Plant and Cell Physiology Advance Access published June 12, 2013

PCP
PLANT & CELL PHYSIOLOGY

4-Coumarate: CoA Ligase Partitions Metabolites for Eugenol Biosynthesis

Shubhra Rastogi[1], Ritesh Kumar[1], Chandan S. Chanotiya[2], Karuna Shanker[2], Madan M. Gupta[2], Dinesh A. Nagegowda[1] and Ajit K. Shasany[1,*]

[1]Biotechnology Division, Central Institute of Medicinal and Aromatic Plants (CSIR), PO CIMAP, Lucknow-226015, UP, India
[2]Analytical Chemistry Division, Central Institute of Medicinal and Aromatic Plants (CSIR), PO CIMAP, Lucknow-226015, UP, India
*Corresponding author: E-mail, ak.shasany@cimap.res.in; akshasany@yahoo.com; Fax: +91-522-2342666.
(Received July 5, 2012; Accepted May 9, 2013)

Biosynthesis of eugenol shares its initial steps with that of lignin, involving conversion of hydroxycinnamic acids to their corresponding coenzyme A (CoA) esters by 4-coumarate:CoA ligases (4CLs). In this investigation, a *4CL* (*OS4CL*) was identified from glandular trichome-rich tissue of *Ocimum sanctum* with high sequence similarity to an isoform (*OB4CL_ctg4*) from *Ocimum basilicum*. The levels of *OS4CL* and *OB4CL_ctg4*-like transcripts were highest in *O. sanctum* trichome, followed by leaf, stem and root. The eugenol content in leaf essential oil was positively correlated with the expression of *OS4CL* in the leaf at different developmental stages. Recombinant OS4CL showed the highest activity with *p*-coumaric acid, followed by ferulic, caffeic and *trans*-cinnamic acids. Transient RNA interference (RNAi) suppression of *OS4CL* in *O. sanctum* leaves caused a reduction in leaf eugenol content and trichome transcript level, with a considerable increase in endogenous *p*-coumaric, ferulic, *trans*-cinnamic and caffeic acids. A significant reduction in the expression levels was observed for *OB4CL_ctg4*-related transcripts in suppressed trichome compared with transcripts similar to the other four isoforms (*OB4CL_ctg1*, *2*, *3* and *5*). Sinapic acid and lignin content were also unaffected in RNAi suppressed leaf samples. Transient expression of OS4CL–green fluorescent protein fusion protein in *Arabidopsis* protoplasts was associated with the cytosol. These results indicate metabolite channeling of intermediates towards eugenol by a specific *4CL* and is the first report demonstrating the involvement of *4CL* in creation of virtual compartments through substrate utilization and committing metabolites for eugenol biosynthesis at an early stage of the pathway.

Keywords: 4 • Coumarate:CoA ligase • Eugenol • Lignin • *Ocimum sanctum* • Phenylpropene • RNAi.

Abbreviations: CAAT, coniferyl alcohol acetyl transferase; CAD, cinnamyl alcohol dehydrogenase; CAF, caffeic acid; CaMV, *Cauliflower mosaic virus*; CCOMT, caffeoyl-CoA *O*-methyl transferase; CCR, cinnamoyl-CoA reductase; C3′H, *p*-coumaroyl shikimate 3′-hydroxylase; CIN, *trans*-cinnamic acid; COU, *p*-coumaric acid; EGS, eugenol synthase; EST, expressed sequence tag; FER, ferulic acid; GFP, green fluorescent protein; LC-MS, liquid chromatograpy–mass spectrometry; OB4CL, *Ocimum basilicum* 4-coumarate:CoA ligase; ORF, open reading frame; OS4CL, *Ocimum sanctum* 4-coumarate:CoA ligase; OB4CL_ctg1–5, *Ocimum basilicum* 4-coumarate:CoA ligase contigs 1–5; RNAi, RNA interference; RQ, relative quantity; RT–PCR, reverse transcription–PCR; SBP, substrate-binding pocket; SIN, sinapic acid.

The nucleotide sequences reported in this paper have been submitted to NCBI under accession numbers *OS4CL*, HM990148; and *OB4CL*, KC576841. Other gene accession numbers used in this investigation are provided in **Supplementary Tables S3** and **S7**.

Introduction

The phenylpropenes synthesized in plant vegetative parts as defense compounds against herbivores and pathogens (Obeng-Ofori and Reichmuth 1997, Koeduka et al. 2008) are also important for human nutrition (Nahrstedt 1990, Prasad et al. 2004). In angiosperms, the biosynthetic pathway of phenylpropenes (eugenol, isoeugenol and related compounds) is shared with lignin/lignan biosynthesis during the initial steps, up to monolignol formation (Koeduka et al. 2008). Although, the biosynthesis of eugenol is localized in the glandular trichomes of basil (Gang et al. 2001, Gang et al. 2002), and the enzymes transforming coniferyl acetate to eugenol have already being described in *Ocimum basilicum* trichome (Koeduka et al. 2008), still much remain to be understood about the initial steps of this complex pathway. The knowledge of genes involved in lignin biosynthesis is available as a combination of reports from diverse plant species and this knowledge is always compared with phenylpropene biosynthesis to derive conclusions on the basis of assumptions. However, the preferred mechanisms of regulation and diversion of the carbon skeleton towards phenylpropenes in different tissues are less understood and are not described in the literature.

Plant Cell Physiol. 0(0): 1–15 doi:10.1093/pcp/pct073, available online at www.pcp.oxfordjournals.org
© The Author 2013. Published by Oxford University Press on behalf of Japanese Society of Plant Physiologists.
All rights reserved. For permissions, please email: journals.permissions@oup.com

Regular Paper

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

The enzyme 4-coumarate:CoA ligase (4CL; EC 6.2.1.12) is required for the biosynthesis of a diverse array of plant natural phenylpropanoid products (Hahlbrock and Scheel 1989, Whetten and Sederoff 1995, Douglas 1996). Being the third enzyme of the general phenylpropanoid pathway, it may also play the central role in regulating overall flux of the hydroxycinnamic acids into subsequent biosynthetic pathways (Dixon and Paiva 1995). This is emphasized in the literature based on expression patterns and phylogenetic analysis, but has never been demonstrated conclusively except in the case of lignin. Comparison of the predicted 4CL protein sequences of cloned genes has revealed the presence of conserved peptide motifs of probable functional significance (Stuible et al. 2000). In many plants, several isoenzymes of 4CL, preferentially accepting cinnamic acid and hydroxylated cinnamic acid derivatives, e.g. 4-coumaric acid, caffeic acid, ferulic acid or sinapic acid, have been described (Lindermayr et al. 2002). The position of *4CL* at the metabolic branch point connecting general phenylpropanoid metabolism with different end-product-specific pathways also makes this enzyme a promising target for intervention (**Fig. 1**). The suitability of *4CL* for pathway engineering is also suggested due to different isoforms, substrate specificities, expression patterns and speculated channeling of phenolic precursors to either lignin or flavonoid biosynthesis (Stuible and Kombrink 2001). However, to date, the role of this enzyme in channeling intermediates towards phenylpropenes has not been demonstrated . Since *4CL* genes are yet to be described in *Ocimum* species, this study reports a *4CL* isoform from enriched peltate glandular trichomes of holy basil (*O. sanctum* L.) with differential substrate specificity, enabling it to divert carbon towards phenylpropene biosynthesis. Though a number of 4CLs have been described in the literature, this is the first report of a functional demonstration of a *4CL* isoform, predominantly involved in the phenylpropene biosynthesis pathway leading to eugenol in *Ocimum* by creating virtual compartments through substrate utilization and committing metabolites to the pathway.

Results

Characterization of the *OS4CL* cDNA and its isoforms

The *OS4CL* cDNA (*O. sanctum 4CL*; GenBank accession No. HM990148, 1,704 bp) clustered with *4CL* genes from *O. basilicum* (*OB4CL*; GenBank accession No. KC576841, 1,704 bp) and *Agastache rugosa* (AY587891, Ar4CL). *OS4CL* and *OB4CL* showed 81% amino acid similarity and 86% and 84% nucleotide similarity, respectively, to *A. rugosa Ar4CL* (**Fig. 2**). The comparison of the full-length nucleotide and amino acid sequences of *OS4CL* (HM990148) and *OB4CL* (KC576841) has been provided in **Supplementary Fig. S1**, showing 92% similarity at the nucleotide level and 96% similarity at the amino acid level. This sequence has two main

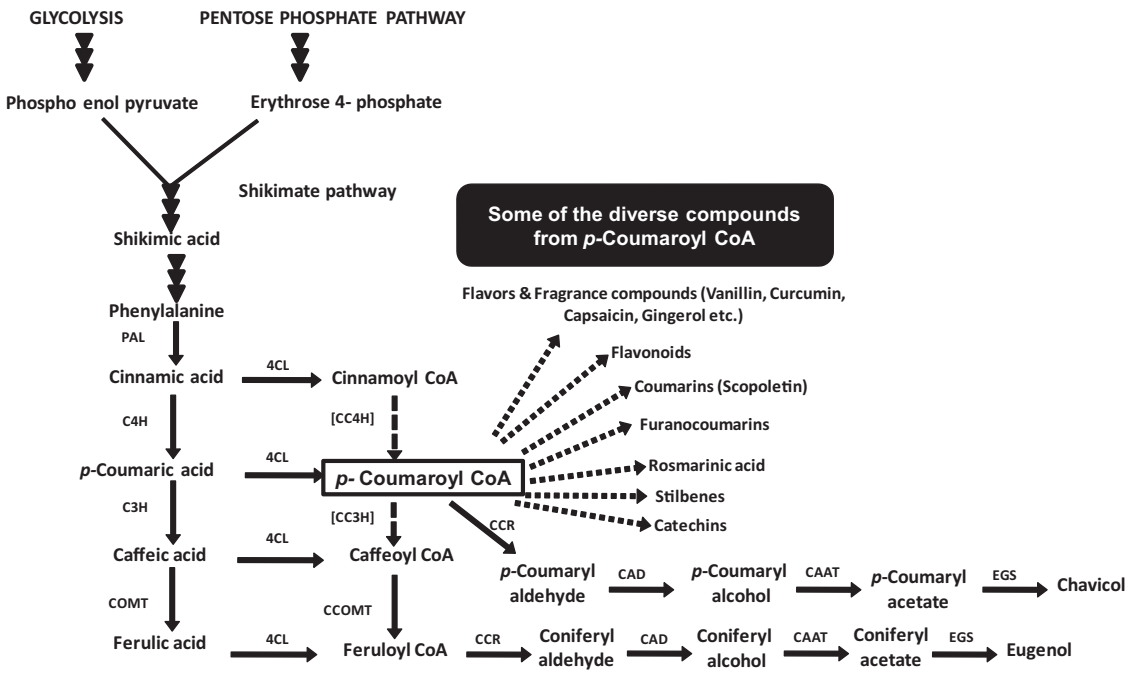

Fig. 1 Proposed biosynthetic pathway of eugenol/chavicol synthesis in *Ocimum sanctum*. [C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate: CoA ligase; C3H, *p*-coumarate 3-hydroxylase; COMT, caffeoyl *O*-methyl transferase; CC4H, cinnamoyl-CoA 4-hydroxylase; CC3H, *p*-coumaroyl-CoA 3-hydroxylase; CCOMT, caffeoyl-CoA *O*-methyl transferase; CCR, cinnamoyl-CoA reductase; CAD, cinnamyl alcohol dehydrogenase; CAAT, coniferyl alcohol acetyl transferase; EGS, eugenol (and chavicol) synthase. The enzymes for 4- and 3-hydroxylation of CoA esters have not been demonstrated yet and hence [CC4H] and [CC3H] are indicated in brackets to indicate their probable role (Gang et al 2001).

