Heliyon 6 (2020) e05237

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Characterization of the acyl-ACP thioesterases from *Koelreuteria paniculata* reveals a new type of FatB thioesterase

R. Martins-Noguerol, C. DeAndrés-Gil, R. Garcés, J.J. Salas, E. Martínez-Force, A.J. Moreno-Pérez *

Group of Genetics and Biochemistry of Seed Lipids, Department of Biochemistry and Molecular Biology of Plants Products, Instituto de la Grasa (CSIC), Ctra. de Utrera km 1, Building 46, 41013 Sevilla, Spain

ARTICLE INFO

Keywords:
Acyl-ACP thioesterases
FatA
FatB
Koelreuteria paniculata
Gondoic acid
Fatty acid synthesis
Agricultural engineering
Plant genetics
Plant physiology
Biochemistry
Molecular biology
Plant biology

ABSTRACT

Koelreuteria paniculata is a deciduous tree, popular in temperate regions for its ornamental value, which accumulates unusual cyanolipids in its seeds. The seed oil of this plant is rich in the unusual cis-11-eicosenoic fatty acid (20:1, or gondoic acid), a monounsaturated oil of interest to the oleochemical industry. In higher plants, de novo fatty acid biosynthesis takes place in the plastids, a process that is terminated by hydrolysis of the thioester bond between the acyl moiety and the ACP by acyl-ACP thioesterases. The specificity of acyl-ACP thioesterases is fundamental in controlling the fatty acid composition of seed oil. To determine the mechanisms involved in fatty acid biosynthesis in *K. paniculata* seeds, we isolated, cloned and sequenced two cDNAs encoding acyl-ACP thioesterases in this plant, *Kp*FatA and *Kp*FatB. Both of them were expressed heterologously in *Escherichia coli* and characterized with different acyl-ACP substrates. The *K. paniculata* FatB2 displayed unusual substrate specificity, so that unlike most FatB2 type enzymes, it displayed preference for oleoyl-ACP instead of palmitoyl-ACP. This specificity was consistent with the changes in *E. coli* and *N. benthamiana* fatty acid composition following heterologous expression of this enzyme. *Kp*FatB also showed certain genetic divergence relative to other FatB-type thioesterases and when modelled, its structure revealed differences at the active site. Together, these results suggest that this thioesterase could be a new class of FatB not described previously.

1. Introduction

Although oil crops produce triacylglycerols containing a limited number of fatty acids, there is a tremendous diversity among these compounds across the plant kingdom. Thus, in different plants it is possible to find a wide range of enzymatic activities that could be of interest to the chemical industry, particularly as the distinct speciality oil crops that produce them represent renewable sources of these chemicals. The activities of these fatty acids or lipids include hydroxyl, acetylene, cyclopropane or epoxide functions. In this regard, trees from the Sapindaceae family contain very unusual cyanolipids (Aichholz et al., 1997) and for example, the oil extracted from mature seeds of the tree *Koelreuteria paniculata* is made up of about 40% cyanolipids plus regular triacylglycerides. The cyanolipid fraction consists of a mixture of monoester and diester of fatty acids with a dihydroxynitrile (1-cyano-2-hydroxymethylpropylene-3-ol; Mikolajczak et al., 1970). The fatty acids esterified to both cyanolipids and triacylglycerols are mostly monounsaturated, and the content of 11-eicosenoic or gondoic acid ($20:1^{\Delta 11}$) is especially remarkable, accounting for 40–48% of the total fatty acids. Gondoic acid is an interesting fatty acid due to its properties for oleochemical applications, providing higher viscosity and lubricity than oleic acid but with a similar stability, making it an interesting base for biolubricants (Zorn et al., 2018).

The synthesis of gondoic acid takes place through extraplastidial elongation of oleoyl-CoA by specific fatty acid elongases (Leonard et al., 2004). The *de novo* synthesis of fatty acids takes place in the chloroplast of vegetative tissues and in the plastids of non-photosynthetic ones. This process produces most of the oleate and palmitate in the cell, and most of the precursors of the acyl moieties of acyl lipids that form part of the organelle membranes or that accumulate in storage organs. The pathway involves the formation of the acyl chain through successive cycles of two carbon condensation, reduction, dehydration and dehydrogenation of the acyl-acyl carrier protein (ACP) derivatives until a C18 saturated chain is reached. A very potent acyl-ACP desaturase activity in those organs mostly desaturates 18.0-ACP to 18:1-ACP. Intraplastidial reactions are terminated by specific acyl-ACP thioesterases that hydrolyse the

* Corresponding author.
E-mail address: ajmoreno@ig.csic.es (A.J. Moreno-Pérez).

https://doi.org/10.1016/j.heliyon.2020.e05237
Received 16 June 2020; Received in revised form 3 September 2020; Accepted 8 October 2020

acyl-ACP derivatives, supplying acyl molecules to the so-called eukaryotic path of glycerolipid synthesis (Ohlrogge and Jaworski, 1997). Acyl-ACP thioesterases are enzymes specific to plants and they can be classified into two well-defined families, FatA and FatB, with divergent amino acid sequences and substrate specificity (Jones et al., 1995). The FatA family are soluble enzymes located in the plastid stroma, and their high specificity for hydrolysing oleoyl-ACP makes them essential for plant viability (Salas and Ohlrogge, 2002; Sánchez-García et al., 2010; Moreno-Pérez et al., 2011). The FatB enzymes are bound to the plastidial membranes and they display greater diversity than the FatA thioesterases. These enzymes can be classified into two sub-families, the FatB1 enzymes that are present in all plants studied and that show preference for palmitoyl-ACP (Dörmann et al., 2000; Salas and Ohlrogge, 2002; Aznar-Moreno et al., 2016), and FatB2 enzymes that are only present in certain species and that hydrolyse short-medium chained acyl-ACP derivatives (Jones et al., 1995). The FatB1 enzymes are involved in providing tissues with appropriate levels of saturated fatty acids, which are necessary to ensure membranes achieve the correct fluidity, and with precursors of essential metabolites like sphingoid bases (Chen et al., 2008). As such, these enzymes are important in fatty acid metabolism and they have been used in plant biotechnology to alter oil composition. Expression of FatB1 thioesterases usually increases saturated 18:0 and 16:0 fatty acids, while FatB2 can induce the accumulation of large amounts of C10 or C12 fatty acids.

The structure of these enzymes has been characterized by molecular docking, based on analogous enzymes from bacteria (Mayer and Shanklin, 2005; Serrano-Vega et al., 2005), and also by X-Ray diffraction of a FatB2 from *Umbellularia californica* (Feng et al., 2017). The active enzyme is a homodimer, whose subunits adopt hot-dog folding, and it has a papain like catalytic triad. The amino acids involved in the substrate-enzyme interaction have been studied by site-directed mutagenesis and structural docking, defining the Met-197, Arg-199 and Thr-231 residues as determinants for the kinetic parameters of the enzyme (Yuan et al., 1995; Feng et al., 2018).

The importance of acyl-ACP thioesterases in fatty acid metabolism and oil crop biotechnology has made them the focus of efforts to find divergent forms of these enzymes that might be of interest for the production of specialist oils. Here, for the first time we have investigated the thioesterase system in *K. paniculata*. The Fat genes were cloned from this species, sequenced and then expressed in *E. coli* for characterization. The FatA enzyme from this plant, *Kp*FatA, displayed the expected specificity profile, with high activity towards oleoyl-ACP. However, the FatB form had an unusual specificity profile, which curiously was more similar to that of FatA thioesterases. The specificity of *Kp*FatB was investigated in relation to its enzyme structure using techniques of molecular docking. Its impact on the fatty acid composition of *K. paniculata* was discussed in the light of these results.

