The specificities of plant enzymes and their effects on the seed oil quality

Simon Jeppson
The specificities of plant enzymes and their effects on the seed oil quality

Simon Jeppson
Faculty of Landscape Architecture, Horticulture and Crop Production Science
Department of Plant Breeding
Alnarp

Doctoral thesis
Swedish University of Agricultural Sciences
Alnarp 2020
“The madder something is, the better it works," the Cheshire cat said. "Which is of course, why you fit in very well.” - Cover photo: Simon Jeppson
The specificities of plant enzymes and their effects on the seed oil quality

Abstract
Triacylglycerols, glycerides with three fatty acids, are the chemical constituent of seed oils. The fatty acid composition determines the chemical and physical characteristics of an oil. Conventional vegetable oils contain a few reoccurring fatty acids, but plants can produce several hundreds of different fatty acids. A wide range of plant oils and their fatty acids with specific characteristics are suitable replacements for many fossil-oil based commodities. Plant breeders and researcher are thus, modifying the fatty acid composition in oil crops to enhance such properties. Erucic acid is a fatty acid, with extensive use within industrial applications. Fatty acids originate from the plastids in plants, but the assembly of triacylglycerols occurs in the endoplasmic reticulum, where integral membrane-bound acyltransferases catalyse the formation of triacylglycerols. Modifications of fatty acids occur while they are associated with specific lipid molecules. Different types of fatty acid modifications require association with different lipid molecules, and several unusual fatty acids require multiple sequential modifications. The intricate system of synthesis, modification and assembly thus requires channelling of fatty acids and substrates from different subcellular compartments and concurrent transfer of fatty acids from various lipid molecules, often at multiple occasions. The enzymes governing these processes often have very distinct substrate specificities, which we investigate here. Re-engineering of a seed oil’s fatty acid composition or the introduction of an exotic fatty acid is a complex process and must take into consideration plants endogenous enzymes, their substrate specificities and the available substrate pools. Detailed biochemical characterisations of crucial enzymes are thus essential for a successful outcome of such efforts. We have thoroughly investigated a range of acyltransferases and other enzymes involved in lipid metabolism, mainly focused towards erucic acid in several plant species, including Arabidopsis, Camelina sativa, Brassica napus, Crambe hispanica ssp. abyssinica and Tropaeolum majus. DGAT2 catalyses the acyl-CoA dependent formation of triacylglycerol and is thus one of the acyltransferases involved in oil synthesis. The assessed substrate specificities of several DGAT2 indicate that its substrate specificity is a bottleneck in the further increase of erucic acid. We have also investigated the amino acid sequence motifs governing the specificity towards erucic acid in DGAT2 enzymes.

Keywords: Plant lipids, TAG, DGAT, acyltransferase, lipid synthesis, enzyme specificities, erucic acid, oil crop

Author’s address: Simon Jeppson, SLU, Department of Plant Breeding, P.O. Box 101, 230 53 Alnarp, Sweden

Nyckelord: Växtlipider, TAG, DGAT, acyltransferase, lipidsyntes, enzym specificitet, erukasyra, oljegröda

Författarens adress: Simon Jeppson, SLU, Institutionen för växtförädling, Box 101, 230 53 Alnarp, Sverige
Dedication

To all the giants upon whose shoulders I stand

Рукописи не горят - Manuscripts do not burn
Mikhail Afanasyevich Bulgakov
List of publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:


II Jeppson, Simon, Mattisson, Helena, Demski, Kamil, Lager, Ida. ”A predicted transmembrane region in plant diacylglycerol acyltransferase 2 regulates the specificity towards very long chain acyl-CoAs.”, Submitted manuscript.


IV Jeppson, Simon, Rivani, Meta, Lager, Ida. ”Identification of DGAT2 and characterisation of DGAT enzymes from Tropaeolum majus.”, Manuscript.

V Lager, Ida, Jeppson, Simon, Stymne, Sten, Marmon, Sofia. "Acyltransferases regulate oil quality in Camelina through both acyl donor and acyl acceptor specificities.”, Submitted manuscript.
The contribution of S.J to the papers included in this thesis was as follows:

I  K.D., S.J., I.L., S.S., and A.B. conceived the research plans; K.D., S.J., I.L., and A.B. designed the experiments; I.L., S.S., and A.B. supervised the experiments; A.M. and M.W. designed and performed the phylogenetic analyses; K.D., S.J., I.L., and K.J.-G. performed the experiments; K.D., S.J., and I.L. analyzed the obtained data; K.D., S.J., A.B., I.L., and S.S. wrote the article; K.D & S.J contributed equally

II  I.L., S.J., and K.D. perceived the concept; I.L. and S.J. designed the experiment; I.L., H.M., and S.J. conducted the experiments; S.J. wrote the original draft; I.L., S.J., H.M., and K.D. reviewed and edited the final version.

III  I.L. and S.S. envisioned this project; S.J., I.L., S.S., and K.D. coordinated and designed the experiments; S.J., K.D., and I.L. drafted the manuscript; A.B., A.C., L.-H.Z., and S.S. contributed to the final version; K.D. & S.J. contributed equally.

IV  I.L. and S.J. perceived the concept; I.L. and S.J. designed the experiments; I.L., S.J., and M.R. conducted the experiments; S.J. and I.L. wrote the manuscript; I.L., S.J., and M.R. reviewed and edited the final version.

V  S.M, S.J., I.L., I.F. and S.S. conceived the original research plan; I.L., S.J., A-L.G., S.M., and S.S. performed the experiments; S.S. wrote the article with contributions from all the authors.
List of figures

Figure 1  The chemical structures for TAG, oleic acid and wax ester……18

Figure 2  Schematic overview of the cyclic reactions catalysing acyl elongation in FAS.................................................................22

Figure 3  Overview of the reactions catalysed by the KAS enzymes during the condensation in FAS.......................................................23

Figure 4  Overview of the synthesis of FAs in the plastid and the following export to the cytosol through proposed mechanisms.............25

Figure 5  Simplified overview of TAG assembly in the ER....................27

Figure 6  Overview of two alternative metabolic routes present in Arabidopsis for the synthesis of thylakoid membrane lipids.............32

Figure 7  Structural features of several UFAs......................................35

Figure 8  The proposed biosynthetic pathway, resulting in nebraskanic acid in O. violaceus.................................................................38

Figure 9  Metabolic routes in the aerobic pathway for the synthesis of both EPA and DHA.................................................................44

Figure 10 Principles behind expressing TAG producing acyltransferases in H1246..............................................................................48

Figure 11 Acyl-CoA specificities of two B. napus DGAT2 enzymes and phylogenetic analysis of these, including both ancestral parents DGAT2 sequences.................................................................54
**Figure 12** Parental wildtype enzymes and chimeric DGAT2 enzymes and their specificity towards 18:3Δ⁹,¹²,¹⁵ and 22:1Δ¹³ ......................56

**Figure 13** Illustration of the BnaA.DGAT2.d amino acid sequence with essential features.................................................................57

**Figure 14** Substrate specificities of crambe DGAT enzymes towards 22:1Δ¹³-CoA in combination with the frequently occurring sn-1-22:1Δ¹³-sn-2-18:1Δ⁹-DAG or di-22:1Δ¹³-DAG........................................58

**Figure 15** Substrate specificities of *T. majus* DGAT enzymes towards 22:1Δ¹³-CoA in combination with most preferred DAG species or di-22:1Δ¹³-DAG ........................................................................59

**Figure 16** Acyl-CoA specificity of CsDGAT2 and a phylogenetic analysis including the *B. napus* DGAT2 amino acid sequences ..........61
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>acetyl-CoA carboxylase</td>
</tr>
<tr>
<td>ACP</td>
<td>acyl carrier protein</td>
</tr>
<tr>
<td>ADS</td>
<td>acyl-CoA desaturase-like</td>
</tr>
<tr>
<td>ANOVA</td>
<td>analysis of variance</td>
</tr>
<tr>
<td>BASF</td>
<td>badische anilin- und soda-fabrik</td>
</tr>
<tr>
<td>Cas9</td>
<td>CRISPR-associated protein 9</td>
</tr>
<tr>
<td>CPT</td>
<td>CDP-choline:DAG cholinephosphotransferase</td>
</tr>
<tr>
<td>CRISPR</td>
<td>clustered regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>CSIRO</td>
<td>commonwealth scientific and industrial research organisation</td>
</tr>
<tr>
<td>DAG</td>
<td>diacylglycerol</td>
</tr>
<tr>
<td>DGAT</td>
<td>acyl-CoA:diacylglycerol acyltransferase</td>
</tr>
<tr>
<td>DGDG</td>
<td>digalactosyl diacylglycerol</td>
</tr>
<tr>
<td>DHA</td>
<td>docosahexaenoic acid</td>
</tr>
<tr>
<td>DSM</td>
<td>dutch state mines</td>
</tr>
<tr>
<td>EPA</td>
<td>eicosapentaenoic acid</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>FA(s)</td>
<td>fatty acid(s)</td>
</tr>
<tr>
<td>FAS</td>
<td>fatty acid synthase</td>
</tr>
<tr>
<td>FFA(s)</td>
<td>free fatty acid(s)</td>
</tr>
<tr>
<td>GPAT</td>
<td>glycerol-3-phosphate acyltransferase</td>
</tr>
<tr>
<td>GPC</td>
<td>glycerol-3-phosphocholine</td>
</tr>
<tr>
<td>GPCAT</td>
<td>glycerophosphocholine acyltransferase</td>
</tr>
<tr>
<td>GRDC</td>
<td>grains research and development corporation</td>
</tr>
<tr>
<td>HEAR</td>
<td>high erucic acid rapeseed</td>
</tr>
<tr>
<td>HELP</td>
<td>high erucic low polyunsaturated</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>HPLC</td>
<td>high-performance liquid chromatography</td>
</tr>
<tr>
<td>HSD</td>
<td>honestly significant difference</td>
</tr>
<tr>
<td>KAS</td>
<td>β-ketoacyl-ACP synthase</td>
</tr>
<tr>
<td>KCS</td>
<td>β-ketoacyl-CoA synthase</td>
</tr>
<tr>
<td>LACS</td>
<td>long-chain acyl-CoA synthetases</td>
</tr>
<tr>
<td>LC-PUFA</td>
<td>long-chain ω-3</td>
</tr>
<tr>
<td>LCS</td>
<td>liquid scintillation counter</td>
</tr>
<tr>
<td>LEAR</td>
<td>low erucic acid rapeseed</td>
</tr>
<tr>
<td>LLC</td>
<td>limited liability company</td>
</tr>
<tr>
<td>LPA</td>
<td>lysophosphatidic acid</td>
</tr>
<tr>
<td>LPAT</td>
<td>lysophosphatidic acid acyltransferase</td>
</tr>
<tr>
<td>LPC</td>
<td>lysophosphatidylcholine</td>
</tr>
<tr>
<td>LPCAT</td>
<td>lysophosphatidylcholine acyltransferase</td>
</tr>
<tr>
<td>MCFA</td>
<td>medium-chain fatty acids</td>
</tr>
<tr>
<td>MCMT</td>
<td>malonyl-CoA:acyl carrier protein malonyltransferase</td>
</tr>
<tr>
<td>MGDG</td>
<td>monogalactosyl diacylglycerol</td>
</tr>
<tr>
<td>MSY</td>
<td>maximum sustainable yield</td>
</tr>
<tr>
<td>PA</td>
<td>phosphatidic acid</td>
</tr>
<tr>
<td>PAP</td>
<td>phosphatidic acid phosphatase</td>
</tr>
<tr>
<td>PC</td>
<td>phosphatidylcholine</td>
</tr>
<tr>
<td>PDAT</td>
<td>phospholipid: diacylglycerol acyltransferase</td>
</tr>
<tr>
<td>PDCT</td>
<td>phosphatidylcholine diacylglycerol cholinephosphotransferase</td>
</tr>
<tr>
<td>PE</td>
<td>phosphatidylethanolamine</td>
</tr>
<tr>
<td>PGP</td>
<td>phosphatidyglycerophosphate</td>
</tr>
<tr>
<td>PI</td>
<td>phosphatidylinositol</td>
</tr>
<tr>
<td>PKS</td>
<td>polyketide synthases</td>
</tr>
<tr>
<td>PS</td>
<td>phosphatidylserine</td>
</tr>
<tr>
<td>PUFA(s)</td>
<td>polyunsaturated FA(s)</td>
</tr>
<tr>
<td>RACE</td>
<td>rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>SAD</td>
<td>Δ⁹ stearoyl-ACP desaturase</td>
</tr>
<tr>
<td>SQDG</td>
<td>sulfoquinovosyl diacylglycerol</td>
</tr>
<tr>
<td>TAG</td>
<td>triacylglycerol</td>
</tr>
<tr>
<td>TLC</td>
<td>thin-layer chromatography</td>
</tr>
<tr>
<td>UFA(s)</td>
<td>unusual fatty acid(s)</td>
</tr>
<tr>
<td>VLCFA</td>
<td>very long-chain fatty acids</td>
</tr>
</tbody>
</table>
1 Preamble

Many fatty acids (FAs) have trivial names usually associated with either a property of the FA or a common source from where the FA is derived. Examples of trivial names include, e.g. caproic acid as it smells of goat (Capricornus means horned goat in Latin), palmitic acid (often associated with palm oil) and linolenic acid (linon is Greek for flax, which produces an oil containing linolenic acid). The trivial names are tedious to learn and seldom reveal any essential structural properties of the FAs. Most often the Δ nomenclature is used instead since it is a simple, information-dense, way of naming FAs. The Δ nomenclature consists of several parts; the number of carbons, the number of desaturations, and where these are situated counting from carboxylic carbon. Linolenic acid is thus, denoted as 18:3Δ₉,₁₂,₁₅. As double bonds exist in two configurations, it is common to denote these as either cis or trans. Here all double bonds are of cis configuration unless explicitly stated as trans. Any additional functional group situated on the acyl chain, such as hydroxyl, is denoted before the acyl chain length and includes the carbon(s) where the functional group is situated, e.g. 12-OH-18:1Δ⁹.

