

Identification of Amino Acids Conferring Chain-Length
Substrate Specificities on Fatty Alcohol-Forming
Reductases FAR5 and FAR8 from *Arabidopsis thaliana*

By

Micaëla Chacón

B.Sc., University of Ottawa, 2011

A thesis submitted to the Faculty of Graduate Studies and Postdoctoral Affairs in

partial fulfillment of the requirements for the degree of

Master of Science

in

Biology

Carleton University

Ottawa, Ontario

© 2013, Micaëla Chacón

Abstract

Fatty alcohols play a variety of biological roles in all kingdoms of life. Fatty acyl reductase (FAR) enzymes catalyze the reduction of fatty acyl-coenzyme A (CoA) or fatty acyl-acyl carrier protein (ACP) substrates to primary fatty alcohols. FAR enzymes have distinct substrate specificities with regard to chain length and degree of saturation. FAR5 (At3g44550) and FAR8 (At3g44560) from *Arabidopsis thaliana* are 85% identical at the amino acid level and are of equal length, but possess distinct specificities for 18:0 or 16:0 fatty acyl chain length, respectively. We used *Saccharomyces cerevisiae* as a heterologous expression system to assess FAR substrate specificity determinants. We identified individual amino acids that affect protein levels or 16:0-CoA versus 18:0-CoA specificity by expressing in yeast FAR5 and FAR8 domain-swap chimeras and site-specific mutants. We found that a threonine at position 347 and a serine at position 363 were important for high FAR5 and FAR8 protein accumulation in yeast, and thus are likely important for protein folding and stability. Amino acids at positions 355 and 377 were important for dictating 16:0-CoA versus 18:0-CoA chain length specificity. Simultaneously converting alanine-355 and valine-377 of FAR5 to the corresponding FAR8 residues, leucine and methionine, respectively, nearly fully converted FAR5 specificity from 18:0-CoA to 16:0-CoA. The reciprocal amino acid conversions, L355A and M377V, made in the active FAR8-S363P mutant background converted its specificity from 16:0-CoA to 18:0-CoA. This study is an important advancement in the engineering of highly active FAR proteins with desired specificities for the production of fatty alcohols with significant industrial value.

Acknowledgements

I would like to thank my thesis supervisor, Owen Rowland, for giving me the opportunity and skills to pursue a passion for plant biology, and for his great insight and help with overcoming any unexpected obstacle or forks in the road my research presented. I would also like to thank my M.Sc. graduate committee members, Dr. Gopal Subramaniam and Dr. Ashkan Golshani from Carleton University, for their support and guidance through my research. And a special thank you to my fellow lab member Ian Pulsifer for always being generous with his time and good spirits.

Statement of Contribution

Publication that includes data and text from this thesis:

Chacón, 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 specificity on fatty alcohol-forming reductases FAR5 and FAR8 from *Arabidopsis thaliana*. Journal of Biological Chemistry, <http://www.jbc.org/content/early/2013/09/04/jbc.M113.499715.long>

I, Micaëla Chacón, wrote the draft of this paper, which became the basis for parts of this thesis.

I, Micaëla Chacón, performed all the experiments and generated all the materials reported in this thesis/above manuscript with the exception of the following:

- 1) *FAR8-S363P*, *FAR5-Y238F*, *FAR5-K242I*, *FAR5-P363S*, *FAR5₁₋₂₈₃FAR8₂₈₄₋₄₉₆* *FAR8₁₋₂₈₃FAR5₂₈₄₋₄₉₆*, *FAR5₁₋₃₈₈FAR8₃₈₉₋₄₉₆*, and *FAR8R₁₋₃₈₈FAR5₃₈₉₋₄₉₆* were previously cloned into pYES2-His6x/T7 by Ashley Fournier (Rowland lab, Carleton University) and Ashley obtained preliminary yeast lipid data.
- 2) *FAR5* and *FAR8* were previously cloned into pYES2-His6x/T7 by Frances Tran (Rowland Lab, Carleton University) and Frances obtained preliminary yeast lipid data.
- 3) The *in vitro* assays with *FAR5*, *FAR5-A355L*, and *FAR5-A355L V377M* were performed by Drs. Franziska Dittrich-Domergue and Frédéric Domergue (Laboratoire de Biogenèse Membranaire, CNRS, Bordeaux, France) using the yeast strains that I provided them.
- 4) Maintenance of the GCs was performed by Ian Pulsifer (Rowland lab, Carleton University). Ian also trained me on the operation of the GCs, assisted me with GC method development, and assisted me with GC data interpretation.

Table of Contents

<i>Abstract</i>	2
<i>Acknowledgements</i>	3
<i>Statement of Contribution</i>	4
<i>Table of Contents</i>	5
<i>List of Figures</i>	7
<i>List of Tables</i>	8
<i>List of Abbreviations</i>	9
CHAPTER 1: General Introduction	11
<i>1.1 Fatty Alcohols in Nature - Structure and Occurrence</i>	11
<i>1.2 Biochemistry of Fatty Acyl Reductases</i>	13
<i>1.3 Biological Functions of FARs</i>	19
<i>1.3.1 Pheromones</i>	19
<i>1.3.2 Lubrication / Waterproofing</i>	22
<i>1.3.3 Surface barriers</i>	23
<i>1.3.4 Ether lipids</i>	28
<i>1.3.5 Energy storage</i>	29
<i>1.4 FAR Substrate Specificities</i>	34
<i>1.5 FAR Industrial Applications</i>	35
<i>1.6 Concluding Remarks</i>	36
<i>1.7 Project Rationale</i>	37
<i>1.8 Thesis Objectives</i>	38
CHAPTER 2: Materials and Methods	39
<i>2.1 Construction of yeast expression plasmids containing Arabidopsis FAR5 and FAR8 variants</i>	39
<i>2.2 Expression of FAR5 and FAR8 variants in yeast</i>	42
<i>2.3 Analysis of protein levels in yeast by protein immunoblotting (Western blots)</i>	42
<i>2.4 Analysis of fatty alcohol content of transgenic yeast by gas chromatography (GC)</i>	43
<i>2.5 Yeast microsome preparations and FAR in vitro enzyme assays</i>	44

CHAPTER 3: Results	45
<i>3.1 Amino acids of FAR5 and FAR8 important for protein stability and enzymatic activity</i>	<i>45</i>
<i>3.2 Domain swaps between FAR5 and FAR8</i>	<i>48</i>
<i>3.3 Reciprocal amino acid substitutions in FAR5 and FAR8</i>	<i>50</i>
<i>3.4 In vitro FAR assays using yeast microsomes</i>	<i>52</i>
CHAPTER 4: Discussion.....	64
CHAPTER 5: Future Directions and Concluding Remarks.....	68
<i>5.1 Future Directions.....</i>	<i>68</i>
<i>5.2 Concluding Remarks.....</i>	<i>72</i>
REFERENCES.....	74

List of Figures

Figure 1.1 Generalized structures of free primary fatty alcohols, wax esters, alkyl hydroxycinnamates, and ether lipids.....	12
Figure 1.2 One-enzyme reduction reaction performed by fatty acyl reductase (FAR) enzymes.....	17
Figure 1.3 Schematic of the structural domains of FAR proteins.....	18
Figure 3.1 FAR structural domains and protein sequence alignment of Arabidopsis FAR5 and FAR8.....	54
Figure 3.2 Gas chromatograms of internal lipids of yeast expressing Arabidopsis FAR5, FAR8, or FAR8-S363P.....	55
Figure 3.3 Amino acids important for FAR5 and FAR8 enzyme stability and activity...	56
Figure 3.4 Fatty alcohol production by FAR5, FAR8, FAR8R (FAR8-S363P) and <i>Euglena gracilis</i> FAR (EgFAR) expressed in yeast for four days and fed with 14:0 fatty acid.....	57
Figure 3.5 Amounts of total fatty alcohol produced by FAR5 and FAR8 variants expressed in yeast.....	58
Figure 3.6 Domain swaps between FAR5 and FAR8.....	59
Figure 3.7 Effects of site-specific substitutions involving six amino acid residues 355-357 and 377-379 on the chain-length specificity of FAR5.....	60
Figure 3.8 Reciprocal amino acid substitutions in FAR5 and FAR8.....	61
Figure 3.9 FAR <i>in vitro</i> assays with yeast microsomes containing FAR5 and FAR8 variants.....	62

List of Tables

Table 1.1 Substrate specificities of FAR enzymes from various organisms.....	32-33
Table 2.1 List of primers used in this study.....	41
Table 3.1 Specific activities of FAR5, FAR5-A355L, and FAR5-A355L V377M.....	63

List of Abbreviations

ACP: Acyl Carrier Protein
ATP: Adenosine Triphosphate
bp: Base Pairs
BSTFA: *N,O*-bis (trimethylsilyl) trifluoroacetamide
C: Carbon
CER: Eceriferum
CoA: Coenzyme A
DNA: Deoxyribonucleic Acid
DPW: Defective Pollen Wall
E: Trans
EV: Empty Vector
FA: Fatty Acid
FAR: Fatty Acyl Reductase
GC: Gas Chromatography
H: Hydrogen
hr: Hour(s)
HEK: Human Embryonic Kidney
HCl: Hydrogen Chloride
His: Histidine
HPTLC: High Performance Thin Layer Chromatography
KCS: β -Ketoacyl-CoA Synthase
kDa: Kilodalton
Leu: Leucine
LiAc: Lithium Acetate
min: Minute(s)
mL: Milliliter
mM: Millimolar
MS2: Male Sterility 2
NaCl: Sodium Chloride
NADPH: Nicotinamide Adenine Dinucleotide Phosphate
NADP: Nicotinamide Adenine Dinucleotide
NMR: Nuclear Magnetic Resonance
O: Oxygen
OD: Optical Density
OH: Alcohol
P: Phosphorus
PEG: Polyethylene Glycol
PCR: Polymerase Chain Reaction
pg: Pheromone Gland
rpm: Rotations Per Minute
SD: Synthetic Media Containing Dextrose
Sec: Second(s)
SG: Synthetic Media Containing Galactose
SDR: Short Chain Alcohol Dehydrogenase/Reductase

SDS: Sodium Dodecyl Sulfate

SDS-PAGE: Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

SS: Single Stranded

Sf9: *Spodoptera frugiperda* cells

TBST: Tris Base, Sodium Chloride, Tween-20

Tris-Base: Tris(Hydroxymethyl) Aminomethane Base

Ura: Uracil

v/v: Volume/Volume

VLCFA: Very-Long Chain Fatty Acids

WS: Wax Synthase

w/v: Mass/Volume

X: Times

Z: Cis

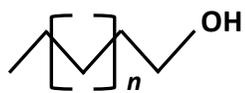
µg: Micrograms

CHAPTER 1: General Introduction

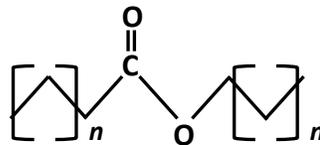
1.1 Fatty Alcohols in Nature – Structure and Occurrence

Fatty alcohols play a variety of biological roles in a diverse number of organisms across several kingdoms. They are characterized by long-chain aliphatic hydrocarbons that contain a hydroxyl group (-OH), typically in the terminal position (Figure 1.1). Primary fatty alcohol compounds differ considerably, with variability in chain length, chain saturation (double bonds), and methyl branching. In nature, these compounds range in chain length from C8 to greater than C32 – usually containing an even number of carbons, as well as being predominantly linear and unsaturated (Kaneda, 1962; Kattner *et al.*, 2003; Mudge, 2005).

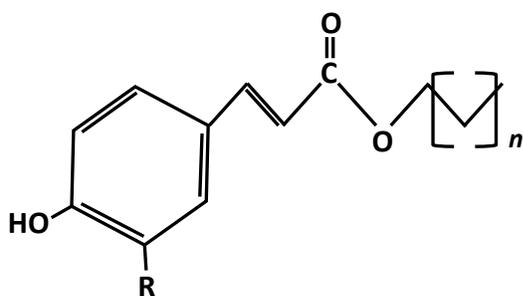
Fatty alcohols can occur in free form or in combined form. Combined forms include wax esters (a fatty alcohol esterified to a fatty acid), alkyl hydroxycinnamates (fatty alcohol linked to hydroxylated derivatives of the phenolic cinnamic acid), and ether lipids (fatty alcohol linked to a glycerophospholipid) (Rowland and Domergue, 2012). As a result, there is considerable structural diversity of fatty alcohol-containing compounds. Additionally, the physiochemical properties of free and combined fatty alcohols vary considerably from one compound to another, consequently allowing them to be utilized in a wide range of biological roles. Free or combined fatty alcohols have been reported in bacteria, protists, fungi, insects, plants and mammals. The presence of fatty alcohols and their derivatives in such evolutionarily distant organisms suggests that they are involved in basal and indispensable functions. Examples of the roles that free and combined fatty alcohols play in nature are: energy storage, communication, abiotic and biotic stress resistance, lubrication, and as a constituent of nervous system tissue (Hajra, 1983; Moto *et al.*, 2003; Rowland *et al.*, 2006; Teerawanichpan *et al.*, 2010; Hellenbrand *et al.*, 2011).



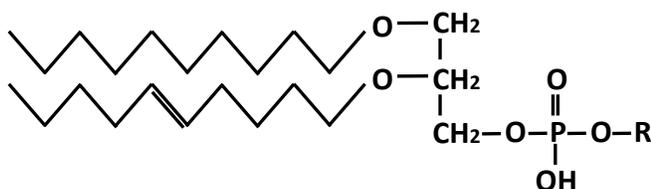
Free Primary Fatty Alcohol



Wax Ester



Alkyl Hydroxycinnamate



Ether Lipid

Figure 1.1 Generalized structures of free primary fatty alcohols, wax esters, alkyl hydroxycinnamates, and ether lipids. R= H, OH, or OCH₃

1.2 Biochemistry of Fatty Acyl Reductases

Fatty alcohols are synthesized through an enzyme-mediated reduction of fatty acyl-coenzyme A (CoA) or fatty acyl-acyl carrier protein (ACP). This process can occur via one of two mechanisms, both of which involve the production of a fatty aldehyde intermediate (Kunst and Samuels, 2003; Doan *et al.*, 2012). The first mechanism is a two-enzyme reaction, in which a fatty acyl reductase (FAR) converts a fatty acyl chain to a fatty aldehyde, and then a separate aldehyde reductase reduces the fatty aldehyde to a fatty alcohol. This two-enzyme mechanism has been observed in the eukaryotic algae *Botryococcus braunii*, the bacterium *Acinetobacter calcoaceticus*, the cyanobacterium *Synechococcus elongatus* PCC7942, and in pea leaves (*Pisum sativum*) (Wang and Kolattukudy, 1995a; Reiser and Somerville, 1997; Vioque and Kolattukudy 1997; Schirmer *et al.*, 2010). The second mechanism requires only one enzyme to catalyze the reduction of a fatty acyl chain to a fatty alcohol via the production of an unreleased aldehyde intermediate (Figure 1.2) (Kolattukudy, 1970). Using cell-free preparations of the protist *Euglena gracilis* or Jojoba (*Simmondsia chinensis*) plant cotyledons, it was determined that in addition to a fatty acyl substrate, this reaction requires the presence of ATP and NADH or NADPH as reductant (Kolattukudy, 1970; Pollard *et al.*, 1979). Because the fatty aldehyde intermediate is not released during this one-enzyme reaction, the alcohol-forming FAR utilized in this process is distinct from the FAR enzyme that carries out the first step of the two-enzyme reaction. Another notable difference between the two FAR types is that the alcohol-forming FARs are believed to be tightly-associated membrane proteins, possibly integral, and have a mass of approximately 56kDa, while the aldehyde-forming FARs are either soluble or have peripheral membrane associations, and have a mass of approximately 30kDa (Khan and Kolattukudy, 1975; Reiser and Somerville, 1997; Metz *et al.*, 2000). The one-step method of

fatty alcohol synthesis occurs more commonly in nature and investigations of the biochemistry of these alcohol-forming FARs have been carried out in a diverse number of organisms. From this point onwards, only this type of FAR protein will be discussed.

The molecular and biochemical characterization of FAR enzymes has been largely carried out in plants, with the first FAR cDNA being cloned from the seed of the desert shrub Jojoba (*Simmondsia chinensis*) (Metz *et al.*, 2000). With the determination of the Jojoba FAR cDNA sequence, identification of FAR homologues in a number of plants, animals and bacteria were made possible. From this information, a general FAR structure was determined in which protein length ranges from approximately 460 to 610 amino acids. Pair-wise protein sequence comparison between plant, animal, protist and bacterial FAR orthologs revealed at least 20% amino acid identity spanning over the entire length of the compared proteins. All FAR enzymes contain two distinct domains: a Rossmann-fold NAD(P)H binding domain at the N-terminus and a fatty acyl-‘CoA’ reductase (FAR_C) domain at the C-terminus (Figure 1.3) (Rowland and Domergue, 2012). The presence of the Rossmann-fold structure is typical of intermediate short-chain dehydrogenase/reductase proteins (Kavanagh *et al.*, 2008). Within the Rossmann-fold domain is a conserved GXXGXX(G/A) motif, which resembles the canonical [I/V/F]-X-[I/L/V]-T-G-X-T-G-F-L-[G/A] ADP binding domain and mediates NAD(P)H-binding. The Rossmann-fold domain contains the active site motif, YXXXXK, where the tyrosine (Y) and lysine (K) residues are predicted to play direct roles in catalysis based on kinetic studies with other reductases (Chen *et al.*, 1993; Fujimoto *et al.*, 2001). The FAR_C domain is unique to the FAR subfamily of reductases and is also annotated as the “male sterility” domain as it was first reported in the MALE STERILITY2 (MS2/FAR2) predicted protein sequence from *Arabidopsis*

thaliana (Aarts *et al.*, 1997; Rowland *et al.*, 2006; Doan *et al.*, 2009). The function of this domain is presently unknown.

In addition, FAR enzymes may also possess other features. Some plant FARs have an N-terminal extension that contains a chloroplast targeting sequence (Doan *et al.*, 2009; Chen *et al.*, 2011; Shi *et al.*, 2011). This cleavable N-terminal transit peptide guides the protein to the chloroplast and acts as an envelope transfer stroma targeting domain. This targeting sequence can range from 20-120 amino acids in length (Von Heijne and Nishikawa, 1991; Soll and Tien, 1998), with the C-terminal portion typically containing a conserved I/V-X-A/C-A consensus sequence proximal to the cleavage site (Gavel and Von Heijne, 1990). As well, mouse (*Mus musculus*) FAR1 has been determined to have a 66 amino acid putative transmembrane sequence at the C-terminus that anchors the protein to the peroxisome (Heilmann *et al.*, 2012). Alignment of the mouse FAR1 with FAR enzymes found in goose, duck, and barn owl demonstrated that they also possess this ~66 amino acid extension and, overall, have high sequence identity to FAR1 from mouse (Figure 1.3) (Hellenbrand *et al.*, 2011).

FAR enzymes can act on either of two types of fatty acyl substrate to produce fatty alcohols: fatty acyl-CoA(s) or fatty acyl-ACP(s) (Rowland *et al.*, 2006; Chen *et al.*, 2011; Doan *et al.*, 2011; Shi *et al.*, 2011). Fatty acyl-CoA is generally the preferred substrate, however some plant FARs prefer acyl-ACP precursors (Chen *et al.*, 2011). These FARs are localized to the plastid, which is the site of *de novo* fatty acid synthesis *in planta*. The affinity of these FARs for acyl-ACP substrates is reasonable as ACP is the main acyl carrier during the process of fatty acid synthesis in the plastid (Samuels *et al.*, 2008). Plastid localized FARs have not been identified in every plant species in which alcohol reduction has been studied, but some examples include FAR2/MS2 and FAR6 from *Arabidopsis thaliana*, and FAR2/DPW from rice (*Oryza sativa*)

(Aarts *et al.*, 1997; Chen *et al.*, 2011; Shi *et al.*, 2011). While it's been demonstrated that Arabidopsis FAR2/MS2 only accepts fatty acyl-ACP (Chen *et al.*, 2011), Arabidopsis FAR2/MS2 and rice FAR2/DPW are able to act on both acyl-ACP and acyl-CoA substrates, though its activity is much reduced when acting on acyl-CoA precursors (Shi *et al.*, 2011).

