Cloning, heterologous expression and biochemical characterization of plastidial sn-glycerol-3-phosphate acyltransferase from *Helianthus annuus*.

Miriam Payá-Milans, Mónica Venegas-Calerón*, Joaquín J. Salas, Rafael Garcés and Enrique Martínez-Force.

Biochemical characterization of recombinant GPAT showed an oleate specific isoform in sunflower plastids, concomitant with plant resistance to chilling temperatures.
Cloning, heterologous expression and biochemical characterization of plastidial sn-glycerol-3-phosphate acyltransferase from *Helianthus annuus*.

Miriam Payá-Milans, Mónica Venegas-Calerón*, Joaquín J. Salas, Rafael Garcés and Enrique Martínez-Force.

Instituto de la Grasa, CSIC, Carretera de Utrera Km1, 41013-Sevilla, Spain.

*To whom correspondence should be addressed

Address Correspondence to: Mónica Venegas-Calerón, Instituto de la Grasa (CSIC), Carretera de Utrera Km1, 41013, Sevilla, Spain.

Tlf: +34 954611550-272

Fax: +34 954616790

Email: mvc@ig.csic.es
Abstract

The acyl-[acyl carrier protein]:sn-1-glycerol-3-phosphate acyltransferase (GPAT; E.C. 2.3.1.15) catalyzes the first step of glycerolipid assembly within the stroma of the chloroplast. In the present study, the sunflower (*Helianthus annuus*, L.) stromal GPAT was cloned, sequenced and characterized. We identified a single ORF of 1344 base pairs that encoded a GPAT sharing strong sequence homology with the plastidial GPAT from *Arabidopsis thaliana* (ATS1, At1g32200). Gene expression studies showed that the highest transcript levels occurred in green tissues in which chloroplasts are abundant. The corresponding mature protein was heterologously overexpressed in *Escherichia coli* for purification and biochemical characterization. In vitro assays using radiolabelled acyl-ACPs and glycerol-3-phosphate as substrates revealed a strong preference for oleic versus palmitic acid, and weak activity towards stearic acid. The positional fatty acid composition of relevant chloroplast phospholipids from sunflower leaves did not reflect the in vitro GPAT specificity, suggesting a more complex scenario with mixed substrates at different concentrations, competition with other acyl-ACP consuming enzymatic reactions, etc. In summary, this study has confirmed the affinity of this enzyme affinity, which would partly explain the resistance to cold temperatures observed in sunflower plants.

Keywords:

*Helianthus annuus*; sunflower; substrate specificity; glycerol-3-phosphate acyltransferase; plastids.
**Abbreviations:**

ACP: acyl-carrier protein  
DAF: days after flowering  
DAG: days after germination  
DGDG: digalactosyldiacylglycerol  
ER: endoplasmic reticulum  
EST: expressed sequence tag  
G3P: glycerol-3-phosphate  
GPAT: \( sn-1 \) glycerol-3-phosphate acyltransferase  
*HaPLSB*: *Helianthus annuus* chloroplast GPAT gene  
LPA: lysophosphatidic acid  
MGDG: monogalactosyldiacylglycerol  
PA: phosphatidic acid  
PC: phosphatidylcholine  
PG: phosphatidylglycerol  
SQDG: sulfoquinovosyldiacylglycerol
1. Introduction

Plants are characterized for being photosynthetic organisms with most of their photosynthetic apparatus embedded in a lipid bilayer. The photosynthetic membranes are largely composed of monogalactosyldiacylglycerol (MGDG), digalactosyldiacylglycerol (DGDG), sulfoquinovosyldiacylglycerol (SQDG) and phosphatidylglycerol (PG) (Moreau et al., 1998). The diacylglycerol backbones required for the biosynthesis of these lipids are provided by two pathways: the eukaryotic pathway, through which glycerolipids are imported from the endoplasmic reticulum (ER) to chloroplasts, mainly in the form of phosphatidylcholine (PC); and the prokaryotic pathway, through which glycerolipids are synthesized de novo within chloroplasts (Ohlrogge and Browse, 1995). Acyltransferases (EC 2.3.1.x) catalyze essential reactions in the biosynthesis of phospholipids and triacylglycerols. Even though the mechanism of catalysis is similar for these acyltransferases, they differ in target substrate and specificity. Glycerol-3-phosphate acyltransferases (GPATs) and lysophosphatidic acid acyltransferases (LPAATs), with \( sn-1 \) and \( sn-2 \) positional specificities respectively, form the first step in the synthesis de novo of glycerolipids, yielding their precursor phosphatidic acid. The high degree of fatty acid specificity of \( sn-2 \) acyltransferases for 16:0 in chloroplasts and 18:1 in the ER may serve as a useful indicator of the molecular species proceeding from each of these two pathways (Ohlrogge and Browse, 1995). Plants can be classified based on the presence or absence of \( cis-7,10,13 \)-hexadecatrienoic acid (16:3) in the \( sn-2 \) position of MGDG (Mongrand et al., 1998). In 16:3 plants, both pathways are functional and produce 18:3/16:3-MGDG in chloroplasts, while in 18:3 plants the biosynthesis of galactolipids and SQDG by the prokaryotic pathway is truncated, and the backbones are transported from the ER-localized PC moiety probably in the form of phosphatidic acid (Roston et al., 2012). Galactolipids and SQDG are derived from both ER and plastid sources, while plastid PG is exclusively synthesized by the prokaryotic pathway (Kunst et al., 1988; Xu et al., 2006) and constitutes the main phospholipid in thylakoid membranes (Moreau et al., 1998).
Glycerol-3-phosphate acyltransferases (GPATs; E.C. 2.3.1.15; PlsB) are enzymes that catalyze the acylation of a molecule of glycerol-3-phosphate (G3P) at the sn-1 position to produce lysophosphatidic acid (LPA). They are found in both soluble and membrane-bound forms in higher plants (Bertrams and Heinz, 1976). Membrane-bound GPATs are mainly involved in the synthesis of cutin, suberin and storage lipids, and they are localized in the ER and mitochondrial membranes (Yang et al., 2012). The soluble forms are found in chloroplasts (green tissues) and plastids (sink tissues), and they participate in the prokaryotic biosynthesis of phospholipids, sulfolipids and galactolipids (Ohlrogge and Browse, 1995). The fatty acid substrates of these enzymes are synthesized within the chloroplast, yielding 16:0-ACP, 18:0-ACP and 18:1-ACP, which can be used by either the soluble GPATs or hydrolyzed by acyl-ACP thioesterases. Following such hydrolysis, the acyl moieties are exported out of the plastids in the form of acyl-CoAs, which are then incorporated into glycerolipids via ER-based pathways (Sanchez-Garcia et al., 2010).

