Sunflower (*Helianthus annuus*) Fatty Acid Synthase complex: Enoyl-[acyl carrier protein]-reductase genes.

Irene González-Thuillier¹, Mónica Venegas-Calderón^{1*}, Rafael Garcés¹, Penny von Wettstein-Knowles² and Enrique Martínez-Force¹

¹Instituto de la Grasa (CSIC), Av. Padre García Tejero, 4, 41012 Sevilla, Spain.

²Biology Department, Copenhagen University, Ole Maaløes Vej 5, 2200-Copenhagen N, Denmark.

Corresponding author:

Mónica Venegas-Calderón

Email: mvc@ig.csic.es

Tlf: +34 954611550

Fax: +34 954616790

Main Conclusion: Enoyl-[acyl carrier protein]-reductases from sunflower

Abstract

A major factor contributing to the amount of fatty acids in plant oils is the first steps of their synthesis. The intraplastidic fatty acid biosynthetic pathway in plants is catalysed by type II fatty acid synthase (FAS), formed by four enzymes. The last step is carried out by the enoyl-[ACP]-reductase, which reduces the dehydrated product of β -hydroxyacyl-[ACP] dehydrase using NADPH or NADH. To determine the mechanisms involved in the biosynthesis of fatty acid in sunflower (*Helianthus annuus*) seeds, two enoyl-[ACP]-reductase genes have been identified and cloned from developing seeds with 75% identity, *HaENR1* (GenBank HM021137) and *HaENR2* (HM021138). The two genes belong to the ENR1 and ENR2 families in dicotyledons, respectively. The genetic duplication originated most likely after the separation of di- and monocotyledons. RT-qPCR revealed distinct tissue specific expression patterns. Highest expression of *HaENR1* was in roots, stems and developing cotyledons whereas that of *HaENR2* was in leaves and early stages of seed development. Genomic DNA gel blot analyses suggest that both are single copy genes. *In vivo* activity of the ENR enzymes was tested by complementation experiments with the JP1111 *fabI(ts)* *E. coli* strain. Both enzymes were functional demonstrating that they interacted with the bacterial FAS components. That different fatty acid profiles resulted infers that the two *Helianthus* genes have different structures, substrate specificities and/or reaction rates. The latter possibility was confirmed by *in vitro* analysis with affinity purified heterologous expressed enzymes that reduced the crotonyl-CoA substrate using NADH with different $V_{\max}$.

Keywords: enoyl-[ACP]-reductase (ENR); fatty acid synthase (FAS); oil biosynthesis; *Helianthus annuus*; substrate specificity.

Introduction

The 'de novo' fatty acids biosynthesis in plants is a process which starts in the plastid where the multienzymatic complexes: Acetyl-CoA Carboxylase (ACCase) and Fatty acid Synthase (FAS) work co-ordinately (Harwood, 2005). We can find two kind of FAS complex in the nature. In mammals, a single gene encodes a polypeptide which is able to produce the fatty acids, called type I FAS (Smith et al., 2003). The type II is present in bacteria and plants and is a dissociated system, where each component is encoded for an independent gene that generates a single protein. These proteins catalyse an individual step in the biosynthetic pathway (Brown et al., 2006). The FAS II generates as final products, palmitoyl-acyl carrier protein (ACP) and stearoyl-ACP and is formed by four enzymes: Ketoacyl-[ACP]-synthase (KASI, II and III), Ketoacyl-[ACP]-reductase (KAR), Hydroxyacyl-[ACP]-hydrase (HAAD) and Enoyl-[ACP]-reductase (ENR) (Li-Beisson et al., 2013). Several studies have been done on the ketoacyl-[ACP]-synthase enzymes, but just a few on the rest of FASII enzymes. The last enzyme of the complex is involved in the reduction of the double bond between the C2 and C3 in the fatty acid during its elongation. It has been described two types of ENR: (I) showing absolute selectivity for a reducing agent (NADH always) or (II) being also able to use NADPH (Harwood, 1988; 1996; 2005). This enzyme has been extensively studied in bacteria, because it is one of the drug targets to fight against tuberculosis (Banerjee et al., 1994). Moreover some bacterial and plant ENRs have been crystallized to understand the relation structure-function, for instance Rafferty et al. (1995), who analysed the structure of *Brassica napus* ENR. Earlier was found that ENR normally works *in vivo* as a tetramer (Slabas et al., 1986). The protein belongs to the Short-Chain Dehydrogenase (SDR) superfamily, showing the typical Rossmann structure. The subunit of ENR is composed by seven β sheet (β 1- β 7) and seven α -helix

($\alpha 1$ - $\alpha 7$), and a determinate number of loops which can be with different lengths modifying the functional characteristics of the enzymes (White et al., 2005). In *Arabidopsis thaliana* and *Petunia hybrida* only has been found one gene encoding for ENR instead of two, as in *Nicotiana tabacum* and *Brassica napus*, allotetraploid species (deBoer et al., 1998). It was demonstrated that ENR from *B. napus* was able to functionally replace one of the components of the bacterial complex. The lipid analysis of the recombinant bacteria expressing the ENR gen from rape showed an apparent variation in the fatty acid profile, indicating that the enzyme carried out a limiting step for the unsaturated fatty acid synthesis; concretely the ratio 16:1/18:1 suffered an increase (Kater et al., 1994). In *Arabidopsis* the mutant *mod1* presents an amino acidic substitution in the *MOD1* gene (At2g05990) encoding for the ENR. That mutation caused a high decrease in the ENR enzymatic activity, and consequently a big reduction in the total lipids production in the mutant (Mou et al., 2000). The enzyme ENR has been biochemically characterized in four different plants, *Spinacia oleracea* (Shimakata et al., 1982a and c), *Brassica napus* (Slabas et al., 1986), *Carthamus tinctorius* (Shimakata et al., 1982b) and *Arabidopsis thaliana* (Dayan et al., 2008). Although the natural substrate of the enzyme is an acyl-ACP, it has been used another analogs as acyl-CoA or acyl-cisteamin, but as expected, the enzyme showed a higher *K_m* with the last two mentioned substrates (Rafferty et al., 1995). The purification of ENR in safflower and rape was carried out from seed extracts using an HPLC, in spinach was from leaves extracts, but the ENR from *Arabidopsis* was heterologous expressed and purified in *Escherichia coli*. The biochemical analysis in rape was done on heterogeneous conditions, but not in the case of safflower and spinach where the two isoforms of the enzyme (ENR I,II) could be purified homogenously. The *K_m* observed for the type I enzyme for the crotonyl-ACP was similar in most of the cases around 1

93 μM , on the other hand the *K_m* observed for analogous substrates as crotonyl-CoA was
94 higher and changed depending on the organism. The type II ENR from both plants was
95 not active with crotonyl joined to any of these cofactors, CoA or ACP, but showed
96 activity with longer acyl chains. Dayan et al. (2008) developed a phytotoxic analysis on
97 the heterologous expressed Arabidopsis ENR, where they found that high
98 concentrations of cyperin could generate a phenotype in the plant similar to the mutant
99 *mod1*. At some concentrations the production of linoleic and linolenic acids was
100 decreased in the plant.

101 In the present study we have looked for different alleles of ENR in sunflower,
102 which is a relevant gene in the fatty acid biosynthesis, to find the best to optimize the
103 total lipid content in developing seeds. Two ENR cDNAs from sunflower have been
104 isolated, cloned and sequenced. The expression levels of these genes in seed and
105 vegetative tissues were examined and genome copy number estimated by Southern blot.
106 The structural model of these proteins was predicted by homology showing some
107 differences between both isoforms. The sunflower proteins, heterologously expressed in
108 an *E. coli fabI(ts)* strain with a temperature sensitive ENR, were able to complement the
109 mutant under restrictive conditions allowing their activity to be tested *in vivo* by
110 studying their fatty acid profiles. The substrate specificity of purified recombinant
111 enzymes was assayed *in vitro*.

Materials and methods

Biological materials

Sunflower (*H. annuus* L.) wild-type line CAS-6 (Sunflower Collection of Instituto de la Grasa, CSIC, Seville, Spain) was grown and materials collected as described by González-Mellado et al. (2010) except for the following. The photon flux density was $250 \mu\text{mol m}^{-2} \text{s}^{-1}$. For the quantitative expression studies seed samples were collected every other day from 12 to 30 days after flowering (DAF) and from cotyledons 1-7 days after germination (DAG).

For the heterologous expression of *HaENR1* and *HaENR2* two different strains of *E. coli* were used as host, XL1-Blue (Stratagene, La Jolla, CA, USA) and JP1111 (Hfr(PO1), *galE45*(GalS), λ^- , *fabI392*(ts), *relA1*, *spoT1*, *thiE1*) from the *E. coli* gene bank. The bacteria were grown in SB medium (2.5% tryptone, 1.5% yeast extract, 0.5% NaCl, pH 7.5), and the liquid cultures of XL1-Blue were shaken vigorously at 37°C. When the plant genes were expressed in JP1111 they were grown either at the permissive temperature of 32°C or the restrictive temperature of 42°C. For plasmid selection 100 $\mu\text{g ml}^{-1}$ of ampicillin was added. The specific growth rate was calculated from the plot generated from the increase in log OD₆₀₀ with time. The mean of three independent experiments was used to calculate each value for the growth rate.