In some cases, especially in a nutritional context, FAs are denoted as ω (omega), e.g. ω-3 or ω-6. Here the index number indicates the last double bond position as counted from the ω carbon in the opposite end of the carboxyl group in the acyl chain. The last double bond of 18:3Δ₉,₁₂,₁₅ is situated three carbons from the ω (Δ¹⁸-Δ¹⁵=3) and is thus a ω-3 FA.

Many plant lipids contain a glycerol backbone, see for example figure 1A. Stereocchemical numbering (sn) is useful to keep track of where an enzyme performs an activity or which FA is esterified to which carbon. Glycerol-3-phosphate, for example, is esterified at the sn-3 carbon with a phosphate, see figure 1C.
2 Introduction

Lipids are vital to plants, and their roles in plants are very diverse. Lipids are a principal constituent of cell membranes, act as signalling molecules, prevents desiccation and infections as a component of the cuticle, and enables an efficient energy and carbon storage. Lipids are not only vital to plants, but plant-derived lipids are of great importance for humans as well, as a nutrient source containing essential FAs and as feedstock for several industrial products.

The perhaps most well-known non-food application of plant oil is biodiesel, a product where FAs are methylated from seed oil or other plant sources such as tall oil and used as an environmentally friendly replacement of fossil-based diesel (Halleux et al., 2008). There is, however, a multitude of other applications where plant oil-derived products are suitable as a replacement for fossil oil-based products, e.g. as lubricants, slipping agents, detergents, precursors of plastics and as a constituent of pharmaceuticals and cosmetics.

Plant seed oil consists of triacylglycerols (TAG), a glycerol backbone esterified with three FAs, see figure 1A. FAs are carboxylic acids with aliphatic chains, see figure 1B, and several industrial applications make use of the FAs rather than the oil as such. The FA composition and the stereochemical position (figure 1C) of the FAs determine the chemical and physical properties of a plant seed oil and thus its suitability in different applications (Neff & El-Agaimy, 1996). Although it may not be obvious, glycerol is a prochiral compound, and TAG species are therefore chiral, the sn-1 and sn-3 positions are thus not equivalent. Jojoba (Simmondsia chinensis (Link) C.K.Schneid.), a bush endemic to the Sonoran desert in North America, is the only known exemption in lipid storage in that it produces a seed oil comprised of mainly wax esters instead of TAG (Busson-Breysse et al., 1994; Gentry, 1958). Wax esters are unlike TAG, molecules consisting of an FA esterified to fatty alcohol, see figure 1D. The jojoba seed oil has superior lubrication properties and is frequently used as a
constituent in cosmetics but less commonly in lubricants due to the high production costs (Shirani et al., 2020).

**Figure 1** A triacylglycerol molecule with three FAs acylated to a glycerol backbone. B Oleic acid, a common FA with a carboxyl group at Δ^1^ and a double bond at Δ^9^. C A glycerol-3-phosphate molecule and the stereochemical numbering (sn) of the glycerol carbon backbone. D Wax ester consisting of an FA and a fatty alcohol moiety.
Fatty acids fall into several different categories based on their characteristics. They may differ in acyl chain-length, degree of unsaturation, the position of desaturation(s) and additional functional groups. Most seed oils intended for human consumption such as soybean (Glycine max (L.) Merr.), canola (edible rapeseed oil, Brassica napus L.), oil palm (Elaeis guineensis Jacq.) and sunflower (Helianthus annuus L.) generally contain saturated and unsaturated FAs that are 16-18 carbons long, mainly palmitic (16:0), stearic (18:0), oleic (18:1Δ⁹), linoleic (18:2Δ⁹,12) and linolenic (18:3Δ⁹,12,15) acids (Dubois et al., 2007). These common FAs are also the dominating FAs species in membrane lipids of most plants. So far, findings describe the presence of several hundred different novel or unusual FAs (UFAs) in vascular plants (Ohlrogge et al., 2018). Unusual FA producing plant species, often accumulate very high levels of UFAs in their seed oil while they are often virtually absent in the plants' membrane lipids (Voelker & Kinney, 2001).

Many UFAs found in plants are suitable for a wide range of industrial applications, but economic reasons often prevent large scale production of these. Several suitable UFAs occurs in the seed oil of wild plants and are thus not fit for agriculture. Domesticated plants that produce UFAs may have additional factors that hamper large scale production.

Castor bean oil, for example, is of great interest for the oleochemical industry as it contains up to 90% of ricinoleic acid, 12-OH-18:1Δ⁹ (Achaya et al., 1964). Industrial applications use ricinoleic acid and its derivatives as a component in a wide range of products, e.g. cosmetics, pharmaceuticals, lubricants, plasticisers, coatings and surfactants (Murphy, 1992). However, the castor bean plant, Ricinus communis L., does not only produce ricinoleic acid but also contains allergens and highly toxic compounds such as ricin and ricinine. These compounds cause health concerns that rightfully discourage farmers from cultivating castor bean even though it has several other advantageous agricultural traits (Ogunniyi, 2006).

Petroselinic acid (18:1Δ⁶) is another UFA with desirable traits for several industrial applications. Several domesticated species within Umbelliferae, e.g. carrot (Daucus carota subsp. sativus (Hoffm.) Schübl. & G. Martens) and coriander (Coriandrum sativum L.), produces seed oils with more than 50% of petroselinic acid (Matthaus et al., 2003; Kleiman & Spencer, 1982). Domesticated varieties of both carrot and coriander are, however, adapted for other purposes than seed oil production and are currently not a very viable option for seed oil production (Iorizzo et al., 2013; Diederichsen & Hammer, 2003).
Rapeseed is a domesticated oil crop which has a seed oil that contains ample amounts of a UFA, erucic acid (22:1Δ13). Erucamide, a derivative of erucic acid, is frequently used as a slipping agent and erucic acid is a highly appreciated renewable commodity (Piazza & Foglia, 2001). Rapeseed varieties bred explicitly for erucic acid production, have a seed oil with very high levels of erucic acid, exceeding 50 mol % (McVetty et al., 2016). A further increase in erucic acid content is desirable not only because it results in a higher yield of erucic acid, but also that an increased concentration of erucic acid in the seed oil drastically reduces the downstream purification and production cost of erucamide (Li et al., 2012; Scarth & Tang, 2006). A higher erucic acid content would thus make biobased erucamide economically more competitive against fossil oil-based options.

A robust understanding of plant lipid metabolism is essential to overcome current limitations in the production of seed oils rich in UFAs and increase the competitiveness of plant-based alternatives to fossil-based products. Lipid metabolism is still surprisingly poorly characterised, albeit their vital and diverse role in living organisms. Some processes within lipid metabolism are probably universal, whereas others differ between organisms. Plants synthesise FAs in plastids, for example, whereas animals and fungi rely on cytosolic synthesis. Even within plants, differences in lipid synthesis are apparent, as will be discussed later.

Kennedy and Weiss elucidated several enzymatic steps, central in lipid metabolism, during the late 1950s, nowadays known as the Kennedy pathway (Weiss et al., 1960; Weiss & Kennedy, 1956). Biochemists have since then identified new enzymatic processes which are essential in lipid metabolism. High throughput methods of genetics and molecular biology have more recently gained a tremendous interest at the expense of more laborious classical biochemical methods. Less than half of the approximately 600 genes in Arabidopsis annotated as involved or putatively involved in lipid metabolism are confirmed in vivo (Allen et al., 2015). Biochemical analyses are still vital for understanding lipid metabolism and has currently regained popularity. The recent identification of an acyltransferase, glycerophosphocholine acyltransferase (GPCAT), is an excellent example of where biochemistry was crucial in order to identify a gene with at the time unknown function (Gląb et al., 2016).
The properties of three critical components, FA synthesis, FA modification, and TAG assembly significantly affect the composition of a seed oil during TAG synthesis. The TAG assembly requires an intricate shuttling of molecular components from different compartments of the cell. TAG synthesis also intertwines with the synthesis of several other lipids, vital for the survival of plants.

2.1 Fatty acid synthesis in plants

Most eukaryotic organisms synthesise FAs in the cytosol, but plants differ in this regard. Instead, plants synthesise de novo FAs in the stroma of plastids (Harwood, 1996; Ohlrogge et al., 1979). Acetyl-CoA carboxylase (ACC) facilitates the formation of malonyl-CoA from bicarbonate and acetyl-CoA (2:0-CoA), and this initiates the dedicated synthesis of FAs (Kannangara & Stumpf, 1972).

The plastidial ACC in most plants are of prokaryotic origin. Grasses, however, possess a plastidial ACC of eukaryotic origin (Konishi et al., 1996). Aryloxyphenoxypropionate and cyclohexanedione are two chemical compounds that efficiently inhibit the eukaryotic ACC of grasses, and thus FA synthesis. The inhibition of FA synthesis is detrimental to plants and these and other similar compounds act as efficient grass specific herbicides (Burton et al., 1989). The loss of prokaryotic plastidial ACC in grasses was the first published example of a eukaryotic nucleus residing genes replacement of a plastidial gene (Konishi et al., 1996).

The malonyl moiety produced by ACC is subsequently transferred from CoA to an acyl carrier protein (ACP) by malonyl-CoA:acyl carrier protein malonyltransferase (MCMT) (Guerra & Ohlrogge, 1986). Long-chain FAs are synthesised cyclically by fatty acid synthase (FAS), an enzymatic complex, where malonyl-ACP provides two carbon for elongation in each cycle (Harwood, 1988). Each cycle in FAS includes several steps; condensation, reduction, dehydration, and reduction, see figure 2 (Harwood, 1988). β-ketoacyl-ACP synthases (KAS) catalyses the condensation in FAS. Plants have several isoforms of KAS, where KAS III catalyses the initial condensation of acetyl-CoA and malonyl-ACP, while the other involved KAS utilise acyl-ACP and malonyl-ACP as substrates (Slabas & Fawcett, 1992). KAS I catalyses the condensation of acyl chains up to 16 carbons length whereas KAS II carries out the condensation of palmitoyl-ACP (16:0-ACP) and malonyl-ACP to β-keto-stearoyl-ACP (Slabas & Fawcett, 1992). The KAS enzymes catalyse several
chemical reactions included in a Claisen condensation, see figure 3 (Huang et al., 1998). In most plant species this cyclic elongation of acyl-ACP carries on until the acyl chains are 16 to 18 carbons long. A plastidial residing desaturase, Δ⁹ stearoyl-ACP desaturase (SAD), introduce a double bond in 18:0-ACP and thus desaturate it, which results in 18:1-Δ⁹-ACP, see figure 4 for an overview of plastidial processes during FA synthesis (McKeon & Stumpf, 1982).

**Figure 2** schematic overview of the cyclic reactions catalysing acyl elongation in FAS. FAS use malonyl-ACP for two carbon elongation during each cycle. The initial condensation catalysed by KASII uses acetyl-CoA instead of acyl-ACP. The condensed β-ketoacyl-ACP is further reduced into 3-hydroxy acyl-ACP and dehydrated into Δ⁹trans-2-enoyl-ACP. A final reduction results in acyl-ACP, which in turn can be further elongated by KAS I and later KAS II.

**Abbreviations:** ACP – acyl carrier protein, FAS – fatty acid synthase
Removal of the acyl group from ACP terminates the FA synthesis; either through transacylation or release of free FAs (FFAs) (Ohlrogge & Browse, 1995). Acyl-ACP thioesterase FatA and FatB facilitate the hydrolysis of FFAs from acyl-ACP (Jones, 1995). The vast majority of FAs released by the thioesterases in Arabidopsis are 18:1Δ9, followed by minor amounts of 16:0 and 18:0 (Li-Beisson et al., 2010).

Transcriptional factors such as WRI1 rigidly control the gross rate of FA synthesis (Grimberg et al., 2015). WRI1 autoregulates itself through a negative feedback loop, which may act as a fail-safe and may enable a faster response time than otherwise possible (Snell et al., 2019). These features potentially allow tuning of the FA synthesis rate. Several post-translational regulations are known to affect FA synthesis as well. ACC which facilitates the first committed step in FA synthesis, for example, is inhibited by 18:1Δ9-ACP, the primary end product in FA synthesis. Further, light, energy level and substrate availability regulates the ACC activity and thus enable a fine-tuning of FA synthesis (Troncoso-Ponce et al., 2016).

Figure 3 overview of the reactions catalysed by the KAS enzymes during the condensation in FAS. Malonyl-ACP is decarboxylated while the acyl-ACP (or acetyl-CoA) is transferred to a Cys residue of a KAS enzyme. The acyl is later condensed to the enolate ion resulting in a β-ketoacyl-ACP. Abbreviations: ACP – acyl carrier protein, Cys - cystine
2.2 Export of acyl-chains from the plastid

The nascent FFAs released from FAS are available for export to other subcellular compartments. Reports describe several processes that facilitate the export of FFAs, but the full process is yet not fully clarified and understood. Several enzymes involved in the acyl export from the plastid are known. Studies of an Arabidopsis transporter protein, fatty acid export 1 (FAX1), situated in the inner envelope of plastids suggest that FAX1 is involved in the transfer of FFA past membranes (Li et al., 2015). Studies also indicate the presence of long-chain acyl-CoA synthetases (LACS) on the surface of plastids that drives the export. LACS enzymes catalyse the formation of acyl-CoA from FFA and coenzyme A (CoA). Several LACS genes are present in Arabidopsis and LACS9, located in the plastid envelope is a significant contributor to acyl-CoA formation in purified plastids (Schnurr, 2002). An additional LACS enzyme, LACS1, localised in the endoplasmic reticulum (ER) also have an overlapping function with LACS9 (Zhao et al., 2010). LACS9, LACS1 and FAX1 are, however, not sole in the export of nascent FAs; other yet unknown concurrent processes facilitate the export of FAs from the plastid as well. The exported acyl-CoAs are then available, for glycerolipid assembly and other processes in the ER.