Several different approaches have been used to characterize stromal GPAT, the earliest using chloroplasts or leaves from pea (Pisum sativum), spinach (Spinacea oleracea) and squash (Cucurbita moschata) to obtain enzyme-enriched fractions by ion-exchange chromatography (Bertrams and Heinz, 1981; Frentzen et al., 1983; Frentzen et al., 1987). Both acyl-CoAs and acyl-ACPs were identified as substrates of these enzymes, which were generally more specific for their physiological substrates (i.e.: acyl-ACPs). The first cDNA nucleotide sequence of a plastidial GPAT was published in 1988 (Ishizaki et al., 1988); it was cloned from squash leaves and later used for the first crystallographic three-dimensional structure description of a recombinant GPAT (Turnbull et al., 2001). In enzymatic studies two different forms of GPATs were depicted: selective GPAT, targeting 18:1 over 16:0 fatty acids (Frentzen et al., 1983); and non-selective GPAT, targeting 16:0 and 18:1 fatty acids at comparable rates (Cronan and Roughan, 1987). Research suggests a correlation between the ability of some plants to tolerate cold temperatures and the presence of a certain form of GPAT. Thus, those carrying the nonselective form of the enzyme are usually chill-sensitive plants, whereas those with selective
forms of the enzyme are usually chill-tolerant (Tamada et al., 2004; Zhu et al., 2009). This is due to the higher capacity of cis unsaturated phosphatidylglycerol (PG) to prevent membrane phase transitions that may take place with saturated PG species at low temperatures (Nishida and Murata, 1996; Sun et al., 2011). An analogous mechanism has been described for high temperature stress, with increased thermotolerance associated with an increase in the level of saturation of thylakoid membrane lipids (Sui et al., 2007a; Yan et al., 2008).

Here, we describe the identification of a cDNA sequence encoding the plastidial GPAT from sunflower and the properties of this recombinant sunflower GPAT when expressed in *E. coli*, which includes a high degree of specificity for oleate. Based on our findings, we discuss the role of this enzyme in the fatty-acid profile of polar lipids in sunflower leaves.
2. Results and discussion

2.1. Identification of the plastidial GPAT from sunflower

A BLAST search with the plastidial GPAT sequence from *Arabidopsis* (ATS1) returned one relevant EST corresponding to its sunflower homologue. Based on this result, an internal fragment was cloned by PCR and sequenced. Primers for PCR were designed within this region and using the RACE technique, a cDNA clone of 1751 bp was obtained that contained an ORF of 1344 bp and UTR regions 74 bp upstream and 333 bp downstream. The gene was designated *HaPLSB*, given its homology with the *Escherichia coli* GPAT encoded by the *plsB* gene.

GPAT is synthesized as a 447 amino acid precursor protein, yet once inside the chloroplast the mature protein presumably lacks the transit peptide located at the N-terminus. Previous studies predicted the cleavage site of GPATs in various plants (Murata and Los, 1997; Slabas et al., 2002), although producing no clear consensus. However, on the basis of TargetP results and previous findings (Slabas et al., 2002), a transit peptide of 57 amino acids was predicted for sunflower GPAT. An alignment of the plastidial GPAT protein sequences from sunflower and representatives of distinct photosynthetic taxa (dicot, monocot, moss and green algae) highlighted the strong conservation of these proteins during evolution (Figure 1). For example, the acyltransferase from the green alga *Chlamydomonas reinhardtii* shares 33% identity and 50% sequence similarity with *Helianthus annuus* stromal GPAT. The amino terminal regions correspond to the transit peptides and present major changes compared with the rest of the sequences to the point that the sequences are totally dissimilar. The cores of the enzymes are highly conserved indicating essential motifs that are required to maintain its activity. Some of these conserved residues determine the substrate binding sites or the catalytic center and will be discussed in the next subsection.

A phylogenetic tree was constructed for the deduced amino acid sequence of *HaPLSB* and for GPAT protein sequences from chloroplasts in a variety of eukaryotic organisms, including
diverse taxa from algae to higher plants (Supplementary data, Figure 1). Since no identifiable homologues exist in cyanobacteria, GPATs from *Chlamydiae* species were used as an outgroup to root the tree. The short distance at the root of the tree suggests a very similar ancestral form of the protein in both prokaryotes and eukaryotes, probably resulting from horizontal gene transfer. The branching pattern of flowering plant families, with very close nodes, is indicative of their rapid evolution. Despite the long period of evolutionary time since the divergence of green algae and higher plants, stromal GPAT has not undergone substantial evolutionary changes, also viewed in the alignment.