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

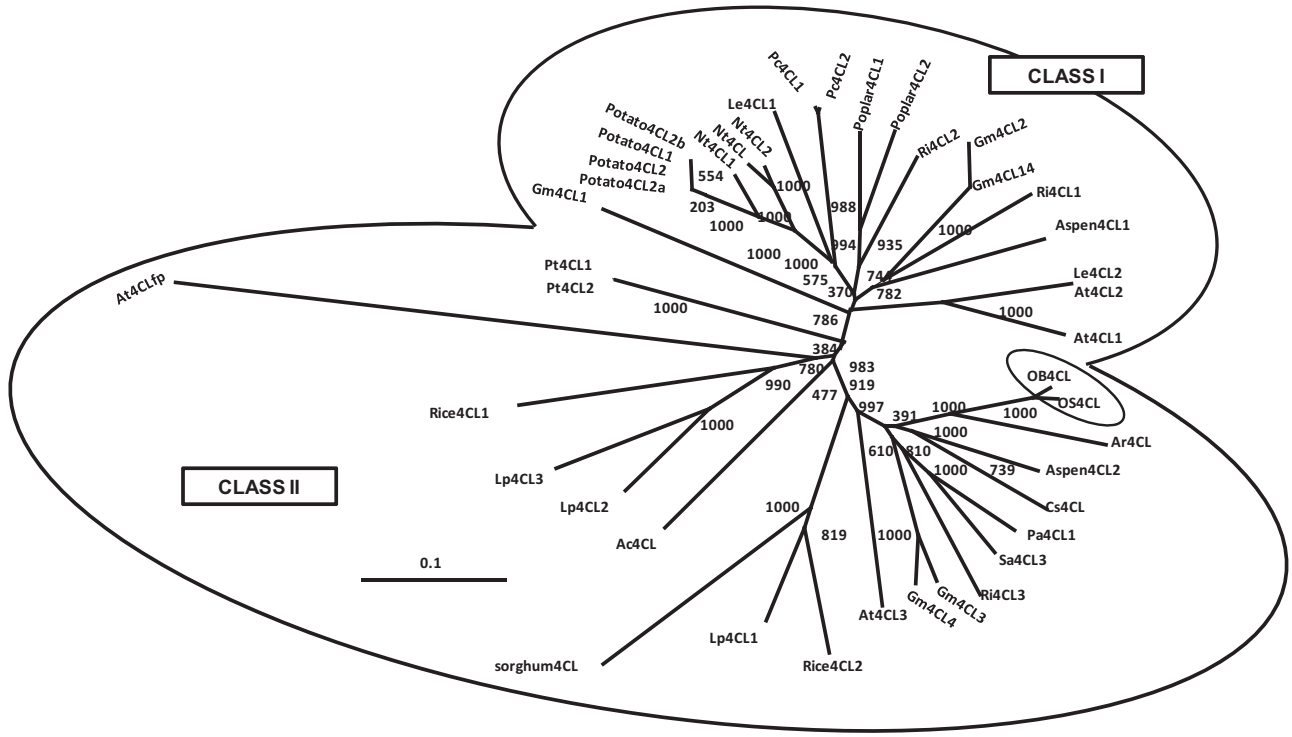

Fig. 2 Unrooted Neighbor–Joining tree comparing the amino acid sequences of *O. sanctum* 4CL (OS4CL) with those of 4CLs reported from other species. GenBank accession numbers for the corresponding amino acid sequences are: *Ocimum sanctum* OS4CL (HM990148); *Ocimum basilicum* OB4CL (KC576841); *Nicotiana tabacum* Nt4CL (D43773); *N. tabacum* Nt4CL1 (U50845); *N. tabacum* Nt4CL2 (U50846); *Oryza sativa* Rice4CL1 (X52623); *O. sativa* Rice4CL2 (L43362); *Petroselinum crispum* Pc4CL1 (X13324); *P. crispum* Pc4CL2 (X13325); *Pinus taeda* Pt4CL1 (U12012); *P. taeda* Pt4CL2 (U12013); *Populus hybrida* Poplar4CL1 (AF008184); *P. hybrida* Poplar4CL2 (AF008183); *Rubus idaeus* Ri4CL1 (AF239687); *R. idaeus* Ri4CL2 (AF239686); *R. idaeus* Ri4CL3 (AF239685); *Solanum tuberosum* Potato4CL1 (M62755); *S. tuberosum* Potato4CL2 (AF150686); *S. tuberosum* Potato4CL2a (AAD40664); Potato4CL2b (AF150687); *Sorghum bicolor* sorghum4CL (U23787); *Populus tremuloides* Aspen4CL1 (AF041049); *P. tremuloides* Aspen4CL2 (AF041050); *Prunus avium* Pa4CL1(GU990523); *Sorbus aucuparia* Sa4CL3 (GU949553); *Allium cepa* Ac4CL (AY541033); *Agastache rugosa* Ar4CL (AY587891); *Camellia sinensis* Cs4CL (DQ194356); *Arabidopsis thaliana* At4CLfp (NM_123172); *A. thaliana* At4CL1 (U18675); *A. thaliana* At4CL2 (AF106086); *A. thaliana* At4CL3 (AF106088); *Glycine max* Gm4CL1 (AF279267); *G. max* Gm4CL2 (AF002259); *G. max* Gm4CL3 (AF002258); *G. max* Gm4CL4 (X69955); *G. max* Gm4CL14 (X69954); *Lithospermum erythrorhizon* Le4CL1 (D49366); *L. erythrorhizon* Le4CL2 (D49367); *Lolium perenne* Lp4CL1 (AF052221); *L. perenne* Lp4CL2 (AF052222); *L. perenne* Lp4CL3 (AF052223).

conserved motifs: Box I and Box II. Box I (SSGTTGLPKGV) is highly conserved across 4CLs, luciferases, acetyl-coenzyme A (CoA) synthetases, long-chain fatty acyl-CoA synthetases and peptide synthetases. The occurrence of this motif has even been used as an important criterion for establishing the superfamily of adenylate-forming enzymes (Ehlting et al. 1999). Box II (GEICIRG) is also conserved in all 4CLs, and its central cysteine residue is suggested to be involved in catalysis (Stuible et al. 2000). Based on structural data obtained by homology modeling, Schneider et al. (2003) postulated the existence of a 4CL signature motif consisting of 12 amino acid residues lining the substrate-binding pocket (SBP), determining 4CL substrate specificity. The amino acid sequence of OS4CL also shows the presence of a similar signature motif (**Supplementary Fig. S2**). The *O. basilicum*, expressed sequence tag (EST) database (http://www.planttrichome.org/trichomedb/estbyspecies_detail.jsp?species=Ocimum%20basilicum) was searched for homologous sequences, and 29 hits falling into five contigs were obtained. Downloaded ESTs were clustered into contigs using the CAP3 Sequence Assembly Program (Huang and Madan 1999). These contigs were assumed to be isoforms of *4CL* (*OB4CL_ctg1–5*) (**Supplementary Table S3a**). The nucleotide similarity matrix between the full-length *4CL* cDNAs and the contigs is presented in **Supplementary Table S3b**, and the primer positions for contigs are indicated in **Supplementary Table S3c**.

Transcript abundance of *OS4CL* and eugenol content at different developmental stages

Quantitative reverse transcription–PCR (RT–PCR) analysis was carried out to measure the *OS4CL* transcript levels at different developmental stages (2, 4 and 6 weeks after seed germination) (**Supplementary Fig. S4a**). The transcript level of 2-week-old plants was taken as the calibrator with a relative quantity (RQ) of '1' and compared with the expression in 4- and 6-week-old plants with RQs of 3.55 and 5.13, respectively. The average

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

eugenol percentage in the essential oil of 2-week-old plants was found to be 25.41% (v/v), which increased to 69.64% and 77.29% in the leaves of 4- and 6-week-old plants, respectively (**Supplementary Fig. S4b**). The increase in eugenol content indicated a positive correlation with the level of *OS4CL* expression.

Tissue-specific expression

The abundance of transcripts similar to all the five isoforms of *4CL* genes reported in *O. basilicum* was analyzed in different tissues of *O. sanctum* by designing the primers from the contigs (*OB4CL_ctg1–5*) generated from the assembled ESTs of *O. basilicum* 'TrichOME Database'. The highest level of amplification was observed with *OB4CL4* contig primers in the trichome of *O. sanctum* followed by leaf, stem and root (**Fig. 3**). A similar trend was also observed for the *OS4CL* transcript. The abundance of transcripts related to all other isoforms (*OB4CL_ctg1*, 2, 3 and 5) was maximal in the root, with less expression in trichomes of *O. sanctum*. *OS4CL* was assumed to be the homologous counterpart of *OB4CL4* based on sequence homology. Hence the full-length cDNA for *OB4CL_ctg 4* was also isolated and characterized from *O. basilicum* (*OB4CL*) for comparison.

Characterization of recombinant 4CL enzymes

The *OS4CL* cDNA was expressed in *Escherichia coli* as a fusion protein containing an N-terminal His_6-tag with a molecular mass of 62.48 kDa. The purified protein was tested for its ability to utilize *trans*-cinnamic acid (CIN), *p*-coumaric acid (COU), caffeic acid (CAF), ferulic acid (FER) and sinapic acid (SIN). Analysis of enzyme activity indicated the utilization of COU, FER, CAF and CIN, but not SIN. Maximal activity was observed with COU as substrate, with an apparent K_m of 1.66 μM (**Table 1**). FER was the next most preferred substrate (K_m = 3.02 μM), followed by CAF (K_m = 4.94 μM) and CIN (K_m = 25.76 μM). Liquid chromatoography–mass spectrometry (LC-MS) analysis confirmed the formation of CoA esters for the substrates COU, FER, CAF and CIN acids (**Supplementary Fig. S5**). New peaks were identified to be the products of each

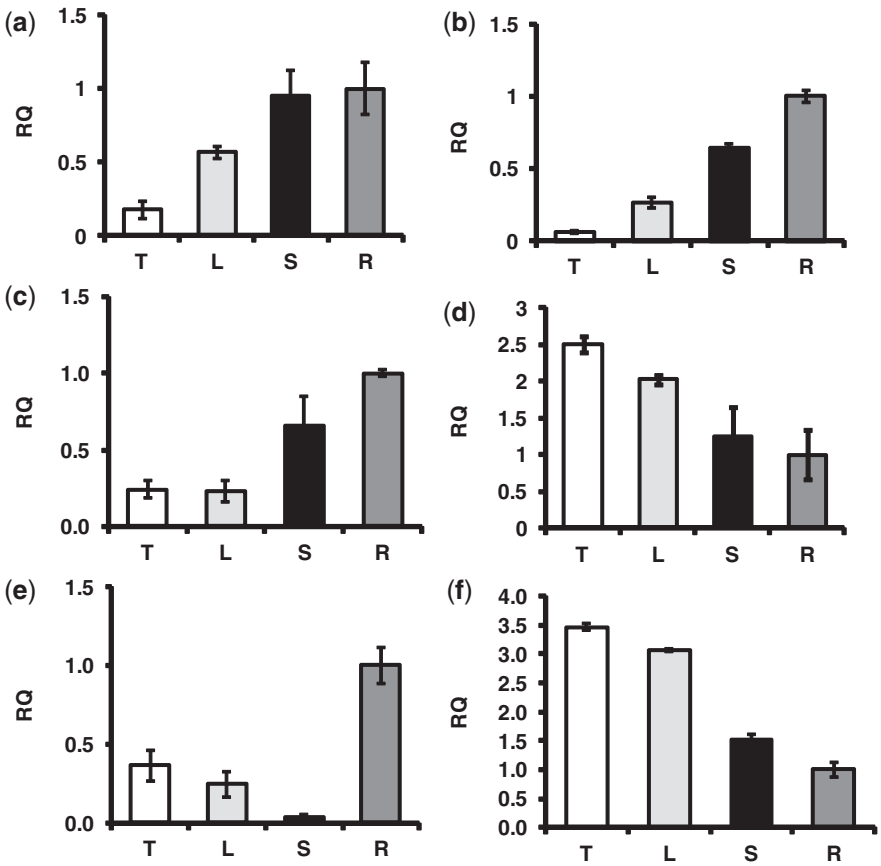

Fig. 3 Expression pattern of transcripts similar to *OB4CL*_ctg*1–5* of *O. basilicum* compared with *OS4CL* in different tissues of *O. sanctum*. RNA samples from different tissues (T, trichome; L, leaf; S, stem; R, root) were converted to cDNA and subjected to quantitative RT–PCR amplification with contig/gene-specific primer pairs. After normalization with endogenous control (actin), the quantification was carried out by calculation of ΔΔCt to determine the fold difference in gene expression [ΔCt target − ΔCt calibrator]. Relative quantity (RQ) was determined as $2^{-\Delta\Delta CT}$. Expression in different tissues was compared with a root RQ value = 1. (a–e) Expression patterns of *OB4CL*_ctg*1–5*, respectively. (f) Expression pattern of *OS4CL*. Data are means ± SD (for at least five biological replicates), and the *y*-axis represents the RQ. Similar expression patterns were observed for *OB4CL_ctg4* and *OS4CL*.

