DOI: 10.1007/s10535-013-0344-y BIOLOGIA PLANTARUM 57 (4): 635-645, 2013

Functional expression and subcellular localization of pea polymorphic isoflavone synthase CYP93C18

M. PIČMANOVÁ[1], D. REŇÁK[1], J. FECIKOVÁ[1], P. RŮŽIČKA[1], P. MIKŠÁTKOVÁ[2], O. LAPČÍK[2], and D. HONYS[1,3]*

Laboratory of Pollen Biology, Institute of Experimental Botany of the ASCR, Rozvojová 263, CZ-16502 Prague, Czech Republic[1]
Department of Chemistry of Natural Compounds, Faculty of Food and Biochemical Technology, Institute of Chemical Technology in Prague, CZ-16628 Prague, Czech Republic[2]
Faculty of Science, Charles University in Prague, CZ-12844 Prague, Czech Republic[3]

Abstract

Isoflavone synthase (IFS; CYP93C) plays a key role in the biosynthesis of phenolic secondary metabolites, isoflavonoids. These compounds, which are well-known for their benefits to human health and plant defence, are produced mostly in legumes. However, more than 200 of them have been described in 59 other plant families without any knowledge of their respective *IFS* orthologue genes (with the sole exception of sugar beet). In this study, we selected IFS from *Pisum sativum* L. (CYP93C18) for functional expression. *CYP93C18* was isolated, cloned, and introduced into *Arabidopsis thaliana*. The presence of the gene was shown by Southern blot analysis and its expression in the transgenic *Arabidopsis* was proven by RT-PCR and Western blots. The functional activity of the heterologous IFS was verified by HPLC-MS analysis of the metabolite levels: the isoflavone genistein and its derivatives tectorigenin and biochanin A were detected in the overexpressing lines. In addition, 35S::CYP93C18::GFP fused proteins were transiently expressed in the leaves of *Nicotiana benthamiana* and the localization of the GFP signal was observed on the endoplasmic reticulum using confocal microscopy which is consistent with the data from the literature and with our *in silico* predictions. The putative mode of attachment of IFS to the endoplasmic reticulum membrane is suggested. The undemanding methodology presented in this paper is applicable to the functional analysis of newly-identified isoflavone synthase genes from various species.

Additional key words: *Arabidopsis thaliana*, cytochrome P_{450}, endoplasmic reticulum, isoflavonoids, *Nicotiana benthamiana*, *Pisum sativum.*

Introduction

Isoflavone synthase (also called 2-hydroxyisoflavanone synthase; IFS or 2-HIS, respectively), a member of the cytochrome P_{450} superfamily, plays a key role in the biosynthesis of a group of plant secondary metabolites, isoflavonoids (Fig. 1). The literature deals extensively with positive effects of these well-known phytoestrogens on human health including cancer prevention and mitigation of menopause symptoms as well as other effects associated with their consumption (see Adams 1995, Ososki and Kennelly 2003, Cornwell *et al.* 2004, Dixon and Pasinetti 2010). In addition, isoflavonoids have considerable importance to the plants themselves, particularly as phytoalexins and chemoattractants in rhizobial symbiosis.

As far as isoflavone synthase is concerned, this enzyme belongs to the subfamily CYP93C of cytochromes P_{450}, as it was established in 1999 - 2000 by three independent groups on the basis of genomic studies (Akashi *et al.* 1999, Steele *et al.* 1999, Jung *et al.* 2000). Yu and McGonigle (2005) called IFS an "intriguing

Received 13 November 2012, *accepted* 3 April 2013.
Abbreviations: IFS - isoflavone synthase; KLH - keyhole limpet hemocyanin; P_{450} - cytochrome P_{450}; GFP - green fluorescent protein.
Acknowledgments: The authors thank Dr. E. Drdová and Dr. J. Petrášek (Institute of Experimental Botany ASCR, Prague) for their assistance with transient expression procedures and confocal microscopy, and Dr. R. Pleskot (Institute of Experimental Botany ASCR, Prague) for the molecular modelling work (Fig. 3F and Fig. 4E,F). This work was supported by the Czech Science Foundation (grant No. 525/09/0994) and the Ministry of Education, Youth, and Sports of the Czech Republic (grant No. OC10054).
* Corresponding author; fax: (+420) 225 106456, e-mail: david@ueb.cas.cz

enzyme“ due to its outstanding ability to catalyze both the hydroxylation of two flavanone precursors – liquiritigenin and naringenin – as well as the critical migration of the aryl group from the C2 to the C3 position on the chromene skeleton of the aforementioned flavanones. To date, 28 individual IFS-amino-acid sequences (including four IFS isoforms), displaying a very high degree of homology (over 95 %), have been described in 17 leguminous plant species (P450 Engineering Database, University of Stuttgart, http://www.cyped.uni-stuttgart.de) including IFS2 from *Trifolium repens* that has been recently expressed and functionally characterized in tobacco (Franzmayr *et al.* 2012) and IFS from *Pueraria montana* var. *lobata* (He *et al.* 2011) and *Pueraria candollei* var. *mirifica* (Wiriyaampaiwong *et al.* 2012). Moreover, *P. montana* var. *lobata* has been extensively screened for the presence of ESTs involved in the isoflavone biosynthesis pathway (He *et al*. 2011). Other additions for legume species have appeared in the GenBank, including *Glycine max, Medicago truncatula*, *Trifolium pratense*, *Vigna radiata* (Jung *et al.* 2000), *Lupinus luteus* (Madrzak and Narozna personal communication), and one IFS sequence from the legume *Cullen corylifolium* (Misra *et al.* 2010). The presence of IFS in non-leguminous species has hitherto been reported only in *Beta vulgaris* from family *Chenopodiaceae*, where two IFS isoforms have been found (Jung *et al.* 2000).

In connection with isoflavone synthase, the issues often discussed in the literature are concerned with its potential applications in the metabolic engineering of isoflavonoid biosynthesis to increase levels of isoflavonoids in human diet or to reduce their content in pastures. Here, we utilized *Arabidopsis thaliana* for the functional study of *Pisum sativum* isoflavone synthase. For this purpose, we chose one of the two known IFS genes from *Pisum sativum* – CYP93C18, GenBank accession No. AF532999 (Cooper *et al.* 2005). The *CYP93C18* cDNA was first discovered due to its up-regulation after the treatment of pea pods with the insect elicitor *Bruchin B* and was named on the basis of sequence similarity to *CYP93C*s known at that time (Cooper *et al.* 2005). Moreover, our further objective was to test a simple method for verifying the function of newly-identified genes orthologous to known *IFS* sequences from leguminous species. Given that there are 60 isoflavonoid-producing plant families known to date (Lapčík *et al.* 2006, Macková *et al.* 2006, Lapčík 2007), our approach was informed by five particular hypotheses: *1*) the number of families producing isoflavonoids might in fact be significantly larger than so far known; *2*) plants producing isoflavonoids must necessarily possess an enzyme catalyzing the migration of the aryl group on the chromene skeleton of flavanones; *3*) the genes encoding this enzyme are likely to display a homology with known IFS genes from leguminous species; *4*) the absence of such orthologues does not always need to signify the absence of the biosynthetic pathway leading to the production of isoflavonoids – and thus the possible existence of some other yet undiscovered enzyme should not be ruled out; and *5*) conversely, the presence of an orthologue to CYP93C does not necessarily guarantee the presence of isoflavonoids.

