CLONING AND PROKARYOTIC EXPRESSION OF A COMPLEMENTARY DNA GENE FOR CYCLOPHILIN FROM CAMELLIA OLEIFERA

XIAOFENG TAN1*, BAOMING WANG1, LIN ZHANG1, MINGWANG SHI1,2, LUSHAN XIE1 AND FANGMING HU1

1The Key Lab. of Non-wood Forest Products of State Forestry Administration, Central South University of Forestry and Technology, Changsha, 410004, P. R. China
2Henan Institute of Science and Technology, Xinxiang, 453003, P. R. China
*Corresponding author Email: tanxiaofengcn@126.com; Tel: +86 731-5623416

Abstract

Cyclophilins (CyPs) are ubiquitous proteins and involved in protein folding and stress response due to peptidyl-prolyl cis-trans isomerase (PPIase) activity. In this study, a 975-base pair (bp) full-length complementary DNA (cDNA) gene, Camellia oleifera CyP (CoCyP, GenBank access no. FJ377540), was isolated from the conducted cDNA library of C. oleifera seeds. This cDNA encodes a polypeptide of 207 amino acids, which has an endoplasmic reticulum-localized signal sequence at its N-terminus, but not the typical signal sequence at its C-terminus. Moreover, its putative amino acid sequence shares highest identity with those of Arabidopsis thaliana CyPB (84%) and Triticum aestivum CyPB (81%), and possesses a highly conserved central core domain including cyclosporin A (CsA)-binding sites, the tetrapeptide Ala-Ala-Pro-Ala-binding sites, residues required for PPIase catalysis and an insertion of seven amino acids of unknown function. Collectively, it is suggested that the CoCyP protein belongs to CyPB and has PPIase activity. An approximately 21 kDa protein was expressed via the recombinant pET-30b+/CoCyP in Escherichia coli. In this study, we tentatively put forward to the hypothesis that the gene expressing during the peak of seed lipid biosynthesis, might correlate with the caused harmful products in the lipid biosynthesis process, and protect cells against reactive oxygen species (ROS) damage; thus it may be crucial during lipid biosynthesis and stress responsiveness.

Introduction

As one of the most important edible oil trees in China, tea-oil tree (Camellia oleifera) is planted in the large areas due to its strong adaptability. Tea-oil is a kind of excellent vegetable oil with the higher edible and medicinal value, and enriches unsaturated acid, such as oleic acid and linoleic acid (Zhuang, 2008). To reveal the mechanism of lipid biosynthesis is increasingly becoming a current crucial issue. One of solving strategies is to gain better understanding of the related lipid biosynthesis genes including enzymes with PPIase activity and molecular chaperones, which assist in the folding and maturation of proteins and the deposition of lipid storage compounds (Galili et al., 1998). CyPs are a conserved class of proteins that bind the immunosuppressive drug CsA (Handschumacher et al., 1984). They belong to the cluster of immunophilins that possess PPIase activity, the rate-limiting step in protein folding (Brandts et al., 1975). They are ubiquitous proteins present in all subcompartments (Galat, 1999), and play significant roles in a wide variety of processes including cellular metabolism (Shieh et al., 1989), transcription regulation (Ryczynand & Clevenger, 2002) and cell signaling (Brazin et al., 2002). CyPB, one class of cyclophilin family, accumulates both in the endoplasmic reticulum (ER) and in complexes on the plasma membrane (Price et al., 1994; Wu et al., 2009). It possesses a signal sequence at the N-terminus, which targets the storage protein to the
lumen of ER, followed by protein folding in it (Wu et al., 2009). Plant CyPs have been isolated and identified from a number of plants, e.g., tomato, maize and *Brassica napus* (Gasser et al., 1990), and *A. thaliana* (Romano et al., 2004). They also are involved in the biotic and abiotic stress responses besides the roles mentioned above and may be to act as chaperon-like proteins to facilitate the folding of the stress-related proteins or to protect these proteins from proteolytic degradation or aggregation (Godoy et al., 2000; Chen et al., 2007; Wang & Tan, 2008). Researches have focused on the roles of CyPs involved in protein folding as PPIase activity in *T. aestivum* (Wu et al., 2009), *Ricinus communis* (Gottschalk et al., 2008); nevertheless, further studies are still necessary due to little attention paying to lipid synthesis.

