Plant Physiol. (1995) 109: 999–1006

Cloning of a Coconut Endosperm cDNA Encoding a 1-Acyl-*sn*-Glycerol-3-Phosphate Acyltransferase That Accepts Medium-Chain-Length Substrates

Deborah S. Knutzon*, Kathryn D. Lardizabal, Janet S. Nelsen, Janice L. Bleibaum, H. Maelor Davies, and James G. Metz

Calgene, Inc., 1920 Fifth Street, Davis, California 95616

Immature coconut (*Cocos nucifera*) endosperm contains a 1-acyl-*sn*-glycerol-3-phosphate acyltransferase (LPAAT) activity that shows a preference for medium-chain-length fatty acyl-coenzyme A substrates (H.M. Davies, D.J. Hawkins, J.S. Nelsen [1995] Phytochemistry 39: 989–996). Beginning with solubilized membrane preparations, we have used chromatographic separations to identify a polypeptide with an apparent molecular mass of 29 kD, whose presence in various column fractions correlates with the acyltransferase activity detected in those same fractions. Amino acid sequence data obtained from several peptides generated from this protein were used to isolate a full-length clone from a coconut endosperm cDNA library. Clone pCGN5503 contains a 1325-bp cDNA insert with an open reading frame encoding a 308-amino acid protein with a calculated molecular mass of 34.8 kD. Comparison of the deduced amino acid sequence of pCGN5503 to sequences in the data banks revealed significant homology to other putative LPAAT sequences. Expression of the coconut cDNA in *Escherichia coli* conferred upon those cells a novel LPAAT activity whose substrate activity profile matched that of the coconut enzyme.

The fatty acyl groups found in membrane phospholipids of higher plants are predominantly 16 or 18 carbons in length. However, the seed TAGs of many plants contain large amounts of fatty acyl groups different from those found in the phospholipids, suggesting that mechanisms exist for partitioning of specific fatty acids into the TAG fraction. Seed storage TAGs are synthesized in the ER from acyl-CoA and glycerol-3-P in a series of reactions termed the Kennedy pathway (reviewed by Stymne and Stobart, 1987; Frentzen, 1993). The first step of this pathway is the acylation of the *sn*-1 position of glycerol-3-P (catalyzed by glycerol-3-P acyltransferase) to form 1-acyl *sn*-glycerol-3-P, also termed LPA. The *sn*-2 position of LPA is subsequently acylated to yield PA in a reaction catalyzed by 1-acyl *sn*-glycerol-3-P acyltransferase (EC 2.3.1.51). This enzyme is more commonly known as LPAAT. Formation of TAG is completed by dephosphorylation of PA to produce diacylglycerol and the transfer of a third acyl group to the *sn*-3 position of the glycerol backbone by diacylglycerol acyltransferase. Synthesis of precursors of membrane phospholipids can also occur via a similar set of enzymatic steps with the exception of the final acylation (reviewed by Frentzen, 1993). Thus, *sn*-3 acylation is the only known enzymatic reaction unique to TAG production.

Although these basic pathways have been characterized, the underlying mechanisms regulating the specific partitioning of fatty acids into TAGs are not well understood. In addition, the distribution of individual fatty acids along the glycerol backbone of both TAGs and phospholipids is not random (Stymne and Stobart, 1987). The specificities of the acyltransferases, particularly LPAAT, are believed to play a key role in this nonrandom distribution (Icihara et al., 1987; Oo and Huang, 1989; Frentzen, 1993). Several studies have demonstrated a correlation between the in vitro substrate specificities of LPAAT assayed in developing seed tissue and the fatty acid composition of the TAGs present in the mature seed (reviewed by Frentzen, 1993). These studies have also indicated that LPAAT is more selective with respect to substrates than the other two acyltransferases of the Kennedy pathway.

The selectivity of the LPAAT involved in TAG synthesis has clear implications for efforts to develop oils with very high levels of specific fatty acids. In particular, the preference of the *Brassica* LPAAT for 18:1-CoA and its lack of activity on 12:0- or erucoyl-CoA (Oo and Huang, 1989; Cao et al., 1990) suggests a theoretical 67 mol% limit for accumulation of these fatty acids in TAGs. Expression of LPAAT from species that can utilize these substrates may permit the incorporation of these fatty acids into the *sn*-2 position of *Brassica* TAGs.

The obstacle in this task has been the lack of cloned genes for LPAAT with the desired specificities. Direct biochemical purification of an LPAAT involved in TAG biosynthesis has been difficult owing to its association with membranes. Although attempts at solubilization have been made (Hares and Frentzen, 1991; Taylor et al., 1992), they have not previously led to identification of proteins that could be associated with the enzyme activity. The only LPAAT sequences so far reported have been obtained by a genetic complementation approach. An *Escherichia coli* mutant defective in LPAAT activity has been isolated (Coleman, 1990), and the corresponding LPAAT gene (*plsC*) has been

* Corresponding author; e-mail knutzon@calgene.com; fax 916–753–1510.

Abbreviations: CHAPS, 3-[(3-cholamidopropyl)-dimethylammonio]-1-propane-sulfonate; HA, hydroxylapatite; LPA, lysophosphatidic acid; LPAAT, lysophosphatidic acid acyltransferase; PA, phosphatidic acid; TAG, triacylglycerol; 12:0, lauroyl; 18:1, oleoyl.

cloned and sequenced (Coleman, 1992). Complementation of this *E. coli* mutant has been used to identify putative LPAAT clones from maize (Brown et al., 1994) and meadowfoam (Hanke et al., 1995). A yeast gene that also complements the *E. coli* mutant has been isolated. This gene, *SLC1*, is thought to encode an *sn*-2-acylglyceride fatty acyltransferase; however, it is not known whether the actual physiological acceptor is lysophosphatidylinositol or LPA (Nagiec et al., 1993).

