. Therefore, the acyl chain incorporation patterns in respect to WE synthesis were analysed for both enzymes.

Upon expression of one of the enzymes in *S. cerevisiae* fed with either 16:0-OH or 18:1-OH and subsequent GC-MS analyses of lipid extracts, unique substrate specificities for WE synthesis were revealed. MmAWAT2 predominantly incorporates 16:1 acyl chains in WEs in quantities of approximately 55 %. Apart from that, 18:1 acyl chains account for approximately 25 %, while 16:0 and 18:0 acyl chains account for approximately 15 % and 5 % in the WE fraction of cultures expressing MmAWAT2, respectively (Figure 34).

In contrast to that, MmDGAT2 incorporates 16:0, 16:1 and 18:1 chains in quantities of 25-35 %, while 18:0 chains are incorporated to approximately 10-15 %. The described substrate preferences of both enzymes seem to be independent from the fed alcohol, since changes in acyl chain compositions of the produced WEs upon feeding different alcohols are only minor (Figure 34). Remarkably, none of the two enzymes synthesises WEs which carry acyl chains longer than 18 carbon atoms under the conditions tested here. The percentage of WEs carrying respective acyl chains is about 1 % and thus negligible (Figure 34). All values represent data derived from samples of at least three independent cultures.

### 5.3.2 Acyl-CoA pool composition in *S. cerevisiae* H1246 upon feeding of different alcohols

In order to evaluate which acyl-CoA substrates are potentially available in the yeast cell, the composition of the *S. cerevisiae* H1246 intrinsic acyl-CoA pool was further analysed (Figure 33). In order to enable a genuine impression of the substrate composition which is offered by the cell, it was chosen to investigate the acyl-CoA pool of *S. cerevisiae* H1246 cultures expressing a pYES2/NT empty vector control, fed with either 16:0-OH or 18:1-OH. Under this condition, no heterologous enzyme is expressed which might alter the pool's composition. Hence, these analyses represent the basic, unaltered acyl-CoA pool composition provided by the cell.

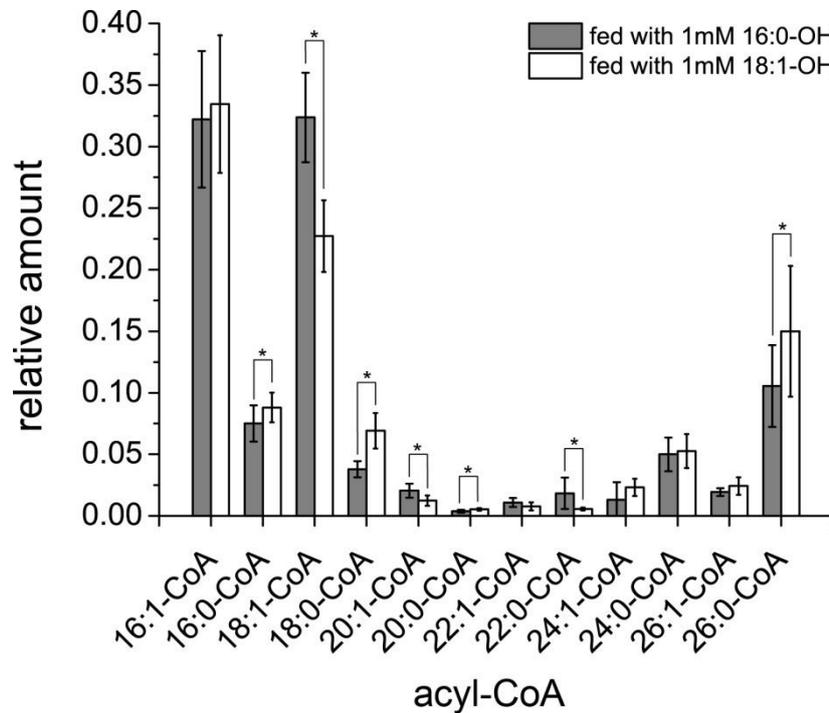
Upon feeding 1 mM 16:0-OH, the most abundant acyl-CoA species were 16:1-CoA and 18:1-CoA, which both accounted for 32 % of all measured acyl-CoA molecules ( $\pm 6$  and  $\pm 4$  %, respectively). The second major species was 26:0-CoA ( $11 \pm 3$  %), followed by 16:0-CoA ( $8 \pm 1$  %). Furthermore, 18:0-CoA, 20:0-CoA, 20:1-CoA, 22:0-CoA, 22:1-CoA, 24:0-CoA, 24:1-CoA and 26:1-CoA were identified. All of these species accounted for 5 % (24:0-CoA) or less.

In case of feeding 18:1-OH to respective cultures, the main acyl-CoA species was 16:1-CoA ( $33 \pm 6$  %), followed by 18:1-CoA ( $23 \pm 3$  %). The third most abundant species was 26:0-CoA again ( $15 \pm 5$  %). It was followed by 16:0-CoA ( $9 \pm 1$  %). The amount of 18:0-CoA accounted for  $7 \pm 1$  % and was approximately twice as high as in cultures fed with 16:1-OH. The species 20:0-CoA, 20:1-CoA, 22:0-CoA, 22:1-CoA, 24:0-CoA, 24:1-CoA and 26:1-CoA were identified in similar amounts to cultures fed with 16:0-OH. All values represent data derived from samples of more than three independent cultures.

In summary, feeding of 16:0-OH resulted in accumulation of similar amounts of 16:1-CoA and 18:1-CoA, whereas feeding of 18:1-OH resulted in a lower abundance of 18:1-CoA than 16:1-CoA. In

## RESULTS

all cultures, 26:0-CoA was the third most abundant species. Acyl-CoAs with fully saturated and monounsaturated acyl chains ranging from chain length of 16-26 carbon atoms were identified in all cultures (Figure 33, Figure 37 C).



**Figure 33: Relative acyl-CoA pool composition in *S. cerevisiae* H1246 expressing pYES2/NT upon feeding either 1 mM 16:0-OH or 1 mM 18:1-OH, respectively.** Asterisks indicate significant differences with a levels of  $p < 0.05$  as derived from a student's t-test. Data show the mean and standard deviation of samples derived from more than three independent cultures.

### 5.3.3 MmAWAT2 and MmDGAT2 show distinct specificities in respect to acyl chain incorporation into WEs

A comparison of the acyl chain distribution in the acyl-CoA pool and in WEs produced by either MmAWAT2 or MmDGAT2 illustrates certain overlaps, but also reveals some differences.

For instance, MmAWAT2 incorporates 16:1 acyl chains in approximately half of all WE molecules, regardless of the fed alcohol. In contrast, the acyl-CoA pool consists of only about one third of 16:1-acyl chains. This matches the amount which is incorporated in WEs by MmDGAT2 (Figure 34).

MmDGAT2 incorporates 16:0-CoA in approximately 25 % of all WE molecules, while the acyl-CoA pool consists of only around 10 %. MmAWAT2, on the other hand, incorporates an intermediate amount of 16:0 chains of approximately 15 % and significantly differs from both, MmDGAT2 and acyl-CoA pool (Figure 34).

The incorporation of 18:1 acyl chains in WEs by MmAWAT2 and MmDGAT2 on the one side and the providence of this acyl chain by the acyl-CoA pool on the other side do not show as drastic differences. Accordingly, the levels of 18:1 chains in WEs from cultures expressing either MmAWAT2 or MmDGAT2 do not significantly differ from each other upon feeding 16:0-OH. However, MmAWAT2 incorporates significantly less 18:1 acyl chains into WEs than are present in the acyl-CoA pool upon feeding 16:0-OH. Upon feeding 18:1-OH, MmDGAT2 incorporates significantly more 18:1

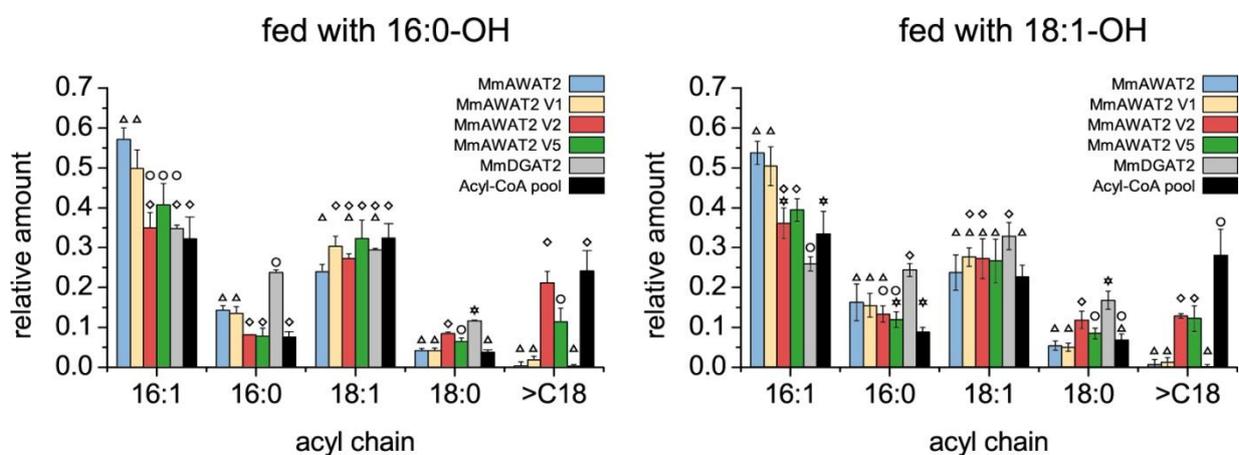
## RESULTS

acyl chains than are present in the CoA-pool, whereas MmAWAT2 and the acyl-CoA pool do not significantly differ from each other (Figure 34).

In case of 18:0 acyl chains, the amount incorporated by MmAWAT2 matches the prevalence of the acyl-CoA pool, while MmDGAT2 incorporates approximately twice as much 18:0. These patterns are independent of the fed alcohol.

Both enzymes only synthesise negligible amounts of WEs carrying acyl chains longer than 18 carbon atom in lengths, although respective species account for around 25 % of the acyl chains in the acyl-CoA pool (Figure 34). However, only weak signals of WEs containing 26:0 acyl chains were found in some of the MmAWAT2 and MmDGAT2 samples fed with either 16:0-OH or 18:1-OH, hardly making up 1 % of all WEs upon GC-FID measurements (Figure 34, Figure 37 A).

In summary, MmAWAT2 and MmDGAT2 show discrete acyl chain incorporation patterns into WEs. These patterns differ from each other, but also from the acyl-CoA pool composition.



**Figure 34: Acyl chain composition of the intrinsic yeast acyl-CoA pool and of WEs synthesised by MmAWAT2, MmDGAT2 and corresponding domain swap variants of both enzymes.** Bars represent the relative amounts of each acyl chain in the WE fraction of *S. cerevisiae* H1246 cultures expressing the respective enzyme. Moreover, the relative amounts of each acyl chain in the yeast acyl-CoA pool is shown (grey bars). Values for > C18 include acyl chains from 20 to 26 carbon atoms. Symbols indicate significant differences between the values at a level of  $p < 0.05$ , as derived from a Tukey's test. Data show the mean and standard deviation of at least three samples derived from independent expression cultures for each construct.

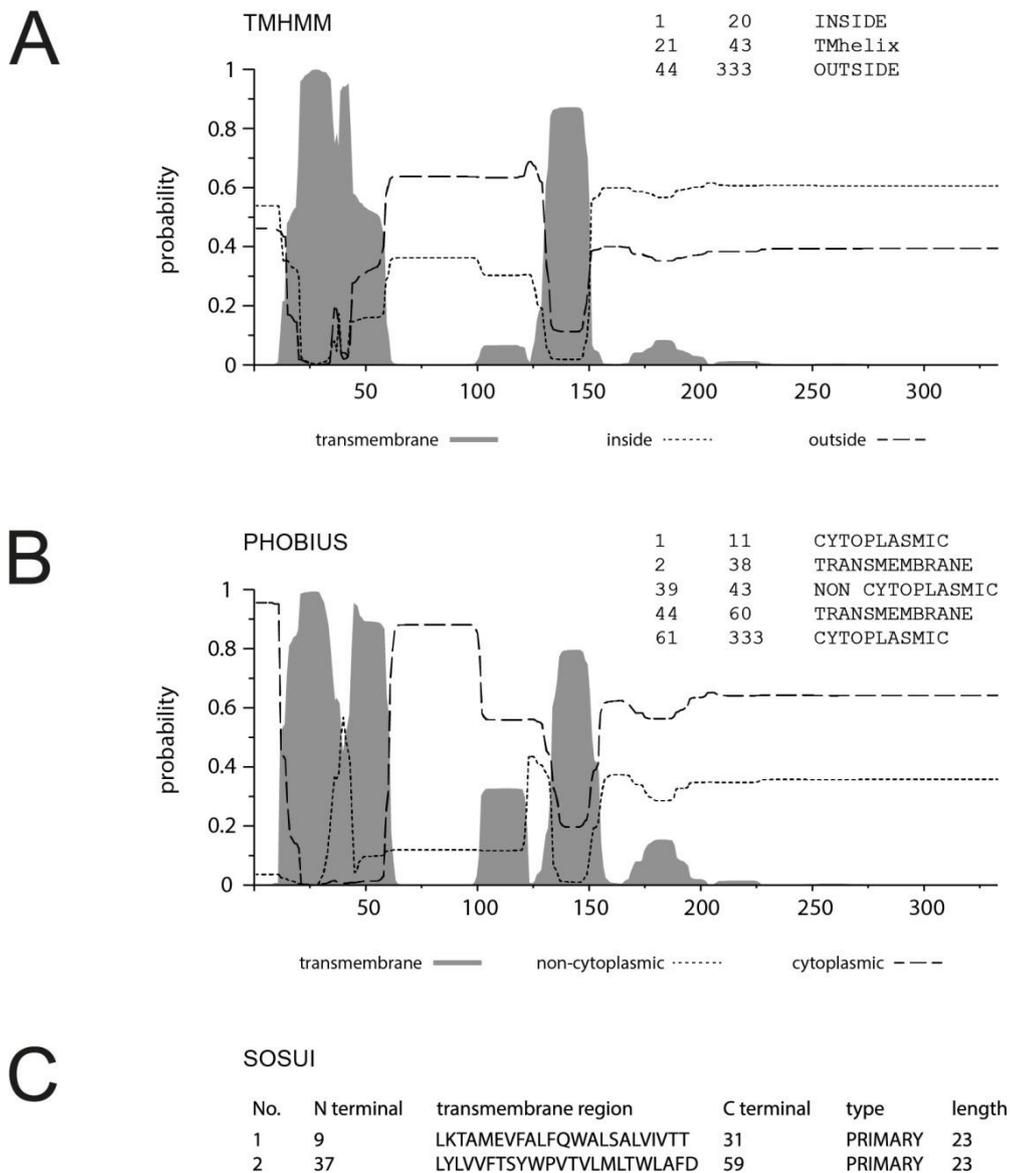
### 5.3.4 Construction of domain swap variants on the basis of MmAWAT2 and MmDGAT2

Although MmAWAT2 and MmDGAT2 share a high similarity on the sequence level, they differ in their substrate specificity regarding acyl acceptors as well as in acyl chain incorporation patterns into WEs (Figure 32, Figure 34). In order to elucidate structures that potentially determine this substrate specificity, domain swapping between both enzymes was carried out. Therefore, the sequences of MmAWAT2 and MmDGAT2 were divided into domains according to their predicted TM topology. In case of MmDGAT2, the respective TM topology was determined before (Stone et al., 2006) and verified recently (McFie et al., 2014). It consists of a short N-terminal cytosolic stretch, which is followed by two TM domains. The TM domains act as a membrane anchor for the enzyme in the ER. The first TM domain within the MmDGAT2 sequence harbours the consensus sequence FLXLXXX, which is thought to be involved in neutral lipid binding (Alam et al., 2006; Au-Young and Fielding, 1992). The part C-terminal to the two TM domains is reported to harbour the active site of the

## RESULTS

enzyme and to reside on the cytosolic side of the ER membrane as well. Consequently, the MmDGAT2 sequence was divided in three parts, consisting of the N-terminus, the two TM domains and the C-terminus (Figure 36 A).

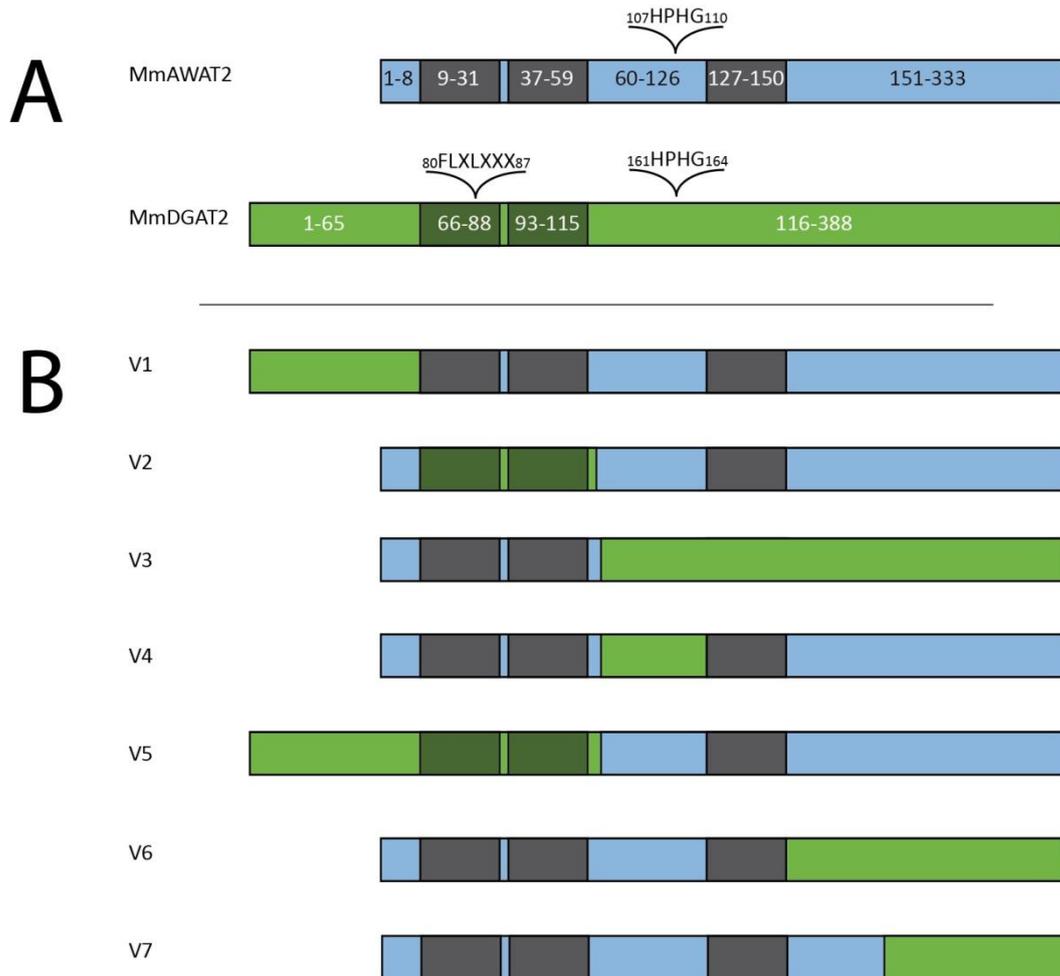
In order to elucidate a possible TM topology and thus identify domains in MmAWAT2, its sequence was submitted to the three web based TM domain prediction servers SOSUI (Hirokawa et al., 1998), TMHMM (Sonnhammer et al., 1998) and Phobius (Käll et al., 2007) (Figure 35). TMHMM predicted the presence of a single N-terminal TM domain (Figure 35 A), whereas Phobius and SOSUI predicted the presence of two N-terminal domains, which are connected by a short stretch of only a few amino acids (Figure 35 B&C). In addition, TMHMM and Phobius determined the presence of a highly hydrophobic patch in the middle of the enzyme. However, this patch was not predicted to form a TM domain (Figure 35 A&B).



**Figure 35: Prediction of transmembrane domains in MmAWAT2. A)** Prediction according to data generated with TMHMM (Sonnhammer et al., 1998), **B)** Phobius (Käll et al., 2007) and **C)** SOSUI (Hirokawa et al., 1998).

## RESULTS

The predicted TM topology of MmAWAT2 is thus similar to MmDGAT2 (illustrated in Figure 40). The MmAWAT2 sequence was divided in five segments, with both predicted TM domains grouped in one domain (Figure 36 A). Seven domain swap variants were constructed on the basis of MmAWAT2. Accordingly the N-terminus of the enzyme (V1) as well as the putative TM domains (V2) and both parts together (V5) were exchanged. Additionally, variants were constructed in which the complete C-terminus (V3), the part between the TM domains and the hydrophobic patch in the middle of the enzyme (V4), as well as the sequence C-terminal from this patch (V6 and V7) were exchanged (Figure 36 B). All variants were cloned in the pYES2/NT yeast expression vector.



**Figure 36: Domain swap variants constructed from MmAWAT2 and MmDGAT2. A)** Arrangement of the putative domain structures of MmAWAT2 and MmDGAT2. The conserved active site motif HPHG and the putative neutral lipid binding site of MmDGAT2 with the consensus sequence FLXLXXX are indicated. **B)** MmAWAT2 and MmDGAT2 were used to construct seven domain swap variants (V1-V7). In each variant, a specific part of MmAWAT2 was exchanged for the respective part of MmDGAT2.

### 5.3.5 Substrate specificities of domain swap variants

Of the seven generated domain swap variants, V1, V2 and V5 showed substantial WS activity upon expression in *S. cerevisiae* H1246 (Figure 31). In contrast, expression of V3, V4, V6 and V7 did not result in WE or TAG accumulation (data not shown). All active variants had a complete, non-modified C-terminus in common. In contrast, all inactive variants carried a C-terminus, which was at least partly exchanged for respective parts from MmDGAT2 (Figure 36).

Upon TLC analyses of lipid extracts from yeast cultures expressing either V2 or V5, additional signals were observed on respective TLC plates. These signals were absent in samples derived from V1, MmAWAT2 or MmDGAT2 expressing cultures. The additional spots migrated above the spots of normal WEs (Figure 31, indicated by black arrows). Moreover, the spots occurred regardless of the fed alcohol (data not shown). To elucidate the nature of the additional signals, respective samples were analysed by GC-MS. In case of samples derived from cultures expressing either V2 or V5 and fed with 16:0-OH, four additional peaks were detected in the GC-MS chromatograms in comparison to samples derived from culture expressing MmAWAT2. In contrast, cultures expressing either V2 or V5 and fed with 18:1-OH showed only a single additional peak (representatively shown in Figure 37 A&B).

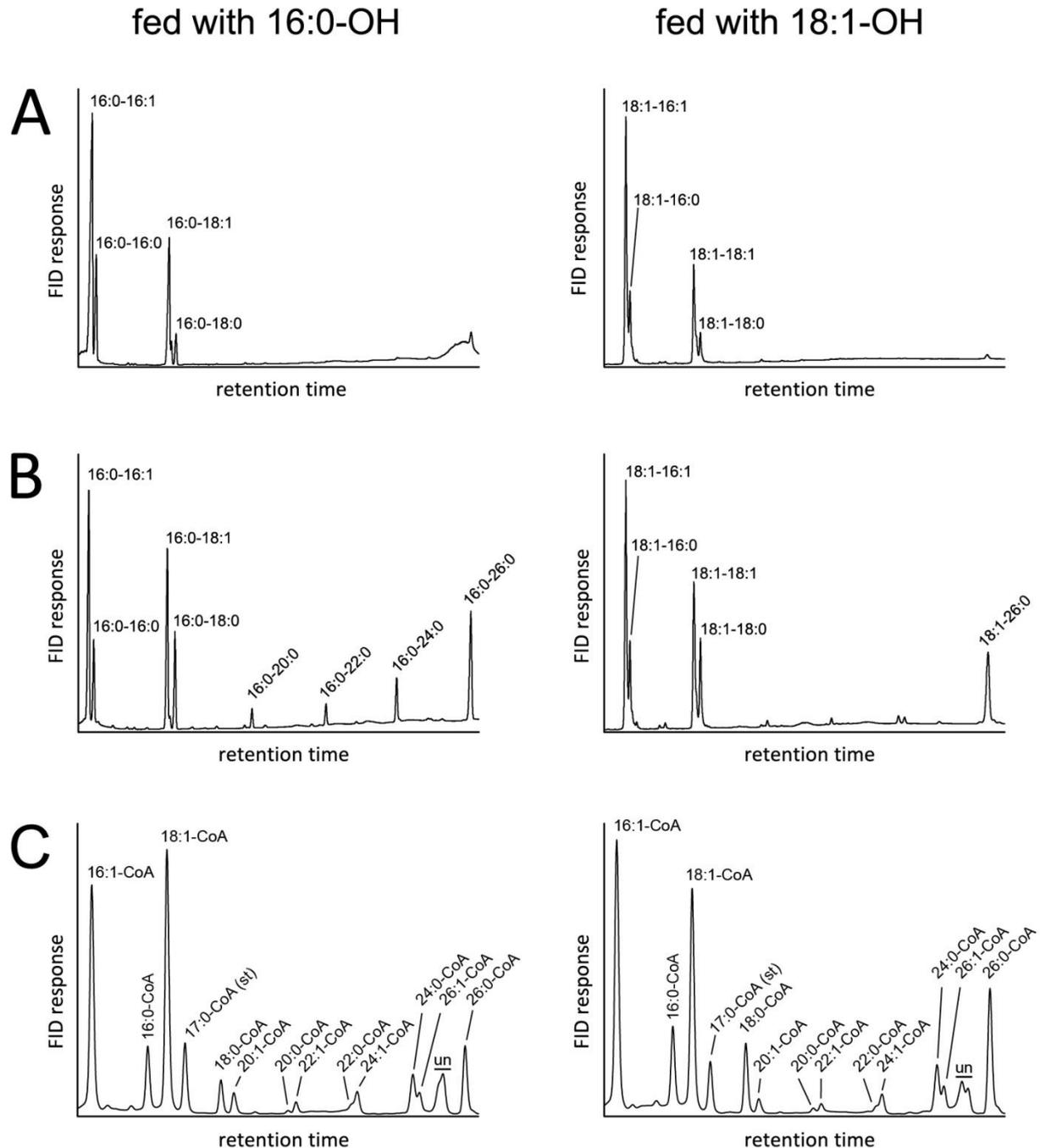
The mass spectra of the four additional peaks from samples of 16:0-OH-fed cultures showed highly similar overall fragmentation patterns (Figure 38 B-E). The mass of the most abundant fragment in each spectrum sequentially differed in a multitude of 28 mass units, which may correspond to a C<sub>2</sub>H<sub>4</sub>-fragment (Figure 38 B-E). Moreover, a second prominent fragment in the higher mass range of each spectrum also sequentially differed in a multitude of 28 mass units among the different spectra. The spectra of the first two additional peaks (peak B and C in Figure 38) match the published fragmentation patterns of the very long chain (VLC) WEs hexadecanoyl icosanoate (16:0 20:0) and hexadecanoyl docosanoate (16:0 22:0) (Urbanova et al., 2012), respectively. According to the published data, the most abundant fragment within these spectra represents the acid chain fragment of the corresponding WE, whereas the second prominent fragment in the higher mass range represents the molecular ion of the corresponding WE. In the spectra of the other two additional peaks (peak D and E in Figure 38), the masses of the respective fragments perfectly fit to the corresponding masses of hexadecanoyl-tetracosanoate (16:0 24:0) and hexadecanoyl-hexacosanoate (16:0 26:0). Thus, the additional peaks in samples derived from cultures expressing either V2 or V5 and fed with 16:0-OH most likely represent WEs carrying saturated VLC acyl chains from 20 to 26 carbon atoms.

Samples of cultures expressing either V2 or V5 and fed with 18:1-OH contained only a single additional peak, which was assigned to 18:1-26:0 (Figure 37 B, mass spectrum not shown). As already mentioned above, neither samples from cultures expressing MmAWAT2 nor from cultures expressing MmDGAT2 contained VLC WEs in higher amounts (Figure 37 A).

The relative contribution of VLC WEs to the total WE fraction was quantified via GC-FID. V2 expressing cultures were determined to synthesise  $21 \pm 3$  % VLC WEs upon feeding of 16:0-OH, while feeding of 18:1-OH resulted in production of  $13 \pm 1$  % VLC WEs. In case of V5, the values were determined to be  $13 \pm 3$  % and  $12 \pm 3$  %, respectively. Samples derived from V1 expressing cultures consistently contained only small amounts of VLC WEs around 1 %, which did not significantly differ from MmAWAT2 or MmDGAT2 (Figure 34). All values represent data derived from at least three samples of independent expression cultures.

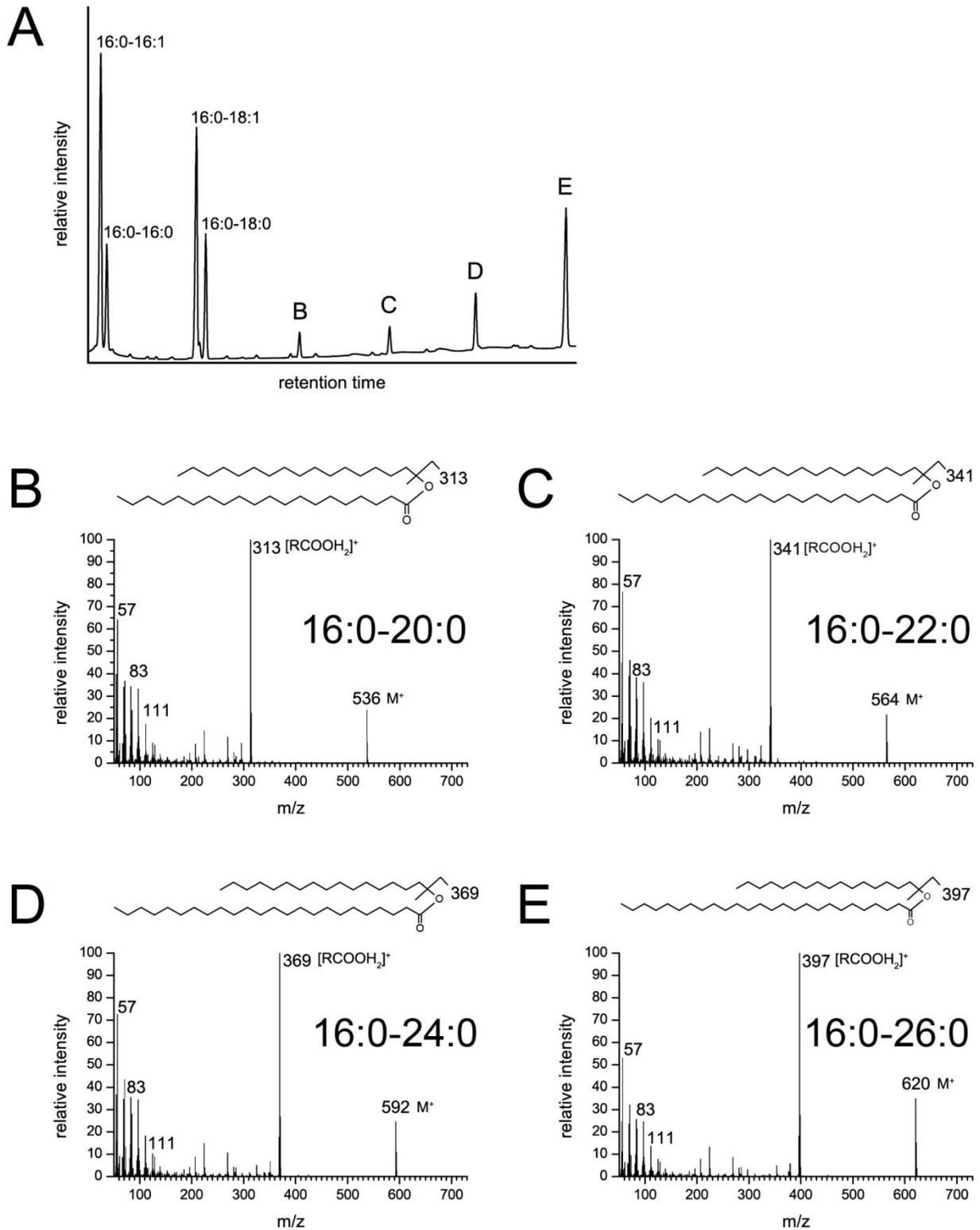
## RESULTS

Interestingly, no considerable amounts of WEs carrying unsaturated VLC acyl chains were identified in any of the samples, although these chains seem to be present within the acyl-CoA pool (compare Figure 37 B&C).



**Figure 37: Representative GC chromatograms of WE and acyl-CoA analyses. A)** GC-FID analyses of WEs produced by MmAWAT2 upon expression in *S. cerevisiae* H1246 either fed with 1 mM 16:0-OH or 1 mM 18:1-OH. **B)** GC-FID analyses of WEs produced by the MmAWAT2 A25F N36R variant upon expression in *S. cerevisiae* H1246 either fed with 1 mM 16:0-OH or 1 mM 18:1-OH. Data are also representative for GC-FID analyses of samples derived from cultures expressing either V2, V5 or the MmAWAT2 N36R variant. **C)** HPLC analyses of acyl-CoA extracts from *S. cerevisiae* H1246 expressing the pYES2/NT empty vector control and fed with either 1 mM 16:0-OH or 1 mM 18:1-OH. un = unknown. Data are representative for at least three samples derived from independent expression cultures.

RESULTS



**Figure 38: Mass spectra of VLC WEs produced by V2, V5, MmAWAT2 N36R and MmAWAT2 A25F N36R upon feeding of 16:0-OH.** **A)** Representative GC-FID chromatogram of a sample derived from a culture expressing MmAWAT2 A25F N36R fed with 16:0-OH. The chromatogram is also representative for samples from cultures expressing either V2, V5 or MmAWAT2 N36R fed with 16:0-OH. Based on the respective Mass spectra, the peaks were assigned to: **B)** hexadecanoyl icosanoate (16:0-20:0), **C)** hexadecanoyl docosanoate (16:0-22:0), **D)** hexadecanoyl-tetracosanoate (16:0-24:0) and **E)** hexadecanoyl-hexacosanoate (16:0-26:0). In cultures expressing the constructs mentioned above fed with 18:1-OH, only 18:1-26:0 was found as a VLC WE (**Figure 37 B**, Mass spectrum not shown).

## RESULTS

Subsequent to the identification of the newly formed VLC WEs it was also addressed, whether the domain swap variants V1, V2 and V5 showed any differences in comparison to MmAWAT2 concerning the incorporation of C16 and C18 acyl chains into WEs.

In comparison to the WE load of cultures expressing MmAWAT2, WEs of cultures expressing either V2 or V5 contained significantly reduced levels of 16:1 acyl chains, regardless of the fed alcohol (Figure 34). Furthermore, WEs of V2 or V5 expressing cultures contained decreased amounts of 16:0 chains in case of feeding 16:0-OH, while upon feeding 18:1-OH, only WEs from cultures expressing V5 contained significantly decreased amounts of 16:0 acyl chains. WEs of cultures expressing V5 contained significantly higher levels of 18:1 chains upon feeding 16:0-OH in comparison to WEs from cultures expressing MmAWAT2, whereas 18:1 chains of cultures expressing V2 and fed with 16:0-OH or cultures expressing either V2 or V5 and fed with 18:1-OH did not significantly differ from the respective MmAWAT2 values. The levels of 18:0 acyl chains were significantly increased in WEs of cultures expressing either V2 or V5 in comparison to WEs from cultures expressing MmAWAT2, regardless of the fed alcohol (Figure 34).

