

Functional identification of rice *syn*-copalyl diphosphate synthase and its role in initiating biosynthesis of diterpenoid phytoalexin/allelopathic natural products

Meimei Xu¹, Matthew L. Hillwig¹, Sladjana Priscic¹, Robert M. Coates² and Reuben J. Peters^{1,*}

¹Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, IA 50011, USA, and

²Department of Chemistry, University of Illinois, Urbana, IL 61801, USA

Received 4 April 2004; revised 29 April 2004; accepted 12 May 2004.

*For correspondence (fax +1 515 294 0453; e-mail rjpeters@iastate.edu).

Summary

Rice produces a number of phytoalexins, and at least one allelopathic agent, from *syn*-copalyl diphosphate (CPP), representing the only known metabolic fate for this compound. Thus, the class II terpene synthase that converts the universal diterpenoid precursor geranylgeranyl diphosphate to *syn*-CPP catalyzes the committed step in biosynthesis of these natural products. Here the extensive sequence information available for rice was coupled to recombinant expression and functional analysis to identify *syn*-copalyl diphosphate synthase (OsCPS_{syn}). In addition, OsCPS_{syn} mRNA was found to be specifically induced in leaves by conditions that stimulate phytoalexin biosynthesis. Therefore, transcription of OsCPS_{syn} seems to be an important regulatory point for controlling the production of these defensive compounds. Finally, alignments carried out with OsCPS_{syn} revealed that class II terpene synthases exhibit a sequence conservation pattern substantially different from that of the prototypical class I enzymes. One particularly notable feature is the specific conservation of the functionally cryptic 'insertional' sequence element in class II terpene synthases, indicating that this region is important for the corresponding cyclization reaction.

Keywords: terpene synthase, labdane-related diterpenoids, phytoalexin, allelopathy, natural products biosynthesis, plant defense.

Introduction

Plants produce large numbers of low-molecular weight organic compounds, the majority of which do not appear to be necessary for growth under optimal conditions and are, thus, considered secondary metabolites (Croteau *et al.*, 2000). Nevertheless, many of these natural products have important ecological roles in plant defense and allelopathy. For example, phytoalexins are produced in response to microbial infection and exhibit potent antimicrobial activity (VanEtten *et al.*, 1994), while allelochemicals are secreted to the rhizosphere to suppress the growth of neighboring plants (Bais *et al.*, 2004). Terpenoids are often found serving in such roles, as these compounds are particularly abundant in plant secondary metabolism and, in fact, comprise the largest class of natural products (Croteau *et al.*, 2000). Notably, a substantial fraction of the known terpenoids can be classified as labdane-related diterpenoids (20 carbon). This is a large

group of over 5000 known compounds defined here as minimally containing the bicyclic hydrocarbon structure found in the labdane class of diterpenoids. In addition, this core structure can be further cyclized and/or rearranged, as in the related/derived structural classes (e.g. gibberellins, kauranes, abietanes, and (iso)pimaranes).

Biosynthesis of labdane-related natural products requires two cyclization steps, the first of which gives rise to the characteristic bicyclic backbone in forming labdadienyl/copalyl diphosphate (CPP). This initial cyclization is catalyzed by class II terpene synthases that selectively produce a specific stereoisomer of CPP from the universal diterpenoid precursor (*E,E,E*)-geranylgeranyl diphosphate (GGPP). CPP is often further elaborated by class I terpene synthases that stereoselectively utilize these bicyclic prenyl diphosphates to produce specific polycyclic compounds (e.g. Figure 1). Thus, class II and class I diterpene

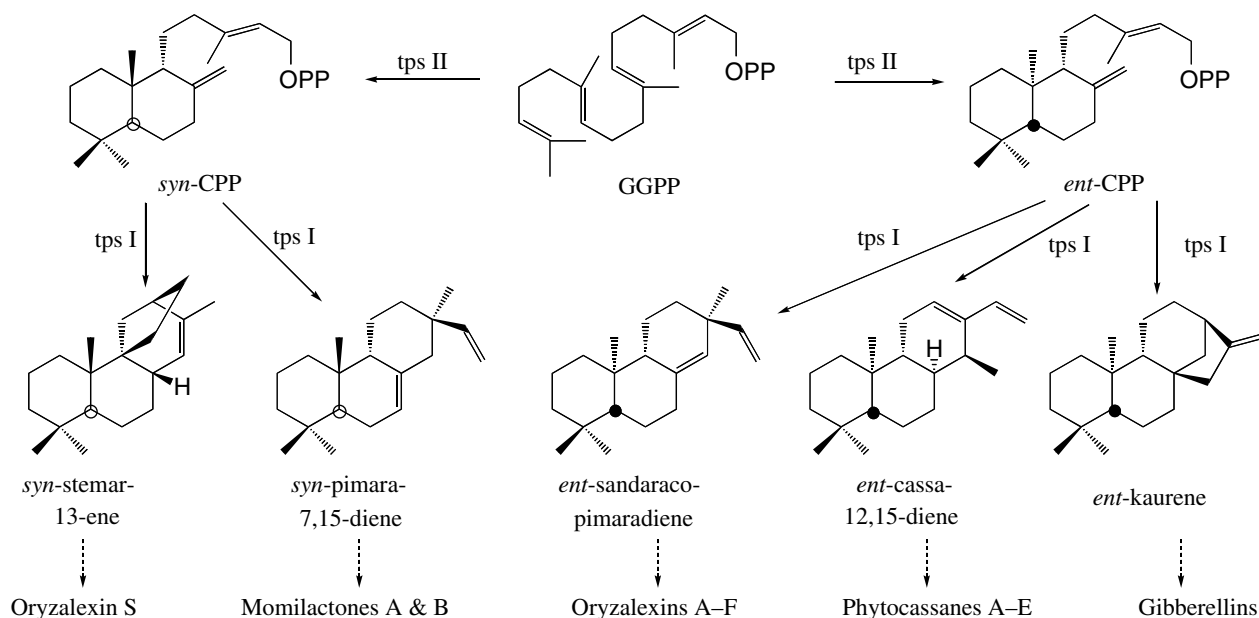


Figure 1. Known cyclization steps in labdane-related diterpenoid biosynthesis in rice.

Reactions catalyzed by class II (tpsII) or class I (tpsI) terpene synthases are indicated, along with the corresponding classes of natural products derived from each of the named polycyclic hydrocarbon structures (dashed arrows indicate multiple enzymatic steps).

synthases act sequentially to form labdane-related skeletal backbones.

The two classes of plant terpene synthases, while clearly related (Bohlmann *et al.*, 1998a) and both catalyzing reactions involving electrophilic cyclization/rearrangement, are biochemically distinct. Each utilizes different active sites and aspartate-rich motifs functionally associated with catalytic activity (Peters *et al.*, 2001). Class I enzymes contain a DDXXD metal binding motif that facilitates the metal ion assisted allylic diphosphate *ionization* reactions commonly associated with terpene synthases (Davis and Croteau, 2000). In contrast, the class II cyclases use a separately placed DXDD motif to mediate *protonation* of a carbon-carbon double bond in an acid/base catalyzed reaction (Peters and Croteau, 2002). When the first class II terpene synthase was cloned (Sun and Kamiya, 1994) the corresponding cyclization mechanism and motif was noted to closely resemble those found in squalene-hopene cyclases (Ochs *et al.*, 1992). This led to the common class II designation for the unrelated diterpene (terpene synthase family) and triterpene (lanosterol synthase family) cyclases (Wendt and Schulz, 1998).