Recently, we reported solubilization of an LPAAT having medium-chain-length substrate specificity from endosperm of immature coconut (*Cocos nucifera*) (Davies et al., 1995), a species containing high levels of medium-chain fatty acids in its oil. The solubilization was evidenced by three independent criteria: the inability to pellet the activity by high speed centrifugation, molecular weight estimation by size-exclusion chromatography, and protein fractionation. In this paper we report the results of our efforts to use chromatographic separations to identify a protein associated with this medium-chain-specific LPAAT activity, to clone the corresponding cDNA from a coconut endosperm library, and to verify the identity of this clone via expression of the enzyme in *E. coli*.

MATERIALS AND METHODS

Materials

Immature coconut (*Cocos nucifera*) seeds of desired developmental stage were obtained from local retail stores as detailed by Davies et al. (1995). Proteolytic enzymes and restriction endonucleases were from Boehringer Mannheim. Unless otherwise indicated, chemicals and chromatography media were purchased from Sigma. Oligonucleotide primers were synthesized on an Applied Biosystems model 394 DNA and RNA synthesizer.

Enzyme Assays

Radioactive substrates were prepared, and the solubilized LPAAT activity was assayed as described by Davies et al. (1995). Unless otherwise stated, coconut LPAAT activity was measured using [1-^{14}C]12:0-CoA and 12:0-LPA.

Synthesis of 12:0-CoA Sepharose

12:0-CoA was covalently coupled to Sepharose 4B with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide HCl and 6-amino hexanoic acid-Sepharose 4B according to the instructions provided with the product (Sigma). The amount of covalently coupled 12:0-CoA was 4 mg/mL of support (determined by GLC analysis after derivatization to produce fatty acid methyl esters).

Chromatography and LPAAT Protein Identification

Preparation of a coconut endosperm membrane fraction and solubilization of LPAAT using CHAPS were carried out as described by Davies et al. (1995). The 250,000*g* supernatant fraction, which contained LPAAT activity, was diluted 2-fold with a buffer to yield a solution with the following components: 50 mM Hepes, 20% (w/v) glycerol, 0.5 M NaCl, 1.125% CHAPS, 5 mM β-mercaptoethanol, 100 μM Pefabloc (Boehringer Mannheim), 1 μM leupeptin, and 1 μM pepstatin, at pH 7.5. The buffering solution used throughout chromatographic separations was 50 mM Hepes, 20% (w/v) glycerol, 1% (w/v) CHAPS, and 5 mM β-mercaptoethanol (buffer A). Where indicated, buffer A was supplemented by the inclusion of NaCl and/or potassium phosphate. In all cases, the final pH of the buffer was adjusted to 7.5 with NaOH.

The diluted supernatant fraction was applied to a 2.5- × 3-cm column of Red 120 agarose equilibrated with 0.5 M NaCl in buffer A. The column was washed with equilibration buffer and LPAAT activity was eluted with 2.5 M NaCl in buffer A. Fractions from the Red 120 agarose column that contained LPAAT activity were pooled and applied to a 1.5- × 6.6-cm column of Bio-Rad HA HT equilibrated with 1 M NaCl in buffer A. The HA column was washed with 2.5 bed volumes of equilibration buffer. LPAAT activity was eluted during application of a 10-bed-volume linear gradient of 0 to 100 mM potassium phosphate in buffer A that also contained 1 M NaCl. Fractions from the HA column that had LPAAT activity were pooled, diluted 2.5-fold with buffer A, and applied to a 1.5- × 1.1-cm column of 12:0-CoA Sepharose 4B equilibrated in 0.4 M NaCl in buffer A. After the column was washed with 10 bed volumes of equilibration buffer, LPAAT activity was eluted with buffer A containing 2.5 M NaCl. Fractions having LPAAT activity were pooled, diluted 2.5-fold in buffer A, and concentrated 65-fold in an Amicon (Beverly, MA) stirred cell fitted with a YM30 membrane. A portion of the concentrate was applied to a Superose 12 (Pharmacia) size-exclusion column equilibrated with 1 M NaCl in buffer A and LPAAT activity was eluted with the same buffer. Proteins in the eluted fractions were concentrated by ultrafiltration (Centricon 30, Amicon). Samples were adjusted to 30 mM DTT and 2% SDS prior to electrophoresis on Tris-Gly 10 to 15% acrylamide gradient gels (Novex, San Diego, CA). Proteins were visualized using a silver-staining protocol (Blum et al., 1987).

Protein Sequencing

Large-scale preparations of protein for sequencing were made from 200 g of coconut endosperm. The proteins were solubilized and chromatographed over Red 120 agarose and HA as described above with no increase in the column bed sizes. Appropriate fractions from the HA column were pooled, and the proteins were prepared as described above and electrophoretically separated on 12% acrylamide gels (Novex, San Diego, CA). Immediately following electrophoresis, protein was transferred (Towbin et al., 1979) either to a nitrocellulose membrane for enzymatic digestion or to ProBlott (Applied Biosystems) for N-terminal sequence analysis. Protein blotted to nitrocellulose was digested with either trypsin or a combination of asparaginyl-N protease and chymotrypsin (Aebersold, 1989). Peptides obtained from enzymatic digestions were separated via reverse-phase HPLC on a C_{18} column (Vydac, Hesperia, CA) with a gradient of 7 to 52.5% acetonitrile in

0.1 mM sodium phosphate, pH 2.2 (Rosner and Robbins, 1982). Sequencing was performed on an Applied Biosystems 477A protein sequencer with an on-line 120A phenylthiohydantoin analyzer.

