Engineering the stereoisomeric structure of seed oil to mimic human milk fat

Harrie van Erp, Fiona M. Bryant, Jose Martin-Moreno, Louise V. Michaelson, Govindprasad Bhutada, and Peter J. Eastmond

Human milk fat substitute (HMFS) is a class of structured lipid that is widely used as an ingredient in infant formulas. Like human milk fat, HMFS is characterized by enrichment of palmitoyl (C16:0) groups specifically at the middle (sn-2 or β) position on the glycerol backbone, and there is evidence that triacylglycerol (TAG) with this unusual stereoisomeric structure provides nutritional benefits. HMFS is currently made by in vitro enzyme-based catalysis because there is no appropriate biological alternative to human milk fat. Most of the fat currently used in infant formulas is obtained from plants, which exclude C16:0 from the middle position. In this study, we have modified the metabolic pathway for TAG biosynthesis in the model oilseed Arabidopsis thaliana to increase the bilized recombinant sn-1/3-structured triacylglycerols and facilitate the formation of C16:0 soaps. The aim of this study was to explore whether the stereoisomeric structure of vegetable fat can be altered by iterative metabolic engineering to oiled seeds to provide a source of HMFS for infant formula.

Significance

In human milk fat, saturated fatty acids are esterified to the middle position on the glycerol backbone giving the triacylglycerol molecules an unusual stereochemistry that assists nutrient absorption in the infant gut. However, the fat used in most infant formulas is derived from plants, which esterify saturated fatty acids to the outer positions. Here, we have engineered the metabolism of an oiled seed plant so that it accumulates triacylglycerol with more than 70% of the saturated fatty acid palmitate in the middle position, thereby mimicking human milk fat stereoisomeric structure. Applying this technology to oilseeds (or oleaginous microorganisms) might provide a source of human milk fat substitute for infant nutrition.

Author contributions: P.J.E. designed research; H.v.E., F.M.B., J.M.-M., L.V.M., G.B., and P.J.E. performed research; H.v.E., L.V.M., and P.J.E. analyzed data; and H.v.E. and P.J.E. wrote the paper. The authors declare no competing interest.

This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

To whom correspondence may be addressed. Email: peter.eastmond@rothamsted.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1907915116/-/DCSupplemental.

First published September 30, 2019.
metabolic engineering, so that it mimics HMF. To our knowledge, no land plant (Embryophyta) produces TAG enriched in C16:0 at the sn-2 (versus sn-1/3 positions) and C16:0 is largely excluded from this position in virtually all cases (4, 5, 8). Even in palm oil that contains ~48% C16:0 in total, only 9% of this occupies the sn-2 position (5). Here, we describe a method for modifying TAG biosynthesis, in the model oilseed Arabidopsis thaliana, that results in a stereoisomeric redistribution of acyl groups such that the amount of C16:0 at the sn-2 position increases more than 20-fold to over 70% of the total; a level of enrichment that is comparable to HMF. Applying this technology to oilseed crops might provide a source of HMFS for infant formula.

Results and Discussion

LPAT1 Can Be Redirected to the ER by Removing Its Chloroplast Targeting Signal.

