Gene isolation, heterologous expression, purification and functional confirmation of sesamin synthase from Sesamum indicum L.

Kishan Chandra, Abhipsa Sinha, Neelakantan Arumugam*

Department of Biotechnology, School of Life Sciences, Pondicherry University, Pondicherry 605014, India

1. Introduction

*Sesamum indicum* is an important ancient oilseed crop that serves as major source of the antioxidant lignan sesamin [1]. The lignans, in turn, are dimers of phenyl propane units formed by a covalent linkage between the central carbons of their propyl side chains. Sesamin exhibits clinically significant properties such as lowering of cholesterol, control of hypertension and induction of apoptosis in cancer cells [2–4]. The compound also protects liver from the damaging effects of ethanol and carbon tetrachloride, exhibits anti-inflammatory effect, inhibits vascular superoxide production and induces nitric oxide production in vasculatures [5–10]. In view of these properties there is a great demand for Sesamin [11]. Chemical synthesis of the compound has not been successful [12]. Even though there are many natural sources identified for sesamin, the oil and seed from *S. indicum* alone continues to be the major source of the compound [13]. Sesamin merely ranges from 0.2% to 0.8% of by the dry weight of seeds in the existing varieties of the sesame crop [14–16]. Genetics of sesamin production indicates that it could be a polygenic trait [17,18]. One of the options available is to enhance sesamin synthesis in live plants of *S. indicum* by metabolic engineering [13]. Sesamin follows a complex biochemical pathway [19]. In Sesamin, lignan biosynthesis starts with phenylalanine which is sequentially converted to sesamin and related lignans. Pinoresinol is one of the intermediate in the pathway synthesized from coniferyl alcohol by a reaction assisted by a ‘dirigent’ protein [20]. Piperitol synthase converts pinoresinol to piperitol which is then converted to sesamin by sesamin synthase [21,22]. Recently, following a reverse genetics approach, it was shown that the two step conversion of pinoresinol to sesamin was catalysed by a single enzyme that has dual catalytic activity and the enzyme accordingly was named as piperitol-sesamin synthase (PSS). The PSS belongs to cytochrome P450 (CYP) group of enzymes and in *Sesamum* it is encoded by the gene CYP81Q1, consequently the enzyme was also termed as CYP81Q1 enzyme [23]. The level of expression of CYP81Q1 often differed with genotype of the plant. For example, the PSS from the sesame variety Gomazou exhibited a higher expression than Kin-goma with concomitant increase in sesamin content. Therefore it was inferred that genotypic difference of CYP81Q1 is one of the important factor affecting sesamin content in sesame [17]. *CYP81Q3*, a homolog of the gene from *S. alatum* lacked sesamin biosynthesis activity and
accordingly the species was designated as a sesamin deficient one. CYP enzymes constitute a super family of enzymes that are involved in diverse primary and secondary metabolic pathways in plants [24]. The divergent CYP enzymes in fact are responsible for much of the phytochemical diversity seen in almost all organisms including prokaryotes, eukaryotes and other higher animals. Many of the CYP enzymes are also involved in degradation of xenobiotics in mammals including human [25]. In order to apply biotechnology for enhancing sesamin content in S. indicum, as a preliminary prerequisite the present study was undertaken to clone, express and characterize CYP81Q1 enzyme in a heterologous system. Earlier, Hemati and Nayeri [26] described amplification and cloning of CYP81Q1 gene from genomic DNA of S. indicum. But its expression and activity was not reported. However, using bioinformatics tools, the authors predicted only the structure but not function of the protein. Gene if amplified from genomic DNA, the CYP81Q1 genes e.g., CYP81Q1, CYP81Q2, CYP81Q3 and CYP81D2in general contained an intron and therefore remain non expressive in prokaryotes [23]. Therefore, we intended to follow the cDNA approach to enable isolation and expression of the gene in prokaryotic system. CYP81Q1 is an amonooxygenase and therefore requires a reduction reaction to proceed prior to its function. PSS meets this requirement by pairing with the membrane bound cytochrome P450 reductase1 (CPR1) that helps transfer of electrons efficiently from the co-factor NADPH to heme unit of the CYP of the CYP81Q1 [27,28]. CPR shavellavin groups that mediate electron transfer in a highly regulated manner [29]. Herewe present our observation on cDNA cloning, expression and characterization of CYP81Q1 and sesamin CPR1 genes from an Indian variety of S. indicum.