In comparison to WEs from cultures expressing MmDGAT2, the levels of 16:1 chains in WEs of cultures expressing either V2 or V5 were unaltered upon feeding 16:0-OH, while they were significantly increased upon feeding of 18:1-OH. The levels of 16:0 chains were significantly reduced in WEs of cultures expressing either V2 or V5, regardless of the fed alcohol. In case of 18:1 chains, only WEs from cultures expressing V5 and fed with 18:1-OH showed a significant decrease in comparison to WEs from cultures expressing MmDGAT2. In contrast, the levels of 18:1 chains in WEs from cultures expressing V2 and fed with 18:1-OH did not significantly differ from WEs of cultures expressing MmDGAT2. Likewise, also WEs of cultures expressing either V2 or V5 and fed with 16:0-OH did not contain significantly different levels of 18:1 chains in comparison to WEs from cultures expressing MmDGAT2. The levels of 18:0 chains were significantly reduced in the WEs from cultures expressing either V2 or V5 in comparison to WEs from cultures expressing MmDGAT2, regardless of the fed alcohol (Figure 34).

The acyl chain distribution within the acyl-CoA pool was rather similar to WEs from cultures expressing either V2 or V5 at the first sight, but also showed significant differences. In comparison to the acyl-CoA pool, 16:1 chains were significantly increased in WEs from cultures expressing V5, regardless of the fed alcohol. In contrast, 16:1 chains were on acyl-CoA pool levels in case of WEs from cultures expressing V2, also regardless of the fed alcohol. In case of 16:0 acyl chains, WEs from cultures expressing either V2 or V5 were on acyl-CoA pool level in case of feeding 16:0-OH, while feeding of 18:1-OH resulted in significantly increased values for 16:0 chains in WEs from cultures expressing V2 only. No significant differences compared to the acyl-CoA pool composition were measured for 18:1 chains in WEs from cultures expressing either V2 or V5, regardless of the fed alcohol. However, 18:0-chains in WEs from cultures expressing either V2 or V5 and fed with 16:0-OH were significantly increased compared to the acyl-CoA pool. The same was true for WEs from cultures expressing V2 fed with 18:1-OH (Figure 34).

Notably, the differences in the acyl chain compositions of WEs from cultures expressing either V2 or V5 compared to the acyl-CoA pool were rather subtle. In contrast, the respective differences in comparison to MmAWAT2 and MmDGAT2 were more pronounced (Figure 34).

In summary, WEs from cultures expressing either V2 or V5 showed significantly different acyl chain incorporation patterns in comparison to WEs from both, cultures expressing either MmAWAT2 or MmDGAT2. Besides that, these patterns were also significantly different from the acyl chain composition of the acyl-CoA pool.

## RESULTS

In contrast, WEs from cultures expressing V1 were similar to WEs from cultures expressing MmAWAT2 and only differed significantly from the corresponding values in respect to incorporation of slightly more 18:1 chains upon feeding 16:0-OH (Figure 34).

Apart from the differences in WE compositions, samples of cultures expressing either V2 or V5 also showed stronger TAG signals in comparison to MmAWAT2, as observed during TLC analyses (Figure 31). Accordingly, the WE/TAG ratios of samples from cultures expressing either V2 or V5 were significantly reduced by approximately 50 % in comparison to the values of MmAWAT2 (Figure 32). In contrast, the WE/TAG ratios of samples from cultures expressing V1 were not significantly different from those of MmAWAT2 (Figure 32). Interestingly, the WE/TAG ratios were approximately twice as high for all constructs upon feeding of 16:0-OH in comparison to feeding of 18:1-OH.

### 5.3.6 The predicted N-terminal transmembrane domain in MmAWAT2 is highly conserved among vertebrates

Since the predicted TM domain of MmAWAT2 was identified to have a strong influence on the enzyme's substrate specificity, a sequence analysis of the respective parts of AWAT2 and DGAT2 enzymes was carried out. Therefore, the UniProt database (The UniProt Consortium, 2014) was searched for respective vertebrate-derived sequences, which were subsequently aligned using the Clustal Omega service (Sievers et al., 2011). In order to predict the TM topology of the corresponding enzymes, the sequences were moreover submitted to the SOSUI online service (Hirokawa et al., 1998). Interestingly, a general domain architecture consisting of two TM domains is predicted to be strongly conserved in both vertebrate enzyme classes, AWAT2 and DGAT2 (Figure 39). According to the predictions, these two TM domains are connected via a short stretch of three to five amino acids in almost all checked sequences. In case of MmDGAT2, this architecture was already experimentally verified (Stone et al., 2006).

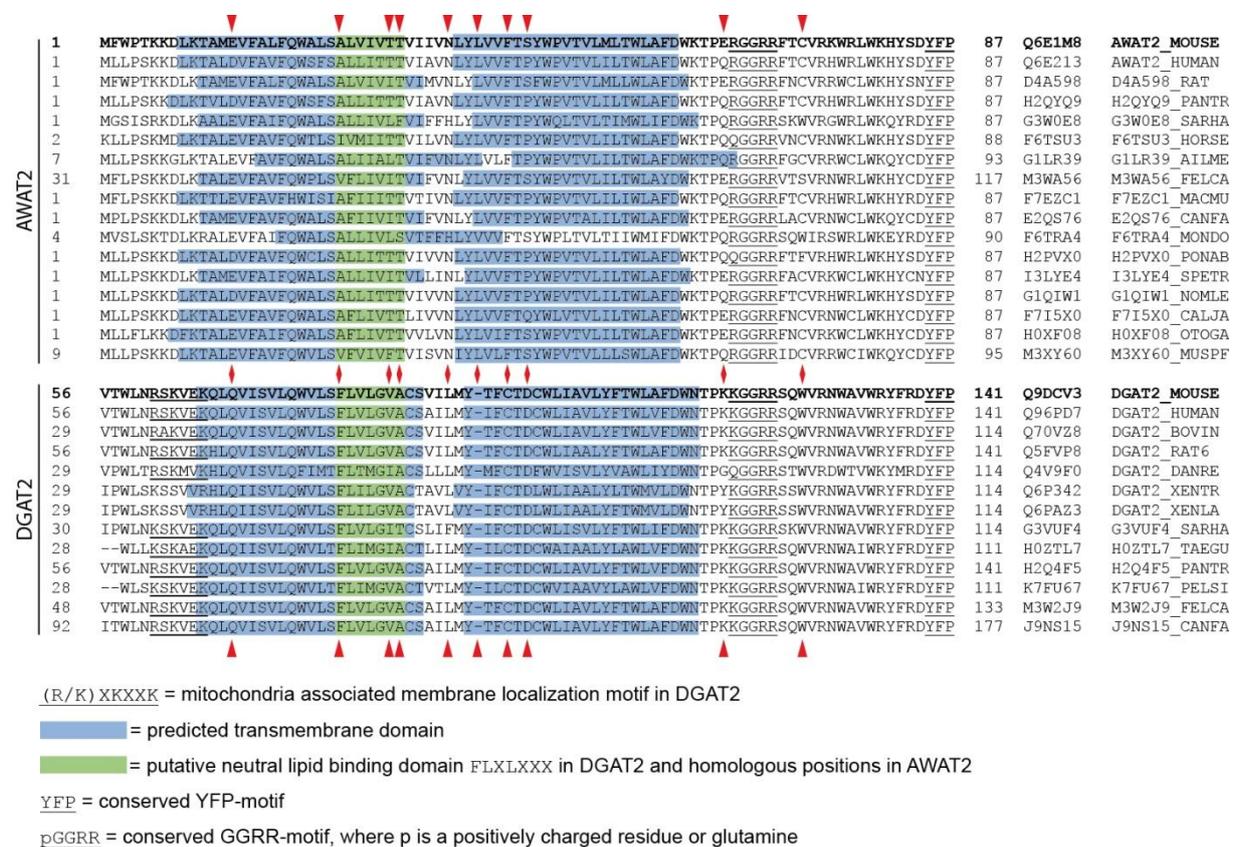
In DGAT2 enzymes, the first TM domain harbours a strongly conserved putative neutral lipid binding site (NLBS) with the consensus sequence FLXLXXX (Alam et al., 2006; Au-Young and Fielding, 1992) (Figure 39). However, in AWAT2 sequences, this motif is not conserved. In case of MmAWAT2, it corresponds to "ALVIVTT" and thus misses the conserved phenylalanine as well as the second conserved leucine. Moreover, the two last positions of the motif are occupied by apolar residues in DGAT2 sequences, whereas in AWAT2 enzymes they are mostly occupied by polar residues (Figure 39). A conserved sequence motif in both, vertebrate AWAT2 and DGAT2 sequences is a pGGRR motif. Here, p stands for a positively charged residue or glutamine. It is located just downstream of the second predicted TM domain in AWAT2 and DGAT2 sequences and represents a cluster of positive charge. To date, it has not been elucidated whether the motif plays a role in the enzymes activity. Another conserved motif in both AWAT2 and DGAT2 sequences is a YFP motif. It is located shortly after the pGGRR motif (Figure 39) and has been shown to cause an almost complete loss of activity upon mutation of all three residues to alanine in a respective enzyme variant from *S. cerevisiae* (Liu et al., 2011). Directly N-terminal to the TM domain of MmDGAT2, a RXKXXK was found to be responsible for mediating localisation to MAM (Stone et al., 2009). This motif is not present in AWAT2. Instead, a relatively conserved KKDLK motif can be found at the corresponding part of AWAT2 sequences (Figure 39). However, it is not known whether this part of AWAT2 sequences is involved into MAM localisation, or if AWAT2 enzymes are capable to localise to MAM at all.

## RESULTS

### 5.3.7 The substrate specificity in MmAWAT2 is influenced by different residues in the predicted transmembrane region

In order to identify residues which are responsible for the observed changes in the domain swap variants V2 and V5, a set of MmAWAT2 single amino acid exchange variants was constructed based on conserved differences within the predicted TM domains of AWAT2 and DGAT2 enzymes (Figure 39, Figure 40).

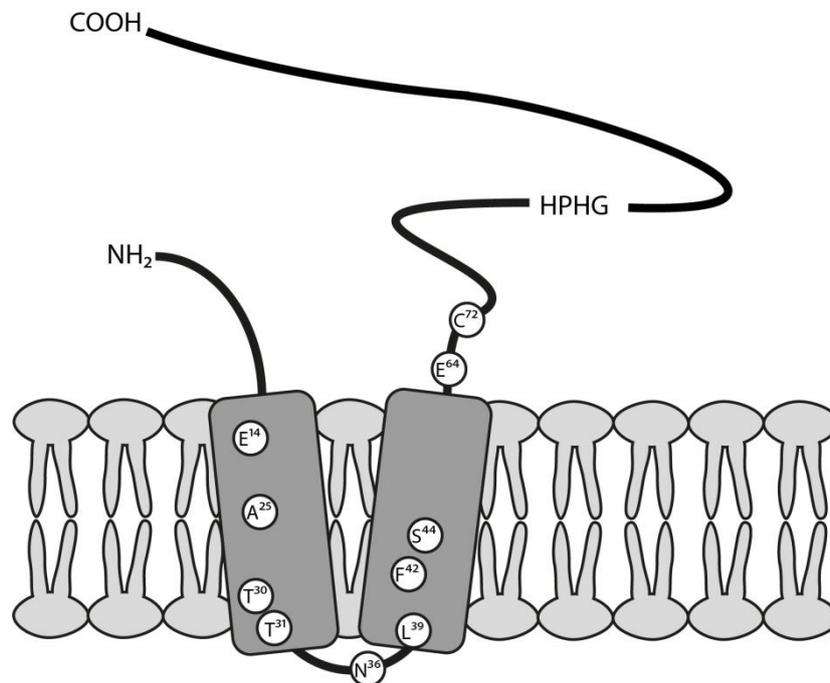
In case of the MmAWAT2 variants L39-, F42C, S44D E64K and C72W, no changes in WE compositions of 18:1-OH fed yeast cultures expressing the respective constructs were observed in comparison to the wild type enzyme (data not shown). Minor changes in acyl chain incorporation specificities were detected for the MmAWAT2 variants E14Q and S44D. Both variants caused significantly decreased levels of 16:1 acyl chains within the WE fraction of respective cultures, as concluded from a student's t-test ( $p < 0.05$ ). The levels of the other acyl chains stayed unaltered in comparison to MmAWAT2 (Appendix 20). However, both constructs were only tested upon feeding 18:1-OH.



**Figure 39: Alignment of vertebrate-derived AWAT2 and DGAT2 enzymes.** G1LR39 = *Ailuropoda melanoleuca* AWAT2, Q70VZ8 = *Bos taurus* DGAT2, F7I5X0 = *Callithrix jacchus* AWAT2, J9NS15 = *Canis familiaris* DGAT2, Q4V9F0 = *Danio rerio* DGAT2, F6TSU3 = *Equus caballus* AWAT2, M3WA56 = *Felis catus* AWAT2, M3W2J9 = *Felis catus* DGAT2, Q6E213 = *Homo sapiens* AWAT2, Q96PD7 = *Homo sapiens* DGAT2, F7E2C1 = *Macaca mulatta* AWAT2, F6TRA4 = *Monodelphis domestica* AWAT2, Q6E1M8 = *Mus musculus* AWAT2, Q9DCV3 = *Mus musculus* DGAT2, M3XY60 = *Mustela putorius furo* AWAT2, G1QIW1 = *Nomascus leucogenes* AWAT2, H0XF08 = *Otolemur garnettii* AWAT2, H2Q4F5 = *Pan troglodytes* DGAT2, H2QYQ9 = *Pan troglodytes* AWAT2, K7FU67 = *Pelodiscus sinensis* DGAT2, H2PX0 = *Pongo abelii* AWAT2, D4A598 = *Rattus norvegicus* AWAT2, Q5FVP8 = *Rattus norvegicus* DGAT2, G3VUF4 = *Sarcophilus harrisii* DGAT2, G3W0E8 = *Sarcophilus harrisii* AWAT2, I3LYE4 = *Spermophilus tridecemlineatus* AWAT2, H0ZTL7 = *Taeniopygia guttata* DGAT2, Q6PAZ3 = *Xenopus laevis* DGAT2, Q6P342 = *Xenopus tropicalis* DGAT2. Alignment of the sequences was done using Clustal Omega (The UniProt Consortium, 2014), while the TM predictions were done with the SOSUI online tool (Hirokawa et al., 1998). Predicted TM domains are

## RESULTS

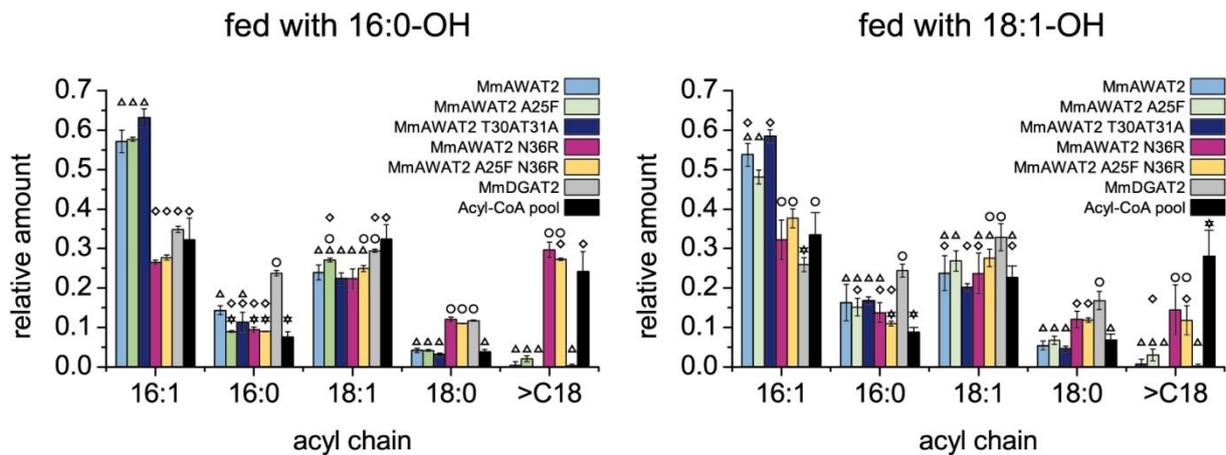
highlighted in blue, whereas parts of the sequences which are homologues to the putative NLBS “FLXLXXX” in DGAT2 are highlighted in green. The conserved motifs pGGRR, YFP as well as (R/K)XKXXX (DGAT2 only) are underlined. The sequences of MmDGAT2 and MmAWAT2, which were used in this study, are printed in bold. Positions within the sequence of MmAWAT2 which were mutated in the course of this study are indicated by red triangles. In case of MmDGAT2, the indicated transmembrane structure resembles the actual topology determined by Stone et al. (2006). Besides an abbreviation for each enzyme, also the UniProt ID for the respective protein is listed.



**Figure 40: Schematic model of the predicted topology of MmAWAT2.** The predicted N-terminal transmembrane domain is indicated as grey boxes. Residues which were mutated during this study are indicated in circles. According to the prediction, both N-terminus and C-terminus as well as the catalytic HPHG motif are facing the cytosolic side of the membrane.

As a part of the putative NLBS, phenylalanine at position 80 in MmDGAT2 was already shown to cause an almost complete loss of enzymatic activity in case of replacement with alanine (Stone et al., 2006). Interestingly, the respective position in MmAWAT2 is occupied by an alanine residue. Hence, the corresponding MmAWAT2 A25F variant was constructed in order to elucidate the effects of partially mimicking the DGAT2 NLBS in MmAWAT2. Upon TLC analyses of respective lipid extracts, an additional band above the normal WE band was observed in some of the samples. However, this band was only faint and far less pronounced than in case of V2 and V5. To further elucidate the presence and quantify the amount of VLC WEs in A25F-derived lipid extracts, respective samples were further analysed via GC-MS. The variant was capable of synthesising small amounts of VLC WEs upon feeding 16:0-OH or 18:1-OH, which constituted to the WE fraction to about 2 % in case of feeding 16:0-OH and approximately 3 % in case of feeding 18:1-OH. However, the values were found to not be significantly higher than those measured for MmAWAT2 and MmDGAT2 in case of conducting a Tukey’s test and a significance level of  $p < 0.05$  (Figure 41). In contrast, a student’s t-test indicated the difference between VLC WEs synthesised by the A25F variant and MmAWAT2 to be significant ( $p < 0.05$ ). Apart from that, significantly decreased levels of 16:0 acyl chains were measured in WEs of cultures expressing MmAWAT2 A25F upon feeding of 16:0-OH. Interestingly, no TAG signals were detected via TLC analyses in respective samples. In contrast, the respective values upon feeding of 18:1-OH were on wild type level (Figure 42).

## RESULTS



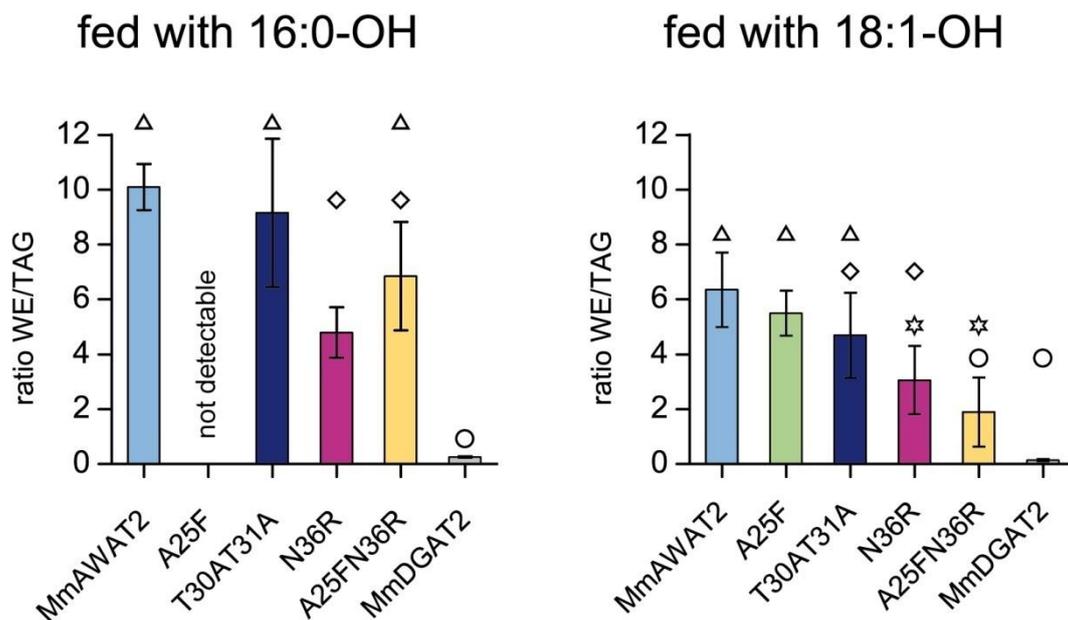
**Figure 41: Acyl chain composition of WEs synthesised by MmAWAT2 and respective single amino acid exchange variants in comparison with MmDGAT2 and the acyl-CoA pool composition.** Bars represent the relative amounts of each acyl chain in the WE fraction of *S. cerevisiae* H1246 cultures expressing the respective enzyme as well as relative amounts of each acyl chain in the in the acyl-CoA pool, respectively. Values for >C18 include acyl chains from 20 to 26 carbon atoms. Symbols indicate significant differences between the shown values at a level of  $p < 0.05$  as derived from a Tukey's test. Data show the mean and standard deviation of at least three samples derived from independent expression cultures.

As already mentioned above, the last two residues of the putative NLBS in AWAT2 sequences are mostly occupied by polar residues, whereas the respective residues in DGAT2 sequences are of highly conserved, non-polar nature. To mimic this situation in MmAWAT2, the T30A T31A variant was constructed. The corresponding WE fractions did not contain VLC WEs, but slightly increased levels of 16:1 acyl chains. According to a Tukey's test, this difference was not significant (Figure 41,  $p < 0.05$ ), whereas it was significant according to a student's t-test (Appendix 20, Appendix 21,  $p < 0.05$ ). The MmAWAT2 T30A T31A N36R variant showed a similar effect in comparison to the MmAWAT2 N36R variant. According to a student's t-test, the level of 16:1 chains in WEs from cultures expressing MmAWAT2 T30A T31A N36R is significantly higher in comparison to the N36R variant (Appendix 20,  $p < 0.05$ ). In contrast, it is not according to a Tukey's test (data not shown). However, this variant was only tested upon feeding 18:1-OH to respective cultures.

Apart from the residues of MmAWAT2 corresponding to the NLBS in DGAT2 sequences, an asparagine at position 36 in MmAWAT2 was further analysed. In vertebrate AWAT2 sequences, exclusively polar residues are found at the corresponding positions, while in DGAT2 sequences, only non-polar ones can be found (Figure 39). A replacement of this arginine with leucine, as it is present in most of the vertebrate DGTA2 sequences, did not result in any changes of WE composition or TAG accumulation upon feeding of 18:1-OH. A replacement with alanine yielded in a slight but significant decrease in 16:1 chain incorporation in WEs in comparison to MmAWAT2, as concluded from a student's t-test (Appendix 20,  $p < 0.05$ ). According to the predicted TM topology of MmAWAT2, N36 is located in a short stretch of five amino acids, which connects both TM domains and resides in the ER lumen. Hence, a side chain at this position which is not only polar, but carries a real charge, may interfere with the enzyme's activity. Therefore, the MmAWAT2 N36R variant was constructed. It showed a pronounced accumulation of VLC WEs upon feeding of 16:0-OH or 18:1-OH (Figure 31). Interestingly, the variant behaved very similar to the domain swap variants V2 and V5. Feeding of 16:0-OH resulted in incorporation of VLC acyl chains in the range from C20:0 to C26:0 in to WEs, whereas upon feeding of 18:1-OH, only 26:0 as VLC acyl chain was incorporated into WEs (representatively shown in Figure 37 and Figure 38). Relative quantification of the contribution of VLC WEs to the total WE fraction of cultures expressing MmAWAT2 N36R resulted in values of  $30 \pm 2\%$  and  $14 \pm 6\%$  upon feeding 16:0-OH and 18:1-OH, respectively ( $n \geq 3$ ) (Figure 41). Aside from

## RESULTS

those remarkably high amounts of VLC WEs, the N36R variant resembled domain swap variants V2 and V5 also in its substrate specificity concerning the incorporation of C16 and C18 acyl chains into WEs. Samples derived from cultures expressing MmAWAT2 N36R and fed with 18:1-OH showed severely decreased levels of 16:1 chains in the WE fraction, which were approximately half of the respective MmAWAT2 value. Upon feeding of 16:0-OH, 16:0 acyl chains were also significantly reduced, though less pronounced. On the other side, incorporation of 18:1-chains was unaltered, whereas 18:0-chains were abundant in amounts twice as high as compared to MmAWAT2 levels. This result was obtained irrespective of the fed alcohol (Figure 41). Just like the domain swap variants V2 and V5, MmAWAT2 N36R also showed decreased WE/TAG ratios, which were approximately 50 % of the MmAWAT2 values (Figure 42).



**Figure 42: Ratios of WE/TAG produced by MmAWAT2, respective single amino acid exchange variants and MmDGAT2.** Ratios were determined densitometrically after separation of respective lipid extracts via TLC. In case of samples derived from cultures expressing the MmAWAT2 A25F variant and fed with 16:0-OH, no TAG signals were detected on the respective TLC plates. Symbols indicate significant differences between the shown values at a level of  $p < 0.05$  as derived from a Tukey's test. Data show the mean and standard deviation of at least three samples derived from independent expression cultures.

Apart from being positively charged, an arginine residue is also rather bulky. In order to elucidate which of those two properties was essential for the observed effects associated with the N36R variant, the two MmAWAT2 variants N36W and N36K were constructed. Tryptophan in N36W is a bulky, but uncharged residue, whereas lysine in N36K carries a positive charge. Both amino acid exchanges resulted in a significant decrease of 16:1 chains as well as in a significant increase of 18:0 acyl chains in the WE fraction of respective expression cultures fed with 18:1-OH (student's t-test  $p < 0.05$ ). Notably, these changes were not as intense as respective changes observed in case of the MmAWAT2 N36R variant (Appendix 20). Moreover, the N36W variant showed a slight production of VLC WEs, which was significantly higher than respective values of MmAWAT2 (student's t-test  $p < 0.05$ ). Both variants were only tested upon feeding 18:1-OH to the cultures. Thus, although both variants exhibited slightly altered substrate specificities in comparison to wild type MmAWAT2, none of the variants was able to mimic the MmAWAT2 N36R phenotype to a full extend.

Since MmAWAT2 A25F and MmAWAT2 N36R were both found to synthesise VLC WEs, MmAWAT2 A25F N36R was constructed to elucidate the effects of both amino acid replacements in a double

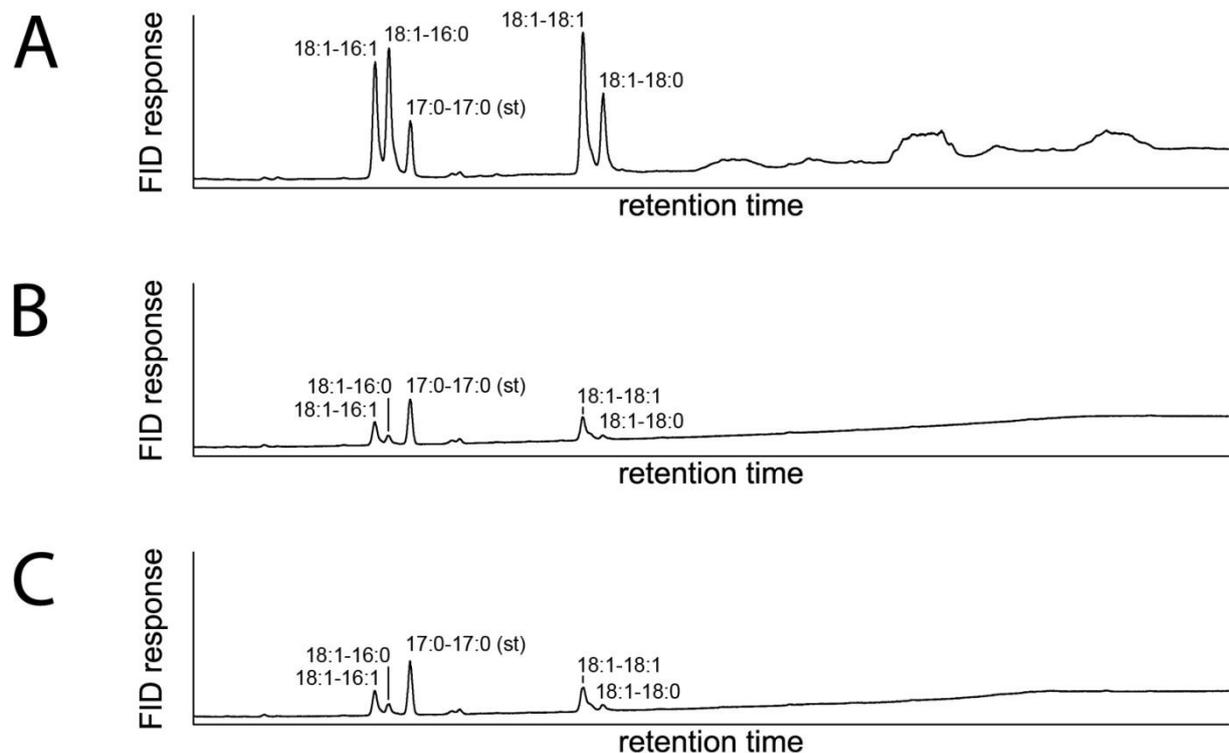
## RESULTS

variant. However, VLC WE amounts as well as acyl chain incorporation in WEs did not significantly differ from the MmAWTA2 N36R single variant, independent from the fed alcohol (Figure 41). The WE/TAG ratio of samples derived from cultures expressing the A25F N36R variant was neither significantly different from MmAWAT2 nor from the MmAWAT2 N36R variant upon feeding of 16:0-OH. In contrast, the WE/TAG ratio of samples derived from cultures expressing MmAWAT2 A25F N36R was on the MmAWAT2 N36R level and significantly different from MmAWAT2 upon feeding 18:1-OH (Figure 42).

### 5.3.8 Activity of MmDGAT2 F80A and MmDGAT2 L91R

Since the MmAWAT2 N36R variant showed a significantly altered substrate specificity in comparison to wild type MmAWAT2, it was elucidated whether the respective position is a general determinant of substrate specificity in MmDGAT2-like acyltransferases. In MmDGAT2, the amino acid corresponding to N36 in MmAWAT2 is a non-polar leucine residue at position 91. Hence, the corresponding MmDGAT2 L91R variant was constructed. Samples derived from respective yeast expression cultures did not contain detectable TAG signals upon TLC analyses (Appendix 25). Upon analyses of the samples via GC-FID to elucidate the presence of WEs (Figure 43 C), only residual amounts were detected, indicating a loss of enzymatic function or severe problems with protein folding.

It was described before that an exchange of phenylalanine at position 80 in MmDGAT2 for alanine results in a severe loss of DGAT activity (Stone et al., 2006). Analyses of lipid extracts from yeast cultures expressing MmDGAT2 F80A fed with 18:1-OH furthermore did not contain WEs either (Figure 43 B).



**Figure 43: Single amino acid exchange variants MmDGAT2 F80A and MmAWAT2 L91R are severely impaired in their enzymatic activity. A)** GC-FID trace of the WE fraction produced by *S. cerevisiae* H1246 expressing MmDGAT2. **B)** GC-FID

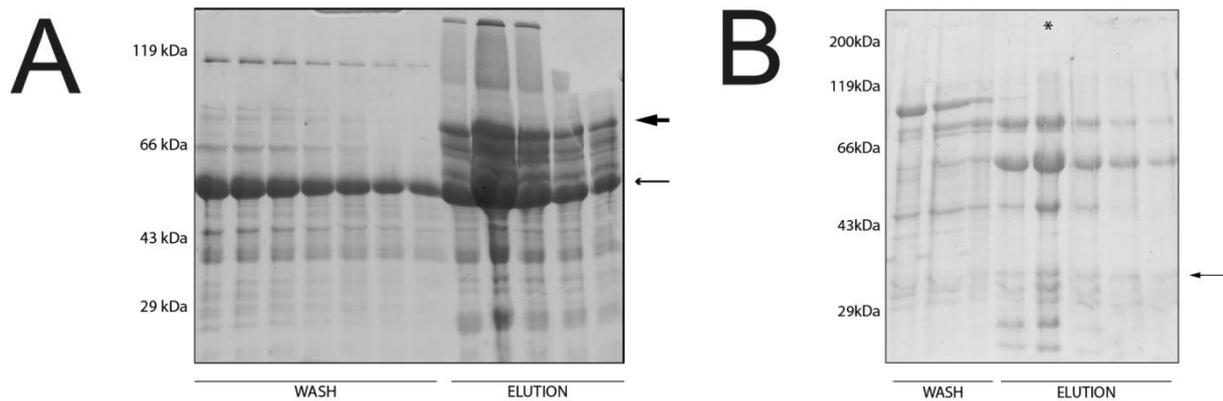
## RESULTS

trace of the WE fraction produced by *S. cerevisiae* H1246 expressing the MmDGAT2 F80A variant or the **C**) MmDGAT2 L91R variant, respectively. Equal amounts of heptadecanoyl-heptadecanoate (17:0-17:0) were added as an internal standard to all samples. All cultures were fed with 1 mM 18:1-OH. Data are representative for three samples derived from independent expression cultures in each case.

### 5.3.9 The putative transmembrane domains in MmAWAT2 might be dispensable for enzyme activity

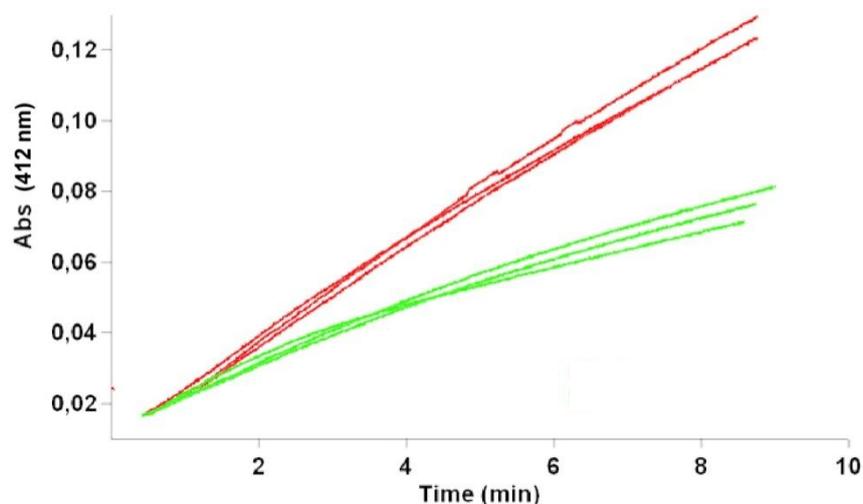
It has been shown earlier, that a truncated version of MmDGAT2, missing its N-terminal TM regions, is still active upon expression in human cells (McFie et al., 2011). In analogy to that, a corresponding truncated MmAWAT2 variant was constructed. The MmAWAT2 $\Delta$ N variant missed the first 63 N-terminal residues including both of the predicted TM domains. Upon expression of the enzyme in *S. cerevisiae* H1246, no signals for WEs or TAGs were detected via TLC analyses or GC-FID measurements (Appendix 22 & Appendix 23). Except for the putative TM domains at its N-terminus, no further membrane spanning structures for MmAWAT2 $\Delta$ N are predicted (Figure 35). Hence, MmAWAT2 $\Delta$ N is most likely a soluble enzyme. In order to exclude a lack of activity caused by a cytoplasmic localisation of the truncated variant, it was cloned in frame to C-terminal to the coding sequence of Oleosin 3 from *A. thaliana* (AtOLE3), resulting in the AtOLE3-MmAWAT2 $\Delta$ N fusion protein. In previous studies it had been shown, that not only AtOLE3, but also a fusion protein of AtOLE3 and the full length version of MmAWAT2 localises to the surface of oil bodies (Heilmann et al., 2012). In this respect, fusion of MmAWAT2 $\Delta$ N to AtOLE3 may result in a relocation of the truncated MmAWAT2 to ER or LD-derived membranes. However, expression of the AtOLE3-MmAWAT2 $\Delta$ N fusion protein in *S. cerevisiae* H1246 did not result in accumulation of WEs above levels of the empty vector control (Appendix 22). Hence, the expression system was changed. The sequence of MmAWAT2 $\Delta$ N was cloned in the bacterial pET28 and the pCold expression vectors. Both, the 6xHis-TF-MmAWAT2 $\Delta$ N fusion protein as well as the N-terminal hexahistidine-tagged version of MmAWAT2 $\Delta$ N (6xHis-MmAWAT2 $\Delta$ N) were expressed in *E. coli* BL21\* (DE3) and the soluble fraction of respective cell lysates was subjected to IMAC purification. In the course of 6xHis-TF-MmAWAT2 $\Delta$ N purifications, high amounts of protein were eluted upon washing of the IMAC column with 30 mM of imidazole after protein loading. The respective fractions contained high amounts of a protein running between the 43 and 66 kDa protein molecular weight marker bands upon SDS-PAGE analyses. This size might correspond to the sole 6xHis-TF, which is 54.4 kDa. Upon final elution of the IMAC column with 300 mM imidazole, an additional prominent protein species was observed in respective samples upon SDS-PAGE analyses. This second prominent protein band ran little above the 66 kDa-band of the protein molecular weight marker and thus would fit to the calculated size of the 6xHis-TF-MmAWAT2 $\Delta$ N fusion protein, which is 83.8 kDa. However, this band was less intense than the putative 6xHis-TF band (Figure 44 A). In case of purification of 6xHis-MmAWAT2 $\Delta$ N, the resulting elution fractions were not pure. The calculated size of 6xHis-MmAWAT2 $\Delta$ N is 34.4 kDa. However, no prominent band at the respective position in the corresponding gels was observed (Figure 44 B).