2.2. Analysis of enzyme structure based on the GPAT sequence

The tertiary structure of recombinant GPAT from squash chloroplasts, with which *HaPLSB* showed a high degree of similarity (67.6% identity), has been determined (Turnbull et al., 2001). Based on this structure, we modeled the tertiary structure of *HaPLSB* in silico and we predicted sunflower GPAT to be composed mainly of alternating $\alpha/\beta$ secondary structures. The *HaPLSB* model contains 15 $\alpha$-helices and 9 $\beta$-strands, which account for 46% and 13% of the sequence, respectively (the relative positions of these secondary structure elements are shown in the polypeptide sequence: Figure 2).

The modeled 3D structure is composed of two domains: a helical domain forming a four-helix bundle (residues 82-154) and an alternating $\alpha/\beta$ domain (residues 162-445) linked by a loop region (residues 155-161: Figure 3A). The $\alpha/\beta$ domain is folded into parallel/antiparallel $\beta$-strands surrounded by the remaining 11 $\alpha$-helices. Pockets and tunnels that may bind substrates and regulators are visible on the protein surface (Figure 3B and C). One pocket (surrounded by the blue patch in Figure 3B) is highly conserved among chloroplast GPATs and it contains a positively charged surface patch formed by residues His217, Lys271, His272, Arg313 and Arg315. This pocket is predicted to contain the G3P binding site, which would attract the phosphate moiety to the basic surface patch and then accommodate the molecule in the pocket at position *sn*-1, located near the catalytic residues His217 and Asp222. These residues, which
belong to the H(X₄)D motif (found in most G3P acyltransferases and that is important for catalytic activity), are located near a predicted hydrophobic patch (highlighted in yellow in Figure 3B) and the positive cluster that bind the acyl chain and G3P molecule, respectively. Considering that a L261P mutation in squash (Slabas et al., 2002; homolog L339 in sunflower shown in brown) in this hydrophobic patch changed enzyme selectivity, specificity may be related to flexibility of the acyl chain when binding to the enzyme surface. Docking experiments using G3P and oleyl-CoA (18:1-CoA) as acyl donor (Figure 3C and D) corroborates the positioning of the G3P molecule as described, while the acyl group fits in a long tunnel interacting with residues A221, P223, Y243, V244, V249, C255, P258, S259, L264, W306 and M360. Docking experiments of palmitoyl-CoA (16:0-CoA) and stearoyl-CoA (18:0-CoA) gave similar results to oleyl-CoA (data not shown), concordant to the suggestion that selectivity may occur prior to entering the tunnel.

2.3. HaPLSB expression

Quantitative RT-PCR was used to study the expression of HaPLSB, normalizing the data to that of actin as a housekeeping gene. In accordance with the role of its protein product in chloroplast lipid synthesis, the HaPLSB gene was predominantly expressed in green tissues with little expression in sink tissues (Figure 4A). Its expression was enhanced during cotyledon development, consistent with the elevated rate of de novo membrane lipid biosynthesis in the chloroplast during the early stages of plant growth, and was maintained at high levels in mature leaves. The expression profile obtained was similar to that of the corresponding plastidial GPAT from Arabidopsis (ATS1; Nishida et al., 1993), with the highest levels of expression found in cotyledons and early leaves (Figure 4B; data from Schmid et al., 2005). One significant difference in the expression of this gene was observed during the development of Arabidopsis and sunflower seeds, reflecting the maintenance of active chloroplasts during seed development in Arabidopsis and their continuous decay in sunflower in favor of oil biosynthesis and accumulation. This difference was expectable due to Arabidopsis producing green seeds that are
able to do photosynthesis during most of their developmental period, whereas sunflower yield white seeds that are sink organs and lack any photosynthetic machinery.

2.4. Heterologous expression in E. coli

A heterologous expression system was prepared by cloning the predicted HaPLSB mature protein coding region of 1140 nucleotides into a pQE-80L inducible expression vector (primers shown in Table 1). This recombinant pQE80::HaPLSB expression vector was used to transform the competent E. coli XL1-Blue strain. Cultures of transformed E. coli cells were grown with IPTG induction for the over-expression of the recombinant HaPLSB gene and without induction as control. The fatty acid composition of the host in both uninduced and induced cultures was analyzed (Table 2).

The differences encountered in the fatty acid composition of both cultures were an indicative of the successful production of an active recombinant enzyme that was altering the fatty acid profile in the host bacteria. Induced bacteria showed increased unsaturated fatty acid proportion (54% compared to 50% in control) concomitant with a decrease of cyclopropane fatty acids (11% compared to 14% in control), which reflects the higher membrane production due to a more rapid growth rate. In this case, the fatty acid species with 16 carbon atoms experienced more relevant changes, varying their proportion of saturated and unsaturated, by diminishing and increasing 9% and 10% respectively, taking into account that the 17:0Δ derives from 16:1.

The flux of intermediates into each branch of saturated and unsaturated fatty acids biosynthesis in bacteria is regulated by their long-chain products, and then the increase of unsaturated fatty acids in induced cultures may respond to their consumption by HaPLSB (Rock and Jackowski, 2002). The lipid content of induced cultures was also slightly lower than that of uninduced cultures, due to the elongation of the exponential phase in GPAT overproducing cultures instead of the decrease in GPAT activity and accumulation of intracellular acyl-ACP when reaching the stationary phase (Heath et al., 1994). Thus, the over-production of the exogenous HaPLSB in E.
coli cells alters the composition and levels of the pool of acyl-ACP, leading to a variation in cell fatty acid profile.

