การโคลนและการแสดงออกของยีนโพลีคีไทด์ซินเทส จากชุมเห็ดเทศ เจตมูลเพลิงแดง และทาทาเรี่ยนรูบาบ

นายศุภชัย สมัปปิโต

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาเภสัชเคมีและผลิตภัณฑ์ธรรมชาติ
คณะเภสัชศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2545

ISBN 974-17-1340-1

CLONING AND EXPRESSION OF POLYKETIDE SYNTHASE GENES FROM *CASSIA ALATA*, *PLUMBAGO INDICA* AND *RHEUM TATARICUM*

Mr.Supachai Samappito

A Dessertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Science

Faculty of Pharmaceutical Sciences

Chulalongkorn University

Academic Year 2002

ISBN 974-17-1340-1

Thesis Title	CLONING AND EXPRESSION OF POLYKETIDE SYNTHASE GENES FROM *Cassia alata, Plumbago indica* and *Rheum tataricum*
By	Mr. Supachai Samappito
Field of study	Pharmaceutical Chemistry and Natural Product
Thesis Advisor	Associate Professor Wanchai De-Eknamkul, Ph.D.
Thesis Co-advisor	Professor Toni M. Kutchan, Ph.D.

Accepted by the Faculty of Pharmaceutical Sciences, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctor's Degree.

…… Dean of Faculty of Pharmaceutical Sciences
(Associate Professor Boonyong Tantisira, Ph.D.)

THESIS COMMITTEE

…… Chairman
(Associate Professor Ekarin Saifah, Ph.D.)

…… Thesis Advisor
(Associate Professor Wanchai De-Eknamkul, Ph.D.)

…… Member
(Associate Professor Wichai Cherdshewasart, D.Sc.)

…… Member
(Associate Professor Nongluksna Sriubolmas, Ph.D.)

…… Member
(Assistant Professor Rutt Suttisri, Ph.D.)

…… Member
(Assistant Professor Arinthip Thamchaipenet, Ph.D.)

ศุภชัย สมัปปิโต : การโคลนและการแสดงออกของยีนโพลีคีไทด์ชินเทสจากชุมเห็ดเทศ เจตมูลเพลิงแดง และทาทาเรี่ยนรูบาบ (CLONING AND EXPRESSION OF POLYKETIDE SYNTHASE GENES FROM *Cassia alata, Plumbago indica* and *Rheum tataricum*) อ.ที่ปรึกษา : รศ.ดร.วันชัย ดีเอกนามกูล, อ.ที่ปรึกษาร่วม : ศ.ดร.โทนี่ เอ็ม คุทชาน, 183 หน้า. ISBN 974-17-1340-1

การศึกษานี้ได้ทำการโคลนและติดตามการแสดงออกของยีนโพลีคีไทด์ชินเทสจากพืชสมุนไพรไทยที่สร้างสารแอนทราควิโนน (ชุมเห็ดเทศและทาทาเรี่ยนรูบาบ) และพืชที่สร้างสารแนพโธรควิโนน (เจตมูลเพลิงแดง) ผลจากการทำ cDNA library ที่ได้จากการใช้อาร์เอ็นเอ (RNA) จากชิ้นส่วนรากของชุมเห็ดเทศ พบว่าได้ cDNA จำนวน 3 ชิ้นที่สมบูรณ์มีลำดับกรดอะมิโนคล้ายคลึงกันแต่ไม่เหมือนกันทั้งหมด เมื่อวิเคราะห์การแสดงออกของยีนด้วยเทคนิค northern blot แสดงให้เห็นว่ายีนดังกล่าวมีการแสดงออกมากในส่วนราก ซึ่งสอดคล้องกับผลการวิเคราะห์สารฟลาโวนอยด์ด้วย HPLC ที่พบสาร quercetin, naringenin และ kaempferol เป็นต้น จากนั้นนำยีนดังกล่าวไปแสดงออกและชักนำให้สร้างโปรตีนใน *Escherichia coli* พบว่าเอนไซม์บริสุทธิ์ที่ได้สามารถใช้สารตั้งต้นโคเอ (substrate-CoA) ได้ทั้ง aromatic และ aliphatic-CoA เมื่อทำปฏิกิริยาร่วมกับ malonyl-CoA ทำให้เกิดผลิตภัณฑ์หลากหลายชนิดในหลอดทดลอง โดยมีสารผลิตภัณฑ์หลักคือ naringenin แสดงให้เห็นว่ายีนที่แยกได้ทำหน้าที่ควบคุมการสร้างเอนไซม์ชาลโคลนชินเทส (chalcone synthase, CHS) ซึ่งเกี่ยวข้องกับชีวสังเคราะห์ของสารฟลาโวนอยด์ (flavonoid) ในรากของชุมเห็ดเทศ ส่วนในกรณีของเจตมูลเพลิงแดงภายหลังการทำ cDNA library โดยใช้ RNA ที่เตรียมจากชิ้นส่วนรากจะพบ cDNA จำนวน 1 ชิ้นที่สมบูรณ์ เมื่อวิเคราะห์การแสดงออกของยีนด้วยเทคนิค northern blot จะเห็นว่าการแสดงออกของยีนนี้จะพบมากในส่วนราก เมื่อนำยีนนี้ไปแสดงออกและชักนำให้สร้างโปรตีนใน *E. coli* จะพบว่าเอนไซม์ที่ได้สามารถใช้ substrate-CoA ได้หลายชนิด ร่วมกับ malonyl-CoA และเกิดผลิตภัณฑ์หลักเฉพาะสารกลุ่มไพโรน (pyrone) แสดงให้เห็นว่ายีนที่แยกได้ทำหน้าที่ควบคุมการสร้างเอนไซม์ไพโรนชินเทส (pyrone synthase, PS) นอกจากนี้ยังค้นพบ cDNA จำนวน 1 ชิ้นที่สมบูรณ์จากชิ้นส่วนไรโซมของทาทาเรี่ยนรูบาบ เมื่อนำยีนที่ได้มาแสดงออกและชักนำให้สร้างโปรตีนใน *E. coli* พบว่าเอนไซม์บริสุทธิ์ดังกล่าวสามารถใช้ substrate-CoA ได้ทั้ง aromatic และ aliphatic-CoA เมื่อทำปฏิกิริยาร่วมกับ malonyl-CoA สามารถเกิดผลิตภัณฑ์หลักเป็นสารเรสเวอราทอล (resveratrol) นอกจากนี้ยังพบว่าเอนไซม์นี้ยังสามารถสร้างสาร naringenin, bis-noryangonin (BNY-type) และ (p-coumroyltriacetic lactone, CTAL) เกิดขึ้นในปฏิกิริยาในปริมาณเล็กน้อยอีกด้วย แสดงให้เห็นว่ายีนที่แยกได้จากรากของทาทาเรี่ยนรูบาบทำหน้าที่ควบคุมการสร้างเอนไซม์สทิลบีนชินเทส (stilbene synthase, STS) และเมื่อทำ northern blot แสดงให้เห็นว่าการแสดงออกของยีนนี้จะพบมากในบริเวณราก ซึ่งสอดคล้องกับผลการตรวจวิเคราะห์สารเรสเวอราทอลด้วย HPLC ที่พบปริมาณมากในส่วนรากของทาทาเรี่ยนรูบาบเช่นกัน

ภาควิชา	-	ลายมือชื่อนิสิต .....
สาขาวิชา	เภสัชเคมีและผลิตภัณฑ์ธรรมชาติ	ลายมือชื่ออาจารย์ที่ปรึกษา .....
ปีการศึกษา	2545	ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....

4176958033 : MAJOR PHARMACEUTICAL CHEMISTRY AND NATURAL PRODUCT
KEY WORD: CLONING, EXPRESSION, PLANT POLYKETIDE SYNTHASE, GENES, *Cassia alata, Plumbago indica* and *Rheum tataricum*
SUPACHAI SAMAPPITO : THESIS TITLE. CLONING AND EXPRESSION OF POLYKETIDE SYNTHASE GENES FROM *Cassia alata, Plumbago indica* and *Rheum tataricum.* THESIS ADVISOR : ASSOC.PROF.WANCHAI DE-EKNAMKUL, Ph.D., THESIS CO-ADVISOR : PROF.TONI M. KUTCHAN, Ph.D. 183 pp.
ISBN 974-17-1340-1

Studies on cloning and expression of polyketide synthases (PKSs) have been performed using Thai medicinal plants producing anthraquinones (*Cassia alata* and *Rheum tataricum*) and naphthoquinones (*Plumbago indica*). Three cDNAs encoding very similar but unique isoforms of chalcone synthase [EC 2.3.1.74] were isolated from a cDNA library prepared from RNA from root tissue of *C. alata*. Gene transcript for these three type III polyketide synthases was found to accumulate predominantly in root. The heterologously expressed enzymes accepted a broad starter-coenzyme A (CoA) and together with the co-substrate malonyl-CoA *in vitro* assays, formed multiple products. *C. alata* accumulates the flavonoids quercetin, naringenin and kaempferol in roots, suggesting that the *in planta* function of these enzymes is the biosynthesis of root flavonoids. For *P. indica,* one cDNA cloned encoding a pyrone synthase, *Pin*PS, was isolated from the root part that correlated with the expression pattern of the *Pin*PS transcripts were found predominantly in roots. Recombinant *Pin*PS was expressed in *E. coli* and assayed with a various starter-CoA as primers of polyketide synthesis and was found to accepted broad substrate specificity in *in vitro* assay and produce pyrone product. For *R. tataricum, a* cDNA encoding a stilbene synthase, *Rt*STS, was isolated from the rhizomes. Recombinant *Rt*STS expressed in *E. coli* and assayed with a various substrate-CoA as primers, could synthesize resveratrol and a trace amount of naringenin chalcone from *p*-coumaroyl-CoA. These suggested that the enzyme is a resveratrol-type stilbene synthase (EC 2.3.1.95). Moreover, bis-noryangonin and *p*-coumaroyl triacetic acid lactone (CTAL)-type pyrones were observed in minor amounts in the reaction with *p*-coumaroyl-CoA. . Using Northern blot analysis, *Rt*STS transcript was found to be highly expressed in *R. tataricum* rhizomes.This expression pattern correlated with the occurrence of resveratrol, which was detected in higher amounts in *R. tataricum* rhizomes compared with leaves and petioles using HPLC.

Program	Phar.Chem and Nat. Prod.	Student's signature ……..
Field of study	Phar.Chem and Nat. Prod.	Advisor's signature ……...
Academic year	2002	Co-advisor's signature ……

ACKNOWLEDGEMENTS

I would like to thank Associate Professor Dr.Wanchai De-Eknamkul, of the Department of Pharmaconognosy, Faculty of Pharmaceutical Sciences, Chulalongkorn University for his excellent guidance, for providing me an opportunity to gain scientific experiences and continual encouragement throughout the course of this work.

I would like to express my sincere thank to Prof.Dr.Toni Kutchan for giving me a chance to learn the techniques of plant molecular biology. Continuing this laboratories at the Institute of Plant Biochemistry (IPB), Halle, Germany, was a real pleasure.

I would like to thank Prof.Dr.Dr.Meinhart H. Zenk, Biocenter, Martin-Luther University, Halle, Germany, for his excellent guidance and for giving me an opportunity to gain scientific experiences.

I would like to thank Dr.Jonathan Page from IPB, Halle, Germany, for introducing and guiding me various techniques of molecular biology. Without him, it would be much more difficult for me to finish this work.

I would like to thank to Dr.Jurgen Schmidt, IPB, Halle, Germany, for his accurrate performance of mass spectrometry analysis.

I would like to thank the Thai Royal Golden Jubilee Fund, Bangkok, the Deutsche Akademische Austausch Dienst, Bonn, the Alexander von Humboldt Foundation, Bonn and by Fonds der Chemischen Industrie, Frankfurt, that supported through this research work.

I would like to thank all members of the Tissue Culture Unit, Faculty of Pharmaceutical Sciences, Chulalongkorn Univesity, for their friendships, encouragements, suggestions and helps.

Finally, I wish to express my indebtedness and grateful thanks to my wife and my family for their financial support and without whose love, understanding and cheerfulness this work would have not been successful.