2. Material and methods

2.1. Biological material and growth conditions

Developing seeds from *K. paniculata* were harvested from 25 to 30 year old trees growing in Seville (Spain: 37°20′53.2″N 5°58′40.0″W) and frozen until use.

The *Escherichia coli* XL1-Blue (Stratagene) strain was used for gene cloning and the BL21(DE3) strains for protein expression. Unless otherwise indicated, the bacteria were grown in either solid or liquid LB medium at 37 °C, supplemented with the appropriate antibiotics. Liquid cultures were shaken vigorously to achieve oxygenation.

2.2. mRNA preparation and cDNA synthesis

mRNA was isolated from 3 seed pools replicates and the cDNA was synthetized as previously described by Moreno-Pérez et al. (2011).

2.3. Cloning of K. paniculata acyl-ACP thioesterase genes

Conserved regions from available FatA and FatB plant gene sequences, and the sequences obtained from RNA-seq experiment of *K. paniculata* developing kernels (data not published), were used to design specific primers in order to amplify the full sequences encoding the *Kp*FatA and *Kp*FatB proteins: KpFatA_F1, KpFatA_R1, KpFatB_F1, and KpFatB_R1 (Table S1). As expected, both proteins corresponding to cloned genes, were predicted to be located to the chloroplast (DeepLoc-1.0; Almagro-Armenteros et al., 2019), and the transit peptides for plastid translocation were predicted using the ChloroP 1.1 (Emanuelsson et al., 1999) and TargetP-2.0 (Almagro-Armenteros et al., 2017) software. Primers with internal *Sal*I and *Pst*I restriction sites for *KpFatA* (KpFatA-SalI_F2 and KpFatAPstI_R2) or *Sac*I and *Sma*I sites for *KpFatB* (KpFatB-SacI_F2 and KpFatBSmaI_R2) were then designed to amplify the coding regions of the mature proteins for their endogenous expression in *E. coli*. In the case of *Kp*FatB, the membrane domain located close to the transit peptide at the N-terminal extreme was also removed (predicted using the HMMTop and PolyPhobius online tools: Tusnády and Simon, 2001; Käll et al., 2007). The resulting PCR products were cloned into the *Sal*I-*Pst*I (*Kp*FatA) and *Sac*I-*Sma*I (*Kp*FatB) restriction sites of the pQE-80L expression vector (Qiagen). The resulting constructs were named pQE80L_KpFatA and pQE80L_KpFatB, and they were used to transform competent bacteria.

2.4. Sequence analysis and phylogenetic tree constructions

K. paniculata thioesterase proteins were confirmed and they were identified as acyl-ACP thioesterases through alignment to the homologous sequences from other plant species using the alignment search tool BLASTp (Camacho et al., 2009). Alignments and phylogenetic tree were performed as described previously (Martins-Noguerol et al., 2020).

2.5. Protein expression and purification

Cells from *E. coli* BL21 (DE3) strain transformed with the pQE80L-KpFatA or pQE80L-KpFatB plasmids were grown overnight at 37 °C in 2 mL liquid LB medium supplemented with ampicillin, and the resulting cultures were used to start the 1 L inoculum of the same medium. These cultures were induced at OD600 0.4–0.5 by adding 0.5 mM IPTG and grown for a further 4 h, the bacteria were then pelleted by centrifugation (15 min at 3000 x g) and washed with distilled water before resuspending them in Binding Buffer (20 mM NaH_2PO_4 [pH 7.4], 20 mM imidazole and 500 mM NaCl). The cells were then disrupted by sonication (Branson, model 4C15). After centrifuging at 30,000 x g and 4 °C for 10 min, the supernatant was recovered and used to purify the $(His)_6$-tagged recombinant proteins using SpinTrap Kit (GE Healthcare), according to the manufacturer's instructions.

2.6. Fatty acid analysis from E. coli cultures

BL21 (DE3) *E. coli* cells transformed with pQE-80L::KpFatA or pQE-80L::KpFatB were grown in 25 mL of LB medium and induced as described above. Cells harbouring the empty vector pQE-80L were also grown in parallel as a control. The cells were harvested 4 h after induction, washed with distilled water and centrifuged for 10 min at 2500 x g. The *E. coli* pellets were then transmethylated using the one-step method reported previously (Garcés and Mancha, 1993), and subsequently modified (Sánchez-García et al., 2010). The experiment was repeated in 3 independent batches.

2.7. Transient expression in Nicotiana benthamiana leaves

For functional transient expression in tobacco, the full *KpFatA* and *KpFatB* cDNAs were cloned into the pBIN19:35S vector. *Agrobacterium tumefaciens* cells were transformed with pBIN19:35S constructs and

Table 1. Typical fatty acid composition of *K. paniculata* seeds measured in this work and reported by other authors.

Fatty acids (mol%)							References
16:0	18:0	$18:1^{\Delta 9}$	$18:2^{\Delta 9,12}$	$18:3^{\Delta 9,12,15}$	20:0	$20:1^{\Delta 11}$	
6	1	32	13	1	1	46	Hopkins and Swingle (1967)
9.7	1.8	25.5	8.5	3.6	2.4	48.5	Khan et al. (2020)
5.3	1.1	31.1	11.6	5.2	0.6	45.1	Present study

cultures with individual clones were co-infiltrated with the P19 gene from the tomato bushy stunt virus onto the leaves of 5-week-old *N. benthamiana* plants (Popescu et al., 2007). Independent disks were punched out of 4 different tobacco leaves 4 days after *Agrobacterium* infiltration and the fatty acid composition was analysed according to the protocol described above.

2.8. Western blots

After protein separation by SDS-PAGE, the proteins were transferred from the gel to a PVDF membrane (Transblot TurboTM Mini PVDF Transfer Packs, Bio-Rad) in a TransBlot TurboTM Transfer Starter System (Bio-Rad). The membrane was then blocked for 1 h with a 5% solution of non-fat dry milk in Tris-buffered saline solution (TBS) and after washing three times with TBS, it was probed for 2 h at 4 °C with an anti-polyHistidine-Peroxidase conjugate (Sigma-Aldrich) diluted 1/2,000 in TBS. The recombinant proteins were then visualized using the ECL Western Blotting Detection Kit (GE Healthcare).

2.9. Holo-ACP activation and preparation of the acyl-ACP substrates

Labelled acyl-ACP substrates for the thioesterases were prepared from the recombinant *E. coli* histidine-tagged ACP and acyl-ACP synthetase constructs kindly provided by Dr Penny Von Wettstein-Knowles (University of Copenhagen, Copenhagen, Denmark) and Dr John Shanklin (Brookhaven National Laboratory, Upton, NY, USA). The recombinant ACP was purified on a NTA-Ni column by affinity chromatography (see Haas et al., 2000) and activated using recombinant *E. coli* holo-ACP synthase as described previously (Lambalot and Walsh, 1995).

Radiolabelled acyl-ACPs were prepared as described by Aznar-Moreno et al. (2016).

