Phytochemistry 107 (2014) 7–15

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Acyl-ACP thioesterases from *Camelina sativa*: Cloning, enzymatic characterization and implication in seed oil fatty acid composition

Manuel Fernando Rodríguez-Rodríguez, Joaquín J. Salas *, Rafael Garcés, Enrique Martínez-Force

Instituto de la Grasa, CSIC, Avda Padre García Tejero 4, 41012 Seville, Spain

ARTICLE INFO

Article history:
Received 21 May 2014
Received in revised form 16 July 2014
Available online 8 September 2014

Keywords:
Camelina sativa
Acyl-ACP thioesterases
FatA
FatB
Fatty acid bio-synthesis

ABSTRACT

Acyl-acyl carrier protein (ACP) thioesterases are intraplastidial enzymes that terminate *de novo* fatty acid biosynthesis in the plastids of higher plants by hydrolyzing the thioester bond between ACP and the fatty acid synthesized. Free fatty acids are then esterified with coenzyme A prior to being incorporated into the glycerolipids synthesized through the eukaryotic pathway. Acyl-ACP thioesterases belong to the TE14 family of thioester-active enzymes and can be classified as FatAs and FatBs, which differ in their amino acid sequence and substrate specificity. Here, the FatA and FatB thioesterases from *Camelina sativa* seeds, a crop of interest in plant biotechnology, were cloned, sequenced and characterized. The mature proteins encoded by these genes were characterized biochemically after they were heterologously expressed in *Escherichia coli* and purified. *C. sativa* contained three different alleles of both the FatA and FatB genes. These genes were expressed most strongly in expanding tissues in which lipids are very actively synthesized, such as developing seed endosperm. The *Cs*FatA enzyme displayed high catalytic efficiency on oleoyl-ACP and *Cs*FatB acted efficiently on palmitoyl-ACP. The contribution of these two enzymes to the synthesis of *C. sativa* oil was discussed in the light of these results.

1. Introduction

Camelina sativa is a member of the Brassicaceae family (Al-Shehbaz, 1987) that has been cultivated in Europe since the Bronze Age (Bouby, 1998). *Camelina* accumulates high levels of oil in its seeds, representing between 15.5% and 41.7% of the seed weight, which are rich in ω-3 and ω-6 fatty acids. Typically, the major fatty acid species in *Camelina* oil are palmitic acid (16:0, 6.8% of the total fatty acids), stearic acid (18:0, 2.7%), oleic acid (18:1, 16.7%), linoleic acid (18:2, 21.9%), linolenic acid (18:3, 29.3%), gondoic acid (20:1, 13.9%) and erucic acid (22:1, 2.8%: Rodríguez-Rodríguez et al., 2013). The high percentage of polyunsaturated fatty acids in *Camelina* oil makes it unstable and inappropriate as a source of biodiesel or as a biolubricant stock (Fröhlich and Rice, 2005; Ciubota-Rosie et al., 2013).

In higher plants, *de novo* fatty acid biosynthesis takes place in the plastids and it is catalyzed by the fatty acid synthase complex (FAS), which produces aliphatic fatty acids through the sequential elongation of acyl-acyl carrier protein (ACP) derivatives (Pidkowitch et al., 2007). The main product of the FAS complexes are 16:0-ACP and 18:0-ACP. The activity of the intraplastidial stearoyl-ACP desaturase enzyme (SAD) desaturates the latter intermediate to produce oleoyl-ACP (18:1-ACP: Shanklin and Cahoon, 1998). An important part of these acyl moieties are exported out of the plastids to be incorporated into the extraplastidial pathways of glycerolipid synthesis in the endoplasmic reticulum (Ohlrogge et al., 2000). Fatty acids export involves the hydrolysis of the thioester bond between ACP and the acyl group, a reaction catalyzed by acyl-ACP thioesterases (Voelker et al., 1997). Acyl-ACP thioesterases fall within the group of thioester active enzymes, which are classified in 25 families. The enzymes FatA and FatB comprise the family TE14 (http://www.enzyme.cbirc.iastate.edu/). The products of this reaction are ACP, that is recycled for the further synthesis of acyl-ACPs, and free fatty acids that are quickly condensed with coenzyme A by long-acyl-CoA-synthetase (LACS: Koo et al., 2004; Aznar-Moreno et al., 2014).

