

L. Tian · R.A. Dixon

## Engineering isoflavone metabolism with an artificial bifunctional enzyme

Received: 13 November 2005 / Accepted: 21 January 2006 / Published online: 16 February 2006  
© Springer-Verlag 2006

**Abstract** Plant secondary metabolism has been a focus of research in recent years due to its significant roles in plant defense and in human medicine and nutrition. A protein engineering strategy was designed to more effectively manipulate plant secondary metabolite (isoflavonoid) biosynthesis. A bifunctional isoflavone synthase/chalcone isomerase (IFS/CHI) enzyme was constructed by in-frame gene fusion, and expressed in yeast and tobacco. The fusion protein was targeted to the endoplasmic reticulum (ER) membrane and the individual enzymatic functions of its component fragments were retained when assayed in yeast. Petals and young leaves of *IFS/CHI* transgenic tobacco plants produced higher levels of the isoflavone genistein and genistein glycosides as a ratio of total flavonoids produced than did plants transformed with *IFS* alone. Thus, through a combined molecular modeling, in vitro protein engineering and *in planta* metabolic engineering approach, it was possible to increase the potential for accumulation of isoflavonoid compounds in non-legume plants. Construction of bifunctional enzymes will simplify the transformation of plants with multiple pathway genes, and such enzymes may find broad uses for enzyme (e.g., cytochrome P450 family) and biochemical pathway engineering.

**Keywords** Chalcone isomerase, Cytochrome P450, Fusion enzyme, Isoflavone synthase, Legume, Metabolic engineering

**Abbreviations** CHI: Chalcone isomerase · CHS: Chalcone synthase · EGFP: Enhanced green fluorescent protein · IFS: Isoflavone synthase

### Introduction

Plants are capable of synthesizing a large variety of low molecular weight organic compounds, which are collectively called secondary metabolites. In contrast to primary metabolites that are common to all plants, many secondary metabolites are differentially distributed among limited taxonomic groups within the plant kingdom. Plant secondary metabolites play important roles in plant–environment interactions, and in human nutrition and medicine (DellaPenna 1999; Dixon 2001; La Camera et al. 2004). Secondary metabolites usually exist in low abundance in plants. Because of the structural complexity of secondary metabolites, their chemical synthesis is not only difficult and expensive, but also often results in low yields. It is therefore desirable to be able to manipulate biosynthetic pathways for large-scale production of targeted secondary metabolites in plants, such as by enhancing the expression of endogenous genes or by introducing foreign genes (Galili and Hofgen 2002; Thelen and Ohlrogge 2002; Dixon 2005).

Phenylpropanoids are one of the largest groups of plant secondary metabolites, and are synthesized from the aromatic amino acids L-phenylalanine and L-tyrosine. Isoflavonoids are derived from the phenylpropanoid pathway and are distributed predominantly in the Leguminosae. They were initially recognized for their roles in plant disease resistance and induction of nodulation (Dixon 2001). They have also received much attention in recent years due to their estrogenic, antioxidant, and anticancer activities in humans (Dixon and Ferreira 2002; Cornwell et al. 2004; Dixon 2004). It is desirable, in the long term, to be able to produce isoflavonoids in a wide range of plants and crops besides legumes for dietary disease prevention.

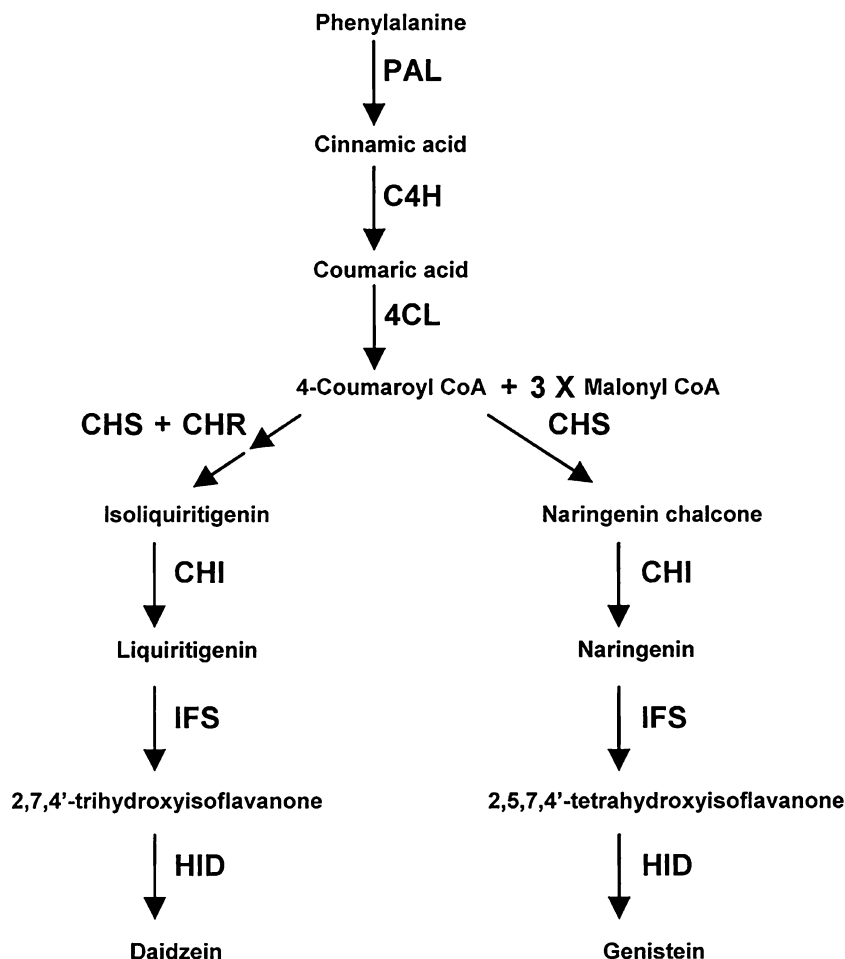
**Electronic Supplementary Material** Supplementary material is available for this article at <http://dx.doi.org/10.1007/s00425-006-0233-0> and is accessible for authorized users.

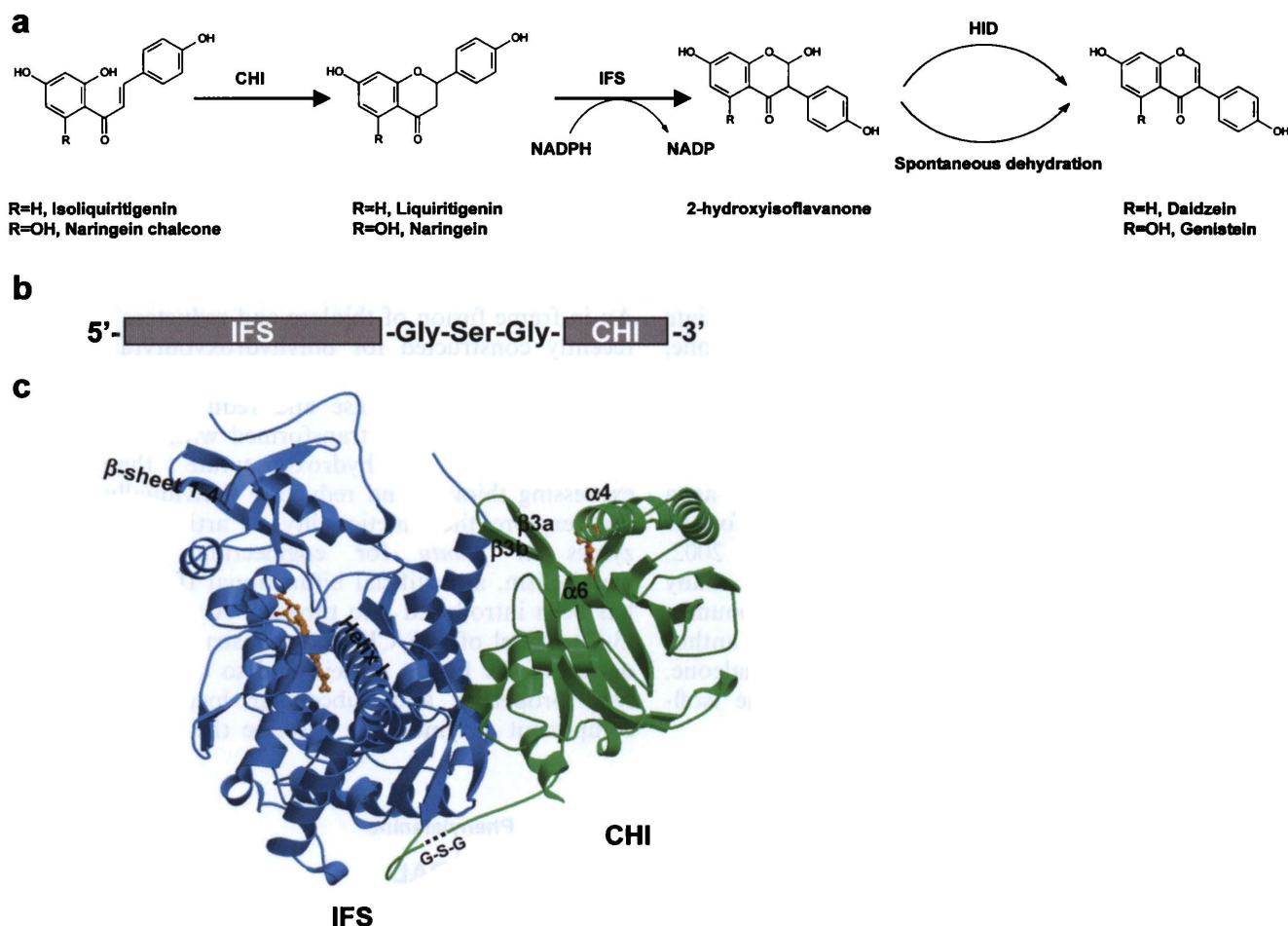
L. Tian · R.A. Dixon (✉)  
Plant Biology Division, Samuel Roberts Noble Foundation, 2510  
Sam Noble Parkway, Ardmore, OK, 73401, USA  
E-mail: [radixon@noble.org](mailto:radixon@noble.org)  
Tel.: +1-580-2246601  
Fax: +1-580-2246692