Materials and methods

Plants and DNA extraction: *Pisum sativum* L. plants were cultivated in a growth chamber at temperature of 21 °C, relative humidity of 50 %, a 16-h photoperiod, and irradiance of 150 μmol m^{-2} s^{-1}). Genomic DNA was extracted from 100 mg of young leaves using the standard CTAB DNA extraction protocol (Weigel and Glazebrook 2002). *Arabidopsis thaliana* L. ecotype Col-0 and *Nicotiana benthamiana* Domin., grown in the same growth chamber, were utilised for stable and transient expressions, respectively.

***CYP93C18* identification:** To isolate *CYP93C18* (GenBank acc. No. AF532999), DNA extracted from *P. sativum*. was used as the template for PCR amplification using *Phusion* polymerase (*Finnzymes*, Espoo, Finland) with *CYP93C18*-specific forward primer 5′-CACCATGTTGGTTGAACTTGCACTTGCTT-3′ and reverse primers 5′- TTAAGAGGAAAGCAATTTAG CTGCT-3′ (for stable expression) or 5′-TGCAGA GGAAAGCAATTTAGCTGCT-3′ where the stop codon was replaced by GCA triplet encoding alanine (for transient expression). PCR conditions were as follows: initial denaturation (94 °C for 7 min) followed by 35 cycles of denaturation (94 °C for 30 s), annealing (temperature gradient 50 - 65 °C for 30 s), extension (72 °C for 30 s), and a final extension (72 °C for 10 min).

Southern hybridization: Genomic DNA from *P. savitum* was digested with *Sac*I at 37 °C overnight and separated on a 1 % (m/v) agarose gel. Cleaved DNA was stained with ethidium bromide (Fig. 3*A*). Genomic DNA was depurinated (250 mM HCl), denatured (0.5 M NaOH + 1.5 M NaCl), neutralized (1.5 M NaCl + 0.5 M Tris-HCl, pH 7.5), and transferred onto nylon membrane. DNA was fixed by UV cross-linker at 1200 μJ for 30 s. The membrane was pre-hybridized in *Dig Easy Hyb* solution (*Roche Applied Science*, Mannheim, Germany) at 42 °C for 1 h. Full-length isoflavone synthase open reading frame (primers: forward 5′-CACCATGTTGGT TGAACTTGCACTTGCTT-3′ and reverse 5′- TTAAGA GGAAAGCAATTTAGCTGCT-3′) was amplified by PCR following the protocol in the *Dig* probe synthesis kit (*Roche Applied Science*) and used as a probe. The membrane was hybridized in *Dig Easy Hyb* solution at 42 °C overnight. After the hybridization, the membrane was washed in low stringency washing buffer [2× saline-

sodium citrate (SSC) + 0.1 % (m/v) sodiumdodecyl sulphate (SDS)] for 2 × 10 min at room temperature and then in high stringency washing buffer (0.1× SSC + 0.1 % SDS) at 65 °C for 2 × 15 min. The membrane was blocked in blocking buffer for 1 h and hybridized with anti-digoxigenin antibody (1:10 000) for exactly 30 min at room temperature. After incubation, the membrane was washed in *Dig* washing buffer at room temperature for 2 × 15 min and after that submerged in detection buffer for 10 min. The membrane was transferred into the hybridization bag for colour development and 1 cm^3 of *CDP Star*® chemiluminescent substrate (*Sigma-Aldrich*, St. Louis, MO, USA)) was added. The signal was exposed on X-ray film for 5 min.

Cloning and over-expression of *CYP93C18*: Full-length and stop-codon-missing *CYP93C18* DNA fragments (both including an intron between positions 903 and 991) were cloned using the *Gateway*™ technology according to the manufacturer´s instructions (*Life Technologies*, Carlsbad, CA, USA). The pENTR™/ D-TOPO® vector (*Invitrogen*, Carlsbad, CA, USA) was used to generate an entry clone; pGWB2 and pGWB5 vectors (obtained from Dr. Nakagawa, Shimane University, Japan) were used for the production of destination binary clones for stable and transient expressions, respectively. Both destination vectors harbour the 35S promoter upstream of the cloning site. The pGWB5 vector contains also the sGFP sequence tag for C-terminal fusion (Nakagawa *et al.* 2007). Expression constructs were transformed into *Agrobacterium tumefaciens* strain GV3101 by electro-poration (2.5 kV) and the transformants were selected on yeast extract broth (YEB)-plates with 50 μg cm^{-3} kanamycin and 100 μg cm^{-3} hygromycin.

Transformation of *Arabidopsis thaliana* and *Nicotiana benthamiana*: For stable expression, *Arabidopsis* plants were transformed with pGWB2::CYP93C18 construct using the floral dip method (Clough and Bent 1998). After 2 - 3 weeks, *ca*. 3 000 seeds of the transformed plants were selected on MS-plates with 50 μg cm^{-3} hygromycin and the recovered transformed seedlings were transferred into *Jiffy* peat pots (*Jiffy Products*, New Brunswick, Canada) for subsequent analysis. The selection was repeated over the next generation to obtain homozygous transformant lines. For transient expression and subsequent IFS subcellular visualization, the pGWB5::CYP93C18 plasmid was infiltrated into the *N. benthamiana* leaves by syringe injection of trans-formed *Agrobacterium* inoculum and analysed 2 - 3 d after the transformation (Li *et al.* 2009). As a positive control, *A. tumefaciens* (strain C58C1) carrying the pBIN m-gfp5-ER plant binary vector (obtained from Dr. Jim Haseloff, University of Cambridge, UK) was used. It encodes for the ER-localized GFP variant mGFP5-ER, a thermo-tolerant derivative of mGFP4-ER (Haseloff *et al.* 1997), and contains a C-terminal ER retention signal sequence (HDEL). As a negative control, *A. tumefaciens* (strain GV3101) carrying pGWG5 without any insert (expression of free sGFP) was employed. Both control proteins were transiently expressed in *N. benthamiana* leaves using the same protocol.

Analysis of transgenic *A. thaliana*: DNA was extracted from 100 mg of leaves individually collected from 10 T1 and 20 T2 randomly-selected transformants of two independent transgenic lines as described above, and control PCR with *CYP93C18*-specific primers and annealing temperature 58.3 °C was carried out. Total RNA was extracted from 100 mg of leaves of the plants with the confirmed presence of the *IFS* transgene using the *RNeasy* plant mini kit (*Qiagen*, Valencia, CA, USA) and DNase-treated. Reverse transcription (RT)-PCR was then performed using *ImProm*-II™ RT system (*Promega*, Madison, WI, USA). Both protocols were carried out according to the manufacturers' instructions.

A peptide sequence derived from an extremely conservative region of IFS was used to prepare an IFS-specific antibody, and its epitope surface localization was confirmed by protein modelling in *Modeller9v5* (Sali and Blundell 1993). The peptide NH_2-DPKYWKRPLEFRPER (according to Yu *et al.* 2000) conjugated with keyhole limpet hemocyanin (KLH) (*Davids Biotechnologie*, Regensburg, Germany) was then used for the immunization of a rabbit (Department of Biological Control, Institute of Physiology, AS CR, Prague, Czech Republic) according to their immuni-zation schedule. For Western blotting, total proteins extracted from 100 mg of leaves from confirmed *Arabidopsis* IFS transformants according to Wang *et al.* (2006) were used. Proteins were separated using 1-D SDS-PAGE and transferred onto a nitrocellulose membrane (*Serva*, Heidelberg, Germany) using a transfer apparatus *Bio-Rad Mini Trans-Blot*® cell (*Bio-Rad*, Hercules, CA, USA). The membranes were subsequently incubated with IFS antibodies collected 30 d after immunization. The primary antibody was diluted 1:200 in a mixture of Tris-buffered saline and *Tween 20* (TBST buffer) with 5 % (m/v) non-fat dry milk; the secondary anti-rabbit goat IgG conjugated to alkaline phosphatase (*Sigma-Aldrich*, St. Louis, MO, USA) was used in 1:15 000 dilution in a TBST buffer with 1 % (m/v) bovine serum albumine. The antibodies bound to IFS were detected by submerging the nitrocellulose membrane in an alkaline phosphatase (AP) buffer with 0.37 mM 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and 0.37 mM nitro blue tetrazolium (NBT) and blue-coloured bands were observed after the membrane dried.