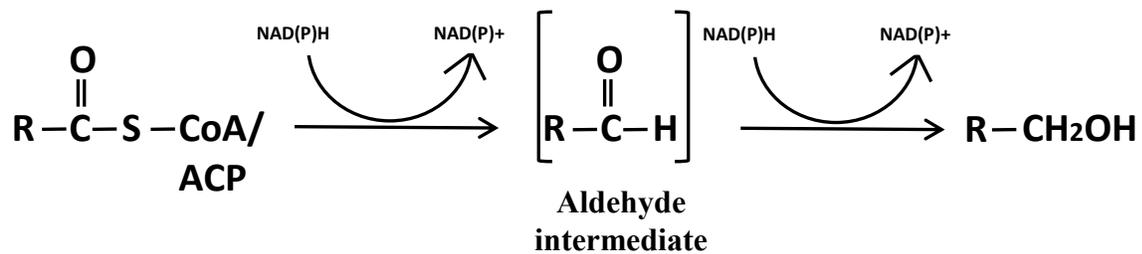


Figure 1.2 One-enzyme reduction reaction performed by fatty acyl reductase (FAR) enzymes. The substrate may be fatty acyl-CoA or fatty acyl-ACP, and the reductant is NADPH. There is an unreleased fatty aldehyde intermediate. R represents a hydrocarbon chain.

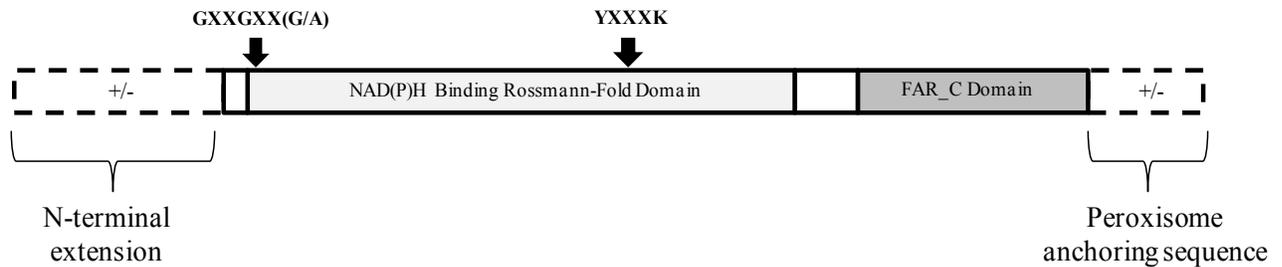


Figure 1.3 Schematic of the structural domains of FAR proteins. FAR enzymes, minus possible (+/-) subcellular localization signals, are approximately 500 amino acids in length and contain a NADPH-binding Rossmann-fold domain (light grey) and a FAR_C domain (dark grey) at the N-terminus and C-terminus, respectively. The GXXGXX(G/A) predicted NADPH binding motif as well as the predicted active site motif YXXXX are indicated (where 'X' represents any amino acid). Some FARs have an N-terminal extension of approximately 20 to 120 amino acids that contains a plastid targeting sequence. Others have a C-terminal extension of approximately 66 amino acids, which is a transmembrane sequence that anchors the protein to the peroxisome.

1.3 Biological Functions of FARs

FAR enzymes are found in a diverse number of organisms that include plants, animals, insects, protists and bacteria. They are involved in a wide number of critical biological functions. Because of this, individual FARs have evolved to fit the specific requirements of the biological role they play. This specificity is reflected in their specific subcellular localization, tissue distribution, temporal expression and substrate preferences. In this section, I discuss FAR activity in relation to its *in vivo* function, which includes: pheromone production, lubrication/waterproofing, surface lipid barrier production, ether lipid production, and energy storage.

1.3.1 Pheromones

Pheromones are secreted chemical mixtures that elicit social responses in members of the same species (Wyatt, 2003). Pheromone-based behaviour is critical throughout the animal kingdom, with pheromones playing roles in sexual signaling, territorial scent marking, and alarm/defense signaling. Pheromone mixtures are often complex blends of multiple compounds, which include lipids and proteins (Khannoon *et al.*, 2011). These mixtures often contain long-chain fatty alcohols, as observed in a number of species of reptile, insect, bird, and primate (Yarger *et al.*, 1976; Moto *et al.*, 2003; Hellenbrand *et al.*, 2011; Khannoon *et al.*, 2011). Recent investigations into pheromone biosynthesis of several species of moth have identified FAR enzymes that are specific to the pheromone gland (pgFARs). An example of this is the pgFAR found in silkworm (*Bombyx mori*), a moth species that produces the sex pheromone bombykol, (*E,Z*)-10,12-hexadecadiene-1-ol, a diunsaturated primary fatty alcohol. The final reaction step in the production of bombykol requires a FAR-mediated reduction of (*E,Z*)-10,12-hexadecanoic

acid precursor to its corresponding fatty alcohol (i.e. bombykol). Functional expression of this pgFAR in *Saccharomyces cerevisiae* reveals that this enzyme is NADPH dependent and possesses broad substrate specificity. pgFAR is able to reduce 14:0-, 16:0-, and 18:0-acyl-CoAs as well as *E*11-16:1- and *Z*11-16:1-acyl-CoAs, but its activity for each of these substrates is variable. It demonstrated strong activity towards the bombykol precursor (*E,Z*)-10,12-16:2-acyl-CoA (Moto *et al.*, 2003). Other examples of pgFARs that exhibit a broad substrate range include those from the small ermine moth family (*Lepidoptera: Yponomeutidae*) and honey bee (*Apis mellifera*) (Liénard *et al.*, 2010; Teerawanichpan *et al.*, 2010). FAR homologues were found in the female pheromone gland of three species of *Yponomeutidae*: *Y. evonymellus*, *Y. padellus*, and *Y. rorellus*, all of which produce a sex pheromone composed primarily of chemically related C14- and C16- fatty alcohols and their derivatives (Liénard *et al.*, 2010). In addition to being localized to the female pheromone gland, it was shown that *yev*-pgFAR (from *Y. evonymellus*) expression is highest during periods of known pheromone emission. The pgFARs found in all three ermine species demonstrated identical substrate specificity when heterologously expressed in *S. cerevisiae*. This activity was for 14:0- and 16:0-acyl-CoA, as well as *Z*9-, *Z*11- and *E*-11-14:1-acyl-CoA and *Z*11-16:1-acyl-CoA (Liénard *et al.*, 2010). Another pheromone gland-specific FAR that exhibits broad substrate specificity is found in honey bee (*A. mellifera*) (Teerawanichpan *et al.*, 2010). The honey bee produces complex communication pheromones that are comprised of a mixture of hydrocarbons, fatty acids, wax esters, fatty alcohols, and fatty aldehydes. A search of the honey bee genome identified six putative FARs, one of which, *AmFAR1*, was characterized in detail. *AmFAR1* was found to be ubiquitously expressed in all body parts, with predominance in the head. When expressed in *S. cerevisiae*, it produced C16-22 saturated primary fatty alcohols as well as 16:1-OHⁿ⁻⁹, 18:1-OHⁿ⁻⁹, and 20:1-OHⁿ⁻⁹, with 18:0-

OH the preferred product. The expression pattern and substrate profile of *AmFAR1* suggests that it is most likely involved in pheromone and/or ether lipid production (Teerawanichpan *et al.*, 2010).

Pheromone gland specific FARs found in silkworm, small ermine moths and honey bee represent reductase systems that reduce a broad set of fatty acyl precursors, suggesting that the structural composition of the pheromone blends depends in part upon specific chemical precursors made by upstream biosynthetic enzymes (Li nard *et al.*, 2010). An opposite example of this would be the European Corn Borer (*Ostrinia nubilalis*), which displays strict FAR substrate specificity (Lassance *et al.*, 2010). *O. nubilalis* consists of two sex pheromone races that use different ratios of a *cis* (*Z*) and *trans* (*E*) isomer of the same acetate compound: (*E/Z*)-11-tetradecenyl acetate. This compound is synthesized from palmitic acid (16:0), which undergoes several modifications to produce either (*E*)-11-tetradecenoyl or (*Z*)-11-tetradecenoyl. These monounsaturated precursors are then reduced to their corresponding fatty alcohol by a FAR enzyme before being acetylated to yield the final acetate product. The difference in the *E/Z* ratio of this compound in the final pheromone mixture of each race has resulted in reproductive isolation and may represent the first step towards speciation (Lassance *et al.*, 2010). The pheromone gland FAR from the two races of *O. nubilalis*, named pgFAR-*E* and pgFAR-*Z*, were heterologously expressed in *S. cerevisiae* supplied with (*E*)-11-tetradecenoyl and (*Z*)-11-tetradecenoyl substrates, and it was observed that pgFAR-*E* almost exclusively converted the *E* isomer to its corresponding fatty alcohol with minimal amounts of the *Z* isomer being reduced, and pgFAR-*Z* almost exclusively converted the *Z* isomer to its corresponding fatty alcohol with minimal amounts of the *E* isomer being reduced. This indicates that the strict substrate specificity of the pgFAR expressed in these two races is responsible for the difference in final

pheromone product, and that this observed phenotypic variation of female pheromone composition has led to race-specific signals (Lassance *et al.*, 2010).

1.3.2 Lubrication / Waterproofing

Wax esters, a fatty alcohol esterified to a fatty acid, are utilized by many bird species as a component of uropygial gland secretions. The uropygial gland is found at the base of the tail between the fourth caudal vertebrae and the pygostyle, and its secretions are generally composed of acidic mucins (heavily glycosylated proteins), neutral lipids, glycolipids, and phospholipids. The secretions contribute to hygiene, lubrication, and waterproofing of the feather coat (Salibian and Montalti, 2009). The type and composition of these wax esters is dependent on the bird species, and can be composed of straight- or branched-chained fatty acids and fatty alcohols (Murray, 1962; Jacob and Poltz, 1974). Identification and isolation of FAR sequences from the uropygial gland of barn owl (*Tyto alba*, *Ta*), domestic chicken (*Gallus gallus domesticus*, *Ta*) and domestic goose (*Anser anser domesticus*, *Ad*) identified two distinct FAR groups, FAR1 and FAR2. The first group is comprised of *GgFAR1* and *GgFAR1db* from chicken, *TaFAR1* from owl, and *AdFAR1* from goose, while the second group contains *GgFAR2* and *TaFAR2* from chicken and owl, respectively (no *FAR2* homologue was found in goose). Gene expression analyses showed that *FAR* transcripts from both groups are expressed in various tissues. *FAR1* expression was mainly found in the uropygial gland, while *FAR2* expression was highest in brain tissue. This indicates that the *FAR2* group of enzymes are most likely involved in ether lipid production in the brain, while the presence of *FAR1* in the uropygial gland suggests that this group is involved in the production of wax esters for secretion. *In vitro* assays from both FAR groups expressed in *S. cerevisiae* demonstrated that the *FAR1* group has preference for 15:0- and 16:0-acyl-CoA substrate chain lengths, while the *FAR2* group has pronounced preference for

18:0-acyl-CoA, with some activity for 17:0-acyl-CoA. The only FAR that displayed relaxed chain length substrate specificity was *AdFAR1* from goose, which was able to produce 14:0-OH, 16:0-OH, and 18:0-OH in similar quantities (Hellenbrand *et al.*, 2011). Branched-chain fatty alcohols had been previously reported in the uropygial wax esters of owl (Jacob and Poltz, 1974). Further investigation into the ability of the avian FARs to reduce branch-chained fatty acids revealed that *GgFAR1* and *TaFAR1* were able to reduce 2-methyl-16:0-CoA, *GgFAR2* and *TaFAR2* were able to reduce 2-methyl-18:0-CoA, and *AdFAR1* demonstrated broader specificity and was able to reduce 2-methyl-14:0-, 2-methyl-16:0-, and 2-methyl-18:0-CoA chain lengths. These avian FARs, however, display distinctly lower activity for the branched acyl chain substrates compared to the unbranched substrates (Hellenbrand *et al.*, 2011).

1.3.3 Surface barriers

A number of plant and insect species possess protective extracellular lipid-based barriers. These barriers serve many functions, such as prevention of water loss, resistance to pathogen attack, and prevention of penetration by chemicals and toxins (Schrieber *et al.*, 1996; Reina-Pinto and Yephremov, 2009; Golebiowski *et al.*, 2011). In insects, these barriers are referred to as cuticular lipids and are generally composed of hydrocarbons, wax esters, sterol esters, fatty ketones, fatty aldehydes, fatty alcohols, and fatty acids (Brucker, 1993). Plants, on the other hand, possess three distinct types of lipid-based barriers that differ in their location and composition. These barriers are referred to as cuticle, suberin and sporopollenin (Pollard *et al.*, 2008; Samuels *et al.*, 2008; Ariizumi *et al.*, 2011).

The cuticle coats the aerial surfaces of all land plants (Kunst and Samuels, 2009) and is composed of two distinct layers: cutin and cuticular waxes. The cutin polyester consists

primarily of C16 and C18 fatty acid derivatives, such as ω - and mid-chain hydroxylated fatty acids, α,ω -dicarboxylic fatty acids, and long-chain fatty alcohols, as well as glycerol (Beisson *et al.*, 2010). Cuticular wax is composed of very-long-chain fatty acids (VLCFA) ranging in chain length from C24-C34 and corresponding derivatives, such as fatty aldehydes, fatty ketones, primary fatty alcohols, alkanes, and wax esters (Samuels *et al.*, 2008). This barrier is important for preventing uncontrolled water and solute loss from interior tissues (Schreiber *et al.*, 1996), resistance to bacterial and fungal pathogen attack (Reina-Pinto and Yephremov, 2009), protection against mechanical damage and UV radiation (Jenks *et al.*, 2002), and reducing water retention on the surface of plants to minimize deposition of spores, dust and air pollutants (Kerstiens, 1996; Reiderer and Muller, 2006). Suberin is a protective hydrophobic barrier with tissue-specific deposition in plant cell walls. Its synthesis is developmentally regulated and is induced by environmental stimuli such as wounding (Ranathunge *et al.*, 2010). Suberin deposition occurs in the cell walls of root endodermal cells, shoot and root periderms (including wound periderm), the bundle sheath of grasses, the chalazae of seeds, and seed coat integuments (Matzke *et al.*, 1991; Bernards, 2002). The suberin polymer is composed of phenolic and aliphatic molecules. The phenolic portion is composed mostly of hydroxycinnamate derivatives, such as ferulic acid, cinnamic acid, and *p*-coumaric acid. The aliphatic layer is made up primarily of ω -hydroxy fatty acid, α,ω -dicarboxylic acids, VLCFAs, and primary fatty alcohols (Schreiber *et al.*, 2005; Molina *et al.*, 2009). Suberin is capable of performing many of the same roles as the cuticle, such as protecting the plant from invasion of fungal and bacterial pathogens, as well as controlling the flow of water, mineral nutrients, and essential gases (Ranathunge *et al.*, 2010). Sporopollenin is a major component of the protective outer exine layer of the pollen wall. This biopolymer is exceptionally stable and is highly resistant to degradation, which makes

determining its chemical composition difficult, although current evidence suggests that the biopolymer contains both aromatic and aliphatic components such as oxygenated aromatics, fatty acids, and primary fatty alcohols (Aarts *et al.*, 1997; Dominguez *et al.*, 1999; Dobrista *et al.*, 2009). The function of the sporopollenin layer is to protect the gametophyte from environmental stresses such as desiccation and UV radiation, as well as mediate interactions between the pollen grain and stigma (Wiermann *et al.*, 1992; Chen *et al.*, 2011).

While fatty alcohols are found in each of these extracellular lipid barriers, the amounts of fatty alcohols present, as well as the composition of fatty alcohols with regards to chain lengths, are variable between the three different barriers and is variable from species to species (Pollard *et al.*, 2008). Investigation into the synthesis of these fatty alcohols has been extensively studied in the model plant *Arabidopsis thaliana*, where a family of eight *FAR* genes (named *FAR1* to *FAR8*) has been identified (Domergue *et al.*, 2010). *FAR* enzymes from bread wheat and rice have also been characterized (Wang *et al.*, 2002; Shi *et al.*, 2011). Each *Arabidopsis* *FAR* enzyme displays distinct substrate specificity for certain saturated fatty acyl chain lengths ranging from C16-C30. Each has a distinct gene expression pattern that typically correlates with a site of surface lipid barrier deposition.

FAR1, *FAR4* and *FAR5* from *Arabidopsis* are expressed at sites of suberin deposition (Domergue *et al.*, 2010). *FAR4* and *FAR5* gene expression patterns are mostly restricted to root endodermal cells, which is a known site of suberin deposition, suggesting that these two encoded enzymes serve only to produce fatty alcohols that are incorporated into the suberin polymer. *FAR1* expression is more ubiquitous, with expression in root endodermal cells, young leaves, flower petals, and cotyledons, which suggests it also has non-suberin functions. All three genes are induced by wounding and salt stress. Domergue *et al.* (2010) also expressed *FAR1*, *FAR4*,

and FAR5 in *S. cerevisiae* confirming that these three proteins are fatty acyl reductases that produce saturated fatty alcohols ranging in length from C18-C22. FAR1 produced mostly 22:0 fatty alcohol, FAR4 mostly 20:0 fatty alcohol, and FAR5 mostly 18:0 fatty alcohol. This yeast data indicated that FAR1, FAR4, and FAR5 have distinct chain length specificities, in agreement with the chain length of fatty alcohols most affected in the aliphatic suberin polyester of *Arabidopsis far1, far4, and far5* mutant plants (Domergue *et al.*, 2010; Rowland and Domergue, 2012).

FAR2, also named MS2 (MALE STERILITY2), is a fatty acyl-ACP reductase from *Arabidopsis* which produces fatty alcohols that are likely incorporated into the sporopollenin polymer of the exine layer of the pollen coat (Aarts *et al.*, 1995, 1997; Chen *et al.*, 2011). *FAR2/MS2* is expressed in the tapetal cells that surround the developing pollen grain in the anther. It is believed that the tapetal cells are responsible for synthesizing and secreting the sporopollenin precursors onto the pollen grain surface where polymerization of the sporopollenin exine layer then occurs (Li and Zhang, 2010). Loss-of-function *ms2/far2* mutants have thin pollen walls and non-functional pollen (Aarts *et al.*, 1997; Chen *et al.*, 2011). Functional characterization of the FAR2/MS2 protein showed that it is localized to the plastid and is highly specific for acyl-ACP substrate (Chen *et al.*, 2011). FAR2/MS2 displays strong activity for 16:0-ACP, trace activity for 18:0-ACP and no activity for 14:0-ACP or 18:1-ACP *in vitro*, indicating that this enzyme has strong preference for palmitoyl-ACP as a substrate (Chen *et al.*, 2011). A rice ortholog of FAR2/MS2, known as DEFECTIVE POLLEN WALL (DPW), has been found to be expressed in the tapetal cells of rice anthers as well as in the microspores (Shi *et al.*, 2011). Loss-of-function *dpw* mutants display a dramatically altered pollen exine layer, as well as abnormalities to the anther surface. The *dpw* mutants also have reduced levels of anther cutin monomers,

including fatty alcohols. There was approximately a 40% reduction of both 16:0 and 18:0 primary fatty alcohols in the total soluble anther lipids of the *dpw* mutant. *In vitro* assays demonstrated that DPW has activity for 16:0-ACP as well as 16:0-CoA and 16:1-CoA (Shi *et al.*, 2011).