Library Construction and Screening

Total RNA was isolated from immature coconut endosperm using a modification of a hexadecyltrimethylammonium bromide DNA isolation procedure as described by Jones et al. (1995). A coconut endosperm cDNA library was prepared using the UniZap system (Stratagene) with the following modifications to the synthesis of first-strand cDNA. Forty micrograms of total RNA from coconut endosperm were heated to 65°C for 20 min and chilled on ice prior to the addition of reaction components. The first-strand synthesis was carried out in a 50-μL reaction volume as recommended by Stratagene with the substitution of Superscript (GIBCO-BRL) first-strand buffer, Superscript reverse transcriptase (600 units), and DTT. First-strand cDNA synthesis was carried out at 45°C for 60 min. The remaining steps of library synthesis were performed as described by the Stratagene UniZap protocol. The primary cDNA library contained 1.42×10^6 clones with an average insert size of 1.25 kb. Library screening and plaque purification were carried out using standard protocols (Sambrook et al., 1989). Hybridization was carried out at 42°C for 18 h in 50% (v/v) formamide, 5× SSC (1 × SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.1% (w/v) SDS, 0.1 mg/mL salmon sperm DNA, 10× Denhardt's solution (1× Denhardt's solution is 0.02% Ficoll, 0.02% PVP, 0.02% BSA), 5 mM EDTA. Filters were washed in 1× SSC, 0.1% (w/v) SDS at 37°C.

Oligonucleotide Design and PCR

Peptide sequences were reverse translated, and degenerate oligonucleotides for PCR primers were synthesized that contained all possible codons corresponding to selected regions of peptide sequence. To reduce complexity, positions of 4-fold degeneracy were synthesized with a 1:1 mixture of deoxyinosine and deoxycytosine. An additional 12 nucleotides were added to the 5′ end of each primer to facilitate subcloning of products using the Cloneamp system (GIBCO-BRL). The sequence of primers successfully used are as follows (degenerate positions are in parentheses): 4865, 5′-CUACUACUACUAAT(T/C/A)AT(T/C/A)TT(T/C)CC(I/C)GA(A/G)GG-3′; 4988, 5′-CAUCAUCAUCAU(A/G)TG(A/G/T)ATCAT(C/T)TC(I/C)AC-3′. Template for PCR was generated by reverse transcription of 10 μg of total coconut endosperm RNA using the primer T-syn (5′-CCAAGCTTCTGCAGGAGCTCTTTTTTTTTTTT TTT-3′) as described above for library construction. PCR was conducted in a Perkin-Elmer Cetus GeneAMP PCR system 9600 instrument in a 50-μL reaction volume containing template derived from 40 ng of total RNA, 0.8 to 2.4 μM each primer, 200 μM each deoxyribonucleotide triphosphate, 60 mM Tris-Cl, pH 8.5, 15 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$. The reaction mixture was heated to 95°C for 10 min and cooled to 72°C, and 1 unit of Amplitaq DNA polymerase (Perkin-Elmer Cetus) was added. Amplification took place using 35 cycles of: 1 min at 94°C, 1.5 min at 48°C, 2 min at 72°C.

DNA Sequencing and Analysis

DNA was sequenced using an Applied Biosystems 373A sequencer. Sequences were analyzed using the Intelligenetics Suite, version 5.3 (Intelligenetics, Inc., Mountain View, CA), and MacVector, version 4.5 (Kodak). Homology searches were performed at the National Center for Biotechnology Information using the BLAST Network Service. DNA sequence alignment was carried out using Megalign software (DNASTAR, Inc., Madison, WI). Protein structure prediction was performed using TopPred II (Claros and von Heijne, 1994).

Expression of Coconut LPAAT in *Escherichia coli*

The *Nde*I, *Nhe*I, and *Bam*HI sites of pET3a (Rosenberg et al., 1987) were replaced by *Sal*I, *Bam*HI, and *Pst*I to create pCGN7645 (M. Lassner, personal communication). The LPAAT cDNA in pCGN5503 was modified by PCR to insert a *Sal*I restriction site immediately upstream of the ATG start codon and a *Bam*HI site immediately downstream of the TAA stop codon. The LPAAT-coding region was inserted as a *Sal*I/*Bam*HI fragment into pCGN7645 to create pCGN5505 for expression in *E. coli* under control of a T7 promoter.

To produce samples for enzyme assay, *E. coli* cultures were grown at 37°C to an A_{600} of 0.5 and induced with isopropyl β-D-thiogalactopyranoside to 0.4 mM for 2 h. Pelleted cells were suspended in 50 mM Hepes, 1 M NaCl, 10 mM EDTA, 100 μM Pefabloc, 1 μM leupeptin, 1 μM pepstatin A, 5 mM β-mercaptoethanol, pH 7.5, and were broken by sonication. The samples were centrifuged at 12,000*g* for 15 min. The supernatant fractions were centrifuged for 2 h at 134,000*g*, and the pelleted membranes were suspended in 50 mM Hepes, 200 mM NaCl, 20% (w/v) glycerol, 5 mM β-mercaptoethanol, pH 7.5. Membrane fractions were assayed for acyl-CoA substrate specificities with 12:0-LPA and various acyl-CoA species as described by Davies et al. (1995).

RESULTS

Identification of a Protein Associated with LPAAT Activity

Starting with the solubilized 12:0-CoA LPAAT activity from coconut endosperm (Davies et al., 1995), several chromatographic methods were screened for utility in enriching LPAAT activity. Inclusion of a high concentration of CHAPS (1%) in the column buffers was necessary to prevent aggregation of the enzyme. Also, at least 400 mM NaCl was required to maintain LPAAT activity. We were unsuccessful in attempts to use either ion-exchange or hydrophobic interaction methods, perhaps because of these detergent and salt requirements. From among several dye resins capable of adsorbing the LPAAT activity under conditions of high ionic strength, Red 120 agarose was chosen for the initial chromatographic step. Approximately 80% of the