Lyophilised shoots and dried seeds were ground in an electric mill (*IKA Werke*, Staufen, Germany) and 400 mg of the plant material with 8 cm^3 of 80 % (v/v) ethanol were left for 14 d at room temperature with occasional agitation. The extract was then centrifuged at 20 000 *g* and filtered through a 17 mm, 0.45 μm *PTFE* syringe filter. The extract (2.5 cm^3) was air-dried and dissolved in 0.3 cm^3 of 40 % (v/v) methanol for HPLC-ESI-MS analysis. The HPLC-ESI-MS tandem consisted of *Hewlett Packard 1100* HPLC series (*HP/Agilent*

Technologies, Santa Clara, USA) and an HP mass selective detector (*G1946A*), and was controlled by means of *ChemStation* software (revision *A 07.01*). A *LiChroCART*® 125-4 mm HPLC-cartridge with *Prospher*® *STAR RP-18* end caped (5 µm) was used as a column (*Merck*, Darmstadt**,** Germany). The mobile phases were used as follows: 40 % methanol (solvent A) and 100 % methanol (solvent B), both supplemented with 0.5 % (m/v) acetic acid. The following gradient was employed: A:B 0 min - 100:0; 5 min - 80:20; 15 min - 55:45; 20 min - 0:100; 27 min - 100:00, and stop at 27 min followed by 3 min post-run. The flow-rate was 0.8 cm^3 min^{-1}, the temperature of the chamber was set to 25 °C and the injection volume was 0.02 cm^3. The mass spectrometer was operating in the positive ESI mode. Individual isoflavones were identified by comparison of their retention times (t_R) and molecular ions $[M+H]^+$ with those of standards. Isoflavonoid standards were obtained from *Indofine* (Hillsborough, NJ, USA - biochanin A, daidzein, daidzin, formononetin, genistein, genistin, glycitein, glycitin, ononin, prunetin, puerarin, and sissotrin) and *PhytoLab* (Vestenbergsgreuth, Germany - tectoridin and tecto-rigenin). Isoformononetin was prepared by selective methylation of daidzein as described elsewhere (Lapčík *et al.* 1999).

The same procedures were also carried out with *P. sativum* and with wild-type *Arabidopsis* as positive and negative controls, respectively.

Analysis of transgenic *Nicotiana bethamiana*: *Agrobacterium*-injected spots on *Nicotiana benthamiana* leaves were excised and subjected to immediate observation by confocal laser scanning microscopy using a confocal microscope (*LSM Zeiss 5 Duo*, Jena, Germany) with the 488 nm laser line of the Ar-Kr laser.

Modelling CYP93C18 protein structure: The appropriate templates for homology modelling the CYP93C18 protein structure were found using the *Psipred* threading algorithm (Bryson *et al.* 2005). Several known 3D structures of mammalian P_{450} (*e.g*. rcsb code: 1tqn, 2hi4, and 3e6i) were utilized as templates. The three-dimensional model was then generated on the basis of the resulting alignment, by *Modeller 9v5* (Sali and Blundell 1993). The *Adaptive Poisson-Boltzmann Solver* (APBS) was used for the evaluation of the electrostatic properties of the modelled structure (Baker *et al.* 2001). Molecular graphics images were produced by the *UCSF Chimera* package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, USA (Pettersen *et al.* 2004).

Results

On the basis of the known *CYP93C18* sequence (Cooper *et al.* 2005), we performed PCR to amplify this gene from the genomic DNA extracted from *P. sativum* using *CYP93C18* gene-specific primers (spanning the open reading frame and adapted for *TOPO*® cloning). The complete sequence obtained was 98.7 % identical to that published by Cooper *et al.* (2005; to distinguish Cooper´s sequence from that obtained by us, the sequences are named CYP93C18a and CYP93C18b, respectively, in this paper). The amplified 1662-bp fragment contained a 87-bp intron localized between the nucleotides 903 and 991 (Fig. 2*A*). However, 21 nucleotides of the DNA sequence were different. These differences resulted in 9 amino acid divergence within the 524-bp long deduced protein (Fig. 2*B*). Seven of the amino acid mismatches discovered could be considered as unimportant due to the similarity in chemical properties of the amino acid pairs (Tyr/Ser, Lys/Arg, Leu/Met, Ile/Met, and Asn/Gln). In the two remaining cases – Val122/Asp and Gly177/Glu – the mismatch could be potentially more serious as it might affect intramolecular interactions in the folded protein and thus its correct catalytic function. However, as established later (see below), the functional activity of IFS was not affected. In order to determine whether *CYP93C18b* represented a newly-described pea IFS isoform or a polymorphism within *CYP93C18* locus, Southern hybridisation was performed (Fig. 3*B*). Full-length IFS probe hybridised to two IFS sequence fragments, thereby confirming that *P. sativum* genome contained two previously published IFS isoforms (Jung *et al.* 2000, Cooper *et al.* 2005). *CYP93C18a* (Cooper *et al.* 2005) and *CYP93C18b* (this study) sequences thus were very likely to belong to the same polymorphic locus CYP93C18.

To verify the activity of pea CYP93C18, the obtained sequence was cloned and introduced into *A. thaliana* plants. In the literature, *Arabidopsis* is generally considered to be an isoflavonoid pathway-free plant mainly due to the absence of a gene homologous to CYP93Cs, but Lapčík *et al*. (2006) detected low but measurable content of several isoflavonoids. Due to the fact that the isoflavonoid precursor naringenin (a flavanone also involved in the biosynthesis of anthocyanins) is present in *Arabidopsis*, this plant has been shown to represent a suitable model for functional tests for isoflavone synthase activity (Jung *et al.* 2000, Yu *et al.* 2000).