According to their sequence and acyl-ACP preference, acyl-ACP thioesterases are usually categorized into two sub-families, FatA and FatB (Jones et al., 1995). FatA-type proteins show high substrate specificity towards 18:1-ACP, whereas FatB enzymes are classified into further two subclasses: FatB1, with preference towards long-chain saturated acyl-ACPs, especially 16:0-ACP

Abbreviations: ACP, acyl carrier protein; DAF, days after flowering; UFA, unsaturated fatty acids; QRT-PCR, quantitative real time PCR; SFA, saturated fatty acids.

* Corresponding author. Tel.: +34 954611550; fax: +34 954616790.
E-mail address: emforce@ig.csic.es (J.J. Salas).

http://dx.doi.org/10.1016/j.phytochem.2014.08.014
0031-9422/

(Dörmann et al., 1995); and FatB2 that prefer short/medium-chain saturated acyl-ACPs. FatA and FatB1 enzymes are present in all plants, whereas FatB2 enzymes are only found in species that accumulate C8–C14 fatty acids in their seed oils (Voelker et al., 1997).

The specificity of the acyl-ACP thioesterases is important in defining the fatty acid profile of seed oil and thus, these enzymes are relevant targets to manipulate the fatty acid composition of seed oils. For example, seeds of a double FatA *Arabidopsis thaliana* mutant in which the expression of both genes was reduced had less triacylglycerides and a modified fatty acid profile (Moreno-Pérez et al., 2012). Furthermore, a FatB *A. thaliana* knock-out mutant had approximately 45% lower palmitic acid content, a similar reduction to that produced by FatB silencing with a miRNA (Bonaventure et al., 2003; Belide et al., 2012).

There is currently increased interest in *C. sativa* as a model plant for applied crop research because it is an oil crop that can be easily transformed using floral dip infiltration under vacuum (Lu and Kang, 2008), and it is genetically similar to *A. thaliana* (Collins-Silva et al., 2011). Moreover, this plant has several agronomic advantages, such as a short vegetative period of about 90–120 days (Putnam et al., 1993), low nutritional requirements (Bramm et al., 1990; Zurb, 1997) and the secretion of phytoalexins that inhibit the development of pathogens (Lovett and Jackson, 1980). In the present work, we have cloned and characterized two acyl-ACP thioesterases from *C. sativa*, *Cs*FatA and *Cs*FatB. Furthermore, we have studied their distribution in the plant and their phylogenetic relationship with thioesterases from other species. Finally, the contribution of these enzymes to oil synthesis in *Camelina*, as well as their possible biotechnological applications, are discussed in the light of our results.

2. Results and discussion

2.1. Isolation and sequence analysis of acyl-ACP thioesterases from C. sativa

Using two different degenerate primer pairs, Deg_FatA_F plus Deg_FatA_R and Deg_FatB_F plus Deg_FatB_R (Supplementary data, Table 1), DNA fragments of 458 and 633 bp were amplified from developing *C. sativa* seed cDNA, corresponding to internal fragments of the *CsFatA* and *CsFatB* genes, respectively. Afterwards, we obtained the full-length *CsFatA* cDNA clone of 1110 bp and *CsFatB* cDNA clone of 1251 bp by RACE, as described in the Material and Methods (Supplementary data, Table 1). The alignment of the deduced amino acid sequences showed a high degree of identity with the internal coding regions of known FatA and FatB thioesterase protein sequences. Moreover, specific primers were designed from these sequences to amplify all possible *CsFatA* alleles (CsFatASphI_F3 and CsFatAXmaI_R3, Supplementary data, Table 1), identifying three different alleles: *CsFatA1*, 1110 bp (GenBank accession number AFQ60947.1); *CsFatA2*, 1110 bp (GenBank accession number AFQ60948.1); and *CsFatA3*, 1107 bp (GenBank accession number AFQ60946.1). Similarly, another pair of primers were designed based on the *CsFatB* sequences (CsFatB_F3 and CsFatB_R3, Supplementary data, Table 1), identifying the same number of alleles for this gene too: *CsFatB1*, 1251 bp (GenBank accession number AFQ60949.1); *CsFatB2*, 1251 bp (GenBank accession number AFQ60951.1); and *CsFatB3*, 1251 bp (GenBank accession number AFQ60950.1).

