AN ABSTRACT OF THE THESIS OF

Jeffrey M. Leonard for the degree of Doctor of Philosophy in Botany and Plant Pathology presented on December 16, 1996. Title: Acyl-ACP Thioesterases and Beta-Ketoacyl-ACP Synthase Regulate Fatty Acid Chain Length in Seed Oils of *Cuphea wrightii*

Abstract approved: Redacted for Privacy

Steven J. Knapp

Cuphea wrightii seed oils contain 30% caprate (10:0) and 54% laurate (12:0), whereas most seed oils contain 16 or 18 carbon fatty acids. The objectives of this thesis were to isolate and characterize the genetic determinants of this unusual phenotype. During fatty acid synthesis, the acyl-ACP thioesterase-catalyzed hydrolysis of the thioester bond linking the acyl chain and ACP may regulate acyl chain length. Two cDNAs, encoding thioesterases *Cw* FatB1 and *Cw* FatB2, were isolated from *C. wrightii* embryos and expressed in *Arabidopsis*. *Cw FatB2* seeds produced 10:0, 12:0 and 14:0, fatty acids not found in wildtype *Arabidopsis*. Homozygous *Cw FatB1* seeds produced 12:0 and 14:0. Both transgenic lines produced ~2.5-fold more 16:0 than wildtype and decreased levels of unsaturated fatty acids. Because 10:0 and 12:0-ACP specific thioesterases were expected, the high levels of 14:0- and 16:0-ACP activity were puzzling. The possibility that gene dosage effects might shift distribution towards 10:0 and 12:0 was considered but not observed when hemizygous and homozygous phenotypes were compared. Thus *Cw FatB1* and *Cw FatB2* may be necessary but insufficient determinants of the *C. wrightii* phenotype. β-ketoacyl-ACP synthases (KAS) may also regulate fatty acid chain length. A

C. wrightii cDNA encoding KASII homologue *Cw* Kas21 was isolated and expressed in *Arabidopsis*. The wildtype 16:0 content of 8.2 mol% dropped to 6.2 mol% in transgenic *Arabidopsis*. In the presence of the KASI inhibitor cerulenin, transgenic seed extracts extended 6:0- and 8:0-ACP *in vitro* indicating medium-chain activity; cerulenin reduced this activity in wildtype seed extracts. Seeds homozygous for *Cw Kas21* and *Cw FatB1* produced 3-fold more 12:0 than the *FatB1* parent with simultaneous decreases in 14:0 and 16:0. Seeds expressing *Cw Kas21* and *Cw FatB2* produced significantly more 10:0 and 12:0 than the *Cw FatB2* parent while 14:0 and 16:0 accumulation declined. We hypothesize that the dramatic shift towards shorter chains in double transgenics resulted from increased pools of medium-chain acyl-ACPs produced by *Cw* KAS21 activity. The combination of *Cw* Kas21 and *Cw* FatB thioesterases appears to determine the *C. wrightii* phenotype. This synergistic effect can be exploited in genetic engineering of oilseeds.

Acyl-ACP Thioesterases and Beta-Ketoacyl-ACP Synthase Regulate Fatty Acid Chain Length in Seed Oils of *Cuphea wrightii*

by

Jeffrey M. Leonard

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirement for the
degree of

Doctor of Philosophy

Presented December 16, 1996
Commencement June 1997

Doctor of Philosophy thesis of Jeffrey M. Leonard presented on December 16, 1996

APPROVED:

Redacted for Privacy

Major Professor, representing Botany and Plant Pathology

Redacted for Privacy

Chair of Department of Botany and Plant Pathology

Redacted for Privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Jeffrey M. Leonard, Author

ACKNOWLEDGEMENTS

I offer my sincerest thanks to Dr. Steven J. Knapp. His wide ranging knowledge generated a constant flow of ideas that broadened my own perspectives far beyond the material in this thesis. His continual encouragement to pursue my interests and to develop a professional career spurred my curiosity and fostered a modest level of self confidence. I am forever indebted to Dr. Mary B. Slabaugh, who taught me the techniques of molecular biology and how to apply those skills to questions. Her steadfast encouragement, direction, respect, and prodding are certainly as responsible as my own efforts for what success I have had in my research. In countless discussions (generally with good humor), Steve and Mary have guided my education and taught me how to do science.

I wish also to thank the members of my committee, who have offered encouragement and advice regarding both science and my professional development. Jimmie Crane and Daryl Ehrensing kept the GC running. I was always able to find space in a greenhouse or a growth chamber to propagate plants by asking either Jimmie or Daryl.

To my wife Sandy and son Ian, I owe more than can be expressed in words. They continually remind me of which things in life are important.

CONTRIBUTION OF AUTHORS

Dr. Steven J. Knapp proposed the subject of this research and contributed substantially to all genetic and statistical analyses. Dr. Mary B. Slabaugh, made the *C. wrightii* cDNA library, isolated *Cw Kas21*, constructed the binary vector for transgenic expression of the clone, and performed the *in vitro* assays used to characterize the extension capabilities of *Cw* KAS21. Both Drs. Knapp and Slabaugh assisted in the interpretation of the data and made many revisions and suggestions to the manuscripts.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

ACYL-ACP THIOESTERASES AND BETA-KETOACYL-ACP SYNTHASE REGULATE FATTY ACID CHAIN LENGTH IN SEED OILS OF *CUPHEA WRIGHTII*

Chapter 1

INTRODUCTION

Progress in plant transformation techniques and plant molecular biology opened the doors for production of genetically-engineered crops. In addition to efforts towards engineering useful agronomic traits such as herbicide or insect resistance, there is also interest in manipulating biosynthetic pathways. Plants represent a tremendous genetic resource of enzymes that catalyze the synthesis of over 20,000 different terpenoids, flavonoids, alkaloids, and fatty acids (Ohlrogge, 1994). Commercial utilization of this germplasm has been targeted towards production of novel compounds and alterations of the relative amounts of naturally occurring substances. Fatty acid metabolism is a particularly valuable target for genetic engineering because of the potential usage of plant oils as surfactants, plasticizers, and lubricants in industrial applications and the production of edible vegetable oils with modified nutritional values or consistencies.

Current world production of vegetable oils is ~65 million tons per year (Murphy, 1994) with an approximate market value of ~20 billion dollars (Battey *et al.*, 1989). Approximately one third of these oils are used for industrial, rather than food applications (Ohlrogge, 1994). The unique physical and chemical qualities of plant oils are derived from their constituent fatty acids (Figure 1.1). Although more than 500 different fatty acids have been identified in plants (Stobart *et al.*, 1993), most plants synthesize only a

very limited subset which are stored in seed oils (Table 1.1) Because of this phenomenon, just six fatty acids account for more than 95% of world oil production (Schmid, 1987). Nonetheless, some uncommon fatty acids such as laurate (recovered from palm kernel), erucic acid (recovered from rapeseed), and ricinoleic acid (recovered from castor) are commercially exploited. (Figure 1.1) Other fatty acids have potential uses but are presently unavailable in commercial quantities (Battey *et al.*, 1989), because the plants which produce them are not viable crop species.

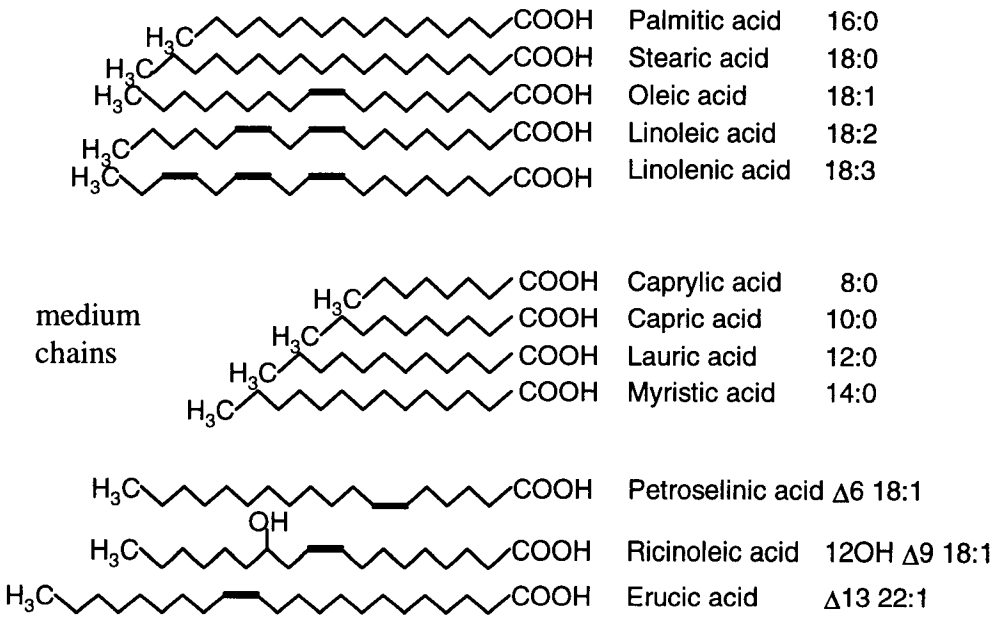

Figure 1.1. The structure and nomenclature of some common fatty acids. The trivial names have often been derived from a commmon plant source; eg., palmitic acid. The first number in the numerical nomenclature refers to the number of carbons in the chain and the second number refers to the number of double bonds. Fatty acids with 8 to 14 carbons are termed medium-chain.

The potential for genetically engineering high market value oil seed crops positively impacted research of plant fatty acid metabolism over the past decade. A relevant example is laurate (12:0), which is produced from palm kernel oil and used as an ingredient in many soaps, detergents, and shampoos. The U.S. imports approximately 640,000 tons per year (Battey *et al.*, 1989) with an estimated market value of 350 million dollars (Ohlrogge, 1994). Desire to produce a genetically-engineered domestic source of laurate motivated research into mechanisms that regulate acyl chain length in plants that produce medium-chain (C8 to C14) fatty acids

Table 1.1. Fatty acid composition of common vegetable oils.
Data represents percentage of each fatty acid by weight. The oils from the six species listed (there are two varieties of rapeseed) account for over 85% of total commercial oil production.

	8:0	10:0	12:0	14:0	16:0	18:0	18:1	18:2	18:3
soybean	-	-	-	-	9	2	32	53	3
palm	-	-	-	2	42	4	42	10	--
rapeseed	-	-	-	-	3	1	24	15	8
rapeseed	-	-	-	-	4	2	60	20	10
sunflower	-	-	-	-	-5	2	35	57	-
cottonseed	-	-	-	-	21	2	25	50	-
peanut					12	4	47	31	-

The initial reactions of fatty acid synthesis are catalyzed by a series of dissociable enzymes within plant plastids (Figure 1.2) (reviewed by Browse and Sommerville, 1991; Ohlrogge and Browse, 1995). Because enzymatic activities of fatty acid synthesis are separable (Shimikata and Stumpf, 1982), plant fatty acid synthase (FAS) is similar to the prokaryotic Type II FAS and not the multienzyme Type I complex found in yeast and

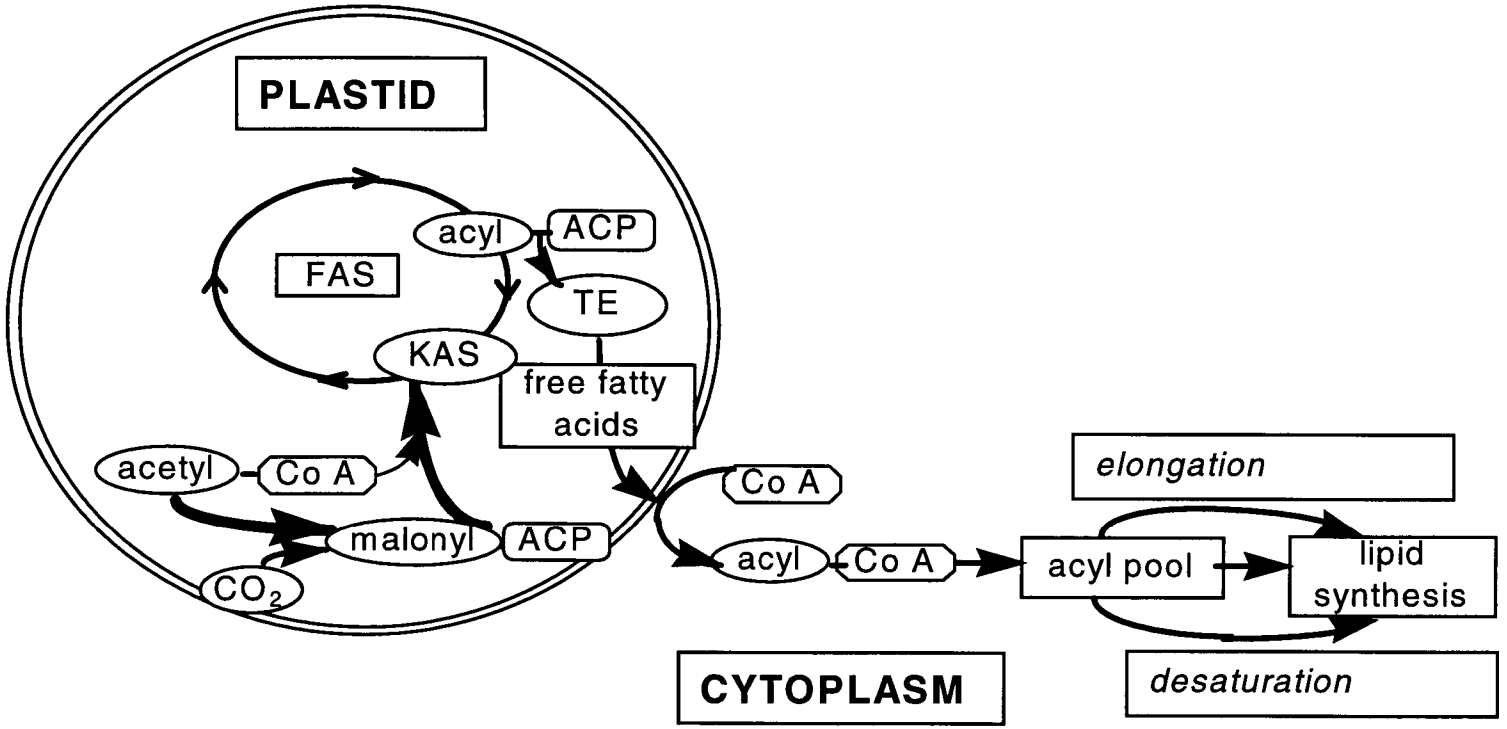

Figure 1.2. Diagram of fatty acid synthesis in plants. The initial stages of fatty acid synthesis in plants occur within the plastid. During elongation, the acyl chain is esterified to an acyl-carrier protein (ACP). Condensation of two carbon units donated from malonyl-ACP is catalyzed by β-ketoacyl-ACP synthase (KAS). A cyclic pattern of reduction, dehydration, and reduction is catalyzed by three separate enzymes of fatty acid synthase (FAS). In most plants, the acyl chain is elongated to 16 or 18 carbons. A soluble acyl-ACP thioesterase (TE) hydrolyzes the acyl-ACP thioester bond releasing free fatty acids. The free fatty acids are reesterified to Co A on the surface of the plastid membrane. The acyl-Co A molecules may be further modified prior to assembly into triglycerides

animals. A β-ketoacyl-acyl carrier protein synthase (KAS) catalyzes the sequential addition of 2 carbon units from malonyl-acyl carrier protein (ACP) to an acyl chain which is also esterified to ACP. The resultant β-ketoacyl-ACP molecule undergoes a cycle of reduction, dehydration, and reduction prior to additional condensation reactions. Most fatty acyl-ACP chains are elongated to 16 or 18 carbons in length. Within most plants, the majority of the stearoyl-ACP (18:0)-ACP is desaturated to oleoyl-ACP (18:1) by a soluble stearoyl desaturase. Because palmitoyl-ACP (16:0-ACP) is not a substrate for stearoyl desaturase, 16:0 and 18:1 are the major products of plastidal fatty acid synthesis.

Within the plastid, acyl groups may be transferred from ACP to glycerol-3-phosphate or to monoacylglycerol-3-phosphate, which are then incorporated into plastid membranes via the "prokaryotic" pathway. Alternatively, fatty acids may enter the "eukaryotic" pathway if hydrolyzed from ACP by an acyl-ACP thioesterase (TE). Free fatty acids exit the plastid by an unknown mechanism and are esterified to CoA on the outer membrane of the chloroplast envelope. The resultant acyl-CoA may be further modified in the ER and assembled into phospholipids or triacylglycerols.

As portions of the plant biosynthetic pathway were elucidated, several mechanisms were proposed by which a plant might produce medium-chain fatty acids (Harwood, 1989), including the possibility of a substrate-specific TE which would prematurely terminate elongation of an acyl-ACP chain (Stumpf, 1987). This unusual substrate preference was presumed to form an enriched pool of medium-chain free fatty acids that would subsequently be incorporated into triglycerides. An alternative model invoking the action of a substrate specific condensing enzyme (KAS) was also proposed

(Harwood, 1988). These proposals were speculative as there were limited data to support either.

Pollard *et al.* (1991) researched California bay (*Umbellularia californica*), a species that accumulates primarily 10:0 and 12:0 in seed storage lipids, and succeeded in partially purifying a 12:0-ACP-specific thioesterase from embryo extracts. Their work was the first direct evidence of a mechanism which might produce medium-chain fatty acids in plants. A 12:0-ACP TE was subsequently purified and characterized as a monomer with a molecular weight of 34 kDa (Davies *et al.*, 1991). Purification and sequencing of protein fragments allowed design of degenerate oligonucleotides and cloning of a cDNA encoding the 12:0-specific TE (Voelker *et al.*, 1992). When the TE was expressed in *Arabidopsis*, seeds accumulated a significant amount of 12:0.