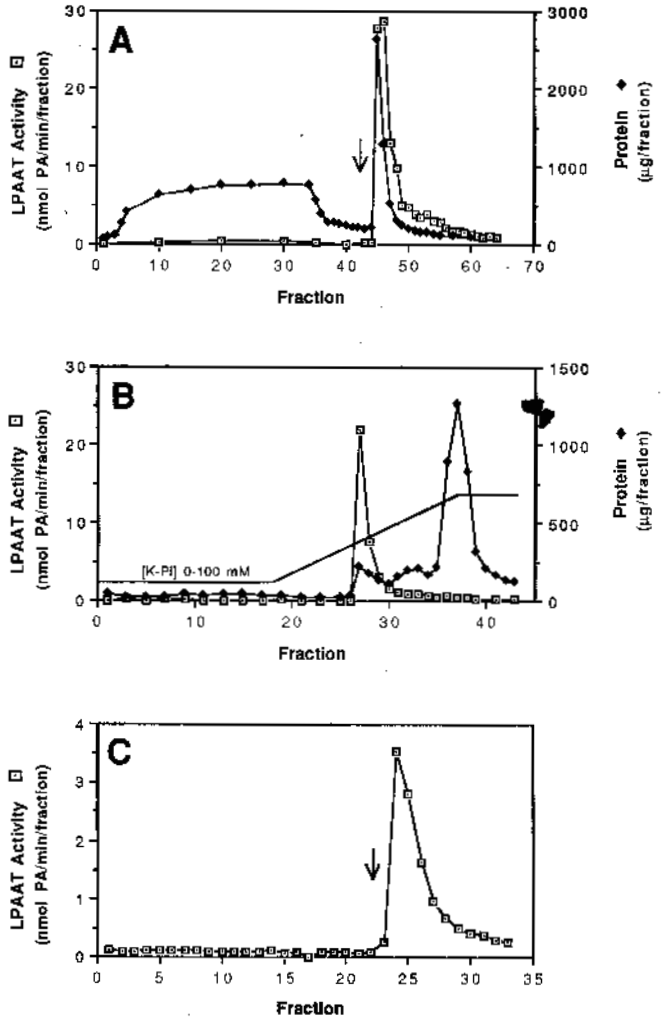

Figure 1. Representative profiles of chromatography of solubilized coconut endosperm LPAAT. A, Dye-ligand chromatography. Membranes were treated with CHAPS, and solubilized material was applied to Red 120 agarose. The column buffer during the load and wash contained 0.5 M NaCl. The arrow indicates the point at which buffer containing 2.5 M NaCl was applied to the column. B, HA chromatography. LPAAT-active fractions from a Red 120 column were loaded directly onto HA and eluted with a potassium phosphate gradient. C, 12:0-CoA Sepharose chromatography. Salt concentrations in the active fractions from an HA column were decreased by dilution, and the sample was applied to a 12:0-CoA Sepharose column. LPAAT activity was eluted by application of a buffer containing 2.5 M NaCl (indicated by the arrow).

CHAPS-solubilized protein was excluded from the Red 120 agarose at 0.5 M NaCl, whereas the LPAAT activity was completely adsorbed. Subsequent elution with 2.5 M NaCl resulted in recovery of 80 to 100% of the applied LPAAT activity (Fig. 1A). All of the LPAAT activity bound to the HA and was eluted with a potassium phosphate gradient (Fig. 1B). Three protein peaks were always obtained during the phosphate gradient elution, but the relative proportions of these peaks varied among preparations, depending on the maturity of the coconut tissue. The LPAAT activity was always associated with the first of the three protein peaks. The combination of chromatography over Red 120 agarose and HA separated the LPAAT activity from approximately 98% of the protein. Recovery of activity from the membrane fraction through the HA column was approximately 45%. The activity of the HA-purified fraction had the same specificity for medium-chain acyl-CoA substrates as did the original solubilized membrane fraction.

From SDS-PAGE analysis of the HA-purified fraction, it was not possible to unambiguously associate one of the remaining proteins with LPAAT activity. Chromatography over 12:0-CoA Sepharose (Fig. 1C) allowed recovery of LPAAT activity while removing several contaminating proteins. Active fractions from the 12:0-CoA column were pooled, concentrated, and applied to a Superose 12 size-exclusion column. The LPAAT activity profile and SDS-PAGE protein-banding patterns of the fractions from this column are shown in Figure 2. Band patterns were compared from fraction to fraction to identify bands whose intensities increased and decreased in concert with LPAAT activity. LPAAT activity was most closely correlated with a

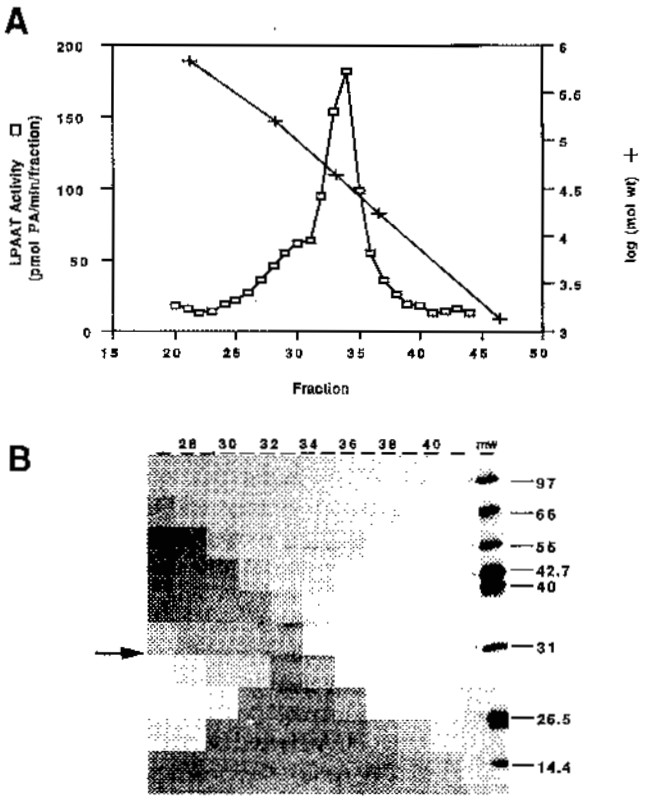

Figure 2. Identification of a 29-kD LPAAT candidate. A, Size-exclusion chromatography. LPAAT-active fractions from the 12:0-CoA Sepharose column were concentrated and applied to a Superose 12 column. The elution pattern of molecular mass standards, chromatographed under the same column and buffer conditions, is also indicated (670 kD, thyroglobulin; 158 kD, γ-globulin; 44 kD, ovalbumin; 31 kD, carbonic anhydrase; 1.4 kD, vitamin B12). B, SDS-polyacrylamide gel analysis of proteins in fractions from the Superose 12 column. The number above each lane of the gel corresponds to the fraction number loaded in that lane. The arrow indicates a 29-kD protein we associated with LPAAT activity. Molecular mass standards (mw lane) are indicated to the right of the gel.

protein band having an electrophoretic mobility of 29 kD. The correlation between LPAAT activity and the presence of the 29-kD protein in specific fractions was also observed upon chromatography on Blue A agarose, thiophilic affinity agarose, and rechromatography on HA (data not shown).