Prototypical class I terpene synthases are similar in size and are composed of two structurally defined domains, which are largely divided between amino and carboxyl halves of their sequences and have simply been called the N- and C-terminal domains (Starks *et al.*, 1997; Whittington *et al.*, 2002). However, the class II enzymes invariably contain a large amount of additional N-terminal sequence (approximately 240 residues; Peters and Croteau, 2002),

which has been termed the 'insertional' element (Figure 2). Similar sequences can be found in a few class I enzymes as well, most notably those involved in cyclization of CPP to labdane-related structures (e.g. kaurene synthase; Yamaguchi *et al.*, 1996).

In addition to the ubiquitous gibberellin growth hormones (Phinney *et al.*, 1957), rice (*Oryza sativa*) produces a number of other labdane-related diterpenoids that serve as phytoalexins and/or allelochemicals (Figure 1). In particular, four other structurally distinct classes have been defined, momilactones A and B (Cartwright *et al.*, 1981; Kato *et al.*, 1973), oryzaalexins A-F (Akatsuka *et al.*, 1985; Kato *et al.*, 1993, 1994; Sekido *et al.*, 1986), oryzaalexin S (Kodama *et al.*, 1992), and phytocassanes A-E (Koga *et al.*, 1995, 1997). All of these natural products exhibit

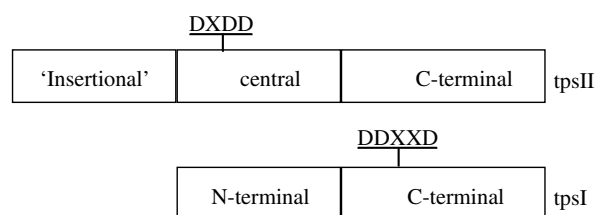


Figure 2. Domain schematic for prototypical class I (tpsI) and class II (tpsII) terpene synthases (tpsII), as modeled on the structures determined for the typical class I *epi*-aristolochene and bornyl diphosphate synthases (Starks *et al.*, 1997; Whittington *et al.*, 2002).

Approximate locations of the aspartate-rich motifs and the tpsII-associated 'insertional' element are as indicated. For clarity, the transit sequence required in mono- and di-terpene synthases for their plastidal location in planta is not shown.

antimicrobial properties and are produced in leaves in response to infection with the blast pathogenic fungus *Magnaportha grisea* and, thus, can be defined as phytoalexins (VanEtten *et al.*, 1994). In addition, momilactones A and B were originally identified as dormancy factors from rice seed husks (Kato *et al.*, 1973), and momilactone B has recently been shown to be constitutively secreted from the roots of rice seedlings (Kato-Noguchi and Ino, 2003), where it acts as an allelopathic agent (Kato-Noguchi *et al.*, 2002). The secretion of antimicrobial agents to the rhizosphere may also provide a competitive advantage for root establishment through local suppression of soil microorganisms (Bais *et al.*, 2004). Notably, rice leaves produce all of these compounds after UV-irradiation as well as blast fungal infection, providing a convenient method for inducing biosynthesis of these natural products and, presumably, transcription of the corresponding enzymatic machinery (Kodama *et al.*, 1988).

UV-irradiation has been found to induce terpene synthase activity producing polycyclic diterpene hydrocarbons corresponding to the putative precursors of momilactones A and B, oryzalexin S, and oryzalexins A–F [*syn*-pimara-7,15-diene, *syn*-stemar-13-ene, and *ent*-sandaracopimaradiene, respectively (Wickham and West, 1992)]. Enzymatic activities specifically converting *ent*- or *syn*-CPP to these identified polycyclic hydrocarbons were found in cell-free extracts from chitin-elicited suspension-cultured rice cells (Mohan *et al.*, 1996). A number of other diterpene hydrocarbons were also produced in this cell-free system, presumably one of these corresponds to *ent*-cassa-12,15-diene, the putative precursor to phytocassanes A–E (Yajima and Mori, 2000). Recently, synthesis of *ent*-cassa-12,15-diene (Yajima *et al.*, 2004) has enabled identification of this compound and cloning of the relevant class I terpene synthase from rice (Cho *et al.*, 2004). However, the conversion of GGPP to *syn*-CPP in rice has not been directly demonstrated. Nevertheless, in catalyzing the committed step in biosynthesis of *syn*-labdane-related diterpenoids, all of which, at a minimum, act as phytoalexins in rice, the corresponding class II terpene synthase would be expected to be a key regulatory target for control of this important metabolic process. Additionally, the corresponding enzyme represents a novel stereochemical outcome relative to the identified class II terpene synthases, as only enzymes producing normal- or *ent*-CPP are known. Therefore, we utilized a functional genomics approach to identify *syn*-copalyl diphosphate synthase (*OsCPS_{syn}*) from rice (*Oryza sativa* L. ssp. *indica*). Expression analysis was then used to demonstrate that transcription of *OsCPS_{syn}* is an important point at which the corresponding biosynthetic processes are controlled. We also report here insights from sequence comparisons of *OsCPS_{syn}* with other known terpene synthases and speculate about possible functional implication.

Results

Isolation of a class II terpene synthase cDNA from UV-irradiated rice leaves

Our interest in the functionally distinct class II terpene synthases as potentially significant targets for metabolic engineering and biochemical analysis motivated this work to identify the *syn*-CPP synthase that has been previously implicated in rice phytoalexin/allelopathic natural product biosynthesis. The extensive genomic (Goff *et al.*, 2002; Yu *et al.*, 2002) and full-length cDNA (Kikuchi *et al.*, 2003) sequence information available for rice was used to identify a number of putative class II terpene synthases, as defined by the presence of a DXDD motif and 'insertional' element. One of these was readily cloned from mRNA prepared from UV-irradiated rice and verified by complete sequencing. However, alignment of the corresponding amino acid sequence with other diterpene synthases indicated the absence of the usual substantial N-terminal transit peptide required for the plastidial location of diterpene synthases in plants (see Figure 5a). The presence of additional N-terminal coding sequence was verified by 5' rapid amplification of cDNA ends (RACE). The additional N-terminal coding sequence (46 aa) clearly forms at least part of the transit peptide, with an alkaline pI of 11.9, relative to the overall pI of 5.5, which is a common feature of plastid targeting sequences (Keegstra *et al.*, 1989). Thus, a full-length cDNA sequence was determined, which has been deposited into the various nucleotide sequence databases under accession number AY530101. The full-length open reading frame was then also cloned and verified as described above for the partial cDNA.