The demonstration of the power of a single enzyme to divert carbon flux within the fatty acid pathway and produce a novel phenotype immediately led to a search for similar enzymes from other species; however, it was unknown if this mechanism was universal to medium-chain-producing species or unique to California bay. Davies (1993) assayed TE activity in seed extracts from elm (*Ulmus americana*), palm (*Cocos nucifera*) and camphor (*Cinnamomum camphora*); each of these medium-chain species represented separate plant families. All three extracts were found to have hydrolytic activities on various medium-chains acyl-ACPs, although the range of substrates hydrolyzed was, in each case, somewhat broader than predicted from *in vivo* accumulation patterns. This survey suggested that medium-chain specific TEs were likely to be found in other plants with medium-chain phenotypes.

The new world genus *Cuphea* is unique in the number of its ~250 species that produce medium-chain seed oils (Table 1.2) (Graham *et al.*, 1981). The potential of *Cuphea* as a genetic resource for unusual enzymes has been widely recognized, and a number of species have been investigated. *C. wrightii* is an herbaceous annual, found largely in the highlands of Mexico (Graham, 1988). The species is an interesting model in chain-length regulation because it produces seed oils containing 30% caprate (10:0) and 54% laurate (12:0). (Graham *et al.*, 1981). We hypothesized that the unusual deposition patterns seen in seed oils of this plant reflected substrate specificities of acyl-ACP TEs, as had been demonstrated in California bay. The primary objectives of this research were to isolate TEs from *C. wrightii*, to characterize their function in fatty acid metabolism, and to determine if they regulated the *C. wrightii* phenotype.

Table 1.2. Fatty acid composition of selected *Cuphea* species; percentage by weight.

	8:0	10:0	12:0	14:0	16:0	18:0	18:1	18:2	18:3
C. lanceolata	-	83	2	2	3	-	3	5	-
C. denticulata	-	-	-	-	33		10	53	4
C. pulcherimma	94	3	-	-	1	-	-	1	-
C. wrightii	-	30	54	5	2	-	3	5	-

This thesis consists of two manuscripts as chapters and a third manuscript as an appendix. The first chapter describes isolation of two cDNAs encoding TEs from C. *wrightii* and characterization of their activities in transgenic *Arabidopsis*. The second chapter describes experiments that tested the interaction of a *C. wrightii* KAS with each of the two *C. wrightii* TEs in *Arabidopsis* expressing both classes of transgenes. Four

other cDNAs encoding TEs isolated from *C. wrightii* and a clustering analysis of *Cuphea* TEs are presented in the appendix.

Chapter 2

CUPHEA WRIGHTII THIOESTERASES HAVE UNEXPECTED BROAD SPECIFICITIES ON SATURATED FATTY ACIDS

Jeffrey M. Leonard, Mary B. Slabaugh, and Steven J. Knapp

ABSTRACT

Cuphea wrightii A. Gray is an herbaceous annual that accumulates 30% caprate (10:0) and 54% laurate (12:0) in seed storage lipids. We investigated the role of acyl-acyl carrier protein (ACP) thioesterases (TE) in acyl chain length regulation in *C. wrightii*. Two embryo-derived cDNAs, encoding the TEs *Cw* FatB1 and *Cw* FatB2, were isolated. Both proteins were detected in developing embryos (~41 kDa) and mature seeds (~33 kDa) but not in other tissues, suggesting involvement in seed oil synthesis. Although expected to be 10:0/12:0-ACP specific, both genes produced a broad range of fatty acids (12:0, 14:0, and 16:0) in transgenic *Arabidopsis* with the greatest accumulation at 14:0. *Cw FatB2* transformants also accumulated small amounts of 10:0. Because *C. wrightii* accumulates only ~5% 14:0 and ~2% 16:0, we tested the possibility that gene dosage effects might significantly alter the overall kinetics of the pathway. Phenotypic comparisons of progeny segregating for the transgenes individually and in a hybrid population demonstrated that increased enzyme pools *in vivo* had a minor effect on diverting fatty acid production to shorter chains. We propose that *Cw* FatB1 and *Cw* FatB2 may be necessary but not sufficient determinants of the *C. wrightii* phenotype.

INTRODUCTION

Plants synthesize fatty acids (FA) through a series of discrete enzymes localized to plastids (Browse and Sommerville, 1991). Acyl chains esterified to acyl carrier protein (ACP) are elongated by the sequential addition of two-carbon units donated by malonyl-ACP. Hydrolysis of the acyl-ACP thioester bond by an acyl-ACP thioesterase (TE)

terminates elongation and releases ACP and free fatty acids. These fatty acids cross the plastid membrane and may be further modified during assembly into membrane lipids or triacylglycerols in seeds. C_{16} and C_{18} fatty acids predominate in the seed storage lipids of most plants; however, a number of plant species are interesting for their unusual accumulation of medium-chain (C_8 - C_{14}) fatty acids in seeds. Biologists pursuing the engineering of seed oils in crop species are particularly interested in the mechanism that allows developing seeds of these species to switch from long-chain to medium-chain fatty acid production (Ohlrogge, 1994).

The simplest explanation for this switch is induction of a TE with specificity for medium-chain acyl-ACPs, such that acyl elongation terminates prematurely. Acyl transferases would then assemble triacylglycerols from the increased medium-chain fatty acid pool. Expression in *Arabidopsis* of a thioesterase isolated from California bay, a species that accumulates caprate (10:0) and laurate (12:0) in storage lipids, causes production of significant amounts of 12:0 in seed triglycerides (Voelker *et al.*, 1992). To date, medium-chain specific TEs have been cloned from three *Cuphea* species (Dehesh *et al.*, 1996a, Dehesh *et al.*, 1996b, Martini *et al.*, 1995) indicating a chain-length regulatory mechanism similar to California bay exists in *Cuphea*. Expression in canola of a *Cuphea hookeriana* TE, *Ch* FatB2, produces up to 11 mol% 8:0 and 27 mol% 10:0 in seeds (Dehesh *et al.*, 1996b). Although *C. hookeriana* seed oils contain twice as much 8:0 as 10:0, *Ch* FatB2 is probably a significant determinant of the phenotype. Two TEs have been cloned from *C. palustris* (Dehesh *et al.*, 1996a), a species that accumulates 20% 8:0 and 64% 14:0 in seeds. Assays done *in vitro* and in *E. coli* show *Cp* FatB1 has both 8:0- and 10:0-ACP activity, and *Cp* FatB2 has both 14:0- and 16:0-ACP activity. Although

the reported specificities are broader than suggested by the *C. palustris* phenotype, these enzymes are indicated as major factors regulating chain length in this species.

Many *Cuphea* species have seed oils rich in medium-chain fatty acids (Graham *et al.*, 1981) and represent a potential source of both economically important medium-chain fatty acids as well as a genetic resource for the engineering of seed oils in existing crops. Two TEs were partially purified from *C. wrightii* (Dörmann *et al.*, 1993), whose seeds contain 29% 10:0 and 54% 12:0. One TE is active on 18:1- ACP, while the other has broad medium-chain specificity. Interestingly, 10:0- and 12:0-ACP specific TEs were not found. Since *Cuphea* FatB TEs are members of a small gene family with at least four members (Jones *et al.*, 1995, Voelker, 1996), partial purification may have failed to resolve individual TEs with differing specificities. To explore this possibility, we isolated and characterized two FatB TEs from *C.wrightii.* We also investigated the possibility of an interaction between the two thioesterases by combining their activities in transgenic plants.

MATERIALS AND METHODS

Library Construction and Clone Isolation

Embryos were excised from immature seeds of greenhouse grown *Cuphea wrightii* A. Gray. Poly(A^+) RNA was purified from total RNA isolated from developing embryos by oligo-dT cellulose spin chromatography. cDNA was synthesized and an aliquot ligated into Lambda ZAPII (Stratagene) using the manufacturer's protocol.

Approximately 140,000 plaques were blotted onto nylon membranes. Degenerate oligonucleotides (*TE 1* and *TE 3*, Figure 2.1) were designed to regions conserved in the amino acid sequences of the California bay medium-chain TE (Voelker *et al.*, 1992) and two oleoyl-ACP TEs from safflower (Knutzon *et al.*, 1992). Deoxyinosine was included at all positions of four-fold degeneracy. The oligonucleotides primed a PCR amplification of embryo-derived cDNA producing a ~540-bp PCR product that was labeled with [α-^{32}P]dCTP using random hexamer primers. The membranes were hybridized at 42^0 in 50% formamide with a final wash at 68^0 in 1 x SSC and 0.1% SDS. Sixteen of the thirty four plaques that hybridized at high stringency were screened a second time with the same protocol. Phagemids excised from ten plaques isolated in the second screen were separated into three classes by restriction analyses. Three clones, cwTE11, cwTE48 , and cwTE55, were selected as representative of the classes and analyzed further. Sequencing was done on an Applied Biosystems autosequencer using vector-based primers and custom oligonucleotide primers. A suite of GCG computer programs (Genetics Computer Group Software, Madison, WI) were used for analysis of sequence data and multiple sequence alignment.

Production of Antibodies

A 560-bp *Nde*I-*Xho*I fragment from the 3′ end of cwTE11 was ligated into vector pET15B (Novagen, Madison WI) cut with the same enzymes, creating pET11-7. PCR was used to amplify a 615-bp fragment from cwTE11 beginning at the putative start of the mature protein (Figure 2.1) and to add a 5′ *Nde*I site. A 460-bp *Nde*I fragment from

the cloned PCR product was ligated into *Nde*I-linearized pET11-7. A clone with the correct orientation of the *Nde*I fragment (pTE11E1) was transformed into *E. coli* strain BL21(DE3) (Novagen).

A 5′ *Nde*I site was added during amplification of a 780 bp fragment from cwTE48, beginning near the putative start of the mature protein (Figure 2.1). After ligation into pBluescript SK II, a 640-bp *Nde*I-*Xho*I fragment was ligated into pET15B cut with the same enzymes, creating p48P. A 470-bp *Xho*I fragment from the 3′ end of cwTE48 was ligated into the *Xho*I site of the intermediate p48P. A clone with the *Xho*I insert in the correct orientation (pTE48E1) was transformed into *E. coli* strain BL21(DE3).

Cultures of BL21(DE3) harboring plasmids pTE48E1 and pTE11E1 were grown four hours at 30^0 C, induced with 0.1 mM isopropyl β-D-thiogalactopyranoside (IPTG) four additional hours at 30^0 C, pelleted by centrifugation, and stored at -80^0 C. Histidine-tagged recombinant proteins were partially purified using a nickel affinity resin (Ni-NTA, Quiagen) according to the manufacturer's protocol. Briefly, the frozen pellets were resuspended in Buffer A (6 M GuHCl, 0.1 M Na-Phosphate 0.01 M Tris/HCl (pH 8.0)), stirred at room temperature for one hour, then sonicated. Recombinant proteins were bound to the resin and washed per the protocol except the final wash, which included 30 mM imidazole. Proteins eluted with 250mM imidazole were further separated on a 10% SDS polyacrylamide gel. Gel fragments containing recombinant proteins were homogenized in complete Freund's adjuvant and used to raise antibodies in rabbits.

Western Analyses

Seed materials for the developmental study analyses were crushed under liquid N_2, then homogenized in 1:10 w/v of tissue extraction buffer (50 mM Tris-HCl (pH 7.5), 2 mM EDTA, 2% [w/v] SDS, 2 mM phenylmethysulfonyl fluoride (PMSF), 1 mg/l pepstatin, 1 mg/l leupeptin, 1 mM benzamide-HCl, and 1% [w/v] polyvinylpolypyrrolidone (PVPP). The homogenate was incubated for one hour on ice then centrifuged for 15′ at 4^0 C. 1.5 µl samples at each time point were electrophoresed on a 10% polyacrylamide gel, blotted, and immunostained with TE antisera (1:1000) using standard methods.

Transformation of *Arabidopsis*

PCR was used to amplify a ~1260 bp fragment of cwTE11 and ~1300 bp fragment of cwTE48, including 10 bp of the pBluescript vector at the 5′ end in both cases. 3′ primers were designed to create *Bgl*II sites 36 bp downstream of the stop codon of cwTE11 and 30 bp downstream of the stop codon of cwTE48. PCR products were cloned into pBluescript SKII cut with *Eco*RV and dT-tailed (Marchuk *et al.*, 1991). Clones were selected such that the insert was in the opposite orientation to the vector-encoded *lacZ* in order to add a 5′ *Sal*I site to the cDNAs. Digestion with *Sal*I/*Bgl*II released the entire coding regions of both genes that were subsequently inserted into *Sal*I/*Bgl*II-cut pCGN3223 (Kridl *et al.*, 1991) containing a seed-specific promoter and termination fragment from *Brasssica rapa*. *Kpn*I fragments containing the chimeric genes were cloned into the binary plant transformation vector pCGN1557 (McBride and Summerfelt,

1990). Vacuum infiltration (Bechtold *et al.*, 1993; Bent *et al.*, 1996) was used to transform *Arabidopsis thaliana* (ecotype Wassilewskija) with *Agrobacterium* strain LBA4404 harboring the plasmids.

T_1 transformed seeds (3-5,000 per plate) were screened on media containing 50 mg/l kanamycin while T_2 and T_3 seeds (1-200 per plate) were screened at 20 mg/l kanamycin. Fatty acid profiles for single T_2 seeds or pooled homozygous seeds were determined using gas chromatography according to Brandt and Knapp (1993).

RESULTS

Isolation and Characterization of two TE cDNA Clones from *C. wrightii*

Degenerate oligonucleotides were designed to two conserved regions of the medium-chain TE from California bay (Voelker *et al.*, 1992) and the oleoyl-ACP TE from safflower (Knutzon *et al.*, 1992) (Figure 2.1). A TE specific probe was PCR amplified using a template of cDNA derived from developing seeds of *C wrightii*. The probe was used to screen a cDNA library derived from the same tissue. Seven clones were isolated with inserts large enough to encode a TE, and were separated into three classes by restriction digests. Two clones, cwTE11 and cwTE55, representing two classes, were partially sequenced and found to be 96% identical for the first 250 predicted amino acids. These are most likely alleles or orthologs since *C. wrightii* is an allotetraploid (Graham 1988). cwTE11 was completely sequenced, but no further work

was done with cwTE55. A representative of the third class, cwTE48, was also completely sequenced.

The complete nucleotide sequence of cwTE11 is 1296 bp. Beginning at an initiation codon at position 20, an open reading frame of 1197 bp encodes a 398-amino acid protein with a predicted mol wt of 44,201. A 77 residue transit peptide and 36.3-kD mature protein were predicted based on the putative N terminus of the TE from *C. hookeriana* (Jones *et al.*, 1995). cwTE48 contains a 1,449-bp cDNA with a 1,227-bp open reading frame beginning at the initiation codon at position 23. The predicted 408 amino acid preprotein had a calculated mol wt of 45,843. The mature protein was estimated to have a mol wt of 36,451 after cleavage of an 89-amino acid transit peptide.

Deduced amino acid sequences of the preprotein and mature forms of cwTE11 and cwTE48 were 78% identical. The cDNAs were 84% identical overall and within the coding sequence. Predicted amino acid sequences of both clones were approximately 48% identical to the amino acid sequence of the California bay TE (Voelker *et al.*, 1992), but only ~38% identical to the safflower TEs (Knutzon *et al.*, 1992). Relative to the safflower TE, both cwTE11 and cwTE48 have deletions in the predicted mature proteins characteristic of the recently designated FatB group of TEs (Jones *et al.*, 1995). Therefore, cwTE11 was designated *Cw FatB1*, and cwTE48 was designated *Cw FatB2*.

We used the computer program Pileup (Genetics Computing Group, Madison, WI) to compare the deduced amino acid sequences of the *C. wrightii* TEs with five other *Cuphea* FatB TEs (Figure 2.1), including representatives of four clades of *Cuphea* FatB genes (Voelker 1996). Within the putative mature proteins, 171 residues (53%) were identical, and 73 of the remaining sites had conservative substitutions. The two active-

site motifs identified by Yuan *et al.* (1996) were completely conserved. *Cw* FatB1 clustered with the 14/16:0-ACP specific FatB2 from *C. palustris* (Dehesh *et al.*, 1996a). *Cw* FatB2 clustered with the 8:0/10:0-ACP specific FatB1 from *C. palustris* TE (Dehesh *et al.*, 1996a) and the 8:0/10:0-ACP specific FatB2 from *C. hookeriana* TE (Dehesh *et al.*, 1996b). Even with inclusion of the putative transit peptide sequences, the UPGMA-derived clustering order reflected the same topology as a phylogenetic tree derived by maximum parsimony analysis of homologous residues (Voelker, 1996). The clustering of *Cw* FatB1 and *Cw* FatB2 with medium-chain specific TEs and not the 16:0 specific *Ch.* FatB1 (Jones *et al.*, 1995), was circumstantial evidence for their involvement in storage lipid synthesis.

When three residues of the 12:0-ACP specific California bay TE are mutagenized (M197R/R199H/T231K), the enzyme becomes 14:0-ACP specific (Yuan *et al.*, 1995). However, the homologous residues (Met-211, Arg-213, and Lys-245 of *Cw* FatB1) were completely conserved in the *Cuphea* TEs (Figure 2.1), even though the *Cuphea* TEs exhibit a range of specificities. Therefore, substrate specificity is not determined solely by these residues in *Cuphea*. Interestingly, one of the California bay TE mutants (T231K) created by Yuan *et al.* (1995) is identical to the *Cuphea* TEs at these three sites but retains wildtype (12:0-ACP) activity. Unfortunately, without three-dimensional structures, it is impossible to distinguish mutations that alter substrate specificity from neutral changes that reflect only a common evolutionary history.

