Substrate Specificities of Plant Alcohol-Forming Fatty Acyl Reductases

By

Ashley Fournier

B.Sc. (Honours), Carleton University, 2009

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

Your file Votre référence
ISBN: 978-0-494-91545-5

Our file Notre référence
ISBN: 978-0-494-91545-5

Abstract

Primary fatty alcohols are found throughout the biological world, either in free form or in a combined state (e.g. wax esters). Alcohol-forming fatty acyl reductases (FARs) are responsible for the NADPH-dependent reduction of fatty acyl-coenzymeA (CoA) or fatty acyl-acyl carrier protein (ACP) to primary fatty alcohols via an unreleased fatty aldehyde intermediate. The genome of *Arabidopsis thaliana* contains eight genes encoding FAR enzymes (FAR1-FAR8). I assessed the substrate specificities of seven of these FARs by heterologous expression in the yeast *Saccharomyces cerevisiae*. Arabidopsis FAR1, FAR2/MS2, FAR3/CER4, FAR4, FAR5, FAR6, and FAR8, as well as versions of FAR2 and FAR6 with truncations of their predicted N-terminal chloroplast targeting sequences were characterized. Analysis of yeast internal and secreted lipids revealed that the Arabidopsis FAR protein family produce primary fatty alcohols ranging from C16:0 to C26:0 using the endogenous yeast acyl-CoA pool, with each FAR having a distinct chain length specificity. FAR5 and FAR8 were found to have strict specificity for C18:0 and C16:0 acyl chain length, respectively. A serine to proline substitution at position 363 of FAR8 was found to greatly increase its ability to produce C16:0 primary fatty alcohol in yeast. Domain swaps of FAR5 and FAR8 were characterized and a 72 amino acid region between residues 312 and 383 was found to be important for dictating C16:0 versus C18:0 chain length specificity. The function of Arabidopsis FAR6, which specifically produces C16:0 fatty alcohol in yeast, is currently unknown. Characterization of transgenic Arabidopsis plants expressing a *GUS* reporter gene under the control of the *FAR6* promoter indicated that *FAR6* is expressed in various tissues of Arabidopsis: stem (outer cell layers), anther, silique, and root tips. This is a distinct gene expression pattern

compared to the other Arabidopsis *FAR* genes, which suggests a specialized function for FAR6 *in planta*.

Acknowledgments

I would like to thank Dr. Owen Rowland for being a superb mentor throughout my thesis project and providing me with excellent guidance and supervision that provided me with a new array of valuable skills. I would also like to thank my M.Sc. graduate committee members, Dr. John Vierula and Dr. Gopal Subramaniam of Carleton University, for their helpful suggestions and guidance throughout the course of this work.

Table of Contents

Abstract ..... ***i***

Acknowledgments ..... ***iii***

List of Figures ..... ***vii***

List of Tables ..... ***ix***

List of Abbreviations ..... ***xi***

CHAPTER 1: General Introduction *1*

1.1 Fatty Alcohols: Free and Combined Forms *1*

1.2 Fatty Alcohols in Plants *4*

1.3 Biochemistry of Alcohol-Forming Fatty Acyl Reductase Enzymes *9*

1.4 Function and Substrate Specificities of FARs *14*

1.5 Jojoba (Simmondsia chinensis) FAR *14*

1.6 Arabidopsis thaliana FAR Family *15*

1.7 FARs in Non-Plant Organisms *22*

1.7.1 Phytoflagellate Protist (Euglena gracilis) *22*

1.7.2 Moths *23*

1.7.3 Honey bees (Apis mellifera) *25*

1.7.4 Copepod (Calanus finmarchicus) *26*

1.7.5 Mouse and Human *27*

1.8 Effect of Heterologous Expression Host on Apparent FAR Substrate Specificities *30*

1.9 Industrial Applications of FARs *31*

1.10 Thesis Objectives *32*

2.0 Materials and Methods ***33***

2.1 Plant Material and Growth Conditions *33*

2.2 RNA Extraction *33*

2.3 cDNA Synthesis *34*

2.4 Yeast Expression Constructs for Heterologous Expression of FAR Proteins *35*

2.5 Plasmid Transformation in DH5α Escherichia coli cells and Positive Clone Selection *48*

2.6 Yeast Transformation *50*

2.7 SDS-PAGE and Western Blot *50*

2.8 Gas Chromatography *53*

2.9 Promoter::GUS Fusions and GUS Histochemical Assay *55*

3.0 CHAPTER 3: Substrate Specificities of Arabidopsis FARs and Jojoba FAR ***59***

3.1 Introduction *60*

3.2 Materials and Methods *70*

3.3 Arabidopsis FAR6 Gene Expression Pattern *70*

3.4 Heterologous Yeast Expression System for Characterization of Plant FARs *74*

3.5 Fatty Alcohol Production in Yeast Expressing Arabidopsis FARs or Jojoba FAR *82*

3.6 Mutation of Putative Active Site Residues in Arabidopsis FAR5 ..... *95*

3.7 Discussion *97*

4.0 CHAPTER 4: Amino Acids Important for Chain-Length Substrate Specificities of Arabidopsis FAR5 and FAR8 ..... ***103***

4.1 Introduction *103*

4.2 Materials and Methods *104*

4.3 Arabidopsis FAR5 and FAR8 proteins *105*

4.4 Expression of Arabidopsis FAR5 and FAR8 variants in Saccharomyces cerevisiae *107*

4.5 Discussion *120*

5.0 CHAPTER 5: General Discussion ***124***

5.1 Future Directions *125*

5.2 Concluding Remarks *128*

6.0 References ***130***

List of Figures

Figure 1.1 Generalized structures of primary fatty alcohols, wax esters, alkyl hydroxycinnamates, and ether lipids *3*

Figure 1.2 Fatty alcohols and their derivatives *8*

Figure 1.3 FAR domain structure *12*

Figure 1.4 Phylogenetic tree of related FAR protein sequences *20-21*

Figure 3.1 The Arabidopsis *FAR* gene family on chromosome I-V *63*

Figure 3.2 Phylogram of the Arabidopsis FAR protein family and seed-expressed Jojoba FAR *65*

Figure 3.3 Alignment of Arabidopsis FAR and Jojoba FAR protein sequences *66-67*

Figure 3.4 Schematics of the FAR2 and FAR6 proteins *68*

Figure 3.5 Gene expression pattern of *FAR6* using a reporter GUS promoter fusion *72*

Figure 3.6 Gene expression pattern of *FAR6* using semi-quantitative RT-PCR *73*

Figure 3.7 SDS-PAGE of yeast pellets collected after 0, 6, 9, 12, 24, and 48 hours of induction by galactose *77*

Figure 3.8 Gas chromatograms of lipids extracted from transgenic yeast expressing tagged versions of Arabidopsis FAR5 and FAR8 ..... *78*

Figure 3.9 Gas chromatograms of lipids extracted from transgenic yeast expressing untagged and tagged versions of Arabidopsis FAR5 ..... *79*

Figure 3.10 *Arabidopsis thaliana* FAR protein family and Jojoba FAR expressed in yeast *85*

Figure 3.11 Gas chromatograms of lipids extracted from transgenic yeast expressing tagged Arabidopsis FAR1, FAR2, FAR2ΔN, FAR3 and FAR4 ..... *86*

Figure 3.11 continued. Gas chromatograms of lipids extracted from transgenic yeast expressing tagged Arabidopsis FAR5, FAR6, FAR6ΔN, FAR8, and Jojoba FAR *87*

Figure 3.12 Zoomed in gas chromatograms of lipids extracted from transgenic yeast expressing tagged Arabidopsis FAR1, FAR2, FAR2ΔN, FAR3, and FAR4 ..... *88*

Figure 3.12 continued. Zoomed in gas chromatograms of lipids extracted from transgenic yeast expressing tagged Arabidopsis FAR5, FAR6, FAR6ΔN, FAR8, and Jojoba FAR *89*

Figure 3.13 *Arabidopsis thaliana* FAR protein family and Jojoba FAR expressed in yeast: Internal and Secreted Fatty Alcohols *90*

Figure 3.14 Percent of fatty alcohols secreted by transgenic yeast strains expressing Arabidopsis and Jojoba FARs *94*

Figure 3.15 Mutation of tyrosine and lysine residues in the putative active site of FAR5 ..... *96*

Figure 4.1 Alignment of Arabidopsis FAR5 and FAR8 protein sequences *106*

Figure 4.2 Influence of reciprocal amino substitutions at positions 337 and 363 in Arabidopsis FAR5 and FAR8 expressed in yeast *113*

Figure 4.3 Expression of Arabidopsis FAR5/FAR8 domain swaps in yeast *114*

Figure 4.4 Expression of Arabidopsis FAR8 variants and FAR5/FAR8 domain swaps in yeast *115*

Figure 4.5 Expression of Arabidopsis FAR8 variants containing site-specific substitutions in yeast *116*

List of Tables

Table 1.1 Domains Predicted in the Arabidopsis Family of FARs and Jojoba FAR using the Conserved Domain Database *13*

Table 1.2 Substrate Specificities of FARs from Various Organisms..... *29*

Table 2.1 Primers *46-47*

Table 3.1 Summary Table of Amino Acid Sequence Identity between the Arabidopsis FAR family and Jojoba FAR *64*

Table 3.2 Chloroplast target peptide predictions for the Arabidopsis FAR family and Jojoba FAR using ChloroP 1.1 Server..... *69*

Table 3.3 Primary fatty alcohol compositions of the pellets of untagged and tagged versions of Arabidopsis FAR5 expressing yeast strains..... *80*

Table 3.4 Primary fatty alcohol compositions of the supernatants of untagged and tagged versions of Arabidopsis FAR5 expressing yeast strains..... *81*

Table 3.5 Internal Fatty Alcohol Composition of the pellet of transgenic yeast strains expressing Arabidopsis and Jojoba FARs..... *91*

Table 3.6 Secreted fatty alcohol composition of the supernatant of transgenic yeast strains expressing Arabidopsis and Jojoba FARs..... *92*

Table 3.7 Fatty alcohol composition of the combined pellet and supernatant fractions, of transgenic yeast strains expressing Arabidopsis and Jojoba FARs *93*

Table 4.1 Internal Fatty Alcohol Composition of the pellet of transgenic yeast strains expressing Arabidopsis FAR5 and FAR8 site directed mutants and domain swaps *117*

Table 4.2 Secreted fatty alcohol composition of the supernatant of transgenic yeast strains expressing *Arabidopsis thaliana* FAR5 and FAR8 site directed mutants and domain swaps..... ***118***

Table 4.3 Internal and Secreted Internal fatty alcohol composition, of the pellet and supernatant combined, of transgenic yeast strains expressing *Arabidopsis thaliana* FAR5 and FAR8 site directed mutants and domain swaps..... ***119***

List of Abbreviations

ACP: Acyl Carrier Protein
AHCS: Alkyl Hydroxycinnamate Synthase
APS: Ammonium Persulfate
AVG: Average
BSTFA: N,O-bis (trimethylsilyl) trifluoroacetamide
C: Carbon
CER: Eceriferum
CoA: Coenzyme A
DEPC: Diethylpyrocarbonate
DNA: Deoxyribonucleic Acid
dNTP: Deoxyribonucleotide Triphosphate
DPW: DEFECTIVE POLLEN WALL
DTT: Dithiothreitol
EDTA: Ethylenediaminetetraacetic Acid
EV: Empty Vector
ER: Endoplasmic Reticulum
FA: Fatty Acid
FAE: Fatty Acid Elongase
FAR: Fatty Acyl-CoA Reductase
GAPC: Glyceraldehyde-3-Phosphate Dehydrogenase C Subunit
GC: Gas Chromatography
GFP: Green Fluorescent Protein
GUS: β-glucuronidase
HEK: Human Embryonic Kidney
his: Histidine
$K_3Fe(CN)_6$: Potassium ferricyanide
$K_4Fe(CN)_6$: Potassium ferrocyanide
KCl: Potassium Chloride
KCS: a β-ketoacyl-CoA synthase
kDa: Kilodalton
kV: Kilovolt
leu: Leucine
LB: Lysogeny Broth
LiAc: Lithium Acetate
$MgCl_2$: Magnesium Chloride
MS2: MALE STERILITY2
NaCl: Sodium Chloride
NADB Domain: Nicotinamide Adenine Dinucleotide Binding Domain
NADH: Nicotinamide Adenine Dinucleotide

NADPH: Nicotinamide Adenine Dinucleotide Phosphate
$NaPO_4$: Sodium Phosphate
$(NH_4)_2SO_4$: Ammonium Sulphate
OD: Optical Density
-OH: Alcohol
ORF: Open Reading Frame
PEG: Poly(Ethylene Glycol)
PCR: Polymerase Chain Reaction
PMSF: Phenylmethylsulfonyl Fluoride
RED: Reductases/Epimerases/dehydrogenases
RNA: Ribonucleic Acid
RT-PCR: Reverse Transcription Polymerase Chain Reaction
SD: Synthetic Media Containing D-glucose
SG: Synthetic Media Containing Galactose
SDR: Short-Chain Alcohol Dehydrogenase/Reductase
SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SS: Single Stranded
ST DEV: Standard Deviation
STE: Sodium Chloride, Tris-HCl, EDTA
t: Time
Taq: *Thermus aquaticus* DNA Polymerase
TBST: Tris Base, Sodium Chloride, Tween-20
TCA: Trichloroacetic acid
TE: Tris-EDTA
TEMED: Tetramethylethylenediamine
Tris-HCl: Tris(Hydroxymethyl) Aminomethane) Hydrochloride
trp: Tryptophan
ura: Uracil
UV: Ultraviolet
WS: Wax Synthase
WSD: Wax Synthase/Diacylglycerol Acyltransferase
X-Gluc: 5-bromo-4-chloro-3-indolyl β-D-glucuronide
YPAD: Yeast Extract, Peptone, Adenine Hemisulphate, D-glucose

CHAPTER 1: General Introduction

1.1 Fatty Alcohols: Free and Combined Forms

Primary fatty alcohols are aliphatic hydrocarbons that contain a hydroxy group at the terminal carbon. Long-chain (C16 and C18) and very-long-chain (>C18) primary fatty alcohols are found in nature either as free alcohols or in a combined state (Figure 1.1) (Metz *et al.*, 2000). Fatty alcohols can be combined, for example, with fatty acids to give wax esters or with hydroxycinnamic acids to give alkyl hydroxycinnamates. They can also be combined with glycerophospholipids to give ether lipids (Hajra, 1983; Wang and Kolattukudy, 1995). Free long-chain alcohols are not known to accumulate to high levels within livings cells and it is thought that they are detrimental to cell viability (Metz *et al.*, 2000). Modifying or combining the fatty alcohols (i.e. wax esters) may thus be necessary if functioning intracellularly.

Primary fatty alcohols and derivatives partake in biological roles that are essential in a wide range of living organisms (Liénard *et al.*, 2010). Some important functions of fatty alcohols and/or wax esters in non-plants are: (1) fatty alcohols and wax esters act as energy reserves when stored in some microorganisms (e.g. *Acinetobacter* sp. and *Euglena gracilis*), (2) wax esters are chemical constituents of bee hives to protect larvae and store food, (3) wax esters fill the spermaceti organ, along with triglycerides, in the heads of sperm whales to allow for echolocation sensing and possibly regulating buoyancy, and (4) fatty alcohols are precursors of sex pheromones in moths (Wang and Kolattukudy, 1995; Spencer, 1979; Granier *et al.*, 2002; Moto *et al.*, 2003; Samuels *et al.*, 2008; Lassance *et al.*, 2010; Liénard *et al.*, 2010; Teerawanichpan *et al.*, 2010[b]). Fatty alcohols are also precursors of ether lipids, which are found in various animal tissues (Hajra, 1983;

Wang and Kolattukudy, 1995). Plasmalogens are a subclass of ether phospholipids that are widely distributed in the animal kingdom, as well as in certain anaerobic microorganisms. Plasmalogens are present in nervous tissue, cardiac tissue, and inflammatory and immunological cells. They may play a role in membrane dynamics, intracellular signalling, and a protective role during oxidant-induced stress by functioning as endogenous anti-oxidants that protect against reactive oxygen species (Nagan and Zoeller, 2001).

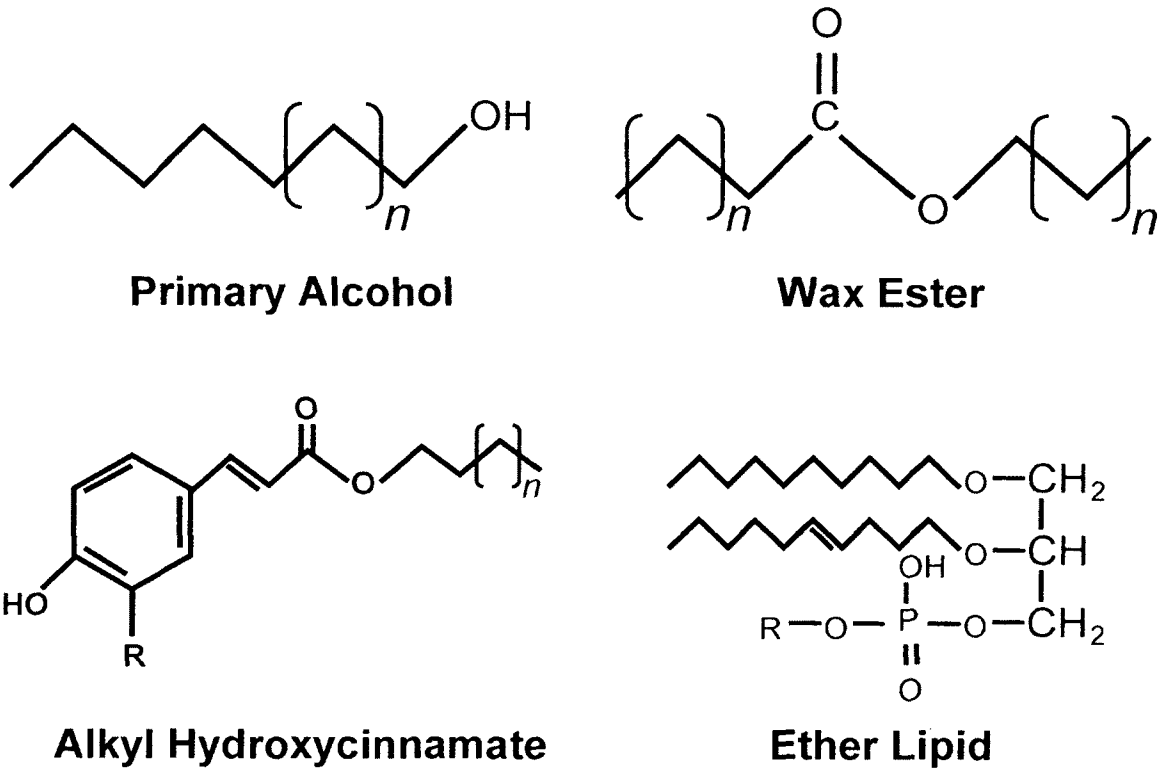


Figure 1.1 Generalized structures of primary fatty alcohols, wax esters, alkyl hydroxycinnamates, and ether lipids. R = H, OH, or OCH_3.

1.2 Fatty Alcohols in Plants

In plants, fatty alcohols and derivatives are common components of extracellular lipid barriers, such as cuticle, suberin and sporopollenin (Pollard *et al.*, 2008). Fatty alcohols are also found in the seeds of the Jojoba plant in the form of wax esters as a seed lipid energy reserve (Metz *et al.*, 2000). The chemical compositions and functions of these lipids are described in further detail below and summarized in Figure 1.2.

Plant cuticle is the lipidic layer coating the majority of aerial surfaces of land plants (Samuels *et al.*, 2008). It forms a continuous seal over the outer walls of epidermal cells, including pavement, guard and trichome cells. The cuticle is crucial for the survival of plants in a terrestrial environment providing protection against various stresses, such as desiccation and pathogen attack (Schreiber, 2010). The plant cuticle is composed of two layers: (1) a cuticle membrane layer composed of polymerized fatty acids (cutin) and intracuticular waxes, and (2) an epicuticular wax layer that covers the cuticle membrane layer (Samuels *et al.*, 2008). Primary fatty alcohols are typical components of plant cuticular waxes along with alkanes, fatty aldehydes, secondary fatty alcohols, fatty ketones, and wax esters. The wax load on wild-type Arabidopsis stems and leaves contains 10-15% and 15-25% primary fatty alcohols, respectively, either as free alcohols or as wax esters (Jenks *et al.*, 1995; Rowland *et al.*, 2006). The primary fatty alcohols that have been reported to exist in Arabidopsis cuticle waxes are saturated and range from C20-C32, with C26-C30 being the dominating chain lengths (Rowland *et al.*, 2006; Lai *et al.*, 2007).

Suberin is a lipid and phenolpropanoid-based barrier constitutively deposited in the cell walls of various external and internal tissue layers (Pollard *et al.*, 2008).

Underground tissue layers with suberized cell walls are root endodermis (Casparian Strip), root exodermis, and tuber peridermis (Matzke and Riederer, 1991; Bernards, 2002). Suberin is also synthesized at abscission zones to seal tissues and in response to wounding (Dean and Kolattukudy, 1976). Suberin can be found constitutively deposited in some specialized aerial tissue layers, such as bundle sheath cells, bark tissue, cotton (*Gossypium hirsutum*) fibers, the chalazae plug of seeds, seed coat integuments, sheaths around idioblasts, pigment strands of grains, and in the boundary between the plant and its secretory organs (glands, trichomes) (Kolattukudy, 2001; Bernards, 2002). Suberin has poly(aliphatic) and poly(phenolic) domains, which are covalently linked together on the inner face of cell walls (Bernards, 2002). The suberin poly(aliphatic) domain is assembled from a series of aliphatic components including primary fatty alcohols, ω-hydroxyalkanoic acids, α,ω-dioic acids, mid-chain epoxide- as well as di- and tri-hydroxy-substituted octadecanoates, and glycerol (Bernards, 2002). The fatty alcohols associated with suberin are usually saturated and are typically C18, C20 and C22 in length (Franke *et al.,* 2005; Schreiber *et al.*, 2005; Beisson *et al.*, 2007; Domergue *et al.*, 2010). The suberin poly(phenolic) domain consists of monolignols (*p*-coumaryl, coniferyl and sinapyl alcohols), hydroxycinnamic acids (*p*-coumaric, caffeic, and ferulic acids), ferulic acid esters of long-chain alcohols (ferulates), and feruloyltyramine (Bernards, 2002). Hydroxycinnamic acid conjugates have been shown to accumulate at infection sites and are important in protecting plants against pathogens (Matern and Kneusel, 1988). Non-polymerized "waxes" are often associated with suberin with alkyl hydroxycinnamates (a fatty alcohol linked with a hydroxycinnamic acid) often being a major component of these waxes (Bernards, 2002). Alkyl hydroxycinnamates have

various biological activities, such as anti-insect, anti-viral, anti-bacterial, and anti-fungal properties (Elliger *et al.*, 1981; Cheminat *et al.*, 1988; Ravn and Brimer, 1988; Ravn *et al.*, 1989; Andary, 1993; Hohlfeld *et al.*, 1996). Suberin wax alkyl ferulates represent about 25% of the hexane and chloroform extracts (non-polymerized fractions) of *Pseudotsuga menziesii* bark cork and 60% of the solvent-extractable material of potato wound periderm (Adamovics *et al.*, 1977; Laver and Fang, 1989; Schreiber *et al.*, 2005). In the suberin-associated root waxes of Arabidopsis, esters of C18:0-C22:0 saturated fatty alcohols linked with *p*-coumaric, caffeic, and ferulic acids (alkyl hydroxycinnamates) are the predominating chemical components (47% w/w) (Li *et al.*, 2007).

Sporopollenin is the major component of the outer walls in the exine of pollen (Dobritsa *et al.*, 2009). Sporopollenin is insoluble in aqueous and organic solvents, and highly resistant to non-oxidative physical, chemical, and biological degradative treatments. These attributes account for preservation of pollen grains, but it makes it difficult to study the chemical composition (Dobritsa *et al.*, 2009). Sporopollenin accumulation is not restricted to the outer pollen wall alone; sporopollenin-containing elements include ubisch bodies (orbicules), tapetal and peritapetal cell walls, and viscin threads (Wiermann and Gubatz, 1992). It is thought that C16:0 fatty alcohols are components of sporopollenin in Arabidopsis and rice (Chen *et al.*, 2011; Shi *et al.*, 2011), although this has not been directly measured due to lack of material and resistance to depolymerisation for analysis by gas chromatography or other analytical methods.

The Jojoba (*Simmondsia chinensis*) plant, native to American Southwest deserts, produces a seed lipid energy reserve mostly made up of wax esters, which are esters of long-chain alcohols and fatty acids (40-60% of dry seed weight) (Ohlrogge *et al.*, 1977;

Metz *et al.*, 2000). This is a unique example as all other plants described thus far store their lipid reserves as triacylglycerols. The wax ester chain lengths of Jojoba oil are C38 (6.23%), C40 (30.56%), C42 (49.50%), and C44 (8.12%), with other esters occurring in amounts lower than 1% (Bassam, 1998). The fatty alcohol moiety is dominated by C20:1-OH and C22:1-OH, together representing nearly 90% of the total fatty alcohol in almost equal proportions, and the fatty acid moiety is mostly C20:1 (~70% of the total fatty acid) and the remainder mostly C22:1 (~15%) and C18:1 (~10%) (Miwa, 1971).

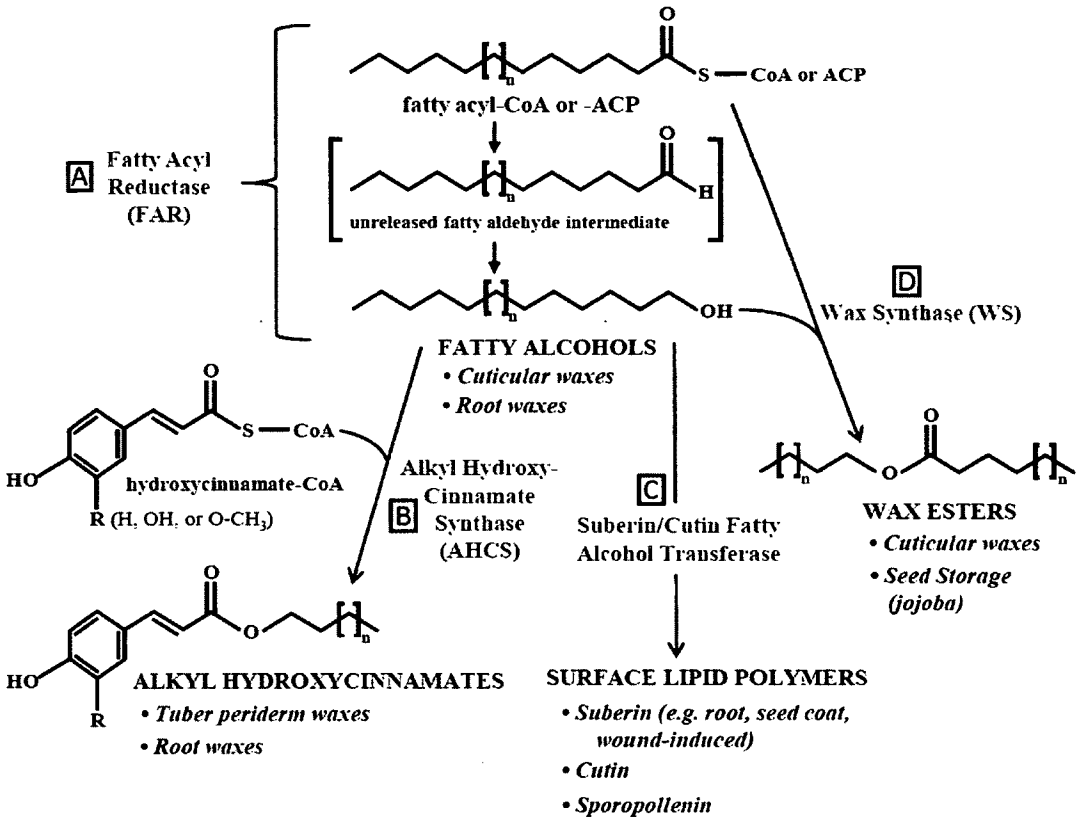


Figure 1.2 Fatty alcohols and their derivatives. [A] Fatty acyl-CoAs or fatty acyl-ACPs are reduced by fatty acyl reductases (FARs; EC:1.2.1.n2) generating fatty alcohols (C16-C30 chain lengths in plants). [B] Fatty alcohols may be linked with phenolics to yield alkyl hydroxycinnamates by an alkyl hydroxycinnamate synthase (AHCS). [C] Fatty alcohols can be incorporated into surface polymers (suberin, cutin, sporopollenin) by an as yet uncharacterized suberin/cutin alcohol transferase. [D] Fatty alcohols can be linked with fatty acids to yield wax esters by a wax synthase (WS). Grey = enzyme activities, uppercase = major products, italics = location where the compounds accumulate. Taken from Rowland and Domergue, 2012, Plant Science, in press, reproduced with permission.

1.3 Biochemistry of Alcohol-Forming Fatty Acyl Reductase Enzymes

The reduction of fatty acyl-CoAs or fatty acyl-ACPs to fatty alcohols occurs via a fatty aldehyde intermediate and is catalyzed by either one or two enzymes (Kunst and Samuels, 2003; Doan *et al.*, 2011). For the two-enzyme reaction, a fatty-acyl reductase first converts the fatty acyl chain to a fatty aldehyde, and then an aldehyde reductase reduces the fatty aldehyde to a fatty alcohol (Kolattukudy 1971). Kolattukudy (1970) showed that the reduction of fatty acyl-CoA to fatty alcohols can proceed alternatively via an unreleased aldehyde intermediate by the action of a single enzyme (Figure 1.2). It was shown that it still goes through an aldehyde intermediate by trapping the aldehyde with semicarbazide or phenylhydrazine carbonyl reagent (Kolattukudy 1970). The single enzyme reaction occurs more commonly than the two-enzyme reduction in nature (Kolattukudy, 1970). FARs (EC:1.2.1.n2) in plants exhibit the single enzyme reaction and are dependent on NAD(P)H for activity (Kolattukudy, 1970; Doan *et al.*, 2011).

The seed-expressed Jojoba FAR has served as the model for plant FARs with biochemical characterization starting in the 1970s (Pollard *et al.*, 1979) and then the cloning of the cDNA in the late 1990s (Metz *et al.*, 2000). The Jojoba FAR cDNA sequence encoding a predicted 56.2 kDa protein has homologs in all land plants whose genomes have been sequenced to date, as well as a homolog in the moss *Physcomitrella patens*. FARs, including those in non-plant organisms, contain two distinct domains: a Rossmann fold NAD(P)H-binding domain and a FAR_C domain (Figure 1.3). The Rossmann fold structure has an α/β folding pattern (Rossmann, 1974; Richardson, 1981) and is found in the short-chain alcohol dehydrogenase/reductase (SDR) family of proteins, also known as the single-domain reductases/epimerases/dehydrogenases (RED)

family. The conserved GXXGXX(G/A) motif found near the N-termini of FARs is similar to the consensus sequence for ADP binding, [I/V/F]-X-[I/L/V]-T-G-X-T-G-F-L-[G/A] (Wierenga *et al.*, 1986; Eggink *et al.*, 1990). SDR proteins also have the catalytic motif YXXXK within the Rossmann fold domain (Jörnvall *et al.*, 1981; Ghosh *et al.*, 1995; Denessiouk *et al.*, 2001; Kavanagh *et al.*, 2008) and this is conserved in FAR enzymes (Figure 1.3). The FAR_C domain is also annotated as "Male Sterility 2" or "Sterile" domain since it was first reported in the MALE STERILITY2 (MS2/FAR2) predicted protein from *Arabidopsis thaliana* (Aarts *et al.*, 1997). This domain is unique to the FAR subfamily of reductases (Aarts *et al.*, 1997; Rowland *et al.*, 2006; Doan *et al.*, 2009). The FAR_C domain is located at the C-termini of FARs, but the function of this domain is presently unclear (Figure 1.3; Table 1.1).

Some FARs have an N-terminal extension containing a chloroplast target sequence (Figure 1.3) (Doan *et al.*, 2009; 2011; Chen *et al.*, 2011; Shi *et al.*, 2011). The majority of chloroplast proteins are encoded in the nuclear genome, translated in the cytosol, and post-translationally imported into the chloroplast; therefore, these proteins require an N-terminal chloroplast transit peptide to guide them to the chloroplast stroma. This cleavable N-terminal transit sequence of the precursor proteins functions as an envelope transfer stroma targeting domain, which ranges from 20 to 120 amino acids in length (von Heijne and Nishikawa, 1991; Soll and Tien, 1998). In general, the N-proximal portion (targeting sequence) does not contain positively charged residues, glycine or proline (von Heijne *et al.*, 1989; Soll and Tien, 1998). The central domain is rich in hydroxylated amino acids (such as serine and threonine), and does not usually contain acidic residues (von Heijne *et al.*, 1989; Soll and Tien, 1998). The C-terminal

portion of this targeting sequence usually has the conserved consensus sequence I/V-X-A/C-A near the cleavage site (Gavel and von Heijne, 1990; Soll and Tien, 1998). The N-terminal target sequence commences proteolytical processing concurrently with or shortly after translocation, by a soluble stromal processing peptidase (Soll and Tien, 1998).

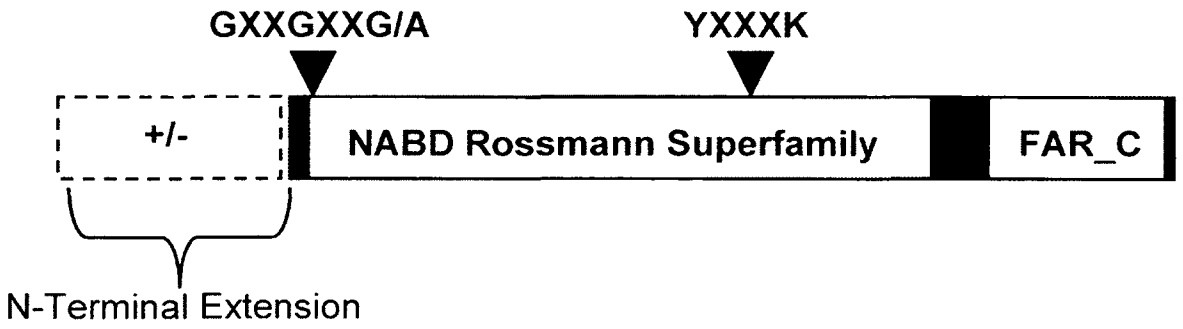


Figure 1.3 FAR domain structure. The Rossmann fold NAD(P)H-binding domain generally found in reductases is highlighted, with the GXXGXX(G/A) sequence motif for ADP/NAD(P)H binding indicated as well as the predicted catalytic residues YXXXK ('X' represents any amino acid). The FAR_C domain common to alcohol-forming fatty acyl reductases is also indicated. The regions encompassed by these two domains, including the region between, is about 490 amino acids. Some FARs have an N-terminal extension ranging in size from ~70 to ~120 amino acids and contain a chloroplast targeting sequence.

Table 1.1 Domains Predicted in the Arabidopsis Family of FARs and Jojoba FAR using the Conserved Domain Database (Marchler-Bauer *et al.*, 2011)

FAR*	Domain	Amino Acids	E-Value**
Jojoba FAR	NADB Rossmann Fold	14-364	3.53721e-100
	FAR_C	396-492	2.9951e-14
AtFAR1	NADB Rossmann Fold	12-360	2.99262e-113
	FAR_C	389-489	7.006e-17
AtFAR2/MS2	NADB Rossmann Fold	131-486	5.9136e-127
	FAR_C	550-612	1.86833e-22
AtFAR3/CER4	NADB Rossmann Fold	17-365	1.08414e-131
	FAR_C	394-493	1.52834e-17
AtFAR4	NADB Rossmann Fold	12-362	9.16733e-122
	FAR_C	391-491	2.28386e-17
AtFAR5	NADB Rossmann Fold	12-363	1.10411e-124
	FAR_C	420-494	6.37991e-14
AtFAR6	NADB Rossmann Fold	84-426	4.90168e-126
	FAR_C	471-539	3.31579e-20
AtFAR7	NADB Rossmann Fold	12-362	1.13646e-104
	FAR_C	390-475	1.445e-10
AtFAR8	NADB Rossmann Fold	12-363	7.62569e-105
	FAR_C	420-494	6.87928e-13

* The full length *Arabidopsis thaliana* CDSs obtained from the TAIR website: AtFAR1 (At5g22500), AtFAR2 (At3g11980), AtFAR3 (At4g33790), AtFAR4 (At3g44540), AtFAR5 (At3g44550), AtFAR6 (At3g56700), AtFAR7 (At5g22420), and AtFAR8 (At3g44560). The GenBank accession number of Jojoba FAR is AF149917.

** The E-value is a parameter that describes the number of hits expected when searching for matches against position-specific scoring matrices (PSSMs) in the database. PSSM profiles various amino acids present in a given position of a multiple sequence alignment for a domain model and the frequency of each one was observed. The closer the E-Value is to zero, the higher the significance is to the match. NADB (Nicotinamide Adenine Dinucleotide Binding) Rossmann Superfamily NCBI Accession is cl09931 and FAR_C Superfamily NCBI Accession is cl03838.

1.4 Function and Substrate Specificities of FARs

Evolutionarily related FARs have been identified in plants, mammals, insects, and microorganisms, each with a distinct substrate specificity (Figure 1.4; Table 1.2). I will first describe our knowledge of plant FARs, with regard to function and substrate specificity. The focus will be on Jojoba FAR and the Arabidopsis FAR family. As described above, there has been considerable biochemical characterization of the seed-expressed Jojoba FAR and it was the first to have its corresponding DNA coding sequence cloned. Considerable knowledge of FARs has since been gained from studies on Arabidopsis FARs. I will follow with a description of FAR function and specificities from non-plant organisms (summarized in Table 1.2).