2.5. In vitro substrate specificity of recombinant GPAT

Specificity assays were performed in vitro with the recombinant HaPLSB enzyme purified from *E. coli* (Figure 5) to identify the form (selective or nonselective) of the enzyme expressed by this sunflower line. Radioactive 16:0, 18:0 and 18:1 acyl-ACPs were used as separate substrates, and activity was measured in function of the production of radiolabelled lysophosphatidic acid (LPA) from unlabelled G3P. The reaction mixture contained varying concentrations of acyl-ACP and G3P (with 5 mg/mL BSA), at a pH of 7.4, that of the plastid stroma (Shen et al., 2013). According to previous calculations in spinach chloroplasts, the physiological concentration of acyl-ACP is around 1-2 µM (Soll and Roughan, 1982) while the stromal G3P concentration in *Amaranthus* is estimated at 450-620 µM (Cronan and Roughan, 1987).

At identical acyl-ACP concentrations, oleoyl-ACP was the most active acyl donor to sunflower GPAT (Figure 6A), and at the predicted physiological acyl-ACP concentrations, the rates of 18:1-LPA, 16:0-LPA and 18:0-LPA synthesis were 0.2, 0.1 and 0.01 nkat per mg protein, respectively. In these conditions, the rate of 18:1-ACP incorporation was twice that of 16:0-ACP, suggesting that the products contained mainly 18:1, some 16:0 and traces of 18:0 in position sn-1.

GPAT displayed a lower $K_M$ value for 16:0-ACP than for 18:1-ACP at the predicted physiological concentration, and a 9-fold lower $V_{max}$ (Table 3). By contrast, the $K_M$ for 18:1-ACP was 4-fold higher than the predicted stromal 18:1-ACP concentration. The $k_{cat}/K_M$ ratio was used to compare the catalytic efficiencies of HaPLSB towards its different substrates, revealing greater specificity and catalytic efficiency of HaPLSB towards 18:1 than 16:0, as expected for a chilling-resistant plant (Zhu et al., 2009). Moreover, the low activity levels
towards 18:0-ACP suggest limited incorporation of this fatty acid into LPA via the intraplastidial GPAT enzyme.

Due to the better availability and stability, as well as not leading to saturating levels during the assays, palmitoyl-ACP was used as an acyl donor to study the kinetic parameters of G3P (Figure 6B). The $K_m$ value for this substrate was 659 µM, which was in the same order of magnitude reported for *Amaranthus* (Cronan and Roughan, 1987), suggesting similar conditions in sunflower chloroplasts. As expected, the turnover under saturating conditions of G3P was similar to that obtained under saturating conditions of 16:0-ACP.

2.6. Effect of pH on HaPLSB activity

The pH inside the chloroplast ranges from 7.4 at night to 8.0 during the day and it regulates the activity of many enzymes. We studied the response of the *HaPLSB* enzyme to pH using palmitoyl-ACP as an acyl donor, observing maximum *HaPLSB* activity at pH 6.9 that decreased significantly up to pH 8.0. This pH curve (Figure 6C) is similar to that reported for the pea chloroplast GPAT (Bertrams and Heinz, 1981), for which the optimal pH value is 7.4. A recent study reported that the activity of the plastidial GPAT from eight different plant species at pH 7.4 is lower than at pH 8.0 (Zhu et al., 2009), showing that plastidial GPAT activity responds to pH changes.

2.7. Lipid analysis of sunflower leaves

After characterizing sunflower GPAT, we analyzed the species and fatty acid composition from sunflower leaf chloroplast lipids, paying special attention to MGDG, DGDG and PG. Unlike other organelles, chloroplast membranes are rich in galactolipids and contain low levels of phospholipids (Moreau et al., 1998). Thus, some phospholipid species like phosphatidyl ethanolamine (PE), which accounts for 20 to 40% of the total species present in extraplastidial membranes, are absent in chloroplast ones. The polar lipid extract prepared from sunflower leaves contained 4% of PE (Table 4A) and 7% of phosphatidic acid (PA), which are probably
derived from the ER and the plasma membranes (with an estimated PA level of 5-20%). As expected, almost half of the lipids analyzed corresponded to MGDG, being the levels of sulfolipids below the 6% estimated for chloroplast membranes.

The sunflower belongs to a plant group known as 18:3 plants (Mongrand et al., 1998) in which only PG is exclusively synthesized inside the chloroplast, all other plastid lipids being synthesized via the eukaryotic pathway (see Introduction). To confirm this lipid profile, MGDG, DGDG and PG were isolated and their positions analyzed. These lipids were treated with lipozyme from *Mucor miehei*, which specifically liberates the acyl group in position sn-1. The PG purified, mainly produced via the prokaryotic pathway, provides an indication of the specificity of the plastid acyltransferase. Assuming that 18:2 and 18:3 are derived from an 18:1 precursor, and 16:1 from a 16:0 precursor, this phospholipid had 52% 18:1, 29% 16:0 and 19% 18:0 at position sn-1, and 18% 18:1, 77% 16:0 and 5% 18:0 in position sn-2. The sn-1 proportions do not appear to tally with the specificity of the *HaPLSB* enzyme. Previous reports usually measured very low levels of stearic acid at sn-1 position in leaf lipids with the exception of *Cucumis sativus* with 12% 18:0 (Murata et al., 1982; Whitaker, 1986). Regarding these data, the proportion of stearic acid in sunflower is relatively high, but the final PG composition must be influenced by the effect of intraplastidial substrate concentrations in the enzymatic activity, as well as other possible regulatory mechanisms. Moreover, differences in the degree of unsaturation of 16 carbon atoms molecules were detected between the two positions of the PG molecule. A similar finding was reported previously in tomato (Sui et al., 2007), where 16:0 desaturation occurs asymmetrically between positions sn-1 and -2, enriching the sn-1 position in saturated fatty acids and the sn-2 in delta 3-trans-hexadecenoyl (16:1(3t)).