Protein Sequencing

Although purification over the 12:0-CoA and Superose 12 columns was necessary for initial identification of the 29-kD band, only the Red 120 and HA chromatographic steps, in combination with SDS-PAGE, were used for preparation of the protein for sequencing. These two chromatographic steps removed most of the contaminating proteins that migrated near the 29-kD candidate on a 12% polyacrylamide SDS gel. It was not possible to obtain an N-terminal sequence from the intact protein, which suggested that the N terminus was blocked. Amino acid sequence was obtained from a total of 11 peptides (Fig. 3).

Generation of an LPAAT DNA Probe

Degenerate oligonucleotides to be used as PCR primers were designed to regions of sequenced peptides obtained from the 29-kD protein. The relative orientation and spacing of the individual peptides within the protein were not known; therefore, many combinations of primer pairs were used under several reaction conditions. PCR amplification of coconut first-strand cDNA with primers 4865 and 4988 (Fig. 3) resulted in a 271-bp product. Subcloning and sequencing of this PCR product revealed a sequence encoding the known peptides at each end, as well as three additional peptide sequences obtained from the 29-kD protein.

Isolation of pCGN5503

The 271-bp PCR product was radiolabeled and used to probe a coconut endosperm cDNA library. Approximately 240,000 phage plaques were screened, and 29 hybridizing plaques were purified. The cDNAs in each were recovered as excised plasmids. The single-strand DNA sequence of

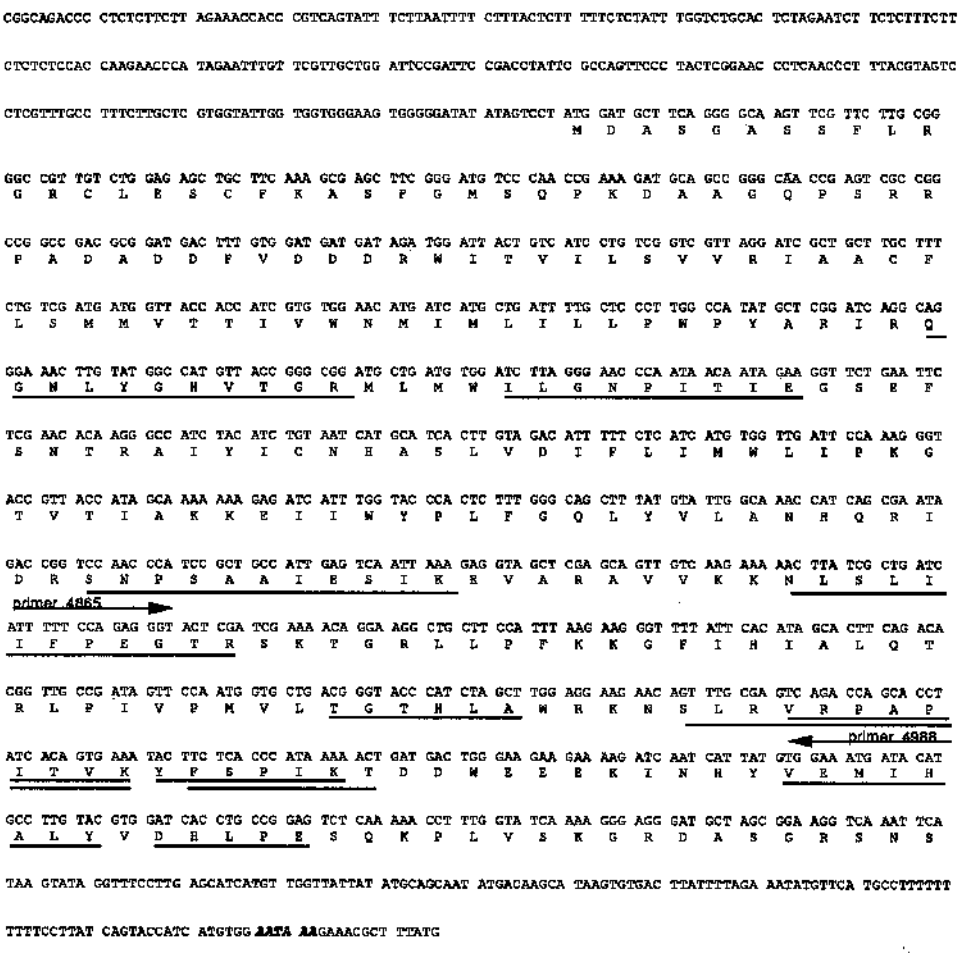

Figure 3. Nucleotide and deduced amino acid sequence of the coconut cDNA clone pCGN5503. Sequenced peptides are underlined. Location of the degenerate primers successfully used in PCR amplificaton are overlined. A consensus polyadenylation signal (Proudfoot and Brownlee, 1976) is italicized and in bold. Double underlining indicates overlapping peptides.

the 3′ end of each cDNA insert indicated that, although various sites of polyadenylation were used, all clones represented transcripts from a single gene.

The complete sequence of one of the longest clones, pCGN5503, is shown in Figure 3. Beginning with the ATG codon at position 259, the 1325-bp cDNA contains an open reading frame encoding 308 amino acids. This ATG is believed to be the translational start site because of the presence of stop codons in all three frames upstream and the similarity to plant consensus start sites (Lütcke et al., 1987). There is a consensus poly(A) addition signal (Proudfoot and Brownlee, 1976) 18 bp 5′ to the site of polyadenylation in this cDNA. The long open reading frame contains the PCR product of primers 4865 and 4988 as well as the other peptide sequences indicated in Figure 3. The calculated molecular weight of the protein encoded by pCGN5503 is 34,806 and the estimated pI is 9.79.