Isoflavonoids share common biosynthetic precursors, the chalcones, with other phenylpropanoid compounds such as aurones, flavones, and anthocyanins. Chalcones are polyketides that are synthesized via the addition of three malonyl groups to coumarate by chalcone synthase (CHS) (Fig. 1). Chalcone isomerase (CHI) catalyzes the cyclization of chalcone to form flavanone. The first committed step for isoflavonoid biosynthesis is a unique aryl migration reaction catalyzed by a cytochrome P450 enzyme CYP93C, the isoflavone synthase (IFS) (Fig. 2a). The immediate product of the IFS reaction is 2-hydroxyisoflavanone, which can be dehydrated to isoflavones either spontaneously or through catalysis by 2-hydroxyisoflavanone dehydratase (Fig. 2a; Akashi et al. 2005). Genistein is the end product of the IFS and 2-hydroxyisoflavanone dehydratase reactions when naringenin is used as a substrate (Fig. 1). Genistein is well known for its phytoestrogenic activities (Dixon and Ferreira 2002; Barnes 2004), but is also a building block for many structurally more complicated isoflavonoid compounds. Most plant species accumulate flavonoids and anthocyanins; hence, their biosynthetic precursor, chalcone, is available as a substrate for introducing the isoflavonoid pathway into non-legume plants.

In nature, enzymes have often evolved as multi-domain proteins in order to perform tasks that require more than one function. Nature's strategy has been adopted by scientists from different disciplines, in particular *E. coli* and yeast researchers, to develop recombinant multi-functional proteins (Bülow 1990; Nixon et al. 1998; James and Viola 2002). However, this approach has not yet been broadly applied to plants, whose secondary metabolism presents a more complicated biochemical context than yeast or *E. coli*. An in-frame fusion of thiolase and reductase genes was recently constructed for polyhydroxybutyrate biosynthesis in *Arabidopsis* (Kourtz et al. 2005). The fusion protein exhibited thiolase and reductase activities in *E. coli*, though plants transformed with this construct produced less polyhydroxybutyrate than plants expressing thiolase and reductase individually. To further explore the functionality of artificial fusion enzymes *in planta* for engineering of secondary metabolism, an artificial bifunctional IFS/CHI enzyme has been introduced into tobacco. A three-dimensional (3-D) model of IFS/CHI was constructed to guide the design of the fusion protein and to optimize the catalytic properties and subcellular localization of the component enzymes. We describe the *in vitro* activity

**Fig. 1** Isoflavone biosynthesis in legumes. Enzymes are as follows: *PAL*, L-phenylalanine ammonia-lyase; *C4H*, cinnamate 4-hydroxylase; *4CL*, 4-coumarate CoA ligase; *CHS*, chalcone synthase; *CHR*, chalcone reductase; *CHI*, chalcone isomerase; *IFS*, isoflavone synthase; *HID*, 2-hydroxyisoflavanone dehydratase





**Fig. 2** Construction of the IFS/CHI fusion protein. **a** Enzyme catalyzed conversion of chalcone to isoflavone. *CHI*, chalcone isomerase; *IFS*, isoflavone synthase; *HID*, 2-hydroxyisoflavanone dehydratase. **b** Schematic illustration of the IFS/CHI fusion protein. The 5' end of the alfalfa *CHI* coding sequence was fused in-frame to the 3' end of the soybean *IFS* coding sequence via a

Gly-Ser-Gly linker. **c** Graphic simulation of the 3-D structure of the IFS/CHI fusion protein. Strands are shown as *arrows* with *arrow heads* pointing to their C-termini. Helices are represented as *ribbons*. The heme group of IFS and the product of CHI (2*S*-naringenin) are shown as ball-and-stick models. The location of the Gly-Ser-Gly (G-S-G) linker is indicated

of the IFS/CHI fusion enzyme in a yeast assay system, the metabolic consequences of expression of the fusion protein in different tissues of tobacco, and the relative effectiveness of the fusion protein compared to IFS alone for isoflavonoid pathway engineering in a non-legume plant.

## Materials and methods

### Chemicals and plant materials

Isoliquiritigenin, liquiritigenin, naringenin, genistein, and genistin were purchased from Indofine Chemical Co. (Hillsborough, NJ). Naringenin chalcone was prepared from naringenin as previously described (Shimokoriyama 1957). 2-Hydroxyisoflavanone was previously synthesized by Dr. B. Deavours (Deavours and Dixon 2005). Genistein malonyl glucoside was identified by LC-MS as described (Huhman et al. 2005).

*Nicotiana tabacum* cv Xanthi NN plants were grown in soil at  $64\text{--}93\ \mu\text{mol m}^{-2}\text{ s}^{-1}$  under an 18-h light/6-h dark cycle (24°C and 50–75% humidity) in the greenhouse. Young leaves were used for transient expression studies and as starting material for tobacco transformation. Transgenic tobacco plants were grown in the greenhouse under the above-described conditions. Petal tissue and young leaves were collected for total flavonoid analysis.

Plasmid constructs for yeast transformation, plant transformation, and transient expression in tobacco.

Sequences of the primers used for subcloning are listed in Electronic Supplementary Material Table 1. All of the polymerase chain reaction (PCR) reactions were performed using high-fidelity Pfu DNA polymerase (Stratagene, La Jolla, CA). The PCR amplified regions in all constructs were sequenced to ensure that no mutations were introduced during subcloning.

For construction of the IFS/CHI fusion enzyme for expression in yeast, the soybean IFS open reading frame was amplified from the plasmid carrying soybean IFS2

(GenBank accession number AF135484; Steele et al. 1999) with primers containing KpnI (5' end) and BamHI (3' end) sites, and inserted into the KpnI and BamHI sites of pBlueScript KS+ vector. A 650 bp fragment containing the alfalfa CHI open reading frame was amplified from the plasmid carrying alfalfa CHI (GenBank accession number M91079; Liu et al. 2002) with primers containing a BamHI site and a Gly-Ser-Gly linker sequence, GGATCCGGA, at the 5' end, and a SacI site at the 3' end. PCR products were digested with BamHI and SacI and ligated to pBlueScript KS+ -IFS predigested with BamHI and SacI. The resulting *IFS/CHI* fusion fragment was inserted into the KpnI and SacI sites of the yeast pYeDP60 vector. Soybean IFS and alfalfa CHI open reading frames were also subcloned into the pYeDP60 vector singly and used as controls in the yeast expression studies.

For plant transformation, the  $\beta$ -glucuronidase coding region was removed from the binary vector pBI121 (Clontech, Mountain View, CA) and replaced by the coding regions of *CHI*, *IFS*, or *IFS/CHI* fusion, respectively. Primers used for constructing plant transformation plasmids are listed in Electronic Supplementary Material Table 1. Gene expression was under control of the cauliflower mosaic virus 35S promoter and nos terminator.