2. Material and method

2.1. Plant material

Seed of Sesamum indicum varietyTMV3, a black seeded one, used in the present study was obtained from National Bureau of Plant Genetic Resource (NBGPR), New Delhi. The crop was maintained by cultivation in the experimental garden of our Department. For cultivation of sesame the procedure as described in Handbook of Agriculture (1996), ICAR, New Delhi was followed. Sesame seeds were sown in rows by hand during the Rabi season (December) and harvested in April. A row to row distance of 50 cm and a plant to plant distance of 20 cm were maintained. Weeds were physically removed and the crop was irrigated on daily routine till seed germination and at regular interval during flowering till pod formation. Fruits of the plant, known as pods, were harvested at different stages of development. Based on the size and colour, the pods were grouped into five classes to correspond to their five developmental stages (Table 1). Seeds were collected separately from each of the stages and split into two sets. One set of the seed was utilized for lignan analysis by HPLC and the other set was used for RNA isolation.

2.2. Lignan analysis in developing seeds

200 mg of seeds harvested from different stages of pods were pulverized in liquid nitrogen and the powder obtained was extracted in 1 ml of 80% methanol in an eppendorf tube. The mixture was centrifuged at 5585xg for 3 min and the supernatant was collected to serve as crude lignan extract. The crude samples thus obtained were analysed using Shimadzu HPLC system equipped with C-18 reverse phase column and UV-VIS detector at 290 nm. The mobile phase consisted of 65:35 acetonitrile - water with a flow rate of 0.8 ml/min [30]. Authentic standards of Sesamin (Cat No. SMB00705) and Pinoresinol (Cat No. 40674) were purchased from Sigma-aldrich, USA. HPLC grade acetonitrile (Cat No. AS029) and Water (Cat No. AS077) were purchased from Himedia, India. The lignan content of the harvested sesame seed were quantified following the method described in Dar et al. [16].

2.3. Isolation of RNA, first strand synthesis, cDNA amplification and cloning of CYP81Q1 and CPR1

RNA was isolated from developing seeds collected from stage-4 pods, using RNeasy® Plant Mini Kit (Qiagen) and converted to cDNA using the First strand cDNA synthesis kit (Invitrogen) following the manufacturer’s protocol. Prior to RNA isolation about 100 mg of seeds were crushed using liquid nitrogen into powder in a pestle and mortar. The powder was solubilized in RLC buffer given in the kit. Based on the published reports specific primers were designed and synthesized for amplification of CYP81Q1 and CPR1 genes from the cDNA [23,31]. Gene specific forward and reverse primers were designed in such a way as to include BamH1 and Xhol restriction sites respectively, at 5’ and 3’ end of the genes to facilitate their cloning in right direction in the expression vector. The restriction sites are indicated by an underline in the primer sequence. The forward primer for amplification of CYP81Q1 gene consisted of 5’-ATCGGGATCCATGGAACCCA- and the reverse primer consisted of 5’-CTAGCTCGAGAACGTTG-GAAAACCTGAGCAAG-3’. Similarly for amplification of cDNA of CPR1, the forward primer consisted of 5’-ATCGGGATCCATGGAACCCA-CATCGGGAAAGC-3’ and the reverse primers consisted of 5’-CTAGCTCGAGTACACATCACAGGAACT-3’. Independent PCR’s were performed for the two genes using Q5 high fidelity Taq polymerase (NEB) for 30 cycles of 98°C for 30 s (denaturation), 60°C for 15 s (annealing) and 72°C for 1 min (extension). PCR product was purified using QIA quick PCR Purification Kit (QIAGEN) and quantified using a nanodrop. In two independent experiments the PCR products were cloned into the pET28a (Novogen) expression vector to obtain pET28a-CYP81Q1and pET28a-CPR1 constructs. While CPR1 construct was designed to have only N-terminal His tag, the CYP81Q1 was designed to have both N- as well as C- terminal His tags. They were then independently transformed into the expression host BL21-(DE3) codon plus (Stratagene) following manufacturer’s protocol. pET28a vectors have been designed to have 6xHis tags that get co-expressed with