Comparison of the deduced 308-amino acid sequence from pCGN5503 to protein sequence data bases revealed significant homology to the *E. coli* LPAAT encoded by the *plsC* gene (Coleman, 1992) and the closely related sequence from *Salmonella typhimurium* (Luttinger et al., 1991) and also to a putative LPAAT from yeast, *SLC1* (Nagiec et al., 1993). Figure 4 displays some of these regions of homology. In addition, smaller localized homologies were seen to other acyltransferases from *E. coli*, such as 2-acylglycerophosphoethanolamine acyltransferase (Jackowski et al., 1994) and glycerol-3-P acyltransferase (Lightner et al., 1983). The only putative higher plant LPAAT, pMAT1, recently cloned from maize (Brown et al., 1994), was not in the top 20 matches to the coconut sequence.

The hydrophobicity profile of the protein encoded by pCGN5503 is shown in Figure 5; two transmembrane stretches are predicted, consisting of amino acids 64 to 84 and 131 to 151.

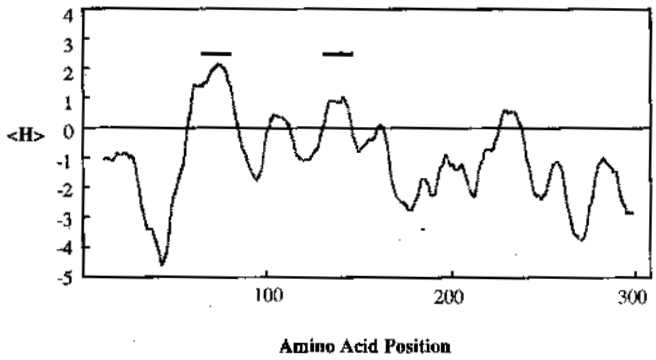

Figure 5. Hydrophobicity (<H>) profile of protein encoded by pCGN5503 as determined by TopPredII (Claros and von Heijne, 1994) using the Goldman, Engleman, Steitz scale (Engleman et al., 1986). Hydrophobic regions are designated as positive numbers. Predicted transmembrane domains are indicated by bars.

Expression of pCGN5503 in *E. coli*

To verify the identification of the cDNA insert of pCGN5503 as the coconut medium-chain-specific LPAAT, pCGN5505 was constructed to express the open reading frame in *E. coli* using a T7 RNA polymerase-based system. Membrane fractions of cells expressing the coconut cDNA had higher activity with medium-chain substrates, especially 12:0-CoA, than did those of control *E. coli*, which preferred 18:1-CoA (Fig. 6A). When the *E. coli* background activity was subtracted from pCGN5505 cultures, the resulting activity profile with medium-chain acyl-CoA substrates was similar to that obtained from membrane fractions of immature coconut endosperm supplied with these substrates (Fig. 6B). In addition, the LPAAT activity of the *E. coli*-expressed enzyme was found to be specific for acyl-CoA versus acyl-ACP substrates (data not shown).

DISCUSSION

In this paper we present the direct biochemical purification and molecular cloning of a higher plant LPAAT involved in TAG biosynthesis. Chromatographic purification of hydrophobic enzymes involved in plant lipid synthesis has traditionally been unsuccessful. However, we were able to establish conditions that maintained the coconut LPAAT in a soluble and active state (Davies et al., 1995). Although we did not purify the LPAAT to homogeneity, we did achieve sufficient enrichment to permit correlation of enzyme activity with the staining intensity of a particular polypeptide band on SDS gels. Cloning of pCGN5503 depended on the amino acid sequence data obtained from proteolytic fragments generated from this candidate polypeptide. Verification that pCGN5503 encodes a coconut medium-chain length preferring LPAAT was demonstrated by detection of that activity in membrane fractions obtained from *E. coli* cultures expressing the enzyme.

The calculated mass of the protein encoded by pCGN5503 is 34.8 kD, whereas the corresponding partially purified protein migrates on SDS gels with an apparent mass of 29 kD. This discrepancy could be due to anomalous migration of the protein during SDS-PAGE; however, preliminary experiments using polyclonal antibodies raised

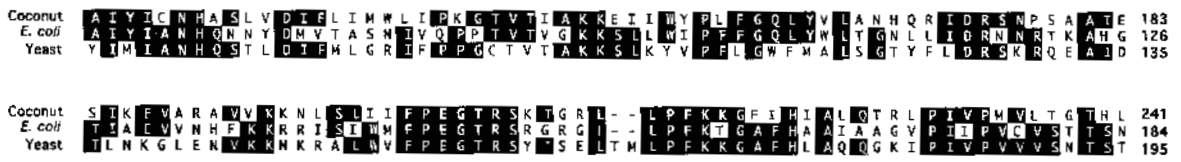

Figure 4. Sequence alignment of coconut LPAAT to other putative acyltransferases. Numbers correspond to amino acid residues. Boxes indicate residues present in at least two sequences. Coconut, pCGN5503; *E. coli*, *plsC* gene (Coleman, 1992); yeast, *SLC1* (Nagiec et al., 1993).