Functional expression and characterization demonstrated the specific production of syn-copalyl diphosphate

Full-length class II terpene synthase preprotein (tpsII) was expressed in *Escherichia coli* either alone or as a C-terminal fusion to glutathione-S-transferase (GST-tpsII). In addition, the partial cDNA sequence, which presumably resembles the mature native protein, was expressed alone as a potential pseudomature protein. Enzymatic assays with partially purified recombinant preparations were carried out with GGPP as substrate. After the reaction was allowed to proceed for 1 h at room temperature, the solution was extensively extracted with organic solvent to remove any relatively non-polar compounds (i.e. hydrocarbon or alcohol formed from GGPP). Phosphatase treatment was then employed to remove the pyrophosphate from GGPP or any enzymatically formed derivative, to enable straightforward extraction of the resulting alcohol into organic solvent. Enzymatic turnover of GGPP was analyzed by gas chromatography-mass spectrometry

(GC-MS) of the resulting extracts. No hydrocarbons or alcohols were produced by any construct. However, conversion of GGPP into an altered prenyl diphosphate structure was detected, although only with the pseudo-mature construct. The lack of activity with the full-length preprotein provides an extreme example of the previously observed deleterious effect of plant transit peptides on recombinant expression and folding (Williams *et al.*, 1998). Comparison of the enzymatically formed compound to similarly dephosphorylated authentic samples of *ent*- and *syn*-CPP demonstrated that the enzyme produces *syn*-CPP (Figure 3). Therefore, we have functionally identified rice *syn*-CPP synthase (OsCPS_{syn}).

Expression pattern of OsCPS_{syn} mRNA

The metabolic fate of *syn*-CPP is thought to be limited to phytoalexin/allelopathic natural products. Therefore, regulation of OsCPS_{syn} activity, which catalyzes the committed biosynthetic step, provides a logical point for controlling production of these natural products. Previous review of the relevant literature has been used to suggest that plant secondary metabolism is most often regulated at the level

of transcription (Peters and Croteau, 2004). Transcriptional control is manifested by upregulation of enzymatic mRNA levels, with subsequent phytochemical accumulation, in response to the appropriate environmental conditions. The convenience of UV-irradiation for induction of phytoalexin biosynthesis in rice leaves has long been appreciated (Cartwright *et al.*, 1977). In fact, this was utilized in the recently reported large-scale cDNA project from which the original sequence for OsCPS_{syn} is derived (Kikuchi *et al.*, 2003). Further, quantitative analysis of phytochemical accumulation for the detached leaf UV-irradiation induction method used here has been previously reported (Kodama *et al.*, 1988). Therefore, the ability of UV-irradiation to induce expression of OsCPS_{syn} mRNA in rice leaves was characterized, demonstrating transcriptional upregulation prior to phytoalexin accumulation (Figure 4). In addition, OsCPS_{syn} transcription is induced by methyl jasmonate, an important plant defense signaling molecule (Farmer and Ryan, 1990), which, as free jasmonic acid, has been previously demonstrated to induce phytoalexin biosynthesis in rice cell culture (Nojiri *et al.*, 1996). Also consistent with a role in defense, five of the six expressed sequence tags (EST) associated with OsCPS_{syn} in the TIGR Gene Index are from blast pathogen-infected rice EST projects (<http://www.tigr.org>; riceTC205530). Finally, as expected from its requisite role in constant production of an allelochemical (Kato-Noguchi and Ino, 2003), OsCPS_{syn}

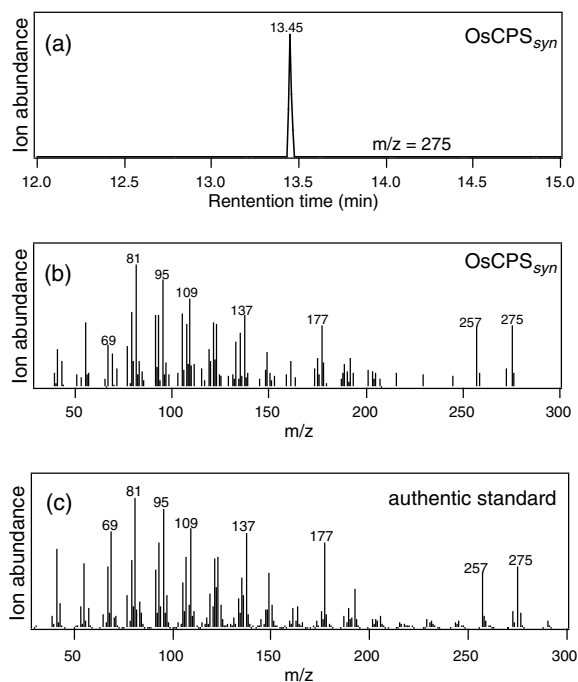


Figure 3. Production of *syn*-copalyl diphosphate (CPP).

(a) GC-MS analysis (275 *m/z* extracted ion chromatograph) of the dephosphorylated reaction products from GGPP. (b) Mass spectrum of GC-MS 275 *m/z* chromatograph peak (RT = 13.45 min). (c) Mass spectrum from authentic dephosphorylated *syn*-CPP (i.e. *syn*-copalol), which also exhibits RT = 13.45 min and is clearly separable from dephosphorylated *ent*-CPP (i.e. *ent*-copalol; RT = 13.61 min) and GGPP (i.e. geranylgeraniol, which is seen as a distinct peak in the total ion chromatograph; RT = 13.37 min).

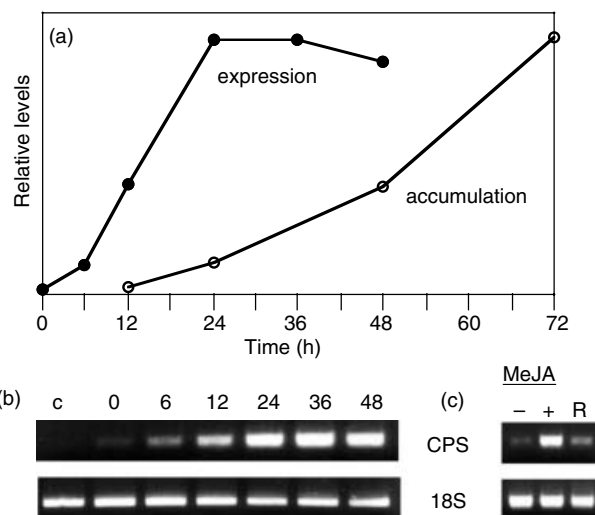


Figure 4. *syn*-Copalyl diphosphate synthase (OsCPS_{syn}) expression analysis. (a) Graphical comparison of OsCPS_{syn} mRNA levels (closed circles) and phytoalexin accumulation [open circles; as described by Kodama *et al.* (1988)], in UV-irradiated detached leaves. Quantitative RT-PCR analysis of OsCPS_{syn} mRNA expression levels is shown in (b) and (c). Specific bands corresponding to the 18S rRNA internal control and OsCPS_{syn} are indicated. (b) Expression in response to UV-irradiation. Time (h) after exposure is indicated (c = control leaves after approximately 18 h). (c) Expression in untreated 4-week-old plant roots (R), or in germinated seedlings in response to application of 0.5 mM methyl jasmonate (+MeJA) or water control (−MeJA).