FAR3, also named CER4 (ECERIFERUM 4), is a fatty acyl reductase responsible for producing primary fatty alcohols that are incorporated into the cuticle of Arabidopsis aerial tissues. Rowland *et al.* (2006) found that *FAR3/CER4* is expressed only in the epidermal cells of the stem, which is consistent with this enzyme's specific role in cuticle formation. Wax analysis of Arabidopsis *far3/cer4* mutants revealed that this enzyme is responsible for the production of 24:0-30:0 fatty alcohols. Consistent with this finding, expression of FAR3/CER4 in *S. cerevisiae* results in accumulation of 24:0 and 26:0 saturated primary alcohols (Rowland *et al.*, 2006). *Far3/cer4* mutants have a significant decrease in both primary fatty alcohols and wax esters (fatty alcohols combined with fatty acids) in the cuticle, although the mutant's total wax load remained nearly the same as the wild-type's wax load due to flux into the alkane-forming pathway (Rowland *et al.*, 2006). A FAR enzyme in bread wheat (*Triticum aestivum*), named TAA1, was shown to be involved in the production of fatty alcohols for incorporation into the cuticular wax layer that coats the pollen grain, known as the tryphine (Wang *et al.*, 2002). TAA1 is anther filament-specific and is expressed at the time of microspore thickening. Heterologous expression of TAA1 in tobacco seed resulted in the production of 18:1, 20:1, 22:1, 24:0 and 26:0 primary fatty alcohols. While TAA1 is specific to the pollen grain, it is likely not contributing to sporopollenin formation. Because this FAR does not possess a chloroplast targeting sequence and is more related to Arabidopsis FAR3/CER4 than FAR2/MS2, as well as its observed

activity for very-long-chain fatty acyl-CoAs in tobacco seed, it is likely that TAA1 is involved in tryptophan production (Rowland and Domergue, 2012).

The functions of Arabidopsis FAR6, FAR7 and FAR8 have not yet been elucidated. FAR6 is expressed in the epidermis of the stem as well as the root cap and anthers (Doan *et al.*, 2011). In the same study, FAR6 was localized to the chloroplast and produces 16:0 primary fatty alcohols within this organelle. FAR6 is able to accept both acyl-ACP and acyl-CoA as substrates. In addition to producing fatty alcohols, this enzyme has been shown to produce fatty aldehydes and fatty acids under certain *in vitro* conditions (Doan *et al.*, 2011). FAR6 displays a distinct expression pattern compared to other Arabidopsis FARs, suggesting a unique role but that role is currently unknown. FAR7 is most likely a pseudogene because an in-frame stop codon results in a truncated 95 amino acid predicted protein and no fatty alcohols are produced upon expression in *E. coli*. (Doan *et al.*, 2009). Expression of FAR8 in *E. coli* produces small amounts of 16:0 primary alcohols with trace amounts of 14:0-OH and 18:1-OH, while expression of FAR8 in yeast produces trace amounts of 16:0-OH (see Results section of this thesis). Analysis of the tissue specific gene expression levels of FAR1-8 demonstrated that the level of expression of FAR7 and FAR8 are very low (Domergue *et al.*, 2010), suggesting that they are either expressed at very specific developmental stages or that they are pseudogenes.

1.3.4 Ether lipids

Ether lipids (fatty alcohol linked to a glycerophospholipid), a sub category of phospholipids, are a ubiquitous and sometimes a major component of bacterial and mammalian cell membranes (Paltauf, 1994). Ether lipids may play a role in the lipid secondary messenger system as well as act as antioxidants (Spector and Yorek, 1985; Brosche and Platt, 1998). They

account for approximately 20% of phospholipids in the human body (Cheng *et al.*, 2004). Two mouse (*Mus musculus*) and human (*Homo sapien*) FAR isozymes, named FAR1 and FAR2, have been identified that share 58% sequence identity. Both FAR1 and FAR2 have been localized to the peroxisome, which is consistent with their roles in ether lipid synthesis as this organelle contains three of the seven enzymes required for ether lipid production. It has been found that *FAR1* mRNA transcripts display broad distribution in mouse tissues with highest levels seen in the preputial gland, while *FAR2* is expressed at lower levels in a smaller number of tissues with highest expression in eyelids and skin. An *in vitro* assay of mouse FAR1 expressed in insect ovarian cells (*Spodoptera frugiperda* (Sf) 9 cells) produced 16:0, 18:0, 18:1 and 18:2 fatty alcohols, while expression of mouse FAR2 resulted in 16:0 and 18:0 fatty alcohols. Both FARs required NADPH, but not NADH, as a cofactor. The broad tissue distribution of FAR1 and its flexible substrate specificity suggests that it plays a general role in fatty alcohol synthesis, while the narrow distribution and strict substrate specificity of FAR2 indicates more specialized function. Support for this hypothesis is the observed variety of ether lipids found in tissues expressing FAR1; the ether lipids found in these tissues have diverse structures, consistent with the incorporation of a variety of fatty alcohol constituents (Cheng *et al.*, 2004). *GgFAR2* and *TaFAR2* from chicken and owl, respectively, may also be involved in ether lipid biosynthesis, as both of these avian FARs are highly expressed in brain tissue (see above in lubrication/waterproofing) (Hellenbrand *et al.*, 2011).

1.3.5 Energy storage

It has been demonstrated that wax esters can serve as an energy storage reserve, as opposed to the more commonly used triacylglycerol, in a number of species of bacteria,

unicellular phytoflagellates, marine fish and crustaceans, and in the seed of the dessert shrub plant Jojoba. Investigation of wax ester synthesis in bacteria has revealed that fatty alcohol formation occurs more frequently through the two-enzyme method than the more well-characterized one-enzyme reaction – examples of this include *Acinetobacter calcoaceticus* and *Synechococcus elongatus* PCC7942 (Reiser *et al.*, 1997; Willis *et al.*, 2011). However, a FAR enzyme from the gram-negative marine bacterium *Marinobacter aquaeolei* VT8 has been identified that carries out the one-enzyme reduction of fatty acyl chains. This *M. aquaeolei* FAR is NADPH dependent and only uses fatty acyl-CoA substrate. It displays broad substrate specificity as it is able to reduce 8:0, 12:0, 14:0, 16:0, 16:1, 18:0, 18:1, and C20:4 saturated and unsaturated fatty acyl-CoAs, with highest activity for 16:0 acyl-CoA (Willis *et al.*, 2011). The unicellular phytoflagellate protist *Euglena gracilis* accumulates wax esters as an energy storage reserve during dark or anaerobic conditions. Wax esters can make up to 62% of the total lipid content, with myristyl myristate (14:0–14:0) being the major species (Teerawanichpan and Qiu, 2010). Experiments with cell-free preparations of the associated *Euglena* FAR demonstrated that it is NADH dependent and works on fatty acyl-CoA substrates (Kolattukudy, 1970). Heterologous expression of this FAR in *S. cerevisiae* results in production of 14:0 and 16:0 fatty alcohols, with 14:0-OH being the major product (Teerawanichpan and Qiu, 2010). This is in agreement with myristyl myristate being the major wax ester constituent in *E. gracilis* (Teerawanichpan and Qiu, 2010). Wax esters have also been found to be synthesized *de novo* in a number of myctophid fish and crustaceans (Kayama and Nevenzel, 1974). The marine crustacean *Calanus finmarchicus* comprises a large proportion of Arctic and North Sea biomass, and it accumulates wax esters in which 20:1ⁿ⁻⁹ and 22:1ⁿ⁻¹¹ represent the major fatty alcohol constituents and 16:0, 16:1ⁿ⁻⁹, 18:1ⁿ⁻⁹ and 18:1ⁿ⁻⁶ fatty alcohols represent minor fatty alcohol

constituents (Teerawanichpan and Qiu, 2011). Three *C. finmarchicus* FAR enzymes, named CfFAR1, CfFAR2, and CfFAR3, have been identified. These three FARs display distinct substrate specificities when expressed in *S. cerevisiae*: CfFAR1 is active towards 18:0-26:0 fatty acyl-CoA chain lengths, CfFAR2 displays narrower specificity for 24:0 and 26:0 acyl-CoAs only, and CfFAR3 is active towards 16:0, 16:1, 18:0 and 18:1 fatty acyl-CoAs. These data suggest that CfFAR3 is in part responsible for the fatty alcohol profile found in *C. finmarchicus* wax esters, while the roles of CfFAR1 and CfFAR2 are unclear (Teerawanichpan and Qiu, 2011). Jojoba (*Simmondsia chinensis*) is a shrub found in the south-western region of North America. It is the only known plant that utilizes wax esters in its Jojoba seed as the major energy reserve. These wax esters play an equivalent role to the triacylglycerols found in other oilseed plants such as *Brassica napus*, and provide the majority of the energy required by the seed to achieve post-germination growth prior to the establishment of photosynthesis (Metz *et al.*, 2000). Jojoba oil is composed of 98% wax ester, in which C40-C44 are the major wax ester chain lengths. The fatty alcohol component of these wax esters is predominated by equal proportions of 20:1-OH and 22:1-OH (Rowland and Domergue, 2012). Expression of the seed-specific Jojoba FAR in *E. coli* results in the production of 12:0, 14:0, 16:0 and 18:1 fatty alcohols; however these chain lengths are not in agreement with the fatty alcohol species found *in planta*, which are predominantly 20:1 and 22:1 (Miwa, 1971; Doan *et al.*, 2009). Expression of Jojoba FAR in the developing embryos of *B. napus* resulted in a series of wax esters in the seed, the most abundant chain-length being 42:2 (22:1 fatty alcohol esterified to a 20:1 fatty acid). This observed formation of wax esters in the seed of *B. napus* is a result of endogenous wax synthase (WS) activity, which catalyzes the esterification of a fatty alcohol to a fatty acid (Lardazibal *et al.*, 2000; Metz *et al.*, 2000). Jojoba FAR has also been expressed in *Arabidopsis* seeds in

combination with a β -ketoacyl-CoA synthase (KCS; component of fatty acid elongase) from *Lunaria annua*, which was chosen for its activity in elongating 18:1 acyl-CoA to 20:1-, 22:1- and 24:1-fatty acyl-CoAs, the required substrates for Jojoba FAR (Lassner, 1997; Lardazibal *et al.*, 2000). Co-expression of these two proteins resulted in 20:1, 22:1 and 24:1 free fatty alcohols, as well as a small amount of wax ester (Lardazibal *et al.*, 2000).

Table 1.1 Substrate specificities of FAR enzymes from various organisms. Modified from Ashley Fournier's M.Sc. thesis, Carleton University, 2011.

Organism	Substrate Specificity (using endogenous acyl lipid pools unless otherwise stated)			Publication
	<i>Escherichia coli</i>	<i>Saccharomyces cerevisiae</i>	Plant/Mammalian cells	
Marine Bacteria (<i>Marinobacter aquaeolei</i> VT8)	8:0, 12:0, 14:0, 16:0, 18:0, 18:1, 20:4 (<i>in vitro</i> assays)			Willis <i>et al.</i> , 2011
Phytoflagellate protist (<i>Euglena gracilis</i>)		14:0, 16:0		Teerawanichpan and Qiu, 2010b
Copepod (<i>Calanus finmarchicus</i>)		CjFAR1: 18:0-26:0 CjFAR2: 24:0, 26:0 CjFAR3: 16:0, 18:0, 18:1, 20:1		Teerawanichpan and Qiu, 2011
Silk moth (<i>Bombyx mori</i>)		14:0, 15:0, 16:0, 17:0, 18:0, E,Z-10,12- hexadecadienoic acid		Moto <i>et al.</i> , 2003
Small Ermine Moth (<i>Yponomeutidae</i> sp.)		12:0, 14:0, 16:0		Liénard <i>et al.</i> , 2010
Corn Borer Moth (<i>Ostrinia nubilalis</i>)		pgFAR-Z: (Z)-11- tetradecenoyl pgFAR-E: (E)-11- tetradecenoyl		Lassance <i>et al.</i> , 2010
Honey Bee (<i>Apis mellifera</i>)	18:0, 20:0, 22:0 (<i>in vitro</i> assay)	16:0, 18:0, 20:0, 22:0, 12-OH-18:1n-9, 16:1n- 9, 18:1n-9, 20:1n-9		Teerawanichpan <i>et al.</i> , 2010a
Goose (<i>Anser anser domesticus</i>)		AdFAR1: 15:0, 16:0, 17:0, 18:0, 2-methyl- 14:0-18:0-CoA		Hellenbrand <i>et al.</i> , 2011
Barn Owl (<i>Tyto alba</i>)		TaFAR1: 15:0, 16:0, 2- methyl-16:0-CoA TaFAR2: 17:0, 18:0, 2- methyl-18:0-CoA		Hellenbrand <i>et al.</i> , 2011
Chicken (<i>Gallus gallus domesticus</i>)		GgFAR1: 15:0, 16:0, 2- methyl-16:0-CoA GgFAR2: 17:0, 18:0, 2- methyl-18:0-CoA		Hellenbrand <i>et al.</i> , 2011
Jojoba (<i>Simmondsia chinensis</i>)	12:0, 14:0, 16:0, 18:1		22:1 (HEAR <i>Brassica napus</i>)	Metz <i>et al.</i> , 2000; Doan <i>et al.</i> , 2009; Lardizabal <i>et al.</i> , 2000
Bread Wheat (<i>Triticum aestivum</i>)	14:0, 16:0, 18:1		18:1, 20:1, 22:1, 24:0, 26:0 (N.t.)	Wang <i>et al.</i> , 2002
Rice (<i>Oryza sativa</i>)	16:0-CoA, C16:0- ACP, C16:1-ACP (<i>in vitro</i> assays)			Shi <i>et al.</i> , 2011
Thale Cress (<i>Arabidopsis thaliana</i>)	FAR1: 14:0, 16:0, 18:0, 18:1 FAR2: 14:0, 16:0, 18:0, 18:1 (ACP) (16:0 ACP specificity determined by <i>in vitro</i> assay) FAR3: 14:0, 16:0, 18:1 FAR6: 12:0-ACP,	FAR1: 16:0, 18:0, 20:0, 22:0 FAR3: 24:0, 26:0 FAR4: 18:0, 20:0 FAR5: 18:0 FAR6: 16:0, 18:0	FAR1: 22:0 (A.t.) FAR2: 16:0, 18:0 (N.b.) FAR3: 24:0, 26:0, 28:0 (A.t.) FAR4: 20:0 (A.t.) FAR5: 18:0 (A.t.) FAR6: 16:0, 18:0 (N.b.)	Rowland <i>et al.</i> , 2006; Doan <i>et al.</i> , 2009; Domergue <i>et al.</i> , 2010; Chen <i>et al.</i> , 2011; Doan <i>et al.</i> , 2011

	14:0 (CoA, ACP), 16:0 (CoA, ACP), 18:0-CoA, 18:1- ACP, 20:0-CoA (fatty acyl-CoA specificity determined <i>by in vitro</i> assay) FAR8: 14:0, 16:0, 18:1			
Pea leaves (<i>Pisum sativum</i>)			16:0-CoA	Vioque and Kolattukudy, 1997
Mouse (<i>Mus musculus</i>)			FAR1: 16:0, 18:0, 18:1, 18:2, 20:3, 20:4 FAR2: 16:0, 18:0 (HEK 293/Sf9 cells)	Cheng and Russell, 2004
Human (<i>Homo sapiens</i>)			FAR1: 16:0 FAR2: 16:0 (HEK 293 cells)	Cheng and Russell, 2004

A.t. = *Arabidopsis thaliana*

N.b. = *Nicotiana benthamiana* leaves

N.t. = *Nicotiana tabacum* cv. *Xanthi* seeds

HEAR = *Brassica napus* (high erucic acid rapeseed)

HEK 293 = Human Embryonic Kidney (HEK) 293 cells

Sf9 = *Spodoptera frugiperda* (Sf) 9 cells

1.4 FAR Substrate Specificities

FAR enzymes across nature display diverse substrate specificities, tailored to the fatty alcohol requirements of the given species. FARs have distinct substrate specificities with regards to both acyl chain-length and acyl chain saturation. This specificity can either be broad, where an individual FAR may have activity for a number of acyl chains, or it may be specific for a single substrate. Determining the biologically relevant specificity of an individual FAR is often difficult as a single FAR may display alternate activities depending on the heterologous system in which it is expressed. An example of this is TAA1 from wheat, which when expressed in *E. coli* produces 14:0, 16:0 and 18:1 fatty alcohols from the endogenous acyl pool, but when expressed in transgenic tobacco seeds produces 20:1, 22:1, 24:0, and 26:0 fatty alcohols from the endogenous acyl pool (Wang *et al.*, 2002). This indicates that in addition to a given FARs substrate specificity, the pool of available acyl precursors is also in part responsible for dictating the composition of fatty alcohols produced (Rowland and Domergue, 2012). Nonetheless, FAR enzymes display distinct substrate specificities and it has been observed that small differences to the FAR protein sequence can result in changes to its substrate preference. This can be observed in the pheromone gland specific FARs (pgFAR) found in the moth species *Ostrinia furnacalis* and *Ostrinia nubilalis*. These two pgFARs possess high sequence similarity, only differing at 16 amino acid positions, but they display distinct substrate specificities. Both pgFARs have high activity for Z11-14:1 acyl chain; however the *O. furnacalis*-pgFAR is also able to reduce Z9-14:1 and Z12-14:1 acyl chains while *O. nubilalis*-pgFAR-Z cannot. Site directed mutagenesis of the 16 differing amino acids between these two pgFARs revealed that a single amino acid at position 453 affects the enzymes affinity for Z9-14:1 versus Z12-14:1 acyl substrates (Lassance *et al.*, 2013). This indicates that it is possible to alter FAR substrate specificity with minor changes to

the amino acid sequence, opening the door to the possibility of engineering FARs that produce fatty alcohols with desired properties for industrial/commercial use.

1.5 FAR Industrial Applications

Free and combined fatty alcohols are valuable in a number of commercial applications. Long-chain free fatty alcohols are found as ingredients in food products, cosmetics, and pharmaceutical/agrochemical formulations, typically acting as thickening agents, emollients and emulsifiers (Mudge *et al.*, 2008; Rowland and Domergue, 2012). Fatty alcohols esterified to a fatty acid, i.e. wax esters, have also been employed for a variety of applications. Historically wax esters were derived from the spermaceti organ in the head of sperm whales (*Physeter macrophalus*), which is primarily composed of 32:0 wax ester (16:0 fatty alcohol esterified to 16:0 fatty acid). An adult male sperm whale contains approximately 2000 L of spermaceti oil, which is believed to be involved in the whale's sonar system (echolocation) and possibly helps to regulate buoyancy (Clark, 1970). Spermaceti oil was used extensively as an ingredient in cosmetics, candles, lubricants, and as a pharmaceutical excipient (for cerates and ointments); however, overexploitation of this resource over many years resulted in a significant reduction in the sperm whale population. This led to a ban on whale oil use in 1972 followed by a general ban on whaling several years later. Since then, alternative sources of wax esters have been explored, such as Jojoba seed oil. However, Jojoba is not amenable to large scale cultivation, which means that its use is generally limited to cosmetics, such as shampoos and creams (Naqvi and Ting, 1990).

Because wax esters intrinsically possess excellent lubrication properties and are highly resistant to hydrolysis and oxidation at high temperatures, they are sought after for a number of

industrial applications, such as high performance machinery lubricants and automobile transmission fluids (Carlsson, 2006). These intrinsic qualities, as well as the fact that wax esters derived from plants are a renewable resource, has generated interest in producing natural wax esters as a replacement to current fossil oil-derived lubricants mostly being used by industry today. Recently, there has been interest in engineering wax esters with qualities that are customized to fit specific industrial roles, likely this would involve producing short-chain and unsaturated wax esters as they possess very low melting temperatures. One way to develop such designer wax esters would be to engineer FAR enzymes to specifically produce fatty alcohols of desired physical properties (such as short chain lengths or unsaturated chains), which can then be incorporated into these ‘tailor-made’ wax esters. Such genetic engineering of FAR proteins would be aimed at both optimizing total fatty alcohol production, which can be done through enhancement of the catalytic site, and modifying substrate chain-length/saturation specificity. These engineered FAR enzymes could then be expressed in transgenic microbes or transgenic oilseed crops in conjunction with a wax synthase for mass production of these designer wax esters. For this to be possible, however, a greater understanding of FAR protein structure must be developed. There is currently no X-ray crystal structure of a FAR, which makes it difficult to predict which amino acids contribute to substrate preference. In the meantime, site directed mutagenesis can be used to identify residues that contribute to chain-length and chain-saturation specificity.