Rapid kinetic labelling of Arabidopsis cell suspension cultures with $[^{14}C]$acetate reveals a surprisingly short lag time between incubation and synthesis of radiolabelled phosphatidylcholine (PC). A hypothesis that explains such a short lag time includes channelling of acyl-CoA from LACS more or less instantaneously to PC via lysophosphatidylcholine acyltransferase (LPCAT) (Tjellström et al., 2012). LPCAT is an enzyme that catalyses the formation of PC from lysophosphatidylcholine (LPC) and acyl-CoA. For more details on this phenomenon, see pages 26-29.

Generally, ER is the subcellular organelle ascribed for LPCAT localisation, but a few studies also suggest a plastidial localisation in leek (Allium porrum L.) and pea (Pisum sativum L.) (Kjellberg et al., 2000; Bessoule et al., 1995). An extensive network of ER surrounds chloroplasts in Arabidopsis, and several membrane contact sites are apparent (Andersson et al., 2007). Protein-protein interactions probably maintain these contact sites. The observed LPCAT activity in the plastids may thus be contamination of remnant membranes from ER or ER localised proteins. Thus, LPCAT may be present in the ER, or both in plastids and ER. In any case, it is at least associated with contact sites between plastids and ER. There are strong indications that LPCAT mediated transfer of FAs between plastids and ER occurs irrespective of LPCAT’s actual subcellular membrane localisation.
Figure 4 overview of the synthesis of FAs in the plastid and the following export to the cytosol through proposed mechanisms. ACC produces malonyl-CoA from HCO$_3^-$ and acetyl-CoA. The malonyl-CoA is later thioesterifed to ACP. The first condensation catalysed by KAS III use malonyl-ACP and acetyl-CoA and results in 4:0-ACP after the first cycle of FAS. The FAS complex further elongates 4:0-ACP into acyl chains of up to 18 carbons length in a cyclic manner. KAS I facilitate the condensing reaction from 4:0-ACP to 16:0-ACP, while KAS II carries out the final condensation of 16:0-ACP to 18:0-ACP. SAD desaturate a large proportion of the nascent 18:0-ACP into 18:1-ACP. The elongated acyl chains are later used either by the prokaryotic lipid assembly which occurs in 16:3 plant species or released as FFA by FatA and FatB. FAX1 facilitates the transfer of FFA across the plastidial membranes where the FFA are thioesterifed with CoA by LACS9. LPACT may play an
essential role in many plant species and incorporate the nascent acyl-CoAs into PC. **Abbreviations:** CoA – coenzyme A, ACC - acetyl-CoA carboxylase, MCMT - malonyl-CoA:acyl carrier protein malonyltransferase, ACP – acyl carrier protein, KAS(I, II & III) - β-ketoacyl-ACP synthase, FAS - fatty acid synthase, SAD - Δ⁹ stearoyl-ACP desaturase, FatA/B – thioesterase, FFA – free fatty acid, FAX1 – fatty acid export protein 1, LACS9 - Long-chain acyl-CoA synthase 9, LPCAT - lysophosphatidylcholine acyltransferase, LPC – lysophosphatidylcholine, PC - phosphatidylcholine

2.3 TAG assembly in the endoplasmic reticulum

In developing oilseeds, a significant proportion of the nascent FAs end up as a constituent of TAG. Assembly of glycerolipids such as TAG in ER is dependent on FAs produced and exported from the plastids. TAG is produced in the ER by a range of acyltransferases and other enzymatic processes, see figure 5 for an overview. Diacylglycerol (DAG) is the final precursor of TAG, and the Kennedy pathway is a metabolic route that gives rise to *de novo* synthesised DAG subsequently converted into TAG (Weiss et al., 1960; Weiss & Kennedy, 1956). All known acyltransferases in the Kennedy pathway utilises acyl-CoA as acyl-donor. The first step in the Kennedy pathway is the acylation of the *sn*-1 position of glycerol-3-phosphate (G3P, figure 1C) carried out by glycerol-3-phosphate acyltransferase (GPAT), which results in lysophosphaticid acid (LPA) (Zheng, 2003). LPA is further acylated into phosphatidic acid (PA) at the *sn*-2 position by lysophosphatidic acid acyltransferase (LPAT) (Knutzon et al., 1995). Phosphatidic acid phosphatase (PAP) subsequently decouples the phosphate group at *sn*-3 and converts PA into DAG (Eastmond et al., 2010). DAG is besides being the precursor of TAG, also a precursor of PC and other phospholipids. CDP-choline:DAG cholinephosphotransferase (CPT) produces *de novo* PC from DAG and CDP-choline (Dewey et al., 1994).

PC is of central interest in TAG assembly, as many acyl desaturations and other FA modifications occur while acyl-chains are associated with PC. Further, there are indications as previously mentioned that most nascent FA exported from the plastid might not always enter lipid assembly directly through acyltransferases in the Kennedy pathway but initially though the acyl modification cycle of PC. The modification cycle of PC facilitates incorporation, modification and release of FAs from PC. Several enzymes incorporate FAs into PC in a range of different ways besides the previously described CPT. Phosphatidylcholine diacylglycerol cholinephosphotransferase (PDCT), for example, interconvert DAG and PC but gives no net gain of PC (Lu et al., 2009). The forward reaction of LPCAT incorporates FAs from the acyl-CoA pool at the
sn-2 position of LPC. LPC, in its turn, may be synthesised by another acyltransferase, GPCAT, which catalyses the formation of LPC from glycerol-3-phosphocholine (GPC) and acyl-CoA (Glab et al., 2016).

Figure 5 illustrates a simplified overview of TAG assembly in the ER. The Kennedy pathway includes the sequential acylation of G3P into TAG. GPAT acylates the sn-1 position of G3P with an acyl group derived from acyl-CoA into LPA which in turn is similarly acylated by LPAT into PA. PAP removes the phosphate group at sn-3 and converts the PA into DAG. DAG is a critical branch point in lipid metabolism since the production of both TAG and de novo synthesised PC utilises DAG. DGAT catalyses the final acylation at sn-3, which results in TAG. PDAT is also concurrently promoting the acyl-CoA independent synthesis of TAG, recruiting an acyl moiety from PC which transesterifies DAG into TAG. Acyl groups associated with PC are often modified, which is an essential feature in many plants’ lipid metabolism. Acyl groups can enter PC through several processes. PDCT interconverts the phosphocholine headgroup between PC and DAG enabling acyl groups from DAG to enter PC for modification. LPCAT enables rapid incorporation of acyl
groups into PC via the acylation of LPC and the reverse reaction of LPCAT releases potentially PC modified acyl-CoAs into the acyl-CoA pool. The eukaryotic pathway also utilises PC derived intermediates for the synthesis of thylakoid membrane lipids. **Abbreviations:** G3P - glycerol-3-phosphate, CoA - coenzyme A, GPAT - glycerol-3-phosphate acyltransferase, LPA - lysophosphatidic acid, LPAT- lysophosphatidic acid acyltransferase, PA - phosphatidic acid, PAP - phosphatidic acid phosphatase, DAG – diacylglycerol, DGAT - acyl-CoA:diacylglycerol acyltransferase, TAG – triacylglycerol, LPC – lysophosphatidylcholine, LPCAT - lysophosphatidylcholine acyltransferase, PC – phosphatidylcholine, PDCT - phosphatidylcholine diacylglycerol cholinephosphotransferase, PDAT - phospholipid: diacylglycerol acyltransferase

Several important modifications of FA occurs, and many UFAs gain their functional groups, while associated with PC. Several of the examples described below on page 35, figure 7, which illustrate modifications of FAs into UFAs while they are a moiety of PC.

As with the incorporation of FAs into PC, multiple pathways facilitate the release of FAs from PC. The reverse reaction of LPCAT releases FAs to the acyl-CoA pool (Lager et al., 2013). Phospholipase C and D, PDCT and potentially also the reverse reaction of CPT, facilitate the incorporation of FAs into substrates used in the Kennedy pathway for TAG synthesis (Bates et al., 2013; Slack et al., 1985). Further, another acyltransferase, phospholipid: diacylglycerol acyltransferase (PDAT), produces TAG through transacylation of acyl groups from PC to DAG (Dahlqvist et al., 2000).

Metabolic flux analyses of de novo synthesised FAs reveal that plants may have different strategies to incorporate nascent FA into glycerolipids as previously mentioned. The flux of nascent FAs in developing soybean embryos, for example, is far higher into the acyl modification cycle than into de novo lipid assembly via the Kennedy pathway (Bates et al., 2009). The soybean seed oil reflects the high flux of nascent FAs into PC modification cycle, where a majority of the FAs are of PC modified origin, such as the polyunsaturated FAs 18:2Δ9,12 and 18:3Δ9,12,15. Studies reveal similar results in several other species, e.g. safflower (Carthamus tinctorius L.), flax (Linum usitatissimum L.), and Arabidopsis (Bates & Browse, 2012; Griffiths et al., 1988; Slack et al., 1978). These are all species that boast a TAG FA composition that is rather high in FAs modified in the PC modification cycle. Other species, e.g. avocado (Persea americana Mill.), cigar flower (Cuphea ignea A. DC.), and cacao tree (Theobroma cacao L.) are species with a relatively low content of FAs edited during the PC modification cycle in their seed oil (Griffiths & Harwood, 1991).
Studies of these species indicate that these are more likely to incorporate the nascent FAs in the Kennedy pathway directly instead of through the PC modification cycle (Griffiths & Harwood, 1991; Bafor et al., 1990; Griffiths et al., 1988). Interestingly, coriander, containing high levels of UFAs not modified while associated with PC, still incorporate most UFAs into PC before TAG synthesis (Cahoon & Ohlrogge, 1994). In conclusion, these studies reveal that different plants rely on different routes for incorporating nascent FAs into glycerolipid synthesis.

Acyltransferases facilitate the final acylation of TAG regardless of which route that gives rise to the DAG available for TAG synthesis, as seen in figure 5. PDAT, as mentioned earlier, facilitates acyl transfer from PC to DAG, while acyl-CoA:diacylglycerol acyltransferase (DGAT) facilitates TAG formation via the Kennedy pathway. DGAT is the sole enzyme in the Kennedy pathway exclusively directed to TAG assembly and utilises acyl-CoA as acyl-donor, unlike PDAT.

DGAT is of great importance, both quantitatively and qualitatively in TAG synthesis (Shockey et al., 2006). As of yet, researchers have identified at least three types of DGAT. DGAT1 first identified in mice, and DGAT2 first cloned from an oleaginous fungus, and both are present in plants (Lardizabal et al., 2001; Cases et al., 1998). DGAT2 does not seem to contribute to TAG accumulation in developing seeds of Arabidopsis, whereas DGAT1 null mutants exhibit a decreased TAG accumulation (Zhang et al., 2009; Routaboul et al., 1999; Katavic et al., 1995). Consistent with these findings, overexpression of DGAT1 in Arabidopsis during seed development enhances the amount of TAG deposition in seeds (Jako et al., 2001). The DGAT2 of Arabidopsis is, however, an active catalytic DGAT enzyme and exhibits a substrate specificity that is different from that of the DGAT1 form (Ayme et al., 2014; Zhou et al., 2013). DGAT2 enzymes appear to be involved in the incorporation of UFAs into TAG in plant species expressing high levels of DGAT2 in their seeds. The DGAT2 enzymes of the tung tree (Vernicia fordii (Hemsl.) Airy Shaw), castor bean and ironweed (Vernonia galamensis (Cass.) Less.) are all capable of introducing UFAs into seed oil (Li et al., 2010a; Burgal et al., 2008; Shockey et al., 2006).

Both DGAT1 and DGAT2 enzymes are membrane-bound, located in the ER (Shockey et al., 2006). DGAT1 and DGAT2 differ from each other in many regards, albeit their related catalytic activity. The two classes distinguish from each other in membrane-topology, amino acid length, and they belong to different protein families (Turchetto-Zolet et al., 2011; Yen et al., 2008).
third class of DGAT present in plants, DGAT3, is soluble and is yet to be confirmed to be associated with seed oil accumulation (Saha et al., 2006). The three classes of DGAT are probably the result of convergent evolution and share no significant sequence homology (Aymé et al., 2018).

The TAG composition in oilseeds is hence, dependent on the abundance of the various acyltransferases, as well as their specificity towards acyl-donors and acyl-acceptors and the substrate availability and composition. The final acylation of DAG, which results in TAG, is facilitated by either DGAT or PDAT, and these are the final determinants of the TAG composition. PDAT and DGAT may concurrently contribute to the accumulating TAG pool. Their relative contribution differs not only between different plant species, but temporal and spatial variations co-exist as well (Li et al., 2010b). Spatial differences in TAG composition is present, not only within a single seed, but lipid composition heterogenicity is also present at tissue-level (Sturtevant et al., 2016).

It is crucial to keep in mind that many other lipids beside TAG, are assembled in the ER. Both PA and DAG produced in the Kennedy pathway are precursors for several other lipids, phosphatidylinositol (PI), phosphatidyglycerophosphate (PGP) and PC, phosphatidylethanolamine (PE), phosphatidylserine (PS), respectively. Further, FAs from PC are also substrates for the synthesis of thylakoid membrane lipids.

