

A diverse family of type III polyketide synthases in *Eucalyptus* species†

Sheryl B. Rubin-Pitel,^{‡a} Yunzi Luo,^{§a} Jung-Kul Lee^b and Huimin Zhao^{*a}

Received 9th April 2010, Accepted 14th May 2010

First published as an Advance Article on the web 1st June 2010

DOI: 10.1039/c004992a

Eucalyptus species synthesize a wealth of polyketide natural products, but no relevant biosynthetic enzyme has been identified. Degenerate primers designed from conserved regions of fourteen chalcone synthase superfamily enzymes were used to isolate gene fragments from at least five different Type III polyketide synthases (PKSs) in *E. camaldulensis* and *E. robusta*.

Medicinal plants have been used for thousands of years in the treatment of numerous ailments.¹ Modern analytical techniques have been used to identify bioactive compounds in plant extracts and approximately 25% of today's medicines are derived from plant natural products.² Plant-derived polyketides such as curcuminoids (curcumin) and flavonoids (genistein, quercetin, resveratrol) are well-known complementary therapies, with a growing body of evidence vis-à-vis their usefulness as anti-inflammatory and anticancer agents.^{3,4} However, acylphloroglucinols, which make up the bioactive constituents of both modern and traditional therapies, have received comparatively little attention.

A number of plant species produce a wide variety of phloroglucinol derivatives. The diversity of natural phloroglucinols comes both from the priming acyl-CoAs and elongating acyl-CoAs used, and the transformations occurring after polyketide synthesis. For example, phloroglucinol natural products may be monomeric, dimeric, trimeric, or modified by prenylation or *O*- or *C*-glycosylation. These varied phloroglucinols have many known bioactivities including anti-feedant, anti-fungal, antibiotic, antitumor, antimalarial, antidepressant, anti-dementia, antiviral, antioxidant, and anti-inflammatory or anti-allergy activities.⁵

The phloroglucinol components of bioactive plants were generally discovered by analyzing organic solvent fractions of plant material. Searching databases of scientific literature uncovered a number of target plants across a wide range of species, and in nearly all cases, chemical characterization of the phloroglucinol natural products was not followed by investigation to discover the related biosynthetic enzymes. Acylphloroglucinol compounds may be formylated or acylphloroglucinol-terpene adducts (e.g. euglobals and macrocarpals). *Eucalypt* polyketide natural products are known to be antiviral,^{6,7} antibacterial,⁸ anticancer,⁹ and cariostatic (inhibit the growth

of bacteria that cause tooth decay),^{10,11} and *Eucalypts* have been used by the Aboriginal people of Australia for the treatment of colds, influenza, toothache, snakebites, fevers, and other ailments.¹² However, there has been no cloning of relevant genes. Here we report the development of an efficient strategy for cloning and characterizing Type III PKS genes that are likely responsible for the synthesis of these acylphloroglucinols.

The Type III PKSs are the smallest and simplest PKS enzymes. They exist as homodimers with a molecular weight of approximately 80 000–90 000 Dalton.¹³ In spite of their simplicity, Type III PKSs produce a variety of aromatic products with single or multi-ring structures including stilbenes, pyrones, naphthalenes, phloroglucinols, chromones and curcuminoids. The best-studied plant Type III PKSs are the chalcone synthases, which are involved in the synthesis of precursors to defensive compounds and pigments. Among the various Type III PKSs, only two acylphloroglucinol-producing enzymes from plants have been studied in significant detail: isovalerophenone synthase from *Humulus lupulus*¹⁴ and PKS1 from *Hypericum perforatum*,¹⁵ which synthesize isovalerophenone and the precursor to hyperforin, respectively. Trees of the genus *Eucalyptus* represent a group of plants in which acylphloroglucinols are particularly abundant.

Two *Eucalyptus* species were chosen as initial targets for Type III PKS discovery. *Eucalyptus robusta* produces the dimeric and monomeric acylphloroglucinols Robustal A and B, respectively, as well as the acylphloroglucinol-terpene adducts Robustal A and B (Fig. 1).¹² The leaves of *E. robusta* are used to produce an antimalarial medicine, “Da Ye An,” in China, and also to treat bacterial illnesses and dysentery.^{12,16} *Eucalyptus camaldulensis*, or red gum, was chosen as the second target species. It is an abundant tree also known to produce acylphloroglucinol-terpene adducts.¹⁷

Detection of genes by PCR is a commonly used strategy in environmental and clinical molecular biology.¹⁸ By incorporating genetic diversity into a primer, it can hybridize with a variety of related target gene sequences, in this case unknown members of the chalcone synthase (CHS) family from various medicinal plants. Fourteen enzymes in the plant Type III PKS family, representing a broad spectrum of catalytic activities (Fig. S1), were used as the templates for degenerate primer design (details on degenerate primer design are given in the ESI†).