Gallactolipids are major components of chloroplast membranes and play an active role during cold acclimation (Moellering et al., 2010). In sunflower, both MGDG and DGDG are synthesized using phosphatidic acid precursors imported from the ER (Roston et al., 2012) and we found that MGDG molecules contained mainly 18:1 in position sn-1, while DGDG also contained 16:0 and 18:0 (Table 4B). This asymmetry suggests a possible selection against the
galactosylation of 18:1/18:1 MGDG by DGDG synthase in favor of molecular species with other fatty acids. In both galactolipids, the proportion of 18:1 derivatives in sn-2 was over 90%, mainly in the 18:3 form. Being the galactolipids the most abundant membrane lipids in plastids, we suggest that the different fatty acid distribution in both MGDG and DGDG moieties facilitates membrane remodelling, as MGDG molecules are consumed during freezing tolerance to form oligogalactolipids and diacylglycerol (Moellering et al., 2010).
3. Conclusions

G3P acyltransferases play an essential role in membrane synthesis by catalyzing the first step in glycerolipid assembly. Enzymatic specificity determines membrane fluidity and in chloroplasts, the degree of unsaturation of PG molecules determines their ability to protect the photosynthetic apparatus from damage caused by high or low temperatures. Sunflower chloroplast GPAT displays a specificity profile that confers resistance to cold temperatures due to its specificity for unsaturated substrates as oleate. This characteristic is reflected in the profile of fatty acids in position sn-1 in leaf PG, the majority of which are oleoyl-derived. Concerning the nature of this enzyme, the sequence of this selective GPAT form is now available for future studies seeking to transform sensitive plants to overcome temperature sensitivity.

4. Experimental

4.1. Biological material and growth conditions

Plants from the standard sunflower line CAS-6 were cultivated in growth chambers at 25/15°C (day/night) with a 16 h photoperiod (300 µE m⁻² s⁻¹). RNA was extracted from 0.5 g of mature functional leaves from two-month-old plants to produce cDNA, as described previously (Sanchez-Garcia et al., 2010). The E. coli strain XL1-Blue (Stratagene, California) was used as a host for the expression of sunflower plastidial GPAT (HaPLSB), growing the bacteria at 37°C with vigorous shaking in LB medium (1% tryptone, 0.5% yeast extract, 1% NaCl). For plasmid selection, ampicillin (100 µg/mL) was added to the medium.

4.2. Cloning and sequencing of the sunflower GPAT gene

The sunflower EST database was searched against the plastid-located GPAT protein sequence from Arabidopsis thaliana (L.) Heynh. (ATS1, accession no. NP_174499) using the tBLASTn search tool (Altschul et al., 1990) of the NCBI. The EST with GenBank accession no. CD847532.1 was used to design primers. Using cDNAs from sunflower leaves as a template, an
internal fragment of 665 bp was amplified using the specific HaPLSB-F1 and HaPLSB-R1 primers (Table 1). All primers were synthesized by Eurofins MWG Operon (Germany). The 5’ cDNA end was amplified using the BD SMART™ RACE cDNA Amplification Kit (Clontech, Japan) and the specific internal primers HaPLSB-R1 and HaPLSB-R2 (Table 1). The 3’ end was amplified using the external oligonucleotide FA2Z (Table 1), complementary to the sequences incorporated during the initial cDNA synthesis, and the specific internal primers HaPLSB-F1 and HaPLSB-F2 (Table 1). The amplification products were cloned into the pMBL-T cloning vector (Dominion MBL, Spain), sequenced (Secugen, Spain) and the sequences were verified by comparison with other GPATs. Finally, using the specific HaPLSB-F and HaPLSB-R PCR primers (Table 1), the full HaPLSB gene (1344 bp) was amplified, cloned and sequenced, and the sequence was deposited at GenBank under accession no. HM490309.

4.3. Overexpression and purification of the HaPLSB protein in E. coli

The coding region of the putative mature HaPLSB protein was amplified with the SphIHaPLSB-F and KpnIHaPLSB-R primers (Table 1) and subcloned into the SphI-KpnI sites of the pQE-80L expression vector (Qiagen, Germany) to produce a fusion protein with a 6xHis tag at the amino terminus. Correct in frame insertion was verified by sequencing and the resulting construct was designated pQE::HaPLSB. E. coli XL1-Blue cells were transformed with pQE::HaPLSB or the empty vector (as a control) and grown in LB medium at 37ºC with shaking. Protein over-production and cell lysis were performed as described previously (Sanchez-Garcia et al., 2010). The HaPLSB protein was purified from the resulting soluble fraction using the His SpinTrap Kit (GE Healthcare, UK) and the enzyme was stored at -20ºC in glycerol-containing buffer. The concentration of the purified protein was quantified by the Bradford method (Bradford, 1976) using the Bio-Rad Protein Assay (Bio-Rad, California) and the proteins were visualized by SDS-PAGE (Sanchez-Garcia et al., 2010).

4.4. Synthesis of radioactive acyl-ACPs
Radiolabelled acyl-ACP substrates were synthesized from commercial [1-\(^{14}\)C] palmitic, [1-\(^{14}\)C] stearic and [1-\(^{14}\)C] oleic acids (ARC, Missouri; 55 Ci/mol), as described elsewhere (Sanchez-Garcia et al., 2010).

4.5. Determination of GPAT substrate specificity

Substrate specificity of GPAT was determined as described previously (Frentzen et al., 1987; Zhu et al., 2009) with some minor modifications. The reaction was performed in a total volume of 100 \(\mu\)L containing 0.25 M HEPES buffer (pH 7.4), 5.0 mg/mL BSA and 0.4 mM G3P, with variable amounts of radioactive acyl-ACP. During the acyl-ACP specificity assays, substrate concentrations ranged from 0.5 to 24 \(\mu\)M, while the concentration of G3P ranged from 50 to 800 \(\mu\)M, with 3 \(\mu\)M 16:0-ACP. To study the influence of pH, the system was buffered with a mixture of 0.1 M Tris, 0.05 M MES and 0.05 M acetic acid, with 2 \(\mu\)M 16:0-ACP. Assays began with the addition of 200 ng of the purified protein and after a 5 min incubation at 25ºC, the reaction was stopped by adding 100 \(\mu\)L acetic acid (0.15 M).