## RESULTS



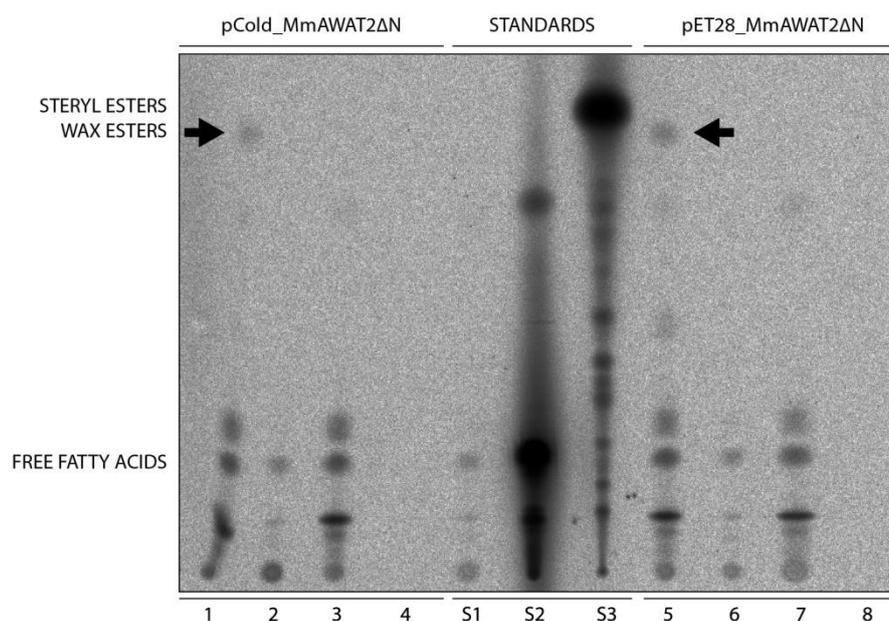
**Figure 44: IMAC purification of heterologously produced MmAWAT2 $\Delta$ N.** **A)** Purification of 6xHis-TF-MmAWAT2 $\Delta$ N. After protein loading, the IMAC column was washed with buffer containing 30 mM imidazole. Bound protein was eluted with buffer containing 500 mM imidazole. The bold arrow indicates a protein band which corresponds to the size of 6xHis-TF-MmAWAT2 $\Delta$ N (83.3 kDa), whereas the thinner arrow indicates the band corresponding to the size of 6xHis-TF (54.4 kDa). Purification of 6xHis-TF-MmAWAT2 $\Delta$ N was done twice. **B)** Purification of 6xHis-MmAWAT2 $\Delta$ N. After protein loading, the IMAC column was washed with buffer containing 30 mM imidazole. Bound protein was eluted with buffer containing 500 mM imidazole. The arrow indicates a band which might correspond to the size of 6xHis-MmAWAT2 $\Delta$ N (34.4 kDa). The fraction marked with an asterisk was used for the activity assay shown in Figure 45. Purification of 6xHis-TF-MmAWAT2 $\Delta$ N was done twice.

The *in vitro* activity of elution fractions derived from the IMAC purifications of both, 6xHis-TF-MmAWAT2 $\Delta$ N and 6xHis-MmAWAT2 $\Delta$ N, was further elucidated. Therefore, DTNB-based *in vitro* experiments with 18:1-OH and 16:0-CoA as substrates were performed. Respective experiments resulted in slight differences between samples containing protein derived from a 6xHis-MmAWAT2 $\Delta$ N purification and negative controls lacking fatty alcohol (Figure 45). In case of 6xHis-TF-MmAWAT2 $\Delta$ N, respective experiments did not show significant differences between negative controls and actual samples (data not shown). In order to further elucidate the activity of the elution fractions, *in vitro* assays using [ $^{14}$ C] labelled acyl-CoAs and unlabelled fatty alcohol as substrates were performed. Elution fractions derived from IMAC purification of both, 6xHis-MmAWAT2 $\Delta$ N as well as 6xHis-TF-MmAWAT2 $\Delta$ N, were capable of mediating WE synthesis. In contrast, negative controls did not show WE signals (Figure 46).



**Figure 45: DTNB-based *in vitro* activity tests with partially purified, heterologously produced 6xHis-MmAWAT2 $\Delta$ N.** Samples contained 0.5 mM DTNB, 20  $\mu$ M oleyl alcohol (18:1-OH), 100  $\mu$ l MmAWAT2 $\Delta$ N protein solution (asterisk marked fraction shown in Figure 44 B) and 12.5  $\mu$ M palmitoyl-CoA (16:0-CoA) in a total volume of 1 ml TBS. Samples were pipetted without 16:0-CoA, mixed thoroughly and placed in a spectrophotometer. Absorption was monitored at 412 nm. Once a stable baseline was formed, 16:0-CoA was added and samples were mixed again. Red traces show samples with addition of 18:1-OH, while green traces show samples without the addition of 18:1-OH. Data are representative for a single experiment.

## RESULTS



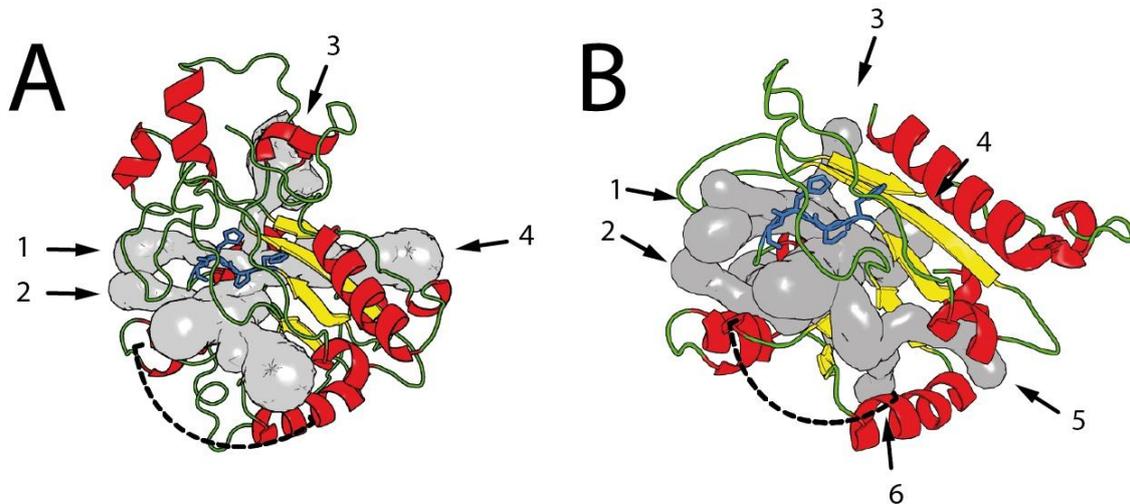
**Figure 46: MmAWAT2ΔN *in vitro* assay with [<sup>14</sup>C] labelled substrates.** Samples contained 10 μg protein in case of 6xHis-MmAWAT2ΔN and 20 μg of protein in case of 6xHis-TF-MmAWAT2ΔN, 100 μM 18:1-OH and 50 μM [<sup>14</sup>C] labelled oleyl-CoA (18:1-CoA) in a total volume of 50 μl TBS. Samples were incubated for 30 min at 30 °C. The reaction was stopped by addition of 100 μl 2 M HCl. Afterwards, 800 μl hexane were added and samples were vortexed for 15 min at full speed. Phase separation was done by centrifugation for 1 min at 16000 x g. Organic phase was recovered and evaporated under a stream of nitrogen. Samples were resolved in 50 μl of hexane and loaded on a TLC plate. The TLC plate was developed in hexane:diethyl ether:acetic acid (80:20:2, v/v/v) and placed on an image plate (Fuji, Japan) for 2 days. Afterwards, signals were analysed using a Fuji FLA-3000 phosphor imager. Black arrows indicate WE spots. 1-4: samples on the basis of MmAWAT2ΔN. 1 = complete sample, 2 = no enzyme, 3 = no 18:1-OH, 4 = no [<sup>14</sup>C] 18:1-CoA, S1 = [<sup>14</sup>C] 18:1-CoA standard, S2 = [<sup>14</sup>C] 18:1-COOH standard, S3 = [<sup>14</sup>C] steryl oleate standard. 5-8: samples on the basis of 6xHis-TF-MmAWAT2ΔN. 5 = complete sample, 6 = no enzyme, 7 = no 18:1-OH, 8 = no [<sup>14</sup>C] 18:1-CoA. Data for both, 6xHis-MmAWAT2 and 6xHis-TF-MmAWAT2 are representative for more than three independent samples from two independent experiments with protein from a single purification, respectively.

### 5.3.10 Modelling of MmAWAT2 and MmDGAT2

Modelling of the full length version of MmAWAT2 by the Phyre server (Kelley and Sternberg, 2009) resulted in unsatisfactory results. The service was not able to model any residue of the sequence with over 90 % accuracy. Moreover, 218 of 333 residues were modelled by *ab initio*, which, according to the Phyre service, is highly unreliable. As a template, the structure of triosephosphate isomerase from *C. reinhardtii* (PDB accession code 4MKN, (Zaffagnini et al., 2014) was used. However, the confidence in real homology between the MmAWAT2 sequence and the used template was 87.3 % and thus relatively low. Hence, the sequence of the truncated version of MmAWAT2, which lacks the predicted N-terminal TM domains, was submitted. This time, the Phyre server identified a glycerol-3-phosphate acyltransferase from *Cucurbita moschata* (PDB accession code 1IUQ, Tamada et al., 2004) as the best fitting template and confidence in real homology was 91.2 %. Moreover, 73 % of all residues were modelled with over 90 % confidence. The resulting structure shows five central β-sheets, which are arranged in a half barrel form. This structure is further surrounded by three α-helices and nine smaller helical structures. In order to visualise potential tunnels or caves for substrate binding, the PyMol Caver plugin 2.1.2 (Kozlikova et al., 2014) was used. Caver identified four tunnels, which all have access to the proteins surface and meet in proximity to the “HPHG”

## RESULTS

active site motif of the protein. Additionally, a rather wide pocket on the surface of the model was identified, which has access to the HPHG motif as well (Figure 47 A).



**Figure 47: Models of MmAWAT2 $\Delta$ N and MmDGAT2 $\Delta$ N.** A) Model of MmAWAT2 $\Delta$ N. B) Model of MmDGAT2 $\Delta$ N. Black arrows indicate tunnels which have access to the surface of the enzyme. Numbers indicate corresponding tunnels in both models. The dashed line in both models indicates a rather wide pocket with access to the proteins surface.  $\alpha$ -helices are shown in red,  $\beta$ -sheets are shown in yellow. Connecting loops are shown in green, while the HPHG active site motif of both enzymes is shown as blue sticks. The models were generated with the Phyre server (Kelley and Sternberg, 2009), while the tunnels (grey structures within the enzymes) were predicted using the Caver-plugin for PyMol (Kozlikova et al., 2014).

The sequences of MmDGAT2 and MmDGAT2 $\Delta$ N were also submitted to the Phyre server in order to elucidate if there were obvious structural differences in both proteins. As for the full length version of MmAWAT2, also the predicted structure of MmDGAT2 was derived from 274 out of 388 residues which were modelled by *ab initio*, and thus highly unreliable. For the truncated version of MmDGAT2, the Phyre server chose the structure of triosephosphate isomerase from *C. reinhardtii* as the best fitting template. However, to better compare the models for MmAWAT2 $\Delta$ N and MmDGAT2 $\Delta$ N, a model which was computed according to the glycerol-3-phosphate acyltransferase from *Cucurbita moschata* was chosen for comparison with MmAWAT2 $\Delta$ N. The confidence for real homology between MmDGAT2 $\Delta$ N and the glycerol-3-phosphate acyltransferase was 86.7 %. Phyre threaded 82 % of the residues with a confidence over 80 %. The resulting model was organised similar to the respective model of MmAWAT2 $\Delta$ N. The centre of the model consists of six  $\beta$ -sheets, which are surrounded by four  $\alpha$ -helices and five small, helical structures. Caver predicted the presence of six tunnels. Four tunnels are analogous to the model of MmAWAT2 $\Delta$ N, while two tunnels are exclusive for the model of MmDGAT2 $\Delta$ N. All tunnels have access to the surface of the protein. Additionally, a pocket similar to the MmAWAT2 $\Delta$ N was predicted (Figure 47 B). Interestingly, the central cavity in proximity to the HPHG motif appears to be bigger in comparison to the model of MmAWAT2 $\Delta$ N. Since acceptable results for the modelling of the truncated versions of both MmAWAT2 and MmDGAT2 to the structure of glycerol-3-phosphat acyltransferase from *Cucurbita moschata* were obtained, the full length version of both enzymes was resubmitted to Phyre in order to model them on the same structure using the one on one threading mode. However, one-on-one threading of MmAWAT2 and MmDGAT2 to the *Cucurbita moschata* enzyme did not yield convincing results either (data not shown).

## 6 DISCUSSION

Previous studies indicate a broad substrate range for many FARs and WSs. Furthermore, the substrate specificities of certain FARs and WSs were shown to heavily alter upon expression in different hosts. These circumstances can be problematic for the biotechnological production of fatty alcohols or WEs, where defined blends are generally preferred over undefined mixtures of respective substances. In order to optimise the substrate specificities of FARs and WSs, knowledge about structure-function relationships in these enzymes is necessary. However, this knowledge is scarce to date.

The aim of this work was therefore to contribute to the understanding of structural aspects in FARs and WSs. In this respect, attempts to purify and crystallise different FARs and WSs were performed and succeeded in case of the bacterial bifunctional WS/DGAT from *A. baylyi*. Obtained results and issues associated with this project are discussed in the following sections.

A second focus was laid on the elucidation of substrate specificity determinants in the murine MmAWAT2. For this, comparative studies on the basis of MmAWAT2 and MmDGAT2, followed by the construction of enzyme chimeras and single amino acid exchange variants, have been performed. The experiments indicated an important role of two predicted TM domains in substrate specificity determination of MmAWAT2. Respective findings are discussed in the following section.

### 6.1 Purification of FARs and WSs

#### 6.1.1 Enzymatic activity of DmFAR1

The substrate specificity of the insect-derived DmFAR1 was first described in the course of this study. Interestingly, the specificity was shown to severely differ upon expression in either *S. cerevisiae* or *E. coli* (Figure 15 & Figure 16). While expression of DmFAR1 in yeast yielded in the detection of 24:0-OH and 26:0-OH, expression in *E. coli* resulted in the detection of 14:0-OH, 16:0-OH, 16:1-OH and 18:1-OH. Since *E. coli* synthesises vaccenic acid ((11Z)-18:1-COOH) instead of oleic acid ((9Z)-18:1-COOH), the detected 18:1-OH is most probably (11Z)-18:1-OH. All of the acyl-CoA substrates for the production of the alcohols found in *E. coli* except for 14:0-CoA and (11Z)-18:1-CoA were detected in *S. cerevisiae* in the course of this study. However, *S. cerevisiae* synthesises (9Z)-18:1-COOH, which in principle could serve as a substrate for the production of (9Z)-18:1-OH. Thus, the yeast cell might be able to produce 16:0-OH, 16:1-OH and 18:1-OH as well. A possible explanation for the lack of these alcohols upon expression of DmFAR1 in yeast might be an insufficient sensitivity of the GC-MS method which was utilised in the present study. Yet, respective fatty alcohols might be present in the samples, but in quantities which are below the detection limit. A missing access of DmFAR1 to the respective acyl-CoA substrates *in vivo* might be another explanation for the lack of alcohols with shorter chain lengths in samples derived from yeast cultures expressing DmFAR1. In this respect, DmFAR1 might only have access to a specific substrate pool which is primarily composed from VLC acyl-CoAs, but is missing acyl-CoAs of shorter chain length. This might be due to the localisation of DmFAR1 in respective microdomains of the yeast membrane. In contrast, a lack of 24:0-OH and 26:0-OH in samples derived from *E. coli* can be explained by the absence of the required substrates within this organism. Feeding of 26:0-COOH to expression cultures did not result in 26:0-OH synthesis

## DISCUSSION

either. This might be due to a missing ability of *E. coli* to activate VLC fatty acids to the corresponding acyl-CoA or acyl-ACP esters.

DmFAR1 is an excellent example for the difficulties accompanied with *in vivo* substrate specificity determination of FARs. Like AtFAR3 from *Arabidopsis thaliana* (Domergue et al., 2010), it shows completely different substrate specificities in yeast and *E. coli*. One possibility to elucidate the correct substrate specificities of DmFAR1 may be *in vitro* assays based on purified enzyme. However, first experiments in this respect did not yield detectable activity (Appendix 6). Due to time constraints, these experiments were performed under non-optimised standard conditions. Also, only NADPH was utilised as a reducing equivalent in the assays. Hence, one explanation for missing *in vitro* activity of DmFAR1 might be that DmFAR1 is dependent on NADH rather than on NADPH. However, additional work in order to optimise DmFAR1 *in vitro* assays is necessary for further characterisations of its substrate specificities.

FARs showing a similar substrate specificity like DmFAR1 upon expression in *E. coli* and yeast have been described from the copepod *Calanus finmarchicus* (Teerawanichpan and Qiu, 2011) and *Arabidopsis thaliana* (Domergue et al., 2010). The respective enzymes were termed CfFAR2 and AtFAR3, respectively, and yielded in accumulation of only 24:0-OH and 26:0-OH upon expression in *S. cerevisiae*. Similar to DmFAR1, AtFAR3 produced 14:0-OH, 16:0-OH and 18:1-OH in *E. coli*. Its substrate specificity it thus highly similar to DmFAR1. However, the sequences of AtFAR3 and CfFAR2 do not cluster with DmFAR1 in phylogenetic analyses (Figure 4), indicating a lack of relationship between substrate specificity on the one hand and sequence homology on the other.

### 6.1.2 Purification of DmFAR1

Purification of the DmFAR1 protein without any additional protein tag yielded different results for the N- and C-terminal 6xHis-tagged version of the protein. Purification of N-terminal 6xHis-tagged DmFAR1 resulted in two prominent bands approximately matching the expected molecular weight of the enzyme. On the one hand, one of these bands could represent an unrelated impurity. On the other hand, it could also represent a degradation product of DmFAR1. According to the modelled structure of DmFAR1, the N-terminus is highly exposed to the solvent (Figure 13). Fusing of the 6xHis-tag to the N-terminus might result in an even higher susceptibility of this part of the enzyme to proteases, which might result in protease-mediated degradation of the N-terminus. Partially degraded protein without its N-terminus may nevertheless be purified as a quaternary complex, in which truncated variants interact with full length DmFAR1. As a result, SDS-PAGE analyses would show a mixture from full length and degraded protein. In contrast, the N-terminus of 6xHis-TF-DmFAR1 and MBP-DmFAR1 might be protected from protease mediated degradation by the respective fusion partners. Accordingly, no degradation products have been observed during respective purifications (Appendix 2 & Appendix 3).

Purification of the C-terminal 6xHis-tagged version of DmFAR1 only yielded one prominent, but smeary band upon SDS-PAGE analysis. The position of this band within the gel is essentially consistent with the calculated size of the protein. Unusual migration patterns of membrane proteins were described before (Rath et al., 2009). According to the theory presented by Rath et al., altered SDS binding at TM domains of proteins may cause a mixture of protein molecules with different amounts of bound SDS. This in turn is reflected by different migration properties of respective molecules upon SDS-PAGE analysis. Fittingly, it is quite likely that DmFAR1 is an integral membrane protein. Two predicted TM domains in the enzyme's sequence as well as results indicating an ER

## DISCUSSION

localisation of the protein (Jaspers et al., 2014) support this idea. Hence, the smeary band upon SDS-analysis may indeed represent a mixture of DmFAR1 molecules with different amounts of bound SDS. However, to verify the presence of the DmFAR1 in the elution fractions, immunostaining experiments or analyses of the respective protein bands via mass spectrometry have to be conducted.

### 6.1.3 TcFAR1

Expression of the *Tribolium*-derived TcFAR1 in *S. cerevisiae* did not yield any detectable fatty alcohols in respective extracts. Since no attempts were undertaken in order to detect the protein in cell extracts via immunostaining experiments, it is possible that transcription, translation or folding of the enzyme are erroneous. The simplest explanation for the missing activity of the enzyme upon expression in yeast would thus be a non-existent or misfolded protein.

Phylogenetic analyses of different FAR sequences indicated a relatively high similarity between DmFAR1 and TcFAR1 among the analysed sequences (Figure 4). Thus, it might well be that TcFAR1 is furthermore similar to DmFAR1 in its substrate specificity for VLC acyl-CoA. Following this, it is possible that TcFAR1 possesses a substrate specificity, which is inconsistent with the available substrate pool in yeast. Since there are no polyunsaturated acyl-CoAs or acyl-CoAs longer than C26 present in yeast, it remains elusive if the enzyme is capable of reducing respective molecules. In contrast, CdFAR2 from the copepod *Calanus finmarchicus* and AtFAR3 from *Arabidopsis thaliana* possess a similar substrate specificity to DmFAR1, although they do not cluster with DmFAR1 and TcFAR1 upon phylogenetic analyses. Thus, clustering of DmFAR1 and TcFAR1 in phylogenetic analyses can't be projected on a similar substrate specificity.

Furthermore, the present study showed a severely different substrate specificity of DmFAR1 upon expression in either *S. cerevisiae* or *E. coli*. In analogy to that, also TcFAR1 might exhibit an altered specificity upon expression in *E. coli* as well. However, expression of TcFAR1 in *E. coli* was not performed.

### 6.1.4 Interaction of TF with MaFAR1 and DmFAR1

In contrast to the fusion protein from 6xHis-TF and AbWSD1, expression of 6xHis-TF-DmFAR1 and 6xHis-TF-MaFAR1 did not result in defined multimeric states, but in aggregates (representatively shown in Figure 19). Possible reasons for such different solubility promoting effects of the TF are diverse. TF is a member of the bacterial cell's chaperone machinery. *In vivo*, it is mostly located at the exit of the ribosome and binds to protein L23 of the 50S ribosomal subunit via its N-terminal ribosomal binding domain (Kramer et al., 2002). It is the first protein which interacts with nascent peptide chains and shields them from the surrounding solvent in order to promote correct folding (Ferbitz et al., 2004). Moreover, it conducts folding of the peptides via a peptidyl-prolyl isomerase domain (Stoller et al., 1995). TF binds other proteins by hydrophobic interactions, conducted by four specific patches of the enzyme. All these patches are located in a sort of cradle, which is formed by the N-terminal and the C-terminal domain of the protein as well as by the middle part of the sequence (Ferbitz et al., 2004). One possible reason for a missing ability of the TF to efficiently prevent aggregate formation upon being fused to either MaFAR1 or DmFAR1 may be due to steric effects. Since the TF and its binding partner are fused to each other by a 34 amino acid linker, their position relative to each other is not flexible. Hence, TF might not be able to reach certain surface

## DISCUSSION

exposed, hydrophobic parts of MaFAR1 or DmFAR1. On the other hand, the TF could also exhibit a negative effect on its fusion partner. Since the TF is positioned at the N-terminal end of its fusion partner, it is likely that it is already folded and active while its C-terminal fusion partner is still being translated at the ribosome. In this respect, interaction of the TF with hydrophobic parts of the currently translated fusion partner might disturb its correct folding and promote aggregate formation. Also, heterologous TF may block the activity of intrinsic TF molecules of the host *E. coli*, since it may occupy their binding sites at the ribosome. One observation which might support the idea of a negative effect of the TF towards its binding partner is the fact that 6xHis-TF-DmFAR1 was only purifiable in the form of protein aggregates, whereas preliminary results indicate that 6xHis-tagged DmFAR1 lacking TF does not seem to form aggregates during purification.

A third explanation for aggregate formation despite of using TF as a fusion partner for protein expression could be an insufficient stoichiometric ratio of TF and the desired protein. It has been shown, that numerous TF molecules can bind to a single unfolded protein molecule in order to facilitate its correct folding (Saio et al., 2014). Given the fact that DmFAR1 and MaFAR1 may exhibit multiple hydrophobic patches on their surfaces, a 1:1 ratio of TF and the fusion partner might not be sufficient to cover all patches, resulting in aggregate formation. An observation which might support this idea is the fact that MaFAR1-6xHis, not fused to TF, precipitated upon elution from the IMAC column under non-optimised conditions, whereas 6xHis-TF-MaFAR1 was present in soluble aggregates. Thus, the fusion of MaFAR1 to TF seems to prevent the precipitation of MaFAR1, but the solubility enhancing properties of TF might not be sufficient to prevent 6xHis-TF-MaFAR1 aggregate formation. A strategy in order to overcome issues caused by the stoichiometry of MaFAR1 and TF might be the additional expression of the single TF in *E. coli*. The ratio of TF/target protein could be increased like this and heterologously produced proteins might be able to bind more than one TF.

### 6.1.5 MaFAR1

As described in 5.1.1.4.1, SEC analyses of purified 6xHis-TF-MaFAR1 resulted in elution of the protein in the void volume of the column, indicating the presence of protein aggregation (Figure 19). Purification of the sole MaFAR1 without the fused TF led to precipitation of the protein under non-optimised conditions. Upon analyses of respective fractions via SDS-PAGE, several bands in the expected height, which migrated in close proximity to each other, were observed (Figure 18, Figure 19). This migration pattern might indicate degradation of the enzyme. In the modelled enzyme structure, the 23 N-terminal amino acids are predicted to adopt an unordered loop (Figure 13 D), which might be highly accessible for proteases. Interestingly, this potential degradation pattern was only observed for MaFAR1-6xHis, but not for 6xHis-TF-MaFAR1. In 6xHis-TF-MaFAR1, the N-terminus is fused to the TF, which thus might prevent an access of proteases to this part of the enzyme. However, since 6xHis-TF-MaFAR1 also forms aggregates, a possible partial degradation of the enzyme cannot be the only reason for aggregate formation. Analyses of the MaFAR1-sequence for TM domains using TMHMM (Sonnhammer et al., 1998) did not yield probable TM regions or hydrophobic regions at all (Figure 14). However, hydrophobicity analyses according to Kyte & Doolittle (Kyte and Doolittle, 1982) indicated the presence of four rather hydrophobic motifs in the sequence (Appendix 8). Analyses of these motifs in the modelled MaFAR1 structure indicated that all of them are exposed to the surface of the model. The C-terminal one of these four motifs is moreover predicted to form an exposed  $\alpha$ -helix (data not shown). Thus, aggregation of the MaFAR1

## DISCUSSION

protein upon purification from *E. coli* may be caused by interactions of hydrophobic patches of the protein with each other in case of both, MaFAR1-6xHis and 6xHis-TF-MaFAR1.

Interestingly, *in vitro* experiments with purified MaFAR1 enzyme did only yield fatty alcohols, when at least 10 mM NADPH were added to the reaction mixture (Figure 21). The acyl-CoA concentration in respective experiments was 50  $\mu$ M and thus 200-fold smaller. In order to reduce one molecule of acyl-CoA to the corresponding alcohol, two molecules of NADPH are necessary. However, upon addition of 100  $\mu$ M of NADPH, no fatty alcohols were detected. Thus, NADPH seems to have an additional effect on the MaFAR1 enzyme besides being the reduction equivalent. Since the original publication on MaFAR1 also describes the use of such high amounts of NADPH, it seems unlikely that the here observed requirement is specifically associated with the employed purification procedure of MaFAR1 or the setup of the *in vitro* assay. Based on the experimental results obtained so far, no plausible hypothesis explaining the requirement of this massive NADPH excess in the samples can be derived. One possible explanation may be a kind of stabilizing effect, resulting in prevention of protein aggregates, which would otherwise interfere with the enzyme's activity.

Another interesting observation was the fact that precipitation of the MaFAR1-6xHis protein was avoided by supplementation of respective cultures with 300 mM NaCl and 1 mM betaine. Moreover, aggregation of the MaFAR1 protein was avoided by the use of a basic buffer. Betaine may act as an osmolyte. It is known to be accumulated by the cell in order to counteract the high osmolarity of the media caused by high salt concentrations. Upon addition to media containing high salt, betaine is preferentially taken up by the cell (Roberts, 2005). To date, it is not fully understood how osmolytes prevent proteins from misfolding and aggregation. Theories span from an unfavourable, direct interaction of osmolytes with hydrophobic or unfolded parts of proteins over osmolyte induced changes in the water structure to osmolyte induced prevention of enzyme water penetration under high osmotic conditions (Roberts, 2005). It is thus reasonable to assume that betaine is able to counteract misfolding tendencies of the MaFAR1 protein. On the other hand, the high pH of the utilised CHES buffer might induce an altered surface charge of MaFAR1. This might lead to a partial destabilization of the protein aggregates, resulting in the observed elution pattern upon SEC analyses (Figure 19). However, no activity tests at pH = 9.5 were performed. Thus it remains elusive if the basic conditions have a negative effect on the enzymes activity.

### 6.1.6 Models of FARs

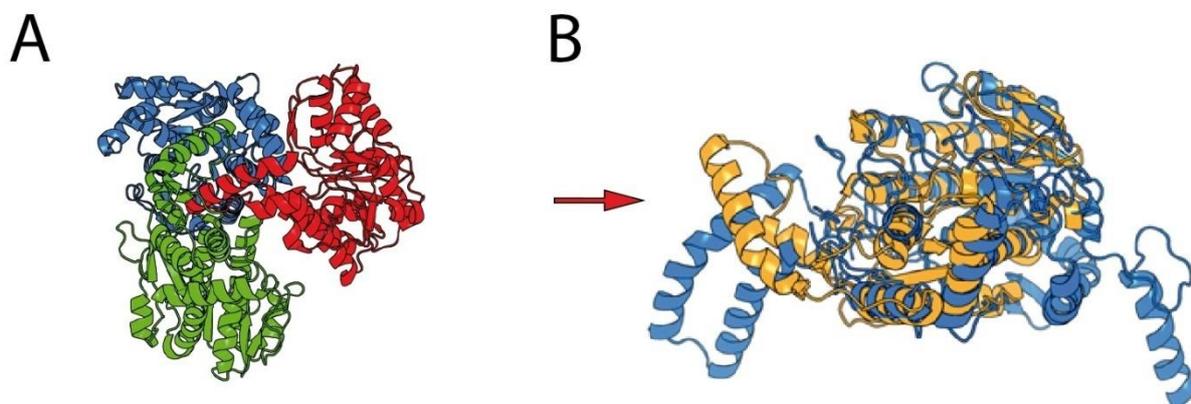
Three of the four FARs analysed within this study are predicted to contain a characteristic hairpin-like structure. It is composed of two  $\alpha$ -helices, which are encoded in the C-terminal quarter of the sequences (Figure 13). The hairpin is present in DmFAR1, TcFAR1 and MmFAR1, but not in MaFAR1. The reductase domain of a non-ribosomal peptide synthetase from *M. tuberculosis*, which was used as template structure for modelling in all cases, does not contain such a hairpin either. Interestingly, one of the two hairpin building helices is predicted to form a TM domain in TcFAR1 and DmFAR1. The corresponding helix is encoded approximately 100 amino acids away from the C-terminus in both enzymes. In MmFAR1, however, none of the hairpin building helices is predicted to form a TM domain. The probability for the hairpin building helices in MmFAR1 to form TM domains is only 30%. Apart from that, DmFAR1 and TcFAR1 are predicted to contain one additional TM domain at their very C-terminus, while MmFAR1 is predicted to contain two (Figure 14).

In case of DmFAR1 and MmFAR1 it was already shown, that the respective enzymes are located at the ER (Jaspers et al., 2014) and at peroxisomes (Cheng, 2004b; Heilmann et al., 2012), respectively.

## DISCUSSION

A possible function of the hairpin-like structures could thus be to serve as an anchor for the proteins in the membrane. As already mentioned, the two potential TM domains at the very C-terminus of MmFAR1 do not contribute to the hairpin structure of MmFAR1. Deletion of these domains led to a loss of peroxisomal localisation of the enzyme anyway, although the hairpin structure still was present. Thus, membrane localisation of MmFAR1 is unlikely to be mediated by the hairpin, but rather by the predicted C-terminal TM regions in MmFAR1. It remains elusive, if membrane association of DmFAR1 might likewise be caused by its predicted C-terminal TM domain.

A recently published article by Yin et al. describes the structure of Gox2253 from *Gluconobacter oxydans*, which is a fatty aldehyde reductase of the SDR enzyme family. Since reduction of fatty aldehydes basically corresponds to the second step of the FAR reaction, both enzyme classes are mechanistically closely related. The structure of Gox2253 was solved in a homo-trimeric form (Yin et al., 2014). Interestingly, Gox2253 trimerises with the help of a hairpin-like structure (Figure 48 A), which is fairly similar to the one in the modelled FAR structures of DmFAR1, TcFAR1 and MmFAR1. Upon alignment of a Gox2253 monomer with the modelled structure of DmFAR1 it becomes furthermore visible, that respective domains are localised on the same side of the enzyme (Figure 48 B). The same holds true for alignments of Gox2253 with TcFAR1 and MmFAR1, respectively (data not shown). Hence, the hairpin-like structure might resemble a potential oligomerisation domain of the FAR enzymes. This theory can be supported by the fact that initial SEC analyses of purified DmFAR1 resulted in elution of protein at volumes corresponding to a dimeric and trimeric state of the native enzyme.



**Figure 48: Gox2253 trimerises with the help of a hairpin motif. A)** Trimeric structure of the aldehyde reductase Gox2253 according to Yin et al. (2014) (PDB accession code 3WJ7). **B)** Alignment of the structures of a Gox2253 monomer (orange) and the modelled DmFAR1 (blue). Structures were aligned using the PyMol software. The exposed hairpin structures of both enzymes are indicated by the arrow.