A. thaliana plants were transformed with CaMV 35S::*CYP93C18* constructs by floral dipping. Hygromycin-resistant transgenic plants and their progeny were then analysed at four different levels: DNA, RNA, proteins, and metabolites. The *CYP93C18* gene was present in genomic DNA extracted from transgenic *Arabidopsis* and *P. sativum* (positive control) but not in wild-type *Arabidopsis* (negative control, Fig. 3*C*). The introduced CYP93C18 transgene was actively transcribed in *Arabidopsis* as demonstrated by RT-PCR of CYP93C18 mRNA (Fig. 3*D*). To confirm the correct

expression of the isoflavone synthase protein, immunoblotting with an IFS-specific polyclonal peptide antibody according to Yu *et al.* (2000) was performed, The presence of non-specific bands notwithstanding,

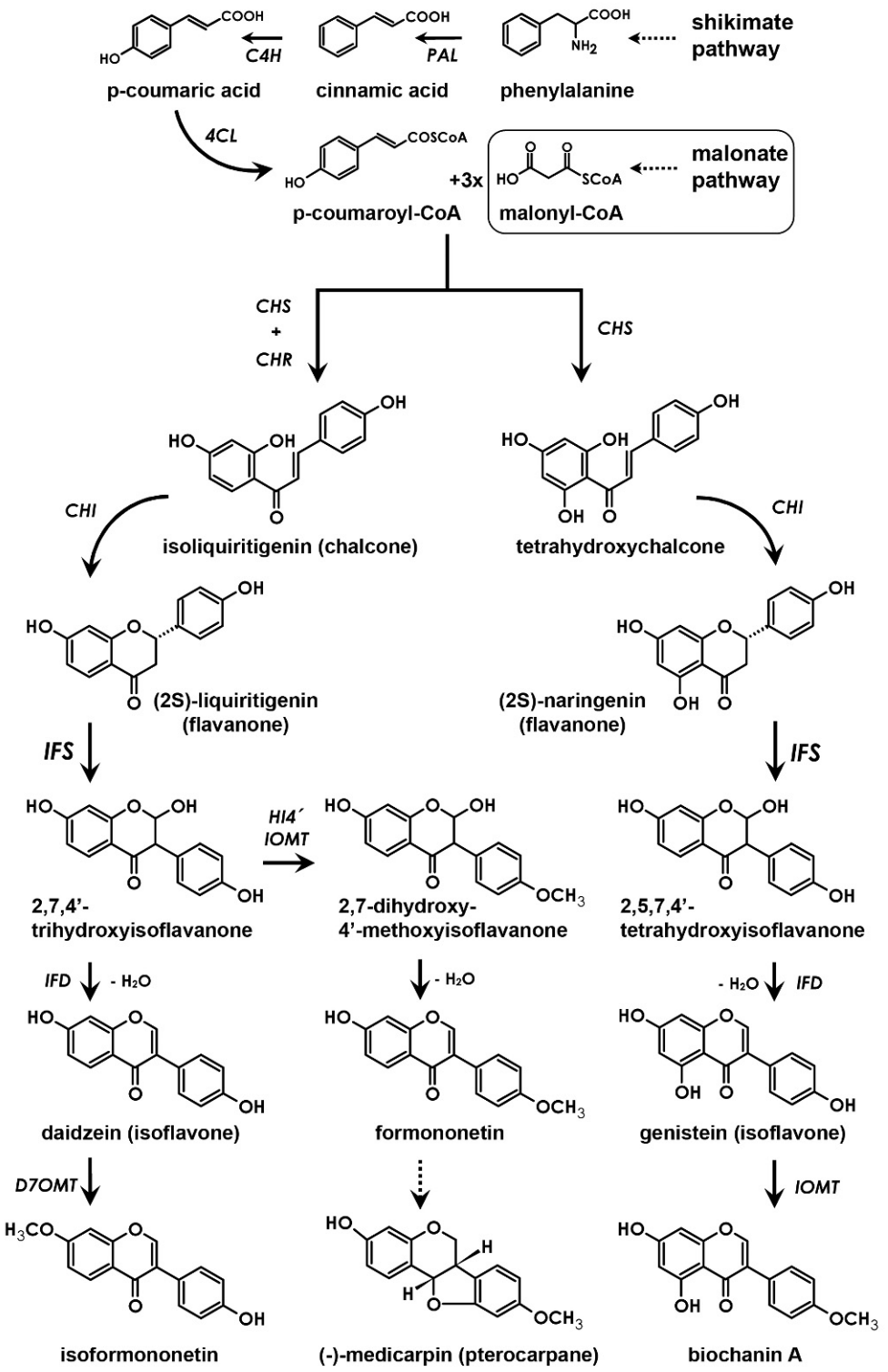


Fig. 1. Schematic description of the isoflavonoid biosynthetic pathway. PAL - phenylalanine ammonia lyase, C4H - cinnamate-4-hydroxylase, 4CL - 4-coumaroyl:CoA-ligase, CHS - chalcone synthase, CHI - chalcone isomerase, IFS - isoflavone synthase, IFD - isoflavone dehydratase, IOMT - isoflavone *O*-methyltransferase, HI4´IOMT - hydroxyisoflavanone 4´-*O*-methyltransferase, D7OMT - daidzein 7-*O*-methyltransferase (adapted from Crozier *et al.* 2006 and Veitch 2007).

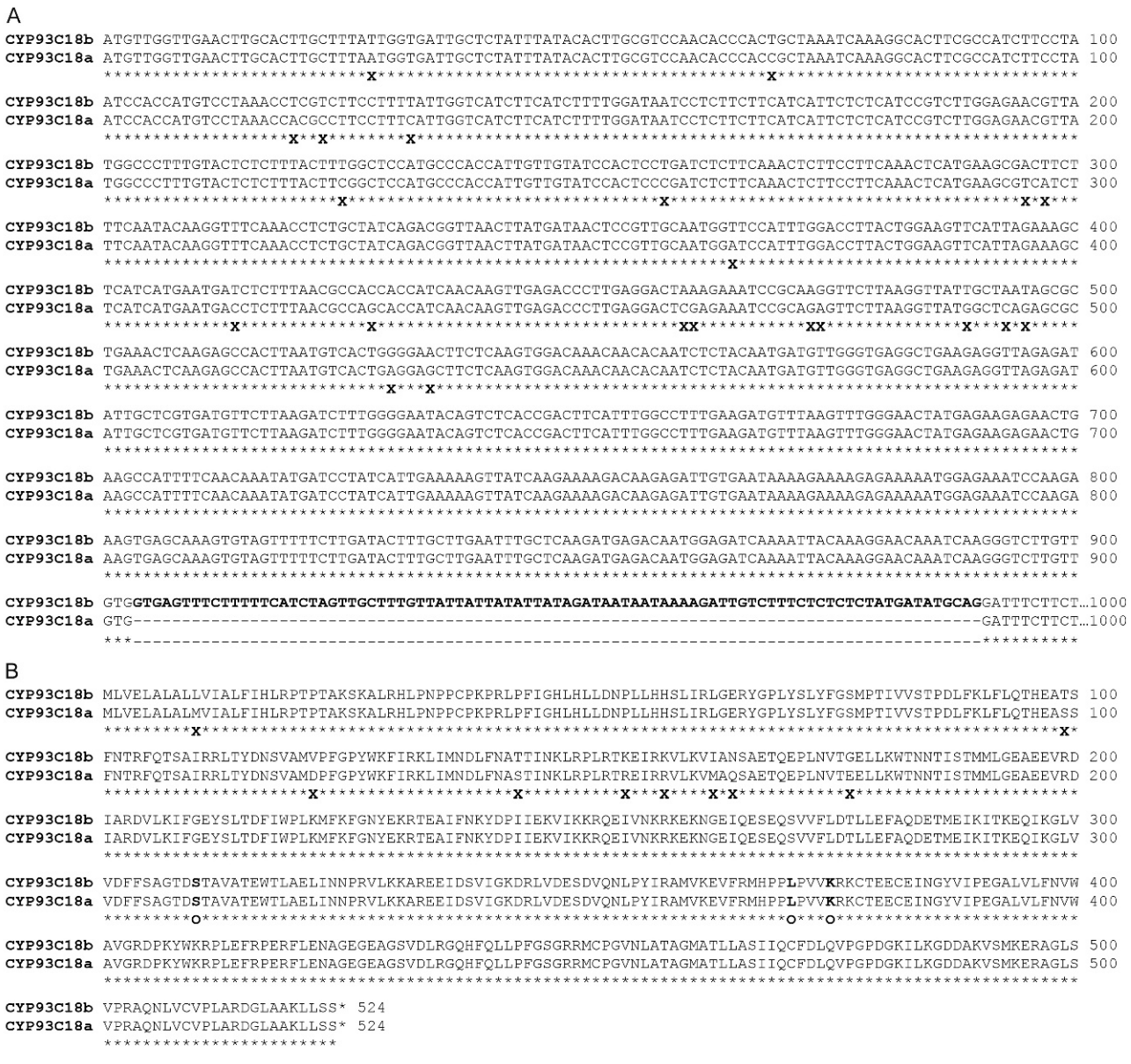


Fig. 2. *ClustalW* alignment of CYP93C18a and Cooper´s CYP93C18b nucleotide (*A*) and deduced amino-acid (*B*) sequences. The alignment is given for the first 1 000 nucleotides. The remaining parts of both sequences were identical. The discovered mismatches are marked with "x", the 87-nucleotides long intron is highlighted *in bold* and the three amino-acid residues essential for aryl migration (Sawada and Ayabe 2005) are marked with "o".