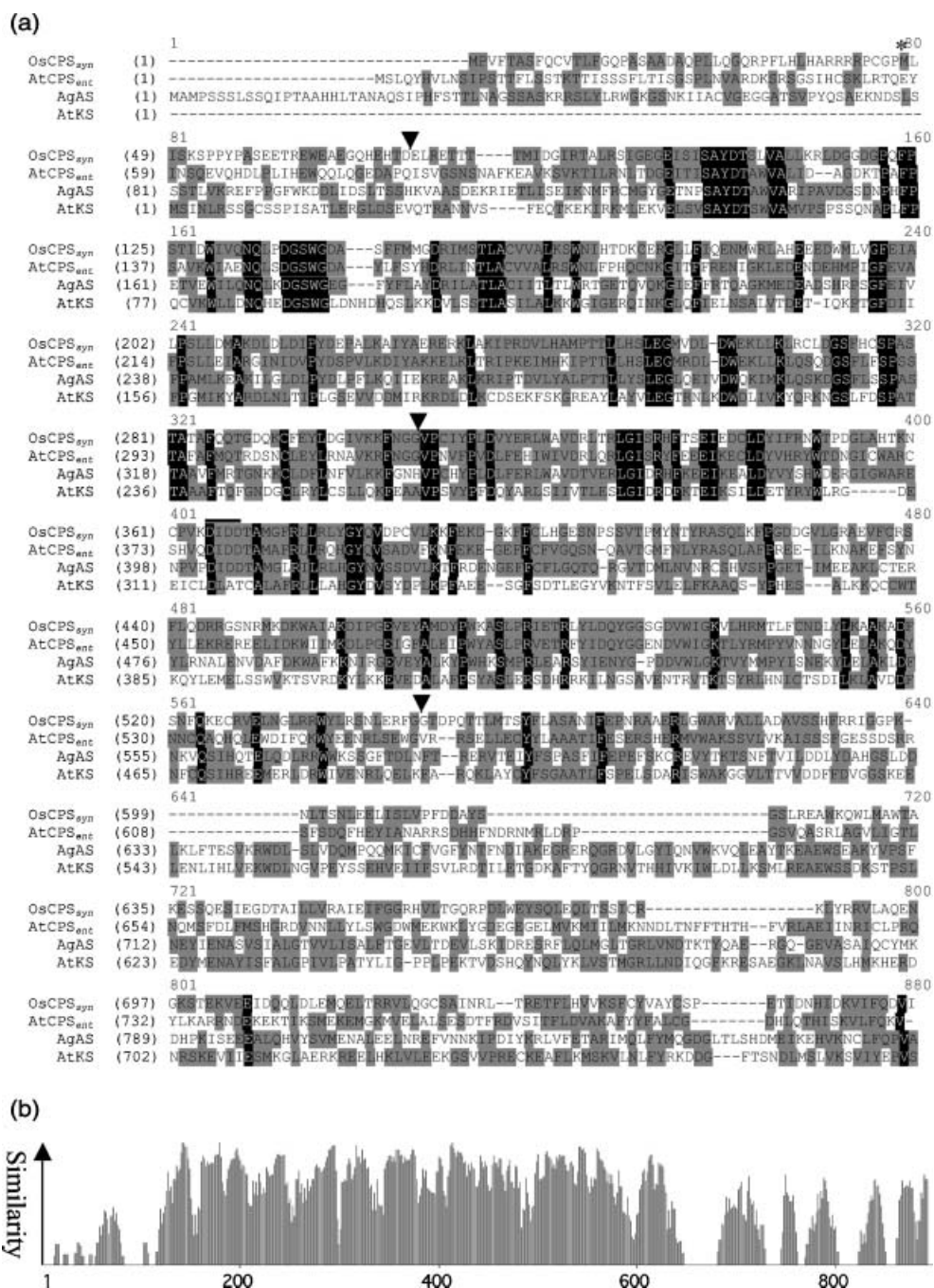


Figure 5. Sequence comparison of *OsCPS_{syn}* with other terpene synthases.

(a) Alignment with a representative class II terpene synthase (*AtCPS_{ant}*) and 'insertional' element containing class I terpene synthase, kaurene synthase (*AtKS*), both from *Arabidopsis*, as well as bifunctional class II/I cyclase, abietadiene synthase from grand fir (*AgAS*). The originally predicted start codon is marked with an '*', the boundaries used for domain comparisons are marked with arrowheads, and the requisite class II DXDD motif marked with an overhead line.

(b) Plot of sequence similarity over the entire sequence length for all the known terpene synthases exhibiting class II activity (numbering based on the consensus sequence).

mRNA seems to be constitutively present in roots. These results strongly indicate that biosynthesis of the corresponding syn-labdane-related diterpenoid natural products is controlled, at least in part, by transcriptional regulation of *OsCPS_{syn}*.

Sequence comparison of *OsCPS_{syn}* with other class II terpene synthases

As selected for by the identification strategy, *OsCPS_{syn}* contains not only the requisite DXDD motif, but also

the class II synthase associated 'insertional' element (Figure 2). In addition, *OsCPS_{syn}* clearly does not contain the C-terminal domain DDXXD metal ion binding motif required for class I catalysis (Davis and Croteau, 2000), explaining its inability to catalyze the corresponding diphosphate ionization reaction. Amino acid alignment of *OsCPS_{syn}* with the other known class II terpene synthases from plants demonstrates the expected homology (Figure 5). Specifically, *OsCPS_{syn}* is 38–44% identical with *ent*-copalyl diphosphate synthases (*CPS_{ent}*) (required for gibberellin plant growth hormone biosynthesis) from various angiosperms (Ait-Ali *et al.*, 1997; Bensen *et al.*, 1995; Rebers *et al.*, 1999; Richman *et al.*, 1999; Smith *et al.*, 1998; Sun and Kamiya, 1994) and approximately 28% identical with the known gymnosperm bifunctional class II/I terpene synthases (Schepmann *et al.*, 2001; Stofer Vogel *et al.*, 1996), which produce normal CPP. Significantly, this homology is not evenly distributed throughout the entire sequence, as can be seen in Figure 5(b), where similarity is plotted over the length of these proteins. On the basis of previous modeling and truncation studies of abietadiene synthase (AgAS) from *Abies grandis* (grand fir), approximate boundaries for the 'insertional' element (AgAS aa107–344), central region (AgAS aa345–582), and C-terminal domain (AgAS aa583–868) have been assigned (Peters *et al.*, 2003). Given these definitions it can be readily appreciated that the class II enzymes are most similar in the area of the 'insertional' element and central region, and least similar in the C-terminal domain. In particular, the 'insertional' element and central region are $\geq 40\%$ identical, while the C-terminal domain is much less conserved ($< 20\%$ identity in some cases), in terpene synthases with class II activity. This contrasts to the situation with class I terpene synthases, where those few that do contain an 'insertional' element show lower levels of conservation across this portion of their sequences, exhibiting identities as low as 25%.

Discussion

Previous cell-free biosynthetic tracer studies have implied that rice produces *syn*-CPP as the first committed intermediate in the production of *syn*-labdane-related diterpenoids (Mohan *et al.*, 1996). These compounds all act as phytoalexins or allelopathic agents. Therefore, the corresponding class II terpene synthase is expected to be a key regulatory point in biosynthesis of these natural products. We are also interested in the class II terpene synthases as model enzymatic systems for the analysis of electrophilic cyclization/rearrangement reactions, and none of the previously identified class II enzymes results in this specific stereochemical outcome. This is of particular biochemical interest because, whereas the bicyclization to normal- and *ent*-CPP catalyzed by the known class II terpene synthase proceeds from a

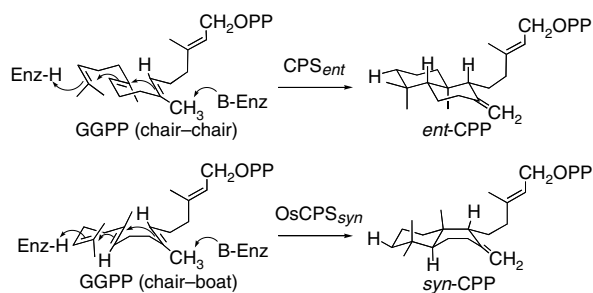


Figure 6. Illustration of the chair-chair and chair-boat conformations required of GGPP to undergo cyclization to *ent*- and *syn*-CPP, respectively.

chair-chair conformation, GGPP must be folded in the energetically less stable chair-boat conformation to form *syn*-CPP (Figure 6). In addition, the extensive sequence information available for rice was expected to assist functional identification of the gene for this putative enzymatic activity.