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

substrate by comparing the molecular mass obtained from mass spectra read in both positive $(M+H)^+$ and negative $(M\text{-}H)^-$ ionization modes. The results of LC-MS analysis indicated the utilization of COU, FER, CAF and CIN as substrates by OS4CL. Similarly, OB4CL was expressed and analyzed, and was observed to have comparable specific activity and similar substrate specificity to OS4CL (**Fig. 4**).

Table 1 Substrate specificity of *E. coli*-expressed OS4CL recombinant protein

Substrate	*Km* (μM)	V_{max} (μmol min^{-1} mg^{-1})	k_{cat} (s^{-1})
p-Coumaric acid (COU)	1.66 ± 0.06	703.65 ± 10.85	0.71
Ferulic acid (FER)	3.02 ± 0.64	770.55 ± 3.45	1.56
Caffeic acid (CAF)	4.94 ± 0.83	430.65 ± 17.55	0.44
trans-Cinnamic acid (CIN)	25.76 ± 1.2	572.3 ± 25.5	0.58
Sinapic acid (SIN)	ND	ND	ND

Specific activities are given as means ± SD (at least five independent assays). ND, not detected.

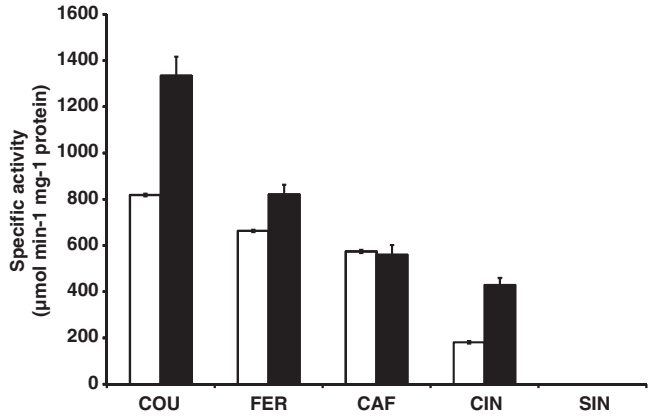

Fig. 4 Specific activities of OS4CL and OB4CL. *trans*-cinnamic acid (CIN), *p*-coumaric acid (COU), caffeic acid (CAF), ferulic acid (FER) and sinapic acid (SIN). Open bars represent the specific activity of OS4CL and filled bars represent the specific activity of OB4CL. Neither of the enzymes shows activity with SIN. Data are means ± SD (at least five replicates).

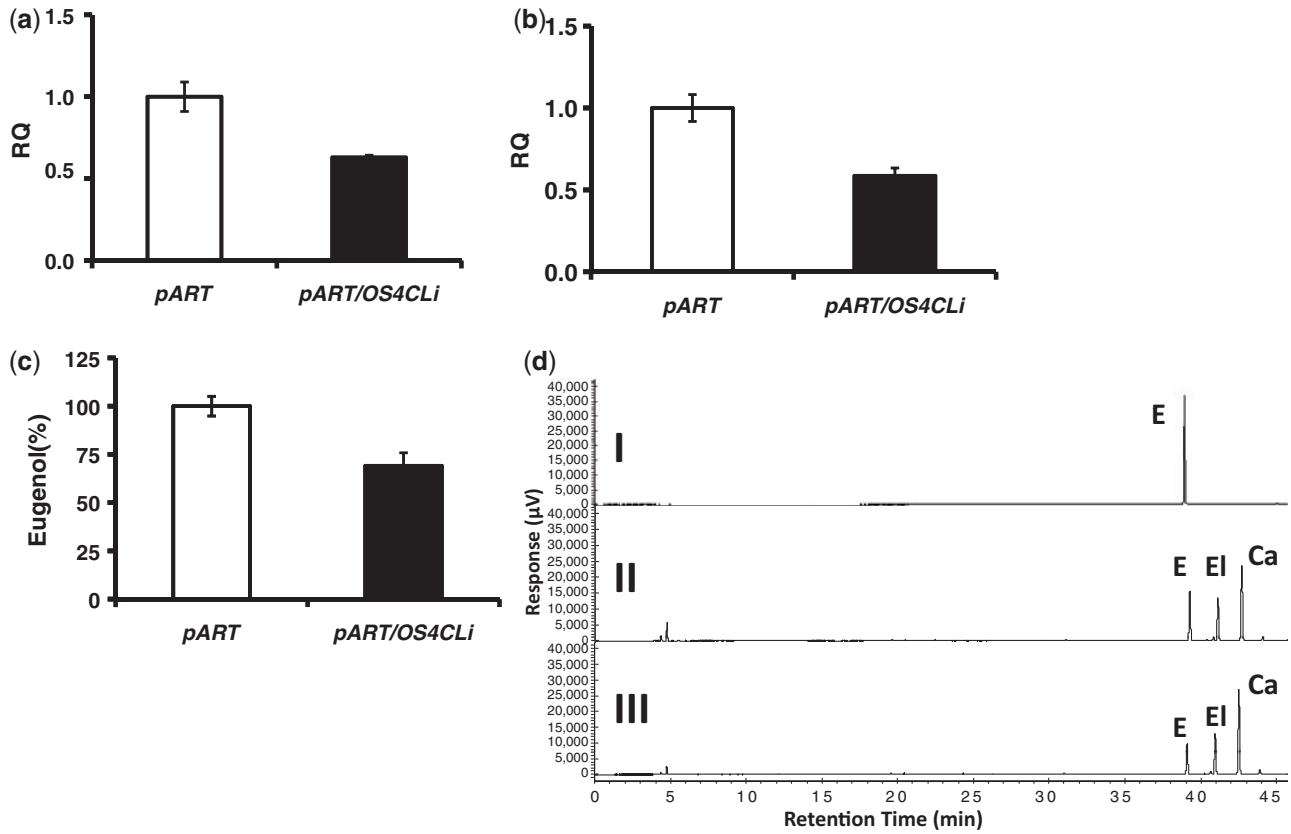

Fig. 5 Effect of pART/*OS4CLi* (RNAi construct) on the *OS4CL* transcript and eugenol level. (a and b) The relative abundance of *OS4CL* transcript in pART/*OS4CLi*-infiltrated samples compared with the pART vector using quantitative RT–PCR of cDNA from leaf trichome and whole leaf, respectively. Data are means ± SD (at least five biological replicates), and the *y*-axis represents relative quantity (RQ). (c) Relative eugenol percentage estimated by static headspace gas chromatography and compared with samples infiltrated with vector pART27. Data are means ± SD (at least five biological replicates). (d) Representative gas chromatograms for eugenol standard (I); eugenol in pART vector-infiltrated leaf samples (II); diminished peak of eugenol in pART/*OS4CLi*-infiltrated leaf samples (III). Other terpene volatiles, β-elemene (El) and β-caryophyllene (Ca), were also detected along with eugenol (E), the characteristic volatiles of *O. sanctum*.

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

Transient silencing of the *OS4CL* gene in planta

Since transformation of *O. sanctum* using *Agrobacterium* was not successful, the leaf infiltration approach was employed for transient RNA interference (RNAi) to suppress *OS4CL*. *Ocimum sanctum* twigs were infiltrated with *Agrobacterium tumefaciens* carrying pART/*OS4CLi* [pART27 carrying *OS4CL* RNAi sense and antisense strands under the *Cauliflower mosaic virus* (CaMV) 35S promoter] construct or with agrobacteria carrying a vector (pART27) with only the CaMV 35S promoter (control). Quantitative RT–PCR analysis of *OS4CL* from the mRNA showed a reduced expression in trichome (37%) and whole leaf (40%) (**Fig. 5a, b**). Quantitative analysis of eugenol showed a reduction of 31% in the infiltrated leaves with respect to the vector control (**Fig. 5c, d**). Analysis of endogenous levels of five phenolic acids in the acidified methanolic extract showed an increase in COU (~10-fold), FER (3.6-fold), CIN (84%) and CAF (21%), but the SIN content remained unchanged in both pART/*OS4CLi* and vector control leaf samples (**Fig. 6**). In addition, no significant effect of *OS4CL* suppression on the total lignin content of *O. sanctum* leaf was observed in both pART/*OS4CLi* and vector control samples, showing 26.85% and 26.77% (of leaf dry weight) of the total lignin content, respectively (**Supplementary Fig. S6**). No significant difference was also observed in the transcript levels related to *O. basilicum 4CL* isoforms, except for *OB4CL_ctg4* in the trichome of pART/*OS4CLi*-infiltrated plant samples compared with control (**Fig. 7**). The downstream genes after *4CL* in the eugenol biosynthetic pathway were analyzed for expression in RNAi-suppressed plant samples. Decreased expression of the genes *C3′H* (*p*-coumaroyl shikimate 3′-hydroxylase), *CCOMT* (caffeoyl-CoA *O*-methyl transferase), *CCR* (cinnamoyl-CoA reductase), *CAD* (cinnamyl alcohol dehydrogenase), *CAAT* (coniferyl alcohol acetyl transferase) and *EGS* (eugenol synthase) was observed in the suppressed samples compared with the control (**Fig. 8**).