Cloning of cDNAs encoding sunflower ENR

Approximately 0.1 g of each tissue was ground in liquid N₂ with precooled sterile mortar and pestle. Total RNA was extracted using a Spectrum Plant Total RNA kit (Sigma-Aldrich, St. Louis, USA), and mRNA was isolated from total RNA (1 μg) using the GenElute mRNA Miniprep kit (Sigma-Aldrich). The corresponding cDNA

was synthesized using a Ready-To-Go T-Primed First Strand Kit (GE Healthcare Life Science, UK) from DNA-free mRNA.

The *A. thaliana* ENR1 protein sequence encoded by the At2g05990 gene was used to search sunflower expressed sequence tags (ESTs) publicly available in the NCBI database (<http://www.ncbi.nlm.nih.gov/>) to find putative mRNAs encoding ENR homologs using the TBLASTN algorithm (Altschul et al., 1997). ESTs corresponding to the 5' and 3' ends were identified. These ESTs were aligned using the ClustalX v.2.0.10 program (Larkin et al., 2007) to group those with similar or identical sequences. Polymerase chain reaction (PCR) fragments were amplified with two pairs of primers designed from these groups, F-ENR1/R-ENR1 and F-ENR2/R-ENR2 (Table 1; all the primers were synthesized by Eurofins MWG Operon, Ebersberg, Germany). The PCR fragments were cloned into the pMBL-T vector (Dominon, Rhode Island, USA), transformed into XL1-Blue, and several clones were sequenced on both strands by SECUGEN SL (Madrid, Spain). The identity of the clones was confirmed using the BLASTX algorithm (Altschul et al., 1997), distinguishing two different types coding for ENR isoforms *HaENR1* and *HaENR2* (1179 bp in both cases). These cDNA sequences were deposited in GENBANK under Accession numbers HM021137 |gi:317373786| and HM021138 |gi:317373788|, respectively.

cDNA and protein sequence analyses

Protein sequences homologous to the predicted sequences of *HaENR1* and *HaENR2* were identified using the BLASTP algorithm (Altschul et al., 1997). Alignment of the amino acid sequences, including the transit peptides, for plastid ENR proteins publicly available at GENBANK was performed using the ClustalX v.2.0.10 program with the default settings (Larkin et al., 2007). These entire alignments were

used to generate a phylogenetic tree based on the neighbour-joining algorithm (Saitou and Nei 1987), and the resulting ‘phenogram’ was drawn using the MEGA program 4.0.2 (Tamura et al., 2007). Transit peptides were identified through alignment with known plastid ENR sequences and using the network-based program TargetP V1.1 (Emanuelsson et al. 2000).

Genomic DNA Southern blot analysis

Genomic DNA from 2.5 g leaves of sunflower was isolated as described previously (González-Mellado et al., 2010). DNA samples were digested with different restriction enzymes (all from New England Biolabs, UK) and electrophoresed in a 0.8% agarose gel. The gel was soaked in 250 mM HCl for 30 min, then washed three times in distilled water and finally blotted onto a Hybond-N+ transfer membrane (GE Healthcare). The filter was probed with [α -³²P]dCTP-labelled *HaENR1* and *HaENR2* gene-specific DNA probes. The gene-specific probes with a size of 1232 bp and 1208 bp, respectively, were obtained by PCR amplification with the following pairs of primers: F-ENR1 plus R-ENR1XmaI for *HaENR1* and F-ENR2 plus R-ENR2XmaI for *HaENR2* (Table 1). Hybridization was performed in 0.2 mM sodium phosphate buffer (pH 7.2), 250 mM sodium dodecyl sulfate (SDS) and 1 mM ethylenediaminetetraacetic acid (EDTA) overnight at 65°C. The filter was washed twice in 2X saline-sodium citrate (SSC) buffer, followed by 0.1% SDS for 10 min at the same temperature. Images of radioactive filters were obtained as described by González-Mellado et al. (2010).

Real time quantitative PCR (RT-qPCR)

The cDNAs obtained from developing seeds and different vegetative tissues as described above were subjected to RT-qPCR with the gene-specific pair of primers: F-

ENR1qpcr plus R-ENR1qpcr for *HaENR1* and F-ENR2qpcr plus R-ENR2qpcr for *HaENR2* (Table 1) using SYBR Green (QuantiteTect™ SYBR® Green PCR Kit, Qiagen, Germany) in an MiniOpticon system (Bio-Rad, CA, USA). The reaction mixture was heated to 50°C for 2 min and then to 95°C for 15 min before subjecting it to 40 PCR cycles of: 94°C for 15 s, 58°C for 30 s and 72°C for 15 s while monitoring the resulting fluorescence. The calibration curve was done using sequential dilutions of developing seed cDNA. The Livak and Schmittgen (2001) method was applied to calculate comparative expression levels between samples. The sunflower actin gene *HaACT1* (GenBank Accession number FJ487620) was used as the internal reference to normalize the relative amount of cDNAs for all samples using primers QHaActin-F4 plus QHaActin-R4 (Table 1). Biological triplicates of each seed stage and vegetative tissues samples were obtained and analysed from three independent plants. Two technical replicates from each sample were also run in the PCR machine.

Expression of sunflower recombinant ENR proteins in *E. coli*

The used expression vector pQE80L contains an N-terminal 6x histidine affinity tag to facilitate protein purification. The genes *HaENR1* and *HaENR2* were amplified using the primer pairs F-ENR1SacI plus R-ENR1XmaI and F-ENR2SacI plus R-ENR2XmaI (Table1), respectively. The PCR products, corresponding to both genes without the peptide signal, contained the restriction sites for the enzymes *SacI* and *XmaI* at each end for cloning them into pQE80L after their digestion with such enzymes. The constructed vectors pQE-80L::*HaENR1* and pQE-80L::*HaENR2* were transformed into competent *E. coli* JP1111 and XL1-Blue strains, respectively. Heterologous expression of these sunflower genes and purification of the recombinant proteins was carried out as previously described (McGuire et al., 2001). Isopropyl β -D-1-thiogalactopyranoside

(IPTG) to a final concentration of 1 mM, and the cultures grown for 3 h at 37°C. Cells were harvested by centrifugation (15 min at 3200 g) and resuspended in 1 ml of purification buffer (25 mM TRIS-HCl, 300 mM NaCl, pH 7.8) plus 10 mM of DTT and 5 mM of EDTA. The cells were incubated at 4°C for 30 min with 1 mg ml⁻¹ lysozyme and then broken by sonication.

Protein gel electrophoresis and Western blots

Protein samples were combined with SDS-polyacrylamide gel electrophoresis (PAGE) loading buffer (63 mM TRIS-HCl pH 6.8, 10% (v/v) glycerol, 2% (p/v) SDS, 2.5-5% (v/v) β-mercaptoethanol and 0.0025% (p/v) bromophenol blue) and heated at 95°C for 5 min. The samples were separated by electrophoresis on 4–15% Mini-PROTEAN TGX Precast Gels (BioRad, California, USA) in MES buffer (Invitrogen California, USA) and the gels were fixed and stained with 0.1% Coomassie R-250 in 40% ethanol, 10% acetic acid. Molecular weight standards were obtained from GE Healthcare. Once the gels were run proteins were transferred using the kit Trans-Blot Turbo™ Mini PVDF Transfer Packs (Bio-Rad, California, USA) and a Trans-Blot® Turbo™ Transfer Starter System Bio-Rad with the program Mixed MW, according to the manufacturer. The polyvinylidene difluoride (PVDF) membrane was rinsed with the block solution [milk powder, 5% (w/v) in phosphate-buffered saline (PBS) wash solution] and shaken slowly for 1 h. The solution was discarded and the membrane washed four times for 15 min with PBS wash solution (0.1% (v/v) Tween 20, 137 mM NaCl, 10 mM sodium phosphate, 2 mM potassium phosphate, pH 7.4). For immune detection the blot was incubated with anti-polyHistidine (Sigma-Aldrich, Missouri, USA) (1/5000 dilution) overnight at 4°C. The membrane was washed with PBS wash solution as described above. The SuperSignal West Dura Extended Duration Substrate

Kit (Thermo Scientific Massachusetts, USA) was used to fuse the peroxidase to the antibody. The blot was revealed using Kodak autoradiography in the dark. To estimate the relative amount of *HaENR1* recombinant protein in purified fractions, stained protein gels were scanned and processed using ImageJ software (version 1.44; <http://rsb.info.nih.gov/ij/>). The concentration of protein was measured using a BCA Protein Assay Kit Pierce from Thermo Scientific (Massachusetts, USA).