Most FAR enzymes described today are strictly NADPH dependent and show no or only weak *in vitro* activity with NADH as the reduction equivalent (Doan et al., 2012; Pollard et al., 1979; Shi et al., 2011; Willis et al., 2011). Structural reasons for this preference in FARs are not elucidated until now. However, in several other extended SDR enzymes, respective determinants have already been identified (Yin et al., 2014; Yuan et al., 2013). In the strictly NADPH dependent bacterial aldehyde reductase Gox2253 it was shown, that replacement of an arginine at position 44 by a lysine or tryptophan severely changed the specificity for NADPH as the cofactor. Respective variants were able to utilise NADH and NADPH in equal efficiencies, while the wild type enzyme did not show any activity with NADH. The authors of the study stated that the cofactor preference in the wild type

## DISCUSSION

enzyme is achieved by interaction of NADPH with Gox2253 via a hydrogen bond between the 2'-phosphate of the NADPH molecule and the enzyme. In case of NADH, this stabilization is missing, resulting in an insufficient binding of NADH to the enzyme. In order to achieve a stabilization of NADH molecules in the enzyme, the authors conducted saturated virtual mutation analyses. They identified a position within the enzyme in close proximity to the adenosine moiety of NADPH, which was occupied by the earlier mentioned arginine residue. Upon mutation to lysine or tryptophan, the corresponding variants mediated NADH stabilization through an anticipated, additional hydrogen bond towards the 2'-hydroxy group of the adenosine of NADH (Yin et al., 2014).

In order to investigate to what extent this determinant might also play a role in the determination of cofactor specificity in FARs, the structure of Gox2253 was aligned with the models of DmFAR1, TcFAR1, MmFAR1, MaFAR1 and EgFAR1. EgFAR1 from *E. gracilis* is described to be strictly NADH dependent (Khan and Kolattukudy, 1973; Teerawanichpan and Qiu, 2010). The structures of MaFAR1 and MmFAR1 did not superimpose with Gox2253 to a degree suitable for direct comparison of the NADPH binding site of the enzymes. However, the NADPH binding sites of TcFAR1, DmFAR1 and Gox2253 showed a high similarity (Appendix 9 A). However, it is not known whether DmFAR1 is reliant on NADPH or NADH as a reducing equivalent. In order to elucidate possible structural differences at the NAD(P)H binding sites of Gox2253 and EgFAR1, the structure of EgFAR1 was modelled and superimposed to Gox2253. EgFAR1 from *E. gracilis* also possessed the arginine residue shown to be crucial in Gox2253. Apart from that, it also showed two additional residues (K44 and R52) in near proximity, which possibly could form hydrogen bonds with the adenosine moiety of NADH in order to stabilise it (Appendix 9 B). Although the validity of the generated FAR models is naturally low, it would still be interesting to see if a mutation similar to the Gox2253 R42K variant in DmFAR1 and TcFAR1 or an EgFAR1 variant with mutations in R42, K44 or R52 would alter the cofactor specificity of these enzymes.

### 6.1.7 AtDGAT3

The class of soluble DGAT, termed DGAT3, was established in 2006 with the discovery of a DGAT3 in developing cotyledons of *Arachis hypogaea* (peanut) (Saha et al., 2006). Since then, several other members of this class were identified and analysed from *Rhodotorula glutinis* (RgDGAT3) (Rani et al., 2012) and *Arabidopsis thaliana* (AtDGAT3 and AtDCR) (Hernandez et al., 2012; Rani et al., 2010). AtDGAT3 was identified during a study on TAG homeostasis in *Arabidopsis* mutant seedlings which were impaired in TAG breakdown. In the respective study, an increase in 18:3 and 18:2 fatty acids within the TAG pool was observed together with an up regulation of the AtDGAT3 transcript in the cotyledons of the seedlings. To verify a possible DGAT activity of AtDGAT3, the gene was transiently expressed in *N. benthamiana*. The authors reported increased TAG contents in *N. benthamiana* leaves transformed with AtDGAT3, especially of TAG species carrying 18:2 and 18:3 fatty acids. In contrast, a knockout of the gene in the TAG breakdown deficient background led to decreased 18:2 and 18:3 acyl chain levels in the TAG pool. Although these data are strong hints towards a DGAT activity of AtDGAT3, a direct enzymatic proof for a DGAT activity of AtDGAT3 is lacking (Hernandez et al., 2012). For several other DGAT3, this evidence has already been provided. In the course of the respective studies, recombinant enzymes were produced in *E. coli* and purified to homogeneity. The activity of the enzymes was shown by *in vitro* TAG formation on the basis of [<sup>14</sup>C] labelled substrates (Rani et al., 2010, 2012; Saha et al., 2006). Moreover, the ability to restore TAG synthesis in the

## DISCUSSION

quadruple knockout mutant strain *S. cerevisiae* H1246, which is devoid of neutral lipid synthesis, was shown (Chi et al., 2014; Rani et al., 2010, 2012).

In the present study, the same experiments were done with AtDGAT3. Expression of AtDGAT3 in *S. cerevisiae* H1246 did not result in any detectable TAG formation (Appendix 12). In contrast, expression of AtDGAT3 in the corresponding wild type *S. cerevisiae* W303 indicated elevated TAG levels in expression cultures, as concluded from a TLC analysis (Appendix 13). A possible explanation for this observation might be a mandatory interaction of AtDGAT3 with other enzymes involved in TAG biosynthesis. Since in the quadruple knockout mutant *S. cerevisiae* H1246, genes for the synthesis of neutral lipids are deleted, necessary interaction partners of AtDGAT3 might be lacking. On the other hand, a mandatory interaction of AtDGAT3 with a certain structure might not be restricted to other enzymes. In this respect, an interaction of AtDGAT3 with LDs may be essential for activity. In *S. cerevisiae* H1246, LDs are not present due to the lack of neutral lipids. In contrast, LDs as well as neutral lipid synthesising enzymes as potential interaction partners of AtDGAT3 are present in the W303 wild type yeast, potentially allowing AtDGAT3 to be active. This feature would be unique for AtDGAT3, since other DGAT3 were shown to mediate TAG synthesis also upon expression in *S. cerevisiae* H1246. However, no experiments addressing the presence of AtDGAT3 in yeast by immunodetection have been performed. Hence, a missing activity caused by the absence of the AtDGAT3 protein can't be ruled out.

In case of DGAT2 from *Arabidopsis* (AtDGAT2), the functionality of the enzyme has been questioned for a long time. Expression of AtDGAT2 in yeast did not yield TAG synthesis as well. Furthermore, AtDGAT2 was not able to complement the phenotype of AtDGAT1 knockout plants, and a knockout of AtDGAT2 had no negative effect on the TAG content of the seeds of respective plants (Zhang et al., 2009). Activity of AtDGAT2 was finally shown upon transient expression in *N. benthamiana* (Zhou et al., 2013). Shortly thereafter, activity in yeast was demonstrated by expression of a codon optimised version of AtDGAT2. Thus, one possibility for the missing activity of AtDGAT3 in yeast could be due to an unfavourable codon usage. To rule out this possibility, the codon usage of AtDGAT3 in *S. cerevisiae* was analysed using the graphical codon usage analyser online tool (Fuhrmann et al., 2004) (Appendix 11). None of the determined codon usage frequency values was below 20 %, making a successful translation of AtDGAT3 in *S. cerevisiae* likely.

*In vitro* activity of AtDGAT3 was not detectable in the course of this study (Appendix 16). Since increasing amounts of DAG caused a decreasing activity in *in vitro* assays with DGAT3 from peanut (Saha et al., 2006), AtDGAT3 *in vitro* assays were performed with different DAG concentrations. However, none of the utilised conditions led to detectable AtDGAT3-mediated TAG formation. However, unfavourable assay conditions might still be a possible reason for the lack of activity, since optimisation of parameters like pH, salt content etc. has not been addressed.

BLAST analyses (Altschul et al., 1997) indicate the presence of a "thioredoxin (TRX)-like [2Fe-2S] ferredoxin-domain" within the sequences of both AtDGAT3 (Appendix 10) and AhDGAT3, but not in the sequences of AtDCR and RgDGAT3 (data not shown). Appropriately, a brown colour and an absorption spectrum which might indicate the presence of a [2Fe2S]-cluster within the protein were observed during analyses of purified AtDGAT3. In contrast, DGAT3 from *A. hypogaea* (AhDGAT3) was not reported to appear brownish after purification (Saha et al., 2006). The presence of a metal cluster within an acyltransferase hasn't been reported before, and there is no obvious need for a metal cluster during the DGAT reaction.

Interestingly, AhDGAT3 and AtDGAT3, both carrying a predicted [2Fe2S]-cluster binding domain, were isolated from cotyledons. While AhDGAT3 was isolated from developing cotyledons from immature seeds, AtDGAT3 was isolated from growing seedlings five days after imbibition (Hernandez

## DISCUSSION

et al., 2012; Saha et al., 2006). A direct participation of the cluster in the catalytic mechanism of AtDGAT3 seems unlikely. Instead, it may have regulatory purposes. As such, the cluster's redox state might activate or inactivate ATDGAT3. Suggesting that the enzyme is active with an intact, reduced cluster, it would be reliant on reducing conditions in order to be active. Reducing conditions may especially occur during the first days of growth of a seedling. Seedlings are reliant on the catabolism of cotyledon-derived TAG through  $\beta$ -oxidation. If  $\beta$ -oxidation would be impaired in a highly reducing environment due to environmental stresses like oxygen deficiency or likewise, free fatty acids, originating from TAG breakdown, might accumulate. Free fatty acids are toxic for the cell due to their detergent character and their ability to alter membrane properties (Sikkema et al., 1995), but can be detoxified by esterification to CoA and subsequent channelling into TAG. A redox state sensitive DGAT might thus represent a safety mechanism in order to prevent the accumulation of free fatty acids by channelling them into TAG under highly reducing conditions or oxygen deficiency. Under conditions of sufficient oxygen, TAG formation is counterproductive for the seedlings growth and has to be prevented. Thus, redox state dependent regulation of the DGAT3 might be an efficient mechanism in order to control DGAT3-mediated TAG formation.

### 6.1.8 Interaction of TF and AbWSD1

In order to obtain pure AbWSD1 protein for crystallisation studies, extensive attempts of separating 6xHis-TF and AbWSD1 after protease-mediated cleavage of the fusion protein were conducted. Despite of the numerous different techniques which were applied, no satisfactory separation of 6xHis-TF and AbWSD1 was achieved. Calculation of the isoelectric point (IP) values of both proteins according to Bjellqvist et al. (1993) yielded values of 5.05 for 6xHis-TF and 9.17 for AbWSD1, respectively. According to these values, both proteins should have a different surface charge. Thus, chromatographic separation of both proteins via ion exchange chromatography should be possible. However, computed IP values are calculated on the basis of the total amino acid content of a protein. Thus, the calculation does not take into account that amino acids which are buried inside the proteins core have no impact on the actual IP of the protein. As a result, respective values might give a first hint of a proteins surface charge, but might also be misleading. Similar surface charges of AbWSD1 and 6xHis-TF might explain the difficulties faced upon separation of both enzymes via ion exchange chromatography (Figure 26).

Alternatively, separation of both proteins via affinity chromatography should be possible. Since it was also not possible to separate both proteins by affinity chromatography (Figure 26), it is reasonable to assume that both proteins tightly interact over various non-covalent bonds and form a stable quaternary structure. Such an interaction could be unspecific, for example due to a partial misfolding in one of the peptides. Hydrophobic parts of the respective protein might be present on its surface and promote aggregate formation. On the other hand, interaction of 6xHis-TF and AbWSD1 may result from a specific interaction of both proteins. As outlined above, TF is a member of the bacterial cell's chaperone machinery and thus capable of binding hydrophobic patches of other proteins. TF is reported to exhibit also a weak holdase activity when it is not bound to the ribosome (Saio et al., 2014). Apparently, this activity is strong enough to support correct folding of misfolded proteins by transient binding to unfolded parts of the peptide chain (Saio et al., 2014). The latter might explain the severe difficulties experienced during attempts to separate 6xHis-TF and AbWSD1. TF may bind to exposed hydrophobic parts of AbWSD1, thereby preventing separation of the two molecules by chromatographic methods (Figure 27 B).

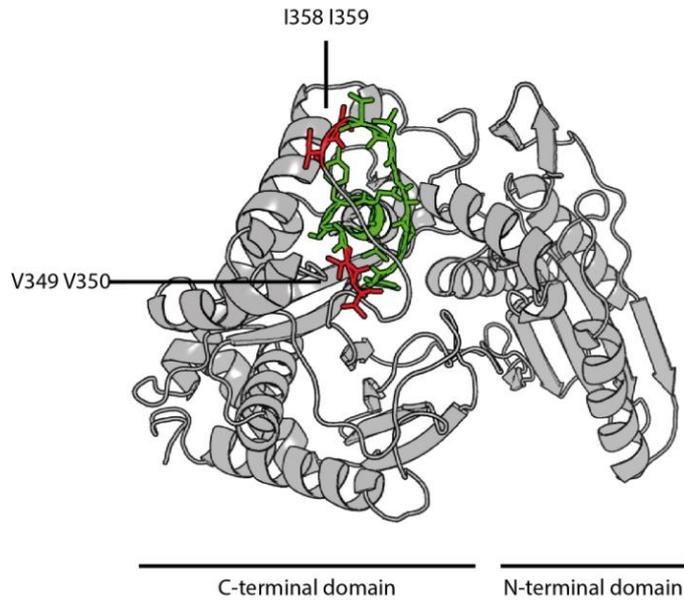
## DISCUSSION

*In silico* analyses of the AbWSD1 amino acid sequence in respect to its hydrophobicity resulted in the prediction of a hydrophobic stretch within the sequence. However, the probability of this prediction was rather low. Within the modelled structure of AbWSD1, the corresponding residues are found to be exposed on the structures surface in the form of a hydrophobic patch (Figure 49, marked in green). Thus, TF may specifically bind to this patch, resulting in a complex of AbWSD1 and 6xHis-TF. This theory can be further strengthened by the fact that purification of the sole AbWSD1 protein without the TF results in precipitation of the protein upon elution from an IMAC column. Since the hydrophobic patch on the protein's surface cannot be shielded from interaction with other AbWSD1 monomers under these conditions, AbWSD1 monomers aggregate and precipitate. This phenomenon was avoided upon exchanging two particularly exposed isoleucine residues to threonines. Purification of the respective AbWSD1 I358S I359S variant from cell lysates of *E. coli* expression cultures did not result in precipitation of the protein. Taken together, it is very likely that the TF plays an important part in the purification process of the 6xHis-TF-AbWSD1 fusion protein by shielding an aggregation prone, hydrophobic part of AbWSD1 (Figure 27 B). The physiological role of such a hydrophobic patch on the surface of AbWSD1 might be linked to a membrane binding of the enzyme. This theory can be supported by the fact that the majority of AbWSD1 activity is associated with the membrane fraction of its natural host *A. baylyi* (Stöveken et al., 2005), indicating membrane binding *in vivo*.

It was shown before, that the WS/DGAT Ma2 from *Alcanivorax borkumensis*, which is highly similar to AbWSD1, is cleaved between its N- and C-terminal domains upon trypsin treatment, resulting in two fragments of 20 kDa and 28 kDa, respectively. The authors of the respective study proposed a partially unordered linker between the two domains to be the main target for the trypsin on the basis of the enzyme's modelled structure (Villa et al., 2013). Interestingly, similar fragments were observed upon SDS-PAGE analyses of 6xHis-AbWSD1 purification fractions (Figure 28). This might indicate protease-mediated cleavage of the protein during the purification procedure or already before. In case of 6xHis-TF-AbWSD1 purifications, respective bands were only faintly visible (Figure 25). The hydrophobic patch mentioned above and the proposed linker between the N- and C-terminal domains of AbWSD1 are predicted to be located on the same site of the protein (data not shown). Ergo, it might well be that the TF does not only shield the hydrophobic patch on the surface of AbWSD1 from interaction with other proteins, but also the linker region between both domains, thereby preventing protease-mediated cleavage of the linker.

Interestingly, the modelled structure of Ma2 does not contain the hydrophobic patch which is present in the model of AbWSD1 (data not shown/Figure 49). Fittingly, the authors describe an easy and straight forward purification for the enzyme without aggregation (Villa et al., 2013). Although only based on models, these findings allow additional hints towards a main role of I358 and I359 in the aggregation pattern of the AbWSD1 enzyme.

## DISCUSSION



**Figure 49: Structure model of AbWSD1.** Model was generated by the Phyre server (Kelley and Sternberg, 2009) according to a surfactin synthetase subunit 3 (PDB accession code 2VSQ) (Tanovic et al., 2008). The C-terminal domain of the enzyme is visible on the left side of the picture, while the N-terminal domain is shown on the right side. A patch predicted to be hydrophobic (Figure 22 A) is shown in green, while V349 and V350 as well as I358 and I359 are shown in red.

Despite of its interaction with the ribosome, TF is also present in a soluble form in the cytosol. Here, it is thought to alternate between a monomeric and a dimeric state (Patzelt et al., 2002). This fact may explain the observed peak upon initial SEC analyses of 6xHis-TF-AbWSD1. There, a peak corresponding to approximately 130 kDa was observed. The respective elution fractions contained protein monomers of approximately 60 kDa, as concluded from SDS-PAGE analyses (Figure 25 A). The 6xHisTF-AbWSD1 may thus be cleaved in solution, resulting in the presence of the sole, dimeric TF. This would be surprising, since separation of AbWSD1 and TF was not possible *in vitro*. Also, no band which could represent the AbWSD1 part of the cleaved fusion protein was detected during SEC (Figure 25). On the one hand, this may be reasoned by a rapid degradation of the single AbWSD1. On the other hand, the accumulation of TF might be a result of a translational stop of the 6xHis-TF-AbWSD1 fusion protein after the 6xHis-TF part due to problems with correct protein folding.

## 6.2 Substrate specificities of MmAWAT2

MmAWAT2 was shown to exhibit unique substrate specificities towards the acyl chain acceptors fatty alcohol and DAG in comparison to MmDGAT2 in the course of this study. It clearly favours acylation of fatty alcohols over acylation of DAG (Figure 32). In respect to the acyl chains which are esterified to these fatty alcohols, MmAWAT2 shows a preference for unsaturated over saturated and C16 over C18 acyl chains (Figure 34). In order to elucidate the structural basis of this specificity, comparative studies between MmAWAT2 and MmDGAT2 were carried out in combination with the construction of domain swap variants of the two enzymes and single amino acid exchange variants of MmAWAT2.

### 6.2.1 Altered WE/TAG ratios in MmAWAT2 variants

Some of the generated MmAWAT2 variants showed an altered affinity in respect to the acyl acceptors DAG and fatty alcohol, as concluded from ratios of WE/TAG production. In all cases (wild type MmAWAT2 as well as variants), feeding of 16:0-OH to the cultures resulted in approximately twice as high WE/TAG ratios in comparison to respective values for feeding of 18:1-OH. Hence, feeding of 16:0-OH leads to the production of relatively more WEs and/or less TAGs in comparison to feeding of 18:1-OH. This phenomenon is unlikely to be caused by the physical properties of the alcohols. Due to its fully saturated acyl chain, 16:0-OH is less soluble in aqueous milieus in comparison to 18:1-OH. Assuming an equal uptake of both alcohols by the cell, 18:1-OH should be available in higher quantities than 16:0-OH. Under the premise that MmAWAT2 and the respective variants have an equal preference for both fatty alcohols, feeding of 18:1-OH should thus result in higher WE productions than feeding of 16:0-OH. Since the exact opposite is the case, 16:0-OH may be taken up by the cell more efficiently than 18:1-OH. Alternatively, all enzymes might have a higher specificity for 16:0-OH in comparison to 18:1-OH.

Surprisingly, despite of the fact that V2, V5 and MmAWAT2 N36R show changes in respect to general acyl acceptor preferences and acyl chain incorporation specificities, it seems like that the higher specificity for 16:0-OH in these variants is not affected. As described above, all tested constructs showed approximately twice as high WE/TAG ratios upon feeding of 16:0-OH in comparison to the respective values for feeding of 18:1-OH. If one of the variant would exhibit an altered affinity to one of the alcohols, this relationship should change. However, since only two different types of fatty alcohols were fed, it cannot be ruled out that there are changes in the fatty alcohol preferences which were not detected.

Domain swap variants V2 and V5 as well as MmAWAT2 N36R show a relative increase of TAG production (Figure 31, Figure 32 and Figure 42). In contrast, exchange of alanine at position 25 in MmAWAT2 to phenylalanine (MmAWAT2 A25F) seems to have a negative effect on the enzymes' affinity for DAG in case of feeding 16:0-OH. In respective samples, no TAG signals were detected. Consequently, also the MmAWAT2 A25F N36R variant shows less TAG accumulation in comparison to the MmAWAT2 N36R variant. However, the differences between these two variants are not significant (Figure 42). Taken together, exchange of alanine 25 towards phenylalanine in MmAWAT2 may have a negative effect on the enzyme's affinity for DAG upon feeding 16:0-OH.

### 6.2.2 Incorporation of VLC acyl-CoA by MmAWAT2 variants

In case of VLC WE producing MmAWAT2 variants, clear differences in VLC acyl chain incorporation into WEs upon feeding either 16:0-OH or 18:1-OH were observed. Upon feeding of 18:1-OH, only 26:0-CoA was utilised, while upon feeding 16:0-OH, also 20:0-CoA, 22:0-CoA and 24:0-CoA were used for WE synthesis (Figure 37). In contrast, the specificity for C16- and C18-CoAs did not alter with feeding of different alcohols (Figure 34, Figure 41). It was proposed earlier, that the fatty alcohol is the substrate which binds first to the WS upon WE formation (Barney et al., 2012a; Stöveken et al., 2009). Hence, one possible explanation for the altered incorporation of VLC acyl chains could be a differential change of the enzymes' conformation upon binding of different fatty alcohols. That means, that the acyl-CoA binding site of VLC WE producing variants would be capable of binding 20:0-CoA, 22:0-CoA, 24:0-CoA and 26:0-CoA in combination with 16:0-OH, whereas 18:1-OH would result in a conformational change which would only allow binding of 26:0-CoA.

Another explanation for the phenomenon may be an altered substrate availability in the cell upon feeding of different fatty alcohols. On the other hand, analyses of the acyl-CoA pool of *S. cerevisiae* H1246 cultures expressing pYES2/NT upon feeding different alcohols did not reveal drastic differences in total VLC acyl-CoA distributions (Figure 33). However, these measurements are only able to represent the total acyl-CoA pool of the cell. It might thus well be, that different alcohols influence local acyl-CoA sub-pools, resulting in the observed effects.

VLC WE producing enzyme variants also differed from MmAWAT2 in C16 and C18 acyl chain substrate specificities. These specificities were different from respective patterns of MmDGAT2 and MmAWAT2. Although not completely, they best match the respective patterns of the acyl-CoA pool (Figure 34, Figure 41). These observations may indicate a partial loss of substrate specificity of these enzymes. As a consequence, acyl chain incorporation into WEs would reflect the acyl chain distribution in the acyl-CoA pool. However, analyses of the total acyl-CoA pool indicated the presence of monounsaturated VLC acyl-CoAs in the cell. In contrast, these species were not found in the WE fraction of VLC WE producing enzymes. Thus, a complete loss of substrate specificity of these enzymes is unlikely. However, the observed effects represent strong evidence towards a major role of the predicted TM domain of MmAWAT2 in determination of the substrate specificity of the enzyme.

### 6.2.3 Role of a putative neutral lipid binding sequence in MmAWAT2

In MmDGAT2, the first of its two TM domains was already suggested to be involved in substrate binding before. It contains a FLXLXXX consensus sequence which represents a putative NLBS. A deletion (Au-Young and Fielding, 1992) or mutation of the consensus sequence in several enzymes led to decreased activities (Alam et al., 2006; Stone et al., 2006). In the present study, slight effects caused by amino acid exchanges in this part of the MmAWAT2 sequence were observed. Alanine 25 in MmAWAT2 corresponds to phenylalanine 80 in MmDGAT2, which in turn represents the F of the MmDGAT2 FLXLXXX motif (Figure 39). Exchange of the alanine towards phenylalanine in MmAWAT2 caused a slight alteration of the acyl chain incorporation specificity of the corresponding variant, whereas mutation of F80A in MmDGAT2 led to an almost completely abolished enzymatic activity (Stone et al., 2006). Furthermore, threonine 30 and threonine 31 of MmAWAT2 were exchanged in the present study. They correspond to the last two positions of the FLXLXXX motif in MmDGAT2. Exchange of the two threonines towards alanines caused a slight increase of 16:1 acyl chains in WEs

## DISCUSSION

from cultures expressing MmAWAT2 T30A T31A (Figure 41, Appendix 20). In summary, the respective sequence in MmAWAT2 does not comply with the proposed FLXLXXX consensus sequence, as it is found in MmDGAT2, as the neutral lipid binding site. Nevertheless, corresponding amino acid exchange variants of MmAWAT2 showed slight alterations in their substrate specificities. Hence, it is debatable if integrity of the FLXLXXX sequence is crucial for substrate specificity determination. In contrast to MmAWAT2 and MmDGAT2, the FLXLXXX site in the human cholesteryl ester transfer protein does neither form a possible TM domain nor an  $\alpha$ -helix. Instead, it is present on a  $\beta$ -sheet in the crystal structure of the respective enzyme (Qiu et al., 2007). Hence, the sequence motif is likely not restricted to a certain form of secondary structure.

### 6.2.4 VLC WE production might be explained by altered substrate specificities of MmAWAT2 variants

VLC WE production by MmAWAT2 variants might be explained by a change of substrate specificity of respective enzymes. These changes would result from the combination of the MmDGAT2-derived TM domains or the N36R exchange in the TM domains from MmAWAT2 with the MmAWAT2-derived catalytic centre. In this theory, the two TM domains of MmAWAT2 and MmDGAT2 might act as a kind of substrate acquiring structure, which guards the entrance in front of the active site cavity. In case of MmDGAT2, the TM domains would exhibit a higher affinity towards DAG in comparison to the respective sequence in MmAWAT2. In MmAWAT2, the sequence would in contrast have a higher affinity towards fatty alcohols than the respective sequence in MmDGAT2. According to this theory, swapping of the MmDGAT2 TM domains in front of the active site of MmAWAT2 or the N36R exchange in MmAWAT2 would result in an increased acquisition of DAG molecules. Due to unfavourable properties of the MmAWAT2 active site pocket in respect to DAG acylation, the acquired DAG could not be utilised as good as in case of MmDGAT2. The modelled structures of MmAWAT2 $\Delta$ N and MmDGAT2 $\Delta$ N partly support this theory. MmAWAT2 $\Delta$ N shows a smaller cavity in direct proximity to the designated active HPHG motif in comparison to MmDGAT2 $\Delta$ N (Figure 47). Hence, it is conceivable that proper binding of DAG in the rather small active site pocket is hindered. However, the predicted structures of MmAWAT2 $\Delta$ N and MmDGAT2 $\Delta$ N are not modelled with the highest confidence, making this theory highly speculative. In summary, swapping of the TM domains of MmDGAT2 to the active site of MmAWAT2 or the MmAWAT2 N36R mutation might result in an easier access of DAG towards the MmAWAT2 active site, resulting in a relative increase of TAG productions in comparison to WEs. Nevertheless, the active site would not be as suitable for TAG production as the respective structure in MmDGAT2, explaining the differences between domain swap variants and MmDGAT2.

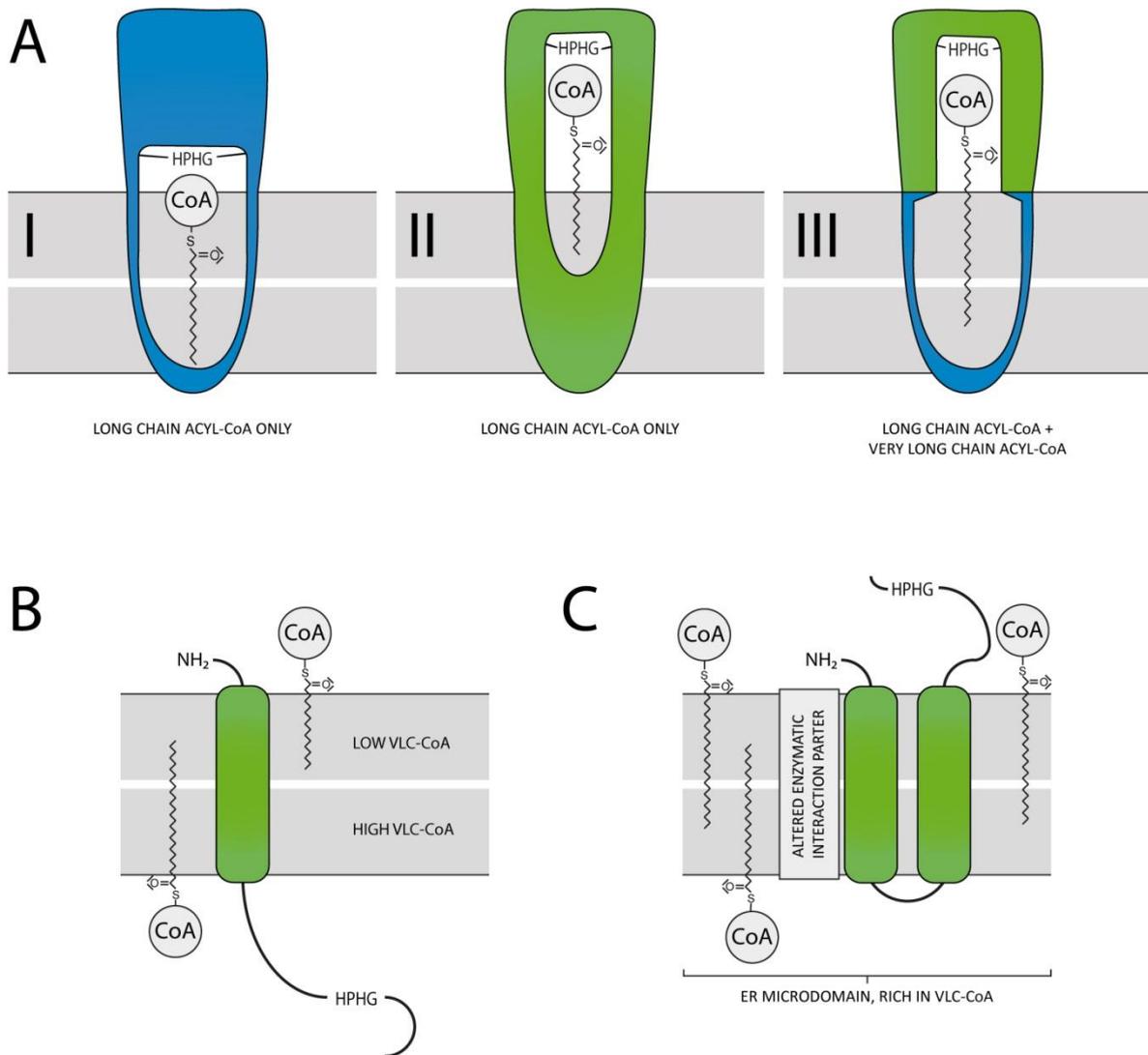
In respect to VLC WE production, this model works similar. Swapping of the TM domain as well as replacement of asparagine 36 with arginine in MmAWAT2 may result in a loss of the guarding activity of the TM domains of MmAWAT2. VLC acyl-CoAs would not be discriminated any longer from the active site, and are thus utilised for WE synthesis (Figure 50 A). However, the loss of the guarding activity against VLC acyl chains in case of V2 and V5 would have to be special for the combination of MmDGAT2-derived TM domains and MmAWAT2-derived catalytic centre, since MmDGAT2-derived TM domains alone are not able to mediate VLC WE synthesis in MmDGAT2 (Figure 50 A). Yet, it would be interesting to know whether MmDGAT2 and MmAWAT2 would have the same specificity without their respective TM domains.

### 6.2.5 VLC WE production might be explained by access of MmAWAT2 variants to new substrate pools

Besides the theory concerning alterations in substrate specificities of VLC WE producing enzymes, other explanations for the observed effects are possible. It was shown earlier that both, MmAWAT2 and MmDGAT2, are integral membrane enzymes and are ER as well as LD associated (Cheng, 2004a; Heilmann et al., 2012; McFie et al., 2014; Stone et al., 2006). In case of MmDGAT2, the TM topology was already determined. MmDGAT2 possesses two TM domains which act as a kind of membrane anchor. Its active site is localised in the cytosol. The TM prediction for MmAWAT2 indicates a similar topology. Furthermore, this particular TM architecture is predicted to be widely conserved among vertebrate AWAT2 and DGAT2 sequences. Thus, it is reasonable to assume that MmAWAT2 adopts a similar TM topology like MmDGAT2, with a cytosol localised active site. Domain swapping and amino acid exchange in VLC WE producing MmAWAT2 variants might result in impaired formation of one TM domain. As a consequence, the active site of these enzymes would be located on the opposite side of the membrane in comparison to the active site of wild type MmAWAT2. Consequently, the respective MmAWAT2 variants would have access to a different substrate pool in comparison to the wild type MmAWAT2 (Figure 50 B). This different substrate pool might harbour more VLC acyl chain substrates in comparison to the substrate pool wild type MmAWAT2 has access to. In order to fathom this theory, mapping of the TM topologies of MmAWAT2 and VLC WE producing enzymes would be of great help. One argument against an altered membrane topology of MmAWAT2 upon introducing amino acid exchanges in its TM domains can be taken from a study of Stone et al. (2006). The group showed that mutations in the first TM domain of MmDGAT2 resulted in an altered enzymatic activity, but in an unaltered topological orientation of the enzyme.

A similar explanation for the observed facts could be the interaction of VLC WE producing MmAWAT2 variants with protein partners which wild type MmAWAT2 does not interact with (Figure 50 C). It has been shown before, that proteins involved in lipid metabolism can interact with each other via a defined stretch of a TM domain (Zhang et al., 2014). It is furthermore known, that TAG synthesising enzymes can cluster in defined ER micro domains (Shockey, 2006) and form protein complexes (Lehner and Kuksis, 1995; Wilfling et al., 2013). Hence, changes in the part of MmAWAT2 which is spanning the ER membrane could on the one side result in prevention of interaction with specific enzymes MmAWAT2 interacts with. On the other side, it could also result in the formation of new interactions with enzymes MmAWAT2 does not interact with. As a quintessence of both, these effects would again result in access to different substrate pools of the VLC WE producing variants of MmAWAT2.

## DISCUSSION



**Figure 50: Models explaining the observed phenomena of VLC WE producing MmAWAT2 variants. A)** A differential substrate specificity of VLC WE producing variants might result from an altered substrate binding pocket and accompanied with that, the ability to bind VLC acyl-CoAs. **I)** Situation in MmDGAT2. Active site and TM domains favour DAG binding. Due to steric reasons, long chain (LC) acyl CoAs, but not VLC acyl-CoAs can bind. **II)** Situation in MmAWAT2. Active site cavity and TM domains favour binding of fatty alcohols. Again, VLC acyl-CoA molecules cannot bind due to steric reasons. **III)** VLC WE producing variants do not discriminate VLC acyl-CoA anymore. TM domains enable better binding of DAG, resulting in an altered WE/TAG ratio in comparison to MmAWAT2. **B)** Distribution of lipids in both leaflets of the membrane might be uneven. Mutations and domain swapping may cause an altered membrane topology of the enzyme, resulting in an altered orientation of the catalytic centre of the enzyme. This differential topology of the enzymes might then result in access to different substrates in comparison to the wild type enzyme. **C)** VLC WE producing variants might interact with different enzymatic partners in the membrane than MmAWAT2 and MmDGAT2 do. This may result in an access of the variants to different substrate pools containing VLC acyl-CoAs.

### 6.2.6 Truncation of MmAWAT2

It has been shown earlier, that a truncated version of MmDGAT2 is active without its TM domains (McFie et al., 2011) upon expression in human cell culture. In the present study, hints towards similar results for a truncated version of MmAWAT2 were obtained. Activity of a truncated version of MmAWAT2 would be in line with the theory, that the presence of the TM domains is able to determine substrate specificity. In contrast, absence of the TM domains would not abolish the enzyme activity, but might result in a loss of substrate specificity. To test this hypothesis, it would be of great value to compare substrate specificities of either MmAWAT2 or MmDGAT2 with and without its TM domains.