the antibody reacted with the 59.4 kDa polypeptide corresponding to the estimated size of the CYP93C18 protein in transgenic *Arabidopsis* and *P. sativum* but not in wild-type *Arabidopsis* (Fig. 3*E*,*F*). A similar problem associated with the non-specific binding the aforementioned IFS antibody has already been described by Jaganath (2005) and can be ascribed to the fact that polyclonal antibodies raised against the KLH:peptide conjugates can cross-react with various plant antigens in Western blot analysis, thus causing the non-specificity of the immunodetection (Oulehlová *et al.* 2009).

In order to determine the isoflavone synthase activity of CYP93C18, extracts from whole shoots and seeds of the plants were analysed by HPLC-MS for the presence of isoflavonoids in the transgenic *Arabidopsis* despite the absence of previous stress induction. This test represented the main indicator of the correct function of introduced CYP93C18. Using 15 different standards, three isoflavonoids were detected in transgenic *Arabidopsis* shoots and seeds: genistein and its methylated derivatives, tectorigenin (4′,5,7-trihydroxy-6-methoxy-isoflavone) and biochanin A (Table 1). Similarly, isoflavonoids glycitin, daidzein, daidzin, genistein, genistin, and prunetin were found in *P. sativum* (Table 1). Low HPLC-MS signals with t_R and $[M+H]^+$ corresponding to the standard of prunetin were observed

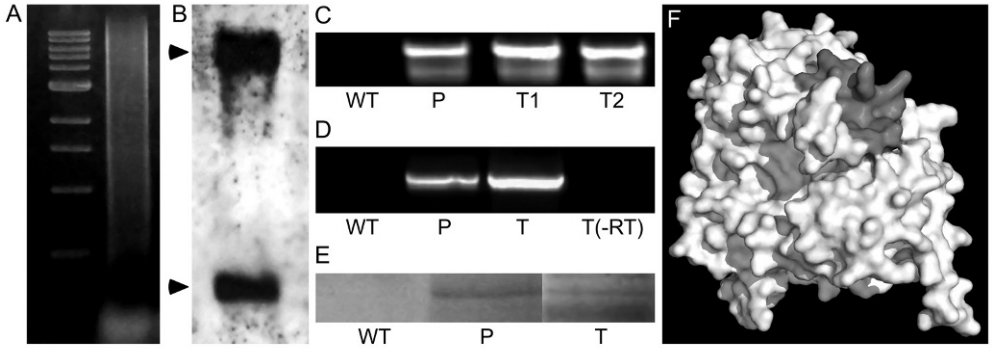
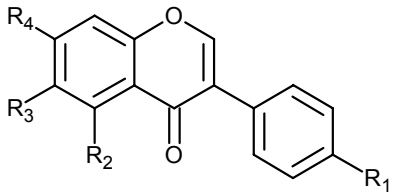


Fig. 3. Southern hybridisation analysis of *P. sativum* isoflavone synthase genes and the analysis of transgenic *A. thaliana*. *A* - Ethidium bromide stained pea genomic DNA (15 μg of genomic DNA were digested with *Sac*I and electrophoresed on 1 % agarose gel; left column - DNA molecular mass marker, right column - cleaved genomic DNA used for Southern blotting). *B* - Southern hybridization analysis of pea genomic DNA – presence of two IFS isoforms (15 μg of genomic DNA was digested, separated on 1 % agarose gel and transferred onto membrane. The full-length isoflavone synthase gene was used as a probe and the hybridization was performed at 42 °C. *C* - Analysis of transgenic *Arabidopsis thaliana*: *CYP93C18* (1662 bp long) is present in transgenic *Arabidopsis* and pea but not in wild-type *Arabidopsis*. *D* - RT-PCR of *CYP93C18* mRNA. The cDNA (1575 bp long) is present in transgenic *Arabidopsis* and pea but not in wild-type *Arabidopsis*. *E* - 1-D SDS-PAGE immunoblot of the total protein of the plants examined. The CYP93C18 protein (59.4 kDa) can be assumed to be present in transgenic *Arabidopsis* and pea but not in wild-type *Arabidopsis*. *F* - The model of CYP93C18 with the grey region indicating the epitope of the peptide NH_2-DPKYWKRPLEFRPER. This peptide, used for IFS-specific antibody production, is localized in a highly-accesible site of CYP93C18. WT - wild-type *Arabidopsis,* P - pea, T1, T2 - two generations of transgenic *Arabidopsis,* T - transgenic *Arabidopsis,* T(-RT) - transgenic *Arabidopsis*, negative control (no RT).

Table 1. HPLC-MS analysis of isoflavonoid content in the ethanol extracts from shoots and seeds of transgenic *A. thaliana* and wild-type *A. thaliana,* and from seeds of *P. sativum* (m/z - mass to charge ratio, Glc - glucose). Selective ion monitoring was used as follows: SIM1 - the molecular ions $[M+H]^+$ 433, 463, and 447 were monitored from 2 to 11.5 min, and $[M+H]^+$ 301 was monitored from 11.5 to 20 min; SIM2 - the molecular ions $[M+H]^+$ 417 and 431 were monitored from 2 to 8.5 min, and $[M+H]^+$ 255, 269, 271, and 285 were monitored from 8.4 to 20 min. Isoflavonoid standards used (MS method; $[M+H]^+$; retention time): puerarin (SIM2; 417; 2.436 min), daidzin (SIM2; 417; 3.235 min), ononin (SIM2; 431; 7.739 min), daidzein (SIM2; 255; 9.354 min), glycitein (SIM2; 285; 9.867 min), genistein (SIM2; 271; 11.974 min), isoformononetin (SIM2; 269; 14.429 min), formononetin (SIM2,; 269; 15.624 min), prunetin (SIM2; 285; 18.789 min), biochanin A (SIM2; 285; 19.270 min), glycitin (SIM1; 447; 3.420 min), genistin (SIM1; 433; 4.809 min), tectoridin (SIM1; 463; 5.054 min), sissotrin (SIM1; 447; 11.041 min), and tectorigenin (SIM1; 301; 11.840 min). * see below the table

Plant	Tissue	Isoflavonoid	R_1*	R_2*	R_3*	R_4*	m/z	SIM	Retention time [min]	Content [μg g^{-1}(d.m.)]
Transgenic	shoot	tectorigenin	OH	OH	OCH_3	OH	301	SIM1	11.974	1.28
A. thaliana		biochanin A	OCH_3	OH	H	OH	285	SIM2	19.166	0.15
	seed	genistein	OH	OH	H	OH	271	SIM2	11.991	0.50
Wild-type	shoot	-	-	-	-	-	-	-	-	-
A. thaliana	seed	prunetin	OH	OH	H	OCH_3	285	SIM2	18.749	0.02
P. sativum	seed	glycitin	OH	H	OCH_3	Oglc	447	SIM1	3.433	0.33
		daidzein	OH	H	H	OH	255	SIM2	9.385	0.55
		daidzin	OH	H	H	Oglc	417	SIM2	3.253	6.63
		genistein	OH	OH	H	OH	271	SIM2	11.985	0.37
		genistin	OH	OH	H	Oglc	433	SIM1	4.824	3.22
		prunetin	OH	OH	H	OCH_3	285	SIM2	18.801	0.27

*



also in wild type *Arabidopsis* extracts, but their content was very low (Table 1).