Figure 2.1 Alignment of the deduced amino acid sequences of *Cw* FatB1 and *Cw* FatB2 with TEs from *C. hookeriana* (Dehesh *et al.*, 1996b, Jones *et al.*, 1995), *C. lanceolata* (Toepfer *et al.*, 1995) and *C. palustris* (Dehesh *et al.*, 1996a). Dendogram on first line represents the UPGMA clustering order determined by the computer program PILEUP. Boxes around sequence names indicate orthologous genes (Voelker, 1996). Completely conserved residues are printed in reverse contrast. Horizontal arrows indicate location of degenerate primers *TE 1* and *TE 3*. Putative start of the mature FatB proteins is indicated by an arrowhead (Jones *et al.*, 1995). Starting points of the histidine-tagged fusion proteins of *Cw* FatB1 (Ala-95) and *Cw* FatB2 (Asp-115) are boxed. The active site motifs (Yuan *et al.*, 1996) are marked with asterisks. Filled circles indicate position of residues mutated in California bay TE to confer 14:0 specificity (Yuan *et al.*, 1995).

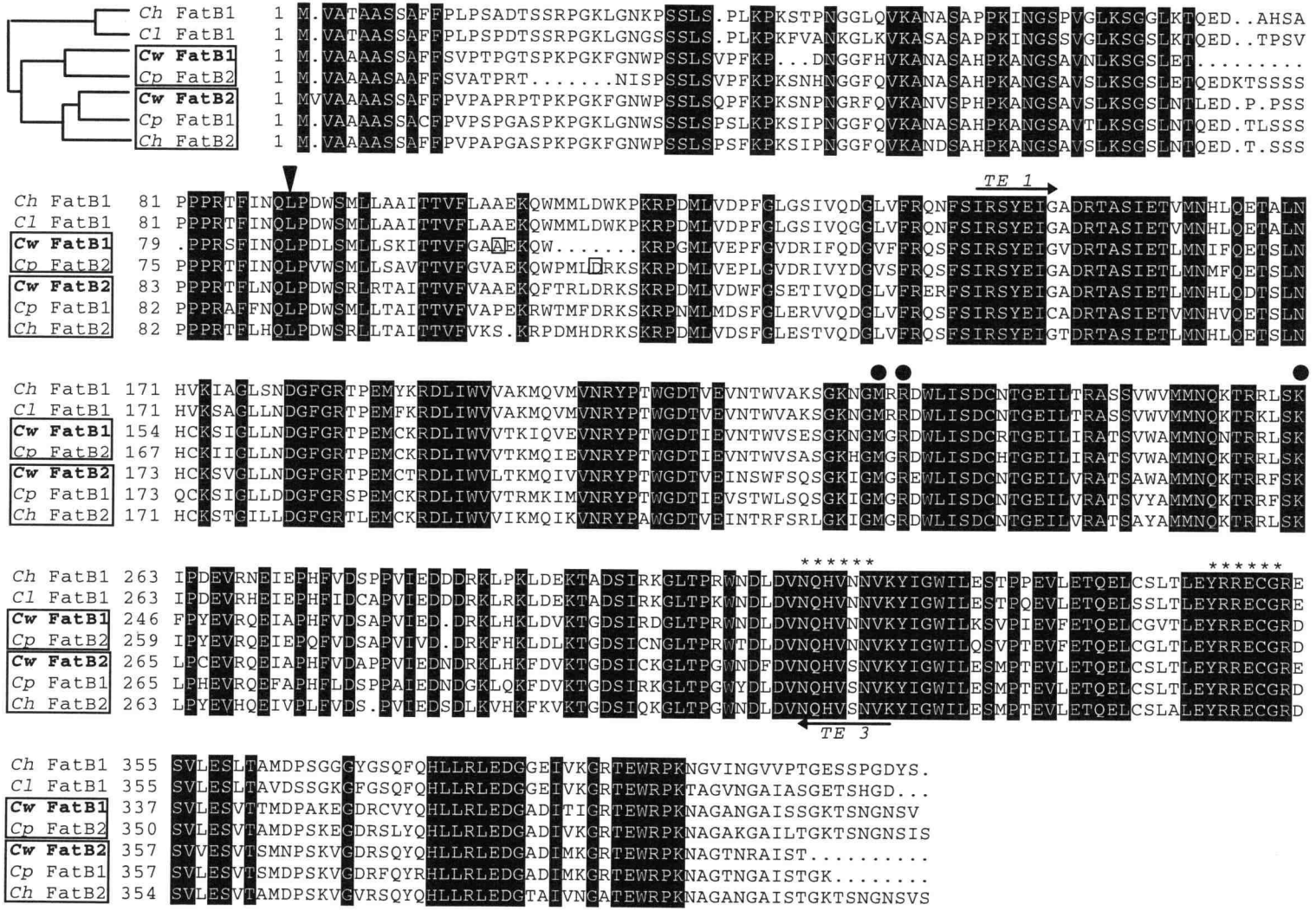
Ch FatB1
Cl FatB1
Cw FatB1
Cp FatB2
Cw FatB2
Cp FatB1
Ch FatB2
TE 1
TE 3

Figure 2.1

Tissue-specific expression of *Cw FatB1* and *Cw FatB2*

In order to characterize the expression patterns of *Cw FatB1* and *Cw FatB2*, we raised antibodies to recombinant fusion proteins. Regions of *Cw FatB1* and *Cw FatB2* encoding the putative mature proteins (Figure 2.1) were cloned into pET15B (Novagen),and polyclonal antibodies raised to the histidine-tagged fusion proteins. When antisera were assayed for specificity against the recombinant proteins, anti-*Cw* FatB1 serum recognized both recombinant proteins, whereas anti-*Cw* FatB2 serum detected only the *Cw* FatB2 fusion protein (data not shown).

We expected thioesterases that regulate seed oil phenotype to be primarily expressed in developing embryos. When various tissues of *C. wrightii* were assayed, anti-*Cw* FatB1 serum reacted to 41.9 and 41.0-kD proteins in developing seeds and three proteins of ~33 kD in mature seeds. Anti-*Cw* FatB2 serum reacted to a 41.9-kD protein in developing seeds and two ~33 kD proteins in mature seeds. Neither antiserum produced major bands in samples from flowers, cotyledons, roots, leaves, or stems (data not shown). Because anti-*Cw* FatB1 serum recognized both recombinant proteins, the 41.9-kD band probably represents *Cw* FatB2, whereas the 41.0-kD band is the *Cw FatB1* gene product.

The two size classes apparent in developing (~41 kD) and mature (~33 kD) seeds were investigated further by Western analyses of seed samples collected at two day intervals during the course of seed development (Figure 2.2). Both anti-*Cw* FatB1 and anti-*Cw* FatB2 sera detected ~41-kD proteins at 7 days postanthesis. The concentration of the FatB proteins was maximal from 9 through 15 days postanthesis. Beginning 13

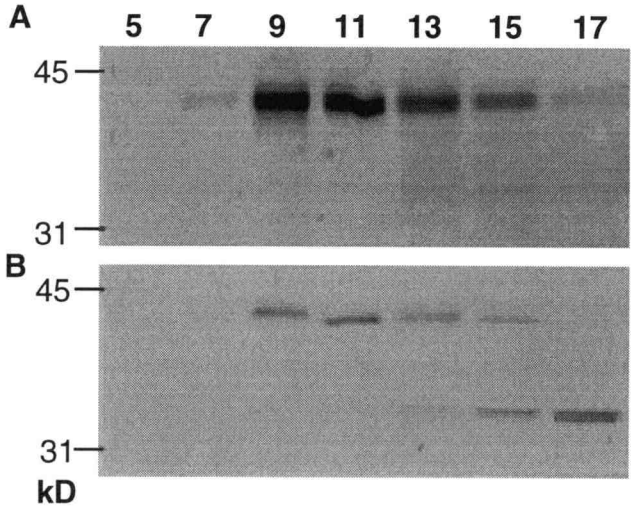

Figure 2.2 Expression of thioesterases during seed development. Seeds of *C. wrightii* were collected at two-day intervals from flowers tagged at anthesis. Proteins from crushed, frozen seeds were solubilized in extraction buffer (1:10, w/v). Each lane was loaded with 1.5 µl of crude lysate. Lane numbering indicates the number of days postanthesis of each sample. **A**. Immunoblots probed with anti-*Cw* FatB1. **B.** Immunoblots probed with anti-*Cw* FatB2 serum.

days postanthesis, the smaller class of proteins (~33 kD) appeared and a concomitant reduction in the larger class of proteins (~41 kD) was observed, suggesting that the smaller bands were degradation products of the larger bands.

Expression of *Cw FatB1* and *Cw FatB2* in *Arabidopsis*

No measureable thioesterase activity was detected during repeated attempts to measure the *in vitro* activity of the histidine-tagged fusion proteins in either crude lysates of *E. coli* expressing the fusion proteins or in assays of partially purified recombinant proteins. To assess the *in planta* function of the *Cuphea* TEs, we transformed *Arabidopsis thaliana* (WS) with the coding sequences of *Cw FatB1* and *Cw FatB2* under

the control of the seed-specific napin promoter (Kridl *et al.*, 1991). Five kanamycin resistant T_1 plants transformed with *Cw FatB1* and six kanamycin resistant T_1 plants transformed with *Cw FatB2* were identified. To confirm expression of the TE transgenes, fatty acid profiles were determined for 10 to 12 individual T_2 seeds from each of the eleven transformants. Approximately 88% of the seeds tested from four of five plants transformed with *Cw FatB1* had dramatically altered fatty acid phenotypes. Increases in medium-chain fatty acids ranged from 2.4 to 9.2% 12:0 and 10.7 to 23.3% 14:0, while decreasing amounts of unsaturated long-chains were recorded (data not shown). Accumulation of 16:0 in these seeds was ~20% compared to 8% in the wildtype. Only one of 11 seeds tested from the fifth *Cw FatB1* transformant exhibited an altered fatty acid phenotype. Similarly, ~75% of seeds from three plants transformed with *Cw FatB2* also had altered fatty acid profiles. Increases of medium-chain fatty acids ranged from 1.6 to 3.0% 10:0, 6.0 to 10.5% 12:0, and 10.4 to 12.8% 14:0 (data not shown). Accumulation of 16:0 in these seeds ranged from 19.5 to 26.7%. Seeds from three other kanamycin resistant plants transformed with *Cw FatB2* did not accumulate medium-chain fatty acids

We isolated a *Cw FatB1* line and a *Cw FatB2* line with single transgene loci by characterizing segregation of kanamycin resistance in T_2 progeny. T_2 progeny from Line 11-1 (*Cw FatB1*) segregated 3:1 for kanamycin resistance as expected for a single dominant gene ($\chi^2 = 0.298$, $P = 0.585$). The T_2 progeny from line 481-3 (*Cw FatB2*) also segregated 3:1 for kanamycin resistance ($\chi^2 = 0.228$, $P = 0.633$). T_2 seeds from both of these lines also accumulated medium-chain fatty acids indicating expression of the transgenic TEs; therefore both lines were analyzed further.

Because of the unexpected broad fatty acid profiles of the transformants, we considered gene dosage effects by analyzing the fatty acid profiles of both hemizygous and homozygous classes of the T_2 progeny from both lines. Three phenotypic classes were identified in T_2 seeds from plant 11-1 (*Cw FatB1*). One class (15 progeny) accumulated no medium chains and was presumed to carry no transgenes. A second class (15 progeny), presumed to be hemizygous, produced 5.1 mol% 12:0, 15.1 mol % 14:0, and 20.1 mol% 16:0. The third class (11 progeny) produced 8.5 mol% 12:0, 22.6 mol % 14:0, and 21.1 mol% 16:0 and was presumed to be homozygous for *Cw FatB1*. The observed segregation ratio was not significantly different than expected for a single segregating locus ($\chi^2 = 2.930$, $P = 0.087$). Doubling the gene dosage in the homozygotes caused a 67% increase in 12:0 and a 50% increase in 14:0 relative to the hemizygotes (Figure 2.3A).

Three phenotypic classes were also identified among the 45 T_2 progeny of 481-3 (*Cw FatB2*). One class (19 progeny) accumulated no medium chains and was presumed to carry no transgenes. A second class (19 progeny), presumed to be hemizygous for *Cw FatB2,* produced 1.7 mol% 10:0, 11.4 mol% 12:0, 13.7 mol % 14:0, and 18.6 mol% 16:0. The third class (7 progeny) produced 2.0 mol% 10:0, 11.3 mol% 12:0, 15.5 mol % 14:0, and 20.8 mol% 16:0 and was presumed homozygous for *Cw FatB2*. The observed segregation ratio was significantly different than expected for a single locus ($\chi^2 = 7.119$, $P = 0.008$). *Cw FatB2* was dominant for 10:0 and 12:0 accumulation; the mean accumulations of these fatty acids was the same for hemizygotes and homozygotes. However, both 14:0 and 16:0 increased by ~ 11% in the homozygotes relative to the hemizygotes (Figure 2.3B)

We confirmed the identification of the putative T_2 homozygotes by isolating a T_3 line homozygous for *Cw FatB1* (111-C) and a T_3 line homozygous for *Cw FatB2* (4813-H). The T_3 progeny from both lines were 100% kanamycin resistance. Fatty acid profiles of pooled seed samples (Table 2.1) were not significantly different than the values of

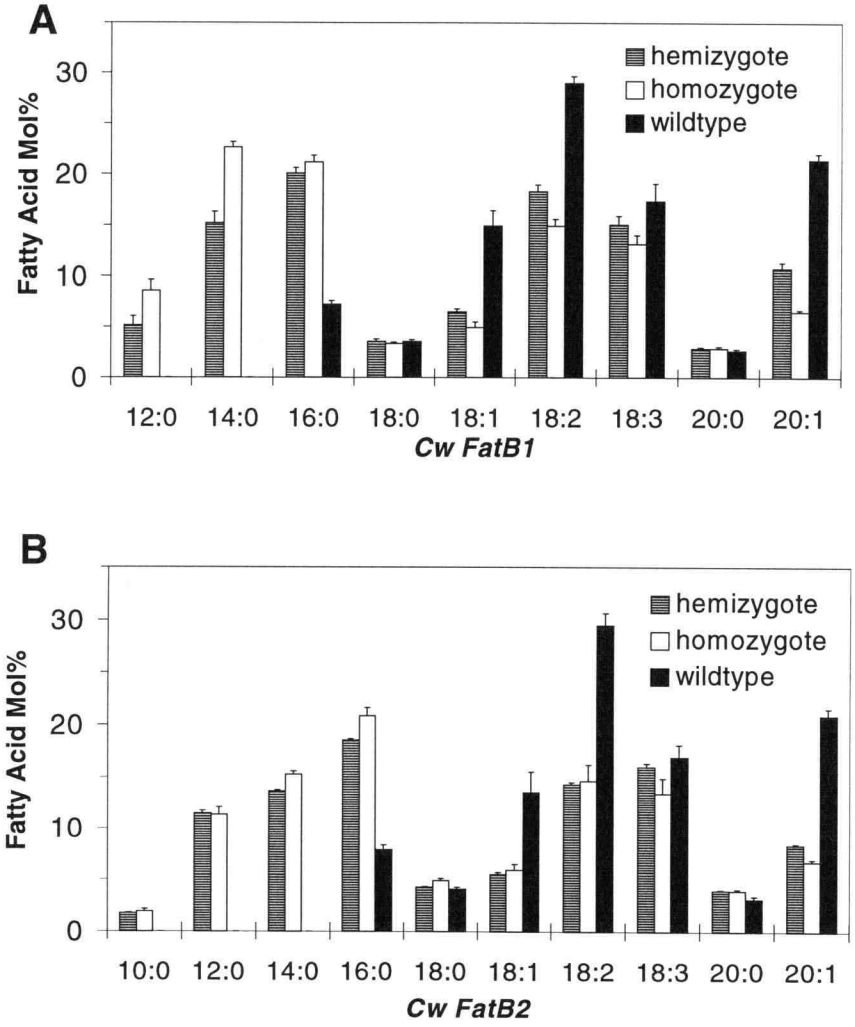

Figure 2.3. Accumulation of fatty acids in T_2 seeds of transformed *Arabidopsis*. Approximately 40 individual seeds from a T_1 plant with a single *Cw FatB1* locus or *Cw FatB2* locus were analyzed by GC. Seeds were grouped and assigned putative zygosities based on accumulation of 14:0 (*Cw FatB1*) or 14:0 and 16:0 (*Cw FatB2*). Data represent means and error bars indicate standard deviation. **A**. Seeds from plant 11-1(*Cw FatB1*) **B**. Seeds from plant 481-3 (*Cw FatB2*).

the T_2 homozygous classes. Both transgenes caused a ~45% increase in saturated fatty acids of which ~30% were medium-chains (C_{10} to C_{14}).

Analysis of F_2 progeny segregating for *Cw FatB1* and *Cw FatB2*

We further tested dosage effects by crossing a *Cw FatB1* homozygous plant with a *Cw FatB2* homozygous plant and assaying the segregating F_2 progeny. Four phenotypic classes were identified among 101 F_2 progeny (Figure 2.4A). One class (59 progeny) produced 10:0 as well as 12:0 and 14:0. Because *Cw FatB2* was dominant for 10:0 production, these progeny were presumed to carry one or two *Cw FatB2* loci. Within this group, we could not distinguish between those progeny with no *Cw FatB1* locus and those progeny carrying one or two *Cw FatB1* loci. A second class (15 progeny) lacked 10:0 but produced 5.3 mol% 12:0 and 18.3 mol% 14:0, similar to the *Cw FatB1* hemizygotes. This class was presumed to carry no *Cw FatB2* loci but was presumed hemizygous for *Cw FatB1*. A third class (16 progeny) accumulated no 10:0 and had a phenotype similar to the *Cw FatB1* homozygous parent. This group was presumed to carry no *Cw FatB2* loci but was presumed homozygous for the *Cw FatB1* transgene. The fourth class (10 progeny) produced no medium-chain fatty acids and was presumed to carry no transgenes.