1.5 Jojoba (Simmondsia chinensis) FAR

In Jojoba embryos, very-long-chain acyl-CoAs (mostly C20:1 and C22:1) are reduced to fatty alcohols by the seed-specific FAR, and this alcohol is combined with a fatty acid by the action of a fatty acyl-CoA:fatty alcohol acyltransferase (wax synthase) (Pollard *et al.*, 1979; Lardizabal *et al.*, 2000; Metz *et al.*, 2000). For Jojoba FAR, NADPH is the preferred reductant, although NADH was also found to be acceptable *in vitro* (Pollard *et al.*, 1979). Metz *et al.*, (2000) purified the FAR enzyme from developing Jojoba embryos and then cloned the corresponding Jojoba *FAR* cDNA (Figure 1.4). Expression of Jojoba FAR in *E. coli* results in accumulation of C12:0-OH, C14:0-OH, C16:0-OH and C18:1-OH fatty alcohols, which does not match the chain lengths found in Jojoba embryos (Metz *et al.*, 2000; Doan *et al.*, 2009). Jojoba FAR cDNA expressed in developing embryos of *Brassica napus* results in a series of wax esters (C38, C40, C42, C44) in the seeds (Lardizabal *et al.*, 2000). The most abundant wax ester found was

C40:2, comprised of C22:1 alcohol and C18:1 fatty acid (Table 1.2) (Lardizabal *et al.*, 2000). The presence of wax esters in *B. napus* seeds only expressing the Jojoba FAR is due to endogenous wax synthase activity. The fatty acid/fatty alcohol composition of the wax esters suggests that the Jojoba FAR has substrate specificity for very-long-chain acyl-CoAs *in planta* (Metz *et al.*, 2000). The Jojoba FAR along with a Jojoba wax synthase (WS) was expressed in Arabidopsis seeds along with a β-ketoacyl-CoA synthase (KCS; component of fatty acid elongase complex) from *Lunaria annua* (Lardizabal *et al.*, 2000). The *L. annua* KCS was chosen since it elongates C18:1 fatty acyl-CoA to C20:1, C22:1 and C24:1 fatty acyl-CoAs (Lassner, 1997), which would provide relevant substrates for Jojoba FAR. Co-expression of *L. annua* KCS and Jojoba FAR resulted in a small amount of wax esters, and C20:1, C22:1, and C24:1 free fatty alcohols (mostly C22:1-OH and C24:1-OH). Co-expression of *L. annua* KCS, Jojoba FAR, and Jojoba WS resulted in no free alcohols and significant amounts of wax esters (Lardizabal *et al.*, 2000). The large amount of C22:1-OH found in the wax esters of native Jojoba seeds suggests that this reflects the substrate preference of the Jojoba FAR (Lardizabal *et al.*, 2000).

1.6 Arabidopsis thaliana FAR Family

After the cloning of the seed-expressed Jojoba FAR, other FARs were found in different species, including an eight-membered family in Arabidopsis (*FAR1-8*) (Figure 1.4). The Arabidopsis *FAR* gene family members have diverse gene expression patterns and the encoded FAR enzymes have diverse substrate specificities as described below.

Arabidopsis *FAR3* (*At4g33790*), also known as *CER4*, encodes an alcohol-forming fatty acyl-CoA reductase that is involved in cuticular wax biosynthesis (Rowland

et al., 2006). FAR3/CER4 is the major and likely only FAR involved in generating the very-long-chain primary fatty alcohols found in Arabidopsis stem and leaf cuticular waxes. *FAR3/CER4* is expressed in epidermal cells along the entire length of the inflorescence (Rowland *et al.*, 2006). It is also expressed in flowers, siliques, leaves and roots, although its role in roots is currently unclear. FAR3/CER4 is localized to the ER when expressed in yeast cells (Rowland *et al.*, 2006), but ER localization *in planta* has not yet been verified. C24:0, C26:0 and C28:0 primary fatty alcohols are absent in Arabidopsis *cer4* mutants and C30 primary alcohol are significantly reduced but nonetheless detectable (Hannoufa *et al.*, 1993; Jenks *et al.*, 1995; Jenks *et al.*, 2002; Rowland *et al.*, 2006). The biochemical source of these residual C30 primary alcohols is currently unknown. *cer4* mutants also have 8-fold reduced levels of wax esters compared to wild type (Rowland *et al.*, 2006). The much lower density of epicuticular wax crystals with altered morphology in *cer4* mutants is likely specifically due to loss of free fatty alcohols because mutation of the wax synthase, WSD1, which results in no wax esters but wild-type fatty alcohol levels, is not altered in epicuticular wax crystal density or morphology (Li *et al.*, 2008). CER4 expressed in yeast produces C24:0 and C26:0 primary alcohols using endogenous acyl-CoA substrates (Rowland *et al.*, 2006).

Arabidopsis FAR1, FAR4, and FAR5 are responsible for generating C18:0-C22:0 fatty alcohols associated with suberin (Domergue *et al.*, 2010). *FAR1, FAR4,* and *FAR5* are all specifically expressed in root endodermal cells, which is a known location of suberin deposition. The transcripts of these three FARs are also upregulated in response to wounding and salt stress (Domergue *et al.*, 2010). This is in agreement with suberin reported as being wound induced (Dean and Kolattukudy, 1976; Kolattukudy, 2001).

FAR1, FAR4 and FAR5 heterologously expressed in transgenic yeast produce saturated fatty alcohols ranging from C18:0 to C22:0 (Domergue *et al.*, 2010). FAR1 mostly produces C22:0-OH and C18:0-OH, FAR4 mostly produces C20:0-OH and C18:0-OH, and FAR5 produced almost exclusively C18:0-OH (Table 1.2). This indicates that FAR1, FAR4, and FAR5 have distinct chain length specificities. T-DNA insertion lines of FAR1, FAR4, and FAR5 were used to investigate their roles in fatty alcohol production in suberin of roots, seed coats, and wounded leaf tissue (Domergue *et al.*, 2010). The *far1*, *far4* and *far5* mutant lines were deficient in C22:0, C20:0 and C18:0 suberin-associated fatty alcohols, respectively. The fatty alcohol chain-length profile of suberin-associated root waxes in Arabidopsis matches the substrate specificities of FAR1, FAR4, and FAR5 (C18:0-C22:0 alcohols) (Li *et al.*, 2007), which suggests that these three FARs also provide the fatty alcohols for root waxes.

FAR2 (*At3g11980*), also known as *MALE STERILITY2* (*MS2*), encodes an anther-specific protein involved in the synthesis of fatty alcohols that are thought to be associated with sporopollenin, an important constituent of the outer pollen wall (exine) (Figure 1.4) (Aarts *et al.*, 1995; Aarts *et al.*, 1997; Chen *et al.*, 2011). Pollen exine is essential for pollen fertility, protecting gametophytes against pathogen attack, dehydration, UV irradiation, and facilitating pollen recognition and adhesion to the stigma (Aarts *et al.*, 1997; Jung, 2006; Doan *et al.*, 2009). FAR2/MS2 encodes a fatty acyl-acyl carrier protein (ACP) reductase that requires NADPH or NADH for activity (Chen *et al.*, 2011). *FAR2/MS2* is expressed in the tapetum shortly after the microspore is released from the tetrad (Aarts *et al.*, 1997). Cross-sections of mature anthers revealed that *ms2* mutants do not have a pollen exine layer, and the tapetal layer in young anthers

was abnormal resulting in a high proportion of pollen abortion shortly after release from tetrads (Aarts *et al.*, 1993). The pollen walls of *ms2* mutants are collapsed and have a shrunken appearance and they are also sensitive to acetolysis treatment, indicating that the exine layer has been severely compromised (Aarts *et al.*, 1997; Chen *et al.*, 2011). FAR2/MS2 is localized to plastids where C16 and C18 long-chain fatty acids are synthesized *de novo* (Chen *et al.*, 2011). FAR2/MS2 has strong activity for C16:0-ACP *in vitro*, trace activity with C18:0-ACP, and no activity for C14:0-ACP, C16:1-ACP, and C18:1-ACP or any the equivalent acyl-CoA substrates (C14:0/1-C18:0/1-CoA) (Chen *et al.*, 2011). Doan *et al.* (2009) found that FAR2/MS2 produces C14:0-OH, C16:0-OH and C18:1-OH using endogenous acyl-ACP pools in *E. coli* (Table 1.2). When FAR2/MS2 was transiently expressed in leaves of *Nicotiana benthamiana,* significant amounts of C16:0-OH and C18:0-OH were generated (Chen *et al.*, 2011). Rice *DEFECTIVE POLLEN WALL* (*DPW*) is the ortholog of Arabidopsis *FAR2/MS2* (Shi *et al.*, 2011). *DPW* encodes a fatty acyl reductase that is expressed in tapetal cells and microspores. It shares 59% amino acid identity with Arabidopsis MS2. Rice *dpw* mutants display defective anther development, degenerated pollen grains, and irregular exine on the pollen grains (Shi *et al.*, 2011). DPW is localized to plastids similar to FAR2/MS2. Purified recombinant DPW enzyme converts C16:0-ACP, C16:0-CoA and C16:1-CoA to their corresponding fatty alcohols, with the highest affinity for C16:0-ACP and no activity towards C14:0-, C16:1-, C18:0-ACP, C10:0-, C12:0-, C14:0-, C18:0-, C18:1-, C18:2-CoA, or C16 and C16:1 free fatty acids. DPW activity required NADPH, and there was no activity detected in the presence of NADH. Three anther-expressed FARs genes have been identified in bread wheat (*Triticum aestivum*) called *TAA1a*, *TAA1b*, and

TAA1c (Wang *et al.*, 2002). *TAA1a* transcript was shown to be specifically expressed in the tapetum during the formation of the outer cell wall of pollen grains (after the tetrads resolved into young, free microspores). Transgenic expression of TAA1a produced long-chain and very-long-chain fatty alcohols in tobacco seeds (*Nicotiana tabacum* cv. *Xanthi*) (C18:1, C20:1, C22:1, C24:0, C26:0) and in *E. coli* (C14:0, C16:0, C18:1) (Wang *et al.*, 2002). The predicted TAA proteins do not have a predicted chloroplast-targeting sequence and are thus likely extraplastidic. They are also not in the phylogenetic clade containing MS2 and DPW (Figure 1.4) and are likely functioning in anther and pollen exine development but in a manner different than that of the Arabidopsis and rice plastid FARs.

The roles of Arabidopsis FAR6, FAR7, and FAR8 are unclear (Figure 1.4). The only cloned *FAR7* (*At5g22420*) cDNA contains an early stop codon and encodes a truncated, inactive protein (Doan *et al.*, 2009). *FAR7* is thus likely a pseudogene. FAR8 has very low activity (See Chapter 3 and 4 of this thesis) with near undetectable gene expression (Domergue *et al.*, 2010). *FAR8* may also be a pseudogene. *FAR6* is highly expressed in stems and enriched in the epidermis (Suh *et al.*, 2005; Domergue *et al.*, 2010). FAR6 characterization is further described in Chapter 3 of this thesis.

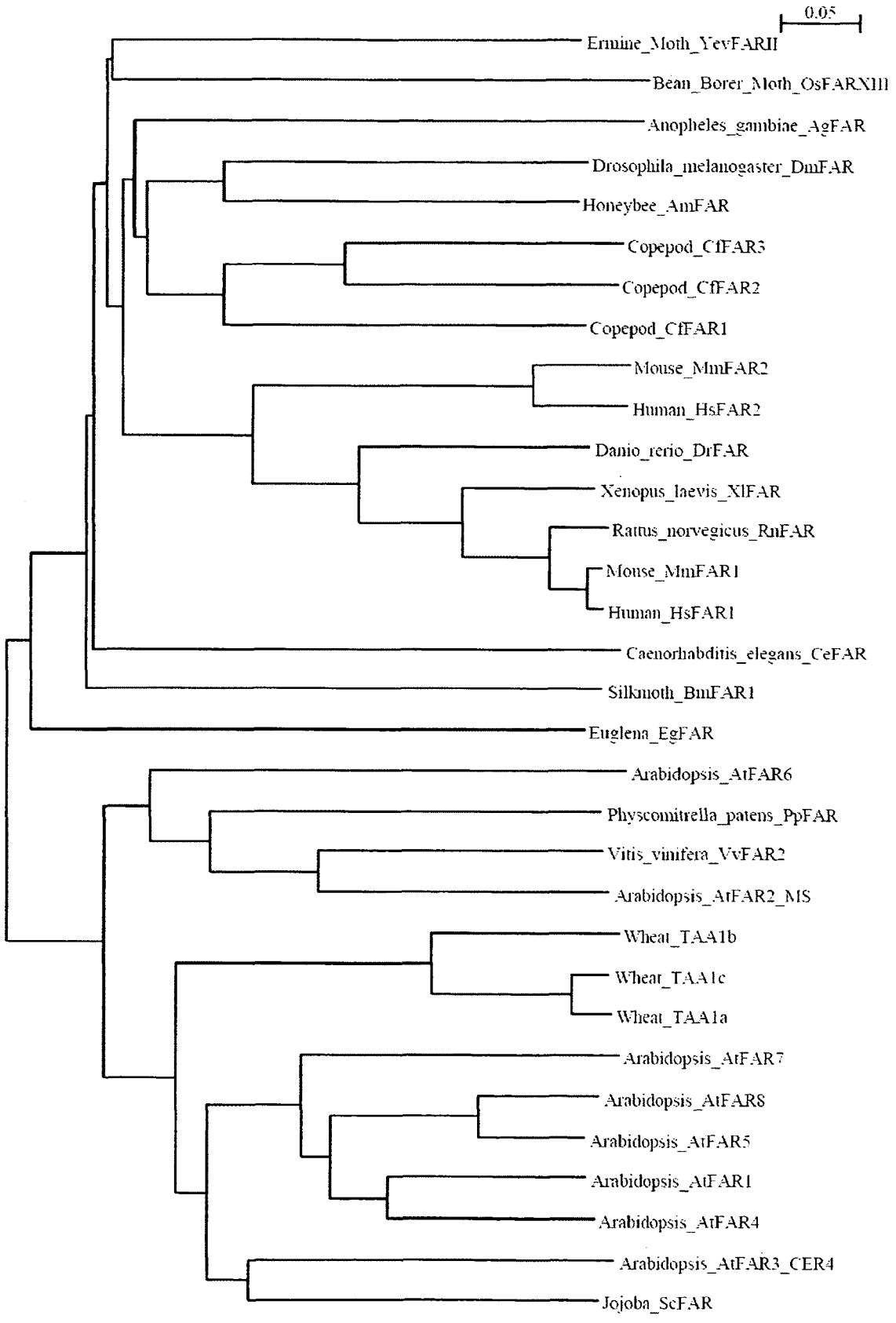

0.05
Ermine_Moth_YevFARII
Bean_Borer_Moth_OsFARXIII
Anopheles_gambiae_AgFAR
Drosophila_melanogaster_DmFAR
Honeybee_AmFAR
Copepod_CfFAR3
Copepod_CfFAR2
Copepod_CfFAR1
Mouse_MmFAR2
Human_HsFAR2
Danio_rerio_DrFAR
Xenopus_laevis_XlFAR
Rattus_norvegicus_RnFAR
Mouse_MmFAR1
Human_HsFAR1
Caenorhabditis_elegans_CeFAR
Silkmoth_BmFAR1
Euglena_EgFAR
Arabidopsis_AtFAR6
Physcomitrella_patens_PpFAR
Vitis_vinifera_VvFAR2
Arabidopsis_AtFAR2_MS
Wheat_TAA1b
Wheat_TAA1c
Wheat_TAA1a
Arabidopsis_AtFAR7
Arabidopsis_AtFAR8
Arabidopsis_AtFAR5
Arabidopsis_AtFAR1
Arabidopsis_AtFAR4
Arabidopsis_AtFAR3_CER4
Jojoba_ScFAR

Figure 1.4 Phylogenetic tree of related FAR protein sequences. The GenBank accession numbers* of the sequences are as follows: *Bean Borer Moth OsFARXIII, EU817405; Copepod (Calanus finmarchicus) CfFAR1, JN243755; Copepod (Calanus finmarchicus) CfFAR2, JN243756; Copepod (Calanus finmarchicus) CfFAR3, JN243757; Ermine Moth (Yponomeutidae evonymellus) YevFARII, GQ907232; Fruit Fly (Drosophila melanogaster) DmFAR, GI24654209; Grape (Vitis vinifera) VvFAR, XP 002276588; Honeybee (Apis mellifera) AmFAR, NM_001193290; Human (Homo Sapien) HsFAR1, AY600449; Human (Homo Sapien) HsFAR2, BC022267; Jojoba (Simmondsia chinensis) ScFAR, AF149917; Mosquito (Anopheles gambiae) AgFAR, XP_307899.1; Mouse (Mus Musculus) MmFAR1, BC007178; Mouse (Mus Musculus) MmFAR2, BC055759; Nematode (Caenorhabditis elegans) CeFAR, NP_508505.1; Physcomitrella patens, XP 001758118; Phytoflagellate (Euglena gracilis) EgFAR, GU733919; Rat (Rattus norvegicus) RnFAR, NP_001011933.1; Silkmoth (Bombyx mori) BmFAR, AB104896; Toad (Xenopus laevis) XlFAR; Wheat FAR (TaTAA1a), AJ459249, GI28277293, and the full length Arabidopsis thaliana CDSs from the TAIR website for AtFAR1 (At5g22500), AtFAR2 (At3g11980), AtFAR3 (At4g33790), AtFAR4 (At3g44540), AtFAR5 (At3g44550), AtFAR6 (At3g56700), AtFAR7 (At5g22420), AtFAR8 (At3g44560)* *obtained from NCBI (Carnegie Institution of Washington Department of Plant Biology, 2008; U.S. National Library of Medicine, 2009). Corresponding branch lengths are proportional to the amount of inferred evolutionary change. The tree was constructed using CLUSTALW2 and NJ plotWIN95

1.7 FARs in Non-Plant Organisms

1.7.1 Phytoflagellate Protist (Euglena gracilis)

Euglena gracilis is a unicellular phytoflagellate protist that accumulates large amounts of medium and long-chain wax esters (C20-C36) under anaerobic growth conditions (62%), compared to aerobic conditions (28%) (Tucci *et al.*, 2009). Temperature, nutrients, oxygen and light influence the production and composition of the wax esters that accumulate (Tucci *et al.*, 2009). Two genes in *E. gracilis* encoding enzymes involved in wax ester biosynthesis have been identified and characterized: a fatty acyl-CoA reductase (EgFAR) catalyzing the conversion of fatty acyl-CoAs to fatty alcohols and a wax synthase (EgWS) catalyzing the esterification of fatty acyl-CoAs and fatty alcohols (Figure 1.2) (Kolattukudy, 1970; Teerawanichpan and Qiu, 2010[b]). Biochemical assays showed that EgFAR requires ATP, CoA, and NADH as cofactors (Kolattukudy, 1970). EgFAR expressed in *S. cerevisiae* converts exogenously fed C14:0 and C16:0 fatty acids to the corresponding primary fatty alcohols, with C14:0 as the preferred substrate (Table 1.2). No fatty alcohols were detected when yeast were fed the saturated fatty acids C10:0, C12:0, or C18:0 or the unsaturated fatty acids C16:1n-9, C18:1n-9, or C18:2n-6. EgWS is able to use a wide range of fatty acyl-CoAs and fatty alcohols as substrates, although C14:0-CoA and 16:0-OH are the preferred substrates (Teerawanichpan and Qiu, 2010[b]). Reconstituting the wax biosynthetic pathway in yeast strain H1246, which contains four disrupted acyltransferase genes involved in triacylglycerol and sterol ester biosynthesis, by co-expressing EgFAR and EgWS in the presence of exogenously fed C14:0 yields C14:0-C14:0, C14:0-C16:1, C14:0-C16:0, and

C16:0-C14:0 wax esters, accounting for ~23%, ~22%, ~32%, and ~23% of the total wax esters produced, respectively (Teerawanichpan and Qiu, 2010[b]).

1.7.2 Moths

FARs are required for sex pheromone biosynthesis in certain moth species. Female moths produce the sex pheromones and the major classes are made up of C10-C18 unsaturated acyclic, aliphatic compounds containing an oxygenated functional group (acetate ester, alcohol or aldehyde) (Moto *et al.*, 2003).

Small ermine moths (*Lepidoptera: Yponomeutidae*) have sex pheromones that are synthesized *de novo* in a specialized gland in female moths and are derived from fatty alcohols. There are three distinct FARs that have been identified in *Y. evonymellus*: *Yev*-FARI, *Yev*-FARII, and *Yev*-FARIII (Figure 1.4) (Liénard *et al.*, 2010). They are 32-36% identical at the amino acid level. *Yev*-FARI and *Yev*-FARIII are broadly expressed throughout tissues, whereas *Yev*-FARII is female-specific and expressed solely in the pheromone gland (*Yev*-FARII = *Yev*-pgFAR). Heterologous expression of *Yev*-pgFAR in *S. cerevisiae* supplied with saturated substrates ranging from C8:0-C24:0 resulted mostly in the production of C14:0 and C16:0 fatty alcohols with minor amounts of C12:0 fatty alcohol (Liénard *et al.*, 2010). *Yev*-pgFAR is unable to reduce any saturated fatty acyl substrates ranging from C8:0-C10:0 and C18:0-C24:0 (Table 1.2). Since *Yev*-pgFAR produces alcohols solely in the pheromone gland, it is predicted that it plays a critical role in producing multi-component pheromones of *Yponomeutids* (Liénard *et al.*, 2010).

The silkmoth (*Bombyx mori*) contains a pheromone gland-specific long-chain FAR (BmFAR), which is the key enzyme that converts the fatty-acyl pheromone

precursor to bombykol (*E,Z*-10,12-hexadecadien-1-ol), the final active product of the biosynthetic pathway (Figure 1.4) (Moto *et al.*, 2003). Bombykol is biosynthesized from *Z*11-16:Acyl, desaturated to *E,Z*-10,12-16:Acyl, and reduced to *E,Z*-10,12-16:0-OH (bombykol). BmFAR expressed in *S. cerevisiae* fed with saturated C14:0-C20:0 long chain fatty acids preferentially reduces C15 and C16 acids to fatty alcohols, C17 and C18 acids less well, C14 barely at all, and C19 and C20 not at all (Table 1.2). Using mono- and di-unsaturated C16 fatty acids, BmFAR strictly recognizes the position and isomeric nature of the double bonds and displays a strong preference for the pheromone precursor fatty acid (*E,Z*-10,12-hexadecadienoic acid). Transgenic yeast expressing the silkmoth FAR fed with the pheromone precursor fatty acid evokes mating behaviour in male *B. mori* moths (Moto *et al.*, 2003).

The European corn borer (*Ostrinia nubilalis*) consists of two sex pheromone races that is a result of the utilization of different ratios of *cis* (*Z*) and *trans* (*E*) acetate pheromone isomer components. The *E* race uses a 98:2 blend of (*E*)-11-tetradecenyl acetate and (*Z*)-11-tetradecenyl acetate, whereas the *Z* race uses a 3:97 *E/Z* blend (Klun, 1975). These components are synthesized from palmitic acid (C16:0), which is chain-shortened by β-oxidation into myristate (C14:0) followed by Δ11 desaturation resulting in (*E*)-11-tetradecenoyl and (*Z*)-11-tetradecenoyl. These acids are subsequently reduced to fatty alcohols by FARs and acetylated to give (*E*)-11-tetradecenyl and (*Z*)-11-tetradecenyl acetates. The pheromone gland FAR (pgFAR) from *O. nubilalis* from the two races, pgFAR-E and pgFAR-Z, were expressed in yeast supplied with (*E*)-11-tetradecenoyl and (*Z*)-11-tetradecenoyl precursors to determine if their divergence affected substrate specificity. pgFAR-Z almost exclusively converts the *Z* isomer with

minimal amounts of the *E* isomer being reduced, and pgFAR-E almost exclusively converts the *E* isomer with minimal amounts of the *Z* isomer being reduced (Table 1.2). This shows that allelic variation in a *FAR* gene involved in pheromone biosynthesis is responsible for phenotypic variation in the production of female pheromones, which led to race-specific signals and reproductive isolation in moths (Lassance *et al.*, 2010).

1.7.3 Honey bees (Apis mellifera)

Honey bees produce varied amounts of long-chain aliphatic fatty alcohols that range from C18-C32 with variation between each body segment. The honey bee FAR (AmFAR1) converts a wide range of fatty acyl-CoAs to their corresponding alcohols in the presence of NADPH (Figure 1.4). These fatty alcohols are constituents of pheromones, comb wax, and ether lipids. The comb wax contains wax esters that consist mainly of C16:0 and 15-hydroxy-C16:0 (15-OH-C16:0) fatty acids esterified with a very-long-chain fatty alcohol ranging from C24-C34 (Granier *et al.*, 2002). Honey bees also produce alcohols shorter than 24 carbons for other biological functions, such as the cuticular layer of drone cocoon (elicits strong arrestment response in *Varroa* mites) consisting of C17:0-C22:0 fatty alcohols and C19:0-C22:0 aldehydes (Donzé *et al.*, 1998). AmFAR1 was found to be expressed ubiquitously in all body segments with predominance in the head. AmFAR1 heterologously expressed in yeast produces C16:0-OH, C18:0-OH, C20:0-OH and C22:0-OH from the endogenous fatty acyl pool. Yeast expressing AmFAR1 was also exogenously supplied with saturated fatty acids (C14:0-C22:0), unsaturated fatty acids (C16:1, C18:1n-9, C18:2n-6, C18:3n-3), and hydroxyl fatty acids (12-OH-C18:1n-9, 12-OH-C18:1n-9) to test for substrate specificity. The

highest conversion efficiency was for C18:0, followed by C20:0, C22:0, C20:1n-9, C16:0, respectively, and some activity for 12-OH-C18:1n-9, C16:1n-9 and C18:1n-9 (Table 1.2). A wax biosynthetic pathway was reconstituted by expressing AmFAR1 and *Euglena* wax synthase (EgWS) in the neutral-lipid deficient yeast strain H1246 (Teerawanichpan *et al.*, 2010[a]). The transgenic yeast were grown in the presence of a wide range of fatty acids: C14:0, C16:0, C16:1, C18:0, C18:1, C20:0, and C22:0. The C14:0-OH, C16:0-OH and C18:0-OH produced by AmFAR1 became esterified with C14:0, C16:0, C16:1 and C18:0 fatty acids yielding C14:0-C14:0 (~23% of total wax esters), C16:0-C14:0 (~15% of total wax esters), C16:0-C16:0 (~6.5% of total wax esters), C18:0-C14:0 (~26% of total wax esters), C18:0-C16:0 (~19% of total wax esters), and C18:0-C16:1 (~11% of total wax esters) (Teerawanichpan *et al.*, 2010[a]).

1.7.4 Copepod (Calanus finmarchicus)

Three genes were identified in the marine copepod (*Calanus finmarchicus*), *CfFAR1, CfFAR2* and *CfFAR3*, which encode enzymes likely responsible for a series of fatty alcohol moieties present in wax esters that accumulate in this organism (Figure 1.4). This small crustacean, which constitutes a considerable amount of the Arctic and Northern seas biomass, accumulates high levels of wax esters mainly comprising C20:1n−9 and C22:1n−11 fatty alcohols linked with various fatty acids (i.e. n−3 polyunsaturated fatty acids). These three FARs all produce fatty alcohols, but with distinct substrate specificities: CfFAR1 is active toward saturated C18:0-C26:0 fatty acyl-CoAs, CfFAR2 is active toward C24:0 and C26:0 saturated very-long-chain fatty acyl-CoAs, and CfFAR3 is active towards both saturated (C16:0 and C18:0) and unsaturated

(C18:1 and C20:1) fatty acyl-CoAs (Table 1.2). It is likely that the activity of CfFAR3 is at least in part responsible for the observed fatty alcohol profile of *C. finmarchicus* wax esters consisting of both saturated and monounsaturated alcohols; C20:1n−9Alc (52.4% of total fatty alcohols) and C22:1n−11Alc (31.9% of total fatty alcohols) are the major fatty alcohols, with C16:0Alc, C16:1n−9Alc, C18:1n−9Alc, and C18:1n−6Alc being minor components (12.8% (Teerawanichpan and Qiu, 2011). The contributions of CfFAR1 and CfFAR2 are unclear, but they may be involved in the production of other less abundant metabolites.

1.7.5 Mouse and Human

Fatty alcohols in mammals have thus far been found either incorporated into ether lipids or wax esters. There are two described fatty acyl-CoA reductase isozymes, FAR1 and FAR2, each in mouse (m) and human (h) (Figure 1.4) (Cheng and Russell, 2004). Mouse *FAR1* is expressed the highest in the preputial gland and *FAR2* is expressed at lower levels in a smaller number of tissues with the highest levels in the eyelid and skin. Both *FAR1* and *FAR2* are expressed in the brain where large quantities of ether lipids are synthesized (Cheng and Russell, 2004). Mouse and human FAR1 and FAR2 expressed in Human Embryonic Kidney (HEK) 293 cells convert C16:0 acyl-CoA to C16:0-OH, with FAR1 being more active than FAR2 in both human and mouse (Cheng and Russell, 2004). Mouse FAR1 and FAR2 were also expressed in insect ovarian cells (*Spodoptera frugiperda* (Sf) 9 cells), which produced higher levels of enzyme activity than that obtained with enzymes expressed in the HEK 293 cells. Again, mouse FAR1 activity was 5-10 fold higher than FAR2 using this expression system. In the Sf9 cells, FAR1

preferentially reduces C16:0, C18:0, C18:1, and C18:2 fatty acyl-CoAs, has less activity for C20:3 and C20:4, and weak activity is observed for C10-C14. FAR2 preferentially reduces saturated C16 and C18 fatty acyl-CoAs and shows very weak activity for C10:0-C14:0 (Table 1.2). In summary, mouse FAR1 acts on fatty acyl-CoAs of different chain lengths and degrees of saturation, while mouse FAR2 prefers saturated C16 and C18 fatty acyl-CoAs as substrates suggesting that FAR1 plays a general role in fatty alcohol synthesis and FAR2 has a more specialized function. Also, it was found that NADPH was required as a cofactor for FAR1 and there was no activity using NADH (FAR2 was not tested) (Cheng and Russell, 2004). Confocal light microscopy of the transfected cells revealed that FAR1 and FAR2 were localized to the peroxisome, which is consistent with these FARs being required to produce fatty alcohols for ether lipid biosynthesis (Cheng and Russell, 2004).

Table 1.2 Substrate Specificities of FARs from Various Organisms

Organism	Substrate Specificity			Publication
	Escherichia coli	*Saccharomyces cerevisiae*	Plant/Mammalian Cells	
Phytoflagellate Protist *(Euglena gracilis)*		14:0, 16:0		Teerawanichpan and Qiu, 2010[b]
Small Ermine Moths *(Yponomeutidae sp.)*		12:0, 14:0, 16:0		Liénard *et al.*, 2010
Silkmoth *(Bombyx mori)*		14:0, 15:0, 16:0, 17:0, 18:0, E,Z10,12–hexadecadienoic acid		Moto *et al.*, 2003
Honey bees *(Apis mellifera)*	18:0, 20:0, 22:0	16:0, 18:0, 20:0, 22:0, 12-OH-18:1n-9, 16:1n-9, 18:1n-9, 20:1n-9		Teerawanichpan *et al.*, 2010
Copepod *(Calanus finmarchicus)*		ClFAR1: 18:0-26:0 ClFAR2: 24:0, 26:0 ClFAR3: 16:0, 18:0, 18:1, C20:1		Teerawanichpan and Qiu, 2011
Mouse (*Mus musculus*)			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 (*Homo sapiens*)			FAR1: 16:0 FAR2: 16:0 (HEK 293 cells)	Cheng and Russell, 2004
Jojoba *(Simmondsia chinensis)*	12:0, 14:0, 16:0, 18:1		22:1 (HEAR *Brassica napus*)	Metz *et al.*, 2000; Doan *et al.*, 2009; Lardizabal *et al.*, 2000
Thale Cress *(Arabidopsis thaliana)*	FAR1: 14:0, 16:0, 18:0, 18:1 FAR2: 14:0, 16:0, 18:0, 18:1 (ACP) FAR3: 14:0, 16:0, 18:1 FAR6: 12:0-ACP, 14:0 (CoA, ACP) 16:0 (CoA, ACP), 18:0-CoA, 18:1-ACP, 20:0-CoA FAR8: 14:0, 16:0, 18:1	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 *et al.*, 2006; Doan *et al.*, 2009; Domergue *et al.*, 2010; Chen *et al.*, 2011; Doan *et al.*, 2011
Bread Wheat (*Triticum aestivum*)	14:0, 16:0, 18:1		18:1, 20:1, 22:1, 24:0, 26:0 (N.t.)	Wang *et al.*, 2002
Pea Leaves *(Pisum sativum L.)*			16:0-CoA	Vioque and Kolattukudy, 1997
Corn Borer Moth *(Ostrinia nubilalis)*		pgFAR-Z: (Z)-11-tetradecenoyl pgFAR-E: (E)-11-tetradecenoyl		Lassance *et al.*, 2010

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

1.8 Effect of Heterologous Expression Host on Apparent FAR Substrate Specificities

One obstacle that exists in elucidating substrate specificities of FARs is that the fatty alcohols produced by a FAR can vary depending on what host the FAR is expressed in, whether it is expressed *in planta*, in *S. cerevisiae*, or in *E. coli* systems. The composition of the acyl pools between the host systems can have a powerful effect on the apparent substrate profile of the enzyme. For example, the acyl pools in *E. coli* and yeast differ significantly with regard to chain length profile and acyl linkage. The acyl chains in *E. coli* are mostly linked to ACP and it contains fatty acyl chains only up to C18, whereas in yeast the acyl chains are mostly linked to CoA and it contains fatty acyl chains up to C26. In plants, the acyl pool in plastids (i.e. chloroplasts and leucoplasts) is mostly acyl-ACP, whereas the acyl pool outside the plastid is mostly acyl-CoA. Therefore, unless the FAR of interest is plastid localized, expression in yeast is typically a better indicator of *in planta* substrate specificity than the *E. coli* system (Wang *et al.*, 2002; Rowland *et al.*, 2006; Doan *et al.*, 2009; Teerawanichpan *et al.*, 2010[a]; Doan *et al.*, 2011). Transient expression of FARs in leaves of *Nicotiana benthamiana* is an alternative approach that has been recently reported (Doan *et al.*, 2011), but it is potentially complicated by endogenous FAR activities and it still doesn't necessarily reflect the acyl pools of the plant cells (e.g. root, flower, seed) that express the FAR under study. *In vitro* assays with FARs have been difficult to develop, especially with the extraplastidial FARs that are likely membrane localized. Also, acyl-CoAs longer than C16 and C18 are not commercially available and very expensive to synthesize.

1.9 Industrial Applications of FARs

Fatty alcohols and wax esters are valuable, as they have a variety of commercial applications such as in detergents and other cleaning products, high-performance industrial lubricants, cosmetics, and pharmaceutical formulations (Kalscheuer and Steinbüchel, 2003). Fatty alcohols are currently produced from reduction of vegetable oils or synthesized from petrochemicals (Kreutzer, 1984). Wax esters were extracted in the past from the spermaceti organ in the heads of sperm whales, but this led to the near extinction of sperm whales and hunting is now banned. Metabolic engineering of high-yielding oilseed crops is an alternative method of production that may provide a rapid and cost-effective approach to produce commercially valuable fatty alcohols and wax esters. This can be accomplished by expressing FARs and wax synthases of desired specificities in seeds of target oilseed crops. The substrate specificity of a FAR has been shown to have a great influence on the final product, with regards to the chain length and degree of saturation. The chain length and degree of saturation of the primary alcohols significantly influence the physical properties of wax esters produced by wax synthases. Similar transgenic approaches may be feasible with a microbe, such as *E. coli* or an industrial yeast (i.e. *Yarrowia lipolytica*).

1.10 Thesis Objectives

The two main objectives of this research were to: (1) characterize the substrate specificities of the Arabidopsis FAR enzyme family as well as Jojoba FAR using yeast as an expression system, and (2) to dissect in detail the substrate specificities of FAR5 and FAR8 by determining which amino acids are important for activity and substrate specificity. This information is important for protein engineering of FARs with high activities and desired substrate specificities for industrial applications.

2.0 Materials and Methods

2.1 Plant Material and Growth Conditions

Arabidopsis thaliana ecotype Columbia-0 (Col-0) seeds were obtained from Dr. Owen Rowland (Carleton University, Ottawa, Canada). All seeds were surface sterilized in a microcentrifuge tube using 100% ethanol, mixed by inversion 5 times, followed by incubation in 50% bleach/0.5% SDS solution, mixed by inversion 5 times, rinsed with sterile deionized water 3 times with mixing by inversion 10 times, and then suspended in sterile 0.1% agarose. The seeds were then plated on AT-agar (2 mM $MgSO_4$, 2.5 mM KH_2PO_4, 50 μM Fe(EDTA), 2 mM $Ca(NO_3)_2$, 5 mM KNO_3, micronutrients 10μM H_3BO_3, 14μM $MnCl_2$, 0.5μM $CuSO_4$, 1μM $ZnSO_4$, 0.2 μM $NaMoO_4$, 10μM NaCl, and 10nm $CoCl_2$, and solidified with 0.7% (w/v) agar in sterile Petri plates). The seeds were stratified on AT-agar for 4 days at 4°C and in the dark. Plates were sealed with micropore tape and placed in a growth chamber (Conviron, model TC 26) at 22°C under continuous light (90 μmol m^{-2} s^{-1}) conditions for 7-10 days (until first true leaves appeared). Seedlings were transplanted to soil (Pro-Mix MPV Multipurpose Growing Medium) containing 20-20-20 fertilizer, covered with cellophane for 5 days and grown in a growth chamber (Percival, model AR66L) at 21°C and 80% relative humidity under continuous light (110 μmol m^{-2} s^{-1}) conditions.

2.2 RNA Extraction

Flower and stem tissues were harvested from 6-week-old *Arabidopsis thaliana* Col-0 plants. Plant tissue was wrapped in aluminum foil and immediately submerged in

liquid nitrogen, then stored at -80°C. 1.5ml microcentrifuge tubes and micropestles were pre-chilled in liquid nitrogen and remain chilled throughout the experiment in a rack submerged in liquid nitrogen. The frozen plant tissue was placed in the chilled microcentrifuge tubes and ground briefly using the chilled micropestle. 600µl of TRIzol® reagent (Invitrogen), which is a proprietary mixture of guanidinium thiocyanate (15-40%) and phenol (30-60%), was added to the ground-up plant tissue, and the mixture ground to form a fine homogenate. A further addition of 400µl TRIzol reagent was added and briefly ground, and then the homogenized samples were incubated at room temperature for 5 minutes. After the incubation, 200µl of chloroform was added, followed by vigorous shaking for 15 seconds and then incubated at room temperature for 3 minutes. Following the incubation, the samples were centrifuged at 12 000 g for 15 minutes at 4°C. The aqueous phase was transferred to a fresh 1.5ml microcentrifuge tube containing 0.5ml of isopropyl alcohol, mixed and incubated at room temperature for 10 minutes. After the incubation, the samples were centrifuged at 12 000g for 10 minutes at 4°C. The supernatant was poured off leaving the RNA pellet. 1ml of 75% ethanol was added to the RNA pellet and mixed by vortexing for 15 seconds. The sample was centrifuged at 7 500g for 5 minutes at 4°C, the ethanol was poured off and the RNA pellet was allowed to air dry for 10 minutes at room temperature in the fume hood. The dried RNA was dissolved in 50µl of diethylpyrocarbonate (DEPC)-treated water and stored at -80°C.