1.6 Concluding Remarks

Additional work needs to be done to fully understand the process of fatty alcohol synthesis and the role these fatty alcohols play biologically. To achieve this, we need to: (1)

Investigate the presence and biological roles of FARs in previously unexplored organisms. For example, database searches reveal putative mouse FAR orthologs in many species, including toad (*Xenopus laevis*), mosquito (*Anopheles gambiae*), rat (*Rattus norvegicus*), zebra fish (*Danio rerio*), and fruit fly (*Drosophila melanogaster*) (Cheng *et al.*, 2004). These homologous sequences range from 33% (mosquito) to 96% (rat) identical with the mouse FAR, suggesting there are many more FAR enzymes that can be investigated, (2) Elucidate the three-dimensional FAR protein structure by X-ray crystallography or NMR, and (3) Identify the amino acid residues that contribute to activity and substrate specificity through site-directed mutagenesis and/or directed evolution. A better understanding of FAR protein structure and activity will allow for the engineering of proteins with desired qualities for industrial use.

1.7 Project Rationale

The *Arabidopsis thaliana* genome encodes for eight FAR proteins (FAR1-FAR8), each with a distinct substrate specificity for saturated fatty acyl precursors with chain lengths ranging from C16-C30 (Rowland *et al.*, 2006; Doan *et al.*, 2009; Domergue *et al.*, 2010; Chen *et al.*, 2011; Doan *et al.*, 2011). *FAR5* (At3g44550) and *FAR8* (At3g44560) are located in tandem on chromosome 3 and encode for proteins that are 85% similar at the amino acid level (Figure 3.1B). *FAR5* generates 18:0 fatty alcohols (18:0-OH) found in the suberin polymer of roots, seed coats, and wounded leaves (Domergue *et al.*, 2010). *FAR5* also generates the 18:0 fatty alcohol moiety found in alkyl hydroxycinnamates of *Arabidopsis* root waxes (Kosma *et al.*, 2012). *FAR8* produces low levels of fatty alcohols from endogenous lipid pools when expressed in *E. coli* (Doan *et al.*, 2009). However, *FAR8* gene expression is nearly undetectable *in planta*, at least under non-stressed conditions (Domergue *et al.*, 2010). When expressed in yeast, *FAR5* and

FAR8 produce nearly exclusively 18:0-OH and 16:0-OH, respectively, although the amount of 16:0-OH produced by FAR8 is very low (see Results below). Since FAR5 and FAR8 have high sequence similarity but strict chain-length specificity for different fatty acyl precursors, they are an excellent platform for probing amino acids responsible for dictating substrate specificity. We have engineered chimeric proteins of FAR5 and FAR8 as well as single point mutants of FAR5 and FAR8, and have evaluated their activities by heterologously expressing them in yeast.

1.8 Thesis Objectives

The objective of this research was to identify which amino acids contribute to C16:0 versus C18:0 chain-length substrate specificity of the alcohol-forming fatty acyl reductases FAR5 and FAR8 from *Arabidopsis thaliana*. This information will be used to engineer FAR proteins with optimal activity for industrial applications.

CHAPTER 2: Materials and Methods

2.1 Construction of yeast expression plasmids containing Arabidopsis FAR5 and FAR8 variants

The coding regions of *FAR5* and *FAR8* were cloned between the *Bam*HI and *Xho*I restriction enzyme sites of a modified version of pYES2 (Invitrogen) called pYES2-His(6x)/T7 (Pulsifer *et al.*, 2012). This vector allows for galactose inducible expression in *Saccharomyces cerevisiae* (yeast) of proteins fused with an N-terminal T7 epitope tag for detection by protein immunoblotting (Western blots). To facilitate domain swaps, a silent mutation was introduced into the *FAR8* coding sequence to create a *Sal*I restriction site, which was present at the equivalent position in the *FAR5* coding sequence. The common *Sal*I restriction site was used to generate a pair of reciprocal domain swaps centered at positions corresponding to amino acid residues 283/284 of *FAR5* and *FAR8*. All other *FAR5* and *FAR8* domain swap constructs and site directed mutants were made by overlap extension PCR consisting of two sequential DNA amplifications. For each plasmid construct, PCR was first used to amplify two overlapping DNA fragments (<1000 bp) using the primer sequences found in Table 2.1. The cycling conditions of these PCRs were: 25 cycles of 98 °C for 10 sec, 52 °C for 15 sec, 72 °C for 35 sec, with a final single step of 72 °C for 10 min. The second amplification step used the two overlapping DNA fragments as templates and primers, along with flanking primers *FAR5_Bam*HI_Forward or *FAR8_Bam*HI_Forward and *FAR5_Xho*_Reverse or *FAR8_Xho*_Reverse such that the entire *FAR* open reading frames were amplified and flanked by restriction enzyme sites. The cycling conditions of the second PCRs were: 5 cycles of 98 °C for 15 sec, 52 °C for 25 sec, 72 °C for 45 sec, followed by 25 cycles of 98 °C for 15 sec, 52 °C for 15 sec, 72 °C for 45 sec, with a final extension at 72 °C for 10 min. All PCRs were carried out with iProof™ High-Fidelity DNA

polymerase (Bio-Rad). The amplified products were digested with *Bam*HI and *Xho*I and ligated between the same restriction sites in pYES2-His(6x)/T7.

Table 2.1 List of primers used in this study

Primer	Sequence
FAR5/FAR8 amplification	
FAR5_BamHI_For	GAGGGATCCATGGA ACTCAATTGTGTTCAAT
FAR5_XhoI_Rev	GCGCTCGAGTCACTTCTTAAGCACGTGTG
FAR8_BamHI_For	GAGGGATCCATGGAATTCAGTTGTGTTCA
FAR8_XhoI_Rev	GCGCTCGAGTTACTTCTTAAGCACGTGAG
Domain swap constructs	
FAR8-FAR5_388_For	CTGGCTTTGTT CAGCCTCTACATGACCCTA
FAR8-FAR5_388_Rev	TAGGGTCATGTAGAGGCTGAACAAAGCCAG
FAR5-FAR8_388_For	ATGGCTTTGTT CAGCCTCTACATGACCATA
FAR5-FAR8_388_Rev	TATGGTCATGTAGAGGCTGAACAAAGCCAT
FAR8-FAR5_353_For	GGAGAGATCCGTGAAATTGCGGTTTCGTTAC
FAR8-FAR5_353_Rev	GTAACGAACCGCAATTTACGGATCTCTCC
FAR5-FAR8_353_For	GGAGAGATTCATGAG ATTTTGT TTTGTTAC
FAR5-FAR8_353_Rev	GTAACAAAACAAAATCTCATGAATCTC
FAR8-FAR5_344_For	TCATCTCACCAAACCCAGTCACATTTGGA
FAR8-FAR5_344_Rev	TCCAAATGTGACTGGGTTTTGGTGAGATGA
Point mutation constructs	
FAR8_S363P_For	TCACCAAAAACCCGTTGCGCAGT
FAR8_S363P_Rev	ACTGCGCAACGGGTTTTTGGTGA
FAR5_P363S_For	TTACGAAAAACTCTTTGCGAAGT
FAR5_P363S_Rev	ACTTCGCAAAGAGTTTTTCGTAA
FAR8_I347T_For	GGATCTCTCCATATGTTATTGGGTTTTGGT
FAR8_I347T_Rev	ACCAAACCCAATAACATATGGAGAGATCC
FAR5_T347I_For	CAGAACCCAGTCATATTTGGAGAGATTCAT
FAR5_T347I_Rev	ATGAATCTCTCCAAATATGACTGGGTTCTG
FAR5_Y238F_For	GCCTAACACATTTGTTTTACCA
FAR5_Y238F_Rev	TGGTGAAAACAAATGTGTTAGGC
FAR5_K242I_For	TGTTTTACCATATCAATGGGAG
FAR5_K242I_Rev	CTCCATTGATATGGTGAAAACA
FAR5_A355L_For	ATTCATGAGATTTTGGTTCGTTACTTTACGAAA
FAR5_A355L_Rev	TTTCGTAAAGTAACGAACCAAATCTCATGAAT
FAR5_V377M_For	ACCGTCTCAAAAATGAGGTTTCATACCAACCATG
FAR5_V377M_Rev	CATGGTTGGTATGAACCTCATTTTTGAGACGGT
FAR8_M377V_For	ACTGTCTCGAAAGTGAAGCTGATACCA
FAR8_M377V_Rev	TGGTATCAGCTTCACTTTTCGAGACAGT
FAR8_L355A_For	ATCCGTGAAATTGCGTTTTGTTACTTC
FAR8_L355A_Rev	GAAGTAACAAAACGCAATTTACGGAT

2.2 Expression of FAR5 and FAR8 variants in yeast

The pYES2-His(6x)/T7 plasmids containing wild-type and mutant *FAR5* and *FAR8* coding regions were transformed into *Saccharomyces cerevisiae* yeast strain W303-1A (*MATa his3Δ1 leu2 trp1-289 ura3-52*) using the method of Gietz and Woods (2002). Yeast transformants were selected on synthetic dropout (SD) media plates (2% *D*-glucose, 0.67% yeast nitrogen base, 0.01% adenine, 0.002% histidine, 0.002% tryptophan, 0.01% leucine, and 2% agar) lacking uracil (SD^{-Ura}). Four yeast colonies for each construct, including empty vector, were separately inoculated into 3 ml of SD^{-Ura} liquid media and grown for 24 hr at 30 °C and 250 rpm. The OD₆₀₀ for each culture was measured and a volume of culture required to inoculate 3 ml of induction SG^{-Ura} liquid media (2% galactose, 0.67% yeast nitrogen base, 0.01% adenine, 0.002% histidine, 0.002% tryptophan, 0.01% leucine) at an OD₆₀₀ of 0.4 was removed and centrifuged. The resulting yeast pellets were washed twice with 1.5 ml of sterile water and then resuspended in 3 ml of SG^{-Ura} media. The yeast strains were then grown for 2 days at 30 °C and 250 rpm prior to lipid extraction and analysis of protein levels by Western blotting.

2.3 Analysis of protein levels in yeast by protein immunoblotting (Western blots)

A volume of each galactose-induced yeast culture equal to an OD₆₀₀ of 2.5 was harvested by centrifugation. The supernatants were removed and the pellets stored at -80 °C. Protein samples were prepared according to Kushnirov (2000) using phosphate-containing loading buffer (62.5 mM sodium phosphate buffer pH 7.0, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.0001% bromophenol blue) prior to separation on a 12% SDS polyacrylamide gel at 150 V in 1X running buffer (25 mM Tris pH 8.3, 186 mM glycine, 0.1% SDS). The proteins were transferred to nitrocellulose membranes by wet-transfer electrophoresis. The nitrocellulose

membranes were blocked overnight at 4 °C in blocking solution consisting of 5% fat free skim milk in TBST (137 mM NaCl, 25 mM Tris Base pH 7.6, 0.1% Tween-20). A 1:50 000 dilution of T7 Tag Monoclonal Mouse Antibody (catalog number 69522-3, EMD4Biosciences) was added to detect the T7-epitope tagged proteins and the membranes were washed four times for 5 min each with TBST. The membranes were then incubated with a 1:50 000 dilution of horseradish peroxidase-conjugated anti-mouse secondary antibody (EMD4Biosciences) for 1 hr, and then rinsed four times with TBST for 5 min each. The blots were then incubated with a 1:1 mixture of lumigen TMA-6 Solution A (contains Tris Buffer in 3.2% v/v ethanol) and lumigen TMA-6 Solution B (proprietary substrate in Tris Buffer) (GE Healthcare) for 5 min in the dark and imaged with a FluorChemQ (Alpha Innotech). Western blots with yeast microsomes were done with 15 µg of protein and revealed with the ECL Western blotting detection kit (Amersham Biosciences).

2.4 Analysis of fatty alcohol content of transgenic yeast by gas chromatography (GC)

Yeast cells from 2 ml of each galactose-induced culture were harvested by centrifugation and the supernatant was poured into a separate GC tube. 10 µg of 1-pentadecanol (15:0-OH) was added as an internal standard to both the pellet and supernatant. The yeast supernatant was extracted twice with 1 ml of 2:1 chloroform/methanol and once with 1 ml of chloroform, the organic phases were combined and washed with 2.5 ml of 0.9% NaCl (w/v) before being evaporated under nitrogen gas at 37 °C. Both the yeast pellet and supernatant samples were resuspended in 3 ml of 1 M methanolic-HCl and lipids transmethylated at 80 °C for 90 min. 0.9% NaCl was added to each sample and the lipids were extracted twice into 500 µl of hexane. The pooled hexane extracts were dried under nitrogen gas at 37 °C. The lipids were then

resuspended in 100 μ l of *N,O*-bis (trimethylsilyl) trifluoroacetamide (BSTFA) plus 100 μ l of pyridine and incubated at 110 $^{\circ}$ C for 15 min for silylation of hydroxyl groups. The samples were then re-dried down as before and resuspended in a 1:1 (v/v) mixture of hexane:toluene. Lipids were quantified with a Varian GC450 equipped with a split-splitless injector, HP-1 column (30 m length, 0.25 mm inner diameter, 0.10 μ m film thickness), and a flame ionization detector. One μ l of each sample was injected using a 10:1 split ratio. The carrier gas was helium with a constant flow rate of 2 mL/min. The column oven was held initially at 150 $^{\circ}$ C for 5 min, then ramped at 10 $^{\circ}$ C/min to 300 $^{\circ}$ C, and held for 8 min, for a total run time of 28 min.

2.5 Yeast microsomes preparations and FAR in vitro enzyme assays

Yeast microsomes were prepared according to Yang *et al.* (2010) and FAR activity was assayed with 5 μ g of protein for 5 min at 30 $^{\circ}$ C in 50 μ l sodium phosphate buffer (50 mM, pH 6.5) containing 5 mM NADPH, 20 μ M [14 C]-fatty acyl-CoA (palmitoyl- or stearoyl-CoA), 1 mM MgCl₂, and 2 mg/ml fatty acid-deficient bovine serum albumin. Reactions were stopped by adding 600 μ l of 1% HClO₄ and lipids were extracted by the successive addition of 2 ml of chloroform / methanol (1:2; v/v), 1 ml of chloroform, and 0.8 ml of 1% HClO₄. The lower organic phases were dried under nitrogen, and analyzed by thin-layer chromatography using HPTLC Silica Gel 60 plates (Merck) and hexane/ether/acetic acid (70:30:2, v/v/v) as the solvent. Radiolabeled products were identified by co-migration with unlabeled standards and quantified by autoradiography using an Instant Imager (Packard).

CHAPTER 3: Results

3.1 Amino acids of FAR5 and FAR8 important for protein stability and enzymatic activity

We examined the activities and substrate specificities of the FAR5 and FAR8 enzymes using *Saccharomyces cerevisiae* as a heterologous system. The open reading frames for both genes were cloned into the yeast expression vector pYES2-His(6x)/T7 downstream of the *GALI* promoter to allow high-level protein induction by galactose. The expressed FAR proteins contained a T7 epitope tag at the N-terminus to enable detection using Western blots. The empty pYES2-His(6x)/T7 vector was used as a negative control (Figures 3.3, 3.5, 3.6, 3.7, 3.8). The endogenous acyl-CoA pool of *S. cerevisiae* has high and nearly equal amounts of 16:0-CoA and 18:0-CoA (Domergue *et al.*, 2003), thus providing relevant substrates for heterologously expressed FAR5 and FAR8 variants.

Analysis of the internal lipid content of 2-days old galactose induced yeast cultures by gas chromatography (GC) demonstrated that FAR5 produced high levels of 18:0 primary fatty alcohol (2.07 μg per unit OD) and that the protein was highly expressed in yeast as determined by Western blot analysis (Figure 3.2, 3.3). In contrast, FAR8 did not produce any fatty alcohols under the same conditions and the protein was barely detected in the Western blot (Figure 3.2, 3.3). However, when transgenic yeast expressing FAR8 were grown for 4 days prior to lipid analysis, low levels of 16:0 fatty alcohol were detected (Figure 3.4). Similarly, very small amounts of 16:0-OH were sometimes detected in yeast expressing FAR5 for 2 days (Figure 3.6A) and these amounts increased after 4 days of FAR5 expression albeit to levels less than 4% of the accumulated 18:0-OH (Figure 3.4).

With the exception of Arabidopsis FAR8, all plant FAR proteins reported have a proline at relative position 363 of FAR5 and FAR8. Instead, FAR8 has a serine at this position (Figure 3.1B). To investigate the significance of the conserved proline residue, serine-363 from FAR8 was converted to proline (FAR8-S363P). Unlike native FAR8 protein, the FAR8-S363P variant generated substantial amounts of 16:0-OH (0.85 μg per unit OD) (Figure 3.2, 3.3, 3.4). A small amount of 18:0-OH was sometimes detected, but this quantity was negligible compared to that of 18:0-OH produced by FAR5 (<1%) or that of 16:0-OH produced by FAR8-S363P. Analysis of protein levels in yeast revealed that FAR8-S363P accumulated to higher levels than its native counterpart (Figure 3.3B). For much of our subsequent mutational analyses, this active FAR8-S363P mutant was used as the model of 16:0 fatty alcohol production. We subsequently refer to this active FAR8 variant as FAR8*R*, where *R* stands for ‘*resurrected*’. The reciprocal amino acid conversion in FAR5, proline to serine conversion at position 363, resulted in a 30% decrease in 18:0-OH production (Figure 3.3A), although the FAR5-P363S and native FAR5 protein levels in yeast were comparable (Figure 3.3B). It was possible that FAR8 or FAR8*R* had activity for a shorter acyl chain length than 16:0, with the most likely possibility being 14:0. Therefore, we also tested whether FAR5, FAR8 and/or FAR8*R* could generate 14:0-OH by feeding yeast 14:0 fatty acid. This approach was previously used to demonstrate activity in yeast of *Euglena gracilis* FAR for 14:0 acyl chain length (Teerawanichpan and Qiu, 2010). No 14:0-OH was produced by FAR5 or FAR8, and only a very small amount by FAR8*R* in our feeding experiments (Figure 3.4). We cannot rule out other potential substrates for FAR8, but the most likely scenario is that FAR8 is indeed a low activity FAR due to the serine rather than proline at position 363 and that the specificity of FAR8*R* is nearly strict for 16:0-CoA.

The 18:0 and 16:0 fatty alcohols produced by the FAR5 and FAR8R variants, respectively, were found not only within the yeast cells but also secreted into the media (Figure 3.5). Of the total fatty alcohols produced by FAR5 and FAR8R in yeast, 47% and 38%, respectively, were found secreted into the media (Figure 3.5). For ease of screening purposes (below), we only quantified the fatty alcohols found in the yeast pellet. We verified that no bias was introduced by excluding the secreted fatty alcohols in our screens by measuring the internal and secreted fatty alcohols of FAR5 and FAR8R, as well as some key FAR5 and FAR8 variants that alter substrate specificity as described below (Figure 3.5). The chain lengths distributions of fatty alcohols produced by these FARs expressed in yeast were nearly identical whether examining just internal fatty alcohols or examining total fatty alcohols (internal plus secreted).

We next investigated the proposed roles of tyrosine-238 and lysine-242 in the predicted YXXXXK active site motif of the Rossmann-fold in FAR proteins (Figure 3.1) (Doan *et al.*, 2009; Rowland and Domergue, 2012). Two separate site-specific substitutions were made in the FAR5 coding sequence to convert tyrosine-238 to a phenylalanine (FAR5-Y238F) and separately lysine-242 to an isoleucine (FAR5-K242I) in the encoded protein. These same amino acid substitutions were used to investigate the active site of alcohol dehydrogenase from *Drosophila melanogaster* (Chen *et al.*, 1993), revealing the importance of the hydroxy group of tyrosine and the amino group of lysine in catalysis. Phenylalanine differs from tyrosine only by the absence of a hydroxyl group on the aromatic ring. Lysine and isoleucine are similarly sized amino acids, but isoleucine lacks the positive charge conferred by the amino group on lysine. Lipid analysis of both the FAR5-Y238F and FAR5-K242I variants expressed in yeast revealed no fatty alcohol production (Figure 3.3A), although Western blot analysis of the mutant proteins indicated that they were expressed to the same levels as the wild-type FAR5 (Figure 3.3B).