It has been stated in the literature that isoflavone synthase is localized on the membrane of the endoplasmic reticulum as a component of the isoflavonoid metabolone (Liu and Dixon 2001). The correct localization of IFS is crucial for isoflavonoid biosynthesis. Mislocalization of IFS in the cytoplasm

makes it impossible to access IFS by its substrate naringenin which can then be channelled towards the IFS competitor, flavonol-3-hydroxylase (F3H) instead, thereby leading to the production of flavonols (Crozier *et al.* 2006).

Using the *SignalP 3.0* on-line tool (Bendtsen *et al.* 2004), the N-terminal signal peptide in CYP93C18 was predicted with 99.6 % probability (Hidden Markov models). The cleavage site was determined between the 18th and 19th amino acid residues with not a very high probability of 39.2 %. The prediction of ER targeting was verified by the transient expression of the 35S::CYP93C18::GFP fusion protein in *Nicotiana benthamiana* leaves. The fluorescent signal was observed in the endoplasmic reticulum as predicted (Fig. 4*A*,*B*). The CYP93C18::GFP localization was identical to that of the GFP::HDEL (pBIN *m-gfp5*-ER) fusion protein marker for ER localization (Fig. 4*D*). These results were very different from the diffused localization of free GFP in the cytoplasm and nucleus (Fig. 4*C*) and confirmed the previously shown ER localization of IFS (Liu and Dixon 2001).

The question as to how IFS is bound to the ER membrane, however, remains still to be answered. In IFS, there was no signal anchor (*i.e.* uncleaved signal peptide) recognised by the *HMM*. Similarly, no potential modification site was found within the CYP93C18 sequence. However, further investigation using freely available on-line tools *SOSUI* (http://bp.nuap.nagoya-u.ac.jp/sosui/; Mitaku Group, Department of Applied Physics, Nagoya University) and *TMpred* (Hofmann and Stoffel 1993) revealed the presence of the N-terminal hydrophobic transmembrane helix with an inner-outer membrane orientation – a property of type 1 trans-membrane proteins. CYP93C18 3D structure was modelled using the *Modeller9v5* program. As templates, 1tqn, 2hi4, and 3e6i were utilized since they were found to be the most suitable by the *Psipred* threading algorithm. This model fits with previous predictions as it proposes a mechanism of CYP93C18 membrane-binding with the N-terminal helix embedded in a membrane (Fig. 4*E*). As to the manner in which IFS is associated to the ER membrane, support is also lent to this suggestion by the electrostatic charge distribution on the protein computed by the linear Poisson-Boltzmann equation (Fig. 4*F*).

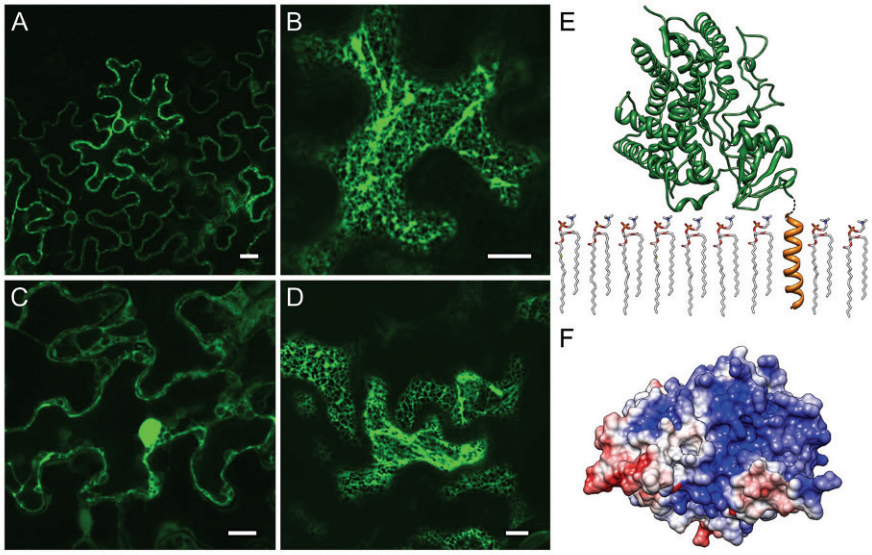


Fig. 4. IFS subcellular localisation and a model of the CYP93C18 3-D structure. Transient expression of 35S::IFS::GFP fusion proteins (*A*; *B*), of the GFP itself (*C*), and of the 35S::GFP::HDEL fusion protein (*D*) in the epidermal cells of *Nicotiana benthamiana* leaves. The GFP signal was observed within the endoplasmic reticulum (confocal microscopy, *bar* =10 μm). Models of CYP93C18 3-D structure. *In silico*-modelled 3D structure of CYP93C18 (*E*) and its association with the endoplasmic reticulum membrane on the cytosolic side. The N-terminal hydrophobic helix (indicated in orange) is anchored in the membrane. Electrostatic charge distribution on the CYP93C18 model (*F*), viewed from the bottom (membrane-associated) side. *Red colour* indicates -5 kT/e, *blue colour* stands for +5 kT/e. The positively-charged and almost flat region represents a putative site of the IFS-membrane association.

Discussion

In a previous study, Cooper *et al.* (2005) demonstrated the obvious connection between the up-regulation of the newly-discovered "putative isoflavone synthase gene", CYP93C18, and an increase in isoflavonoid pisatin content detected in pea pods treated with insect elicitor. However, direct evidence for the functional activity of CYP93C18 was not reported. In this work, we provide direct evidence for the isoflavone synthase catalytic

activity of CYP93C18, as well as demonstrate its subcellular localization. We also demonstrate the applicability of an undemanding PCR method with appropriately-designed primers to identify isoflavone synthase genes in the genomic DNA of leguminous plant species. Equally encouraging results were achieved upon the application of this strategy on other leguminous species, namely *Phaseolus vulgaris* L. and *Pachyrhizus tuberosus* (Lam.) Spreng. whose complete *IFS* sequences were also obtained (data not shown). On the other hand, the same method failed to work for non-leguminous species which we also investigated (data not shown). However, these findings need to be interpreted with caution. In some non-leguminous plants, isoflavone synthesis may be absent and thus the gene itself is likely to be absent. Primers designed on the basis of our knowledge of the *IFS* genes from leguminous plants and from sugar beet might not necessarily be appropriate to phylogenetically distant plant species. It is possible, albeit improbable, that a hitherto unknown protein responsible for isoflavone synthesis through a different enzyme pathway may exist instead of IFS; and if so, this pathway still awaits discovery.