For construction of a chimeric *IFS/CHI*/enhanced green fluorescent protein (*EGFP*; Clontech, Palo Alto, CA) gene, the following two-step recombinant PCR strategy was applied (Higuchi 1990). In the first step, the *IFS/CHI* fusion open reading frame was amplified using pBluescript KS+ -*IFS/CHI* fusion as a template, with a forward primer to introduce an XhoI restriction site and a reverse primer containing reverse complementary sequences from the end of *CHI* and the start of the *EGFP* open reading frames. Similarly, *EGFP* was amplified using a forward primer with reverse complementarity to *CHI* and a reverse primer with an introduced XbaI restriction site. The resulting *IFS/CHI* and *EGFP* open reading frame fragments were recovered from an agarose gel and served as templates in a second PCR reaction using the *IFS* forward (with an XhoI site) and *EGFP* reverse (with an XbaI site) primers. After digestion with XhoI and XbaI, the resulting chimeric cDNA was inserted into the corresponding sites of the shuttle vector pRTL2 (Restrepo et al. 1990) under the control of a double 35S promoter. The DNA fragment encoding free *EGFP* was amplified with a forward primer containing an XhoI restriction site and the *EGFP* reverse primer (with an XbaI site) and was also inserted in pRTL2. pRTL2-*C4H/MA-EGFP* construct (Achnine et al. 2004) was used as a positive control for endoplasmic reticulum (ER)-membrane localization.

#### Transient expression assay and confocal microscopy

Plasmid DNAs (5  $\mu$ g) harboring *EGFP*, *C4H/MA-EGFP*, or *IFS/CHI-EGFP*, under the control of the

double 35S promoter, were transiently expressed in young tobacco leaf epidermal cells as previously described (Liu and Dixon 2001). Briefly, plasmid DNA coated gold particles were fired at 900 p.s.i., and tobacco leaves were examined 12 h after bombardment using a Bio-Rad 1024ES confocal imaging system. Confocal images were collected and assembled using Adobe Photoshop 5.0 L.E. (Adobe Systems, San Jose, CA).

#### Expression and assay of recombinant proteins in yeast

Transformation of yeast WAT11 cells with the empty pYeDP60 vector and pYeDP60-*CHI*, pYeDP60-*IFS*, or pYeDP60-*IFS/CHI* fusion constructs and preparation of microsomes were conducted as previously described (Liu et al. 2003). Microsomal protein concentrations were determined by the Bradford assay (Bradford 1976).

Isoliquiritigenin, liquiritigenin, naringenin chalcone, or naringenin (8  $\mu$ M) were incubated with 0.8 mg of recombinant yeast microsomes in the presence or absence of 1 mM NADPH. The reaction mixture was incubated at 16°C for 12 h and extracted twice with an equal volume of ethyl acetate. The ethyl acetate fractions were pooled, dried under N<sub>2</sub>, resuspended in methanol and analyzed by reverse phase high performance liquid chromatography (HPLC) using the gradient described in Akashi et al. (1999). The identity of the peaks was confirmed by the retention times and by comparison with authentic standards.

#### Plant transformation and molecular characterization of transgenic tobacco plants

Binary vector constructs were transformed into *Agrobacterium tumefaciens* strain LBA4404 by electroporation. Transformed agrobacteria were confirmed by colony PCR, and the constructs transformed into *N. tabacum* cv. Xanthi NN leaf discs as previously described (Horsch et al. 1988). T1 transformants were selected on MS media containing kanamycin, and the kanamycin resistant plantlets were transferred to soil after 14 days.

Genomic DNA was extracted from transgenic tobacco plants using a DNeasy Plant mini kit (Qiagen, Valencia, CA). The presence of the transgene was confirmed by amplification with *NPTII* primers and the PCR products were subjected to 1% agarose gel electrophoresis.

Transgene expression in leaf tissue was determined by semi-quantitative RT-PCR. Total RNA was extracted from young transgenic tobacco leaf tissue with TRI reagent (Molecular Research Center, Inc., Cincinnati, OH) and treated with RQ1 RNase-free DNase (Promega, Madison, WI) following the manufacturer's instructions. Two micrograms of total RNA was used as template for first-strand cDNA synthesis with an oligo-dT(16) primer using RT-beads (Amersham Pharmacia

Biotech, Piscataway, NJ). Primers used for second strand cDNA synthesis are listed in Electronic Supplementary Material Table 2. Linearity of the PCR reaction was monitored by comparing relative amounts of PCR products after 22, 24, 26, and 30 cycles. The optimized PCR cycling conditions were as follows: 94°C for 3 min, 24 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 2 min, and 1 cycle of 72°C for 10 min.

#### HPLC analysis of total flavonoids in transgenic tobacco plants

Transgenic tobacco petals and young leaves were ground to a fine power in liquid N<sub>2</sub>. Tissue (0.1 g) was extracted with 2 ml 80% methanol overnight at 4°C. The extract was centrifuged to remove tissue debris and the supernatant dried under a stream of N<sub>2</sub>. For  $\beta$ -glucosidase hydrolysis, 3 ml of 5 mg/ml  $\beta$ -glucosidase in citric acid buffer (pH 5.5) was added to the dried samples and the reaction was incubated at 37°C for 12 h and extracted twice with 3 ml ethyl acetate. For acid hydrolysis, the dried samples were incubated with 3 ml of 1 N HCl at 90°C for 2 h and extracted twice with 3 ml of ethyl acetate. For both  $\beta$ -glucosidase and acid hydrolyses, the ethyl acetate extracts were pooled and dried under N<sub>2</sub> and resuspended in 100  $\mu$ l methanol. Ten microliters of the methanolic solution was used for HPLC analysis.

The separation of hydrolyzed and non-hydrolyzed leaf extracts was carried out by reverse-phase HPLC as previously described (Achnine et al. 2005). The separation of petal flavonoids was performed using the following modified gradient: solvent A (1% phosphoric acid) and B (acetonitrile) at 1 ml/min flow rate: 0–5 min, 5% B; 5–10 min, 5–10% B; 10–25 min, 10–17% B; 25–30 min, 17–23% B; 30–39 min, 23–30% B; 39–59 min, 30–35% B; 59–85 min, 35–50% B; 89–99 min, 50–100% B; 99–100 min, 100–5% B. Data were collected at 254, 270, 280, 315, and 524 nm.

## Results

#### Computer modeling of an IFS/CHI fusion enzyme

Soybean IFS is an ER-localized cytochrome P450 enzyme whereas alfalfa CHI is localized in the cytoplasm (Kochs and Grisebach 1988; Bednar and Hadcock 1988). To ensure the ER membrane localization of IFS and its required interaction with the membrane-associated NADPH-cytochrome P450 reductase, IFS was placed as the N-terminus of the fusion enzyme to preserve its ER-targeting membrane anchor (Fig. 2b). As shown from the sequence alignment (Electronic Supplementary Material Fig. 1), IFS is 15 amino acids longer than the mammalian microsomal cytochrome P450 enzyme CYP2C5 at the C-terminus. This gives IFS extra flexibility at the C-terminus. In order to predict

whether the IFS/CHI protein could be folded correctly in vivo, the 3-D structures of IFS (modeled, X. Wang, unpublished data) and CHI (experimentally determined, Jez et al. 2000) were linked *in silico* at the C-terminus of IFS and N-terminus of CHI to simulate the in vivo in-frame fusion of the two enzymes (Fig. 2c). The 3-D structure of IFS was obtained by homology modeling using CYP2C5 as a template (Williams et al. 2000). The tertiary structure of the fusion protein shows a flexible connection between IFS and CHI, because of which the polypeptide chain can fold into distinct compact regions (Fig. 2c). A three amino acid linker peptide, Gly-Ser-Gly, was added between IFS and CHI to reinforce the flexibility of this connection (Fig. 2b), and this further favors correct conformation and independent action of the joined functional domains.

#### ER-localization of the IFS/CHI fusion protein in tobacco epidermal cells

To investigate whether the IFS/CHI fusion protein locates correctly to the ER, EGFP was attached to the C-terminus of the IFS/CHI fusion. Cinnamate 4-hydroxylase (C4H) is a biosynthetic cytochrome P450 enzyme in the early phenylpropanoid pathway and has been shown to be ER-membrane bound (Ro et al. 2001). The C4H membrane anchor region (C4H MA), which directs the enzyme to the cytosolic face of the ER, was fused to the N-terminus of EGFP and used as a positive control for ER-localization. Open reading frames of *IFS/CHI-EGFP*, *C4H MA-EGFP*, and free *EGFP* were subcloned into the pRTL2 vector under the control of a double 35S promoter and independently transfected into young tobacco (*N. tabacum*) leaves through particle bombardment. Green fluorescence was visualized by laser scanning confocal microscopy. Confocal images of transfected tobacco leaf epidermal cells showed that both the C4H MA-EGFP protein and the IFS CHI-EGFP protein were localized to the ER, as indicated by the fine reticulate localization of fluorescence, whereas free EGFP protein was localized to broader cytoplasmic strands (Fig. 3).