Interestingly, the truncated version of MmAWAT2, missing the N-terminal TM domains, did not show activity upon expression in *S. cerevisiae* H1246 or the corresponding wild type strain W303. In a previous study, the likewise truncated version of MmDGAT2 did not localise to the ER anymore, but showed a LD localisation upon oleate treatment of the cells (McFie et al., 2011). McFie (2014) suggested that the LD localisation of MmDGAT2 is mediated by a helical structure in MmDGAT2, which also harbours the active site. However, in the modelled enzymatic structures of MmAWAT2 $\Delta$ N and MmDGAT2 $\Delta$ N, the active site motif is neither predicted to be located in a helical structure, nor to be exposed to the surface of the protein. In order to test a mandatory presence of LD for MmAWAT2 $\Delta$ N activity in yeast, the truncated version was expressed in the *S. cerevisiae* wild type strain W303 supplemented with 18:1-COOH and 18:1-OH. Supplementation of 18:1-COOH should lead to formation of LD. MmAWAT2 $\Delta$ N might localise to these LD and utilise 18:1-OH for WE production. However, no elevated levels of WEs or TAGs synthesis upon TLC analyses of lipid extracts from respective cultures were detected (data not shown).

Another possible hypothesis is that MmAWAT2 $\Delta$ N may be expressed in *S. cerevisiae* H1246 in a soluble, cytosolic form, but does not reveal its activity due to restricted access towards its substrates, which are not present in the yeast's cytosol. Without its N-terminus, MmAWAT2 is predicted to contain no further TM domains and is thus likely present in a soluble state. Hence, upon localisation of the protein in the cytosol, membrane-bound substrates might not be accessible for the protein. These problems could be circumvented by the *in vitro* assays with the partly purified MmAWAT2 $\Delta$ N protein. However, no immunodetection experiments were done in order to elucidate the presence of MmAWAT2 $\Delta$ N in yeast or *E. coli* cell extract. Hence, another possibility for missing activity in yeast might be problems with translation or transcription of the protein.

## 7 SUMMARY

Wax esters (WEs) are esters of fatty acids and fatty alcohols. They can cover a broad range of physical properties, which makes them especially interesting for industrial applications, including additives in cosmetics and high class lubricants. Biosynthesis of WEs is a widespread feature in nature and is carried out by two essential classes of enzymes. One of them is the class of fatty acyl reductases (FARs), which reduce acyl-CoAs or acyl-acyl carrier proteins (acyl-ACPs) to the corresponding fatty alcohols. The second one is the class of wax synthases (WSs), which esterify fatty alcohols with acyl CoAs, yielding WEs. The majority of FARs and WSs described to date exhibit a broad substrate range, resulting in wax blends of heterogeneous compositions. With respect to a commercial production of WEs in genetically modified plants, a defined WE blend is desired rather than a mixture of WEs. Enzymes with improved substrate specificities, tailored for the production of individual WE blends, are one way to overcome these issues. However, the required knowledge about structure-function relationships in FARs and WSs for the construction of respective enzymes is lacking to date. The aim of the present thesis was thus to elucidate structural determinants of substrate specificity in FARs and WSs.

To date, crystal structures are neither available from FARs nor from WSs. Hence, four FARs (DmFAR1 from *Drosophila melanogaster*, MaFAR1 from *Marinobacter aquaeolei*, MmFAR1 from *Mus musculus* and TcFAR1 from *Tribolium castaneum*), two WSs (AbWSD1 from *Acinetobacter baylyi* and MmAWAT2 from *Mus musculus*) and a soluble diacylglycerol O-acyltransferase (AtDGAT3 from *Arabidopsis thaliana*) were studied in order to obtain structural insights. The activity of DmFAR1 was first described in the present work. It produces tetracosanol and hexacosanol upon expression in yeast, while expression in *E. coli* yielded tetradecanol, hexadecanol, hexadecenol and octadecenol. DmFAR1, MaFAR1, MmFAR1, AbWSD1, MmAWAT2 and AtDGAT3 were expressed in and purified from *E. coli*. Although this resulted in aggregated proteins in most cases, AbWSD1 was obtained in a quaternary structure corresponding to a trimer, when expressed as a fusion protein with the bacterial trigger factor. The respective fusion protein reproducibly formed crystals, which diffracted to 2.1 Å. Attempts to soak the crystals with iodine led to disruption of the crystals. Hence, the respective selenoprotein was produced and applied to crystallisation screens, which were ongoing at the end of this thesis.

In order to elucidate substrate specificity determining structures in MmAWAT2, comparative studies of MmAWAT2 and MmDGAT2, which share ~ 70 % homology, were carried out. Both enzymes are acyltransferases and are capable of synthesising both, TAGs and WEs. Furthermore, both enzymes show distinct substrate specificities with respect to WE synthesis. Two predicted, neighboured hairpin forming transmembrane (TM) domains were identified to have an influence on the substrate specificity of MmAWAT2. Chimeric enzyme variants of MmAWAT2 carrying the respective section of the MmDGAT2 sequence showed a severely altered acyl chain incorporation pattern into WEs as compared to both, MmAWAT2 and MmDGAT2. Furthermore, respective variants showed an altered ratio of produced WEs and TAGs. This phenotype was also exhibited by the MmAWAT2 single amino acid exchange variant N36R, carrying a mutation in the part of the sequence which encodes the two predicted TM domains. Thus, the predicted TM domains of MmAWAT2 seem to have a role in substrate specificity determination of the enzyme.

This work provides further insights into structure function relationships concerning substrate specificity in DGAT2-type acyltransferases. Furthermore, the successful crystallisation of a WS might pave the way for an extensive comprehension of this class of enzymes.

## 8 OUTLOOK AND PERSPECTIVE

In order to obtain further insights in structure-function relationships of FARs and WSs, several experimental lines of this work could be continued.

### DmFAR1

In the present study it was shown that DmFAR1 exhibits severely different product blends upon expression in either *E. coli* or *S. cerevisiae*. In order to determine the substrate specificity of DmFAR1 independent of an expression system, *in vitro* assays on the basis of the purified enzyme may be done. Respective experiments would likely solve the question regarding the cofactor preference of DmFAR1 (NADH, NADPH or both). To test the influence of R72 on cofactor binding, single amino acid exchange variants of DmFAR1 could be generated in analogy to Yin et al. (2014). In order to verify the molecular nature of the protein species obtained during IMAC purifications of 6xHis-DmFAR1 and DmFAR1-6xHis, analyses of respective samples via mass spectrometry or immunostaining would be a feasible method. In order to further optimise the purification procedure of DmFAR1, a truncated version of the protein without the predicted TM domain at its C-terminus (Figure 14) could be generated. Since it was shown that the respective part in MmFAR1 mediates the membrane localisation of the enzyme (Heilmann et al., 2012), a truncated version of DmFAR1 might contribute to the solubility of the enzyme. Finally, crystallisation experiments on the basis of DmFAR1 might yield the first structure of a FAR.

### TcFAR1

Expression of TcFAR1 in *S. cerevisiae* did not result in the detection of fatty alcohols. Apart from verification of the presence of the TcFAR1 in yeast by immunodetection, the protein could be expressed in *E. coli* in order to elucidate its activity. Alternatively, expression cultures of TcFAR1 in either *E. coli* or *S. cerevisiae* or respective cell lysates could be supplemented with substrates which are not intrinsically offered by the host in order to elucidate a corresponding substrate specificity of TcFAR1. Respective molecules include VLC acyl chains longer than 26 carbon atoms or polyunsaturated acyl species.

### MmFAR1

Expression of MmFAR1 in *E. coli* did not result in detectable production of the enzyme. Since two predicted TM domains at the C-terminus of MmFAR1 might lead to misfolding and protein degradation, expression of a truncated version of the enzyme might be tried instead.

### MaFAR1

*In vitro* activity of MaFAR1 was dependent on high amounts of NADPH. It was speculated that NADPH might have an influence on the aggregated state of purified MaFAR1. In this respect, aggregation patterns of MaFAR1 in relation to NADPH concentrations could be examined.

### AtDGAT3

Activity of AtDGAT3 was neither detectable *in vivo* nor *in vitro*. To verify the presence of the AtDGAT3 protein in yeast, immunostaining experiments should be conducted. Lacking activity of the purified enzyme *in vitro* might be due to unfavourable assay conditions. Hence, attempts to optimise the assay could be carried out. The spectral properties of AtDGAT3 indicated the presence of a metal cluster within the protein. The exact identity of the cluster could be elucidated by x-ray absorption spectroscopy on the basis of purified enzyme. In order further explore the theory of a mandatory interaction of AtDGAT3 with other enzymes for activity, pull down assays with AtDGAT3 and *A. thaliana* cell lysate could be done.

### AbWSD1

In order to solve the crystal structure of AbWSD1, highest efforts should be directed to the generation of AbWSD1 selenoprotein crystals. In case of difficulties with crystallisation, the AbWSD1 I358S I359S variant is a promising candidate for further crystallisation studies. Additional experiments could be conducted in order to soak AbWSD1 crystals with its two substrates acyl-CoAs and fatty alcohols. Alternatively, co-crystallisation of AbWSD1 with substrates could be done. A structure containing bound substrates might help to identify crucial residues for substrate specificity determination. In a further step, mutational studies on respective residues might be able to alter the substrate specificity of AbWSD1.

Apart from that, it would be highly interesting to compare the structures of the wild type AbWSD1 and the AbWSD1 I358S I359S variant in order to verify the position as well as the potential role of I358 and I359 in the quaternary structure of AbWSD1. In this respect, the AbWSD1 I358S I359S variant could furthermore be expressed as a fusion protein with the TF. If separation of both proteins after protease cleavage would be possible in this case, this would be an additional hint towards a major role of I358 and I359 in the quaternary structure determination of 6xHis-TF-AbWSD1.

The modelled structure of AbWSD1 indicates the presence of an unordered linker between its N- and C-terminal domains. This might be a target for protease-mediated cleavage in 6xHis-AbWSD1, whereas 6xHis-TF-AbWSD1 seemed to be less susceptible in this respect. To elucidate a potential protective role of the TF for this part of AbWSD1, a comparative protease treatment of AbWSD1 and the 6xHis-TF-AbWSD1 fusion protein might provide further insights into the quaternary structure of a potential 6xHis-TF-AbWSD1 complex.

### MmAWAT2

To validate the theories of VLC WE production of MmAWAT2 variants due to a differential access of the respective enzymes towards VLC acyl chain substrates, *in vitro* experiments on the basis of purified MmAWAT2 and the according variants could be performed. If the availability of VLC acyl-CoAs as substrates would be the only bottleneck for VLC WE production, MmAWAT2 and its variants would synthesise equal amounts of VLC WEs under non-limiting substrate conditions. In contrast, a real change in substrate specificity due to the enzymatic structures of MmAWAT2 and the respective variants would result in an absence of VLC WEs in case of MmAWAT2, while the variants would produce VLC WEs in amounts comparable to respective *in vivo* results. In an ideal case, such experiments would be done with purified, membrane free protein in order to ensure an equal access of the substrates to the enzymes. However, production of full length MmAWAT2 in *E. coli* for

## OUTLOOK AND PERSPECTIVE

respective experiments was not successful during the present study. Instead, utilization of microsomal preparations of yeast cultures for such experiments may work likewise.

## 9 LITERATURE

- Aaij, C., and Borst, P. (1972). The gel electrophoresis of DNA. *BBA-GENE STRUCT EXPR* 269, 192–200.
- Aarts, M.G., Hodge, R., Kalantidis, K., Florack, D., Wilson, Z.A., Mulligan, B.J., Stiekema, W.J., Scott, R., and Pereira, A. (1997). The *Arabidopsis* male sterility 2 protein shares similarity with reductases in elongation/condensation complexes. *PLANT J* 12, 615–623.
- Alam, M., Gilham, D., Vance, D.E., and Lehner, R. (2006). Mutation of F417 but not of L418 or L420 in the lipid binding domain decreases the activity of triacylglycerol hydrolase. *J LIPID RES* 47, 375–383.
- Aleklett, K., Höök, M., Jakobsson, K., Lardelli, M., Snowden, S., and Söderbergh, B. (2010). The peak of the oil age – analyzing the world oil production reference scenario in world energy outlook 2008. *ENERG POLICY* 38, 1398–1414.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *NUCLEIC ACIDS RES* 25, 3389–3402.
- Altschul, S.F., Wootton, J.C., Gertz, E.M., Agarwala, R., Morgulis, A., Schäffer, A.A., and Yu, Y.-K. (2005). Protein database searches using compositionally adjusted substitution matrices. *FEBS J* 272, 5101–5109.
- Alvarez, H.M., and Steinbüchel, A. (2002). Triacylglycerols in prokaryotic microorganisms. *APPL MICROBIOL BIOT* 60, 367–376.
- American Fuel & Petrochemical Manufacturers (2014). Wax facts. Retrieved October 5, 2014, from <http://www.afpm.org/wax-Facts/>.
- American Oil Chemists' Society (1988). Seventh international conference on jojoba and its uses: proceedings (Champaign, Illinois: American Oil Chemists' Society).
- Antony, B., Fujii, T., Moto, K., Matsumoto, S., Fukuzawa, M., Nakano, R., Tatsuki, S., and Ishikawa, Y. (2009). Pheromone-gland-specific fatty-acyl reductase in the adzuki bean borer, *Ostrinia scapularis* (Lepidoptera: Crambidae). *INSECT BIOCHEM MOLEC* 39, 90–95.
- Ash, G.J., Albiston, A., and Cother, E.J. (2005). Aspects of jojoba agronomy and management. *ADV AGRON*, pp. 409–437.
- Aslan, S., Sun, C., Leonova, S., Dutta, P., Dörmann, P., Domergue, F., Stymne, S., and Hofvander, P. (2014). Wax esters of different compositions produced via engineering of leaf chloroplast metabolism in *Nicotiana benthamiana*. *METAB ENG*, DOI: 10.1016/j.ymben.2014.07.001.
- Ausubel F.M., Brent, R.E., Kingston, D.D., Seidmann, J.R., Smith, J.A., Struhl, K. (1993). *Current Protocols in Molecular Biology*, New York: Green Publishing Associates and John Wiley and Sons Inc.
- Bagge, L.E., Koopman, H.N., Rommel, S.A., McLellan, W.A., and Pabst, D.A. (2012). Lipid class and depth-specific thermal properties in the blubber of the short-finned pilot whale and the pygmy sperm whale. *J EXP BIOL* 215, 4330–4339.
- Barney, B.M., Wahlen, B.D., Garner, E., Wei, J., and Seefeldt, L.C. (2012a). Differences in substrate specificities of five bacterial wax ester synthases. *APPL ENVIRON MICROB* 78, 5734–5745.

## LITERATURE

- Barney, B.M., Mann, R.L., and Ohlert, J.M. (2012b). Identification of a residue affecting fatty alcohol selectivity in wax ester synthase. *APPL ENVIRON MICROB* 79, 396–399.
- Bart, J.C.J., Gucciardi, E., and Cavalloro, S. (2013). *Biolubricants: science and technology* (Woodhead Publishing Limited).
- Bassam, N.E. (1997). *Energy plant species: their use and impact on environment and development* (London: James & James).
- BBC (2007). BBC NEWS | Business | Will biofuel leave the poor hungry? [News.bbc.co.uk http://news.bbc.co.uk/2/hi/business/7026105.stm](http://news.bbc.co.uk/2/hi/business/7026105.stm).
- Bergendahl, V., Linne, U., and Marahiel, M.A. (2002). Mutational analysis of the C-domain in nonribosomal peptide synthesis. *EUR J BIOCHEM* 269, 620–629.
- Bernard, A., and Joubès, J. (2013). *Arabidopsis* cuticular waxes: advances in synthesis, export and regulation. *PROG LIPID RES* 52, 110–129.
- Bianchi, G., Avato, P., and Salamini, F. (1984). Surface waxes from grain, leaves, and husks of maize (*Zea mays* L.). *CEREAL CHEM* 61, 45–47.
- Biester, E.-M., Hellenbrand, J., and Frentzen, M. (2012a). Multifunctional acyltransferases from *Tetrahymena thermophila*. *LIPIDS* 47, 371–381.
- Biester, E.-M., Hellenbrand, J., Gruber, J., Hamberg, M., and Frentzen, M. (2012b). Identification of avian wax synthases. *BMC BIOCHEM* 13, 4.
- Birnboim, H.C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *NUCLEIC ACIDS RES* 7, 1513–1523.
- Bisht, R.P.S., Sivasankaran, G.A., and Bhatia, V.K. (1993). Additive properties of Jojoba oil for lubricating oil formulations. *WEAR* 161, 193–197.
- Bjellqvist, B., Hughes, G.J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J.C., Frutiger, S., and Hochstrasser, D. (1993). The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *ELECTROPHORESIS* 14, 1023–1031.
- Blomquist, G.J., Soliday, C.L., Byers, B.A., Brakke, J.W., and Jackson, L.L. (1972). Cuticular lipids of insects: V. Cuticular wax esters of secondary alcohols from the grasshoppers *Melanoplus packardii* and *Melanoplus sanguinipes*. *LIPIDS* 7, 356–362.
- Boyer, H.W. (1971). DNA restriction and modification mechanisms in bacteria. *ANNU REV MICROBIOL* 25, 153–176.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *ANAL BIOCHEM* 72, 248–254.
- Brändén, C.-I. (1980). Relation between structure and function of  $\alpha/\beta$ -proteins. *Q REV BIOPHYS* 13, 317.
- Bryksin, A., and Matsumura, I. (2010). Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *BIOTECHNIQUES* 48, 463–465.

## LITERATURE

- Buckner, J.S., Pitts-Singer, T.L., Guédot, C., Hagen, M.M., Fatland, C.L., and Kemp, W.P. (2009). Cuticular lipids of female solitary bees, *Osmia lignaria* say and *Megachile rotundata* (f.) (hymenoptera: megachilidae). *COMP BIOCHEM PHYS B* 153, 200–205.
- Buglino, J., Onwueme, K.C., Ferreras, J.A., Quadri, L.E.N., and Lima, C.D. (2004). Crystal structure of PapA5, a phthiocerol dimycocerosyl transferase from *Mycobacterium tuberculosis*. *J BIOL CHEM* 279, 30634–30642.
- Bundesministerium für Wirtschaft und Energie (2014). Das Erneuerbare-Energien-Gesetz 2014. Retrieved October 5, 2014, from <http://www.bmwi.de/BMWi/Redaktion/PDF/Publikationen/das-erneuerbare-energien-gesetz-2014>.
- Burdett, K., Larkins, L.K., Das, A.K., and Hajra, A.K. (1991). Peroxisomal localisation of acyl-coenzyme A reductase (long chain alcohol forming) in guinea pig intestine mucosal cells. *J BIOL CHEM* 266, 12201–12206.
- Carlsson, A.S. (2006). Production of wax esters in crambe: outputs from the EPOBIO project November 2006 (Newbury: CPL).
- Carlsson, A.S., Yilmaz, J.L., Green, A.G., Stymne, S., and Hofvander, P. (2011). Replacing fossil oil with fresh oil - with what and for what? *EUR J LIPID SCI TECH* 113, 812–831.
- Carrier, D.R., Deban, S.M., and Otterstrom, J. (2002). The face that sank the Essex: potential function of the spermaceti organ in aggression. *J EXP BIOL* 205, 1755–1763.
- Chacon, M.G., Fournier, A.E., Tran, F., Dittrich-Domergue, F., Pulsifer, I.P., Domergue, F., and Rowland, O. (2013). Identification of amino acids conferring chain length substrate specificities on fatty alcohol-forming reductases FAR5 and FAR8 from *Arabidopsis thaliana*. *J BIOL CHEM* 288, 30345–30355.
- Chen, W., Yu, X.-H., Zhang, K., Shi, J., De Oliveira, S., Schreiber, L., Shanklin, J., and Zhang, D. (2011a). male sterile2 encodes a plastid-localised fatty acyl carrier protein reductase required for pollen exine development in *Arabidopsis*. *PLANT PHYSIOL* 157, 842–853.
- Chen, X., Snyder, C.L., Truksa, M., Shah, S., and Weselake, R.J. (2011b). sn-glycerol-3-phosphate acyltransferases in plants. *PLANT SIGNAL BEHAV* 6, 1695–1699.
- Cheng, J.B. (2004a). Mammalian wax biosynthesis: II. Expression cloning of wax synthase cDNAs encoding a member of the acyltransferase enzyme family. *J BIOL CHEM* 279, 37798–37807.
- Cheng, J.B. (2004b). Mammalian wax biosynthesis: I. identification of two fatty acyl-coenzyme A reductases with different substrate specificities and tissue distributions. *J BIOL CHEM* 279, 37789–37797.
- Chhabra, A., Haque, A.S., Pal, R.K., Goyal, A., Rai, R., Joshi, S., Panjekar, S., Pasha, S., Sankaranarayanan, R., and Gokhale, R.S. (2012). Nonprocessive [2 + 2]<sub>e</sub>- off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. *P NATL ACAD SCI USA* 109, 5681–5686.
- Chi, X., Hu, R., Zhang, X., Chen, M., Chen, N., Pan, L., Wang, T., Wang, M., Yang, Z., Wang, Q., et al. (2014). Cloning and functional analysis of three diacylglycerol acyltransferase genes from peanut (*Arachis hypogaea* L.). *PLOS ONE* 9, e105834.

## LITERATURE

- Chisti, Y. (2013). Constraints to commercialization of algal fuels. *J BIOTECHNOL* 167, 201–214.
- Clarke, M.R. (1978). Buoyancy control as a function of the spermaceti organ in the sperm whale. *J MAR BIOL ASSOC UK* 58, 27.
- Committee on world food security (2013). Biofuels and food security - HLPE-Report-5. [http://www.fao.org/fileadmin/user\\_upload/hlpe/hlpe\\_documents/HLPE\\_Reports/HLPE-Report-5\\_Biofuels\\_and\\_food\\_security.pdf](http://www.fao.org/fileadmin/user_upload/hlpe/hlpe_documents/HLPE_Reports/HLPE-Report-5_Biofuels_and_food_security.pdf).
- Craig, J.P., and Tomlinson, A. (1997). Importance of the lipid layer in human tear film stability and evaporation. *OPTOMETRY VISION SCI* 74, 8–13.
- De Crécy-Lagard, V., Marlière, P., and Saurin, W. (1995). Multienzymatic non ribosomal peptide biosynthesis: identification of the functional domains catalysing peptide elongation and epimerisation. *CR ACAD SCI III-VIE* 318, 927–936.
- Davis, M.S. (2000). Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in *Escherichia coli*. *J BIOL CHEM* 275, 28593–28598.
- Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesised proteins. *NATURE* 400, 693–696.
- Dieterichs, E.E.F. (1916). A practical treatise on friction, lubrication, fats and oils (Philadelphia, H. C. Baird & co.). <https://archive.org/details/apracticaltreat01dietgoog>
- Ding, B.-J., Hofvander, P., Wang, H.-L., Durrett, T.P., Stymne, S., and Löfstedt, C. (2014). A plant factory for moth pheromone production. *NATURE Communications* 5.
- Doan, T.T.P., Domergue, F., Fournier, A.E., Vishwanath, S.J., Rowland, O., Moreau, P., Wood, C.C., Carlsson, A.S., Hamberg, M., and Hofvander, P. (2012). Biochemical characterisation of a chloroplast localised fatty acid reductase from *Arabidopsis thaliana*. *BBA-MOL CELL BIOL* L 1821, 1244–1255.
- Domergue, F., Vishwanath, S.J., Joubes, J., Ono, J., Lee, J.A., Bourdon, M., Alhattab, R., Lowe, C., Pascal, S., Lessire, R., et al. (2010). Three *Arabidopsis* fatty acyl-coenzyme A reductases, FAR1, FAR4, and FAR5, generate primary fatty alcohols associated with suberin deposition. *PLANT PHYSIOL* 153, 1539–1554.
- Downing, D.T., and Strauss, J.S. (1974). Synthesis and composition of surface lipids of human skin. *J INVEST DERMATOL* 62, 228–244.
- Du, X., Herrfurth, C., Gottlieb, T., Kawelke, S., Feussner, K., Rühling, H., Feussner, I., and Maniak, M. (2014). *Dictyostelium discoideum* DGAT2 can substitute for the essential function of DGAT1 in triglyceride production but not in ether lipid synthesis. *EUKARYOT CELL* 13, 517–526.
- Duan, Y., Zhu, Z., Cai, K., Tan, X., and Lu, X. (2011). *De novo* biosynthesis of biodiesel by *Escherichia coli* in optimised fed-batch cultivation. *PLOS ONE* 6, e20265.
- Duncan, C.C., Yermanos, D.M., Kumamoto, J., and Levesque, C.S. (1974). Rapid ethanolysis procedure for jojoba wax analysis by gas liquid chromatography. *J AM OIL CHEM SOC* 51, 534–536.
- Dyer, J.M., Stymne, S., Green, A.G., and Carlsson, A.S. (2008). High-value oils from plants. *PLANT J* 54, 640–655.

## LITERATURE

- El Kinawy, O. (2004). Comparison between jojoba oil and other vegetable oils as a substitute to petroleum. *ENERG SOURCE* 26, 639–645.
- Elbahloul, Y., and Steinbüchel, A. (2010). Pilot-scale production of fatty acid ethyl esters by an engineered *Escherichia coli* strain harboring the p(microdiesel) plasmid. *APPL ENVIRON MICROB* 76, 4560–4565.
- Elias, N., and O’Brodivich, H. (2006). Clearance of fluid from airspaces of newborns and infants. *NEOREVIEWS* 7, e88–e94.
- Endlein, E., and Peleikis, K.-H. (2011). Natural waxes – properties, compositions and applications. Retrieved October 5, 2014, from [http://www.kahlwax.de/fileadmin/downloads/SOFW\\_Journal\\_4\\_2011.pdf](http://www.kahlwax.de/fileadmin/downloads/SOFW_Journal_4_2011.pdf).
- Ericsson, U.B., Hallberg, B.M., Detitta, G.T., Dekker, N., and Nordlund, P. (2006). Thermofluor-based high-throughput stability optimisation of proteins for structural studies. *ANAL BIOCHEM* 357, 289–298.
- European Union (2006). Studie über die Durchführung der Gap-Massnahmen für Energiepflanzen und den Bioenergiemarkt. Retrieved October 5, 2014, from [http://ec.europa.eu/agriculture/eval/reports/bio\\_energy/sum\\_de.pdf](http://ec.europa.eu/agriculture/eval/reports/bio_energy/sum_de.pdf)
- Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *NATURE* 431, 590–596.
- Forbes, W.G. (1943). *Lubricants and cutting oils for machine tools* (J. Wiley & sons, inc.).
- Förster, T.D., and Woods, H.A. (2013). Mechanisms of tracheal filling in insects. *BIOL REV CAMB PHILOS* 88, 1–14.
- Fuhrmann, M., Hausherr, A., Ferbitz, L., Schödl, T., Heitzer, M., and Hegemann, P. (2004). Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *PLANT MOL BIOL* 55, 869–881.
- Gallagher, I.H. (1971). Occurrence of waxes in *Acinetobacter*. *J GEN MICROBIOL* 68, 245–247.
- Gentry, H.S. (1958). The natural history of jojoba (*Simmondsia chinensis*) and its cultural aspects. *ECON BOT* 12, 261–295.
- Gietz, R.D., and Schiestl, R.H. (2007). High-efficiency yeast transformation using the LiAc/ss carrier DNA/PEG method. *NAT PROTOC* 2, 31–34.
- Gimpel, J.A., Specht, E.A., Georgianna, D.R., and Mayfield, S.P. (2013). Advances in microalgae engineering and synthetic biology applications for biofuel production. *CURR OPIN CHEM BIOL* 17, 489–495.
- Green, +Nastassia (2013). Nail oil with jojoba and vitamin E: Natural Care for Manicure. Retrieved October 5, 2014, from <http://oilypedia.com/nail-oil-with-jojoba-and-vitamin-e-natural-care-for-manicure>
- Greene, R.A., and Foster, E.O. (1933). The liquid wax of seeds of *Simmondsia californica*. *BOT GAZ* 94, 826.

## LITERATURE

- Di Guan, C., Li, P., Riggs, P.D., and Inouye, H. (1988). Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *GENE* 67, 21–30.
- Hagemann, J.W., and Rothfus, J.A. (1979). Oxidative stability of wax esters by thermogravimetric analysis. *J AM OIL CHEM SOC* 56, 629–631.
- Hagström, Å.K., Liénard, M.A., Groot, A.T., Hedenström, E., and Löfstedt, C. (2012). Semi-selective fatty acyl reductases from four heliothine moths influence the specific pheromone composition. *PLOS ONE* 7, e37230.
- Heilmann, M., Iven, T., Ahmann, K., Hornung, E., Stymne, S., and Feussner, I. (2012). Production of wax esters in plant seed oils by oleosomal cotargeting of biosynthetic enzymes. *J LIPID RES* 53, 2153–2161.
- Hellenbrand, J., Biester, E.-M., Gruber, J., Hamberg, M., and Frentzen, M. (2011). Fatty acyl-CoA reductases of birds. *BMC BIOCHEM* 12, 64.
- Hernandez, M.L., Whitehead, L., He, Z., Gazda, V., Gilday, A., Kozhevnikova, E., Vaistij, F.E., Larson, T.R., and Graham, I.A. (2012). A cytosolic acyltransferase contributes to triacylglycerol synthesis in sucrose-rescued *Arabidopsis* seed oil catabolism mutants. *PLANT PHYSIOL* 160, 215–225.
- Herok, G.H., Mudgil, P., and Millar, T.J. (2009). The effect of meibomian lipids and tear proteins on evaporation rate under controlled *in vitro* conditions. *CURR EYE RES* 34, 589–597.
- Hicks, S.C., and Siemer, S.R. (2001). Method of controlling powdery mildew infections of plants using jojoba wax. U.S. Patent US6174920 (B1) issued January 16 2001
- Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *BIOINFORMATICS* 14, 378–379.
- Hofvander, P., Doan, T.T.P., and Hamberg, M. (2011). A prokaryotic acyl-coA reductase performing reduction of fatty acyl-coA to fatty alcohol. *FEBS LETT* 585, 3538–3543.
- Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K., and Pease, L.R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *GENE* 77, 61–68.
- Engineeringtoolbox.com (2014). Liquids - kinematic viscosities. Retrieved October 5, 2014, from [http://www.engineeringtoolbox.com/kinematic-viscosity-d\\_397.html](http://www.engineeringtoolbox.com/kinematic-viscosity-d_397.html)
- Hutchison, C.A., 3rd, Phillips, S., Edgell, M.H., Gillam, S., Jahnke, P., and Smith, M. (1978). Mutagenesis at a specific position in a DNA sequence. *J BIOL CHEM* 253, 6551–6560.
- Ingram, L.O., Conway, T., Clark, D.P., Sewell, G.W., and Preston, J.F. (1987). Genetic engineering of ethanol production in *Escherichia coli*. *APPL ENVIRON MICROB* 53, 2420–2425.
- Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *GENE* 96, 23–28.
- Intergovernmental Panel on Climate Change (1996). Climate change 1995: the science of climate change (Cambridge ; New York: Cambridge University Press).

## LITERATURE

- International Centre for Trade and Sustainable Development (2011). The impact of us biofuel policies on agricultural price levels and volatility. CHINA AGR ECON REV, DOI: 10.1108/17561371211284786.
- Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983). Transformation of intact yeast cells treated with alkali cations. J BACTERIOL 153, 163–168.
- Iven, T., Herrfurth, C., Hornung, E., Heilmann, M., Hofvander, P., Stymne, S., Zhu, L.-H., and Feussner, I. (2013). Wax ester profiling of seed oil by nano-electrospray ionization tandem mass spectrometry. PLANT METHODS 9, 24.
- Japanese Conference on the Biochemistry of lipids (JCBL) (2014). Wax (wwa2301). Retrieved October 5, 2014, from <http://www.lipidbank.jp/cgi-bin/detail.cgi?id=WWA2301#0016>.
- Jaspers, M.H.J., Pflanz, R., Riedel, D., Kawelke, S., Feussner, I., and Schuh, R. (2014). The fatty acyl-CoA reductase waterproof mediates airway clearance in *Drosophila*. DEV BIOL 385, 23–31.
- Käll, L., Krogh, A., and Sonnhammer, E.L.L. (2007). Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. NUCLEIC ACIDS RES 35, W429–432.
- Kallberg, Y., Oppermann, U., and Persson, B. (2010). Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models. FEBS J 277, 2375–2386.
- Kalscheuer, R. (2002). A novel bifunctional wax ester synthase/acyl-CoA:diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in *Acinetobacter calcoaceticus* adp1. J BIOL CHEM 278, 8075–8082.
- Kalscheuer, R., Uthoff, S., Luftmann, H., and Steinbüchel, A. (2003). *In vitro* and *in vivo* biosynthesis of wax diesters by an unspecific bifunctional wax ester synthase/acyl-CoA:diacylglycerol acyltransferase from *Acinetobacter calcoaceticus* ADP1. EUR J LIPID SCI TECH 105, 578–584.
- Kalscheuer, R., Stölting, T., and Steinbüchel, A. (2006a). Microdiesel: *Escherichia coli* engineered for fuel production. MICROBIOLOGY+ (Reading, Engl.) 152, 2529–2536.
- Kalscheuer, R., Stoveken, T., Luftmann, H., Malkus, U., Reichelt, R., and Steinbüchel, A. (2006b). Neutral lipid biosynthesis in engineered *Escherichia coli*: Jojoba oil-like wax esters and fatty acid butyl esters. APPL ENVIRON MICROB 72, 1373–1379.
- Kavanagh, K.L., Jörnvall, H., Persson, B., and Oppermann, U. (2008). Medium- and short-chain dehydrogenase/reductase gene and protein families: the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. CELL MOL LIFE SCI 65, 3895–3906.
- Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the web: a case study using the Phyre server. NAT PROTOC 4, 363–371.
- Khan, A.A., and Kolattukudy, P.E. (1973). A microsomal fatty acid synthetase coupled to acyl-CoA reductase in *Euglena gracilis*. ARCH BIOCHEM BIOPHYS 158, 411–420.
- Kim, E.-K., Moon, J.C., Lee, J.M., Jeong, M.S., Oh, C., Ahn, S.-M., Yoo, Y.J., and Jang, H.H. (2012). Large-scale production of soluble recombinant amyloid- $\beta$  peptide 1-42 using cold-inducible expression system. PROTEIN EXPRES PURIF 86, 53–57.