Fatty acids analysis in *E. coli*

Cultures of 50 ml of *E. coli* JP1111 *fabI(ts)* carrying either pQE80L::*HaENR1* or pQE80L::*HaENR2* were grown at 32°C to OD₆₀₀ 0.8 when the expression of the sunflower proteins was induced by adding IPTG to a final concentration of 1 mM. The cultures were then placed at 32 or 42°C and grown for 2 h. Cells were harvested by centrifugation (15 min at 3200 g) and resuspended in water (1/10 from the initial volume), after two washes with water. 150 µg of heptadecanoic acid was added to the samples to serve as an internal standard for quantification. Extraction of the lipids, preparation of the methyl esters and their analysis by gas chromatography is described in González-Mellado et al. (2010).

In vitro characterization of *HaENR1* and *HaENR2*

The ability of *HaENR1* and *HaENR2* to reduce the crotonyl-CoA (Sigma, Missouri, USA) substrate was monitored by following the decrease in absorbance at 340 nm as NADH is oxidized, similar to the procedure described Dayan et al. (2008) with some modifications. Crotonyl-CoA rather than the true substrate crotonyl-ACP was used since the initial velocity of the reaction with the former was too rapid to measure in case of the *B. napus* ENR (Slabas et al., 1986). The purified proteins *HaENR1* or

HaENR2 were added at 0.4 $\mu\text{g ml}^{-1}$ or 8.3 $\mu\text{g ml}^{-1}$, respectively, to a buffer (17.9 mM potassium phosphate, 1.6 mM DTT, 300 μM NADH or NADPH (Sigma, Missouri) at pH 6.8 to give a final volume of 600 μL . The reaction was started by adding the crotonyl-CoA substrate at different concentrations between 20-250 μM . The blank lacked the substrate. The reaction was carried out at 28°C and monitored each minute using the spectrophotometer Ultrospec 3300pro (Amersham-Biosciences, Nueva York) and the software Datrys Life Science (2.2.0.0 Version). The extinction coefficient of NADH $\epsilon_{340}=6.22 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ (Shimakata et al., 1982a) was used to convert absorbance measurements to concentration of NADH. For each enzymatic assay three independent experiments were done with the purified protein.

Kinetic parameters analysis

The spectrophotometric data were analysed in MSEXcel (Microsoft, Washington) and OriginPro 8 (Originlab, Massachusetts). The non-linear function in Origin allowed us to find the kinetic parameters of the Hill Equation (Michaelis constant, K_m , and Maximum velocity, $V_{\max}$) that fitted best to the experimental data.

Modelling of the three dimensional structures of *HaENR1* and *HaENR2*

Homology modelling studies were performed as detailed in González-Mellado (2010). The structure of the *B. napus* ENR sequence with accession number UniProtKB P80030 (Rafferty et al., 1995) was used as a model for both sunflower sequences, being the most homologous ENR for which X-ray structure information is available in the database RCSB PDB (Berman et al., 2000) (PDB Entry: 1eno). This *BnENR* has 89 and 88% sequence identity over 299 residues to the *HaENR1* and *HaENR2* mature proteins, respectively.

Results

Cloning and sequence analysis of two sunflower enoyl-[ACP]-reductase cDNAs

Two groups of ESTs for ENR plastid enzymes in sunflower were identified on the basis of their homology to the Arabidopsis *ENR* gene (At2g05990). Using full length sequences for both groups two cDNAs were amplified from 15 DAF developing seed mRNA. These PCR products, consisting of 1261 base pairs (bp) for *HaENR1* and 1228 bp for *HaENR2*, were sequenced and found to be identical with the EST *HaENR1* and *HaENR2* sequences. The protein sequences, deduced from the cDNAs of both genes using bioinformatics tools consisted of 392 residues. Through alignment with known ENR sequences and using the network-based method TargetP V1.1 to identify chloroplast transit peptides (Emanuelsson et al., 2007), Ser46 and Ser75 of the sunflower ENR1 and ENR2 sequences, respectively, were the best candidates for the N-termini of the mature proteins (Fig. 1). In the case of *HaENR1* proteolytic processing of the transit peptide would produce a 346 residue protein with a putative molecular mass of 36.87 kDa and a pI of 8.43, while *HaENR2* would have 317 residues with a mass of 33.58 kDa and a pI of 5.36. These two sunflower amino acid sequences with 75% identity were compared with homologous proteins from different phylogenetic groups such as *A. thaliana*, *O. sativa* or *E. coli* (Fig. 1). This alignment shows, excluding the transit peptide, which is the less conserved region, identical or highly conserved domains between the different type II ENRs (Rafferty et al., 1995; White et al., 2005). In these conserved domains homology predicts the residues involved in the recognition of NADH as Gly100, Val101, Ala102, Gly106, Tyr107, Trp127, Leu163, Asp164, Ala165, Ser211, Leu212, Ala213, Asn214, Ala236, Leu261, Thr262, Tyr263, Tyr273, Lys281, Ala308, Gly309, Pro310, Leu311, Ser313, Arg314, Ala315 and Ala316 for *HaENR1* and Gly99, Ile100, Ala101, Gly105, Tyr106, Trp126, Leu162, Asp163,

Ala164, Ser210, Leu211, Ala212, Asn213, Ala235, Leu260, Thr261, Tyr262, Tyr272, Lys280, Ala307, Gly308, Pro309, Leu310, Ser312, Arg313, Ala314 and Ala315 for *HaENR2*. The predicted catalytic residues of *HaENR1* are Tyr263 and Lys281 assisted by Ser237 and the equivalents in *HaENR2* are Tyr262 and Lys280 and Ser236.

Whole sequenced genome availability of different organisms from many phylogenetic groups together with bioinformatic tools permits a better comparative analysis, identifying groups potentially sharing a common ancestor with the ENR from plants (Fig. 2). As the dendrogram shows, in some species from the Plantae Kingdom, such as *Vitis vinifera*, *Populus trichocarpa*, *Zea mays* or *O. sativa*, ENRs belong to two separate families, Type I and II, as in the case for *HaENR1* and *HaENR2* (Fig. 2) with a common phylogenetic origin. *HaENR1* has a $\geq 79\%$ identity with other Type I ENRs, such as *V. vinifera*, *P. trichocarpa* or *Capsicum annuum* and also with some ENRs from *B. napus* or *Z. mays*, which have not been classified as either Type I or II. The same relationship to these species is true for *HaENR2* except that identity is $\geq 77\%$. The dendrogram (Fig. 2) also reveals that plant ENRs are relatively close to their homologous from the Chlorophyta and Cyanobacteria.

Tertiary structure prediction of *HaENR* proteins

The sequences of *HaENR1* and *HaENR2* are 89 and 88% identical, respectively, to that of *BnENR* (Fig. 3). Using the known structure for the latter (UniProtKB P80030), structures for the sunflower ENRs were modelled. Both have a single domain formed by seven α helices ($\alpha 1$ - $\alpha 7$) and seven β sheets ($\beta 1$ - $\beta 7$) plus intervening loops, as is shown in the ribbon structures in Fig. 4a, b. They belong as does *BnENR* to the short-chain dehydrogenase/reductase superfamily (SDR), showing the Rossman fold (White et al., 2005) an open alpha/beta structure with twisted β sheets surrounded on both sides

by α helices. The distribution of α helices and β sheets is very similar in the sunflower proteins with a few minor differences in length. A three residue loop is present in the sunflower sequences between $\beta 2$ and $\alpha 2$ that is absent in *B. napus* ENR in which $\beta 2$ and $\alpha 2$ are contiguous. The residues involved in catalysis and NADPH/NADH binding are localized in the most highly conserved areas of the three proteins. The catalytic residues are distributed between the $\alpha 4$ and $\alpha 5$ helices plus the $\beta 5$ strand with the latter being covered by the former (Fig. 4 a-d). By comparison, the cofactor binding residues are localized in loops or nearby them. The location of the catalytic sites plus substrate and cofactor binding residues in the two sunflower enzymes are shown in Fig. 4c-d as viewed from the surface and internally. A minor difference between the two generated sunflower models is that the *HaENR1* carboxyl-terminus ends in a two residue α helix that is absent in *HaENR2* (Fig.4a and b) which is also not present in *BnENR*. A second minor difference is the organization of the loop region between the $\alpha 2$ helix and the $\beta 3$ strand. That is, in *HaENR1* the former is followed by 3 loop forming residues and 4 α helix forming residues, whereas in *HaENR2* 5 loop forming residues are followed by 2 α helix forming residues. This, together with the change of G312 in *HaENR1* to R311 in *HaENR2*, results in a modification of the orientation of the Y107 residue (Y106 in *HaENR2*), see circle in Fig. 4a and b, and the modification of the reducing agent binding site (Fig. 4 c-f).