2.3 cDNA Synthesis

Complementary DNA (cDNA) was made using the flower and stem RNA extracted from *Arabidopsis thaliana* Col-0 plants. The RNA samples were diluted to a

concentration of 0.12 μg/μl using DEPC-treated water. For each sample, 1μg (8.5 μl) of RNA was heated at 70°C for 5 minutes and snap chilled on ice. For a final volume of 20μl, the RNA samples (8.5μl) were then added to 11.5μl of the reverse transcriptase reaction mixture (3μM oligo-dT primer, 0.5mM dNTPs, 5U RNaseOUT™ (Invitrogen), 10mM DTT, 50 mM Tris-HCl (pH 8.3, 75 mM KCl, 3 mM $MgCl_2$) on ice. 50 units of SuperScript III® Reverse Transcriptase (Invitrogen) were added to each reaction and held at 50°C for 60 minutes, 72°C for 10 minutes, cooled down to 4°C and stored at -20°C.

2.4 Yeast Expression Constructs for Heterologous Expression of FAR Proteins

pYES2 (Invitrogen) is a yeast expression vector used to express proteins of interest in *Saccharomyces cerevisiae*. It contains a *URA3* gene for selection in yeast and a 2μ origin of replication for high-copy maintenance. It also contains the *GAL1* promoter that is used for high level inducible protein expression and is repressed by glucose. A modified version of pYES2 was used that contained a histidine (x6) and T7 epitope tag for protein purification and detection, respectively.

The full length open reading frames for FAR1, FAR2(MS2), FAR3(CER4), FAR4, FAR5, FAR6, FAR8, Jojoba FAR, as well as versions of FAR2 and FAR6 with truncations of their N-terminal extension sequences (denoted FAR2ΔN and FAR6ΔN, respectively), were cloned into the modified version of pYES2 (pYES2-His6x/T7). FAR2ΔN and FAR6ΔN lacked the first 119a.a. and 71a.a., respectively, at the N-terminus. *FAR5* was previously cloned into pYES2-His6x/T7 using *Bam*HI and *Xho*I by Frances Tran (Rowland Lab, Carleton University) and *FAR1* and *FAR8* was previously cloned into pYES2-His6x/T7 using *Bam*HI and *Xho*I by Sollapura Vishwanath (Rowland

Lab, Carleton University). The *FAR* family, including *FAR2ΔN* and *FAR6ΔN*, as well as domain swaps, site directed mutants and the Jojoba FAR were amplified by PCR using iProof™ High-Fidelity DNA polymerase (Bio-Rad). Each PCR contained 1X High-Fidelity iProof Buffer (containing 1.5mM $MgCl_2$), 0.2mM dNTPs, 1-5ng DNA template, 0.5μM of each Primer (Forward and Reverse), and 0.5 Units of iProof DNA Polymerase. *FAR2*, *FAR2ΔN*, *FAR3* and *FAR4* were amplified using flower cDNA as a template, the short versions of FAR6 and FAR6ΔN were amplified using stem cDNA as a template, the long versions of FAR6 and FAR6ΔN were amplified using pET15b:At3G56700 plasmid (Doan *et al.*, 2009) as a template, and the Jojoba FAR was amplified using pBINGlyRed-FWS3-Jojoba FAR/WS/KCS plasmid as a template (kindly provided by Ed Cahoon, University of Nebraska-Lincoln). PCR cycling conditions included an initial denaturation step of 98°C for 30 seconds, followed by 30 cycles of 98°C for 15 seconds (denaturation), 57°C for 30 seconds (annealing), and 72°C for 1 minute (extension), with a final single step of 72°C for 5 minutes to ensure fully elongated products. DNA products were amplified using the primers FAR1_BamHI_Forward and FAR1_XhoI_Reverse, FAR2_BglII_Forward and FAR2_SphI_Reverse, FAR2N_BglII_Forward and FAR2_SphI_Reverse, FAR3_BamHI_Forward and FAR3_XhoI_Reverse, FAR4_BamHI_Forward and FAR4_XhoI_Reverse, FAR5_BamHI_Forward and FAR5_XhoI_Reverse, FAR6_EcoRI_Forward and FAR6_SphI_Reverse, FAR6N_EcoRI_Forward and FAR6_SphI_Reverse, FAR8_BamHI_Forward and FAR8_XhoI_Reverse, Jojoba_BamHI_FAR_Forward and Jojoba_XhoI_FAR_Reverse.

The site directed mutants and many of the domain swaps were made by overlap PCR, consisting of two separate PCR reactions detailed below for each construct. Two PCR programs were designed depending on the size of amplicon. For shorter (<1000 bp) fragments, the cycling conditions, denoted iProof PCR 1 Program, were: initial denaturation step at 98°C for 30 seconds, 30 cycles at 98°C for 15 seconds, 57°C for 20 seconds, and 72°C for 30 seconds, with a final single step of 72°C for 5 minutes. For longer (>1000 bp) fragments, the cycling conditions, denoted iProof PCR 2 Program, were: initial denaturation step at 98°C for 30 seconds, 30 cycles at 98°C for 15 seconds, 57°C for 30 seconds, and 72°C for 1 minute, with a final single step of 72°C for 5 minutes.

For the first PCR:

FAR5 H337Q was generated by first amplifying DNA products using pYES2-His6x/T7::FAR5 as template, using primers FAR5_BamHI_Forward and FAR5_H337Q_B in one reaction (FAR5 H337Q AB Fragment), and FAR5_H337Q_C and FAR5_XhoI_Reverse in another reaction (FAR5 H337Q CD Fragment), and using iProof PCR 1 Program.

FAR5 K242I was generated by first amplifying DNA products using pYES2-His6x/T7::FAR5 as template, using primers FAR5_BamHI_Forward and FAR5_K242I_B in one reaction (FAR5 K242I AB Fragment) and FAR5_K242I_C and FAR5_XhoI_Reverse in another reaction (FAR5 K242I CD Fragment), and using iProof PCR 1 Program.

FAR5 P363S was generated by first amplifying DNA products using pYES2-His6x/T7::FAR5 as template, using primers FAR5_BamHI_Forward and FAR5_P363S_B in one reaction (FAR5 P363S AB Fragment) and FAR5_P363S_C and FAR5_XhoI_Reverse in another reaction (FAR5 P363S CD Fragment), and using iProof PCR 1 Program.

FAR5 Y238F was generated by first amplifying DNA products using pYES2-His6x/T7::FAR5 as template, using primers FAR5_BamHI_Forward and FAR5_Y238F_B in one reaction (FAR5 Y238F AB Fragment) and FAR5_Y238F_C and FAR5_XhoI_Reverse in another reaction (FAR5 Y38F CD Fragment), and using iProof PCR 1 Program.

FAR8 Q337H was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8 as template, using primers FAR8_BamHI_Forward and FAR8_Q337H_B in one reaction (FAR8 Q337H AB Fragment) and FAR8_Q337H_C and FAR8_XhoI_Reverse in another reaction (FAR8 Q337H CD Fragment), and using iProof PCR 1 Program.

FAR8 S363P was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8 as template, using primers FAR8_BamHI_Forward and FAR8_S363P_B in one reaction (FAR8 S363P AB Fragment) and FAR8_S363P_C and FAR8_XhoI_Reverse in another reaction (FAR8 S363P CD Fragment), and using iProof PCR 1 Program.

FAR8 Q337H S363P was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8 S363P as template, using primers FAR8_BamHI_Forward and FAR8_Q337H_B in one reaction (FAR8 Q337H S363P AB Fragment) and

FAR8_Q337H_C and FAR8_XhoI_Reverse in another reaction (FAR8 Q337H S363P CD Fragment), and using iProof PCR 1 Program.

FAR8-*Sal*I is a mutant that has the restriction site for *Sal*I introduced (nucleotide exchange A→C at position 852 relative to the start codon). However, it is a silent mutation that retains the valine at the corresponding amino acid position (a.a. 285). There is a *Sal*I site present in FAR5, but it is not present in FAR8. Creating the *Sal*I site in the FAR8 open reading frame allowed for domain swaps to be easily done between the two ORFs using restriction enzymes. FAR8-*Sal*I was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8 as the template, using primers FAR8_BamHI_Forward and FAR8_SalI_Site_B in one reaction (FAR8-*Sal*I AB Fragment) and FAR8_SalI_Site_C and FAR8_XhoI_Reverse in another reaction (FAR8-*Sal*I CD Fragment), and using iProof PCR 1 Program.

FAR8-*Sal*I Q337H was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8-*Sal*I as the template, using primers FAR8_BamHI_Forward and FAR8_Q337H_B in one reaction (FAR8-*Sal*I Q337H AB Fragment) and FAR8_Q337H_C and FAR8_XhoI_Reverse in another reaction (FAR8-*Sal*I Q337H CD Fragment), and using iProof PCR 1 Program.

FAR8-*Sal*I S363P was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8-*Sal*I as the template, using primers FAR8_BamHI_Forward and FAR8_S363P_B in one reaction (FAR8-*Sal*I S363P AB Fragment) and FAR8_S363P_C and FAR8_XhoI_Reverse in another reaction (FAR8-*Sal*I S363P CD Fragment), and using iProof PCR 1 Program.

FAR8-*Sal*I Q337H S363P was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8-*Sal*I S363P as the template, using primers FAR8_BamHI_Forward and FAR8_Q337H_B in one reaction (FAR8-*Sal*I Q337H S363P AB Fragment) and FAR8_Q337H_C and FAR8_XhoI_Reverse in another reaction (FAR8-*Sal*I Q337H S363P CD Fragment), and using iProof PCR 1 Program.

FAR8-*Sal*I Q337H I347T S363P was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8-*Sal*I Q337H S363P as the template, using the primers FAR8_BamHI_Forward and FAR8_I347T_B in one reaction (FAR8-*Sal*I Q337H I347T S363P AB Fragment) and FAR8_I347T_C and FAR8_XhoI_Reverse in another reaction (FAR8-*Sal*I Q337H I347T S363P CD Fragment), and using iProof PCR 1 Program.

FAR8-*Sal*I Q337H S363P L383M was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8-*Sal*I Q337H S363P as the template, using primers FAR8_BamHI_Forward and FAR8_L383M_B in one reaction (FAR8-*Sal*I Q337H S363P L383M AB Fragment) and FAR8_L383M_C and FAR8_XhoI_Reverse in another reaction (FAR8-*Sal*I Q337H S363P L383M CD Fragment), and using iProof PCR 1 Program.

FAR8-*Sal*I Q337H S363P S410W was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8-*Sal*I Q337H S363P as the template, using primers FAR8_BamHI_Forward and FAR8_S410W_B in one reaction (FAR8-*Sal*I Q337H S363P S410W AB Fragment) and FAR8_S410W_C and FAR8_XhoI_Reverse in another reaction (FAR8-*Sal*I Q337H S363P S410W CD Fragment), and using iProof PCR 1 Program.

FAR8-*Sal*I Q337H Y348F S363P was generated by first amplifying DNA products using pYES2-His6x/T7::FAR8-*Sal*I Q337H S363P as the template, using primers FAR8_BamHI_Forward and FAR8_Y348F_B in one reaction (FAR8-*Sal*I Q337H Y348F S363P AB Fragment) and FAR8_Y348F_C and FAR8_XhoI_Reverse in another reaction (FAR8-*Sal*I Q337H Y348F S363P CD Fragment), and using iProof PCR 1 Program.

$FAR5_{1-387}$-$FAR8_{388-497}$ was generated by first amplifying the following two DNA products: (1) $FAR5_{1-387}$-$FAR8_{388-497}$ AB product was amplified using pYES2-His6x/T7::FAR5 as the template, using primers FAR5_BamHI_Forward and FAR5-FAR8_1166_B, and using iProof PCR 2 Program, and (2) $FAR5_{1-387}$-$FAR8_{388-497}$ CD product was amplified using pYES2-His6x/T7::FAR8-*Sal*I Q337H S363P as the template, using primers FAR5-FAR8_1166_C and FAR8_XhoI_Reverse, and using iProof PCR 1 Program.

FAR8-*Sal*I Q337H S363P $_{1-387}$-$FAR5_{388-497}$ was generated by first amplifying the following two DNA products: (1) FAR8-*Sal*I Q337H S363P $_{1-387}$-$FAR5_{388-497}$ AB product was amplified using pYES2-His6x/T7::FAR8-*Sal*I Q337H S363P as the template, using primers FAR8_BamHI_Forward and FAR8-FAR5_1166_B, and using iProof PCR 2 Program, and (2) and FAR8-*Sal*I Q337H $S363P_{1-387}$-$FAR5_{388-497}$ CD product was amplified using pYES2-His6x/T7::FAR5 as the template, using primers FAR8-FAR5_1166_C and FAR5_XhoI_Reverse, and using iProof PCR 1 program.

$FAR5_{1-442}$-$FAR8_{443-497}$ was generated by first amplifying the following two DNA products: (1) $FAR5_{1-442}$-$FAR8_{443-497}$ AB product was amplified using pYES2-His6x/T7::FAR5 as the template, using primers FAR5_BamHI_Forward and FAR5-

FAR8_1332_B, and using iProof PCR 2 Program, and (2) $FAR5_{1-442}$-$FAR8_{443-497}$ CD product was amplified using pYES2-His6x/T7::FAR8–*Sal*I Q337H S363P as the template, using primers FAR5-FAR8_1332_C and FAR8_XhoI_Reverse, and using iProof PCR 1 Program.

FAR8-*Sal*I Q337H S363P $_{1-442}$-$FAR5_{443-497}$ was generated by first amplifying the following two DNA products: (1) FAR8-*Sal*I Q337H S363P $_{1-442}$-$FAR5_{443-497}$ AB product was amplified using pYES2-His6x/T7::FAR8–*Sal*I Q337H S363P as the template, using primers FAR8_BamHI_Forward and FAR8-FAR5_1332_B, and using iProof PCR 2 Program, and (2) FAR8-*Sal*I Q337H $S363P_{1-442}$-$FAR5_{443-497}$ CD product was amplified using pYES2-His6x/T7::FAR5 as the template, using primers FAR8-FAR5_1332_C and FAR5_XhoI_Reverse, and using iProof PCR 1 program.

All of the PCR products described above were run on a 0.8% agarose gel (0.8% agarose, 40mM Tris Base, 20mM glacial acetic acid, 1mM EDTA pH 8.0) at 120V for 20 minutes. The bands were excised on a UV light box and inserted in a 1.5ml microcentrifuge tube for the DNA to be isolated by the "Freeze and Squeeze" method, which is described as follows. The gel band was placed in the barrel of a 3ml BD (Becton, Dickinson and Company) Luer-Lok syringe and crushed by quickly forcing the plunger down to push the band through the syringe tip into a 1.5ml microcentrifuge tube. An equal volume of TE-saturated phenol (1µl per mg of gel mass) was added to the crushed gel and the tube was vortexed upside down and flicked repeatedly to ensure the gel pieces were dispersed in the phenol. The microcentrifuge tube was then placed at -80°C for 5 minutes, proceeded by centrifugation at 21 100 g for 15 minutes at room temperature. The supernatant was transferred to a fresh 1.5ml microcentrifuge tube to be

extracted with an equal amount of phenol:chloroform:isoamyl alcohol (25:24:1), vortexed for 1 minute and centrifuged at 21 100 g for 5 minutes at room temperature. The supernatant was transferred to a fresh 1.5ml microcentrifuge tube and the DNA was precipitated by adding: mussel glycogen to a final concentration of 0.09µg/µl, $1/10^{th}$ the supernatant volume of 3M sodium acetate (pH 5.3), and 2.5 times the supernatant volume of 100% ethanol. The mixture was vortexed and placed on ice for 10 minutes. The samples were then centrifuged for 15 minutes at 21 100 g at room temperature. The supernatant was removed and the pellet was washed with 400 µl of cold 70% ethanol. The pellet was dried at 37°C for 10 minutes, and then re-suspended in 10 µl of nanopure water.

For the second PCR:

FAR5 H337Q, FAR5 K242I, FAR5 P363S and FAR5 Y238F were generated from their corresponding AB and CD fragments from the first PCR reactions using primers FAR5_BamHI_Forward and FAR5_XhoI_Reverse, and using iProof PCR 2 Program.

FAR8 Q337H, FAR8 S363P, FAR8 Q337H S363P, FAR8-*Sal*I, FAR8-*Sal*I Q337H, FAR8-*Sal*I S363P, FAR8-*Sal*I Q337H S363P, FAR8-*Sal*I Q337H I347T S363P, FAR8-*Sal*I Q337H S363P L383M, FAR8-*Sal*I Q337H S363P S410W, and FAR8-*Sal*I Q337H Y348F S363P were generated from their corresponding AB and CD fragments from the first PCR reactions using primers FAR8_BamHI_Forward and FAR8_XhoI_Reverse, and using iProof PCR 2 Program.

$FAR5_{1-387}$-$FAR8_{388-497}$ and $FAR5_{1-442}$-$FAR8_{443-497}$ were generated from their corresponding AB and CD fragments from the first PCR reactions using primers FAR5_BamHI_Forward and FAR8_XhoI_Reverse, and using iProof PCR 2 Program.

FAR8-*Sal*I Q337H S363P $_{1-387}$-$FAR5_{388-497}$ and FAR8-*Sal*I Q337H S363P$_{1-442}$-$FAR5_{443-497}$ were generated from their corresponding AB and CD fragments from the first PCR reactions using primers FAR8_BamHI_Forward and FAR5_XhoI_Reverse, and using iProof PCR 2 Program.

The products from the second PCR reactions were isolated by the "Freeze and Squeeze" method as described above.

FAR1, FAR3, FAR4, FAR5, FAR8, the FAR5 and FAR8 site directed mutants, and the overlap PCR domain swaps were digested with *Bam*HI and *Xho*I. FAR2 and FAR2ΔN were digested with *Bgl*II and *Sph*I. FAR6 and FAR6ΔN were digested with *Eco*RI and *Sph*I. Empty pYES2-His6x/T7 was also digested with corresponding pairs of restriction enzymes. All restriction digests were for approximately 2 hours at 37°C using 10 units of enzyme per μg of DNA. The digested DNA was then isolated by the "Freeze and Squeeze" method. The digested insert DNA molecules were ligated into digested pYES2-His6x/T7 in a reaction containing 1x ligase buffer (50mM Tris-HCl, 10mM $MgCl_2$, 1mM ATP, 10mM DTT pH 7.5), ~50ng of vector, ~35ng of insert and 1 Unit of T4 DNA Ligase) for 1 hour at room temperature.

There were also domain swaps created using restriction sites, rather than by overlap PCR. The restriction site domain swaps were created by swapping the 3' end of the open reading frame (corresponding to the C-terminal end of the protein), while the 5' end of the open reading frame (corresponding to the N-terminal end of the protein) was

still associated with the vector backbone. $FAR5_{1-141}$-$FAR8_{142-497}$ was made by ligating the FAR8 *Sac*I and *Xho*I fragment into the pYES2-His6x/T7::FAR5 vector with the *Sac*I and *Xho*I portion removed. $FAR8_{1-141}$-$FAR5_{142-497}$ was made by ligating the FAR5 *Sac*I and *Xho*I fragment into the pYES2-His6x/T7::FAR8 vector with the *Sac*I and *Xho*I portion removed. $FAR5_{1-284}$-$FAR8_{285-497}$ was made by ligating the FAR8 *Sal*I and *Xho*I fragment into the pYES2-His6x/T7::FAR5 vector with the *Sal*I and *Xho*I portion removed. $FAR8_{1-284}$-$FAR5_{285-497}$ was made by ligating the FAR5 *Sal*I and *Xho*I fragment into the pYES2-His6x/T7::FAR8 vector with the *Sal*I and *Xho*I portion removed. $FAR5_{1-284}$-FAR8 Q337H $_{285-497}$, $FAR5_{1-284}$-FAR8 $S363P_{285-497}$, and $FAR5_{1-284}$-FAR8 Q337H $S363P_{285-497}$ were made by ligating the FAR8-*Sal*I Q337H, FAR8—*Sal*I S363P and FAR-*Sal*I Q337H S363P *Sal*1 and *Xho*1 fragments into the pYES2-His6x/T7::FAR5 vector with the *Sal*I and *Xho*I portion removed. pYES2-His6x/T7::FAR5 and pYES2-His6x/T7::FAR8 were digested separately with *Sac*I and *Xho*I for 2 hours at 37°C using 1 unit of enzyme per µg of DNA. pYES2-His6x/T7::FAR5, pYES2-His6x/T7::FAR8, pYES2-His6x/T7::FAR8-*Sal*I Q337H, pYES2-His6x/T7::FAR8-*Sal*I S363P and pYES2-His6x/T7::FAR8-*Sal*I Q337H S363P were digested with *Sal*I and *Xho*I for 2 hours at 37°C using 1 unit of enzyme per µg of DNA. The relevant digested DNA molecules were isolated by the "Freeze and Squeeze" method described above. The digested insert DNA molecules were ligated into digested pYES2-His6x/T7 in a reaction containing 1x ligase buffer (50mM Tris-HCl, 10mM $MgCl_2$, 1mM ATP, 10mM DTT pH 7.5), ~50ng of vector, ~35ng of insert and 1 Unit of T4 DNA Ligase) for 1 hour at room temperature.

Table 2.1 Primers

Primer Name	Primer (5`→ 3`)
CYC_Terminator_Reverse	ACCTAGACTTCAGGTTGTCT
FAR1_BamHI_Forward	GAGGGATCCATGGAATCCAATTGTGTTCAAT
FAR1_XhoI_Reverse	GCGCTCGAGTTATTGTTTAAGCACATGGGTGA
FAR2_BglII_Forward	GTCAGATCTATGGAGGCTCTCTTCTTGAGT
FAR2N_BglII_Forward	GTCAGATCTATGGGACTTGGCATAATCAGTTTC
FAR2_SphI_Reverse	GTCGCATGCTTAAGCTCTTCCTTTCAAGACA
FAR3_BamHI Forward	GTCGGATCCATGTCGACAGAAATGGAGGTC
FAR3_XhoI Reverse	GTCCTCGAGTTAGAAGACATACTTAAGCAGC
FAR4_BamHI Forward	GAGGGATCCATGGACTCCAATTGCATTCAG
FAR4_XhoI Reverse	GCGCTCGAGTTATTTTTTGAGTACATAGGTGAT
FAR5_BamHI_Forward	GAGGGATCCATGGAACTCAATTGTGTTCAAT
FAR5_XhoII_Reverse	GCGCTCGAGTCACTTCTTAAGCACGTGTG
FAR5_H337Q_B	GATGAGCCGACTTGGTACACCAT
FAR5_H337Q_C	ATGGTGTACCAAGTCGGCTCATC
FAR5_K242I_B	CTCCCATTGATATGGTGAAAACA
FAR5_K242I_C	TGTTTTCACCATATCAATGGGAG
FAR5_P363S_B	ACTTCGCAAAGAGTTTTTCGTAA
FAR5_P363S_C	TTACGAAAAACTCTTTGCGAAGT
FAR5_Y238F_B	TGGTGAAAACAAATGTGTTAGGC
FAR5_Y238F_C	GCCTAACACATTTGTTTTCACCA
FAR5-FAR8_1166_B	TATGGTCATGTAGAGGCTGAACAAAGCCAT
FAR5-FAR8_1166_C	ATGGCTTTGTTCAGCCTCTACATGACCATA
FAR5-FAR8_1332_B	GTATTCCTATCGTCGAATATTCCCTTGAAG
FAR5-FAR8_1332_C	CTTCAAGGGAATATTCGACGATAGGAATAC
FAR6_EcoRI_Forward	GTCGAATTCATATGGCTACCACAAATGTCCTC
FAR6N_EcoRI_Forward	GTCGAATTCATATGAGTGACGGAATTGGAATCGTC
FAR6_SphII_Reverse	GTCGCATGCTTACTCAGTCTTCTTCTTAGAAAG
FAR6_qPCR_LP	TGTGGTGTCCCAGAGTTCAA
FAR6_qPCR_RP	TCCAATGGAAAGTCACACAGA
FAR8_BamHI_Forward	GAGGGATCCATGGAATTCAGTTGTGTTCA
FAR8_XhoI_Reverse	GCGCTCGAGTTACTTCTTAAGCACGTGAG
FAR8_I347T_B	GGATCTCTCCATATGTTATTGGGTTTTGGT
FAR8_I347T_C	ACCAAAACCCAATAACATATGGAGAGATCC
FAR8_L383M_B	GCTGAACAAAGCCATAGTTGGTATCAGCTT
FAR8_L383M_C	AAGCTGATACCAACTATGGCTTTGTTCAGC
FAR8_Q337H_B	GATGAACCAACATGGTACACCAT
FAR8_Q337H_C	ATGGTGTACCATGTTGGTTCATC
FAR8_S363P_B	ACTGCGCAACGGGTTTTTGGTGA
FAR8_S363P_C	TCACCAAAAACCCGTTGCGCAGT
FAR8_S410W_B	TCGTCTCCTTCCCTCCAAGGATATATTATA
FAR8_S410W_C	TATAATATATCCTTGGAGGGAAGGAGACGA
FAR8 SalI Site B	ATTACACTGTCGACAGTTCTTAG
FAR8 SalI Site C	CTAAGAACTGTCGACAGTGTAAT
FAR8_Y348F_B	TTTCACGGATCTCTCCAAATATTATTGGGT
FAR8_Y348F_C	AACCCAATAATATTTGGAGAGATCCGTGAA
FAR8-FAR5_1166_B	TAGGGTCATGTAGAGGCTGAACAAAGCCAG

FAR8-FAR5_1166_C	CTGGCTTTGTTCAGCCTCTACATGACCCTA
FAR8-FAR5_1332_B	GTATTTCTATCGTCGAATATGCCCTTGAAG
FAR8-FAR5_1332_C	CTTCAAGGGCATATTCGACGATAGAAATAC
GAPC_RT_Forward	TCAGACTCGAGAAAGCTGCTAC
GAPC_RT_Reverse	GATCAAGTCGACCACACGG
Jojoba_BamHI_FAR_Forward	GAGGGATCCATGGAGGAAATGGGAAGCAT
Jojoba_XhoI_FAR_Reverse	GCGCTCGAGTTAGTTAAGAACGTGCTCTA
T7_Forward	TAATACGACTCACTATAGGG
FAR6_Prom_forward	AGAGTCGACACGGAGGGACTTTTTCTGCT
FAR6_Prom_reverse	AGAGGATCCGAGGACATTTGTGGTAGCCAT

2.5 Plasmid Transformation in DH5α Escherichia coli Cells and Positive Clone Selection

Ligations described above were transformed into DH5α *Escherichia coli* (F– Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) U169 *recA1 endA1 hsdR17* (rK–, mK+) *phoA supE44* λ– *thi-1 gyrA96 relA1*). 5µl of the ligation reaction was added to 50µl of ultra-competent DH5α *E. coli* cells were thawed on ice, gently mixed by flicking, and incubated on ice for 30 minutes. After the incubation, the cells were heat shocked at 42°C for 40-45 seconds, then cooled on ice for 2 minutes. 1ml of Lysogeny Broth (LB) (10g/L tryptone, 5g/L yeast extract, 10g/L NaCl) was added to the cells and incubated at 37°C for 1 hour while rotating on a test tube rotator. The cells were then plated on LB plates (1.5% agar) containing 100µg/ml of ampicillin and incubated for approximately 16 hours.

Colony PCR was performed to identify *E. coli* colonies containing the correct ligation product. Selected colonies were picked with a sterile pipette tip from the LB plate containing 100µg/ml of ampicillin and then inserted into 20 µl of PCR mix (10mM KCl, 10mM $(NH_4)_2SO_4$, 20mM Tris-HCl pH 8.3, 0.1% Triton X-100, 2mM $MgCl_2$, 0.25mM dNTPs, 0.25mM T7 Forward Primer, 0.25mM CYC Terminator Reverse Primer, 0.1µl *Taq* Polymerase). The PCR was conducted using T7_Forward primer (5'-TAATACGACTCACTATAGGG-3') and CYC_Terminator_Reverse primer (5'-ACCTAGACTTCAGGTTGTCT-3'). The PCR products were run on a 0.8% agarose gel and the correct products had bands 1642-2101bp in length. Empty vector produces a band of 314 bp using T7 Forward and CYC Terminator Reverse primers.

The positive colonies were inoculated into 3ml of LB containing 100µg/ml of Ampicillin. A mini-preparation of the plasmids from *E. coli* was then performed. 1.5ml

of the culture was pipetted into a microcentrifuge tube and centrifuged at 21 100 g for 30 seconds at 4°C. The supernatant was poured off and the pellet was resuspended in 800µl of STE (0.1M NaCl, 0.01M Tris-HCl pH 8.0, 0.001M EDTA) and centrifuged at 14 800 rpm for 1 minute. The supernatant was removed and resuspended in resuspension buffer (0.5M Tris-HCl pH 8.0, 0.01M EDTA, 1µg/ml RNase A) and vortexed to mix. Lysis buffer (0.2M NaOH, 1% SDS) was then added, inverted 5 times, and incubated at room temperature for 5 minutes. Neutralization buffer (3.0M Potassium Acetate pH 5.5) was then added and inverted 5 times. The samples were then centrifuged at 21 100 g for 10 minutes at 4°C. The supernatant was poured into a fresh microcentrifuge tube. 800µl of isopropanol was added and incubated for 5 minutes at room temperature, centrifuged at 21 100 g for 15 minutes at room temperature. The supernatant was removed and the pellet was rinsed with 500µl cold 70% ethanol, dried at 37°C for 15 minutes, and then resuspended in 100µl of water.

A diagnostic digest was performed on the potentially correct clones using the restriction enzymes used to insert the gene of interest into the vector, as well as unique restriction sites present within the gene. For sequencing, the inserts were amplified by PCR in a 50µl reaction (10mM KCl, 10mM $(NH_4)_2SO_4$, 20mM Tris-HCl pH 8.3, 0.1% Triton X-100, 2mM $MgCl_2$, 0.25mM dNTPs, 0.25mM T7 Forward Primer, 0.25mM CYC Terminator Reverse Primer, 0.25µl *Taq* Polymerase and 10ng of Template) using T7_Forward primer (5'-TAATACGACTCACTATAGGG-3') and CYC_Terminator_Reverse primer (5'-ACCTAGACTTCAGGTTGTCT-3'). Sequencing was performed at Eurofins MWG|Operon (Hunstville, Alabama) using T7 and CYC primers, as well as gene specific primers.

2.6 Yeast Transformation

The yeast expression constructs were transformed into yeast using the lithium acetate (LiAc)/single strand (SS) carrier DNA/polyethylene glycol (PEG) method (Gietz and Woods, 2002). 5ml of YPAD media (1% yeast extract, 2% peptone, 2% D-glucose, 0.01% adenine hemisulphate, 1.5% agar) was inoculated with the yeast strain W303-1A (*MATa his3Δ1 leu2 trp1-289 ura3-52*) and incubated at 30°C and 250rpm for 16 hours. The yeast cells were harvested the following day by centrifugation at 4000 rpm for 5 minutes at room temperature. The pellet was resuspended in 240µl of sterile 50% (w/v) PEG 3500, and then transferred to a sterile 1.5ml microcentrifuge tube. 36µl sterile 1M LiAc, 50µl boiled SS carrier DNA (2.0mg/ml), 5µl plasmid DNA (100ng-1µg), and 29µl of sterile H_2O were then added to the yeast suspension. The mixture was mixed by flicking until resuspended, and then incubated in a 42°C water bath for 1 hour while being mixed by flicking periodically during the incubation period to prevent separation of the constituents. After the incubation period, the tubes were centrifuged at 21 100 g for 30 seconds at room temperature, resuspended in 200µl of sterile water, and 100µl was plated onto SD Ura$^-$ plates (0.674% yeast nitrogen base, 2% D-glucose, 0.01% adenine, 0.002% histidine, 0.002% tryptophan, 0.01% leucine, 2% agar) to select for transformants. The plates were incubated at 30°C for 3 days.

2.7 SDS-PAGE and Western Blot

Individual transformants for FAR5 and FAR8 (including mutants and domain swaps) were selected on the SD Ura$^-$ plates and inoculated in triplicate cultures in 20ml SD Ura$^-$ media (0.674% yeast nitrogen base, 2% D-glucose, 0.01% adenine, 0.002% histidine, 0.002% tryptophan, 0.01% leucine) and grown for 2 days at 30°C and 250rpm.

The OD_{600} was measured for each culture and the amount was calculated in order to get an OD_{600} of 0.4 to make 100ml of SG Ura^- induction media (0.674% yeast nitrogen base, 2% galactose, 0.01% adenine, 0.002% histidine, 0.002% tryptophan, 0.01% leucine). This calculated amount was centrifuged at 2880 g for 10 minutes at room temperature, the pellets washed with 5ml of water, vortexed, and then centrifuged again at 2880 g at room temperature. The pellet was resuspended in 100 ml of SG Ura- media and induced for 0, 6, 9, 12, 24, and/or 48 hours at 30°C and 250 rpm for protein expression. The yeast was harvested to have a pellet with 15 OD_{600} units of cells by centrifuging at 2880 g at room temperature for 10 minutes. The pellet was resuspended in 1ml of water, transferred to a 1.5ml microcentrifuge tube, centrifuged at 21 100 g for 1 minute, and then stored at -80°C. It was found that 6 hours of induction was sufficient to readily see induced protein by Western blotting, and therefore all subsequent inductions were carried out for 6 hours for sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot analysis.

The yeast pellets were prepared for SDS-PAGE and Western blot analysis using the trichloroacetic acid (TCA) method. The yeast pellets were thawed on ice for 10 minutes, resuspended in 200µl of ice-cold TCA buffer (20mM Tris-HCl pH 8.0, 50mM ammonium acetate, 2mM EDTA, 50µl/ml protease inhibitor solution, 1µM PMSF), and then placed on ice. The cell suspension was transferred to a 2ml microcentrifuge tube containing 200µl of 425-600µm acid-washed glass beads and 200µl of ice-cold 20% TCA. The cells were vortexed for 1 minute, incubated on ice for 30 seconds, and this was repeated 4 times. The tubes were placed on ice and the supernatants were transferred to fresh 1.5ml microcentrifuge tubes. The glass beads were then washed with 500 µl of ice

cold 1:1 mixture of 20% TCA and TCA buffer, vortexed for 1 minute, and then incubated on ice for 30 seconds. This was repeated 2 times. The supernatant was transferred to the tube containing the first cell extract. The proteins were pelleted by centrifugation at 21 100 g for 10 minutes at 4°C. The supernatant was removed and discarded, and the pellet was resuspended in 150µl of Laemmli loading buffer (120mM Tris-Base, 8mM EDTA, 14% glycerol v/v, 3.5% SDS w/v, spatula tip of bromophenol blue) with freshly added β-mercaptoethanol (5% final concentration), 0.002mM PMSF, and 20µl complete mini roche protease inhibitor cocktail solution (0.03µg chymotrypsin, 0.016µg thermolysin, 0.02µg papain, 0.03µg pronase, 0.3µg pancreatic extract, 0.004µg trypsin). The samples were boiled for 10 minutes and centrifuged at 21 100 g for 10 minutes at room temperature and were ready for loading onto the gel.

The samples were run on SDS-PAGE gels in duplicates with a 12% resolving gel (12% acrylamide:bis-acrylamide 37.5:1, 0.375M Tris-HCl pH 8.8, 0.1% SDS, 0.05% APS, 0.005% TEMED) and 4% stacking gel (4% acrylamide:bis-acrylamide 37.5:1, 0.1M Tris-HCl pH 6.8, 0.1% SDS, 0.05% APS, 0.005% TEMED). 20µl of each sample was run on the gel for 30 minutes at 100V and 1.5 hours at 150V in 1X running buffer (25 mM Tris, 186 mM glycine, 0.1% SDS, pH8.3).

One of the gels was stained with Coomassie blue stain (0.25% Coomassie Brilliant Blue R-250, 10% glacial acetic acid, 50% methanol) for 1 hour while shaking on a nutator, then de-stained overnight in de-staining solution (25% methanol, 7% glacial acetic acid). The other gel was transferred to 0.45µm nitrocellulose membrane in chilled transfer buffer (25mM Tris Base, 192mM Glycine, 20% methanol) for 1 hour at 350mA at 4°C. The nitrocellulose membrane was then blocked overnight in blocking solution

(5% fat free skim milk, TBST pH 7.6 (25mM Tris Base, 137mM NaCl, 0.1% Tween-20)) at 4°C. The following day the membrane was incubated with a 1:50 000 dilution of T7 Tag Monoclonal Mouse Antibody (Novagen) in 25ml blocking solution for 1 hour while shaking on the Fisher Scientific Clinical Rotator (Model # 2314FS). The membrane was washed with TBST for 1x 15minutes and then 3 x 5minutes, replacing with fresh TBST each time. The membrane was then incubated with 1:50 000 dilution of horseradish peroxidase (HRP)-conjugated Anti-Mouse Secondary antibody (Novagen) in 25ml blocking solution for 1 hour while shaking on a Fisher Scientific Clinical Rotator (Model # 2314FS). The membrane was then washed again with TBST for 1x 15minutes and then 3 x 5minutes. The membrane was covered with 1:1 Lumigen TMA-6 Solution A (Solution containing Tris Buffer in 3.2% v/v ethanol) and Lumigen TMA-6 Solution B (Proprietary substrate in Tris Buffer), which are part of the Amersham™ ECL™ Advance Western Blotting Detection Kit (GE Healthcare). Immunodetection was performed on the FluorChem Q Mulitimage III imaging system (Alpha Innotech) using the Chemi-Super Protocol (Excitation = none, Emission = Chemi, Speed/Resolution = Normal/Ultra) and the auto-exposure time.