Altogether, these data showed that FAR5 and FAR8 nearly exclusively produce 18:0-OH and 16:0-OH, respectively, that proline-363 is important for protein stability, and that tyrosine-238 and lysine-242 are necessary for catalytic activity.

3.2 Domain swaps between FAR5 and FAR8

A series of domain swap chimeras involving FAR5 and FAR8 were then made to investigate the protein region responsible for conferring substrate specificity. An initial reciprocal pair of domain swaps were made between FAR5 and FAR8R at positions corresponding to amino acid residues 283/284 (Figure 3.1B, 3.6A). The FAR5₁₋₂₈₃FAR8R₂₈₄₋₄₉₆ chimera [1] produced predominantly 16:0-OH. Conversely, the FAR8₁₋₂₈₃FAR5₂₈₄₋₄₉₆ chimera [2] produced substantial amounts of 18:0-OH and very little 16:0-OH (Figure 3.6A). This indicated that C16/C18 chain-length specificity is dictated by amino acids between position 284 and the C-terminus of the FAR5/FAR8 proteins. To further delineate this region, domain swaps between FAR5 and FAR8 were made at amino acid residues 388/389 (Figure 3.1B, 3.6A). We found that the FAR5₁₋₃₈₈FAR8₃₈₉₋₄₉₆ chimera [3] produced nearly exclusively 18:0-OH, while the FAR8R₁₋₃₈₈FAR5₃₈₉₋₄₉₆ chimera [4] had low activity, but it produced more 16:0-OH than 18:0-OH (Figure 3.6A). Collectively, the data from these four domain swaps indicate that C16/C18 chain-length specificity is dictated by amino acids between positions 284 and 388 of the FAR5 and FAR8 proteins.

To examine this further, we made two FAR5/FAR8 internal domain swap mutants spanning residues 284 to 388, inclusively (Figure 3.6A). The FAR5₁₋₂₈₃FAR8R₂₈₄₋₃₈₈FAR5₃₈₉₋₄₉₆ [5] and FAR8₁₋₂₈₃FAR5₂₈₄₋₃₈₈FAR8₃₈₉₋₄₉₆ [6] internal domain swap chimeras produced nearly exclusively 16:0-OH and 18:0-OH, respectively (Figure 3.6A). This verified that the amino acid residues responsible for 16:0 fatty acyl-CoA versus 18:0 fatty acyl-CoA substrate specificity are

between residues 284 through 388. To further narrow down the protein region responsible for conferring chain length specificity, two more reciprocal internal domain swap constructs were made spanning residues 354 to 388, inclusively (Figure 3.6A). The FAR5₁₋₃₅₃FAR8R₃₅₄₋₃₈₈FAR5₃₈₉₋₄₉₆ chimera [7] produced nearly exclusively 16:0-OH. The FAR8₁₋₃₅₃FAR5₃₅₄₋₃₈₈FAR8₃₈₉₋₄₉₆ chimera [8] appeared to have very low activity for either fatty acyl-CoA substrate, but it produced more 18:0-OH than 16:0-OH. This suggests that the amino acid residues responsible for dictating 16:0 versus 18:0 acyl chain-length specificity in FAR5 and FAR8 are found within the 35 amino acids interval spanning residues 354 to 388.

The chimeras that contained FAR5 sequence at the N-terminal end [1 + 3] accumulated to higher levels in yeast than the chimeras that contained FAR8 at the N-terminal end [2 + 4] (Figure 3.6B). The internal domain swap chimeras were generally expressed to similar levels as the FAR8R protein. An exception to this is FAR8₁₋₃₅₃FAR5₃₅₄₋₃₈₈FAR8₃₈₉₋₄₉₆ [8], which gave very little fatty alcohol production and had near undetectable protein accumulation, and FAR5₁₋₃₅₃FAR8R₃₅₄₋₃₈₈FAR5₃₈₉₋₄₉₆ [7], which generated high amounts of fatty alcohols and was expressed at high levels similar to that of FAR5 (Fig. 4B). Additionally, this latter chimera [7] accumulated to higher levels than the FAR5₁₋₂₈₃FAR8R₂₈₄₋₃₈₈FAR5₃₈₉₋₄₉₆ chimera [5], which had FAR8 sequence between residues 284 and 353. We therefore speculated that some residues immediately N-terminal to position 354 are critical for protein stability. To test this, another FAR8-FAR5-FAR8 domain swap chimera was expressed that contained residues 345 through 388 of FAR5. This FAR8₁₋₃₄₄FAR5₃₄₅₋₃₈₈FAR8₃₈₉₋₄₉₆ chimera [9] accumulated to relatively high levels (Figure 3.6B), indicating that one or more amino acids between positions 345 and 353 are important for protein stability. Of the four amino acids that are different between FAR5 and FAR8 in this region, only position 347 is a non-conservative amino acid difference (Figure

3.1B). Residue 347 is a threonine in FAR5 whereas it is an isoleucine in FAR8. We found that converting the isoleucine present at position 347 in FAR8 to the reciprocal threonine of FAR5 caused the resultant FAR8R-I347T protein to accumulate to higher levels in yeast than the parent FAR8R (Figure 3.3B). The amount of 16:0-OH produced by FAR8R-I347T was 1.9 times that of FAR8R (Figure 3.3A). The presence of a threonine at position 347 in FAR5 thus likely promotes protein stability. To further investigate the role of threonine at position 347, we made the reverse mutation in FAR5 converting threonine-347 to isoleucine. The resultant FAR5-T347I mutant had a 40% decrease in 18:0-OH production (Figure 3.3A). This was paralleled by decreased FAR5-T347I protein levels compared to wild-type FAR5 (Figure 3.3B), again indicating that threonine-347 plays an important role in protein stability.

3.3 Reciprocal amino acid substitutions in FAR5 and FAR8

To determine individual amino acid residues of FAR5 and FAR8 responsible for dictating chain-length substrate specificity, site specific substitutions were made within the 44 amino acid region (residues 345 to 388) of FAR5 and FAR8 identified by internal domain swaps to ‘flip’ substrate specificity and be important for protein stability (Figure 3.7, 3.8). Within this region, there are 13 residues that are different between FAR5 and FAR8, and we targeted 8 of these residues for site-specific mutagenesis. Positions 347 and 363 had already been examined (above), and found to be important for protein stability/activity but did not influence chain-length specificity. In all cases, we converted a given amino acid found at a certain position in one FAR to the reciprocal amino acid found at that same position in the other FAR.

Within the 345 to 388 region, there are two triplet blocks of amino acids, residues 355-357 (triplet I) and 377-379 (triplet II), that are not conserved between FAR5 and FAR8 (Figure

3.7A). A triple mutation of FAR5 converting the alanine-valine-arginine (AVR) of triplet I to the corresponding FAR8 residues leucine-phenylalanine-cysteine (LFC) resulted in a protein, FAR5-AVR→LFC, which produced both 16:0 and 18:0 fatty alcohols (Figure 3.7B). Compared to wild-type FAR5, the amount of 18:0-OH produced by FAR5-AVR→LFC was about 40% less and it produced similar levels of 16:0-OH as FAR8R. Similarly, a triple mutation of FAR5 converting the valine-arginine-phenylalanine (VRF) of triplet II to the corresponding FAR8 residues methionine-lysine-leucine (MKL) resulted in a protein that produced both 16:0 and 18:0 fatty alcohols at similar levels to FAR5-AVR→LFC (Figure 3.7B). We then individually mutated each of the six amino acids in triplets I and II of FAR5 to the corresponding FAR8 amino acid, and found that amino acid conversions at positions 355 and 377 of triplet I and II, respectively, resulted in the most 16:0-OH production with the other four conversions causing little or no 16:0-OH production (Figure 3.7B, 3.8B). The FAR5-A355L single mutant produced 16:0-OH to levels similar to that of FAR8R and produced 18:0-OH to about 66% of that observed with wild-type FAR5 (Figure 3.8B). The FAR5-V377M single mutant also produced both 16:0-OH and 18:0-OH, but the amount of 16:0-OH produced was about 25% of FAR8R whereas the amount of 18:0-OH produced was similar to wild-type FAR5 (Figure 3.8B). By combining the mutations, we were able to generate a FAR5 variant, FAR5-A355L V377M, which produced mostly 16:0-OH (Figure 3.8B). The FAR5-A355L V377M mutant had 37% greater 16:0-OH yield than FAR8R, although it produced a small amount of 18:0-OH (approximately 6% of the amount produced by wild-type FAR5). These single and double FAR5 variants were all expressed to similar levels as wild-type FAR5 in yeast (Figure 3.8B).

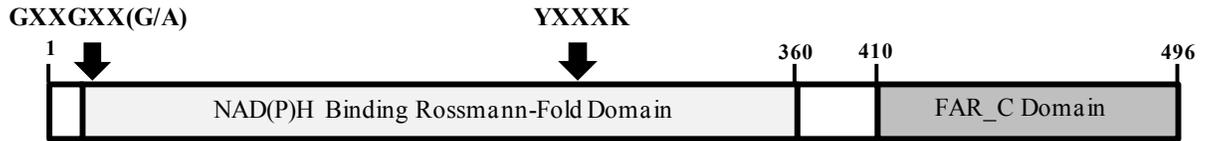
We then examined the effects of reciprocal amino substitutions at positions 355 and 377 of FAR8 (Figure 3.8C). The FAR8R-L355A and FAR8R-M377V mutants had greatly reduced

levels of 16:0 fatty alcohol production, approximately 8% and 18% of FAR8R levels, respectively. These mutants also produced 18:0-OH, albeit at very low levels compared to FAR5. Since the I347T mutation had a major effect on FAR8 protein stability/activity (above), we combined this change with the M377V change. The resultant FAR8R-I347T M377V double mutant produced about twice the amount of 16:0-OH than the parent FAR8R, although slightly less than FAR8R-I347T (Figure 3.3A, 3.8C). FAR8R-I347T M377V additionally produced high levels of 18:0-OH (0.63 μg per unit OD). As predicted, FAR8R-I347T M377V protein accumulated at much higher levels in yeast than FAR8R-M377V, likely accounting for the increased fatty alcohol production. In an attempt to further convert FAR8 substrate preference from 16:0-CoA to 18:0-CoA, we added a third L355A mutation to create FAR8R-I347T L355A M377V. This triple mutation nearly completely converted FAR8R specificity to 18:0-CoA. The FAR8R-I347T L355A M377V mutant produced 18:0-OH to levels about 60% of wild-type FAR5. The FAR8R-I347T L355A M377V mutant produced a small amount of 16:0-OH, but it was only about 30% of that found with FAR8R. In summary, we altered FAR8 chain length specificity from 16:0-CoA nearly completely to 18:0-CoA by changing amino acids at two positions, 355 and 377, to that of the corresponding FAR5 amino acids, but amino acids changes were also required at positions 347 and 363 to allow for high protein accumulation.

3.4 In vitro FAR assays using yeast microsomes

To provide additional support of our findings, we prepared microsomes from different transgenic yeast strains and performed *in vitro* assays using either [1-¹⁴C]-palmitoyl-CoA (16:0-CoA) or [1-¹⁴C]-stearoyl-CoA (18:0-CoA) as substrate. After 5 min of incubation in the presence of 5 μg microsomal proteins, lipids were extracted and analyzed by thin layer chromatography.

Microsomes containing FAR5 produced labeled fatty alcohols in the presence of [1-¹⁴C]-C18:0-CoA, whereas no radioactivity associated with fatty alcohols was observed using microsomes from yeast transformed with empty vector, regardless of the labeled fatty acyl-CoA used (Figure 3.9A). In the presence of [1-¹⁴C]-C16:0-CoA, FAR5 produced very little fatty alcohol. Conversely, FAR5-A355L V377M produced significant amounts of fatty alcohols with [1-¹⁴C]-C16:0-CoA as substrate, in agreement with the results obtained *in vivo*. In contrast to the microsomes containing FAR5 variants, no fatty alcohols were detected when using microsomes from yeast transformed with FAR8R, FAR8R-I347T M377V or FAR8R-I347T L355A, whatever the labeled fatty acyl-CoA used (Figure 3.9A). The reason that these microsomes were inactive *in vivo* was most probably due to protein stability, as none of the FAR8 variants could be detected in the microsomes by Coomassie blue staining or by Western blots in contrast to the FAR5 variants (Figure 3.9B). The FAR8 variants may also be deficient in their ability to associate with membranes. These *in vitro* assays nevertheless allowed for quantification of the specific activities of FAR5, FAR5-A355L and FAR5-A355L V377M for 16:0-CoA and 18:0-CoA (Table 3.1). FAR5 was 20 times more active *in vitro* with 18:0-CoA than with 16:0-CoA. The FAR5-A355L single mutant was able to convert both 16:0-CoA and 18:0-CoA to the corresponding fatty alcohols, with its specific activity for 16:0-CoA being about half of that for 18:0-CoA. The FAR5-A355L V377M double mutant had a clear preference for 16:0-CoA, with its specific activity for 16:0-CoA being 4.7 times that for 18:0-CoA (Table 3.1). Altogether, the *in vitro* assays using FAR5, FAR5-A355L, and FAR5-A355L V377M were in good agreement with the results obtained *in vivo*.

A**B**

FAR5	1	ME	LN	CV	QF	LR	NK	TIL	VT	GAT	GFL	AK	VF	VE	KIL	RV	QPN	VK	KL	YL	LV	VR	AS	DN	EA	AT	KR	LR	TE	V	FE	KE	L	FK	VL	R	ON	LG
FAR8	1	ME	FS	CV	HF	LQ	NK	TIL	VT	GAT	GFL	AK	VF	VE	KIL	RV	QPN	VN	KL	YL	LV	VR	AS	DN	EA	AT	KR	LR	TE	A	FE	KD	L	FK	VL	R	DN	LG

GXXGXX(G/A) NAD(P)H binding motif

FAR5	76	DE	KL	NT	LL	Y	E	K	V	V	S	V	P	G	D	I	A	T	D	O	L	G	I	N	D	S	H	L	R	E	R	M	O	K	E	I	D	I	V	V	N	V	A	A	T	T	N	F	D	E	R	Y	D	V	G	L	G	I	N	T	F	G	A	L	N	V	L	N	F	A	K
FAR8	76	DE	KL	NT	LL	S	E	K	V	V	P	V	A	G	D	I	A	M	D	H	L	G	M	K	D	S	N	L	R	E	R	M	O	K	E	I	D	I	V	V	N	V	A	A	T	T	N	F	D	E	R	Y	D	I	G	L	G	I	N	T	F	G	A	L	N	V	L	N	F	A	K

YXXXK active site

FAR5	151	CV	K	V	Q	L	L	H	V	S	T	A	Y	V	C	G	E	K	P	G	L	I	P	E	K	P	F	I	M	E	E	T	R	N	E	N	G	L	O	L	D	I	N	L	E	R	E	L	M	Q	R	L	K	E	L	N	E	Q	D	C	S	E	E	D	I	T	L	S	M	K	E	L	G
FAR8	151	CV	K	A	Q	L	L	H	V	S	T	A	Y	V	C	G	E	K	P	G	L	I	P	E	K	P	F	V	M	E	E	T	C	N	E	N	G	L	O	L	D	I	N	L	E	R	E	L	M	Q	R	L	K	E	L	N	E	Q	G	C	S	E	E	G	T	F	Y	M	K	E	L	G	

FAR5	226	M	E	R	A	K	L	H	G	W	P	N	T	Y	V	F	T	K	S	M	G	E	M	L	L	G	K	H	K	E	N	L	P	L	V	I	R	P	T	M	I	T	S	T	S	E	P	F	P	G	W	I	E	G	L	R	T	V	D	S	V	I	A	Y	G	K	G	V	L	K	C	F	L
FAR8	226	M	E	R	A	K	L	H	G	W	P	N	T	Y	V	F	T	K	S	M	G	E	M	L	L	G	N	H	K	E	N	L	P	L	V	I	R	P	T	M	I	T	S	T	S	E	P	F	P	G	W	I	E	G	L	R	T	V	D	S	V	I	A	Y	G	K	G	V	L	K	C	F	L

FAR5	301	V	D	V	N	S	V	C	D	M	I	E	V	D	M	V	A	N	A	M	T	A	A	A	K	H	A	G	G	S	G	V	H	M	V	Y	H	V	G	S	S	H	Q	N	P	V	T	E	G	E	I	H	E	L	A	V	R	Y	F	T	K	N	S	L	R	S	R	N	G	S	I	T	V	S		
FAR8	301	V	D	V	N	S	V	C	D	M	I	E	P	A	D	M	V	A	N	A	M	T	A	A	A	T	H	A	G	G	S	K	V	H	M	V	Y	H	V	G	S	S	H	Q	N	P	I	I	Y	G	E	I	R	E	I	L	F	C	Y	F	T	K	N	S	L	R	S	R	N	G	S	M	I	T	V	S

FAR5	376	K	V	R	F	I	P	T	M	A	L	F	S	L	Y	M	T	L	R	Y	K	L	P	Q	L	L	K	L	V	D	I	I	Y	P	W	R	N	G	D	K	Y	G	D	K	N	R	K	I	E	L	V	M	R	L	V	E	L	Y	E	P	Y	V	L	F	K	G	I	F	D	D	R	N	T	K	S
FAR8	376	K	M	K	L	I	P	T	L	A	L	F	S	L	Y	M	T	L	R	Y	K	L	P	Q	L	L	K	L	V	D	I	I	Y	P	S	R	E	G	D	E	Y	K	N	K	R	K	I	D	M	V	M	R	L	V	E	L	Y	E	P	Y	V	L	F	K	G	I	F	D	D	R	N	T	K	N	

FAR5	451	L	C	A	N	O	K	E	E	I	K	N	T	E	K	L	M	F	D	F	D	P	K	I	N	W	G	D	Y	L	T	N	I	H	I	S	G	L	V	T	H	V	L	K	K	
FAR8	451	L	C	A	R	O	K	E	E	I	N	R	N	S	E	N	F	M	F	D	F	D	P	K	I	K	W	K	D	Y	L	I	N	V	H	I	P	G	L	I	T	H	V	L	K	K

Figure 3.1 FAR structural domains and protein sequence alignment of Arabidopsis FAR5 and FAR8. A. Schematic of the structural domains of FAR proteins. FAR enzymes, minus possible subcellular localization signals (e.g. plastid targeting/peroxisome anchoring sequence), are approximately 500 amino acids in length and contain a NAD(P)H binding Rossmann-fold domain (light grey) and a FAR_C domain (dark grey) at the N-terminus and C-terminus, respectively. The GXXGXX(G/A) predicted NADPH binding motif as well as the predicted active site motif YXXXK are indicated (where ‘X’ represents any amino acid). B. Protein sequence alignment of FAR5 and FAR8, which are 85% similar at the amino acid level. Identical amino acids are highlighted in black and physiochemically similar amino acids are highlighted in grey. Arrows indicate domain swap sites and asterisks indicate site-specific mutagenesis sites.