In plant cells, triacylglycerol (TAG) is formed by a cytosolic glycerolipid biosynthetic pathway situated on the endoplasmic reticulum (ER) and the enzyme responsible for acylation of the sn-2 position is lysophosphatidic acid acyltransferase (LPAT) (9) (Fig. 1). ER-resident isoforms of LPAT commonly discriminate against C16:0-CoenzymeA (CoA) as a substrate, and this may be why C16:0 is excluded from the sn-2 position (9, 10). To overcome this limitation, we decided to express an LPAT with specificity for C16:0-CoA (Fig. 1). Several candidate transgenes have been described from cyanobacteria (11), mammals (12), and algae (13, 14). However, plants already possess an LPAT with the appropriate selectivity that resides in the chloroplast (15, 16) (Fig. 1). This LPAT uses a C16:0-acyl carrier protein (ACP) substrate but will also accept C16:0-CoA in vitro (17, 18). We therefore decided to test whether chloroplast LPAT could be retargeted to the ER (Fig. 1). Chloroplast LPAT is an integral membrane protein that is nuclear encoded and contains an N-terminal chloroplast targeting signal (CTS) (19). CTS deletion has previously been used to alter protein localization (20). Using transient expression in Nicotiana benthamiana epidermal cell transiently expressing RFP-ΔCTS-LPAT1 and m-GFP5-ER marker, (Scale bar, 20 μm.) (B) Effect of seed-specific ΔCTS-LPAT1 expression in Arabidopsis on the percentage of C16:0 esterified to the sn-2 position of TAG, verses sn-1+3. WT, wild type; L30, L6, and L11, 3 independent homozygous ProGLY:ΔCTS-LPAT1 lines. Values are the mean ± SE of measurements made on separate seed batches from 3 plants of each genotype (n = 3). a, b and c denote values significantly (P < 0.05) different from WT (ANOVA + Tukey HSD test).
To determine whether ΔCTS-LPAT1 functions in plants and can enable C16:0 to be incorporated into the sn-2 position of TAG, we expressed this truncated protein under the control of the seed-specific soybean glycinin-1 promoter (ProGLY) in the model oilseed Arabidopsis (21). We selected more than 40 primary transformants (T1) using a DsRed fluorescent marker system (21) and analyzed the total fatty acyl composition of T2 seed batches. We found that several lines exhibited an increase in total C16:0 content, which suggested that the transgene was promoting C16:0 incorporation into TAG (Supplemental Appendix, Table S1). We selected 3 independent single copy T2 lines (L30, L6, and L11) with high C16:0 content and obtained homozygous T3 seed. When we purified TAG from these homozygous seed batches and determined its stereochemistry using lipase digestion (22), we found that the percentage of C16:0 at the sn-2 position (vs. sn-1/3), had increased more than 16-fold, from only ~2% in wild type to values ranging between ~32 and ~39% in the 3 independent ProGLY:ΔCTS-LPAT1 lines (Fig. 2B and Supplemental Appendix, Table S2). ΔCTS-LPAT1 expression was therefore sufficient to allow incorporation of C16:0 into the sn-2 position of TAG, but not to achieve positive enrichment at this position versus the sn-1/3 positions, which can already incorporate a low proportion of C16:0 (9) (Fig. 1).

**Disruption of LPAT2 Enhances C16:0 Incorporation into the sn-2 Position of TAG.** Competition between heterologous and native acyltransferases is one factor that may limit the incorporation of specific fatty acyl groups into TAG (23). We therefore investigated whether ΔCTS-LPAT1-dependent incorporation of C16:0 into the sn-2 position of TAG could be enhanced by disrupting the function of the native ER-resident LPAT; believed to be LPAT2 in Arabidopsis (10) (Fig. 1). The Ipata-1 null mutant is embryo lethal (10). However, T-DNA insertions in noncoding regions of essential genes can be used to produce viable hypomorphic alleles (24, 25). We therefore isolated 2 T-DNA mutants (Ipata-2 and Ipata-3) with insertions 302 and 139 bp 5′ of the LPAT2 translational start site (Fig. 3A). We then crossed ProGLY:ΔCTS-LPAT1 L11 into each of the new Ipata alleles and recovered homozygous seed batches. When we purified TAG from these seed batches and performed positional analysis, we found that the percentage of C16:0 at the sn-2 position had increased from ~33% in the parental ProGLY:ΔCTS-LPAT1 line to ~51% in the Ipata-3 background, whereas the effect in the Ipata-2 background was not significant (P > 0.05) (Fig. 3B and Supplemental Appendix, Table S3). qRT-PCR analysis showed that LPAT2 expression is reduced by ~83% in developing Ipata-3 siliques, but only by ~24% in Ipata-2. (Fig. 3B). These data support the hypothesis that LPAT2 contributes to TAG biosynthesis in Arabidopsis seeds (10) and that it competes with ΔCTS-LPAT1. The level of C16:0 enrichment at sn-2 also appears to respond to the strength of LPAT2 repression and achieving a greater reduction than ~83% might therefore lead to even stronger enrichment.