CONTENTS

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

ABBREVIATIONS

ACS	acridone synthase
BAS	benzalacetone synthase
BBS	bibenzyl synthase
bp	base pair
BPS	benzophenone synthase
BNY	bis-noryangonin
CalCHS	*Cassia alata* chalcone synthase
cDNA	complementary DNA
CID	collision-induced decomposition
CHI	chalcone isomerase
CHS	chalcone synthase
Ci	curie
CoA	coenzyme-A
cpm	count per minute
CTAB	cetyltrimethyl ammonium bromide
CTAL	p-coumaroyl triacetic acid lactone
dNTP	2'-deoxynucleoside 5'-triphosphate
DTT	dithiothreitol
E.coli	*Escherichia coli*
h	hour
I	inosine
IPTG	isopropylthio-β-D-galactoside
kD	kilodalton
LB-medium	Luria-Bertani medium
M	molar,mol.l^{-1}
min	minute
MMLV-RT	moloney murine leukemia virus-reverse transcriptase
MOPS	morpholinopropane sulphonic acid
mRNA	messenger ribonucleic acid

ABBREVIATIONS (continued)

OD	optical density
PCR	polymerase chain reaction
Pfu	plaque forming unit
PIBP	phloroisobutyrophenone
*Pin*PS	*Plumbago indica* pyrone synthase
PIVP	phloroisovalerophenone
PKS	plant polyketide synthase
PS	pyrone synthase
RACE	rapid amplification of cDNA ends
rpm	rotations per minute
RT	room temperature
*Rt*STS	*Rheum tataricum* stilbene synthase
SDS	sodium dodecyl sulfate
sec	second
SPS	styrylpyrone synthase
SSC	saline sodium citrate
STS	stilbene synthase
TEMED	*N,N,N',N'*-tetramethyethylenediamine
THN	1,3,6,8-tetrahydroxynaphthalene
TLC	thin layer chromatography
Tris	tris(hydroxymethyl)aminomethane
U	unit
UPM	universal primer mix
V	voltage
VPS	valerophenone synthase
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

CHAPTER I
INTRODUCTION

Plant polyketide synthases (PKSs) are encoded by a multigene family that has chalcone synthase as the prototype. Gene family evolution in plants appears to occur through gene duplication, followed by nucleotide substitution which can lead to biochemical diversity. Plant PKSs are presumably derived from a common ancestor which diverged to perform different reactions. This is clearly demonstrated by the highly similar genes that encode chalcone synthase and resveratrol synthase found in a single species, for example, in *Vitis vinifera* (Melchior and Kindl, 1990; Sparvoli *et al*., 1994), or by the occurrence of the closely related, but enzymatically unique chalcone and acridone synthases in *Ruta graveolens* (Springob *et al*., 2000). Detailed crystal structure and site-directed mutagenesis studies have provided insight into the structural control of polyketide formation in plant homodimeric iterative (type III) polyketide synthases (Jez *et al*., 2000a; Suh *et al*., 2000). The active site cavity can serve as a size-based filter to sterically exclude bulkier starter molecules (Jez *et al*., 2000a). This explains why pyrone synthase cannot accept *p*-coumaryl CoA as a starter molecule, whereas chalcone synthase can use acetyl CoA as a starter molecule, albeit inefficiently, and forms a small amount of methylpyrone (Schuz *et al*., 1983). Modification of as few as three amino acids can change chalcone synthase into pyrone synthase (Jez *et al*., 2000a). This demonstrated how readily plant type III polyketide synthases can be modified by nucleotide substitution to produce new natural products. Plants produce a variety of polyketide-derived secondary metabolites such as naphtha- and anthraquinones that could potentially be biosynthesized from type III polyketide synthases, such as has been demonstrated in the formation of tetrahydroxy-naphthalene in the microorganism *Streptomyces griseus* (Funa *et al*., 1999).

Plant PKSs are catalytically flexible, usually forming several products *in vitro*, and accepting a variety of physiological and non-physiological substrates. Both chalcone and stilbene synthases convert *p*-coumaroyl-CoA to the major products naringenin chalcone or resveratrol, respectively, as well as to the byproduct pyrones bis-noryangonin (BNY) and *p*-coumaroyl triacetic acid lactone (CTAL) (Yamaguchi *et al*., 1999). The latter are formed by derailment after two and three condensation reactions with malonyl-CoA, respectively. Additionally, cross reactivity between CHS and STS has been demonstrated with CHS forming resveratrol and STS forming naringenin in assays with *p*-coumaroyl-CoA, albeit in small amounts (Yamaguchi *et al*., 1999). Both CHSs and STSs accept aliphatic CoA esters as reaction primers. For example, *Scutellaria baicalensis* CHS converts isovaleryl-CoA to phloroacylphenones (via a correct Claisen cyclization) (Morita *et al*., 2000) while *Arachis hypogaea* STS forms BNY-type pyrones rather than the aromatic products from the same substrate (Morita *et al*., 2001). CHS and STS also utilize thiophene, furan and halogenated analogues of *p*-coumaroyl-CoA to form unnatural polyketides (Abe *et al*., 2000; Morita *et al*., 2001). The ability of plant PKSs to accept different CoA primers and carry out various condensation and cyclization reactions (Jez *et al*., 2002) makes this enzyme class, along with terpene cyclases (Greenhagen and Chappell, 2001), one of the major generators of carbon skeleton diversity in natural products.

As part of a study of the biosynthesis of polyketides present in Thai medicinal plants, we investigated the CHS superfamily members present in *Cassia alata* L. (Leguminosae), *Plumbago indica* L. (Plumbaginaceae) and *Rheum tataricum* L. (Polygonaceae).

Cassia alata L. (ringworm bush, Leguminosae), distributed mainly in the tropics and subtropics, is used to treat superficial mycosis of the skin (Crockett *et al*., 1992). Anthraquinones accumulate in aerial plant parts as well as in roots (Kelly *et al*., 1994;

Yadav and Kalidhar, 1994) and are the principal laxative constituents in *Cassia* species used as purgatives (Elujoba *et al*., 1989). *C. alata* also finds use in preventative veterinary medicine and is used to deworm dogs (Lans *et al*., 2000).

Plumbago indica L. (Plumbaginaceae), is an important medicinal plant, distributed mainly in the tropical regions. It is used in traditional treatments for leprosy, dyspepsia, diarrhea, skin diseases (Kirtikar and Basu, 1975). The roots and aerial parts of this plant contain several alkaloids and alkaloid derivatives such as plumbagin, droserone, isoshinanolone, eliptinone, chitranone, zelanone and isozelanone (Kamel *et al*., 1983;Santha Kumari *et al*.,1978) Pharmacological and phytochemical properties of the *Plumbago* spp., especially the antifertility and anticancer activities of extracts plumbagin, have been extensively investigated (Bhatia and Lal, 1933; Jayaraman, 1987; Krishnaswamy and Purusottaman, 1980; Van der Vijver, 1972).

Rheum tataricum (Polygonaceae), commonly know as tartarian rhubarb is a medicinal plant that contains considerable amount of aromatic polyketides such as anthraquinones (Okabe *et al*., 1973), stilbenes (Kashiwada *et al*., 1988), naphtalenes (Tsuboi *et al*., 1977), and phenylbutanones (Nonaka *et ,al*., 1981; 1983). Part of this plant is used as a pharmaceutical for its laxative, purgative and chathartic activities attributed to the anthraquinones and their derivatives which are usually contained in various species of rhubarb, have been considered as the active principles (Stoll *et al*., 1950; Miyamoto *et al*., 1967; and Oshio *et al*., 1974).

In an effort to characterize polyketide synthase genes from medicinal plants native to Thailand. The objective of the doctoral work described herein aims to provide a basic understanding of the secondary metabolic routes leading to polyketide formation in plants. The member of the CHS superfamily presented in *Cassia alata* L. (Leguminosae), *Plumbago indica* L. (Plumbaginaceae) and *Rheum tataricum* L. (Polygonaceae) are specifically investigated. A homology-based cloning strategy was

employed using degenerated primers to amplify conserved regions in plant PKS genes to obtain partial cDNA clones of PKS. The amplification products were subsequently used for screening cDNA libraries and RACE (Random Amplification cDNA End) PCR technique was also used to obtain full-length clones. The function and characterization of the plant polyketide synthases *in vitro* reaction is undertaken.

CHAPTER II

LITERATURE REVIEW

1. The reactions catalyzed by plant polyketide synthases

1.1 Overview of the reactions catalyzed by plant polyketide synthases

The principle reactions catalyzed by a polyketide synthase as deduced from functional and structural analyses are summarized in **Figure 1** (Ferrer *et al*., 1999; Schroder, 1999).

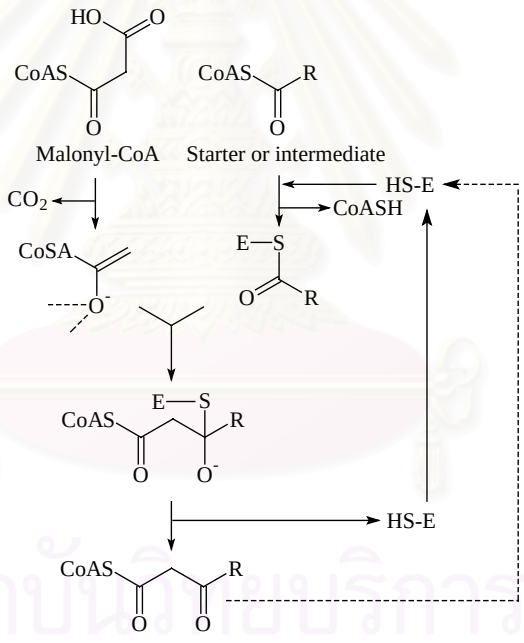


Figure 1 Model of the condensing reaction performed by CHS and related enzymes. Only the initial condensation is shown; the diketide product is the substrate for the next chain extension. HS-E is the active site cystein (Cys 169) that covalently binds the starter residue prior to the condensing reaction.

The reaction begins with the binding of the starter substrate CoA-ester to the active site containing cystein and the release of the free CoASH. The binding of malonyl-CoA is the next step, and chain extension proceeds with an acetyl-CoA anion produced by decarboxylation of malonyl-CoA. The anion is stabilized by various interactions with residues on the protein and by tautomerization between the keto and enol species, but apparently not by covalent binding of the residue to the enzyme protein (Ferrer *et al*, 1999). When compared to other polyketide synthases type I and type II, the distinctive features of the reactions catalyzed by plant polyketide synthase are the direct use of CoA-ester (no involvement of acyl carrier proteins or 4'-phosphopanthethein arms), the presence of only one essential cystein in the protein (covalent binding of the starter residue, or of the intermediate to be used in the next elongation cycle), and the chain extension without covalent binding of the extender unit to the protein (Schroder, 2000).

In terms of classification, the family of CHS-related enzymes can be classified based on their functional diversity such as substrate specificities, condensation reactions, types of ring closure of the release products, and in some cases by modification of reaction intermediates by additional enzymes. **Figure 2** gives an overview of the intermediates after one, two, and three condensation reactions, and the products identified from various activities of the enzymes. All complex reactions are carried out by homodimeric proteins with subunits of 40-44 kDa.

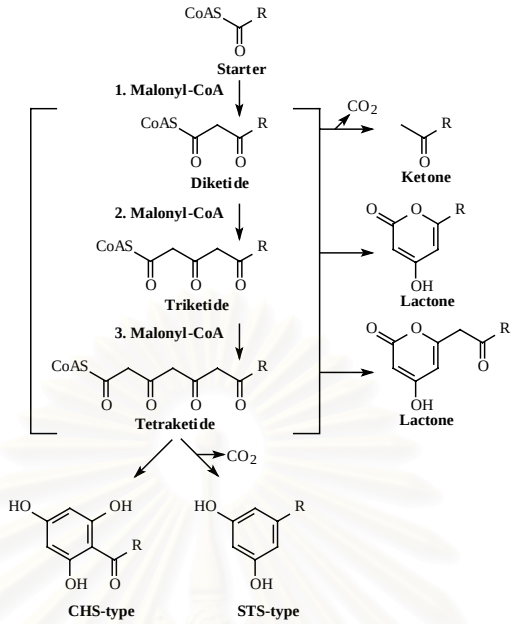


Figure 2 Overview of reaction intermediates (in brackets) and products involved with CHS-related proteins.

1.2 Chalcone synthase- type and stilbene synthase-type reactions

CHS enzymes use a starter CoA-ester from the phenylpropanoid pathway and perform three sequential condensation reactions with C_2 units from decarboxylated malonyl-CoA, to form a linear tetraketide intermediate which then folded to form a new aromatic ring system **(Figure 2).** The intermediates cannot be demonstrated directly due to their instability. The reactions of CHS and STS are identical up to the tetraketide stage. The difference is in the formation of the aromatic ring systems: the tetraketide must be folded differently in order to connect different carbon atoms for the ring closure to form either the chalcone or stilbene. The reaction mechanisms are also formally different (Claisen or Aldol condensation). All STSs analyzed *in vitro* remove the terminal carboxy group of the tetraketide and it is not known whether this occurs before or after the ring closure. The existence of stilbenoids retaining the carboxy group (Gorham, 1995) indicates that the decarboxylation is not a necessary consequence or

an essential part of the STS-type ring closure. The enzymes are either CHS or STS, and intermediate forms possessing both high CHS and STS activities are not known. Such forms have not been observed during attempts to convert a CHS into a STS by site-directed mutagenesis (Tropf *et al*., 1994), or in attempts to produce hybrids between CHS and STS (Tropf *et al*., 1995). This indicates an either/or switch mechanism in the CHS- and STS-type ring folding. However, data suggest that the mutual exclusion is probably not perfect (at least *in vitro*), because a few percent of the CHS products are stilbenes, and likewise STS may synthesize a very low percentage of the chalcone (Schroder, 1999). These small amounts are not detected in routine assays and cross-reactions between CHS and STS have only been confirmed by dilution analysis after extensive purification with HPLC. CHS and STS activities are usually measured by using [2-^{14}C] malonyl-CoA as substrate, followed by ethylacetate extraction of the products, and quantification after TLC analysis. The ethylacetate extracts may be counted directly (Schroder *et al*., 1997).