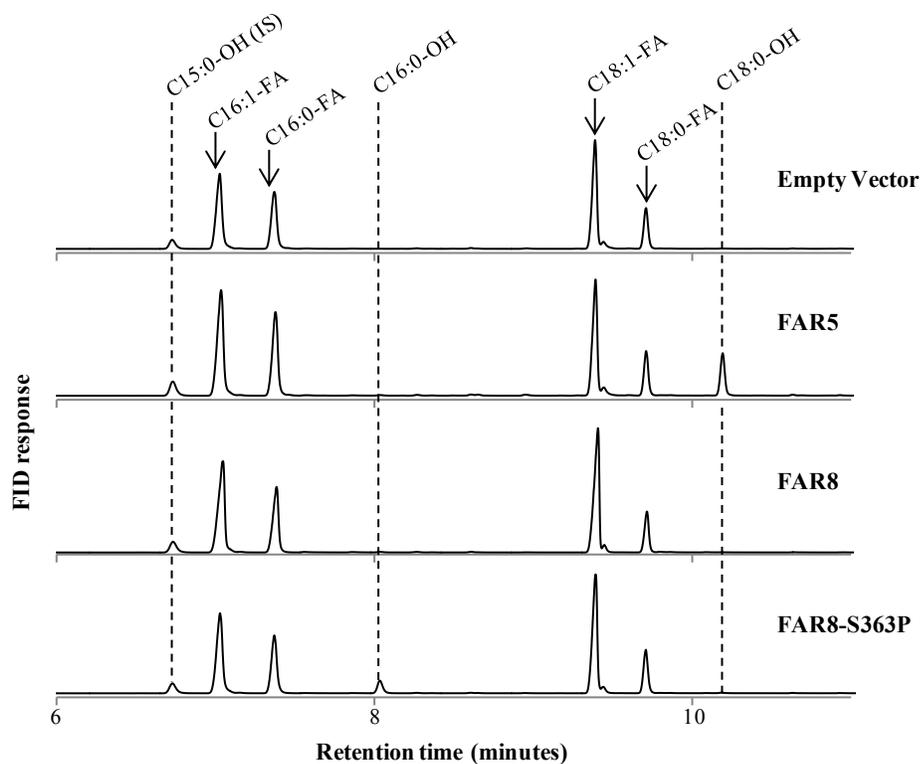


Figure 3.2 Gas chromatograms of internal lipids of yeast expressing Arabidopsis FAR5, FAR8, or FAR8-S363P. The empty vector pYES2-His6x/T7tag acted as a negative control. Transformants were cultured in galactose media to induce protein expression. Fatty acids were transmethylated to their corresponding methyl esters and fatty alcohol hydroxyl groups were derivitized to trimethylsilyl ethers before separation by GC and detection by flame ionization. The peaks corresponding to pentadecanol (C15:0-OH) internal standard (IS), C16:0-OH and C18:0-OH are indicated, as well as saturated and monounsaturated C16 and C18 fatty acids (FA).

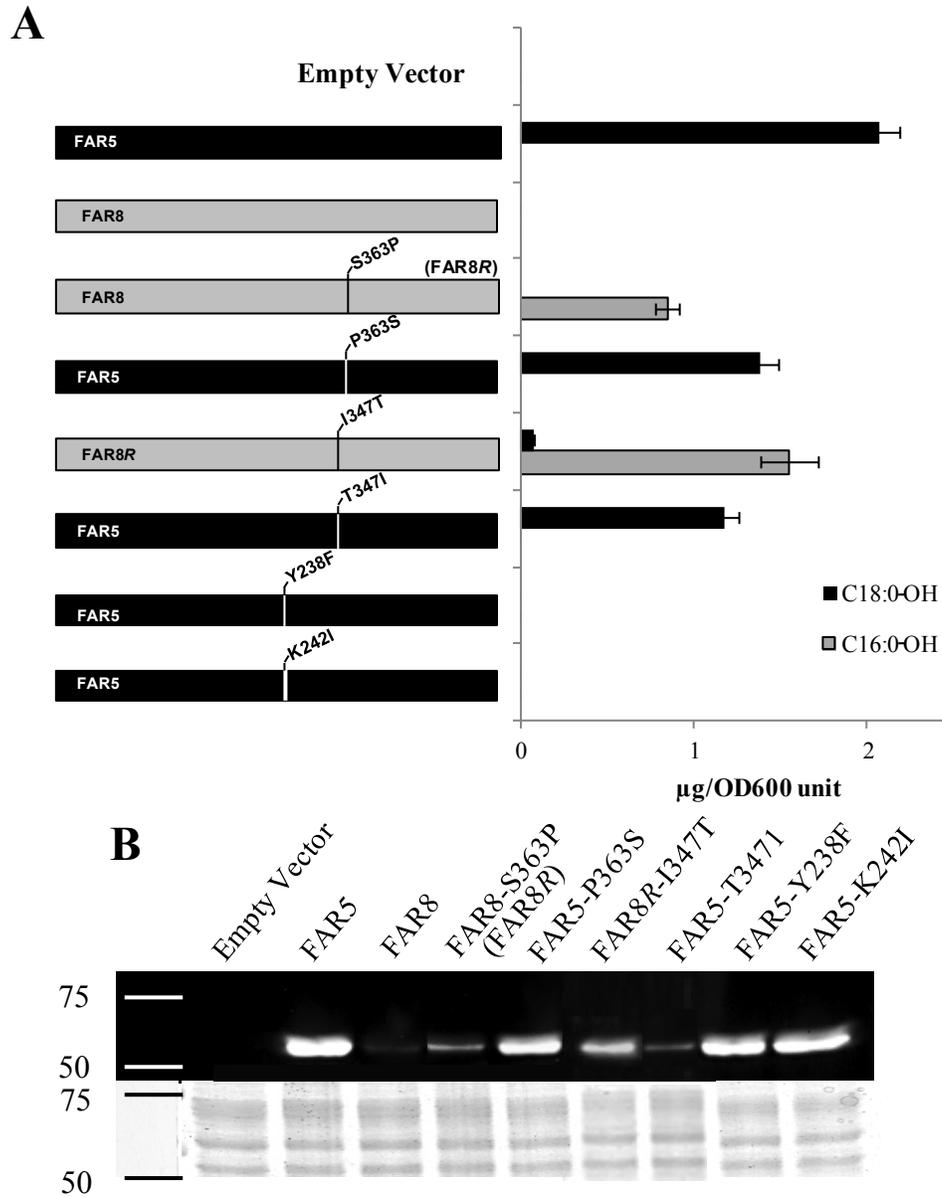


Figure 3.3 Amino acids important for FAR5 and FAR8 enzyme stability and activity. A. At left, schematics of FAR5 and FAR8 variants. The FAR5 variants are in black and FAR8 variants are in grey, with positions of site-specific mutations indicated above each protein schematic. At right, the amounts of non-secreted fatty alcohols produced by yeast-expressing FARs, where values are expressed in $\mu\text{g}/\text{OD600 unit}$ + standard deviation ($n=4$). B. Protein levels of FAR5 and FAR8 variants expressed in yeast. Western blots (top half) were performed using anti-T7 mouse antibody to detect the amino-terminal T7 epitope in the protein fusions. The Coomassie-stained gel is shown in the bottom half to indicate equal loading. The positions of the protein size markers (in kDa) are indicated to the left of the Western blot and stained gel.

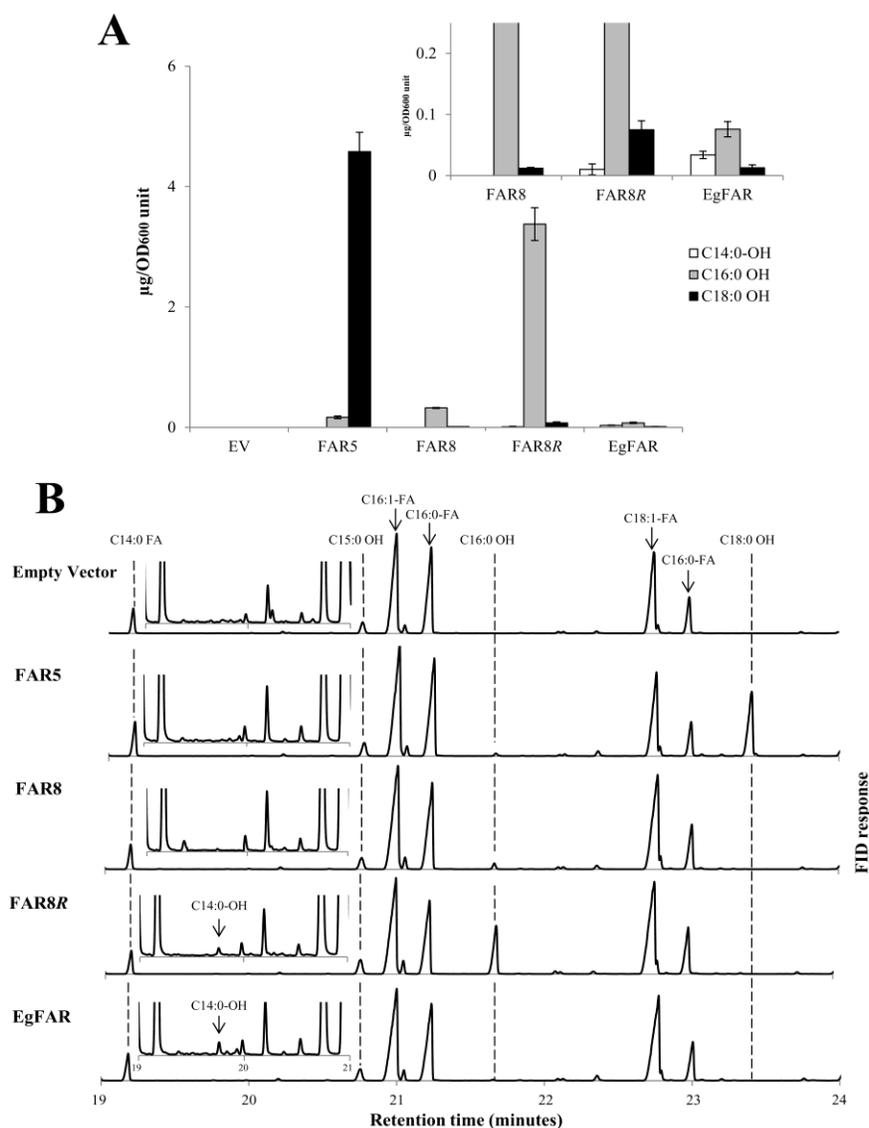


Figure 3.4 Fatty alcohol production by FAR5, FAR8, FAR8R (FAR8-S363P) and *Euglena gracilis* FAR (EgFAR) expressed in yeast for four days and fed with 14:0 fatty acid. A. Quantification of internal fatty alcohols produced by yeast expressing each of the FAR enzymes. The negative control was empty vector (EV). Yeast were grown, induced to express a FAR using galactose, and fed 14:0 fatty acid according to Teerawanichpan and Qiu (2010). Values are expressed in $\mu\text{g}/\text{OD}_{600}$ Unit + standard deviation ($n=4$). No 14:0-OH was produced by FAR8 in these feeding experiments. A low amount of 14:0-OH was produced by FAR8R in 14:0-fed yeast after 4 days of expression, but this was negligible compared to the amount of 16:0-OH produced by FAR8R. The amount of 14:0-OH produced by FAR8R was about 30% that of the *Euglena gracilis* FAR in the side-by-side comparison. B. Representative gas chromatograms of internal lipid content of yeast expressing each of the FAR enzymes. Fatty acids were transmethylated to their corresponding methyl esters and fatty alcohol hydroxyl groups were derivatized to trimethylsilyl ethers before separation by GC. The peaks corresponding to 14:0 fatty acid (C14:0 FA), 15:0-OH (internal standard), 16:0-OH and 18:0-OH are indicated by dashed lines. The small peak of 14:0-OH in the FAR8R and EgFAR chromatographs is indicated by an arrow in the zoomed section of the trace, which corresponds to the 19 to 21 minute retention time interval.

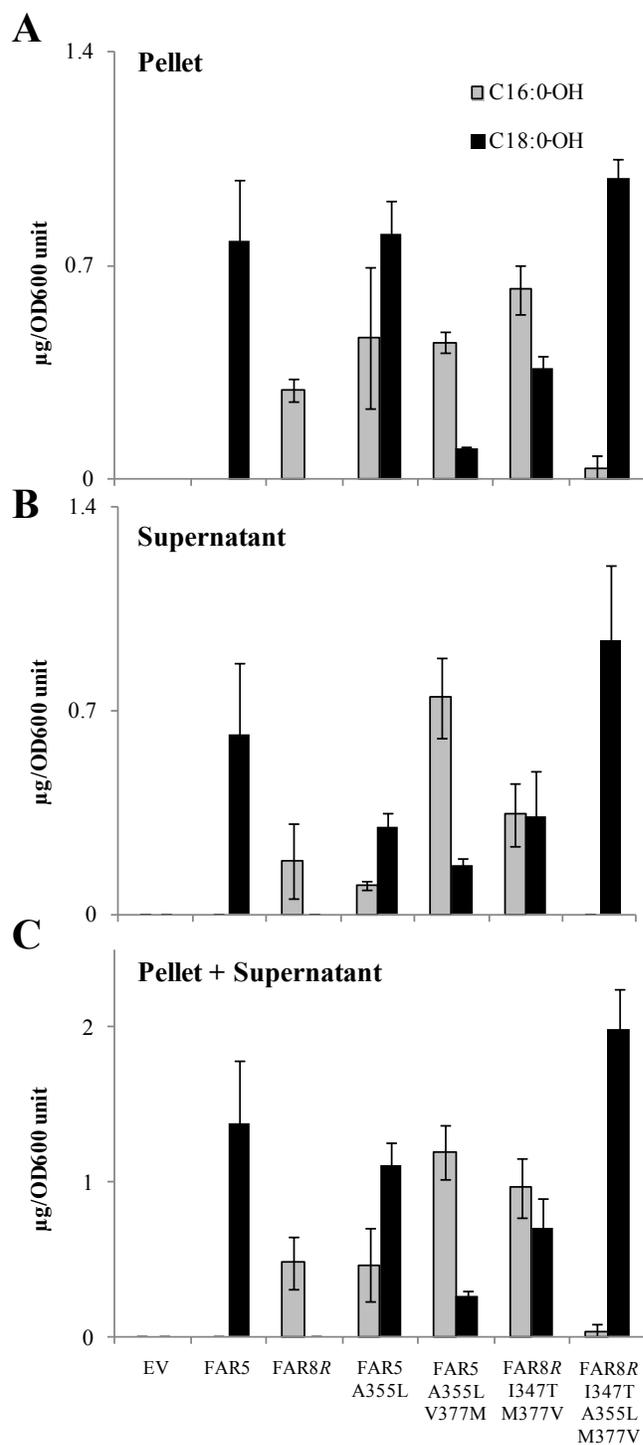


Figure 3.5 Amounts of total fatty alcohols produced by FAR5 and FAR8 variants expressed in yeast. A. Internal fatty alcohol content of yeast cells (non- secreted fatty alcohols). B. Fatty alcohols found in supernatant of yeast cultures (secreted fatty alcohols). C. Total fatty alcohols produced by yeast expressing a FAR variant (combined non-secreted and secreted fatty alcohol content). Values are expressed in $\mu\text{g}/\text{OD}_{600}$ Unit + standard deviation ($n=4$). Yeast cultures were grown for 48 hr in galactose media for protein induction before lipid extraction and analysis.

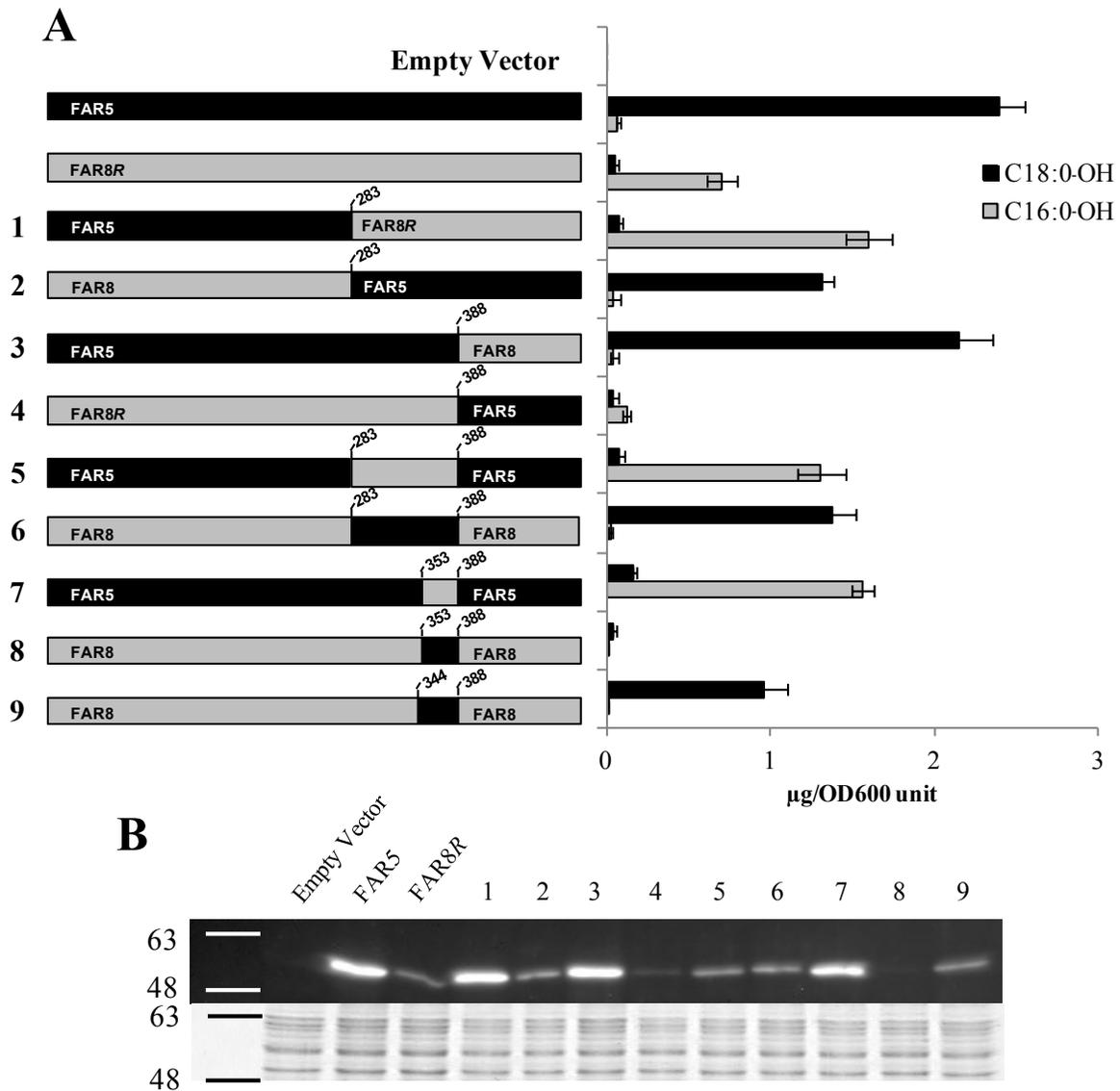


Figure 3.6 Domain swaps between FAR5 and FAR8. A. At left, schematics of FAR5 and FAR8 domain swap chimeras, with the portions from FAR5 in black and the portions from FAR8 in grey. The active FAR8-S363P mutant, denoted as FAR8R, was used in all domain swaps. At right, amounts of non-secreted fatty alcohols produced by yeast expressing FAR5, FAR8, or a FAR5/FAR8 chimera, where values are expressed in $\mu\text{g}/\text{OD600 unit} + \text{standard deviation}$ ($n=4$). B. Protein levels of FAR5, FAR8, and FAR5/FAR8 chimeras expressed in yeast. Western blots (top half) and Coomassie-stained gel (bottom half) are as described in Fig. 3.3 legend.