The expected segregation ratio for the four classes was 12:2:1:1. We could not test for linkage between the transgenes because the first class was a mixture of all three *Cw FatB1* genotypes. For the same reason, only the last three classes could be used to check for segregation of *Cw FatB1*. The observed segregation ratio for the *Cw FatB1*

Table 2.1 Fatty acid profiles of seeds from homozygous plants. Data are the means of five replicates of ~100 seeds from a wildtype plant, a plant homozygous for *Cw FatB1*, or a plant homozygous for *Cw FatB2*.

	Fatty Acid Species (mol%)											
	10:0	12:0	14:0	16:0	18:0	18:1	18:2	18:3	20:0	20:1	20:2	22:1
wildtype	0	0	0	8.2	3.8	12.9	29.0	18.5	2.6	20.8	2.2	1.8
Cw FatB1	0	7.1	24.4	22.8	3.3	4.5	14.1	12.9	3.0	6.0	1.1	0.9
Cw FatB2	2.3	11.9	16.0	21.6	4.4	6.9	16.7	10.2	3.0	5.7	0.7	0.6

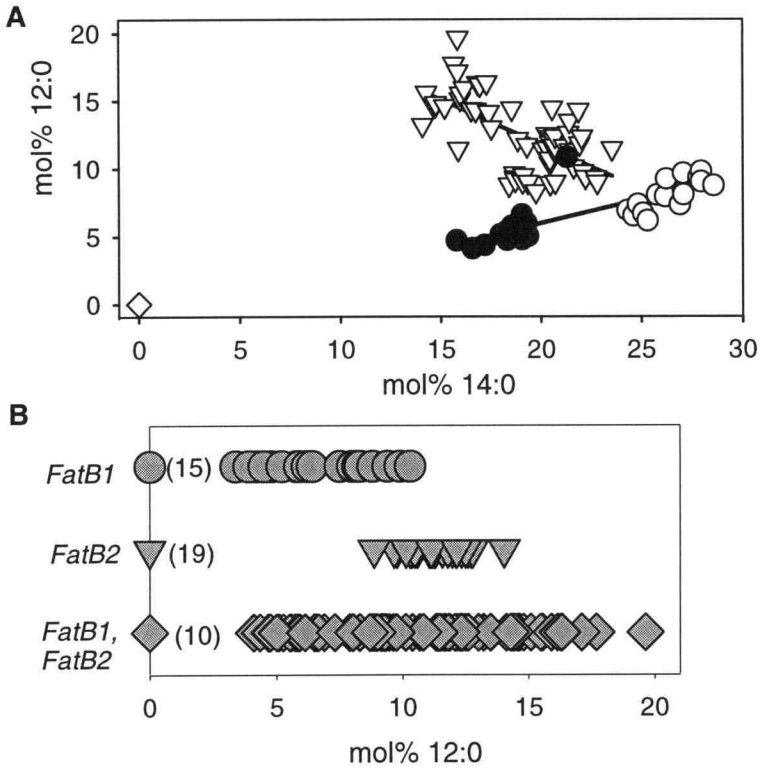

Figure 2.4. **A**. Accumulation of 12:0 and 14:0 in 101 F_2 progeny segregating for both *Cw FatB1* and *Cw FatB2* loci. Putative zygosities were assigned based on accumulation of 10:0, 12:0 and 14:0. Open diamond indicates progeny presumed to carry no transgenes. Filled circles represent progeny presumed hemizygous for *Cw FatB1* but carrying no *Cw FatB2* loci. Open circles indicate progeny presumed homozygous for *Cw FatB1* but carrying no *Cw FatB2* loci. Open triangles indicate progeny presumed to carry one or two *Cw FatB2* loci and 0, 1, or 2 *Cw FatB1* loci. **B**. 12:0 accumulation in T_2 progeny segregating for *Cw FatB1* (circles), *Cw FatB2* (triangles), or F_2 progeny segregating for both *Cw FatB1* and *Cw FatB2* (diamonds). Numbers in parentheses indicate the number of seeds that accumulated no 12:0.

gene (31 *FatB1*$^+$: 10 *FatB2*$^-$) was not significantly different from the 3:1 expected ratio (χ^2 = 0.008 P = 0.929). The whole population was used to check for segregation of *Cw FatB2*. The observed segregation ratio for the *Cw FatB2* gene (59 *FatB2*$^+$: 42 *FatB2*$^-$) was significantly different than the 3:1 expected ratio (χ^2 = 14.815, P = 0.000). We observed a lack of seedling vigor in plants carrying *Cw FatB2* that may be correlated with the apparent segregation distortion against *Cw FatB2* (data not shown).

The additive effects of the genes could not be quantified because of the inability to separate the genotypes in the first class. The mean fatty acid profile of this class was not significantly different than the homozygous *Cw FatB2* parent. However, the upper range of 12:0 values in the F_2 progeny exceeded both parental lines (Figure 2.4B). The upper range of 16:0 values (maximum value = 24.5 mol%) slightly exceeded the *Cw FatB2* parent (maximum value = 22.3 mol%), but 14:0 accumulations never exceeded the *Cw FatB1* parent (maximum value = 23.5 mol%). Interestingly, the correlation between 12:0 and 14:0 production, which was positive in both homozygous parents, became negative in the mixed genotype class (Figure 2.4A). We presumed that those progeny with the highest 12:0 to 14:0 ratio were expressing three or four FatB alleles. The mixed-genotype class accumulated 62.1 mol% saturated fatty acids, similar to the 60.6 mol% total for a *Cw FatB1* homozygote and 59.2 mol% total for a *Cw FatB2* homozygote.

DISCUSSION

The broad specificities of both *Cw* FatB1 and *Cw* FatB2 when expressed in *Arabidopsis* were completely unexpected. TE activity is a key determinant of fatty acid

chain length in California bay (Voelker *et al.*, 1992), *C. hookeriana* (Dehesh *et al.*, 1996b) and in *C. palustris* (Dehesh *et al.*, 1996a). We hypothesized that the seeds of *C. wrightii* accumulate 30% 10:0 and 54% 12:0 through the action of medium-chain specific TEs. The broad substrate specificities for *C. wrightii* TEs we observed did not confirm this assumption.

We cannot explain the lack of *in vitro* activity of the histidine-tagged clones. Measureable activities have been reported from other *Cuphea* TEs cloned with an N-terminal histidine tag (Dehesh *et al.*, 1996a). However, *in vitro* activity levels of a *C. palustris* TE are 900 fold less than those measured in California Bay (Yuan *et al.*, 1995). Therefore, it is possible that the activities of the *C. wrightii* TE recombinant proteins were below the level of detection. Nonetheless, *in vitro* measurements of both the range and relative activity levels of FatB TE on acyl-ACP substrates have closely matched transgenic phenotypes in every instance of which we are aware. This includes *Uc FatB1* expressed in *Arabidopsis* (Voelker *et al.*, 1992) and in canola (Voelker *et al.*, 1996), *Ch FatB2* expressed in canola (Dehesh *et al.*, 1996b) and *Ch FatB1* expressed in canola (Jones *et al.*, 1995). In the case of *Ch* FatB2, small amounts of laurate accumulated in transgenic canola although no 12:0-ACP activity was detected by *in vitro* assay (Dehesh *et al.*, 1996b). It therefore seems likely that the phenotypes reported in our study reflect the specific activities of both *Cw* FatB1 and *Cw* FatB2.

We considered the possibility that the impact on the overall pathway kinetics of *Cw* FatB1 12:0 activity and *Cw* FatB2 10:0 and 12:0 activity might explain the *C. wrightii* phenotype. Specifically, we wondered if the activities of these enzymes on 10:0- and 12:0-ACP pools would sufficiently reduce the pools of 14:0- and 16:0-ACP as to

make these substrates unavailable to the *C. wrightii* TEs. A somewhat similar model was proposed by Davies (1993) to explain the discrepancy between *in vitro* measurements of TE activity from California bay embryo extracts and the actual seed oil phenotype. Computer simulation demonstrates that increasing the concentration of a 10:0 specific TE sufficiently reduces the carbon flux through the pathway to redirect fatty acid accumulation primarily to 10:0 even in the presence of constant concentration of a 12:0 specific TE.

The segregating T_2 populations allowed us to consider the phenotypic effect of doubling 12:0 activity (*Cw FatB1*) or 10:0 and 12:0 activities (*Cw FatB2*) by comparing hemizygous and homozygous phenotypes (Figure 2.3). The Davies model (Davies, 1993) varied the concentration of one monospecific enzyme while holding the concentration of a different monospecific enzyme constant. Because the *C. wrightii* TEs have broad specificities, the concentration of the enzymatic activities for each substrate was presumably doubled in the homozygote as compared to the hemizygote. Both 12:0 and 14:0 increased in the *Cw FatB1* homozygotes relative to the hemizygotes. Therefore the increased diversion at 12:0-ACP did not reduce the 14:0-ACP pool sufficiently to make it unavailable to *Cw* FatB1. *Cw* FatB2 was dominant for 10:0 and 12:0 production; doubling the gene dosage did not increase the accumulation of either fatty acid in the homozygotes. Neither of these studies indicated that the *C. wrightii* phenotype could be directly derived by increasing the concentrations of *Cw* FatB1 or *Cw* FatB2.

When *Cw FatB1* and *Cw FatB2* were combined, some F_2 progeny of the mixed hemizgous class exhibited 12:0 accumulations above either parent together with a decrease in 14:0 (Figure 2.4). The 12:0 increase appears to be due to the addition of *Cw*

FatB1 alleles because *Cw FatB2* was dominant for a 12:0 phenotype of ~11 mol%. The negative correlation between 12:0 and 14:0 accumulation in the mixed class progeny implies that, as suggested in Davies' model (Davies, 1993), increased hydrolysis of shorter acyl-ACPs may reduce longer acyl-ACP pools sufficiently to make them less available for TE hydrolysis. To further test this model in our system would require adding a third FatB locus.

In transgenic canola expressing the 12:0 bay TE, 12:0 and 14:0 accumulations remain positively correlated even as 12:0 production increases due to increased gene dosage (Voelker *et al.*, 1996). This suggests that the 14:0-ACP pool is not sufficiently reduced by 12:0 activity to impact 14:0 production. The substantial levels of 14:0 (~15%) and 16:0 (~18%) we found in the high 12:0 F_2 progeny of the *Cw FatB1/Cw FatB2* cross does not strongly support the hypothesis that even higher levels of *Cw* FatB1 and *Cw* FatB2 could yield the *C. wrightii* phenotype.

We propose two alternative interpretations of these results. First, there may be uncharacterized *FatB* genes in *C. wrightii* with 10:0- and 12:0-ACP specific activity. In *Cuphea*, four FatB genes have been identified in *C. hookeriana* (Voelker 1996), *C. lanceolata* (Martini *et al.*, 1995, Slabaugh *et al.*, in press) and *C. viscossisima* (Slabaugh *et al.*, in press). Based on chromosome number and pollen morphology, *C. wrightii* is hypothesized to be an allotetraploid (Graham, 1988), suggesting the possibility of eight *FatB* genes in this species.

There are two objections to this interpretation. First, the 16:0 TEs from both *Arabidopsis* and *C. hookeriana* are expressed in all tissues. Expression of both *Cw FatB1* and *Cw FatB2* was seed specific, implying a role in seed oil deposition as opposed to

housekeeping. Second, within the four clades of *Cuphea FatB* genes (Voelker, 1996), *Cw FatB2* appears orthologous to the 8:0- and 10:0-ACP specific TEs from *C. hookeriana* and *C. palustris*, (Dehesh *et al.*, 1996a, Dehesh *et al.*, 1996b) which are implicated in chain length regulation of storage lipids. *Cw FatB1* appears orthologous to a 14:0-ACP specific TE from *C. palustris*, which also appears to determine storage lipid phenotype. All gene members of both clades are seed specific (Dehesh *et al.*, 1996a, Dehesh *et al.*, 1996b, Slabaugh *et al*, in press). The conservation of function and pattern of expression in these clades further suggests that *Cw* FatB1 and *Cw* FatB2 participate in chain length regulation in *C. wrightii*. If function and patterns of expression have been conserved in the clade of the ubiquitously expressed 16:0-ACP *Ch* FatB1, it is probable that an uncharacterized 16:0-ACP TE occurs in *C. wrightii* as well.

A second interpretation is that monospecific TEs do not occur in *C. wrightii*. This implies that chain length is determined by a combination of factors including possible interactions with components of fatty acid synthase. Two pieces of evidence support this model. First, Dörmann *et al.* (1993) partially purified a medium-chain TE from *C. wrightii* seeds. Approximately equal activity was demonstrated at 12:0-, 14:0- and 16:0-ACP and slightly less 10:0-ACP activity. In addition to the lack of 10:0/12:0-ACP preference expected in this species, the *in vitro* data mimics the fatty acid phenotype of the transgenic *Arabidopsis* we reported. Second, cerulenin inhibition of fatty acid synthesis in intact plastids of *C. wrightii* (Heise and Fuhrmann, 1994) causes an increase in caprate accumulation. The only known cerulenin insensitive KAS enzyme is KASIII, catalyst of the initial condensation reaction. However, addition of cerulenin to the system decreases total fatty acid synthesis by only 10 to 20%, while greatly decreasing elongation

of 10:0-ACP. These results suggest the possibility of an uncharacterized cerulenin-insensitive KAS capable of medium-chain synthesis. It is interesting to speculate that increased pools of medium-chain acyl-ACPs produced by this type of enzyme might lead to the *C. wrightii* phenotype when coupled with the medium-chain activity of *Cw* FatB1 and *Cw* FatB2.

This is the first report of a correlation between seed development and *in vivo* degradation of FatB TEs (Figure 2.2). Similar size classes were reported when antisera to *Uc* FatB1, the 34-kD TE protein purified from California bay (Davies *et al.*, 1991), reacted with a 40-kD product in canola transformed with *Uc FatB1* (Jones *et al.*, 1995). It was suggested that loss of an N-terminal hydrophobic domain occurred during purification of the 34-kD form (Jones *et al.*, 1995). Although unnecessary for activity, this domain (residues 78 - 99 in *Cw* FatB1) is highly conserved in all FatB TEs and may serve to anchor the FatB genes to a membrane or other fatty acid synthase enzymes (Jones *et al.*, 1995). The physiological significance of the *Cw* FatB1 and *Cw* FatB2 degradation is unknown. If loss of the N-terminal domain in bay is related to the proteolysis of *C. wrightii* TEs, one possibility is that removal of this domain terminates association of FatB with an as yet uncharacterized site *in vivo*. The expression patterns of the 41-kD proteins in developing seeds of *C. wrightii* (Figure 2.2) parallel the expression of acetyl-CoA carboxylase in the same species (Deerberg *et al.*, 1990), suggesting that this is the physiologically relevant form.

Isolation of medium-chain specific TEs from three *Cuphea* species (Dehesh *et al.*, 1996a, Dehesh *et al.*, 1996b, Slabaugh *et al.*, in press) suggest that FatB duplication and divergence may be primarily responsible for the diversity of fatty acid phenotypes in this

genus. However, the *C. wrightii* FatB TEs we isolated have broad medium-chain specificities that suggest a more complex mechanism in this plant. The patterns of expression in *C. wrightii* as well as high levels of activity and broad specificities in transgenic *Arabidopsis* lead us to conclude that *Cw* FatB1 and *Cw* FatB2 may be necessary but not sufficient determinants of the *C. wrightii* phenotype.

ACKNOWLEDGMENTS

This research was supported by grants from Proctor and Gamble, the USDA-NRICGP (#9137300-6569), and the USDA-ARS (#58-5114-9-1002). We thank the following individuals from Calgene Inc.: Vic Knauf for providing access to the binary vectors used in plant transformation, Jean Kridl for supplying the vectors and technical advice, and Toni Voelker for sharing phylogenetic analyses of thioesterases prior to publication. We also thank Elizabeth Yoon for teaching us how to do the crosses in *Arabidopsis* The Genbank accession numbers for the sequences reported in this article are U56103 (*Cw FatB1*) and U56104 (*Cw FatB2*).

REFERENCES CITED

Bechtold N, Ellis J, Pelletier G (1993) *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. Paris, Life Sciences 316: 1194-1199.

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudet J, Leung J, Staskawicz BJ (1994) RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265: 1856-1860.

Brandt T, Knapp SJ (1993) Gamete selection and fatty acid synthesis at marker loci and meiosis of *Cuphea lanceolata* x *Cuphea* populations. Crop Sci. 33: 1138-1143.

Browse J, Sommerville CR (1991) Glycerolipid synthesis: Biochemistry and regulation. Annu. Rev. Plant Physiol. Plant Mol. Biol. 42: 467-506.

Davies HM (1993) Medium chain acyl-ACP hydrolysis activities of developing oilseeds. Phytochemistry 33: 1353-1356.

Davies HM, Anderson L, Fan C, Hawkins DJ (1991) Developmental induction, purification and further characterization of 12:0-ACP thioesterase from immature cotyledons of *Umbellularia californica*. Arch. Biochem. Biophys. 290: 37-45.

Deerberg S, von Twickel J, Förster H-H, Cole T, Fuhrmann J, Heise K-P (1990) Synthesis of medium-chain fatty acids and their incorporation into triacylglycerols by cell-free fractions from *Cuphea* embryos. Planta 180: 440-444.