Southern blot analysis of *HaENR1* and *HaENR2*

Two DNA gel blot analyses were performed on genomic DNA extracted from sunflower leaves to evaluate the number of copies of *HaENR1* and *HaENR2* in the genome. An aliquot of the gDNA was digested with restriction enzymes having one (*AleI*) or no (*XbaI*, *SacI* and *NcoI*) cutting sites within the cDNA of *HaENR1*. Fig. 5a

shows the single bands obtained with *Xba*I and *Nco*I and two bands generated by *Ale*I, indicating that *HaENR1* is a single copy gene. With *Sac*I two bands (around 10.5 and 12.8 Kb) are seen instead of the expected one, pointing to the existence of introns including a *Sac*I restriction site in the gene. This possibility was strengthened by carrying out a search in the TAIR database (<http://www.arabidopsis.org/index.jsp>) which revealed that *AtENR* (At2g05990) has ten introns within the coding region of the gene. In a similar way another aliquot of the gDNA was digested with *Hind*III, having one target within the *HaENR2* cDNA, and others lacking targets in the cDNA (*Sph*I, *Bam*HI, *Eco*RI and *Ale*I). Fig. 5b shows multiple bands except for the *Sph*I lane with a single band. The latter again suggests a single gene. The 2-4 hybridizing fragments for the other four enzymes when only 0-1 are predicted again suggests the presence of introns. Given that the digests were complete in all cases the results infer that both *HaENR1* and *HaENR2* are single copy genes. Moreover, a Blast search of available *Helianthus annuus* ESTs performed with the *HaENR1* and *HaENR2* sequences reported here retrieved only nucleotide fragments corresponding to both sequences supporting our contention.

Tissue expression profiles of sunflower enoyl-[ACP]-reductases

The expression of *HaENR1* and *HaENR2* was analysed by RT-qPCR in roots, stems, 1-7 days after germination (DAG) cotyledons, leaves of 20-day-old seedlings and 12–30 days after flowering (DAF) seeds in different developmental stages (Fig. 6). Different patterns were observed for each sunflower gene. *HaENR2* was highly expressed in seeds although the level decreased with age. Likewise the level decreased during cotyledon development. While *HaENR1* had its highest levels of expression in 1 DAG cotyledons which thereafter decreased with age although not so drastically, its

level of expression increased with seed age before falling at the end. Only traces of *HaENR2* were expressed in roots and older cotyledons.

Complementation of JP1111 with the sunflower *ENR* genes

The constructions pQE80L::*HaENR1* and pQE80L::*HaENR2*, which lacked the 138 and 225 bases constituting the deduced plastid targeting sequences, respectively, which were replaced by a 6xHis-tag, were transformed into *E. coli* strain JP1111 in which the endogenous *fabI(ts)* encoded enoyl-[ACP]-reductase is non-functional at 42°C as fatty acids can't be synthesized. Figure 7 a and b confirms this phenotype. When IPTG was added to induce expression of the *HaENR* genes, the bacteria were able to grow at 42°C (Figure 7 c and d). These results demonstrate that not only are these recombinant enzymes from sunflower functional, but also that they are able to interact with the other moieties of the bacterial FAS complex.

Fatty acid modifications in *E. coli* cells expressing recombinant *HaENR1* and *HaENR2*

The growth rate of the cells expressing *HaENR1* was monitored under different conditions at 32°C without IPTG (functional *EcFabI*), at 32°C with IPTG (functional *EcFabI* and expressed *HaENR1* or *HaENR2*) and 42°C with IPTG (non-functional *EcFabI* and expressed *HaENR1* or *HaENR2*). The resulting values of 0.54/0.53/0.53 h⁻¹, respectively, are not significantly different. The same experiment with *HaENR2* gave an analogous result, 0.68/0.71/0.65 h⁻¹. The minor difference between the growth rates in the two experiments is presumably caused by slightly different environmental conditions. Similar experiments with XL1-BI and other genes in our laboratory have given a range of 0.47 to 0.55 h⁻¹ (González-Mellado et al., 2010). While fewer fatty acids were present in the second of the two experiments with only the pQE80L vector

present, their fatty acid profiles were almost identical (Table 2). The presence of the *HaENR1* and *HaENR2* genes, however, resulted in a decrease of the total fatty acid content to circa 25% and 32%, respectively, compared to when only the vector was present (Table 2). Reductions in fatty acid amount were also observed in the cells expressing the plant enzymes together with the endogenous *EcFabI*, albeit in lower proportions. When both *EcFabI* and *HaENR1* were expressed the relative amounts of the various fatty acids was the same as when only *EcfabI* was expressed. If only *HaENR1* was expressed, however, marked changes took place. That is, the proportion of saturated fatty acids increased from 50 to 70%, the ratio of 16:0 to 18:0 decreased from 38 to 19% and the proportion of unsaturated fatty acids converted to cyclopropanes increased, with the amount of C17 tripling (Table 2). The latter is characteristic for *E. coli* cells under stress. Unlike what was observed with *HaENR1*, when *HaENR2* was expressed together with *EcfabI* the proportion of saturated fatty acids decreased from 50 to 30%, the ratio of 16:0 to 18:0 decreased from 37 to 29% and the ratio of 16:1 to 18:1 decreased from 80 to 50% (Table 2). At the same time the proportion of unsaturated fatty acids converted to cyclopropanes underwent little change, suggesting the cells were not under stress. If only *HaENR2* was expressed little change took place in the fatty acid profile compared to its absence. The ratio of C16:0 to 18:0 decreased from 37 to 26% and not quite a doubling of the amount of C17 cyclopropane from its 16:1 precursor took place inferring that *HaENR2* is more compatible with the *E. coli* FAS than is *HaENR1*.

Purification of the recombinant *HaENR1* and *HaENR2* proteins and their biochemical characterization

Since the complementation experiments were successful in JP1111, it was tested as a host strain for expression of the two sunflower ENRs at the permissive and restrictive temperatures of 32 and 45 °C, respectively, as well as testing XL1-BI at 37 °C. For *HaENR2* the best results were obtained using XL1-BI. This is illustrated in Suppl. Fig. 1b which shows a Coomassie stain of the total proteins in the XL1-BI strain after expression, the starting soluble fraction as well as affinity purified protein of the expected 36.87 kDa. By comparison, *HaENR1* was not detectable when expressed in XL1-Blue, and was very poorly expressed in JP1111. The Western blot in Suppl. Fig. 1a was required to localize the *HaENR1* in the gel, corresponding to a band of 45.75 kDa. After affinity purification *HaENR1* was only 6% of the proteins.

The capacity of the two affinity purified protein fractions from sunflower to reduce the substrate crotonyl-CoA was determined. ENRs require a pyridine nucleotide cofactor for activity. Neither *HaENR1* nor *HaENR2* was active with 300 µM NADPH, but instead both depended on NADH. Figure 8 shows that for both enzymes NAD⁺ production from NADH was dependent on concentration and time of the reaction. The slopes generated from the graphs and the Hill equation were used to calculate K_m and V_{max} values. This resulted in a K_M for crotonyl-CoA of 119 µM and V_{max} of 9.75 µmol min⁻¹ mg⁻¹ of protein for *HaENR1* and a K_M of 121 µM and V_{max} of 2.37 µmol min⁻¹ mg⁻¹ of protein for *HaENR2*.

Discussion

Two enoyl-[ACP]-reductase genes, *HaENR1* and *HaENR2*, were isolated and cloned from sunflower developing seed cDNA. The deduced proteins had transit peptides targeting the mature protein to the plastid or chloroplast. Surprisingly, neither of the transit peptides showed phosphorylation motifs specific for binding the 14-3-3 protein that together with HSP70 can form a cytosolic guidance complex towards the chloroplast (May and Soll, 2000). This was the case for other previously studied sunflower proteins involved in fatty acid biosynthesis, such as FatA-type thioesterases (Serrano-Vega et al., 2005), the ω 3-desaturase *HaFAD7* (Venegas-Calderón et al., 2006) and the condensing enzyme *HaKASIII* (González-Mellado et al., 2010). Southern blot analysis of *HaENR1* and *HaENR2* indicated that both are most likely single copy genes as was observed in *O. europaea* (Poghosyan et al., 2005). Our Southern analysis also suggests the presence of introns in both genes which is expected due to the existence of conserved introns in other plant homologous ENR genes from different phylogenetic groups.

The phylogenetic analysis of the sunflower ENRs with related proteins from different taxonomic groups revealed the presence of two groups of these proteins in some families from dicotyledons (type I and type II), the protein *HaENR1* belongs to type I and *HaENR2* to type II. However, the duplicated proteins of the monocotyledons do not belong to either group (Figure 2). This infers that the duplication of the ENR gene in the mono- and dicotyledons occurred independently from one another. Why the Brassica ENR proteins do not belong to either type I or II of the other dicotyledons remains to be explained. The existence of a common ancestor with cyanobacterial enoyl-[ACP]-reductases concurs with the cyanobacterial origin of the chloroplast via endosymbiosis (Goksøyr, 1967), as was observed for the KASIII protein from the

sunflower FAS complex (González-Mellado et al., 2010). During evolution the plastid/chloroplast ENR gene sequences have acquired transit peptides, the major difference to the cyanobacterial ENR genes.