#### Functional analysis of recombinant proteins in yeast

To determine whether the IFS/CHI fusion protein is functional in vitro, open reading frames encoding *IFS*, *CHI*, and *IFS/CHI* fusion were each subcloned into the yeast expression vector pYedP60 under the control of a galactose inducible promoter. The yeast WAT11 strain has been engineered to express ATR1, the Arabidopsis NADPH-cytochrome P450 reductase, in place of the native NADPH-cytochrome P450 reductase (Pompon et al. 1996), and was used as a host for expression of the single and fused proteins.

Microsomal fractions from WAT11 transformed with pYedP60 vector control, soybean *IFS*, alfalfa *CHI*, or

soybean *IFS* alfalfa *CHI* fusion were isolated and tested for enzymatic activities using isoliquiritigenin, liquiritigenin, naringenin chalcone, and naringenin as substrates, in the presence (Fig. 4a–d) or absence (Fig. 4e–h) of NADPH. Reaction products were separated by HPLC and their identity confirmed by retention time and UV spectrum compared to those of authentic standards. When incubated with isoliquiritigenin (a substrate for alfalfa *CHI*) without NADPH, liquiritigenin (the product of the *CHI* reaction) was produced in microsomes from cells expressing the *IFS/CHI* fusion protein (Fig. 4d), indicating that the *CHI* enzyme in the *IFS CHI* fusion protein was functional. Microsomes from WAT11 expressing the pYeDP60-*CHI* construct also converted isoliquiritigenin to liquiritigenin (Fig. 4b), whereas microsomes from cells harboring pYeDP60-*IFS* or pYeDP60 did not exhibit any activity towards isoliquiritigenin (Fig. 4a, c).

When microsomes from pYeDP60- *IFS CHI* transformed yeast cells were incubated with isoliquiritigenin and NADPH (a required reductant for the *IFS* reaction), a major additional product, daidzein, was observed, along with a smaller amount of 2,7,4'-trihydroxyisoflavanone (Fig. 4h). This indicated that liquiritigenin produced by the *CHI* reaction was further converted to 2,7,4'-trihydroxyisoflavanone by the functional *IFS* enzyme in the *IFS/CHI* fusion protein construct and that this latter compound was then non-enzymatically dehydrated to daidzein. No daidzein was observed on incubation of isoliquiritigenin and NADPH with microsomes expressing *IFS* alone, *CHI* alone, or empty vector (Fig. 4e–g). Similar levels of daidzein were produced when isoliquiritigenin was incubated with microsomes from yeast cells transformed with pYeDp60-*IFS CHI* fusion or with a mixture of *IFS* and *CHI* independently expressed in yeast (Electronic Supplementary Material Fig. 2).

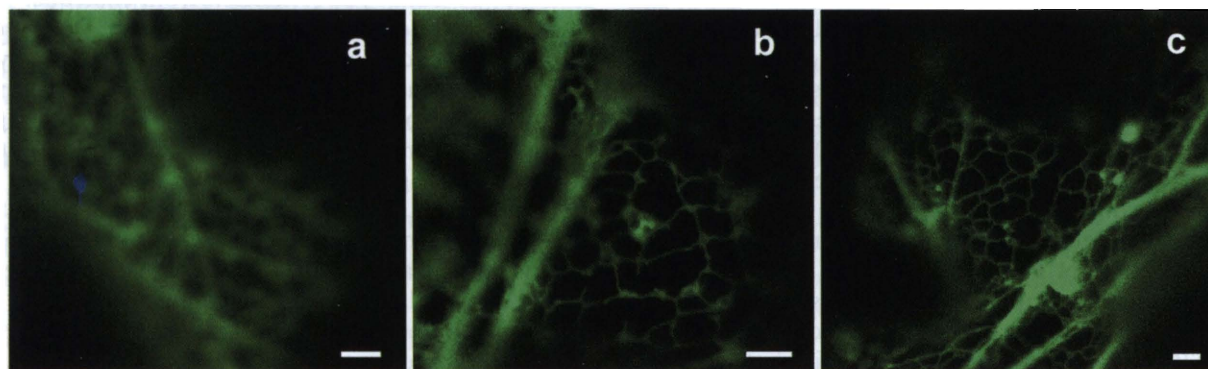
When liquiritigenin was used as substrate in the absence of NADPH, very low levels of daidzein were observed in microsomes from cells expressing pYeDp60-*IFS* and pYeDP60-*IFS/CHI* (Fig. 5c, d), but none was detected in microsomes expressing empty

vector or *CHI* alone (Fig. 5a, b). However, when NADPH was included in the reactions, pYeDp60-*IFS* and pYeDP60-*IFS/CHI* showed similar levels of conversion of liquiritigenin to daidzein, with the intermediate 2,7,4'-trihydroxyisoflavanone also being detected (Fig. 5g, h).

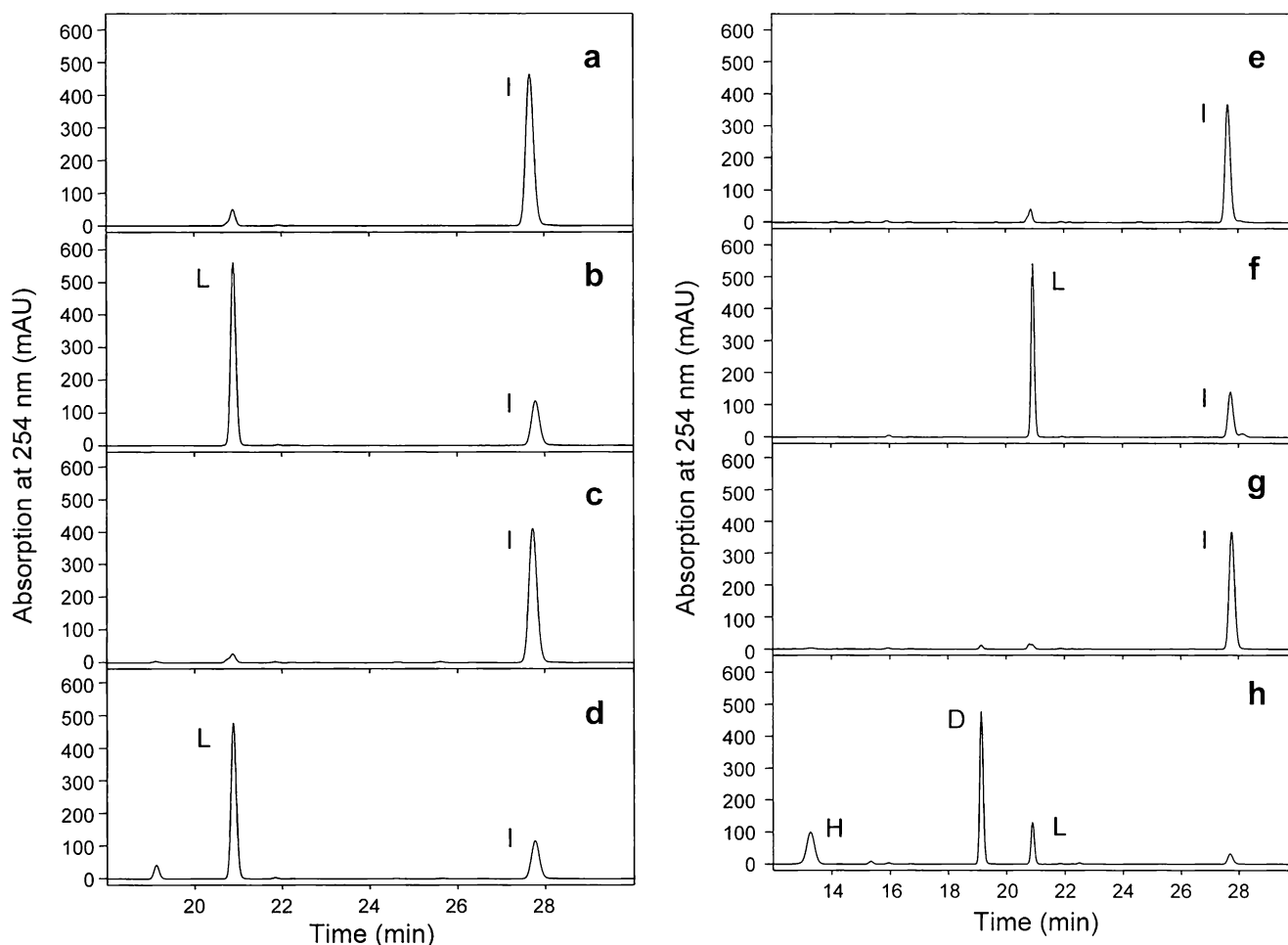
Naringenin chalcone and naringenin were also tested as substrates for the enzymes expressed in recombinant yeast microsomes in the presence or absence of NADPH. Activities of the *IFS/CHI* fusion protein towards these substrates were similar to those observed with isoliquiritigenin and liquiritigenin, respectively (Electronic Supplementary Material Figs. 3 and 4).