A predicted class II terpene synthase was found in the rice large-scale full-length cDNA project (Kikuchi *et al.*, 2003). Corresponding sequence was readily amplified from rice leaves that had been UV-irradiated, which has been shown to induce phytoalexin biosynthesis (Kodama *et al.*, 1988). This putatively full-length cDNA was cloned, and then extended by 5' RACE. Expression and biochemical analysis of a recombinant pseudomature construct demonstrated the production of *syn*-CPP from GGPP (Figure 3) and we have designated the gene *syn*-CPP synthase (*OsCPS_{syn}*). In addition, mRNA expression analysis (Figure 4) demonstrated that *OsCPS_{syn}* is constitutively transcribed in roots, but is only transcribed in leaves under conditions that stimulate phytoalexin production (i.e. UV-irradiation or exposure to methyl jasmonate). This matches the pattern of biosynthesis of the known *syn*-labdane-related diterpenoid natural products. Thus, transcription of *OsCPS_{syn}* seems to be required for phytochemical production, strongly indicating that the transcriptional control of *OsCPS_{syn}* represents an important regulatory point in these important metabolic processes.

Putative class II terpene synthases were initially identified by the presence of the catalytically requisite DXDD motif, as well as the functionally cryptic N-terminal 'insertional' element that has been associated with class II enzymes. Thus, *OsCPS_{syn}* contains both, with the DXDD motif then falling within a central region that is homologous to the N-terminal domain of typical class I terpene synthases (Figure 2). Alignment of *OsCPS_{syn}* with the other known plant class II terpene synthases reveals striking conservation of both the central region and 'insertional' element (Figure 5b). In contrast, the C-terminal domain, which contains the class I active site in the corresponding synthases (Starks *et al.*, 1997), is relatively divergent. As the class II central region contains the DXDD motif and other residues

important for class II activity (Peters and Croteau, 2002), its conservation is not surprising. However, the role of the 'insertional' element is not well defined. Modeling and mutational analysis of AgAS was used to suggest a role for this region in catalysis (Peters and Croteau, 2002). Unfortunately, truncation analysis was only able to demonstrate that the 'insertional' element is required for correct folding (Peters *et al.*, 2003). In particular, it was not possible to separate the class II and I activities into separate, or even overlapping, polypeptides, as was demonstrated with a, at best, very distantly related bifunctional class II/I fungal *ent*-kaurene diterpene synthase (Kawaide *et al.*, 2000). Nevertheless, recognition of the highly conserved nature of the 'insertional' region, specifically in class II terpene synthases, strengthens the suggestion that this sequence element serves a functional role in class II catalysis.

In conclusion, we have used a functional genomics approach to identify *syn*-CPP synthase (OsCPS_{syn}), a previously putative enzymatic activity, from rice. Notably, production of *syn*-CPP represents a novel stereochemical outcome relative to the previously known class II terpene synthases. In addition, OsCPS_{syn} initiates production of *syn*-labdane-related diterpenoids and its transcription seems to be an important control point for production of these natural products. Specifically, OsCPS_{syn} mRNA expression is induced in leaves by conditions that stimulate phytoalexin biosynthesis, but is constitutive in roots where it is required for the constant production of at least one allelochemical. Finally, alignment of OsCPS_{syn} with other known terpene synthases revealed that the class II enzymes exhibit a sequence conservation pattern substantially different from that of the class I synthases. In particular, the selective conservation of the 'insertional' element in class II terpene synthases seems to indicate a functional role in catalysis for this region.

Experimental procedures

Chemicals

The preparation of GGPP, and *ent*- and *syn*-CPP, from the corresponding copalol stereoisomers (Yee and Coates, 1992), has been previously described (Mohan *et al.*, 1996). Unless otherwise noted, all other chemicals were purchased from Fisher Scientific (Fairlawn, NJ, USA).

Plant material

Rice plants (*Oryza sativa* L. ssp. *indica* cv. IR24) were cultivated from seed under standard greenhouse conditions. After 4 weeks the plants were harvested for analysis. Detached leaves were UV-irradiated for 25 min using a Spectroline UV lamp (set to emit at 254 nm wavelength) suspended 15 cm above the leaf surfaces, then incubated 18 h at 30°C with high humidity in the dark. These tissues were frozen by liquid N₂ and total RNA extracted using Concert Plant Reagent, following the manufacturers instructions (Invitrogen,

Carlsbad, CA, USA), and stored at -80°C. Where indicated, purification of mRNA was performed with Dynabeads Oligo(dT)₂₅, again following the recommended procedure (DynaL Biotech, Oslo, Norway) and also stored at -80°C.

Isolation of cDNA

Three putative class II terpene synthases were identified by BLAST queries of the rice genomic and cDNA databases with the nucleotide sequence of *ent*-CPP synthase from *Arabidopsis* (Sun and Kamiya, 1994), with manual screening for the presence of the DXDD motif and 'insertional' element. RT-PCR reactions were performed to generate 5' fragments of the corresponding open reading frames, which exhibit limited homology to each other in this region (<55% nucleotide identity). However, for only one of these genes was it possible to amplify this region (efforts are underway to clone the remaining two sequences using differentiated internal sequences). The obtained 5' terminal fragment (599 bp) was cloned into pCR-Blunt II-TOPO (Invitrogen) and sequenced to verify its identity. To amplify the corresponding (partial) cDNA it was necessary to perform a 3' RACE reaction, which was done following the recommended protocol for the utilized GeneRacer kit (Invitrogen) and using the forward primer from the original 5' fragment isolation. This was also cloned into pCR-Blunt II-TOPO and verified by sequencing. The corresponding (partial) open reading frame was then subcloned, by PCR amplification for directional topoisomerization, into pENTR/SD/D-TOPO (Invitrogen). This was again verified by complete sequencing, demonstrating approximately 99% identity to the expected coding sequence (nucleotide sequence database accession number AK100631). Presumably the few observed differences reflect the intersubspecies variation between the ssp. *japonica* used by Kikuchi *et al.* (2003) and the ssp. *indica* used here. The true full-length cDNA was then identified through a 5' RACE reaction, using the reverse primer from the original 5' fragment isolation. Identical sequence was found in the 599 bp overlap between the originally cloned partial cDNA and the 5' RACE product. The full-length open reading frame was simply defined as running from the first occurrence of ATG to an in-frame stop codon (a total of 2304 bp). 3' RACE was again used to clone the open reading frame and 3' untranslated region into pCR-Blunt II-TOPO. The full-length open reading frame was then also subcloned into pENTR/SD/D-TOPO and the resulting construct verified by complete sequencing. This was then transferred, by directional recombination into the pDEST14 and pDEST15 expression vectors, while the partial cDNA was transferred into pDEST14 only (Gateway Technology, Invitrogen).