**Disruption of PDCT Also Enhances C16:0 Incorporation into the sn-2 Position of TAG.** In developing Arabidopsis seeds >90% of the glycerol backbone in TAG is derived from the membrane lipid phosphatidylcholine (PC), owing to rapid diacylglycerol (DAG)-PC interconversion (26), catalyzed mainly by the plant-specific head group exchange enzyme PC:DiAG cholinephosphotransferase (PDCT) (27, 28) (Fig. 1). Although LPAT is responsible for the initial acylation of glycerolipids at sn-2, once these acyl groups are in PC they may be removed and replaced by acyl editing activities (26, 29, 30) (Fig. 1). To determine whether bypassing glycerolipid flux through PC (Fig. 1) might increase ΔCTS-LPAT1-dependent incorporation of C16:0 into the sn-2 position of TAG, we crossed ProGLY:ΔCTS-LPAT1 L11 into the pdct (reduced olate desaturation1) mutant (27). When we purified TAG from ProGLY:ΔCTS-LPAT1 pdct seed batches and performed positional analysis, we found that the percentage of C16:0 at sn-2 had increased from ~30% in the parental ProGLY:ΔCTS-LPAT1 line to ~56% in the pdct background (Fig. 4A and Supplemental Appendix, Table S4). These data suggest that a more direct flux of newly made DAG into TAG (28) (Fig. 1) favors C16:0 incorporation and/or retention at the sn-2 position. In WT seeds, it is...
conceivable that C16:0 entering the sn-2 position of PC might either be edited from it by the action of lysophosphatidylcholine acyltransferase (LPCAT) or a phospholipase A2 (28). Interestingly, Lager et al. (29), have provided in vitro evidence that the reverse activities of Arabidopsis LPCAT1 and LPCAT2 can selectively remove certain fatty acyl groups from PC, but C16:0 was not tested. Although rapid DAG-PC interconversion occurs in Arabidopsis seeds (26), it is noteworthy that interspecific variation has been reported in this flux (31) and so the effect of PDCT disruption on C16:0 enrichment at the sn-2 of TAG may differ between oilseeds.

**Disruption of LPAT2 and PDCT Has an Additive Effect on Incorporation of C16:0 at sn-2.** To determine whether the combination of reducing LPAT competition and bypassing flux through PC would have an additive effect on ΔCTS-LPAT1-dependent incorporation of C16:0 into the sn-2 position of TAG (Fig. 1), we crossed ProGLY:ΔCTS-LPAT1 lpat2-3 with ProGLY:ΔCTS-LPAT1 pdct. When we purified TAG from homozygous seed batches and performed positional analysis, we found that the percentage of C16:0 at sn-2 had increased from ~56% in ProGLY:ΔCTS-LPAT1 pdct to ~71% in ProGLY:ΔCTS-LPAT1 lpat2-3 pdct (Fig. 4A and SI Appendix, Table S4). The combination of just 3 modifications to the TAG biosynthetic pathway in Arabidopsis (i.e., ΔCTS-LPAT1 expression, plus LPAT2 and PDCT suppression) is therefore sufficient to replicate the level of C16:0 enrichment at the sn-2 position (vs. sn-1+3) that is found in HMF (1–3). Analysis of TAG composition in ProGLY:ΔCTS-LPAT1 lpat2-3 pdct (All) seeds using high resolution/accurate mass (HR/AM) lipidomics (32) also confirmed the presence of C16:0 groups at the sn-2 position, since tripalmitin was 27-fold more abundant than in WT (SI Appendix, Fig. S2A). By contrast, no dipalmitoyl PC was detected in ProGLY:ΔCTS-LPAT1 lpat2-3 pdct seeds and molecular species of PC containing one C16:0 group were not increased (SI Appendix, Fig. S2B). These data suggest that an asymmetrical distribution of saturated and unsaturated fatty acyl groups in PC is maintained in ProGLY:ΔCTS-LPAT1 lpat2-3 pdct seeds and this may be important to prevent membranes assuming the gel phase at physiological temperatures (33, 34).