2.10. Acyl-ACP thioesterase assay

Thioesterase activity was assayed by monitoring the production of radiolabelled free fatty acids from the [1–^{14}C] acyl-ACP substrates as indicated above. The reaction medium contained 50 mM Tris-HCl [pH 8.0], 5 mM DTT, 50–2500 Bq of [1–^{14}C] acyl-ACP and 0.025–1.0 ng of the recombinant enzyme in a final volume of 0.1 mL. The reactions were carried out at room temperature for 5 min and quenched by adding 0.25 ml of 1 M acetic acid in 2-propanol. Hydrolysed free fatty acids were then extracted twice with 0.3 ml hexane and the radioactivity was measured in a liquid scintillation counter (Rackbeta II; LKB).

2.11. Modelling of K. paniculata thioesterase proteins

Putative *Kp*FatA and *Kp*FatB protein structures were modelled by homology-modeling using the Swiss Model Workspace (Waterhouse et al., 2018; https://swissmodel.expasy.org/), based on their sequences, the 12:0-ACP thioesterase crystal structure from *U. californica* as a template (Protein Data Bank accession 5 × 04: Feng et al., 2017) and default primary parameters. Furthermore, molecular docking was performed with SwissDock (Grosdidier et al., 2011a, 2011b; http://www.swissdock.ch/) using palmitic, stearic and oleic acids as substrates (with default values for Docking type, region of interest and flexibility), and confirmed through I-TASSER (Yang et al., 2015; http://zhanglab.ccmb.med.umich.edu/I-TASSER/) using option I to specify those fatty acids as templates and without any contact or distance restraints. The UCSF Chimera package was applied for critical residue mapping and model visualization (Pettersen et al., 2004).

2.12. Statistical analysis

The data represented in this work are presented as the means of 3 or 4 measurements and the standard deviation (SD). One-way ANOVA was used to establish significant differences between the measurements, combined with a Tukey post hoc analysis at a significance level of 0.05%. All calculations were made using Sigma Plot 14.0 (Systat 340 Software Inc.).

3. Results and discussion

3.1. Fatty acid composition of K. paniculata seeds

The *K. paniculata* tree produces seeds that accumulate oil from the beginning of spring to the end of summer. The oil produced by several species of Sapindaceae has been studied previously, with the characteristic presence of large proportions of cyanolipids. The synthetic pathways leading to the generation of these compounds are poorly understood, although it is thought that they are produced by condensation of acyl-CoA derivatives with leucine, followed by certain modifications (Seigler and Butterfield, 1976). In this study, we specifically addressed the fatty acid composition of this seed oil (see Table 1) and how it is produced. Apart from its cyanolipid content, *K. paniculata* oil is remarkable for its low saturated fatty acids content (7%) and the high levels of gondoic acid ($20:1^{\Delta 11}$), which were as high as 45.1% of the total acyl composition, as seen previously (Hopkins and Swingle, 1967; Khan et al., 2020). Gondoic acid is present in the oil of many species, especially in cruciferous like rapeseed or *Arabidopsis*, although it is generally stored at low proportions due to its subsequent elongation or desaturation, events that do not occur to the same extent in *K. paniculata* where it is stored as the main very long chain fatty acid, VLCFA, species. Gondoic acid is synthesized by extraplastidial elongation of oleoyl-CoA due to the action of the fatty acid elongase, a type 2 enzymatic complex located in the endoplasmic reticulum (ER: Leonard et al., 2004). Thus, it is hypothesised that developing *K. paniculata* seeds have developed a high oleate flux out of plastids, molecules that are then elongated to oleoyl-CoA by a highly specific elongase that does not continue the elongation of the initial reaction product, gondoyl-CoA, to longer fatty acid derivatives. Gondoyl-CoA is then used in glycerolipid and cyanolipid synthesis by the enzymes expressed by this plant. Here we have studied the acyl precursor supply for gondoic acid synthesis, which takes place through the activity of acyl-ACP thioesterases, enzymes that terminate intrapastidial fatty acid synthesis, and that determine the flux and composition of the *de novo* fatty acids synthesized in developing seeds.

3.2. Cloning and sequence characterization

The *K. paniculata* thioesterases were cloned from cDNA transcribed from the developing seed mRNA isolated using oligonucleotides designed from the consensus sequences of different plant FatA and FatB enzymes (Supplementary material, Table S1). The cloning procedure involved