### 2.4 Acyl import to the plastids

Acyl groups are not only exported from the plastids to the ER as recently described, but FAs are, also, imported from the ER to the plastid. Thylakoid membranes within plastids consist mainly of the galactolipids, monogalactosyl diacylglycerol (MGDG) and digalactosyl diacylglycerol (DGDG) and the sulfolipid sulfoquinovosyl diacylglycerol (SQDG) (Dorne et al., 1990). These are lipids synthesised by two separate pathways, the prokaryotic and the eukaryotic (Heinz & Roughan, 1983). Plants fall into two categories; those that contain 16:3Δ⁷,10,13, e.g. spinach (Spinacia oleracea L.) and Arabidopsis, hence called 16:3 species while those devoid of 16:3Δ⁷,10,13 in photosynthetic tissues are called 18:3 plants, reflecting a mix of prokaryotic and eukaryotic and only eukaryotic pathways, respectively. The prokaryotic pathway depends on de novo synthesised FAs recruited from the plastidial acyl-ACP pool, while the eukaryotic pathway utilises FAs first exported to the ER and then returned to the plastid. The individual LPAT specificity within these two pathways produces a
hallmark that differentiates the lipids synthesised by either pathway (Roughan & Slack, 1984). The plastidial LPAT specificity favours acylating saturated acyl-chains of 16 carbons length while the ER located LPAT favours unsaturated 18 carbon acyl-chains. Lipids synthesised through the eukaryotic pathway, therefore, contain an 18 carbon acyl group at the sn-2 position while those from the prokaryotic contains a 16 carbon acyl group. The FAs in lipids produced by the eukaryotic pathway are, thus, synthesised in plastids and exported to the ER. In the ER these FAs are associated with PC, desaturated and reintroduced to the plastid, potentially further desaturated and used for assembly of thylakoid membrane lipid species as depicted in figure 6.

About 12 % of the angiosperms produce MGDG, DGDG, and SQDG through both pathways, while the majority seems to rely on the eukaryotic pathway alone (Mongrand et al., 1998). There are strong indications that a transporter system, constituted by the TGD1-5 proteins, facilitate the reintroduction of PC derived FAs into plastids in Arabidopsis (Karki et al., 2019). It is still however unclear which PC derived intermediate that is transported by this complex (Karki et al., 2019).
Figure 6 is an overview of two alternative metabolic routes present in Arabidopsis for the synthesis of thylakoid membrane lipids. The prokaryotic pathway use acyl-ACP substrates, where the substrate specificities of GPAT and LPAT generate sn-1-18:1-sn-2-16:0-DAG. Thylakoid membrane lipids derived from the prokaryotic pathway therefore always contain a 16:X acyl moiety at sn-2. The specificity of the eukaryotic acyltransferases, on the other hand, give rise to PC species that exclusively include an 18:X acyl moiety at sn-2. These are imported to the plastid by the transporter system TGD1-5. Thylakoid membrane lipids made from the eukaryotic pathway will thus have an 18:X acyl moiety at sn-2 instead of 16:X. Abbreviations: FAS – fatty acid synthase, SAD - Δ9 stearoyl-ACP desaturase, ACP – acyl carrier protein, G3P
- glycerol-3-phosphate, **GPAT** - glycerol-3-phosphate acyltransferase, **LPAT** - lysophosphatidic acid acyltransferase, **PAP** - phosphatidic acid phosphatase, **DAG** – diacylglycerol, **MGDG** - monogalactosyl diacylglycerol, **DGDG** - digalactosyl diacylglycerol, **SQDG** - sulfoquinovosyl diacylglycerol, **CoA** – coenzyme A, **LPC** – lysophosphatidylcholine, **LPCAT** - lysophosphatidylcholine acyltransferase, **PC** – phosphatidylcholine, **PDCT** - phosphatidylcholine diacylglycerol cholinephosphotransferase, **PA** - phosphatidic acid, **TGD(1-5)** – trigalactosyldiacylglycerol, **ER** – endoplasmic reticulum

### 2.5 Unusual fatty acid modifications

Plants can produce several hundreds of different UFAs. Modification of FAs into UFAs occurs at several different stages during FAs synthesis and lipid assembly.

Several UFAs are synthesised directly in connection to the fatty acid synthesis in the plastids. California bay, *Umbellularia californica* (Hook. & Arn.) Nutt., is a species known to accumulate high levels of medium-chain fatty acids (MCFA) (Hopkins et al., 1966). Transgenic canola (*Brassica napus* L.) expressing *FatB1* from California bay exhibit a significantly altered FA composition in the accumulated seed oil with drastically increased levels of laurate (12:0), which increases from 0.02 mol% in the wildtype to up to 60 mol% in transgenic lines (Voelker et al., 1996).

Another species with high levels of MCFA in its seed oil is Wright's waxweed, *Cuphea wrightii* A. Gray (Thompson & Kleiman, 1988). **KAS A1**, as well as **FatB1** and **FatB2** from *C. wrightii*, have been cloned and expressed in transgenic Arabidopsis. Expression of either *CwFatB1* or *CwFatB2* enhances levels of MCFA (10:0 and 12:0) in the Arabidopsis seed oil while the expression of *CwKAS A1* does not. However, if either of the two *CwFatB* is co-expressed with *CwKAS A1*, a further significant increase of 10:0 and 12:0 in the seed oil is observed. The additional increase of MCFA indicates that efficient production of MCFA may not only depend on fatty acid thioesterases alone but also, on specialised KAS enzymes during FA synthesis, see figure 4 for context (Leonard et al., 1998).

Many plant species contain plastidial acyl-ACP desaturases similar to the previously mentioned Δ⁹ SAD. These acyl-ACP desaturases are in several cases capable of utilising palmitoyl-ACP instead of stearoyl-ACP and introduces double bonds at novel positions compared to SAD. Unusual FAs produced through such deviating acyl-ACP desaturases include Δ⁶ palmitic acid (16:1Δ⁶)
in black-eyed Susan vine, *Thunbergia alata* Bojer ex Sims (Cahoon et al., 1994) and Δ⁴ palmitic acid (16:1Δ⁴) in coriander (Cahoon & Ohlrogge, 1994). Δ⁴ palmitic acid is the precursor of petroselinic acid, which coriander produces through subsequent elongation of Δ⁴ palmitoyl-ACP into petroselinoyl-ACP (18:1Δ⁵). Petroselinic acid with its deviating double bond position is suitable as a precursor of adipic acid (6:0 dicarboxylic), a widely used monomer for polymer production, and 12:0, a common constituent of surfactants and detergents (Cahoon et al., 1992).

A wide range of enzymes that modify FAs into UFAs is associated with the ER. The common FAs 18:2Δ⁹,₁₂ and 18:3Δ⁹,₁₂,₁₅, are desaturated while acylated at sn-2 of PC during the PC modification cycle. Fatty acid desaturase 2 (FAD2) desaturate 18:1Δ⁹ at Δ₁₂ into 18:2Δ⁹,₁₂ while FAD3 catalyse the further desaturation of 18:2Δ⁹,₁₂ into 18:3Δ⁹,₁₂,₁₅. FAD2 enzymes contain a conserved motif which is vital for enzymatic activity. The FAD2 motif consists of eight conserved histamine residues in three histamine boxes (Shanklin & Cahoon, 1998). This motif is also present in FAD2 from several organisms belonging to other phyla as well. The sequence homology and the FAD2 motif has aided in search of a wide range of FAD2-like enzymes responsible for modifications that produce UFAs (Cahoon & Kinney, 2005).

Castor bean produces high levels of 12-OH-18:1Δ⁹ in its seed oil. The Δ¹² hydroxylase (FAH12) that converts olate into ricinoleate in castor bean and the desaturation of 18:1Δ⁹ into 18:2Δ⁹,₁₂ catalysed by FAD2 share not only the modification on the same delta position but also the substrate, an oleoyl moiety of PC. Functional characterisation of castor bean FAD2-like cDNA in transgenic plants confirms that this FAD2-like enzyme, FAH12, indeed catalyses the hydroxylation of olate in castor bean (Van De Loo et al., 1995). Several other UFAs, e.g. vernolic acid (12,13-epoxy-18:1Δ⁹), crepenynic acid (12-alkyne-18:1Δ⁹), and several conjugated FAs (polyunsaturated FAs with conjugated double bonds, e.g. α-eleostearic acid, 18:3Δ⁹,trans¹¹,trans¹³) are all modified by FAD2-like enzymes, see figure 7 (Dyer et al., 2002; Cahoon et al., 2001; Lee et al., 1998). Many of these UFAs are of industrial interest. FAD2-like enzymes modify FAs into UFAs while associated with PC, but the PC pool contains minor levels of these UFAs. The UFAs modified by FAD2-like enzymes are efficiently removed from PC and rapidly incorporated into TAG (Bafor et al., 1991; Moreau & Stumpf, 1981).
Accumulating evidence indicates that yet another form of acyl desaturases exists in plants. These desaturases may utilise acyl-CoA instead of acyl-ACP or glycerolipids such as PC. Acyl-CoA desaturase-like genes (ADS) encode these desaturating enzymes which resemble those found in mammals and yeast (Smith et al., 2013). Meadowfoam, Limnanthes alba Benth. and several closely related species produce an oil quality that contains an eicosenoic acid with an unusual desaturation position, 20:1Δ5. Time course studies using radioactive substrates and microsomal preparations from developing seeds of meadowfoam reveal that this unusual desaturase activity is probably utilising acyl-CoA as substrate (Moreau et al., 1981). An expressed ADS gene cloned from Douglas meadowfoam, Limnanthes douglasii R. Br., desaturates a range of FAs at the unusual Δ5 position in somatic embryos of soybean (Cahoon et al., 2000). Anemone L. is another genus that contains an example of an ADS encoded enzyme that potentially utilises acyl-CoA. Sciadonic acid (20:3Δ5,11,14) represent nine per cent of the FAs in the seed oil of Anemone rivularis Buch.-Ham. ex DC. var. rivularis. Transgenic Arabidopsis expressing two A. rivularis var. rivularis derived ADS genes produce significant amounts of sciadonic acid. Further, there are strong indications that these enzymes indeed utilise acyl-CoA as substrate (Sayanova et al., 2007). Not all enzymes encoded by ADS genes in plants utilise acyl-CoA however, as at least one Arabidopsis ADS gene, AtADS3, encodes for FAD5 which catalyses the Δ7 desaturation on the 16 carbon sn-2 moiety of MGDG in Arabidopsis (Heilmann et al., 2004).

**Figure 7** Structural features of several UFAs, unusual features are highlighted in red. A Crepenyic acid B α-eleostearic acid, C vernolic acid and D ricinoleic acid.
Many plants produce seed oils containing FAs that are longer than that produced by plastidial FAS. Elongation beyond the acyl-chain lengths provided by the plastidial FAS results in very long-chain fatty acids (VLCFA), such as 22:1Δ13, abundant in many species within Brassicales. The elongase, a membrane-bound enzyme-complex which closely resembles that of FAS in the plastid, elongate FA into VLCFA (von Wettstein-Knowles, 1982). The initial step of elongation catalysed by β-ketoacyl-CoA synthase (KCS) condensate malonyl-CoA and a long-chain acyl-CoA into β-ketoacyl-CoA. The elongase complex further reduces and dehydrates the β-ketoacyl-CoA, which results in a two carbon elongation of the initial long-chain acyl-CoA.

Albeit several similarities between the FAS and elongase complex, see figure 2, they differ in a few aspects. The ER localised elongase complex utilise moieties associated with CoA while the plastidial FAS requires that at least one of the substrates is a moiety of ACP. Both the KAS enzymes that facilitate condensation in the FAS and the KCS of the elongase complex belong to the thiolase superfamily. In FAS, KASIII catalyses the initial condensation and is the only condensing enzyme that utilises acetyl-CoA as substrate. Both KCS and KAS III share the same active site architecture, which is different from that of KAS I and KAS II. Phylogenetic analyses also indicate that KCS and KAS III are more similar to each other than KAS III is to the KAS I and KAS III (Jiang et al., 2008).

The FAE1 gene in Arabidopsis encodes a seed-specific KCS responsible for the condensation initiating the elongation of FA into VLCFA. Studies reveal that the activity of KCS is rate-limiting in the elongase complex. The KCS enzyme also determines the end product acyl length through its acyl-CoA specificity (Millar & Kunst, 1997). The identified FAE1 homologues of T. majus (Mietkiewska et al., 2004) and Crambe hispanica subsp. abyssinica (Hochst. ex R.E. Fr.) Prina (Mietkiewska et al., 2007) are capable of initiating the elongation of various acyl-CoA into erucic acid, which is a significant FA in these species seed oils.

The recent identification of two UFAs, nebraskanic acid 7,18-OH-24:1Δ15 and wuhanic acid 7,18-OH-24:2Δ15,21, and the following elucidation of the biosynthetic synthesis of these, presents yet another mechanism, for the synthesis of UFAs (Li et al., 2018). Nebraskanic and wuhanic FAs comprise up to 50 wt% of the seed oil from the Brassicaceae species Orychophragmus violaceus (L.) O.E. Schulz. Several genes identified in O. violaceus are most likely involved in the synthesis of these VLCFAs that contains two hydroxyl
groups at Δ⁷ and Δ¹⁸. The synthesis of both these FAs starts with an 18 carbon moiety associated with PC and includes a Δ¹² hydroxylase and the following elongation process, which also introduces the second hydroxyl group. An identified FAD2-like gene, OvFAD2-2, encodes for an enzyme that contains an amino acid substitution situated before the first histamine box known to occur in FAD2-like enzymes, e.g. the castor bean FAH12, which catalyses the introduction of hydroxyl groups at Δ¹² of FAs (Cahoon & Kinney, 2005).

