Journal Of Oil Palm Research (Special Issue) p. 81-87

ISOLATION OF A cDNA CLONE ENCODING AN AWL-AWL CARRIER PROTEIN THIOESTERASE FROM THE MESOCARP OF OIL PALM (*Elaeis guineensis*)

Keywords: oil palm, acyl-ACP thioesterase, cDNA.

ABRIZAH, 0"; LAZARUS, C M and STOBART, A K****

* Palm Oil Research Institute of Malaysia. P.O. Box 10620. 50720 Kuala Lumpur, Malaysia.

** School of Biological Sciences. University of Bristol, Woodland Road. Bristol BS8 1UG, United Kingdom.

***A** 1450 base pair (bp) cDNA encoding an acyl-ACP thioesterase was isolated from a developing mesocarp cDNA library. The cDNA sequence corresponding to the mature protein is 66% identical to* Cuphea hookeriana Ch *FatB1, and 35% identical to* Garcinia mangostana Garm *FatA1, indicating that this clone belongs to the FatB type of acyl-ACP thioesterases. Although the sequence showed the cDNA to be incomplete, lacking the 5' untranslated region and most of the transit peptide, the mature protein is intact and will be expressed in* Escherichia coli *to investigate the substrate specificity of the acyl-ACP thioesterase.*

INTRODUCTION

In plants, fatty acids are synthesized *de novo* in chloroplasts and in other plastids of non-photosynthetic tissues. Fatty acid biosynthesis proceeds by the sequential addition of malonyl units on a growing acyl chain bound to an acyl carrier protein (ACP) until such time as it reaches a chain length of Cl6 or C18. The 16- and la-carbon acyl-ACP products can act as substrates for other enzymes including desaturases, thioesterases and acyltransferases. Acyl-ACP thioesterase catalyzes the hydrolysis of acyl-ACP to ACP and free fatty acids, which are transported to the cytoplasm for further modifications and incorporation into storage triacylglycerols (Stymne and Stobart, 1987). Acyl-ACP thioesterases play an essential role in chain termination during de *novo* fatty acid synthesis and in the channelling of carbon flux between two lipid biosynthetic pathways in plants. A number of thioesterase cDNAs and genes have been isolated and sequenced from various plants (Topfer and Martini, 1995) including elm (*Ulmus americana*) (Voelker *et al.,* 1997) and nutmeg (*Myristica fragrans*) (Voelker *et al.,* 1997). The acyl-ACP thioesterases found

in seeds can be separated into two distinct but related classes known as FatA and FatB (Jones et al., 1995). FatA thioesterases are known to be specific for 18:1 substrates, The FatB enzymes tend to be specific for saturated substrates, with some acting on C14-C18 substrates and others on medium chain fatty acids. Little is known about fatty acid biosynthesis in oil palm, which is the second most important source of plant oils. In this paper, we report the isolation of a cDNA encoding a FatB thioesterase from the developing mesocarp of oil palm.

MATERIALS AND METHODS

Production of Thioesterase Gene Probes

An amplified oil palm cDNA library in λZAP II (Stratagene, La Jolla, CA) prepared from 15-week mesocarp was kindly provided by Siti Nor Akmar Abdullah (PORIM, Malaysia). The polymerase chain reaction (PCR) was used to amplify thioesterase gene fragments from the cDNA library. Initial amplification involving degenerate oligonucleotide primers based on conserved regions of known thioesterase genes generated a 130bp fragment of a FatB type acyl-ACP thioesterase (Abrizah et al., 1995). Probes for screening libraries were then obtained by high stringency PCR using gene-specific primers (ACO-1 and ACO-2) and flanking vector primers (T3 and T7). ACO-1 (5'-TTGGCTGGATCCTTGAGAGTG-3') was used in conjunction with the standard vector primer T7 to generate 3' fragments of thioesterase cDNAs and a 750bp fragment was generated. ACO-2, the complement of ACO-1, was used in conjunction with T3 to generate 5' fragments. PCR (25µl) was with 0.3µM primers in 10mM Tris-HCl, pH 9.0, 50mM KCl, 1.5mM $MgCl_2$ and 0.2mM each deoxynucleotide triphosphate, with 1µl (approximately 25 000pfu) cDNA library as template. The PCR reactions were carried out with 1.0 unit of Taq DNA Polymerase (Pharmacia) using an amplification programme of 3min denaturation at 94°C, followed by 32 cycles of 30sec at 94°C, 30sec at 55°C, 2min at 72°C, and terminated by 6min extension at 72°C. PCR products were fractionated on an agarose gel and purified using a High Pure PCR Product Purification Kit (Boehringer Mannheim) and labelled with [^{32}P]dCTP using a random primed labelling kit according to the instructions of the supplier (Pharmacia).