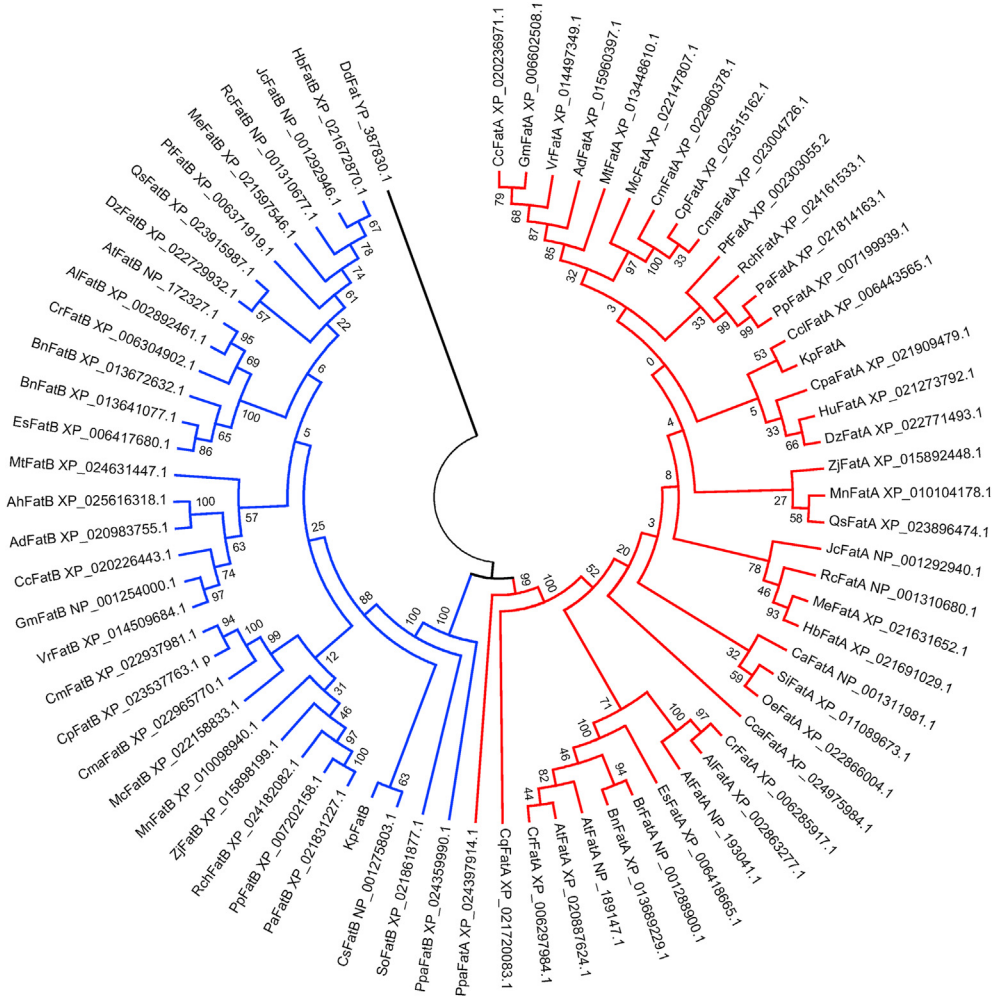

Figure 1. Phylogenetic comparison of the plant acyl-ACP thioesterase genes. The red subtree represents the FatA thioesterases and the blue subtree the FatB thioesterases. The plant species included in the phylogenetic tree were: Ad *Arachis duranensis*, Ah *Arachis hypogaea*, Al *Arabidopsis lyrata*, At *Arabidopsis thaliana*, Bn *Brassica napus*, Br *Brassica rapa*, Ca *Capsicum annuum*, Cc *Cajanus cajan*, Cca *Cynara cardunculus*, Ccl *Citrus clementina*, Cm *Cucurbita moschata*, Cma *Cucurbita máxima*, Cp *Cucurbita pepo*, Cpa *Carica papaya*, Cq, *Chenopodium quinoa*, Cr *Capsella rubella*, Cs *Citrus sinensis*, Da *Desulfovibrio alaskensis*, Dz *Durio zibethinus*, Es *Eutrema salsugineum*, Gm *Glycine max*, Hb *Hevea brasiliensis*, Hu *Herrania umbrática*, Jc *Jathropa curcas*, Kp *Koelreuteria paniculata*, Mc *Momordica charantia*, Me *Manihot esculenta*, Mn *Morus notabilis*, Mt *Medicago truncatula*, Oe *Olea europea*, Pa *Prunus avium*, Pp *Prunus pérsica*, Ppa *Physcomitrella patens*, Pt *Populus trichocarpa*, Qs *Quercus suber*, Rc *Ricinus comunis*, Rch *Rosa chinensis*, Si *Sesamun indicum*, So *Spinacia oleracea*, Vr *Vigna radiata*, Zj *Ziziphus jujuba*. The *Desulfovibrio* gene was used as outgroup.

isolating the gene fragment targeted and then amplifying the full sequence through RACE technology, the approach used to clone the thioesterase genes of many species (Sánchez-García et al., 2010; Moreno-Pérez et al., 2011). The sequences cloned were deposited in GeneBank under the codes MG594388 and MK341067, and they encode two pre-proteins of 371 and 416 amino acids, respectively, and with a molecular weight of 42.3 kDa for *Kp*FatA and 46.2 kDa for *Kp*FatB. Both were basic proteins, with a pI of 8.41 and 6.14, like all other thioesterases described previously (Sánchez-García et al., 2010; Moreno-Pérez et al., 2011). Moreover, they were both predicted to be chloroplast/plastid proteins with signal peptides, which when removed gave rise to the mature *Kp*FatA and *Kp*FatB proteins of 317 (36.18 kDa) and 368 (41.43 kDa) amino acids, respectively. In addition, and as described previously for most FatB1 proteins, *Kp*FatB had a short membrane binding sequence at its N-terminus for binding to the inner plastid membrane (supplementary material, Figure S1: Aznar-Moreno et al., 2018).

3.3. Phylogenetic study

The amino acid sequences of *Kp*FatA and *Kp*FatB were used to construct a phylogenetic tree that included the FatA and FatB thioesterases from 40 different species, with the tree rooted in the bacterial thioesterase from *Desulfovibrio desulfuricans* as an outgroup protein (Figure 1). *Kp*FatA clustered together with the form of *Citrus clementina* and close to the FatAs of other tree species like *Carica papaya*, *Hevea brasiliensis* and *Durio zibethinus*. The tree showed that while belonging to evolutionarily distant species, *Kp*FatA maintained a high degree of conservation, indicating an essential function in plant metabolism as it is the supply of oleate for complex lipid synthesis. Oleic acid is the precursor of linoleate and linoleate, the most abundant fatty acids in all plant tissues. These fatty acids are required for the correct arrangement of membranes and they are precursors necessary for plants to respond correctly to environmental stress (Kachroo and Kachroo, 2009). A higher degree of divergence was observed among the different forms of FatB. The *Kp*FatB also clustered together with the FatB from *C. clementina* in a branch that was very distant from that containing the FatB enzymes present in the most common crops and oilseeds. This suggests the early differentiation of this FatB and weaker evolutionary pressure on the conservation of this enzyme, which is indeed known to display a more ample range of substrate specificities and is involved in saturated fatty acid metabolism. This facet indicates that *Kp*FatB could have diverged both functionally and in other terms from many of the previously studied FatB1 enzymes, an issue that is dealt with in more detail below.

3.4. Structural modelling

The structure of plant thioesterases was initially modelled through analogy to bacterial proteins (Mayer and Shanklin, 2005; Serrano-Vega et al., 2005) and later, using the structure of the 12:0-ACP thioesterase from *U. californica* (*Uc*FatB1) as a template (Aznar-Moreno et al., 2018), a structure previously determined through X-ray diffraction studies (Feng et al., 2017). These enzymes are homodimers whose subunits contain an active site that adopts a hot-dog arrangement and that contains a papain like catalytic triad. The C-terminal domains of the FAT enzymes

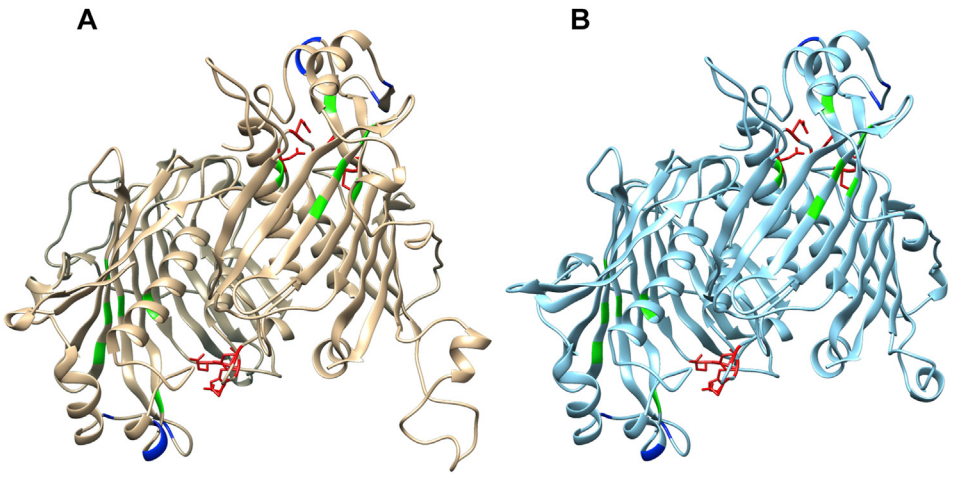

Figure 2. Proposed structural models for the *Kp*FatA (A) and *Kp*FatB (B) homodimers. In the ribbon diagrams, the hypothetical residues involved in catalysis are in red (FatA D263, N265, H267, E301 and C302; FatB D311, N313, H315, E349 and C350), those involved in acyl interaction in green (FatA G115, I141, T175, R177, W179 and S195; FatB A167, I193, M227, R229, W231 and S247) and the positive charged residues involved in the interaction with ACP in blue (FatA R136, K137, R205, R206 and K209; FatB K189, R257, R258 and K261).