Dehesh K, Edwards P, Hayes T, Cranmer AM, Fillatti J (1996a) Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. Plant Physiol. 110: 203-210.

Dehesh K, Jones A, Knutzon DS, Voelker TA (1996b) Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. Plant J. 9: 167-172.

Dörmann P, Spener F, Ohlrogge JB (1993) Characterization of two acyl-acyl carrier protein thioesterases from developing *Cuphea* seeds specific for medium-chain and oleoyl-acyl carrier protein. Planta 189: 425-432.

Dörmann P, Voelker TA, Ohlrogge JB (1995) Cloning and expression in *Escherichia coli* of a novel thioesterase from *Arabidopsis thaliana* specific for long-chain acyl-acyl carrier proteins. Arch. Biochem. Biophys. 316: 612-618.

Graham SA (1988) Revision of *Cuphea* section Heterodon (Lythraceae). Sys. Bot. Monogr. 20: 1-168.

Graham SA, Hirsinger F, Röbbelen G (1981) Fatty acids of *Cuphea* (Lythraceae) seed lipids and their systematic significance. Amer. J. Bot. 68: 906-917.

Heise K, Fuhrmann J (1994) Factors controlling medium-chain fatty acid synthesis in plastids from *Cuphea* embryos. Prog. Lipid Res 33: 87-95.

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

Knutzon DS, Bleibaum JL, Nelsen J, Kridl JC, Thompson GA (1992) Isolation and characterization of two safflower oleoyl-acyl carrier protein thioesterase cDNA clones. Plant Physiol. 100: 1751-1758.

Kridl JC, McCarter DW, Rose RE, Scherer DE, Knutzon DS, Radke SE, Knauf VC (1991) Isolation and characterization of an expressed napin gene from *Brassica rapa*. Seed Sci. Res. 1: 202-219.

Marchuk D, Drumm M, Saulino A, Collins FS (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res. 19: 1154.

Martini N, Schell J, Töpfer R (1995) Expression of acyl-ACP thioesterase in *Cuphea lanceolata* and in transgenic rapeseed. In Kader J-C, Mazliak P (eds), Plant Lipid Metabolism, pp. 495-498. Kluwer Academic Publishers, Dordrecht.

Mcbride KE, Summerfelt KR (1990) Improved binary vectors for *Agrobacterium*-mediated plant transformation. Plant Mol. Biol. 14: 269-276.

Ohlrogge JB (1994) Design of new plant products: engineering of fatty acid metabolism. Plant Physiol. 104: 821-26.

Schuch R, Brummel M, Spener F (1994) β-ketoacyl-acyl carrier protein (ACP) synthase III in *Cuphea lanceolata* seeds: identification and analysis of reaction products. J. Plant Physiol. 143: 556-560.

Slabaugh MB, Huestis GM, Leonard J, Holloway JL, Rosato C, Hongtrakul V, Martini N, Töepfer R, Voetz M, Schell J, Knapp SJ: Sequence-based genetic markers for genes and gene families: single-strand conformational polymorphisms for the fatty acid synthesis genes of *Cuphea*. Theor. Appl. Genet. (in press).

Töepfer R, Martini N, Schell J (1995) Modification of Plant Lipid Synthesis. Science 268: 681-686.

Voelker T (1996) Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis. In Setlow JK (ed), Genetic Engineering, Vol. 18, pp. 111-133. Plenum Press, NY.

Voelker TA, Hayes TR, Cranmer AM, Turner JC, Davies HM (1996) Genetic engineering of a quantitative trait: metabolic and genetic parameters influencing the accumulation of rapeseed. Plant J. 9: 229-241.

Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke SE, Davies HM (1992) Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257: 72-74.

Yuan L, Nelson, BA, Caryl G (1996) The catalytic cysteine and histidine in the plant acyl-acylcarrier protein thioesterases. J. Biol. Chem. 271:3417-3419.

Yuan L, Voelker TA, Hawkins, DJ (1995) Modification of the substrate specificity of an acyl-acylcarrier protein thioesterase by protein engineering. Proc. Natl. Acad. Sci. 92: 10639-10643.

Chapter 3

CUPHEA β-KETOACYL-ACP SYNTHASE SHIFTS THE SYNTHESIS OF FATTY ACIDS TOWARDS SHORTER CHAINS IN *ARABIDOPSIS* SEEDS EXPRESSING *CUPHEA* FATB THIOESTERASES

Jeffrey M. Leonard, Steven J. Knapp, Mary B. Slabaugh

ABSTRACT

Acyl-acyl carrier protein thioesterases (TE) with specificities on medium-chain substrates (C8 to C14) are requisite enzymes in plants that produce 8:0, 10:0, 12:0, and 14:0 seed oils, but they may not be the sole enzymatic determinants of chain length. We investigated the contribution to chain length regulation of a β-ketoacyl-ACP synthase, *Cw Kas21*, isolated from *Cuphea wrightii*, a species that accumulates 30% 10:0 and 54% 12:0 in seed oils. Expression of *Kas21* in *Arabidopsis* led to a 26% reduction in 16:0 in homozygous seeds, consistent with expected KAS II activity on 14:0- and 16:0-ACP. In the presence of the KAS I inhibitor cerulenin, however, transgenic seed extracts extended 6:0- and 8:0-ACP *in vitro,* whereas this activity was greatly reduced in wildtype seed extracts. This medium-chain activity was tested by combining *Kas21* with the *C. wrightii* medium-chain specific thioesterases, *Cw FatB1* or *Cw FatB2,* in crosses of transformed plants. Fatty acid synthesis shifted towards shorter chains in progeny expressing both classes of enzymes. *Kas21/FatB1* homozygotes produced 16.3 mol% more 12:0 than the *FatB1* parent but 7.5 mol% less 14:0, and 11.9 mol% less 16:0. F_2 progeny expressing *Kas21* and *FatB2* produced 4.8 mol% more 10:0, 7.2 mol% more 12:0, 4.3 mol% less 14:0, and 8.2 mol% less 16:0 than progeny expressing *FatB2* alone. We hypothesize that the shift towards production of shorter chains resulted from increased pools of medium-chain acyl-ACPs resulting from *Kas21* activity. The combined activities of KAS21 and FatB thioesterases appear to determine the *C. wrightii* phenotype

INTRODUCTION

The potential industrial applications of uncommon seed oils has resulted in rapid advances in plant lipid metabolism research. The most visible accomplishment has been the isolation (Davies *et al.*, 1991) and expression of the California bay 12:0-ACP thioesterase (TE) in rapeseed (Voelker *et al.*, 1992) resulting in the production of laurate and the development of the first genetically engineered oil crop. The ability of a single enzyme to divert carbon from the common C16 and C18 fatty acids to medium chain fatty acids (C8 to C14) has spurred efforts to isolate similar enzymes from other plant species with unusual fatty acid phenotypes. TEs with the potential to confer different medium-chain phenotypes have now been isolated from a number of species, especially from the genus *Cuphea* (reviewed by Voelker, 1996).

We are investigating chain length regulation in the herbaceous annual *Cuphea wrightii*, a species that accumulates primarily caprate (30%) and laurate (54%) in seed oils (Graham, 1981). Two FatB TEs isolated from *C. wrightii, Cw FatB1* and *Cw FatB2* have unexpected, broad specificities (Leonard *et al.*, 1997). Both produce 12:0, 14:0, and 16:0 in transformed *Arabidopsis*, while only one, *Cw FatB2,* produces 10:0 as well. Although the embryo-specific expression patterns of these TEs suggest that they function in seed oil metabolism, there is a wide discrepancy between the fatty acid phenotypes of the transgenic plants and that of *C. wrightii.*

We speculated that the FatB TEs are necessary determinants of fatty acid phenotype but that other contributory enzymatic activities might exist in *C. wrightii.* Likely candidates for chain-length regulation activities are β-ketoacyl-acyl-ACP

synthases (KAS), the condensing enzymes responsible for the cyclic two carbon elongations. Of three known classes of plant KAS enzymes, only KAS I elongates substrates from 4:0-ACP to 14:0-ACP (Shimikata and Stumpf, 1983); however, even when KAS I is inhibited by cerulenin, intact *C. wrightii* plastids still produce 10:0 (Fuhrmann and Heise, 1993). This suggests the presence of a medium-chain specific condensing enzyme in this species.

In a companion paper (Slabaugh *et al.*, 1997), we report that *C. wrightii* embryo extracts are capable of synthesizing 6:0-, 8:0-, and 10:0-ACP in the presence of 10 μM cerulenin. A KAS II type condensing enzyme is implicated because the activity is lost at higher concentrations of cerulenin. We isolated an embryo-derived cDNA encoding a KAS II homologue, *Cw* KAS21. Polyclonal antibodies raised to a recombinant *Cw* KAS21 protein detect a 46 kDa protein band in six medium-chain producing *Cuphea* species but not in a *Cuphea* species lacking medium-chain fatty acids (MCFAs). Although circumstantial evidence suggests that a KAS II might contribute to the *C. wrightii* phenotype, we were unable to directly assay the activity of recombinant *Cw* KAS21.

We reasoned that a condensing enzyme with the ability to extend medium-chain acyl-ACPs might complement the broadly specific *C. wrightii* FatB TEs to produce the high levels of 10:0 and 12:0 found in the seed oils. In this study, we report the results of two approaches we used to investigate the function of *Cw* KAS21 in seed lipid metabolism. First, by expressing *Cw Kas21* in transgenic *Arabidopsis,* we were able to assay activity *in vitro* as well as directly assess the phenotypic perturbation. Second, we tested the possibility of interaction with the *C. wrightii* TEs by combining the genes in

pairwise crosses of transgenic *Arabidopsis*. The results clearly suggest that both types of enzymes are involved in regulation of acyl chain length in *C. wrightii*.

MATERIALS AND METHODS

Plant transformation and analysis of *Cw Kas21* transformants

Using a combination of PCR and restriction digests, the complete coding sequence of *Cw* Kas21, including 13 5' untranslated nucleotides, was placed under the control of the seed-specific napin promoter (Kridl *et al.*, 1991) in the vector pCGN3223. A *Kpn*I fragment containing the chimeric gene was then cloned into the binary plant transformation vector pCGN1557 (McBride and Summerfelt, 1990). The *npt*II gene carried on this vector allowed kanamycin selection of transformed plants. The completed binary vector was introduced into *Agrobacterium tumefaciens* strain LBA 4404. *Arabidopsis* (WS) plants were transformed using a modification (Bent *et al.*, 1994) of the vacuum infiltration method developed by Bechtold *et al.* (1993). Primary transformants (3 - 5000 seeds per plate) were detected on media containing 50 mg/l kanamycin. Eight primary transformants were isolated and grown to maturity. The segregating T_2 progeny were assayed for expression of *npt*II on media containing 20 mg/l kanamycin (100 -200 seeds per plate).

In vitro analysis of *Arabidopsis* extracts

Siliques from wildtype *Arabidopsis* plants and from transformed plants homozygous for *Cw Kas21* were harvested and stored at -80^0 C. Preparation of extracts and extension assays were performed as described by Slabaugh *et al.* (1997). Briefly, whole siliques were homogenized in buffer, and proteins in the soluble fraction were precipitated with 65% ammonium sulfate prior to use in extension assays. Reactions were performed in the presence of 1 mm NADH, 2 mm NADPH, 50 µM malonyl-CoA, 10 µM [1-^{14}C] acetyl-CoA (50 mCi/mmol), and 50 µM recombinant spinach ACP1. Crude seed protein was preincubated in 10 µM cerulenin (23^0 C for 10 min.) to inhibit KAS I in specified reactions. Reactions contained 120 µg protein/120 µl reaction mix. Aliquots of 36 µl were removed at 15, 30, and 45 minutes and stopped by TCA precipitation. Acyl-ACP products were analyzed on a 2.25M urea/18% polyacrylamide gel. Spinach acyl-ACP1 standards detected by anti-ACP serum were included to determine the size of the reaction products. Radioactivity in the acyl-ACP bands was detected by phosporimaging and measured using ImageQuaNT software (Molecular Dynamics, Sunnyvale, CA).

Construction of hybrid populations

Arabidopsis plants homozygous for *Cw FatB1* or *Cw FatB2* are described by Leonard *et al.* (1997). *Cw FatB1* and *Cw FatB2* homozygotes were used as the female plants in the crosses with a *Cw Kas21* homozygote. F_2 seed was collected from naturally selfed F_1 plants in the greenhouse.

RESULTS

Cw Kas21 reduces 16:0 in transgenic *Arabidopsis*

The isolation and characterization of *Cw Kas21* is described elsewhere (Slabaugh *et al.*, 1997). To test the activity of *Cw* KAS21, we expressed the *Cuphea* enzyme in transgenic plants. The complete coding sequence of *Cw Kas21* was placed under the control of the seed-specific napin promoter (Kridl *et al.*, 1991) in the binary vector pCGN1557 (McBride and Summerfelt, 1990), which encodes kanamycin resistance. We transformed *Arabidopsis* with this construct and identified plants with single transgenic loci by assaying for kanamycin resistance in samples of T_2 progeny from eight primary transformants. The observed segregation ratio of resistance in the T_2 progeny of primary transformant K1 (120 resistant: 27 sensitive) was not significantly different than the expected ratio (3:1) for a single dominant gene ($\chi^2 = 0.003$, $P = 0.956$). K1 T_2 progeny fell into two non-overlapping 16:0 phenotypic classes. Progeny from the largest class (n = 31) had 6.4 mol% 16:0 and were presumed to carry one or two copies of the transgene, while progeny from the smaller class (n = 15) had wildtype 16:0 percentages (8.6 mol%) and were presumed to carry no transgenes. The segregation ratio of the 16:0 phenotype (31 *Kas*21/_: 15 -/-) (minus sign denotes lack of transgene) was not significantly different than expected (3:1) for a single dominant gene ($\chi^2 = 1.420$, $P = 0.233$). Two other primary transformants produced T_2 progeny segregating for a single transgenic locus. Twelve T_2 seeds from each of these two transformants were assayed, and the seeds from both plants formed two 16:0 phenotypic classes identical to the classes identified in line

K1. This was circumstantial evidence that the reduced 16:0 phenotype was the result of *Cw Kas21* expression and not an artifact of transformation. We produced a T_3 line homozygous for *Cw Kas21*. The T_3 progeny were 100% kanamycin resistant and produced a mean of 6.2 mol% 16:0 (Table 3.1). This confirmed the putative genotypic classification of the segregating T_2 progeny.

Table 3.1. Fatty acid profiles of seeds from homozygous plants.
The fatty acid profiles of homozygous plants were determined by GC. Homozygosity was determined by lack of segregation of kanamycin resistance in the singly transgenic plants and by lack of segregation of fatty acid phenotype in the doubly transgenic line.
[a] Data are the means of five replicates of ~100 seed samples.
[b] Leonard *et al.* 1997
[c] Data are the means of 41 single seeds from B1KAS-13

	Fatty Acid Species (mol%)									
	10:0	12:0	14:0	16:0	18:0	18:1	18:2	18:3	20:0	20:1
wildtype[a]	0	0	0	8.2	3.8	12.9	29.0	18.5	2.6	20.8
FatB1[ab]	0	7.1	24.4	22.8	3.3	4.5	14.1	12.9	3.0	6.0
Kas21[a]	0	0	0	6.2	4.3	16.6	29.2	16.7	2.7	20.5
FatB1 Kas21[c]	0	23.4	16.9	10.9	2.9	6.1	15.8	12.1	2.4	7.7

The reduced production of 16:0 in transgenic plants gave some clues to *Cw* KAS21 activity. We reasoned that *Cw Kas21* might encode a protein with 16:0-ACP extending activity, thereby reducing the steady state pool of 16:0-ACP available for hydrolysis by the endogenous *Arabidopsis* TE (Dörmann *et al.*, 1995). It was also possible that competition for malonyl-ACP between a transgenic medium-chain-specific KAS and endogenous KAS I might reduce the 14:0-ACP extending activity of the

Arabidopsis KAS. This scenario could also reduce the 16:0-ACP pool resulting in the same phenotype.

Cw Kas21 promotes extension of 6:0-ACP and 8:0-ACP in extracts of developing *Arabidopsis* siliques

The effect of *Cw* KAS21 on fatty acid synthesis in developing *Arabidopsis* seeds was investigated using a biochemical assay in which silique extracts synthesized fatty acid chains from [1-^{14}C] acetyl-CoA and malonyl-ACP generated *in situ* from malonyl-CoA and ACP. By supplying exogenous spinach ACP1 to the system, the reaction products could be displayed on urea-containing acrylamide gels as a discrete set of acyl-ACP bands (Post-Beittenmiller *et al.*, 1991) and identified by comparison to authentic standards.

Siliques were harvested 7-12 days after flowering from both wildtype and *Cw Kas21*-transformed plants grown at the same time under identical conditions. Preliminary experiments, in which silique extracts were analyzed by immunoblotting, established that expression of the napin promoter-*Kas21* construct was detectable six days after flowering and accumulation was maximal for approximately a one-week period (data not shown). Assays were performed with and without inhibition of KAS I activity by preincubation of the extract with 10 μM cerulenin.

When we compared the relative distribution of chain lengths produced over a 45-minute time course, the reactions containing untreated transgenic plant protein displayed more rapid extension kinetics relative to the untreated wildtype extracts. At 15 minutes, 17% of the acid-precipitable radioactivity in the wildtype reaction and 26% in the

transgenic reaction was in the 10:0- and 12:0-ACP bands, a 1.5 fold difference. By 30 minutes, the bulk of the radioactivity in both reactions (59% in wildtype and 71% in transgenic) was in 10:0- and longer acyl-ACPs, a 1.3-fold difference. There were negligible differences in product distribution by 45 minutes (Figure 3.1).