**Redistribution of C16:0 Reduces Seed Oil Content, But Not Germination or Establishment.** Many studies have shown that modifying fatty acyl composition can reduce TAG accumulation in oilseeds and in some cases can also impair seed germination and seedling establishment (35, 36). Our primary objective in this study was not to alter fatty acyl composition per se, but to change

---

**Fig. 4.** Bypassing flux through PC increases C16:0 incorporation into the sn-2 position of TAG. (A) Effect of pdct mutant background on percentage of C16:0 esterified to the sn-2 position of TAG in ProGLY:ΔCTS-LPAT1 and ProGLY:ΔCTS-LPAT1 lpat2-3 seeds. WT, wild type; L11, homozygous ProGLY:ΔCTS-LPAT1 line. (B) Seed weight and (C) percentage oil content of WT and ProGLY:ΔCTS-LPAT1 lpat2-3 pdct (All). Values are the mean ± SE of measurements on separate seed batches from between 3 and 6 plants in A and 5 plants in B and C of each genotype (n = 3–6). a and b denote values significantly (P < 0.05) different from L11 and pdct, respectively (ANOVA + Tukey HSD test) and c from WT (2-tailed Student’s t test).

---

**Fig. 5.** Effect of genetic modifications on seed vigor at 20 °C. Percentage (A) seed germination, (B) cotyledons expanded by day 4, and (C) true leaves developing by day 7. (D) Representative images of seedlings with expanded cotyledons and developing true leaves. (E) Seed/seedling TAG content at days 0 and 4. WT, wild type; All, ProGLY:ΔCTS-LPAT1 lpat2-3 pdct. Values are the mean ± SE of measurements made on separate seed batches from 3 plants of each genotype (n = 3). In D, Scale bar, 2 mm. a and b denote values significantly (P < 0.05) different from WT (2-tailed Student’s t tests).
the stereoisomeric structure of TAG. To examine the physiological impact of C16:0 enrichment at the sn-2 position of TAG, we compared seed batches from wild type and ProGLY:ACTS-LPAT1 lpat2-3 pdct plants that had been grown together under standard laboratory conditions. We found no significant difference (P > 0.05) in seed weight between the 2 genotypes (Fig. 4B). However, the fatty acid content of ProGLY:ACTS-LPAT1 lpat2-3 pdct seeds was ca. 0.05% lower than that of wild type, when expressed as a percentage of seed weight (Fig. 4C). These data suggest that the modifications leading to incorporation of C16:0 into the sn-2 position reduce TAG biosynthetic flux. This finding is consistent with previous studies in which seed TAG composition has been modified either using genetic engineering or mutant breeding methods (35, 36). In warm conditions (20°C), ProGLY:ACTS-LPAT1 lpat2-3 pdct seed germination, scored as radicle emergence (Fig. 5A) and seedling establishment, scored as cotyledon expansion (Fig. 5B) and true leaf development (Fig. 5C), did not appear to be significantly (P < 0.05) impaired, relative to wild type. TAG breakdown also was not impeded in ProGLY:ACTS-LPAT1 lpat2-3 pdct seeds following germination in warm conditions (Fig. 5D), and this contrasts with some studies where seeds have been modified to incorporate uncommon fatty acyl groups into TAG (12). In cool conditions (10°C), ProGLY:ACTS-LPAT1 lpat2-3 pdct seed germination and seedling establishment also appeared not to be significantly (P < 0.05) impaired, relative to wild type (SI Appendix, Fig. S3). Finally, although ProGLY:ACTS-LPAT1 lpat2-3 pdct carries a hypomorphic allele of the essential gene LPAT2 (10) (Fig. 3), this does not appear to adversely affect growth and morphology at the rosette stage (SI Appendix, Fig. S4).