Screening the cDNA Library

Approximately 100 000pfu were mixed with *E. coli* KW 251 plating cells and plated on 24 x 24cm plates. The plaques were transferred to nitrocellulose filters in duplicate and the filters were prehybridized for 15min at 65°C in 7% SDS, 0.25M sodium phosphate buffer, pH 7.2. Subsequent hybridization was carried out for 16hr under the same conditions; duplicate filters were hybridized to the 5' or 3' probes. The membranes were washed under low-stringency conditions: twice for 15min at 65°C in 2x SSC, 0.1% SDS. The membranes were exposed to Hyperfilm™-MP x-ray film overnight. Plaques giving rise to signals on duplicate filters were removed and subjected to further rounds of plating and hybridization until pure. After plaque purification, the phagemids of positive plaques were excised in *vivo* with the ExAssist Helper Phage (Stratagene) and *E. coli* XLOLR. The cDNA was sequenced completely on both strands using an automated ABI 377 sequencer (Applied Biosystems).

RESULTS AND DISCUSSION

Isolation of Thioesterase cDNA Clones from Oil Palm Mesocarp

Initially, degenerate primers were synthesized to conserved regions of thioesterase genes. Using the cDNA library as template, a 130bp fragment was amplified, cloned and sequenced. The sequence had 64% homology with a FatB thioesterase from *Cuphea hookeriana* (Abrizah et al., 1995). Complementary gene-specific oligonucleotides (ACO-1 and ACO-2) were synthesized to a sequence within the cloned 130bp fragment, These were used in combination with the common vector primers, T3 and T7, to amplify 5' (ACO-2 and T3) and 3' (ACO-1 and T7) fragments of a thioesterase cDNA from the library (Figure I). PCR fragments of 800bp and

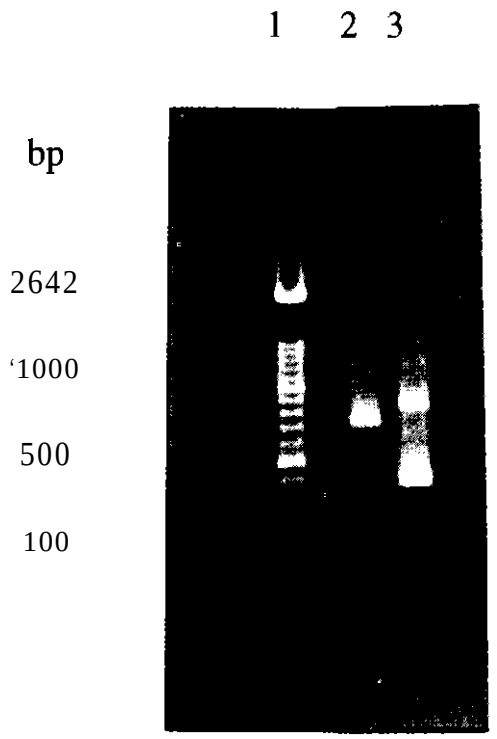

Figure 1. PCR amplification of 5' and 3' fragments of a thioesterase cDNA from the 15-week library. Lane I is a 100bp ladder. Lane 2 shows a 750bp fragment using primers ACO-1 and T7 and lane 3 shows the 8OOobp and 400bp fragments amplified with primers ACO-2 and T3.