In spite of such important potential roles of CyPs. To date, no CyP genes have been cloned and expressed from *C. oleifera*. As an initial step toward understanding the roles of CyP in the pathway of the lipid biosynthesis and stress response, this study describes the isolation, identification and prokaryotic expression of the *CoCyP* gene, and aims to explore its roles in the processes of lipid biosynthesis, stress response and seed protein synthesis (Tan et al., 2006).

**Materials and Methods**

**Materials:** The cDNA library of *C. oleifera* seeds, *E. coli* stain BL21 (DE3) and the expression vector pET-30b (+) were stored in our laboratory. Gel Extraction Kit was bought from Ambiogen Life Science Technology Co. Ltd. *Taq* DNA polymerase (TaKaRa), Nde, EcoR, T4 DNA ligase, isopropylthio-β-D-galactoside (IPTG), Sodium dodecyl sulfate (SDS) and other chemicals were of reagent grades or the highest purity commercially available.

**Cloning identifying and sequencing cDNA of CyP gene:** The MonoClone (rzots0_000167.y1.scf) from the *C. oleifera* cDNA library was chosen and incubated in Luria broth (LB) agar plate containing containing 50 μg/ml ampicillin, and then grown in LB medium containing 50 μg/ml ampicillin overnight at 37°C. Plasmid DNA was prepared by SDS-alkaline lysis (Minipreparation) (Sambrook et al., 1999). The polymerase chain reaction (PCR) was performed with vector-based primers T3 (5'-ATTAACCCTCACTACGGGAGA-3') and T7 (5'-GCTGTTACAATACGACTCACTATAG-3') in a 20μL reaction containing 1×PCR buffer, 0.5 mM dNTPs, 1.5 mM MgCl₂, 0.5mM of each primer, 1U Taq polymerase and 10 ng DNA template. Cycling conditions consist of precycling at 94°C for 5 min, and then 35 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 45 s and elongation at 72°C for 1 min 30 s, and final extension at 72°C for 7 min. The PCR product was analyzed by 1.5% (w/v) agarose gel. DNA sequencing was performed by Shanghai Boshang biotechnology Co. Ltd (data not shown).

**Analysis of sequence data:** Database searches and the protein analysis (including molecular weight and signal sequence prediction) were performed with the NCBI server (http://www.ncbi.nlm.nih.gov), the ProtParam tool (http://www.expasy.ch/tools/protparam.html) and the SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP/). Subcellular location and predicted scores were conducted using the TargetP service (http://www.cbs.dtu.dk/services/TargetP/). Multiple sequence alignment was carried out by Vector NTI 10.0, and displayed by GeneDoc.

**Plasmid construction and expression:** The coding region of *CoCyP* was prepared by PCR using the cDNA product as template. The forward primer was chosen as 5'-
CGGCATATGGCCAGAACTCGATCG-3’ and reverse primer as 5’-CGGGAATTCTT-A

CACCAAGGAAGTCGCG-3’. The two primers harbored a recognition site for Nde

(underlined) and E. coR (underlined), respectively. The reaction volume and cycling

were as described above. After purification according to Gel Extraction Kit (Ambiogen),

the PCR product was digested with Nde and E. coR, and ligated into the pET-30b (+)

at the corresponding restriction sites. The mixture was transformed into the E. coli BL21

(DE3) strain and the new recombinant with CoCyP gene in the pET-30b (+) (referred to

as pET-30b (+)/CoCyP) were verified by DNA sequencing (data not shown) and the PCR

amplification. The nine recombinants chosen randomly were performed PCR, and the

reaction volume, conditions and primers were as the preparation of the coding region of

CoCyP described. Transformed cells were grown in LB medium containing 100 μg/ml

kanamycin to an OD600 of 0.6 at 37°C; IPTG was added to a final concentration of 1 mM

and the incubation was continued for an additional 4 h. After centrifugation, the pellet

was resuspended in 20μL sterilized H2O, and added to equal volume 2×SDS gel-loading

buffer, and then heated it to 100°C for 10 minutes. The suspension was analyzed on 15%

denaturing polycrylamide gel (Sambrook et al., 1999).