^a Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave. Urbana IL 61801, USA. E-mail: zhao5@illinois.edu; Fax: (001) 217-333-5052

^b Department of Chemical Engineering and Institute of SK-KU Biomaterials, Konkuk University, Seoul, 143-701, South Korea

† Electronic supplementary information (ESI) available: Experimental, supplementary figures and tables. See DOI: 10.1039/c004992a

‡ Present address: ExxonMobil Research and Engineering Company, 600 Billingsport Road, Paulsboro NJ 08066, USA.

§ These authors contributed equally to this work.

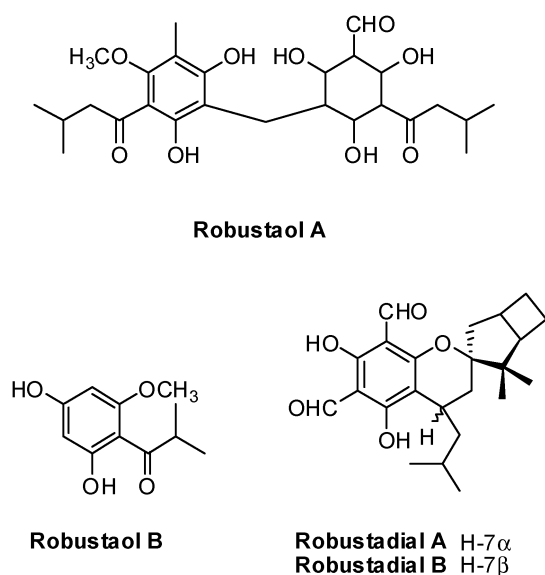


Fig. 1 Polyketides and polyketide-terpene adducts from *E. robusta*.

Six regions of homology were identified from Clustal W alignment¹⁹ of the Type III PKS mRNA sequences and used to design three sets of nested degenerate primers. The primers for amplifying the smallest fragment of approximately 400 nt were also designed such that the important residues Thr-197 and Gly-256 (*M. sativa* CHS2 numbering) would be represented (Fig. S2). The amino acids at those positions modulate active site volume and their identities suggest the substrate preference and elongation number of a Type III PKS.²⁰

To validate the degenerate primers designed in this study, Type III PKSs from *Hypericum perforatum*, also known as St. John's Wort, were isolated (see details in the ESI†). St. John's Wort is a medicinal plant and extracts of its aerial parts are still commonly used for the treatment of mild to moderate depression, and four Type III PKSs were identified from this plant.²¹ Nested degenerate PCR was performed on a cDNA library prepared from *H. perforatum* var. Topas seedlings germinated on Murashige-Skoog media, and resulted in gene fragments from three enzymes: *H. perforatum* chalcone synthase, benzophenone synthase, and a previously uncharacterized chalcone synthase family enzyme.

Specimens of the target plants *Eucalyptus camaldulensis* var. Silverton Province and *Eucalyptus robusta* were obtained for study. Briefly, the collected leaf tissue was flash-frozen in liquid nitrogen and crushed, and total RNA was extracted from the plant tissue and used to create cDNA libraries. This cDNA was used as the template for nested PCR. At each stage, PCR products were analyzed by agarose gel electrophoresis, and in the initial step, an amplification of a fragment of the abundant mRNA for RuBisCo (Ribulose-1,5-bisphosphate carboxylase oxygenase) was also used as a control. The first and second amplifications did not result in bands of the expected size, with the exception of the RuBisCo control, but the final amplification gave a fragment of the appropriate size (~400 nt). The fragments were then digested and ligated into a vector for sequencing. The absence of product bands until the third nested PCR indicates the need for this strategy due to low abundance of the mRNAs of interest in *Eucalyptus*.