2.8 Gas Chromatography

Individual transformants for FAR5 and FAR8 (including mutants and domain swaps) were selected on the SD Ura$^-$ plates and were inoculated in triplicate cultures in 20ml SD Ura$^-$ media (0.674% yeast nitrogen base, 2% D-glucose, 0.01% adenine, 0.002% histidine, 0.002% tryptophan, 0.01% leucine) and grown for 2 days at 30°C and 250rpm. The OD_{600} was measured for each culture and the amount was calculated in

order to get an OD_{600} of 0.4 to induce 20ml of SG Ura$^-$ media (0.674% yeast nitrogen base, 2% galactose, 0.01% adenine, 0.002% histidine, 0.002% tryptophan, 0.01% leucine). This amount was centrifuged at 2880 g for 10 minutes at room temperature, the pellets were washed with 5ml of water, vortexed, and then centrifuged again at 2880 g at room temperature. The pellet was resuspended in 20ml of SG Ura$^-$ media and induced for 1-6 days at 30°C and 250rpm for protein expression.

Lipids were extracted from these yeast cultures as described below. It was found that 4 days of induction was best, since small amounts of C16:0-OH by FAR8 could be detected, and therefore all yeast cultures were induced for 4 days prior to lipid analysis by gas chromatography. 2ml of each yeast culture was transferred to a glass 13mm x 100mm tube and centrifuged at 4500 g for 10 minutes. The supernatant was poured into a fresh glass tube. The supernatant was extracted once with 1ml 2:1 chloroform/methanol (v/v) containing 10µg 1-tridecanol (C13:0-OH) as an internal standard, once more with 1ml 2:1 chloroform/methanol (v/v), and once with 1ml chloroform. All of the organic phases were combined, washed with 2.5ml 0.9% NaCl (w/v), and evaporated to dryness under a gentle stream of nitrogen gas at 37°C for 15-20 minutes. The supernatant extractions and the yeast pellets from the centrifugation step were resuspended in 3ml of 1M methanolic-HCl and 3ml of 1M methanolic-HCl containing 10µg 1-tridecanol (C13:0-OH), respectively, by vortexing and incubated at 80°C for 90 minutes. 1ml of 0.9% NaCl was added to the samples after they were cooled to room temperature and then the lipids were extracted with 500µl of hexane. The top organic phase was transferred to a new glass tube. The liquid-liquid extraction was repeated with an additional 500µl of hexane and the top organic phase was transferred to the first extraction. These were evaporated to

dryness under nitrogen gas and heated at 37°C for 15-20 minutes. The lipids were derivatized with 50µl of *N,O*-bis (trimethylsilyl) trifluoroacetamide (BSTFA) at 80°C for 90 minutes. 1µl of sample was injected into a Varian 3900 Gas Chromatograph system equipped with a Varian CP-8400 autosampler, an HP-1 column (15m length x 0.25mm internal diameter x 0.25µm film thickness), and a flame ionization detector. The initial temperature of 150°C was held for 5 minutes, increased at 10°C/minute to 300°C, and then held at 300°C for 5 minutes for a total of 28 minutes. Quantification of fatty alcohols was based on peak areas of 1-tridecanol (C13:0-OH) as the internal standard. Helium was used as the carrier gas (2ml/min) and it was run with a split-ratio of 50:1.

2.9 Promoter::GUS Fusions and GUS Histochemical Assay

Previously, a pBI101 (Clontech) derived plant binary vector was generated containing a T-DNA harboring the promoter of the *FAR6* gene (2187 bp region upstream of the start codon and the first seven codons of the *FAR6* coding region) fused in frame with the coding region of the β–glucuronidase (GUS) reporter gene creating a translational fusion (constructed by Adel Al-Shammari, Rowland Lab, Carleton University).

This construct was then introduced into wild type *Arabidopsis thaliana* Col-0 plants by Agrobacterium-mediated transformation using a modified version of the floral dip method (Clough and Bent, 1998). 1 µl (2-10 ng) of plasmid DNA was added to 50 µl of *Agrobacterium tumefaciens* cells that were thawed on ice and 40 µl were added to a pre-cooled electroporation cuvette. An electric pulse was applied to the cells (field strength = 2.5 kV, capacitance = 25 µF, resistance = 400-600 ohms, pulse length = 8-12

ms). 1ml of LB was added immediately to the cells in the cuvette, chilled on ice for 2 minutes, transferred to a 1.5ml microcentrifuge tube and incubated at 29°C for 3 hours. After the 3 hour incubation, 10 µl and 100 µl aliquots of the cells was put onto LB selection plates (Rif^{100} $Gent^{25}$ Kan^{50}) and incubated for 3 days at 29°C. 2 ml of LB (Rif^{100} $Gent^{25}$ Kan^{50}) was inoculated with a single Agrobacterium colony and incubated overnight at 27°C with moderate rotation (200 rpm). 50 ml of LB (Rif^{100} $Gent^{25}$ Kan^{50}) was then inoculated with 10 µl of the overnight culture and incubated overnight at 27°C with moderate rotation (200 rpm). When the OD_{600} reading of the culture reached 0.2-0.8 it was spun down in a 250ml centrifuge bottle at 7438 g for 10 minutes. The supernatant was discarded and the pellet was resuspended in 500ml of 5% sucrose containing 0.05% Silwet L-77. *Arabidopsis thaliana* ecotype Columbia-0 (Col-0) plants with unopened flowers were dipped into the Agrobacterium sucrose solution for 2-3 seconds. The plants were laid on their sides, placed in a plant tray covered with a dome, to retain humidity, and kept in the dark for 12-24 hours. Plants were then grown in a chamber under continuous long-day conditions (16 hour light/8 hours of dark). After the plants matured and the seeds were harvested the transformants were selected on AT-agar plates containing kanamycin (50µg/ml). Various plant tissues of the third generation lines (T3) were analyzed for β-glucuronidase (GUS) activity at different stages of the plant's life cycle. Plant tissues were collected in 24-well multiwell culture plates, submerged in cold heptane, incubated on ice for 5 minutes, and then allowed to air dry for 5 minutes. The tissues were rinsed with staining buffer (50mM $NaPO_4$ pH 7.0, 0.5mM $K_4Fe(CN)_6$, 0.5mM $K_3Fe(CN)_6$, and 0.1% TritonX-100) and then covered in staining solution containing 1.12 mM 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-Gluc). The samples

were covered with aluminum foil and incubated at 37°C. The time of incubation varied on the tissue to give the optimal amount of blue 'staining': root tissues were incubated for 1-2 hours, stem tissues were incubated for 14-17 hours, and aerial tissues were incubated for 5-7 hours. The stained tissues were imaged using a Zeiss Discovery V20 stereomicroscope equipped with an Axiocam MRC camera.

For stem sections, the stained stems were collected in glass scintillation vials and fixed in FAA fixative (50% Ethanol, 5% Glacial Acetic Acid, 3.7% Formaldehyde) for 4 hours at room temperature. After fixation, the plant tissues were rinsed 3 times with 50% ethanol and then put through a dehydration series (50% x 2, 60%, 70%, 85%) changed at 30 minute intervals, and left in 95% ethanol containing 0.1% Eosin stain overnight. The plant tissues were transferred to new glass scintillation vials containing 100% ethanol and incubated for 1 hour, and then twice it was changed for fresh 100% ethanol and incubated for 30 minutes each. The plant tissues were treated with increasing xylene concentrations, and decreasing ethanol concentrations, which were changed at 30 minute intervals (25% xylene: 75% ethanol, 50% xylene: 50% ethanol, 75% xylene: 25% ethanol, and 100% xylene twice). The plant tissues were infiltrated with paraffin by adding 20 Paraplast® Plus paraffin wax chips (Sigma-Aldrich) to each scintillation vial and left overnight. The wax chips were melted in the scintillation vial during a 1 hour incubation period in a 42°C water bath, transferred to a 60°C oven for 4 hours then the vial solution was changed to 100% wax (pre-melted wax chips) and left overnight at 60°C. The 100% wax in the vial was changed twice a day for three days at least 6 hours apart, always being kept in a 60°C oven. The wax moulds were prepared by emptying the vial contents into a Petri plate and the plant tissues were positioned. The Petri dish was stored at 4°C to allow

the wax to solidify for sectioning. Transverse sections were prepared in 14μm thick slices with a rotary microtome (Leica Rotary Microtome, model HM325). The sections were imaged using a Zeiss Discovery V20 stereomicroscope equipped with an Axiocam MRC camera.

CHAPTER 3: Substrate Specificities of Arabidopsis FARs and Jojoba FAR

Publication that includes data from this chapter:

The *FAR6* promoter::GUS gene expression analysis (Figure 3.5) was published in:

Thuy T.P. Doan, Frédéric Domergue, **Ashley E. Fournier**, Sollapura J. Vishwanath, Owen Rowland, Patrick Moreau, Craig C. Wood, Anders S. Carlsson, Mats Hamberg, and Per Hofvander (2011) Biochemical characterization of a chloroplast localized fatty acid reductase from *Arabidopsis thaliana*. Biochimica et Biophysica Acta, in press, http://dx.doi.org/10.1016/j.bbalip.2011.10.019

Statement of contribution:

I performed all the experiments and generated all the materials reported in this data chapter with the exception of the following:

The cross section of an anther shown in Panel E of Figure 3.4 was done by Sollapura J. Vishwanath (Rowland Lab, Carleton University).

FAR5 was previously cloned into pYES2-His6x/T7 by Frances Tran (Rowland Lab, Carleton University).

FAR1 and *FAR8* were previously cloned into pYES2-His6x/T7 by Sollapura J. Vishwanath (Rowland Lab, Carleton University).

FAR6 was previously cloned into pBI101 by Adel Al-Shammari (Rowland Lab, Carleton University).

3.1 Introduction

Arabidopsis has a relatively small genome made up of 5 chromosomes (~130Mb) (Kapitonov and Jurka, 1999). The Arabidopsis *FAR* genes are distributed on chromosomes III, IV, and V (Figure 3.1). *FAR2*, *FAR4*, *FAR5*, and *FAR8* are located on chromosome III, with the *FAR4*, *FAR5*, and *FAR8* genes arranged directly in tandem. *FAR3* is a singleton located on chromosome IV. *FAR1* and *FAR7* are in close proximity to each other on chromosome V.

The predicted Arabidopsis FAR proteins share 32%-85% amino acid identity between each other (Table 3.1). A phylogram of Arabidopsis FAR1-FAR8 and the seed-expressed Jojoba FAR reveal distinct clades (Figure 3.2). The two plastid-localized FARs, FAR2 and FAR6, form one clade and have 41% amino acid identity. Jojoba FAR and FAR3 form another distinct clade and have 54% amino acid identity. The FARs associated with suberin, FAR1, FAR4, and FAR5, have diverged from a common ancestor along with FAR8. FAR1 and FAR4, and FAR5 and FAR8 belong to separate subclades and have 74% and 85% amino acid identity, respectively (Table 3.1). FAR7 forms a clade of its own, but it is not an active FAR and is thought to derive from a pseudogene (Doan *et al.*, 2009).

The primary amino acid sequence alignment of Arabidopsis FAR1-6, FAR8 and Jojoba FAR reveals a high degree of amino acid sequence identity over a ~500 amino acid region (Figure 3.3; Table 3.1). These FARs range in length from 491 to 616 amino acids. FAR7 was not included in the line-up since the only cloned cDNA sequence contains an early stop codon (Doan *et al.*, 2009). A conserved NAD(P)H-binding motif

[I/V/F]-X-[I/L/V]-T-G-X-T-G-F-L-[G/A] was previously noted in the predicted FAR2/MS2 protein and is found at the N-terminus of the Nicotinamide Adenine Dinucleotide Binding (NABD) Rossmann fold superfamily domain (Aarts *et al.*, 1997). This NAD(P)H binding motif is found in the corresponding N-terminal region of all Arabidopsis FARs and the Jojoba FAR (Figure 3.3, and also see Figure 1.3 of the General Introduction). FARs also possess a strictly conserved predicted active site motif containing a tyrosine and lysine (YXXXK) found in other enzymes of the short-chain alcohol dehydrogenase/reductase (SDR) superfamily (Figure 3.3) (Jörnvall *et al.*, 1981; Ghosh *et al.*, 1995; Denessiouk *et al.*, 2001; Kavanagh *et al.*, 2008). This active site motif is located within the NABD Rossmann fold superfamily domain (see Figure 1.3 of the General Introduction).

FAR2 (616 a.a.) and FAR6 (548 a.a.) are noticeably longer than the other Arabidopsis FARs (~491-496a.a.), having N-terminal extensions of 119 a.a. and 71 a.a., respectively, compared to the other six FAR enzymes (Figures 3.3 and 3.4). These extensions are predicted to contain chloroplast targeting sequences according to the ChloroP 1.1 Server (http://www.cbs.dtu.dk/services/ChloroP; Emanuelsson *et al.*, 1999). FAR2 and FAR6 are predicted to have chloroplast target peptides of 12 and 47 amino acids in length, respectively, and with CS-scores of 3.897 and 2.286, respectively (Table 3.2).

The eight-membered Arabidopsis FAR family and Jojoba FAR are well-characterized relative to other plant FARs (see General Introduction). The Arabidopsis FARs are reported to have chain-length specificities that range from C16:0-C30:0. However, these specificities have been reported using various methodologies and

sometimes the substrate specificities reported are conflicting (see Table 1.2 of General Introduction). Arabidopsis FARs have been expressed in *E. coli* (Doan *et al.*, 2009) and *S. cerevisiae* (Rowland *et al.*, 2006; Domergue *et al.*, 2010; Doan *et al.*, 2011) and the composition of fatty alcohols produced from endogenous acyl pools used to infer substrate specificity. Specificities have also been inferred from analysis of mutant Arabidopsis plants (Rowland *et al.*, 2006; Domergue *et al.*, 2010). *In vitro* assays have been done using FAR2 and FAR6 since they can be purified in soluble form (Chen *et al.*, 2011; Doan *et al.*, 2011), but *in vitro* assays have proven difficult to develop with the membrane-associated FARs.

I chose to more directly compare the substrate specificities of Arabidopsis FAR1-6 and 8 as well as Jojoba FAR using yeast as a heterologous host. Tagged versions were used to monitor protein levels, which had not been done previously. Also, the gene expression pattern of *FAR6* from Arabidopsis was characterized in detail using a promoter::GUS reporter gene, which had not been done previously.

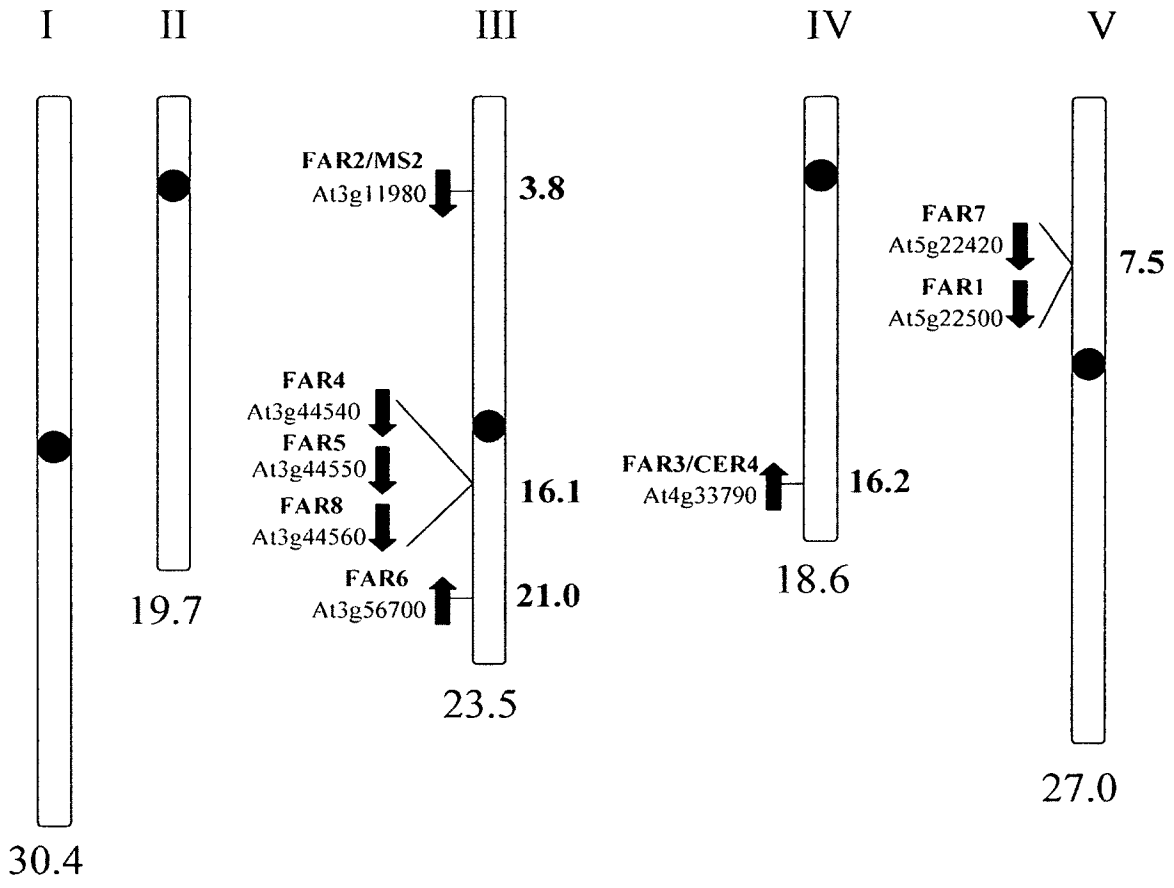


Figure 3.1 The Arabidopsis *FAR* gene family on chromosomes I-V. The 8 positions of the Arabidopsis *FAR* genes are shown on chromosomes I, II, III, IV and V. The length of the chromosome is indicated below in mega-basepairs. The black circles represent centromeres. Sequences are from the TAIR website (www.arabidopsis.org). FAR1: AT5G22500, FAR2: AT3G11980, FAR3: AT4G33790, FAR4: AT3G44540, FAR5: AT3G44550, FAR6: AT3G56700, FAR7: AT5G2420, FAR8: AT3G44560.

Table 3.1 Summary Table of Amino Acid Sequence Identity between the Arabidopsis FAR family and Jojoba FAR

% Identity	FAR1	FAR2	FAR3	FAR4	FAR5	FAR6	FAR7	FAR8	Jojoba FAR
FAR1	-	39	51	74	68	34	61	64	51
FAR2	39	-	39	38	39	41	38	38	39
FAR3	51	39	-	50	50	33	47	49	54
FAR4	74	38	50	-	69	35	61	65	51
FAR5	68	39	50	69	-	35	62	**85**	51
FAR6	34	41	33	35	35	-	32	35	32
FAR7	61	38	47	61	62	32	-	60	47
FAR8	64	38	49	65	**85**	35	60	-	49
Jojoba FAR	51	39	54	51	51	32	47	49	-

* Based on the full length CDSs from the TAIR website for FAR1 (At5g22500), FAR2 (At3g11980), FAR3 (At4g33790), FAR4 (At3g44540), FAR5 (At3g44550), FAR6 (At3g56700), FAR7 (At5g22420), FAR8 (At3g44560) and the protein sequence for the Jojoba FAR (AAD38039) obtained from NCBI (Carnegie Institution of Washington Department of Plant Biology, 2008; U.S. National Library of Medicine. (2009); European Bioinformatics Institute, 2010).

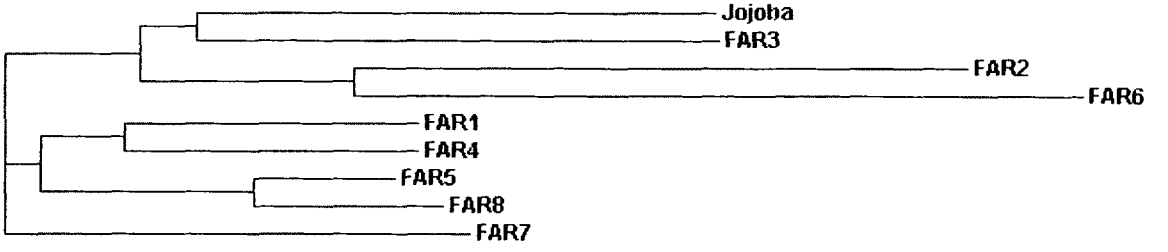
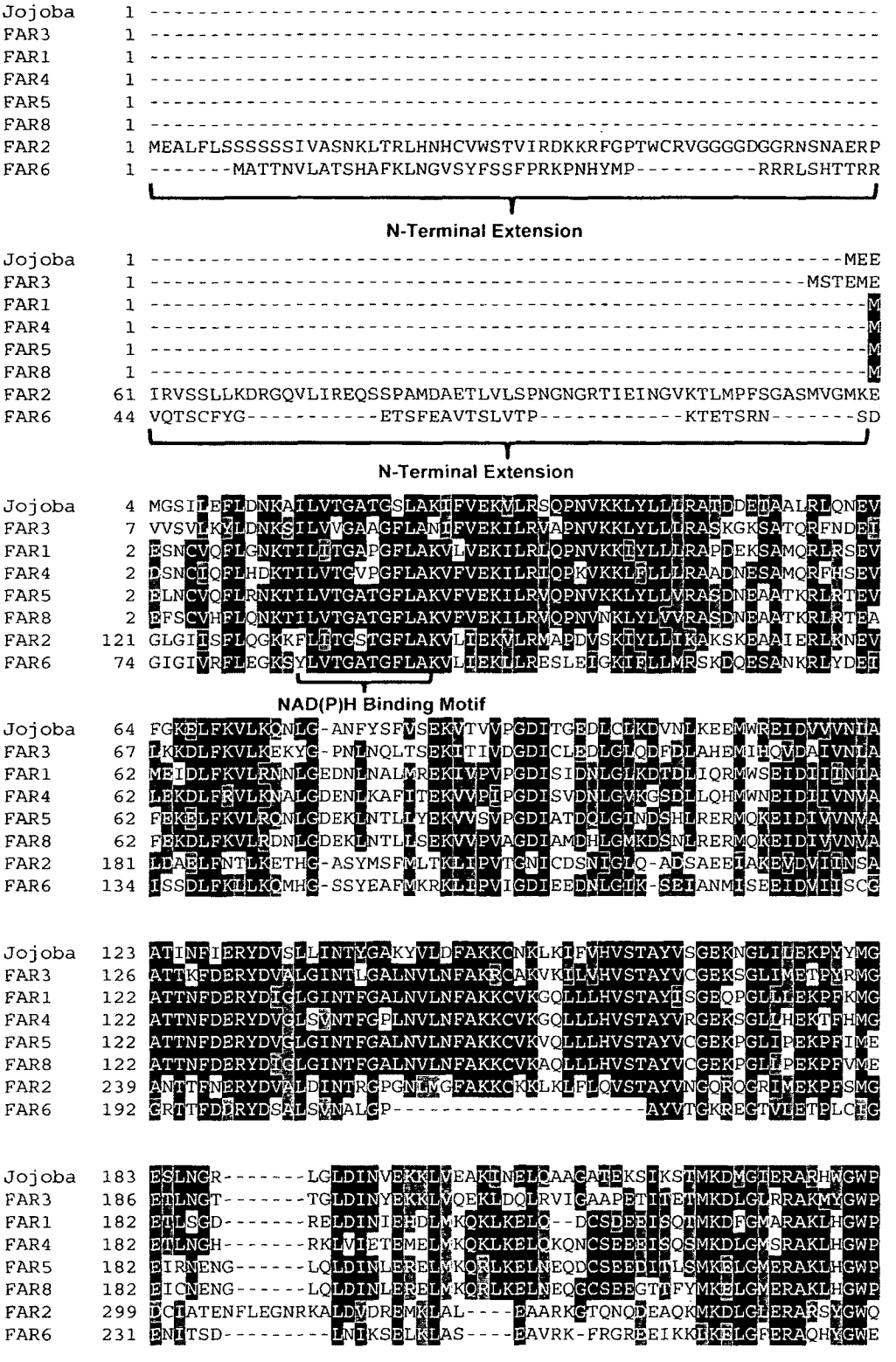


Figure 3.2. Phylogram of the Arabidopsis FAR protein family and seed-expressed Jojoba FAR. The tree was constructed using CLUSTALW2 and NJ plotWIN95. The branch lengths are proportional to the amount of inferred evolutionary change. The predicted Arabidopsis FAR protein sequences were obtained from the TAIR website: FAR1 (At5g22500), FAR2 (At3g11980), FAR3 (At4g33790), FAR4 (At3g44540), FAR5 (At3g44550), FAR6 (At3g56700), FAR7 (At5g22420), and FAR8 (At3g44560). The GenBank accession number of Jojoba *(Simmondsia chinensis)* FAR is AF149917 (obtained from National Center for Biotechnology Information, NCBI).


N-Terminal Extension
N-Terminal Extension
NAD(P)H Binding Motif

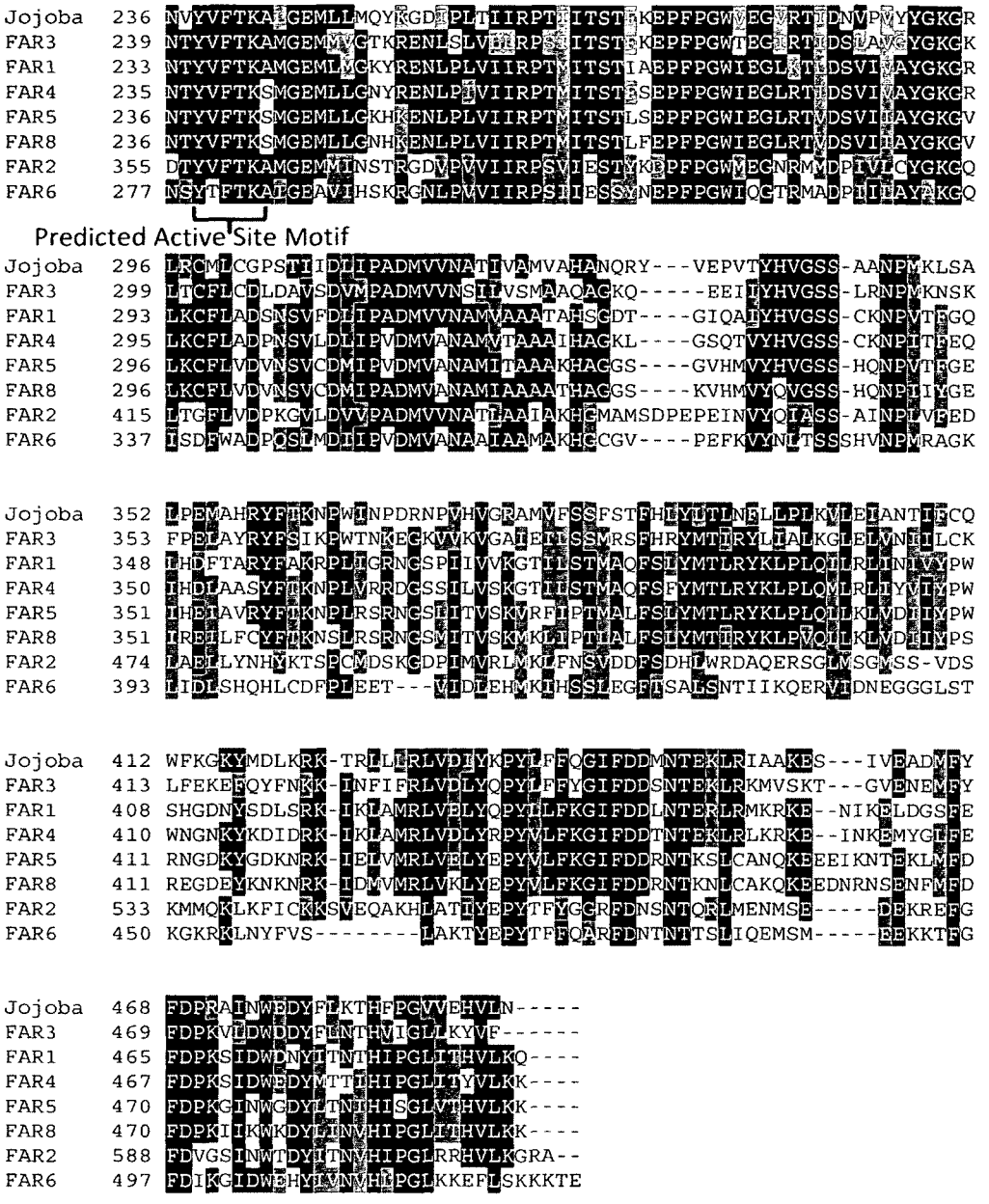


Figure 3.3 Alignment of Arabidopsis FAR and Jojoba FAR protein sequences. Identical residues = black, physicochemical similar residues = grey. The alignment was constructed using ClustalW2 and BOXSHADE 3.21Server. Full length CDSs were obtained from the TAIR website for FAR1 (At5g22500), FAR2 (At3g11980), FAR3 (At4g33790), FAR4 (At3g44540), FAR5 (At3g44550), FAR6 (At3g56700), and FAR8 (At3g44560). The protein sequence for the seed-expressed Jojoba FAR (AAD38039) was obtained from NCBI.

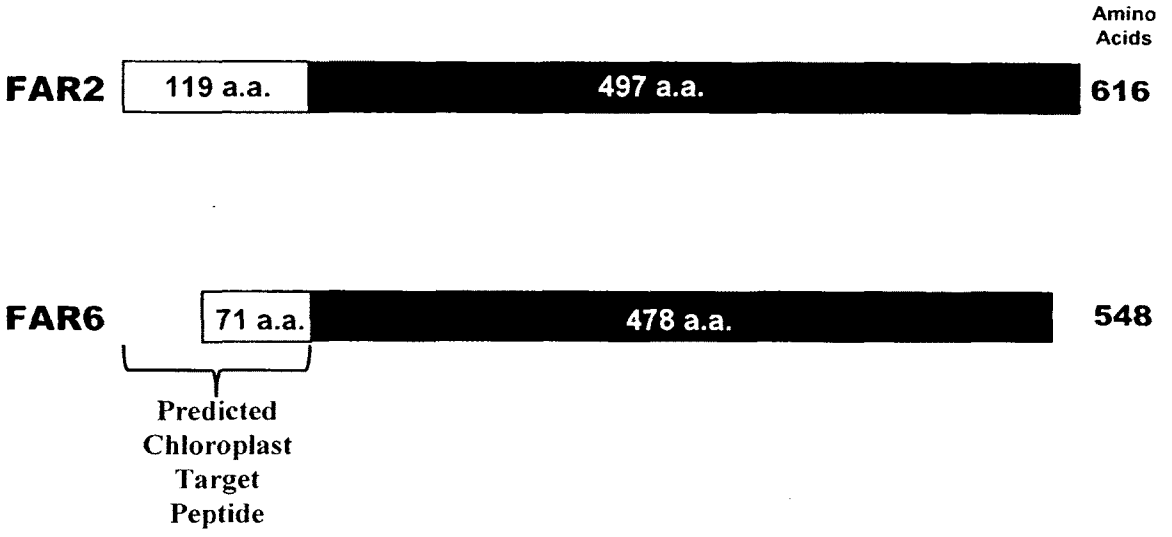


Figure 3.4 Schematics of the FAR2 and FAR6 proteins. FAR2 and FAR6 contain N-terminal extensions of 119 a.a. and 71 a.a., respectively, relative to other Arabidopsis FARs and the seed-expressed Jojoba FAR.

Table 3.2 Chloroplast target peptide predictions for the Arabidopsis FAR family and Jojoba FAR using ChloroP 1.1 Server

Name	Length (a.a.)	Score	Chloroplast Target Peptide	CS-score	Chloroplast Target Peptide Length (a.a.)
Jojoba FAR	493	0.434	NO	1.242	48
FAR1	491	0.443	NO	-0.777	23
FAR2/MS2	**616**	**0.502**	**YES**	**3.897**	**14**
FAR3/CER4	493	0.439	NO	10.329	51
FAR4	493	0.445	NO	4.614	46
FAR5	496	0.454	NO	6.852	46
FAR6	**548**	**0.569**	**YES**	**2.286**	**47**
FAR7	409	0.442	NO	-0.529	15
FAR8	496	0.465	NO	6.852	46

***Score** is the output score from the second step network. The prediction cTP/no cTP is based solely on this score.

***CS-score** is the MEME scoring matrix score for the suggested cleavage site.

***cTP-length** is the predicted length of the pre-sequence

*3.2 **Materials and Methods:** See Chapter 2*

Results

3.3 Arabidopsis FAR6 Gene Expression Pattern

The Arabidopsis *FAR* family members have diverse gene expression patterns. *FAR1*, *FAR4*, and *FAR5* are expressed at sites of suberin deposition (root endodermal cells, micropyle region of the seed, and wounded leaf tissue) (Domergue *et al.*, 2010). *FAR2* (*MS2*) is expressed in the tapetum of anthers at the time of microspore release from tetrads (Aarts *et al.*, 1997; Chen *et al.*, 2011). *FAR3* (*CER4*) is expressed specifically in the epidermal cells of aerial tissues, which is consistent with its role in cuticular wax metabolism (Rowland *et al.*, 2006). *FAR7* has nearly undetectable gene expression in root and aerial tissues (Domergue *et al.*, 2010), and does not encode a functional FAR because of an early STOP codon in the transcript (Doan *et al.*, 2009). *FAR8* also has near undetectable gene expression in root and aerial tissues (Domergue *et al.*, 2010), and the encoded protein has very low activity (Doan *et al.*, 2009; Chapter 3 and 4 of this thesis). Previous work has shown that *FAR6* is expressed in the stem using quantitative RT-PCR (Domergue *et al.*, 2010) and a DNA microarray experiment indicated that it is upregulated in the epidermis (Suh *et al.*, 2005). However, a detailed characterization of the gene expression pattern of *FAR6* was lacking.

To analyze the tissue specific gene expression pattern of *FAR6* in more detail, a promoter-reporter fusion was constructed using the pBI101 plant binary vector by fusing 2187 bp of the upstream sequence and the first seven codons of the *FAR6* coding region

in frame with the β-Glucuronidase (GUS) reporter gene. The T-DNA harbouring the *FAR6* promoter::GUS fusions was incorporated into *Arabidopsis thaliana* Columbia-0 ecotype (wild-type) by means of *Agrobacterium*-mediated transformation. Five independent T2 lines that showed representative expression patterns were selected for histochemical staining (276-3, 277-3, 277-5, 278-2 and 278-3). *FAR6* promoter::GUS gene expression was found in various tissues of the plant. There was high expression in the bottom of the stem (1st internode) (Figure 3.5, A). The cross section of the stem revealed that there is expression in the stem epidermal layer and the underlying few cell layers, but not in the inner cortex or vascular bundles (Figure 3.5, B-C). In flowers, the *FAR6* promoter was active in the epidermis, endothecium and tapetum of anthers, but not in the microspores or vascular strand (Figure 3.5, D-E). The *FAR6* promoter also drove GUS expression in the replum and receptacle of siliques, but not the seeds or carpels (Figure 3.5, F). In roots, there was expression in the emerging root primordia and it remained in the root cap throughout primary and lateral root development (Figure 3.5, G-I).

The tissue-specific gene expression patterns of *FAR6* were further examined by semi-quantitative RT-PCR using RNA extracted from 6-week-old *Arabidopsis thaliana* Columbia-0 ecotype plants (wild-type). The GUS histochemical assay revealed that *FAR6* was expressed principally in the stem and flowers; therefore, these regions were focused on for the semi-quantitative RT-PCR. The analysis indicated that *FAR6* is expressed the highest in the bottom region of the stem (1st internode and node), and moderately expressed in the middle and top regions of the stem (2nd to 3rd internodes) (Figure 3.6).

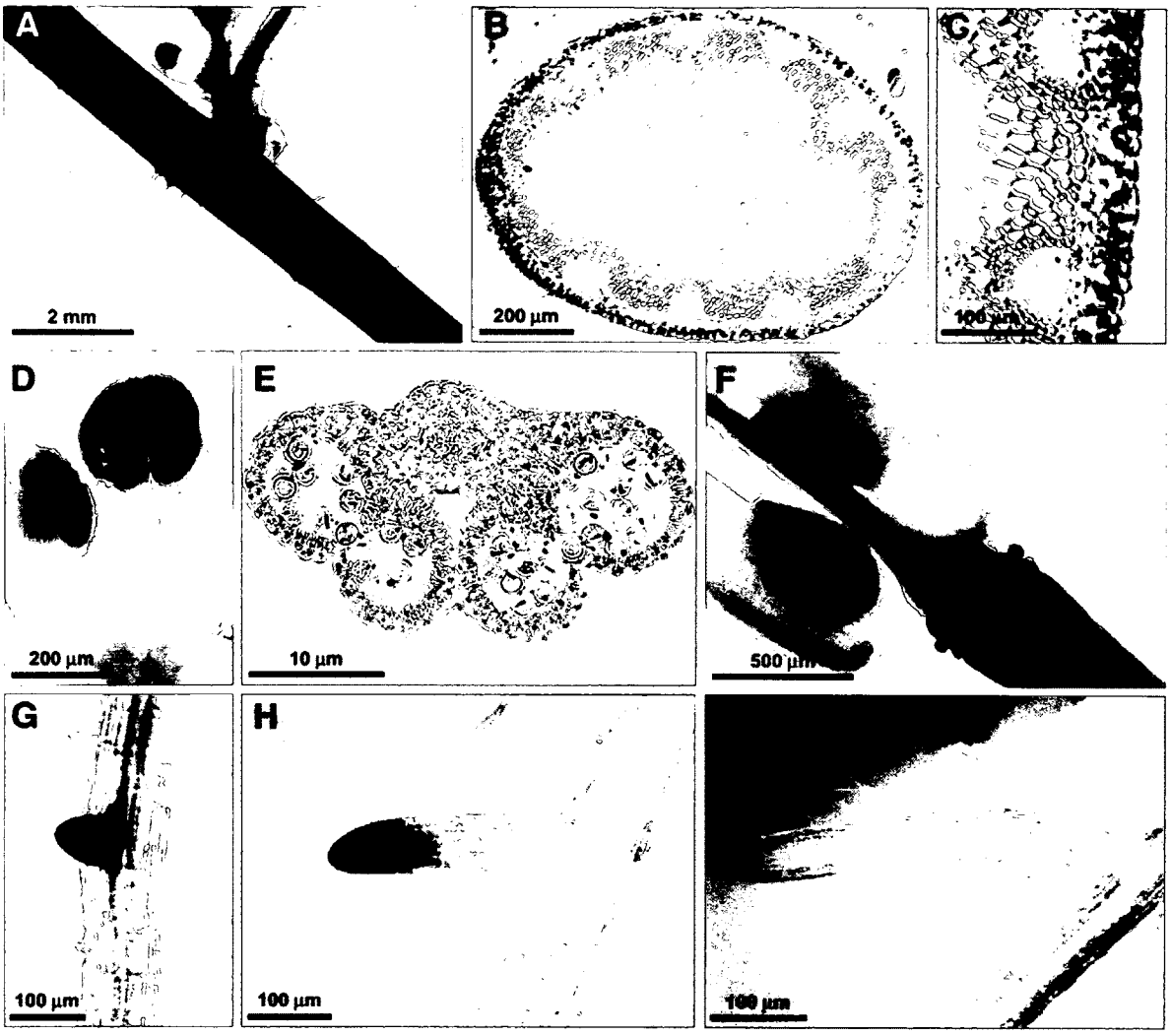


Figure 3.5 Gene expression pattern of *FAR6* using a reporter GUS promoter fusion. The fusion containing 2187 bp upstream region, relative to the start codon, of *FAR6*. A representative transgenic line is shown. The panels are the first internode of the stem (A) cross section of the first internode of the stem (B-C), anther (D), anther cross section (E), silique receptacle (F), lateral emerging root (G), and elongated lateral root (H-I).