Monohydroxy FAs, e.g. 12-OH-18:1Δ⁹ and 12-OH-18:2Δ⁹,¹⁵, are present in the seed oil of transgenic Arabidopsis expressing OvFAD2-2, but the seed oil is still devoid of dihydroxy FA (Li et al., 2018). An elongase gene, OvFAE1-I, encoding for a divergent KCS enzyme is capable of condensing the monohydroxy 18 carbon FAs with malonyl-CoA. Reduction of the resulting 3-ketotacyl-CoA into 3-hydroxyacyl-CoA seems to occur as in the generic elongation process catalysed by the elongase complex. At this point, the divergent OvFAE1-I enzyme probably discontinues the following processes occurring by the generic elongase complex, which includes dehydration and the final reduction, and the elongase process restarts with the novel 20 carbon dihydroxyacyl-CoA, probably catalysed by another FAE (Li et al., 2018). The subsequent two rounds of elongations from 20 into 24 carbon, seem to follow the complete cycle of the generic elongase process in O. violaceus and results in nebraskanic FA, see figure 8. The discontinued elongation thus represents yet an additional way to introduce hydroxyl groups at novel positions compared to FAD2-like hydroxylases. The synthesis of wuhanic acid (7,18-OH-24:2Δ¹⁵,²¹) follow the same route as for nebraskanic acid except that 12-OH-18:1Δ⁹ is desaturated into densipolic acid, 12-OH-18:2Δ⁹,¹⁵, before elongation with the divergent KCS.

Transgenic Arabidopsis expressing both OvFAD2-2 and OvFAE1-I produces nebraskanic FA while wuhanoic FA production requires incorporation on an additional elongase, OvFAE1-2. O. violaceus may also contain additional specialised enzymes facilitating the reductions and dehydration in the elongase complex (Li et al., 2018).
The diversity of UFAs in plants is vast, and even though scientists have elucidated how several plants can modify them into their final configuration, the major part of their biosynthesis are yet unresolved. Researchers have also made significant progress on, albeit not yet wholly clarified, how plants can channel these UFAs into TAG through acyltransferases with pronounced specificities. Plants producing high levels of UFAs have often developed mechanisms excluding UFAs in membrane lipids, probably to avoid potentially detrimental effects in the membranes (Millar et al., 2000). The evolutionary advantage of preserving a significant proportion of UFAs in the seed oil, or derivatives of these, might come from their novel properties. Some seed oils high in UFAs are known to be laxative, emetic and poisonous and may thus prevent herbivores from consuming the plant. Other UFAs are a suitable precursor for substances with protective effects against pathogens.

**Figure 8** depicts a proposed biosynthetic pathway, resulting in nebraskanic acid in *O. violaceus*. OvFAD2-2 introduce a hydroxyl group at Δ\(^{12}\) on oleate acid associated with PC. A discontinued elongation cycle incorporates a second hydroxyl group at Δ\(^{3}\). Two additional complete elongation cycles give rise to nebraskanic acid. The individual steps of the elongation cycle resemble that of fatty acid synthase where CoA substrates are used instead of ACP substrates, see figure 2. **Abbreviations:** FAD2 - Fatty acid desaturase 2, CoA – coenzyme A, LPC – lysophosphatidylcholine, PC – phosphatidylcholine.

The diversity of UFAs in plants is vast, and even though scientists have elucidated how several plants can modify them into their final configuration, the major part of their biosynthesis are yet unresolved. Researchers have also made significant progress on, albeit not yet wholly clarified, how plants can channel these UFAs into TAG through acyltransferases with pronounced specificities. Plants producing high levels of UFAs have often developed mechanisms excluding UFAs in membrane lipids, probably to avoid potentially detrimental effects in the membranes (Millar et al., 2000). The evolutionary advantage of preserving a significant proportion of UFAs in the seed oil, or derivatives of these, might come from their novel properties. Some seed oils high in UFAs are known to be laxative, emetic and poisonous and may thus prevent herbivores from consuming the plant. Other UFAs are a suitable precursor for substances with protective effects against pathogens.
2.6 Engineered seed oil quality through conventional plant breeding.

*B. napus*, is a dual-purpose crop where varieties belonging to oilseed rape contains high levels of erucic acid whereas canola varieties are virtually free from erucic acid. Breeders developed canola or low erucic acid rapeseed (LEAR), in response to raised concerns about potentially detrimental health issues of consuming erucic acid (Beare *et al.*, 1957). The developing seed embryo of *B. napus* constitutes a substantial part of the *B. napus* seed, and it is the embryonic genotype, not the maternal, that determines the oil quality in *B. napus* (Downey & Harvey, 1963). The erucic acid content in *B. napus* seed oil is to a great extent determined by two loci, E1 and E2, which have a quantitative additive effect (Harvey & Downey, 1964). Breeders developed lines with two Mendelian recessive alleles, e1 and e2, which decreases the erucic acid content drastically. Two FAE1 genes, *FAE1.1* and *FAE1.2*, are colocalised with E1 and E2, respectively. Both *FAE1.1* and *FAE1.2* encodes KCS enzymes catalysing the rate-limiting condensation in the elongase complex (Fourmann *et al.*, 1998). *B. napus* probably arose in a spontaneous cross between *Brassica rapa* L. and *Brassica oleracea* L. and *B. napus* is thus an amphidiploid. The amino acid sequence of FAE1.1 is similar to that of FAE1 from *B. rapa* and FAE1.2 to that of *B. oleracea* (Fourmann *et al.*, 1998). The BnFAE1.2 allele from canola contains a frameshift leading to an incomplete protein due to a premature stop codon (Fourmann *et al.*, 1998). Comparison of FAE1.1 amino acid sequences from the rapeseed cultivar Askari and the canola variety Dakkar reveal that a single amino acid substitution renders the FAE1.1 enzyme of the canola FAE1 inactive (Han *et al.*, 2001). Continuous improvement of canola through breeders’ efforts during almost 60 years development has led to an oil crop with seeds free from erucic acid and with a low glucosinolate content.

Breeders have, in parallel to the development of canola, intentionally increased the erucic acid content in the seed oil of high erucic acid rapeseed (HEAR). The increased demand for erucic acid and its derivates has intensified the development of HEAR varieties. Traditional *Brassica* oilseed crops produce seed oils with approximately 45 mol % of erucic acid. Conventional breeding efforts have increased the erucic acid content up to 50-55 mol % in commercial varieties (McVetty & Scarth, 2002). Stereopositional distribution of erucoyl moieties in seed oil TAG within Brassicaceae reveals that erucic acid is in most cases virtually absent on the *sn*-2 position (Taylor *et al.*, 1994). LPAT acylates the *sn*-2 position of LPA into PA in the Kennedy pathway, see figure 5. Determination of *B. napus* LPAT substrate specificity strongly indicate that it is incapable of using 22:1Δ¹³-CoA as acyl donor. The inability of *B. napus* LPAT

39
to acylate the sn-2 position with erucic acid efficiently restricts the theoretical maximum concentration of erucic acid in TAG to 2/3 (Cao et al., 1990).

The inherent theoretical maximum concentration of erucic acid in *B. napus* is challenging to circumvent with conventional breeding without breeding material containing a suitable LPAT. Breeders and researchers have thus resorted to modern biotechnological approaches instead of trying to further increase the erucic acid content by conventional breeding.

### 2.7 Engineered seed oil quality through a biotechnological approach

Genetic engineering provides researchers and breeders with tools and opportunities that enable the production of novel oil qualities in oil crops. Many of these oil qualities are impossible or extremely time-consuming to achieve with conventional breeding.

Researchers have screened the substrate specificity of LPATs from several species with a high content of VLCFA to identify candidate genes capable of circumventing the inherent inability of *B. napus* to acylate the sn-2 position of LPA with erucic acid. Two LPATs from species within the genus *Limnanthes*, *L. alba* and *L. douglasii*, LdLPAT and LaLPAT respectively, have the catalytic capability to facilitate the incorporation of erucic acid at the sn-2 position of LPA (Löhden & Frentzen, 1992; Cao et al., 1990). Species within the genus *Limnanthes* produce seed oils with a TAG composition comprised almost exclusively of VLCFA, far beyond the 2/3 limit of *B. napus*. Δ⁵-eicosenoic acid (20:1Δ⁵) comprises the majority of the VLCFA in the seed oil in *Limnanthes* species, but the two LPATs accepts erucoyl-CoA nonetheless (Miller et al., 1964). Transgenic *B. napus* expressing the LdLPAT or LaLPAT in seeds during development does not gain an increase erucic acid content compared to the wildtype (Weier et al., 1997; Brough et al., 1996; Lassner et al., 1995). The erucic acid is, however, more evenly distributed at all three sn positions (Brough et al., 1996; Lassner et al., 1995). Suspicions that the elongase complex in *B. napus* is unable to provide sufficient amounts of erucic acid to the acyltransferases in these transgenic plants led scientists to implement a more elaborate approach.

The elongase complex in *B. napus* utilises 18:1Δ⁹ for elongation. As do FAD2 and indirectly FAD3 for desaturation during the synthesis of 18:2Δ⁹,₁₂ and 18:3Δ⁹,₁₂,₁₅. The concurrent elongation and desaturation processes thus compete
for the same substrate, 18:1Δ⁹. Scientists have cloned \textit{FAE1.1} and \textit{FAE1.2} from a HEAR cultivar and co-expressed them with \textit{LdLPAT} in a breeding line named ‘HELP’. Some of these transgenic plants managed to produce seed oils with up to 72 % erucic acid in greenhouse trials. On average, there was an increase of 9 % of erucic acid in the transgenic plants compared to the breeding line ‘HELP’ (Nath \textit{et al.}, 2009).

‘HELP is an acronym for ‘High Erucic Low Polyunsaturated’, and the seed oil of this breeding line contains approximately 50 % erucic acid and 7 % polyunsaturated FAs (PUFA), predominantly 18:2Δ⁹,₁₂ and 18:3Δ⁹,₁₂,₁₅. Another transgenic \textit{B. napus} produce a seed oil which has a fatty acid composition that resembles that of ‘HELP’. The seed-specific downregulation of \textit{FAD2} in these transgenic plant decreases the amount of 18:2Δ⁹,₁₂ and 18:3Δ⁹,₁₂,₁₅ in the seed oil to less than 5% compared to approximately 20 % in the wildtype \textit{B. napus}. The erucic acid content increases at the same time, from 42 % to 45 % (Shi \textit{et al.}, 2017). While the seed-specific downregulation of \textit{FAD2} through an RNAi construct significantly affect the seed oil composition, it does so without other severe phenotypic changes of the plant (Shi \textit{et al.}, 2017).

Crambe is an alternative oil crop suitable for the production of erucic acid. Crambe has several benefits compared to HEAR. Both the seeds and the seedpods are distinctly different from that of the cultivated \textit{Brassica} ssp., reducing the risk of misidentification between a seed lot intended for food and high erucic containing material. Further, crambe does not seem to hybridise with cultivated \textit{Brassica} crops, which is an advantage compared to HEAR (FitzJohn \textit{et al.}, 2007). Cultivation of HEAR inevitably includes risks of erucic acid contamination of canola. HEAR cultivated in the vicinity of canola, for example, can facilitate cross-pollination (Warner & Lewis, 2019). The embryonic genotype determines the FAs composition in \textit{B. napus} and the absence of erucic acid in canola relies, at least partially, on Mendelian recessive alleles. No such consideration must be taken into account while producing erucic acid in crambe as it does not cross-pollinate with cultivated \textit{Brassica} ssp. Crambe is currently already regarded as a source for industrial feedstock rather than food, and the erucic acid content in the seed oil is comparable to that of HEAR (Lazzeri \textit{et al.}, 1997).

An attempt to increase the erucic acid content in a crambe cultivar ‘Galactica’ has sofar been modestly successful. The \textit{sn-2} position in ‘Galactica’ seed oil TAG is, as in \textit{B. napus}, very seldom acylated with an erucoyl moiety. The total erucic acid content is, however, almost 60 %, and the amount of PUFA
constitutes 14 % (Li et al., 2012). Li et al. (2012) implemented an approach based on the gained experience of transgenic HEAR; including a down-regulation of CaFAD2 through an RNAi in concert with the overexpression of BnFAE1 and LdLPAT. Analysis of the seed oil from the transgenic crambe reveal a significant increase in erucic acid content, up to 73 %, similar to the best-performing transgenic HEAR lines (Li et al., 2012; Nath et al., 2009). A more even distribution of erucic acid between the three sn positions is apparent in the seed oil, albeit low amounts of trierucin (Li et al., 2012). There is, however, no net gain of erucic acid in the transgenic crambe; due to a yield penalty on an oil per seed basis (Ortiz et al., 2020). The erucic acid content in the DAG pool is much higher in developing seeds of the transgenic lines compared to the cultivar ‘Galactica’ throughout the seed development. Further, the difference in erucic acid content between ‘Galactica’ and the transgenic lines is more articulated in DAG than it is in TAG during seed development (Guan et al., 2014).

The highest level of erucic acid content recorded is from, a perhaps unexpected plant, Tropaeolum majus L., garden nasturtium (Ohlrogge et al., 2018; Taylor et al., 2010). T. majus has gained much attention due to the seed oil that contains almost 80 % erucic acid. The seeds of T. majus are however not very suitable for oil production as they contain only ~10 % oil. Studies of three identified genes from T. majus involved in lipid metabolism, TmLPAT, TmDGAT1 and TmFAE indicate that these are potential candidates for increasing the erucic acid content in transgenic crops (Taylor et al., 2010; Xu et al., 2008; Mietkiewska et al., 2004). Expression of TmDGAT1 in a HEAR breeding line increases the oil content by 3-8 %, but there is, unfortunately, no significant increase of erucic acid content (Xu et al., 2008).

Another area of interest where scientists have identified opportunities for improvement with engineered oil qualities is for feed in aquaculture and the production of long-chain ω-3 PUFA (LC-PUFA) in oil crops. ω-3 in this context refers to a group of long-chain PUFAs where the last double bond is situated three carbons from the ω carbon. One of the outcomes during the United Nations world summit on sustainable development in 2002 regarding fisheries was to:

‘Maintain or restore stocks to levels that can produce the maximum sustainable yield (MSY) with the aim of achieving these goals for depleted stocks on an urgent basis and where possible not later than 2015’ (Ye et al., 2013).