4.6. Lipid extraction and analysis

Lipids formed in the reactions were extracted as described previously (Ruiz-Lopez et al., 2010). Polar lipids were analyzed on TLC plates, previously dried and activated by heating in an oven at 120ºC for 1 h, using chloroform-methanol-acetic acid-water (85:15:10:3.5) as the solvent. The radioactive product was visualized by autoradiography on a Packard Instant Imager (Packard Instrument Co, Connecticut). LPA was scraped from the plate and the radioactivity incorporated was determined using a LS-6500 Multipurpose Scintillation Counter (Beckman Coulter, California) Ecoscint scintillant (National Diagnostics, Georgia).

4.7. Total lipid extraction from leaves

The total lipid fraction was extracted using modified conditions of the Hara and Radin (1978) method, from mature sunflower leaves (0.5 g) maintained at 80ºC in 6 mL isopropanol for 15 min. The tissue was then ground with sea sand and the lipids extracted after the addition of 9
mL of hexane and 7.5 mL of 6.7% Na₂SO₄. The organic phase was transferred to a fresh tube and the aqueous phase was re-extracted with 10 mL hexane-isopropanol (7:2). After combining the organic phases, the solvent was evaporated under vacuum in a rotary evaporator (VWR by IKA®, Germany). The lipid residue was dissolved in 3 mL of chloroform and different aliquots of this extract were used for to assay the polar lipids by HPLC and fatty acid analysis.

4.8. Isolation and lipolysis of polar lipids from chloroplasts

Total lipids were loaded onto a TLC plate coated with 0.15 M (NH₄)₂SO₄ and the polar lipids were separated with a mixture of acetone-toluene-water (70:23:6). The lipids were visualized by exposure to iodine vapor and identified using standards. MGDG, DGDG and PG were scraped off the plate and eluted from the silica with chloroform-methanol (1:1). The lipids were then dried under nitrogen gas and dissolved in acetonitrile. Lipolysis was performed in a final volume of 200 µL by adding an equal volume of 40 mM HEPES (pH 7.4) and 10 U of immobilized lipozyme from *Mucor miehei* (Sigma-Aldrich, Missouri) and incubating the mixture overnight at room temperature. Next, 1.4 mL chloroform-methanol (1:1) and 560 µL H₃PO₄ (0.2 M) in 1 M KCl was added to the reaction, mixed vigorously and centrifuged to separate the phases. The lower phase (containing the lipids) was removed and dried under nitrogen gas. The lipolytic products were separated by chromatography in two-stages: first in hexane/diethyl ether-acetic acid (40:60:1) to recover the non-polar lipids and then with chloroform-methanol-acetic acid-water (85:15:10:3.5) to obtain the polar lipids. The lipids were recovered from the TLC plate and used to prepare fatty acid methyl esters.

4.9. Preparation and analysis of fatty acid methyl esters

The fatty acids in transformed *E. coli* cells expressing the recombinant *HaPLSB* protein were analyzed. Cells taken from 50 mL of induced and non-induced cultures were harvested at OD₆₀₀ ~ 2, washed with distilled water and 3 mL of the methylation mixture (~1.25 M hydrogen chloride in methanol; Fluka, Germany), and 0.15 mg of the internal standard heptadecanoic acid (17:0) was added to the pellet. The samples were methylated for 1 h at 80°C and heptane (1 mL)
was added after cooling to extract the methyl esters. The upper phase was transferred to a fresh tube and washed with 2 mL Na$_2$SO$_4$ (6.7%) and the solvent was then evaporated under nitrogen gas. The methyl esters were finally re-suspended in 200 µL heptane.

Fatty acid methyl esters from lipolysis were obtained by heating the lipid residues dissolved in 2 mL methanol-toluene-sulfuric acid (170:30:5) at 80ºC for 1 h. After cooling, 2 mL Na$_2$SO$_4$ (6.7%) and 3 mL hexane were added, the solution was mixed vigorously and the upper layer was transferred to a fresh tube and the solvent was evaporated under nitrogen gas. The residues containing methyl esters were dissolved in 300 µL heptane.

Methyl esters were analyzed by gas-liquid chromatography (GLC) using a Hewlett-Packard 6890 gas chromatograph (California) with a Supelco SP-2380 fused-silica capillary column (length, 30 m; i.d., 0.25 mm; film thickness, 0.20 µm: Supelco, Pennsylvania). Hydrogen was used as the carrier gas at 28 cm/s, while the temperature of the flame ionization injector and detector was set at 200ºC, the oven temperature at 170ºC, and the split ratio was 1:50.