As far as CYP93C18a and CYP93C18b protein sequences are concerned (Cooper *et al.* 2005 and this study), nine different amino acids were discovered – all were upstream of the intron. Judging from the Southern blot analysis, in which the probe hybridized to two IFS sequence fragments most probably corresponding to the two published isoforms of IFS from *P. sativum* (Jung *et al.* 2000, Cooper *et al.* 2005), the observed mismatches can be ascribed to the CYP93C18 allelic variation or the genomic polymorphism of the individual pea cultivars. Possible sequencing errors can be ruled out as in our case the gene was sequenced three times from both 5′and 3′ends, where each sequence run was based on an independent amplification reaction. Moreover, a multiple alignment of *CYP93C18*b with 30 known isoflavone synthase deduced proteins – including that described by Cooper (*P450 Engineering* database, University of Stuttgart, Germany) – revealed that five of the mismatched amino-acids in our sequence corresponded to the consensus; three of the mismatches in Cooper′s sequence were consensual, and the mismatched amino-acid residue at the position 155 was the only one to fail to correspond to the consensual sequence both in Cooper′s and our case. These findings support the polymorphism theory. The ascertained polymorphic nucleotides were not localized in the catalytic centre of the enzyme and did not affect the conserved amino-acid residues Ser 310, Lys 375, and Leu 371 critical for aryl migration (see Fig. 2; Sawada *et al.* 2002, Sawada and Ayabe 2005). Thus, the described polymorphism has no apparent influence over the localization and function of pea IFS as it has been demonstrated in the present study.

The present functional study of *P. sativum* protein CYP93C18 provides an evidence that: *1*) CYP93C18 is a functional isoflavone synthase – isoflavonoids genistein, tectorigenin, and biochanin A were detected in IFS-transformed *Arabidopsis,* and *2*) *Arabidopsis* is an appropriate plant-expression system for IFS functional assays as it contains the IFS substrate naringenin. The elimination of one H_2O molecule to yield isoflavones from 2-hydroxyisoflavanone is likely to be spontaneous or the dehydratation could possibly be catalysed by 2-hydroxyisoflavanone dehydratase (Akashi *et al.* 2005), although no evidence of the presence of this enzyme in non-leguminous species has been reported. Moreover, isoflavone-modifying enzymes, namely hydroxylase and *O*-methyltransferase, are clearly present and are able to recognise and modify the novel substrate genistein as indicated by the presence of high concentrations of tectorigenin (4′,5,7-trihydroxy-6-methoxyisoflavone). The data presented also demonstrate that the introduced IFS successfully competed in transgenic plants with the endogenous enzymes involved in flavonoid biosynthesis which also utilize the flavanone substrate naringenin – for instance, flavanone-3-hydroxylase or flavone synthase (Yu and McGonigle 2005). In addition, the unexpected presence of the methylated isoflavonoid tectorigenin, a well-known anti-cancerogenic isoflavonoid (Thelen *et al.* 2005), in the IFS-transgenic *Arabidopsis* plant is, to the best of our knowledge, reported here for the first time (*cf.* the results reported in Yu *et al.* 2000 and Liu *et al.* 2002). All in all, the simple methodology presented here may be usefully applied to the investigation of the function of newly-identified genes encoding for isoflavone synthases from various species.

The fact that the function of the N-terminal signal peptide in CYP93C18 causes its corresponding ER localization was demonstrated by the transient expression of the CYP93C18::GFP fusion protein in tobacco leaves, where the GFP signal was apparently localized on the endoplasmic reticulum as predicted. Further to the earlier demonstration of the association of *Medicago truncatula* isoflavone synthase over-expressed in *Medicago sativa* seedlings (Liu and Dixon 2001), our results confirmed the correct subcellular localization of CYP93C18.

The observed localization of the CYP93C18::GFP fusion protein was consistent with the published concept of isoflavonoid metabolone (Yu and McGonigle 2005). According to this concept, membrane-bound IFS ensures correct metabolic channelling by interacting with enzymes upstream and downstream in the pathway. However, the manner in which the IFS is attached to the ER membrane is still not entirely clear. It is generally acknowledged that membrane-bound proteins of P_{450}-family have their N-terminal signal sequence anchored in the membrane (type 1 membrane proteins) and it has recently been demonstrated in the case of CYP93C20 (IFS from *Medicago truncatula*) that deletion of the N-terminal anchor leads to protein solubilization (Chang *et al.* 2010). Moreover, membrane-associated P_{450} have further hydrophobic residues in the F-G loop (between F and G helices) associated with the cytosolic side of the ER membrane (Baudry *et al.* 2006). Other possible interaction domains, such as segments before and after the α-helix, and amino acid residues in β strand

2-2, have also been described (Williams *et al.* 2000). *In silico* predictions involving pea isoflavone synthase are consistent with the above-mentioned features of the membrane-bound P_{450}-family proteins. Dai *et al.* (1998) proposed a model of mammalian P_{450} 2B1 and clarified its attachment to the endoplasmic reticulum membrane. The N-terminus is predicted as the sole transmembrane domain of the protein. Nonetheless, its stabilization by other possible hydrophobic regions is highly probable. Such regions would be localized on the protein surface facing the ER membrane as it does the N-terminal helix; they may be found within the F-G loop and the pre-A region as well as in the β sheet 2-2. Although the general localization of IFS with the ER membrane has already been proven (Liu and Dixon 2001), the mode of proposed IFS membrane association remains to be experimentally confirmed.

Collectively, our data provide direct evidence for the proposed isoflavone synthase enzymatic activity of *Pisum sativum* L. CYP93C18 protein *in vivo*. Moreover, we have demonstrated that the pea isoflavone synthase is associated with the endoplasmic reticulum membrane on its cytosolic side and according to the *in silico* model it is anchored there by N-terminal helical hydrophobic transmembrane domain.

References

Adams, N.R.: Detection of the effects of phytoestrogens on sheep and cattle. - J. anim. Sci. **73**: 1509-1515, 1995.

Akashi, T., Aoki, T., Ayabe, S.: Cloning and functional expression of a cytochrome P450 cDNA encoding 2-hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice. - Plant Physiol. **137**: 821-828, 1999.

Akashi, T., Aoki, T., Ayabe, S.: Molecular and biochemical characterization of 2-hydroxyisoflavanone dehydratase. Involvement of carboxylesterase-like proteins in leguminous isoflavone biosynthesis. - Plant Physiol. **137**: 882-891, 2005.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A.: Electrostatics of nanosystems: application to microtubules and the ribosome. - Proc. nat. Acad. Sci. USA **98**: 10037-10041, 2001.

Baudry, J., Rupasinghe, S., Schuler, M.A.: Class-dependent sequence alignment strategy improves the structural and functional modeling of P450s. - Protein Engn. Design Select. **19**: 345-353, 2006.

Bendtsen, J.D., Nielsen, H., Von Heijne, G., Brunak, S.: Improved prediction of signal peptides: SignalP3.0. - J. mol. Biol. **340**: 783-795, 2004.

Bryson, K., McGuffin, L.J., Marsden, R.L., Ward, J.J., Sodhi, J.S., Jones, D.T.: Protein structure prediction servers at University College London. - Nucl. Acids Res. **33** (Suppl.): W36-38, 2005.

Chang, Z., Wang, X., Wei, R., Liu, Z., Shan, H., Fan, G., Hu, H.: Functional expression and purification of CYP93C20, a plant membrane-associated cytochrome P450 from *Medicago truncatula*. - Protein Expres. Purif. http://dx.doi.org/ 10.1016/j.pep.2010.11.012

Clough, S.J., Bent, A.F.: Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. - Plant J. **16**: 735-743, 1998.