Nevertheless, assessments estimate that more than 2/3 of monitored fish populations accounting for 80 % of the global fish populations in 2008 were overharvested (Ye et al., 2013). The annual catch by global fisheries is stable
since the 1990s and yields approximately 90 million tonnes of fish per year. There is a continuously increased demand for fish due to a growing human population, and approximately 16% of the total animal-protein intake of humans are of marine origin. The yield of fish in aquaculture has already increased tenfold from 1980 to 2010. Predictions estimate that the yields of aquacultured fish will exceed that of the global fisheries by 2030 as a response to meet human’s demand for more fish (Kobayashi et al., 2015).

Several studies advocate dietary intake of oily fish, due to the beneficial effects of LC-PUFAs, especially docosahexaenoic acid (DHA, 22:6Δ4,7,10,13,16,19), and eicosapentaenoic acid (EPA, 20:5Δ5,8,11,14,17) (Harris, 2004). Most animals, including fish species, are unable to de novo synthesise PUFAs, with a few exceptions within invertebrates. PUFAs are, therefore, an essential part of the dietary intake for most animals (Malcicka et al., 2018). Successful aquaculture, therefore, relies on PUFA supplemented feed. The LC-PUFAs in wild fish originates from single-celled marine organisms that fish eat (Harris, 2004). Salmonid fishes are predatory and popular within aquaculture. Salmonids can, to some extent, synthesise both EPA and DHA from plant seed oils containing 18 carbon PUFAs (Tocher, 2015). Fish fed with PUFA containing plant oils are, however, not able to maintain the high levels of LC-PUFA as fish fed with fish oil or wild living relatives (Tocher, 2015). Fish feed for aquacultures, therefore, contains fish oils derived from wild fish populations. The price of these commodities has recently increased drastically, and predictions indicate that they will continue to increase (Kobayashi et al., 2015). Feeding aquaculture with marine-based feed products also counteracts one of the benefits with aquaculture as it endangers the marine populations through potential overharvesting.

Scientists have, for the last two decades, investigated the possibility to produce LC-PUFA in oil crops to meet the increasing demand (Abbadi et al., 2004; Qi et al., 2004). Small aquatic organisms, e.g. microalgae, bacteria and diatoms can synthesise LC-PUFA via two distinct pathways, the aerobic and the anaerobic. The anaerobic pathway occurs predominantly in bacteria, which produces LC-PUFA cyclically from malonyl-CoA and acetyl-CoA. PUFA synthase is a multi-enzyme complex that catalyses the two-carbon elongations and desaturations in the anaerobic pathway. Amino acid alignment of PUFA synthase domains reveals that PUFA synthase represents a hybrid between polyketide synthases (PKS) and FAS (Metz et al., 2009). The aerobic pathway is more similar to the FAs modification and synthesis that occurs in oil crops than the anaerobic (Napier et al., 2015). Several oil crops already produce
significant amounts of precursors to the aerobic pathway. In the aerobic pathway, consecutive desaturation and elongation of 18:2Δ⁹,₁₂ results in DHA and EPA, see figure 9. Two intertwined routes, ω-3 and ω-6, exists within the aerobic route. The ω-3 route relies on 18:3Δ⁹,₁₂,₁₅ while the ω-6 route uses 18:2Δ⁹,₁₂. The aerobic pathway also contains additional alternative routes. Desaturation of 18 carbon substrates by Δ⁶ desaturases and subsequent elongation predominates over the alternative route where elongation occurs before desaturation of a Δ⁸ desaturase, see figure 9 (Napier et al., 2015).

Figure 9 illustrates several possibilities for the synthesis of both EPA and DHA. Multiple alternating desaturations (pink) and elongations (blue) of the plant available 18:2Δ⁹,₁₂ and 18:3Δ⁹,₁₂,₁₅ mimics the aerobic pathway of many microalgae (Napier et al., 2015; Petrie & Singh, 2011). Abbreviations: DHA - docosahexaenoic acid, EPA - eicosapentaenoic acid
Scientists have identified all the genes coding for the enzymes responsible for the synthesis of LC-PUFA within both the aerobic and anaerobic pathway (Petrie & Singh, 2011). Extensive and laborious efforts have resulted in a successful production of LC-PUFA in transgenic oil crops. Dow AgroSciences LLC and DSM Nutritional Products LLC presented in 2016 canola synthesising LC-PUFA through the anaerobic pathway. Field-grown transgenic canola produces a seed oil that contains up to 4.4 % DHA and 0.6 % EPA (Walsh et al., 2016).

Other research groups have successfully implemented the aerobic pathway in both canola and Camelina sativa (L.) Crantz (Napier et al., 2019). The production of LC-PUFA through the aerobic pathway in oil crops is by no means a small feat; multiple components have to function coherently for efficient synthesis of LC-PUFA and subsequent incorporation into TAG, as seen in figs. 5 and 9. Two groups, BASF/Cargill and Nuseed/CSIRO/GRDC, have in parallel developed transgenic canola varieties using slightly different approaches, but with similar amounts, ~12 % LC-PUFA in the seed oil (Napier et al., 2019). The two canola varieties differ however in the FA composition, one with high levels of EPA and the other with high levels of DHA. The two groups used rather large constructs, consisting of more than 32 kb, containing either seven or 11 genes, dedicated to LC-PUFA that originates from a wide range of organisms (Connelly & MacIntosh, 2018; Sottosanto et al., 2018).

Transgenic C. sativa is also an up-and-coming candidate for producing LC-PUFA and can produce a seed oil containing very high levels of these FAs, exceeding the publically announced levels achieved in transgenic canola. The seed oil of transgenic C. sativa lines can, so far, contain up to 24 % EPA (Ruiz-Lopez et al., 2014). The highest published amount of DHA produced in a C. sativa line is currently ~12 %, and this line contains ~4 % of EPA as well (Petrie et al., 2014). A transgenic line of C. sativa, worth mentioning in this context is not producing the highest amounts of LC-PUFA, but the composition of both EPA and DHA, 11 % and 8 % respectively, and the overall of FA composition of the seed oil roughly resembles that of bulk fish oil (Ruiz-Lopez et al., 2014).
The examples presented above illustrate that it is possible to engineer an oil quality with desirable traits using a biotechnological approach. It is not merely a vision with future applications. At present, we see several instances where transgenic oil crops already can replace harmful practices and replace less environmentally friendly alternatives. A robust understanding of plant lipid metabolism and the biochemical characterisation of involved enzymes underpins the future production of renewable oleochemical commodities.
3 Principles and justification of reoccurring methods

The papers included in this thesis make use of radiolabelled $[^{14}C]$ substrates and microsomal preparations. The inherent nature of integral enzymes such as those in the Kennedy pathway precludes structural determination by protein X-ray crystallography. While biochemical in vivo studies can reveal essential fluxes and processes, parallel processes can make it virtually impossible to interpret the data for a single enzyme. In B. napus, for example, eight DGATs concurrently facilitate the acyl-CoA dependent formation of TAG, while PDAT synthesises TAG from PC and DAG simultaneously. To determine the substrate specificity of a single DGAT in such a complex environment is, to say the least very challenging, if not impossible.

Introduction, overexpression or downregulation of a gene in stable transformants can provide valuable information but is time-consuming. Transient gene expression of Agrobacterium tumefaciens Smith & Townsend in, for example, Nicotiana benthamiana Domin is another alternative but is dependent on the endogenously available substrates. Co-expression of additional transgenes facilitating the synthesis of the desired substrate may thus be necessary.

We base most of the investigations we carry out as described in this thesis on microsomal preparations derived from either developing plant embryos or yeast cultures. Carefully selected substrates and appropriate buffers enable detailed studies of integral proteins present in the microsomal membranes. Microsomal preparations consist of reconstituted vesicles derived from membranes of the ER. We recover the microsomal membranes through ultracentrifugation of homogenised cells after removal of larger cell debris. We frequently express integral plant-derived lipid metabolism involved genes in yeast. Yeast is capable
of producing TAG, and the four genes \textit{ARE1}, \textit{ARE2}, \textit{DGA1} and \textit{LRO1}, contribute to the synthesis of TAG. Disruption of these four genes results in yeast, still viable but devoid of TAG. We, therefore, often utilise such a TAG devoid yeast, yeast strain H1246 (Sandager \textit{et al.}, 2002), to study lipid metabolism enzymes see, figure 10. The TAG produced by recombinant H1246 expressing, for example, a \textit{DGAT} is thus exclusively produced by the transgene in question. Yeast cultures are also facile to transform, cultivate and maintain.

![Diagram of yeast cultures transfection](image)

\textit{Figure 10} illustrates the principles behind expressing TAG producing acyltransferases in H1246. \textbf{A} Wildtype yeast is fully capable of producing TAG. \textbf{B} The disruptions of four genes in H1246 results in a viable yeast strain, H1246, devoid of TAG. \textbf{C} Introduction of a DGAT restores the capability to produce TAG, \textbf{D} once expressed. The generated TAG is hence the product of the introduced enzyme, enabling detailed biochemical characterisation of such acyltransferases

Most lipid substrates, both radiolabelled and unlabelled, are meticulously synthesised in our lab from FAs based on Sten Stymne lifelong experience within biochemistry. Many of these substrates are currently unavailable on the market or prohibitively expensive to procure. We produce acyl-CoA, various DAG-, PC-, LPC-, LPA- TAG-, and unfortunately, now and then, inadvertently MAG-species on a routine basis. These substrates are an absolute necessity for the methods we apply to characterise the enzymes involved in lipid metabolism.

Our assays consist of defined amounts of necessary substrates and microsomal preparations diluted in an appropriate buffer. We incubate the assays during a defined period and the combination of microsomal protein amount, time and substrate are usually adjusted to ensure close to linear conditions during the assays. After we have performed the assays, we analyse the lipids (extracted into chloroform), both remaining, degraded and produced during the incubation. We use either thin-layer chromatography (TLC) or high-performance liquid chromatography (HPLC) to separate the products in parallel with a liquid
scintillation counter (LCS), which determines the total radioactivity in an aliquot of the chloroform extract. A Packard InstantImager Electronic Autoradiography system determines the relative contribution of the total radioactivity as determined by LCS from each lipid class separated by TLC. The radioactivity detector Raytest Ramona Star determines the relative contribution of lipid classes in samples separated by HPCL.

We have previously generally used [14C] di-6:0-DAG as acyl-acceptor during characterisation of DGAT enzymes, due to the water-solubility of this DAG species. Longer acyl-chain-DAG species such as di-18:1Δ⁹-DAG does not readily dissolve in buffers and is thus challenging to add to the microsomal preparation in a way that makes them accessible for the enzymes. We have recently developed a robust method to conduct assays with various long-chain-DAG species, which uses DAG species dissolved in a small volume of DMSO.
4 Results and insights derived from the papers

Results at a glance

Paper I
The DGAT2 enzymes from *B. napus* with similar amino acid sequences fall into two categories based on their drastically different acyl-CoA specificities when assessed with di-6:0-DAG.

Paper II
The acyl-CoA specificity of DGAT2 enzymes is significantly affected by a specific region residing in the first third of the amino acid sequence. The identified segment contains two predicted transmembrane helices which are abundant in a wide range of plant-derived DGAT2 enzymes.

Paper III
Several DGAT genes, of both DGAT1 and DGAT2 present in the crambe genome, produce enzymes that lack or exhibit severely reduced capacity to use di-22:1Δ13-DAG as a substrate. The reduced ability to acylate di-22:1Δ13-DAG by these DGAT enzymes may pose a significant drawback in the pursuit to achieve a seed oil with a very high content of erucic acid in crambe.

Paper IV
We elucidate the elusive DGAT2 sequence of *T. majus*, which is a potential candidate for circumventing the bottleneck identified in paper III. biochemical characterisation of both TmDGAT1 and TmDGAT2 reveal that these too, just as the crambe DGAT enzymes, are either unable or possess a reduced capacity to acylate di-22:1Δ13-DAG.

Paper V
Here we characterise several essential enzymes involved in lipid metabolism from *C. sativa*. The characterised DGAT2 exhibit a similar specificity as one of the DGAT2 categories of *B. napus* assessed in paper I. Further, the *C. sativa* PDCT is unable to use ricinoyl containing DAG species. The PDCT is not only unable to use ricinoyl containing DAG species, but the presence of ricinoleoyl containing DAG species also seem to inhibit head group interchange by PDCT of other DAG species into PC as well.
Careful considerations regarding the compatibility of the intrinsic lipid synthesis are essential when introducing the synthesis of exogenous FAs or in other ways re-engineering the TAG FA composition in an oil crop. Several factors within the intrinsic lipid synthesis will affect the outcome. The modification of FAs into UFAs often require that a specific FA is associated with different lipids throughout its modification to its final, desired configuration. For example, many, but not all desaturases, utilise FA-moieties of PC as substrates while the extraplastidic elongase complexes in plants utilise acyl-CoA as a substrate which requires an efficient interchange of FAs between PC and acyl-CoA. The substrate specificity of the endogenous enzymes involved in lipid metabolism will thus inevitably affect the result. Concurrent modifications of FAs and lipid assembly may also compete for the initial substrate or its intermediates. Spatial and temporal variations of lipid species occur within developing seeds of several species, as is the case with seeds of, e.g. B. napus and C. sativa (Marmon et al., 2017; Usher et al., 2017; Woodfield et al., 2017). Further, the enzymatic capacity of acyltransferases within microsomal preparations from developing seeds is known to vary drastically between different developmental stages (Furmanek et al., 2014).

All of these aspects, and several additional, affects the final FA composition in the final seed oil. Adjustments that counteract any potential incompatibility between the endogenous lipid metabolism and the desired products may very well add to the complexity of an approach to engineer the seed oil composition. Rigorous biochemical characterisation of the lipid metabolism within a plant provides vital pieces of information, aiding the design process of how to alter the FA composition of a seed oil successfully.