Table 1

| Name of the Trait | Developmental Stage and Characteristic Feature | 1 | 2 | 3 | 4 | 5 |
|------------------|-----------------------------------------------|---|---|---|---|---|
| Pod colour       |                                               | Light Green | Light Green | Light Green | Dark Green | Brown |
| Pod length (cm)  |                                               | 1.2 ± 0.20  | 1.7 ± 0.15  | 2.2 ± 0.10  | 2.2 ± 0.15 | 2.2 ± 0.15 |
| Lignans (mg/g seed) |                                   | 2.4 ± 0.30  | 4.6 ± 0.50  | 3.8 ± 0.40  | 0.38 ± 0.20 | 0.29 ± 0.10 |
| Pinoresinol      |                                               | 0            | 0            | 0            | 0.92 ± 0.30 | 0.85 ± 0.10 |
| Sesamin          |                                               | 0            | 0            | 0            | 0.43 ± 0.10 | 0.27 ± 0.04 |
| Sesamolin        |                                               | 0            | 0            | 0            | 0.43 ± 0.10 | 0.27 ± 0.04 |
the target protein to facilitate purification and identification of the target proteins (Novagen, TB055 8th Edition 02/99).

2.4. Culture of recombinant bacterium and protein induction

A Single colony of the transformed BL21-(DE3) codon plus harbouring recombinant plasmids for CYP81Q1 or CPR1 were grown overnight at 37 °C in 10 ml LB broth containing 50 μg/l kanamycin in a 50 ml test tube at 200 rpm. The cell suspension of the overnight grown cultures was inoculated in to 1 L fresh LB broth containing the antibiotic at the same concentration taken in a 2 L conical flask and incubated at 37 °C on a shaker set at 200 rpm. The OD600nm was monitored until it reached a value of 0.8 OD. To this culture 1 mM IPTG (Himedia Cat No. RM2578) was added and incubated for 6 h at 22 °C for induction of the protein production. Cells were then harvested by centrifugation at 3000×g for 15 min at 4 °C.

2.5. Isolation of recombinant protein and purification

The harvested cells were lysed by sonication in a buffer of pH 7.5 containing 50 mM Tris, 500 mM NaCl, 20 mM imidazole, 0.5 mM DTT and 5% glycerol with 0.5 mM PMSF. The lysed suspension was centrifuged at 16,000×g for 30 min. Supernatant was passed through 1 ml Ni-NTA packed affinity column (GE healthcare) and with 100 volumes of washing buffer of same composition as sonication buffer but with 40 mM imidazole. The bound protein was eluted using 20 mM HEPES pH 7.5, 250 mM NaCl, 5% glycerol, 300 mM imidazole, 1 mM DTT and 1 mM PMSF. Purified protein was dialysed in 20 mM HEPES buffer of same composition as that of elution buffer but devoid of imidazole [32].