Recombinant expression and preparation

Expression was carried out with the BL21-derived C41 strain (Miroux and Walker, 1996), using freshly transformed cells. Initial 1 ml NZY (1% NZ amine, 1% NaCl, 0.5% yeast extract, 0.2% MgSO₄) cultures, with 50 µg ml⁻¹ ampicillin, were inoculated from three to five colonies to average over previously observed variations in expression level from single colony inoculation. The initial cultures were grown overnight (200 rpm at 37°C) to saturation and 0.5 ml used to inoculate 50 ml NZY cultures, with 50 µg ml⁻¹ ampicillin. These expression cultures were grown to A₆₀₀ approximately 0.6, transferred to 16°C (200 rpm) for 1–2 h prior to induction with 1 mM IPTG, then incubated a further 16–20 h. The cells were harvested by centrifugation (4000 g, 20 min, 4°C), resuspended in 1 ml of lysis buffer (50 mM Bis-Tris, pH 6.8, 150 mM KCl, 10 mM MgCl₂, 1 mM DTT) and lysed by mild sonication on ice (Branson sonifier 450: 10 sec, continuous, output, setting 5). The resulting lysates were

cleared by centrifugation (15 000 *g*, 30 min, 4°C) and filtration (0.45 µ) to yield soluble extracts. The recombinant protein was then partially purified as previously described (Peters *et al.*, 2003). Briefly, the soluble extract was adsorbed to 0.2 ml ceramic hydroxyapatite type II beads (Bio-Rad, Hercules, CA, USA), which was washed four times with 1 ml buffer (50 mM Bis-Tris, pH 6.8) to remove unbound material. Recombinant terpene synthase was then eluted with 0.5 ml of 200 mM sodium phosphate (pH 6.8), glycerol added to 10% (v/v), DTT to 5 mM, and the resulting solution spin filtered (0.2 µ).

Enzymatic assay and analysis

Enzymatic analysis was carried out in assay buffer (50 mM HEPES, pH 7.2, 100 mM KCl, 7.5 mM MgCl₂, 5% glycerol, and 5 mM DTT), with 25 µl of recombinant soluble extract, and initiated by the addition of GGPP to 200 µM for a final volume of 0.2 ml. Reactions were incubated for 1 h in the dark at room temperature and then extracted three times with 0.5 ml of hexanes. Residual organic solvent was removed by evaporation under N₂ to enable enzymatic dephosphorylation by calf intestinal phosphatase (10 U; New England Biolabs, Beverly, MA, USA), which was allowed to proceed approximately 4 h at 37°C, largely as described (Wise *et al.*, 2001). Similar dephosphorylation was carried out with authentic *ent*- and *syn*-CPP. The dephosphorylated compounds were then extracted two times with 0.5 ml of hexanes. All extracts were completely dried under N₂ and then resuspended in 100 µl of hexanes for GC-MS analysis. GC-MS was performed using an HP-1 column on an Agilent (Palo Alto, CA, USA) 6890N GC instrument with 5973N mass selective detector. Samples (5 µl) were injected at 40°C in the splitless mode. After a 3 min isothermal hold at 40°C, the temperature was increased at 20°C min⁻¹ to 300°C, where it was held for an additional 4 min. MS data from 40 to 500 *m/z* were collected during the temperature ramp and high isothermal hold.

Expression analysis

OscPS_{syn} mRNA transcription in response to UV-irradiation was measured using *ssp. japonica* rice, also grown from seed for 4 weeks under standard greenhouse conditions. Roots were directly harvested, while detached leaves either underwent the UV-irradiation treatment described above (Plant material), or the same procedure without exposure to UV-irradiation (control leaves). At the indicated times leaves were frozen in liquid N₂ and stored at -80°C prior to RNA extraction. Transcription in response to methyl jasmonate was measured in *ssp. indica* seedlings. Seeds were sterilized for 2–3 h in 10% bleach, rinsed five times in sterile water, then germinated in Petri dishes on three layers of filter paper moistened with 3 ml sterile water (this was maintained by the addition of 1 ml of sterile water every other day) at 30°C in the dark for a week. Methyl jasmonate treatment was based on previous reports (Farmer and Ryan, 1990; Martin *et al.*, 2002). Briefly, one plate (seven seedlings) was then sprayed with 2 ml 0.5 mM (±)-methyl jasmonate (Bedoukian Research, Danbury, CT, USA) in carrier solution (0.1% Tween 20 in sterile deionized water), while another (control plate) was only treated with 2 ml of the carrier. These plates were then incubated in separate air-tight containers for 48 h, also at 30°C in the dark, prior to RNA extraction from the whole seedlings. Quantitative RT-PCR was performed on 0.5 µg total RNA (isolated as described above) from each of the indicated tissues, using OscPS_{syn} specific primers (from the 5' fragment isolation) or QuantumRNA 18S standard primers (Ambion, Austin, TX, USA). Appropriate preliminary studies were performed to ensure that the final amplification conditions were in the linear response range.

RT-PCR reaction products were separated on a 2.5% agarose gel, stained with ethidium bromide, visualized over a UV-light source, and quantified (where indicated and with normalization against the 18S standard) with NIH Image software. Notably, our preliminary results with the other two putative rice class II terpene synthases indicate that these each exhibit a different expression pattern than OscPS_{syn} (i.e. one is non-inducible while the other is inducible but not expressed in roots).

Sequence analysis/alignments

Sequence alignments and identity/similarity comparisons were performed with the AlignX program in the Vector NTI software package (Invitrogen), using standard parameters. OscPS_{syn} was the reference sequence in all cases. CPS_{ent} sequences were those from *Arabidopsis thaliana* (Sun and Kamiya, 1994), *Zea mays* (Bensen *et al.*, 1995), *Pisum sativum* (Ait-Ali *et al.*, 1997), *Cucurbita maxima* (Smith *et al.*, 1998), *Lycopersicon esculentum* (Rebers *et al.*, 1999), and *Stevia rebaudiana* (Richman *et al.*, 1999). Bifunctional class II/I sequences for abietane synthases were those from *Abies grandis* (Stofer Vogel *et al.*, 1996) and *Ginkgo biloba* (Schepmann *et al.*, 2001). Kaurene synthase sequences were those from *Cucurbita maxima* (Yamaguchi *et al.*, 1996), *Arabidopsis thaliana* (Yamaguchi *et al.*, 1998), and *Stevia rebaudiana* (Richman *et al.*, 1999). Other strictly class I sequences included in sequence comparisons were specifically those containing 'insertional' elements; linalool synthase from *Clarkia breweri* (Dudareva *et al.*, 1996), taxadiene synthase from *Taxus brevifolia* (Wildung and Croteau, 1996), and bisabolene synthase from *Abies grandis* (Bohlmann *et al.*, 1998b). Domain level comparisons were carried out by manually trimming the relevant sequences based on whole sequence alignment with abietadiene synthase from grand fir (e.g. Figure 5), for which approximate domain boundaries have been defined (Peters *et al.*, 2003).