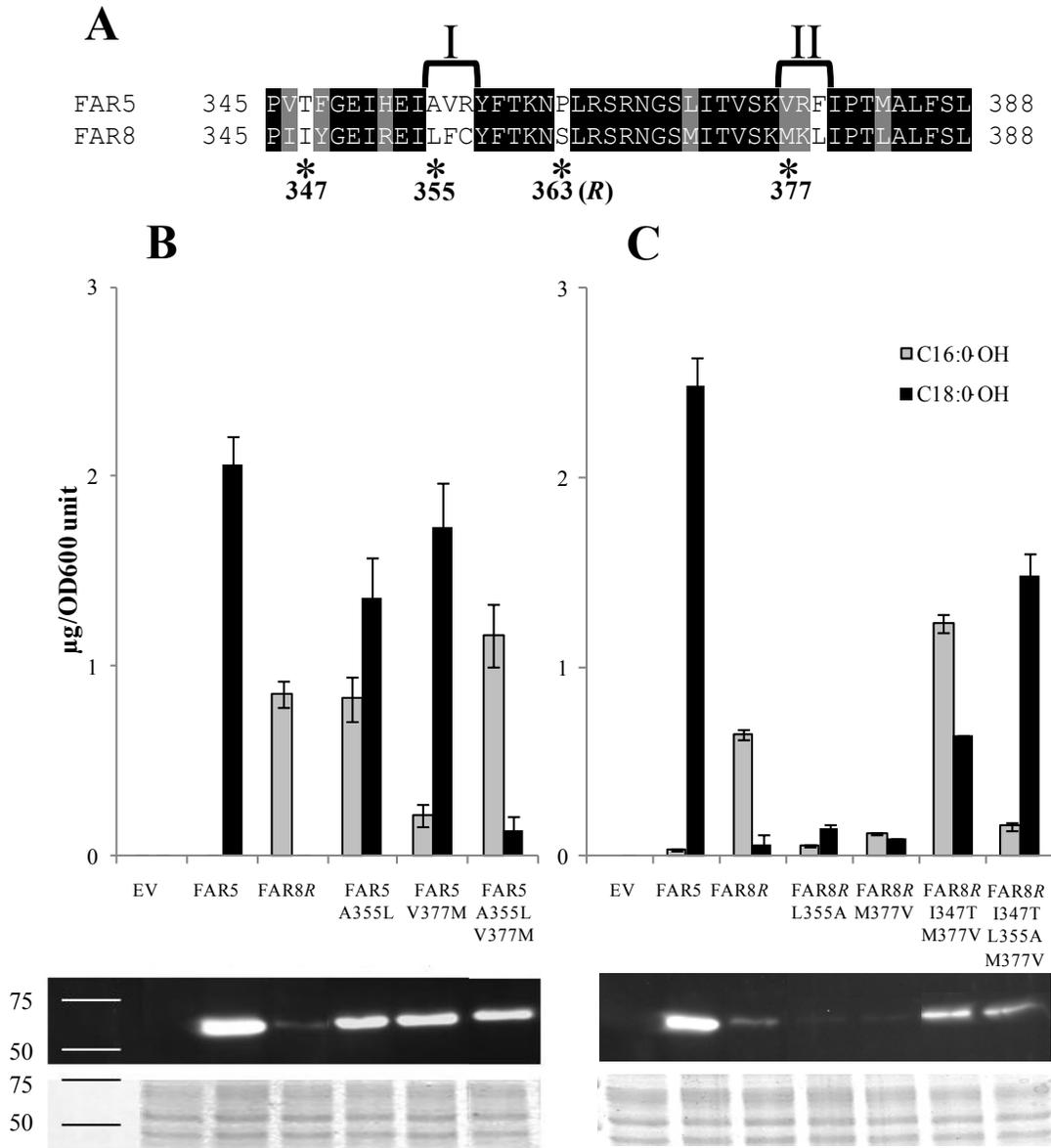


Figure 3.8 Reciprocal amino acid substitutions in FAR5 and FAR8. A. Protein alignment of amino acid residues 345 to 388 of FAR5 and FAR8. Identical amino acids are highlighted in black and physiochemically similar amino acids are highlighted in grey. Asterisks indicate the four amino acids (347, 355, 363, and 377) targeted for analysis in this region by site-specific mutagenesis. B. Analysis of amino acids 355 and 377 affecting substrate specificity of FAR5. C. Analysis of amino acids 355 and 377 affecting substrate specificity of FAR8. The S363P mutation in FAR8, denoted as FAR8R, was present in all FAR8 variants. The top parts of (B) and (C) are graphs reporting the amounts of non-secreted fatty alcohols produced by FAR5 and FAR8 variants, where values are expressed in $\mu\text{g}/\text{OD}_{600}$ unit + standard deviation ($n=4$). The bottom parts of (B) and (C) report the protein levels of FAR5 and FAR8 variants. Western blots (top half) and Coomassie-stained gel (bottom half) are as described in Fig. 3.3 legend.

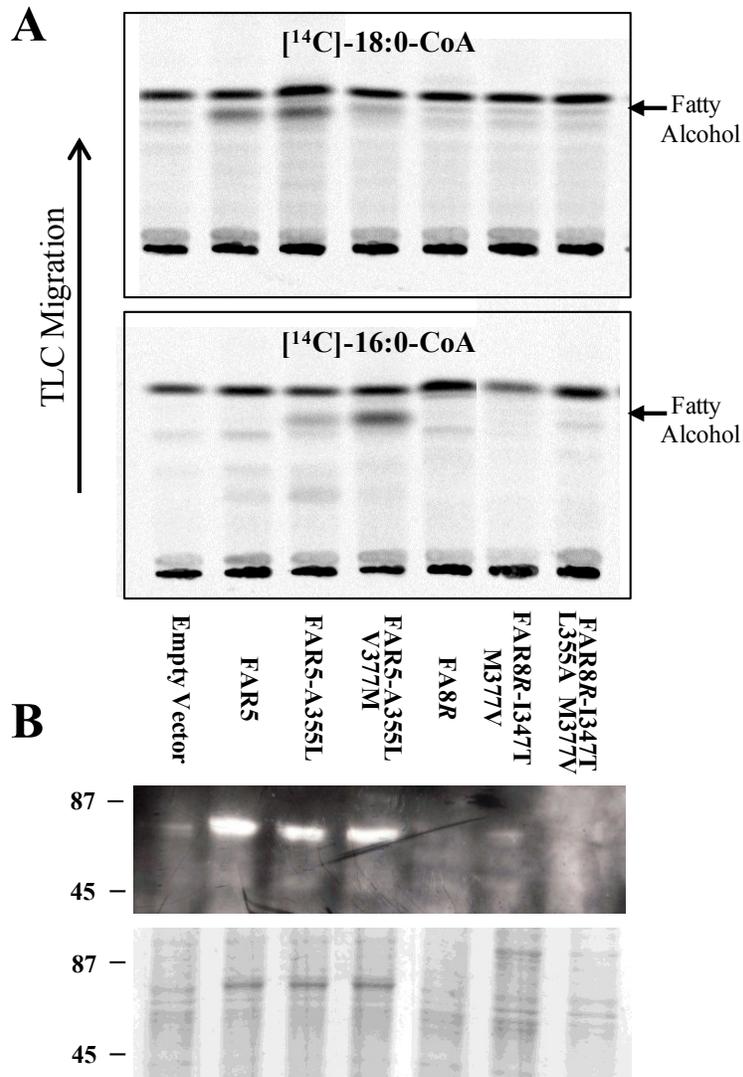


Figure 3.9 FAR *in vitro* assays with yeast microsomes containing FAR5 and FAR8 variants. A. Representative experiment showing separation of lipids recovered from *in vitro* FAR assays by thin layer chromatography (TLC). Assays were conducted in the presence of ¹⁴C radiolabelled 18:0-CoA (top) or ¹⁴C radiolabelled 16:0-CoA (bottom) and microsomes extracted from yeast expressing the FAR indicated below the plate images. EV = empty vector negative control. The radiolabeled fatty alcohols were identified by co-migration with unlabeled standards. B. Protein levels of FAR5 and FAR8 variants in yeast microsomes. Western blots (top half) and Coomassie-stained gel (bottom half) are as described in Fig. 3.3 legend.

Table 3.1 Specific activities of FAR5, FAR5-A355L, and FAR5-A355L V377M. *In vitro* FAR assays were done using yeast microsomes in the presence of NADPH and either [¹⁴C]-16:0-CoA or [¹⁴C]-18:0-CoA. Data are presented as nmol/mg/min ± standard deviation (n=5).

FAR5 variant	[¹⁴C]-C16:0-CoA	[¹⁴C]-C18:0-CoA
FAR5 (wild-type)	0.49 ± 0.39	10.14 ± 1.61
FAR5-A355L	3.67 ± 0.97	7.23 ± 1.99
FAR5-A355L V377M	1.26 ± 1.63	2.7 ± 0.94

CHAPTER 4: Discussion

FAR5 and *FAR8* are located in tandem on the Arabidopsis genome and the encoded proteins are 85% identical in sequence. It is likely that *FAR5* and *FAR8* are the result of a recent genome duplication event. We therefore speculate that the low native activity of FAR8 is a result of recently acquired mutations through its evolutionary history and that it may be on the way to being a pseudogene. The serine at position 363 of FAR8 rather than the conserved proline at this position may be the result of such a mutation. Although FAR5 and FAR8 are highly similar, they have distinct and non-overlapping substrate specificities. Expression of Arabidopsis FAR5 in yeast produced relatively large quantities of 18:0-OH and very small amounts of 16:0-OH, while expression of the native Arabidopsis FAR8 produced exclusively 16:0-OH in very small quantities. To date, 16:0 fatty alcohols have not yet been detected, either in free or combined form, in Arabidopsis (Doan *et al.*, 2011), although they are expected to be part of pollen exine based on the high activity of MS2/FAR2 for 16:0-ACP (Chen *et al.*, 2011). MS2/FAR2 is essential for pollen exine development (Aarts *et al.*, 1997; Chen *et al.*, 2011). FAR8 may contribute to 16:0-OH in pollen exine as low levels of *FAR8* transcript were detected in developing and mature pollen using DNA microarray analysis (Schmid *et al.*, 2005). Regardless of the endogenous role of FAR8, it provided an ideal partnership with FAR5 to investigate amino acids important for dictating their distinct chain length substrate specificities.

As with other NAD(P)H-dependent oxidoreductases, the FAR enzymes possess a predicted YXXXXK active site motif (Labesse *et al.*, 1994). The tyrosine and lysine residues present at this position, along with a serine residue that comes from another part of the protein, have been shown to play an important role in catalysis (Chen *et al.*, 1993; Fujimoto *et al.*, 2001). Deletion of the YXXXXK motif in MS2/FAR2 results in a protein that cannot complement the

pollen exine defects of *ms2/far2* mutants (Chen *et al.*, 2011). However, the individual roles of the tyrosine and lysine predicted to be directly involved in catalysis had not been previously investigated. We made two site-specific FAR5 mutants in which we converted tyrosine-238 to phenylalanine and lysine-242 to isoleucine (FAR5-Y238F and FAR5-K242I). The same substitutions were used to investigate the catalytic residues of *Drosophila* alcohol dehydrogenase (Chen *et al.*, 1993). We found that both FAR5-Y238F and FAR5-K242I produced no fatty alcohol, but were expressed to the same level as wild-type FAR5 in yeast. This confirmed the essential functions of these two residues for FAR enzyme activity; however a detailed catalytic mechanism for the FAR enzymes is not yet known.

We found that substitution of the serine at position 363 to proline in FAR8 (FAR8-S363P) greatly increased its production of 16:0 fatty alcohol in yeast. We hypothesized that a proline at this position plays an important structural role influencing protein stability and activity. Proline is commonly found in tight turns within proteins and such a turn of the backbone may be important for the structural integrity of FAR8. The ‘resurrected’ FAR8 was expressed to higher levels than the native FAR8, likely due to enhanced protein stability. The reciprocal conversion of proline-363 to serine in FAR5 (FAR5-P363S) reduced, but did not abolish 18:0-OH production in yeast. This suggests that while proline-363 is conserved and strongly influences FAR protein activity and stability, it is not essential to all FAR enzymes. The presence of threonine at position 347 also had a positive effect on protein accumulation, although this is not a strictly conserved amino acid in the FAR enzyme family. Native FAR8 has an isoleucine at this position and conversion to threonine significantly enhanced FAR8 accumulation in yeast and correspondingly the amount of 16:0-OH produced in yeast. Conversely, replacement of threonine in FAR5 with isoleucine dramatically decreased its protein

levels. However, the amount of FAR8-S363P I347T that was expressed in yeast was still far less than native FAR5. There are likely other amino acids in FAR8 that would need to be altered to increase protein accumulation further.

Through domain swaps between FAR5 and FAR8, we found that an internal span of 35 residues (354 to 388) dictated the 16:0-CoA versus 18:0-CoA specificity of these two enzymes. Using site-specific substitutions, we identified residues 355 and 377 as the two key residues dictating 16:0-CoA versus 18:0-CoA chain length substrate specificity in FAR5 and FAR8. Individual reciprocal substitution at either of these positions broadened enzyme specificity of either FAR5 or FAR8 such that both 16:0 and 18:0 fatty alcohols were produced in yeast. While single amino acid conversions at position 355 and 377 were not sufficient to fully convert substrate specificity of FAR5 or FAR8, we found that combining the two amino acid conversions nearly completely shifted substrate preference. For FAR5, this phenomenon was confirmed using *in vitro* enzyme assays with radiolabelled 16:0-CoA and 18:0-CoA substrates. Therefore, to fully convert FAR5/FAR8 specificity from one chain length to another, the enzyme's preference for a particular chain length must be 'switched off', otherwise the enzyme exhibits activity for both substrates.

In summary, the results of this study indicate that specific amino acids at the junction of the Rossmann-fold domain and the FAR_C domain govern FAR chain-length specificity and that it is possible to alter this substrate specificity by just two amino acid changes. However, without the benefit of an X-ray or NMR structure of the FAR substrate binding site, it is difficult to determine precisely how amino acids 355 and 377 in FAR5 and FAR8 control binding of 16:0-CoA versus 18:0-CoA substrates. We attempted to use comparative homology modeling of FAR5 and FAR8 to gain further insights into the roles of these residues, but this approach was

not feasible due to the lack of a similar enough modeling template. Since fatty alcohols are used in a variety of human applications, such as in cleaning detergents, cosmetic and pharmaceutical formulations, food products, textiles and coatings, or high performance industrial lubricants (Rowland and Domergue, 2012), this study is a step forward in the engineering of FAR enzymes to produce fatty alcohols with desired specificities and to allow the production of renewable fatty alcohol-containing lipid products of high commercial value.

CHAPTER 5: Future Directions and Concluding Remarks

5.1 Future Directions

The results from this investigation indicate that the amino acids present at position 355 and 377 in FAR5 and FAR8 are responsible for dictating 16:0 versus 18:0 chain length substrate specificity. The domain swapping and point mutation approach applied in this study could be used to investigate amino acid residues important for conferring other chain lengths or to investigate amino acids that affect chain saturation specificities (e.g. 18:0 vs. 18:1). Short-chained and unsaturated wax esters possess low melting temperatures that are ideal for some industrial applications (Patel *et al.*, 2001). For this reason, the engineering of highly active FAR enzymes that have activity for monounsaturated medium chain length fatty alcohols (ranging from C12-C18) that can be incorporated into wax esters is an important biotechnological goal for the future use of wax esters in industrial applications. To investigate the amino acid residues responsible for dictating FAR specificity for saturated chain lengths shorter than 16:0, one could design domain swaps and point mutations between FAR8R from *Arabidopsis* and the *Euglena* FAR (*EgFAR*); both of which demonstrate specificity for 16:0 acyl chain length, however *EgFAR* is also able to work on 14:0 acyl chain length with higher efficiency (Teerawanichpan and Qiu, 2010b; this study). However, this venture would be more difficult than the FAR5/FAR8 substrate specificity investigation considering that *EgFAR* and FAR8R are more diverged than FAR5 and FAR8R; only sharing approximately 27% amino acid sequence identity. Additionally, as *EgFAR* demonstrates low product yield in yeast (Figure 3.4), it would be important to establish whether this is a result of protein activity or substrate availability – 14:0 fatty acid is fed to the yeast and it is unclear whether, once internalized, the 14:0 is activated and available to

the FAR enzyme. If it is determined that the low product yield from *EgFAR* is a result of low protein activity, one could investigate the amino acid residues influencing protein activity in an attempt to increase this product yield. If *EgFAR* could be engineered to produce 14:0 fatty alcohol at similar levels to *Arabidopsis FAR5*'s production of 18:0 fatty alcohol, it could then be of significant industrial value. Alternatively, one could investigate the presence of a higher yielding 14:0-OH producing FAR in other organisms. A potential example of this is from *Brassica napus*, in which 14:0 fatty alcohol has been found to be a constituent in the lipid polyester of the seed (Molina *et al.*, 2008), indicating the presence of a yet unidentified seed-specific *B. napus* FAR. Isolation and characterization of this FAR could potentially provide us with a 14:0-specific reductase more highly related to FAR8R than the *Euglena* FAR. A more similar FAR enzyme, such as a *B. napus* FAR, would more easily facilitate domain swap/point mutation investigations between the two proteins, making identification of the amino acid residues dictating 14:0 acyl chain substrate specificity more feasible.

In addition to investigating FAR chain-length substrate specificity, it would be beneficial to investigate chain saturation specificity because the presence of a double bond significantly decreases the melting temperature of a wax ester. To do this, the domain-swapping and point mutation approach could be applied to *Arabidopsis FAR3/CER4* and the Jojoba seed-expressed FAR. These two proteins share 54% amino acid sequence identity in addition to having overlapping substrate specificity. FAR3/CER4 is specific for C24-C30 saturated fatty acyl precursors (Rowland *et al.*, 2006), while Jojoba FAR is specific for C20-C24 monounsaturated fatty acyl substrates (Metz *et al.*, 2000). The engineered FAR3/CER4 and Jojoba FAR proteins could be expressed in yeast fed with 24:1 fatty acid and the chimeras could be evaluated on their ability to reduce either the endogenous 24:0 or the fed 24:1 fatty acyl substrates. Upon

identification of an amino acid region that controls saturation specificity, individual amino acid conversions in FAR3/CER4 could then be used to identify specific residues dictating saturation preference. Ideally, the information gathered from the chain-length and saturation specificity investigations could then be combined to engineer a FAR that produces medium chain-length monounsaturated fatty alcohols, such as 12:1, 14:1, 16:1, and 18:1 fatty alcohols, for incorporation into industrial wax esters. These wax esters could then be mass produced for industry by expressing engineered FARs in conjunction with a wax synthase (WS) either in an oilseed crop or in microbes.

An alternative method for identifying residues important for dictating FAR chain length or chain saturation specificities would be a directed evolution approach. Directed evolution is a technique that mimics natural evolution with the purpose of generating proteins with novel activities by selecting for protein variants with desirable functions (Tao and Cornish, 2002). This method has been used to modify enzymes for increased thermal stability, catalytic activity, altered substrate specificity, and/or create novel enzyme function. This directed evolution approach could be applied to FAR proteins in an attempt to increase enzyme activity –ex. of the *Euglena* FAR – or to modify substrate specificity – ex. by converting *Arabidopsis* FAR5 specificity from 18:0 acyl-CoA to 18:1 acyl-CoA. With the latter example, the method used to achieve such a goal would be to develop a library of randomly mutated individual sequences of the *FAR5* coding region through error prone PCR. These mutated FARs would then be heterologously expressed in the yeast strain H1246, which contains knockouts in four genes involved in triacylglycerol synthesis (*DGAI1*, *LROI1*, *ARE1*, and *ARE2*) (Sandager *et al.*, 2001), and grown on a media supplemented with 18:1 fatty acid. The H1246 yeast strain cannot convert fatty acids to neutral lipids. Hence, media supplemented with certain fatty acids (e.g. 18:1 fatty

acid) results in significantly reduced growth (Siloto *et al.*, 2009). Instead, these fatty acids accumulate within the yeast cell and become toxic and inhibit growth. Studies have shown, however, that cellular health can be rescued by artificially restoring triacylglycerol biosynthesis. An example of this has been demonstrated by expressing plant triacylglycerol synthesizing enzymes, DGAT1 or PDAT, in the H1246 yeast strain (Siloto *et al.*, 2009). Such rescue may also be possible by expressing a FAR enzyme capable of converting 18:1 fatty acyl-CoA to 18:1 fatty alcohol. Thus, should any rescued H1246 colonies heterologously expressing a mutated FAR5 be observed, it would suggest that this mutation mediates FAR5's ability to utilize 18:1-CoA as a substrate for fatty alcohol synthesis. This directed evolution approach offers an alternative method for engineering FAR enzymes that produce fatty alcohols with desired qualities for industry. While this method is less deliberate or calculated than the domain-swap/point mutation technique, screening libraries containing hundreds of mutants of the targeted protein allows for many more amino acids and combinations of amino acids to be explored and yield the desired result.