The main focus in this thesis is on acyltransferases, DGATs in particular, and especially the DGAT2 forms. DGAT is a significant contributor to the accumulating seed oil in many plants. The acyl-donor specificity directly affects the FAs composition at the sn-3 position of TAG. The acyl acceptor specificity of DGAT, however, may act as a filter and exerting a distinct selection of DAG species to be acylated. Thus, DGAT specificity may affect the FAs composition on all three sn positions in TAG. Several plants producing UFAs contains DGAT2 enzymes that possess specificities prone to prefer or at least readily accept the abundant UFA. The systematic biochemical characterisation of DGAT enzymes aids the fundamental understanding of lipid metabolism. Further, an increasing number of characterised enzymes enables the recognition of patterns of motifs related to the specificity emerging within the DGAT2 amino acid sequences.
In paper I, we provide a robust characterisation of several DGAT enzymes derived from *B. napus*. The *B. napus* genome as an amphidiploid contains four genes for each DGAT. We identified four *DGAT1* alleles and five *DGAT2* alleles during a screening of DGAT genes present in the genome of the canola variety ‘Monolit’ at the HEAR variety ‘Maplus’. Two of the *DGAT2* alleles representing *BnaA.DGAT2.c* differed slightly on the amino acid level between the two cultivars but behaved similarly in later acyl-CoA specificity determinations. We assessed the acyl-CoA specificity for each of the enzymes encoded by the identified alleles with the artificial acyl acceptor di-6:0-DAG in recombinant yeast. The four *DGAT1*, all exhibited a similar acyl-CoA specificity while the *DGAT2* specificities differentiate. The *DGAT2* enzymes fall into two categories based on their acyl-donor specificities, those that primarily accept 18:3Δ9,12,15-CoA and those that accept 22:1Δ13-CoA at similar rates as 18:3Δ9,12,15-CoA, see figure 11A and B. These two groups of *DGAT2* enzymes that differ starkly in the acyl-CoA specificity have however very similar amino acid sequences. Phylogenetic analyses indicate that both ancestral parent of *B. napus*, *B. rapa* and *B. oleracea*, provided one of each type of the *DGAT2* categories, see fig 11C. All *DGAT2* genes are, however, annotated on chromosomes from the A genome, derived from *B. rapa*. 
Figure 11  

A The acyl-CoA specificity of *BnaA.DGAT2.b* is representing the category of *B. napus* DGAT enzymes that prefer 18:3-CoA.  

B The acyl donor specificity of *BnaA.DGAT2.d* represents the category of *B. napus* DGAT enzymes that prefers both 18:3Δ⁹,12,15-CoA and 22:1Δ¹³-CoA at a similar level.  

We assessed both DGAT2 enzymes activities with di-6:0-DAG as acyl donor under close to linear conditions. Error bars denote standard deviation, n= 3.  

C A phylogenetic analysis of the DGAT amino acid sequences of *B. napus* and its ancestral parents DGAT sequences is based on the maximum likelihood method algorithm in mega X with 1000 bootstraps. Pink dot indicates that a DGAT belonging to the DGAT2 category presented in A while blue dot indicates that a DGAT2 belongs to the category represented in B.
The DGAT2 enzymes of *B. napus* provide an excellent opportunity to study motifs/amino acid segments that influence the acyl-CoA specificity of DGAT2 enzymes as they differ in their acyl-CoA specificity despite their similar amino acid sequences. **Paper II** describes the approach we implemented to elucidate the potential presence of a confined region of the amino acid sequence of DGAT2 that affect the acyl donor specificity. The iterative approach includes the design of chimeric enzymes and the screening process of the chimeric enzymes for an altered acyl donor specificity. The characterisation of the first chimeric enzymes, where large segments of amino acids from a 22:1Δ13-CoA accepting enzyme substitutes the corresponding sequence of an enzyme predominantly accepting 18:3Δ9,12,15-CoA enables a rough estimation of where regions of interest for substrate specificity are situated. The screening, conducted on microsomal preparations of recombinant yeast include separate assays with [14C]-labelled di-6:0-DAG and either 18:3Δ9,12,15-CoA or 22:1Δ13-CoA. A shift in the proportion of *de novo* formed TAG during assays between 18:3Δ9,12,15-CoA and 22:1Δ13-CoA compared to the parental 18:3Δ9,12,15-CoA preferring enzyme, indicate a change in acyl-donor specificity, see figure 12A and B. New generations of chimeric enzymes with a design based on the results from the screening process of the previous generation of enzymes allowed us to decrease the size of the amino acid substitution gradually. We managed to identify a region that significantly affects the acyl donor specificity while implementing this strategy, see figures 12A-E and 13.

The identified region, situated in the first third of the DGAT2 enzymes contains two predicted transmembrane helices. Predictions indicate that these two transmembrane regions reoccur in a wide range of plant-derived DGAT2 enzymes. We designed a new chimeric enzyme to elucidate whether the importance of this region confines to *B. napus* DGAT2 enzymes or whether it may be applicable in other species as well. The new chimeric enzyme contains a backbone of Arabidopsis DGAT2 substituted in the identified region by a 22:1Δ13-CoA accepting *B. napus* DGAT2. The acyl-CoA specificity of the chimeric *AtBnaDGAT2* differs significantly from that of the wildtype Arabidopsis DGAT2, and it exhibits a drastic reduction in the preference for 18:3Δ9,12,15-CoA, down to a level comparable to 22:1Δ13-CoA, see figure 12D and E. This change in specificity indicates that the identified region may not only be of importance in *B. napus* but also for other plant-derived DGAT2 enzymes.
Figure 12 Specificity towards 18:3Δ⁹,¹²,¹⁵-CoA and 22:1Δ¹³-CoA with supplementing [¹⁴C]-di-6:0-DAG of parental wildtype enzymes from *B. napus* (A and B) and Arabidopsis (D) and chimeric enzymes (C and E) A *B. napus* parental wildtype DAGT2 form that prefers 18:3Δ⁹,¹²,¹⁵-CoA predominantly. B *B. napus* parental wildtype form preferring both 18:3Δ⁹,¹²,¹⁵-CoA and 22:1Δ¹³-CoA at a similar rate. C chimeric *B. napus* enzyme where B substitute the identified delimited region in the A enzyme. D Arabidopsis wildtype parental enzyme. E A chimaera where a region of the *B. napus* enzyme B replace the region of interest in the Arabidopsis DGAT2 amino acid sequence. n=3, error bars denote the standard deviation.
Figure 13 depicts the amino acid sequence of BnaA.DGAT2.d, predicted transmembrane helices, previously described motifs and the region affecting the acyl donor specificity, identified in paper II (semitransparent blue ellipse). A putative neutral lipid-binding motif (yellow) is present, however in reversed order compared to Stone et al. (2006). YFP and EPHS (green) are both essential for DGAT catalytic activity (Liu et al., 2011). Highlighted pink residues represent a putative acyl acceptor pocket (Lu et al., 2020; Turnbull et al., 2001).
The third paper, paper III, focuses on crambe and a potential bottleneck posed by DGAT substrate specificities. A previous investigation elucidating potential bottlenecks of erucic acid accumulation in transgenic high erucic crambe reveal that developing seeds possess a much higher accumulation of di-22:1Δ^{13}-DAG compared to the increase of erucic acid in TAG. A hypothesis arose that the intrinsic acyltransferases, primarily DGAT, that catalyse the final step in the synthesis of TAG are unable to utilise di-22:1Δ^{13}-DAG as a substrate efficiently. We successfully identified and cloned eight alleles encoding DGAT in crambe, four of each DGAT1 and DGAT2. Acyl-CoA substrate specificity determined when using di-6:0-DAG revealed that 22:1Δ^{13}-CoA was one of the most favoured of all the acyl-CoA tested by all enzymes but one. The latter enzyme, a DGAT2, deviate and strongly prefers 18:3Δ^{9,12,15}-CoA over 22:1Δ^{13}-CoA. The deviating DGAT2 enzyme differs not only in the specificity of acyl donors but also in acyl acceptor preference.

We conducted a more detailed study of substrate specificities of one DGAT1 and one DGAT2, which both prefers 22:1-CoA as a substrate in combination with di-6:0-DAG. These more detailed studies include the use of DAG species likely to occur in developing seeds of crambe unlike the artificial di-6:0-DAG. The most frequently occurring DAG species utilised for TAG synthesis in wildtype crambe is sn-1-22:1Δ^{13}-sn-2-18:1Δ^{9}-DAG. Both of the tested DGAT

![Figure 14](image.png)

Figure 14 substrate specificities of crambe DGAT enzymes towards 22:1Δ^{13}-CoA in combination with the frequently occurring sn-1-22:1Δ^{13}-sn-2-18:1Δ^{9}-DAG or di-22:1Δ^{13}-DAG. A CaDGAT1 D B CaDGAT2 I, n=3 error bars denote standard deviation, significance determined by one-way ANOVA with posthoc Tukey's HSD test.
enzymes readily accept sn-1-22:1Δ^{13}-sn-2-18:1Δ^{9}-DAG as acyl acceptor and can acylate the sn-3 position of this DAG with 22:1Δ^{13}-CoA, see figure 14. The situation is quite different with di-22:1Δ^{13}-DAG as acyl acceptor. The DGAT2 is unable to utilise the di-erucic acyl-acceptor while the DGAT1 is capable of catalysing the formation of TAG from di-22:1Δ^{13}-DAG but at a much-reduced rate compared to sn-1-22:1Δ^{13}-sn-2-18:1Δ^{9}-DAG, see figure 14. We also conducted similar long-chain DAG assays with the deviating DGAT2 enzyme as well. This enzyme was unable to use both sn-1-22:1Δ^{13}-sn-2-18:1Δ^{9}-DAG and di-22:1Δ^{13}-DAG. Thus, a substantial part of the endogenous TAG catalysing enzymes in crambe are unable to or at least very poor at, producing trierucin despite the presence of di-22:1-DAG in the transgenic crambe (Guan et al., 2014).

The poor capability of the crambe DGAT enzymes to produce trierucin encouraged us to identify a suitable DGAT gene candidate to aid the enrichment of erucic acid in transgenic crambe. The identification of DGAT2 from *T. majus*, with its record-high levels of erucic acid in the seed oil, has remained elusive despite repeated attempts to identify it. We re-evaluated publically available transcriptomic data in paper IV and found a partial nucleotide sequence potentially belonging to a DGAT2 enzyme. We retrieved a full-length sequence through a 5′/3′ RACE conducted on cDNA derived from developing seeds of *T. majus* with primers based on the partial sequence from the transcriptome. The identified DGAT2 enzyme exhibit an extreme specificity towards

![Substrate Specificity Graph](image-url)

*Figure 15* substrate specificities of *T. majus* DGAT enzymes towards 22:1Δ^{13}-CoA in combination with most preferred DAG species or di-22:1Δ^{13}-DAG. **A** TmDGAT1 **B** TmDGAT2, n=3 error bars denote standard deviation
18:3\textsuperscript{\text{9,12,15}}-CoA when supplemented with di-6:0-DAG, which is more than sixfold larger than the second most favoured acyl donor. We assessed the acyl-CoA specificity of both TmDGAT1 and the newly identified TmDGAT2 with di-22:1\Delta^{13}-DAG and again, as in crambe, TmDGAT2 is unable to utilise di-22:1\Delta^{13}-DAG, while the TmDGAT1 enzyme can, see figures 14 and 15.

Transgenic C. sativa is, as demonstrated in the production of EPA and DHA, very suitable for the production of novel oil qualities. As seen in figure 9, synthesis of, e.g. EPA and DHA in an oil crop requires a wide range of introduced genes, where sequential alternating desaturations and elongations modify 18:2\Delta^{9,12} and 18:3\Delta^{9,12,15} into the final product (Connelly & MacIntosh, 2018; Sottosanto et al., 2018). The characterisation of enzymes involved in lipid metabolism in C. sativa is still very valuable, despite the successful production of EPA and DHA as the incorporation of other FAs may be less compatible with the endogenous lipid metabolism. We, therefore, characterised several enzymes from C. sativa involved in lipid metabolism in paper V. Assessment of the acyl-donor specificity of CsDGAT1 and CsDGAT2 while providing di-6:0-DAG indicate that the CsDGAT1 exhibit very broad and indiscriminate specificity and readily accepts all assessed acyl-CoA substrates. CsDGAT2, on the other hand, express a specificity profile that resembles that of the B. napus BnaA.DGAT2.b, where 18:3\Delta^{9,12,15} is highly preferred compared to the other tested acyl-CoAs, see figures 16A and 11A. Phylogenetic analyses also indicate a closer kinship of CsDGAT2 to the BnaA.DGAT2 enzymes with similar specificity profiles rather than the two BnaA.DGAT2 enzymes that readily accept 22:1\Delta^{13}-CoA as substrate, see figure 16B. Characterisation of the C. sativa PDCT enzyme reveals that acyl donor specificity towards PC species is dependant of the acyl acceptor DAG species and changes significantly, just as with the DGAT enzymes in crambe.