4.10. Quantitative PCR

Sunflower cDNAs from vegetative tissues and seeds at different developmental stages were subjected to quantitative real time PCR (qRT-PCR) using a pair of specific primers (QHaPLSB-F and QHaPLSB-R, Table 1) and SYBR Green I (QuantiTect® SYBR® Green PCR Kit: Qiagen, Crawley, UK) using a MiniOpticon system (Bio-Rad, California). The reaction mixture was heated to 95ºC for 15 min before subjecting it to 40 PCR cycles of 94ºC for 15 s, 58ºC for 30 s and 72ºC for 30 s, while monitoring the resulting fluorescence in each cycle. The calibration curve was drawn up using sequential dilutions of pQE::HaPLSB and the comparative expression between samples was calculated using the Livak method (Livak and Schmittgen, 2001). The sunflower actin gene HaACT1 (GenBank accession no. FJ487620) was used as an internal reference to normalize the amount of cDNA in each sample relative to the specific primer (QHaAct-F4 and QHaAct-R4, Table 1).
4.11. Bioinformatics

GPAT protein sequences from public databases were aligned to identify regions of homology using the ClustalX v2.1 program (Larkin et al., 2007). The sequences were aligned using BioEdit (Hall, 1999) and a phylogenetic tree was constructed using MEGA 5.0 software (Tamura et al., 2011). The amino terminal transit peptide was predicted using TargetP 1.1 (Emanuelsson et al., 2000). The tertiary structure of sunflower GPAT was modelled using the SWISS-MODEL server (Guex and Peitsch, 1997) based on the structure of *Cucurbita moschata* (squash) GPAT (accession no. BAB17755) from residues 82 to 445. Molecular docking experiment was performed using Vina with AutoDock Tools and PyMOL (Trott and Olson, 2010).
Acknowledgements

Thanks are due to A. Gonzalez-Callejas and B. Lopez-Cordero for skillful technical assistance. This work was supported by the MINECO and FEDER [project AGL2011-23187].
Figure legends and Tables

**Figure 1.** Alignment of deduced *HaPLSB* amino acid sequence with PLSB proteins from various species. Identical amino acids are shaded in black. The reference sequences are: *HaPLSB* (HM490309), *Helianthus annuus*; ATS1 (GI: 15222600), *Arabidopsis thaliana*; *CmPLSB* (BAB17755), *Cucurbita moschata*; *OsPLSB* (GI: 115483650), *Oryza sativa*; *PpPLSB* (GI: 168037616), *Physcomitrella patens* subsp. *patens*; *CrPLSB* (GI: 159473711), *Chlamydomonas reinhardtii*.

**Figure 2.** Alignment of *HaPLSB* with squash (*Cucurbita moschata*) GPAT. The secondary structure elements from modeled sunflower GPAT are represented by cylinders (α helices) and arrows (β strands) above the aligned sequences. The black arrow indicates the putative processing site of the transit peptide. The H(X)₄D motif conserved among glycerol-3-phosphate acyltransferase enzymes is underlined. Identical amino acids are shaded in black.

**Figure 3.** Modelled tertiary structure of *HaPLSB*. (A) Ribbon diagram. Helices and strands are colored in succession from blue at the amino terminus to red at the carboxyl terminus. (B) Surface computing. Predicted patches to attract and bind both the G3P and acyl molecules are highlighted in blue and yellow respectively; L339 marked in brown may play a role in acyl selectivity. Substrate binding site is framed. (C) Framed binding site with 20% transparency showing the best scored positions for ligands G3P (sticks) and oleyl-CoA (spheres) predicted by docking experiments. (D) Rotated view of the cluster interacting with the docked G3P molecule at maximum distance of 4 Å. The atom type color scheme is the following: carbon, white; oxygen, red; nitrogen, blue; phosphate, orange.

**Figure 4.** Expression of plastid GPAT gene in *Helianthus annuus* and *Arabidopsis thaliana* tissues at different developmental stages. (A) *HaPLSB* gene expression data normalized with *HaACT1*. (B) *Arabidopsis thaliana ATS1* gene data from microarray [33]. DAF: days after flowering; DAG: days after germination; Stage 3: mid-globular to early heart embryos; Stage 4: early to late heart embryos; Stage 5: late heart to mid-torpedo embryos; Stage 6: mid to late torpedo embryos; Stage 7: late torpedo to early walking-stick embryos; Stage 8: walking-stick
to early curled cotyledons embryos; Stage 9: curled cotyledons to early green cotyledon embryos. Leaves 1 + 2 are aged 7 DAG. The data in panel A represent the mean values ± SD of three independent seedling samples.

**Figure 5.** Coomasie Brilliant Blue stained SDS-PAGE of recombinant sunflower GPAT purified by affinity column after heterologous expression. Lane 1, soluble fraction of pQE80::HaPLSB transformed *E. coli* lysate. Lane 2, purified His-tagged HaPLSB protein. *HaPLSB* represents 6.9% of the soluble fraction and 84.0% after purification, as estimated using ImageJ (http://imagej.nih.gov/ij/; Schneider et al., 2012).

**Figure 6.** *HaPLSB* was assayed to determine the specificity and kinetic parameters using 16:0 (○), 18:0 (□) and 18:1 (△) acyl-ACPs as substrates. Enzymatic activity is expressed as the amount of radiolabelled LPA formed in reactions as a function of the acyl-ACP concentration (A), G3P concentration (B) or pH (C). Values are means ± S.D.
Table 1. Oligonucleotides used in this study. * Restriction sites are indicated in bold.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence 5' → 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>HaPLSB-F1</td>
<td>AGAGAGGTAGAAGCTGGAGCAC</td>
</tr>
<tr>
<td>HaPLSB-R1</td>
<td>TGATTTGTGACGGATCGAG</td>
</tr>
<tr>
<td>FA2Z</td>
<td>AACTGGAAGAAATTCGCGG</td>
</tr>
<tr>
<td>HaPLSB-F2</td>
<td>TGAACGATGGTTCCTGAGCTTG</td>
</tr>
<tr>
<td>HaPLSB-R2</td>
<td>TGCTTCGGATTTGGTGTTTGAG</td>
</tr>
<tr>
<td>HaPLSB-F</td>
<td>TGTCGATTCTCCGATCTTCTCTTCTCC</td>
</tr>
<tr>
<td>HaPLSB-R</td>
<td>TCAACAGGTTGCAGCAATGAGACAC</td>
</tr>
<tr>
<td>pQE_HaPLSB-F</td>
<td>ATTTTGTGATGCAGCAGACGTTTTGAGACAC</td>
</tr>
<tr>
<td>pQE_HaPLSB-R</td>
<td>TGGTACCCCTGCTGATATACTAGTTCAACAGG</td>
</tr>
<tr>
<td>QHaPLSB-F</td>
<td>TGAACGATGGTTCCTGAGCTTG</td>
</tr>
<tr>
<td>QHaPLSB-R</td>
<td>TGATTTGTGACGGATCGAG</td>
</tr>
<tr>
<td>QHaAct-F4</td>
<td>GCTAACAGGAAAAGATGACTC</td>
</tr>
<tr>
<td>QHaAct-R4</td>
<td>ACTGGCATAAAGAGAAAGACG</td>
</tr>
</tbody>
</table>
Table 2. Fatty acid composition (mol%) of bacterial cells that do or do not produce the GPAT protein. Cells were harvested at OD = 2. The data correspond to average values (<10% variation). 16:0, palmitic acid; 16:1, palmitoleic acid; 17:0Δ, cis-9,10-methylenehexadecanoic acid; 18:0, stearic acid; 18:1, cis-vaccenate; 19:0Δ, cis-9,10-methyleneoctadecanoic acid.