While these molecular approaches will help us to better understand FAR substrate specificity and activity, they are hindered by our lack of knowledge of a FAR three-dimensional structure. Currently, there is no X-ray crystallographic or NMR structure for a FAR protein, and the highest sequence identity for a solved protein structure that could be used to model FAR structure is 26%, which is too low for reliable modeling. Also, this potential structural template only spans the Rossmann-fold domain and lacks the FAR_C domain. My research indicates that amino acids at the junction of the Rossmann-fold domain and the FAR_C domain are involved in chain-length substrate specificity, and thus possessing a full structure to model is critical. In the future, efforts should be made to identify an X-ray structure of a FAR so that specific amino

acids responsible for chain-length and saturation specificity would be more easily identifiable. Such structure identification would be most feasible with a soluble FAR, such as the plastidial localized ones in plants (e.g. FAR2 or FAR6 from *Arabidopsis*); then, from such a structure, the membrane-associated FARs (such as FAR5 or FAR8) may be modeled.

5.2 Concluding Remarks

In conclusion, the findings from this study represent the first step in identifying amino acid residues that contribute to FAR substrate specificity. With this knowledge, additional research can be done to identify residues that dictate specificity for other chain-lengths/degrees of saturation. This information could then be applied to the engineering of FARs that produce fatty alcohols of industrial value. These designer FARs could be expressed in tandem with a wax synthase either in a microbial host or an oilseed crop for mass production of industrial wax esters. Several microbial host options exist, such as *E. coli*, which has already been engineered for the production of certain fatty alcohols and wax esters (Steen *et al.*, 2010). Other heterologous host systems could that be employed include the oleaginous industrial yeast *Yarrowia lipolytica* or a microalgae such as cyanobacteria (Beopoulos *et al.*, 2009). Engineered FARs expressed in an oilseed crop for industrial production of wax ester could utilize alternate oilseed plants such as *Crambe abyssinica*, *Camelina sativa*, or *Brassica carinata*. These three crops are ideal for industrial fatty alcohols or wax ester production because none are generally used for human consumption nor do they readily intercross with related *Brassica* food crops or vegetables. While both the microbial host and oilseed crop options represent excellent platforms on which to develop the mass production of designer wax esters or free fatty alcohols with commercial value, this stage is still many years down the road and immediate efforts should

continue to be aimed at further investigating FAR enzyme structure and function so that we may further elucidate the amino acid residues contributing to specific chain length and chain saturation specificities.

REFERENCES

- Aarts, M.G.M., Hodge, R., Kalantidis, K., Florack, D., Wilson, Z.A., Mulligan, B.J., Stiekema, W.J., Scott, R., Pereira, A. (1997). The Arabidopsis MALE STERILITY 2 protein shares similarity with reductases in elongation/condensation complexes. *Plant J.* 12: 615-623.
- Ariizumi T., Toriyama, K. (2011). Genetic regulation of sporopollenin synthesis and pollen exine development. *Annu. Rev. Plant Biol.* 62: 437-460.
- Li-Beisson, Y. *et al.* (2010). The Arabidopsis Book: Acyl-Lipid Metabolism – biosynthesis of cutin polyesters. *American Society of Plant Biologists*. Doi:10.1199/tab.0133:
<http://www.aspb.org/publications/arabidopsis/>
- Beopolous, A., Cescus, J., Haddouche, R., Uribelarrea, J.L., Molina-Jouve, C., Nicaud, J.M., (2009). *Yarrowia lipolytica* as a model for bio-oil production. *Prog. Lipid Res.* 48: 375-387
- Bernards, M.A. (2002). Demystifying suberin. *Can. J. Bot.* 80: 227–240.
- Brosche, T., Platt, D. (1998). The biological significance of plasmalogens in defense against oxidative damage. *Exp. Gerontol.* 33: 363-369.
- Brucker, J.S. (1993). Cuticular polar lipids of insects. In *Insect Lipids: Chemistry, Biochemistry, and Biology*, Stanley-Samuelson, D. W.; Nelson, D. R., Eds.; University of Nebraska Press: Lincoln, NE; pg. 227–270.
- Carlsson, A. (2006). Production of wax esters in Crambe. CLC Press, Newbury, UK.
- Chen, W., Yu, X.-H., Zhang, K., Shi, J., De Oliviera, S., Schreiber, L., Shanklin, J., Zhang, D., (2011). *Male Sterile 2* encodes a plastid-localized fatty acyl-ACP reductase required for pollen exine development in *Arabidopsis thaliana*. *Plant Physiol.* 157: 842-853
- Chen, Z., Jiang, J.C., Lin, Z.G., Lee, W.R., Baker, M.E., Chang, S.H. (1993). Site-specific mutagenesis of Drosophila alcohol dehydrogenase: evidence for involvement of Tyrosine- 152 and Lysine- 156 in catalysis. *Biochemistry* 32: 3342-3346.
- Cheng, J.B., Russell, D.W. (2004). 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.
- Dobritsa, A.A., Shrestha, J., Morant, M., Pinot, F., Matsuno, M., Swanson, R., Møller, B.L., Preuss, D. (2009). CYP704B1 is a long-chain fatty acid α -hydroxylase essential for sporopollenin synthesis in pollen of *Arabidopsis*. *Plant Physiol.* 151: 574-589.
- Doan, T.T.P., Carlsson, A.S., Hamberg, M., Bulow, L., Stymne, S., Olsson, P. (2009). Functional expression of five Arabidopsis fatty acyl-CoA reductase genes in *Escherichia coli*. *J. Plant Physiol.* 166: 787-796.

Doan, T.T.P., Domergue, F., Fournier, A.E., Vishwanath, S.J., Rowland, O., Moreau, P., Wood, C.C., Carlsson, A.S., Hamberg, M., Hofvander, P. (2011). Biochemical characterization of a chloroplast localized fatty acid reductase from *Arabidopsis thaliana*. *Biochim. Biophys. Acta*, 1821: 1244-1255.

Domergue, F., Abbadi, A., Ott, C., Zank, T.K., Zähringer, U., Heinz, E. (2003). Acyl carriers used as substrates by the desaturases and elongases involved in very long-chain polyunsaturated fatty acids biosynthesis reconstituted in yeast. *J. Biol. Chem.* 278: 35115-35126.

Domergue, F., Vishwanath, S. J., Joubes, J., Ono, J., Lee, J.A., Bourdon, M., Alhattab, R., Lowe, C., Pascal, S., Lessire, R., Rowland, O. (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.

Dominguez, E., Mercado, J.A., Quesada, M.A., Heredia, A. (1999). Pollen Sporopollenin: degradation and structural elucidation. *Sex Plant Reprod.* 12: 171-178.

Fujimoto, K., Hara, M., Yamada, H., Sakurai, M., Inaba, A., Tomomura, A., Katoh, S. (2001). Role of the conserved Ser-Tyr-Lys triad of the SDR family in sepiapterin reductase. *Chem. Biol. Interact.* 130-132: 825-832.

Gavel, Y., von Heijne, G. (1990). A conserved cleavage-site motif in chloroplast transit peptides. *FEBS Letters.* 261(2): 455-458.

Gietz, R.D. and Woods, R.A. (2002). Transformation of yeast by the LiAc/ss carrierDNA/PEG method. *Methods Enzymol.* 350: 87-96.

Hajra, A.K. (1983). in *Ether lipids: biochemical and biomedical aspects*, eds. Mangold, H. K. and Paltauf, F. NY: Academic Press pg. 85-106.

Heilmann, M., Iven, T., Ahmann, K., Hornung, E., Stymne, S., 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-V., Gruber, J., Hamberg, M., Frentzen, M. (2011). Fatty acyl-CoA reductases of birds. *BMC Biochem.* 12: 1-12.

Jacob, J., Poltz, J. (1974). Chemical composition of uropygial secretions of owls. *J. Lipid Res.* 15: 243-248.

Kaneda, T. (1962). Biosynthesis of branched chained fatty acids: Microbial synthesis of branched long chain fatty acids from certain short chain fatty acid substrates. *J. Biol. Chem.* 238: 1229-1235.

- Kattner, G., Albers, C., Graeve, M., Schnack-Schiel, S.B. (2003). Fatty acid and alcohol composition of the small polar copepods, *Oithona* and *Oncaea*: indication on feeding modes. *Polar Biol.* 26: 666-671.
- Kavanagh, K.L., Jörnvall, H., Persson, B., 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.
- Kayama, M., Nevenzel, J.C. (1974). Wax ester biosynthesis by midwater marine animals. *Marine Biology* 24: 279-285.
- Kerstiens, G. (1996). Signalling across the divide: A wider perspective of cuticular structure-function relationships. *Trends Plant Sci.* 1: 125-129.
- Khan, A.A., Kolattukudy, P.E. (1975). Solubilization of fatty acid synthetase, acyl-CoA reductase, and fatty acyl-CoA alcohol transacylase from the microsomes of *Euglena gracilis*. *Arch. Biochem. Biophys.* 170: 400-408.
- Khannoon, E.R., El-Gendy, A., Hardege, J.D. (2011). Scent marking pheromones in lizards: cholesterol and long chain alcohols elicit avoidance and aggression in male *Acanthodactylus boskianus* (*Squamata: Lacertidae*). *Chemoecology* 21: 143-149.
- Kolattukudy, P.E. (1970). Reduction of fatty acids to alcohols by cell-free preparations of *Euglena gracilis*. *Biochemistry* 9: 1095-1102.
- Kolattukudy, P.E. (1971). Enzymatic synthesis of fatty alcohols in *Brassica oleracea*. *Arch. Biochem. Biophys.* 142: 701-709.
- Kolattukudy, P.E., Croteau, R., Buckner, J.S. (1976). Biochemistry of plant waxes. *Chemistry and Biochemistry of Natural Waxes*. P. E. Kolattukudy. Amsterdam, Elsevier: pg. 289-347.
- Kosma, D.K., Molina, I., Ohlrogge, J.B., Pollard, M. (2012). Identification of an Arabidopsis fatty alcohol:caffeoyl-coenzyme A acyltransferase required for the synthesis of alkyl hydroxycinnamates in root waxes. *Plant Physiol.* 160: 237-248.
- Kunst, L., Samuels, A.L. (2003). Biosynthesis and secretion of plant cuticular wax. *Prog. Lipid Res.* 42: 51-80.
- Kunst, L., Samuels, L. (2009). Plant cuticles shine: advances in wax biosynthesis and export. *Curr. Opin. Plant Biol.* 12: 1-7.
- Kushnirov, V.V. (2000). Rapid and reliable protein extraction from yeast. *Yeast* 16: 857-860.

- Labesse, G., Vidal-Cros, A., Chomilier, J., Gaudry, M., Mornon, J.-P. (1994). Structural comparisons lead to the definition of a new superfamily of NAD(P)(H)-accepting oxidoreductases: the single-domain reductases/epimerases/dehydrogenases (the 'RED' family). *Biochem. J.* 304: 95-99.
- Lardizabal, K.D., Metz, J.G., Sakamoto, T., Hutton, W.C., Pollard, M.R., 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: 645-655.
- Lassner, M. (1997). Transgenic oilseed crops: a transition from basic research to product development. *Lipid Technology* 9: 5-9.
- Lassance, J.M., Groot, A.T., Lienard, M.A., Antony, B., Borgwardt, C., Andersson, F., Hedenstro, E., Heckel, D.G., Löfstedt, C. (2010). Allelic variation in a fatty-acyl reductase gene causes divergence in moth sex pheromones. *Nature* 466: 486-489.
- Li, H., Zhang, D. (2010). Biosynthesis of anther cuticle and pollen exine in rice. *Plant Signal. Behav.* 5: 1121-1123.
- Lienard, M.A., Hagstrom, A.K., Lassance, J.M., Lofstedt, C. (2010). Evolution of multicomponent pheromone signals in small ermine moths involves a single fatty acyl reductase gene. *Proc. Natl. Acad. Sci. U.S.A.* 107: 10955-10960.
- Matzke, K., Riederer, M. (1991). A comparative study into the chemical constitution of cutins and suberins from *Picea abies* (L.) Karst., *Quercus robur* L., and *Fagus sylvatica* L.. *Planta* 185: 233-245.
- Metz, J.G., Pollard, M.R., Anderson, L., Hayes, T.R., 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.
- Miwa, T.K., (1971). Jojoba oil wax esters and derived fatty acids and alcohols: Gas chromatographic analyses. *J. Am. Oil Chem. Soc.* 48: 259-264.
- Molina, I., Li-Beisson, Y., Beisson, F., Ohlrogge, J.B., Pollard, M. (2009). The lipid polyester composition of *Arabidopsis thaliana* and *Brassica napus* seeds. *Phytochemistry* 67: 2597-2610.
- Molina, I., Ohlrogge, J.B., Pollard, M. (2008). Deposition and localization of lipid polyester in developing seeds of *Brassica napus* and *Arabidopsis thaliana*. *Plant J.* 53: 437-449.
- Moto, K., Yoshiga, T., Yamamoto, M., Takahashi, S., Okano, K., Tetsu, A., Nakata, T., Matsumoto, S. (2003). Pheromone gland-specific fatty-acyl reductase of the silkworm, *Bombyx mori*. *Proc. Natl. Acad. Sci. U.S.A.* 100: 9156-9161.
- Mudge, M.S. (2005). Fatty Alcohols- A review of their natural synthesis and environmental distribution. *The Soap and Detergent Association* pg. 1-152.

Mudge, M.S., Belanger, S.E., Nielsen, A.M., (2008). Fatty alcohols: anthropogenic and natural occurrence in the environment. The Royal Society of Chemistry, Cambridge, UK pg. 1-19

Murray, K.E., (1962). Studies in Waxes XXI. The branched-chain acids of the preen gland wax of goose. *Australian Journal of Chemistry* 15: 510-520.

Paltauf, F. (1994). Ether lipids in biomembranes. *Chem. Phys. Lipids* 74: 101-139.

Patel, S., Nelson, D.R., Gibbs, A.G. (2001). Chemical and physical analysis of wax ester properties. *J. Insect Sci.* 1.4 (e-publication)

Pollard, M.R., McKeon, T., Gupta, L.M., 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.

Pollard, M., Beisson, F., Li, Y., Ohlrogge, J.B. (2008). Building lipid barriers: biosynthesis of cutin and suberin. *Trends Plant Sci.* 13: 236-246.

Pulsifer I.P., Kluge S., Rowland, O. (2012). Arabidopsis LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1), LACS2, and LACS3 facilitate fatty acid uptake in yeast. *Plant Physiol. Biochem.* 51: 31-39.

Ranathunge, K., Schreiber, L., Franke, R. (2010). Suberin research in the genomics era – new interests for an old polymer. *Plant Sci.* 180: 399-413.

Reiser, S., 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.

Rowland, O., 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., Kunst, L. (2006). CER4 encodes an alcohol-forming fatty acyl-coenzymeA reductase involved in cuticular wax production in Arabidopsis. *Plant Physiol.* 142: 866-877.

Salibian, A., Montalti, D. (2009). Physiological and biochemical aspects of avian uropygial gland. *Braz. J. Biol.* 69: 437-446.

Samuels, L., Kunst, L., Jetter, R. (2008) Sealing plant surfaces: Cuticular wax formation by epidermal cells. *Annu. Rev. Plant Biol.* 59: 683–707.

Sandager, L., Gustavsson, M.H., Ståhl, U., Dahlqvist, A., Wiberg, E., Banas, A., Lenman, M., Ronne, H., Stymne, S. (2001). Storage lipid synthesis is non-essential in yeast. *J. Biol. Chem.* 277: 6478-6482.

Schirmer, A., Rude, M.A., Li, X., Popova, E., Del Cardavre, S.B. (2010). Microbial biosynthesis of alkanes. *Science* 329: 559-562.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., Lohmann, J.U. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* 37: 501-50635.

Schreiber, L., Franke, R., Hartmann, K., (2005). Wax and suberin development of native and wound periderm of potato (*Solanum tuberosum* L.) and its relation to peridermal transpiration. *Planta* 220: 520-530.

Shi, J., Tan, H., Yu, X-H., Liu, Y., Liang, W., Ranathunge, K., Franke, R.B., Schreiber, L., Wang, Y., Kai, G., Shanklin, J., Ma, H., Zhang, D. (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.

Siloto, R.M.P., Truksa, M., Brownfield, D., Good, A.G., Weselake, R.J. (2009). Directed evolution of acyl-CoA:diacylglycerol acyltransferase: Development and characterization of *Brassica napus* DGAT1 mutagenized libraries. *Plant Physiol. Biochem.* 47: 456-461.

Soll, J., Tien, R. (1998). Protein translocation into and across the chloroplastic envelope membranes. *Plant Mol. Biol.* 38: 191-207

Spector, A., Yorek, M. (1985). Membrane lipid composition and cellular function. *J. Lipid Res.* 26: 1015-1035.

Spencer, G.F. (1979). Alkoxy-acyl combinations in the wax esters from winterized sperm whale oil by gas chromatography-mass spectrometry. *J. Am. Oil Chem. Soc.* 56: 642-646.

Steen, E.J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., Del Cardavre, S.B., Keasling, J.D., (2010). Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* 463: 559-562

Tao, H., Cornish, V.W. (2002). Milestones in directed enzyme evolution. *Curr. Opin. Chem. Biol.* 6: 858-864.

Teerawanichpan, P., Robertson, A.J., Qiu, X. (2010a). 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. Mol. Biol.* 40: 641-649.

Teerawanichpan, P., Qiu, X. (2010b). Fatty acyl-CoA reductase and wax synthase from *Euglena gracilis* in the biosynthesis of medium-chain wax esters. *Lipids* 45: 263–273.

- Teerawanichpan, P., Qiu, X. (2011). Molecular and functional analysis of three fatty acyl-CoA reductases with distinct substrate specificities in Copepod *Calanus finmarchicus*. *Marine Biotechnol.* 14: 227-236
- Vioque, J., Kolattukudy, P.E. (1997). Resolution and purification of an aldehyde-generating and an alcohol-generating fatty acyl-CoA reductase from Pea Leaves (*Pisum sativum* L.). *Arch. Biochem. Biophys.* 340: 64-72.
- Von Heijne, G., Nishikawa, K. (1991). Chloroplast transit peptides. *FEBS Letters* 278: 1-3.
- Wang, X., Kolattukudy, P.E. (1995). Solubilization, purification and characterization of fatty acyl-CoA reductase from duck uropygial gland. *Biochem. Biophys. Res. Commun.* 208: 210-215.
- Wang, A., Xia, Q., Xie, W., Dumonceaux, T., Zou, J., Datla, R., Selvaraj, G. (2002). Male gametophyte development in bread wheat (*Triticum aestivum* L.): molecular, cellular, and biochemical analyses of a sporophytic contribution to pollen wall ontogeny. *Plant J.* 30: 613-623.
- Wiermann, R., Gubatz, S., (1992). Pollen wall and sporopollenin. *Sexual reproduction in flowering plants*. Academic press, Inc. London, England: pp. 35-52.
- Willis, R.M., Wahlen, B.D., Seefeldt, L.C., Barney, B.M. (2011). Characterization of a fatty acyl-CoA reductase from *Marinobacter aquaeolei* VT8: A bacterial enzyme catalyzing the reduction of fatty acyl-CoA to fatty alcohol. *Biochemistry* 50: 10550-10558.
- Wyatt, T.D., (2003). Pheromones and animal behaviour: communication by smell and taste. Book. Cambridge, United Kingdom: Cambridge University Press
- Yang, W., Pollard, M., Li-Beisson, Y., Beisson, F., Feig, M., Ohlrogge, J. (2010). A distinct type of glycerol-3-phosphate acyltransferase with *sn*-2 preference and phosphatase activity producing 2-monoacylglycerol. *Proc. Natl. Acad. Sci. U.S.A.* 107: 12040-12045.
- Yarger, R.G., Smith, A.B., Preti, G., Epple, G. (1976). The major volatile constituents of the scent mark of a South American primate *Saguinus fuscicollis*, *Callitrichidae*. *J. Chem Ecol.* 3: 45-46.