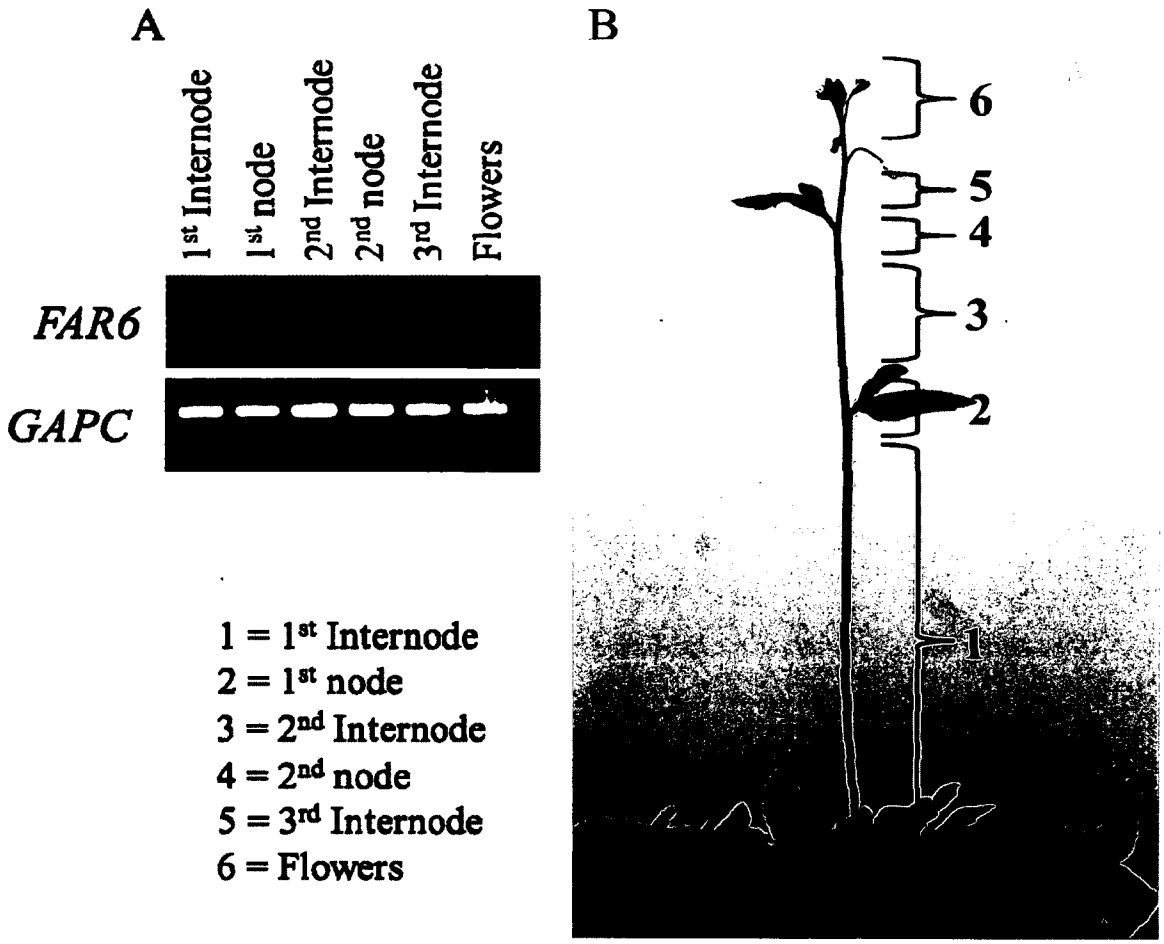


Figure 3.6 Gene expression pattern of *FAR6* using semi-quantitative RT-PCR. A: Semi-quantitative RT-PCR of *FAR6* transcript levels in Arabidopsis Col-0, using RNA isolated from 6 different tissues (#1-6) (Top) and *GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE C SUBUNIT* (*GAPC*) as the corresponding constitutively expressed control (Bottom). B: Diagram of locations of RNA isolation from Arabidopsis Col-0 plants.

3.4 Heterologous Yeast Expression System for Characterization of Plant FARs

A modified version of pYES2 (Invitrogen), pYES2-His6x/T7, was used for heterologous expression in yeast. It contains coding regions for N-terminal histidine (x6) and T7 epitope tags for protein purification and detection, respectively. pYES2 is a yeast expression vector used to express proteins of interest in *Saccharomyces cerevisiae*. It contains a *URA3* gene for selection in yeast, a 2μ origin for high-copy maintenance, and a *GAL1* promoter for high level inducible protein expression in yeast by galactose. The coding regions of the Arabidopsis FARs and Jojoba FAR were cloned into pYES2-His6x/T7 in frame with the tags and downstream of the *GAL1* promoter.

The induction conditions were initially monitored using tagged FAR5 and FAR8 over 48 hours with yeast cell cultures harvested at six time points (t= 0, 6, 9, 12, 24, and 48 hours). FAR5 and FAR8 were used to establish induction conditions because preliminary experiments indicated that FAR5 is expressed at high levels and FAR8 is expressed at low levels, and so they represent two extremes. Protein was extracted and run on duplicate SDS-PAGE gels: one stained with Coomassie blue and the other transferred to nitrocellulose membrane for Western blot analysis. The Coomassie stained gels revealed that there was protein present in each lane. The Western blot revealed 6 hours was sufficient time to induce FAR5 and that the signal remained high through 48 hours of induction (Figure 3.7). FAR8 was expressed at very low levels, but signal was consistently detected at 6 and 9 hours (Figure 3.7).

For lipid analysis, the induction conditions were initially monitored using tagged FAR5 and FAR8 over six days and yeast cell culture was harvested at 24 hr intervals (t= 1, 2, 3, 4, 5, and 6 days). This time course was performed with FAR8 to ensure that the

small amount of C16:0-OH produced could be detected. FAR5 mostly generated C18:0-OH, but a relatively small amount of C16:0-OH was also produced (Figure 3.8). Both fatty alcohols were detectable after 24 hours of induction, but maximal amounts were produced after 4 days of induction. FAR8 only produced C16:0-OH, but this was in very small amounts and even less than the C16:0-OH produced by FAR5 (Figure 3.8). This is consistent with the relatively low amounts of FAR8 protein detected in the Western blots. Nonetheless, the C16:0-OH produced by FAR8 induction was observable after 24 hours of induction. It was determined that 4 days was the most favourable time to induce FAR proteins for reliable detection of fatty alcohols, even when produced in small amounts.

In order to determine if the N-terminal His6x/T7 tag affected FAR enzyme function, the un-tagged and tagged versions of empty vector (EV) and FAR5 were induced for 4 days and the fatty alcohols found in the yeast pellets and supernatants were analyzed using gas chromatography. The chromatograms were examined by comparing the percent areas of the peaks. Percent area was calculated using the area of the peak of interest out of the total area of the peaks for C16:0-FA, C16:1-FA, C16:0-OH, C18:0-FA, C18:1-FA, C18:0-OH, and C26:0-FA (FA = fatty acid, OH = fatty alcohol). For transgenic yeast expressing untagged FAR5 (11.79± 0.93% area C18:0-OH, 0.26 ± 0.02% area C16:0-OH) and tagged FAR5 (13.17 ± 1.10% area C18:0-OH, 0.52 ± 0.04% area C16:0-OH), the chromatograms revealed they produced relatively the same amount of C16:0-OH and C18:0-OH in the pellet (Figure 3.9A, 3.9B, Table 3.3). Untagged FAR5 (38.62± 9.27% area C18:0-OH, 0.67 ± 0.21% area C16:0-OH) and tagged FAR5 (38.10 ± 8.83% area C18:0-OH, 0.38 ± 0.14% area C16:0-OH) also produced relatively the same amount of C16:0-OH and C18:0-OH in the supernatant (Table 3.4). For the transgenic

yeast expressing untagged FAR8 (0.37 ± 0.02% area C16-OH) and tagged FAR8 (0.26 ± 0.02% area C16-OH), the chromatograms revealed they produced relatively the same amount of C16-OH in the pellet (Figure 3.9A, 3.9B, Table 3.3). . There were no detectable amounts of C16-OH or C18-OH from the transgenic yeast expressing untagged FAR8 and tagged FAR8 (Table 3.4). These results showed that there is no discernible difference between activities of untagged and tagged FAR5. It was presumed that this would also be the case for the other FARs, hence the tagged versions were used for the remainder of the experiments to monitor protein levels.

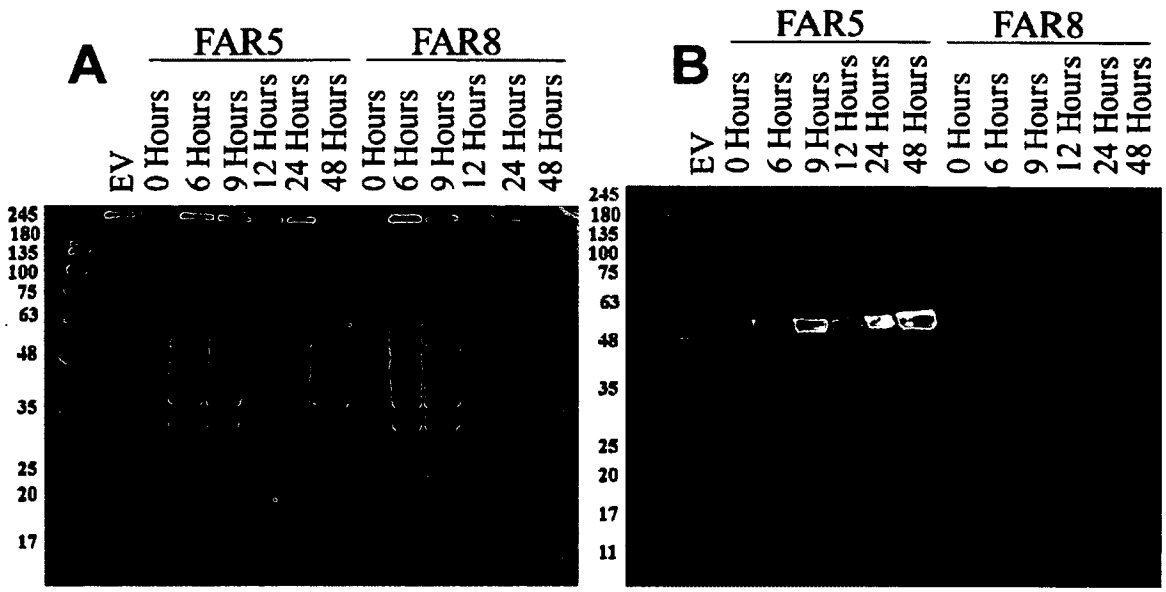


Figure 3.7 SDS-PAGE of yeast pellets collected after 0, 6, 9, 12, 24, and 48 hours of induction by galactose. (A) SDS-PAGE stained with Coomassie Brilliant Blue, (B) Western blot analysis of proteins detected using T7 Tag monoclonal mouse primary antibody and horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibody.

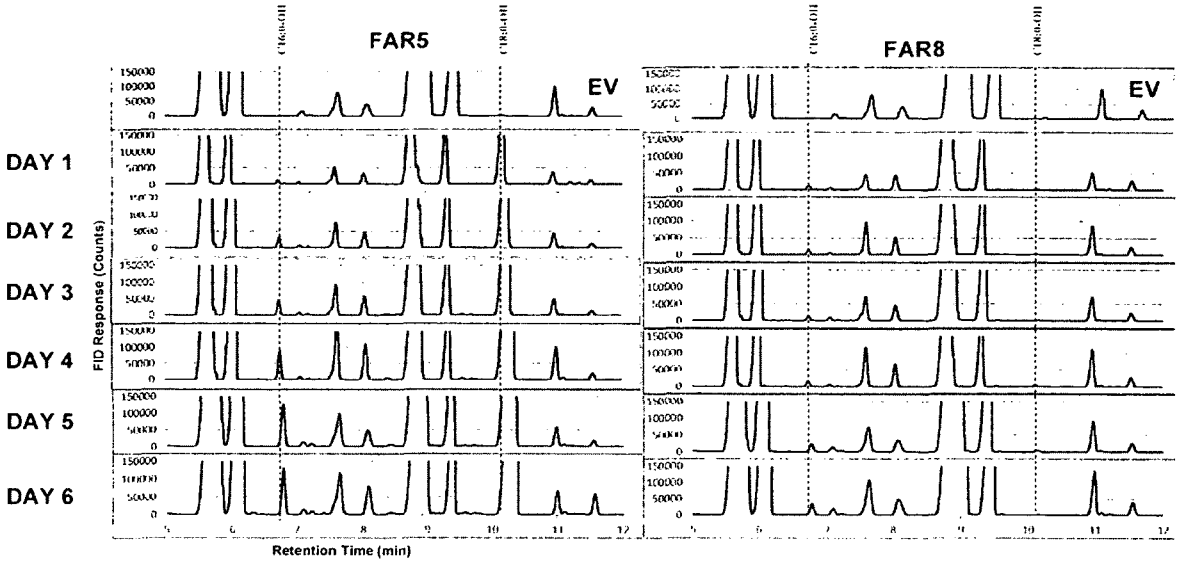


Figure 3.8 Gas chromatograms of lipids extracted from transgenic yeast expressing tagged versions of Arabidopsis FAR5 and FAR8. The top chromatograms are empty vector (EV) control, and the bottom six are FAR5 (left) and FAR8 (right) expressing yeast. Fatty alcohols were extracted from the yeast cells and analyzed by gas chromatography after 1-6 days of induction with galactose. The primary alcohol peaks of interest are indicated by the dashed lines and denoted above the chromatograms. The EV chromatogram is after 6 days of induction with galactose. FID = Flame Ionization Detection.

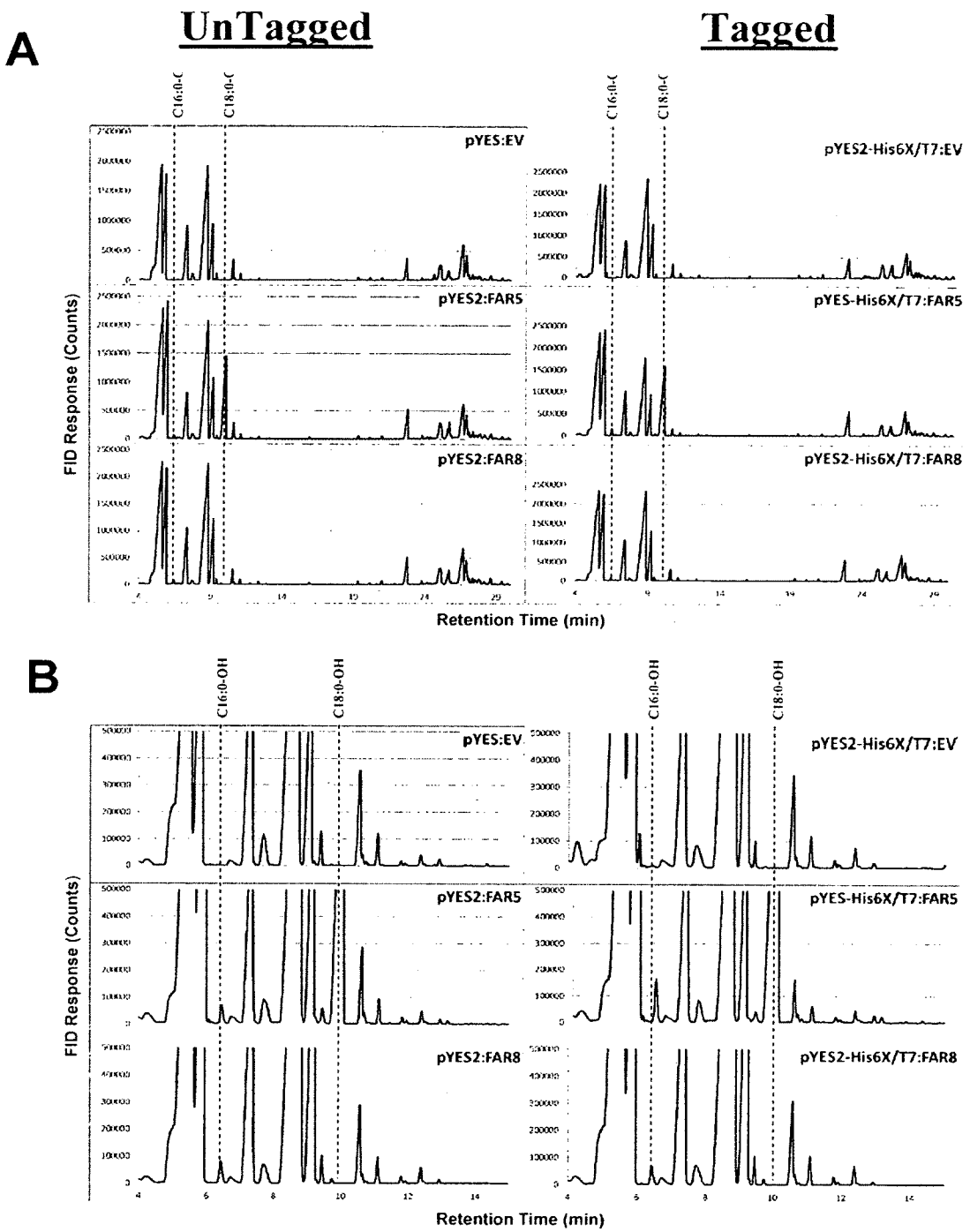


Figure 3.9 Gas chromatograms of lipids extracted from transgenic yeast expressing untagged and tagged versions of Arabidopsis FAR5 and FAR8. (A) Zoomed out chromatograms, (B) Zoomed in chromatograms. Lipids were extracted from the yeast cells and analyzed by gas chromatography after 4 days of induction. The primary fatty alcohol peaks of interest are indicated by the dashed lines and denoted above the chromatograms. FID = Flame Ionization Detection.

Table 3.3 Primary fatty alcohol compositions of the pellets of untagged and tagged versions of Arabidopsis FAR5 expressing yeast strains (Values are displayed in percent area)

	Untagged		Tagged	
Sample	**16:0-OH (% area)**	**18:0-OH (% area)**	**16:0-OH (% area)**	**18:0-OH (% area)**
Empty Vector #1	-	-	-	-
Empty Vector #2	-	-	-	-
Empty Vector #3	-	-	-	-
AVG	-	-	-	-
St Dev	-	-	-	-
FAR5 #1	0.25	11.18	0.57	14.36
FAR5 #2	0.26	11.34	0.51	12.98
FAR5 #3	0.29	12.86	0.48	12.19
AVG	**0.26**	**11.79**	**0.52**	**13.17**
St Dev	0.02	0.93	0.04	1.10
FAR8 #1	0.36	-	0.27	-
FAR8 #2	0.35	-	0.24	-
FAR8 #3	0.39	-	0.26	-
AVG	**0.37**	-	**0.26**	-
St Dev	0.02	-	0.02	-

Table 3.4 Primary fatty alcohol compositions of the supernatants of untagged and tagged versions of Arabidopsis FAR5 expressing yeast strains (Values are displayed in percent area)

	Untagged		Tagged	
Sample	**16:0-OH (% area)**	**18:0-OH (% area)**	**16:0-OH (% area)**	**18:0-OH (% area)**
Empty Vector #1	-	-	-	-
Empty Vector #2	-	-	-	-
Empty Vector #3	-	-	-	-
AVG	-	-	-	-
St Dev	-	-	-	-
FAR5 #1	0.60	31.10	0.82	58.35
FAR5 #2	0.91	48.98	0.28	31.86
FAR5 #3	0.50	35.78	0.48	44.34
AVG	**0.67**	**38.62**	**0.53**	**44.85**
St Dev	0.21	9.27	0.27	13.25
FAR8 #1	-	-	-	-
FAR8 #2	-	-	-	-
FAR8 #3	-	-	-	-
AVG	-	-	-	-
St Dev	-	-	-	-

3.5 Fatty Alcohol Production in Yeast Expressing Arabidopsis FARs or Jojoba FAR

The activities and substrate specificities of Arabidopsis FARs, as well as the seed-expressed Jojoba FAR, were analyzed by means of the *Saccharomyces cerevisiae* heterologous expression system. The full length open reading frames for *FAR1, FAR2 (MS2), FAR3 (CER4), FAR4, FAR5, FAR6,* and *FAR8* as well as versions of *FAR2* and *FAR6* lacking their N-terminal extension sequences were cloned into pYES2-His6x/T7. FAR2ΔN and FAR6ΔN were designed to have truncations of 119 and 71 amino acids, respectively (Figure 3.4, 3.10A). These truncations were made based on the alignment with the other Arabidopsis FARs, such that the unique N-terminal extensions of FAR2 and FAR6 were removed (Figures 3.3 and 3.4). The protein levels in transgenic yeast were measured by Western blot analysis and the lipid contents were measured by gas chromatography. Empty vector (EV) was used as a negative control.

For detection of protein levels by Western blot analysis, transgenic yeast cells were induced for 6 hours at 30°C in galactose-containing media. Western blotting confirmed that all of the proteins were expressed, and there was no expression for empty vector control (Figure 3.10 C). FAR1, FAR2, FAR3, FAR4, FAR5 accumulated to relatively high levels in yeast and had the expected molecular weights of 55.5kDa, 68.4kDa, 56.0kDa, 56.2kDa, and 56.4kDa, respectively. The Jojoba FAR accumulated to lower levels than FAR1-5 and had the expected molecular weight of 56.2kDa. FAR2ΔN, FAR6, FAR6ΔN and FAR8 accumulated to much lower levels than FAR1-5. FAR6ΔN and FAR8 had the expected molecular weights of 53.7kDa and 56.4kDa, respectively. FAR2ΔN and FAR6, however, ran slightly higher than what was expected, which were supposed to be 55.6kDa and 61.6kDa, respectively.

For the lipid analysis by gas chromatography, transgenic yeast cells were induced for 4 days at 30°C in media containing galactose. The fatty alcohols were produced using the endogenous acyl-CoAs pools in *S. cerevisiae* (no feeding of fatty acids performed). The fatty alcohols retained in the yeast pellets (internal) and those found in the media (secreted) were separately extracted and analyzed. The fatty alcohols were quantified using C13:0-OH as an internal standard. The GC analysis revealed that the FAR family produced primary alcohols ranging from C16:0 to C26:0 in yeast, with each FAR having a distinct substrate specificity (Figures 3.10 D, 3.11-3.13; Tables 3.5-3.7). The quantitated results are shown graphically in Figures 3.10 D and 3.13, and tabulated in Tables 3.5-3.7. The gas chromatograms of lipids extracted from pellets of yeast are shown in Figures 3.11 (zoomed out) and 3.12 (zoomed in). FAR1 mainly produced C22:0-OH, and to a lesser extent C18:0-OH and C20:0-OH, which is in agreement with previous results obtained by Domergue *et al.* (2010). FAR2 produced very little, if any, fatty alcohols above background. FAR2ΔN, which lacks the N-terminal extension, was in contrast highly active and mainly produced C16:0-OH and to a lesser extent C18:0-OH, which is in agreement with the results obtained by Chen *et al.* (2011). FAR3 produced C24:0-OH and C26:0-OH but mostly C26:0-OH, which is in agreement with results obtained by Rowland *et al.* (2006) and Domergue *et al.* (2010). FAR4 mainly produced C20:0-OH and to a lesser extent C18:0-OH and C22:0-OH, which is in agreement with results obtained by Domergue *et al.* (2010). FAR5 mainly produced C18:0-OH and to a lesser extent C16:0-OH, which is in agreement with results obtained by Domergue *et al.* (2010). FAR6 produced very little, if any, fatty alcohols above background. FAR6ΔN, which lacks the N-terminal extension, did generate C16:0-OH

alcohols above background (apparent in Figure 3.12 zoomed in chromatogram), but these are in very low abundance especially when compared to the amount of C16:0-OH produced by FAR2ΔN expression in yeast. FAR8 also generated very little fatty alcohol relative to empty vector, but close inspection of the chromatograms did reveal a consistent peak corresponding to C16:0-OH (Figure 3.8). Jojoba FAR produced small amounts of C18:0-OH above background (Figures 3.10D, 3.11, and 3.12).

Separate analysis of internal and secreted primary fatty alcohols produced in the transgenic yeast strains revealed large differences in the amounts of fatty alcohol secreted into the media (Figures 3.13 and 3.14; Tables 3.5-3.7). The C16:0-OH produced by FAR2ΔN and the C18:0-OH produced by FAR5 were found in relatively large amounts in the media (Figure 3.13, B). The percentage of C16:0-OH produced by FAR2ΔN that was secreted was about 20%, whereas the percentage of C18:0-OH produced by FAR5 that was secreted was about 40% (Figure 3.14). Fatty alcohols produced by other FARs were found to at least some degree in the media, with the exception of the C24:0 and C26:0 fatty alcohols produced by FAR3. The percentage of C22:0-OH secreted by yeast expressing FAR1 was very low and barely detectable. The C20:0-OH produced by yeast expressing FAR1 and FAR4 were secreted at 28 and 12%, respectively.

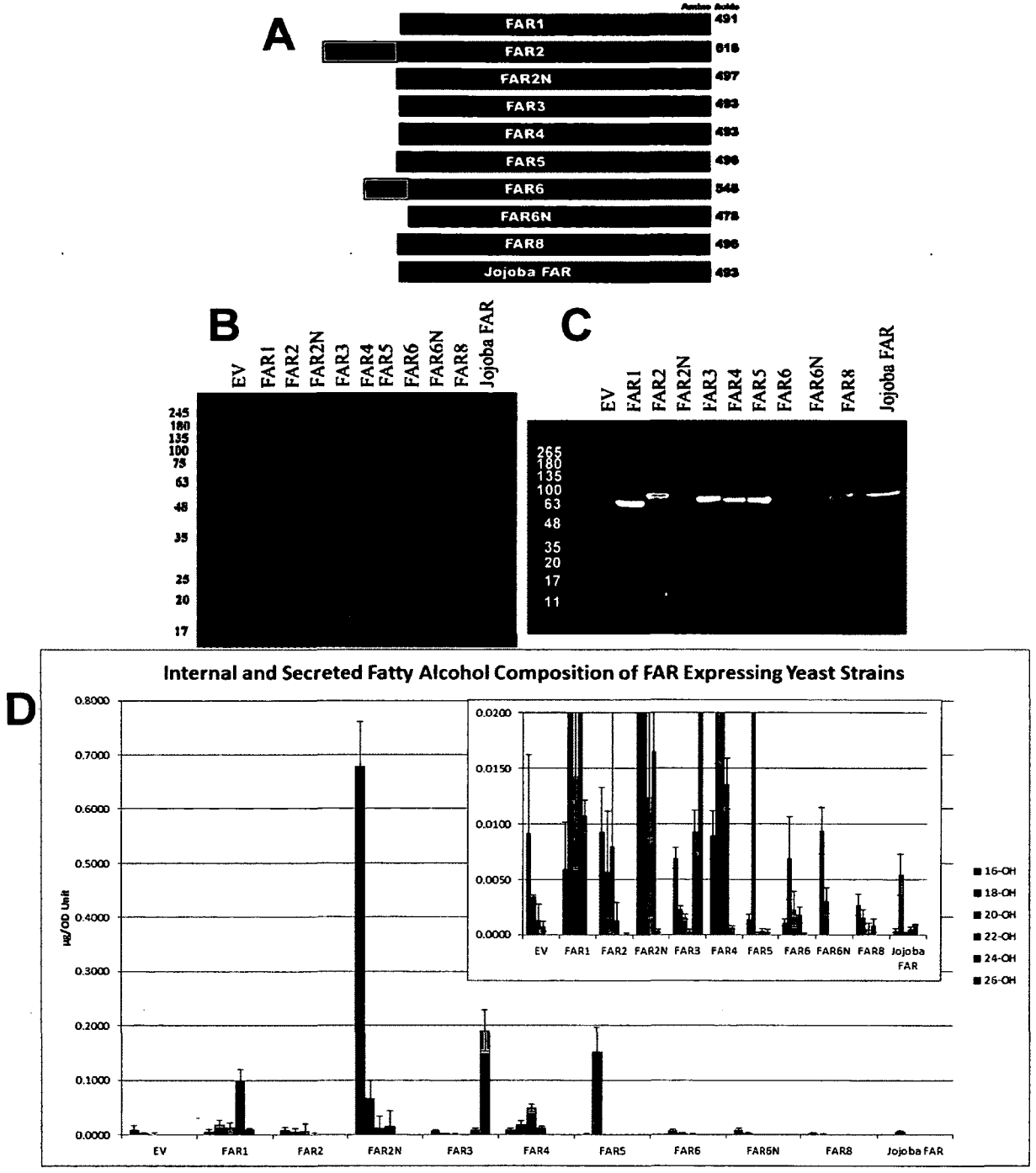


Figure 3.10 *Arabidopsis thaliana* FAR protein family and Jojoba FAR expressed in yeast. A. Structures of FAR1, FAR2, FAR3, FAR4, FAR5, FAR6, FAR8, FAR2 and FAR6 with truncations of their predicted N-terminal chloroplast targeting sequences, and Jojoba FAR. B. SDS-PAGE stained with Coomassie Brilliant Blue. C. Western blot analysis of proteins detected with T7 Tag monoclonal mouse primary antibody and horseradish peroxidase (HRP)-conjugated Anti-Mouse Secondary antibody. D. Analysis of total fatty alcohols produced by gas chromatography, done in triplicates, where values are expressed in µg/OD600 Unit (error bars are ± standard deviation)

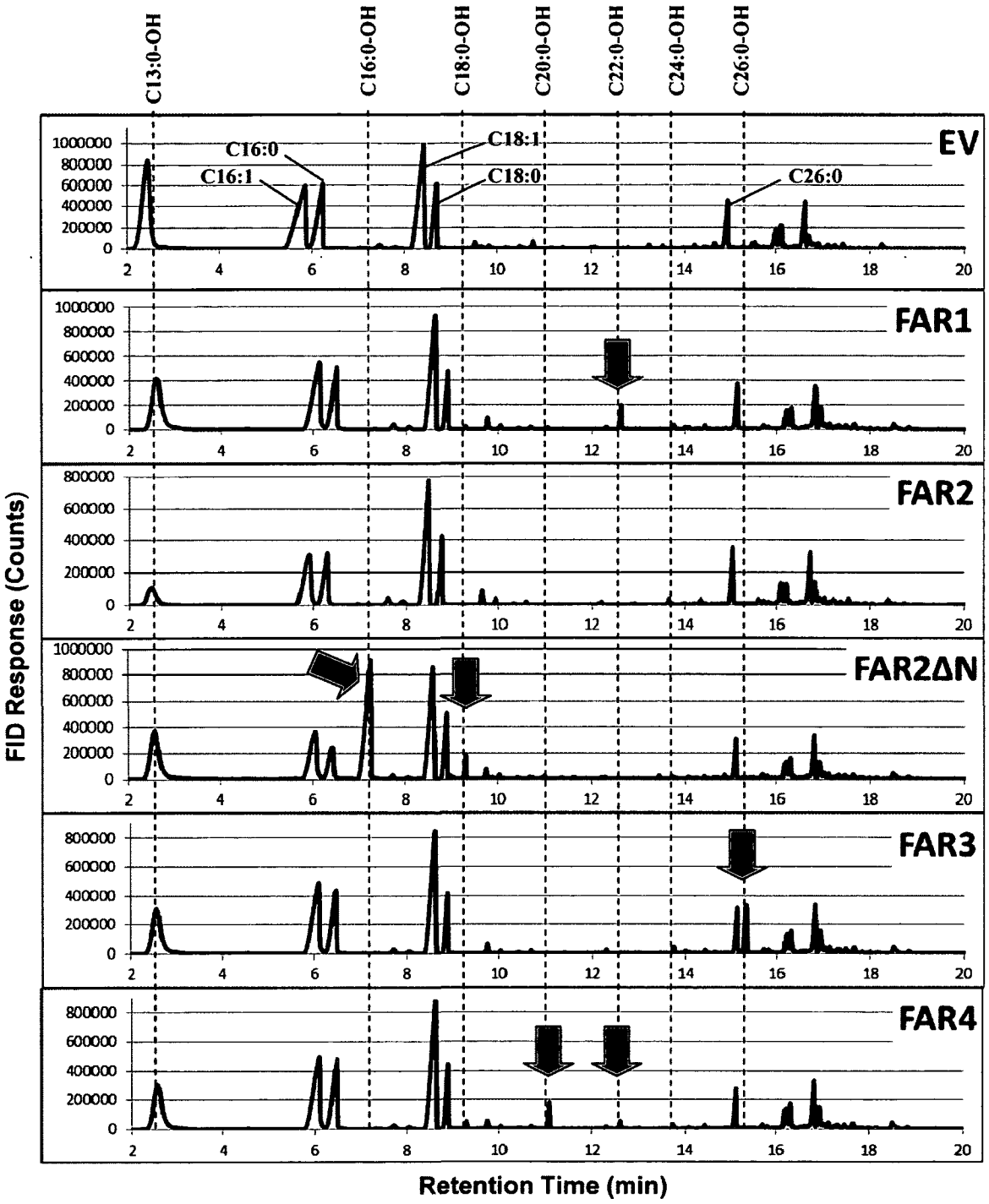


Figure 3.11 Gas chromatograms of lipids extracted from transgenic yeast expressing tagged Arabidopsis FAR1, FAR2, FAR2ΔN, FAR3, and FAR4. Lipids were extracted from the yeast cell pellets and analyzed by gas chromatography. The primary alcohol peaks of interest are indicated by the dashed lines and denoted above the chromatograms. The peak indicated by C13:0-OH is the internal standard (IS) (1-tridecanol). FID = Flame Ionization Detection.

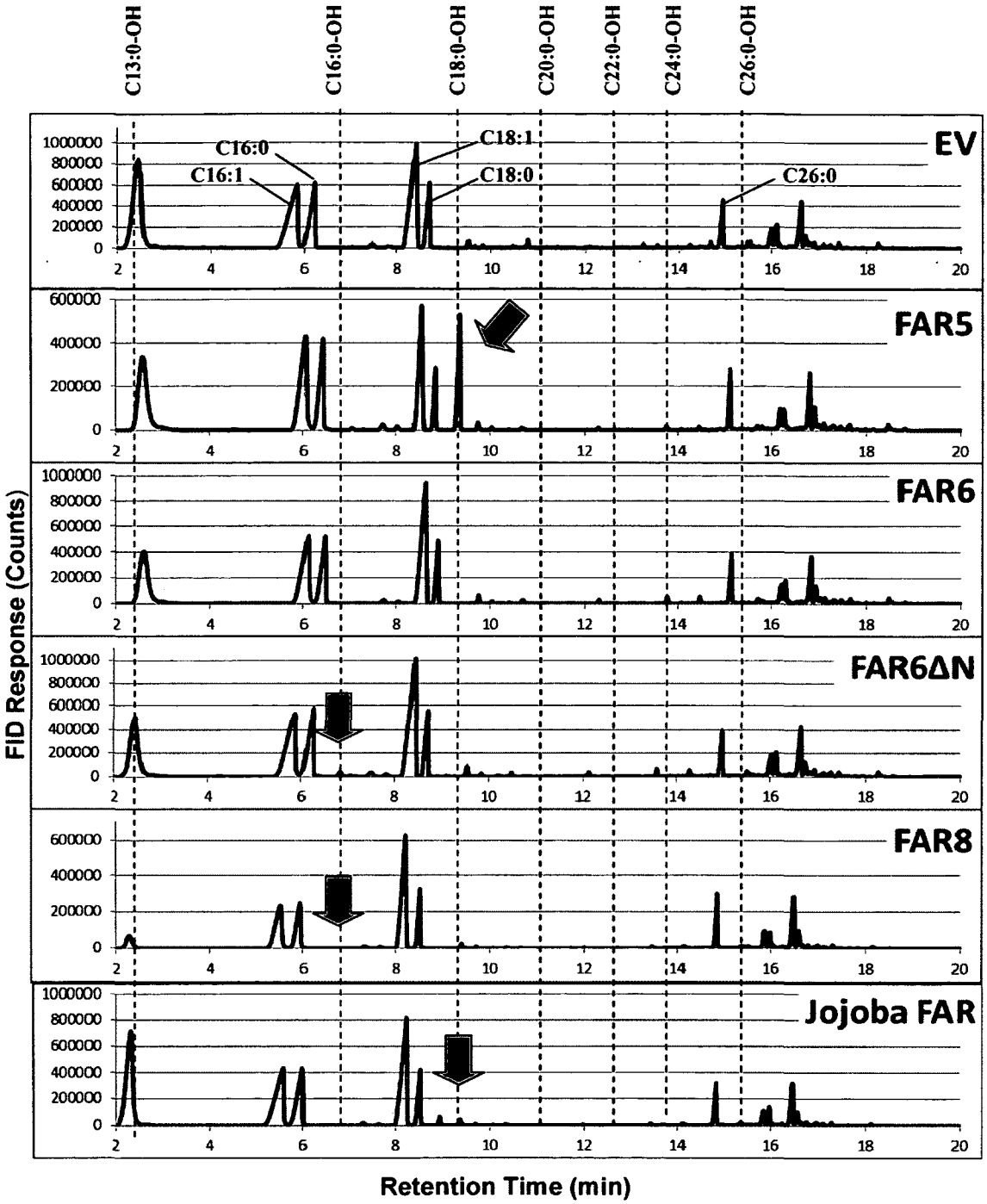


Figure 3.11 continued. Gas chromatograms of lipids extracted from transgenic yeast expressing tagged Arabidopsis FAR5, FAR6, FAR6ΔN, FAR8, and Jojoba FAR. Lipids were extracted from the yeast cell pellets and analyzed by gas chromatography. The primary alcohol peaks of interest are indicated by the dashed lines and denoted above the chromatograms. The peak indicated by C13:0-OH is the internal standard (IS) (1-tridecanol). FID = Flame Ionization Detection.