Acknowledgements

The authors thank Dr Adam Bogdanove and his laboratory for supplying rice plants and seeds, Dr Robert Thornburg for methyl jasmonate, Dr Nathan K. N. Yee for synthesis and Yinghua Jin for assistance with the *syn*- and *ent*-CPP samples used in this study, and Dana J. Morrone and P. Ross Wilderman for critical reading of the manuscript. This work was generously supported by funds from Iowa State University and grant no. 03-190 from the Roy J. Carver Charitable Trust.

References

- Ait-Ali, T., Swain, S.M., Reid, J.B., Sun, T. and Kamiya, Y. (1997) The LS locus of pea encodes the gibberellin biosynthesis enzyme *ent*-kaurene synthase A. *Plant J.* **11**, 443–454.
- Akatsuka, T., Kodama, O., Sekido, H., Kono, Y. and Takeuchi, S. (1985) Novel phytoalexins (oryzaalexins A, B, and C) isolated from rice blast leaves infected with *Piricularia oryzae*. Part I: Isolation, characterization and biological activities of oryzaalexins. *Agric. Biol. Chem.* **49**, 1689–1694.
- Bais, H.P., Park, S.W., Weir, T.L., Callaway, R.M., Vivanco, J.M. (2004) How plants communicate using the underground information superhighway. *Trends Plant Sci.* **9**, 26–32.
- Bensen, R.J., Johal, G.S., Crane, V.C., Tossberg, J.T., Schnable, P.S., Meeley, R.B. and Briggs, S.P. (1995) Cloning and characterization of the maize An1 gene. *Plant Cell*, **7**, 75–84.

- Bohlmann, J., Meyer-Gauen, G. and Croteau, R. (1998a) Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl Acad. Sci. USA*, **95**, 4126–4133.
- Bohlmann, J., Crock, J., Jetter, R. and Croteau, R.B. (1998b) Terpeneoid-based defenses in conifers: cDNA cloning, characterization, and functional expression of wound-inducible (E)-alpha-bisabolene synthase from grand fir (*Abies grandis*). *Proc. Natl Acad. Sci. USA*, **95**, 6756–6761.
- Cartwright, D.W., Langcake, P., Pryce, R.J., Leworthy, D.P. and Ride, J.P. (1977) Chemical activation of host defence mechanisms as a basis for crop protection. *Nature*, **267**, 511–513.
- Cartwright, D.W., Langcake, P., Pryce, R.J., Leworthy, D.P. and Ride, J.P. (1981) Isolation and characterization of two phytoalexins from rice as momilactones A and B. *Phytochemistry*, **20**, 535–537.
- Cho, E.-M., Okada, A., Kenmoku, H. *et al.* (2004) Molecular cloning and characterization of a cDNA encoding *ent*-cassa-12,15-diene synthase, a putative diterpenoid phytoalexin biosynthetic enzyme, from suspension-cultured rice cells treated with a chitin elicitor. *Plant J.* **37**, 1–8.
- Croteau, R., Kutchan, T.M. and Lewis, N.G. (2000) Natural Products (Secondary Metabolites). In *Biochemistry and Molecular Biology of Plants* (Buchanan, B., Gruissem, W. and Jones, R., eds). Rockville, MD, USA: American Society of Plant Biologists, pp. 1250–1318.
- Davis, E.M. and Croteau, R. (2000) Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. *Top. Curr. Chem.* **209**, 53–95.
- Dudareva, N., Cseke, L., Blanc, V.M. and Pichersky, E. (1996) Evolution of floral scent in *Clarkia*: novel patterns of S-linalool synthase gene expression in the *C. breweri* flower. *Plant Cell*, **8**, 1137–1148.
- Farmer, E.E. and Ryan, C.A. (1990) Interplant communication: airborne methyl jasmonate induces synthesis of proteinase inhibitors in plant leaves. *Proc. Natl Acad. Sci. USA*, **87**, 7713–7716.
- Goff, S.A., Ricke, D., Lan, T.H. *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science*, **296**, 92–100.
- Kato, T., Kabuto, C., Sasaki, N., Tsunagawa, M., Aizawa, H., Fujita, K., Kato, Y. and Kitahara, Y. (1973) Momilactones, growth inhibitors from rice, *Oryza sativa* L. *Tetrahedron Lett.* **39**, 3861–3864.
- Kato, H., Kodama, O. and Akatsuka, T. (1993) Oryzaalexin E, a diterpene phytoalexin from UV-irradiated rice leaves. *Phytochemistry*, **33**, 79–81.
- Kato, H., Kodama, O. and Akatsuka, T. (1994) Oryzaalexin F, a diterpene phytoalexin from UV-irradiated rice leaves. *Phytochemistry*, **36**, 299–301.
- Kato-Noguchi, H. and Ino, T. (2003) Rice seedlings release momilactone B into the environment. *Phytochemistry*, **63**, 551–554.
- Kato-Noguchi, H., Ino, T., Sata, N. and Yamamura, S. (2002) Isolation and identification of a potent allelopathic substance in rice root exudates. *Physiologia Plantarum*, **115**, 401–405.
- Kawaide, H., Sassa, T. and Kamiya, Y. (2000) Functional analysis of the two interacting cyclase domains in *ent*-kaurene synthase from the fungus *Phaeosphaeria* sp. L487 and a comparison with cyclases from higher plants. *J. Biol. Chem.* **275**, 2276–2280.
- Keegstra, K., Olsen, L.J. and Theg, S.M. (1989) Chloroplastic precursors and their transport across the envelope membranes. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 471–501.
- Kikuchi, S., Satoh, K., Nagata, T. *et al.* (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from *japonica* rice. *Science*, **301**, 376–379.
- Kodama, O., Suzuki, T., Miyakawa, J. and Akatsuka, T. (1988) Ultraviolet-induced accumulation of phytoalexins in rice leaves. *Agric. Biol. Chem.* **52**, 2469–2473.
- Kodama, O., Li, W.X., Tamogami, S. and Akatsuka, T. (1992) Oryzaalexin-S, a novel stemarane-type diterpene rice phytoalexin. *Biosci. Biotechnol. Biochem.* **56**, 1002–1003.
- Koga, J., Shimura, M., Oshima, K., Ogawa, N., Yamauchi, T. and Ogasawara, N. (1995) Phytocassanes A, B, C, and D, novel diterpene phytoalexins from rice, *Oryza sativa* L. *Tetrahedron*, **51**, 7907–7918.
- Koga, J., Ogawa, N., Yamauchi, T., Kikuchi, N., Ogasawara, N. and Shimura, M. (1997) Functional moiety for the antifungal activity of phytocassane E, a diterpene phytoalexin from rice. *Phytochemistry*, **44**, 249–253.
- Martin, D., Tholl, D., Gershenzon, J. and Bohlmann, J. (2002) Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Plant Physiol.* **129**, 1003–1018.
- Miroux, B. and Walker, J.E. (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289–298.
- Mohan, R.S., Yee, N.K.N., Coates, R.M., Ren, Y.Y., Stamenkovic, P., Mendez, I. and West, C.A. (1996) Biosynthesis of cyclic diterpene hydrocarbons in rice cell suspensions: conversion of 9,10-syn-labda-8(17),13-dienyl diphosphate to 9 β -pimar-7,15-diene and stemar-13-ene. *Arch. Biochem. Biophys.* **330**, 33–47.
- Nojiri, H., Sugimura, M., Yamane, H., Nishimura, Y., Yamada, A., Shibuya, N., Kodama, O., Murofushi, N. and Omori, T. (1996) Involvement of jasmonic acid in elicitor-induced phytoalexin production in suspension-cultured rice cells. *Plant Physiol.* **110**, 387–392.
- Ochs, D., Kaletta, C., Entian, K.-D., Beck-Sickinger, A. and Poralla, K. (1992) Cloning, expression, and sequencing of squalene-hopene cyclase, a key enzyme in triterpenoid biosynthesis. *J. Bact.* **174**, 298–302.
- Peters, R.J. and Croteau, R.B. (2002) Abietadiene synthase catalysis: conserved residues involved in protonation-initiated cyclization of geranylgeranyl diphosphate to (+)-copalyl diphosphate. *Biochemistry*, **41**, 1836–1842.
- Peters, R.J. and Croteau, R.B. (2004) Metabolic Engineering of Plant Secondary Metabolism. In *Handbook of Plant Biotechnology: Applications of Plant Biotechnology in Agriculture, the Pharmaceutical Industry, and Other Industries*, Vol. 2 (Kishore, G., ed.). London: John Wiley & Sons Ltd, pp. 609–628.
- Peters, R.J., Ravn, M.M., Coates, R.M. and Croteau, R.B. (2001) Bifunctional abietadiene synthase: free diffusive transfer of the (+)-copalyl diphosphate intermediate between two distinct active sites. *J. Am. Chem. Soc.* **123**, 8974–8978.
- Peters, R.J., Carter, O.A., Zhang, Y., Matthews, B.W. and Croteau, R.B. (2003) Bifunctional abietadiene synthase: mutual structural dependence of the active sites for protonation-initiated and ionization-initiated cyclizations. *Biochemistry*, **42**, 2700–2707.
- Phinney, B.O., West, C.A., Ritzel, M. and Neely, P.M. (1957) Evidence for gibberellin-like substances from flowering plants. *Proc. Natl Acad. Sci. USA*, **43**, 398–403.
- Rebers, M., Kaneta, T., Kawaide, H., Yamaguchi, S., Yang, Y.Y., Imai, R., Sekimoto, H. and Kamiya, Y. (1999) Regulation of gibberellin biosynthesis genes during flower and early fruit development of tomato. *Plant J.* **17**, 241–250.
- Richman, A.S., Gijzen, M., Starratt, A.N., Yang, Z. and Brandle, J.E. (1999) Diterpene synthesis in *Stevia rebaudiana*: recruitment and up-regulation of key enzymes from the gibberellin biosynthetic pathway. *Plant J.* **19**, 411–421.