The RT-qPCR data revealed different expression patterns for the sunflower enoyl-[ACP]-reductase genes. While initially *HaENR2* was very highly expressed in young seeds and *HaENR1* not so, during development the former decreased while the latter increased up to 25 DAF at which time the levels of the two were quite similar (Figure 6). Taking in account that the maximum rate of lipid accumulation in developing sunflower seeds occurs around 18-19 DAF (Martínez-Force et al., 2000), these results suggests that after an initial high expression of *HaENR2* a “basal” level, enough to maintain the high oil biosynthetic rate, is maintained with the expression of both genes (Fig. 6). A similar behaviour was described by Fawcett et al. (1994) when studying the mRNA expression levels of enoyl-ACP reductase isoforms in *Brassica napus* developing seeds. The expression changes of *HaENR1* during seed development mimic that of the previously analyzed *Arabidopsis* At1g24630 gene, the unique gene encoding for the plastidial ENR activity (Schmid et al., 2005). During cotyledon development expression of both genes declined, especially of *HaENR2*. The decrease for total ENR in cotyledons presumably reflects the initial change from dark/reserve tissues to a photosynthetically active one requiring rapid synthesis of large quantities of polar membrane lipids. Likewise, the leaves studied were young and in a phase of rapid expansion that also requires lots of polar lipids for the membranes and the developing cuticular apoplast. Judging from Figure 6 *HaENR* makes the major contribution toward this synthesis, but at later stages in cotyledon development and in stems and roots makes little if any contribution (see below).

The published model of *BnENR* (Rafferty et al., 1995) permitted the prediction of the *HaENR1* and *HaENR2* protein structures, where α helices and β strands form the Rossman fold of the sunflower proteins analogous to those in the *B. napus* structure. The functional residues, catalytic and cofactor binding sites were localized in conserved regions of these proteins. Main differences between the secondary structures of the sunflower ENRs were the presence of a 2 residue α -helix in the C-terminus of *HaENR1*, and the organization of the loop region between the $\alpha 2$ helix and the $\beta 3$ strand (Fig. 4 a and b). The influence of the last difference, together with the change of a glycine residue (G312 in *HaENR1*) to arginine (R311 in *HaENR2*), on Y107 position (Y106 in *HaENR2*) on the tertiary structure can be seen in Figure 4 when comparing the circle area in a vs b and the corresponding regions in c/d and e/f panels. Possibly the latter difference influences catalytic activity due to the modification of the substrate binding site(Fig. 4 c/d and e/f).

The results of the complementation experiments in the *E. coli*(ts) strain JP1111 with *HaENR1* and *HaENR2* demonstrates, first that these heterologous proteins were active and second were able to supply the basic functions of the bacterial protein *FabI* interacting with the FAS complex. This is in accord with the results of Kater et al. (1994) in related complementation experiments with *BnENR*. Expression of the sunflower ENRs did not affect growth rate in contrast to *BnENR*, but they decreased lipid synthesis in the bacteria whether they replaced the bacterial enzyme or were expressed in cells producing *EcFabI*. The decrease was larger in the absence of *EcFabI*. Presumably the decrease results from a structural incompatibility with one or more members of the bacterial FAS complex or the bacterial ACP. The effect on lipid synthesis was not investigated with the *BnENR* (Kater et al., 1994).

Both expressed sunflower proteins were also active *in vitro* converting crotonyl-CoA to butyryl-CoA in the presence of NADH. The K_M 's for crotonyl-CoA of the two enzymes were very similar circa 120 μ M. This is within the previously measured K_M values for crotonyl-CoA (40 to 667 μ M) for ENRs from spinach, rape, Arabidopsis and *Carthamus* (Dayan et al., 2008; Shimakata et al., 1982a; Shimakata et al., 1982b; Slabas et al., 1986). Interestingly, heterologously expressed *Ha*ENR1 had a higher $V_{max} = 9.75$ versus 2.37 μ mol min⁻¹ mg⁻¹ for *Ha*ENR2. Combining this with the fact that during early development the seeds had less *Ha*ENR1 than *Ha*ENR2 mRNA infers that the former may contribute more to total seed ENR activity than is apparent from the RT-qPCR measurements. To determine the real contribution of the two ENRs to fatty acid synthesis requires isolation and quantification of the two isozymes from the same tissues in which their mRNA was quantified.

In bacteria the separation between unsaturated and saturated fatty acid biosynthesis occurs at 10 carbons atoms when isomerization is carried out by the β -hydroxyacyl-[ACP]-dehydrase enzyme (FabA) (Fig. 9) (Brock et al., 1967). FabB and FabI compete for the *cis*-3- and *trans*-2-decenoyl-[ACP] products of *FabA*. If the ratio of FabB to FabI is greater than one then synthesis of unsaturated fatty acids increases (Chan et al., 2010). The normal result is circa 50% each of saturated and monoenoic fatty acids (Fig. 9) (Garwin et al. 1980). In the case of the large excess of overexpressed *Ha*ENR2 replacing the *in vivo* amount of *Ec*FabI, the saturated:unsaturated fatty acid ratio did not change (Fig. 9b). This implies that the affinity of *Ha*ENR2 for *trans*-2-decenoyl-[ACP] is the same as that of *Ec*FabI. Since replacing *Ec*FabI with overexpression of *Ha*ENR1, albeit only detectable in a Coomassie gel with the aid of an antibody, led to an increase of saturated versus unsaturated fatty acids (Fig. 9a), *Ha*ENR1 may have a greater affinity for *trans*-2-decenoyl-[ACP] than *Ec*FabI. This

would affect the equilibrium of the FabA isomerase activity resulting in generation of more *trans*-2-decenoyl from the *cis*-3-decenoyl leading to the observed increase in saturated fatty acids. The observed differences in the fatty acid profiles resulting from expressing the sunflower genes probably do not result from their activities, but rather from the stress these overexpressed proteins instigate which induces changes in the fatty acid biosynthetic pathway (DiRusso et al., 1999).

The existence of two phylogenetically separated genes coding for the enoyl-[ACP]-reductase activity in sunflower, their differences regarding the expression levels in the plant tissue besides their different structural and biochemical characteristics as well as their different effects in the bacteria, suggest that *HaENR1* and *HaENR2* have overlapping as well as different functions in sunflower fatty acid synthesis. This knowledge could be used to increase and optimize fatty acid production in the sunflower seed in the future, changing, for instance, their levels of expression in the seed.

Acknowledgements

We thank Dr. Alicia M. Muro-Pastor for help with southern blot analysis and A. González-Callejas and B. Lopez-Cordero for skilful technical assistance. This work was supported by the “Ministerio de Economía y Competitividad” and FEDER, project AGL2011-23187. IGT was supported by a JAE-CSIC contract in part financed by European Social Fund.

References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402
- Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um KS, Wilson T, Collins D, de Lisle G, Jacobs WR (1994) inhA, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science* 263: 227-230
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28: 235-242
- Brock DJH, Kass LR, Bloch K (1967) β -hydroxydecanoyl thioester dehydrase. II. Mode of action. *J Biol Chem* 242: 4432-4440
- Brown AP, Affleck V, Fawcett T, Slabas AR (2006). Tandem affinity purification tagging of fatty acid biosynthetic enzymes in *Synechocystis* sp PCC6803 and *Arabidopsis thaliana*. *J Exp Bot* 57: 1563–1571
- deBoer GJ, Kater MM, Fawcett T, Slabas AR, Nijkamp HJJ, Stuije AR (1998) The NADH-specific enoyl-[ACP]-reductase: characterization of a housekeeping gene involved in storage lipid synthesis in seed of *Arabidopsis* and other plant species. *Plant Physiol Biochem* 36: 473-486
- Chan DI, Vogel H (2010) Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem J* 430: 1-19
- Dayan FE, Ferreira D, Wang Y-H, Khan IA, McInroy JA, Pan Z (2008). A pathogenic fungi diphenyl ether phytotoxin targets plant enoyl (acyl carrier protein) reductase. *Plant Physiol* 147: 1062-1071
- DiRusso CC, Black PN, Weimar JD (1999) Molecular inroads into the regulation and metabolism of fatty acids, lessons from bacteria. *Prog Lipid Res* 38: 129-197

Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell
 using TargetP, SignalP, and related tools. *Nat Protoc* 2: 953-971

Fawcett T, Simon WJ, Swinhoe R, Shanklin J, Nishida J, Christie WW, Slabas AR
 (1994) Expression of mRNA and steady-state levels of protein isoforms of enoyl-
 ACP reductase from *Brassica napus*. *Plant Mol Biol* 26: 155-163

Fisher M, Kroon JT, Martindale W, Stuitje AR, Slabas AR, Rafferty JB (2000) The X-
 ray structure of *Brassica napus* β -keto acyl carrier protein reductase and its
 implications for substrate binding and catalysis. *Structure* 8: 339-347

Garwin JL, Klages AL, Cronan JE (1980) β -Ketoacyl-acyl carrier protein synthase II of
Escherichia coli. *J Biol Chem* 255: 3263-3265.