Subcellular localization of the OS4CL protein

Green fluorescent protein (GFP) fluorescence was observed for both the full-length OS4CL coding sequence fused to the GFP reporter gene and for a truncated N-terminal OS4CL4 fraction

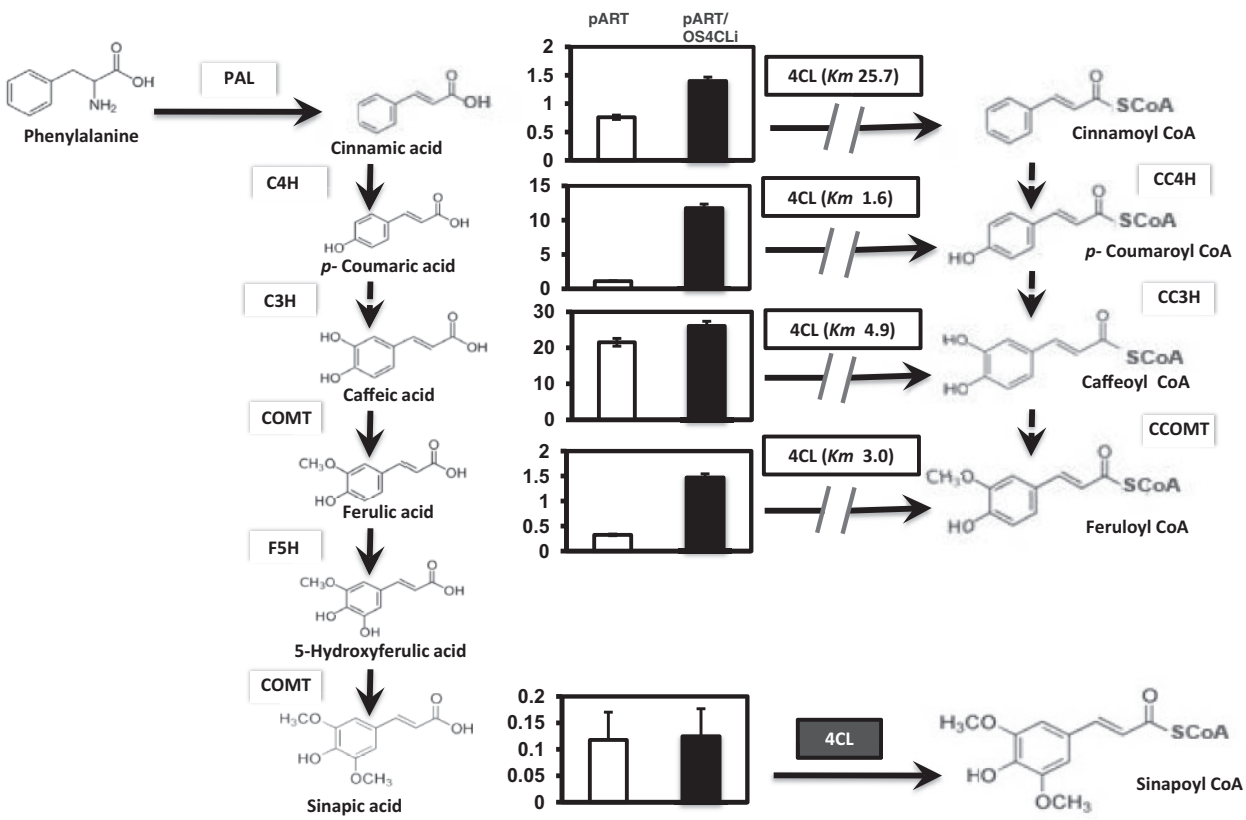

Fig. 6 Effect of pART/*OS4CLi* on the different substrates of the phenylpropanoid pathway. Schematic representation of major ligase reactions occurring in the phenylpropene pathway by 4CL. Broken arrows indicate inhibition of reactions by transient RNAi. The graphs at each reaction position indicate an increase in substrate (filled bars, pART/*OS4CLi*) compared with the control sample (open bars, pART27). The graph also shows no change in sinapic acid whereas the other (hydroxy) cinnamic acids increase. The *y*-axis represents the concentration in mM g^{-1} dry leaf tissue. Data are means ± SD (at least five biological replicates). K_m values are provided for each substrate as μM. PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; C3H, *p*-coumarate 3-hydroxylase; COMT, caffeic *O*-methyltransferase; F5H, ferulate-5-hydroxylase; CC4H, cinnamoyl-CoA 4-hydroxylase; CC3H, *p*-coumaroyl-CoA 3-hydroxylase; CCOMT, caffeoyl-CoA *O*-methyltransferase; 4CL, 4-coumarate: CoA ligase.

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

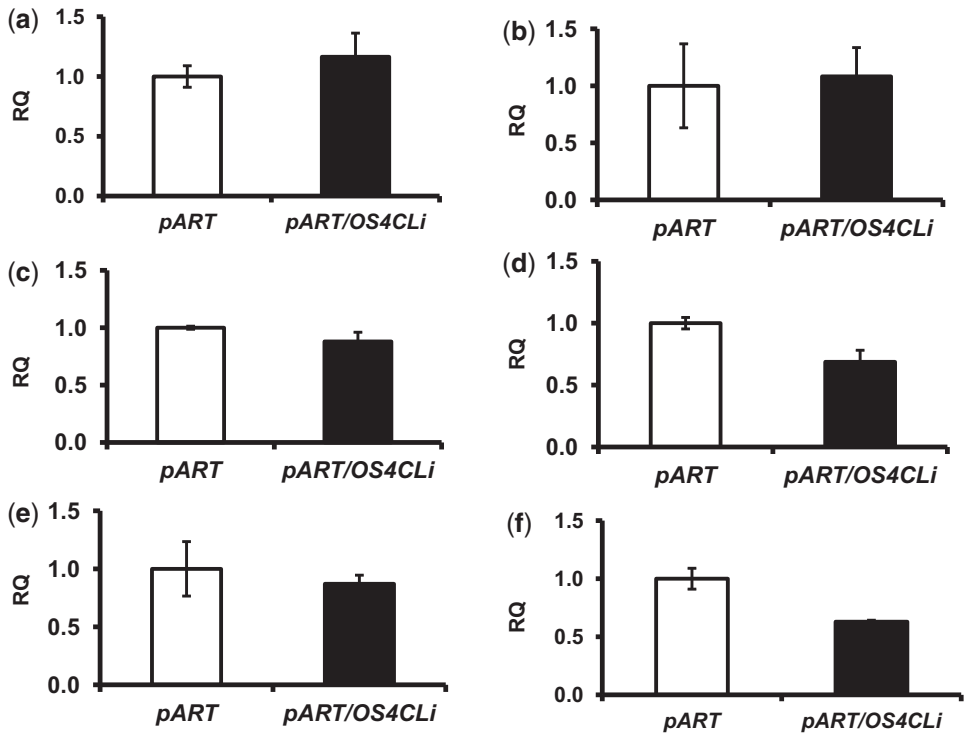

Fig. 7 Effect of pART/*OS4CLi* (RNAi construct) on the expression pattern of transcripts similar to isoforms *OB4CL_ctg1–5* of *O. basilicum* and *OS4CL* in the trichome of *O. sanctum*. (a–e) Expression patterns of *OB4CL_ctg1–5*, respectively. (f) Expression pattern of *OS4CL*. The *y*-axis represents relative quantity (RQ), and data are means ± SD (at least five biological replicates). Reduction of the *OB4CL_ctg4*-like transcript and *OS4CL* was observed in the trichome of pART/*OS4CLi*-infiltrated leaves.

fused to GFP after transient expression in *Arabidopsis* leaves. GFP was detectable as a diffuse signal in the cytosol of *Arabidopsis* protoplasts (**Fig. 9**). The pattern of fluorescence was similar to the fluorescence due to cytosolic localization of a GFP-alone construct.

Discussion

The enzyme 4CL is an important enzyme responsible for generating intermediates for phenylpropanoid biosynthesis. In this investigation, the genotype of *O. sanctum* 'CIM AYU' (Lal et al. 2003) having 83.56% eugenol in its essential oil was selected with the assumption that there was an active eugenol biosynthesis pathway in the trichome. Multiple isoforms of *4CL* genes have been identified and cloned in many plant species (Lee et al. 1995, Allina et al. 1998, Hu et al. 1998, Ehlting et al. 1999, Cukovic et al. 2001, Costa et al. 2005, Shi et al. 2010). Phylogenetic reconstruction studies from several plants indicate two large, evolutionarily ancient classes of *4CL* genes (Class I and Class II) (Ehlting et al. 1999). Class I *4CL* genes are more closely associated with the biosynthesis of lignin, as demonstrated by antisense or RNAi suppression studies (Kajita et al. 1996, Kajita et al. 1997, Lee et al. 1997, Li et al. 2003, Wagner et al. 2009, Gui et al. 2011). The likely involvement of Class II *4CL* genes is predicted for the biosynthesis of other phenolic compounds on the basis of compartmentalized expression in various plant organs and tissues (Hu et al. 1998, Ehlting et al. 1999), but this has never been demonstrated biochemically. A recent RNAi suppression study in *Petunia* indicated that the isoform Ph-*4CL1* has no effect on the benzenoid scent profile (Klempien et al. 2012). Hence, it is important to determine the role of 4CL and the route of metabolite channeling for the biosynthesis of commercially important phenylpropanoids such as eugenol.

OS4CL expression in the trichome correlates with phenylpropene biosynthesis

Although *4CL* genes have been characterized from many plants, the true size of the gene family is variable. Four members of the 4CL family has been identified in the moss *Physcomitrella patens* (Silber et al. 2008), whereas five have been found in rice (*Oryza sativa*) (Gui et al. 2011) and four in *Arabidopsis* (Ehlting et al. 1999). In addition nine *4CL*-like genes were predicted in the genome of *Arabidopsis* (Raes et al. 2003). In the aspen (*Populus trichocarpa*) genome, 17 genes are found to share sequence similarity with known *4CL* genes (Souza et al. 2008, Shi et al. 2010). Of these, five genes are classified as *4CL* genes and the rest as *4CL-like* genes due to their difference/similarity in structure (Souza et al. 2008). In gymnosperms and dicotyledonous angiosperms, *4CL* isoforms have been studied extensively. In *O. basilicum*, five isoforms were detected based on their sequence homology. Although *OB4CL_ctg1* and *4* are

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

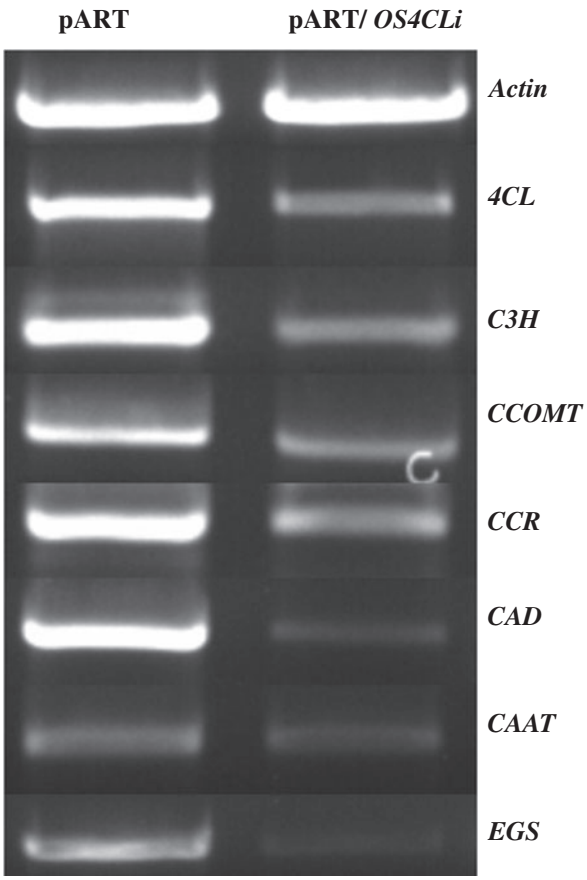

Fig. 8 Effect of *OS4CL* suppression on the expression of downstream genes. Expression of *4CL* (4-coumarate: CoA ligase), *C3′H* (*p*-coumaroyl shikimate 3′-hydroxylase), *CCOMT* (caffeoyl-CoA *O*-methyl transferase), *CCR* (cinnamoyl-CoA reductase), *CAD* (cinnamyl alcohol dehydrogenase), *CAAT* (coniferyl alcohol acetyl transferase) and *EGS* (eugenol synthase) was decreased in the *4CL*-supressed samples compared with the vector control.