Results

Cloning and identifying of C. oleifera cyclophilin gene: PCR amplification of the

MonoClone (rzots0_000167.y1.scf) with T3 and T7 primers led to a product of size

approximately 1100 bp (Fig. 1). After the 110 bp vector sequence being subtracted, an

about 1000 bp cDNA fragment was acquired. It shares the high identity (96%) with

Camellia sinensis CyP in the nucleotide level, and the high degree of identity in the

amino acid level with plant CyPs from GenBank data, thus we proposed that the clone

might be a full-length cDNA. After DNA sequencing, the 975-bp full-length cDNA gene

of CyP was determined, and designated as CoCyP (GenBank access no. FJ377540). It

contains a 621-bp (35-655) opening reading frame (ORF) encoding a polypeptide of 207

amino acids, 34 bp 5'-untranslated region (5' UTR) and a 317 bp of 3'-untranslated

region (3' UTR) with polyadenylation. Behind the terminated code, there is a

polyadenylation signal AATAAA (752-757) (Fig. 2). Within the sequence of the CoCyP

cDNA, the 3'UTR is (A+T)-rich (67%, not contained polyadenylation), and 5UTR' is

much less (A+T)-rich (52.94%); there are no distinct differences between the UTRs and

the coding region (52.08%).

Analysis of the deduced CoCyP amino acid sequence: The prediction results of the

ProtParam tool and the SignalP 3.0 server showed that the CoCyP polypeptide molecular

weight was 22.24 kDa. The protein has a signal sequence (Residues 1-26) including six

Leu residues (Leu 9, Leu 11, Leu 15, Leu 17, Leu 21 and Leu 25). The putative cleavage

site of the signal sequence of the CoCyP protein is suggested between Ala26 and Lys27,

identical to that for the Arabidopsis ER-localized CyP5 (containing a potential cleavage

site between the signal peptide Ala23 and Lys24) (Price, et al., 1991; Saito et al., 1999).

The result of Subcellular location using the TargetP service indicated that the CoCyP

protein was a secretory pathway protein, similar to those of A. thaliana (NP_180557),

potato (ABB86261) and tomato (AAW22880); and their obtained scores were 0.960,

0.981, 0.984 and 0.978, respectively. The results support the point that a potentially

significant feature of the N-terminal signal peptides of the predicted secretory pathway

(SP) CyPs is the presence of Leu-rich residues (Price et al., 1994; Romano et al., 2004).
Fig. 1. The PCR result of the cDNA clone
M: 100 bp plus DNA ladder; 1: the MonoClone rzots0_000167.y1.scf

Fig. 2. The nucleotide sequence and the deduced amino acid sequence of the full-length CyP cDNA of *C. oleifera*.
The initiated nucleotide sequences, the translational termination codon and the polyadenylation signal are boxed. The 5'-untranslated region and the 3'-untranslated region are underlined. The waving line indicates polyadenylation, and the stop codon is indicated by asterisk (*).
Fig. 3. Alignment of the amino acid sequence of CoCyP with sequences of other organism CyPs. Abbreviations for them are: CoCyP (this study), AtCyP5 (A. thaliana CyP5, CyPB; NM128850), TaCyPB (T. aestivum; ACF49500), tomato (M55019), maize (M55021), B. napus (Brassica napus; M55018), hCyPB (human CyPB; M60857) and hCyP-18a (human CyPA; NP_066953). High conserved amino acids are highlighted, and hyphens indicate gaps. Residues involved in CsA contact are indicated in Asterisks (*) (Pflügl et al., 1993), including the key W residue (Bossard et al., 1991); residues in close contact with a tetrapeptide Ala-Ala-Pro-Ala (pep) are marked with a cross (+) (Hirtzlin et al., 1995). Residues involved in enzyme activity are indicated with # (Romano et al., 2004). The seven-amino-acid insertion is boxed.