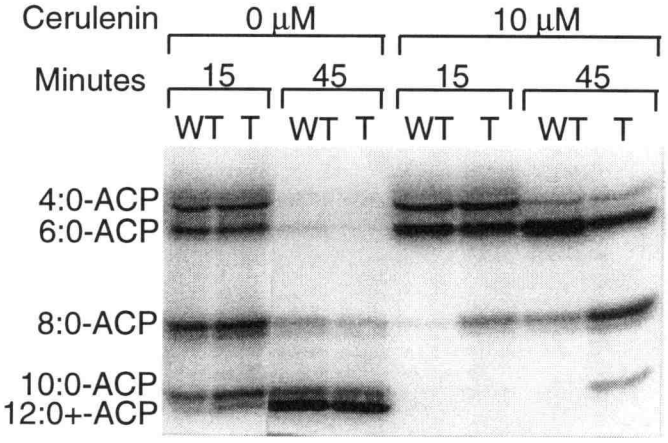

Figure 3.1. Elongation activities of wildtype and transgenic *Arabidopsis* silique extracts. Extracts from siliques of wildtype (WT) and *Cw Kas21* transformed (T) plants were assayed for the ability to synthesize acyl-ACPs *in vitro*. ^{14}C-labeled products were separated on urea-containing acrylamide gels. Bands were identified by comparison to acyl-ACP standards. Extracts were pre-treated with 10 μM cerulenin where indicated.

Preincubation of silique extracts with 10 μM cerulenin, however, revealed a pronounced difference in resistance to the inhibitor. The transgenic extract produced several times the amount of 8:0-ACP relative to the wildtype extract (4.5-fold greater at 15 minutes and 3.6-fold greater at 45 minutes). The transgenic extract also extended acyl-chains to 10:0-ACP, a capability barely detected in wildtype extract. These results suggested that *Cw* KAS21 is a condensing enzyme with preference for shorter chains and

resistance to cerulenin concentrations reported to completely inactivate KAS I (Shimikata and Stumpf, 1982).

Kas21 triples 12:0 production in transgenic progeny expressing *Cw FatB1*

The possibility of an interaction between *Cw* KAS21 and *Cw* FatB1 was tested by producing an *Arabidopsis* F_2 population segregating for both transgenes. We identified four non-overlapping phenotypic classes among 105 F_2 progeny by simultaneously

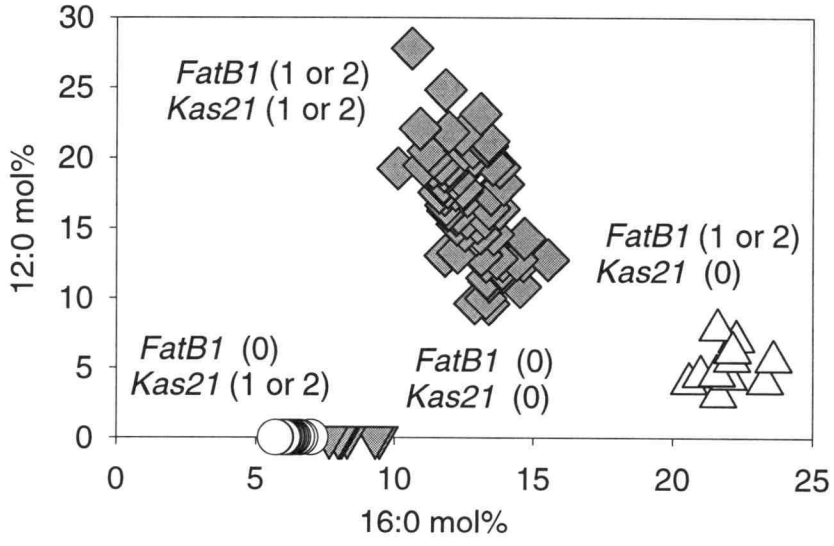

Figure 3.2 F_2 progeny segregating for *Cw FatB1* and *Cw Kas21*.
The fatty acid profiles of 105 single seeds from a hybrid F1 plant were determined by GC. Four classes were identified by simultaneous comparison of 12:0 and 16:0 phenotypes. Genotypes were inferred by similarity to parental or wildtype phenotypes. Numbers in parentheses indicates the putative number of transgene loci in each group. Progeny of the largest class (filled diamonds) were presumed to be expressing both *Cw Kas21* and *Cw FatB1*

comparing 12:0 and 16:0 percentages (Figure 3.2). The fatty acid phenotypes of progeny of three of the classes were similar to the wildtype (n = 8), the *Cw Kas21* parent (n = 18), or the *Cw FatB1* parent (n = 11). Eight progeny had wildtype 16:0 percentages and no MCFAs, and were presumed to carry no transgenes (-/-, -/-). Eighteen progeny produced less 16:0 than the wildtype and no MCFAs. Because *Cw FatB1* is dominant for progeny production of MCFAs (Leonard *et al.*, 1997), and *Cw* KAS21 decreases 16:0, these were presumed hemizygous or homozygous for *Cw Kas21* and lacking *Cw FatB1*. Eleven progeny had 12:0 and 16:0 percentages similar to the *Cw FatB1* parent and were presumed hemizygous or homozygous for *Cw FatB1* and lacking *Cw Kas21*.

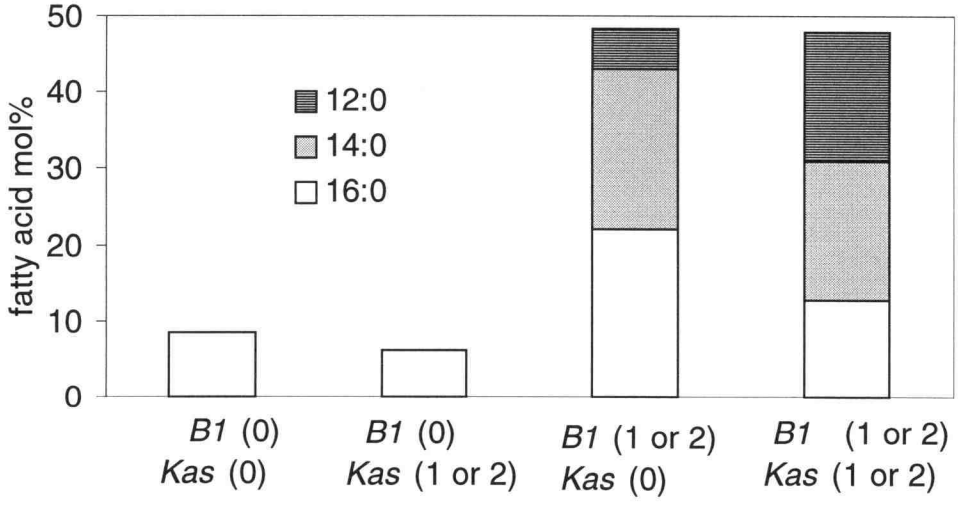

Figure 3.3 Major effects of *Cw FatB1* and *Cw Kas21* on saturated fatty acid production. Data represent the mean accumulations of 12:0, 14:0, and 16:0 in four phenotypic groups identified in an F_2 population segregating for both transgenes. Numbers in parentheses refer to the putative number of loci of each gene in each group.

The fatty acid phenotypes of progeny in the largest class (n = 68) were strikingly different from those of the other three classes (Figure 3.3). These progeny had more than double the 12:0 of the *Cw FatB1* parent. (16.8 versus 7.2 mol%) and were presumed to be hemizygous or homozygous for both transgenes (*FatB1/_*, *Kas21/_*). While 12:0 increased in these progeny, 14:0 decreased by one quarter (24.4 to 18.4 mol%) and 16:0 decreased by one half (22.8 to 10.6 mol%) relative to the *Cw* FatB1 parent.

The observed segregation ratio for *FatB1* (79 FatB1/_: 26 -/-) was not significantly different (χ^2 = 0.003, P = 0.956), and the observed segregation ratio for the *Kas21* gene (86 *Kas21/_*: 19 -/-) was not significantly different (χ^2 = 2.669, P = 0.102) than the expected segregation ratio (3:1) for a single dominant gene. The two transgenes segregated independently (χ^2 = 2.972, P = 0.085).

Because we could not distinguish between hemizygotes and homozygotes for either transgene, we used progeny tests to screen for and develop an F_3 line (B1KAS-13) homozygous for both transgenes. The fatty acid phenotypes of 41 progeny within this line were similar; thus, B1KAS-13 did not segregate and was presumed to be fixed for both transgenes. This line had three-fold more 12:0, one third less 14:0, and one half less 16:0 than the *Cw FatB1* parent (Table 3.1). The 12:0 content of B1KAS-13 was similar to the 12:0 content of progeny from the upper end of the F_2 phenotypic distribution (Figure 3.2). This suggests there was a gene dosage response to one or both of the transgenes.

The production of medium chains in the *Cw FatB1* homozygote is accompanied by a decrease of longer unsaturated chains, 18:1, 18:2, 18:3, and 20:1 (Leonard *et al.*, 1997). The unsaturated long-chain fatty acid contents of B1KAS-13 (41.7 mol%) and the

Cw FatB1 parent (37.5 mol%) were not substantially different. The increase in 12:0 production in the *FatB1/Kas21* seeds came at the expense of 14:0 and 16:0.

Kas21 doubles 10:0 production in transgenic progeny expressing *FatB2*

The possibility of an interaction between *Cw* FatB2, which has 10:0 activity (Leonard *et al.*, 1997), and *Cw* KAS21 was tested by developing an *Arabidopsis* F_2 population segregating for both transgenes. As before, four classes were identified among 97 F_2 progeny. The phenotypes of progeny from three of the classes were similar to the wildtype (n = 9), the *Cw Kas21* parent (n = 25), or the *Cw FatB2* parent (n = 15). The progeny of the largest class (n = 48) produced 2-fold more 10:0 and 1.4-fold more 12:0 than the *Cw FatB2* parental class, whereas 14:0 decreased by one quarter and 16:0 decreased by almost one half (Figure 3.4). Because this phenotype was markedly different from either parental phenotype, we presumed that both transgenes were being expressed in this class.

Progeny were assigned to phenotypic classes by simultaneously comparing 12:0 and 16:0 content. Nine progeny had wildtype 16:0 percentages and no MCFAs and were presumed to be carrying no transgenes (-/-, -/-). Twenty five progeny had less 16:0 than the wildtype (6.2 versus 8.4 mol%) and no MCFAs. These progeny were presumed to lack *Cw FatB2* and were hemizygous or homozygous for *Cw Kas21* (-/-, *Kas21*/_). Fifteen progeny had MCFA percentages similar to the *Cw FatB2* parent and were presumed to be hemizygous or homozygous for *Cw FatB2* and lacking the *Cw* KAS21 transgene (*FatB2*/_, -/-). The 48 progeny of the largest class had the novel phenotype and

were presumed to be hemizygous or homozygous for both transgenes (*FatB2*/_, *Kas21*/_). The 18:1, 18:2 ,18:3, and 20:1 percentages were less than the wildtype, but similar to the *Cw FatB2* parental class (Table 3.2).

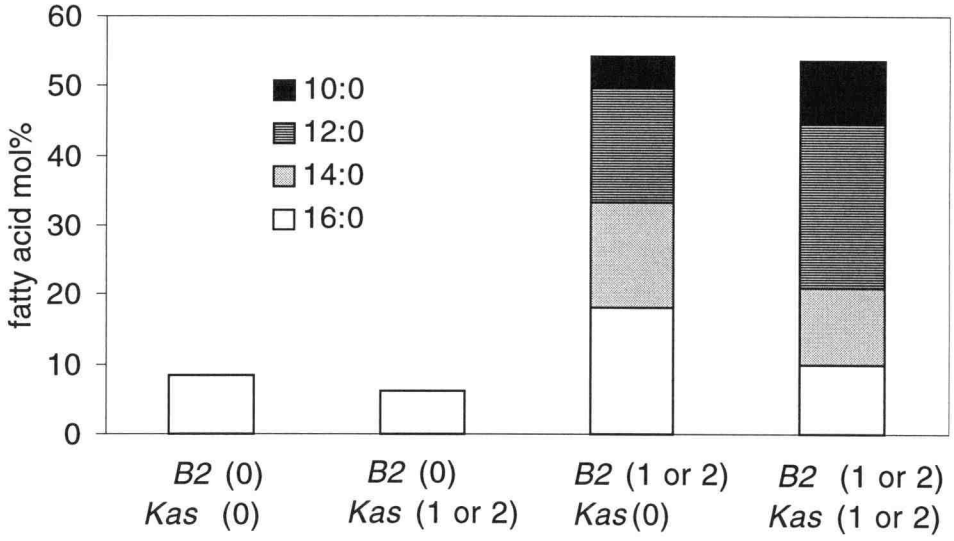

Figure 3.4 Major effects of *Cw FatB2* and *Cw Kas21* on fatty acid production. Data represent the mean accumulations of 10:0, 12:0, 14:0, and 16:0 in four phenotypic groups identified in an F_2 population segregating for both transgenes. Numbers in parentheses refer to the putative number of loci of each gene in each group.

Among F_2 progeny in the *FatB2*/_, *Kas21*/_ class, there were a range of 10:0 (6.3 to 11.7 mol%) and 12:0 (17.7 to 27.8 mol%) phenotypes; however, we could not distinguish between hemizygotes and homozygotes for either transgene and were unable to isolate a doubly homozygous line. Because of the gene dosage effects apparent in the *FatB1/Kas21* cross, it is probable that progeny with phenotypes at the upper ends of the 10:0 and 12:0 ranges were *FatB2/Kas21* homozygotes

The segregation of *FatB2* (63 *FatB2*/_: 34 -/-) was distorted ($\chi^2 = 5.226$, P = 0.022). We have observed a lack of seedling vigor in plants carrying this transgene that

may explain this segregation distortion, but have not determined if this is an effect of the transgene or a position effect resulting from the transformation. The observed segregation ratio for *Kas21* (73 *Kas21*/_: 24 -/-) was not significantly different than expected for a single dominant gene (χ^2 = 0.003, P = 0.956), and the two transgenes segregated independently (χ^2 = 0.092, P = 0.762).

Table 3.2. F_2 progeny of *Cw FatB2* x *Cw Kas21* cross.
The fatty acid profiles of 97 F_2 progeny of a cross between a *Cw FatB1* homozygous plant and a *Cw Kas21* homozygous plant were determined by GC. Four phenotypic groups identified by 12:0 and 16:0 content were assigned putative genotypes based on comparison to both parental and wildtype phenotypes. The numbers of transgenic loci are those presumed to be present in each class.

transgenic loci		Fatty Acid Species (mol%)									
FatB2	*Kas21*	10:0	12:0	14:0	16:0	18:0	18:1	18:2	18:3	20:0	20:1
0	0	0.0	0.0	0.0	8.4	3.8	13.0	29.2	20.1	2.4	19.3
0	1 or 2	0.0	0.0	0.0	6.2	4.0	15.1	30.9	18.6	2.5	19.1
1 or 2	0	4.4	16.4	15.3	18.1	3.9	4.9	13.9	13.6	2.8	5.7
1 or 2	1 or 2	9.2	23.6	11.0	9.9	3.9	7.2	14.2	12.0	2.5	6.2

DISCUSSION

The marked increase in 12:0 among *FatB1/Kas21* transformants and 10:0 and 12:0 in the *FatB2/Kas21* transformants demonstrated that *Cw* KAS21 and FatB thioesterases regulate the distribution of chain lengths in *C. wrightii,* and that they can be used in concert to regulate the fatty acid profiles of oilseeds. The simplest explanation we envision for the synergistic effect between the combined enzymes is that the observed phenotypes directly reflected the size of the substrate pools available to the FatB

thioesterases. This kinetic model assumes that changes in the relative concentrations of acyl-ACPs were a direct effect of the expression and activity of *Cw* KAS21.

The *in vitro* extension assay (Figure 3.1) demonstrated the ability of *Cw* KAS21 to elongate 6:0- and 8:0-ACP in the presence of cerulenin with a subsequent accumulation of 8:0- and 10:0-ACP. We hypothesize that these measurements reflect the situation *in planta*, and that *Cw* KAS21 activity resulted in enlarged pools of 10:0-ACP. Based upon work reported by Post-Beittenmiller *et al.* (1991), we estimate the concentration of 10:0-ACP in spinach seeds to be less than 1.0 μM, well below the 5.1 μM K_m of purified spinach KAS I (Shimikata and Stumpf, 1982). If substrate concentrations and enzyme kinetics in *Arabidopsis* and spinach are similar, the activity of *Arabidopsis* KAS I on increased concentrations of 10:0-ACP in the transgenic plants may have resulted in larger pools of 12:0-ACP available for hydrolysis by *Cw* FatB thioesterases.

These results support the experiments of Fuhrmann and Heise (1993) wherein intact *C. wrightii* plastids inhibited by cerulenin increased 10:0 at the expense of 12:0 and longer chains. The cerulenin induced depression of 12:0 suggests that a cerulenin-sensitive KAS I elongates 10:0-ACP substrate in *C. wrightii,* as we presume occurs in the *Arabidopsis* transformants. Fuhrmann and Heise (1993) speculate that the acyl-ACP equilibrium in *C. wrightii* may be primarily controlled by a cerulenin insensitive condensing enzyme. The similarity between their findings and our extension assay results with *Cw Kas21* transformants leads us to suspect that *Cw* KAS21 may be the responsible enzyme.