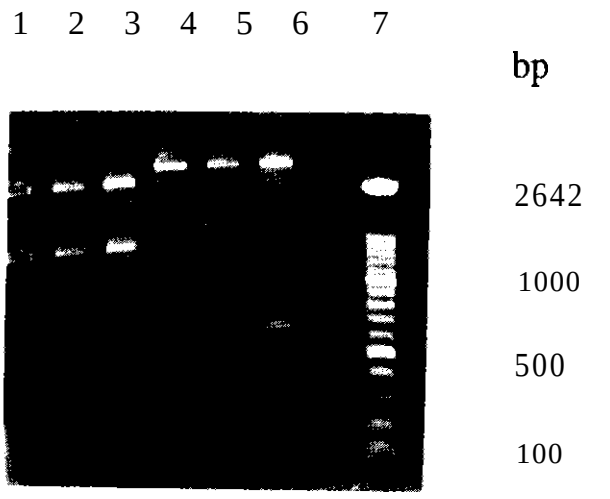

Figure 2. Restriction digestions of the clones with* EcoRI *and* XhoI*: Lane I pHA-1; Lane 2, pHA-2; Lane 3, pHA-3. Lanes 4, 5 and 6 show the digestion of the clones with* EcoRI *and* HindIII*: Lane 4, pHA-I; Lane 5, pHA-2; Lane 6, pHA-3. Lane 7 is a 100bp ladder.

400bp were generated when using ACO-2 and T3, whereas only a single PCR product of 750bp was obtained with ACO-1 and T7. Observation of two PCR products indicated that there were two clones in the library that extended as far as the ACO-l/2 region, and that there were no clones present with inserts longer than 1500bp. The 750bp and 800bp fragments were used to probe the 15-week mesocarp cDNA library.

Approximately 100 000 phage plaques were screened by blotting onto duplicate filters and hybridization to 5' and 3' thioesterase PCR products. Three plaques hybridized to both probes. They were purified by further rounds of screening and converted to plasmids in *E. coli* DH5α. Digestion with *Xho*I and EcoRI showed that they all contained inserts of approximately 1450bp. Digestion with EcoRI and. *Hind*III revealed the presence of a common *Hind*III site approximately 600bp from the 5' end of the cDNA (*Figure* 2). This showed them to be derived by amplification from a single original cDNA clone. This insert was fully sequenced on both strands. The sequence showed the cDNA to be incomplete, extending from within the coding region to the poly(A) tail, with the 5' end of the insert located within the transit peptide sequence. The 3'-untranslated region is 360bp long. The 1002bp of coding region sequence (including the stop codon) are presented in Figure 3, together with the deduced amino acid sequence. Comparison of the deduced amino acid sequences of pHA-3 and *Cuphea hookeriana*, *Ch* FatB1 (Jones *et* al., 1995) indicates that the mature proteins are 66% identical (Figure *4*), whereas there is only 35% amino acid identity to the *Garm* FatA1 sequence of *Garcinia mangostana* (Hawkins and Kridl, 1998). In all probability, therefore, pHA-3 encodes an oil palm FatB thioesterase.