The *CsFatA1*, *CsFatA2* and *CsFatA3* open reading frames (ORFs) identified were predicted to encode proteins of 369, 369 and 368 amino acid residues (Fig. 1), corresponding to calculated molecular masses of 41.47, 41.49 and 41.40 kDa, and with pI values of 7.44, 7.41 and 7.19, respectively. Despite the fact that *CsFatA1*, *CsFatA2* and *CsFatA3* sequences do not appear to contain the consensus sequence for translation initiation in plants (Joshi et al., 1997), they conserve the T at the position +4 from the ATG, which is required for fidelity of translation initiation (Kozak, 1991). The ORFs of *CsFatB1*, *CsFatB2* and *CsFatB3* identified were each predicted to encode proteins of 416 amino acid residues (Fig. 2), with estimated molecular masses of 46.15, 46.11 and 46.11 kDa, and similar pI values of 8.83, 8.82 and 8.82 respectively. The nucleotides around the methionine at the start codon of *CsFatB1*, *CsFatB2* and *CsFatB3* resemble the consensus sequence for translation initiation in plants (Joshi et al., 1997). The papain-like catalytic triad required for the reaction catalyzed by thioesterase enzymes is formed by asparagine, histidine and glutamine (Mayer and Shanklin, 2007). These residues were conserved in all the alleles identified, occupying the position Asn-273, His-275 and Gln-311 in *Cs*FatA and Asn-318, His-320 and Gln-384 in *Cs*FatB (Figs. 1 and 2).

WoLF PSORT, a program that predicts subcellular localization, was used to identify the signal peptide and the cleavage site where the mature protein begins (Horton et al., 2007), with the information available for acyl-ACP thioesterases from *A. thaliana* (Dörmann et al., 1995), *Helianthus annuus* (Martínez-Force et al., 2000; Serrano-Vega et al., 2005) *Gossypium hirsutum* (Huynh et al., 2002), *Diploknema butyracea* (Jha et al., 2006), *Madhuca longifolia* (Ghosh and Sen, 2007), *Ricinus communis* (Sánchez-García et al., 2010) and *Macadamia tetraphylla* (Moreno-Pérez et al., 2011). Val-59 was the best candidate to represent the N-terminal amino acid of the mature *Cs*FatA protein, corresponding to a signal peptide of 58 amino acid residues (Fig. 1). In the case of the *Cs*FatB protein, we considered Leu-93 to be the first amino acid of the mature enzyme, which would correspond to a signal peptide of 92 amino acids (Fig. 2). Taking into account the presence of a signal peptide in both classes of thioesterases, the mature *Cs*FatA and *Cs*FatB proteins were predicted to contain 310 and 324 amino acids, with molecular masses of 34.8 and 33.4 kDa, and pI values of 5.5 and 6.9, respectively. The *Cs*FatB enzyme has a hydrophobic domain between residues Leu-93 to His-128, which is absent in *Cs*FatA thioesterases and that is probably involved in anchorage to the membrane rather than affecting its activity or affinity for different substrates (Jones et al., 1995; Facciotti and Yuan, 1998).

A phylogenetic tree was generated for the novel *Camelina* thioesterase genes based on their deduced amino acid sequences and in relation to all other known plant thioesterase sequences (Supplementary data, Fig. 1). Each set of three *Camelina* thioesterase proteins were situated in the FatA or FatB thioesterase group. Thioesterases from *Camelina* grouped very closely with those from *A. thaliana*, *A. lyrata*, *Capsella rubella*, *Brassica juncea*, *Brassica napus* and *Brassica rapa*, all species that belong to the Brassicaceae family. Other closely related sequences were those from members of the Fabaceae family, *Glycine max*, *Arachis hypogaea*, *Cicer arietinum* and *Medicago truncatula*.

2.2. Genomic organization of C. sativa acyl-ACP thioesterase genes

To analyze the genomic organization of the *CsFatA* and *CsFatB* genes, two genomic DNA fragments at the locus were amplified using two different primer pairs. Clones of 1214 and 1427 nucleotides were obtained and sequenced for *CsFatA* and *CsFatB*, respectively. The intron and exon organization of the three *CsFatA* and *CsFatB* alleles were found by comparing their cDNA and genomic DNA sequences (Supplementary data, Table 2). The *CsFatA1* allele was 1743 bp long, the *CsFatA2* allele was 1745 bp and the *CsFatA3* allele was 1773 bp. All *CsFatA* alleles had six introns and thus, they each contained seven exons, which were of similar length (Supplementary data, Table 2). Intron 3 of the *CsFatA3* allele differed most in length as it contained an insertion of 32 nucleotides at the beginning of the sequence (169 bp in i3CsFatA1; 171 bp in i3CsFatA2; and 200 bp in i3CsFatA3: Supplementary data, Table 2). The