Further, we investigated the modus operandi of the CsPDCT enzyme, catalysing the interconversion between DAG and PC. CsPDCT is unable to utilise DAG-species containing ricinoleoyl. Not only is CsPDCT unable to utilise these DAG species, but they also seem to inhibit the PDCT catalysed interconversion of other DAG species devoid of ricinoleoyl into PC as well. The inhibited interconversion of CsPDCT caused by ricinoleoyl containing DAG species is an excellent example of a potential incompatibility issue within the endogenous lipid metabolism if scientists intend to produce a seed oil rich in ricinoleic acid by gene technology. Additional measurements with increased complexity, e.g. in this case recruitment of a PDCT, not inhibited in this way, may circumvent such an incompatibility issue.
Figure 16 A the acyl-CoA specificity of CsDGAT2, which is similar to the acyl donor specificity of BnaA.DGAT2b (figure 11A). We assessed CsDGAT2 with di-6:0-DAG as acyl donor under close to linear conditions. Error bars denote standard deviation, n = 3. B phylogenetic analysis of CsDGAT2 and B. napus DGAT2 amino acid sequences based on the maximum likelihood method algorithm in mega X with 1000 bootstraps. A pink dot indicates a DGAT which predominantly prefers 18:3Δ9,12,15-CoA and a blue dot indicate an enzyme which expresses a similar preference for both 18:3Δ9,12,15–CoA and 22:1Δ13-CoA.
5 Future perspectives and discussion

Several important questions remain unresolved despite the efforts put into this thesis, as customary. The interpretation of results from an experiment is a sometimes daunting and perplexing task for an inquisitive doctoral student, yet it is also a very inspiring and creative process. On more than one occasion, the acquired data added more questions to an ever-growing pile than it provided answers. Discussions about results and interpretation of these in the lab, with fellow researchers at other universities and scientific conferences, have always been enriching and often provided insights, previously overlooked.

We had not yet developed the methodology for assays supplemented with long-chain-DAG species during the experimental phase of paper I. The DGAT specificity is, therefore, merely characterised with the artificial and water-soluble acyl acceptor di-6:0-DAG. We can currently only speculate if the stark difference in acyl donor specificity will remain or not when assessed with frequently occurring natural DAG species. It would thus be beneficial to re-examine the established, starkly contrasting specificities exhibited by B. napus DGAT2 enzymes, with long-chain-DAG abundant in the DAG pool of developing B. napus seeds and the desired di-22:1Δ13-DAG. These DGAT2 enzymes might not only prefer different acyl donors but may very well acylate contrasting DAG species as well, as we observed in crambe (paper III).

Similarly, it would be interesting to re-evaluate the chimaeras produced during paper II with additional DAG-species. The acyl donor specificity is without a doubt affected by substitutions in the identified region, but what about the acyl acceptor specificity? Are the two specificities regulated by two separate regions, or by partially overlapping regions, or even within the same region? The substrate specificities, acyl acceptor and acyl donor, may also affect each other even if these reside in two separate regions as well. A stretch of amino acids
separate the putative acyl acceptor pocked motif from the predicted transmembrane helices region, and this indicates that a separate region governs the acyl acceptor specificity, see figure 13. It would also be valuable to proceed with new chimaeric enzymes to dissect the region further in the search for a more precise motif governing 22:1Δ^{13}-CoA acceptance. Examinations of additional known DGAT amino acid sequences with determined acyl donor specificities can aid in search of similar motifs promoting acylation of other acyl-CoAs as well. Such motifs may aid future in silico screening of candidate DGAT2 enzymes with a suitable specificity for tailormade purposes. In addition, such an identification, in turn, dependant on the size and nature of the motifs, could enable genetic engineering with, for example, CRISPR/Cas9 to change endogenous DGAT2 encoding genes altering the substrate specificity of the enzyme.

The previous biochemical characterisation of DGAT2 enzymes from several plants producing UFAs, e.g. castor bean, the tung tree and ironweed indicate that these are highly capable of incorporating their respective abundant UFA into TAG. DGAT2 enzymes are thus often mentioned as enzymes specialised in forwarding UFAs into TAG, at least in plants producing UFAs. The DGAT2 enzymes characterised in this thesis exhibit a more distinct specificity towards one or a few specific acyl-CoAs as opposed to the DGAT1 enzymes that instead accept a broader range of the assessed acyl donors at similar rates. Some of the DGAT2 enzymes investigated, prefers common PUFAs rather than abundant erucic acid, however. The extraordinary affinity towards 18:3Δ^{9,12,15}-CoA coupled with the poor acceptance of 22:1Δ^{13}-CoA by the identified *T. majus* DGAT2 enzyme is therefore very interesting. What role does a DGAT2 enzyme have which strongly prefers 18:3Δ^{9,12,15} in *T. majus* which is virtually devoid of PUFAs in its seed oil?

The seed oil of *T. majus* contains the highest recorded levels of erucic acid within vascular plants, and the predominant TAG species is trierucin. A prerequisite for the synthesis of trierucin through yet known mechanisms is the acylation of di-22:1Δ^{13}-DAG with 22:1Δ^{12}-CoA. The DGAT2 enzyme of *T. majus* is entirely unable to utilise di-22:1Δ^{13}-DAG as acyl acceptor when assessed with the acyl donors included in paper IV. The same is true for the two investigated crambe DGAT2 in paper III. To our surprise, the results thus efficiently rule out DGAT2 as the enzyme catalysing the synthesis of trierucin and more or less any erucic acid transfer to TAG at all in *T. majus*.
Our data indicate that either TmPDAT, TmDGAT1, or a combination of both, are responsible for the synthesis of the abundant trierucin in *T. majus*. TmDGAT1 as concluded here, capable of producing trierucin, albeit at reduced rates compared to the best-preferred substrates, so is the investigated crambe DGAT1 enzyme. The DGAT1 of *T. majus*, when expressed in seeds of a transgenic HEAR does increase the oil yield but not the erucic acid content and thus earlier disregarded as the producer of trierucin in *T. majus*. This may, however, be a misinterpretation as HEAR most probably already contain DGATs capable of efficiently acylating the sn-3 position of abundant DAG species with erucoyl-CoA. An increase of erucic acid in HEAR by expressing TmDGAT1 would have to provide at least two critical substrates, di-22:1Δ13-DAG and abundant 22:1Δ13-CoA. The DAG pool in common HEAR varieties are devoid of di-22:1Δ13-DAG due to the inherent inability of the BnLPAT to produce di-22:1Δ13-PA. Transgenic HEAR without a suitable LPAT is thus an improper system to detect the catalytic capability of producing trierucin.

The definition of UFAs is an artificial categorisation and generalisations is to be treated cautiously. UFAs are structurally diverse molecules, and their incorporation into TAG may be just as varied. DGAT2 enzymes may thus not always be the main actor in UFAs incorporation into TAG, despite the generalised assumptions based on a few observations made where DGAT2 catalyse the incorporations of UFAs into TAG.

Which mechanisms can enable such high levels of erucic acid in the seed oil of *T. majus*, as we deem the DGAT enzymes as unsuitable? TmPDAT is to the best of our knowledge still not yet characterised. Although the PC-pool in crambe, providing the acyl donor for PDAT, have deficient levels of erucic acid, it may be different in *T. majus*. How *T. majus* manage to achieve the record-high levels of erucic acid is, therefore, unfortunately, still unresolved. It is, therefore, encouraging to continue to characterise additional enzymes involved in lipid metabolism such as PDAT in order to fully understand how *T. majus* achieves this high level of erucic acid in its seed oil. Once clarified how *T. majus* is channelling erucic acid into TAG, it may be possible to implement this strategy in other oil crops such as crambe, in the pursuit of very high levels of erucic acid. The mechanisms allowing for the high levels of erucic acid in *T. majus* seed oil may, however, not be applicable in conventional oil crops. The mechanisms allowing the high levels of erucic acid might be associated with the rather low oil content in the seeds of *T. majus*, which is approximately 10%.
The complementary specificities of the CsDGAT1 and CsDGAT2 enzymes determined in paper V correlates well with the seed oil composition in the transgenic lines where CsDGAT2 is down-regulated. The *in vitro* characterisation of these acyltransferases explain thus to a great extent the observed *in vivo* phenotypic changes in the transgenic *C. sativa*. Most of our characterisations of acyltransferases investigate one acyl donor and acyl acceptor combination at the time. The acyl-CoA- and DAG-pools in developing seeds contain a range of different species and are not homogenous as in our assays. It is yet to be determined if the acyltransferases selectivity and specificity are coherent, or whether the abundance of various substrate species affects the selectivity and specificity through competition. The characterisation of CsPDCT indicates that the presence of ricinoleoyl in DAG species affect the interconversion of other DAG species into PC. Are there similar, yet unaccounted for, effects that affect the selectivity or specificity for acyltransferases *in vivo* as well? It would, therefore, be valuable to conduct complex assays with a mixture of various DAG-species and acyl-CoA species. How to separate all the produced lipids, determine their configuration and quantify the minute quantities of them is, however a whole other story.

Living organisms are fantastic and more complex than we can comprehend. Here we investigate a small, well-defined area of interest, yet the complexity within this narrow topic is mindboggling, and there is a magnificent beauty to this complexity. Still, after 70 years of researchers’ effort to understand a phenomenon, there is at least just as much waiting to be discovered.


References


optimising omega-3 long chain polyunsaturated fatty acid accumulation in transgenic Camelina sativa. Scientific reports, 7(1), pp. 1-12.


Plants are incredible organisms that can produce seed oils suitable for human consumption. Seed oils can also replace fossil oil-based products as an environmentally friendly alternative. Plant breeders and researchers continually improve not only the oil yield but also the oil quality. Industrial applications demand particular oil qualities and the more suitable oil a plant can produce, the more competitive the environmentally friendly alternative becomes. Plants produce oils with small building blocks, and these are later modified and put together to oil. Specific proteins or enzymes are responsible for the production of the seed oil. These enzymes often prefer specific building blocks, and there are many of these in a plant. Enzymes preferences for building blocks is called substrate specificity. Each plant has its variants of enzymes, and these enzymes often have slightly different specificities. To improve or change the oil quality in plants, one has to know how the oil-producing enzymes will react in order to be successful. We have therefore investigated the preference of several enzymes specifically for erucic acid, a constituent in some seed oils, used in many industrial applications. During our investigations, we have understood the importance of enzymes specificities and that it due to these may be challenging to enhance the content of erucic acid in plant oils. DGAT2 is an enzyme that is important in the final production of oil. Enzymes are large molecules made by subunits called amino acids connected in a long chain. These amino acids determine the behaviour of enzymes, and thus, their substrate specificity. Through changing some of the amino acids in DGAT2, we have managed to change enzymes that dislike erucic acid into enzymes that readily accept erucic acid. The detailed studies of enzymes involved in the production of oil will aid the pursuit of better-adapted oils.
Acknowledgements

This thesis is in no way a solo endeavour even though my name is the only one occurring on the front cover. I owe everybody in my surrounding my gratitude. My contribution to the scientific community summarised here would be impossible without them and all the efforts made by other researchers.

This work would not have existed at all without the kind support and encouragement from my group of supervisors, which consists of Ida Lager, Sten Styme, Anders S. Carlsson and Li-Hua Zhu. Ida, my main supervisor, thank you for believing in me when things were looking bleak and for supporting me through thick and thin. You went that extra mile for me and you will always be an inspiration to me. You have always encouraged me to develop my ideas, no matter how ridiculous they were. You have sincerely affected me, and have made me become a better person. My co-supervisor Sten is also a great source of inspiration, your genuine interest, kindness, knowledge, curiosity, critical thinking and your strong belief in cooperation are just a few of the characters, for which I admire you.

I am very grateful to have become a part of the close collaboration between our group and the professor Antoni Banaś and Kamil Demski, initiated by Sten and Antoni decades ago, and many of the findings in this thesis is a direct result of this collaboration. It has been an honour to work and spend time with Kamil, who is a brilliant aspiring young researcher and a fantastic friend, whom I've grown very fond of. Antoni has always been encouraging and very keen on welcoming us as guests and made us feel like a part of his family.

I would also like to thank my office mates Joel, Magnus and Per, as you endured my exaggerated exaltation every single time I believed that I had solved the mysteries of lipid metabolism and for rescuing me during my downfall, once I realised I was wrong. Per Snell is one of those, unique personalities, that I
nowadays recognise as one of my closest friends. Despite his fierce appearance, he is one of the kindest, supportive, generous persons I have ever met, who always do his uttermost for the collective well-being.

Thank you, Roland, for becoming my senpai long ago during our adventures somewhere in the arctic archipelago. With his excessively entertaining stories, he inadvertently (or deliberately, I'm quite unsure) convinced me to pursue for an academic profession, and may very well be the reason for me becoming a doctoral student.

Rui, Meta and Per Hofvander thank you for being fantastic people that share my passion for old Studio Ghibli classics. I would also like to show my gratitude towards Helena, Helle, Helén and Mirela, four energetic individuals who spread joy and warmth in the lab and make me look forward to show up at work every time.

I would also like to show my greatest gratitude to the plant lipid research community. I have a hard time to believe that there is a more open-minded warm-hearted bunch of people within the academic world. I would especially like to thank Ed Cahoon, who has a beautiful mindset and Kent Chapman with his kind heart and generous personality. A wide range of beverages has been devoured, on numerous occasions and different continents, during endless discussions with Phil Bates whether the specificity of acyltransferases can provide a mechanism for the different acyl pools he so passionately advocates, or whether they exist at all. Further, his brothers in arms, Jay Shockey and Jay Thelén, both prominent researchers, treated me as an equal instantaneously and provided us with more pitchers than I would like to recall. That night, was also the night when our head of department persistently complained, that the Germans are unable to construct straight roads, while he was singing cheerfully. Thank you, Till Ischebeck, for encouraging your doctoral students not to fear authorities and include them in your companionship.

Finally, I would like to thank my closest family Annika, Abbe and Edith for all the encouragement, support and love during this whole process. Thank you, my closest relatives, Rolf, Winnie and Marie for being there when I needed you the most and for all the support, help and warmth.
Acyltransferases catalyse the acylation of glycerol into triacylglycerol, the main constituent in seed oil. The substrate specificities of acyltransferases, responsible for triacylglycerol synthesis, affect the fatty acid composition, which determines the properties of the seed oil. Here, we have identified a wide range of acyltransferases from several plant species and assessed their substrate specificities. Further, we have identified an amino acid region affecting the acyl donor specificity of an essential acyltransferase, acyl-CoA:diacylglycerol acyltransferase 2.

Simon Jeppson received his graduate education at the Department of Plant Breeding, SLU, Alnarp and obtained his MSc degree in Biology at SLU, Alnarp.