Assignment of Amino Terminus of the Mature Protein

The transit peptide cleavage site of FatB thioesterases is not known, although it has been suggested to be near to the beginning of a highly conserved hydrophobic region unique to the FatB sequences (Jones *et* al., 1995). In

```
ttctataatcaattgcctgactggagcgtgcttcttgccgccgtaacgaccatcttcttggcggcggaga      70
F  Y  N  Q  L  P  D  W  S  V  L  L  A  A  V  T  T  I  F  L  A  A  E  K
                  ↑
agcagtggacccttcttgattggaagccgaagcgtcccgacatgcttgctgatgcatttggcctggggaa     140
   Q  W  T  L  L  D  W  K  P  R  R  P  D  M  L  A  D  A  F  G  L  G  K

aatcgtgcaggatggactagtttttaagcaaaacttttccatcaggtcgtatgagatcggggctgatcgg     210
 I  V  Q  D  G  L  V  F  K  Q  N  F  S  I  R  S  Y  E  I   GAD   R

actgcttctatagaaacgctaatgaatcatttacaggaaacagcacttaatcatgtgaggagtgctgggc     280
T  A  S  I  E  T  L  M  N  H  L  Q  E  T  A  L  N  H  V  R  S  A  G  L

tcatggggcgatggctttggtgctacaccagagatgagcaaaagaaatttgatctgggttgtcaccaaaat     350
   M  G  D  G  F  G  A  T  P  E  M  S  K  RN  L  I  W  V  V  T  K  M

gcgggttctgatcaagcactatccttcctgggggatgttgttgaagtagatacttgggttggtccaact     420
 R  V  L  I  K  H  Y  P  S  W  G  D  V  V  E  V  D  T  W  V  G  P  T

ggaaaaaatgggatgcgtcgtgattggcatgttcgtgaccaccgaacaggccaaaccatcttgagagcta     490
G  K  N  G  M  R  R  D  W  H  V  R  D  H  R  T  G  Q  T  I  L  R  A  T

ccagtgtgtgggtgatgatgaataagaacactaggaaattgtctaaagtccctgaagaagtcagggcaga     560
   S  V  W  V  M  M  N  K  NT  R  K  L  S  K  V  P  E  E  V  R  A  E

attaggcccttactttgtggaacgtgctgcaattgtggatgaggacagcagaaagctcccaaagcttgat     630
 I  G  P  Y  F  V  E  R  A  A  I  V  D  E  D  S  R  K  L  P  K  L  D

gaggatactacagattatatcaaaaagggcctaactcctcgatggagcgatttagatgtcaatcagcatg     700
E  D  T  T  D  Y  I  K  KG  L  T  P  R  W  S  D  L  D  V  N  Q  H  V

tgaacaatgtcaaatatattggctggattcttgagagtgctccaatatcattcctggagaatcatgagct     770
   N  N  V  KY  I  GW  I  L  E  S  A  P  I  S  F  L  E  N  H  E  L

tgcaagcatgtctctggaatataggagggagtgtgggagggacagcgtgttgcaatccctcactgccgtc     840
 A  S  M  S  L  E  Y  R  R  E  C  G  R  D  S  V  L  Q  S  L  T  A  V

tcgaatgacttaactgatggcttaccagaagctggcattgagtgccagcatctgctgcagctggaatgtg     910
S  N  D  L  T  D  G  L  P  E  A  G  I  E  C  Q  H  L  L  Q  L  E  C  G

ggaccgaacttgtgaagggacggacagaatggaggcccaagcattccctggctctcagaaacatggggcc     980
   T  E  L  V  K  G  R  T  E  W  R  P  K  H  S  L  A  L  R  N  M  G  P

aactccaggtggtagtgcatga                                                    1002
 T  P  G  G  S  A  *
```

Figure 3. Nucleotide and deduced amino acid sequence of the coding region of cDNA pHA-3 coding for an acyl-ACP thioesterase from the oil palm Elaeis guineensis. *The derived amino acid sequence is presented in single-letter code. The presumptive NH_2-terminus of the mature protein is marked by an arrow. The stop codon is denoted by*.*