highly similar in sequence, the highest expression of the *OB4CL_ctg4* ortholog was observed in the *O. sanctum* trichome (**Fig. 3**) whereas the *OB4CL_ctg1* ortholog had the highest expression in the root of *O. sanctum*. *OS4CL* of *O. sanctum* also showed an expression pattern similar to that of the *OB4CL_ctg4* ortholog in *O. sanctum*, indicating probable homologous counterparts in these two different species. Hence, for comparison with *OS4CL*, the contig *OB4CL_ctg4* was converted to full-length cDNA and named *OB4CL*. *OB4CL_ctg4* shared 80% and 85% nucleotide similarity with *OS4CL* and *OB4CL*, respectively (**Supplementary Fig. S3b**). The contig *OB4CL_ctg4* is incomplete and is the result of the assembly of 10 available raw ESTs from the database with probable sequencing errors/genotypic differences (**Supplementary Fig. S3c**), explaining the difference between *OB4CL* and *OB4CL_ctg4* sequences. However, involvement of *OB4CL_ctg 1* and *5* in trichome eugenol biosynthesis is unlikely, as these two orthologs of *O. basilicum* are expressed more in the root of *O. sanctum*, which is not the site of eugenol biosynthesis, compared with the *OB4CL_ctg4* (similar to *OS4CL*) ortholog expression pattern (**Fig. 3**). In basil, the EGS1 (eugenol synthase1) protein appears to be restricted to glands where eugenol, chavicol and their biosynthetic derivatives such as methyleugenol and methylchavicol are synthesized and stored (Gang et al. 2001). As transcript levels contribute significantly to the production of enzymes, the level of mRNA and protein may be correlated with the metabolic profiles for several enzymes residing at important regulatory points of the pathway (Xie et al. 2008). Since 4CL catalyzes the formation of *p*-coumaroyl-CoA, representing the branch point for the central phenylpropanoid pathway in plants, the detection of a higher transcript level of *OS4CL* in the trichome in combination with the highest enzyme activity with COU may be correlated with routing of the carbon flux towards high eugenol accumulation/biosynthesis in these oil glands. From sequence similarity and conserved domain analysis, *OS4CL* of this investigation was assigned to the Class II *4CL* genes. This group of genes are described for their likely involvement in the biosynthesis of different phenolic compounds. Hence, this implies the probable role of *OS4CL* in metabolite diversion towards eugenol in *Ocimum*. This relationship was also reinforced by the positive correlation between the *OS4CL* quantitative gene expression profile and eugenol content in essential oil at the plant developmental stages. The level of supplied substrate in the cell for a multistep pathway depends upon the initial steps regulating the carbon flux in the direction of the end-products. The efficiency of the intermediate enzymes also depends upon the availability of the flux to utilize the channeled carbon (Dudareva and Pichersky 2000).

Substrate specificity and specific role of OS4CL

The preference of substrates for the recombinant OS4CL was in the order of COU, FER, CAF and CIN. A similar order of preference was observed for OB4CL from *O. basilicum* (**Fig. 4**). Non-utilization of SIN was also confirmed by LC-MS analysis. Similar K_m values were reported earlier for these substrates in *Arabidopsis* (Ehlting et al. 1999, Costa et al. 2005). Lack of SIN activity and high COU activity does not conclusively suggest involvement of OS4CL in facilitating the biosynthetic pathway leading towards eugenol rather than S-lignin. However, as the trichome tissue is active in eugenol biosynthesis, diversion of the carbon towards eugenol by *OS4CL* and *OB4CL* may not be obscured. Further, these types of enzymes are shown to be channeling carbon towards flavonoid biosynthesis (Ehlting et al. 1999). As reported earlier, substitution or deletion of either Val355 or Leu356 generates a SIN-converting enzyme retaining the activity against COU, CAF and FER (Schneider et al. 2003) in *Arabidopsis*, whereas deletion of both inactivates the enzyme. OS4CL (HM990148) has both the valine and leucine residues, unlike the SIN-converting 4CL enzymes of *Glycine max* (AF279267) and *Arabidopsis* (AAM19949). In Gm4CL1 (AF279267), a proline residue was present in place of valine, while leucine was replaced by valine in At4CL4 (AAM19949) (**Supplementary Fig. S2**). Hence, the SBPalso proves OS4CL to be a specialized enzyme excluding SIN as a substrate. The amino acid sequence of *O. basilicum OB4CL* has a similar SBP which like

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

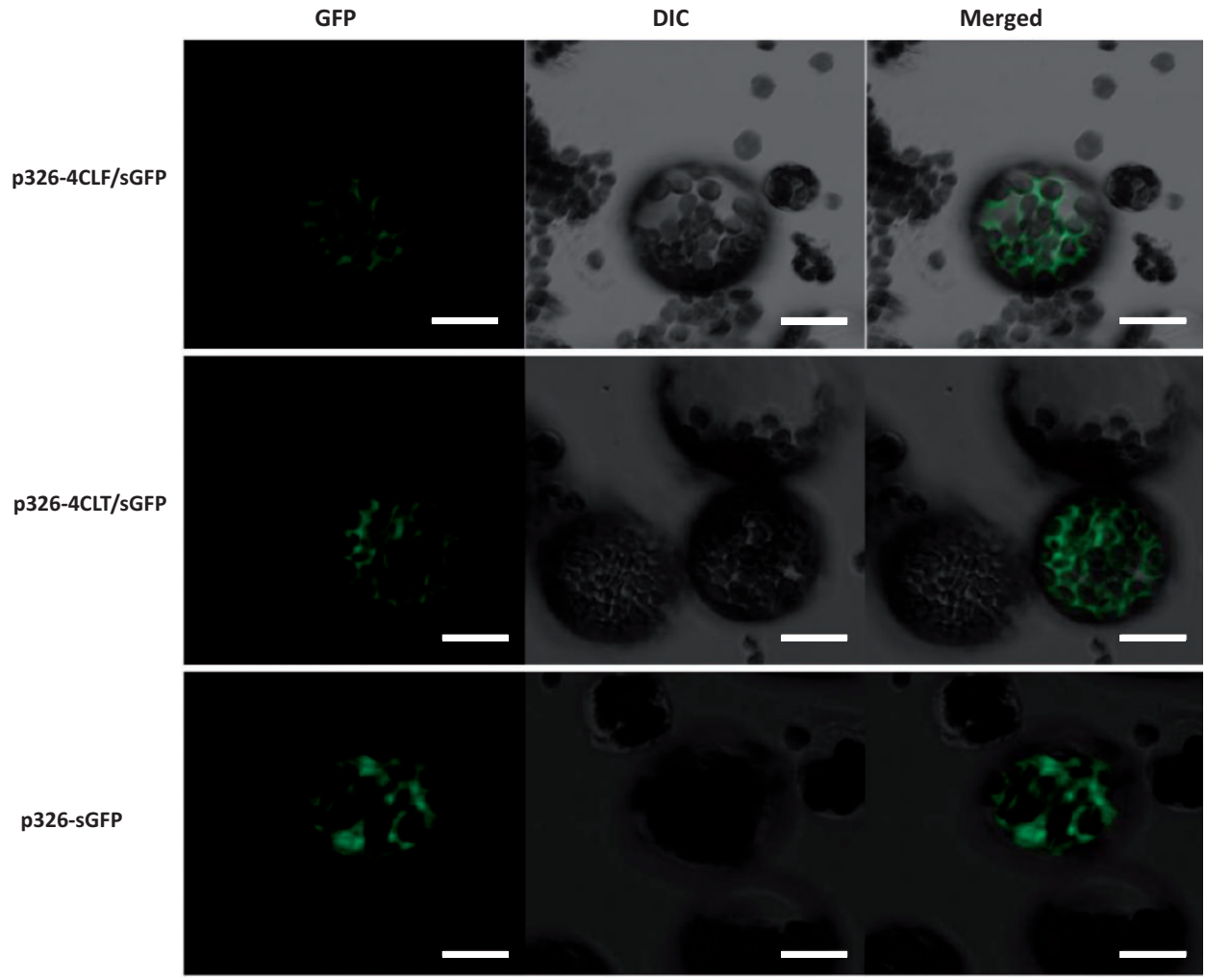

Fig. 9 The OS4CL protein was localized in the cytosol of *Arabidopsis* protoplasts in the transient expression assay. Diffuse GFP fluorescence localized outside the plastids is the same for the 4CLF (full length), 4CLT (truncated signal) and p326-sGFP constructs, indicating a cytosolic localization of the protein. Left, GFP image; center, differential interference contrast (DIC) images; right, merged image of GFP and DIC.

OS4CL does not show any activity for SIN. These results indicate a strong possibility of direct/indirect correlation of the *OS4CL* carbon flux towards eugenol biosynthesis.

In planta transient RNAi analysis for metabolite channeling

The transient expression of genes is an efficient alternative for some functional analyses where the study of gene function becomes limited by the low efficiency/recalcitrant transformation systems and the long time span needed for the generation of transgenic plants such as in rose (Yasmin and Debener 2010). In the absence of a perfect transformation protocol in *Ocimum*, an infiltration assay was preferred to demonstrate the function of *OS4CL* in planta and its involvement in channeling carbon towards eugenol biosynthesis. The enzyme 4CL is predicted to be the point of divergence from general phenylpropanoid metabolism to several major branch pathways (Hamberger and Hahlbrock 2003), including lignin and flavonoids (Cukovic et al. 2001). Transient suppression of *OS4CL* reduced the transcript levels for *OS4CL* in leaf trichome without affecting other isoforms, indicating a knock-down effect. Similarly, reduction of eugenol in RNAi construct-infiltrated leaf samples shows the probable effect of *OS4CL* transcript knock down, modulating its biosynthesis. The eugenol biosynthetic pathway includes several enzymes such as C3′H, CCOMT, CCR, CAD, CAAT and EGS. Initial commitment of intermediates at the level of 4CL channels the metabolites towards the action of these enzymes. Inhibition of an upstream enzyme affects the downstream product several steps away, as at each step the efficiency of the enzymes may decrease due to reduced availability of substrates. Song and Wang (2011) while studying the rosmarinic acid pathway in *Salvia miltiorrhiza* also observed a 20–70% decrease in the total phenolic content in PAL (phenylalanine ammonia lyase) suppressed lines, accompanied by low PAL activity. Also, Maeda et al. (2010) have shown the reduction of downstream benzoid compounds in the petals of petunia RNAi

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

lines for the arogenate dehydratase isoform *ADT1* responsible for phenylalanine biosynthesis. Because of low yield, difficulty in infiltration and estimation of metabolites from intact trichomes, leaf samples were analyzed for the metabolites after transient suppression. However, the possibility of other tissues in the leaf feeding the excess intermediate to the trichome for eugenol biosynthesis cannot be ruled out. This could be due to the huge amount of eugenol biosynthesized and stored in the trichome. The accumulation of COU, CIN, FER and CAF, but not SIN, in the acidified methanolic extract of infiltrated samples is consistent with the observation that OS4CL does not activate SIN. An increase in accumulation of CIN compared with CAF could be due to the utilization and conversion of CAF for another biosynthetic pathway in the trichome and leaf. As reported by Rubin and Jensen (1985), CAF is a non-competitive inhibitor of 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase, the first enzyme in the shikimate pathway. An alternative route could exist for SIN biosynthesis from CIN leading to S-lignin in plants. However, if these are operating to their full potential with metabolites from other isoforms already expressed in the trichome at a low level, they may not accept the accumulated metabolites beyond a threshold limit. This explains the accumulation of metabolites in transiently suppressed leaf. Though eugenol was analyzed from the leaf, considering the biosynthesis and storage of this compound in the trichome (Gang et al. 2001), the decrease in eugenol in RNAi-suppressed samples could be attributed to the eugenol from the trichome. Similarly, a decrease in *OS4CL* gene expression in the trichome also correlates well with the decrease in eugenol. As reported by Boerjan et al. (2003), down-regulation of *4CL* resulted in the reduction of lignin and cell wall-bound hydroxycinnamic acids (COU, FER and SIN) in tobacco and poplar. This differential effect on lignin and other compound is due to down-regulation of specific isoforms of 4CL with distinct substrate specificities (Lee et al. 1997, Allina et al. 1998, Hu et al. 1998, Hu et al. 1999, Harding et al. 2002). In this investigation, no significant difference was observed in either soluble or acid-insoluble lignin content in the transiently suppressed leaf for *OS4CL*, which immediately indicates that the suppression has no effect on biosynthesis of lignin. The leaves used in this analysis were at the growing stage with developing trichomes, and were expected to be active in de novo lignin production. However, considering the short duration of transient suppression, it is difficult to ascertain the effect of *OS4CL* on overall lignin biosynthesis.