1.2.1 Properties of plant chalcone synthases

The first demonstration of CHS activity *in vitro* was reported in 1972 with crude enzyme extracts from parsley (*Petroselinum crispum*) cell suspensions (Kreuzaler and Hahlbrock, 1972). The properties of the enzyme from parsley and more than 30 other plants have subsequently been investigated (Heller and Forkmann, 1994). Many of the studies also investigated the induction kinetics of CHSs under various conditions, but the turnover of the protein has rarely been investigated. A study with parsley cell cultures indicated that the activity disappeared faster than the protein detectable in immunoreactions. The inactive protein revealed no difference in the size of the subunits, suggesting that the inactivation was not simply the result of proteolytic degradation (Schroder and Schafer, 1980).

Initially, the CHS product was identified as the isomeric flavanone which *in vivo* is the result of the chalcone/flavanone isomerase (CHI) activity on the chalcone. Later experiments with the parsley enzymes (Heller and Hahlbrock, 1980) showed that the apparent *in vitro* product was caused either by the presence of very stable CHI or by the rapid and nonenzymatic isomerization of the chalcone to the flavanone (Mol *et al*, 1998) at the pH of the assay (pH 8). The isomerization is pH-dependent, and about 50% conversion is observed even at pH 6 and in very short incubations. The complication of quantifying two products can be avoided in routine incubation. At the end of incubation, it is sufficient to raise the pH to 9 for 10 min to obtain the flavanone as the only product. The K_m for malonyl-CoA is in the range of 30 μM, and most of the enzymes prefer 4-coumaoryl-CoA as the starter unit (K_m usually below 10 μM). Other phenylpropanoid starters (*e.g*., cinnamoyl-CoA, caffeoyl-CoA) are also accepted *in vitro*, but usually at lower efficiency. The possibility that CHSs in some plants synthesize eriodictyol *in vivo* from caffeoyl-CoA has been suggested (Heller and Forkmann, 1994). The two enzymes cloned from *Hordeum vulgare* indicate that CHS2 prefers caffeoyl-CoA rather than 4-coumaroyl-CoA (Chistensen *et al*., 1998), and this appears to be the first direct evidence that 4-coumaroyl-CoA is not the only physiological substrate in some plants. The characterization of the proteins showed that the enzymes are dimers of identical subunits (41-45 kDa). Several properties indicated similarities with the condensing enzymes of fatty acid biosynthesis, for example, CO_2 exchange at the malonate moiety (Kreuzaler *et al*., 1978), inhibition by cerulenin (Kreuzaler and Hahlbrock, 1975), but there is no evidence for an acyl carrier protein or a 4'-phosphopantetheine arm being involved in the reactions (Kreuzaler *et al*., 1979). Two other points emerged from the experiments *in vitro*; (a) the substrate specificity is not confined to phenylpropanoid starters, because the enzyme also accepted alkoxy CoA-esters (Schuz *et al*., 1983), and (b) CHS can synthesize products

that are the result of the release of intermediates after one (benzalacetone, aryldihydropyrone) or two condensation reactions (styrylpyrones) (Kreuzaler and Hahlbrock, 1975; Hrazadina *et al*., 1976).

The first CHS sequence was published in 1983 for the enzyme from parsley (*Petroselinum crispum*) (Reimold *et al*., 1983), and many more sequences have been reported in subsequent years. However, the primary interest has often been in the investigation of transcriptional regulation or of the promoter properties, and the possibility that the cloned cDNAs/genes might encode another protein function has often been ignored (Schroder *et al*., 2000).

1.2.2 Other enzyme activities with CHS-type reactions

1.2.2.1 Acridone synthase (ACS)

Acridone alkaloids are only found in some genera of the family Rutaceae, and about 100 of these substances, with a remarkable variety in structure are known. The enzymatic formation of the acridone backbone was investigated with cell cultures from *Ruta graveolens*. Crude extracts showed that the acridone was synthesized from *N*-methylanthraniloyl-CoA and three malonyl-CoA (Maier *et al*., 1993) and the new aromatic ring system was obviously formed by a CHS-type ring folding **(Figure 3).** The product, (1,3-dihydroxy-*N*-methylacridone), detected *in vitro* is the result of second ring closure, and it is not known whether this is an intrinsic activity of the enzyme or it is a nonenzymatic process. The purified enzyme has K_m of about 11 μM and 32 μM for the starter CoA and malonyl-CoA, respectively, (Baumert *et al*., 1994). Its size is about 69 kDa when measured by gel permeation chromatography, and its subunits of about 40 kDa can cross-react with CHS antisera. Aanalysis of cloned cDNA has proved that ACS is a member of the protein family (>65% overall identical with CHSs) and its function was unambiguously identified after heterologous

expression in *E. coli* (Junghanns *et al*., 1995). Although early data suggested that the enzyme was monomeric (Junghanns *et al*., 1995), later analyses with recombinant protein and also with a cloned from *R. graveolens* have shown that the enzyme are dimer like all CHS-related proteins (Lukacin *et al*., 1999). The ACS is different CHS in that it uses an *N*-methylanthranioyl-CoA as the starter, and the product belongs to an entirely different pathway. The relationship tree indicates that it is not clearly separated from other sequences designated as CHSs (Schroder *et al*., 2000).

1.2.2.2 Benzophenone synthase (BPS)

Xanthones are a group of natural products with interesting pharmaceutical properties, and the majority occur in two plant families (Gentianaceae and Hypericaceae) (Bennett and Lee, 1989). Precursor feeding experiments have indicated that the backbone of the xanthone ring system is synthesized from shikimate-derived precursors and *via* the acetate/malonate route (Sultanbawa, 1980). Cell cultures of *Centaurium erythraea* and *C. litorale* have been used to investigate the xanthones and the regulation of their formation (Beerhues and Berger, 1994;1995). The proposed biosynthetic pathway suggested the condensation of a benzoyl-CoA derivative with three acetate units derived from malonyl-CoA to form the benzophenone, the precursor of the xanthones **(Figure 3)** (Bennett and Lee, 1989; Beerhues and Berger, 1995). The reaction has been demonstrated *in vitro* with a partially purified protein preparation from *C. erythraea* incubated with 3-hydroxybenzoyl-CoA and malonyl-CoA (Beerhues, 1996). The enzyme has also been shown to accept benzoyl-CoA as starter (44% efficiency), but is completely inactive with 2-hydroxybenzoyl-CoA and 4-hydroxybenzoyl-CoA, specificity for only one position of the hydroxy group. The ring closure shows that the BPS is a CHS-type enzyme using a benzoyl-CoA instead of phenylpropanoid-CoA derivative as the starter (Schroder, 1999).

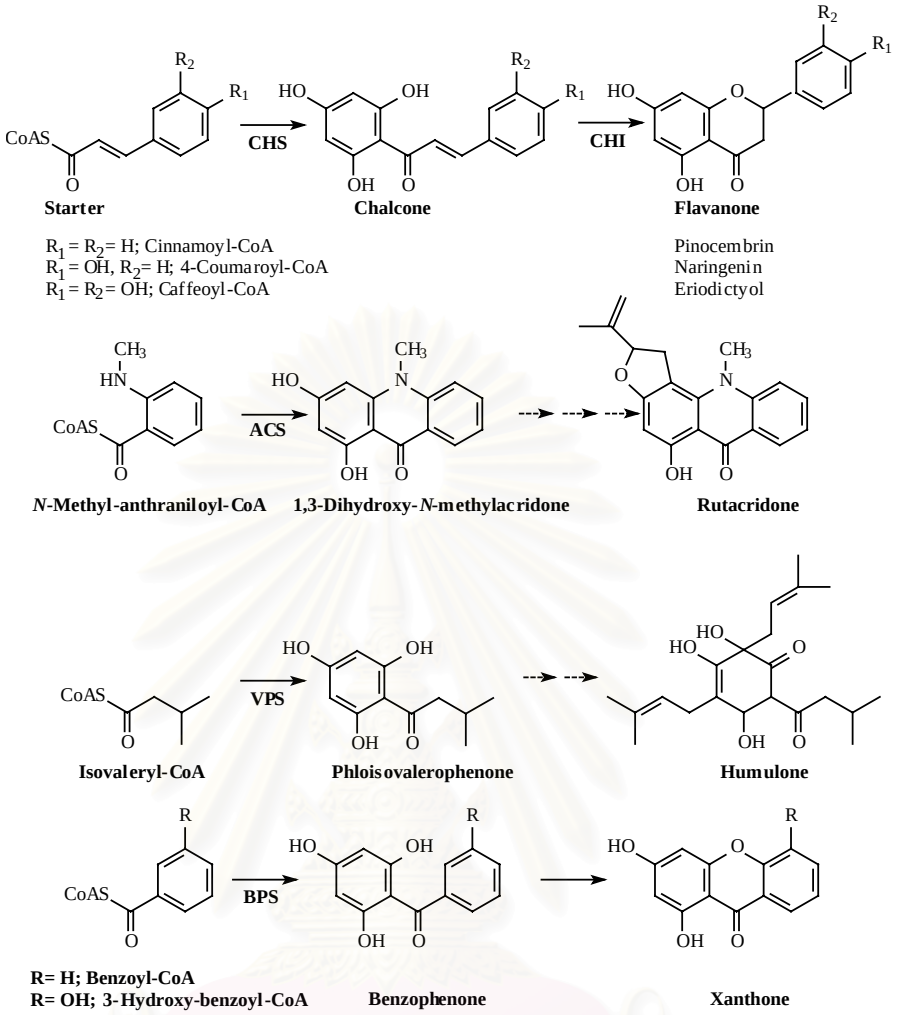


Figure 3 Reactions of CHS-type enzymes. CHS, Chalcone synthase, CHI, chalcone isomerase; ACS, acridone synthase; VPS, valerophenone synthase; BPS, benzophenone synthase.

1.2.2.3 Valerophenone synthase (VPS)

The ripe cones of hop (*Humulus lupulus*, Cannabinaceae) contain up to 20% of bitter acids, *e.g*. humulone and cohumulone, which are converted during the brewing process to the isoforms which are important for the flavor and taste of beer. It has been proposed based on the detection of phlorisovalerophenone (PIVP) and phloisobutyrophenone (PIBP) as intermediates (Zuurbier *et al*., 1995), that the first

aromatic intermediates in the pathway are produced *via* prenylated long-chain acids(Drawert and Beier, 1976). Verpoorte's group has suggested, instead, that the formation of the aromatic ring could precede all of the prenylation steps. Previous experiments have shown that CHS from parsley accepts alkoxy CoA-ester as starter substrate (Schuz *et al*., 1983) and therefore it seems possible that PIVP and PIBP are synthesized *via* a CHS-type reaction using isovaleryl-CoA and isobutyryl-CoA as starters and three malonyl-CoAs.

The hypothesis has been confirmed by a demonstration of the presence of the predicted enzyme activities in crude extracts (Zuurbier *et al*., 1995). Immunoblots with an antiserum against parsley CHS and denatured proteins revealed a band of 45 kDa, which is the expected size for CHSs and CHS-related proteins. CHS activity was also prefer 4-coumaryl-CoA as the starter CoA, but the activity with isovaleryl-CoA was always higher than with 4-coumaryl-CoA at each development stage, and CHS and PIVPS/PIBPS followed different kinetics during the development from flower buds to ripe cones. The presence of different enzymes needs to be confirmed by the separation of CHS from other activities, and ideally by cloning and heterologous expression of single proteins, because CHSs cloned from plants not producing these particular secondary products do accept isovaleryl-and isobutyryl-CoA as substrates (Zuurbier *et al*., 1998).