#### Production of isoflavones in transgenic tobacco petals and young leaves

*CHI*, *IFS*, and *IFS/CHI* open reading frames were subcloned into the plant transformation vector pBI121. These, and empty vector control construct, were transformed into tobacco (*N. tabacum* cv. Xanthi NN) by *Agrobacterium*-mediated transformation. Transformants were screened for the presence of the transgene by selection for kanamycin resistance and by PCR amplification of the *NPTII* gene on the pBI121 vector (data not shown). The expression level of the transgenes was measured by semi-quantitative RT-PCR (Electronic Supplementary Material Fig. 5), and expression of active protein was confirmed by measuring the *CHI* activity in microsomal extracts from leaves expressing the alfalfa *CHI* or the *IFS/CHI* fusion transgenes using isoliquiritigenin as substrate (Electronic Supplementary Material Fig. 6). This assay specifically measures the activity of the *CHI* encoded by the transgenes, since endogenous tobacco *CHI* cannot use the 6'-deoxy chalcone isoliquiritigenin as substrate. Several independent lines for each construct showed high expression levels of the various transgenes as determined by RT-PCR (Electronic Supplementary Material Fig. 5), and were further characterized for isoflavone production in leaf and petal tissues.



**Fig. 3** Confocal images of tobacco leaf epidermal cells transfected with EGFP fusion proteins by particle bombardment. **a** EGFP. **b** Cinnamate 4-hydroxylase transmembrane domain (C4H MA)-EGFP. **c** *IFS/CHI*-EGFP. Bars = 5  $\mu$ m



**Fig. 4** CHI and coupled CHI/IFS activity of enzymes expressed in yeast. Microsomes were isolated from cells expressing pYeDP60 vector (**a**, **e**), alfalfa CHI (**b**, **f**), soybean IFS (**c**, **g**), and soybean IFS/alfalfa CHI fusion protein (**d**, **h**). Isoliquiritigenin was used as

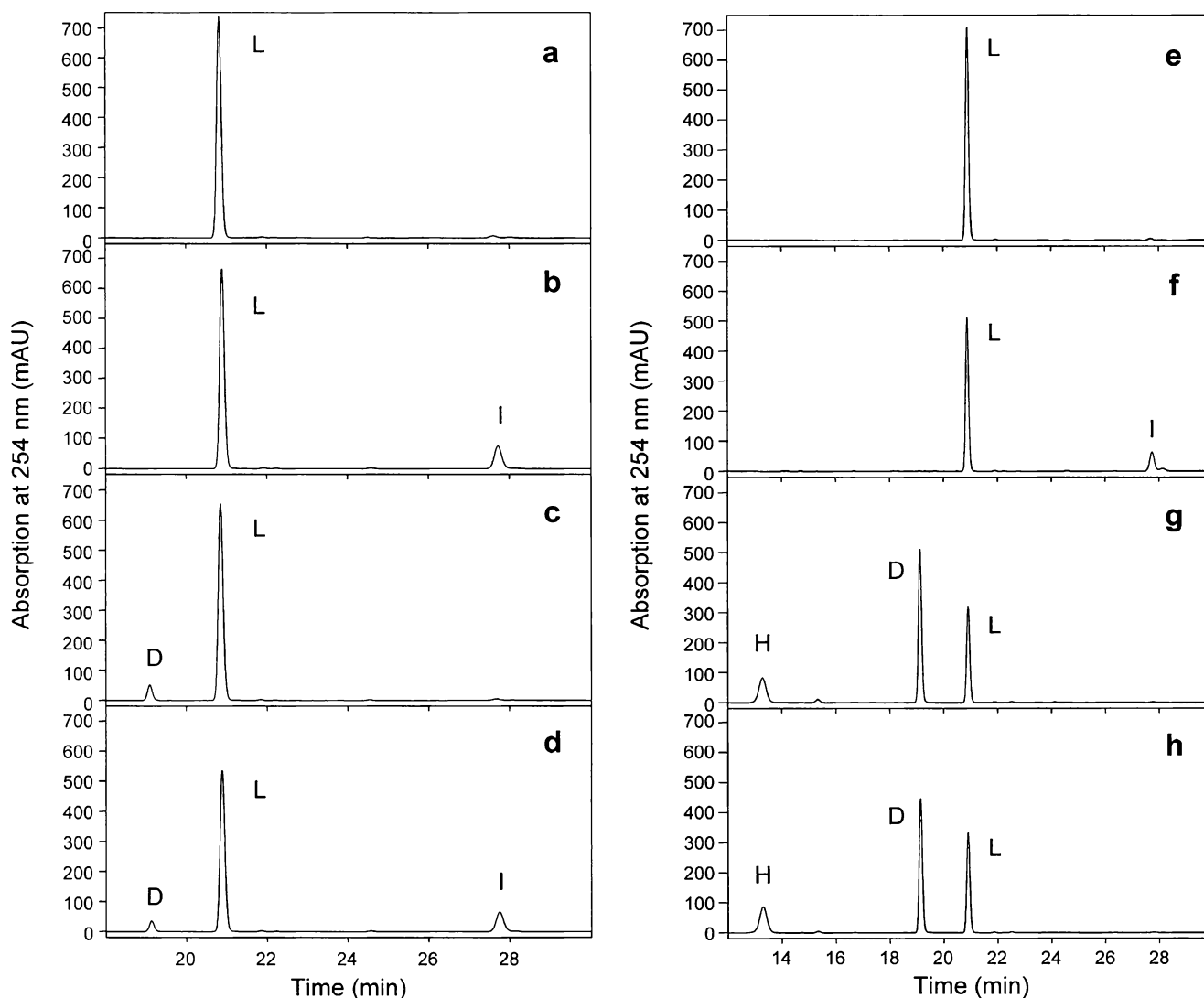
substrate and the reaction mixture was incubated with (**e**–**h**) or without (**a**–**d**) NADPH. Products were analyzed by HPLC. *I*, isoliquiritigenin; *L*, liquiritigenin; *D*, daidzein; *H*, 2,7,4'-trihydroxyisoflavanone

Total flavonoids from transgenic tobacco petals were extracted with 80% methanol, hydrolyzed with 1 N HCl and separated by reverse phase HPLC (Fig. 6a–d). Genistein was detected in extracts from both *IFS* and *IFS/CHI* transgenic plants, but not in extracts from vector control or *CHI* transgenic plants. Non-hydrolyzed extracts from petals were analyzed by HPLC, and showed the presence of genistein (aglycone), genistin (genistein 7-*O*-glucoside, which comigrates with the peak at 34.583 min) and malonyl genistin in *IFS*- and *IFS/CHI*-expressing plants (Electronic Supplementary Material Fig. 7). Hydrolysis of the 80% methanol extracts with almond  $\beta$ -glucosidase released genistein from the glucose conjugates (Fig. 6e–f).

Flavonol levels were quantified from petal extracts hydrolyzed in 1 N HCl, and genistein from  $\beta$ -glucosidase hydrolyzed samples, by comparison to standard curves (Table 1). Petals from transgenic plants expressing *IFS/CHI* fusion produced more genistein than petals from plants expressing the *IFS* transgene alone. However, total flavonol levels were not significantly different

among the transgenic plants, including empty vector controls (Table 1), indicating that expression of alfalfa CHI does not increase flux into flavonoid biosynthesis in tobacco flowers, and suggesting that the increased production of genistein in *IFS/CHI* compared to *IFS* transgenics is most likely the result of improved in vivo efficiency of isoflavone formation from chalcone.

Total flavonoids were also extracted from young transgenic tobacco leaves with 80% methanol, hydrolyzed with  $\beta$ -glucosidase and analyzed by reverse phase HPLC (Electronic Supplementary Material Fig. 8). Genistein accumulated in both *IFS*- and *IFS/CHI*-expressing young leaves, although at a much lower level than in petals (Electronic Supplementary Material Fig. 8; Electronic Supplementary Material Table 3). Surprisingly, leaves of *IFS*- and *IFS/CHI*-expressing plants accumulated more flavonols, and total UV absorbing compounds, than leaves from vector controls or *CHI* transgenic plants (Electronic Supplementary Material Fig. 8; Electronic Supplementary Material Table 3). As with petals, leaves of plants expressing the



**Fig. 5** IFS activity of enzymes expressed in yeast. Microsomes were isolated from cells expressing pYcDP60 vector (**a**, **e**), alfalfa CHI (**b**, **f**), soybean IFS (**c**, **g**), and soybean IFS/alfalfa CHI fusion protein (**d**, **h**). Liquiritigenin was used as substrate and the reaction

mixture was incubated with (**e**–**h**) or without (**a**–**d**) NADPH. Products were analyzed by HPLC. *I*, isoliquiritigenin; *L*, liquiritigenin; *D*, daidzein; *H*, 2,7,4'-trihydroxyisoflavanone

*IFS/CHI* fusion produced more genistein than leaves of plants expressing *IFS* alone. Genistein accumulation could not be detected in mature leaves of any of the transgenic plants (data not shown).