González-Mellado D, von Wettstein-Knowles P, Garcés R and Martínez- Force E
 (2010) The role of β -ketoacyl-acyl carrier protein synthase III in the condensation
 steps of fatty acid biosynthesis in sunflower. *Planta* 231:1277-1289

Goksøyr J (1967) Evolution of eukaryotic cells. *Nature* 214: 1161

Harwood JL (1988) Fatty acid metabolism. *Annu Rev Plant Physiol Mol Biol* 39: 101-
 138

Harwood JL (1996) Recent advances in the biosynthesis of plant fatty acids. *Biochim*
Biophys Acta 1301: 7-56

Harwood JL (2005) Fatty acid biosynthesis. In: Murphy DJ (ed) *Plant Lipids: biology,*
utilisation and manipulation. Blackwell Publishing, Oxford, pp 27-66

Jörnvall H, Persson B, Krook M, Atrian S, González-Duarte R, Jeffery J, Ghosh D
 (1995) Short-chain dehydrogenases/reductases (SDR). *Biochemistry* 34: 6003-
 6013

627 Kater MM, Koningstein GM, Nijkamp HJJ, Stuitje AR (1991) cDNA cloning and
628 expression of *Brassica napus* enoyl-acyl carrier protein reductase in *Escherichia*
629 *coli*. Plant Mol Biol 17: 895-909

630 Kater MM, Koningstein GM, Nijkamp HJJ, Stuitje AR (1994) The use of a hybrid
631 genetic system to study the functional relationship between prokaryotic and plant
632 multi-enzyme fatty acid synthetase complexes. Plant Mol Biol 25: 771-790

633 Kavanagh KL, Jernvall H, Persson B, Oppermann U (2008) The SDR superfamily:
634 functional and structural diversity within a family of metabolic and regulatory
635 enzymes. Cell Mol Life Sci 65: 3895-3906

636 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H,
637 Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins
638 DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948

639 Li-Beisson Y, Shorrosh B, Beisson F, Andersson MX, Arondel V, Bates PD, Baud S,
640 Bird D, DeBono A, Durrett TP, Franke RB, Graham IA, Katayama K, Kelly AA,
641 Larson T, Markham JE, Miquel M, Molina I, Nishida I, Rowland O, Samuels L,
642 Schmid KM, Wada H, Welti R, Xu C, Zallot R, and Ohlrogge J (2013) Acyl-
643 Lipid Metabolism. The Arabidopsis Book. American Society of Plant Biologists.
644 10.1199/tab.0161.

645 Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-
646 time quantitative PCR and the 2- $\Delta\Delta$ CT method. Methods 25: 402-408

647 Martinez-Force E, Cantisan S, Serrano-Vega MJ and Garces R (2000) Acyl-acyl carrier
648 protein thioesterase activity from sunflower (*Helianthus annuus L.*) seeds. Planta
649 211:673-678

650 May T, Soll J (2000) 14-3-3 proteins form a guidance complex with chloroplast
651 precursor proteins in plants. Plant Cell 12: 53-63

652 McGuire KA, Siggaard-Andersen M, Bangera MG, Olsen JG, von Wettstein-Knowles P
 653 (2001) β -ketoacyl-[acyl carrier protein] synthase I of *Escherichia coli*: Aspects of
 654 the condensation mechanism revealed by analyses of mutations in the active site
 655 pocket. *Biochemistry* 40: 9836-9845
 656 Martinez-Force E,
 657 Mou Z, He Y, Dai Y, Liu X, Li J (2000) Deficiency in Fatty Acid Synthase leads to
 658 premature cell death and dramatic alterations in plant morphology. *Plant Cell* 12:
 659 405-417
 660 Oppermann U, Filling C, Hult M, Shafqat N, Wua X, Lindh M, Shafqat J, Nordling E,
 661 Kallberg Y, Persson B, Jörnvall H (2003) Short-chain dehydrogenases/reductases
 662 (SDR):the 2002 update. *Chem Biol Interact* 143-144: 247-253
 663 Persson B, Kallberg Y, Oppermann U, Jörnvall H (2003) Coenzyme-based functional
 664 assignments of short-chain dehydrogenases/reductases (SDRs). *Chem Biol*
 665 *Interact* 143-144: 271-278
 666 Poghosyan ZP, Giannoulia K, Katinakis P, Murphy DJ, Hatzopoulos P (2005) Temporal
 667 and transient expression of olive enoyl-[ACP]-reductase gene during flower and
 668 fruit development. *Plant Physiol Biochem* 43: 37-44
 669 Rafferty JB, Simon JW, Baldock C, Artymiuk PJ, Baker PJ, Stuitje AR, Slabas AR, Rice DW (1995)
 670 Common themes in redox chemistry emerge from the X-ray structure of oilseed
 671 rape (*Brassica napus*) enoyl acyl carrier protein reductase. *Structure* 3: 927-38
 672 Saitou N, Nei M (1987) The neighbour-joining method: a new method for
 673 reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425
 674 Serrano-Vega MJ, Garcés R and Martínez-Force E (2005) Cloning, characterization and
 675 structural model of a FatA-type thioesterase from sunflower seeds (*Helianthus*
 676 *annuus* L). *Planta* 221: 868-880

677 Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B,
 678 Weigel D, Lohmann JU (2005) A gene expression map of *Arabidopsis thaliana*
 679 development. Nat Genet 7: 501-506

680 Shimakata T, Stumpf PK (1982a) Purification and characterizations of β -Ketoacyl-
 681 [Acyl-Carrier-Protein] reductase, β -Hydroxyacyl-[Acyl-Carrier-Protein]
 682 dehydrase, and Enoyl-[Acyl-Carrier-Protein] reductase from *Spinacia oleracea*
 683 leaves. Arch Biochem Biophys 218: 77-91

684 Shimakata T, Stumpf PK (1982b) The prokaryotic nature of the Fatty Acid Synthetase
 685 of developing *Carthamus tinctorius* L. (safflower) seeds. Arch Biochem Biophys
 686 217: 144-154

687 Shimakata T, Stumpf PK (1982c) Fatty Acid Synthetase of *Spinacia oleracea* leaves.
 688 Plant Physiol 69: 1257-1262

689 Slabas AR, Sidebottom CM, Hellyer A, Kessell RMJ, Tombs MP (1986) Induction,
 690 purification and characterization of NADH-specific enoyl acyl carrier protein
 691 reductase from developing seed of oil seed rape (*Brassica napus*). Biochim
 692 Biophys Acta, 877: 271-280

693 Smith S, Witkowski A, Joshi AK (2003) Structural and functional organization of the
 694 animal fatty acid synthase. Prog Lipid Res 42: 289-317

695 Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: Molecular Evolutionary
 696 Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599

697 Venegas-Calderón M, Muro-Pastor AM, Garcés R, Martínez-Force E (2006) Functional
 698 characterization of a plastidial omega-3 desaturase from sunflower (*Helianthus*
 699 *annuus*) in cyanobacteria. Plant Physiol Biochem 44: 517-525

700 Ward WHJ, Holdgate GA, Rowsell S, McLean EG, Pauptit RA, Clayton E, Nichols
 701 WW, Colls JG, Minshull CA, Jude DA, Mistry A, Timms D, Camble R, Hales NJ,

702 Britton CJ, Taylor IWF (1999) Kinetic and Structural Characteristics of the
703 Inhibition of Enoyl (Acyl Carrier Protein) Reductase by Triclosan. *Biochemistry*
704 38: 12514-12525
705 White SW, Zheng J, Zhang YM, Rock CO (2005) The structural biology of type II fatty
706 acid biosynthesis. *Annu Rev Biochem* 74: 791-831
707

708 **Electronic supplementary material** The online version of this article (doi:)
709 contains supplementary material, which is available to authorized users.

710

711

712

Figure Legends

Fig. 1. Alignment of the deduced amino acid sequences of *H. annuus* enoyl-[ACP]-reductases *HaENR1* and *HaENR2*, with the closely related sequences from *A. thaliana*, *AtENR1* (gi|13265591|), rice, *OsENR* (gi|115478314|) and *E. coli*, *EcFabI* (gi|16129249|). Identical residues are highlighted as *black boxes*, highly conserved residues as *dark grey boxes* and weakly conserved residues as *light grey boxes*. Residues participating in the substrate binding site are indicated with black triangles, those in the reducing agent binding site with white triangles and those in both sites with black striped triangles. The putative N-terminal plastid transit peptides of *HaENR1* and *HaENR2*, are underlined, 46 and 75 amino acids, respectively.

Fig. 2. Phylogenetic comparison of plant and algae enoyl-[ACP]-reductase homologues of the *E. coli* FabI protein. *H. annuus* sequences are shown in bold.

Fig. 3. Comparison between the deduced amino acid sequences of *H. annuus* enoyl-[ACP]-reductases, *HaENR1* and *HaENR2*, and *B. napus*, BnKAR (Rafferty et al., 1995). Identical residues are highlighted as *black boxes*, highly conserved residues as *dark grey boxes* and weakly conserved residues as *light grey boxes*. BnKAR structural elements are named as in Rafferty et al. (1995); *h*: α -helix; *s*: β -sheet.

Fig. 4. Proposed structural models for sunflower enoyl-[ACP]-reductases, *HaENR1* (a, c and e) and *HaENR2* (b, d and f), modelled from that known for BnKAR (1eno) (Rafferty et al., 1995). a and b, ribbon diagrams; c and d, views of the molecular surfaces and e and f, internal views of the active sites. Residues in α -helices are red,

those in β -sheets are yellow, those in the substrate binding site are blue, those in the reducing agent binding site are green, and those in the catalytic site are pink. The circle shows the tyrosine residue (Y107 in *HaENR1* and Y106 in *HaENR2*) and the surroundings showing differences between both proteins.