1.2.3 Properties of stilbene synthase

Stilbene synthase (STS) are characterized by three condensation reactions, but they are followed by a ring closure reaction that is clearly different from that of CHS (acylation versus aldol condensation). With all characterized STS activities, the formation of the released product involved the removal of the terminal carboxyl group of the tetraketide as CO_2 **(Figure 2).** The STS and CHS ring-folding types appear to be

mutually exclusive because enzymes producing both stilbenes and chalcones in comparable amounts are not known, and the products have not been detected in experiments coverting a CHS into a STS by site-directed mutagenesis (Tropf *et al*., 1994) or in attempts to produce hybrids between subunits of CHS and STS (Tropf *et al*., 1995). More detailed experiments have indicated, however, that small quantities of stilbenes (at most a few percent) are detectable in CHS incubations, and, likewise, small amounts of chalcones have been recovered from STS reactions (Schroder, 2000). These are not detected in standard incubations.

The first reaction *in vitro* was demonstrated in 1978 in crude extracts from rhizomes of rhubarb (*Rheum rhaponticum*), with 4-coumaroyl-CoA as starter CoA ester and resveratrol as product (Rupprich and Kindl, 1978). STS activities have also been demonstrated in extracts from Scots pine (*Pinus sylvestris*) (Schoppner and Kindl, 1979), groundnut (*Arachis hypogaea*) (Rolfs *et al*., 1981), and *Vitris* species (Fritzemeier and Kindl, 1981). Characterization of the enzymes has shown that they accept various CoA-esters from the phenylpropanoid pathway, but as with CHS, the normal precursor for the natural products occuring in the plants is preferred (Rupprich *et al*., 1980). The purified proteins appear to be homodimerics (90-95 kDa) with subunits of 43-47 kDa (Rolfs and Kindl, 1984). The pH optima and the K_m for the preferred starter CoA-ester and malonyl-CoA have been shown to be in the same range as determined for CHS (Schoppner and Kindl, 1984).

1.2.3.1 Cloned enzymes: Stilbene and Bibenzyl synthases

Both STS and bibenzyl synthase (BBS) use starter CoA-esters originating from the phenylpropanoid pathway **(Figure 4).**

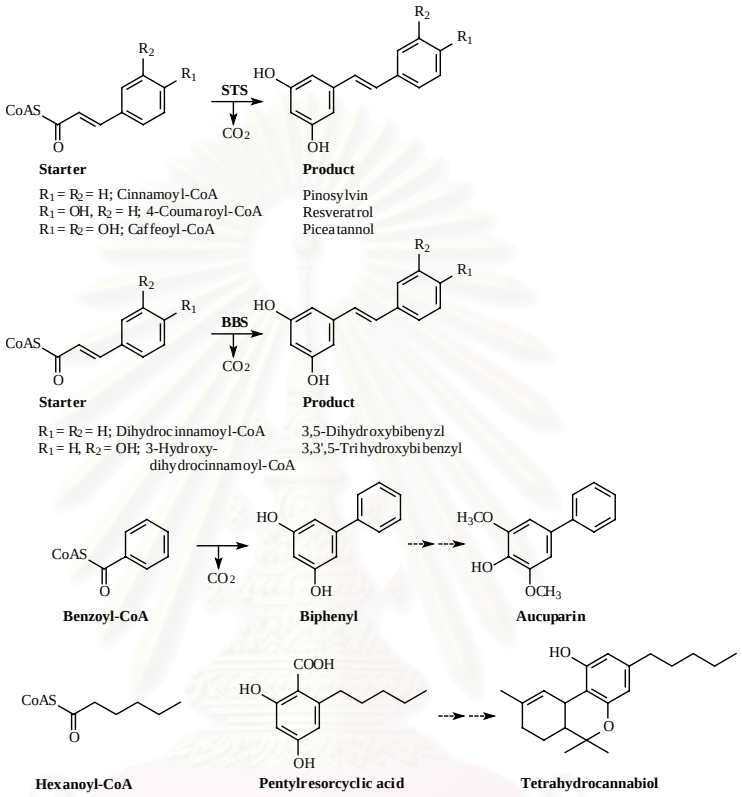


Figure 4 Substrates and products of stilbene (STS) and bibenzyl (BBS) synthase type enzymes. The reactions with benzoyl-CoA and hexanoyl-CoA are proposed based on precursor feeding studies.

The first description of a STS demonstration *in vitro* was published in 1978 (Rupprich and Kindl, 1978). The activities of STSs and BBS have been characterized from several plants, often after heterologous expression of the cloned proteins. Apart from some differences in the substrate specificities, there seems to be no obvious mechanistic differences between STS and BBS, and, therefore, this review will simply use the name STS for both enzymes. The cDNA or genomic sequences show that

STS/BBS enzymes share about 65-75% identity with many CHSs and also with ACS. The STS-type enzymes, thus, belong in the family of CHS-related proteins (Schroder, 1999). In contrast to CHS-derived products, stilbenes and their derivatives appear to be rare in higher plants, in particular in important crop plant cultivars, with the exception of wine and peanut. A large number of interesting bibenzyl-derived molecules (stilbenoids) occur in liverworts (Gorham, 1995). However, stilbenes or stilbene-derived substance may be much more widespread in other plants than previously suspected. The literature from the last few years shows, for example, that several monocotyledonous plants contain stilbenes (Brinker and Seigler, 1991; Harborne *et al*., 1993; Grayer and Harborne, 1994; Powell *et al.*, 1994). One of the important functions of stilbenes in plants is probably their role as phytoalexins in disease resistance. Transgenic STS expression has been used several times to introduce the capacity to produce new phytoalexins in species that otherwise do not synthesize stilbenes (Hain *et al*., 1990; 1993; Fischer *et al*., 1994; 1997; Thomzik *et al*., 1997), and these plants now include the important monocots: rice, barley, and wheat (Strark-Lorenzen *et al*., 1997; Leckband and Lorz, 1998). STS appears to be one of the simplest possibilities for such experiments because the substrates are the same as for the ubiquitous CHS, and there is no need to pay additional attention to their availability in the transgenic plants.

1.2.3.2 Proposals for STS-type reactions

STS-type enzyme activities with compounds other than phenylpropanoid substrates have not yet been described. Precursor feeding experiments, however, suggest several candidates. Two will be mentioned briefly here, mainly because the same or similar substrates are known from CHS-type reactions. One example is the biosynthesis of the biphenyl backbone in the aucuparin-type secondary product that

are phytoalexins in Mountain ash (*Sorbus aucuparia*), apple (*Malus x domestica*), and other plants (Burkhardt *et al*., 1992; Widystuti *et al*., 1992; Kokubun *et al*., 1994; 1995; Kokubun and Harborne, 1995; ; Borejsza *et al*., 1999). In 1980, Sultanbawa proposed that the backbone is synthesized from benzoyl-CoA (or a derivative) *via* three condensation reactions with malonyl-CoA, followed by a STS-type ring closure accompanied by removal of the terminal carboxyl group, as shown in **Figure 4**. The same reaction type, but with an aliphatic starter substrate, has been proposed for the formation of the backbone in tetrahydrocannabinol in Indian hemp (*Cannabis sativa*) **(Figure 4).** It is interesting to note that the proposal suggests a STS-type reaction with retention of the carboxyl group that is removed in the reactions of the known STS, but this needs to be demonstrated *in vitro*.

1.3 Polyketide synthase enzymes with only one or two condensation reactions

1.3.1 Pyrone synthase (2PS)

The cDNA for this protein was originally discovered in *Gerbera hybrida* while screening for CHS-related sequences (Helariutta *et al*., 1995). Three different types of clones have been identified. Functional studies after heterologous expression in *E. coli* have shown that two of the proteins (GCHS1 and GCHS3) are typical CHS whereas the third protein (previously called GCHS2, now renamed to 2PS) is different. It shares 74% identity with the CHS from the same plant and with other CHSs, but was inactive with any of the CHS substrates (Helariutta *et al*., 1995). Enzyme activity has been demonstrated with benzoyl-CoA, but the reaction product has not been identified (Helariutta *et al*., 1995; 1996). A reverse genetic approach with antisense cDNA constructed for the CHS-related protein has been shown to completely suppress the enzyme activity with benzoyl-CoA, but only partially the formation of flavonoids, indicating that the enzyme functions in another pathway (Elomaa *et al*., 1996).

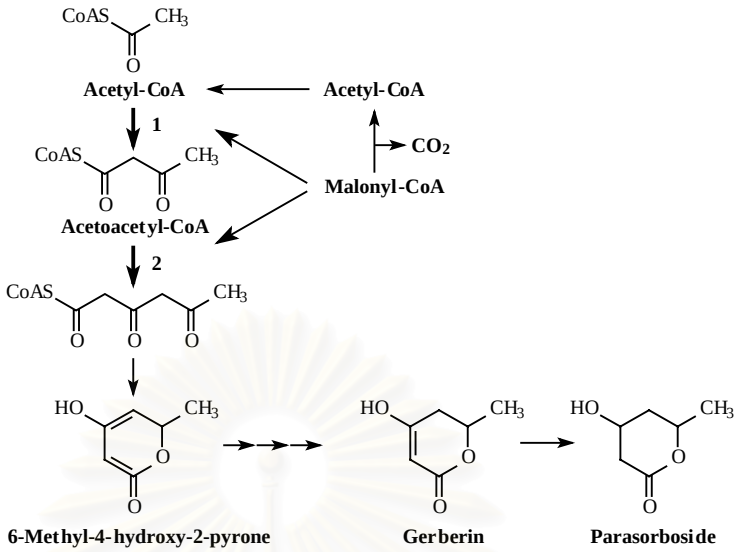


Figure 5 Reactions of the *Gerbera hybrida* pyrone synthase (2PS) in the biosynthesis of the aglycone in gerberine and parasorboside. In absence of the starter substrate, the enzyme decarboxylates malonyl-CoA to acetyl-CoA which is then used as starter.

Subsequent investigation of transgenic GCHS2 antisense plants have revealed that they lack two compounds that are dominant in the control plants: gerberin and parasorboside **(Figure 5)** (Eckermann *et al*., 1998). As the name suggested, gerberin was originally isolated from *Gerbera* (Nagumo *et al*., 1989). Parasorboside was subsequently found in *Gerbera*, but it had been described earlier from cranberry and *Osmunda japonica.* It is also a major constituent in the berries of mountain ash (*Sorbus aucuparia*) (Eckermann *et al*., 1998). The backbone of gerberin and parasorboside had been proposed to be synthesized by a CHS-related enzyme that uses acetyl-CoA as the starter substrate, performs two condensation reactions with malonyl-CoA, and releases 6-methyl-4-hydroxy-2-pyrone (also known as triacetic acid lactone) as the product **(Figure 5).** The prediction has been confirmed recently with the purified enzyme after heterologous expression in *E. coli* (Eckermann *et al*., 1998). This

pyrone synthase is the first proven case for a CHS-related protein that is programmed for only two condensation reactions. It is also the first example that the CHS family contains members that, by using acetyl-CoA as substrate, are functionally more similar to all other types of polyketide synthases than to CHS. It cannot be considered as 'ancient, ancestor-like' CHS, however, because it probably evolved fairly recently from the CHSs in *Gerbera* (Eckermann *et al*., 1998). As already mentioned, another interesting feature of the enzymes is the self-priming of the reaction sequence catalyzed by the decarboxylation of malonyl-CoA to acetyl-CoA.

Like most other CHS-related proteins, the pyrone synthase is not substrate-specific, and it readily accepts other small hydrophobic substrates. The product from benzoyl-CoA has been identified as phenylpyrone (Eckermann *et al*., 1998). This compound or its derivatives are not known from *Gerbera*, but it contains the aglycone backbone of the psilotins, which are considered as chemical markers of the Psilitaceae (far relatives of ferns) (McInnes *et al*., 1965; Takahashi *et al*.,1990). However, precursor feeding studies *in vivo* suggest that the psilotins are not synthesized from benzoyl-CoA, but from phenylpropanoid precursors, with the pyrone ring formed after one condensation reaction (Leete *et al*., 1982).

1.4 Three condensations with other products

1.4.1 Modification of intermediates by reduction

The CHS reaction does not involve a modification of reaction intermediates prior to the ring closure. The only demonstrated example is the formation of 6'-deoxychalcone, the precursor for a large number of biologically and medically important secondary products (Bingham *et al*., 1998; Dixon, 1999; Dixon and Steele, 1999). The modification is performed by an enzyme that uses NADPH to reduce a specific carbonyl group of the intermediate either after the second or third condensation

reactions. Available evidence (Heller and Forkmann, 1994) indicates that the reduction occurs at the polyketide level, prior to the formation of the new aromatic ring **(Figure 6).** Therefore, the name polyketide reductase is still often used although it is to some extent misleading (Heller and Forkmann, 1994). The enzyme activities have been characterized from various plants (Ayabe *et al*., 1988, Hakamatsuka *et al*., 1988; Welle and Grisebach, 1989; Harano *et al*., 1993).