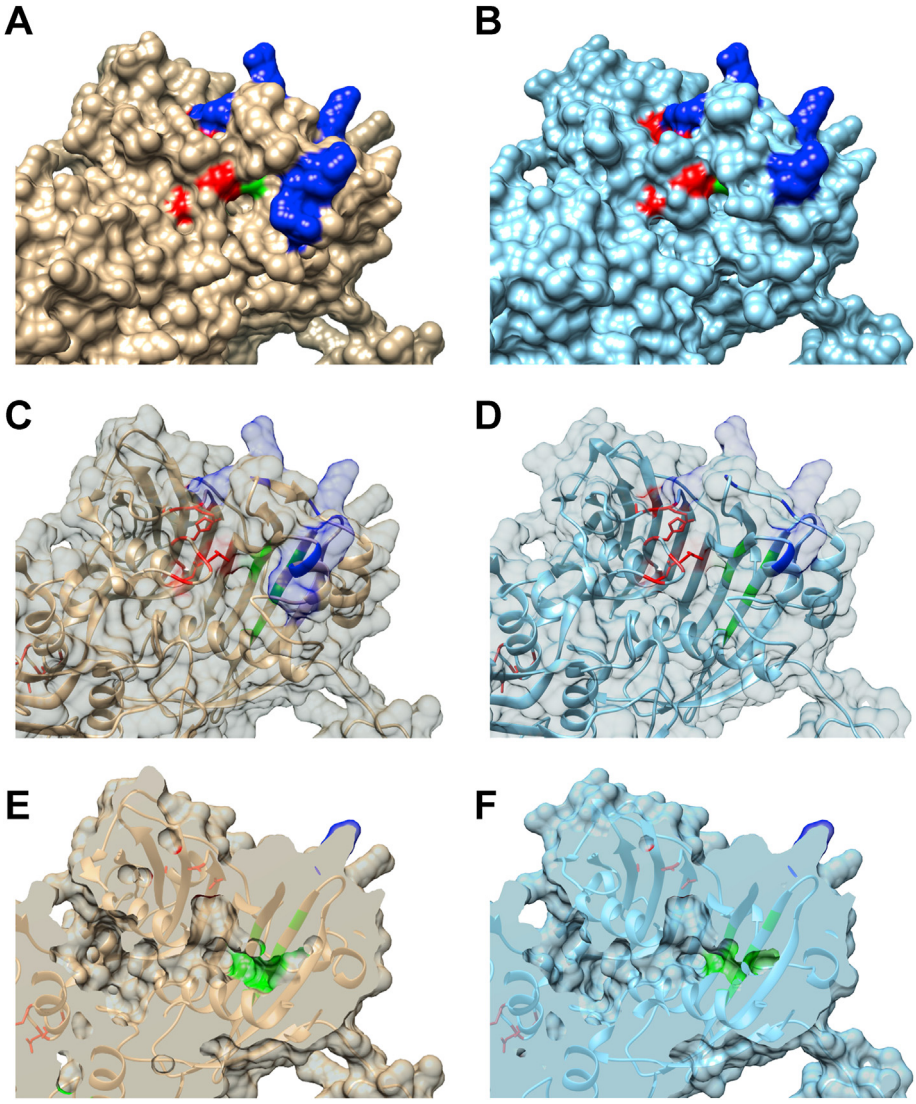

Figure 3. Molecular surfaces of the proposed *Kp*FatA (A, C and E) and *Kp*FatB (B, D and F) thioesterase structures showing the residues involved in the interaction with the ACP molecule (blue), in the acyl chain contact forming part of the hydrophobic substrate pocket (green) and those involved in catalysis (red). Panels E and F represent slab views of the *Kp*FatA (E) and *Kp*FatB (F) substrate binding pockets.

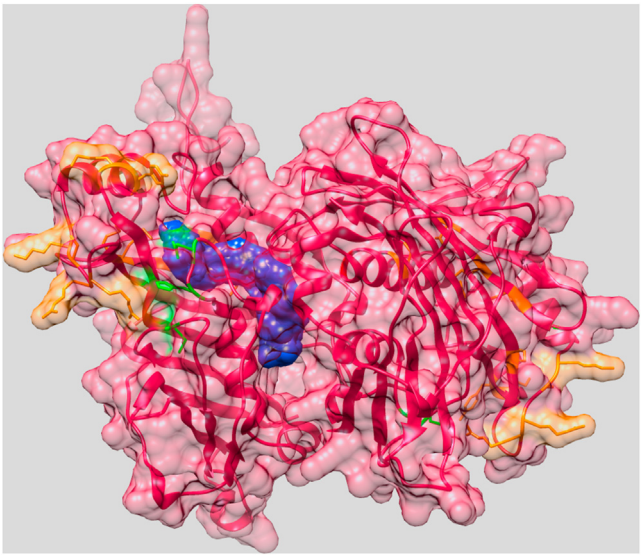

Figure 4. Docking of oleic acid in the *Kp*FatB binding pocket. The polar positive charged residues surrounding the opening of the substrate pocket are in orange, the catalytic residues are in green and the oleic acid surface is in dark blue.

discovered retain strong homology and they contain the residues essential for catalysis, whereas the N-terminal domains are more variable and they determine the enzyme's substrate specificity (supplementary material, Figure S2).

For the structural modelling of *K. paniculata* thioesterases the transit peptide of *Kp*FatA (1–54) and the transit peptide plus the membrane anchoring domain of *Kp*FatB (1–104) were removed. The degree of identity of *Kp*FatA with regards the *Uc*FatB2 template was 40.82%, with a higher level of conservation at the C-terminal domain. Residues D263, N265, H267, E301 and C302 of this enzyme corresponded to highly conserved catalytic residues. As expected, the intermediate region between the C-terminal and N-terminal end displayed much weaker conservation. A three-dimensional model of *Kp*FatA was generated (Figure 2A), demonstrating the differences among different thioesterases despite their generally well-conserved structure. In the region involved with substrate interaction there was an external area containing positively charged residues (Serrano-Vega et al., 2005) that interacts with the ACP protein, and a hydrophobic pocket that interacts with the acyl moiety of the substrate (Figure 3A). The cavity of *Kp*FatA was similar to that of other FatA enzymes (Figure 3C and E), and larger in volume than that of FatB to accommodate unsaturated fatty acids (Aznar-Moreno et al., 2018).

In the case of *Kp*FatB, the structural model (Figure 2B) fitted well with the structural details of other FatB enzymes. Thus, this thioesterase displayed 61.19% identity with *Uc*FatB1, with the residues D311, N313, H315, E349 and C350 participating in catalysis. However, a more detailed analysis of the *Kp*FatB sequence and a model of its tertiary structure pointed to some differences in the area corresponding to the binding pocket relative to other enzymes of the same family (Figure 3B, D and F). Thus, the residues G226 and V248 in the binding pocket area (close to M227 and S247, respectively) are amino acids with smaller lateral chains and they are not present in other FatB emzymes, such as *Ha*FatB (Figure S2). These residues conferred a larger volume to the binding pocket (at least 48.4 $Å^3$), which probably affects its interaction with acyl-ACP substrates and alters its specificity, better adapting to oleoyl acyl chains (as shown in Figure 4).