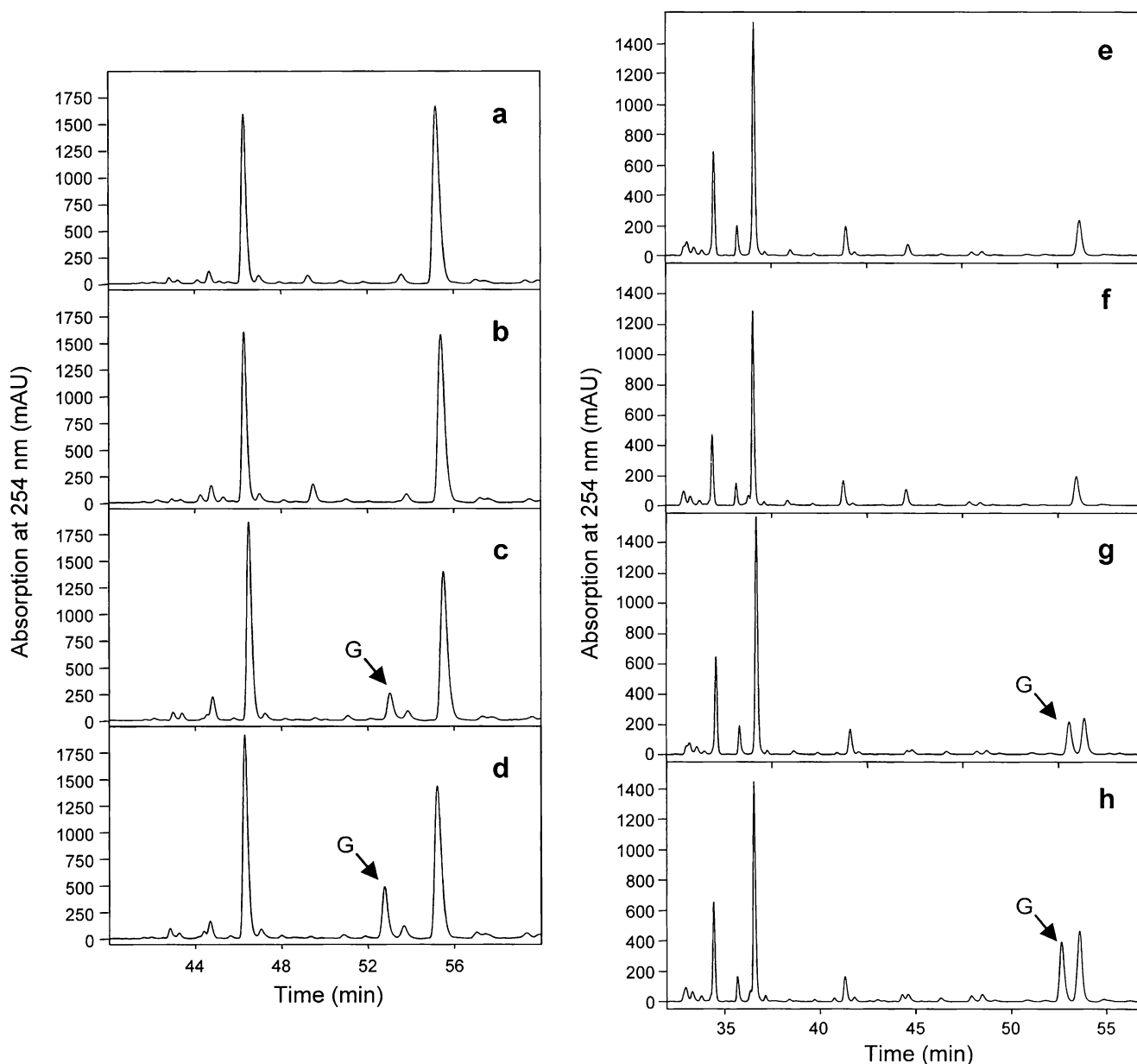
## Discussion

Rational design of a bifunctional fusion protein as directed by the structural/functional properties of the components

A primary consideration for protein engineering is the targeted delivery of the recombinant protein to the functional compartment where its substrates and/or co-enzymes are accessible (Ptashne and Gann 2002).

Although *CHI/IFS* and *IFS/CHI* fusions would be functionally equivalent *in vitro*, *IFS* protein is required to be at the N-terminus (i.e., *IFS/CHI*) in order to guide the fusion protein to the outer (cytosolic) face of the ER membrane, where its co-acting enzyme, NADPH-cytochrome P450 reductase, is located. The transient expression study in tobacco leaf epidermal cells confirmed that the *IFS/CHI* fusion protein localized correctly to the ER.

A key factor in generating a fusion protein is maintaining the functionality of the individual components. A direct head to tail fusion of two proteins together often restrains the correct folding of the individual components, and hence leads to non-functional or mal-functional enzymes (Netzer and Hartl 1997). The availability of 3-D structures of *IFS* (modeled, based on



**Fig. 6** Production of isoflavones in transgenic tobacco. Traces show HPLC profiles of extracts from petals of transgenic tobacco plants after hydrolysis with 1N HCl (**a–d**) or  $\beta$ -glucosidase (**e–h**). Plants were transformed with (**a, e**) pBI121 vector; (**b, f**) pBI121-

CHI (line 2042-5, see Electronic Supplementary Material Supplementary Fig. 4); (**c, g**) pBI121-IFS (line 2043-10); (**d, h**) pBI121-IFS/CHI (line 2044-9). *G*, genistein

Williams et al. 2000) and CHI (experimentally determined) (Jez et al. 2000) facilitated the design of the fusion protein. Molecular modeling of the IFS/CHI fusion indicated a flexible linkage between the two proteins, and a three-amino-acid linker of glycine and serine residues was added to confer additional flexibility to the linkage without interfering with function.

The relative spatial orientations of the active sites were also considered when designing the fusion protein. The substrate-binding site of CHI locates at the bottom of the upside-down bouquet structure and consists of the residues from  $\beta$ -strands  $\beta$ 3a and  $\beta$ 3b and  $\alpha$ -helices  $\alpha$ 4

and  $\alpha$ 6 (Fig. 2c) (Jez et al. 2000). Previous site-directed mutagenesis of the IFS (CYP93C2) protein defined the substrate-binding pocket as involving predicted helix I and  $\beta$ -sheet 1–4 (Sawada et al. 2002). Based on the 3-D model of the IFS/CHI fusion protein, the active sites of the two enzymes are well exposed to catalyze sequential reactions (Fig. 2b, c). When the recombinant IFS/CHI fusion protein was expressed in yeast, the refolded protein was able to catalyze conversion of chalcones to their corresponding isoflavone derivatives, indicating that the fusion protein retained the function of the component modules.

**Table 1** Genistein and flavonol levels in transgenic tobacco petals<sup>a</sup>

Plant line <sup>b</sup>	Quercetin (nmol/g FW)	Kaempferol (nmol/g FW)	Total flavonols (nmol/g FW)	Genistein (nmol/g FW)
Control (pBI121 vector)	2,472.89 ± 80.68	3,549.06 ± 228.09	6,021.95 ± 155.78	n.d.
2042-2 (CHI)	2,774.43 ± 76.22	3,237.16 ± 124.11	6,011.59 ± 191.60	n.d.
2042-3 (CHI)	2,760.48 ± 112.52	3,343.73 ± 92.89	6,104.21 ± 191.92	n.d.
2042-5 (CHI)	2,575.40 ± 135.33	3,626.41 ± 140.38	6,201.81 ± 182.18	n.d.
2042-12 (CHI)	2,536.81 ± 218.30	3,683.25 ± 111.52	6,220.06 ± 194.85	n.d.
2042-14 (CHI)	2,671.69 ± 131.19	3,284.57 ± 108.89	5,956.27 ± 87.11	n.d.
2043-4 (IFS)	3,013.30 ± 119.25	3,081.45 ± 186.81	6,094.75 ± 140.47	225.62 ± 40.92
2043-10 (IFS)	3,099.63 ± 84.25	2,929.42 ± 74.36	6,029.05 ± 148.52	235.59 ± 24.36
2043-12 (IFS)	3,024.22 ± 98.36	3,129.49 ± 143.53	6,153.71 ± 65.10	174.03 ± 6.32
2043-13 (IFS)	2,841.16 ± 102.60	3,127.29 ± 104.85	5,968.45 ± 196.43	190.14 ± 19.78
2043-15 (IFS)	3,090.13 ± 105.44	3,017.18 ± 71.68	6,107.31 ± 112.07	161.49 ± 6.31
2044-5 (IFS CHI fusion)	2,808.65 ± 63.38	3,211.31 ± 215.73	6,019.97 ± 154.45	309.57 ± 35.65
2044-7 (IFS CHI fusion)	2,911.99 ± 190.63	3,118.65 ± 184.21	6,030.63 ± 348.35	284.86 ± 31.11
2044-9 (IFS CHI fusion)	2,947.96 ± 136.24	3,176.63 ± 197.89	6,124.59 ± 168.46	415.65 ± 33.01
2044-13 (IFS CHI fusion)	3,112.16 ± 106.83	3,080.86 ± 238.81	6,193.01 ± 151.25	382.39 ± 16.46
2044-20 (IFS CHI fusion)	2,848.47 ± 98.33	3,144.77 ± 109.26	5,993.24 ± 60.16	278.01 ± 23.93