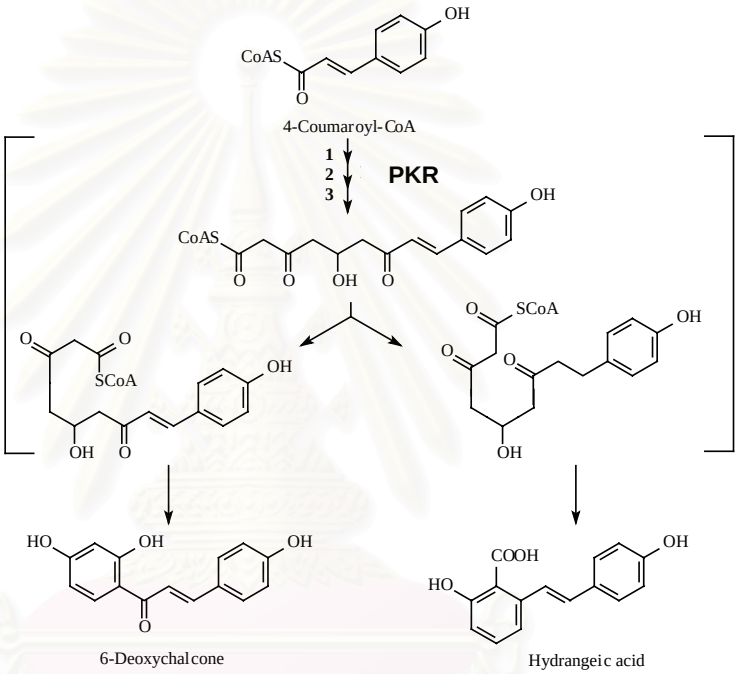


Figure 6 Modification of reaction intermediates. Biosynthesis of 6'-deoxychalcone (CHS-type) and reaction proposed for the biosynthesis of hydrangeic acid (STS-type). PKR, polyketide reductase which reduces with NADPH a specific carbonyl group either at the tri- or tetraketide intermediate stage prior to the ring closure.

Sequences for cDNAs and a functional identification after heterologous enzyme expression were first described for *Glycine max* (Welle *et al*., 1991), and related sequences are known from *Medicago sativa* (Balance and Dixon, 1995; Sallaud *et al.,* 1995a; 1995b), *Sesbania rostrata* (Goormachtig *et al.,* 1995; 1999), *Pueraria lobata*

(Sankawa *et al*, 1995), *Glycyrrhiza echinata* (Akashi *et al*., 1996), *Glycyrrhiza glabra* (Hayashi *et al*., 1996), and *Fragraria ananassa* (Manning, 1998). These proteins belong to the aldo/keto-reductase superfamily (Jez *et al*., 1997). The members share three highly conserved protein sequence elements, but the diversity of functions within the family is at least as large as in the family of CHS-related proteins. A functional identification should, therefore, be mandatory, but this has not been done with all sequences cloned from plants. Interestingly, the transgenic expression of the *Medicago sativa* PKR in petunia has led to a new yellow flower color by redirection of the flavonoid biosynthesis (Davies *et al*., 1998).

With STS, a comparable interaction with a reductase has not yet been demonstrated *in vitro*. However, the formation of the backbone in stilbenoids is likely to involve such a co-action, and a reduction of the same carbonyl group as in the CHS/PKR interaction can be postulated. **Figure 6** shows an example of the reactions predicted for the formation of hydrangeic acid in the garden hortensia (*Hydrangea macrophylla*) (Gorham, 1995). The reaction has not yet been demonstrated *in vitro*, nor has its gene been cloned. It is interesting to note that the product of the reaction is a stilbenecarboxylic acid, *i.e*. it retains the carboxyl group that is usually removed during the known STS reactions. Early precursor feeding experiments strongly suggest that stilbenecarboxylic acids and their many derivatives in liverworts, ferns and some higher plants are derived from phenylpropanoid starters and STS-type reaction (Gorham, 1995).

1.4.2 Derailment product with three condensation

A recent publication showed that the family of CHS-related proteins still has some surprises in store with respect to possible products (Akiyama *et al*., 1999). The work has described the characterization of cDNA clones for CHS-related proteins from

Hydrangea macrophylla var. *thunbergii*. One of the predicted proteins has been shown to share about 90% identity with many CHSs, but the others have only 70-74% identity with any other enzymes from the protein family. Heterologous functional expression in *E. coli* has shown that the first protein is a CHS, but the second protein synthesizes two products from 4-coumaroyl-CoA that are not those expected from CHS (naringenin chalcone), STS (resveratrol), or a stilbenecarboxylic acid synthase (like the STS-type postulated in hydrangeic acid biosynthesis **(Figure 6)** (Akiyama *et al*., 1999). Identification by NMR shows that the products are two lactones: bisnoryangonin and 4-coumaroyltriacetic acid lactone (CTAL) **(Figure 7)** (Akiyama *et al*., 1999).

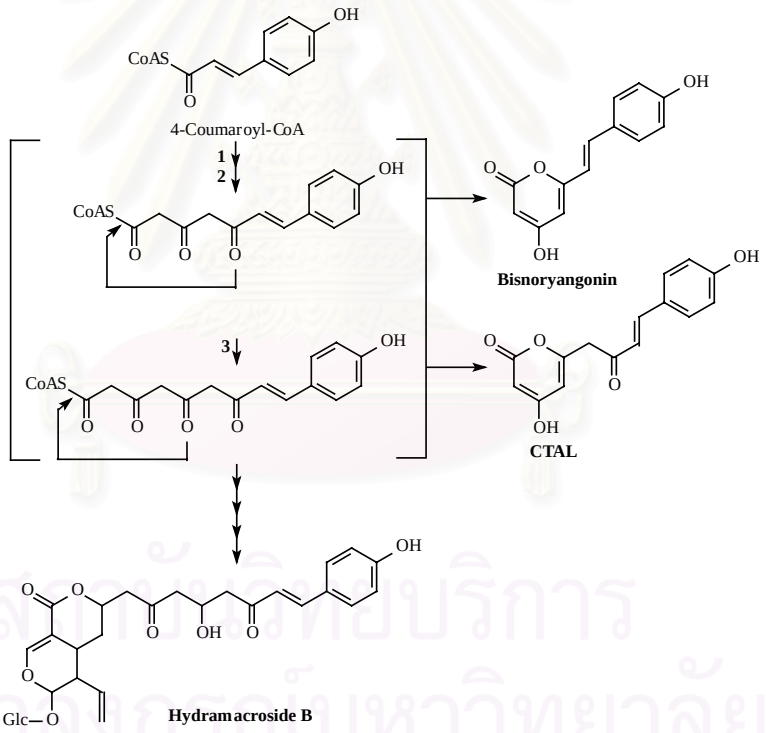


Figure 7 Reaction of coumaroyltriacetic acid synthase (CTAS) and the derailment products detected in vital: bisnoryangonin and coumaroyltriacetic lactone (CTAL). It is proposed that the enzyme does not perform a cyclization, and that the linear tetraketide is precursor of hydramacroside B.

Bisnoryangonin is the product of the styrylpyrone synthase in ferns, but it is also well-known as a derailment product of CHS reactions under non-optimal assay conditions. CTAL is the corresponding lactone product after three condensation reactions. Bisnoryangonin, CTAL, or their derivatives are seemingly not known in *Hydrangea*, but the linear tetraketide (coumaroyltriacetic acid) could be a precursor of hydramacroside B **(Figure 7),** a secondary plant product in *Hydrangea* that contains a secologanin moiety (Yoshikawa *et al*., 1994). Therefore, it has been proposed that coumaroyltriacetic acid synthase (CTAS) activity is a new function in the family of CHS-related proteins which perform three condensations but no subsequent cyclization reaction. The CTAL detected *in vitro* is interpreted as a derailment product that is released in the absence of the proteins that can carry out further reactions (Akiyama *et al*., 1999).

1.5 The portulations of CHS-related enzyme

1.5.1 One condensation reaction : Benzalacetone synthase (BAS)

The characteristic aroma of raspberries is caused by 4-hydroxyphenylbutan-2-one (pHPB, "raspberry ketone") (Larsen *et al*., 1991). Two enzymes are involved in its biosynthesis as investigated in raspberry fruits and tissue cultures (Borejsza-Wysocki and Hrazdina, 1994). The first, benzalacetone synthase (BAS), uses a phenylpropanoid starter CoA-ester, performs one condensation reaction, and a decarboxylated product is detected *in vitro* (Borejsza-Wysocki and Hrazdina, 1996). The second enzyme reduces the double bond in the propenoyl moiety with NADPH to form the aroma component **(Figure 8).**

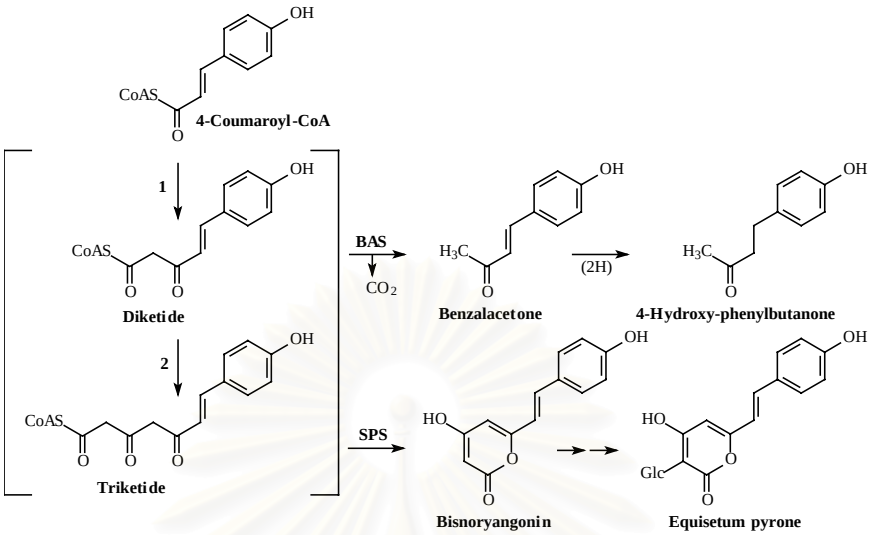


Figure 8 Enzymes with only one or two condensation reactions. BAS, benzalacetone synthase; SPS, styrylpyrone synthase. Only the reactions with 4-coumaroyl-CoA are shown; the enzymes also have high activities with caffeoyl-CoA.

BAS has been purified and characterized from raspberry fruits. All of its properties and the differential induction argue for a proposal that BAS and CHS are different enzymes (Borejsza-Wysocki and Hrazdina, 1996). This, however, needs unambiguous proof by cloning and heterologous expression because it is known that purified CHS can produce benzalacetone *in vitro* as a by product (Hrazdina *et al*., 1976; Saleh *et al*., 1978). BAS should be present in other plants as well, because benzalacetone or its glycoside (Pabst *et al*., 1990) have also been found in several other plants (Schroder, 1999). Precursor feeding studies suggest that enzymes performing one condensation reaction to a diketide intermediate may be involved in the formation of other complex substances of considerable interest, e.g. gingerol, curcumin, and other secondary products (Schroder, 1997).

1.5.2 Two condensation reactions : Styrylpyrone synthase (SPS)

SPS performs two condensations of malonyl-CoA to form styrylpyrone. Styrylpyrones occur in Pteridophytes and in Angiosperms, and they are common constituents in fungi, mainly in the Basidiomycetes (Wat and Towers, 1979). Precursor feeding experiments suggested that they can be synthesized from the precursors used by CHS, but with only two condensation reactions that are followed by ring closure of the triketide intermediate to the styrylpyrone **(Figure 8)** (Perrin and Towers, 1973; Wat and Towers, 1979).

An enzyme from the horsetail fern (*Equisetum arvense*) has been the only example where the predicted enzyme activities have been demonstrated *in vitro* (Beckert *et al*., 1997; Herderich *et al*., 1997). The partially purified styrylpyrone synthase (SPS) accepted both 4-coumaroyl-CoA and caffeoyl-CoA, with bisnoryangonin and hispidin as products, respectively. The tissue used in the experiments (gametophytes) does not express CHS activity, and therefore, SPS must be a different enzyme. There is no clear evidence so far that the SPS belongs to the family of CHS-related proteins. *E. arvense* is an interesting plant system to investigate the differential expression and relation of SPS and CHS, because it shows a developmental switch: gametophytes and rhizomes accumulate styrylpyrones but no flavonoids as major phenolic constituent, while green sprouts (sporophytes) contain various flavonoid glycosides, but no styrylpyrones (Veit *et al*., 1993; Beckert *et al*., 1997).