A total of 94 clones and 27 clones were sequenced from the cDNA libraries created from the leaves of *E. camaldulensis* and *E. robusta*, respectively. Depending on the threshold (*i.e.* the number of different amino acids in the cloned DNA fragments), the number of different genes could range from 5 to 97. If the threshold in the validation experiment was used as a reference (*i.e.* genes with 80% sequence identity are considered the same), the identified DNA fragments could be classified into 11 groups, representing 11 unique putative Type III PKSs genes (Table S4). To make it more stringent, these 11 groups can be further consolidated into 5 groups (PKS1-5) according to the characteristic residues (see below). Representation of each group ranged from 1 to approximately 50 clones. Also, in each group, one of the three conserved catalytic triad residues seen in all known Type III PKSs (His) was present, as well as a highly conserved threonine residue (CHS Thr-197 equivalent).

Based on the translated sequence, specifically on the Gly-256 (*M. sativa* chalcone synthase numbering) equivalent and nearby residues, the activity of a new Type III PKS could be inferred. Enzymes utilizing coumaroyl-CoA as the priming acyl-CoA (chalcone synthase and stilbene synthase enzymes) typically possess a glycine at the equivalent of CHS Gly-256, and a DGH (Asp-255, Gly-256, His-257) pattern, indicating that PKS4 likely represents CHS or STS enzymes. Benzophenone synthases possess a TAH pattern, which is shared by PKS2.²²

PKS1	AKDVAENNKGARVLVVC	AEITVCTFRGPTET	YLDNLVGQALFGDGASSVIVGSDPVP	GV-ERPLYEIVS
PKS2	AKDVAENNANARVLVVC	AENTAMTFHAPNESHLDVIVGQAMFSDGAAALITGANPD	TAAGERAVFNILS	
PKS3	AKDVAENNAGARVLVVC	SEITVTTFRGSPSAENLANLVGQALIFGDGAAALITGSDPEIPA	-ERPIFQLVW	
PKS4	-KDVAENNKGARVLVVC	SEITAVTFRGPSDTHLDSLVGQALFGDGAAALIVGSDPIPEI	-ERPLFEMVW	
PKS5	---AENNKGARVLVVC	SEITAVTFRGPSDTRLNLVGQALFGDGAAALIVGSDPIPEI	-EKPSFEMVW	
PKS1	TSQTLTP-NSEGAIAGHVREIGLTIHLLKDVPGLISK	NVEKNMTEAFHPLGITDWN	SLFFIVHPGGTGI	
PKS2	ASQTIVP-GSDGAI	TAHFYGMGMSYFLKEDV	IPLFRDNIADVMKEAFSP	LDVSDWN
PKS3	TAETVLP-NSEGAI	EGQLRDVGLTFGLSDGVPHLI	ANHIEKRLKEAFDSL	GISDWN
PKS4	TAQTILP-DSEGAI	DGHLREVGLTFHLLKDVPGLISK	NIKALVEAFQPLGISD	YNSIFWIAHPGGTGI
PKS5	TAQTIFPHDGEDITAAHFREIGLTIHLLHRDIPRLVSK	NIKAMIEAFQPLGISD	YNSIFWIAHPGGRGI	

Fig. 2 Alignment of Type III PKS fragments from five distinct enzymes discovered in this work. Catalytic histidine residue is marked in red. The CHS Thr-197 and CHS Gly-256 equivalent residues are marked in blue.

Two enzymes possessed unusual residues at those locations: AGH was seen in PKS1, and AAH for PKS5.

The translated sequence of PKS3 shows that it possesses the residues EGQ at the CHS Asp-255, Gly-256, His-257 equivalent positions. Clustal W alignment of representative translated sequences from each group is shown in Fig. 2, and highlights the diversity present among these five putative Type III PKS enzymes. The phylogenetic relationships among the Type III PKS fragments is given in ESI Fig. S6†.

To further validate the cloned putative Type III PKS fragments, RACE-PCR (rapid amplification of cDNA ends-polymerase chain reaction) was used to clone the full-length gene of a member in the PKS1 group, EC2 (see ESI†). The resulting full-length gene consists of 1170 nucleotides, which corresponds to 389 amino acid residues. The gene was cloned into the pET26b(+) vector and expressed as a soluble protein in *E. coli*. Bioinformatic analysis indicates that the catalytic triad residues (Cys, His, Asn) of Type III PKSs are all present in this protein. Further biochemical characterization is in progress.