In a companion paper (Slabaugh *et al.*, 1997), we demonstrate that *C. wrightii* embryo extracts inhibited by 10 μM cerulenin elongate acyl-ACPs up to 10:0-ACP. The *in vitro* extension assays reported here confirmed our suspicion that *Cw Kas21* encodes this activity. Antibodies to *Cw* KAS21 detect a 46 kDa protein band in six *Cuphea* species that produce medium-chain seed oils, although the 46 kDa protein is barely visible in extract from *C. denticulata,* a species lacking MCFAs (Slabaugh *et al.*, 1997). The correlation between the presence of this protein and the medium-chain phenotypes led us to wonder if the 46 kDa band represents homologues of *Cw* KAS21, whose activities regulate fatty acid chain length. The *in vitro* and *in planta* tests of *Cw* KAS21 function strongly support this hypothesis.

Because both FatB enzymes are very active on 14:0- and 16:0-ACP substrates (Leonard *et al.*, 1997), it is puzzling that the accumulation of these acyl groups declined so dramatically in the *FatB/Kas21* seeds (Tables 3.1 and 3.2). Extending the kinetic model implies that the decreased accumulation was the result of a reduction in the 14:0- and 16:0-ACP substrate pools available for TE hydrolysis. To account for this decline, we can only speculate that overexpression of a 6:0-/8:0-ACP KAS decreased the pool of malonyl-ACP available to the endogenous KAS I responsible for 10:0-/12:0-/14:0-ACP elongation. This hypothesis may also explain the phenotype of the *Kas21* homozygote as follows. Carbon flux to 16:0-ACP decreased as the malonyl-ACP pools available to the endogenous *Arabidopsis* KAS I were reduced by overexpression of *Cw Kas21*. The smaller 16:0-ACP pool available for hydrolysis by the *Arabidopsis* TE resulted in the reduction in 16:0 in the *Cw Kas21* transformants.

It is also possible that the substrate specificity range of *Cw* KAS21 is not a continuum of acyl-ACP chain lengths but rather bimodal. A KAS II cloned from castor (Genez *et al.*, 1991) and expressed in *Arabidopsis* causes a similar reduction of 16:0 accumulation in the transgenic progeny attributed to 16:0-ACP elongation activity of the cloned enzyme (reviewed by Ohlrogge, 1994). The deduced amino acid sequence of the mature *Cw* KAS21 protein is 75% identical (Slabaugh *et al.*, 1997) to the mature castor KAS II. Therefore, it is possible that *Cw* KAS21 might be active on 16:0-ACP as well as 6:0- and 8:0-ACP. This type of activity pattern could explain both the 16:0 decrease in the *Cw Kas21* homozygotes and the phenotype of the transgenic seeds expressing both *Cw Kas21* and a *Cuphea* FatB.

The structure of plant KAS II enzymes is uncertain. *E. coli* KAS II is a homodimer of a ~45 kDa polypeptide (Garwin *et al.*, 1980). A report that castor KAS II exists as a heterodimer of 46 kDa and 50 kDa polypeptides (Nelsen *et al.*, 1994) raises the interesting question of whether *Cw* KAS21 is heterodimeric or homodimeric in *C. wrightii*. This also introduces the possibility that substrate specificity is an effect of the type of dimer formed.

The activity profile of *Cw* KAS21 suggested by our experiments does not match the 14:-/16:0-ACP substrate preference of the partially purified spinach KAS II (Shimakata and Stumpf, 1982); however, because spinach KAS II clones have not yet been isolated, the relationship between the spinach KAS II and *Cw Kas21* is not clear. The depression of 16:0 in *Arabidopsis* seeds expressing the castor KAS II is the only functional assay of a KAS II clone of which we are aware. We can only speculate whether the novel activity we observed in *Cw Kas21* transformants will be found in other

members of the KAS gene family. *Cw Kas21* may have evolved from the same gene lineage as the spinach KAS II and acquired new specificities. Alternatively, it may represent a different class of KAS enzymes. Isolation of KAS II clones from other species and testing of their substrate preferences will answer this question.

Besides a strictly kinetic explanation, we considered the intriguing possibility of interactions between *Cw* KAS21 and FatB thioesterases. Some provocative experiments recently reported by Roughan and Ohlrogge (1996) demonstrate that semi-permeabilized spinach plastids incorporate acetate into lipids under conditions in which other chloroplast functions are impaired. This is indirect evidence that the enzymes of fatty acid synthesis in spinach are arranged into a membrane-associated complex capable of channeling substrates. It may be possible that the *Cuphea* enzymes form an arrangement not available with their *Arabidopsis* counterparts. Jones *et al.* (1995) speculated that a conserved hydrophobic region of the FatB thioesterases might be involved in a protein complex or membrane interaction. Because the substrates for both *Cw* KAS21 and the FatB thioesterases are the same, it is not unlikely that they would be positioned closely if such a multienzyme assembly existed.

This is the first demonstration of the positive effect of a condensing enzyme on medium-chain production. The combined activities of *Kas21* with both *Cw FatB1* and *Cw FatB2* are significant determinants of the *C. wrightii* phenotype and may also be exploited in the engineering of oilseeds to increase the range of medium-chain phenotypes. We are presently testing some of the hypotheses we have put forth to clarify the role of KAS II and the FatB thioesterases in regulation of seed oil deposition.

ACKNOWLEDGMENTS

We would like to thank Vic Knauf and Jean Kridl of Calgene Inc. for providing the binary vectors used in plant transformation and Elizabeth Yoon for teaching us how to perform the crosses in *Arabidopsis*. We also thank Jan Jaworski and John Ohlrogge for the gift of the FAS assay reagents. This research was supported by grants from Proctor and Gamble, the USDA-NRICGP (#9137300-6569), and the USDA-ARS (#58-5114-9-1002).

REFERENCES CITED

Bechtold N, Ellis J, Pelletier, G. (1993) *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C. R. Acad. Sci. Paris, Life Sciences* 316, 1194-1199.

Bent AF, Kunkel, BN, Dahlbeck D, Brown KL, Schmidt R, Giraudet J, Leung J, Staskawicz, BJ (1994) RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265, 1856-1860.

Davies HM, Anderson L, Fan C, Hawkins DJ (1991) Developmental induction, purification and further characterization of 12:0-ACP thioesterase from immature cotyledons of *Umbellularia californica*. Arch. Biochem. Biophys. 290, 37-45.

Dörmann P, Voelker TA, and Ohlrogge JB (1995) Cloning and expression in *Escherichia coli* of a novel thioesterase from *Arabidopsis thaliana* specific for long-chain acyl-acyl carrier proteins. Arch. Biochem. Biophys. 316, 612-618.

Fuhrmann J, Heise K (1993) Factors controlling medium-chain fatty acid synthesis in plastids from maturing *Cuphea* embryos. Z. Naturforsch. 48c, 616-622.

Garwin JL, Klages AL, Cronan JE (1980) Structural, enzymatic, and genetic studies of β-ketoacyl-acyl carrier protein synthases I and II of *Escherichia coli*. J. Biol. Chem. 255, 11949-11956.

Genez A, McCarter D, Nelson J, Bleibaum J, Thompson GA, Knauf, VC, Stalker D (1991) Molecular cloning and characterization of β-ketoacyl-ACP synthase genes from *Ricinus communis*. *Abstract 719, Third International Congress on Plant and Molecular Biology,* Tucson, AZ.

Graham SA, Hirsinger F, Röbbelen G (1981) Fatty acids of *Cuphea* (Lythraceae) seed lipids and their systematic significance. Amer. J. Bot. 68, 906-917.

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

Kridl J, McCarter, DW, Rose RE, Scherer DE, Knutzon DS, Radke SE, Knauf VC (1991) Isolation and characterization of an expressed napin gene from *Brassica rapa*. Seed Sci. Res. 1, 202-219.

Leonard JL, Slabaugh MB, Knapp SJ (in review) *Cuphea wrightii* thioesterases have unexpected broad specificities on saturated fatty acids. Plant Mol. Biol.

McBride KE, Summerfelt KR (1990) Improved binary vectors for *Agrobacterium*-mediated plant transformation. Plant Mol. Biol. 14, 269-276.

Nelsen JS, Crum LS, Foxall S, Thompson GA (1994) Purification and characterization of β-ketoacyl-ACP synthases I and II from developing endosperm of castor seeds. *Abstract 17, 11th International Meeting on Plant Lipids*, Paris.

Ohlrogg JB (1994) Design of new plant products: engineering of fatty acid metabolism. Plant Physiol. 104, 821-826.

Post-Beittenmiller D, Jaworski J, Ohlrogge JB. (1991) *In vivo* free pools of free and acylated acyl carrier proteins in spinach. J. Biol. Chem. 266, 1858-1865.

Roughan PG, Ohlrogge JB (1996) Evidence that isolated chloroplasts contain an integrated lipid-synthesizing assembly that channels acetate into long-chain fatty acids. Plant Physiol. 110, 1239-1247.

Shimikata, T, Stumpf PK (1983) Purification and characterization of β-ketoacyl-ACP synthetase I from *Spinacia oleracea* leaves. Arch. Biochem. Biophys. 220, 39-45.

Shimikata T, Stumpf PK (1982) Isolation and function of spinach leaf β-ketoacyl-[acyl-carrier-protein] synthases. Proc. Natl. Acad. Sci. 79, 5808-5812.

Slabaugh MB, Leonard JM, Knapp SJ (in review) Condensing enzymes from *Cuphea* wrightii associated with medium chain fatty acid biosynthesis. Plant J.

Voelker T (1996) Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis in *Genetic Engineering, Vol. 18* (Setlow, J.K., ed), NY:Plenum Press, pp. 111-131.

Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke SE, Davies HM. (1992) Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257, 72-74.

Chapter 4

CONCLUSION

The initial objectives of this research were to isolate acyl-acyl carrier protein (ACP) thioesterases (TE) from *Cuphea wrightii* and characterize their role in determination of the *C. wrightii* phenotype. We isolated cDNAs encoding *Cw FatB1* and *Cw FatB2*. When expressed in *Arabidopsis, Cw* FatB1 had 12:0-ACP activity, *Cw* FatB2 had 10:0- and 12:0-ACP activity, while both had 14:0- and 16:0-ACP activity. These broad specificities were inconsistent with the *C. wrightii* phenotype. Combining a *C. wrightii* KAS clone, *Cw Kas21,* with the *C. wrightii* TEs caused a dramatic shift towards accumulation of shorter chains, indicating that the joint activities of both enzyme classes are determinants of the *C. wrightii* phenotype. In addition to enhancing our understanding of fatty acid synthesis in plants, this demonstration of a synergistic effect makes a significant contribution towards development of genetically engineered oil crops.

When we initiated this research, the California bay 12:0-ACP TE, *Uc FatB1* (Voelker *et al.*, 1992), was the only TE clone demonstrated to be medium-chain specific; it was unknown if substrate specific TEs existed in other medium-chain plants or if other phenotype conferring mechanisms existed. Experiments done by Davies (1993) with seed extracts from three medium-chain species strongly suggested that substrate specific TEs were widespread within the plant kingdom. This was confirmed by the isolation of medium-chain specific TEs from *C. palustris* (Dehesh *et al.*, 1996a) and *C. hookeriana* (Dehesh *et al.*, 1996b). Our isolation of *Cw* FatB TEs confirmed this mechanism in *C. wrightii*. The compilation of *Cuphea* medium-chain TE sequences with differing

substrate specificities may allow insights into mechanism of substrate recognition through rational testing of the primary structure of TE proteins.

The substrate specificities of the *Cuphea* TEs, however, were not as restricted as expected. For example, *Ch* FatB1 (Dehesh *et al.*, 1996b) has 2-fold more activity on 10:0-ACP than 8:0-ACP although *C. hookeriana* produces 2-fold more 8:0 than 10:0. The activities of the *C. palustris* TEs are similarly skewed in that *Cp* FatB1 hydrolyzes 8:0- and 10-ACP while *Cp* FatB2 hydrolyzes 14:0- and 16:0-ACP; *C. palustris* primarily produces 8:0 and 14:0 (Dehesh *et al.*, 1996a). The discrepancies between FatB TE activities and *Cuphea* phenotypes suggest that homologues of *Cw* KAS21 exist in other *Cuphea* species and perform functions similar to those in *C. wrightii*. This presumption is also evidenced by a 46 kD protein recognized in six medium-chain producing *Cuphea* species by polyclonal antibodies raised to recombinant *Cw* KAS21, but not in a *Cuphea* species lacking medium-chain fatty acids (Slabaugh *et al.*, in review).

Lack of congruence between the *in vitro* activities of seed extracts from elm, coconut, and camphor and the distribution of fatty acids in seeds of those plants led Davies (1993) to propose a model whereby the relative abundance of substrate specific TEs might produce phenotypes not predicted by measured activity profiles. Expression of transgenic TEs in *Arabidopsis* allowed us to address the effects of enzyme concentration by comparing hemizygous and homozygous transgenic populations. We concluded in Chapter 1 that the *C. wrightii* phenotype could not be recreated through increased gene dosage, and therefore that the dynamics proposed in the Davies model are probably not significant in *Cuphea*. The results presented in the Chapter 2 confirmed this interpretation. Whether substrate specific KASs occur in medium-chain species of other

genera and what role they play in phenotype determination are interesting questions derived from this work.

Expressing *C. wrightii* cDNAs in *Arabidopsis* allowed us to circumvent the insolubility of recombinant *Cw* KAS21 and the unexplained inactivity of recombinant *Cw* FatB1 and FatB2. There are several problems, however, inherent in using transgenic plants, including background activity of the *Arabidopsis* enzymes. We were only able to measure incorporation of fatty acids into triglycerides and could not directly measure enzymatic activity. This could be misleading if *Arabidopsis* assembly enzymes were unable to incorporate significant amounts of smaller fatty acids, or if large amounts of free medium-chain fatty acids were produced but rapidly turned over through the β-oxidation pathway. We were also unable to measure or control concentrations of any substrates consumed during fatty acid synthesis as would have been possible *in vitro*.

On the other hand, creation of transgenic plants allowed us to combine activities in a manner impossible to recreate *in vitro* and has given us tools for future experiments. Homozygous lines developed from unique primary transformants will allow the construction of F_2 populations with 1, 2, 3, or 4 transgenes facilitating the study of dosage effects. This is of particular interest in the case of *Cw Kas21* as increased gene dosage appeared to stimulate 12:0 production at the expense of 14:0 (data not shown). We can also combine *Cw FatB1*, *Cw FatB2*, and *Cw Kas21* in a single plant to characterize their joint effects. As additional KAS clones become available, they can be transformed into *Arabidopsis* to allow similar studies.

BIBLIOGRAPHY

Battey JF, Schmid KM, Ohlrogge JB (1989) Genetic engineering for plant oils: potential and limitations. Tibtech 7, 122-126.

Bechtold N, Ellis J, Pelletier G (1993) *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. Paris, Life Sciences 316: 1194-1199.

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudet J, Leung J, Staskawicz BJ (1994) RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265: 1856-1860.

Brandt T, Knapp SJ (1993) Gamete selection and fatty acid synthesis at marker loci and meiosis of *Cuphea lanceolata* x *Cuphea* populations. Crop Sci. 33: 1138-1143.

Browse J, Sommerville CR (1991) Glycerolipid synthesis: Biochemistry and regulation. Annu. Rev. Plant Physiol. Plant Mol. Biol. 42: 467-506.

Davies HM (1993) Medium chain acyl-ACP hydrolysis activities of developing oilseeds. Phytochemistry 33: 1353-1356.

Davies HM, Anderson L, Fan C, Hawkins DJ (1991) Developmental induction, purification and further characterization of 12:0-ACP thioesterase from immature cotyledons of *Umbellularia californica*. Arch. Biochem. Biophys. 290: 37-45.

Deerberg S, von Twickel J, Förster H-H, Cole T, Fuhrmann J, Heise K-P (1990) Synthesis of medium-chain fatty acids and their incorporation into triacylglycerols by cell-free fractions from *Cuphea* embryos. Planta 180: 440-444.

Dehesh K, Edwards P, Hayes T, Cranmer AM, Fillatti J (1996a) Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. Plant Physiol. 110: 203-210.

Dehesh K, Jones A, Knutzon DS, Voelker TA (1996b) Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. Plant J. 9: 167-172.

Dörmann P, Spener F, Ohlrogge JB (1993) Characterization of two acyl-acyl carrier protein thioesterases from developing *Cuphea* seeds specific for medium-chain and oleoyl-acyl carrier protein. Planta 189: 425-432.

Dörmann P, Voelker TA, Ohlrogge JB (1995) Cloning and expression in *Escherichia coli* of a novel thioesterase from *Arabidopsis thaliana* specific for long-chain acyl-acyl carrier proteins. Arch. Biochem. Biophys. 316: 612-618.

Fuhrmann J, Heise K (1993) Factors controlling medium-chain fatty acid synthesis in plastids from maturing *Cuphea* embryos. Z. Naturforsch. 48c, 616-622.

Garwin JL, Klages AL, Cronan JE (1980) Structural, enzymatic, and genetic studies of β-ketoacyl-acyl carrier protein synthases I and II of *Escherichia coli*. J. Biol. Chem. 255, 11949-11956.

Genez A, McCarter D, Nelson J, Bleibaum J, Thompson GA, Knauf VC, Stalker D (1991) Molecular cloning and characterization of β-ketoacyl-ACP synthase genes from *Ricinus communis*. *Abstract 719, Third International Congress on Plant and Molecular Biology,* Tucson, AZ.

Graham SA (1988) Revision of *Cuphea* section Heterodon (Lythraceae). Sys. Bot. Monogr. 20: 1-168.

Graham SA, Hirsinger F, Röbbelen G (1981) Fatty acids of *Cuphea* (Lythraceae) seed lipids and their systematic significance. Amer. J. Bot. 68: 906-917.

Harwood JL (1988) Fatty acid metabolism. Ann. Rev. Plant Physiol. Plant Mol. Biol. 39, 101-138.