2. Three dimensional structure of plant polyketide synthases

In 1999, Joseph Noel and co-worker has determined the crystal structures of CHS2 alone and complexed with substrate and product analoques (Ferrer *et al*., 1999). These structures have been the first ever of any PKS component. They have found that

CHS2 is a homodimer (MW 84 kDa) whose overall fold resembles that of β-ketoacyl synthase (KAS) II of *E. coli* (involved in fatty acid biosynthesis) as well as the thiolase from *Saccharomyces cerevisiae* (Schroder, 2000). The structure also suggested that four residues, Cys 164, Phe 215, His 303, and Asn 336 which are conserved among all known CHS-related enzymes, are the primary participants in catalysis. The role of these residues have been pinpointed by structural and functional characterisation of 16 CHSs point mutants-the Cys serves as an active-site nucleophile, the His and Asn are responsible for malonyl-CoA decarboxylation, and the Phe may help to orient substrates during chain elongation (Jez *et al*., 2000). The His has also been shown to stabilise the active site nucleophile through formation of thiolate-imidazolium ion pair (Jez *et al*., 2000). The crystal structure has further revealed three interconnected cavities that intersect with the four catalytic residues and form the active site architecture of the CHS: a CoA-binding tunnel, a coumaroyl-binding pocket and a cyclisation pocket (Ferrer *et al*.,1999). This division of the active site into discrete pockets provides a structural basis for the ability of the CHSs to orchestrate the multiple reactions of chalcone synthesis.

Noel and co-workers have also determined the crystal structure of a related enzyme, 2-pyrone synthase (2-PS), which forms a triketide methylpyrone from acetyl-CoA and two units of malonyl-CoA (Jez *et al*., 2000). They have found that while 2-PS and CHS share a common fold, the same catalytic residues and a similar CoA binding site, the active site cavity of 2-PS is smaller than that of CHS. Armed with this information, they have shown functional conversion of the CHS into a 2-PS by reducing the size of its initiation/elongation cavity. This experiment indicates that cavity volume influences both the choice of starter molecule and the final length of the polyketide (Jez *et al*., 2000a).

Although the sequences of CHS-related enzyme have no obvious similarity to the type I and II PKSs (Schroder, 1999), the homology of active site residues as well as shared structural features (Jez *et al*., 2000) suggest that observations about CHSs may well have relevance to the productive reengineering of bacterial PKSs.

3. Genetic evolution of plant polyketide synthases

3.1 Functional diversification of CHS in plants.

CHS is the polyketide synthase that is ubiquitous in plants. It is thought that it appeared early in land plants, i.e. in Charophyceae or in simple Bryophytes (Kubitzki, 1987; Markham, 1990; Stafford, 1991; Koes *et al*, 1994). It has been argued for the present-day STS (Trof *et al*., 1994) and for the 2PS (Eckermann *et al*., 1998) that they evolved fairly recently from CHS by gene duplication and mutation. The main lines of reasoning can also be applied to acridone synthase and CTAS; these enzymes are also of limited distribution and share about 70% identity with each other and with CHS. It has been suggested that CHSs are the basic for the known functional diversifications (Schroder, 2000). However, with sequences for predicted plant proteins of unknown functions that share only about 40% or less identity with typical CHS. Such cases have been described from several plants (e.g. *Nicotiana sylvestris*, *Arabidopsis thaliana*, *Oryza sativa*, *Pinus radiata*, and *Brassica napus*) (Atanassov *et al*., 1998), and it may be necessary to consider other possibilities. Unfortunately, up to now, no CHS or CHS-related sequences are available from the postulated early land plants and their possible ancestors, and little is known about products of secondary metabolism that could have been synthesized by CHS-related proteins in such plants.

3.2 Genetic relationships of plant polyketide synthase with other polyketide synthases

CHS-related proteins have been identified in serveral bacteria. The examples include such diverse organisms as *Pseudomonas fluorescens* (Bangera and Thomashow, 1996), *Streptomyces griseus* (Ueda *et al*., 1995), *Bacillus subtilis* (Capuano *et al*., 1996), *Mycobacterium tuberculosis* (accession no. Z81011, Z85952, Z95617), *Amycolatopsis orientalis* (accession no. AJ223998), and *Rhodospirillum centenum* (Jiang *et al*., 1999). The predicted polypeptides constitute a rather heterologous group that shares 25-30% identity with plant proteins, and certain typical elements are clearly conserved (e.g. a cystein in the position expected for the active site of the condensation reactions). The functions of the proteins are unknown in most cases, but the obvious relation suggests that the evolution of the basic structural and functional properties of these condensing enzyme predated the emergence of land plants.

A recent report has identified the function of one of the bacterial proteins (Funa *et al*., 1999). *S. griseus* contains a gene for a protein that was previously shown to be sufficient for the production in the foreign host *E. coli* of an unknown red-brown pigment (Ueda *et al*., 1995). The new results are surprising in several aspects. After correction of the nucleotide sequence, the joining of two small open reading frames has yielded a polypeptide (RppA, 40 kDa) that shares about 30% identity with CHSs. An His-tagged protein has been expressed in *E. coli*, and the function was tested with the purified enzyme. The assays have shown that the protein uses malonyl-CoA as the starter substrate, performs condensation reactions with malonyl-CoA, and releases the pentaketide to form 1,3,6,8-tetrahydroxynaphthalene (THN). The final ring closure to THN involved the removal of the carboxyl group of the malonyl-CoA used as the starter substrate **(Figure 9).**

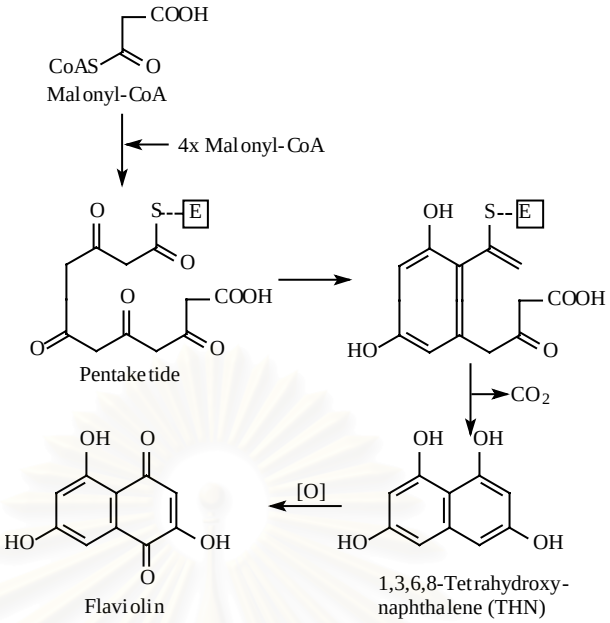


Figure 9. Reaction of the CHS-related enzyme from *Streptomyces griseus*.

This appears to be the first functional identification of a CHS-related bacterial protein. The use of malonyl-CoA as starter and the performance of four condensations are novel and unexpected properties for CHS-related enzymes. Indeed, previous findings with a polyketide synthase cloned from the fungus *Collectotrichum lagenarium* has suggested that THN is synthesized by proteins that have no similarity with the family of CHS-related proteins (Fujii *et al*., 1999). The example of *S. griseus* suggests the possibility that diversity of the biosynthetic capacities in the protein family may be much larger than previously expected, at least in bacteria. It will be interesting to see whether the same applies to plants, or whether the activities in plants represent only a subset of the possibilities realized in bacteria (Schroder, 2000).

4. Naphthoquinones and anthraquinones from polyketide biosynthetic pathways

One of the remarkable features of quinone biosynthesis in higher plants is that they are derived from a variety of different precursors and by different pathways (Bentley and Campbell, 1974; Bentley, 1975; Leistner, 1980).

4.1 Acetate/malonate derived quinone

The most common biosynthetic pathways leading to quinones is the polyacetate or acetate-polymalonate pathway leading to compounds called acetogenins. These types of compounds often exhibit a characteristic substitution pattern reflecting their biosynthesis from acetyl-CoA and malonyl-CoA. Acetyl-CoA is the starting molecule for most polyketides. Linear Claisen condensation with several acetyl residues derived from malonyl-CoA leads, with concomitant loss of CO_2 ,to the polyketide (acetogenin) structures [-$(CH_2$-$CO)_n$-]. Direct condensation (without reduction) and cyclization give various aromatic structures. The aromatic polyketide biosynthesis has to be distinguished from the fatty acid biosynthetic pathway, where the polyketides undergo reduction and dehydration to form aliphatic hydrocarbons. The enzyme-bound postulated polyketide structures are very reactive, and must be temporarily stabilized by hydrogen bonding or chelation on the enzyme surface. Different cyclization mechanisms are probably guided by the formation of different enzymes that catalyze the formation of different phenolics (Leistner, 1981). The aromatic ring systems are either formed by aldol condensation, i.e. reaction of a carbonyl group with an CH_2-group, or Claisen condensation, i.e. reaction of the thioester group with an CH_2-group **Figure 10**. A hypothetic polyketomethylene compound is postulated as an intermediate between the CoA esters and the phenols or quinones. Such a compound, which would be unstable, has never been isolated from a plant source. In an attempt to obtain clearer insight into the nature of such an

intermediate, Franck *et al*. (1974) synthesized a tetraketone which cyclized under mild conditions. A biomimetic synthesis of emodin (1,6,8-trihydroxy-3-methyl-9,10-anthraquinone) from a polyketomethylene precursor has also been accomplished (Leistner, 1981).

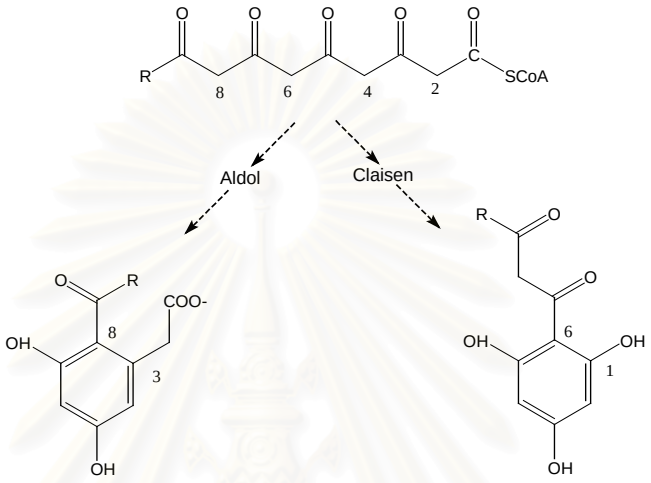


Figure 10 Two examples of polyketide cyclization by Aldol condensation or Claisen condensation

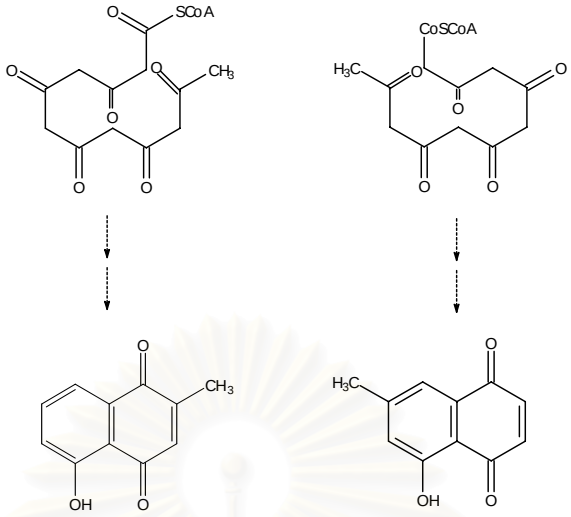
The variation among polyketide-derived phenolic structures derives from (I) the chain initiating unit, e.g. use of malonyl-CoA, propionyl-CoA, butyryl-CoA, or hydroxycinnamoyl-CoA (e.g. 4-coumaroyl-CoA in flavonoid biosynthesis) instead of acetyl-CoA, (ii) the number of malonyl-CoAs involved, (iii) further modification by oxidation or reduction or (iv) introduction of substituents and conjugations (e.g. O- and C- glycosylation) (Torssell, 1997).

The majority of quinones found in plants are relatively simple benzoquinones, naphthoquinones and anthraquinones. Although they derive from different biosynthetic routes, their structures contain phenolic moieties (Thomson, 1971). Most of the biosynthetic sequences in higher plants have been established from isotope incorporation studies and still await enzymatic proof. However, some of the key

enzymes leading to the naphthoate intermediate are well described from bacterial sources. Recent advances of plant cell culture techniques have had an important impact on current studies about quinone biosynthesis (Van den Berg and Labadie, 1989).