Conclusions

In this study we developed a PCR-based screening method able to identify novel Type III PKSs in medicinal plants. Initial results from screening *Eucalyptus* species have been promising, and led to the identification of gene fragments from at least five putative Type III PKS enzymes between *E. robusta* and *E. camaldulensis*, including two enzymes which appear to be conserved in both species. Bioinformatic studies allowed us to identify one putative benzophenone synthase and three putative chalcone or stilbene synthases. RACE-PCR was used to obtain the full-length gene for one *Eucalyptus* Type III PKS and further biochemical characterization is under way. In the future the acylphloroglucinol synthase enzymes present in *E. robusta* and *E. camaldulensis* may be isolated by biochemical characterization of the Type III PKS enzymes identified in this work.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (220-2009-1-D00033).

Notes and references

- 1 R. A. Halberstein, *Ann. Epidemiol.*, 2005, **15**, 686–699.
- 2 L. G. Zhou and J. Y. Wu, *Nat. Prod. Rep.*, 2006, **23**, 789–810.
- 3 K. Bisht, K. H. Wagner and A. C. Bulmer, *Toxicology*, 2009.
- 4 N. Khan, F. Afaq and H. Mukhtar, *Antioxid. Redox Signaling*, 2008, **10**, 475–510.
- 5 I. P. Singh and S. B. Bharate, *Nat. Prod. Rep.*, 2006, **23**, 558–591.
- 6 M. Takasaki, T. Konoshima, T. Shingu, H. Tokuda, H. Nishino, A. Iwashima and M. Kozuka, *Chem. Pharm. Bull. (Tokyo)*, 1990, **38**, 1444–1446.
- 7 K. Umehara, I. P. Singh, H. Etoh, M. Takasaki and T. Konoshima, *Phytochemistry*, 1998, **49**, 1699–1704.
- 8 Y. Yamakoshi, M. Murata, A. Shimizu and S. Homma, *Biosci., Biotechnol., Biochem.*, 1992, **56**, 1570–1576.
- 9 S. Yin, J. J. Xue, C. Q. Fan, Z. H. Miao, J. Ding and J. M. Yue, *Org. Lett.*, 2007, **9**, 5549–5552.
- 10 H. Nagata, Y. Inagaki, Y. Yamamoto, K. Maeda, K. Kataoka, K. Osawa and S. Shizukuishi, *Oral Microbiol. Immunol.*, 2006, **21**, 159–163.
- 11 K. Osawa, H. Yasuda, H. Morita, K. Takeya and H. Itokawa, *Phytochemistry*, 1995, **40**, 183–184.
- 12 E. L. Ghisalberti, *Phytochemistry*, 1996, **41**, 7–22.
- 13 K. Watanabe, A. P. Praseuth and C. C. Wang, *Curr. Opin. Chem. Biol.*, 2007, **11**, 279–286.
- 14 Y. Okada and K. Ito, *Biosci., Biotechnol., Biochem.*, 2001, **65**, 150–155.
- 15 H. P. Bais, R. Vepachedu, C. B. Lawrence, F. R. Stermitz and J. M. Vivanco, *J. Biol. Chem.*, 2003, **278**, 32413–32422.
- 16 S. Schwikkard and F. R. van Heerden, *Nat. Prod. Rep.*, 2002, **19**, 675–692.
- 17 B. M. Eschler, D. M. Pass, R. Willis and W. J. Foley, *Biochem. Syst. Ecol.*, 2000, **28**, 813–824.
- 18 O. J. Jabado, G. Palacios, V. Kapoor, J. Hui, N. Renwick, J. Zhai, T. Briese and W. I. Lipkin, *Nucleic Acids Res.*, 2006, **34**, 6605–6611.
- 19 J. D. Thompson, D. G. Higgins and T. J. Gibson, *Nucleic Acids Res.*, 1994, **22**, 4673–4680.
- 20 M. B. Austin and J. P. Noel, *Nat. Prod. Rep.*, 2003, **20**, 79–110.
- 21 P. Klingauf, T. Beuerle, A. Mellenthin, S. A. El-Moghazy, Z. Boubakir and L. Beerhues, *Phytochemistry*, 2005, **66**, 139–145.
- 22 B. Liu, H. Falkenstein-Paul, W. Schmidt and L. Beerhues, *Plant J.*, 2003, **34**, 847–855.