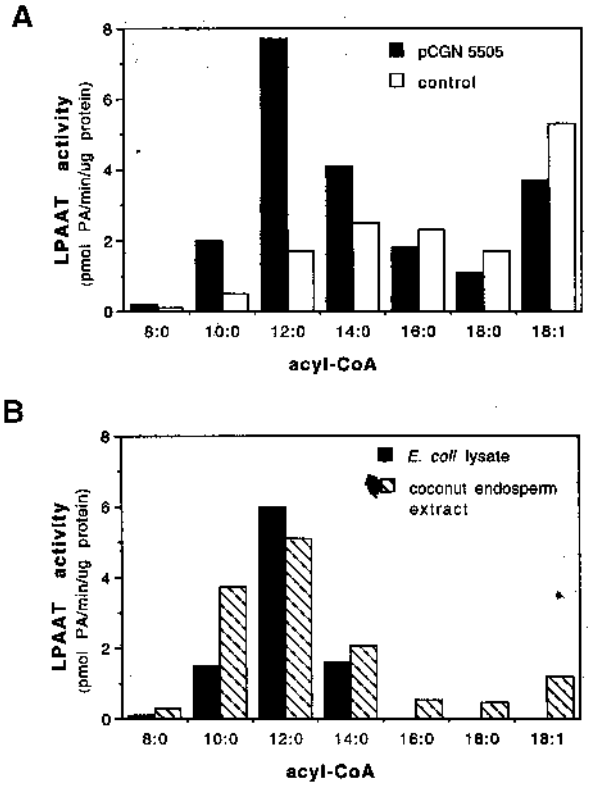

Figure 6. Expression of coconut LPAAT cDNA in *E. coli*. A, Membrane fractions of cultures expressing either pCGN5505 (■) or control plasmid (□) were assayed for LPAAT activity using 12:0-LPA and various acyl-CoA donors. B, Specificity profile of cloned coconut LPAAT (■) obtained by subtracting control background from A is compared to the substrate-specificity profile of membrane fraction of immature coconut endosperm (▨). Activity values for coconut extract have been divided by 100 to allow display on the same scale as those of the *E. coli* lysate. 8:0, Capryloyl; 10:0, caproyl; 14:0, myristoyl; 16:0, palmitoyl; 18:0, stearoyl.

against a recombinant portion of the LPAAT protein suggest that the 29-kD protein has undergone a limited proteolysis during the purification process (data not shown). Further experiments using these LPAAT-specific antibodies to compare the electrophoretic migration of the protein in coconut endosperm extracts with the partially purified samples may resolve this question. The N-terminal sequence encoded by pCGN5503 does not show the characteristics of either a transit peptide of nuclear-encoded chloroplast proteins (de Boer and Weisbeek, 1991; von Heijne and Nishikawa, 1991) or those of a signal peptide (von Heijne, 1990), suggesting that the size discrepancy is not due to in vivo processing.

The protein encoded by pCGN5503 contains two hydrophobic regions predicted to be membrane-spanning domains using the TopPredII program, consistent with an integral membrane protein. Also, as has been observed for several proteins associated with lipid metabolism, the coconut LPAAT has a large net positive charge, which is thought to be important in interactions with both the acidic phospholipid membrane and the LPA substrate (Coleman, 1992). Several plant microsomal enzymes associated with lipid synthesis contain a motif of two Lys residues positioned three and five residues from the C terminus that is thought to be sufficient for retention of transmembrane ER proteins (Jackson et al., 1990). However, since not all ER proteins contain this motif, its significance is not clear (Yadav et al., 1993). The coconut LPAAT described in this paper represents another example of a microsomal protein that lacks this motif.

The preference of the LPAAT activity encoded by pCGN5503 for medium-chain-length acyl-CoA versus acyl-ACP substrates is consistent with that observed in developing coconut endosperm (Davies et al., 1995), a tissue that preferentially incorporates laurate (from 12:0-CoA) into the *sn*-2 position of TAG (Litchfield, 1970). We take this as evidence that we have cloned a cDNA encoding an LPAAT involved in TAG synthesis in coconut. This cDNA would apparently not account for the small amount of LPAAT activity with 18:1-CoA observed in coconut endosperm extract (Fig. 6B). Although LPAAT enzymes involved in cytoplasmic and plastidial membrane lipid synthesis, presumably having a preference for longer-chain fatty acids and/or acyl-ACP substrates, should also be present in developing coconut endosperm, we did not follow these activities in the chromatographic separations of medium-chain LPAAT activity. All of the cDNA clones obtained in this work appear to be derived from a single gene; under the conditions used in our screening, we did not detect homologous clones that could be associated with these other LPAAT activities.

The deduced amino acid sequence of the protein encoded by pCGN5503 contains significant homology to known or suspected LPAATs from bacteria and yeast. Over the region shown in Figure 4, the coconut sequence is 41% identical with both the *E. coli* and yeast sequences. In contrast, the homology of the protein encoded by the maize clone, pMAT1, to these same bacterial and yeast sequences is much more limited (Brown et al., 1994). In addition to the lack of sequence similarity between the proteins encoded by pCGN5503 and pMAT1, the hydrophobicity profiles of the proteins are also very different. Therefore, we believe that we have cloned a cDNA encoding a different enzyme than that encoded by pMAT1.

Transgenic rapeseed plants have been produced in which up to 50 mol% of the fatty acids present in the seed TAG are of medium-chain length (C12 and C14) (Kridl et al., 1993). As anticipated by analysis of rapeseed LPAAT specificities, it appears that the medium chains have been essentially excluded from the *sn*-2 position of the TAG in these lines (data not shown). With the availability of a cDNA encoding the medium-chain-specific LPAAT reported here, it should now be possible to test its effect on the *sn*-2 composition of transgenic rapeseed oil.

ACKNOWLEDGMENTS

We thank Cheryl Eriqat for assistance with the protein purification, Dr. Toni Voelker for preparation of coconut endosperm RNA, Jennifer Yeager for DNA sequencing and oligonucleotide

synthesis, Ann Cranmer for construction of the LPAAT *E. coli* expression vector, and colleagues at Calgene for comments and critical reading of this manuscript.

Received June 28, 1995; accepted August 22, 1995.

The GenBank accession number for the sequence reported in this paper is U29657.