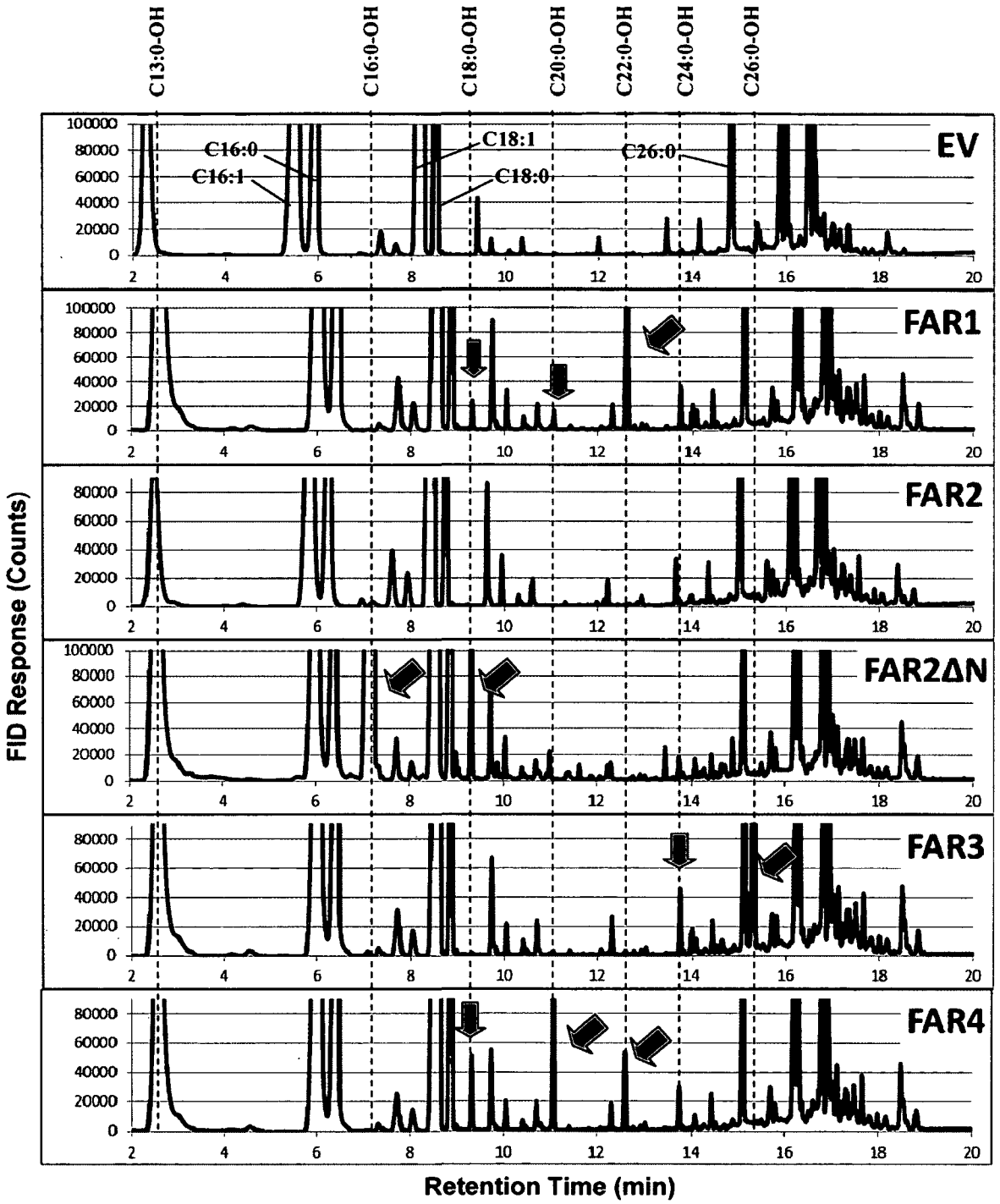


Figure 3.12 Zoomed in gas chromatograms of lipids extracted from transgenic yeast expressing tagged Arabidopsis FAR1, FAR2, FAR2ΔN, FAR3, and FAR4. Lipids were extracted from the yeast cell pellets and analyzed by gas chromatography. The primary alcohol peaks of interest are indicated by the dashed lines and denoted above the chromatograms. The peak indicated by C13:0-OH is the internal standard (IS) (1-tridecanol). FID = Flame Ionization Detection.

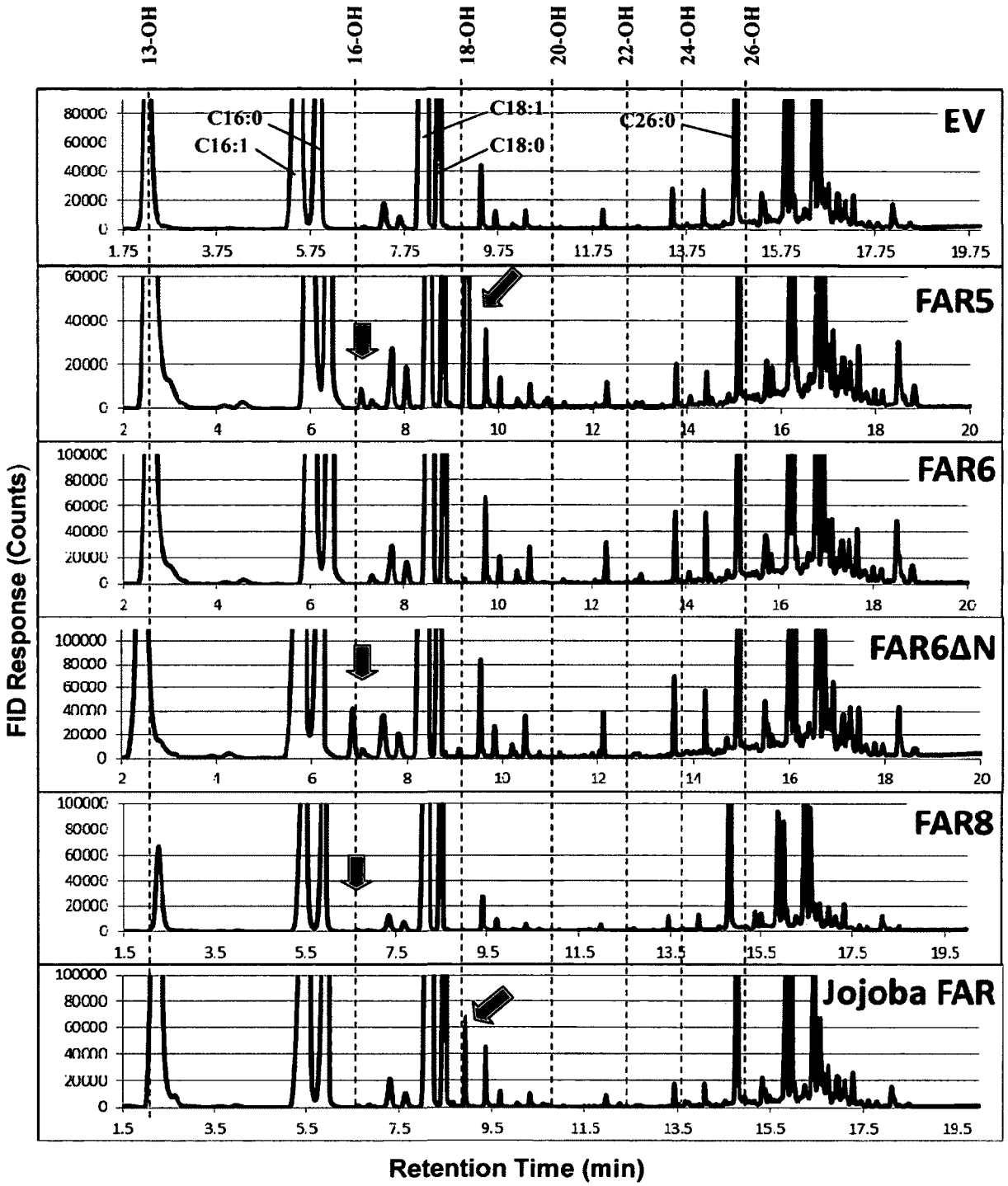


Figure 3.12 continued. Zoomed in gas chromatograms of lipids extracted from transgenic yeast expressing tagged Arabidopsis FAR5, FAR6, FAR6ΔN, FAR8, and Jojoba FAR. Lipids were extracted from the yeast cell pellets and analyzed by gas chromatography. The primary alcohol peaks of interest are indicated by the dashed lines and denoted above the chromatograms. The peak indicated by C13:0-OH is the internal standard (IS) (1-tridecanol). FID = Flame Ionization Detection.

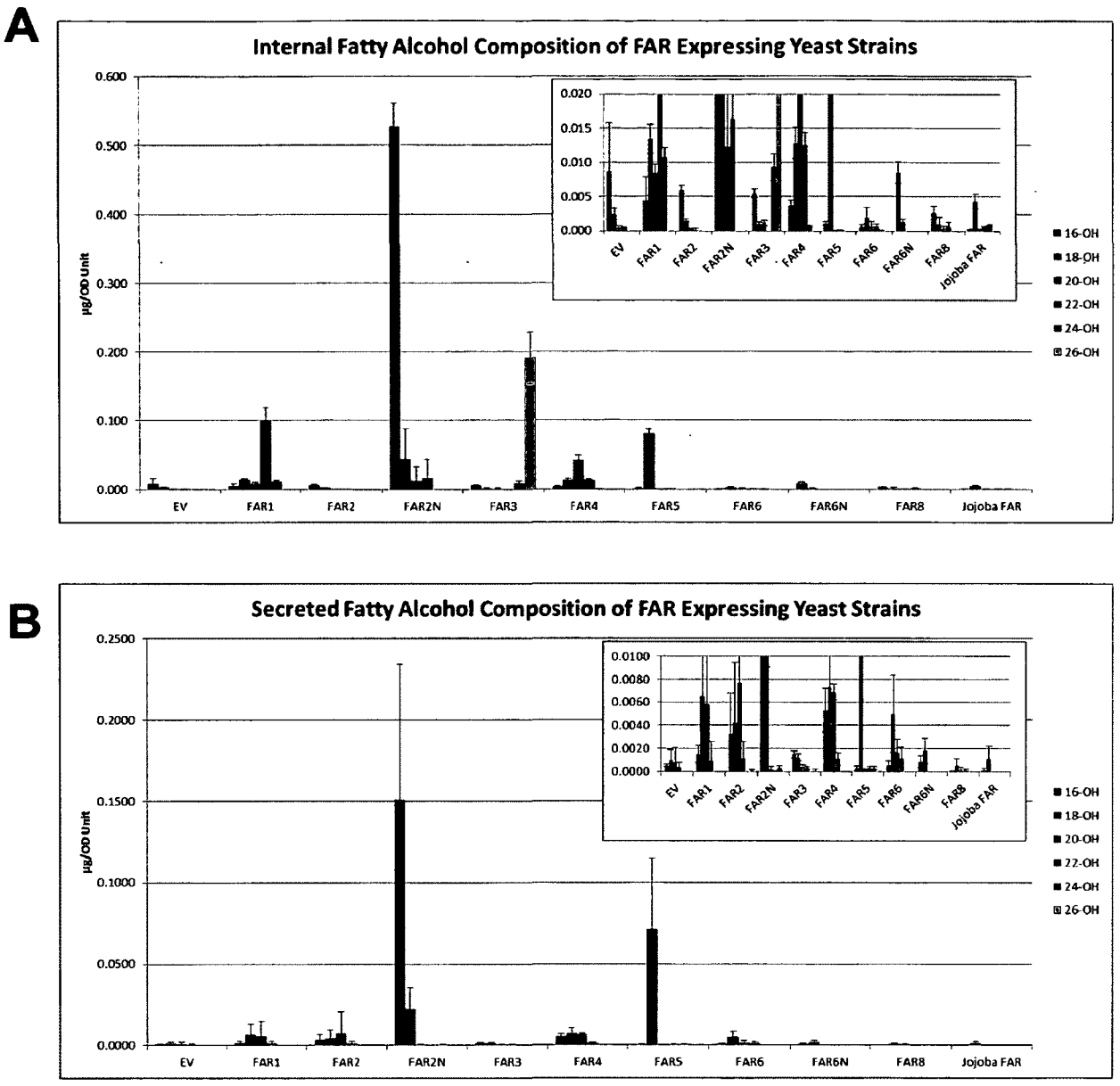


Figure 3.13 *Arabidopsis thaliana* FAR protein family and Jojoba FAR expressed in yeast: Internal and Secreted Fatty Alcohols. A. Analysis of internal fatty alcohols produced in yeast by gas chromatography. B. Analysis of secreted fatty alcohols produced in yeast by gas chromatography. Done in triplicates, where values are expressed in µg/Unit OD_{600} (error bars are ± standard deviation).

Table 3.5 Internal Fatty Alcohol Composition of the pellet of transgenic yeast strains expressing Arabidopsis and Jojoba FARs (Error is ± standard deviation (SD))

	C16:0-OH (μg/OD_{600} ±SD)	C18:0-OH (μg/OD_{600} ±SD)	C20:0-OH (μg/OD_{600} ±SD)	C22:0-OH (μg/OD_{600} ±SD)	C24:0-OH (μg/OD_{600} ±SD)	C26:0-OH (μg/OD_{600} ±SD)
FAR1		0.0134±0.0021	0.0084±0.0013	0.0995±0.0192		
FAR2	0.0060±0.0006	0.0015±0.0002				
FAR2ΔN	0.5277±0.0335	0.0444±0.0444				
FAR3					0.0093±0.0019	0.1911±0.0376
FAR4		0.0128±0.0023	0.0426±0.0065			
FAR5	0.0011±0.0003	0.0815±0.0059				
FAR6	0.0005±0.0004					
FAR6ΔN	0.0085±0.0016					
FAR8	0.0027±0.0009					
Jojoba FAR		0.0043±0.0010				

Table 3.6 Secreted fatty alcohol composition of the supernatant of transgenic yeast strains expressing Arabidopsis and Jojoba FARs (Error is ± standard deviation (SD))

	C16:0-OH ($\mu g/OD_{600}$ ±SD)	C18:0-OH ($\mu g/OD_{600}$ ±SD)	C20:0-OH ($\mu g/OD_{600}$ ±SD)	C22:0-OH ($\mu g/OD_{600}$ ±SD)	C24:0-OH ($\mu g/OD_{600}$ ±SD)	C26:0-OH ($\mu g/OD_{600}$ ±SD)
FAR1		0.0065±0.0062	0.0059±0.0087	0.0010±0.002		
FAR2	0.0033±0.0035	0.0042±0.0053				
FAR2ΔN	0.1508±0.0834	0.0222±0.0131				
FAR3					N.D.	N.D.
FAR4		0.0073±0.0033	0.0069±0.0007			
FAR5	0.0003±0.0002	0.0713±0.0435				
FAR6	0.0006±0.0004					
FAR6ΔN	0.0009±0.0005					
FAR8	N.D.					
Jojoba FAR		0.0011±0.0011				

N.D. = None Detected

Table 3.7 Fatty alcohol composition of the combined pellet and supernatant fractions, of transgenic yeast strains expressing Arabidopsis and Jojoba FARs
(Error is ± standard deviation (SD))

	C16:0-OH (μg/OD_{600} ±SD)	C18:0-OH (μg/OD_{600} ±SD)	C20:0-OH (μg/OD_{600} ±SD)	C22:0-OH (μg/OD_{600} ±SD)	C24:0-OH (μg/OD_{600} ±SD)	C26:0-OH (μg/OD_{600} ±SD)
FAR1		0.0200±0.0080	0.0143±0.0083	0.1004±0.0190		
FAR2	0.0093±0.0040	0.0057±0.0054				
FAR2ΔN	0.6784±0.0815	0.0666±0.0328				
FAR3					0.0093±0.0019	0.1912±0.0375
FAR4		0.0201±0.0047	0.0495±0.0067			
FAR5	0.0014±0.0005	0.1528±0.0447				
FAR6	0.0011±0.0003					
FAR6ΔN	0.0094±0.0021					
FAR8	0.0027±0.0009					
Jojoba FAR		0.0055±0.0019				

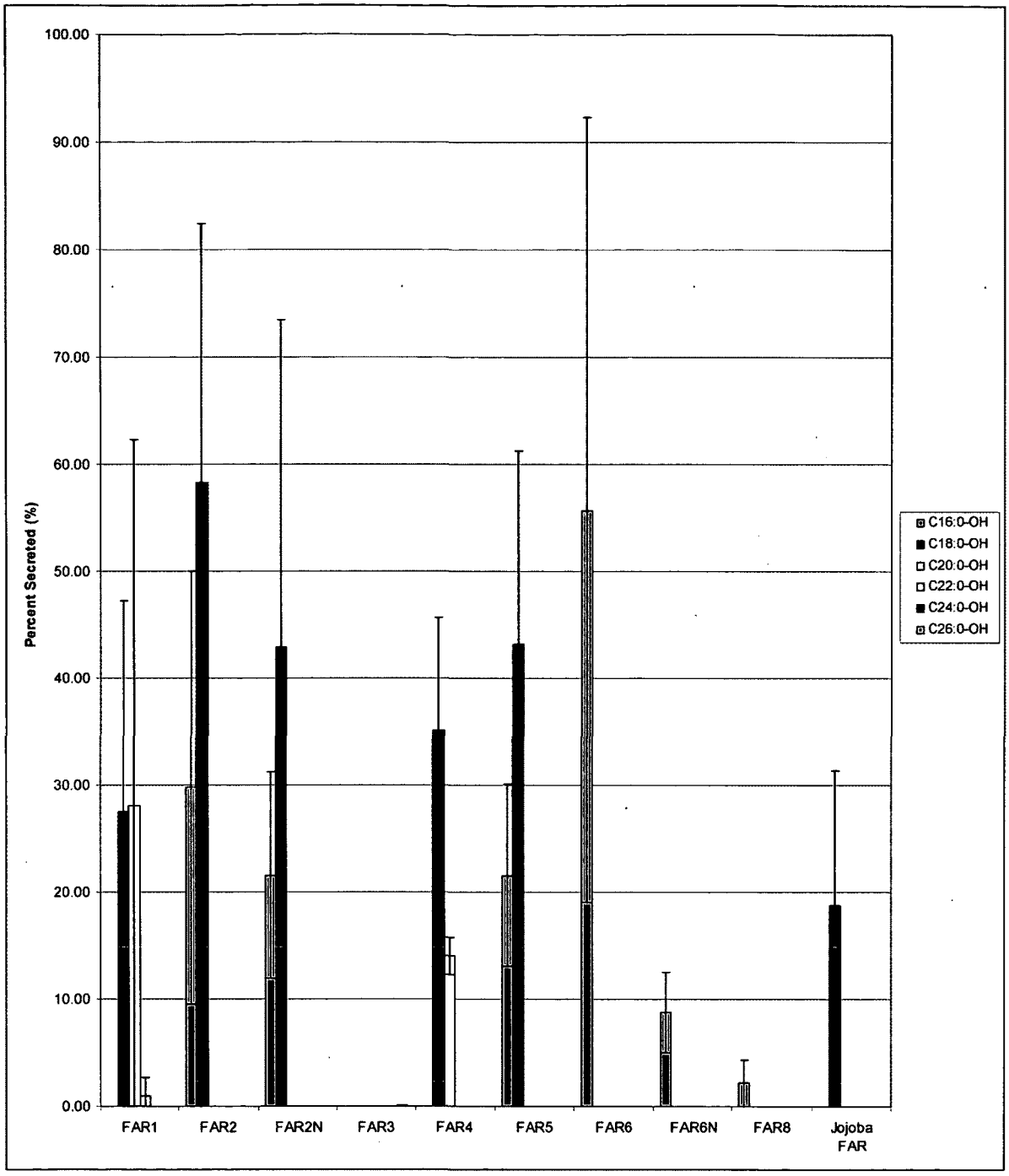


Figure 3.14 Percent of fatty alcohols secreted by transgenic yeast strains expressing Arabidopsis and Jojoba FARs

3.6 Mutation of Putative Active Site Residues in Arabidopsis FAR5

The short-chain alcohol dehydrogenase/reductase (SDR) family members have an active site containing conserved tyrosine and lysine residues (YXXXK) that have important functional roles in catalysis (Jörnvall *et al.*, 1981; Ghosh *et al.*, 1995; Denessiouk *et al.*, 2001; Kavanagh *et al.*, 2008). This motif is present in all FARs and is found within the Rossmann fold domain (Figure 3.3; also see Figure 1.3 of the General Introduction), but the functionality of this motif has not been directly tested in a FAR. To test the role of tyrosine-238 and lysine-156 in FAR5, two mutants were constructed that confer different chemical properties in this region of the protein. Previous studies by Chen *et al.* (1993) on the active site of alcohol dehydrogenase (ADH) in *Drosophila melanogaster* served as the basis for testing the active site in FAR5. Tyrosine was substituted with phenylalanine (Y238F) in FAR5, which was the same substitution made in *Drosophila* ADH demonstrating the importance of the tyrosine phenolic group for catalysis. Phenylalanine lacks a hydroxy group on the aromatic ring. The role of lysine was investigated by mutating it to isoleucine (FAR5-K242I), which again was the same substitution made in *Drosophila* ADH to investigate the role of the conserved lysine at this position. Lysine and isoleucine are similarly sized amino acids, but isoleucine lacks the positive charge conferred by the amino group on lysine. Western blot analysis of FAR5-Y238F and FAR5-K242I showed that they are expressed to the same levels as wild type FAR5 (Figure 3.15C). GC analysis showed that neither mutant produces any fatty alcohols, consistent with essential roles of these amino acids in catalysis (Figure 3.15D).

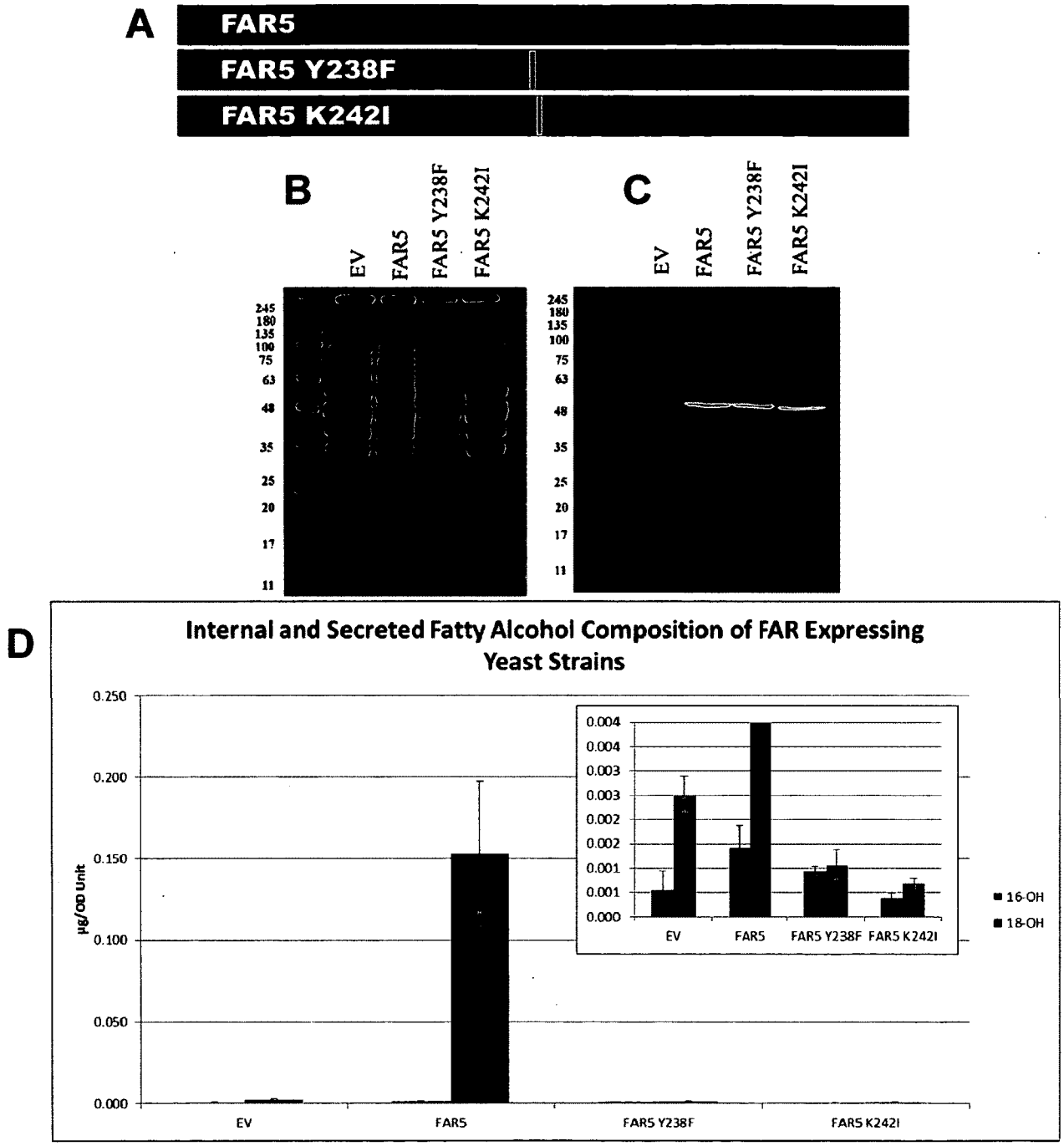


Figure 3.15 Mutation of tyrosine and lysine residues in the putative active site of FAR5. A. Schematics of FAR5, FAR5 Y238F, and FAR5 K242I, showing the relative positions of the active site amino acids. B. SDS-PAGE stained with Coomassie Brilliant Blue. C. Western blot analysis of proteins detected with T7 Tag monoclonal mouse primary antibody and horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibody. D. Gas chromatography analysis of total fatty alcohols produced, done in triplicates, where values are expressed in µg/OD_{600} Unit (error bars are ± standard deviation)

3.7 Discussion

The substrate specificities of eight FARs were assessed by heterologous expression in *Saccharomyces cerevisiae*. Seven of these FARs were from Arabidopsis and one was the well-characterized Jojoba FAR. Quantification of internal and secreted fatty alcohols revealed that FAR2ΔN produced the most fatty alcohol, FAR1, FAR3, FAR4 and FAR5 produced the next most, and FAR2, FAR6, FAR6ΔN, FAR8 and the Jojoba FAR produced little or no fatty alcohols. This variation could be due in part to differing enzyme catalytic efficiencies, but the pool of substrate available in the endogenous acyl pool of yeast also needs to be taken into account. In yeast, the acyl pool has an abundance of C16:0-CoA, C16:1-CoA, C18:0-CoA and C18:1-CoA, a relatively low amount of C26:0-CoA, and very low amounts of C20:0-CoA, C22:0-CoA and C24:0-CoA (Domergue *et al.*, 2003). Thus the apparent high activities of FAR2ΔN and FAR5 generating C16:0 and C18:0 fatty alcohols, respectively, is in part due to the high abundance of relevant acyl-CoA substrate. Nonetheless, the yeast heterologous system was effective in revealing differing substrate specificities between many of the tested FARs. These generally reflected the specificities of the FARs deduced from examination of Arabidopsis *far* mutants. For example, analysis of suberin composition of single mutants with T-DNA insertions in *FAR1*, *FAR4*, or *FAR5* showed that the suberin-associated fatty alcohol composition is specifically altered with regard to chain length in each *far* mutant; C18:0-OH is reduced in *far5*, C20:0-OH is reduced in *far4*, and C22:0-OH is reduced in *far1* (Domergue *et al.*, 2010). These deduced *in planta* specificities match those revealed in this study via heterologous expression in yeast. Loss-of function mutation of *far3/cer4* results in loss of C24:0-C30:0 fatty alcohols normally present in

Arabidopsis stem and leaf cuticular waxes (Rowland *et al.*, 2006). Expression of FAR3/CER4 in yeast generated C24:0 and C26:0 fatty alcohols, but not C28:0 and C30:0 fatty alcohols. Yeast lack C28:0-CoA and C30:0-CoA in its acyl-CoA pool, accounting for this observation. The apparent low or absent activity of a FAR expressed in yeast, such as that observed with FAR2, FAR6, FAR6ΔN, FAR8, and Jojoba FAR may be due to the absence of relevant substrate rather than lack of enzyme activity. For example, Jojoba FAR is expected to be active for monounsaturated acyl chains because the wax esters of jojoba oil are monounsaturated on the alkyl (fatty alcohol) chain. These fatty alcohol moieties are C20:1 and C22:1. It appears that jojoba will not reduce C18:1 acyl-CoA even when substrate is highly abundant. It may be possible to feed C20:1 and C22:1 fatty acids to the yeast or artificially elongate the C18:1 chains by co-expressing a plant condensing enzyme such as Arabidopsis FAE1 (Millar and Kunst, 1997), thus providing relevant substrates for Jojoba FAR. It is also possible that FAR2ΔN, FAR6ΔN, and FAR8, which all generate C16:0 fatty alcohols to varying degrees, could utilize C14:0-CoA substrate if it were present in the yeast acyl pool. C14:0 fatty acid could be fed to determine if C14:0-OH fatty alcohols are produced in yeast expressing these FARs. No feeding experiments were performed in this study due to time constraints. *In vitro* assays are another approach to elucidate substrate specificity and these have been performed with purified FAR2/MS2 and FAR6 demonstrating high activity for C16:0 for both (Chen *et al.*, 2011; Doan *et al.*, 2011), which is consistent with what is observed by heterologous expression in yeast (Doan *et al.*, 2011, and this study). FAR2 and FAR6 are relatively soluble proteins, whereas the other FARs are insoluble under conditions tested thus far, which likely reflects their predicted membrane-associations *in planta*. It is a

challenge to develop *in vitro* assays with insoluble proteins. It is also a challenge to develop *in vitro* assays with acyl substrates longer than C18 as the CoA and ACP derivatives are difficult to synthesize and they are not commercially available.

Comparison of the FAR sequences to other NAD(P)H-dependent oxidoreductases revealed the presence of the canonical YXXXK active site motif (Labesse *et al.*, 1994; Jörnvall *et al.*, 1995). For example, these residues have essential roles in catalysis by *Drosophila* alcohol dehydrogenase (Chen *et al.*, 1993). The tyrosine and lysine residues are part of a catalytic triad of residues, with the third residue being a serine coming from a different part of the protein (Fujimoto *et al.*, 2001). The tyrosine and serine are both directly involved in proton transfer and stabilization of the carbonyl group, whereas the lysine residue is thought to stabilize the nicotinamide moiety of the bound NAD(P)H coenzyme (Fujimoto *et al.*, 2001). The essential functionality of the tyrosine and lysine residues were confirmed by mutagenizing these residues in FAR5. The FAR5-Y238F is a conservative substitution and is not expected to disrupt the overall structure of FAR5, but the hydroxyl function normally present on the tyrosine is absent in the mutant. The lysine to isoleucine substitution (FAR5-K242I) is a less conserved substitution changing a positive charged amino acid to a neutral charged amino acid, but they are similarly sized amino acids. The detailed catalytic mechanism of FAR enzymes remains to be elucidated.

Significant amounts of certain fatty alcohols produced by expression of plant FARs was observed to be secreted into the yeast media, but it appears to be dependent on chain length. A high proportion of C16:0-C20:0 fatty alcohols were found in the media, a very low proportion of C22:0 fatty alcohol, and no C24:0 or C26:0 fatty alcohol. The length of the hydrocarbon chain may thus influence the efficiency of secretion. However,

it is possible that very-long-chain fatty alcohols tightly adhere to the outside of the yeast cell wall and are not easily washed off, or that the very-long-chain fatty alcohols precipitate easily out of solution and come down with the yeast pellet during centrifugation.

FAR2/MS2 has a very large 119 amino acid N-terminal extension relative to other FARs, which is predicted to contain a chloroplast target peptide sequence. It has recently been demonstrated that a FAR2/MS2 fusion to green fluorescent protein (GFP) is indeed targeted to plastids (Chen *et al.*, 2011). Full length FAR2 produced barely detectable amounts of alcohols when expressed in yeast. Since there are no chloroplasts in yeast, FAR2 is not targeted to the chloroplast and the chloroplast target sequence is not cleaved off. Targeting signals may inhibit activity of the enzyme. For example, the *E. coli* heat-labile enterotoxin B subunit (LT-B) protein only assembles into a functional form once the signal peptide is cleaved (Tauschek *et al.*, 2002). However, it was demonstrated that full-length FAR2/MS2 purified from bacteria is highly active for C16:0-ACP substrate (Chen *et al.*, 2011), indicating that the N-terminal extension is likely not inhibitory to enzyme activity. It is possible that the full-length FAR2/MS2 is non-specifically targeted to a subcellular location that does not allow it to access substrate. Some proteins have been reported to be dual-targeted to both plastids and mitochondria with the same N-terminal signal. For example, a *Chlamydomonas* chloroplast protein was targeted to mitochondria when expressed in yeast, and a yeast mitochondrial protein was targeted to both mitochondria and chloroplasts when expressed *in planta* (Hurt *et al.*, 1986; Huang *et al.*, 1990). It is possible that FAR2/MS2 targeted to yeast mitochondria is not active or not able to access substrate. Conversely, expression of FAR2ΔN in yeast produced large

quantities of C16:0-OH. This implies that FAR2ΔN is able to utilize C16:0-CoA, which is in relatively high abundance in the yeast cytosol (the pool of C16:0-ACP is small) (Domergue *et al.*, 2003). Loss-of-function mutation of *FAR2/MS2* results in male sterility and pollen grains with a dramatically altered exine layer (Aarts *et al.*, 1993, 1997; Chen *et al.*, 2011). It is likely that C16:0 fatty alcohols are generated by FAR2/MS2 in the tapetum and then deposited as part of sporopollenin. However, direct measurement of sporopollenin composition has not yet been reported due to its extreme resistance to depolymerization and it thus remains unproven that C16:0 fatty alcohols are in fact part of this extracellular lipid polymer.

Expression of full-length FAR6 in yeast generated negligible amount of fatty alcohols, whereas FAR6 lacking the N-terminal extension generated low, but detectable amounts of C16:0 fatty alcohol. Again, the N-terminal extension could be affecting FAR activity or localization when expressed in yeast. The amount of C16:0 fatty alcohol produced by FAR6ΔN was just above background and dramatically less than that produced by expression of FAR2ΔN in yeast. This low activity contrasts with results reported by Doan *et al.* (2011) who observed high levels of C16:0 fatty alcohol in yeast expressing an N-terminal truncated FAR6 enzyme. However, Doan *et al.* (2011) made a smaller N-terminal truncation of 47 amino acids at the predicted chloroplast targeting peptide cleavage site, whereas the N-terminal truncation in this study was 71 amino acids at the site next to the beginning of the conserved Rossmann fold domain. Amino acids important for enzyme activity were thus missing in the truncation analyzed in this study. Purified mature FAR6 (lacking the N-terminal 47 amino acids) is highly active for both C16:0-ACP and C16:0-CoA *in vitro* (Doan *et al.*, 2011). It is thus likely that FAR6 is

generating C16:0-OH *in planta*, although the role for such a chain length of fatty alcohol is unclear. I demonstrated that *FAR6* is expressed in stems (especially epidermal cells), root tips, silique receptacles, and anthers. However, C16:0 fatty alcohol is not a component of Arabidopsis cuticular waxes (C24:0-C30:0 fatty alcohols) (Rowland *et al.*, 2006) or suberin (C18:0-C22:0 fatty alcohols) (Domergue *et al.*, 2010) indicating that FAR6 has a function other than producing fatty alcohols for these extracellular barriers. It is possible that FAR6 produces C16:0 fatty alcohols that are incorporated into unknown components that have structural and protective properties. The base of the stem needs to be strong to keep the plant erect and sturdy, the root cap needs protection for emerging/growing roots to push through the soil/growth medium, and the base (receptacle) of the silique where floral abscission occurs needs reinforcement after this wounding event. A knock-out mutation of *FAR6* has been examined, but no major phenotype was apparent (Domergue and Rowland, unpublished observations). It is possible that FAR6 is partially redundant with FAR2/MS2 in generating the C16:0 fatty alcohols thought to be part of sporopollenin and this could be investigated using *far2 far6* double mutants. The *in planta* role of FAR6 requires further investigation.

CHAPTER 4: Amino Acids Important for Chain-Length Substrate Specificities of Arabidopsis FAR5 and FAR8

Statement of contribution:

I performed all the experiments and generated all the materials reported in this data chapter with the exception of the following:

FAR5 was previously cloned into pYES2-His6x/T7 by Frances Tran (Rowland Lab, Carleton University).

FAR8 was previously cloned into pYES2-His6x/T7 by Sollapura J. Vishwanath (Rowland Lab, Carleton University).

4.1 Introduction

Evolutionally related FARs have been identified in mammals, insects, microbes, and plants, which contribute to a variety of lipid biosynthetic pathways that are often important for host-environment interactions (see General Introduction). FAR enzymes have distinct substrate specificities with regard to chain length and degree of acyl-chain saturation (Chapters 1 and 3). The specificity of a FAR is often critical for the physical properties of the final chemical product of the biosynthetic pathway. For example, FAR substrate specificity is involved in speciation of some moths by dictating the regiochemistry of species-specific pheromones (Lassance *et al.*, 2010). In addition, fatty alcohol-containing compounds find many uses in human industrial applications (e.g. wax esters) and the chemistry of the fatty alcohol moiety is important for dictating the usefulness and thus value of the chemical product. Little is known about FAR enzymes with regard to the amino acids that are important for determining substrate specificity. A

three-dimensional structure has not yet been reported for a FAR and no other molecular approaches have yet been carried out to determine specific residues influencing chain length or chain saturation specificity.