<sup>a</sup>Flavonol levels were calculated from HPLC traces of acid hydrolyzed samples and genistein levels were calculated from HPLC traces of  $\beta$ -glucosidase hydrolyzed samples. Values are for independent transgenic lines as shown, and represent the average and standard deviation from three replicate determinations

<sup>b</sup>See Electronic Supplementary Material Fig. 5 for transgene expression data

The isoflavonoid biosynthetic pathway is introduced into non-legume plants by expressing the IFS/CHI gene fusion

Expression of the IFS/CHI fusion protein established the early steps of the isoflavonoid biosynthetic pathway in a non-legume plant, tobacco, resulting in the accumulation of genistein and genistein glycosides in petal and young leaf tissues. Tobacco flowers naturally accumulate pink anthocyanin pigments, consisting mainly of conjugates of cyanidin (Xie et al. 2003). Because anthocyanins use the same precursors as isoflavonoid biosynthesis (Winkel-Shirley 2001), tobacco flowers contain the necessary chalcone and flavonoid substrate pools for the engineered isoflavone accumulation.

Tobacco plants expressing the *IFS* transgene alone accumulated genistein and genistein glycosides in both petals and young leaves, although at lower levels than in plants expressing the *IFS/CHI* fusion. In *Arabidopsis*, genistein accumulation in leaf tissue was not further increased by introducing *CHI* into *IFS*-expressing transgenic plants (Liu et al. 2002). It was suggested that competition between the endogenous flavonoid pathway and the introduced isoflavonoid pathway was the limiting step for isoflavonoid production in transgenic *Arabidopsis*. In the present study, expression of the *IFS/CHI* fusion clearly increased the production of genistein glycosides in tobacco compared to expression of the *IFS* transgene alone. One possible reason for this is that the *IFS* enzyme in the *IFS/CHI* fusion is more efficient in vivo because the flavanone substrate for *IFS* (the product of *CHI*) is more readily available in the fusion protein, i.e., the distance that the flavanone has to diffuse to the *IFS* active site is drastically reduced. The observed differences between *Arabidopsis* and tobacco may also be due to the different flavonoid compositions, and therefore potentially different competing pathways,

in the two plants, although both species produce glycosylated flavonols as major components of the leaf phenolic profile.

To more rigorously test whether the apparently improved efficiency of isoflavone production in plants expressing *IFS/CHI* is indeed the result of kinetic factors, it will be necessary to compare metabolite production in plants expressing similar levels of *IFS* and *CHI* as single enzymes or as fusion protein. It will also be important to determine whether the flavonol pathway competes with *IFS* for naringenin substrate in tobacco petals and leaves. To this end, *IFS*- and *CHI*-expressing transgenic plants are being crossed to provide material in which both components of the *IFS/CHI* fusion are expressed but as single enzymes, and the *IFS/CHI* fusion is being introduced into plants expressing an antisense construct targeting flavanone 3-hydroxylase (*F3H*), to block the competing endogenous flavonol pathway.

Free genistein was detected in extracts from young tobacco leaves expressing *IFS* and *IFS/CHI* fusion transgenes, but only after hydrolysis with  $\beta$ -glucosidase. A previous analysis of *IFS* transgenic tobacco plants could not detect free genistein in leaves (Yu et al. 2000). We could not detect genistein in mature transgenic tobacco leaves after hydrolysis (data not shown), suggesting that production and turnover of genistein aglycone and glycosides are developmentally regulated.

Daidzein is, along with genistein, a major soy phytoestrogen (Dixon 2004) and precursor for a range of antimicrobial isoflavonoids in the leguminosae (Dixon 2001), and is used, as its 7-*O*-glucoside, as a treatment for alcoholism (Keung and Vallee 1993). The endogenous *CHI* found in tobacco is not active with isoliquiritigenin, the precursor of 5-deoxyisoflavonoids such as daidzein, which are produced via a distinct form of *CHI* exemplified by the legume *CHI* utilized in the present work (Shimada et al. 2003). The additional

activity of *Medicago* CHI with isoliquiritigenin, yielding liquiritigenin as a substrate for IFS, should allow for daidzein production in the present plants expressing *IFS/CHI*; furthermore, legume CHIs generally show similar kinetics for isoliquiritigenin and naringenin chalcone (Shimada et al. 2003). To achieve this, the plants would have to be further transformed to express chalcone reductase, an enzyme which acts on the near end-product of the CHS reaction to produce the corresponding 6'-deoxychalcone isoliquiritigenin (Fig. 1; Bomati et al. 2005), instead of the product naringenin chalcone formed in the absence of the chalcone reductase.

#### Metabolic engineering of membrane-bound cytochrome P450 proteins

Cytochrome P450s are a large group of enzymes common to animals and plants. Cytochrome P450 enzymes are targeted to various subcellular localizations, such as ER, chloroplast, mitochondria, and plasma membranes (Werck-Reichhart et al. 2002). The present fusion protein approach takes advantage of the subcellular targeting sequence of the cytochrome P450 enzyme, IFS, and adds an additional enzymatic function to it without altering its native activity. In principle, the membrane-targeting sequences of cytochrome P450 enzymes could be used as anchors for hybrid enzymes. Potential examples include fusions between cytochrome P450s and glycosyltransferases, which often glucosylate aromatic hydroxyl groups, introduced by P450 enzymes, to facilitate storage or increase/decrease bioactivity of plant natural products. Plant small molecule glycosyltransferases are operationally soluble enzymes (Vogt and Jones 2000), but may be associated with P450 enzymes, e.g., in the cyanogenic glucoside pathway that utilizes CYP79A1 and CYP71E1 (Jorgensen et al. 2005). Although the components of the cyanogenic glycoside pathway have been engineered into plants as single, independent constructs (Tattersall et al. 2001), it has yet to be determined whether a fusion enzyme approach would facilitate engineering of this or related pathways.

On the other hand, cytochrome P450 enzymes may be engineered to have alternative subcellular localizations, which may lead to the availability of diverse, novel substrates and therefore potential novel functions for these enzymes. The knowledge gained from engineering cytochrome P450 enzymes can be further applied to other membrane-bound proteins, some of which are targets for medicinal drugs and have potential use in biotechnology. One outstanding example is the development of artificial receptor proteins that combine the ligand-binding region of receptors with carrier proteins of favorable scaffolds (Skerra 2003). Such proteins imitate the action of immunoglobulins and have found value in both basic research and pharmaceutical applications (Hey et al. 2005).

Broader application of artificial fusion enzymes in plant metabolic engineering

Engineering of complex metabolic pathways often requires the manipulation of multiple gene products. Current approaches for plant metabolic engineering may require the transformation of several foreign genes simultaneously (Verpoorte and Memelink 2002). Although there have been notable successes in this area (Ye et al. 2000; Li et al. 2003; Wu et al. 2005), there are technical limitations to transforming with multiple genes, such as compatibility of the various vectors for co-transformation, and competition among the different promoters (Lin et al. 2003). An advantage of using fusion proteins in plant metabolic engineering is that the transformation process is simplified, as the two transgenes are fused in frame and therefore share the same promoter for transcription.

In summary, design of a novel fusion protein based on the structure and functionality of its components has allowed us to combine unique properties within a single protein. Such a strategy should find extensive use in biochemical pathway manipulation and plant metabolic engineering.

**Acknowledgements** We thank Dr. Xiaoqiang Wang for molecular modeling of IFS, Dr. Elison Blancaflor for assistance with confocal microscopy, Dr. Lahoucine Achnine for providing pRTL2-C4H MA-EGFP, Dr. Philip Urban for providing the WAT11 strain and pYeDP60 vector, and Drs. Xiaoqiang Wang and Gregory Peel for critical reading of the manuscript. This work was supported by the Samuel Roberts Noble Foundation.