3.5. Heterologous expression in E. coli

The *KpFatA* and *KpFatB* coding sequences, excluding the regions encoding the signal peptides, were subcloned into the pQE-80L vector and expressed in the BL21 (DE3) strain of *E. coli*. In the case of *Kp*FatB, the peptide from residues 87 to 104 that is responsible for membrane anchoring was also removed to favour its accumulation in the soluble phase, as is common when expressing these types of thioesterases heterologously (Sánchez-García et al., 2010; Moreno-Pérez et al., 2011). The transformed bacteria displayed identical growth to those of the control strains carrying the empty pQE-80L plasmid (data not shown). To study the effect of the expression of this enzyme on the host's metabolism and to gain evidence of the enzyme's functionality, the fatty acid composition of the bacteria was analysed after the induction of plasmid expression (Table 2). The expression of *Kp*FatA did not significantly alter the bacteria's fatty acid profile, which remained identical to that of the control line. By contrast, *Kp*FatB expression significantly increased the levels of 14:0, 16:1 and 18:1 at the expense of lower proportions of the 17:0Δ and 19:0Δ cyclopropane fatty acids. Long chain and very long chain saturated fatty acids (16:0, 18:0 and 20:0) remained at the same levels as those found in the control cells, whereas the ratio of unsaturated to saturated fatty acids decreased slightly upon expression of this enzyme. The effect of *Kp*FatB on *E. coli* fatty acids differed to that seen with other forms of this enzyme. Thus, FatB1s generally displayed specificity towards palmitoyl-ACP and they hydrolyse this derivative avoiding further elongation and desaturation, enhancing the amount of palmitate at the expense of unsaturated derivatives in the membranes of bacteria (Voelker and Davies, 1994; Sánchez-García et al., 2010). However,

Table 2. Fatty acid composition of *E. coli* cells transformed with the recombinant plasmid. The data are the average and standard deviation of three independent samples.

	pQE-80L	pQE-80L::*Kp*FATA	pQE-80L::*Kp*FATB
14:0	5.59 ± 0.12	5.58 ± 0.06	*9.77 ± 1.67
16:0	39.54 ± 0.99	39.72 ± 0.14	38.02 ± 0.83
$16:1^{\Delta 9}$	3.19 ± 1.20	3.97 ± 0.16	*6.21 ± 0.11
17:0Δ[a]	21.83 ± 0.10	22.05 ± 0.17	*14.72 ± 0.50
18:0	0.51 ± 0.21	0.67 ± 0.15	0.52 ± 0.23
$18:1^{\Delta 11}$	20.19 ± 0.65	19.67 ± 0.12	*24.43 ± 1.03
19:0Δ[b]	8.40 ± 0.77	7.64 ± 0.21	*5.50 ± 0.23
20:0	0.76 ± 0.02	0.71 ± 0.05	*0.82 ± 0.03
UFA[c]/SFA[d]	1.16 ± 0.05	1.14 ± 0.02	*1.04 ± 0.04
Total mg fatty acids/OD_{600nm} unit	1.03 ± 0.07	1.01 ± 0.05	*1.38 ± 0.07

*Significant differences relative to the control ($p < 0.05$).

[a] cis-9.10-Methylene-hexadecanoic acid, a cyclopropane derivate from $16:1^{\Delta 9}$.

[b] cis-11.12-Methylene-octadecanoic acid, a cyclopropane derivate from $18:1^{\Delta 9}$.

[c] Unsaturated fatty acids and derivatives: $16:1^{\Delta 9}$ + 17:0Δ + $18:1^{\Delta 11}$ + 19:0Δ.

[d] Saturated fatty acids 14:0 + 16:0 + 18:0 + 20:0.

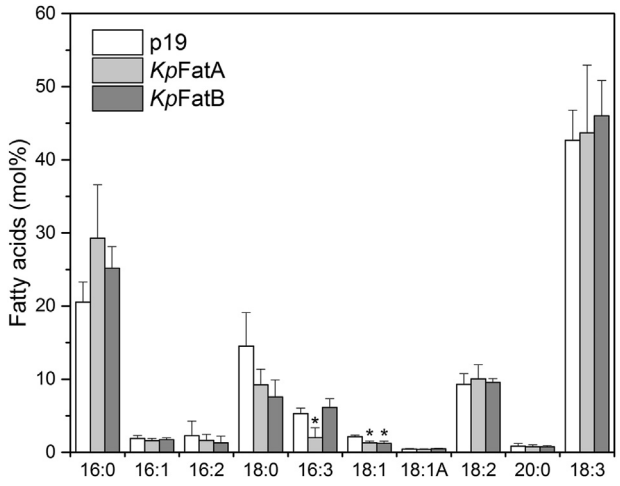

Figure 5. Fatty acid composition of *N. benthamiana* leaves expressing the *Kp*FatA and *Kp*FatB thioesterases. The data represent the average of 3–4 independent replicates ± standard deviation: *Significant differences compared to with the control ($p < 0.05$).

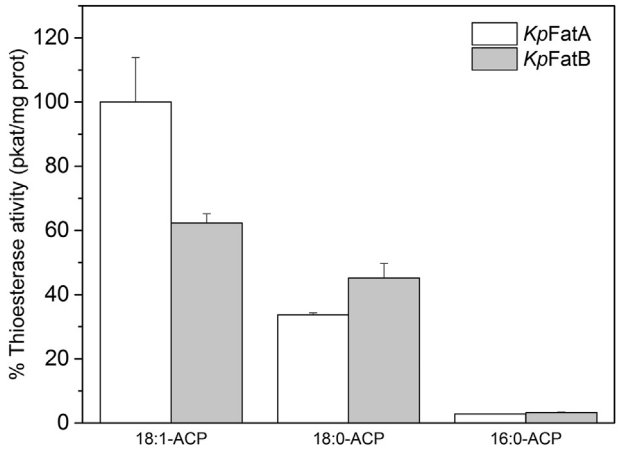

Figure 6. Relative *Kp*FatA and *Kp*FatB activity on different acyl-ACPs substrates. The data represent the average of 3 independent replicates ± standard deviation.

*Kp*FatB seemed to provoke a different effect, which might indicate it possesses a substrate specificity that differs from that most commonly associated with these enzymes.

3.6. Heterologous transient expression in N. benthamiana leaves

The characterization of *K. paniculata* thioesterases was completed by studying their impact on the lipids produced by *N. benthamiana*. This adds information regarding the functionality of these genes in plant cells, a system closer to its physiological environment. The fatty acid composition of tissue transformed with the recombinant plasmids and a p19 viral suppressor was analysed, and in that of control tissue transformed with the p19 construct alone. The *K. paniculata* thioesterases seemed to produce an increase in 16:0 fatty acids in conjunction with a slight decrease in 18:1 fatty acids (Figure 5). The expression of *Kp*FatA also provoked a significant decrease in 16:3 fatty acids. When active thioesterases are transiently expressed in leaf tissue, no important changes in fatty acid composition are usually observed (Moreno-Pérez et al., 2014). The cells in this system already possessed high levels of endogenous thioesterases and the presence of the exogenous ones usually only contributes to the hydrolysis of precursors, enhancing the content of saturated fatty acids while decreasing that of oleate and its derivatives. In this case, only a small decrease in oleate was detected. Moreover, the increase in saturated fatty acids (palmitic) although perceptible was not significant due to the high variability in the concentrations registered in these experiments (Figure 5).