## LITERATURE

- King, A., Nam, J.-W., Han, J., Hilliard, J., and Jaworski, J.G. (2007). Cuticular wax biosynthesis in petunia petals: cloning and characterisation of an alcohol-acyltransferase that synthesises wax-esters. *PLANTA* 226, 381–394.
- Kleanthous, C., Cullis, P.M., and Shaw, W.V. (1985). 3-(bromoacetyl)chloramphenicol, an active site directed inhibitor for chloramphenicol acetyltransferase. *BIOCHEMISTRY-US* 24, 5307–5313.
- Kligman, A.M. (1963). The uses of sebum. *BRIT J DERMATOL* 75, 307–319.
- Kline & Company, Inc. (2010). Global wax industry 2010: market analysis and opportunities. Retrieved October 5, 2014, from <http://www.klinegroup.com/reports/brochures/y635a/brochure.pdf>.
- Kolattukudy, P.E., and Rogers, L. (1986). Acyl-CoA reductase and acyl-CoA: fatty alcohol acyl transferase in the microsomal preparation from the bovine meibomian gland. *J LIPID RES* 27, 404–411.
- Koumanov, A., Benach, J., Atrian, S., González-Duarte, R., Karshikoff, A., and Ladenstein, R. (2003). The catalytic mechanism of *Drosophila* alcohol dehydrogenase: evidence for a proton relay modulated by the coupled ionization of the active site lysine/tyrosine pair and a NAD<sup>+</sup> ribose oh switch. *PROTEINS* 51, 289–298.
- Kozlikova, B., Sebestova, E., Sustr, V., Brezovsky, J., Strnad, O., Daniel, L., Bednar, D., Pavelka, A., Manak, M., Bezdeka, M., et al. (2014). Caver analyst 1.0: graphic tool for interactive visualization and analysis of tunnels and channels in protein structures. *BIOINFORMATICS*, DOI: 10.1093/bioinformatics/btu364.
- Kraftfahrt-Bundesamt (2013). Pressemitteilung Nr. 8/2013 - Der Fahrzeugbestand am 1. Januar 2013. Retrieved October 5, 2014, from [http://www.kba.de/SharedDocs/Publikationen/DE/PM/2013/pm\\_08\\_13\\_bestand\\_2013\\_pdf](http://www.kba.de/SharedDocs/Publikationen/DE/PM/2013/pm_08_13_bestand_2013_pdf)
- Kraftfahrt-Bundesamt (2014). Pressemitteilung Nr. 10/2014 - Der Fahrzeugbestand am 1. Januar 2014. Retrieved October 5, 2014, from [http://www.kba.de/SharedDocs/Publikationen/DE/PM/2014/pm\\_10\\_14\\_bestand\\_2014\\_pdf](http://www.kba.de/SharedDocs/Publikationen/DE/PM/2014/pm_10_14_bestand_2014_pdf).
- Kramer, G., Rauch, T., Rist, W., Vorderwülbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002). L23 protein functions as a chaperone docking site on the ribosome. *NATURE* 419, 171–174.
- Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *J MOL BIOL* 157, 105–132.
- Ladner, C.L., Yang, J., Turner, R.J., and Edwards, R.A. (2004). Visible fluorescent detection of proteins in polyacrylamide gels without staining. *ANAL BIOCHEM* 326, 13–20.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *NATURE* 227, 680–685.
- Lardizabal, K.D., Metz, J.G., Sakamoto, T., Hutton, W.C., Pollard, M.R., and Lassner, M.W. (2000). Purification of a jojoba embryo wax synthase, cloning of its cDNA, and production of high levels of wax in seeds of transgenic *Arabidopsis*. *PLANT PHYSIOL* 122, 645–656.

## LITERATURE

- Lassner, M.W., Lardizabal, K., and Metz, J.G. (1996). A jojoba beta-ketoacyl-CoA synthase cDNA complements the canola fatty acid elongation mutation in transgenic plants. *PLANT CELL Online* 8, 281–292.
- Lee, S.B., Kim, H., Kim, R.J., and Suh, M.C. (2014). Overexpression of *Arabidopsis* myb96 confers drought resistance in *Camelina sativa* via cuticular wax accumulation. *PLANT CELL REP* 33, 1535–1546.
- Lehner, R., and Kuksis, A. (1995). Triacylglycerol synthesis by purified triacylglycerol synthetase of rat intestinal mucosa. Role of acyl-CoA acyltransferase. *J BIOL CHEM* 270, 13630–13636.
- Lesk, A.M. (1995). NAD-binding domains of dehydrogenases. *CURR OPIN STRUC BIOL* 5, 775–783.
- Leslie, A.G., Moody, P.C., and Shaw, W.V. (1988). Structure of chloramphenicol acetyltransferase at 1.75-Å resolution. *P NATL ACAD SCI USA* 85, 4133–4137.
- Li, F., Wu, X., Lam, P., Bird, D., Zheng, H., Samuels, L., Jetter, R., and Kunst, L. (2008). Identification of the wax ester synthase/acyl-coenzyme A:diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in *Arabidopsis*. *PLANT PHYSIOL* 148, 97–107.
- Lienard, M.A., Hagstrom, A.K., Lassance, J.-M., and Lofstedt, C. (2010). Evolution of multicomponent pheromone signals in small ermine moths involves a single fatty-acyl reductase gene. *P NATL ACAD SCI USA* 107, 10955–10960.
- Lippold, F., vom Dorp, K., Abraham, M., Hölzl, G., Wewer, V., Yilmaz, J.L., Lager, I., Montandon, C., Besagni, C., Kessler, F., et al. (2012). Fatty acid phytyl ester synthesis in chloroplasts of *Arabidopsis*. *PLANT CELL* 24, 2001–2014.
- Liu, A., Tan, X., Yao, L., and Lu, X. (2013). Fatty alcohol production in engineered *E. coli* expressing *Marinobacter* fatty acyl-CoA reductases. *APPL MICROBIOL BIOT* 97, 7061–7071.
- Liu, Q., Siloto, R.M.P., Snyder, C.L., and Weselake, R.J. (2011). Functional and topological analysis of yeast acyl-CoA:diacylglycerol acyltransferase 2, an endoplasmic reticulum enzyme essential for triacylglycerol biosynthesis. *J BIOL CHEM* 286, 13115–13126.
- Liu, Q., Siloto, R.M.P., Lehner, R., Stone, S.J., and Weselake, R.J. (2012). Acyl-CoA:diacylglycerol acyltransferase: molecular biology, biochemistry and biotechnology. *PROG LIPID RES* 51, 350–377.
- Ludlam, A.V., Moore, B.A., and Xu, Z. (2004). The crystal structure of ribosomal chaperone trigger factor from *Vibrio cholerae*. *P NATL ACAD SCI USA* 101, 13436–13441.
- Maina, C.V., Riggs, P.D., Grandea, A.G., 3rd, Slatko, B.E., Moran, L.S., Tagliamonte, J.A., McReynolds, L.A., and Guan, C.D. (1988). An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. *GENE* 74, 365–373.
- Mandel, M., and Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *J MOL BIOL* 53, 159–162.
- Mang, T. (1998). Umweltrelevante Kriterien zur Anwendung von Pflanzenölen und deren Derivaten im Schmierstoffbereich. *FETT-LIPID* 100, 524–527.
- Marahiel, M.A., Stachelhaus, T., and Mootz, H.D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *CHEM REV* 97, 2651–2674.

## LITERATURE

- McCarthy, K. (2014). Jojoba oil benefits. Retrieved October 5, 2014, from <http://www.livestrong.com>  
<http://www.livestrong.com/article/87154-jojoba-oil-benefits>
- McFie, P.J., Stone, S.L., Banman, S.L., and Stone, S.J. (2010). Topological orientation of acyl-CoA:diacylglycerol acyltransferase-1 (DGAT1) and identification of a putative active site histidine and the role of the n terminus in dimer/tetramer formation. *J BIOL CHEM* 285, 37377–37387.
- McFie, P.J., Banman, S.L., Kary, S., and Stone, S.J. (2011). Murine diacylglycerol acyltransferase-2 (DGAT2) can catalyse triacylglycerol synthesis and promote lipid droplet formation independent of its localisation to the endoplasmic reticulum. *J BIOL CHEM* 286, 28235–28246.
- McFie, P.J., Jin, Y., Banman, S.L., Beauchamp, E., Berthiaume, L.G., and Stone, S.J. (2014). Characterisation of the interaction of diacylglycerol acyltransferase-2 with the endoplasmic reticulum and lipid droplets. *BIOCHIM BIOPHYS ACTA*, DOI: 10.1016/j.bbali.2014.06.004.
- Metz, J.G., Pollard, M.R., Anderson, L., Hayes, T.R., and Lassner, M.W. (2000). Purification of a jojoba embryo fatty acyl-coenzyme A reductase and expression of its cDNA in high erucic acid rapeseed. *PLANT PHYSIOL* 122, 635–644.
- Miklaszewska, M., Kawiński, A., and Banaś, A. (2013). Detailed characterisation of the substrate specificity of mouse wax synthase. *ACTA BIOCHIM POL* 60, 209–215.
- Morris, R.J. (1973). The lipid structure of the spermaceti organ of the sperm whale (*Physeter catodon*). *DEEP-SEA RES OCEANOGR ABSTR* 20, 911–916.
- Morris, R.J. (1975). Further studies into the lipid structure of the spermaceti organ of the sperm whale (*Physeter catodon*). *DEEP-SEA RES OCEANOGR ABSTR* 22, 483–489.
- Moto, K., Yoshiga, T., Yamamoto, M., Takahashi, S., Okano, K., Ando, T., Nakata, T., and Matsumoto, S. (2003). Pheromone gland-specific fatty-acyl reductase of the silkworm, *Bombyx mori*. *P NATL ACAD SCI USA* 100, 9156–9161.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *COLD SPRING HARB SYM* 51 Pt 1, 263–273.
- Oganesyan, N., Ankoudinova, I., Kim, S.-H., and Kim, R. (2007). Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallisation. *PROTEIN EXPRES PURIF* 52, 280–285.
- Onwueme, K.C., Ferreras, J.A., Buglino, J., Lima, C.D., and Quadri, L.E.N. (2004). Mycobacterial polyketide-associated proteins are acyltransferases: proof of principle with *Mycobacterium tuberculosis* PapA5. *P NATL ACAD SCI USA* 101, 4608–4613.
- Page, R.D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *COMPUT APPL BIOSCI* 12, 357–358.
- Patel, S., Nelson, D.R., and Gibbs, A.G. (2001). Chemical and physical analyses of wax ester properties. *J INSECT SCI* 1, 4.
- Patzelt, H., Kramer, G., Rauch, T., Schönfeld, H.-J., Bukau, B., and Deuerling, E. (2002). Three-state equilibrium of *Escherichia coli* trigger factor. *BIOL CHEM* 383, 1611–1619.

## LITERATURE

- Penkov, S., Ogawa, A., Schmidt, U., Tate, D., Zagoriy, V., Boland, S., Gruner, M., Vorkel, D., Verbavatz, J.-M., Sommer, R.J., et al. (2014). A wax ester promotes collective host finding in the nematode *Pristionchus pacificus*. *NAT CHEM BIOL*, DOI: 10.1038/nchembio.1460.
- Perry, S.L., DeMaster, D.P., and Silber, G.K. (1999). The great whales: history and status of six species listed as endangered under the u.s. endangered species act of 1973. *MAR FISH REV* 1–74.
- Persson, B., Krook, M., and Jörnvall, H. (1991). Characteristics of short-chain alcohol dehydrogenases and related enzymes. *EUR J BIOCHEM* 200, 537–543.
- Pohanish, R.P. (2004). *HazMat data: for first response, transportation, storage, and security* (Hoboken, N.J.: J. Wiley & Sons).
- Pollard, M.R., McKeon, T., Gupta, L.M., and Stumpf, P.K. (1979). Studies on biosynthesis of waxes by developing jojoba seed. II. The demonstration of wax biosynthesis by cell-free homogenates. *LIPIDS* 14, 651–662.
- Qing, G., Ma, L.-C., Khorchid, A., Swapna, G.V.T., Mal, T.K., Takayama, M.M., Xia, B., Phadtare, S., Ke, H., Acton, T., et al. (2004). Cold-shock induced high-yield protein production in *Escherichia coli*. *NAT BIOTECHNOL* 22, 877–882.
- Qiu, X., Mistry, A., Ammirati, M.J., Chrnyk, B.A., Clark, R.W., Cong, Y., Culp, J.S., Danley, D.E., Freeman, T.B., Geoghegan, K.F., et al. (2007). Crystal structure of cholesteryl ester transfer protein reveals a long tunnel and four bound lipid molecules. *NAT STRUCT MOL BIOL* 14, 106–113.
- Rani, S.H., Krishna, T.H.A., Saha, S., Negi, A.S., and Rajasekharan, R. (2010). Defective in cuticular ridges (*dcr*) of *Arabidopsis thaliana*, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. *J BIOL CHEM* 285, 38337–38347.
- Rani, S.H., Saha, S., and Rajasekharan, R. (2012). A soluble diacylglycerol acyltransferase is involved in triacylglycerol biosynthesis in the oleaginous yeast *Rhodotorula glutinis*. *MICROBIOLOGY+* 159, 155–166.
- Rantamäki, A.H., Wiedmer, S.K., and Holopainen, J.M. (2013). Melting points--the key to the anti-evaporative effect of the tear film wax esters. *INVEST OPHTH VIS SCI* 54, 5211–5217.
- Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., and Deber, C.M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *P NATL ACAD SCI USA* 106, 1760–1765.
- Ray, M., Mishra, P., Das, P., and Sabat, S.C. (2012). Expression and purification of soluble bio-active rice plant catalase-A from recombinant *Escherichia coli*. *J BIOTECHNOL* 157, 12–19.
- Razeq, F.M., Kosma, D.K., Rowland, O., and Molina, I. (2014). Extracellular lipids of *Camelina sativa*: characterisation of chloroform-extractable waxes from aerial and subterranean surfaces. *PHYTOCHEMISTRY*, DOI: <http://dx.doi.org/10.1016/j.phytochem.2014.06.018>.
- Reiser, S., and Somerville, C. (1997). Isolation of mutants of *Acinetobacter calcoaceticus* deficient in wax ester synthesis and complementation of one mutation with a gene encoding a fatty acyl coenzyme a reductase. *J BACTERIOL* 179, 2969–2975.
- Riendeau, D., and Meighen, E. (1985). Enzymatic reduction of fatty acids and acyl-CoAs to long chain aldehydes and alcohols. *EXPERIENTIA* 41, 707–713.

## LITERATURE

- Roberts, M.F. (2005). Organic compatible solutes of halotolerant and halophilic microorganisms. *SALINE SYSTEMS* 1, 5.
- Roelofs, W.L., and Wolf, W.A. (1988). Pheromone biosynthesis in lepidoptera. *J CHEM ECOL* 14, 2019–2031.
- Rosendal, J., and Knudsen, J. (1992). A fast and versatile method for extraction and quantitation of long-chain acyl-CoA esters from tissue: content of individual long-chain acyl-CoA esters in various tissues from fed rat. *ANAL BIOCHEM* 207, 63–67.
- Rottem, S. (1980). Membrane lipids of mycoplasmas. *BBA-BIOMEMBRANES* 604, 65–90.
- Röttig, A., and Steinbüchel, A. (2013). Acyltransferases in bacteria. *MICROBIOL MOL BIOL R* 77, 277–321.
- Rottler, A.-M., Schulz, S., and Ayasse, M. (2013). Wax lipids signal nest identity in bumblebee colonies. *J CHEM ECOL* 39, 67–75.
- Rowland, O., and Domergue, F. (2012). Plant fatty acyl reductases: enzymes generating fatty alcohols for protective layers with potential for industrial applications. *PLANT SCI* 193-194, 28–38.
- Rowland, O., Zheng, H., Hepworth, S.R., Lam, P., Jetter, R., and Kunst, L. (2006). CER4 Encodes an alcohol-forming fatty acyl-coenzyme a reductase involved in cuticular wax production in *Arabidopsis*. *PLANT PHYSIOL* 142, 866–877.
- Saha, S., Enugutti, B., Rajakumari, S., and Rajasekharan, R. (2006). Cytosolic triacylglycerol biosynthetic pathway in oilseeds. molecular cloning and expression of peanut cytosolic diacylglycerol acyltransferase. *PLANT PHYSIOL* 141, 1533–1543.
- Saio, T., Guan, X., Rossi, P., Economou, A., and Kalodimos, C.G. (2014). Structural basis for protein antiaggregation activity of the trigger factor chaperone. *SCIENCE* 344, 1250494–1250494.
- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). *Molecular cloning: a laboratory manual*, 2<sup>nd</sup> edn. Cold Spring Harbor NY: Cold Spring Harbor Laboratory.
- Sandager, L. (2001). Storage lipid synthesis is non-essential in yeast. *J BIOL CHEM* 277, 6478–6482.
- Santala, S., Efimova, E., Koskinen, P., Karp, M.T., and Santala, V. (2014). Rewiring the wax ester production pathway of *Acinetobacter baylyi* ADP1. *ACS SYNTH BIOL* 3, 145–151.
- Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *NAT METHODS* 9, 671–675.
- Schrödinger (2010). *PyMOL The PyMOL Molecular Graphics System, Version 1.3*, Schrödinger, LLC.
- Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., Frommer, W.B., Flügge, U.-I., and Kunze, R. (2003). ARAMEMNON, a novel database for *Arabidopsis* integral membrane proteins. *PLANT PHYSIOL* 131, 16–26.
- Sharff, A.J., Rodseth, L.E., and Quijcho, F.A. (1993). Refined 1.8-Å structure reveals the mode of binding of beta-cyclodextrin to the maltodextrin binding protein. *BIOCHEMISTRY-US* 32, 10553–10559.

## LITERATURE

- Shaw, W.V. (1983). Chloramphenicol acetyltransferase: enzymology and molecular biology. *CRC CR REV BIOCH MOL* 14, 1–46.
- Shi, J., Tan, H., Yu, X.-H., Liu, Y., Liang, W., Ranathunge, K., Franke, R.B., Schreiber, L., Wang, Y., Kai, G., et al. (2011). Defective pollen wall is required for anther and microspore development in rice and encodes a fatty acyl carrier protein reductase. *PLANT CELL* 23, 2225–2246.
- Shi, S., Valle-Rodríguez, J.O., Khoomrung, S., Siewers, V., and Nielsen, J. (2012). Functional expression and characterisation of five wax ester synthases in *Saccharomyces cerevisiae* and their utility for biodiesel production. *BIOTECHNOL BIOFUELS* 5, 7.
- Shi, S., Valle-Rodríguez, J.O., Siewers, V., and Nielsen, J. (2014). Engineering of chromosomal wax ester synthase integrated *Saccharomyces cerevisiae* mutants for improved biosynthesis of fatty acid ethyl esters. *BIOTECHNOL BIOENG* 111, 1740–1747.
- Shockey, J.M. (2006). Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localised to different subdomains of the endoplasmic reticulum. *PLANT CELL ONLINE* 18, 2294–2313.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *MOL SYST BIOL* 7, 539–539.
- Sikkema, J., de Bont, J.A., and Poolman, B. (1995). Mechanisms of membrane toxicity of hydrocarbons. *MICROBIOL REV* 59, 201–222.
- Siloto, R.M.P. (2006). The accumulation of oleosins determines the size of seed oilbodies in *Arabidopsis*. *PLANT CELL ONLINE* 18, 1961–1974.
- Sirakova, T.D., Deb, C., Daniel, J., Singh, H.D., Maamar, H., Dubey, V.S., and Kolattukudy, P.E. (2012). Wax ester synthesis is required for *Mycobacterium tuberculosis* to enter *in vitro* dormancy. *PLOS ONE* 7, e51641.
- Smith, K.R., and Thiboutot, D.M. (2008). Thematic review series: skin lipids. sebaceous gland lipids: friend or foe? *J LIPID RES* 49, 271–281.
- Sonnhammer, E.L., von Heijne, G., and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *PROC INT CONF INTELL SYST MOL BIOL* 6, 175–182.
- Spencer, G.F., Plattner, R.D., and Miwa, T. (1977). Jojoba oil analysis by high pressure liquid chromatography and gas chromatography/mass spectrometry. *J AM OIL CHEM SOC* 54, 187–189.
- Stachelhaus, T., Mootz, H.D., Bergendahl, V., and Marahiel, M.A. (1998). Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *J BIOL CHEM* 273, 22773–22781.
- Steen, E.J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., del Cardayre, S.B., and Keasling, J.D. (2010). Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *NATURE* 463, 559–562.

## LITERATURE

- Stoller, G., Rücknagel, K.P., Nierhaus, K.H., Schmid, F.X., Fischer, G., and Rahfeld, J.U. (1995). A ribosome-associated peptidyl-prolyl cis/trans isomerase identified as the trigger factor. *EMBO J* **14**, 4939–4948.
- Stone, S.J., Levin, M.C., and Farese, R.V. (2006). Membrane topology and identification of key functional amino acid residues of murine acyl-CoA:diacylglycerol acyltransferase-2. *J BIOL CHEM* **281**, 40273–40282.
- Stone, S.J., Levin, M.C., Zhou, P., Han, J., Walther, T.C., and Farese, R.V., Jr (2009). The endoplasmic reticulum enzyme DGAT2 is found in mitochondria-associated membranes and has a mitochondrial targeting signal that promotes its association with mitochondria. *J BIOL CHEM* **284**, 5352–5361.
- Stöveken, T., Kalscheuer, R., Malkus, U., Reichelt, R., and Steinbüchel, A. (2005). The wax ester synthase/acyl coenzyme a:diacylglycerol acyltransferase from *Acinetobacter* sp. strain ADP1: characterisation of a novel type of acyltransferase. *J BACTERIOL* **187**, 1369–1376.
- Stöveken, T., Kalscheuer, R., and Steinbüchel, A. (2009). Both histidine residues of the conserved HHXXXDG motif are essential for wax ester synthase/acyl-CoA:diacylglycerol acyltransferase catalysis. *EUR J LIPID SCI TECH* **111**, 112–119.
- Tada, A., Jin, Z.-L., Sugimoto, N., Sato, K., Yamazaki, T., and Tanamoto, K. (2005). Analysis of the constituents in jojoba wax used as a food additive by LC/MS/MS. *J FOOD HYG SOC JPN* **46**, 198–204.
- Takagi, T., and Itabashi, Y. (1977). Random combinations of acyl and alcoholic groups through overall wax esters of sperm whale head oils. *COMP BIOCHEM PHYS B* **57**, 37–39.
- Tamada, T., Feese, M.D., Ferri, S.R., Kato, Y., Yajima, R., Toguri, T., and Kuroki, R. (2004). Substrate recognition and selectivity of plant glycerol-3-phosphate acyltransferases (GPATs) from *Cucurbita moscata* and *Spinacea oleracea*. *ACTA CRYSTALLOGR D* **60**, 13–21.
- Tan, X., Yao, L., Gao, Q., Wang, W., Qi, F., and Lu, X. (2011). Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. *METAB ENG* **13**, 169–176.
- Tanovic, A., Samel, S.A., Essen, L.-O., and Marahiel, M.A. (2008). Crystal structure of the termination module of a nonribosomal peptide synthetase. *SCIENCE* **321**, 659–663.
- Teerawanichpan, P., and Qiu, X. (2010). Fatty acyl-CoA reductase and wax synthase from *Euglena gracilis* in the biosynthesis of medium-chain wax esters. *LIPIDS* **45**, 263–273.
- Teerawanichpan, P., and Qiu, X. (2011). Molecular and functional analysis of three fatty acyl-CoA reductases with distinct substrate specificities in copepod *Calanus finmarchicus*. *MAR BIOTECHNOL* **14**, 227–236.
- Teerawanichpan, P., Robertson, A.J., and Qiu, X. (2010). A fatty acyl-CoA reductase highly expressed in the head of honey bee (*Apis mellifera*) involves biosynthesis of a wide range of aliphatic fatty alcohols. *INSECT BIOCHEM MOLEC* **40**, 641–649.
- Tehlivets, O., Scheuringer, K., and Kohlwein, S.D. (2007). Fatty acid synthesis and elongation in yeast. *BIOCHIM BIOPHYS ACTA* **1771**, 255–270.
- The UniProt Consortium (2014). Activities at the universal protein resource (UniProt). *NUCLEIC ACIDS RES* **42**, D191–D198.

## LITERATURE

- Thiele, J.J., Weber, S.U., and Packer, L. (1999). Sebaceous gland secretion is a major physiologic route of vitamin E delivery to skin. *J INVEST DERMATOL* 113, 1006–1010.
- Thompson, J. (1997). The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *NUCLEIC ACIDS RES* 25, 4876–4882.
- Tower, W.S. (1907). A history of the American whale fishery (Publications of the University of Pennsylvania.).
- Tulloch, A.P. (1970). The composition of beeswax and other waxes secreted by insects. *LIPIDS* 5, 247–258.
- Tulloch, A.P. (1971). Beeswax: structure of the esters and their component hydroxy acids and diols. *CHEM PHYS LIPIDS* 6, 235–265.
- Urbanova, K., Vrkoslav, V., Valterova, I., Hakova, M., and Cvacka, J. (2012). Structural characterisation of wax esters by electron ionization mass spectrometry. *J LIPID RES* 53, 204–213.
- Uthoff, S., Stoveken, T., Weber, N., Vosmann, K., Klein, E., Kalscheuer, R., and Steinbüchel, A. (2005). Thio wax ester biosynthesis utilizing the unspecific bifunctional wax ester synthase/acyl coenzyme a:diacylglycerol acyltransferase of *Acinetobacter* sp. strain ADP1. *APPL ENVIRON MICROB* 71, 790–796.
- Villa, J.A., Cabezas, M., de la Cruz, F., and Moncalian, G. (2013). Use of limited proteolysis and mutagenesis to identify folding domains and sequence motifs critical for wax ester synthase/acyl coenzyme a:diacylglycerol acyltransferase activity. *APPL ENVIRON MICROB* 80, 1132–1141.
- Wagner, H., Luther, R., and Mang, T. (2001). Lubricant base fluids based on renewable raw materials: their catalytic manufacture and modification. *APPL CATAL A-GEN* 221, 429–442.
- Wahlen, B.D., Oswald, W.S., Seefeldt, L.C., and Barney, B.M. (2009). Purification, characterisation, and potential bacterial wax production role of an NADPH-dependent fatty aldehyde reductase from *Marinobacter aquaeolei* VT8. *APPL ENVIRON MICROB* 75, 2758–2764.
- Wältermann, M., and Steinbüchel, A. (2005). Neutral lipid bodies in prokaryotes: recent insights into structure, formation, and relationship to eukaryotic lipid depots. *J BACTERIOL* 187, 3607–3619.
- Wältermann, M., Stöveken, T., and Steinbüchel, A. (2007). Key enzymes for biosynthesis of neutral lipid storage compounds in prokaryotes: properties, function and occurrence of wax ester synthases/acyl-CoA:diacylglycerol acyltransferases. *BIOCHIMIE* 89, 230–242.
- Wang, X., and Kolattukudy, P.E. (1995). Solubilization, purification and characterisation of fatty acyl-CoA reductase from duck uropygial gland. *BIOCHEM BIOPH RES CO* 208, 210–215.
- Wattanachaisaereekul, S., Lantz, A.E., Nielsen, M.L., and Nielsen, J. (2008). Production of the polyketide 6-MSA in yeast engineered for increased malonyl-CoA supply. *METAB ENG* 10, 246–254.
- Wei, H. (2012). An overview of wax production, requirement and supply in the world market. *EUROP CHEM BULL* 1, 266–268.

## LITERATURE

- Weiss, B., and Richardson, C.C. (1967). Enzymatic breakage and joining of deoxyribonucleic acid, I. Repair of single-strand breaks in DNA by an enzyme system from *Escherichia coli* infected with T4 bacteriophage. P NATL ACAD SCI USA 57, 1021–1028.
- Wellendorf, M. (1963). Composition of spermaceti. NATURE 198, 1086–1087.
- White, S.W., Zheng, J., Zhang, Y.-M., and Rock, C.O. (2005). The structural biology of type ii fatty acid biosynthesis. ANNU REV BIOCHEM 74, 791–831.
- Whitehead, H. (2009). Sperm Whale. In Encyclopedia of Marine Mammals, (Elsevier), pp. 1091–1097.
- Wilfling, F., Wang, H., Haas, J.T., Krahmer, N., Gould, T.J., Uchida, A., Cheng, J.-X., Graham, M., Christiano, R., Fröhlich, F., et al. (2013). Triacylglycerol synthesis enzymes mediate lipid droplet growth by relocating from the ER to lipid droplets. DEV CELL 24, 384–399.
- Willis, R.M., Wahlen, B.D., Seefeldt, L.C., and Barney, B.M. (2011). Characterisation of a fatty acyl-CoA reductase from *Marinobacter aquaeolei* VT8: a bacterial enzyme catalyzing the reduction of fatty acyl-CoA to fatty alcohol. BIOCHEMISTRY-US 50, 10550–10558.
- Wisniak, J. (1987). The chemistry and technology of Jojoba oil (Champaign, Illinois: American Oil Chemists' Society).
- Woodman, M.E. (2008). Direct PCR of intact bacteria (colony PCR). In Current Protocols in Microbiology, R. Coico, T. Kowalik, J. Quarles, B. Stevenson, and R. Taylor, eds. (Hoboken, NJ, USA: John Wiley & Sons, Inc.).
- Yeats, T.H., and Rose, J.K.C. (2013). The formation and function of plant cuticles. PLANT PHYSIOL 163, 5–20.
- Yeh, A.P., Chatelet, C., Soltis, S.M., Kuhn, P., Meyer, J., and Rees, D.C. (2000). Structure of a thioredoxin-like [2Fe-2S] ferredoxin from *Aquifex aeolicus*. J MOL BIOL 300, 587–595.
- Yen, C.-L.E., Brown, C.H., Monetti, M., and Farese, R.V. (2005). A human skin multifunctional O-acyltransferase that catalyzes the synthesis of acylglycerols, waxes, and retinyl esters. J LIPID RES 46, 2388–2397.
- Yin, B., Cui, D., Zhang, L., Jiang, S., Machida, S., Yuan, Y.A., and Wei, D. (2014). Structural insights into substrate and coenzyme preference by SDR family protein Gox2253 from *Gluconobacter oxydans*: Crystal Structure of *G. oxydans* Gox2253. PROTEINS, DOI: 10.1002/prot.24603.
- Au-Young, J., and Fielding, C.J. (1992). Synthesis and secretion of wild-type and mutant human plasma cholesteryl ester transfer protein in baculovirus-transfected insect cells: the carboxyl-terminal region is required for both lipoprotein binding and catalysis of transfer. P NATL ACAD SCI USA 89, 4094–4098.
- Yu, K.O., Jung, J., Kim, S.W., Park, C.H., and Han, S.O. (2012). Synthesis of FAEEs from glycerol in engineered *Saccharomyces cerevisiae* using endogenously produced ethanol by heterologous expression of an unspecific bacterial acyltransferase. BIOTECHNOL BIOENG 109, 110–115.
- Yuan, Z., Yin, B., Wei, D., and Yuan, Y.A. (2013). Structural basis for cofactor and substrate selection by cyanobacterium succinic semialdehyde dehydrogenase. J STRUCT BIOL 182, 125–135.
- Yuvaniyama, P., Agar, J.N., Cash, V.L., Johnson, M.K., and Dean, D.R. (2000). NifS-directed assembly of a transient [2Fe-2S] cluster within the NifU protein. P NATL ACAD SCI USA 97, 599–604.

## LITERATURE

- Zaffagnini, M., Michelet, L., Sciabolini, C., Di Giacinto, N., Morisse, S., Marchand, C.H., Trost, P., Fermani, S., and Lemaire, S.D. (2014). High-resolution crystal structure and redox properties of chloroplastic triosephosphate isomerase from *Chlamydomonas reinhardtii*. *MOL PLANT* 7, 101–120.
- Zhang, J., Xu, D., Nie, J., Cao, J., Zhai, Y., Tong, D., and Shi, Y. (2014). Monoacylglycerol acyltransferase-2 is a tetrameric enzyme that selectively heterodimerizes with diacylglycerol acyltransferase-1. *J BIOL CHEM*, DOI: 10.1074/jbc.M113.530022.
- Zhang, M., Fan, J., Taylor, D.C., and Ohlrogge, J.B. (2009). DGAT1 and PDAT1 acyltransferases have overlapping functions in *Arabidopsis* triacylglycerol biosynthesis and are essential for normal pollen and seed development. *PLANT CELL* 21, 3885–3901.
- Zheng, Y.-N., Li, L.-L., Liu, Q., Yang, J.-M., Wang, X.-W., Liu, W., Xu, X., Liu, H., Zhao, G., and Xian, M. (2012). Optimisation of fatty alcohol biosynthesis pathway for selectively enhanced production of C12/14 and C16/18 fatty alcohols in engineered *Escherichia coli*. *MICROB CELL FACT* 11, 65.
- Zhou, X.-R., Shrestha, P., Yin, F., Petrie, J.R., and Singh, S.P. (2013). AtDGAT2 is a functional acyl-CoA:diacylglycerol acyltransferase and displays different acyl-CoA substrate preferences than AtDGAT1. *FEBS LETT* 587, 2371–2376.
- Zimmer, W.M.X., Madsen, P.T., Teloni, V., Johnson, M.P., and Tyack, P.L. (2005). Off-axis effects on the multipulse structure of sperm whale usual clicks with implications for sound production. *J ACOUST SOC AM* 118, 3337–3345.
- Current Protocols in Molecular Biology (2001). (Hoboken, NJ, USA: John Wiley & Sons, Inc.).
- Handbook of Hydrocarbon and Lipid Microbiology (2010). (Berlin, Heidelberg: Springer Berlin Heidelberg).

## 10 LIST OF ABBREVIATIONS

Table 19: List of Abbreviations.

ABBREVIATIONS	MEANING
6xHis	hexahistidine-tag / hexahistidine-tagged
AbWSD1	WSD1 from <i>A. baylyi</i>
AC	affinity chromatography
ACP	acyl carrier protein
APS	ammonium persulfate
AtDGAT3	DGAT3 from <i>A. thaliana</i>
AtOLE3	Oleosin 3 from <i>A. thaliana</i>
AWAT2	acyl-CoA wax alcohol acyltransferase 2
BSTFA	N,O-Bis(trimethylsilyl)trifluoroacetamide
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CoA	coenzyme A
DAG	diacylglycerol
DGAT	diacylglycerol O acyltransferase
DmFAR1	FAR1 from <i>D. melanogaster</i>
DTNB	5-(3-Carboxy-4-nitrophenyl)disulfanyl-2-nitrobenzoic acid
ER	endoplasmic reticulum
FAEE	fatty acid ethyl esters
FAME	fatty acid methyl esters
FAR	fatty acyl reductase
GC-MS	gas chromatography-coupled mass spectrometry
HIC	hydrophobic interaction chromatography
IMAC	immobilized metal ion chromatography
IP	isoelectric point
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
KCS	$\beta$ -ketoacyl-CoA synthase
LD	lipid droplet
MaFAR1	FAR1 from <i>M. aquaeolei</i>
MALS	multiangular light scattering
MAM	mitochondria associated membranes
MBP	maltose binding protein
MmAWAT2	AWAT2 from <i>M. musculus</i>
MmDGAT2	DGAT2 from <i>M. musculus</i>
MPD	2-methyl-2,4-pentadiole
MWM	molecular weight marker

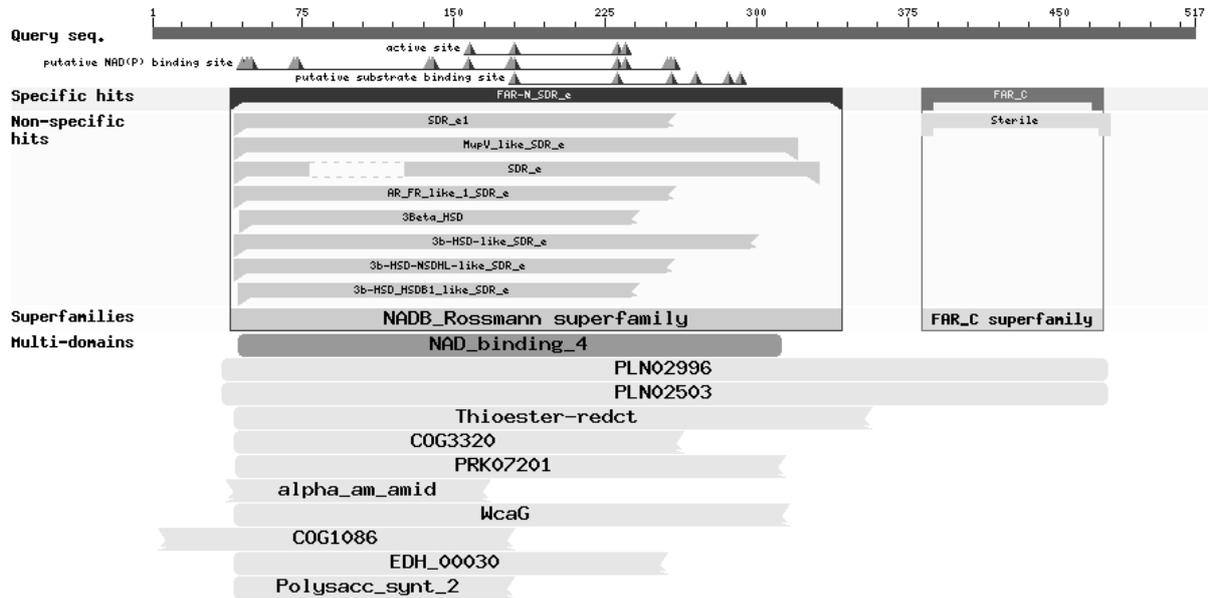
## LIST OF ABBREVIATIONS

NLBS	neutral lipid binding site
PHA	polyhydroxyalkanoate
rpm	rounds per minute
RT	room temperature
SDR	short chain dehydrogenase/reductase
SDS	sodium dodecyl sulphate
SEC	size exclusion chromatography
TEA	triethylamine
TAG	triacylglycerol
TBS	Tris-buffered saline
TCE	2,2,2-trichloroethanol
TEMED	tetramethylethylenediamin
TF	trigger factor from <i>E. coli</i>
TIC	total ion count
TLC	thin layer chromatography
T <sub>m</sub>	melting temperature
TMS	trimethylsilyl
VLC	very long chain
VLC WE	very long chain acyl chain including wax ester
WE	wax esters
WS	wax synthase
WSD	bifunctional wax synthase/diacylglycerol O-acyltransferase

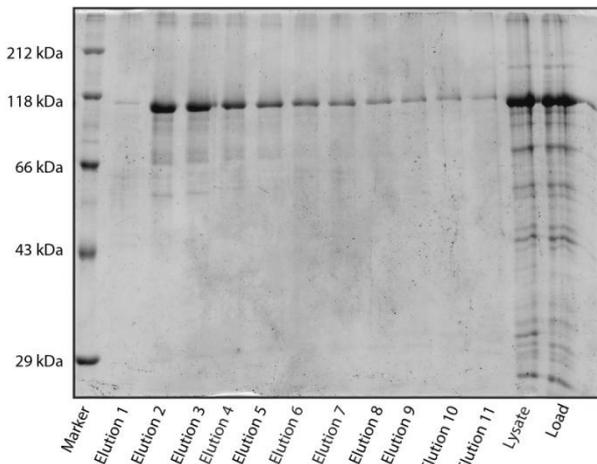
### Acyl chain nomenclature

Acyl chains are abbreviated by a number code in this thesis. In this code, the number in front of the colon stands for the total number of carbon atoms of the acyl chain, while the number behind the colon indicates the number of double bonds within the chain. In case of 18:1, the acyl chain consists of 18 carbon atoms and carries a single double bond.