#### References

- Achnine L, Blancaflor EB, Rasmussen S, Dixon RA (2004) Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. *Plant Cell* 16:3098–3109
- Achnine L, Huhman V, Farag MA, Sumner LW, Blount JW, Dixon RA (2005) Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. *Plant J* 41:875–887
- Akashi T, Aoki T, Ayabe S (1999) Cloning and functional expression of a cytochrome P450 cDNA encoding 2-hydroxyisoflavone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice. *Plant Physiol* 121:821–828
- Akashi T, Aoki T, Ayabe S (2005) Molecular and biochemical characterization of 2-hydroxyisoflavone dehydratase. Involvement of carboxylesterase-like proteins in leguminous isoflavone biosynthesis. *Plant Physiol* 137:882–891
- Barnes S (2004) Soy isoflavones-phytoestrogens and what else? *J Nutr* 134:1225S–1228S
- Bednar RA, Hadcock JR (1988) Purification and characterization of chalcone isomerase from soybeans. *J Biol Chem* 263:9582–9588
- Bomati EK, Austin MB, Bowmann ME, Dixon RA, Noel JP (2005) Structural elucidation of chalcone reductase and implications for deoxychalcone biosynthesis. *J Biol Chem* 280:30496–30503
- Bradford MM (1976) A rapid and sensitive for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254

- Bülow L (1990) Preparation of artificial bifunctional enzymes by gene fusion. *Biochem Soc Symp* 57:123–133
- Cornwell T, Cohick W, Raskin I (2004) Dietary phytoestrogens and health. *Phytochemistry* 65:995–1016
- Deavours BE, Dixon RA (2005) Metabolic engineering of isoflavonoid biosynthesis in alfalfa. *Plant Physiol* 138:2245–2259
- DellaPenna D (1999) Nutritional genomics: manipulating plant micronutrients to improve human health. *Science* 285:375–379
- Dixon RA (2001) Natural products and plant disease resistance. *Nature* 411:843–847
- Dixon RA (2004) Phytoestrogens. *Annu Rev Plant Biol* 55:225–261
- Dixon RA (2005) Engineering of plant natural product pathways. *Curr Opin Plant Biol* 8:329–336
- Dixon RA, Ferreira D (2002) Molecules of interest: genistein. *Phytochemistry* 60:205–211
- Galili G, Hofgen R (2002) Metabolic engineering of amino acids and storage proteins in plants. *Metab Eng* 4:3–11
- Hey T, Fiedler E, Rudolph R, Fiedler M (2005) Artificial, non-antibody binding proteins for pharmaceutical and industrial applications. *Trends Biotechnol* 23:514–522
- Higuchi R (1990) Recombinant PCR. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) *PCR protocols*. Academic, New York, pp 177–182
- Horsch RB, Fry J, Hoffmann N, Neidermeyer J, Rogers SG, Fraley RT (1988) Leaf disc transformation. In: Gelvin SB, Schilperoort RA (eds) *Plant molecular biology manual*, vol A5. Kluwer, Dordrecht, pp 1–9
- Huhman DV, Berhow MA, Sumner LW (2005) Quantification of saponins in aerial and subterranean tissues of *Medicago truncatula*. *J Agric Food Chem* 53:1914–1920
- James CL, Viola RE (2002) Production and characterization of bifunctional enzymes. Substrate channeling in the aspartate pathway. *Biochemistry* 41:3726–3731
- Jez JM, Bowman ME, Dixon RA, Noel JP (2000) Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. *Nat Struct Biol* 7:786–791
- Jorgensen K, Bak S, Busk PK, Sorensen C, Olsen CE, Puonti-Kaerlas J, Møller BL (2005) Cassava plants with a depleted cyanogenic glucoside content in leaves and tubers. Distribution of cyanogenic glucosides, their site of synthesis and transport, and blockage of the biosynthesis by RNA interference technology. *Plant Physiol* 139:363–374
- Keung WM, Vallee BL (1993) Daidzin: a potent, selective inhibitor of human mitochondrial aldehyde dehydrogenase. *Proc Natl Acad Sci USA* 90:1247–1251
- Kochs G, Grisebach H (1986) Enzymic synthesis of isoflavones. *Eur J Biochem* 155:311–318
- Kourtz L, Dillon K, Daughtry S, Madison LL, Peoples O, Snell KD (2005) A novel thiolase–reductase gene fusion promotes the production of polyhydroxybutyrate in *Arabidopsis*. *Plant Biotech J* 3:435–447
- La Camera S, Gouzerh G, Dhondt S, Hoffmann L, Fritig B, Legrand M, Heitz T (2004) Metabolic reprogramming in plant innate immunity: the contributions of phenylpropanoid and oxylipin pathways. *Immunol Rev* 198:267–284
- Li L, Zhou Y, Cheng X, Sun J, Marita JM, Ralph J, Chiang VL (2003) Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. *Proc Natl Acad Sci USA* 100:4939–4944
- Lin L, Liu YG, Xu X, Li B (2003) Efficient linking and transfer of multiple genes by a multigene assembly and transformation vector system. *Proc Natl Acad Sci USA* 100:5962–5967
- Liu CJ, Dixon RA (2001) 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
- Liu CJ, Blount JW, Steele CJ, Dixon RA (2002) Bottlenecks for metabolic engineering of isoflavone glycoconjugates in *Arabidopsis*. *Proc Natl Acad Sci USA* 99:14578–14583
- Liu CJ, Huhman D, Sumner LW, Dixon RA (2003) Regiospecific hydroxylation of isoflavones by cytochrome p450 81E enzymes from *Medicago truncatula*. *Plant J* 36:471–484
- Netzer WJ, Hartl FU (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388:343–349
- Nixon AE, Ostermeier M, Benkovic SJ (1998) Hybrid enzymes: manipulating enzyme design. *Trends Biotechnol* 16:258–264
- Ptashne M, Gann A (2002) *Genes and signals*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Pompon D, Louerat B, Bronine A, Urban P (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol* 272:51–64
- Restrepo MA, Freed DD, Carrington JC (1990) Nuclear transport of plant potyviral proteins. *Plant Cell* 2:987–998
- Ro DK, Mah N, Ellis BE, Douglas CJ (2001) Functional characterization and subcellular localization of poplar (*Populus trichocarpa* × *Populus deltoides*) cinnamate 4-hydroxylase. *Plant Physiol* 126:317–329
- Sawada Y, Kinoshita K, Akashi T, Aoki T, Ayabe S (2002) Key amino acid residues required for aryl migration catalysed by the cytochrome P450 2-hydroxyisoflavanone synthase. *Plant J* 31:555–564
- Shimada N, Aoki T, Sato S, Nakamura Y, Tabata S, Ayabe S (2003) A cluster of genes encodes the two types of chalcone isomerase involved in the biosynthesis of general flavonoids and legume-specific 5-deoxy(iso)flavonoids in *Lotus japonicus*. *Plant Physiol* 131:941–951
- Shimokoriyama M (1957) Interconversion of chalcones and flavanones of a phloroglucinol-type structure. *J Am Chem Soc* 79:4199–4202
- Skerra A (2003) Imitating the humoral immune response. *Curr Opin Chem Biol* 7:683–693
- Steele CL, Gijzen M, Qutob D, Dixon RA (1999) Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean. *Arch Biochem Biophys* 367:146–150
- Tattersall DB, Bak S, Jones PR, Olsen CE, Nielsen JK, Hansen ML, Høj PB, Møller BL (2001) Resistance to an herbivore through engineered cyanogenic glucoside synthesis. *Science* 293:1826–1828
- Thelen JJ, Ohlrogge JB (2002) Metabolic engineering of fatty acid biosynthesis in plants. *Metab Eng* 4:12–21
- Verpoorte R, Memelink J (2002) Engineering secondary metabolite production in plants. *Curr Opin Biotechnol* 13:181–187
- Vogt T, Jones P (2000) Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends Plant Sci* 5:380–386
- Werck-Reichhart D, Bak S, Paquette S (2002) Cytochrome P450. In: Somerville CR, Meyerowitz EM (eds) *The arabidopsis book*, Rockville American Society of Plant Biologists, doi/10.1199/tab.0028, <http://www.aspb.org/publications/arabidopsis>
- Williams PA, Cosme J, Sridhar V, Johnson EF, McRee DE (2000) Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. *Mol Cell* 5:121–131
- Winkel-Shirley B (2001) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol* 126:485–493
- Wu G, Truksa M, Datla N, Vrinten P, Bauer J, Zank T, Cirpus P, Heinz E, Qiu X (2005) Stepwise engineering to produce high yields of very long-chain polyunsaturated fatty acids in plants. *Nat Biotechnol* 23:1013–1017
- Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA (2003) Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. *Science* 299:396–399
- Ye X, Al-Babili S, Kloti A, Zhang J, Lucca P, Beyer P, Potrykus I (2000) Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* 287:303–305
- Yu O, Jung W, Shi J, Croes RA, Fader GM, McGonigle B, Odell JT (2000) Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues. *Plant Physiol* 124:781–793