The role of *C3′H*, *COMT*, *CCR* and *CAD* genes has been extensively investigated during lignification (Zhong et al. 2000, Blee et al. 2006, Rest et al. 2006, Thevenin et al. 2010), whereas their involvement in phenylpropene biosynthesis is still unexplored, except for *OMT* genes, *CAAT* and *EGS*. *OMT* genes are described to be multifunctional enzymes by Pak et al. (2004). These enzymes primarily function in lignin biosynthesis but their role in the synthesis of volatiles can also not be obscured (Wein et al. 2002). Dexter et al. (2007) have reported RNAi suppression of a petunia acyltransferase inhibiting isoeugenol biosynthesis. According to Koeduka et al. (2006), coniferyl acetate is the substrate for both eugenol and isoeugenol synthase. Reduced eugenol in transiently supressed plants may be due to decreased levels of substrates feeding the pathway and is supported by the observation of decreased expression of downstream pathway genes such as *C3′H*, *CCOMT*, *CCR*, *CAD*, *CAAT* and *EGS*. This indicates the cascading effect of *4CL* gene supression on downstream genes (**Fig. 8**).

The cellular localization of OS4CL indicates its function

The enzyme 4CL has commonly been ascribed to be localized in the cytosol. However, one isoform of *Sorbus aucuparia* 4CL is predicted to be localized to the chloroplast (Gaid et al. 2011) and a 4CL-like member of the acyl-activating enzyme gene family of *Arabidopsis* was assigned to the peroxisome (Endler et al. 2008). In this investigation, OS4CL was observed to be located in the cytosol like many other 4CLs.

Although a number of 4CLs have been identified in many plants, the importance of these enzymes in creating virtual compartments through substrate utilization in a tissue-specific manner and diverting intermediates towards phenylpropene (eugenol) biosynthesis has here been investigated and described for the first time. The highest expression of this enzyme occurring in the trichome combined with its substrate specificity towards COU, FER, CAF and CIN, and their accumulation in the transiently suppressed leaf sample for *OS4CL* indicated the role of this enzyme in predominantly channeling metabolites for eugenol biosynthesis. The *OB4CL* gene isolated from *O. basilicum* was also demonstrated to have similar activity and substrate preference, indicating a similar role to *OS4CL*.

Materials and Methods

Plant material

Seeds and plants of *O. sanctum* 'CIM AYU' having 83.56% eugenol in the essential oil (Lal et al. 2003) and *O. basilicum* 'CIM SAUMYA' having 62.8% chavicol (Lal et al. 2004) were acquired from the core collection of basil germplasm stored in the National Gene Bank for Medicinal and Aromatic Plants (NGBMAP) established at CIMAP, Lucknow (India).

Isolation and analysis of full-length *4CL* cDNAs

Glandular trichomes were isolated from young leaves of *O. sanctum* by following the protocol based on the glass bead abrasion technique (McCaskill et al. 1992). About 100 mg of trichome material was used for total RNA isolation using the cetyltrimethyl ammonium bromide (CTAB) method (Shukla et al. 2005), and 5 μg of RNA was taken to synthesize cDNA using the Thermoscript RT PCR System (Invitrogen). From eight different groups based on homology of 4CL sequences, eight pairs of degenerate primers were designed. Of these, only one primer pair corresponding to the group containing the

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

Agastache rugosa (Lamiaceae) *4CL* gene (AY616436) resulted in amplification of a 531 bp fragment, which was cloned in the pGEM-T Easy vector (Promega). Nucleotide sequencing confirmed it as a *4CL* gene, and it was designated as *OS4CL* (*O. sanctum 4CL*). The 5′ and 3′ regions of *OS4CL* were isolated using a Smart RACE cDNA Amplification kit (Clontech) and 3′ RACE System (Invitrogen) from poly(A)$^+$ mRNA, respectively. The full-length cDNA sequence of *OS4CL* was deduced, amplified and cloned using specific primers 4CL_F, 5′-ACGGATCCATGTTGTCCGTGGCGGAAGCTCAGAATTCG-3′ (*Bam*HI site underlined); and 4CL_R, 5′-ACAAGCTTTTAAGAGGTGGAGGAAGGTGCAGCAAG-3′ (*Hin*dIII site underlined) for characterization. Sequence alignments, open reading frame (ORF) translation and molecular mass calculation of the predicted protein were carried out with the Sequence Manipulation Suite 2.0 (http://www.bioinformatics.org/sms2/). BLAST was performed at the NCBI server (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi), whereas structural analysis of the predicted OS4CL was carried out at the Expasy website (http://cn.expasy.org). Similarly, with the help of *OS4CL*-specific primers, the full-length *O. basilicum* 4CL (*OB4CL*) cDNA was amplified, isolated and cloned for further characterization.

Phylogenetic analysis

Amino acid sequences of *O. sanctum 4CL* (*OS4CL*) and *4CL* genes reported from other species were aligned and phylogeny was reconstructed using the bootstrap Neighbor–Joining (N–J) distance method of CLUSTAL X version 2.0.11 (Thompson et al. 1997). TreeView software version 1.6.6 was used to view the resulting bootstrap N–J tree (Saitou and Nei 1987). GenBank accession numbers for the corresponding amino acid sequences used in the alignment are: *Ocimum sanctum* OS4CL (HM990148); *Ocimum basilicum* OB4CL (KC576841); *Nicotiana tabacum* Nt4CL (D43773); *N. tabacum* Nt4CL1 (U50845); *N. tabacum* Nt4CL2 (U50846); *Oryza sativa* Rice4CL1 (X52623); *O. sativa* Rice4CL2 (L43362); *Petroselinum crispum* Pc4CL1 (X13324); *P. crispum* Pc4CL2 (X13325); *Pinus taeda* Pt4CL1 (U12012); *P. taeda* Pt4CL2 (U12013); *Populus hybrida* Poplar4CL1 (AF008184); *P. hybrida* Poplar4CL2 (AF008183); *Rubus idaeus* Ri4CL1 (AF239687); *R. idaeus* Ri4CL2 (AF239686); *R. idaeus* Ri4CL3 (AF239685); *Solanum tuberosum* Potato4CL1 (M62755); *S. tuberosum* Potato4CL2 (AF150686); *S. tuberosum* Potato4CL2a (AAD40664); Potato4CL2b (AF150687); *Sorghum bicolor* sorghum4CL (U23787); *Populus tremuloides* Aspen4CL1 (AF041049); *P. tremuloides* Aspen4CL2 (AF041050); *Prunus avium* Pa4CL1(GU990523); *Sorbus aucuparia* Sa4CL3 (GU949553); *Allium cepa* Ac4CL (AY541033); *Agastache rugosa* Ar4CL (AY587891); *Camellia sinensis* Cs4CL (DQ194356); *Arabidopsis thaliana* At4CLfp (NM_123172); *A. thaliana* At4CL1 (U18675); *A. thaliana* At4CL2 (AF106086); *A. thaliana* At4CL3 (AF106088); *Glycine max* Gm4CL1 (AF279267); *G. max* Gm4CL2 (AF002259); *G. max* Gm4CL3 (AF002258); *G. max* Gm4CL4 (X69955); *G. max* Gm4CL14 (X69954); *Lithospermum erythrorhizon* Le4CL1 (D49366); *L. erythrorhizon* Le4CL2 (D49367); *Lolium perenne* Lp4CL1 (AF052221); *L. perenne* Lp4CL2 (AF052222); *L. perenne* Lp4CL3 (AF052223).

Quantitative and semi-quantitative RT–PCR analysis

Quantitative RT–PCR was carried out following the protocol described by Maeda et al. (2010) using SYBR Green chemistry (Applied Biosystems). *OS4CL*, its isoforms and other gene-specific primers were designed with Primer Express Software version 2.0 (Applied Biosystems) and custom synthesized by Sigma Aldrich, India (**Supplementary Table S7a**). The reactions were carried out in five biological replicates in a '7900HT Fast Real Time PCR System' (Applied Biosystems), and the specificity of the reactions was verified by melting curve analysis with the thermal cycling parameters: initial hold (50°C for 2 min); initial denaturation (95°C for 10 min); and 40 amplification cycles (95°C for 15 s; and 60°C for 1 min) followed by additional steps (60°C for 15 s, 95°C for 15 s and 37°C for 2 min). Relative mRNA levels were quantified with respect to the internal/endogeneous control '*actin* of *O. basilicum* (TCOB40902)' (Raes et al. 2003, Misra et al. 2012). Sequence Detection System (SDS) software version 2.2.1 was used for relative quantification of gene transcripts using the $\Delta\Delta C_T$ method. Threshold cycle (Ct) values obtained after real-time PCR were used for calculation of the ΔCt value (target-endogenous control). The quantification was carried out by calculating $\Delta\Delta$Ct to determine the fold difference in gene expression [ΔCt target – ΔCt calibrator]. The RQ was determined as $2^{-\Delta\Delta CT}$. Semi-quantitative RT–PCR was performed using the protocol: 94°C for 2 min followed by 29 cycles of amplification (94°C for 1 min, 60°C for 1 min, 72°C for 2 min) and final extension at 72°C for 10 min with the actin gene of *O. basilicum* as control. Details of the primers used in the semi-quantitative RT–PCR have been provided in **Supplementary Table S7b**.

Expression of 4CL proteins in *E. coli*

The ORFs of isolated *4CL* cDNAs were cloned in the pET 28a(+) vector in-frame with the C-terminus of the His_6-tag and expressed in *E. coli* [Tuner (DE3)] as described by Koeduka et al. (2006). Protein was extracted from exponentially growing bacteria (3 liters of LB medium) after overnight induction at 15°C with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG; $A_{600} = 0.7$). His_6-tagged 4CL protein was purified using Ni-NTA spin columns (Novagen), analyzed by SDS–PAGE and estimated using Bradford Reagent (Sigma).

Enzyme assays

Enzyme activity was assayed in a standard assay buffer containing 100 mM Tris pH 7.5, 2.5 mM ATP, 2.5 mM $MgCl_2$, 0.2 mM CoA, 0.2 mM substrates (CIN, COU, CAF, FER or SIN) and purified 4CL proteins. The reaction was started by adding CoA at room temperature, and the formation of the CoA esters was monitored at the absorption maxima of 311, 333, 345, 346 and 352 nm for cinnamoyl-CoA, 4-coumaroyl-CoA, feruloyl-CoA,

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

caffeoyl-CoA and sinapoyl-CoA, respectively, for 10 min (Webster et al. 1974, Stockigt and Zenk 1975). Kinetic constants (K_m and V_{max}) for 4CL were analyzed at pH 7.5 and 30°C, with substrate concentrations ranging from 5 to 800 μM for hydroxycinnamic derivatives, and from 50 to 4,000 μM for CIN using 10 μg of purified protein. K_m and V_{max} were estimated from Lineweaver–Burk plots (Costa et al. 2005) from the average of five replicates.