2.6. Western blotting

SDS-PAGE of the purified protein was performed along with appropriate controls and protein bands were transferred onto Nitrocellulose membrane using semi dry transfer unit with 80 mV for 1 h. The membrane was then washed with the buffer for 30 min, transferred to a tray containing anti-His antibody (Cat No. 2365 T) procured from Cell Signaling Technology (CST) was used at 1:1000 dilution and incubated for 2 h. Subsequently the membrane was washed for 1 h in TBST buffer [33], transferred to a tray containing Anti-rabbit secondary antibody (CST 7074P2) and incubated for 2 h in shaking condition. The membrane was then washed in TBST for 15 min 4 times to remove traces of secondary antibody. The bands were visualized using ECL reagent (Biorad) and photographed.

2.7. Enzyme assay for PSS activity

The enzyme assay was performed by estimating the conversion of pinoresinol to sesamin by HPLC. The reaction mixture consisted of 200 μl of pH7.5 PBS buffer (10 mM), 5 μl each of Pinoresinol (1 mg/ml) and NADPH (Cat No. RM576, 50 mM), and 30 μl each of purified enzymes of CPR1 and CYP81Q1 or water instead. The mixture was incubated at 30 °C for 4 h with intermittent shaking. The reaction was terminated by adding an equal volume of ethylacetate. Following a modified protocol described earlier [30], 10 μl of the resulting mixture was injected into HPLC for lignan analysis.

3. Results

3.1. HPLC of seed lignans at different stages of seed development

Expression of many of the plant genes is developmentally regulated. In order to identify asuitable stage of pod for harvesting seeds for isolation of target mRNA, an indirect method of analysis was followed. The pod colour and pod length were first measured followed by lignan content in the seeds from five different stages of pod development analysed by HPLC (Table 1). Pods of the first three stages differed in their length but remained light green in colour. Pod of stages 3–5 showed no change in pod length but their colour changed from light green to dark green and then to brown. Seeds harvested from stages 1 to 4 were white and those of stage 5 were black in colour (Fig. 1). Lignan analysis of methanol extract of seeds from different stages of development by HPLC revealed that there is no trace of sesamin and sesamolin in seeds upto stage III, but showed a prominent peak with a RT (retention time) of 3.7 min which corresponded to the RT of pinoresinol standard indicating its synthesis during these stages (Fig. 2a).

HPLC of the extract from seeds of stages 4 and 5 showed prominent peaks for sesamin (RT 7.5 min) and sesamolin (RT 9.7) with considerable reduction in the area of the peak corresponding to pinoresinol (RT 3.7 min) indicating biosynthesis of the sesamin
and sesamolin taking place during these stages (Fig. 2b). The lignan content of the seeds at different developmental stages coincided with the difference in HPLC peak area observed (Table 1).

3.2. RNA isolation, cDNA synthesis, amplification and sequencing of CYP81Q1 and CPR1

Since sesamin appeared in stage-4, the seeds harvested from stage-4 pods were used for RNA isolation. The total RNA yield was 640 ng/mg of seeds with A260/A280 of 2.0 as recorded using nanodrop. First strand cDNA was synthesized from the isolated RNA using Invitrogen kit. Gene specific primers enabled amplification of CYP81Q1 and CPR1 from cDNA by PCR. Electrophoresis of the PCR products revealed that the size of CYP81Q1 was 1.5 kb and that of CPR1 was 2.1 kb (Fig. 3). Sequencing of the PCR products of the two genes confirmed their identity. Analysis of nucleotide phylogeny of the CYP81Q1 (TMV3) that we have cloned revealed that the sequence had 100% similarity to Sesamum indicum cultivar Karaj cytochrome p450 81Q1 (cyp81q1) gene (the NCBI accession number KP771974.1 Fig. 4) and CPR1 sequence available in NCBI data base (data not shown). The PCR products thus obtained from cDNA were cloned and expressed in heterologous system.