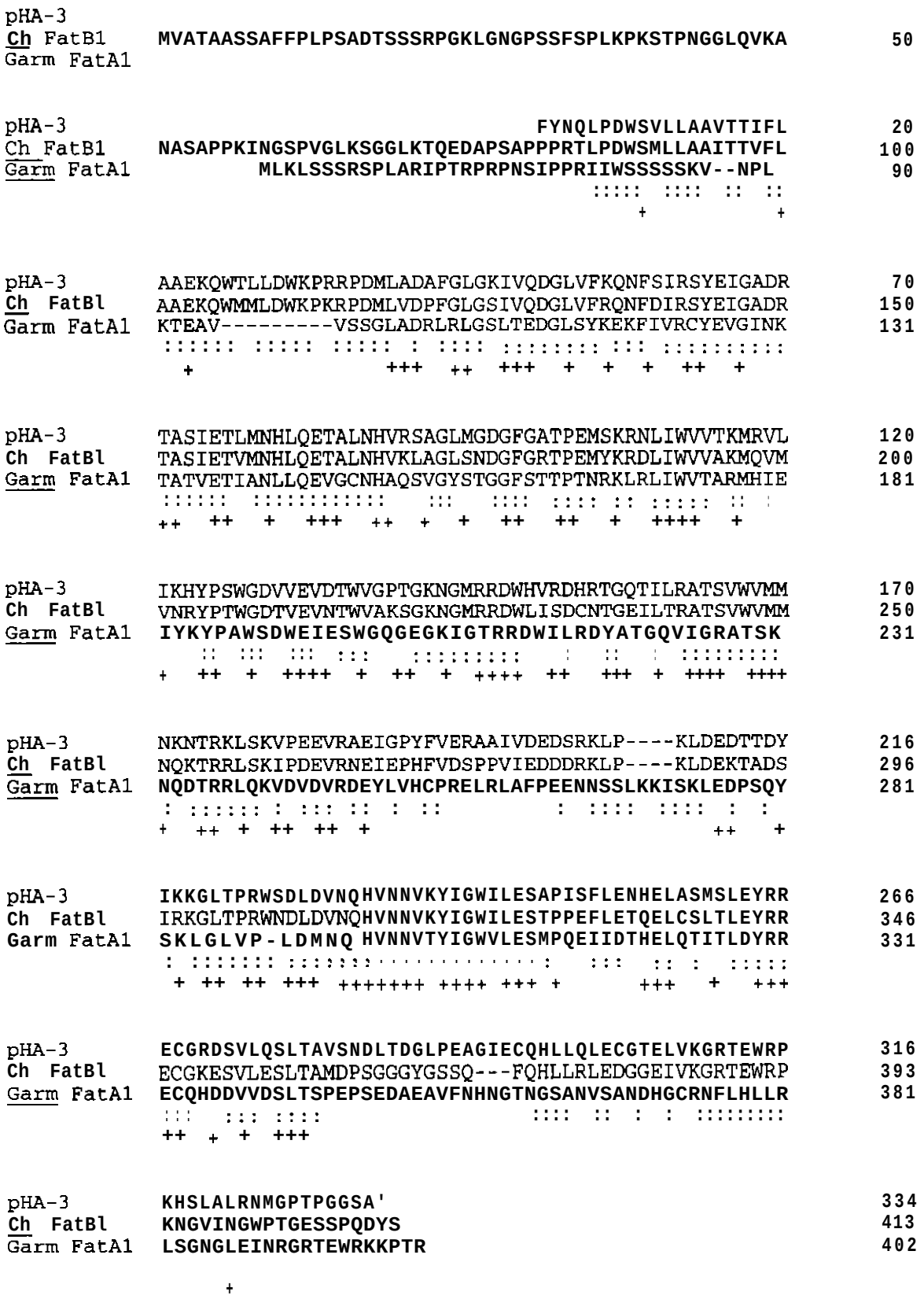

Figure 4. ***Comparison of the deduced amino acid sequences of the oil palm (pHA-3),*** Cuphea hookeriana ***(Ch FatB1) and*** Garcinia mangostana ***(Garm FatA1) acyl-ACP thioesterases. Identical amino acids between oil palm and*** Cuphea hookeriana ***are indicated by a colon; dashes indicate gaps introduced to maximize the alignment and plus signs indicate identical amino acids between pHA-3 and*** Garm ***FatA1 acyl-ACP thioesterases.***

```
pHA-3                          FYNQLPDWSVLLAAVTTIFLAAEKQWTL---   28
Ch FatB1  ----KTQEDAPSAPPPRTLPDWSMLLAAITTVFLAAEKQWMM---   38
At FatB1  ----ETSSHPAPRTFINQLPDWSMLLAAITTIFLAAEKQWMM---   38
```

Figure 5. Alignment of the deduced amino acid sequences of the oil palm (pHA-3), C. hookeriana (Ch *FatB1) and A.* thaliana (At *FatB1) acyl-ACP thioesterases. • Beginning of the FatB homology region - the proposed amino terminus of the mature protein.*

order to identify a putative transit peptide cleavage site, Dormann *et* al. (1995) compared the *Arabidopsis thaliana* FatB cDNA sequence with that of the Californian bay (Voelker *et* al., 1992) and with conserved amino acids in transit peptides (Gavel and von Heijic, 1990). They concluded that the leucine residue at amino acid *88* of the *A. thaliana* sequence was a putative candidate for the N-terminus of the mature protein. The equivalent position in the incomplete cDNA pHA-3 is the fifth residue in the sequence FYNQLPDWSVLL, which marks the beginning of continuous close homology between the *E. guineensis, A. thaliana* and C. *hookeriana* FatB sequences (Figure *5*). The leucine at position 5 of the deduced amino acid sequence of pHA-3 is therefore the likely beginning of the mature protein. Although the results show that the pHA-3 sequence was incomplete and lacking most of the transit peptide, the region encoding the full mature protein can be expressed in *E. coli* for further characterization and to assess substrate specificity.