The reaction was terminated by boiling for 2 min, followed by centrifugation, and the clear supernatant was analyzed by HPLC using the modified protocol of Lee et al. (2007). The flow rate was maintained at 1 ml min^{-1} and UV detection was at 340 nm. LC-MS was performed using an LCMS-2010 EV liquid chromatograph–mass spectrometer (Shimadzu) equipped with an electrospray ionization (ESI) source to confirm the formation of CoA esters.

RNAi gene construct and infiltration for transient suppression

The RNAi construct was prepared by amplifying the sense and antisense gene fragments corresponding to the +24 to +512 region of *OS4CL* from the ATG codon. The primer pair 5′- AGAATTCGGAGCTCTCATCTCAT-3′ and 5′-GGGTACCAAATCCTCACCGAATA-3′ was used to obtain the fragment in the sense orientation, and the pair 5′-CTCTAGAAGAATTCGGAGCTCTC-3′ and 5′-GAAGCTTAAATCCTCACCGAATACG-3′ was used to amplify the antisense fragment. These fragments were cloned sequentially on either side of the intron between *Eco*RI/*Kpn*I and *Xba*I/*Hin*dIII restriction sites of the pHANNIBAL vector (CSIRO), respectively, to obtain the two arms of the hairpin. The complete hairpin cassette was cloned into the pART27 binary vector (Gleave 1992) using the *Not*I restriction site (pART/*OS4CLi*). The initial 489 bp from the 5′ end of *OS4CL* shows a very low sequence similarity with the sinapate-converting *4CL* gene isoforms of *G. max* (52%) and *A. thaliana* (39%). Hence this region was selected to be included in the RNAi construct for specificity. *Agrobacterium* strain GV3103 was transformed separately by the pART27 binary vector with and without the hairpin cassette. Transient RNAi assay in planta was carried out by 'vacuum infiltration', for which a fresh culture of *Agrobacterium* was inoculated into 10 ml of YEP broth with antibiotics (gentamycin 40 μg ml^{-1} and kanamycin 50 μg ml^{-1}) and was grown overnight at 28°C for 18 h. The bacterial pellet was collected by centrifugation at 3,000 r.p.m. (4°C, 15 min), washed with 10 ml of infiltration buffer [50 mM MES pH 5.7, 0.5% (w/v) glucose, 2 mM Na_3PO_4, 100 μM acetosyringone] and resuspended in the same solution to obtain an $OD_{600\,nm}$ of 0.3. The plant samples (6-week-old *O. sanctum* twigs containing 4–5 leaves) were collected after 90 min of spraying with water for stomatal opening. These were submerged in the infiltration solution with *Agrobacterium* in a beaker placed in the vacuum assembly with a vacuum of 15 mmHg until the air bubbles started to come out from the tissue vigorously. The vacuum was slowly released and at this stage the leaves appeared to be becoming dark green and translucent because of the infiltration. After confirming the infiltration by visual inspection, the twigs were dipped in 4% sucrose, covered by cling film pierced at 4–5 places for aeration and incubated in the dark at 24–25°C for 48 h. Finally the dark green and translucent infiltrated leaves were plucked, weighed to analyze the metabolites, and trichomes were isolated for quantitative RT–PCR. Leaves from the same pool were dried to analyze lignin.

Metabolite analysis

Total eugenol was quantified by a Varian CP-3800 GC FID equipped with a Combi Pal static headspace autosampler following the protocol of Sitaramaraju et al. (2008) with slight modification. After standardization for maximum detection of eugenol, extraction from 0.5 g of chopped leaf sample was carried out for 15 min in 20 ml vials (Supelco) sealed hermetically with a PTFE (polytetrafluoroethylene)/white silicone septum and a cap, with oven temperature 100°C and syringe temperature 120°C of the headspace sampler. Samples were injected (500 μl) onto the gas chromatograph equipped with a CP-Sil 8 CB Column (30 m × 0.32 mm i.d., with 1 μm film thickness) and analyzed using a flame ionization detector. The oven temperature was programmed from 60 to 240°C at 3°C min^{-1} and then up to 320°C at 5°C with the flame ionization detection temperature set at 280°C. Hydrogen (constant flow 1 ml min^{-1}) was used as carrier, and identification was based on retention indices, co-injection with a standard and elution order in CP-Sil 8 CB.

The phenolics were extracted from 0.5 g of dried and powdered infiltrated leaves by adding aqueous methanol [10 ml, 70% (v/v)] followed by sonication for 1 h at 60°C. The mixture was filtered, followed by the addition of 10 ml of 6 M hydrochloric acid. In each sample, nitrogen was bubbled for about 40–60 s and refluxed in a water bath at 90°C until dry. Dried extract was resuspended in 1 ml of methanol and centrifuged at 12,000 r.p.m. for 5 min. Supernatant was collected into a fresh tube, filtered through a 0.45 μm membrane filter (Millex-HV) and analyzed (Proestos et al. 2008) similarly to as described for LC-MS analysis.

Lignin extraction and analysis

Lignin content was estimated using the protocol described by Mann et al. (2009). This method involves two-step acid hydrolysis for fractionating carbohydrates and lignin into forms that are easily quantifiable. The analyses were carried out in five replicates. Lignin quantity was estimated as the percentage dry weight of leaves.

Subcellular localization of OS4CL protein

The ORFs of the full length (567 amino acids, OS4CLF) and the first 100 amino acids of the N-terminal region of OS4CL (OS4CLT), i.e. the leader peptide region, were fused separately upstream of GFP in the cloning sites *Xba*I and *Bam*HI of the p326-sGFP vector containing the CaMV 35S promoter (a gift

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

from Dr. I. Hwang, Pohang University of Science and Technology, Korea). Nucleotide sequencing confirmed the accuracy of the in-frame fusion. Each fusion construct, one with the complete gene (p326-OS4CLF/SGFP) and the other with only the leader peptide region of 100 amino acids (p326-OS4CLT/SGFP), and p326-sGFP vector control (15–20 μg of plasmid DNA) were separately transformed into *Arabidopsis* protoplasts to analyze the transient expression of GFP fusion proteins as described by Sheen (2002) and Nagegowda et al. (2008) using a confocal laser scanning microscope (Zeiss LSM 510) with a 40 × 1.55 numerical aperture lens. Fluorescence was collected with a bandpass filter and GFP excitation was measured at 488 nm with an argon laser, while the fluorescence was detected at 520 nm.

Supplementary data

Supplementary data are available at PCP online.

Funding

This work was supported by the Council of Scientific and Industrial Research (CSIR)-Central Institute of Medicinal and Aromatic Plants, Lucknow; CSIR [Senior Research Fellowship to S.R. and Tenth Five Year Plan projects (BSC0203 and BSC0107)].

Acknowledgments

The authors express their sincere gratitude to Professor Ram Rajasekharan, Former Director, CSIR-CIMAP for his keen interest, and for critically going through and editing the manuscript. The plant material provided by NGBMAP and help by CSIR-NBRI for confocal microscopy experiments are also acknowledged. Mr. Michael Long, Purdue University helped in reading and editing the manuscript.

References

Allina, S.M., Pri-Hadash, A., Theilmann, D.A., Ellis, B.E. and Douglas, C.J. (1998) 4-Coumarate:coenzyme A ligase in hybrid poplar. Properties of native enzymes, cDNA cloning, and analysis of recombinant enzymes. *Plant Physiol.* 116: 743–754.

Blee, K.A., Choi, J.W., O'Connell, A.P., Schuch, W., Lewis, N.G. and Bolwell, G.P. (2003) A lignin-specific peroxidase in tobacco whose antisense suppression leads to vascular tissue modification. *Phytochemistry* 64: 163–176.

Boerjan, W., Ralph, J. and Baucher, M. (2003) Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54: 519–46.

Costa, M.A., Bedgar, D.L., Moinuddin, S.G.A., Kim, K., Cardenas, C.L., Cochrane, F.C. et al. (2005) Characterization in vitro and in vivo of the putative multigene 4-coumarate:CoA ligase network in *Arabidopsis*: syringyl lignin and sinapate/sinapyl alcohol derivative formation. *Phytochemistry* 66: 2072–2091.

Cukovic, D., Ehlting, J., VanZiffle, J.A. and Douglas, C.J. (2001) Structure and evolution of 4-coumarate:coenzyme A ligase (*4CL*) gene families. *Biol. Chem.* 382: 645–654.

Dexter, R., Qualley, A., Kish, C.M., Ma, C.J., Koeduka, T., Nagegowda, D.A. et al. (2007) Characterization of a petunia acetyltransferase involved in the biosynthesis of the floral volatile isoeugenol. *Plant J.* 49: 265–275.

Dixon, R.A. and Paiva, N.L. (1995) Stress-induced phenylpropanoid metabolism. *Plant Cell* 7: 1085–1097.

Douglas, C.J. (1996) Phenylpropanoid metabolism and lignin biosynthesis, from weeds to trees. *Trends Plant Sci.* 1: 171–178.

Dudareva, N. and Pichersky, E. (2000) Biochemical and molecular genetics aspects of floral scents. *Plant Physiol.* 122: 627–633.

Ehlting, J., Buttner, D., Wang, Q., Douglas, C.J., Somssich, I.E. and Kombrink, E. (1999) Three 4-coumarate:coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *Plant J.* 19: 9–20.

Ehlting, J., Shin, J.J.K. and Douglas, C.J. (2001) Identification of 4-coumarate:coenzyme A ligase (4CL) substrate recognition domains. *Plant J.* 27: 455–465.

Endler, A., Martens, S., Wellmann, F. and Matern, U. (2008) Unusually divergent 4-coumarate:CoA-ligases from *Ruta graveolens* L. *Plant Mol. Biol.* 67: 335–346.

Gaid, M.M., Scharnhop, H., Ramadan, H., Beuerle, T. and Beerhues, L. (2011) 4-Coumarate:CoA ligase family members from elicitor-treated *Sorbus aucuparia* cell cultures. *J. Plant Physiol.* 168: 944–951.

Gang, D.R., Simon, J., Lewinsohn, E. and Pichersky, E. (2002) Peltate glandular trichomes of *Ocimum basilicum* L. (Sweet Basil) contain high levels of enzymes involved in the biosynthesis of phenylpropenes. *J. Herb Spice Med. Plant* 9: 189–195.

Gang, D.R., Wang, J., Dudareva, N., Nam, K.H., Simon, J.E., Lewinsohn, E. et al. (2001) An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. *Plant Physiol.* 125: 539–555.

Gleave, A.P. (1992) A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol. Biol.* 20: 1203–1207.

Gui, J., Shen, J. and Li, L. (2011) Functional characterization of evolutionarily divergent 4-coumarate:coenzyme A ligases in rice. *Plant Physiol.* 157: 574–586.

Hahlbrock, K. and Scheel, D. (1989) Physiology and molecular biology of phenylpropanoid metabolism. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 40: 347–369.

Hamberger, B. and Hahlbrock, K. (2003) The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proc. Natl Acad. Sci. USA* 101: 2209–2214.

Harding, S.A., Leshkevich, J., Chiang, V.L. and Tsai, C.J. (2002) Differential substrate inhibition couples kinetically distinct 4-coumarate:coenzyme A ligases with spatially distinct metabolic roles in quaking aspen. *Plant Physiol.* 128: 428–438.

Hu, W.J., Kawaoka, A., Tsai, C.J., Lung, J., Osakabe, K., Ebinuma, H. et al. (1998) Compartmentalized expression of two structurally and functionally distinct 4-coumarate:CoA ligase genes in aspen (*Populus tremuloides*). *Proc. Natl Acad. Sci. USA* 95: 5407–5412.

Hu, W.J., Harding, S.A., Lung, J., Popko, J.L. and Ralph, J. (1999) Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nat. Biotechnol.* 17: 808–812.

Huang, X. and Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Res.* 9: 868–877.