Fig. 5. Southern blot analysis of *H. annuus* genomic DNA digested with the indicated restriction enzymes and hybridized with *HaENR1* (a) and *HaENR2* (b) gene specific probes. Hybridizing fragments marked with arrows, and the position of DNA molecular weight standards in Kb is shown on the left.

Fig. 6. Expression levels of enoyl-[ACP]-reductases in developing seeds and vegetative tissues from *H. annuus*, *HaENR1* and *HaENR2*, determined by RT-qPCR from sunflower line CAS-6. Stages in cotyledon development are designated days after germination (DAG) and in seed development days after flowering (DAF). Data are mean values $\pm$ SD of three independent biological samples.

Fig. 7. Sunflower enoyl-[ACP]-reductase proteins complement the *E. coli fabI(ts)* mutant at 42°C. *E. coli fabI(ts)* with pQE-80L::*HaENR1* (top) or pQE-80L::*HaENR2* (bottom) grown with (right) and without (left) IPTG induction.

Fig. 8. Substrate specificity of *H. annuus* enoyl-[ACP]-reductase proteins, *HaENR1* (a) and *HaENR2* (b). *In vitro* assays were performed in 17.9 mM potassium phosphate buffer, 1.6 mM DTT at pH 6.8, 28°C with 300 μ M NADH and *HaENR1* (0.4 μ g ml⁻¹) or *HaENR2* (8.3 μ g ml⁻¹). The reaction was activated by adding the substrate (crotonyl-CoA) at different concentrations (20-250 μ M). The oxidation of NADH was monitored

at 340 nm. Results correspond to the average of three independent determinations as indicated before in Materials and Methods.

Fig. 9. Schematic representation of *E. coli* fatty acid biosynthesis comparing the effect on the ratio of saturated to unsaturated fatty acids overexpressing *H. annuus* enoyl-[ACP]-reductases, by replacing *EcFabI* with *HaENR1* (a) or overexpressing *HaENR2* with functional *EcFabI* (b). Proteins involved in fatty acid biosynthesis in *E. coli* are represented. FabZ, β -hydroxyacyl-[ACP]-dehydrase; FabA, β -hydroxydecanoyl-[ACP]-dehydrase; FabB, β -ketoacyl-[ACP]-synthase I; FabI, enoyl-[ACP]-reductase. Symbols in grey define the bacterial pathway; in black represent the effect of *H. annuus* enoyl-[ACP]-reductases overexpression in the cell.

Tables

Table 1. Sequence of PCR primers used in this work.

Primer name	Sequence ^a
F-ENR1	5'-GAAGGGAAGGTTCAATCAGC-3'
R-ENR1	5'-CTCAACACATACCTTTCAACTTG-3'
F-ENR1SacI	5'- AGAGCT CTCTTCTCATCTGTCATCGGTG-3'
R-ENR1XmaI	5'- TCCCGGG TACTAATGTTTGTCCATTGG-3'
F-ENR1qpcr	5'-TCTGCCATTACGGGTGCTC-3'
R-ENR1qpcr	5'-CTCAACACATACCTTTCAACTTG-3'
F-ENR2	5'-TCGGAGAGGTTTAGGATTGG-3'
R-ENR2	5'-CGGTGTGCTTGTTCACTACC-3'
F-ENR2SacI	5'- TAGAGCT CGCTTCCAGTGACACAAAACC-3'
R-ENR2XmaI	5'- TCCCGGG TTAATGTTTGGGGTCTCTTG-3'
F-ENR2qpcr	5'-TCAGTAGCGGTGCGAAAGG-3'
R-ENR2qpcr	5'-GTTTTGTGTCACTGGAAGCC-3'
QHaActin-F4	5'-GCTAACAGGGAAAAGATGACT-3'
QHaActin-R4	5'-ACTGGCATAAAGAGAAAGCACG-3'

^a Restriction sites are indicated in bold

Table 2. Lipid content and fatty acid composition of *E. coli* JP111 strain *fabI(ts)* bearing the vector pQE80L, or either one of the recombinant plasmids pQE80L::*HaENR1* and pQE80L::*HaENR2* growing on rich medium supplemented with ampicillin in three expression conditions: i) 32°C , non-induction, ii) 32°C , induction with IPTG and iii) 42°C , induction with IPTG. Data are the average of three independent samples having SDs of <5% of mean value. ^a In *E. coli* the monoene fatty acids are C16:1*cis*Δ9 and 18:1*cis*Δ11. ^b C17 and 19 cyclopropanes derived from C16:1 and 18:1, respectively. ^c (16:1 + 17)/(18:1 + 19).

Vectors	Conditions	Genes expressed	Total (μg)	Fatty acids (mol%)						SAT/UNSAT	16:1/18:1 ^c	16:0/18:0
				16:0	16:1 ^a	17 ^b	18:0	18:1 ^a	19 ^b			
pQE80L	32°C	<i>EcfabI(ts)</i>	575	31.4	25.6	5.9	0.8	35.9	0.4	0.5	0.9	38.1
pQE80L:: <i>Ha ENR1</i>	32°C, IPTG	<i>EcfabI(ts)</i> + <i>HaENR1</i>	523	31.3	24.6	5.2	0.9	37.8	0.3	0.5	0.8	36.6
pQE80L:: <i>Ha ENR1</i>	42°C, IPTG	<i>HaENR1</i>	431	39.5	12.5	18.7	2.0	26.4	0.7	0.7	1.1	19.3
pQE80L	32°C	<i>EcfabI(ts)</i>	515	31.5	26.2	4.1	0.9	37.0	0.3	0.5	0.8	36.9
pQE80L:: <i>Ha ENR2</i>	32°C, IPTG	<i>EcfabI(ts)</i> + <i>HaENR2</i>	394	24.9	23.2	2.7	0.9	48.2	0.2	0.3	0.5	28.5
pQE80L:: <i>Ha ENR2</i>	42°C, IPTG	<i>HaENR2</i>	351	32.3	19.9	7.4	1.3	38.8	0.3	0.5	0.7	25.7

<i>At</i> ENR	MAATAASSLQIATRRPSMSSPSKILKAGTYIVGANPGNASWDKLSCTRQLSNLGLCRNHSAVPTCKRPF S FSTRAMSESS	80
<i>Ha</i> ENR1	MMAVGV S ALQITAARPSVFLTQRTL G ATMATFVHESDRILSSSKLTSSSHLSSVLNIK-HKCTTSRPRFQKFVTRSMSEAS	79
<i>Os</i> ENR	MGASVTTGLQMAAARPCIPACQRL L GSRAALPSFGRALSTQTGFASCRKTASAGPFVSLNHKR-----FAVRAMS---	70
<i>Ha</i> ENR2	-MTTNITNMHLAIAKASAFSSHQVSKSSIA D FS S GAKGASCAKLKSSCHIAST-KVSSQNVK F S D SRFERVITKAVSASS	78
<i>Ec</i> ENR	-----	0

	<div> <div>W</div> <div>W</div> <div>▽</div> </div>	
<i>At</i> ENR	ENKAPSGLPIDLRGKRAFIAGIADDNGYGWAIAKSLAAAGAEILVGTWVPALNIFETSLRRGKFDQSRVLPDGSIMEIKK	160
<i>Ha</i> ENR1	GSKSVS Q LSIDLK G KRAFIAGVADDNGYGWAIAKSLAAAGAEILVGTWVPALNIFEMSLRRGKFDES R VLPDGSIMEITK	159
<i>Os</i> ENR	----AQGLPIDLRGKRAFIAGVADDNGYGWAIAKALAAAGAEILVGTWVPALNIFETSLRRGKFDES R KLPDGSIMEITK	146
<i>Ha</i> ENR2	DTKPLPGLPVDLRGKRAFIAGIADDNGYGWAIAKSLAAAGAEILVGTWVPALNIFETSLRRGKFDES R VLPDGSIMEITK	158
<i>Ec</i> ENR	-----MGFLSGKRILVTGVASKLSIAYGIAQAMHREGAELAF T YQNDKL-----KGRVEEF A A	53

	<div> <div>W</div> <div>W</div> <div>W</div> <div>▽</div> </div>	
<i>At</i> ENR	VYALDAVFDNPEDVPEDVKTNKRYAGSSNWTVQEA A ECVKKDFGSI D ILVHSLANGPEVSKPLLET S RGYLA A ISASSY	240
<i>Ha</i> ENR1	VYPLDAVFDNPEDVPEDIKTNKRYAGASNWTVQ E VAESVKQDFGSI D ILVHSLANGPEVT K PLLET S RGYLA A VSASSY	239
<i>Os</i> ENR	VYPLDAVFDSPEDVPDDVKANKRYAGSSNWTVKEVAETVKND F GTIDILVHSLANGPEVKN S LLET S RGYLA A VSASSY	226
<i>Ha</i> ENR2	VYPLDAVYDTPEDVPEDVKTNKRYAGSSKWTVKEVAESVKQDFGSI D ILVHSLANGPEVSKPLLET S RYGYLA A ISASSY	238
<i>Ec</i> ENR	QLGSDIVLQC--DVAEDASIDTMFAELGKVWPKFDG F VHSIGFAPGDQLDGDYVN-----AVTREGFKIAHDISSY	122