Cooper, L.D., Doss, R.P., Price, R., Peterson, K., Olivern, J.E.: Application of Bruchin B to pea pods results in the up-regulation of CYP93C18, a putative isoflavone synthase gene, and an increase in the level of pisatin, an isoflavone phytoalexin. - J. exp. Bot. **56**: 1229-1237, 2005.

Cornwell, T., Cohick, W., Raskin, I.: Dietary phytoestrogens and health. - Phytochemistry **65**: 995-1016, 2004.

Crozier, A., Jaganath, I.B., Clifford, M.N.: Phenols, polyphenols and tannins: an overview. - In: Crozier, A., Clifford, M.N., Ashihara, H. (ed.): Plant Secondary Metabolites: Occurrence, Structure and Role in the Human Diet. Pp. 1-24. Blackwell Publishing, Oxford 2006.

Dai, R., Pincus, M.R., Friedman, F.K.: Molecular modeling of cytochromes P450 2B1: mode of membrane insertion and substrate specifity. - J. Protein Chem. **17**: 121-129, 1998.

Dixon, R.A., Pasinetti, G.M.: Flavonoids and isoflavonoids: from plant biology to agriculture and neuroscience. - Plant Physiol. **154**: 453-457, 2010.

Franzmayr, B.K., Rasmussen, S., Fraser, K.M., Jameson, P.E.: Expression and functional characterization of a white clover isoflavone synthase in tobacco. - Ann. Bot. **110**: 1291-1301, 2012.

Haseloff, J., Siemering, K.R., Prasher, D.C., Hodge, S.: Removal of a cryptic intron and subcellular localisation of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. - Proc. nat. Acad. Sci. USA **94**: 2122-2127, 1997.

He, X., Blount, J.W., Ge, S., Tang, Y., Dixon, R.A.: A genomic approach to isoflavone biosynthesis in kudzu (*Pueraria lobata*). - Planta **233**: 843-855, 2011.

Hofmann, K., Stoffel, W.: TMbase – a database of membrane spanning proteins segments - Biol. Chem. Hoppe-Seyler **374**: 166, 1993.

Jaganath, I.B.: Dietary Flavonoids: Bioavailabilty and Biosynthesis. - PhD Thesis, University of Glasgow, Glasgow 2005.

Jung, W., Yu, O., Lau, S.C., O´Keefe, D., Odell, J., Fader, G., McGonigle, B.: Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes. - Nat. Biotechnol. **18**: 208-212, 2000.

Lapčík, O.: Isoflavonoids in non-leguminous taxa: a rarity or rule? - Phytochemistry **68**: 2909-2916, 2007.

Lapčík, O., Honys, D., Koblovská, R., Macková, Z., Vítková, M., Klejdus, B.: Isoflavonoids are present in *Arabidopsis thaliana* despite the absence of any homologue to known isoflavonoid synthases. - Plant Physiol. Biochem. **44**: 106-114, 2006.

Lapčík, O., Hill, M., Černý, I., Lachman, J., Al-Maharik, N., Adlercreutz, H., Hampl R.: Immunoanalysis of isoflavonoids in *Pisum sativum* and *Vigna radiata*. - Plant Sci. **148**: 111-119, 1999.

Li, J.F., Park, E., Von Arnim, A.G., Nebenführ, A.: The FAST technique: a simplified *Agrobacterium*-based transformation method for transient gene expression analysis in seedlings of *Arabidopsis* and other plant species. - Plant Methods **5**: 6, 2009.

Liu, C.-J., Blount, J.W., Steele, C.L., Dixon, R.A.: Bottlenecks

for metabolic engineering of isoflavone glycoconjugates in Arabidopsis. - Proc. nat. Acad. Sci. USA **99**: 14578-14583, 2002.

Liu, C.-J., Dixon, R.A.: Elicitor-induced association of isoflavone O-methyltransferase with endomembranes prevents the formation and 7-O-methylation of daidzein during isoflavonoid phytoalexin biosynthesis. - Plant Cell **13**: 2643-2658, 2001.

Macková, Z., Koblovská, R., Lapčík, O.: Distribution of isoflavonoids in non-leguminous taxa – an update. - Phytochemistry **67**: 849-855, 2006.

Misra, P., Pandey, A., Tewari, S.K., Nath, P., Trivedi, P.K.: Characterization of isoflavone synthase gene from *Psoralea corylifolia*: a medicinal plant. - Plant Cell Rep. **29**: 747-755, 2010.

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T., Kimura, T.: Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. - J. Biosci. Bioeng. **104**: 34-41, 2007.

Ososki, A.L., Kennelly, E.J.: Phytoestrogens: a rewiev of present state of research. - Phytoterap. Res. **17**: 845-869, 2003.

Oulehlová, D., Hála, M., Potocký, M., Žárský, V., Cvrčková, F.: Plant antigens cross-react with rat polyclonal antibodies against KLH-conjugated peptides. - Cell Biol. Int. **33**: 113-118, 2009.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E.: UCSF chimera – a visualization system for exploratory research and analysis. - Comput. Chem. **25**: 1605-1612, 2004.

Sali, A., Blundell, T.L.: Comparative protein modelling by satisfaction of spatial restraints. - J. mol. Biol. **234**: 779-815, 1993.

Sawada, Y., Ayabe, S.: Multiple mutagenesis of P450 isoflavonoid synthase reveals a key active-site residue. - Biochem. biophys. Res. Commun. **330**: 907-913, 2005.

Sawada, Y., Kinoshita, K., Akashi, T., Aoki, T., Ayabe, S.: Key amino acid residues required for aryl migration catalysed by the cytochrome P450 2-hydroxyisoflavanone synthase. - Plant J. **31**: 555-564, 2002.

Steele, C.L., Gijzen, M., Qutob, D., Dixon, R.A.: Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean. - Arch. Biochem. Biophys. **367**: 146-150, 1999.

Thelen, P., Scharf, J.G., Burfeind, P., Hemmerlein, B., Wuttke, W., Spengler, B., Christoffel, V., Ringert, R.H., Seidlová-Wuttke, D.: Tectorigenin and other phytochemicals extracted from leopard lily *Belamcanda chinensis* affect new and established targets for therapies in prostate cancer. - Carcinogenesis **26**: 1360-1367, 2005.

Veitch, NC: Isoflavonoids of the leguminosae. - Nat. Prod. Rep. **24**: 417-464, 2007.

Wang, W., Vignani, R., Scali, M., Cresti, M.: A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. - Electrophoresis **27**: 2782-2786, 2006.

Weigel, D., Glazebrook, J. (ed.): *Arabidopsis*. A Laboratory Handbook. - Cold Spring Harbor Laboratory Press, Cold Spring Harbor - New York 2002.

Williams, P.A., Cosme, J., Sridhar, V., Johnson, E.F., McRee, D.E.: Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functionaldiversity. - Mol. Cells **5**: 121-131, 2000.

Wiriyaampaiwong, P., Thanonkeo, S., Thanonkeo, P.: Molecular characterization of isoflavone synthase gene from *Pueraria candollei* var. *mirifica*. - Afr. J. agr. Res. **7**: 4489-4498, 2012.

Yu, O., Jung, W., Shi, J., Croes, R.A., Fader, G.M., McGonigle, B., Odell, J.T.: Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues. - Plant Physiol. **124**: 781-794, 2000.

Yu, O., McGonigle, B.: Metabolic engineering of isoflavone biosynthesis. - Adv. Agr. **86**: 147-190, 2005.