4.2 Naphthoquinones

The polyketide route to the naphthalene nucleus and consequently of naphthoquinones, is well established in microorganisms but it appears to be quite rare in higher plants, where other biosynthetic pathways are preferred. Plumbagin (2-methyljuglone) and 7-methyljuglone are the first naphthoquinones in higher plants showed to be formed according to the polyketide pathway (Manitto and Sammes, 1981). It has been demonstrated that the two naphthoquinones, synthesized by plants of the *Drosera* and *Plumbago* genera, arise from a hexaketide **(Figure 11)** rather than the shikimic acid route as occurs for juglone and menadione (Durand and Zenk, 1971). Feeding experiments with young shoots of *P. europaea*, however, showed that neither [7-^{14}C]-shikimate, [$^{14}CH_3$]-L-methionine, [β-^{14}C]-DL-tyrosine, [ring-1-^{14}C]-DL-phenylalanine, nor [5-^{14}C]-DL-mevalonic acid was incorporated into plumbagin to a significant extent. In contrast, [1-^{14}C], [2-^{14}C]-acetate and [2-^{14}C]-malonate labelled this naphthoquinone heavily. This fact suggests that plumbagin is formed by the well known polyketide pathway. In addition, it has been confirmed by the chemical degradation of the labelled plumbagin, either from [1-^{14}C]-or [2-^{14}C]-acetate, that the polyketide pathway is the route used in the biosynthesis of plumbagin. Similar results have been obtained by labelled acetate feeding to different Droseraceae species and degradation of labelled plumbagin. The co-occuring 7-methyl-juglone in *Drosera* plants is also formed *via* the polyketide route (Durand and Zenk, 1971)



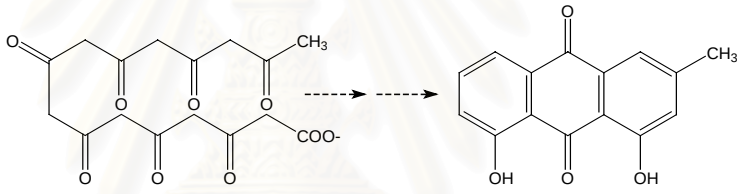
Plumbagin **7-Methyljuglone**

Figure 11 The proposed biosynthetic pathways of some polyketide-derived naphthoquinones.

4.3 Anthraquinones

Anthraquinones are the largest group of quinones. They have been used as purgatives e.g. emodin from *Rheum*, *Rumex*, or *Rhamnus* spp. Anthraquinones are widely spreaded in lower and higher plants. They are present as glycosides in young plants (Torssell, 1997). It is difficult to systematize the biosynthesis of quinones because it reveals such a diversified structure (Thomson, 1971). In the biosynthesis of polyketide anthraquinones, one acetyl CoA is extended by seven malonyl CoA units. Acetyl CoA is produced through the glycolytic pathway, or is synthesised by plants from acetate and coenzyme A by the cytoplasmic enzyme acetyl CoA synthetase. Carboxylation of acetyl CoA by acetyl CoA carboxylase is the major source of malonyl CoA in most organisms. In plants, however, malonyl CoA can also be formed from malonate and coenzyme A, catalysed by malonate thiokinase (van den Berg and Labadie, 1989). Addition of the basic precursors acetate and malonate to suspension

cultures of *Rhamnus purshiana* has shown that the production of anthraquinones and anthrones was significantly stimulated by malonate, while addition of acetate had an inhibitory effect under similar conductions (Van den Berg *et al*., 1988c). It has been suggested that cells of *R. purshiana* and maybe even intact plants are incapable of converting acetyl CoA into malonyl CoA. Malonate thiokinase seems to be of major importance in *Phaseolus vulgaris*, where oxidative decarboxylation of oxaloacetate gives rise to free malonate and acetyl CoA carboxylase is apparently absent (van den Berg and Labadie, 1989). A typical folding mechanism involving aldol condensations of a hypothetical polyketomethylene compound can lead to the anthraquinone chrysophanol in *R. frangula* as shown in **Figure 12** (Leistner, 1971).



Chrysophanol

Figure 12 The proposed biosynthetic pathway of chrysophanol, a polyketide-derived anthraquinone.

5. Plants producing naphthoquinones and anthraquinones derived from polyketide pathway

Quinones are found in bacteria, fungi, lichens, gymnosperms and angiosperms (Thomson, 1971; Leistner, 1981). The majority of quinones found in plants are relatively simple benzoquinones, naphthoquinones or anthraquinones. In higher plants, anthraquinones are found in Leguminosae (in particular *Cassia* sp.), Rhamnaceae, Polygonaceae (in Particular *Rheum* and *Rumex* sp.), Liliaceae (Aloe), Bignoniaceae, Verbenaceae (*e.g*. *Tectona* sp.) and Scrophulariaceae (*Digitalis* sp.). Most naphthoquinones occur in Bignoniaceae (*e.g*. *Catalpa* and *Tabebulia*), Verbenaceae,

Juglandaceae, Plumbaginaceae, Boraginaceae (*e.g*. *Lithospermum* sp.), Proteaceae, Lythraceae, Balsaminaceae, Sterculiaceae, Ulmaceae, Ebenaceae and Droseraceae. (van den Berg and Labadie, 1989).

CHAPTER III

MATERIALS AND METHODS

1. Organisms

1.1 Plant materials

Cassia alata Linn.*, Plumbago indica* Linn. and *Rhuem tataricum* Linn., were grown in the greenhouse of the Institute of Plant Biochemistry, Halle, Germany, under the condition of 24±2 °C with 18 h of light and 50% humidity. Young plant material was used in this study.

1.2 Microorganisms and composition of media

The following microorganisms were used during this work.

***E. coli* DH5α** (Clontech)

Genotype : F^- *ϕ80dlacZΔM15 Δ(lacZYA-argF)U169 deoR recA1 endA1hsdR17* (r_K^-,m_K^+) *phoA supE44λ⁻thi-1 gyrA96 relA1*

***E. coli* XL-Blue MRF'** (Stratagene; part of ZAP-cDNA Synthesis Kit)

Genotype : Δ(*mcrA*)183 Δ(*mcrCB-hsdSMR-mrr*)173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F'*proAB* $lac^qZ\Delta M15$] Tn*10* (Tet^r)]

***E.coli* SOLR™** (Stratagene; part of ZAP-cDNA Synthesis Kit)

Genotype : $e14^-$($McrA^-$)Δ(*mcrCB-hsdSMR-mrr*)171*sbcC recB recJ uvrC* : :Tn5(Kanr) *lac gyrA96 relA1 thi-1 endA1* λ^R [F'*proAB* $lac^qZ\Delta M15$]Su^-(nonsuppressing)

***E. coli* BL21 (DE3) strain** (Novagen)

Genotype : F^-*ompT hsdSB* ($r_B^-m_B^-$) *gal dcm* (DE3)

The following media were used for the growth of microorganisms.

LB broth (Sambrook *et al*., 1989)	1.0 % (w/v) tryptone
	0.5 % (w/v) yeast extract
	1.0 % (w/v) NaCl
	pH 7.5 (adjusted with 1 M NaOH)
LB agar (Sambrook *et al*., 1989)	LB broth with 1.5% (w/v) agar
NZY agar (Sambrook *et al*., 1989)	1.0 % (w/v) NZ-amine
	0.5 % (w/v) yeast extract
	0.5 % (w/v) NaCl
	0.2 % (w/v) $MgSO_4.7H_2O$
	1.5 % (w/v) agar
	pH 7.5 (adjusted with 1 M NaOH)
NZY top agar (Sambrook *et al*., 1989)	
	1.0 % (w/v) NZ-amine
	0.5 % (w/v) yeast extract
	0.5 % (w/v) NaCl
	0.2 % (w/v) $MgSO_4.7H_2O$
	0.7 % (w/v) agarose
	pH 7.5 (adjusted with 1 M NaOH)
SOB medium (Sambrook *et al*., 1989)	
	2.0 % (w/v) tryptone
	0.5 % (w/v) yeast extract
	0.05 % (w/v) NaCl
	2.5 mM KCl, pH 7.0

pH Value of SOB medium was adjusted to 7.0. The medium was subsequently autoclaved and supplemented with sterile $MgCl_2$ and $MgSO_4$ (each 10 mM). SOC medium (Sambrook *et al*., 1989) is SOB medium supplemented with 20 mM glucose (sterile added after autoclaving, together with magnesium salts).

2. Chemicals

All chemicals used were standard commercial products of analytical grade from the companies: Biomol, Boehringer Mannheim, Fluka, Boehringer Ingelheim, Merck, Roth and Sigma.

All starter-CoA ester (Sigma)

[2-^{14}C] Malonyl-CoA (American Radiolabeled Chemical, USA)

Naringenin (Sigma)

Resveratrol (Sigma)

3. Enzymes, kits and special chemicals used in the molecular biology procedures

3.1 Enzymes

The following enzymes were used in this molecular biology work:

Taq DNA-Polymerase (Promega)

T4-DNA Ligase (Promega)

Pfu DNA-Polymerase (Promega)

Klenow fragment (Life Technologies)

M-MLV Reverse Transcriptase (Life Technologies)

Rnase A (Sigma)

3.2 Molecular biology kits:

cDNA Synthesis:	cDNA Synthesis System (Life Technologies)
mRNA Purification:	Oligotex^{TM} mRNA Midi Kit (Qiagen)
DNA Labeling:	RadPrime DNA Labeling System (Life Technologies)
cDNA Library:	cDNA Synthesis Kit, ZAP-cDNA Synthesis Kit and ZAP-cDNA Gigapack III Gold Cloning Kit (Stratagene)
DNA purification:	QIAquick Nucleotide Removal Kit (Qiagen) QIAquick PCR Purification Kit (Qiagen) Qiaquick Gel Extraction Kit (Qiagen) Plasmid Midi Kit (Qiagen) QIAprep Spin Miniprep Kit (Qiagen)

3.3 Special chemicals used in the molecular biology procedures

Blue-white selection:	5-Bromo-4-chloro-3-indolyl-β-D-galctopyranoside (X-gal) (Biomol)
Protein expression:	isopropyl-n-D-thiogalactoside (IPTG) (Biomol)
Protein purification:	TALON Metal Affinity Resin (Clontech) PD-10 Columns (Amersham Pharmacia)
Membranes and filters:	Nitrocellulose Filter 0.45 μm, ϕ132 mm (Sartorius) Nitrocellulose Filter Protran 0.45 μm, ϕ 82 mm (Scheicher & Schuell) Nylon Membran Hybond-XL (Amersham Pharmacia)

Antibiotics: Ampicillin (Sigma)

Kanamycin (Sigma)

3.4 Nucleic acids

3.4.1 DNA

DNA Length standard:

100 bp DNA-Ladder (Life Technologies)

1 kb DNA Ladder (New England Biolabs)

Nucleotides:

deoxynucleotides (dATP, dCTP, dGTP, dTTP) (Life Technologies)

Oligonucleotides:

primers for PCR synthesized at MWG-Biotech AG (Germany)

3.4.2 Plasmids:

pGEM-T Easy Vector System (Promega)

This vector was used to subclone the *Taq* DNA-polymerase amplified DNA fragments. This linear 3015 bp vector has a 3' terminal thymidine overhang. *Taq*-polymerase in the PCR reaction unspecifically adds a 3'-deoxyadenosine overhang to the amplified DNA fragments, which enables a straightforward ligation into pGEM-T Easy vector. The selection of transformed bacteria is ensured by the presence of an additional gene in this plasmid that allows ampicillin resistance.

pBluescript SK(-) (Stratagene)

This 2958 phagemid occurs as the product of *in vivo* excision from a λ-phage. This vector enables a blue-white screening through the complementation of β-galactosidase enzyme and carries a selection marker gene for ampicilin resistance.

pET-14b (Novagen)

pET-14b (4671 bp; selection marker: ampicillin resistance) was used as vector for the expression of recombinant proteins in *E. coli*.

4. Equipment

The following general laboratory equipment was used.