Conclusions

In this study we show that the TAG biosynthetic pathway in plants can be engineered so that the stereoisomeric structure of seed storage oil is altered to mimic that of HMF, with >70% of C16:0 concentrated at the middle (sn-2 or β) position on the glycerol backbone. There is mounting evidence that this configuration is beneficial for infant nutrition (1, 3, 6), but it has not been found to occur naturally in vegetable fats where C16:0 is virtually excluded from the sn-2 position (4, 5, 9). Many infant formulas contain HMFS that are made by restructuring vegetable fats using enzyme-based catalysis, but they are relatively costly to produce; particularly for the manufacture of true mimetics with >70% of C16:0 at the sn-2 position (1, 7). Translation of our technology from the model species Arabidopsis to an oilseed crop might conceivably provide a cheaper and more sustainable source of HMFS for infant formula, but further research would be required to test this supposition. If HMFS could be obtained directly from a vegetable source, this would abrogate the need for enzyme-based catalysis. The infant formula market is currently estimated to use nearly half a million metric tons of vegetable-derived fat per year. Several oilseed crops may be considered as possible hosts for HMFS production, and it is noteworthy that conventional sunflower (Helianthus annuus) and genetically modified oilseed rape (B. napus) varieties have already been developed that have the appropriate fatty acyl composition (37, 38). Even an oilseed crop with more modest C16:0 enrichment at the sn-2 position than we have achieved here may still be desirable since clinical trials have reported benefits with as little as 43% of C16:0 at the sn-2 position (1, 3, 6) and product surveys have found that this level of enrichment is common in infant formulas that are supplemented with HMFS (1).

Materials and Methods

Detailed descriptions of plant material and growth conditions, cloning, and Agrobacterium-mediated transformation, microscopy, mutant genotyping, lipid analysis, qRT-PCR analysis of gene expression, germination and seedling establishment assays, and statistical analysis are provided in SI Appendix, SI Materials and Methods. Primers used are listed in SI Appendix, Table S5.

ACKNOWLEDGMENTS.

We thank Prof. John Browse for pdct seeds and Prof. Edgar Cahoon for the pbnGl5Red3 vector. This work was funded by the UK Biotechnology and Biological Sciences Research Council through Grant BB/P012663/1.