- Schepmann, H.G., Pang, J. and Matsuda, S.P. (2001) Cloning and characterization of *Ginkgo biloba* levopimaradiene synthase, which catalyzes the first committed step in ginkgolide biosynthesis. *Arch. Biochem. Biophys.* **392**, 263–269.
- Sekido, H., Endo, T., Suga, R., Kodama, O., Akatsuka, T., Kono, Y. and Takeuchi, S. (1986) Oryzaalexin D (3,7-dihydroxy-(+)-sandaracopimaradiene), a new phytoalexin isolated from blast-infected rice leaves. *J. Pesticide Sci.* **11**, 369–372.
- Smith, M.S., Yamaguchi, S., Ait-Ali, T. and Kamiya, Y. (1998) The first step of gibberellin biosynthesis in pumpkin is catalyzed by at least two copalyl diphosphate synthases encoded by differentially regulated genes. *Plant Physiol.* **118**, 1411–1419.
- Starks, C.M., Back, K., Chappell, J. and Noel, J.P. (1997) Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science*, **277**, 1815–1820.
- Stofer Vogel, B., Wildung, M.R., Vogel, G. and Croteau, R. (1996) Abietadiene synthase from grand fir (*Abies grandis*). *J. Biol. Chem.* **271**, 23262–23268.
- Sun, T-P and Kamiya, Y. (1994) The Arabidopsis GA1 locus encodes the cyclase *ent*-kaurene synthetase A of gibberellin biosynthesis. *Plant Cell*, **6**, 1509–1518.
- VanEtten, H.D., Mansfield, J.W., Bailey, J.A. and Farmer, E.E. (1994) Two classes of plant antibiotics: phytoalexins versus 'phytoanticipins'. *Plant Cell*, **6**, 1191–1192.
- Wendt, K.U. and Schulz, G.E. (1998) Isoprenoid biosynthesis: manifold chemistry catalyzed by similar enzymes. *Structure*, **6**, 127–133.
- Whittington, D.A., Wise, M.L., Urbansky, M., Coates, R.M., Croteau, R.B. and Christianson, D.W. (2002) Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc. Natl Acad. Sci. USA*, **99**, 15375–15380.
- Wickham, K.A. and West, C.A. (1992) Biosynthesis of rice phytoalexins: identification of putative diterpene hydrocarbon precursors. *Arch. Biochem. Biophys.* **293**, 320–332.
- Wildung, M.R. and Croteau, R.B. (1996) A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis. *J. Biol. Chem.* **271**, 9201–9204.
- Williams, D.C., McGarvey, D.J., Katahira, E.J. and Croteau, R. (1998) Truncation of limonene synthase preprotein provides a fully active 'pseudomature' form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. *Biochemistry*, **37**, 12213–12220.
- Wise, M.L., Pyun, H-J, Helms, G., Assink, B., Coates, R.M. and Croteau, R.B. (2001) Stereochemical disposition of the geminal dimethyl groups in the enzymatic cyclization of geranyl diphosphate to (+)-bornyl diphosphate by recombinant (+)-bornyl diphosphate synthase from *Salvia officinalis*. *Tetrahedron*, **57**, 5327–5334.
- Yajima, A. and Mori, K. (2000) Absolute configuration of phyto-cassanes as proposed on the basis of the CD spectrum of synthetic (+)-2-deoxyphytocassane A. *Tetrahedron Lett.* **41**, 351–354.
- Yajima, A., Mori, K. and Yabuta, G. (2004) Total synthesis of ent-cassa-12,15-diene, a putative precursor of rice phytoalexins, phytocassanes A-E. *Tetrahedron Lett.* **45**, 167–169.
- Yamaguchi, S., Saito, T., Abe, H., Yamane, H., Murofushi, N. and Kamiya, Y. (1996) Molecular cloning and characterization of a cDNA encoding the gibberellin biosynthetic enzyme *ent*-kaurene synthase B from pumpkin (*Cucurbita maxima* L.). *Plant J.* **10**, 101–111.
- Yamaguchi, S., Sun, T., Kawaide, H. and Kamiya, Y. (1998) The GA2 locus of *Arabidopsis thaliana* encodes *ent*-kaurene synthase of gibberellin biosynthesis. *Plant Physiol.* **116**, 1271–1278.
- Yee, N.K.N. and Coates, R.M. (1992) Total synthesis of (+)-9,10-syn- and (+)-9,10-anti-copalol via epoxy trienylsilane cyclizations. *J. Org. Chem.* **57**, 4598–4608.
- Yu, J., Hu, S., Wang, J. et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science*, **296**, 79–92.