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

Kajita, S., Hishiyama, S., Tomimura, Y., Katayama, Y. and Omori, S. (1997) Structural characterization of modified lignin in transgenic tobacco plants in which the activity of 4-coumarate:coenzyme A ligase is depressed. *Plant Physiol.* 114: 871–879.

Kajita, S., Katayama, Y. and Omori, S. (1996) Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate:coenzyme A ligase. *Plant Cell Physiol.* 37: 957–965.

Klempien, A., Kaminaga, Y., Qualley, A., Nagegowda, D.A., Widhalm, J.R., Orlova, I. et al. (2012) Contribution of CoA ligases to benzenoid biosynthesis in *Petunia* flowers. *Plant Cell* 24: 2015–2030.

Koeduka, T., Fridman, E., Gang, D.R., Vassao, D.G., Jackson, B.L., Kish, C.M. et al. (2006) Eugenol and isoeugenol, characteristic aromatic constitutents of spices, are biosynthesized *via* reduction of a coniferyl alcohol ester. *Proc. Natl Acad. Sci. USA* 103: 10128–10133.

Koeduka, T., Louie, G.V., Orlova, I., Kish, C.M., Ibdah, M., Wilkerson, C.G. et al. (2008) The multiple phenylpropene synthases in both *Clarkia breweri* and *Petunia hybrida* represent two distinct protein lineages. *Plant J.* 54: 362–374.

Lal, R.K., Khanuja, S.P.S., Agnihotri, A.K., Misra, H.O., Shasany, A.K., Naqvi, A.A. et al. (2003) High yielding eugenol rich oil producing variety of *Ocimum sanctum*-'CIM-Ayu'. *J. Med. Arom. Plant Sci.* 25: 746–747.

Lal, R.K., Khanuja, S.P.S., Agnihotri, A.K., Shasany, A.K., Naqvi, A.A., Dwivedi, S. et al. (2004) An early, short duration, high essential oil, methyl chavicol and linalool yielding variety of Indian basil (*Ocimum basilicum*) 'CIM-Saumya'. *J. Med. Arom. Plant Sci.* 26: 77–78.

Lee, D., Ellard, M., Wanner, L.A., Davis, K.R. and Douglas, C.J. (1995) The *Arabidopsis* 4-coumarate:CoA ligase (*4CL*) gene, stress and developmentally regulated expression and nucleotide sequence of its cDNA. *Plant. Mol. Biol.* 28: 871–884.

Lee, D., Meyer, K., Chapple, C. and Douglas, C.J. (1997) Antisense suppression of 4- coumarate:coenzyme A ligase activity in *Arabidopsis* leads to altered lignin subunit composition. *Plant Cell* 9: 1985–1998.

Lee, Y.J., Jeon, Y., Lee, J.S., Kim, B.G., Lee, C.H. and Ahn, J.H. (2007) Enzymatic synthesis of phenolic CoAs using 4-coumarate:coenzyme A ligase (4CL) from rice. *Bull. Korean Chem. Soc.* 28: 365–366.

Li, L., Zhou, Y., Cheng, X., Sun, J., Marita, J.M., Ralph, J. et al. (2003) Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. *Proc. Natl Acad. Sci. USA* 100: 4939–4944.

Lindermayr, C., Möllers, B., Fliegmann, J., Uhlmann, A., Lottspeich, F., Meimberg, H. et al. (2002) Divergent members of a soybean (*Glycine max* L.) 4-coumarate:coenzyme A ligase gene family. *Eur. J. Biochem.* 269: 1304–1315.

Maeda, H., Shasany, A.K., Schnepp, J., Orlova, I., Taguchi, G., Cooper, B.R. et al. (2010) RNAi suppression of arogenate dehydratase1 reveals that phenylalanine is synthesized predominantly *via* the arogenate pathway in petunia petals. *Plant Cell* 22: 832–849.

Mann, D.G.J., Labbe, N., Sykes, R.W., Gracom, K., Kline, L., Swamidoss, I.M. et al. (2009) Rapid assessment of lignin content and structure in switchgrass (*Panicum virgatum* L.) grown under different environmental conditions. *Bioenerg. Res.* 2: 246–256.

McCaskill, D., Gershenzon, J. and Croteau, R. (1992) Morphology and monoterpene biosynthetic capabilities of secretory-cell clusters isolated from glandular trichomes of peppermint (*Mentha piperita* L.). *Planta* 187: 445–454.

Misra, A., Chanotiya, C.S., Gupta, M.M., Dwivedi, U.N. and Shasany, A.K. (2012) Characterization of cytochrome P450 monooxygenases isolated from trichome enriched fraction of *Artemisia annua* L. leaf. *Gene* 510: 193–201.

Nagegowda, D.A., Gutensohn, M., Wilkerson, C.G. and Dudareva, N. (2008) Two nearly identical terpene synthases catalyze the formation of nerolidol and linalool in snapdragon flowers. *Plant J.* 55: 224–239.

Nahrstedt, A. (1990) The significance of secondary metabolites for interaction between plants and insects. *Planta Med.* 55: 333–338.

Obeng-Ofori, D. and Reichmuth, C. (1997) Bioactivity of eugenol, a major component of essential oil of *Ocimum suave* (Wild.) against four species of stored-product Coleoptera. *Int. J. Pest Manage.* 43: 89–94.

Pak, F.E., Gropper, S., Dai, W.D., Havkin-Frenkel, D. and Belanger, F.C. (2004) Characterization of a multifunctional methyltransferase from the orchid Vanilla planifolia. *Plant Cell Rep.* 22: 959–966.

Prasad, N.S., Raghavendra, R., Lokesh, B.R. and Naidu, K.A. (2004) Spice phenolics inhibit human PMNL 5-lipoxygenase. *Prostaglandins Leukot. Essent. Fatty Acids* 70: 521–528.

Proestos, C., Kapsokefalou, M. and Komaitis, M. (2008) Analysis of naturally occurring phenolic compounds in aromatic plants by RP-HPLC and GC-MS after silylation. *J. Food Qual.* 31: 402–414.

Raes, J., Rohde, A., Christensen, J.H., de Peer, Y.V. and Boerjan, W. (2003) Genome-wide characterization of the lignification toolbox in Arabidopsis. *Plant Physiol.* 133: 1051–1071.

Rest, B., Danoun, S., Boudet, A.M. and Rochange, S.F. (2006) Down-regulation of cinnamoyl-CoA reductase in tomato (*Solanum lycopersicum* L.) induces dramatic changes in soluble phenolic pools. *J. Exp. Bot.* 57: 1399–1411.

Rubin, J.L. and Jensen, R. (1985) Differentially regulated isozymes of 3-deoxy-D-*arabino*- heptulosonate-7-phosphate synthase from seedlings of *Vigna radiata* [L.] Wilczek. *Plant Physiol.* 79: 711–718.

Saitou, N. and Nei, M. (1987) The neighbor–joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406–425.

Schneider, K., Hovel, K., Witzel, K., Hamberger, B., Schomburg, D., Kombrink, E. et al. (2003) The substrate specificity-determining amino acid code of 4-coumarate:CoA ligase. *Proc. Natl Acad. Sci. USA* 100: 8601–8606.

Sheen, J. (2002) A transient expression assay using *Arabidopsis* mesophyll protoplasts. http://genetics.mgh.harvard.edu/sheenweb/protocols/AtprotoRL04.pdf.

Shi, R., Sun, Y.H., Li, Q., Heber, S., Sederoff, R. and Chiang, V.L. (2010) Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the monolignol biosynthetic genes. *Plant Cell Physiol.* 51: 144–163.

Shukla, A.K., Shasany, A.K. and Khanuja, S.P.S. (2005) Isolation of poly(A)$^+$ mRNA for downstream reactions from some medicinal and aromatic plants. *Indian J. Exp. Biol.* 43: 197–201.

Silber, M.V., Meimberg, H. and Ebel, J. (2008) Identification of a 4-coumarate:CoA ligase gene family in the moss, *Physcomitrella patens*. *Phytochemistry* 69: 2449–2456.

Sitaramaraju, Y., Hul, A., Wolfs, K., Schepdael, A., Hoogmartens, J. and Adams, E. (2008) Static headspace gas chromatography of (semi-)volatile drugs in pharmaceuticals for topical use. *J. Pharmaceut. Biomed. Anal.* 47: 834–840.

Song, J. and Wang, Z. (2011) RNAi-mediated suppression of the phenylalanine ammonia-lyase gene in *Salvia miltiorrhiza* causes

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013

abnormal phenotypes and a reduction in rosmarinic acid biosynthesis. *J. Plant Res.* 124: 183–192.

Souza, CdeA, Barbazuk, B., Ralph, S.G., Bohlmann, J., Hamberger, B. and Douglas, C.J. (2008) Genome-wide analysis of a land plant-specific acyl:coenzyme A synthetase (ACS) gene family in *Arabidopsis*, poplar, rice and *Physcomitrella*. *New Phytol.* 179: 987–1003.

Stockigt, J. and Zenk, M.H. (1975) Chemical syntheses and properties of hydroxycinnamoyl-coenzyme A derivatives. *Z. Naturforsch.* 30: 352–358.

Stuible, H.P., Buttner, D., Ehlting, J., Hahlbrock, K. and Kombrink, E. (2000) Mutational analysis of 4-coumarate:CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate-forming enzymes. *FEBS Lett.* 467: 117–122.

Stuible, H.P. and Kombrink, E. (2001) Identification of the substrate specificity-conferring amino acid residues of 4-coumarate:coenzyme A ligase allows the rational design of mutant enzymes with new catalytic properties. *J. Biol. Chem.* 276: 26893–26897.

Thevenin, J., Pollet, B., Letarnec, B., Saulnier, L., Gissot, L., Grondard, A.M. et al. (2010) The simultaneous repression of CCR and CAD, two enzymes of the lignin biosynthetic pathway, results in sterility and dwarfism in *Arabidopsis thaliana*. *Mol. Plant* 4: 1–13.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876–4882.

Wagner, A., Donaldson, L., Kim, H., Phillips, L., Flint, H., Steward, D. et al. (2009) Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm *Pinus radiata*. *Plant Physiol.* 149: 370–383.

Webster, L.T. Jr., Mieyal, J.J. and Siddiqui, U.A. (1974) Benzoyl and hydroxybenzoyl esters of coenzyme A. *J. Biol. Chem.* 249: 2641–2645.

Wein, M., Lavid, N., Lunkenbein, S., Lewinsohn, E., Schwab, W. and Kaldenhoff, R. (2002) Isolation, cloning and expression of a multifunctional O-methyltransferase capable of forming 2,5-dimethyl-4-methoxy-3(2H)-furanone, one of the key aroma compounds in strawberry fruits. *Plant J.* 31: 755–765.

Whetten, R. and Sederoff, R. (1995) Lignin biosynthesis. *Plant Cell* 7: 1001–1013.

Xie, Z., Kapteyn, J. and Gang, D.R. (2008) A systems biology investigation of the MEP/terpenoid and shikimate/phenylpropanoid pathways points to multiple levels of metabolic control in sweet basil glandular trichomes. *Plant J.* 54: 349–361.

Yasmin, A. and Debener, T. (2010) Transient gene expression in rose petals *via Agrobacterium* infiltration. *Plant Cell Tiss. Organ. Cult.* 102: 245–250.

Zhong, R., Morrison, W.H., Himmelsbach, D.S., Poole, F.L. and Ye, Z.H. (2000) Essential role of caffeoyl coenzyme A *O*-methyltransferase in lignin biosynthesis in woody poplar plants. *Plant Physiol.* 124: 563–577.

Downloaded from http://pcp.oxfordjournals.org/ at Central Inst. Of Medicinal & Aromatic Plants (Cimap) on July 19, 2013