3.7. Enzyme purification and kinetic characterization

The vector pQE-80L allows a 6xHis-tag to be added to the N-terminal end of *K. paniculata* thioesterases, facilitating their purification on a NTA-Nickel affinity column and characterization. Both enzymes were obtained at a high degree of purity through this approach (Supplementary Figure S3) and both of the purified proteins reacted strongly with an antibody against the histidine residues, with an expected molecular weight of 36.4 and 34.0 kDa for *Kp*FatA and *Kp*FatB, respectively (Supplementary Figure S3). The substrate specificity and kinetic parameters of the purified thioesterases were characterized using ACP from *E. coli* substrates in order to determine the affinity in a general way. The specificity profile at a constant substrate concentration (Figure 6) showed the *Kp*FatA enzyme preferred 18:1-ACP, with weaker activity on 18:0-ACP and much lower activity on 16:0-ACP (Figure 6). This was the expected profile for a FatA enzyme, which are responsible for the export of most of the *de novo* synthesized oleate from plastids to the cytosol, as reported for such enzymes isolated from many other species like *H. annuus* (Serrano-Vega et al., 2005; Moreno-Pérez et al., 2014), *R. communis* (Sánchez-García et al., 2010), *M. tetraphylla* (Moreno-Pérez et al., 2011) or *C. sativa* (Rodríguez-Rodriguez et al., 2014). By contrast, the specificity of *Kp*FatB was quite unexpected. This enzyme displayed weak activity on palmitoyl-ACP, whereas it had a clear preference for the hydrolysis of oleoyl- and stearoyl-derivatives (Figure 6). To date, the plant FatB1 thioesterases isolated from different species have displayed an invariable profile of specificity, with high activity towards palmitoyl-ACP and other saturated derivatives like stearoyl-ACP (Sánchez-García et al., 2010; Moreno-Pérez et al., 2011; Rodríguez-Rodriguez et al., 2014). The relative activity on different acyl-ACP substrates can change depending on the ACP protein, although in the case of FatB from sunflower the use of endogenous ACP further enhanced the enzyme's specificity for palmitoyl-ACP (Aznar-Moreno et al., 2016). The specificity profile of *Kp*FatB, differed considerably from that of other FatB1s and it was more similar to that found for FatA enzymes. This profile was consistent with the phylogenetic

Table 3. Kinetic parameters and catalytic efficiencies of histidine-tagged KpFatA and KpFatB expressed in *E. coli* and purified by NTA-Ni affinity chromatography.

	Substrate	V_{max} (pmol/s)/mg prot	K_{cat} (s^{-1})	K_M (μM)	Efficiency (s^{-1} μM^{-1})
*Kp*FatA	16:0-ACP	34.95 ± 1.73	0.40	0.48 ± 0.05	0.83
	18:0-ACP	858.55 ± 201.69	9.85	1.68 ± 0.53	4.91
	18:1-ACP	1377.62 ± 44.15	15.81	0.55 ± 0.03	28.73
*Kp*FatB	16:0-ACP	44.21 ± 4.44	0.56	0.56 ± 0.11	0.99
	18:0-ACP	635.25 ± 130.26	8.07	0.59 ± 0.23	13.59
	18:1-ACP	865.50 ± 151.90	10.99	0.56 ± 0.19	19.39

The data correspond to the average of 3 independent measurements ± standard deviation.

divergence of this thioesterase (Figure 1) with respect to other members of its family and the 3D-structure modelling, which showed it possesses a more voluminous hydrophobic pocket appropriate for larger acyl moieties (Figures 3 and 4).

The kinetic parameters of the *K. paniculata* thioesterases were calculated from the substrate curves by fitting the results to the Lineweaver-Burk equation through a non-linear least-squares regression analysis using OriginPro 2019 software correlated at $P < 0.005$, as determined by a Student's t-test and showed that they displayed sub micromolar K_m values for most of the substrates assayed and the differences in substrate specificities were caused by higher V_{max} values (Table 3). This was common to the other forms of these enzymes studied previously. The highest catalytic efficiencies were achieved with 18:1-ACP in both cases, reaching values of 28.73 and 19.39, respectively. These values were in the order of those found for thioesterases from other species that accumulate oil, such as sunflower, castor bean and macadamia nuts (Moreno-Pérez et al., 2011, 2014; Sánchez-García et al., 2010). The high catalytic efficiency of *Kp*FatB for stearoyl-ACP was also remarkable as it is not usually associated with other FatB1 enzymes.

K. paniculata seeds contain high levels of oleate and of fatty acids derived from it (linoleate, linoleate and gondoic acids), as well as low levels of palmitate relative to the C18 fatty acids. This fatty acid profile fitted well with the specificity of the thioesterases characterized here, which favours the export of C18 fatty acids like stearate and oleate rather than palmitate. The kinetic parameters of *Kp*FatB were in accordance with the effects following its expression in *E. coli*, favouring the accumulation of unsaturated fatty acids rather than saturated derivatives.

4. Conclusions

The present study characterized the thioesterase system present in developing *K. paniculata* seeds, a tree of the Sapindaceae family that accumulates cyanolipids and high levels of gondoic acid in its seeds. This species produces FatA and FatB thioesterases, which were cloned and expressed heterologously in *E. coli* and *N. benthamiana*, the former permitting its purification and characterization. *Kp*FatA displayed a specificity profile similar to other analogous enzymes from other sources, whereas *Kp*FatB had a specificity profile not observed previously, with a stronger preference for oleoyl-ACP and stearoyl-ACP, and weaker activity towards palmitoyl-ACP. Nevertheless, the K_m values and catalytic efficiencies of this *Kp*FatB enzyme were similar to those of other FatB enzymes. The fatty acid profile of the bacteria cultures overexpressing both the *Kp*Fat enzymes was consistent with the data from the enzymatic studies. The phylogenetic study of this *Kp*FatB enzyme clustered it in a group that diverged from the FatB1s in most common oil crops, together with FatB enzymes from tree species like *C. clementina*, *C. papaya*, *H. brasiliensis* and *D. zibethinus*. *Kp*FatB also displayed other structural features that differentiate it from other FatB1 thioesterases, such that it could be considered as a novel type of FatB thioesterase that could be of interest in biotechnology due to its high specificity towards stearoyl-ACP. Accordingly, these enzymes could be used to increase stearate content, a fatty acid that provides oils with a high content of solids and that is considered healthy.

Author contribution statement

Martins-Noguerol, R.: Performed the experiments; Wrote the paper.

DeAndrés-Gil, C.: Performed the experiments.

Garcés, R.: Analyzed and interpreted the data.

Salas, J.J.: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Martínez-Force, E.: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Moreno-Pérez, A.J.: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

This work was supported by the AEI (Agencia Estatal de Investigació) and FEDER Project AGL2017-83449-R.

Competing interest statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2020.e05237.

References

Aichholz, R., Spitzer, V., Lorbeer, E., 1997. Analysis of cyanolipids and triacylglycerols from sapindaceae seed oils with high-temperature gas chromatography and high-temperature gas chromatography-chemical ionization mass spectrometry. J. Chromatogr. A 787, 181–194.

Almagro-Armenteros, J.J., Salvatore, M., Winther, O., Emanuelsson, O., von Heijne, G., Elofsson, A., Nielsen, H., 2019. Detecting sequence signals in targeting peptides using deep learning. Life Sci. Alliance 2, e201900429.

Almagro-Armenteros, J.J., Sønderby, C.K., Sønderby, S.K., Nielsen, H., Winther, O., 2017. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33, 3387–3395.

Aznar-Moreno, J.A., Venegas-Calerón, M., Martínez-Force, E., Garcés, R., Salas, J.J., 2016. Acyl carrier proteins from sunflower (*Helianthus annuus* L.) seeds and their influence on FatA and FatB acyl-ACP thioesterase activities. Planta 244, 479–490.

Aznar-Moreno, J.A., Sánchez, R., Gidda, S.K., Martínez-Force, E., Moreno-Pérez, A.J., Venegas-Calerón, M., Garcés, R., Mullen, R.T., Salas, J.J., 2018. New insights into sunflower (*Helianthus annuus* L.) FatA and FatB thioesterases, their regulation, structure and distribution. Front. Plant Sci. 9, 1496.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinf. 10 (1), 421.