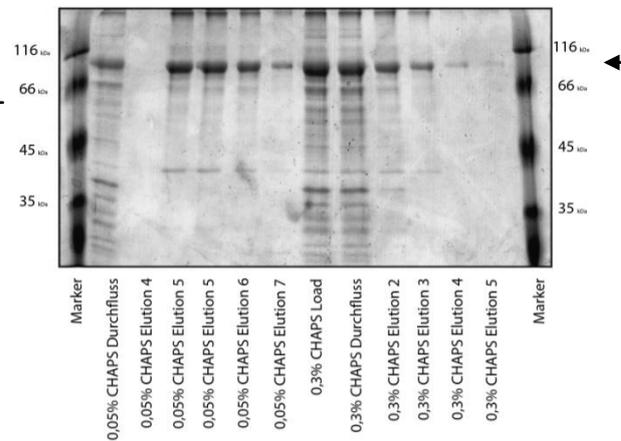
11 APPENDIX



**Appendix 1: Domain prediction of DmFAR1.** DmFAR1 domain prediction (UniProt ID = Q8MS59) was done according to BLAST (Altschul et al., 1997, 2005).

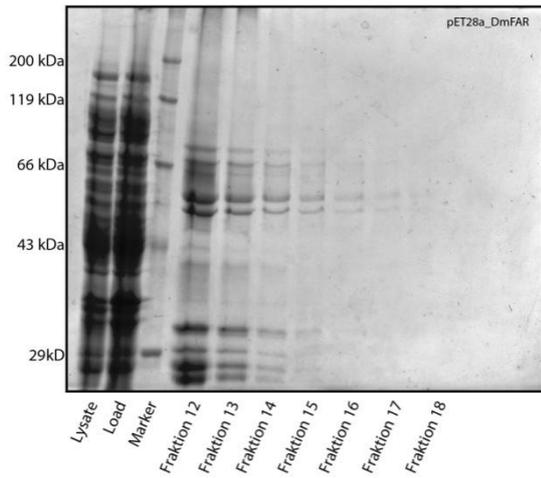


**Appendix 2: IMAC purification of 6xHis-TF-DmFAR1.** Binding buffer: 50 mM Tris/HCl pH = 7.6, 300 mM NaCl, 10 % glycerol, 0.3 %CHAPS. Elution buffer: Binding buffer + 500 mM imidazole. Lysate = lysate of whole cells, Load = soluble fraction of whole cell lysate used for loading of the IMAC column. Elution 1-11 = Elution fractions. Position of 6xHis-TF-DmFAR1 is indicated by the black arrow.

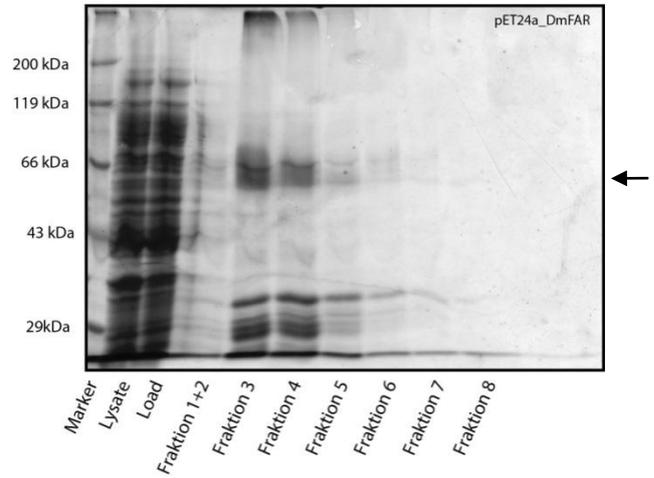


**Appendix 3: Purification of MBP-DmFAR1 via amylose resin affinity chromatography.** Binding buffer: 50 mM Tris/HCl pH = 7.6, 200 mM NaCl, 10 % glycerol, CHAPS as depicted. Elution buffer: Binding buffer + 10 mM maltose. Load = soluble fraction of whole cell lysate used for loading of the IMAC column. Durchfluss = Flow-through after column loading. Elution = Elution fractions. Position of MBP-DmFAR1 is indicated by the black arrow.

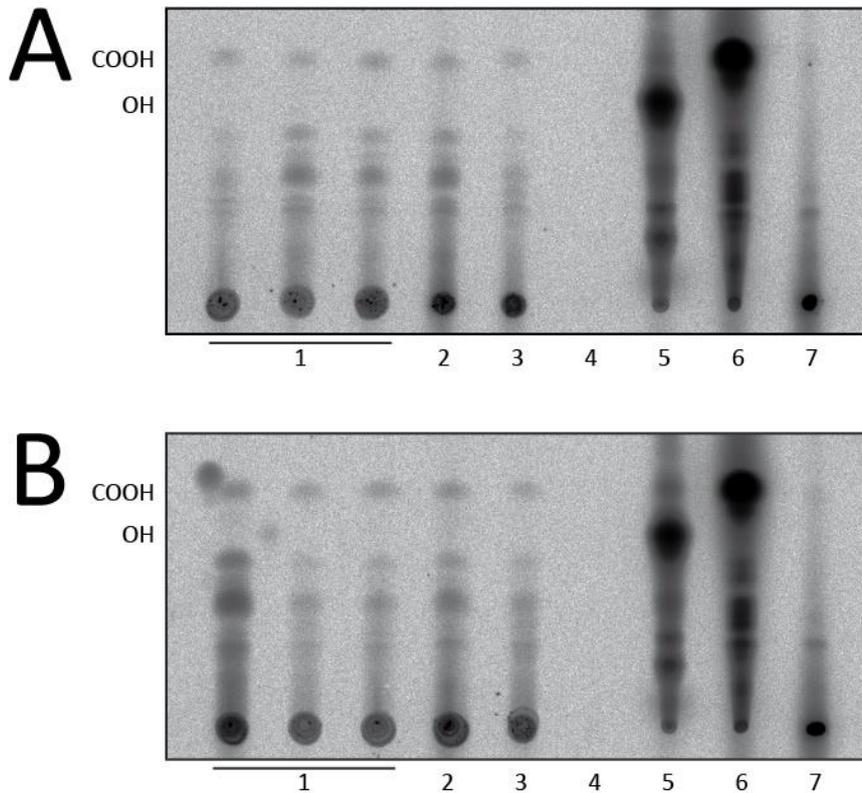
## APPENDIX



**Appendix 4: IMAC purification of N-terminal 6xHis-tagged DmFAR1.** Binding buffer: 50 mM Tris/HCl pH = 8.0, 500 mM NaCl, 10 % glycerol, 0.3 %CHAPS. Elution buffer: Binding buffer + 500 mM imidazole. Lysate = lysate of whole cells, Load = soluble fraction of whole cell lysate used for loading of the IMAC column. Fraktion 12-18 = Elution fractions. Putative position of 6xHis-DmFAR1 is indicated by the black arrow.

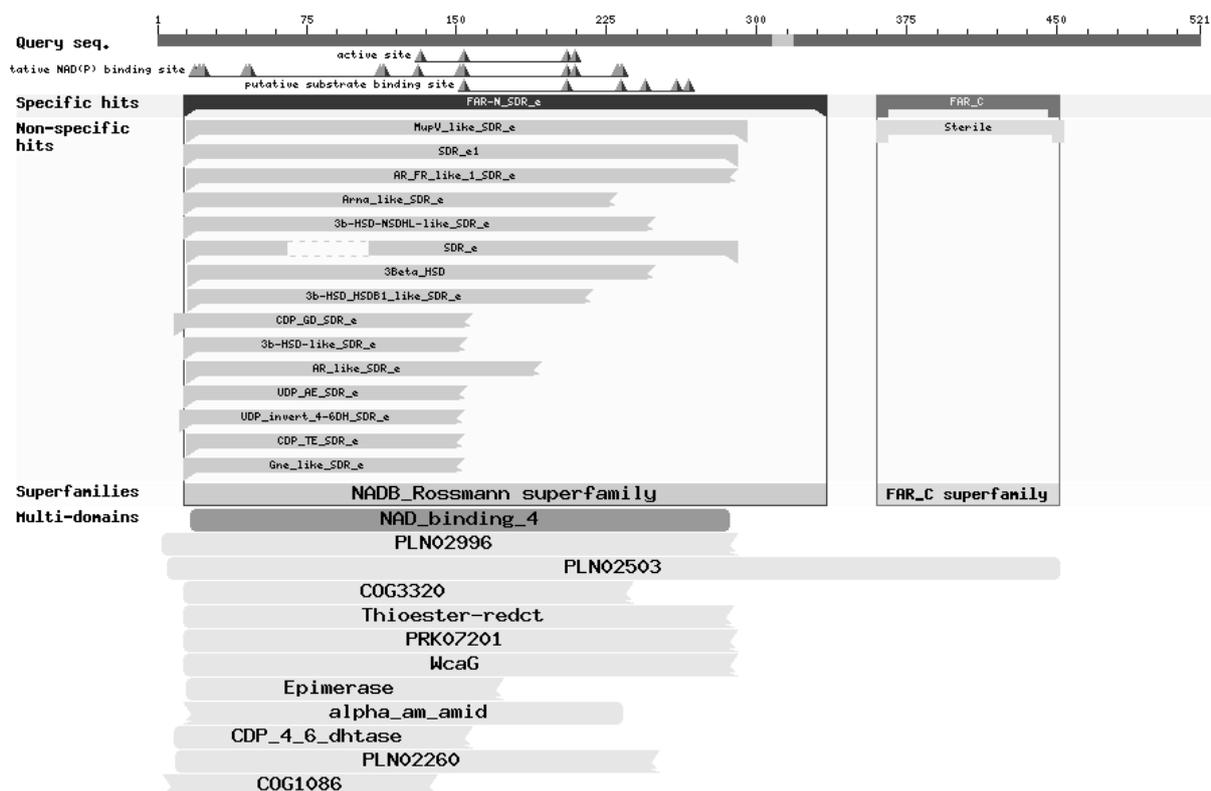


**Appendix 5: IMAC purification of C-terminal 6xHis-tagged DmFAR1.** Binding buffer: 50 mM Tris/HCl pH = 8.0, 500 mM NaCl, 10 % glycerol, 0.3 %CHAPS. Elution buffer: Binding buffer + 500 mM imidazole. Lysate = lysate of whole cells, Load = soluble fraction of whole cell lysate used for loading of the IMAC column. Fraktion 1-8 = Elution fractions. Putative position of DmFAR1-6xHis is indicated by the black arrow.

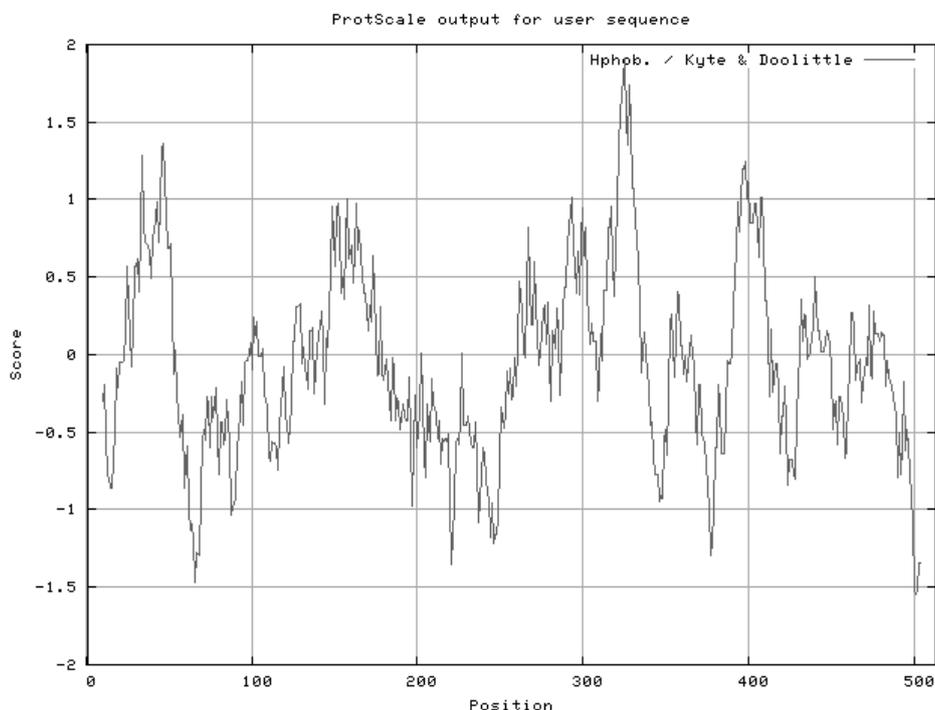


**Appendix 6: DmFAR1 *in vitro* assay with  $[^{14}\text{C}]$ -labelled substrates. A)** Assays on the basis of purified 6xHis-TF-DmFAR1. **B)** Assays on the basis of purified MBP-DmFAR1. 1 = independent samples containing 1  $\mu\text{g}$  DmFAR1, 10 mM NADPH, 50  $\mu\text{M}$   $[^{14}\text{C}]$  18:1-CoA and 3 mg/ml fatty acid free BSA in a total volume of 100  $\mu\text{l}$  TBS. Samples were incubated at 30  $^{\circ}\text{C}$  for 30 min. 2 = negative control containing heat inactivated enzyme (5 min at 95  $^{\circ}\text{C}$ ), 3 = negative control lacking NADPH, 4 = negative control lacking  $[^{14}\text{C}]$  acyl-CoA, 5 =  $[^{14}\text{C}]$  18:1-OH fatty alcohol standard, 6 =  $[^{14}\text{C}]$  18:1-COOH fatty acid standard, 7 =  $[^{14}\text{C}]$  18:1-CoA fatty acyl-CoA standard. TLC plates were developed in hexane:diethyl ether:acetic acid (65:35:1, v/v/v). Afterwards, TLC plates were placed on an imaging plate (FUJI PHOTO FILM CO., LT, Japan) for 2 days and analysed with a Fuji FLA-3000 (Raytest, Germany). Data are representative for a single experiment.

## APPENDIX

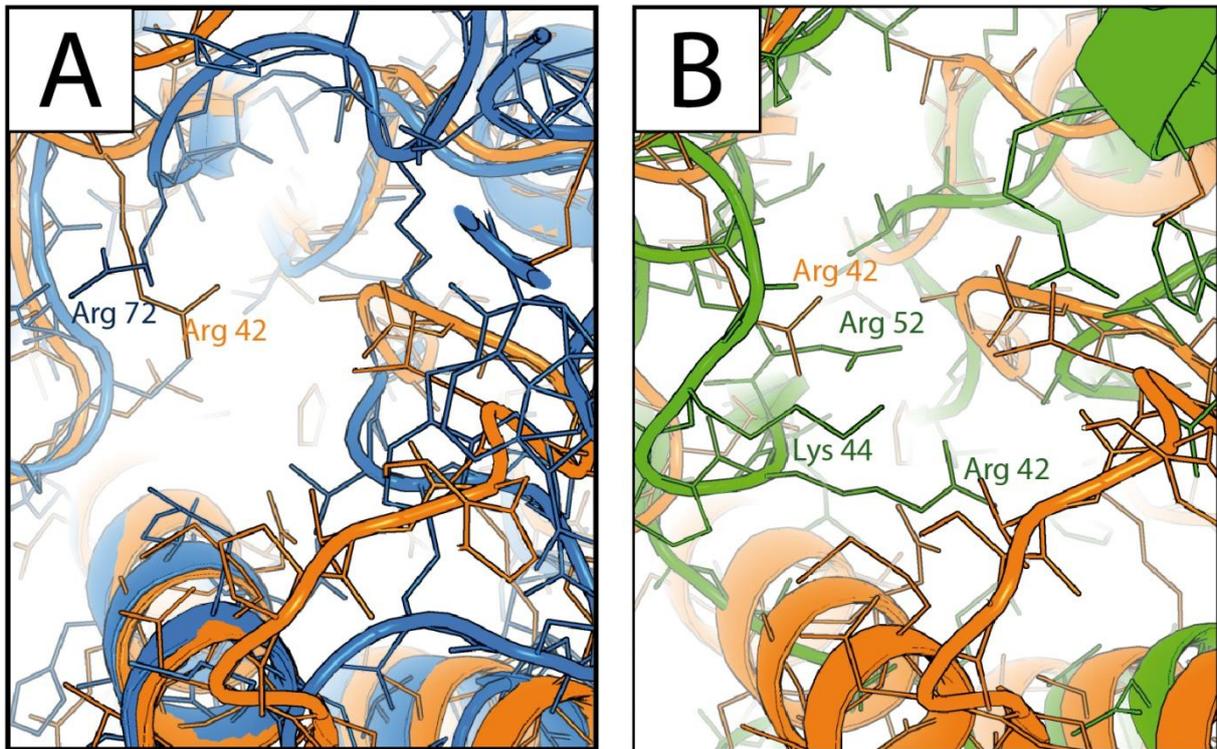


**Appendix 7: Domain prediction of TcFAR1.** TcFAR1 domain prediction (UniProt ID = D2A5A7) was done according to BLAST (Altschul et al., 1997, 2005).

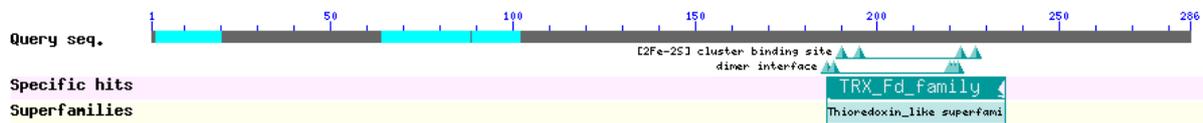


**Appendix 8: Kyte & Doolittle hydrophobicity plot for MaFAR1.** Plot was calculated using a window size of 19 according to Kyte and Doolittle (1982). Peaks with scores greater than 1.6 indicate possible transmembrane regions.

## APPENDIX



**Appendix 9: Models of DmFAR1 and EgFAR1 aligned to the structure of Gox2253.** Models were generated with the Phyre server (Kelley and Sternberg, 2009) in the intense mode and aligned with the PyMol software (Schrödinger, 2010). Arg42 in Gox2253 was shown to have a role in NADPH binding. Residues of DmFAR1 and EgFAR1 with a potential role in NADPH binding are indicated. **A)** Model of DmFAR1 (UniProt ID = Q8MS59) aligned to the structure of Gox2253 (UniProt ID = Q5FNR0). **B)** Model of EgFAR1 (UniProt ID = D7PN08) aligned to the structure of Gox2253.



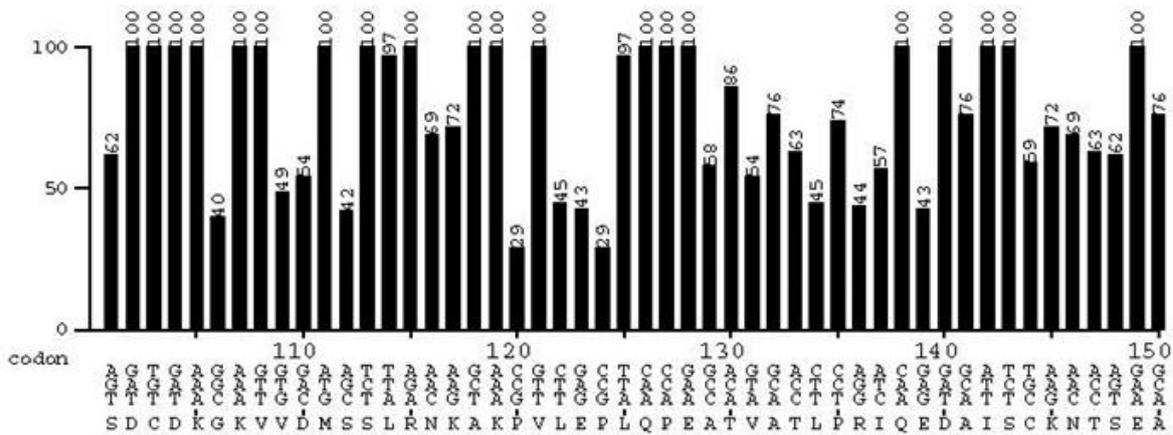
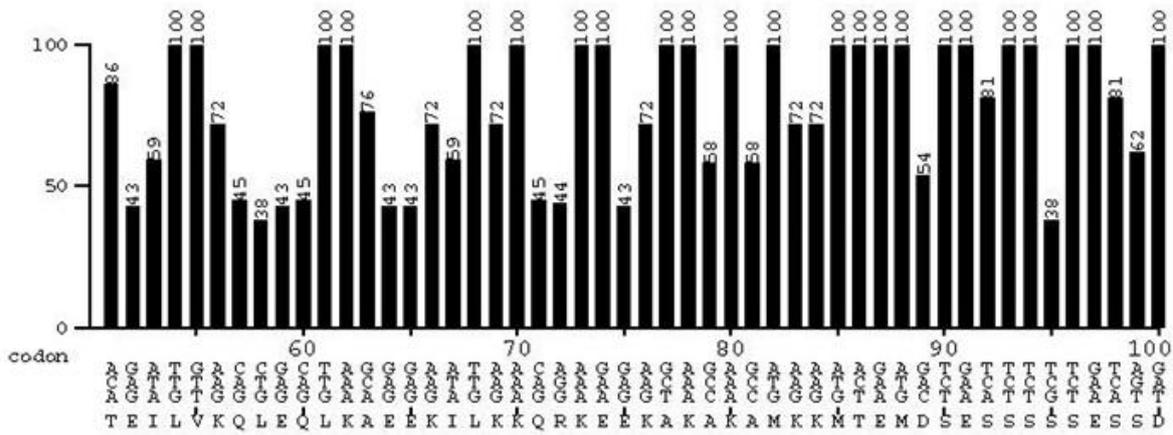
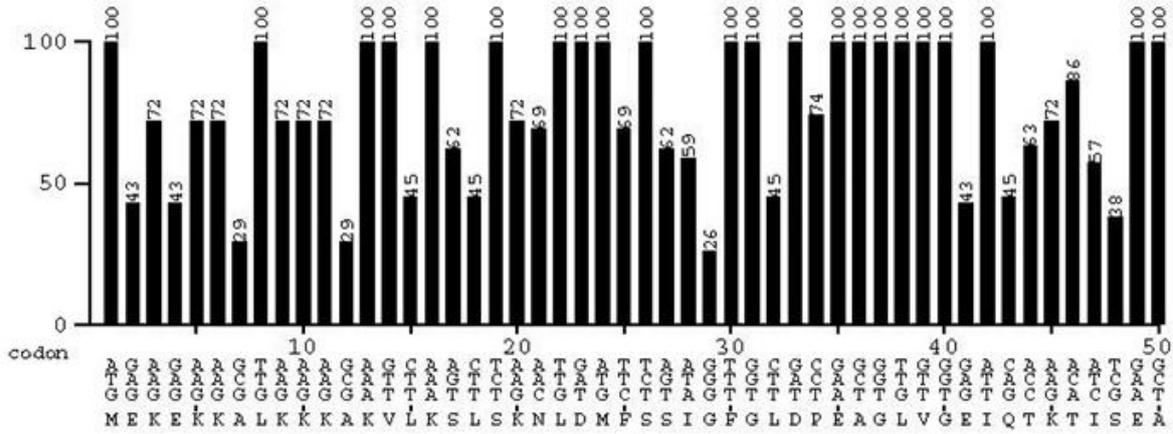
**Appendix 10: Domain structure prediction of AtDGAT3.** AtDGAT3 domain prediction (UniProt ID = Q9C5W0) was done according to BLAST (Altschul et al., 1997, 2005).

APPENDIX

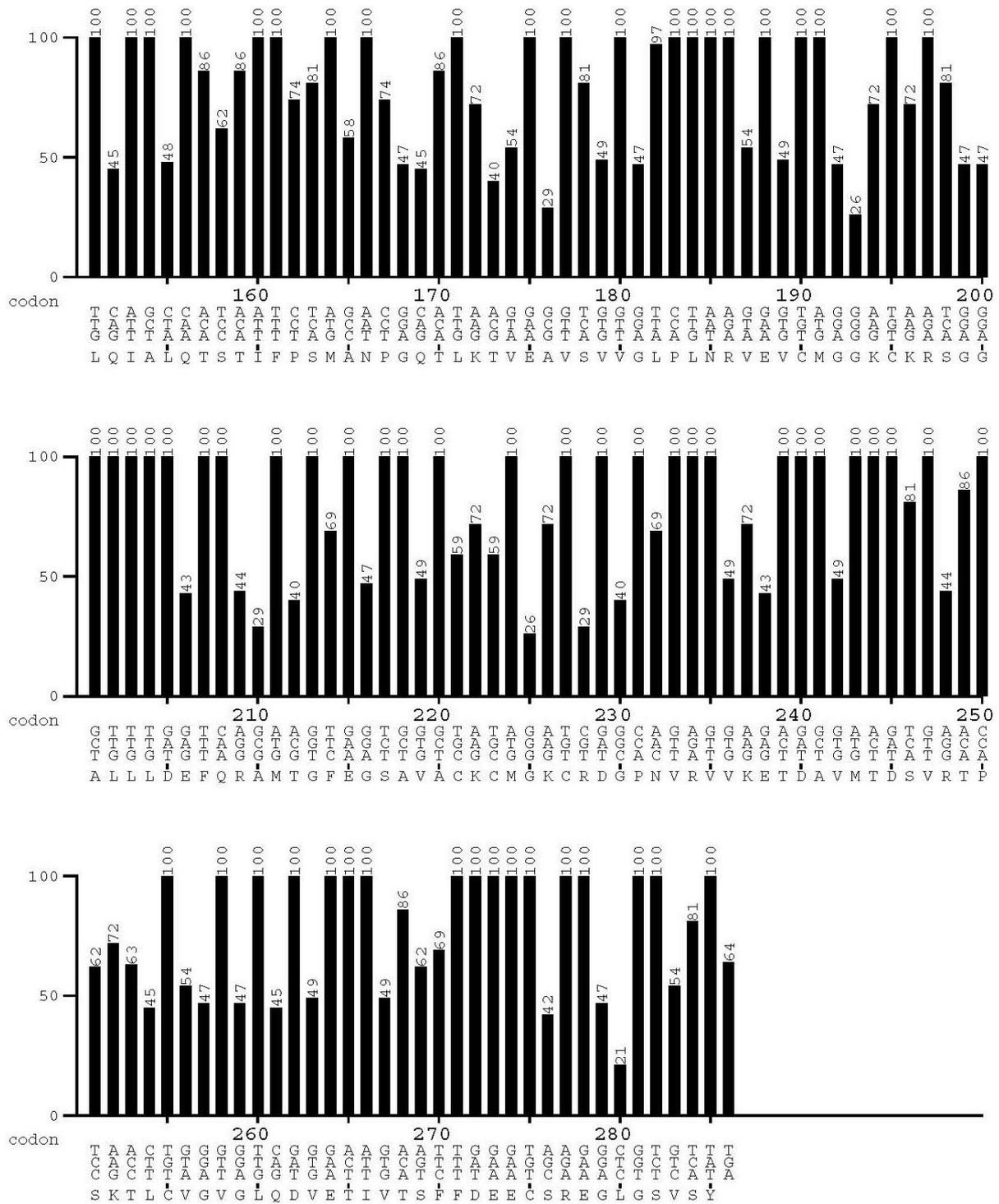
AtDGAT3  
 sequence derived from Arabidopsis\_thaliana

Codontable:  
 Saccharomyces\_cerevisiae

Ordinate (y-axis): relative adaptiveness <20% <10%

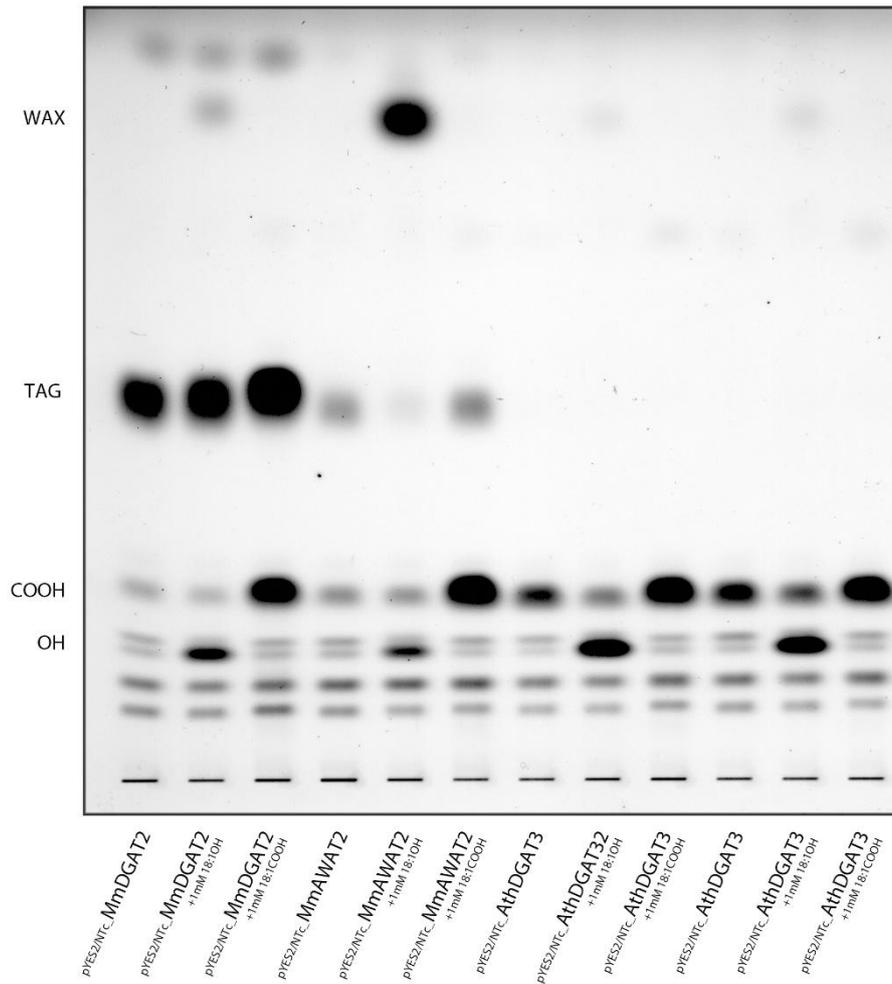


APPENDIX



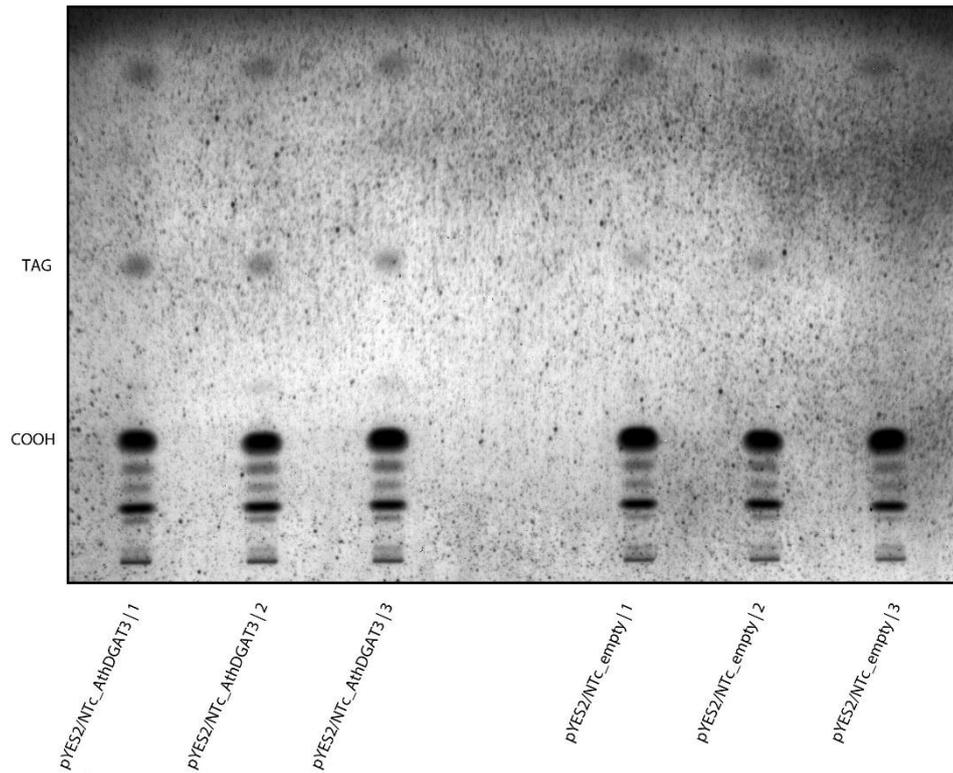
**Appendix 11: Codon usage frequency values of AtDGAT3 coding sequence in *S. cerevisiae*.** Values were determined using the graphical codon usage analyser online tool (Fuhrmann et al., 2004). Originating organism = *Arabidopsis thaliana*, Codon usage table to apply = *Saccharomyces cerevisiae*.