Heise K, Fuhrmann J (1994) Factors controlling medium-chain fatty acid synthesis in plastids from *Cuphea* embryos. Prog. Lipid Res 33: 87-95.

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

Knutzon DS, Bleibaum JL, Nelsen J, Kridl JC, Thompson GA (1992) Isolation and characterization of two safflower oleoyl-acyl carrier protein thioesterase cDNA clones. Plant Physiol. 100: 1751-1758.

Kridl JC, McCarter DW, Rose RE, Scherer DE, Knutzon DS, Radke SE, Knauf VC (1991) Isolation and characterization of an expressed napin gene from *Brassica rapa*. Seed Sci. Res. 1: 202-219.

Leonard JL, Slabaugh MB, Knapp SJ (1997) *Cuphea wrightii* thioesterases have unexpected broad specificities on saturated fatty acids. Plant Mol. Biol. (in review).

Marchuk D, Drumm M, Saulino A, Collins FS (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res. 19: 1154.

Martini N, Schell J, Töpfer R (1995) Expression of acyl-ACP thioesterase in *Cuphea lanceolata* and in transgenic rapeseed. In Kader J-C, Mazliak P (eds), Plant Lipid Metabolism, pp. 495-498. Kluwer Academic Publishers, Dordrecht.

Mcbride KE, Summerfelt KR (1990) Improved binary vectors for *Agrobacterium*-mediated plant transformation. Plant Mol. Biol. 14: 269-276.

Nelsen JS, Crum LS, Foxall S, Thompson GA (1994) Purification and characterization of β-ketoacyl-ACP synthases I and II from developing endosperm of castor seeds. *Abstract 17, 11th International Meeting on Plant Lipids*, Paris.

Ohlrogge JB (1994) Design of new plant products: engineering of fatty acid metabolism. Plant Physiol. 104, 821-826.

Ohlrogge JB, Browse J. (1995). Lipid biosynthesis. Plant Cell 7, 957-970.

Ohlrogge, JB (1994) Design of new plant products: engineering of fatty acid metabolism. Plant Physiol. 104, 821-826.

Pollard MR, Andersen L, Fan C, Hawkins DJ, Davies HM (1991) A specific acyl-ACP thioesterase implicated in medium-chain fatty acid production in immature cotyledons of *Umbellularia californica*. Arch. Biochem. Biophys. 284, 306-312.

Post-Beittenmiller D, Jaworski J, Ohlrogge, JB (1991) *In vivo* free pools of free and acylated acyl carrier proteins in spinach. J. Biol. Chem. 266, 1858-1865.

Roughan PG, Ohlrogge JB (1996) Evidence that isolated chloroplasts contain an integrated lipid-synthesizing assembly that channels acetate into long-chain fatty acids. Plant Physiol. 110, 1239-1247.

Schuch R, Brummel M, Spener F (1994) β-ketoacyl-acyl carrier protein (ACP) synthase III in *Cuphea lanceolata* seeds: identification and analysis of reaction products. J. Plant Physiol. 143: 556-560.

Shimikata T, Stumpf PK (1982) Isolation and function of spinach leaf β-ketoacyl-[acyl-carrier-protein] synthases. Proc. Natl. Acad. Sci. 79, 5808-5812.

Shimikata T, Stumpf PK (1983) Purification and characterization of β-ketoacyl-ACP synthetase I from *Spinacia oleracea* leaves. Arch. Biochem. Biophys. 220, 39-45.

Slabaugh MB, Huestis GM, Leonard J, Holloway JL, Rosato C, Hongtrakul V, Martini N, Töepfer R, Voetz M, Schell J, Knapp SJ: Sequence-based genetic markers for genes and gene families: single-strand conformational polymorphisms for the fatty acid synthesis genes of *Cuphea*. Theor. Appl. Genet. (in press).

Slabaugh MB, Leonard JM, Knapp SJ (in review) Condensing enzymes from *Cuphea* wrightii associated with medium chain fatty acid biosynthesis. Plant J.

Stumpf PK (1987) The biosynthesis of saturated fatty acids. In *The Biochemistry Of Plants*, Volume 9, Lipids: Structure and Function (Stumpf, P.K. ed.) New York: Academic Press, pp. 121-136.

Töepfer R, Martini N, Schell J (1995) Modification of Plant Lipid Synthesis. Science 268: 681-686.

Voelker T (1996) Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis. In Setlow JK (ed), Genetic Engineering, Vol. 18, pp. 111-133. Plenum Press, NY.

Voelker TA, Hayes TR, Cranmer AM, Turner JC, Davies HM (1996) Genetic engineering of a quantitative trait: metabolic and genetic parameters influencing the accumulation of rapeseed. Plant J. 9: 229-241.

Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke SE, Davies HM (1992) Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. Science 257: 72-74.

Voelker, T (1996) Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis in *Genetic Engineering, Vol. 18* (Setlow, J.K., ed), NY:Plenum Press, pp. 111-131.

Yuan L, Nelson, BA, Caryl G (1996) The catalytic cysteine and histidine in the plant acyl-acylcarrier protein thioesterases. J. Biol. Chem. 271:3417-3419.

Yuan L, Voelker TA, Hawkins DJ (1995) Modification of the substrate specificity of an acyl-acylcarrier protein thioesterase by protein engineering. Proc. Natl. Acad. Sci. 92: 10639-10643.

APPENDIX

FOUR NEW MEMBERS OF THE *CUPHEA WRIGHTII* FATB GENE FAMILY

Jeffrey M. Leonard, Mary B. Slabaugh, Steven J. Knapp

INTRODUCTION

Acyl-acyl carrier protein (ACP) thioesterases (TE) are determinants of fatty acid chain length during *de novo* fatty acid biosynthesis in plants. Two classes of plant TEs have been identified; 18:1-ACP specific FatA TEs and saturated acyl-ACP specific FatB TEs (Jones *et al.*, 1995). FatB sequences are characterized by several large deletions relative to the FatA subclass. The duplication of an ancestral Fat gene and subsequent divergence of the two TE classes is believed to have occurred either prior to origination of angiosperms or early in angiosperm radiation (Jones *et al.*, 1995).

Twelve of 22 FatB genes isolated to date have been cloned from four species of *Cuphea*. These include medium-chain (C8 to C14) specific TEs from *C. palustris* (Dehesh *et al.*, 1996a), *C. hookeriana* (Dehesh *et al.*, 1996b), *C. wrightii* (Leonard *et al.*, in review), and *C. lanceolata* (Töepfer *et al.*, 1995), and a 16:0-ACP specific TE isolated from *C. hookeriana* (Jones *et al.*, 1995). A phylogenetic analysis of plant TEs revealed four clades of FatB genes in two *Cuphea* species, *C. lanceolata* and *C. hookeriana* (Voelker, 1996). In addition, we detected four genomic FatB sequences in *C. viscossisima* (Slabaugh *et al.*, in press) which are presumed to be orthologous to those cloned from *C. lanceolata*. Divergence among members of this small gene family is one obvious hypothesis to explain the fatty acid phenotypic diversity in *Cuphea*. Previously, we isolated and characterized two FatB TEs from *C. wrightii* (Leonard *et al.*, in review). Because *C. wrightii* is presumed to be an allotetraploid (Graham, 1988), we hypothesized that eight FatB TEs might be found in *C. wrightii*; one gene from each of two diploid

progenitors in each of the four clades. The objective of this study was to isolate cDNAs for these *C. wrightii* genes.

MATERIALS AND METHODS

An embryo-derived *C. wrightii* cDNA library was constructed as described elsewhere (Leonard *et al.*, in review). Approximately 180,000 plaques were plated and blotted onto nylon membranes. The blots were probed with a ^{32}P labelled *EcoRI/HindIII* fragment of *Cw FatB1* corresponding to the first 837 bp. After stripping, the blots were reprobed with a ^{32}P labelled *EcoRI/XhoI* fragment of *Cw FatB2* corresponding to the first 1008 bp. Hybridizations were performed for 16 h at 50^0 C in 6X SSC, 5X Denhardt's solution, 0.5% SDS, and 100 µg/ml salmon sperm DNA. The final wash was performed for 2 h at 55^0 C in 1X SSC/0.1% SDS. Positive plaques were isolated in a secondary screen and cDNA clones were recovered as excised plasmids Clones were grouped by restriction digests and partial sequences. Sequences were assembled and analyzed using programs included in GCG (Genetics Computer Group, Madison, WI).

RESULTS AND DISCUSSION

In work presented in Chapter 2, we isolated three classes of TEs while probing plaque lifts of a *C. wrightii* cDNA library. Because one probe had hybridized with all three classes, we tested whether we might isolate more members of the gene family by reprobing plaque lifts of the cDNA library with probes derived from the full length TEs. Although we hoped to distinguish gene classes by hybridization strength, both probes

(84% identity) hybridized with approximately equal intensity. Fluctuations in signal strength appeared to be due to variation in plaque size more than target DNA similarity. Six of 54 clones isolated were discarded because the inserts were too small (<500 bp) to encode a TE, or because they were larger than 2500 bp, suggesting a chimera. Eight partially sequenced clones did not encode TEs. Eight clones contained *Cw FatB1* sequences, and seven others contained *Cw FatB2* sequences. Four classes were recognized in the remaining 25 clones and designated *Cw FatB3* (3 clones), *Cw FatB4* (2 clones), *Cw FatB5* (15 clones) and *Cw FatB6* (5 clones). Full length cDNAs were recovered for *Cw FatB3* and *Cw FatB4*, however the plasmids carrying *Cw FatB5* and *Cw FatB6* were chimeric. A 645-bp fragment of unknown DNA lay downstream of the approximate 3' *Cw FatB5* cDNA end as deduced by comparison to *Cw FatB1* and because 13 bp of the clone (beginning at bp 1459) are identical in sequence to the *Eco*RI linkers used during synthesis of the library. A ~425 bp fragment of unknown DNA lay upstream of the true *Cw FatB6* cDNA 5' start based on comparison to another clone from the same class. The reading frames of the cDNAs and the corresponding deduced protein sequences of the four new genes were determined based upon comparison to known TE sequences (Table A.1).

A UPGMA clustering analysis based on similarities between the deduced amino acid sequences of the *Cw* FatB sequences reported here and other sequences recovered from GenBank is represented in the dendrogram in Figure A.1. The topology of the tree is similar to the same portion of a tree derived by maximum parsimony (Voelker, 1996).

The four genes we recovered each represent one of four different clades defined by Voelker (1996). The similarity of *Cw* FatB1 to *Cw* FatB5 (95% identical, 98%

similar) and of *Cw* FatB2 to *Cw* FatB6 (90% identical, 94% similar) suggest that the pairs represent the contribution of the diploid *C. wrightii* progenitors; therefore, the genes are orthologous. The allotetraploid nature of *C. wrightii* was inferred by chromosome number (Graham, 1988). The two pairs of orthologous sequences support that inference. The failure to isolate genes orthologous to *Cw FatB3* and *Cw FatB4* may indicate that

Table A.1 cDNAs isolated from *C. wrightii* encoding FatB acyl-ACP thioesterases. The amino acid sequences were predicted based upon comparison to known FatB thioesterases.

	cDNA (bp)	5' UT (bp)	cds (bp)	3' UT (bp)	preprotein (residues)	predicted mass (kD)
Cw FatB3	1798	318	1239	241	413	45.8
Cw FatB4	1843	299	1248	296	416	46.1
Cw FatB5	1458	16	1227	215	409	45.5
Cw FatB6	1469	28	1251	190	417	46.2

duplication of the *Cw FatB3*/*Cw FatB4* ancestral gene occurred subsequent to the hybridization event that formed *C. wrightii*. Alternatively, the orthologous sequences may not be expressed in embryonic tissue and were not represented in our cDNA library.

Cw FatB1 and *Cw* FatB2 have broad substrate specificities that do not correlate with the *C. wrightii* phenotype. The sequence similarities of *Cw* FatB5 and *Cw* FatB6 to *Cw* FatB1 and *Cw* FatB2, respectively, and their significant representation in our screen suggest that both may contribute to fatty acid chain length regulation.

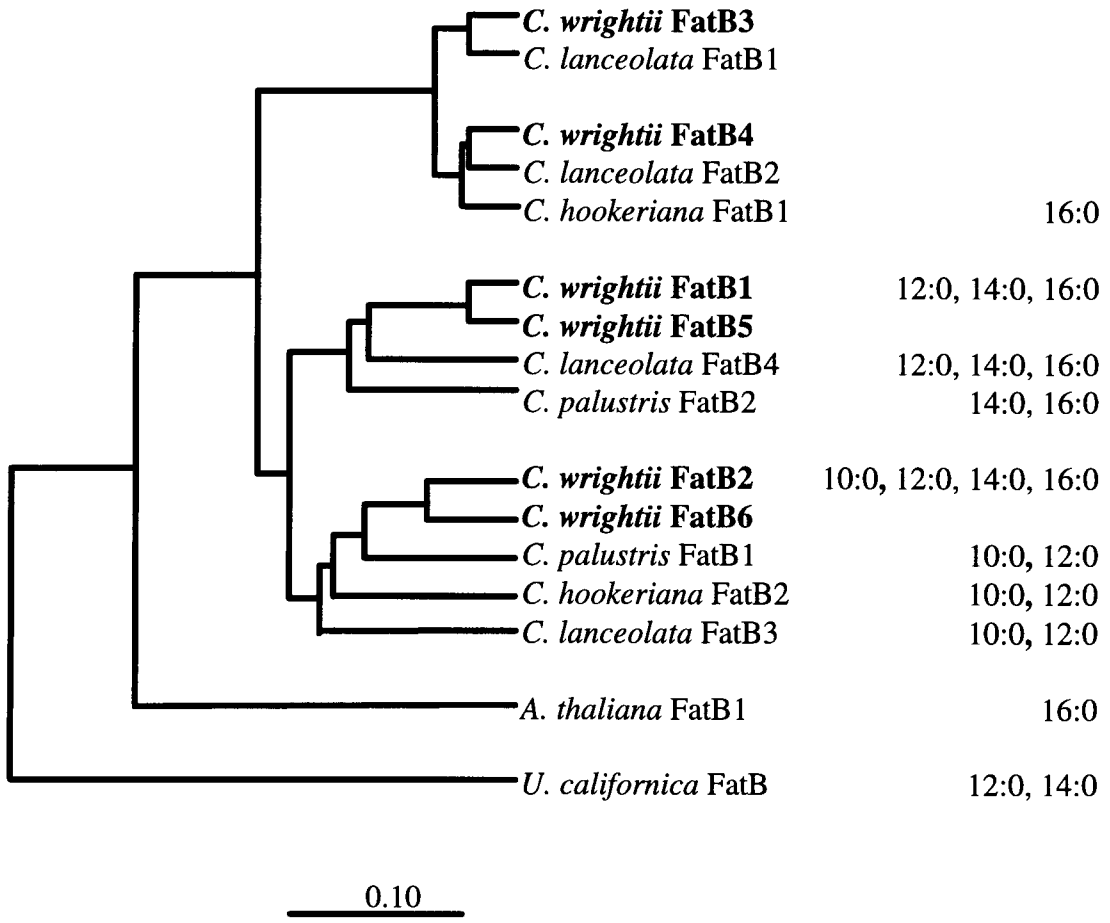

Figure A.1 Dendrogram representing the UPGMA clustering analysis of *Cuphea* FatB genes. The deduced amino acid sequences of the *Cuphea* genes reported here were compared with sequences of other *Cuphea* FatB genes, a 12:0-ACP specific *U. californica* FatB1 (Voelker *et al.*, 1992), and 16:0-ACP specific *A. thaliana* FatB1 (Dörmann *et al.*, 1995). The scale bar represents a 10% difference. Acyl-ACPs for which activities have been shown are listed on the right.

REFERENCES CITED

Dehesh K, Edwards P, Hayes T, Cranmer AM, Fillatti J (1996a) Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil. Plant Physiol. 110: 203-210.

Dehesh K, Jones A, Knutzon DS, Voelker TA (1996b) Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. Plant J. 9: 167-172.

Dörmann P, Voelker TA, Ohlrogge JB (1995) Cloning and expression in *Escherichia coli* of a novel thioesterase from *Arabidopsis thaliana* specific for long-chain acyl-acyl carrier proteins. Arch. Biochem. Biophys. 316: 612-618.

Graham SA (1988) Revision of *Cuphea* section Heterodon (Lythraceae). Sys. Bot. Monogr. 20: 1-168.

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

Leonard JL, Slabaugh MB, Knapp SJ (1997) *Cuphea wrightii* thioesterases have unexpected broad specificities on saturated fatty acids. Plant Mol. Biol. (in review).

Martini N, Schell J, Töepfer R (1995) Expression of acyl-ACP thioesterase in *Cuphea lanceolata* and in transgenic rapeseed. In Kader J-C, Mazliak P (eds), Plant Lipid Metabolism, pp. 495-498. Kluwer Academic Publishers, Dordrecht

Slabaugh MB, Huestis GM, Leonard J, Holloway JL, Rosato C, Hongtrakul V, Martini N, Töepfer R, Voetz M, Schell J, Knapp SJ: Sequence-based genetic markers for genes and gene families: single-strand conformational polymorphisms for the fatty acid synthesis genes of *Cuphea*. Theor. Appl. Genet. (in press).

Töepfer R, Martini N, Schell J (1995) Modification of Plant Lipid Synthesis. Science 268: 681-686

Voelker T (1996) Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis. In Setlow JK (ed), Genetic Engineering, Vol. 18, pp. 111-133. Plenum Press, NY

Voelker, T (1996) Plant acyl-ACP thioesterases: chain-length determining enzymes in plant fatty acid biosynthesis in *Genetic Engineering, Vol. 18* (Setlow, J.K., ed), NY:Plenum Press, pp. 111-131.