Alignment of the putative CoCyP amino acid sequence with those of other CyPs showed that they shared a significant degree of identity with the exception of the N-terminal and C-terminal regions (Fig. 3). The amino acid sequence of CoCyP has an overall identity of 84, 69, 68, 66, 63 and 59% with those of A. thaliana CyP5 (CyPB, NM128850), TaCyPB (T. aestivum; ACF49500), tomato CyP (M55019), B. napus (Brassica napus; M55018), maize CyP (M55021), human CyPA (NP_066953) and human CyPB (M60857), respectively. Noticeably, there was a highly conserved central core domain resided in these sequences. Three consensus sequences FHRl(V)IP(K)XFML(I/C)QGGDFT, PGX-L(V)SMA(R)NA(S)GXDN(T)NGSQQFII and WLDGKHVVFG are found in the middle part of these proteins contain the CsA-binding residues (Pflügl et al., 1993), the tetrapeptide Ala-Ala-Pro-Ala-binding residues (Hirtzlin et al., 1995), and those required for PPIase catalysis (Romano et al., 2004). There is also an insertion of seven amino acids KSGKPLH (residues 82-88) in the putative protein, which was pointed out as a characteristic of organisms containing chloroplasts and cytosolic CyPs, but of unknown function (Gasser et al., 1990; Gottschalk et al., 2008; Wu et al., 2009).
Plasmid construction and expression: The identical PCR amplification products of a size of 621 bp of the nine recombinants were generated using the recombinant plasmid DNA as template (Fig. 4). The results of both the PCR amplification and DNA sequencing (result not shown) verified that the conducted recombinant was correct. As shown in Fig. 5, the pET-30b (+)/CoCyP recombinant was expressed and yielded a protein of about 21kDa below the predict molecular weight of 22.24 kDa. The reason is unknown, and it might be that the N-terminal signal sequence might have been removed in the *E. coli*-expressed CoCyP protein due to the absence of an ER-retention signal KDEL or AIAKE (Saito *et al*., 1999).

Discussion

In this study, we have isolated CoCyP, a gene the product of which contains a hydrophobic N-terminal signal peptide predicted as ER-targeting signal from *C. oleifera* cDNA library. And the cleavage site appears conserved in ER-localized CyPs in plant e.g., Ala-Lys in *C. oleifera* and *Arabidopsis* in the N-terminus. This signal sequence of CoCyP has characteristics common to those found in the ER-localized CyPs (commonly called CyPB) (Saito *et al*., 1999; Wu *et al*., 2009). Moreover, CoCyP shares highest identity of amino acids with AtCyP5 (AtCyPB) (84%) and TaCyPB (81%), and the predicted result by the TargetP service also shows that CoCyP protein is a secretory pathway protein. Taken together, these evidences suggest that the CoCyP protein belongs
to CyPB, which is found within the ER or in distal compartments along the secretory pathway (Price et al., 1991) due to the presence a putative ER-targeting signal sequence at its N-terminus and the absence of an ER-retention signal KDEL or AIAKE at its C-terminus (Saito et al., 1999).

The data presented here (Fig. 3) suggests that the CyP diversity can arise by addition of N-terminal of C-terminal extensions flanking the central CyP core domain (Price et al., 1991). High conservation in the PPlase catalytic sites suggests that PPlase activities are both an active component of its functions and important cellular function for CyPs (Godoy et al., 2000; Roman et al., 2004); and the variation of PPlase activity of the ER-directed CyPB might affect storage protease folding and accumulating and assist in the folding and transport of newly synthesized seed protein (Sharma & Singh, 2003; Johnson et al., 2004; Wu et al., 2009). Furthermore, according to the research of Galili et al., (1998), in the process of oleosin protein and oil bodies synthesis, the localized-ER CyP protein may assist in the folding and maturation of proteins and the deposition of lipid storage compounds owe to PPlase activity and molecular chaperone. Collectively, it could be postulated that the CoCyP protein may play significant roles in the processes of the seed storage and lipid synthesis.

The CoCyP gene might be helpful to improve the activity of peroxidase and play some roles in protecting cells against ROS damage combination with the current studies (Peltier et al., 2002; Motohashi et al., 2003). It might act as chaperone to help protecting proteins of C. oleifera seeds against hydrolysis or aggregation and to participate in response to stresses, perhaps by regulating appropriate folding/repairing damage proteins or in the signal transduction processes due to its PPlase activity (Chen et al., 2007). Additionally, the expressed CoCyP protein in this study can lay a basis for further revealing the relationship between structure and function. The study can be also as an exploring research for immunosuppressant drugs and therapy cancer drugs due to human CyPB peptides play important roles in special immunotherapy of Leukemia and epithelial cancers (Gomi et al., 1999).

Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant No. 31070603) and the Graduate’s Scientific Research Fundation of Central South University of Forestry and Technology (Grant No. 2009bx01).

References


(Received for publication 22 March 2009)