1. W. Wei, Q. Jin, X. Wang, Human milk fat substitutes: Past achievements and current trends. Prog. Lipid Res. 74, 69–86 (2019).
2. W. C. Breckenridge, L. Marai, A. Kukis, Triglyceride structure of human milk fat. Can. J. Biochem. 42, 761–769 (1964).
3. S. M. Innis, Dietary triacylglycerol structure and its role in infant nutrition. Adv. Nutr. 2, 275–283 (2011).
4. H. Brockerhoff, M. Yurkowski, Stereospecific analyses of several vegetable fats. J. Lipid Res. 7, 62–64 (1966).
5. W. W. Christie, B. Novikova-Damyanova, P. Laakso, B. Herslof, Stereospecific analysis of triacyl-sn-glycerols via resolution of diastereomeric diacylglycerol derivatives by high-performance liquid chromatography on silica. J. Am. Oil Chem. Soc. 68, 605–701 (1991).
6. L. Bégin et al., Growth, stool consistency and bone mineral content in healthy term infants fed sn-2-palmitate-enriched starter infant formula: A randomized, double-blind, multicentre clinical trial. Proc. Natl. Acad. Sci. U.S.A. 102, 913–922 (1999).
7. H. U. Kim, A. H. Huang, Plastid lysophosphatidyl ethanolamine acyltransferase is essential for embryo development in Arabidopsis. Plant Physiol. 134, 1206–1216 (2004).
8. B. Yu, S. Wakao, J. Fan, C. Benning, Loss of plastidic lysophosphatidyl ethanolamine acyltransferase causes embryo lethality in Arabidopsis. Plant Cell Physiol. 45, 503–510 (2004).
9. I. Heilmann, M. S. Pidkowicz, T. Girke, J. Shanklin, Switching desaturase enzyme specificity by alternate subcellular targeting. Proc. Natl. Acad. Sci. U.S.A. 101, 10266–10271 (2004).
10. C. Zhang et al., Genetic and biochemical basis for alternative routes of tocotrienol biosynthesis for enhanced vitamin E antioxidant production. Plant J. 73, 628–639 (2013).
11. H. van Erp, P. D. Bates, J. Burgal, J. Shockey, J. Browse, Castor phospholipid:diacylglycerol acyltransferase facilitates efficient metabolism of hydroxy fatty acids in transgenic Arabidopsis. Plant Physiol. 155, 683–693 (2011).
12. H. van Erp, J. Shockey, M. Zhang, N. D. Adhikari, J. Browse, Reducing isozyme competition increases target fatty acid accumulation in seed triacylglycerols of transgenic Arabidopsis. Plant Physiol. 168, 36–46 (2015).
13. B. Lu, C. Xu, K. Aawai, D. A. Jones, C. Benning. A small ATPase protein of Arabidopsis, TG03, involved in chloroplast lipid import. J. Biol. Chem. 282, 35945–35953 (2007).
14. B. Ulker et al., Getting the most out of publicly available T-DNA insertion lines. Plant J. 56, 665–677 (2008).
15. P. D. Bates, J. Browse, The pathway of triacylglycerol synthesis through phosphatidylglycerol in Arabidopsis produces a bottleneck for the accumulation of unusual fatty acids in transgenic seeds. Plant J. 68, 387–399 (2011).
16. C. Lu, Z. Xin, Z. Ren, M. Miquel, J. Browse, An enzyme regulating triacylglycerol biosynthesis is encoded by the ROD1 gene of Arabidopsis. Proc. Natl. Acad. Sci. U.S.A. 106, 18837–18842 (2009).
17. P. D. Bates et al., Acyl editing and headgroup exchange are the major mechanisms that direct polyunsaturated fatty acid flux into triacylglycerols. Plant Physiol. 160, 1530–1539 (2012).
18. J. Lagger et al., Plant acyl-CoA:lysocephatidylcholine acyltransferases (LPCATs) have different specificities in their forward and reverse reactions. J. Biol. Chem. 288, 36902–36914 (2013).
30. A. Dahlqvist et al., Phospholipid:diacylglycerol acyltransferase: an enzyme that catalyzes the acyl-CoA-independent formation of triacylglycerol in yeast and plants. Proc. Natl. Acad. Sci. U.S.A. 97, 6487–6492 (2000).
31. D. K. Allen, P. D. Bates, H. Tjellström, Tracking the metabolic pulse of plant lipid production with isotopic labeling and flux analyses: Past, present and future. Prog. Lipid Res. 58, 97–120 (2015).
32. T. Yamada et al., Development of a lipid profiling system using reverse-phase liquid chromatography coupled to high-resolution mass spectrometry with rapid polarity switching and an automated lipid identification software. J. Chromatogr. A 1292, 211–218 (2013).
33. D. J. Hanahan, H. Brockerhoff, E. J. Barron, The site of attack of phospholipase (lecithinase) A on lecithin: a re-evaluation. Position of fatty acids on lecithins and triglycerides. J. Biol. Chem. 235, 1917–1923 (1960).
34. S. Jain et al., Identification of a novel lysophospholipid acyltransferase in Saccharomyces cerevisiae. J. Biol. Chem. 282, 30562–30569 (2007).
35. D. Lunn, G. A. Smith, J. G. Wallis, J. Browse, Development defects of hydroxy-fatty acid-accumulating seeds are reduced by castor acyltransferases. Plant Physiol. 177, 553–564 (2018).
36. S. Bai et al., Identification, characterization and field testing of Brassica napus mutants producing high-oleic oils. Plant J. 98, 33–41 (2019).
37. J. M. Fernández-Martínez, M. Mancha, J. Osorio, R. Garcés, Sunflower mutant containing high levels of palmitic acid in high oleic background. Euphytica 97, 113–116 (1997).
38. A. Jones, H. M. Davies, T. A. Voelker, Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell 7, 359–371 (1995).