Chen, M., Markham, J.E., Dietrich, C.R., Jaworski, J.G., Cahoon, E.B., 2008. Sphingolipid long-chain base hydroxylation is important for growth and regulation of sphingolipid content and composition in Arabidopsis. Plant Cell 20, 1862–1878.

Dörmann, P., Voelker, T.A., Ohlrogge, J.B., 2000. Accumulation of palmitate in Arabidopsis mediated by the acyl-acyl carrier protein thioesterase FATB1. Plant Physiol. 123, 637–644.

Emanuelsson, O., Nielsen, H., von Heijne, G., 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. 8, 978–984.

Feng, Y., Wang, Y., Liu, J., Liu, Y., Cao, X., Xue, S., 2017. Structural insight into acyl-ACP thioesterase toward substrate specificity design. ACS Chem. Biol. 12, 2830–2836.

Feng, Y., Zhang, Y., Wang, Y., Liu, J., Cao, X., Xue, S., 2018. Tuning of acyl-ACP thioesterase activity directed for tailored fatty acid synthesis. Appl. Microbiol. Biotechnol. 102, 3173–3182.

Garcés, R., Mancha, M., 1993. One-step lipid extraction and fatty acid methyl esters preparation from fresh plant tissues. Anal. Biochem. 211, 139–143.

Grosdidier, A., Zoete, V., Michielin, O., 2011a. Fast docking using the CHARMM force field with EADock DSS. J. Comput. Chem. 32, 2149–2159.

Grosdidier, A., Zoete, V., Michielin, O., 2011b. Swissdock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Res. 39, W270–W277.

Haas, J.A., Frederick, M.A., Fox, B.G., 2000. Chemical and posttranslational modification of *Escherichia coli* Acyl Carrier Protein for preparation of dansyl-acyl carrier proteins. Protein Expr. Purif. 20, 274–284.

Hopkins, C.Y., Swingle, R., 1967. Eicosenoic acid and other fatty acids of *Sapindaceae* seed oils. Lipids 2, 258–260.

Jones, A., Davies, H.M., Voelker, T.A., 1995. Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. Plant Cell 7, 359–371.

Kachroo, A., Kachroo, P., 2009. Fatty acid–derived signals in plant defense. Annu. Rev. Phytopathol. 47, 153–176.

Käll, L., Krogh, A., Sonnhammer, E.L.L., 2007. Advantages of combined transmembrane topology and signal peptide prediction-the Phobius web server. Nucleic Acids Res. 35, W429–432.

Khan, I.U., Yan, Z., Chen, J., 2020. Production and characterization of biodiesel derived from a novel source *Koelreuteria paniculata* seed oil. Energies 13, 791.

Lambalot, R.H., Walsh, C.T., 1995. Cloning, overproduction, and characterization of the *Escherichia coli* Holo-acyl carrier protein synthase. J. Biol. Chem. 270, 24658–24661.

Leonard, A.E., Pereira, S.L., Sprecher, H., Huang, Y.S., 2004. Elongation of long-chain fatty acids. Prog. Lipid Res. 43, 36–54.

Mayer, K.M., Shanklin, J., 2005. A structural model of the plant acyl-acyl carrier protein thioesterase FatB comprised of two helix/4-stranded sheet domains: the N-terminal domain containing residues that affect specificity, the c-terminal domain containing catalytic residues. J. Biol. Chem. 280, 3621–3627.

Mikolajczak, K.L., Smith Jr., C.R., Tjarks, L.W., 1970. Cyanolipids of *Koelreuteria paniculata* Laxm. seed oil. Lipids 5, 672–677.

Martins-Noguerol, R., Moreno-Pérez, A.J., Sebastien, A., Troncoso-Ponce, M.A., Garcés, R., Thomasset, B., Martínez-Force, E., 2020. Impact of sunflower (Helianthus annuus L.) plastidial lipoyl synthases genes expression in glycerolipids composition of transgenic Arabidopsis plants. Sci. Rep. 10 (1), 1–16.

Moreno-Pérez, A.J., Sánchez-García, A., Salas, J.J., Garcés, R., Martínez-Force, E., 2011. Acyl-ACP thioesterases from macadamia (*Macadamia tetraphylla*) nuts: cloning, characterization and their impact on oil composition. Plant Physiol. Biochem. 49, 82–87.

Moreno-Pérez, A.J., Venegas-Calerón, M., Vaistij, F.E., Salas, J.J., Larson, T.R., Garcés, R., Graham, I.A., Martínez-Force, E., 2014. Effect of a mutagenized acyl-ACP thioesterase FATA allele from sunflower with improved activity in tobacco leaves and Arabidopsis seeds. Planta 239, 667–677.

Ohlrogge, J.B., Jaworski, J.G., 1997. Regulation of fatty acid synthesis. Annu. Rev. Plant Biol. 48, 109–136.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera-a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Popescu, S.C., Popescu, G.V., Bachan, S., Zhang, Z., Seay, M., Gerstein, M., Snyder, M., Dinesh-Kumar, S.P., 2007. Differential binding of calmodulin-related proteins to their targets revealed through high-density Arabidopsis protein microarrays. Proc. Natl. Acad. Sci. Unit. States Am. 104, 4730–4735.

Rodríguez-Rodríguez, M.F., Salas, J.J., Garcés, R., Martínez-Force, E., 2014. Acyl-ACP thioesterases from Camelina sativa: cloning, enzymatic characterization and implication in seed oil fatty acid composition. Phytochemistry 107, 7–15.

Salas, J.J., Ohlrogge, J.B., 2002. Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. Arch. Biochem. Biophys. 403, 25–34.

Sánchez-García, A., Moreno-Pérez, A.J., Muro-Pastor, A.M., Salas, J.J., Garcés, R., Martínez-Force, E., 2010. Acyl-ACP thioesterases from castor (*Ricinus communis* L.): an enzymatic system appropriate for high rates of oil synthesis and accumulation. Phytochemistry 71, 860–869.

Seigler, D.S., Butterfield, C.S., 1976. The origin of cyanolipids in *Koelreuteria paniculata*. Phytochemistry 15, 842–844.

Serrano-Vega, M.J., Garcés, R., Martínez-Force, E., 2005. Cloning, characterization and structural model of a FatA-type thioesterase from sunflower seeds (*Helianthus annuus* L.). Planta 221, 868–880.

Tusnády, G.E., Simon, I., 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics 17, 849–850.

Voelker, T.A., Davies, M., 1994. Alteration of the specificity and regulation of fatty acid synthesis of Escherichia coli by expression of a plant medium-chain acyl-acyl carrier protein thioesterase. J. Bacteriol. 176, 7320–7327.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–W303.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2015. The I-TASSER Suite: protein structure and function prediction. Nat. Methods 12, 7–8.

Yuan, L., Voelker, T.A., Hawkins, D.J., 1995. Modification of the substrate specificity of an acyl-acyl carrier protein thioesterase by protein engineering. Proc. Natl. Acad. Sci. Unit. States Am. 92, 10639–10643.

Zorn, K., Oroz-Guinea, I., Brundiek, H., Dörr, M., Bornscheuer, U.T., 2018. Alteration of chain length selectivity of *Candida antarctica* lipase A by semi-rational design for the enrichment of erucic and gondoic fatty acids. Adv. Synth. Catal. 360, 4115–4131.