3.3. Expression and purification of recombinant enzymes

Transcriptional induction of pET28a vector harbouring CYP81Q1-transformed into BL21-(DE3) codon plus yielded 4.8 mg of PSS protein per litre of the bacterial culture. SDS-PAGE of the crude protein revealed presence of 55 kDa band corresponding to the molecular weight of PSS protein. An enriched band of protein was observed on purification with His-tag column (Fig. 5). A similar procedure followed for putative CPR1 yielded 3.9 mg of CPR1 protein (MW 72 kDa) per litre of bacterial culture (Supplementary Fig. 1). Thus the cloning and expression of CYP81Q1 and CPR1 genes from S. indicum was confirmed. Expression of the recombinant genes were further confirmed by western blot analysis of the fractions of bacterial lysates from un-induced, induced cultures and the affinity purified protein fraction. The result showed there is no expression of the recombinant CPR1 protein in the un-induced lysates but confirmed its presence in both induced and purified fractions (Supplementary Fig. 1). Thus the cloning of two genes in fusion with 6xHis, their over expression and purification of the recombinant protein was further confirmed.

3.4. In vitro assay for CYP81Q1 enzyme activity

HPLC of the standard pinoresinol and sesamin showed a RT of 3.7 min and 7.5 min, respectively (Fig. 6a and b). HPLC of the reaction mixture without the enzyme showed no trace of sesamin (Fig. 6c). The sesamin was also undetectable in the presence of CPR1 alone (Fig. 6d). The same is the case in the presence of NADPH alone or NADPH + CPR1 but in the absence of CYP81Q1. However, when CYP81Q1 enzyme (PSS) alone was supplemented along with NADPH and pinoresinol there appeared a small peak corresponding to sesamin indicating occurrence of some base level activity in the synthesis of sesamin (Fig. 6e). This showed that the purified recombinant protein of PSS, obtained from bacteria is capable of converting substrate pinoresinol to the product sesamin. When both CYP81Q1 and CPR1s present together in the reaction mixture, there was considerable increase in CYP81Q1 activity leading to a sharp increase in sesamin produced (Fig. 6f). Thus cloning and expression of functional CYP81Q1 and CPR1 enzymes of sesame is confirmed.

4. Discussion

Sesamin belongs to a group of bioactive compounds known as lignans that are synthesized through phenylpropanoid pathway.
There are several reports available on isolation of sesamin and about fifteen other lignans from *S. indicum* [16,34–36]. Most of these lignans are fat soluble aglycons that on extraction elute in the oil and the remaining is glycosylated that can be isolated mainly from the oil free meal. However, isolation and characterization of enzymes involved in sesamin biosynthesis for biotechnological application has been very few. CYP81Q1 was earlier expressed in heterologous yeast system or in Baculovirus / Sf9 cell system with the enzyme activity assayed using microsomal fractions from the recombinant cells [23,37]. Recently Hemati and Nayeri [26] reported amplification and cloning of CYP81Q1 gene from the genomic DNA isolated from an Iranian *S. indicum* variety Karaj-1. But its expression, either in vitro or in vivo, has not been demonstrated. It may be relevant to mention that we had also used genomic DNA for amplification CYP81Q1 and obtained a 1.7 kb size PCR product that contained a 171bp intron. Expression of genes in bacterial system would be more preferred approach for industrial production of enzymes for which one needs an intron free gene construct. Following cDNA approach we showed here cloning and expression of the intron free CYP81Q1 gene in *E. coli* and demonstrated the functioning of the purified recombinant enzyme by in vitro enzyme assay. We also demonstrated here requirement of a reductase for enhanced functioning of the enzyme. Novelty of our paper lies in the fact that we have established a system for sustainable production of both CYP81Q1 (PSS) and CPR1 (reductase) proteins in a bacterial system for conversion of pinoresinol to sesamin. The system paves the way for production of piperitol-sesamin synthase (PSS) and its product(s) thereof at industrial scale. It is relevant to point out here that CYP enzymes have recently been identified as targets for directed evolution by mutation to impart new and/or enhanced catalytic activities that may be followed by simple expression in *E. coli* [38].