FAR5 and *FAR8* belong to a family of genes encoding fatty acyl reductases (FAR) from Arabidopsis (Chapter 3). *FAR5* and *FAR8* are located in tandem on the Arabidopsis genome and encode proteins that are 85% identical. *FAR5* encodes a highly active FAR generating C18:0 primary fatty alcohol for root, seed coat, and wound-induced suberin (Domergue *et al.*, 2010). The function of *FAR8* is unknown, but it's near undetectable gene expression levels and the apparent low enzyme activity of the encoded protein (Doan *et al.*, 2009; Chapter 3) suggests that it may not have a major metabolic role in Arabidopsis. When expressed in yeast, FAR5 almost strictly produces C18:0-OH with a very small amount of C16:0-OH produced, and FAR8 strictly produces C16:0-OH although the amount of C16:0-OH produced by FAR8-expressing yeast is very low (Chapter 3). The strict and very distinct substrate specificities of FAR5 and FAR8 as well as the high sequence similarity made these two FARs a good platform to identify amino acids important for determining chain-length substrate specificity. Site-specific mutagenesis of FAR8 was first carried out to increase its activity for C16:0, while retaining strict chain-length specificity. Domain swaps and further site specific mutagenesis of FAR5 and FAR8 were carried out and a 72 amino acid region within these proteins was found to be important for dictating C16:0 versus C18:0 chain length specificity.

4.2 Materials and Methods: *See Chapter 2*

4.3 Arabidopsis FAR5 and FAR8 proteins

The amino acid alignment of FAR5 and FAR8 is shown in Figure 4.1. At the amino acid level, the proteins are 85% identical (424/496) and 90% similar (445/496) taking into account amino acids that have similar physiochemical properties (e.g. same charge, similar hydrophobicity). The putative active site motif YXXXK (Doan *et al.*, 2009; Chapter 3) is at amino acid positions 238-242 of the predicted proteins. For this study, swaps were made between the *FAR5* and *FAR8* coding sequences centred at four positions indicated by arrows in Figure 4.1. Site specific mutations were made at 6 positions in the FAR5 and/or FAR8 coding sequences, indicated by asterisks in Figure 4.1, which involved reciprocal exchanges of amino acids in the encoded FAR5 and FAR8 proteins.

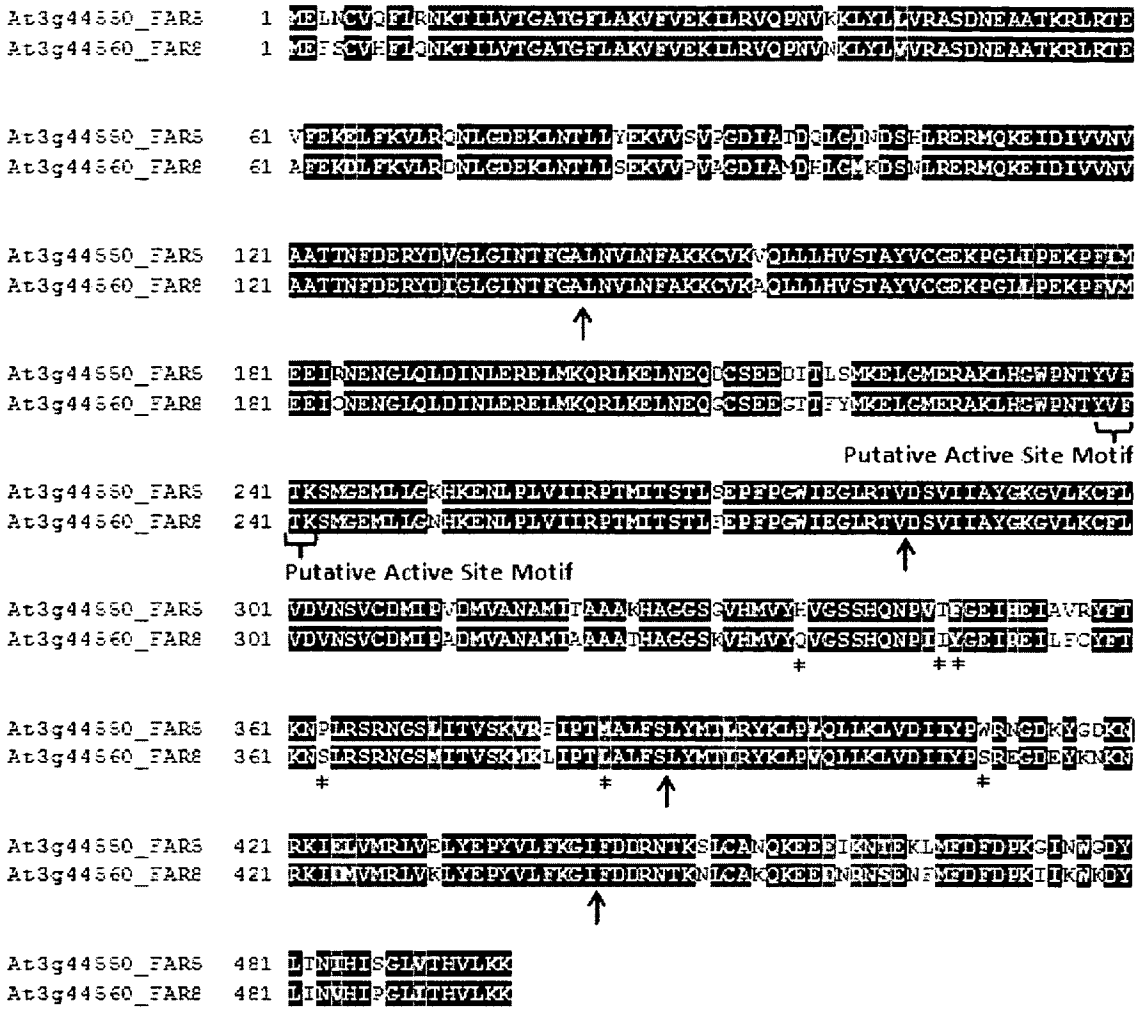


Figure 4.1 Alignment of Arabidopsis FAR5 and FAR8 protein sequences. Identical residues = black, physicochemical similar residues = grey. * = site-specific mutations, ↑ = domain swaps sites. The alignment was constructed using ClustalW2 and BOXSHADE 3.21Server). Full length coding sequences were obtained from the TAIR website for *FAR5* (At3g44550) and *FAR8* (At3g44560) and translated into protein sequences for alignment.

4.4 Expression of Arabidopsis FAR5 and FAR8 variants in Saccharomyces cerevisiae

The coding regions of Arabidopsis *FAR5* and *FAR8* were subcloned into yeast expression vector pYES2-His6x/T7 (see Chapter 3), which places them under the control of the *GAL1* promoter and the expressed proteins are tagged at the N-terminus with T7•tag epitope for detection by Western blot analysis. FAR5 protein accumulated to higher levels than FAR8 in yeast (Figure 4.2 C), and this was generally observed in the site-directed mutants and domain swaps (FAR5/FAR8 chimeras) with the level of accumulation correlating with the FAR chain-length specificity that the protein possessed (FAR5-like yielding predominately C18:0-OH with protein accumulating to higher levels versus FAR8-like yielding C16:0-OH with protein accumulating to higher levels) (Figure 4.2C-4.5 C). Lipid analysis by gas chromatography showed that FAR5 produced mostly C18:0-OH and some C16:0-OH, whereas FAR8 exclusively produced C16:0-OH albeit to very low levels (Figure 4.2 D), consistent with results reported in Chapter 3. The amount of fatty alcohols reported graphically in Figure 4.2D is the combination of fatty alcohols found in the yeast pellet (internal) and the media (secreted). The combined quantity of internal and secreted fatty alcohols is reported in all subsequent analyses described below (graphically in Figures 4.3-4.5 D). The separate internal and secreted quantities of fatty alcohols produced by yeast expressing the FAR variants are found in Tables 4.1 and 4.2, respectively, and the tabulated combination (internal + secreted) found in Table 4.3.

Arabidopsis *FAR5* and *FAR8* genes are located in tandem and are likely the result of a recent genome duplication event (Chapter 3). We speculated that Arabidopsis *FAR8* had recently acquired some mutations in its evolutionary history that would account for its very low activity. Inspection of the deduced FAR8 amino acid sequence in comparison

to other FAR protein sequences revealed two candidate residues that may be affecting FAR8 protein activity and/or stability. FAR8 contains a serine (S) at position 363, which is a proline (P) at the equivalent position in all other active FARs, including FAR proteins from Arabidopsis, wheat, human, mouse, and silkworm (see Figure 1.4 of General Introduction and Figure 3.3 of Chapter 3). FAR8 contains a glutamine (Q) at position 337, which is a histidine (H) in most FARs including FAR5 although not in Arabidopsis FAR2 and FAR6, human FAR1, or mouse FAR1 (less conserved than the proline at position 363). Site specific substitutions were thus made in the *FAR8* coding sequence converting the codons at these two positions to the codons present in FAR5. These variants were called FAR8-S363P and FAR5-Q337H (Figure 4.2 A). The S363P substitution significantly altered fatty alcohol production by FAR8 when expressed in yeast, causing it to produce substantial amounts of C16:0-OH, comparable to the amount of C18:0-OH produced by expression of FAR5 in yeast (Figure 4.2 D). The substrate specificity of FAR8-S363P remained strict for C16:0 acyl chain length. The Q337H substitution in FAR8 did not significantly alter fatty alcohol production when expressed in yeast. The double substitution FAR8-Q337H S363P did result, however, in more C16:0-OH being produced than the single FAR8-S363P mutation alone (Figure 4.2 D). Based on a two-tailed t-test for the triplicate samples of FAR8-S363P and FAR8-Q337H S363P, it was found that the differences were statistically significant (p-value = 0.015). The reciprocal substitutions were made in FAR5 to give FAR5-P363S and FAR5-H337Q (Figure 4.2 A). It was expected that FAR5-P363S would be inactive if the proline residue at this position is critical for FAR enzyme activity, but surprisingly the FAR5-P363S variant was active in making C18:0-OH (Figure 4.2 D). Also, it was unexpectedly

observed that the FAR5-H337Q variant produced about five times more C18:0-OH in yeast than wild-type FAR5 (Figure 4.2 D). All FAR5 and FAR8 variants were expressed in yeast as detected by Western blot analysis with FAR5 accumulating to moderately higher levels than the other proteins (Figure 4.2 C). About equal amounts of protein extract were present in each of the lanes (Figure 4.2 B).

A silent mutation was introduced into the *FAR8* coding sequence that created a *Sal*I restriction site, which is present at the equivalent position in the *FAR5* coding sequence. This common restriction site facilitated the making of domain swaps centred at amino acid position 284 (Figure 4.1). Expression of the original FAR8 and FAR8-*Sal*I in yeast produced similar results, although both are not very active (compare Figure 4.2 D with original FAR8 and Figures 4.3 D and 4.4 D with FAR8-*Sal*I). Versions of FAR8-Q337H, FAR8-S363P and FAR8-Q337H S363P were also generated with the *Sal*I restriction site in the coding region, and these gave similar results as versions lacking *Sal*I (compare Figure 4.2 D and Figure 4.4 D). This indicates that the introduced *Sal*I restriction site does not affect the translation of *FAR8* and thus not the amount of fatty alcohol product produced upon expression in yeast.

Four paired sets of domain swaps were made between *FAR5* and *FAR8* at positions corresponding to amino acid residues 141, 284, 387, and 442 (Figure 4.3 A). The domain swap at position 141 was generated using a common *Sac*I restriction site present at the equivalent position in the *FAR5* and *FAR8* coding sequences and was done using FAR8-*Sal*I template (contains no amino acid substitutions). The domain swap at position 284 was generated using a common *Sal*I restriction site present at the equivalent position in the *FAR5* and *FAR8-Sal*I coding sequences (contains no amino acid

substitutions). The domain swaps made at positions 387 and 442 were made using overlap-PCR and were done using FAR8-Q337H S363P *SalI* template (contains amino acid substitutions that significantly increased the amount of C16:0-OH produced by FAR8 in yeast). These domain swaps were made in order to narrow down the amino acid important for chain-length substrate specificity. The domain swaps $FAR5_{1\text{-}141}FAR8_{142\text{-}496}$ and $FAR8_{1\text{-}141}FAR5_{142\text{-}496}$ (centered at position 141) produced predominately C16:0-OH and C18:0-OH, respectively (Figure 4.3 D). The amount of C16:0-OH produced by $FAR5_{1\text{-}141}FAR8_{142\text{-}496}$ was low, similar to that of wild-type FAR8, and both have the non-conserved serine at position 363 rather than the typical proline. The domain swaps $FAR5_{1\text{-}284}FAR8_{285\text{-}496}$ and $FAR8_{1\text{-}284}FAR5_{285\text{-}496}$ (centered at residue 284) produced predominately C16:0-OH and C18:0-OH, respectively (Figure 4.3 D). Unexpectedly, $FAR5_{1\text{-}284}FAR8_{285\text{-}496}$ produced substantial amounts of C16:0-OH, despite having the non-conserved serine at position 363. Nonetheless, the analysis of this pair of domain swaps indicated that C16/C18 chain-length specificity is conferred by residues C-terminal of position 284 of the FAR proteins. Further domain swaps were tested centred at amino acid positions 387 and 442. The chimeras $FAR5_{1\text{-}387}FAR8_{388\text{-}496}$ and $FAR5_{1\text{-}442}FAR8_{443\text{-}496}$ (FAR5 is N-terminal in both) each produced C18:0-OH when expressed in yeast, whereas FAR8-Q337H $S363P_{1\text{-}387}FAR5_{388\text{-}496}$ and FAR8-Q337H $S363P_{1\text{-}442}FAR5_{443\text{-}496}$ (FAR8 is N-terminal in both) each produced very little fatty alcohol but it was C16:0-OH (Figure 4.3 D including zoomed inset). These domain swaps indicated that C16/C18 chain-length specificity is conferred by residues N-terminal of position 387 of the FAR proteins. The tagged proteins from all FAR5/FAR8 chimeras were detectable by Western blot analysis, although at variable levels (Figure 4.3 C). However, the total amounts of

protein loaded in each lane were also variable (Figure 4.3 B), but this did not correlate with FAR protein signal detected by Western blot. It proved difficult to achieve consistent signals in the Westerns, but it should be noted that the lipid extractions were performed after 4 days of induction allowing for C16:0 or C18:0 fatty alcohols to accumulate internally and externally over a long period.

As described above, it was unexpected that $FAR5_{1-284}FAR8_{285-496}$ generated such high amounts of C16:0-OH since it contained the non-conserved serine at position 363. In the FAR8 protein, this amino acid needed to be changed to proline to achieve high levels of C16:0 fatty alcohol production in yeast (Figure 4.2). Additional $FAR5_{1-284}FAR8_{285-496}$ chimeras were tested that contained substitution(s) at positions 363 (S363P) and/or 337 (Q337H) to test if there was an influence on C16:0-OH production. All three chimeras, $FAR5_{1-284}$FAR8-Q337H$_{285-496}$, $FAR5_{1-284}$FAR8-S363P$_{285-496}$, and $FAR5_{1-284}$FAR8-Q337H S363P$_{285-496}$ produced similar amounts of C16:0-OH (~1 ug internal + secreted C16:0-OH produced per OD unit) as $FAR5_{1-284}FAR_{285-496}$ containing no substitutions (compare Figure 4.3 D and Figure 4.4 D). The amount of C16:0-OH produced by the various $FAR5_{1-284}FAR8_{285-496}$ chimeras (no mutation, S363P, Q337H, and S363P + Q337H) was about five times higher than that produced by FAR8-S363P and three times higher than FAR8-Q337H S363P (Figure 4.4 D). The presence of FAR5 at the N-terminus thus influenced the amount of C16:0-OH produced without affecting chain-length specificity as well as influenced the effect of having a serine at position 363. Western blot analysis indicated that these chimeras were expressed at about similar levels (Figure 4.4 B and C).

Characterization of the FAR5/FAR8 domain swaps described above (Figure 4.3) indicated that C16/C18 chain length specificity of FAR5/FAR8 was dictated by a region

bound by amino acids 285 and 387. Taking into account the high amino acid identity between FAR5 and FAR8 (Figure 4.1), the region of interest was a 72 amino acid region between residues 312 and 383 (the amino acid sequences of FAR5 and FAR8 are identical between residues 285 and 311 as well as 384 and 387). There are 18 amino acid differences between FAR5 and FAR8 in this region. Four site-specific substitutions were made using the FAR8-Q337H S363P (*Sal*I) construct as template: FAR8 Q337H **I347T** S363P, FAR8 Q337H **Y348F** S363P, FAR8 Q337H S363P **L383M**, and FAR8 Q337H S363P **S410W** (Figure 4.5 A). Three of the amino acid substitutions, I347T, Y348F, and L383M, were in the 72 amino acid region indicated to be very important for C16/C18 chain length specificity from the domain swap analysis, and one S410W was just C-terminal to this region. These residues were found to be commonly variable in FARs of varying chain-length specificities (Figure 3.3 of Chapter 3) and therefore perhaps influencing the ability to bind a particular fatty acyl chain length for reduction to fatty alcohol. All four substitutions, however, produced C16:0-OH when expressed in yeast similar to that of the parent FAR8-Q337H S363P protein (Figure 4.5 D, Table 4.2-4.4). All of these FAR8 variants were expressed at similar levels in yeast (Figure 4.5 B and C).

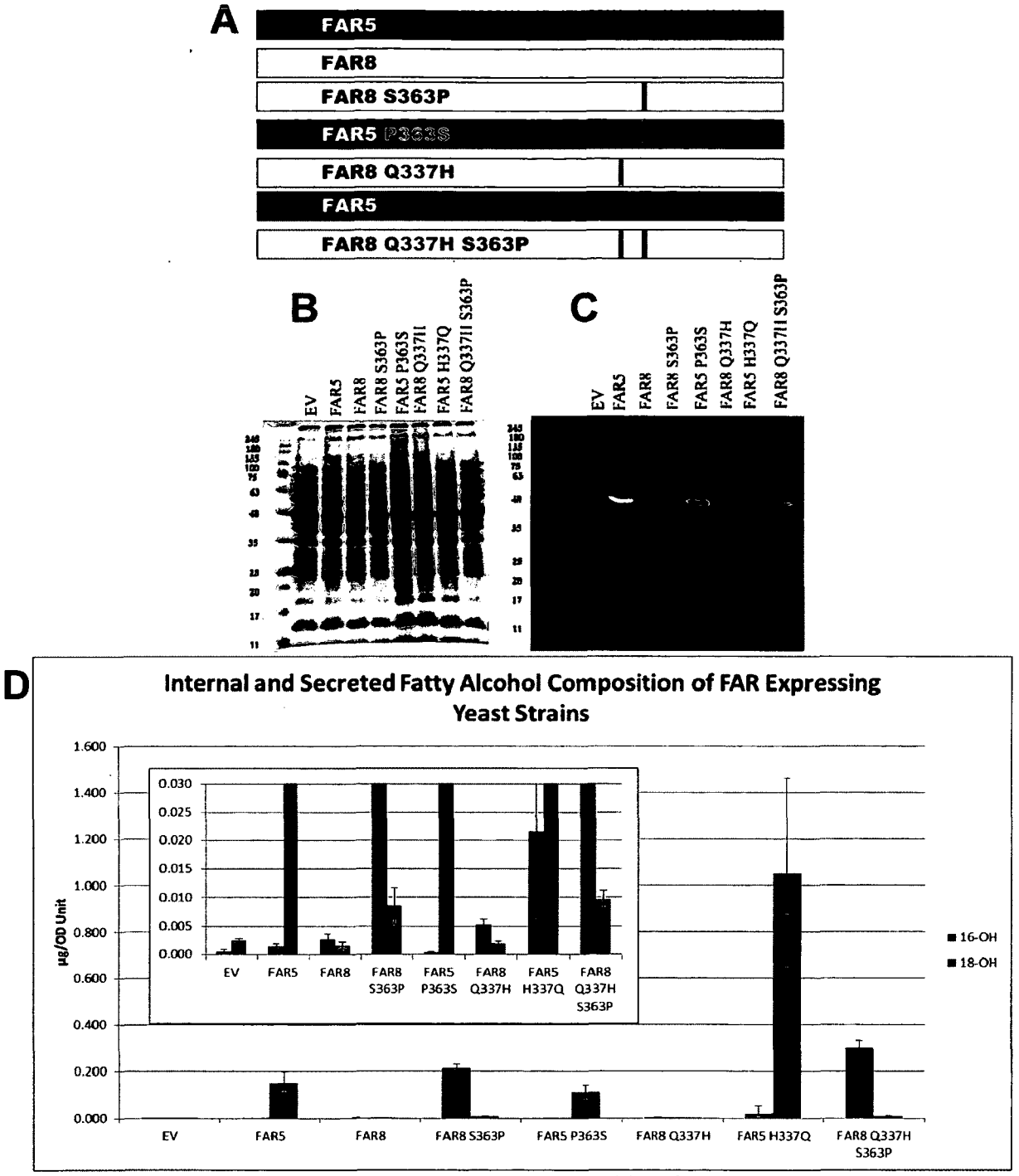
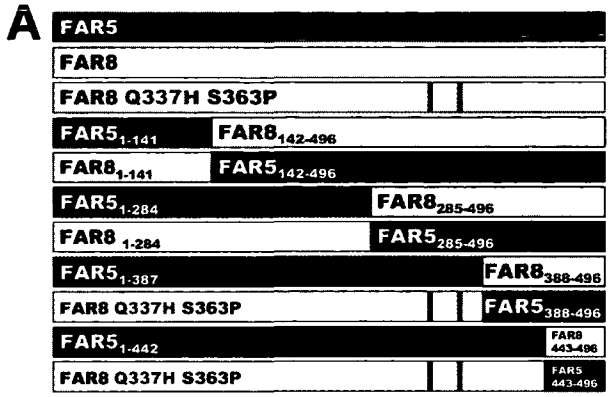
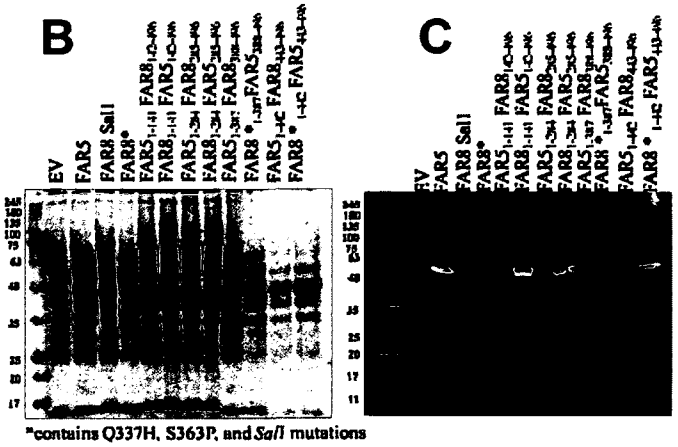


Figure 4.2 Influence of reciprocal amino substitutions at positions 337 and 363 in Arabidopsis FAR5 and FAR8 expressed in yeast. A. One-dimensional schematics of FAR5, FAR8, FAR8-S363P, FAR5-P363S, FAR8-Q337H, FAR5-H337Q, and FAR8-S363P Q337H. B. SDS-PAGE gel stained with Coomassie Brilliant Blue. C. Western blot analysis of proteins expressed in yeast, detected with T7 Tag monoclonal mouse primary antibody and horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibody. D. Total fatty alcohols in lipids produced by yeast expressing FAR5/FAR8 variants measured by gas chromatography. The means were calculated from 3 independent yeast cultures. Values are expressed in $\mu g/OD_{600}$ Unit (error bars are ± standard deviation).





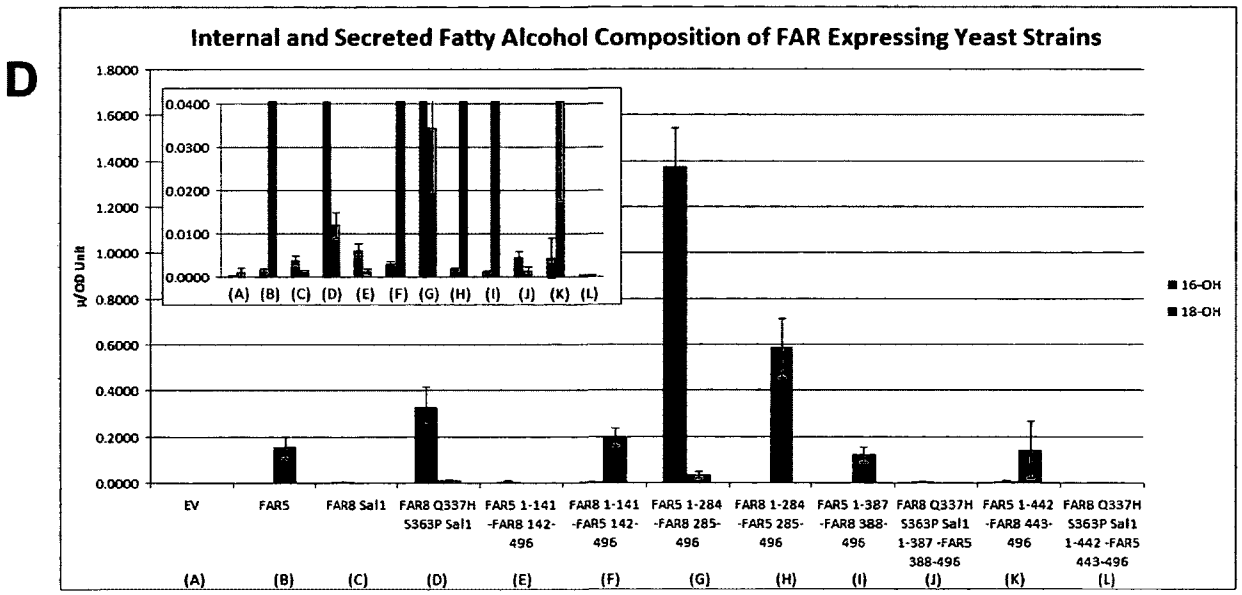
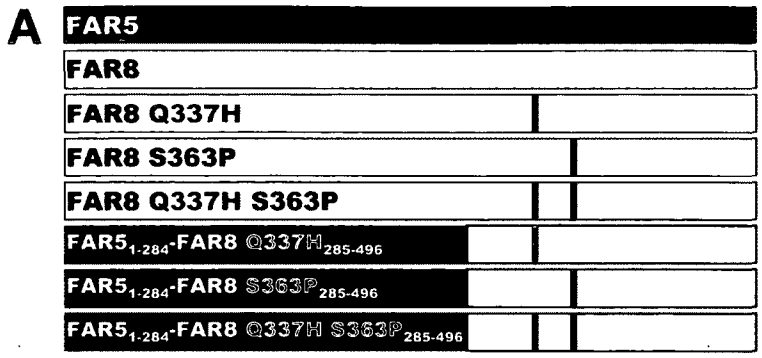
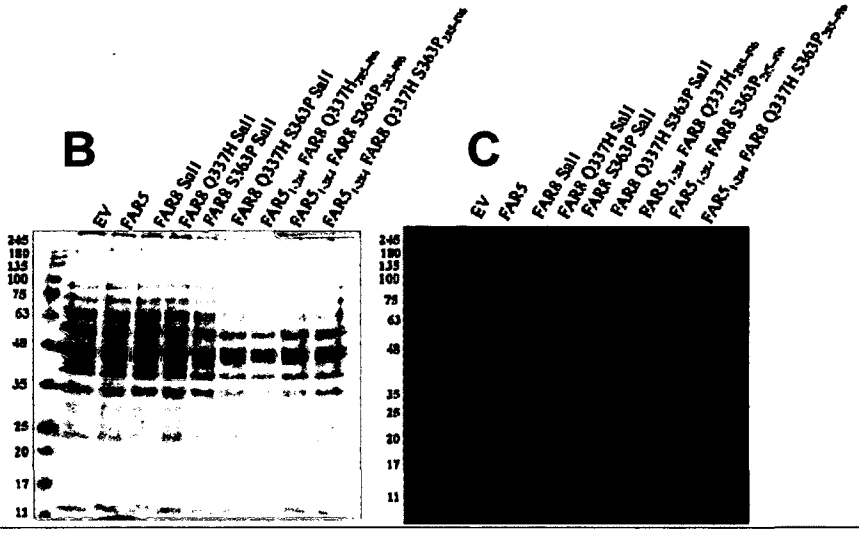


Figure 4.3 Expression of Arabidopsis FAR5/FAR8 domain swaps in yeast. A. One-dimensional schematics of FAR5, FAR8 (*Sal*I), FAR8-Q337H S363P, $FAR5_{1-141}FAR8_{142-496}$, $FAR8_{1-141}FAR5_{142-496}$, $FAR5_{1-284}FAR8_{285-496}$, $FAR8_{1-284}FAR5_{285-496}$, $FAR5_{1-387}FAR8_{388-496}$ $FAR8_{1-387}FAR5_{388-496}$, and $FAR5_{1-442}FAR8_{443-496}$. B. SDS-PAGE gel stained with Coomassie Brilliant Blue. C. Western blot analysis of proteins expressed in yeast, detected with T7 Tag monoclonal mouse primary antibody and horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibody. D. Total fatty alcohols in lipids produced by yeast expressing FAR5/FAR8 variants measured by gas chromatography. The means were calculated from 3 independent yeast cultures. Values are expressed in $\mu g/OD_{600}$ Unit (error bars are ± standard deviation).



*all FAR8s contain *Sal*I site



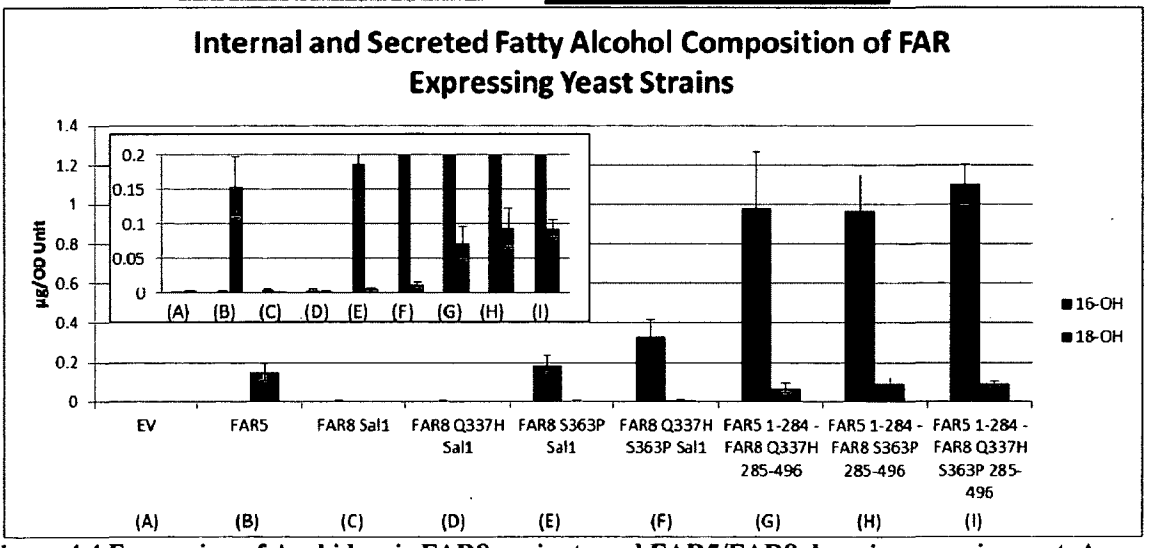


Figure 4.4 Expression of Arabidopsis FAR8 variants and FAR5/FAR8 domain swaps in yeast. A. One-dimensional schematics of FAR5, FAR8 (*Sal*I), FAR8-Q337H *Sal*I, FAR8-S363P *Sal*I, FAR8-Q337H S363P *Sal*I, FAR5$_{1-284}$FAR8 Q337H$_{285-496}$, FAR51.$_{284}$FAR8 S363P$_{285-496}$, and FAR5$_{1-284}$FAR8 Q337H S363P$_{285-496}$. B. SDS-PAGE gel stained with Coomassie Brilliant Blue. C. Western blot analysis of proteins expressed in yeast, detected with T7 Tag monoclonal mouse primary antibody and horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibody. D. Total fatty alcohols in lipids produced by yeast expressing FAR5/FAR8 variants measured by gas chromatography. The means were calculated from 3 independent yeast cultures. Values are expressed in µg/OD$_{600}$ Unit (error bars are ± standard deviation).

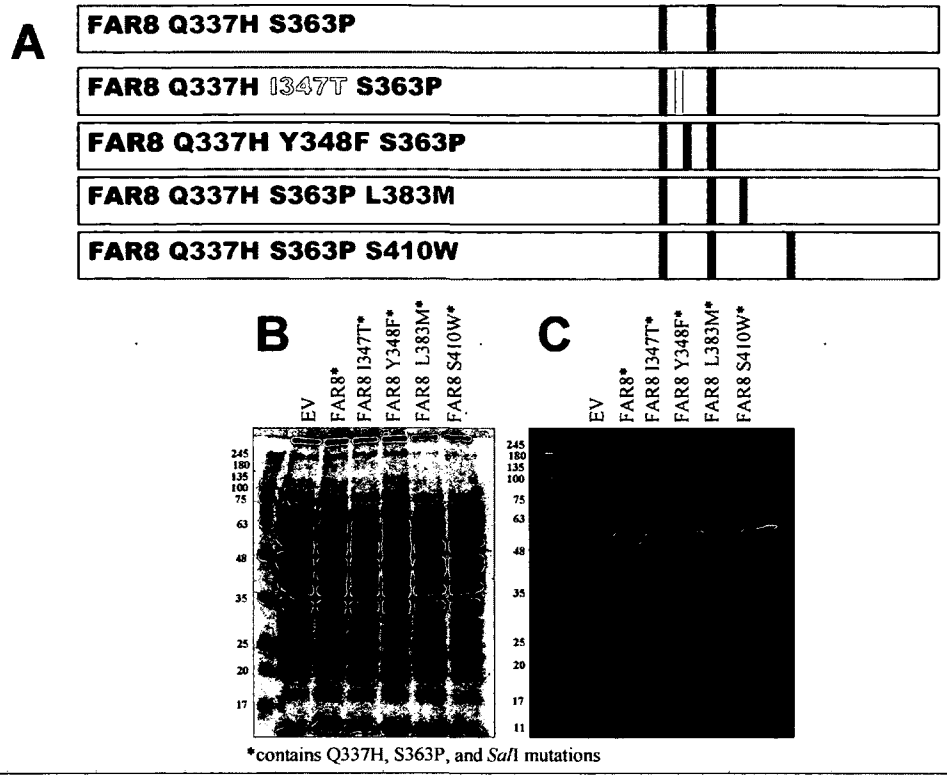


Figure 4.5 Expression of Arabidopsis FAR8 variants containing site-specific substitutions in yeast. A. One-dimensional schematics of FAR8-S363P Q337H *Sal*I, FAR8-S363P I347T Q337H *Sal*I, FAR8-S363P Y348F Q337H *Sal*I, FAR8-S363P Q337H L383M *Sal*I, and FAR8-S363P Q337H S410W *Sal*I. B. SDS-PAGE gel stained with Coomassie Brilliant Blue. C. Western blot analysis of proteins expressed in yeast, detected with T7 Tag monoclonal mouse primary antibody and horseradish peroxidase (HRP)-conjugated anti-mouse secondary antibody. D. Total fatty alcohols in lipids produced by yeast expressing FAR5/FAR8 variants measured by gas chromatography. The means were calculated from 3 independent yeast cultures. Values are expressed in µg/OD_{600} Unit (error bars are ± standard deviation).