	<div> <div>W</div> <div>▽</div> <div>▼</div> <div>▽</div> <div>W</div> <div>W</div> </div>	
<i>At</i> ENR	SFVSLLRHFLPIMNPGGASISLTYIASERIIPGYGGGMSSAKAALES D TRVLAYEAGRKSNIRVNTISAGPLGSRAAKAI	320
<i>Ha</i> ENR1	SYVSLLKHFLPIMNPGGSSISLTYIASERIIPGYGGGMSSAKAALES D TRVLAFEAGRKHKIRVNTISAGPLGSRAAKAI	319
<i>Os</i> ENR	SFISLLQHFLPIMNPGGATISLTYIASERTIPGYGGGMSSAKAALES D TRVLAYEAGRKGKIRVNTISAGPLGSRAAKAI	306
<i>Ha</i> ENR2	SYVSLLQHFAPIMNPGGASISLTYIASERIIPGYGGGMSSAKAGLES D TQVLAFEAGRKHGIRVNTISAGPLRSRAAKAI	318
<i>Ec</i> ENR	SFVAMAKACRSMLNPGSALLTSLYLGAERAI P NYN-VMGLAKASLEANVRYMANAMGPE-GVRVNAISAGPIRTLAASGI	200

<i>At</i> ENR	GFIDTMI E YSYNNGPIQKTLTADEVGNAAAF L ASPLASAITGAT I YVDNGLNAMGVALDSPVFKDLNSKN----	390
<i>Ha</i> ENR1	GFIDTMI D YSFANAPLQKTL S ADEVGN T AAFLASPLASAITGAL I YVDNGLNAMGVGVDSPLFKDLNVPM D KH-	392
<i>Os</i> ENR	GFIEKMI E YSYVNAPLQKELLADEVGNTAAFLASPLASAITGST I YVDNGLNTMGLALD S PTLST-----	371
<i>Ha</i> ENR2	GFIDMMIDYSL E NAPLQKEL S AEEVGNAAAF L ASPLASAITGAVLYVDNGLNAMGVGVDS P VFKDLNIPRDPKH	392
<i>Ec</i> ENR	KDFRKMLAHCEAVTPIRRTVTIEDVGNSAAFLCSDLSAGISGEV V HVDGGFSIAAMNELELK-----	262

Sequence alignment of the *HaENR*, *BnENR*, and *HaENR2* proteins, showing conserved residues and structural elements. The alignment is presented in a table format, with the protein names and residue numbers on the left, and the corresponding amino acid sequences on the right. Conserved residues are highlighted in black boxes. Structural elements are indicated by symbols above the sequences: ∇ for beta-strands, $\nabla\nabla$ for beta-sheets, and $\nabla\nabla\nabla$ for beta-barrels. The alignment is divided into seven regions, labeled $\beta 1$ through $\beta 7$.

Protein	Residue	Sequence	Structural Element
<i>HaENR1</i>	87	LSIDLGKRAFIAGVADDNGYGWAAKSLAAAGAEILVGTWVPALNIFEMSLRRGKFD	∇
<i>BnENR</i>	86	LPIDLGKRAFIAGVADDNGYGWAAKSLAAAGAEILVGTWVPALNIFETSLRRGKFD	∇
<i>HaENR2</i>	86	LPVDLGKRAFIAGVADDNGYGWAAKSLAAAGAEILVGTWVPALNIFETSLRRGKFD	∇
<i>HaENR1</i>		ssssss hhhhhhhhhhhh sssssss hhhhhhhh hhhh	
<i>BnENR</i>		ssssss hhhhhhhhhhhh sssssss hhhhhhhhhh hh	
<i>HaENR2</i>		ssssss hhhhhhhhhhhh sssssss hhhhhhhh hh	
		$\beta 1$ $\alpha 1$ $\beta 2$ $\alpha 2$	
<i>HaENR1</i>	147	RVLDPGSLMEITKVYPLDAVFDNPEDVPEDVKTNKRYAGASNWTVCPEVAESVKQDFGSD	$\nabla\nabla$
<i>BnENR</i>	146	RVLDPGSLMEITKVYPLDAVFDNPEDVPEDVKTNKRYAGASNWTVCPEVAESVKQDFGSD	$\nabla\nabla$
<i>HaENR2</i>	146	RVLDPGSLMEITKVYPLDAVFDNPEDVPEDVKTNKRYAGASNWTVCPEVAESVKQDFGSD	$\nabla\nabla$
<i>HaENR1</i>		sssss hhhh hhh hhhhhhhhhhhh s	
<i>BnENR</i>		sssss hhhh hhh hhhhhhhhhhhh s	
<i>HaENR2</i>		sssss hhhh hhh hhhhhhhhhhhh s	
		$\beta 3$ $\alpha 3$	
<i>HaENR1</i>	207	ILVHSLANGPEVTKPILLETSRKGYLAAVSASSYSVSLLKHELPIMNPGGSISLTYIAS	$\nabla\nabla$
<i>BnENR</i>	206	ILVHSLANGPEVSKPILLETSRKGYLAAVSASSYSVSLLSHELPIMNPGGSISLTYIAS	$\nabla\nabla$
<i>HaENR2</i>	206	ILVHSLANGPEVSKPILLETSRKGYLAAVSASSYSVSLLOHELPIMNPGGSISLTYIAS	$\nabla\nabla$
<i>HaENR1</i>		sssss hhhhhhhhhhhhhhhhhhhhhhh sssssss hh	
<i>BnENR</i>		sssss hhhhhhhhhhhhhhhhhhhhhhh sssssss hh	
<i>HaENR2</i>		sssss hhhhhhhhhhhhhhhhhhhhhhh sssssss hh	
		$\beta 4$ $\alpha 4$ $\beta 5$	
<i>HaENR1</i>	267	ERIIPGYGGGMSSAKAALESDDRVLAFEAGRKEKIRVNTISAGPLGSRAAKAIGFIDTMI	∇ ∇ ∇
<i>BnENR</i>	266	ERIIPGYGGGMSSAKAALESDDRVLAFEAGRKNIRVNTISAGPLGSRAAKAIGFIDTMI	∇ ∇ ∇
<i>HaENR2</i>	266	ERIIPGYGGGMSSAKAALESDDRVLAFEAGRKEGIRVNTISAGPLGSRAAKAIGFIDTMI	∇ ∇ ∇
<i>HaENR1</i>		h hhhhhhhhhhhhhhhhhhhhhhh sssssss hhhhhh	
<i>BnENR</i>		h hhhhhhhhhhhhhhhhhhhhhhh sssssss hhhhhh	
<i>HaENR2</i>		h hhhhhhhhhhhhhhhhhhhhhhh sssssss hhhhhh	
		$\alpha 5$ $\beta 6$ $\alpha 6$	
<i>HaENR1</i>	327	DYSFANAPLOKTLISAEVGNAAAFASPLASAITGALIVYDNGNLNMGVGVDSPLFKDLN	$\nabla\nabla\nabla$ $\nabla\nabla$
<i>BnENR</i>	326	DYSYNAPLOKTLISAEVGNAAAFASPLASAITGATIVYDNGNLNMGVGVDSPLFKDLN	$\nabla\nabla\nabla$ $\nabla\nabla$
<i>HaENR2</i>	326	DYSLENAPLOKTLISAEVGNAAAFASPLASAITGAVIVYDNGNLNMGVGVDSPLFKDLN	$\nabla\nabla\nabla$ $\nabla\nabla$
<i>HaENR1</i>		hhhhh hhhhhhhhhhhh sssss hh	
<i>BnENR</i>		hhhhh hhhhhhhhhhhh sssss hh	
<i>HaENR2</i>		hhhhh hhhhhhhhhhhh sssss hh	
		$\alpha 7$ $\beta 7$	
<i>HaENR1</i>	387	VPMDKH-	
<i>BnENR</i>	386	K-----	
<i>HaENR2</i>	386	IPRDPKH	

a **b** **c** **d** **e** **f**

a **b**

Supplementary Figure 1

Supplementary Figure 1. Purification of *H. annuus* recombinant ENR proteins. (A) Coomassie blue-stained SDS-PAGE (left) and Western blot (right) with anti-His antibodies of recombinant *HaENR1* from XL1-BI grown at 37 °C. Arrow points to expected position of *HaENR1*. (B) Coomassie blue-stained SDS-PAGE of recombinant *HaENR2* from wild type

XL1-BI and the mutant strain JP1111 (*fabIts*) grown at 37 and 32 °C, respectively. T, total *E. coli* extract; S, soluble fraction; AF, affinity purified using one passage through a Ni²⁺-charged HisTrap FF column; kDa, size markers.