LITERATURE CITED

Aebersold R (1989) Internal amino acid sequence analysis of proteins after *in situ* protease digestion on nitrocellulose. *In* P Matsudaira, ed, A Practical Guide to Protein and Peptide Purification for Microsequencing. Academic Press, San Diego, CA, pp 71–88

Blum H, Beier H, Gross HJ (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis **8:** 93–99

Brown AP, Coleman J, Tommey AM, Watson MD, Slabas AR (1994) Isolation and characterization of a maize cDNA that complements a 1-acyl *sn*-glycerol-3-phosphate acyltransferase mutant of *Escherichia coli* and encodes a protein which has similarities to other acyltransferases. Plant Mol Biol **26:** 211–223

Cao Y, Oo K-C, Huang AHC (1990) Lysophosphatidate acyltransferase in the microsomes from maturing seeds of meadowfoam (*Limnanthes alba*). Plant Physiol **94:** 1199–1206

Claros MG, von Heijne G (1994) TopPred II: an improved software for membrane protein structure predictions. Comput Appl Biosci **10:** 685–686

Coleman J (1990) Characterization of *Escherichia coli* cells deficient in 1-acyl-*sn*-glycerol-3-phosphate acyltransferase activity. J Biol Chem **265:** 17215–17221

Coleman J (1992) Characterization of the *Escherichia coli* gene for 1-acyl-*sn*-glycerol-3-phosphate acyltransferase (*plsC*). Mol Gen Genet **232:** 295–303

Davies HM, Hawkins DJ, Nelsen JS (1995) Lysophosphatidic acid acyltransferase from immature coconut endosperm having medium chain length substrate specificity. Phytochemistry **39:** 989–996

de Boer AD, Weisbeek PJ (1991) Chloroplast protein topogenesis: import, sorting and assembly. Biochim Biophys Acta **1071:** 221–253

Engleman DM, Steitz TA, Goldman A (1986) Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu Rev Biophys Biophys Chem **15:** 321–353

Frentzen M (1993) Acyltransferases and triacylglycerols. *In* TS Moore Jr, ed, Lipid Metabolism in Plants. CRC Press, Boca Raton, FL, pp 195–230

Hanke C, Peterek G, Wolter FP, Frentzen M (1995) cDNA clones from *Limnanthes douglasii* encoding an erucoyl-CoA specific 1-acylglycerol-3-phosphate acyltransferase. *In* J-C Kader, P Mazliak, eds, Plant Lipid Metabolism. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 531–533

Hares W, Frentzen M (1991) Substrate specificities of the membrane-bound and partially purified microsomal acyl-CoA:1-acylglycerol-3-phosphate acyltransferase from etiolated shoots of *Pisum sativum* (L.). Planta **185:** 124–131

Icihara K, Asahi T, Fujii S (1987) 1-acyl-sn-glycerol-3-phosphate acyltransferase in maturing safflower seeds and its contribution to the non-random fatty acid distribution of triacylglycerol. Eur J Biochem **167:** 339–347

Jackowski S, Jackson PD, Rock CO (1994) Sequence and function of the *aas* gene in *Escherichia coli*. J Biol Chem **269:** 2921–2928

Jackson MR, Nilsson T, Peterson PA (1990) Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. EMBO J **9:** 3153–3162

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 JC, Davies HM, Lassner MW, Metz JG (1993) New sources of fats, waxes and oils: the application of biotechnology to the modification of temperate oilseeds. Agbiotech News Inform **5:** 121N–126N

Lightner VA, Bell RM, Modrich P (1983) The DNA sequences encoding *plsB* and *dgk* loci of *Escherichia coli*. J Biol Chem **258:** 10856–10861

Litchfield C (1970) Taxonomic patterns in the fat content, fatty acid composition, and triglyceride composition of Palmae seeds. Chem Phys Lipids **4:** 96–103

Lütcke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA (1987) Selection of AUG initiation codons differs in plants and animals. EMBO J **6:** 43–48

Luttinger AL, Springer AL, Schmid MB (1991) A cluster of genes that affects nucleoid segregation in *Salmonella typhimurium*. New Biol **3:** 687–697

Nagiec MM, Wells GB, Lester RL, Dickson RC (1993) A suppressor gene that enables *Saccharomyces cerevisiae* to grow without making sphingolipids encodes a protein that resembles an *Escherichia coli* fatty acyltransferase. J Biol Chem **268:** 22156–22163

Oo K-C, Huang AHC (1989) Lysophosphatidate acyltransferase activities in the microsomes from palm endosperm, maize scutellum, and rapeseed cotyledons of maturing seeds. Plant Physiol **91:** 1288–1295

Proudfoot NJ, Brownlee GG (1976) 3′ Noncoding region sequences in eukaryotic messenger RNA. Nature **263:** 211–214

Rosenberg AH, Lade BN, Chui D, Lin S, Dunn JJ, Studier FW (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56:** 125–135

Rosner MR, Robbins PW (1982) Separation of glycopeptides by high performance liquid chromatography. J Cell Biochem **18:** 37–47

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Stymne S, Stobart AK (1987) Triacylglycerol biosynthesis. *In* PK Stumpf, EE Conn, eds, The Biochemistry of Plants, Vol 9: Lipids. Academic Press, New York, pp 175–214

Taylor DC, Magus JR, Bhella R, Zou J, MacKenzie SL, Giblin EM, Pass EW, Crosby WL (1992) Biosynthesis of triacylglycerols in *Brassica napus* L. cv Reston. Target: Trierucin. *In* SL MacKenzie, DC Taylor, eds, Seed Oils for the Future. American Oil Chemists' Society Press, Champaign, IL, pp 77–102

Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA **76:** 4350–4354

von Heijne G (1990) The signal peptide. J Membr Biol **115:** 195–201

von Heijne G, Nishikawa K (1991) Chloroplast transit peptides: the perfect random coil? FEBS Lett **278:** 1–3

Yadav NS, Wierzbicki A, Aegerter M, Caster CS, Perez-Grau L, Kinney AJ, Hitz WD, Booth R Jr, Schweiger B, Stecca KL, Allen SM, Blackwell M, Reiter RS, Carlson TJ, Russell SH, Feldman KA, Pierce J, Browse J (1993) Cloning of higher plant ω-3 fatty acid desaturases. Plant Physiol **103:** 467–476