<table>
<thead>
<tr>
<th>Fatty acid composition (mol%)</th>
<th>16:0</th>
<th>16:1</th>
<th>17:0Δ</th>
<th>18:0</th>
<th>18:1</th>
<th>19:0Δ</th>
<th>μg FA/mg cel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uninduced</td>
<td>48.17</td>
<td>19.07</td>
<td>13.03</td>
<td>1.07</td>
<td>17.57</td>
<td>1.09</td>
<td>4.79</td>
</tr>
<tr>
<td>Induced</td>
<td>44.14</td>
<td>24.83</td>
<td>10.70</td>
<td>1.57</td>
<td>18.26</td>
<td>0.50</td>
<td>4.23</td>
</tr>
</tbody>
</table>
Table 3. Apparent kinetic parameters calculated for recombinant sunflower chloroplast GPAT. a

Values were calculated using the Hill equation (adjusted R-squared > 0.95). Assays were performed in triplicate as described in the Materials and Methods.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$K_M$ a</th>
<th>$V_{max}$ a</th>
<th>$k_{cat}$</th>
<th>$k_{cat}/K_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:0-ACP</td>
<td>1.96</td>
<td>0.22</td>
<td>9.44 x 10^{-3}</td>
<td>4.82 x 10^{3}</td>
</tr>
<tr>
<td>18:0-ACP</td>
<td>6.12</td>
<td>0.16</td>
<td>6.74 x 10^{-3}</td>
<td>1.10 x 10^{3}</td>
</tr>
<tr>
<td>18:1-ACP</td>
<td>7.80</td>
<td>2.12</td>
<td>9.18 x 10^{-2}</td>
<td>1.18 x 10^{4}</td>
</tr>
<tr>
<td>G3P</td>
<td>659</td>
<td>0.25</td>
<td>1.10 x 10^{-2}</td>
<td>17</td>
</tr>
</tbody>
</table>
Table 4. Polar lipid composition of mature sunflower leaves. The data represent the mean of three separate measurements (<10% variation).

A. Glycerolipids (% of Total)

<table>
<thead>
<tr>
<th></th>
<th>PA</th>
<th>PC</th>
<th>PE</th>
<th>PG</th>
<th>PI</th>
<th>MGDG</th>
<th>DGDG</th>
<th>SQVDG</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>7.53</td>
<td>14.19</td>
<td>4.09</td>
<td>5.14</td>
<td>0.87</td>
<td>47.03</td>
<td>19.24</td>
<td>1.91</td>
</tr>
</tbody>
</table>

B. Positional distribution of fatty acids in galactolipids and phosphatidylglycerol (mol %) \(^a\)16:1(3t)

<table>
<thead>
<tr>
<th></th>
<th>16:0</th>
<th>16:1</th>
<th>18:0</th>
<th>18:1</th>
<th>18:2</th>
<th>18:3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGDG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sn-1)</td>
<td>7.8</td>
<td>0.1</td>
<td>8.8</td>
<td>2.0</td>
<td>7.7</td>
<td>73.7</td>
</tr>
<tr>
<td>(sn-2)</td>
<td>2.1</td>
<td>0.2</td>
<td>3.6</td>
<td>0.3</td>
<td>10.2</td>
<td>83.6</td>
</tr>
<tr>
<td>DGDG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sn-1)</td>
<td>25.8</td>
<td>0.1</td>
<td>15.4</td>
<td>3.4</td>
<td>5.2</td>
<td>50.1</td>
</tr>
<tr>
<td>(sn-2)</td>
<td>6.1</td>
<td>0.2</td>
<td>4.0</td>
<td>0.4</td>
<td>7.5</td>
<td>81.8</td>
</tr>
<tr>
<td>PG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(sn-1)</td>
<td>27.8</td>
<td>1.3</td>
<td>19.4</td>
<td>7.4</td>
<td>17.6</td>
<td>26.5</td>
</tr>
<tr>
<td>(sn-2)</td>
<td>48.1</td>
<td>28.7</td>
<td>5.4</td>
<td>0.4</td>
<td>9.6</td>
<td>7.7</td>
</tr>
</tbody>
</table>
References


substrate specificity, and mechanism of squash glycerol-3-phosphate (1)-acyltransferase. Structure 9, 347-353.


Appendix A. Supplementary data

Supplementary Figure 1. Phylogenetic study of chloroplast GPAT enzymes using GPATs from *Chlamydia*-like bacteria as outgroup to root the tree. The tree was constructed using the neighbor-joining method.