Expression of plant acyl-ACP thioesterases in *E. coli* with subsequent *in vitro* assay has been successful in the assessment of enzyme substrate specificity (Dehesh *et* al., 1996; Knutzon *et* al., 1992; Voelker and Davies, 1994; Hawkins and Kridl, 1998). Functional expression to determine oil palm acyl-ACP thioesterase specificity is important because these enzymes largely determine acyl chain length of the fatty acids that are transported to the endoplasmic reticulum for oil assembly. Understanding the role of acyl-ACP thioesterases in fatty acid biosynthesis is therefore relevant to the modification of oil palm lipid composition. The isolation and characterization of a FatB thioesterase is an important first step in this direction.

CONCLUSION

An acyl-ACP thioesterase has been cloned from an oil palm mesocarp cDNA library. The deduced amino acid sequence indicates that it belongs to the FatB class of acyl-ACP thioesterases and is 66% identical to a 16:0-ACP thioesterase from *Cuphea hookeriana.* Work is in progress to express the cDNA sequence corresponding to the mature protein in *E. coli,* in order to evaluate its acyl specificity.

ACKNOWLEDGEMENTS

Abrizah Othman would like to thank the Director-General of PORIM for permission to publish this work and to acknowledge financial support from PORIM. The authors also thank Siti Nor Akmar Abdullah (PORIM) for the gift of the 15-week oil palm mesocarp cDNA library and Zaini Abdullah for technical assistance.

REFERENCES

ABRIZAH, 0; CHEAH, S C; RAVIGADEVI, S; MEKHEDOV, S and OHLROGGE, J B (1995). Isolation of oil palm acyl-acyl carrier protein thioesterase gene using polymerase chain reaction (PCR). *Proc. of the 20th Annual Conference of the Malaysian Society for Biochemistry and Molecular Biology,* Seremban, Negeri Sembilan, 19-20 September, 1995.

DEHESH, K; EDWARDS, P; HAYES, T; CRANMER, AM and FILLATTI, J (1996). Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. *Plant Physiol., 110*: 203-210.

DORMANN, P; VOELKER, T A and OHLROGGE, J B (1995). Cloning and expression in ***Escherichia coli*** of a novel thioesterase ***from Arabidopsis thaliana*** specific for long-chain acyl-acyl carrier protein. ***Arch. Biochem. Biophys.,*** *316:*612-618.

GAVEL, Y and von HEIJIC, G (1990). A conserved cleavage-site motif in chloroplast transit peptides. *FEBS Lett., 261:*455-458.

HAWKINS, D J and KRIDL, J C (1998). Characterization of acyl-ACP thioesterases of mangosteen (*Garcinia mangostana*) seed and high levels of stearate production in transgenic canola. ***The Plant Journal,*** *13:*743-752.

JONES, A, DAVIES, H M and VOELKER, T A (1995). Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases. ***The Plant Cell,*** 7: 357-371.

KNUTZON D, S; BLEIBAUM, J L; NELSEN, J; KRIDL, J C and THOMPSON, G A (1992). Isolation and characterization of two safflower oleyl-acyl carrier protein thioesterase cDNA clones. ***Plant Physiol.,*** *100:*1751-1758.

STYMNE, S and STOBART, A K (1987). Triacylglycerol biosynthesis. In (ed. Stumpf, P K). ***The Biochemistry of Plants.*** Vol. 9, Academic Press, Orlando, Florida. p. 175-214.

TOPFER, Rand MARTINI, N (1995). Molecular cloning of cDNAs or genes encoding proteins involved in de *novo* fatty acid biosynthesis in plants. ***Journal of Plant Physiology,*** 143:416-425.

VOELKER, T A; WORRELL, AC; ANDERSON, L; BLEIBAUM, J; FAN, C; HAWKINS, D H; RADKE, S E and DAVIES, H M (1992). Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science, *257:*72-74.

VOELKER, T A and DAVIES, H M (1994). Alteration of the specificity and regulation of fatty acid synthesis of ***Escherichia coli*** by expression of a plant medium-chain awl-awl carrier protein thioesterase. ***J. Bacteriol., 176:*** 7320-7327.

VOELKER, T A, JONES, A, CRANMER, A M; DAVIES, H M and KNUTZON, D S (1997). Broad range and binary-range acyl-acyl carrier protein thioesterases suggest an alternative mechanism for medium-chain production in seeds. ***Plant Physiol.,*** *114:*669-677.