## APPENDIX



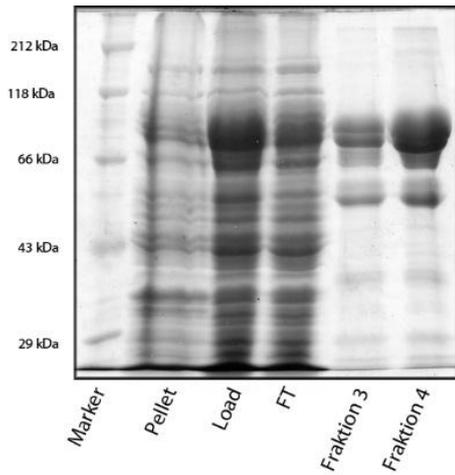
**Appendix 12: TLC analysis of lipid extracts from *S. cerevisiae* strain H1246 expressing AtDGAT3.** Samples of MmDGAT2 and MmAWAT2 serve as positive controls. 20 ml cultures of H1246 expression strains were inoculated with a respective overnight culture to  $OD_{600} = 0.05$ . 1 mM 18:1-OH or 1 mM 18:1-COOH were added to the media of the cultures depicted on the picture (see above). After 5 days of growth, 20  $OD_{600}$ -units were harvested. After centrifugation, 1 ml methanol was added to the cell pellets together with glass beads. The samples were vortexed for 15 min at RT. Subsequently, 2 ml hexane were added, samples were vortexed for another 15 min at RT. The upper phase was removed, evaporated under  $N_2$  and resolved in 200  $\mu$ l of hexane. 10  $\mu$ l of each sample were loaded on a silica TLC plate and developed in hexane:diethyl ether:acetic acid (80:20:1, v/v/v).

## APPENDIX



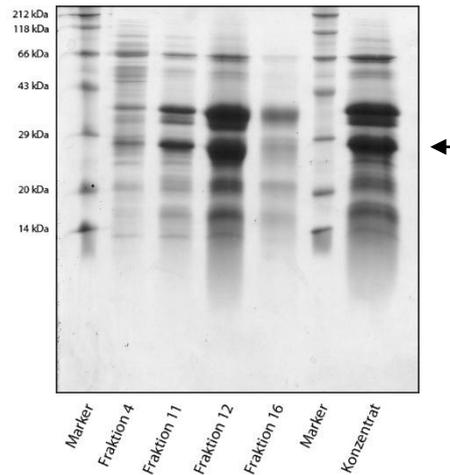
**Appendix 13: *In vitro* activity of AtDGAT3.** TLC analysis of lipid extracts from *S. cerevisiae* strain W303 expressing AtDGAT3. The constructs pYES2/NTC\_AthDGAT3 and pYES2/NT\_empty were transformed into *S. cerevisiae* W303. 10 ml cultures of W303 expression strains were inoculated with a respective overnight culture to  $OD_{600} = 0.05$ . After 5 days of growth, a culture corresponding to 20  $OD_{600}$ -units was harvested via centrifugation. After centrifugation, 1 ml methanol was added to the cell pellets together with glass beads. The samples were vortexed for 15 min at RT. Subsequently, 2 ml hexane was added, samples were vortexed for another 15 min at RT. The upper phase was removed, evaporated under  $N_2$  and resolved in 200  $\mu$ l of hexane. 50  $\mu$ l of each sample were loaded on a silica TLC plate, which was developed in hexane:diethyl ether:acetic acid (80:20:1, v/v/v). Data are representative for a single experiment and shows extracts from three independent clones for each construct.

## APPENDIX



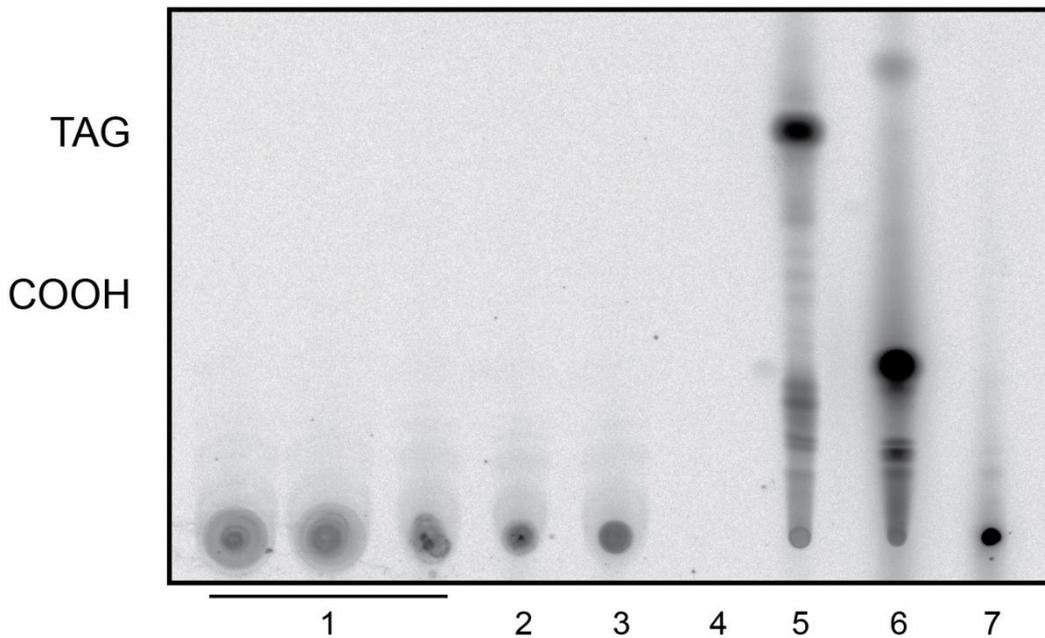
### Appendix 14: Purification of 6xHis-TF-AtDGAT3 via IMAC.

Binding buffer: TBS + 10 % glycerol + 0.3 % CHAPS (w/v). Elution buffer: Binding buffer + 500 mM imidazole. Pellet = insoluble fraction of *E. coli* cell lysate, Load = soluble fraction of *E. coli* cell lysate used for loading of the IMAC column. Fraktion 3 and 4 = elution fractions. Data is representative for a single purification.



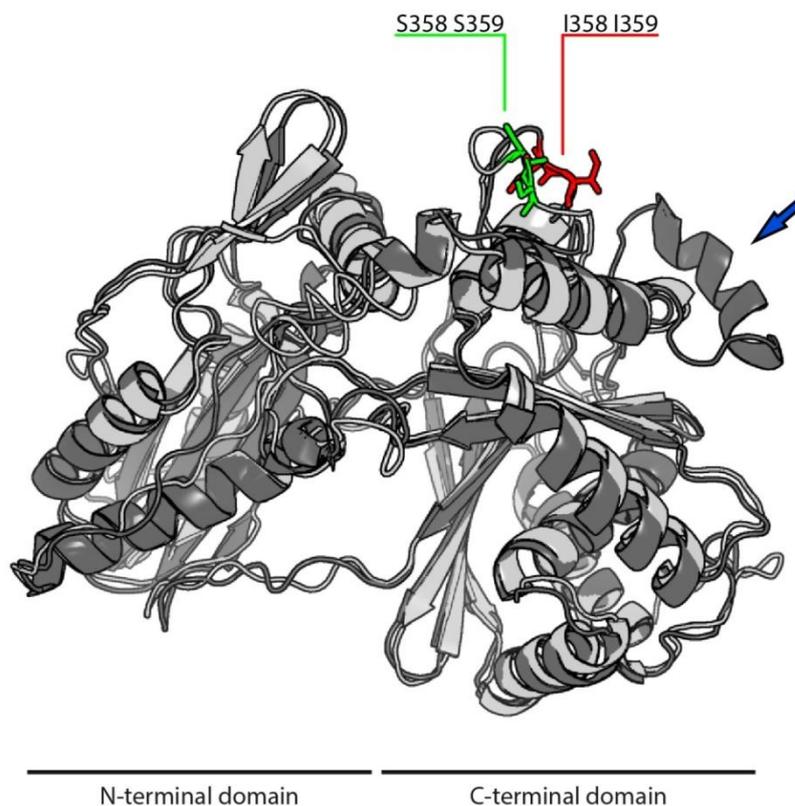
### Appendix 15: Purification of 6xHis-AtDGAT3 via IMAC.

Binding buffer: TBS + 10 % glycerol + 0.3 % CHAPS (w/v). Elution buffer: Binding buffer + 500 mM imidazole. Fraktion 4 = wash fraction upon washing of the column with 5 % elution buffer. Fraktion 11-16 = elution fractions. Konzentrat = concentrated fractions 11-16.



**Appendix 16: *In vitro* activity of AtDGAT3.** Assays were done with 5  $\mu$ g of purified 6xHis-AtDGAT3, 30  $\mu$ M 1,2-dilinoleoylglycerol and 10  $\mu$ M [ $^{14}$ C] linoleoyl-CoA (18:2-CoA) in a total volume of 100  $\mu$ l 30 mM HEPES/HCl pH = 7.5, 500 mM NaCl, 10 % glycerol, 30 mM imidazole. Samples were incubated at 30  $^{\circ}$ C for 30 min. **1** = three independent samples composed as described above. **2** = negative control lacking DAG, **3** = negative control containing heat inactivated enzyme (5 min at 95  $^{\circ}$ C), **4** = negative control lacking [ $^{14}$ C] linoleoyl-CoA, **5** = [ $^{14}$ C] TAG standard ([ $^{14}$ C] triolein), **6** = [ $^{14}$ C] fatty acid standard ([ $^{14}$ C] 18:1-COOH), **7** = [ $^{14}$ C] fatty acyl-CoA standard ([ $^{14}$ C] 18:2-CoA). The TLC plate was developed in hexane:diethyl ether:acidic acid (80:20:2, v/v/v) and exposed to an image plate for 36 h afterwards. Image plate was analysed with a Fuji FLA-3000 (Raytest, Germany). Data are representative for a single experiment under the conditions described above. Furthermore, image is representative for similar single experiments with 60  $\mu$ M and 200  $\mu$ M of 1,2-dilinoleoylglycerol, respectively.

## APPENDIX



**Appendix 17: Overlay of models of AbWSD1 and AbWSD1 I358S I359S.** Both models were generated with Phyre (Kelley and Sternberg, 2009) and aligned with PyMol (Schrödinger, 2010). The dark grey structure shows the model for wild type AbWSD1, whereas the light grey structure shows the model generated for the AbWSD1 I358S I359S variant. The blue arrow indicates an  $\alpha$ -helical part of the structure, which is an unordered loop in the model of the AbWSD1 I358S I359S variant.

ASTRA 5.3.4 Summary Report for 05022014\_MALS\_S200\_200ug.vaf

Experiment name: C:\Documents and Settings\admin\_ms\Desktop\Tom\Felix\05022014\_MALS\_S200\_200ug.vaf Sample: BSA () Processing Operator: admin\_ms Collection Operator: admin\_ms **Collection Astra Version:** 5 . 3 . 4 . 1 4

### CONFIGURATION

Light scattering instrument: miniDAWN TREOS Cell type: Fused Silica  
 Laser wavelength: 658.0 nm  
 Calibration constant: 4.7260e-5 1/(V cm)  
 RI Instrument: n/a UV Instrument: Generic UV instrument Solvent: water Refractive index: 1.331 Flow rate: 0.500 ml/min

### PROCESSING

Processing time: Wednesday February 05, 2014 03:18 PM W. Europe Standard Time Collection time: Wednesday February 05, 2014 12:55 PM W. Europe Standard Time Detectors used: 2 3 Concentration detector: UV Mass results fitting: none (fit degree: n/a) Radius results fitting: none (fit degree: n/a)

	Peak 1	Peak 2	Peak 3	Peak 4
Peak limits (min)	31.095 - 31.775	42.586 - 44.023	54.305 - 56.044	63.453 - 64.587
dn/dc (mL/g)	0.185	0.185	0.185	0.185
A <sub>2</sub> (mol mL/g <sup>2</sup> )	0.000	0.000	0.000	0.000
UV ext. (mL/(g cm))	4.984e+2	4.984e+2	4.984e+2	4.984e+2

APPENDIX

Model	Zimm	Zimm	Zimm	Zimm
Fit degree	1	1	1	1
Injected mass (g)	1.0000e-4	1.0000e-4	1.0000e-4	1.0000e-4
Calc. mass (g)	8.5982e-6	2.6804e-5	1.7469e-5	1.4059e-5
Mass Recovery	8.5982 %	26.8043 %	17.4693 %	14.0590 %
Mass Fraction	12.8464 %	40.0479 %	26.1005 %	21.0052 %

RESULTS

	Peak 1	Peak 2	Peak 3	Peak 4
Polydispersity				
Mw/Mn	1.004(0.2%)	1.000(0.6%)	1.001(5%)	1.030(37%)
Mz/Mn	1.008(0.3%)	1.001(1%)	1.002(8%)	1.063(63%)
Molar mass moments (g/mol)				
Mn	1.689e+6(0.1%)	2.974e+5(0.5%)	7.493e+4(3%)	9.039e+3(26%)
Mp	1.645e+6(0.1%)	2.971e+5(0.4%)	7.193e+4(3%)	8.256e+3(28%)
Mv	n/a	n/a	n/a	n/a
Mw	1.695e+6(0.1%)	2.975e+5(0.5%)	7.500e+4(3%)	9.314e+3(26%)
Mz	1.701e+6(0.3%)	2.976e+5(1%)	7.508e+4(8%)	9.612e+3(57%)
rms radius moments (nm)				
Rn	24.1(0.7%)	9.4(16%)	13.5(61%)	26.4(123%)
Rw	24.0(0.7%)	9.4(16%)	13.5(60%)	27.7(111%)
Rz	24.0(0.7%)	9.4(16%)	13.6(59%)	29.0(100%)

**Appendix 18: MALS measurements of protein species of an IMAC purified TF-AbWSD1-sample eluting from SEC.** Row containing the actual mass weights of each species is coloured in red.

## APPENDIX

CLUSTAL O(1.2.1) multiple sequence alignment

```

DGAT2_MOUSE MKTLIAAYSGVLRGERRAEAAARSENKNKGSALSREGSGRWGTGSSILSALQDIFSVTWLN 60
AWAT2_MOUSE -----MFWPT 5
                                     : * .

DGAT2_MOUSE RSKVEKQLQVLSVLQWVLSFLVLGVACSVILM-YTFCTDCWLIAVLYFTWLAFDWNTPKK 119
AWAT2_MOUSE KKDLKTAMEVFALFQWALSALVIVTVIIIVNLYLVVFTSYWPVTVLMLTWLAFDWKTPER 65
      ..... :*:::*:** ** : : : .. * . * :** :*****:***:

DGAT2_MOUSE GGRRSQWVRNWAVWRYFRDYFPIQLVKTHNLLTTRNYIFGYHPHGIMGLGAFCNFSTEAT 179
AWAT2_MOUSE GGRRFTCVRKWRLWKHYSYFPLKMKVTKDISPDRNYILVCHPHGLMAHSCFGHFATDTT 125
      ****  **:* :*::: *****:*****:  ****:  *****:..* .*:***

DGAT2_MOUSE EVSKKFPGIRPYLATLAGNFRMPVLREYLMSSGGICPVNRDTIDYLLSKNGSGNAIIIVVG 239
AWAT2_MOUSE GFSKTFPGITPYMLTLGAFFWVPFLRDYVMSTGSCSVSRSSMDFLLTQKGTGNMLVVVVG 185
      .**.**** ** : **.. * :*.**:*:** * * *.*.:*:**::*:** :***

DGAT2_MOUSE GAAESLSSMPGKNAVTLKNRKGFVKLALRHGADLVPTYSFGENEVYKQVIFEEGSWGRWV 299
AWAT2_MOUSE GLAECRYSTPGSTTLFLKQRQGFVRTALKHGVSLIPAYAFGETDLYDQHIFTTPGGFVNRF 245
      * ** . * **..: **:*:**: **:*..*:*:*:**:*. * ** *.: . .

DGAT2_MOUSE QKKFQKYIGFAPCIFHGRGLFSSDTWGLVPYSKPITTVVGEPITVPKLEHPTQKDIDLYH 359
AWAT2_MOUSE QKWFQKMVHIYPCAFYGRGLT-KNSWGLLPYSQPVTTTVVGEPLPLPKIENPSEEIVAKYH 304
      ** *** : : ** *:*:** .:***:**:*:*****: **:*. *::: : **

DGAT2_MOUSE AMYMEALVKLFDNHKTKFGLPETEVLEVN 388
AWAT2_MOUSE TLYIDALRKLFDQHKTKFGISETQELVIV 333
      :*:** *****:*****: **: * :
  
```

**Appendix 19: Amino acid sequence alignment of MmDGAT2 and MmAWAT2.** Alignment was done using Clustal Omega, as provided by the UniProt network (The UniProt Consortium, 2014). Predicted (in case of MmAWAT2) and experimentally elucidated (in case of MmDGAT2, Stone et al., 2006)) membrane topologies are indicated by blue boxes. The putative NLBS of MmDGAT2 (Au-Young and Fielding, 1992) and the respective sequence in MmAWAT2 are indicated by green boxes. The conserved pFFRR motif (where p stands for a positively charged residue) in both enzymes is underlined, while the conserved YFP motif is printed in bold letters and the HPHG active motif (Stone et al., 2006) is printed in red letters. Asterisks indicate identical amino acid positions, while colons indicate similar amino acid positions.

APPENDIX

MmAWAT2 and MmAWAT2 derived single amino acid exchange variants fed with 18:1-OH

Chain	MmAWAT2	E14Q	A25F	T30A T31A	N36A	A25F N36R	T30A T31A N36R	N36R	N36L	N36K	N36W	N36R F42R S44D	L39-	F42C	S44D	S44D E64K	C72W	C106Y	MmDGAT2
16:1	0.54 ± 0.03	0.45 ± 0.04	0.48 ± 0.02	0.58 ± 0.02	0.49 ± 0.01	0.38 ± 0.02	0.40 ± 0.03	0.32 ± 0.05	0.56 ± 0.02	0.45 ± 0.05	0.50 ± 0.01	0.38 ± 0.03	0.57 ± 0.02	0.50 ± 0.04	0.48 ± 0.01	0.55 ± 0.06	0.52 ± 0.00	0.54 ± 0.06	0.26 ± 0.02
16:0	0.16 ± 0.05	0.21 ± 0.03	0.15 ± 0.02	0.17 ± 0.01	0.16 ± 0.01	0.11 ± 0.01	0.11 ± 0.01	0.14 ± 0.02	0.18 ± 0.01	0.16 ± 0.01	0.17 ± 0.01	0.15 ± 0.02	0.17 ± 0.02	0.18 ± 0.03	0.19 ± 0.00	0.17 ± 0.03	0.19 ± 0.04	0.17 ± 0.01	0.24 ± 0.02
18:1	0.24 ± 0.04	0.27 ± 0.01	0.27 ± 0.03	0.20 ± 0.01	0.26 ± 0.04	0.28 ± 0.02	0.29 ± 0.02	0.24 ± 0.05	0.21 ± 0.02	0.27 ± 0.03	0.23 ± 0.01	0.23 ± 0.05	0.22 ± 0.00	0.21 ± 0.02	0.27 ± 0.00	0.22 ± 0.03	0.22 ± 0.02	0.20 ± 0.04	0.33 ± 0.03
18:0	0.05 ± 0.01	0.06 ± 0.01	0.07 ± 0.01	0.05 ± 0.01	0.06 ± 0.01	0.12 ± 0.01	0.10 ± 0.00	0.12 ± 0.02	0.05 ± 0.00	0.10 ± 0.03	0.08 ± 0.01	0.12 ± 0.02	0.05 ± 0.00	0.07 ± 0.01	0.05 ± 0.00	0.05 ± 0.01	0.06 ± 0.01	0.07 ± 0.02	0.17 ± 0.02
>C18	0.01 ± 0.01	0.00 ± 0.00	0.03 ± 0.02	0.00 ± 0.00	0.03 ± 0.03	0.12 ± 0.04	0.10 ± 0.02	0.14 ± 0.06	0.00 ± 0.00	0.01 ± 0.01	0.02 ± 0.01	0.09 ± 0.02	0.00 ± 0.00	0.03 ± 0.02	0.00 ± 0.00	0.00 ± 0.00	0.01 ± 0.00	0.01 ± 0.01	0.00 ± 0.00

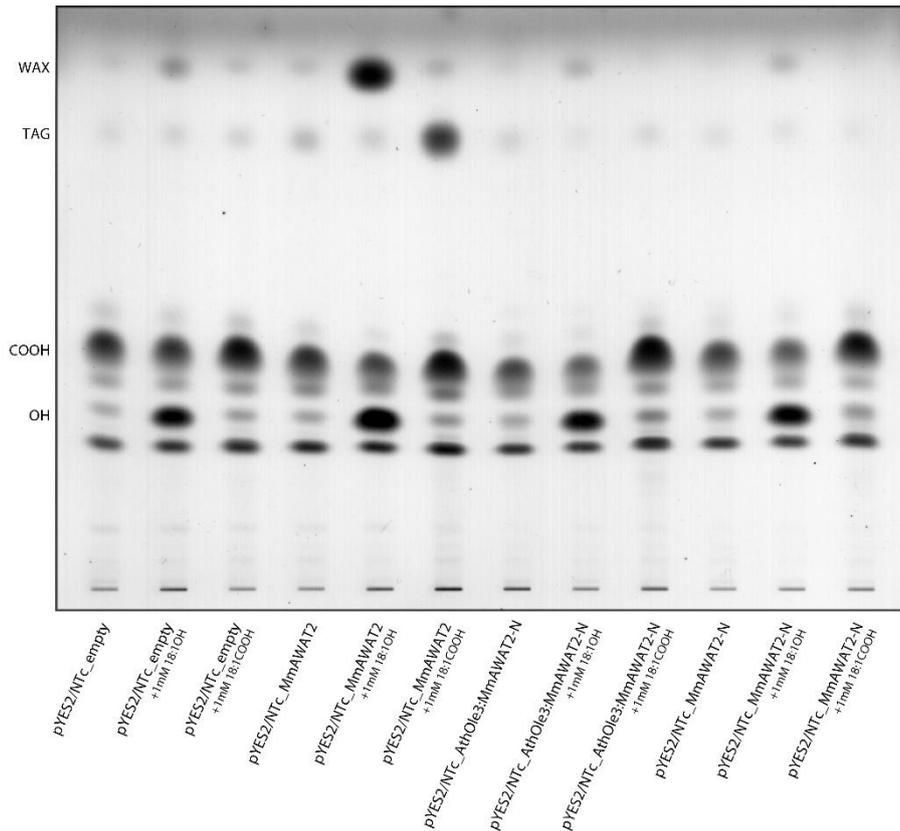
**Appendix 20: Acyl chain composition of WEs synthesised by MmAWAT2 and MmAWAT2-derived variants as well as MmDGAT2 upon feeding of 18:1-OH.** Red boxes indicate values which are significantly lower than respective MmAWAT2-values, whereas green boxes indicate significantly higher values according to a student's t-test ( $p \leq 0.5$ ). Data represent mean and standard deviation derived from at least three samples from independent cultures.

APPENDIX

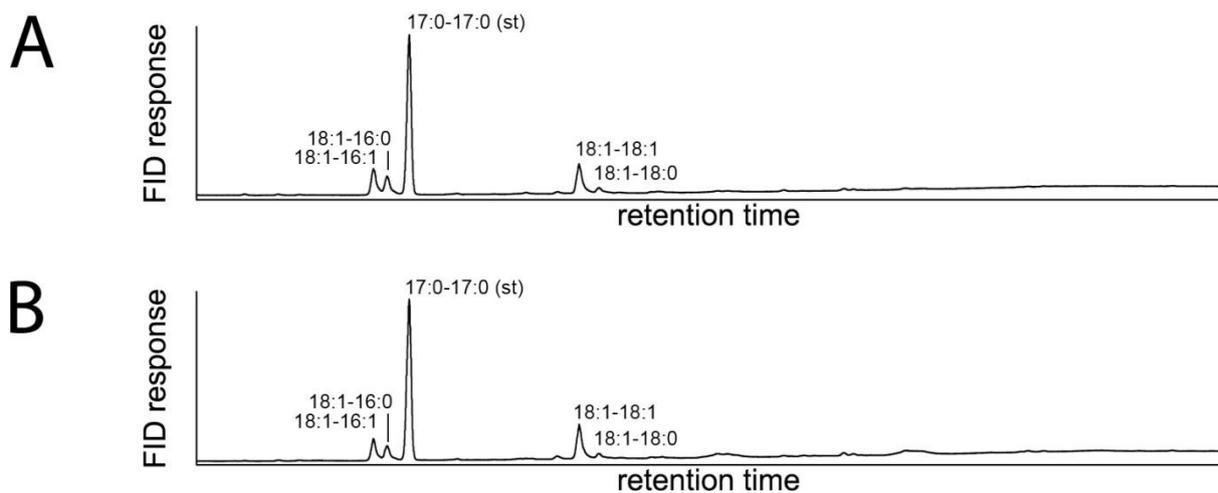
MmAWAT2 and MmAWAT2-derived variants fed with 16:0-OH									
Chain	MmAWAT2	A25F	T30A T31A	A25F N36R	N36R	V1	V2	V5	MmDGAT2
16:1	0.57 ± 0.03	0.58 ± 0.01	0.63 ± 0.02	0.28 ± 0.01	0.27 ± 0.01	0.50 ± 0.05	0.35 ± 0.04	0.41 ± 0.05	0.35 ± 0.01
16:0	0.14 ± 0.01	0.09 ± 0.01	0.11 ± 0.02	0.09 ± 0.00	0.09 ± 0.01	0.14 ± 0.02	0.08 ± 0.00	0.08 ± 0.02	0.24 ± 0.01
18:1	0.24 ± 0.02	0.27 ± 0.01	0.22 ± 0.01	0.25 ± 0.00	0.22 ± 0.03	0.30 ± 0.03	0.27 ± 0.01	0.32 ± 0.05	0.29 ± 0.01
18:0	0.04 ± 0.00	0.04 ± 0.01	0.03 ± 0.01	0.11 ± 0.00	0.12 ± 0.01	0.04 ± 0.01	0.08 ± 0.00	0.06 ± 0.01	0.12 ± 0.01
>C18	0.00 ± 0.00	0.02 ± 0.01	0.00 ± 0.00	0.27 ± 0.00	0.30 ± 0.02	0.02 ± 0.01	0.21 ± 0.03	0.13 ± 0.03	0.00 ± 0.00

**Appendix 21: Acyl chain composition of WEs synthesised by MmAWAT2 and MmAWAT2-derived variants as well as MmDGAT2 upon feeding of 16:0-OH.** Red boxes indicate values which are significantly lower than respective MmAWAT2 values, whereas green boxes indicate significantly higher values as calculated using a student's t-test ( $p \leq 0.5$ ). Data represent mean and standard deviation derived from at least three samples from independent cultures.

APPENDIX

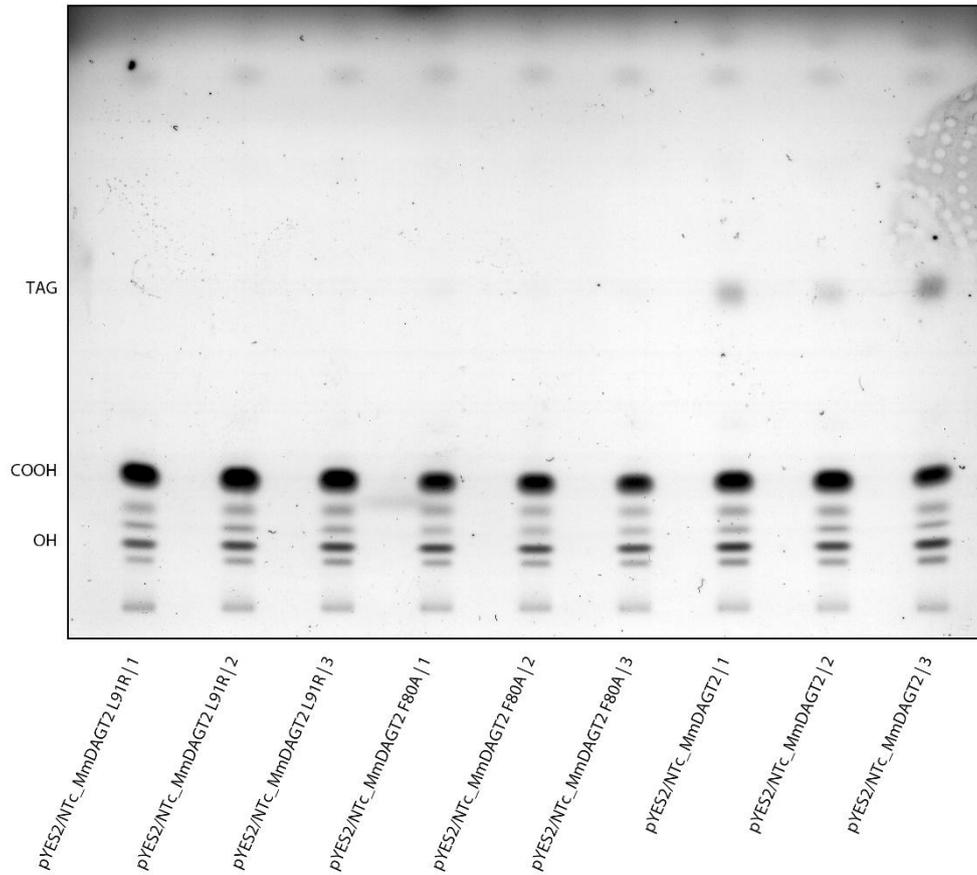


**Appendix 22: TLC analysis of lipid extracts of *S. cerevisiae* H1246 expressing either MmAWAT2, a N-terminal truncated version of MmAWAT2 (MmAWAT2ΔN) as well as a fusion protein from oleosin 3 from *Arabidopsis thaliana* and MmAWAT2ΔN.** 20 ml-cultures of H1246 expression strains were inoculated with a respective overnight culture to an  $OD_{600} = 0.05$ . The cultures were supplemented with either 1 mM 18:1-OH or 1 mM 18:1-COOH, as depicted in the figure. After 5 days of growth, 20  $OD_{600}$ -units of culture were harvested by centrifugation. Afterwards, 1 ml methanol was added to the cell pellets together with glass beads. The samples were vortexed for 15 min at RT. Subsequently, 2 ml hexane were added and samples were vortexed for another 15 min at RT. The upper phase was removed, evaporated under a stream of  $N_2$  and resolved in 200  $\mu$ l of hexane. 20  $\mu$ l of each sample were applied on a silica TLC plate. The TLC plate was developed in hexane:diethyl ether:acetic acid (80:20:1, v/v/v). Data are representative for a single experiment in case of MmAWAT2ΔN and for samples from three independent cultures in case of the AtOle3-MmAWAT2ΔN fusion protein.



**Appendix 23: *In vivo* activity of MmAWAT2ΔN upon expression in *S. cerevisiae* H1246.** **A)** GC-FID analysis of lipid extracts from *S. cerevisiae* H1246 expressing pYES2/NT empty as an empty vector control. **B)** GC-FID analysis of lipid extracts from *S. cerevisiae* H1246 expressing a truncated version of MmAWAT2 missing the predicted two N-terminal transmembrane domains (MmAWAT2ΔN). Equal amounts of heptadecanoyl-heptadecanoate (17:0-17:0) were added as an internal standard to all samples. All cultures were fed with 1 mM 18:1-OH. Data are representative for samples derived from three independent cultures in both cases.

## APPENDIX



**Appendix 24: TLC of lipid extracts from yeast cultures expressing MmDAGT2 and the two variants MmDAGT2 F80A and MmDAGT2 L91R.** 20ml-cultures of *S. cerevisiae* H1246 expression strains were inoculated with a respective overnight culture to an  $OD_{600}=0.05$ . After 5 days of growth, cells corresponding to 20  $OD_{600}$ -units were harvested. After centrifugation, 1 ml methanol was added to the cell pellets together with glass beads. The samples were vortexed for 15 min at RT. Subsequently, 2 ml hexane were added, samples were vortexed for another 15 min at RT. The upper phase was removed, evaporated under a stream of  $N_2$  and resolved in 200  $\mu$ l of Hexane. 25  $\mu$ l of each sample were loaded on a silica TLC plate. The image is representative for a single experiment and shows extracts from three independent cultures for each construct.

## 12 ACKNOWLEDGEMENTS

The work presented here was done in a time course of three years in the Department of Plant Biochemistry of the University of Göttingen under the supervision of Prof. Dr. Ivo Feußner. I am deeply thankful for his support over this time. He contributed important ideas to the project and actively supported my own ones. I am thankful for the many inspiring discussions and his interest in the progress of this project.

I furthermore experienced important guidance from Dr. Florian Brodhun, who accompanied my work especially at its beginning. Scientific discussions with him were and still are fruitful, as well as our everyday conversations about other important topics. I want to express my gratefulness for that.

Another essential factor during this work was the expertise and helpfulness of Dr. Ellen Hornung. She helped me through a lot of problems accompanied with molecular biology. Apart from the multitude of questions she answered during the years, her personality was an enrichment for the everyday laboratory life. And her cakes are exquisite.

Identification of the VLC WEs in the course of the MmAWAT2 project wouldn't have been possible without the expertise of Dr. Tim Iven in the field of WE analyses. Furthermore, Pablo Tarazona contributed his expertise to the analyses of TAG species. I hereby want to thank both of them for their help.

I also want to thank Felix Lambrecht, Laura Mengel and Nele Vollersen, who contributed to my studies in the course of their bachelor theses. While Felix did important work for AtDGAT3 and AbWSD1, Laura and Nele supported my studies on MaFAR1 and DmFAR1, respectively.

Moreover, I am thankful for the support I received from Sabine Freitag and Pia Meyer in the course of all my experiments related to GC and HPLC analyses. Their competent guidance and instructions in respect to operation of the numerous machines probably saved Prof. Feußner a lot of money and Gerhard Marder a lot of work.

Still, many things wouldn't have been possible without the tireless efforts of Gerhard Marder to repair and service our technical equipment. I want to thank him for his mechanical help and his improvising genius, which would probably even leave Angus MacGyver astonished.

The working atmosphere in the group was exceptionally nice and relaxed over the whole time I have been there. I want to thank all members of the group, past and present, for contributing to this nice atmosphere and making the institute a place I liked to go to.

Besides the whole Department of Plant Biochemistry, several collaborators supported this work. The expertise, confidence and patience of Dr. Karin Kühnel from the Department of Neurobiology of the Max Planck Institute for Biophysical Chemistry was the key to the successful crystallisation of AbWSD1. I am deeply grateful for her time, her help and her valuable advises during the last three years.

Furthermore, I would like to thank the group of Prof. Dr. Reinhard Schuh (Department of Molecular Developmental Biology, Max Planck Institute for Biophysical Chemistry, Göttingen) for the successful collaboration in the DmFAR1 project and the group of Prof. Dr. Markus Maniak (Department for Cell Biology, University of Kassel) for the successful collaboration in a project on a DGAT2 from *Dictyostelium discoideum*.

I also want to thank Dr. Blaga Popova from the Department of Molecular Microbiology & Genetics of the University of Göttingen and Dr. Nickels Jensen from the Department of NanoBiophotonics of the Max Planck Institute for Biophysical Chemistry, Göttingen for their contribution to my studies on

## ACKNOWLEDGEMENTS

randomly mutated WSs. Although respective experiments unfortunately didn't make it into this thesis, the collaboration was highly inspiring and an important enrichment for my scientific horizon. I am furthermore grateful for the supervision and guidance of all members of my thesis committee. In the course of our regular meetings, Prof. Dr. Ivo Feußner, Prof. Dr. Jörg Stülke and Dr. Karin Kühnel supplied important ideas and constructive feedback for my work.

This thesis was carefully read by Prof. Dr. Ivo Feußner, Prof. Dr. Michael Göttfert, Dr. Karin Kühnel as well as by Christian Koch and Julia Newie. They all gave valuable comments which supported me in the writing process. I also want to thank the Fonds der chemischen Industrie, the Göttinger Graduiertenschule für Neurowissenschaften and the European ICON project for funding of my work.

Finally, I would not have been able to write this thesis without the help and support of my family, of Susan and of my friends. I want to thank all of you for the happy times I experienced with you outside of the lab, for cheering me up in difficult situations and for standing by my side. Thank you.