Table 4.1 Internal Fatty Alcohol Composition of the pellet of transgenic yeast strains expressing Arabidopsis FAR5 and FAR8 site directed mutants and domain swaps (Error is ± standard deviation (SD))

	C16:0-OH ($\mu g/OD_{600}$ ±SD)	**C18:0-OH ($\mu g/OD_{600}$ ±SD)**
FAR5	0.0011±0.0003	0.0815±0.0059
FAR8	0.0027±0.0009	0.0011±0.0009
FAR5 Y238F	0.0002±0.0003	0.0006±0.0005
FAR5 K242I	0.0001±0.0001	0.0004±0.0001
FAR8 S363P	0.2094±0.0146	0.0054±0.0004
FAR5 P363S	0.0003±0.0000	0.0586±0.0009
FAR8 Q337H	0.0052±0.0009	0.0002±0.0001
FAR5 H337Q	0.0010±0.0008	0.6544±0.5529
FAR8 Q337H S363P	0.2692±0.0302	0.0064±0.0011
FAR8 *Sal*I	0.0037±0.0010	0.0012±0.0001
FAR8 Q337H *Sal*I	0.0029±0.0015	0.0008±0.0009
FAR8 S363P *Sal*I	0.1807±0.0498	0.0043±0.0009
FAR8 Q337H S363P *Sal*I	0.2929±0.0919	0.0085±0.0036
FAR8 Q337H I347T S363P *Sal*I	0.1736±0.0288	0.0058±0.0010
FAR8 Q337H Y348F S363P *Sal*I	0.0847±0.0127	0.0012±0.0003
FAR8 Q337H S363P L383M *Sal*I	0.1238±0.0254	0.0027±0.0007
FAR8 Q337H S363P S410W *Sal*I	0.1269±0.0284	0.0028±0.0010
$FAR5_{1-141}$-$FAR8_{142-496}$	0.0061±0.0015	0.0008±0. 0003
$FAR8_{1-141}$-$FAR5_{142-496}$	0.0026±0.0007	0.1712±0.0439
$FAR5_{1-284}$-$FAR8_{285-496}$	1.3330±0.2316	0.0289±0.0063
$FAR8_{1-284}$-$FAR5_{285-496}$	0.0017±0.0003	0.5368±0.1456
$FAR5_{1-284}$-FAR8 Q337H $_{285-496}$	0.8486±0.2400	0.0462±0.0149
$FAR5_{1-284}$-FAR8 $S363P_{285-496}$	0.6825±0.1324	0.0428±0.0122
$FAR5_{1-284}$-FAR8 Q337H $S363P_{285-496}$	0.9074±0.1361	0.0602±0.0102
$FAR5_{1-387}$-$FAR8_{388-496}$	0.0009±0.0001	0.0869±0.0188
FAR8 Q337H S363P *Sal*I_{1-387}-$FAR5_{388-496}$	0.0042±0.0013	0.0002±0.0001
$FAR5_{1-442}$-$FAR8_{443-496}$	0.0008±0.0007	0.0550±0.0475
FAR8 Q337H S363P *Sal*I_{1-442}-$FAR5_{443-496}$	0.0003±0.0001	0.0001±0.0001

Table 4.2 Secreted fatty alcohol composition of the supernatant of transgenic yeast strains expressing Arabidopsis FAR5 and FAR8 site directed mutants and domain swaps. (Error is ± standard deviation (SD))

	C16:0-OH ($\mu g/OD_{600}$ ±SD)	**C18:0-OH ($\mu g/OD_{600}$ ±SD)**
FAR5	0.0003±0.0002	0.0713±0.0435
FAR8	0.0001±0.0001	0.0005±0.0006
FAR5 Y238F	0.0007±0.0002	0.0005±0.0002
FAR5 K242I	0.0002±0.0002	0.0003±0.0001
FAR8 S363P	0.0075±0.0031	0.0031±0.0030
FAR5 P363S	0.0001±0.0001	0.0510±0.0318
FAR8 Q337H	0.0001±0.0002	0.0017±0.0005
FAR5 H337Q	0.0207±0.0295	0.3982±0.2908
FAR8 Q337H S363P	0.0327±0.0038	0.0032±0.0008
FAR8 *SalI*	0.0001±0.0001	0.0004±0.0001
FAR8 Q337H *SalI*	0.0005±0.0002	0.0014±0.0005
FAR8 S363P *SalI*	0.0061±0.0022	0.0010±0.0002
FAR8 Q337H S363P *SalI*	0.0355±0.0061	0.0035±0.0008
FAR8 Q337H I347T S363P *SalI*	0.0352±0.0045	0.0018±0.0002
FAR8 Q337H Y348F S363P *SalI*	0.0199±0.0023	0.0009±0.0002
FAR8 Q337H S363P L383M *SalI*	0.0501±0.0337	0.0027±0.0022
FAR8 Q337H S363P S410W *SalI*	0.0622±0.0418	0.0031±0.0025
$FAR5_{1-141}$-$FAR8_{142-496}$	0.0001±0.0001	0.0005±0.0004
$FAR8_{1-141}$-$FAR5_{142-496}$	0.0005±0.0002	0.0273±0.0061
$FAR5_{1-284}$-$FAR8_{285-496}$	0.0424±0.0694	0.0053±0.0088
$FAR8_{1-284}$-$FAR5_{285-496}$	0.0001±0.0001	0.0477±0.0214
$FAR5_{1-284}$-FAR8 $Q337H_{285-496}$	0.1347±0.0489	0.0248±0.0110
$FAR5_{1-284}$-FAR8 $S363P_{285-496}$	0.2831±0.0542	0.0509±0.0163
$FAR5_{1-284}$-FAR8 Q337H $S363P_{285-496}$	0.1968±0.0667	0.0322±0.0146
$FAR5_{1-387}$-$FAR8_{388-496}$	0.0002±0.0001	0.0351±0.0145
FAR8 Q337H S363P $SalI_{1-387}$-$FAR5_{388-496}$	0.0003±0.0002	0.0012±0.0009
$FAR5_{1-442}$-$FAR8_{443-496}$	0.0036±0.0052	0.0867±0.0777
FAR8 Q337H S363P $SalI_{1-442}$-$FAR5_{443-496}$	0.0001±0.0001	0.0003±0.0001

Table 4.3 Fatty alcohol composition of the combined pellet and supernatant fractions, of transgenic yeast strains expressing Arabidopsis FAR5 and FAR8 site directed mutants and domain swaps. (Error is ± standard deviation (SD))

	C16:0-OH ($\mu g/OD_{600}$ ±SD)	**C18:0-OH ($\mu g/OD_{600}$ ±SD)**
FAR5	0.0014±0.0005	0.1528±0.0447
FAR8	0.0027±0.0009	0.0016±0.0007
FAR5 Y238F	0.0009±0.0001	0.0011±0.0003
FAR5 K242I	0.0004±0.0001	0.0007±0.0001
FAR8 S363P	0.2168±0.0158	0.0086±0.0032
FAR5 P363S	0.0005±0.0001	0.1097±0.0322
FAR8 Q337H	0.0053±0.0008	0.0019±0.0004
FAR5 H337Q	0.0217±0.0289	1.0526±0.4097
FAR8 Q337H S363P	0.3019±0.0327	0.0097±0.0016
FAR8 *SalI*	0.0037±0.0010	0.0016±0.0001
FAR8 Q337H *SalI*	0.0033±0.0015	0.0022±0.0004
FAR8 S363P *SalI*	0.1868±0.0516	0.0053±0.0011
FAR8 Q337H S363P *SalI*	0.3284±0.0883	0.0120±0.0030
FAR8 Q337H I347T S363P *SalI*	0.2088±0.0309	0.0076±0.0012
FAR8 Q337H Y348F S363P *SalI*	0.1046±0.0148	0.0021±0.0003
FAR8 Q337H S363P L383M *SalI*	0.1740±0.0386	0.0054±0.0023
FAR8 Q337H S363P S410W *SalI*	0.1891±0.0403	0.0059±0.0024
$FAR5_{1-141}$-$FAR8_{142-496}$	0.0063±0.0015	0.0013±0.0005
$FAR8_{1-141}$-$FAR5_{142-496}$	0.0031±0.0005	0.1985±0.0379
$FAR5_{1-284}$-$FAR8_{285-496}$	1.3754±0.1700	0.0343±0.0150
$FAR8_{1-284}$-$FAR5_{285-496}$	0.0018±0.0002	0.5844±0.1280
$FAR5_{1-284}$-FAR8 Q337H $_{285-496}$	0.9833±0.2867	0.0711±0.0242
$FAR5_{1-284}$-FAR8 $S363P_{285-496}$	0.9655±0.1834	0.0938±0.0285
$FAR5_{1-284}$-FAR8 Q337H $S363P_{285-496}$	1.1042±0.0971	0.0924±0.0131
$FAR5_{1-387}$-$FAR8_{388-496}$	0.0011±0.0002	0.1219±0.0330
FAR8 Q337H S363P $SalI_{1-387}$-$FAR5_{388-496}$	0.0045±0.0013	0.0013±0.0009
$FAR5_{1-442}$-$FAR8_{443-496}$	0.0044±0.0046	0.1417±0.1240
FAR8 Q337H S363P $SalI_{1-442}$-$FAR5_{443-496}$	0.0004±0.0001	0.0004±0.0002

4.5 Discussion

Expression of Arabidopsis FAR5 in yeast produced relatively large amounts of C18:0-OH, accumulating both within the yeast cells and in the yeast media, as well as small amounts of C16:0-OH (100 times less than C18:0-OH). Expression of Arabidopsis FAR8 in yeast did not produce any detectable C18:0-OH, but there were small amounts of C16:0-OH generated and slightly more than that produced by FAR5. Substitution of the serine residue at amino acid position 363 with proline "resurrected" the FAR8 enzyme and substantial amounts of C16:0-OH were generated in transgenic yeast with no C18:0-OH detected. A proline is present at the same relative position in all other biochemically active FARs characterized to date and is nearly strictly conserved in all FAR enzymes including non-plant FARs (Arabidopsis FAR8 being the exception). We hypothesized that the proline at this position has a key structural role in FAR enzymes. Proline is unique among the 20 common amino acids because its side chain is connected to the protein backbone twice, forming a five-membered nitrogen-containing ring. Proline is commonly found in tight turns within proteins and such a turn of the backbone may be important for the structural integrity of FAR proteins. A substitution of a proline with a serine would likely eliminate any such turn. It was thus anticipated that changing the proline at position 363 to serine in FAR5 would severely diminish its activity, but this was not the case. It remained highly active for C18:0 chains. A proline at this position may be important for some but not all FARs. The impact of residue 337 was also tested by substituting the glutamine residue with a histidine residue in FAR8. It is a histidine at this position in FAR5 as well as in most other FARs. The FAR8-Q337H variant did not produce any more C16:0-OH than wild-type (i.e. unmutagenized) FAR8, but combining

this substitution with the S363P mutation further enhanced the activity of FAR8 without affecting the specificity. The reciprocal substitution was analyzed in FAR5 (H337Q), and again it was surprising to observe that activity was not diminished but in fact greatly increased. It is difficult to rationalize these results without the aid of three-dimensional information. An X-ray or NMR structure of a FAR has not yet been reported. Nonetheless, the highly active FAR8 variants, which are strictly specific for C16:0 acyl chain, are useful for domain swap and mutagenesis studies in combination with the closely related FAR5 enzyme, which is almost strictly specific for C18:0 acyl chain. Other FARs have broader and overlapping specificities (see Chapter 3), which makes similar analyses more complicated to interpret. Also, C16:0-CoA and C18:0-CoA are relatively easy to chemically synthesize and are commercially available, even in radioactive form, which would make follow-up *in vitro* kinetic assays very feasible compared to FARs with longer chain length specificities.

It should be noted that the *in planta* function and substrate specificity of FAR8 is unknown. We speculate that it is derived from a pseudogene due to its near undetectable gene expression and low apparent activity in *E. coli* and yeast cells (Doan *et al.*, 2009; this study). It is possible, however, that the biologically relevant substrate is not available in these microbes. Other fatty acids could be fed to yeast to test whether FAR8 can use other substrates. However, there are limitations to feeding studies as unusual fatty acids may be rapidly metabolized by the yeast. Also, no fatty alcohols other than saturated even chain-lengths between C16:0 and C30:0 have been found to occur in Arabidopsis to date (Rowland *et al.*, 2006; Domergue *et al.*, 2010; Doan *et al.*, 2011). However, this issue is not really relevant to the goals of this study, which was focussed on determining regions

and specific amino acids of FAR5 and FAR8 that confer C16:0 versus C18:0 chain length specificity.

The information gathered from the FAR5/FAR8 domain swaps narrowed down an important 72 amino acid region between residues 312 and 383 that significantly influences C16/C18 chain-length substrate specificity. This region includes part of the NADB Rossmann fold superfamily domain and part of the region before the annotated FAR_C domain. The Rossmann fold domains of FAR5 and FAR8 are annotated to be between residues 12 and 363 and the FAR_C domains of FAR5 and FAR8 are annotated to be between residues 420 and 494 (Table 1.1). It is possible that the region between the two domains is what controls chain length substrate specificity in FARs and this is a region that is moderately less conserved between plants FARs relative to other regions of the proteins (Figure 3.3 of Chapter 3). Internal domain swaps between FAR5 and FAR8 should be the subject of future studies, exchanging only the central 312 to 383 region of the proteins and assessing for C16:0 versus C18:0 activity. It is possible, however, that amino acids outside this central region can influence chain length specificity as interactions between amino acid residues in the folded up protein likely dictates the precise ability to bind a C16 or a C18 acyl-CoA chain.

Site directed substitutions focusing on the 312 to 383 region of FAR5 and FAR8 can be used to further identify individual amino acids that have an important role in C16/C18 chain-length specificity. There are only 18 amino acid differences between FAR5 and FAR8 in this region. In this study, three site-directed substitutions were made in the 72 amino acid region of FAR8 as well as one slightly C-terminal of this region. All the substitutions involved converting the respective FAR8 residue to a FAR5 residue. All

retained FAR8-like strict C16:0 acyl chain specificity and are thus likely not playing a major role in dictating chain-length specificity. However, it is possible that multiple substitutions are needed to be made in order to "flip" C16/C18 chain length specificity. The remaining 13 amino acids that differ between FAR5 and FAR8 in the 312 to 383 region should be the focus of future investigations. There may also be FAR5/FAR8 variants that no longer have such strict chain-length specificity and produce substantial amounts of both C16:0-OH and C18:0-OH. There are examples of FARs with broader chain length specificities such as FAR4, which produces C18:0-C22:0 fatty alcohols when expressed in yeast (Domergue *et al.*, 2010; and Chapter 3). This study represents an important step forward into characterizing the chain-length substrate binding site of the FAR enzymes.

CHAPTER 5: General Discussion

FARs are responsible for an NAD(P)H-dependent reaction that reduces fatty acyl-CoAs or fatty acyl-ACPs to primary fatty alcohols via an unreleased fatty aldehyde intermediate. Evolutionarily related FARs have been cloned and characterized from mammals (mice, humans), birds (ducks), insects (honeybees, moths), crustaceans (copepods), microorganisms (phytoflagellates), and plants (Arabidopsis, wheat, and Jojoba) (see General Introduction). Primary fatty alcohols and combined derivatives (e.g. wax esters, alkyl hydroxycinnamates, and ether lipids) are utilized in a variety of biological roles. For example, free and/or combined fatty alcohols serve as an energy storage reserve in some microorganisms, as antioxidants and membrane fluidity regulators in mammalian neuronal cells (plasmalogens), as sex pheromones in some insect species, as buoyancy regulators of certain marine organisms, and as important components of protective extracellular lipid barriers in plants (see General Introduction). Fatty alcohol-containing compounds, especially wax esters, are also used in a variety of human applications, such as in cleaning detergents (e.g. fatty alcohol ethoxylates), cosmetic and pharmaceutical formulations, food products (typically acting as thickening agents), textiles, coatings, and high-performance industrial lubricants (Houston, 1984; Kajdas, 1987; HEAR, 2009; Teerawanichpan and Qiu, 2010[b]; Rowland and Domergue, 2012). The current capacity of natural wax production is limited by high production costs and cannot keep up with demand (Jetter and Kunst, 2008). Information pertaining to FAR substrate specificity will be important for engineering FAR enzymes with desired specificities that could enable the production of high value fatty alcohols and wax esters in an industrial oilseed crop or transgenic microbe.

There is currently little information regarding what amino acid residues determine the substrate specificities of FAR enzymes, both with regard to chain length and degree of acyl-chain saturation. The Arabidopsis FAR enzyme family, which have chain-length specificities ranging from C16:0 to C26:0 when expressed in yeast (Chapter 3), provides an excellent platform for identifying protein domains and individual amino acids that are important for catalytic activity as well as substrate specificity. For example, primary amino acid sequence alignments of the Arabidopsis FAR family revealed that there is a high degree of amino acid identity between FAR1 and FAR4 (74%) and FAR5 and FAR8 (85%), but these pairs of enzymes have distinct chain-length substrate specificities (Chapter 3). These high similarities between FARs allow for domain-swapping experiments to be used to identify regions involved in substrate specificity and then site-directed mutagenesis experiments can be performed to further scrutinize these regions, revealing the individual amino acid residues that are responsible. I successfully used such an approach to gain insights into regions important for C16:0 versus C18:0 chain length specificity in Arabidopsis FAR5 and FAR8 (Chapter 4).

5.1 Future Directions

The results of Chapter 4 indicated that a region between amino acids 312 and 383 of Arabidopsis FAR5 and FAR8 is important for determining C16:0 versus C18:0 chain length specificity. Further internal domain swaps between FAR5 and FAR8 need to be carried out as well as site-specific substitutions of the 13 amino acids that are different between the two enzymes in this region, which have not yet been investigated. An *in vitro* FAR assay should be developed to explore in detail the kinetics of these

FAR5/FAR8 variants for C16:0-CoA and C18:0-CoA substrates. Microsome preparations from yeast heterologously expressing these enzymes may be sufficient for such *in vitro* assays. Pure FAR5/FAR8 enzymes, for example prepared from transgenic *E. coli* and affinity purified, would be more difficult to isolate due to the predicted membrane associations of these two FARs.

The domain-swapping approach used in this study could also be repeated with different FARs in order to determine substrate specificities with other characteristics, such as accepting substrates with different degrees of saturation. FAR3/CER4 and the Jojoba FAR share 54% amino acid sequence identity and have overlapping substrate specificity with regards to chain length but differ in acyl chain saturation. FAR3/CER4 is specific for C24-C30 saturated fatty acyl-CoA substrates (Rowland *et al.*, 2006; Chapter 3), whereas Jojoba FAR (jFAR) is specific for C20-C24 monounsaturated fatty acyl-CoA substrates (Metz *et al.*, 2000). Yeast could be fed C24:1 and then various CER4/jFAR chimeras tested for their ability to produced C24:0 or C24:1 fatty alcohols. This would give insights into the specificity of FARs with regard to fatty acyl chain saturation. The introduction of a double bond into the fatty alcohol moiety of a wax ester would decrease the melting temperature of the wax ester, which is important for some commercial applications. For this same reason (low melting point characteristics), the genetic engineering of highly active FAR enzymes with medium chain-length specificity and activity for monounsaturated substrates (e.g. producing C12:1, C14:1, C16:1, or C18:1 fatty alcohols) is an important biotechnology goal for some of the industrial applications described above.

Directed evolution would be another approach for identifying amino acids that are important for understanding FAR enzymology and substrate specificity determinants. Directed evolution mimics natural evolution and its purpose is to generate a protein with novel activity by selecting for protein variants with a desired function (Tao and Cornish, 2002). This technique has been used successfully to modify enzymes to increase their catalytic efficiency, thermal stability, change their substrate specificity, and/or to create a novel function. This may be a very effective approach in generating enzymes with specificity for C18:1 fatty acyl substrate for example. A library of randomly mutated Arabidopsis *FAR5* coding sequences (the encoded protein has C18:0 fatty acyl substrate specificity) could be created through error prone PCR. The yeast heterologous system developed in Chapter 3 could then be used to analyze randomly mutated FAR5 proteins for activity in converting C18:1 (monounsaturated) fatty acyl-CoA to C18:1 fatty alcohol. The yeast strain H1246, which contains knockouts in four genes involved in triacylglycerol synthesis (*DGA1*, *LRO1*, *ARE1*, and *ARE2*), could be used for selection purposes (Sandager *et al.*, 2002). H1246 is incapable of performing neutral (e.g. triacylglycerol) lipid biosynthesis. When this yeast strain is grown on media containing certain fatty acids (e.g. C18:1 fatty acid), it has significantly reduced growth due to its inability to convert these fatty acids into neutral lipids. Instead, the fatty acids accumulate within the yeast cell and inhibit growth. Studies have shown that this phenotype can be rescued by restoring triacylglycerol biosynthesis. For example, this can be achieved by expressing plant triacylglycerol synthesizing enzymes (DGAT or PDAT) in the H1246 yeast strain (Siloto *et al.*, 2009). Rescue may also be possible by introducing a FAR enzyme that converts the C18:1 fatty acyl-CoA to C18:1 fatty alcohol (i.e. a mutated

FAR5). Yeast can tolerate relatively large amounts of fatty alcohol (Chapter 3). The directed evolution approach would provide further insights into the protein domains responsible for conferring substrate saturation specificity (from inspection of active sequences), as well as potentially engineering a FAR enzyme that synthesizes C18:1 primary alcohol for industrial use.

All of these molecular approaches to understanding and manipulating FAR substrate specificity are hampered by the lack of knowledge of FAR three-dimensional structure. There are currently no X-ray crystallography or NMR structures for FAR proteins. The highest sequence identity for a solved protein structure to model FAR protein structure is 26%, matching only a portion of the protein, which is too low to be reliable for modeling. In addition, this structure only includes the Rossmann fold domain and lacks the C-terminal region that I identified as important for determining chain-length substrate specificity. An X-ray or NMR structure of a FAR is very important for elucidating the residues and regions that are important in substrate specificity and should be a future priority in FAR enzyme characterization. This would be most feasible with a soluble, plastid FAR, such as FAR2 or FAR6 from Arabidopsis. Other membrane-associated FARs (e.g. FAR5) may be modelled from the soluble FAR three-dimensional structure.

5.2 Concluding Remarks

In conclusion, the findings from this study provide important information about FAR substrate specificities. Also, a yeast expression/screening system was developed that can be used to extend these findings much further. The future protein engineering of

FARs with desired substrate specificities, based on the fundamental knowledge achieved here and extended in future studies, will lead to the production of renewable fatty alcohol-containing lipid products of high commercial value.

6.0 References

Aarts, M.G.M., Keijzer, C.J., Stiekema, W.J., Pereira, A. (1995) Molecular Characterization of the *CER1* Gene of Arabidopsis Involved in Epicuticular Wax Biosynthesis and Pollen Fertility. *The Plant Cell.* 7: 2115-2127.

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 Journal.* **12**: 615-623.

Adamovics, J.A., Johnson, G., Stermitz, F.R. (1977) Ferulates from Cork Layers of *Solanum tuberosum* and *Pseudotsuga menziesii. Phytochemistry.* **16**: 1089–1090.

Andary, C. (1993) Caffeic acid glycoside esters and pharmacology. *In: Polyphenolic Phenomena.* eds. Scalbert, A. Institut National de la Recherche Agronomique, Paris, France: p 237-245.

Bassam, N.E.. (1998) Energy Plant Species: Their Use and Impact on Environment and Development. James and James (Science Publishers) Ltd, London, UK: p 168

Beisson, F., Li,, Y., Bonaventure, G., Pollard, M., Ohlrogge, J.B. (2007) The Acyltransferase GPAT5 is Required for the Synthesis of Suberin in Seed Coat and Root of Arabidopsis. *The Plant Cell.* **19**:351-368.

Bernards, M.A. (2002) Demystifying suberin. *Canadian Journal of Botany.* **80:** 227–240.

Carnegie Institution of Washington Department of Plant Biology. (2008). *TAIR Gene Search.* Available: http://www.arabidopsis.org. Last accessed 27 January 2010.

Cheminat, A., Zawatzky, R., Becker, H., Brouillard, R. (1988) Caffeoyl Conjugates from Echinacea Species: Structures and Biological Activities. *Phytochemistry.* **27**: 2787-2794.

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.

Chen, W., Yu, X-H., Zhang, K., Shi, J., De Oliveira, S., Schreiber, L., Shanklin, J., Zhang, D. (2011) Male Sterile2 Encodes a Plastid-Localized Fatty Acyl Carrier Protein Reductase Required for Pollen Exine Development in Arabidopsis. *Plant Physiology.* **157**: 842-853.

Cheng, J.B, Russell, D.W. (2004) Mammalian Wax Biosynthesis I. Identification of Two Fatty Acyl-Coenzyme A Recuctases with Different Substrate Specificities and Tissue Distributions. *The Journal of Biological Chemistry.* **279(36)**: 37789-37797.

Clough, Steven, J., and Bent, Andrew, F. (1998) Floral Dip: A Simplified Method for Agrobacterium-Mediated Transformation of *Arabidopsis thaliana. The Plant Journal.* **16(6)**:735-743.

Dean, B.B., Kolattukudy, P.E.. (1976) Synthesis of Suberin during Wound-healing in Jade Leaves, Tomato Fruit, and Bean Pods. *Plant Physiology.* **58**: 411-416.

Denessiouk, K.A., Rantanen V-V., Johnson, M.S. (2001) Adenine Recognition:AMotif Present in ATP-, CoA-, NAD-, NADP-, and FAD-Dependent Proteins. *Proteins: Structure, Function, and Genetics.* **44**:282–291.

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. *Journal of Plant Physiology.* **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* (2011), doi:10.1016/j.bbalip.2011.10.019

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 v-Hydroxylase Essential for Sporopollenin Synthesis in Pollen of Arabidopsis. *Plant Physiology.* **151**: 574-589.

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. *The Journal of Biological Chemistry.* **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 Physiology.* **153**: 1539-1554.

Donzé, G., Schnyder-Candrian, S., Bogdanov, S., Diehl, P-A., Guerin, P.M., Kilchenman, V., Monachon, F. (1998) Aliphatic Alcohols and Aldehydes of the Honey Bee Cocoon Induce Arrestment Behavior in Varroa acobsoni (Acari: Mesostigmata), an Ectoparasite of Apis mellifera. *Archives of Insect Biochemistry and Physiology.* **37(2)**: 129-145.

Eggink, G., Engel, H., Vriend, G., Terpstra, P., Witholt, B. (1990) Rubredoxin Reductase of Pseudomonas oleovorans tructural Relationship to Other Flavoprotein Oxidoreductases Based on One NAD and Two FAD Fingerprints. *Journal of Molecular Biology.* **212(1)**: 135-142.

Elliger, C.A., Lunden, R.E., Haddon, W.F. (1981) Caffeoyl Esters of Glucaric Acid in *Lycopersicon esculentum* leaves. *Phytochemistry.* **20:** 1133-1134.

Emanuelsson, O., Nielsen, H., von Heijne, G. (1999) ChloroP, a Neural Network-Based Method for Predicting Chloroplast Transit Peptides and their Cleavage Sites. *Protein Science.* **8**: 978-984.

European Bioinformatics Institute. (2010) *ClustalW2.* Available: http://www.ebi.ac.uk/Tools/clustalw2/index.html. Last accessed 18 February 2012.

Franke, R., Schreiber, L. (2007) Suberin--a Biopolyester Forming Apoplastic Plant Interfaces. *Current Opinion in Plant Biology.* **10(3):** 252-259.

Gavel, Y., von Heijne, G. (1990) A Conserved Cleavage-Site Motif in Chloroplast Transit Peptides. *FEBS Letters.* **261(2)**: 455-458.

Gietz, R.D., Woods, R.A. (2002) Yeast Transformation by the LiAc/SS Carrier DNA/PEG Method. In: *Yeast Protocol. 2nd ed.*. Totowa, NJ: Humana Press Inc.. p. 107-120.

Ghosh, D., Pletnev, V.Z., Zhu, D.W., Wawrzak, Z., Duax, W.L., Pangborn, W., Labrie, F., Lin, S.X. (1995) Structure of human estrogenic 17 β-hydroxysteroid dehydrogenase at 2.20 Å resolution. *Structure.* **3**:503-513.

Garnier, N., Cren-Olivé, C.,Rolando, C., Regert, M. (2002) Characterization of Archaeological Beeswax by Electron Ionization and Electrospray Ionization Mass Spectrometry. *Analytical Chemistry.* **74(19)**: 4868-4877.

Greer, S., Wen, M., Bird, D., Wu, X., Samuels, L., Kunst, L., Jetter, R. (2007) The Cytochrome P450 Enzyme CYP96A15 is the Midchain Alkane Hydroxylase Responsible for Formation of Secondary Alcohols and Ketones in Stem Cuticular Wax of Arabidopsis. *The Plant Cell.* **19:** 1473-1487.

Hajra, A. K. (1983) in *Ether Lipids: Biochemical and Biomedical Aspects*, eds. Mangold, H. K. and Paltauf, F. NY: Academic Press p. 85-106.

Hannoufa, A., McNevin, J., Lemieux, B. (1993) Epicuticular Waxes of Eceriferum Mutants of Arabidopsis thaliana. *Phytochemistry.* **33(4)**: 851-855.

HEAR. (2009) Human & Environmental Risk Assessment on ingredients of European household cleaning products: Alcohol Ethoxylates Version 2.0. Available: http://www.heraproject.com/files/34-F-09%20HERA%20AE%20Report%20Version%202%20-%203%20Sept%2009.pdf. Last accessed: March 27,2012

Hohlfeld, M., Veit, M., Strack, D. (1996) Hydroxycinnamoyltransferases Involved in the Accumulation of Caffeic Acid Esters in Cametophytes and Sporophytes of *Equisetum arvense*. *Plant Physiol*ogy. **111**:1153-1159.

Huang, J., Hack, E., Thornburg, R.W., Myers, A.M. (1990) A Yeast Mitochondrial Leader Peptide Functions *in vivo* as a Dual Targeting Signal for Both Chloroplasts and Mitochondria. *Plant Cell.* **2**:1249-1260.

Hurt, E.C., Soltanifar, N., Goldschmidt-Clermont, M., Rochaix, J.D., Schatz, G. (1986) The Cleavable Presequence of an Imported Chloroplast Protein Directs Attached Polypeptides into Yeast Mitochondria. *European Molecular Biology Organization Journal.* **5**:1343-1350.

Jetter, R., Kunst, L. (2008) Harnessing Plant Biomass for Biofuels and Biomaterials: Plant Surface Lipid Biosynthetic Pathways and their Utility for Metabolic Engineering of Waxes and Hydrocarbon Biofuels. *The Plant Journal.* **54:** 670-683.

Jenks, M.A., Eigenbrode, S.D., Lemieux, B. (2002) Cuticular Waxes of Arabidopsis: August 12, 2002. The Arabidopsis Book. Rockville, MD: American Society of Plant Biologists. doi: 10.1199/tab.0111, http://www.aspb.org/publications/arabidopsis/

Jenks, M.A., Tuttle, H.A., Eigenbrode, S.D., Feldmann, K.A. (1995). Leaf Epicuticular Waxes of the *Eceriferum* Mutants in *Arabidopsis*. *Plant Physiology.***108:** 369-377

Jörnvall, H., Persson, M., Jeffery, J. (1981) Alcohol and Polyol Dehydrogenases are Both Divided into Two Protein Types, and Structural Properties Cross-Relate the Different Enzyme Activities within each type. *Proceedings of the National Academy of Sciences of the United States of America.* **78**:4226-4230.

Jung, K., Han, M., Lee, D., Lee., Y., Schreiber, L., Franke, R., Faust, A., Yephremov, A., Saedler, H., Kim, Y., Hwang, I., An, G. (2006) Wax-deficient anther1 is Involved in Cuticle and Wax Production in Rice Anther Walls and is Required for Pollen Development. *The Plant Cell.* **18:** 3015-3032.

Kalscheuer, R., Steinbüchel, A.. (2003) A Novel Bifunctional Wax Ester Synthase/Acyl-CoA:Diacylglycerol Acyltransferase Mediates Wax Ester and Triacylglycerol Biosynthesis in *Acinetobacter calcoaceticus* ADP1. *Journal of Biological Chemistry.* **278(10)**: 8075-8082.

Kajdas, C. (1987) About an anionic-radical concept of the lubrication mechanism of alcohols. *Wear.* **116(2)**: 167-180.

Klun, J.A. (1975) Insect Sex Pheromones: Intraspecific pheromonal Variability of *Ostrinia nubilalis* in North America and Europe. *Environmental Entomology.* **4(6)**: 891-894.

Kolattukudy, P.E.. (1970) Reduction of Fatty Acids to Alcohols by Cell-Free Preparations of *Euglena gracilis. Biochemistry.* **9(5)**: 1095-1102.

Kolattukudy, P.E.. (1971) Enzymatic Synthesis of Fatty Alcohols in *Brassica oleracea. Archives of biochemistry and Biophysics.* **142**: 701-709.

Kolattukudy, P.E.. (2001) Polyesters in Higher Plants. *Advances in Biochemical Engineering/Biotechnology.* **71**: 1-49.

Kreutzer, U.R. (1984) Manufacture of Fatty Alcohols Based on Natural Fats and Oils. *Journal of the Oil Chemists' Society.* **61(2)**: 343-348

Kunst, L., Samuels, A.L. (2003) Biosynthesis and secretion of plant cuticular wax. *Progress in Lipid Research.* **42:** 51–80.

Houston, C.A. (1984) Marketing and Economics of Fatty Alcohols. *Journal of the American Oil Chemists' Society.* **61(2)**: 179-184.

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). *Biochemical Journal.* **304**: 95-99.

Lai C., Kunst L., Jetter R. (2007) Composition of Alkyl Esters in the Cuticular Wax on Inflorescence Stems of *Arabidopsis thaliana cer* Mutants. *The Plant Journal.* **50:**189–96.

Lassance, J.M., Groot, A.T., Liénard, M.A., Antony, B., Borgwardt, C., Andersson, F., Hedenström, 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.

Lassner, M. (1997) Transgenic Oilseed Crops: a Transition from Basic Research to Product Development. *Lipid Technology.* **9**: 5–9.

Laver, M.L., Fang, H.H.L. (1989) Ferulic Acid Esters from Bark of *Pseudotsuga menziesii. Journal of Agriculture and Food Chemistry..* **37 (1)**: 114–116.

Li, Y., Beisson, F., Ohlrogge, J., Pollard, M. (2007) Monoacylglycerols are Components of Root Waxes and Can Be Produced in the Aerial Cuticle by Ectopic Expression of a Suberin-Associated Acyltransferase. *Plant Physiology.* **144**: 1267-1277.

Li, F., Wu, X., Lam, P., Bird, D., Zheng, H., Samuels, L., Jetter, R., Kunst, L. (2008) Identification of the Wax Ester Synthase/Acyl-CoA:Diacylglycerol Acyltransferase WSD1 Required for Stem Wax Ester Biosynthesis in *Arabidopsis thaliana. Plant Physiology.* **148(1)**: 97-107.

Liénard, M.A., Hagström, A.K., Lassance, J-M., Löfstedt, C.. (2010) Evolution of Multicomponent Pheromone Signals in Small Ermine Moths Involves a Single Fatty-Acyl Reductase Gene. *Proceedings of the National Academy of Sciences.* **107(24)**: 10955-10960.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., Bryant, S.H. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research.* **39**(Database issue):225-229.

Matern, U., Kneusel, R.E. (1988) Phenolic Compounds in Plant Disease Resistance. *Phytoparasitica.* **16(2)**:153-170.

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.

Millar, A.A., Kunst, L. (1997) Very-Long-Chain Fatty Acid Biosynthesis is Controlled Through the Expression and Specificity of the Condensing Enzyme. *The Plant Journal.* **12(1)**: 121-131.

Nagan, N., Zoeller, R.A. (2001) Plasmalogens: Biosynthesis and Functions. *Progress in Lipid Research.* **40**: 199-229.

Ohlrogge, J.B., Pollard, M.R., Stumpf, P.K. (1977) Studies on Biosythesis of Waxes by Developing Jojoba Seed Tissue. *Lipids.* **13(3)**: 203-210.

Pollard, M., Beisson, F., Li, Y., Ohlrogge, J.B. (2008) Building Lipid Barriers: Biosynthesis of Cutin and Suberin. *Trends in Plant Science.* **13(5)**: 236-246.

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(7)**: 651-662.

Ravn, H., Andary, C., Kovács, G., Melgaard, P. (1989) Caffeic Acid Esters as *in vitro* Inhibitors of Plant Pathogenic Bacteria and Fungi. *Biochemical Systematics of Ecology.* **17**: 175-184

Ravn, H., Brimer, L. (1988) Structure and Antibacterial Activity of Plantamajoside, a Caffeic Acid Sugar Ester from Plantago Major subsp. major. *Phytochemistry.* **27**: 3433-3437.

Richardson, J.S. (1981) The Anatomy and Taxonomy of Protein Structure. *Advances in Protein Chemistry.* **34**:167–339.

Rossmann, M.G., Moras, D., Olsen, K.W. (1974) Chemical and Biological Evolution of a Nucleotide-Binding Protein. *Nature.* **250**: 194-199.

Rowland, O., Domergue, F. (2012). Plant Fatty Acyl Reductases: Enzymes Generating Fatty Alcohols for Protective Layers with Potential for Industrial Applications. *Plant Science.* IN PRESS.

Rowland, O., Zheng, H., Hepworth, S.R., Lam, P., Jetter, R., Kunst, L. (2006) CER4 Encodes an Alcohol-Forming Fatty Acyl-Coenzyme A Reductase Involved in Cuticular Wax Production in Arabidopsis. *Plant Physiology.* **142**: 866-877.

Samuels, L., Kunst, L., Jetter, R. (2008) Sealing Plant Surfaces: Cuticular Wax Formation by Epidermal Cells. *Annual Review of Plant Biology.* **59**: 683-707.

Schreiber, L. (2010) Transport Barriers Made of Cutin, Suberin and Associated Waxes. *Trends in Plant Science.* **15**: 546-553.

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., Zhanga, D. (2011) Defective Pollen Wall Is Required for Anther and Microspore Development in Rice and Encodes a Fatty Acyl Carrier Protein Reductase. *The Plant Cell.* **23(6)**: 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 Physiology and Biochemistry.* **47:** 456-461.

Soll, J., Tien, R. (1998) Protein Translocation Into and Across the Chloroplastic Envelope Membranes. *Plant Molecular Biology.* **38**:191-207

Spencer, G.F. (1979) Alkoxy-Acyl Combinations in the Wax Esters from Winterized Sperm Whale Oil by Gas Chromatography-Mass Spectrometry. *Journal of the American Oil Chemists Society.* **56 (6)**: 642-646.

Suh, M.C., Samuels, L., Jetter, R., Kunst, L., Pollard, M., Ohlrogge, J., Beisson, F. (2005) Cuticular Lipid Composition, Surface Structure, and Gene Expression in Arabidopsis Stem Epidermis. *Plant Physiology.* **139**: 1649-1665.

Tauschek, M., Gorrell, R.J., Strugnell, R.A., Robins-Browne, R.M. (2002) Identification of a Protein Secretory Pathway for the Secretion of Heat-Labile Enterotoxin by an Enterotoxigenic Strain of Escherichia coli. *Proceedings of the National Academy of Sciences.* **99**:7066–7071.

Teerawanichpan, P., Robertson, A.J., Qiu, X. (2010[a] 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 Biochemistry and Molecular Biology.* **40(9)**: 641-649.

Teerawanichpan, P., Qiu, X. (2010[b]) 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 Biotechnology* (2011) doi 10.1007/s10126-011-9406-3

Tucci S., Vacula R., Krajcovic J., Proksch P., Martin W. (2009) Variability of Wax Ester Fermentation in Natural and Bleached *Euglena gracilis* Strains in Response to Oxygen and the Elongase Inhibitor Flufenacet. *Journal of Eukaryotic Microbiology* **57**:63-69.

U.S. National Library of Medicine. (2009) National Center for Biotechnology Information. Available: ttp://www.ncbi.nlm.nih.gov/protein/AAD38039?log$=ACCN. Last accessed: January 3, 2012.

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.). *Archives of Biochemistry and Biophysics.* **340(1)**: 64-72.

von Heijne, G., Nishikawa, K. (1991) Chloroplast transit peptides. *FEBS Letters.* **278:** 1-3.

von Heijne, G., Steppuhn, J., Herrmann, R.G. (1989). Domain structure of mitochondrial and chloroplast targeting peptides. *European Journal of Biochemistry.* **180**: 535-545.

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. *The Plant Journal.* **30(6)**: 613-623.

Wierenga, , R.K., Terpstra, P., Hol, W.G.J. (1986) Prediction of the Occurrence of the ADP-binding βαβ-Fold in Proteins, Using an Amino Acid Sequence Fingerprint. *Journal of Molecular Biology.* **187(1)**: 101-107.

Wiermann, R., Gubatz, S. (1992) Pollen Wall and Sporopollenin *In Sexual Reproduction in Flowering Plants.* eds. S. D. Russell. Academic Press Inc., California, USA: p 52-53.