Recombinant CYP enzymes (monooxygenases and hydroxylases) that are CPR dependant have been assayed in different ways. One of them is to use a mixture of soluble fraction of the lysate from two different recombinant cells, for example in case of catalysis when CYP73A25 (from cotton) and CamCPR (CPR from *Camptotheca acuminata*) were brought together [27]. In cases where the recombinant proteins are expressed as inclusion body, the insoluble fraction of the lysate was solubilized and used as shown for CYP97C27 and CPR1 from *Croton stellatopilosis* [28]. Alternately the recombinant chimeric CYP-CPR is immobilized on nanodisc and assayed as in heterologous expression of chimeric CYP725A4-17α-L-CPR of *Taxus cuspidata* [39]. Earlier the enzyme obtained using heterologous expression of CYP81Q1 of sesame was assayed with the help of yeast microsomal fraction in which some membrane bound native reductase might have served the purpose [23,31]. Here we showed expression of the two plant proteins in...
bacterial system and followed a procedure modified from Qu et al. [27] and Sinupuchee et al. [28] to obtain reproducible results on enzyme activity. We also showed that the presence of reductase is a must for enhanced catalytic conversion of pinoresinol to sesamin in vitro.

In S. indicum var. TMV3, sesamin appears to be synthesized from stage 3 to 5 of the pod development. This correlates with recent report on S. indicum wherein it was shown, by RT-PCR, that maximum amount of CYP81Q1 transcripts accumulated by 25th day of flowering [40]. Unlike the earlier report we found that there is synthesis of pinoresinol in the first three stages as well and there is lack of appearance of piperitol (by HPLC) when sesamin was synthesized from pinoresinol in vitro [23]. We attribute such phenomenon to genotype difference between the varieties as has been suggested earlier [17]. Methylenedioxy bridge formation is an important step in the biosynthesis of several important secondary metabolites [41]. There can be either one (berberine, safrol, yatein) or at the maximum two (sesamin, sanguinarine, podophyllotoxin, stylopine) such bridges that are formed in the molecules with their formation catalyzed by a P450 dependent monooxygenases [42]. If there is more than one such bridge, formation of each bridge is catalyzed by different P450 enzymes in successive reactions [43,44]. The PSS enzyme encoded by CYP81Q1 dealt with herein is the first case where the same enzyme catalyzes the formation of two methylenedioxy bridges in two successive reactions leading to synthesis of sesamin in S. indicum [23]. Further the domestication of the crop appears favour selection of sesame crop against sesamin biosynthesis as there are many wild relatives of the only cultivated S. indicum such as S. malabaricum and S. melayanam characterized by higher sesamin content [40]. Therefore it should be possible to augment sesamin content in cultivated sesame by breeding and further enhanced by metabolic engineering.

Recently by homology modelling and site directed mutagenesis it was shown that residue A308 is important for catalysis leading to formation of characteristic methylenedioxy bridge and a mutation in this residue resulted in inhibition of PSS [37]. Precise catalytic role of the amino acid residues is yet to be known. For the enzymes whose structures have already been elucidated there is a possibility that they can be developed into molecules having clinical or agricultural significance. Such details for the one’s involved in phenylpropanoid pathway are yet to be elucidated [45]. Further analyses such as computational alanine scanning along with comprehensive in vitro studies are required for understanding a complete picture of the mode of action of this enzyme. This may facilitate application of modern methods of genetic engineering such as Crisper/cas9 for increasing sesamin content in the crop plants [46]. In conclusion, it may be stated that the CYP81Q1 and CPR1 genes from S. indicum have been successfully cloned, expressed in E. coli, functional recombinant enzyme successfully isolated from the heterologous system and its activity demonstrated. Through metabolic engineering it should be possible to modify these genes for optimal expression of the desired secondary metabolite for application towards human welfare [44,38,47].

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

NA acknowledges DBT project BUILDER No. BT/PR14554/INF/22/125/2010 and DST-FIST, for infrastructure. KC acknowledges UGC for NFO Research Fellowship.