4.1 Centrifuges

Sorvall RC-5B, RC 26 plus, RC 28S with rotors SS34, GSA and GS3 (DuPont, USA)

Centrifuge 5810 (R) (Eppendorf, Hamburg)

Centrifuge Sigma 4K10 (Sigma, Osterode)

Bench-top centrifuge 5415 and 5415 D (Eppendorf, Hamburg)

4.2 Electrophoresis

Vertical gel electrophoresis apparatus (BioRad, Hercules, CA, USA)

Horizontal gel electrophoresis apparatus (Biometra, Göttingen)

Power supply Phero-stap 500 (Biotec Fischer, Reiskirchen)

Gel Doc 1000 (BioRad, Hercules, CA, USA)

4.3 PCR

Thermal Cycle GeneAmp PCR 9700 (PE Applied Biosystems,USA)

Thermal Cycle 480 (Perkin Elmer, USA)

4.4 Hybridization

Hybridization oven GFL 7601

(Gesellschaft für Labortechnik, Burgwedel)

Cross-Linker UV Stratalinker (Stratagene, La Jolla, CA, USA)

Storage Phospho Screens (Kodak)

Phospho Imager Storm 860 (Molecular Dynamics, USA)

Autoradiography Film Developer 35 Compact (Protec, Oberstenfeld)

Autoradiography Films BioMax MS (Kodax, Rochester, NY, USA)

4.5 Others

Thermoshaker Eppendorf 5436 (Eppendorf, Hamburg)

Vortex Genie 2 (Scientific Industries Inc, USA)

Sonicator Sonorex RX-100 (Bandelin, Berlin)

Balance BP 3100S (Sartorius, Göttingen)

Water Baths T and MT (Lauda, Königshofen)

Water Baths GFL 1083 (Gesellschaft für Labortechnik, Burgwedel)

Cooling Water Bath F12 (Julabo, Seelbach)

pH-Meter pH 526 (WTW, Weilheim)

Photometer Ultrospec 3000 (Pharmacia, Uppsala, Sweden)

Ultraturrax T25 (Janke & Kunkel, IKA-Labortechnik, Staufen)

Membrane Vaccum Pump MZ2C (Vaccubrand, Wertheim)

5. Isolation of nucleic acids

5.1 Isolation of total RNA

Extraction buffer used was composed of 4 M guanidinethiocyanate, 100 mM Tris-HCl, pH 8.0, 25 mM sodium citrate and 0.5 % (w/v) N-lauryl sarcosine. Total RNA isolation was carried out by the method of Salzman, *et al*., (1999). One- gram (fresh weight) of *C*. *alata, P. indica* and *R. tataricum* roots was each frozen in liquid nitrogen and ground into a fine powder in a mortar and pestle. The powdered tissue was transferred into a SS34 centrifuge tube containing 10 ml extraction buffer and mixed strongly by shaking. Ten-ml of chloroform : isoamyl alcohol (24:1) was then added to the tissue solution, mixed and shaked vigorously for 10-20 min at room temperature. The insoluble debris was removed by centrifugation at 10,000 g for 10 min at 4 ^{o}C in a fixed rotor (F28-50). The aqueous phase was transferred into a new SS34 tube, 10 ml of chloroform : isoamyl alcohol (24:1) was then added and mixed by vortexing for 30 sec. The chloroform mixture was separated by 10,000 g centrifugation for 10 min at 4 ^{o}C, and the aqueous phase was transferred into a new SS34 tube. RNAs were precipitated by adding 0.1 volume of 5 M NaCl and 2 volume of ethanol. After being incubated at -20 ^{o}C for at least 3 h or overnight, RNAs were pelleted by centrifugation (10,000 g, 10 min at 4 ^{o}C), and dissolved in 10 ml of steriled water. The RNAs solution was added with an equal volume of Tris-saturated phenol (pH 8.0) : chloroform : isoamyl alcohol (25:24:1), followed by shaking for 10 min at room temperature, then centrifuged at 10,000 g for 10 min at 4 ^{o}C. The upper phase was carefully transferred to a clean tube, avoiding any interphase material. To the supernatant, 2x volume of absolute ethanol and 0.1x volume of 5 M NaCl were added and the RNA was let to the precipitate for at least 3 h or overnight. The solution was then centrifuged at 10,000 g for 10 min at 4 ^{o}C, and the pellet was resuspended in 500 μl of sterile distilled water. The resulting total RNA was used to prepare single stranded cDNA as described in section 6.1.

5.2 Isolation of poly $(A)^+$ RNA (mRNA)

Poly A^+RNA (mRNA) was obtained from total RNA using Oligotex™ kit (Qiagen). The supplier's protocol was precisely followed. The principle of this method is based on the fact that most eukaryotic mRNAs (and some viral RNAs) end in a poly-A tail of 20-250 adenosine nucleotides. In contrast, rRNAs and tRNAs, which account for over 95% of cellular RNAs, are not polyadenylated.

Under conditions that include high-salt concentration, mRNA was purified by hybridizing the poly-A tail to a dT oligomer coupled to a solid-phase matrix. rRNA and tRNA species do not bind to the oligo-dT and were eluted. The mRNA was released from the matrix by lowering the ionic strength and directly used for first strand cDNA synthesis for making the cDNA library as described in section 9.1.

5.3 DNA isolation by agarose gels

DNA bands were separated by electrophoresis in 1x TAE containing 0.8 % (w/v) agarose gel plus ethidium bromide. After the desired DNA band was excised using a clean, sterile razor blade, the agarose slice was transfered to a microcentrifuge tube. The QIAquick Gel Extraction Kit (QIAGEN) was used to isolate the DNA from agarose gels. The principle of this procedure is the selective binding of DNA to a silica-gel membrane. DNA adsorbs to the silica-membrane in the presence of high chaotropic salt while contaminants pass through the column. Buffers provided with the kit enabled 60-80% DNA recovery and removal of gel-based contaminants. The pure DNA was eluted with Tris-HCl buffer pH 8.0. The resulting DNA was used to ligate to plasmid as described in section 7.2.

5.4 Isolation of plasmid DNA from bacteria

Plasmid was purified from 3-ml overnight culture of *E. coli* cells using QIAminiprep Kit (QIAGEN). The *E. coli* cells were pelleted by centrifugation at 10,000 g for 1 min. After the medium was removed, the pellet was resuspened in 250 μl of buffer P1 and transferred to microcentrifuge tube, then mixed with 250 μl of lysis buffer by inversion, and then added 350 μl of neutralized buffer by inversion, and centrifuged at 10,000 g for 10 min, followed by binding of plasmid DNA to an anion-exchange resin under appropriate low salt and pH conditions. RNA, proteins, dyes, and low molecular weight impurities were removed by a medium-salt wash. Plasmid DNA was eluted in high-salt buffer, and then concentrated and desalted by isopropanol precipitation.

In the case of 50-100 ml overnight culture used to prepare a large quantity of plasmid for DNA sequencing, a QIAmidiprep Kit (QIAGEN) was used.

5.5 Determination of nucleic acids

The concentration of nucleic acids was determined photometrically at 260 nm. Depending on the class of nucleic acid, OD_{260}= 1.0 represents the following concentrations:

double-strand DNA	50 μg
RNA	40 μg

A sample containing nucleic acids was diluted in 1 ml water and absorbance at 260 nm was determined with a spectrophotometer. Water was used as a sample blank in the assay.

6. Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) is a multi-step reaction used for DNA amplification. The reaction consists of the repetition of three steps. In the first step, performed at high temperature (94 ^{o}C), the matrix DNA is denatured and the single stranded DNA is formed. During the second step, performed at a lower temperature (45-65 ^{o}C), the more or less specific oligodeoxy-nucleotides (primers) anneal to the individual segments of the single stranded DNA. Finally, at 72 ^{o}C, DNA polymerase synthesizes a complementary DNA strand using the DNA as a template and the annealed primers as the initiation sites. The matrix DNA, limited by the "downstream" and "upstream" primers, is exponentially multiplied by repeating these three steps 30-40 times. Additionally, 10 mM $MgCl_2$ (3 μl per 50 μl assay) was added to the PCR reactions containing *Taq* (*Thermus aquaticus*) DNA polymerase (Promega). The addition of $MgCl_2$ was not necessary when *Pfu* (*Pyrococcus furiosus*) DNA polymerase (Promega) was used. A mineral oil (Sigma) overlay was included for the PCR reaction carried out in a Thermal Cycler 480 to prevent the evaporation of reaction contents. The use of mineral oil was not necessary when the reactions were performed in a cover-heated GeneAmp PCR 9700 thermal cycler.

6.1 Reverse transcription (cDNA synthesis)

First-strand cDNA was synthesized by mixing 10 μg of total RNA of *C. alata*, *P. indica*, and *R. tataricum*, 1.0 μg of oligo (dT) 12-18-primer, and adding water to 10 μl. The reaction mixture was then incubated for 10 min at 70 ^{o}C. After the reaction was chilled on ice, 4 μl of 5x reaction buffer, 2 μl of 2.5 mM dNTPs mix, and 2 μl of 100 mM DTT were added, and incubated at 37 ^{o}C for 5 min to allow the oligo (dT) primer to anneal to RNA template. Then, 200 U of M-MLV-reverse transcriptase was added, and the volume of the reaction was adjusted to 20 μl with sterile water. After incubation at

37 ^{o}C for 1 hour, the enzyme was inactivated by heating the reaction to 94 ^{o}C for 5 min. A 1/10 vol of first-strand cDNA reaction mixture was used as cDNA template for RT-PCR in section 6.2.

6.2 Reverse transcriptase PCR (RT-PCR)

DNA fragments were amplified by mixing 100 ng of template cDNA as described in section 6.1. The following components were combined in PCR reactions:

10x PCR buffer	5 μl
$MgCl_2$	3 μl
dNTP mix (each 2.5 mM)	1 μl
cDNA	1 μl
upstream primer (PK 1 forward, 50 μM)	2 μl
5'-AA(A/G)GC(C/T)AT(A/C)GAIGA(A/G)TGGGG-3	
downstream primer (PK 2.3 reverse, 50 μM)	2 μl
5'-CCACCIGG(A/G)TGI(A/G)CAATCC-3'	
DNA polymerase, 5 U/μl	0.5 μl
H_2O	added to 50 μl

Samples were subjected to 30 cycles of PCR, each incorporating 30 sec of denaturation at 94 ^{o}C, 30 sec of annealing at 48 ^{o}C, and 1 min of extension at 72 ^{o}C. The resulting DNA fragment was separated in 0.8 % (w/v) agarose gel electrophoresis. The DNA fragment was isolated as described in section 5.3 and used for ligation with pGEM –T Easy as described in section 7.2.

7. Enzymatic modifications of nucleic acids

7.1 Restriction analysis

Bacterial restriction endonucleases enzymes are able to specifically hydrolyze DNA after recognition of a specific sequence in double-stranded DNA. These enzymes represent a useful tool to cut the desired fragments from DNA-molecules. The restriction endonucleases from Life Technologies and New England Biolabs were used. DNA digestion was incubated at 37 ^{o}C for more than two hours and the standard reaction conditions were:

DNA	1-10 μg
10x reaction buffer	2 μl
restriction endonuclease enzyme	1 U
sterile water	to final volume of 20 μl

The resulting digested DNA was used for futher ligation with plasmid.

7.2 Ligation reaction

The enzyme T4 DNA ligase catalyses the formation of a phosphodiester bond between the 5'-phosphate and the 3'-hydroxy end of double stranded DNA molecules. T4 DNA ligase from Promega (4U/μl) was used in the following standard ligation reaction:

Vector (pGEM-T Easy)	10-20 ng
Insert DNA	50 ng
10x buffer T4 DNA ligase	2 μl
T4 DNA ligase	4 U
Sterile water	to final volume of 10 μl

A molar ratio of insert to vector of 3:1 was applied in all ligation reactions. The quantity of the insert was calculated with the formula:

$$\frac{\text{ng vector x kb insert size x (molar ratio insert:vector)}}{\text{Kb vector size}} = \text{ng insert}$$

The ligation reactions were performed in eppendorf reaction tubes. The tubes were incubated in a cooling water bath at 16 ^{o}C overnight. Ligated vectors were subsequently transformed into competent bacteria as described in section 8.1.

8. Preparation and transformation of competent *E. coli* cells

8.1 Preparation of competent *E. coli* cells

The acquirition of DNA from the environment into the host cell is termed as transformation. Several plasmids were transformed into *E. coli* during this work. The preparation of competent *E. coli* cells that were used for the transformation was arranged using the calcium chloride method (Hanahan, 1985).

The following buffers were used for the preparation of competent *E. coli* cells:

TfB I Buffer : 30 mM potassium acetate, pH 5.8
10 mM $CaCl_2$
50 mM $MnCl_2$
100 mM KCl
15 % (v/v) glycerol

TfB II Buffer : 10 mM MOPS/NaOH, pH 7.0
75 mM $CaCl_2$
10 mM KCl
15 % (v/v) glycerol

Single bacterial colony was transferred to 3 ml SOB medium and cultivated at 37 ^{o}C overnight. The next day, 10 μl of culture was used to inoculate 100 mL SOB medium (in a 1 liter flask). The culture was grown on a shaker (200 rpm) at 37 ^{o}C until OD_{550} 0.28. The bacteria were harvested by centrifugation (5 min, 1000 x g, 4 ^{o}C) and

were resuspended in 15 ml ice-cold TfB I buffer. After 10 min incubation on ice, the cells were again centrifuged (5 min, 1000 x g, 4 ^{o}C) and slowly (45 min) and gently resuspended in 4 ml ice-cold TfB II buffer. The aliquots of 100 μl were transferred to pre-cooled (liquid nitrogen) 1.5 ml eppendorf tubes. The tubes containing competent bacteria were stored at -80 ^{o}C.

8.2 Transformation of competent *E. coli* cells

A heat-shock method is a simple and an efficient transformation method when small vectors (up to 10 kb in size) are used for transformation. For each transformation, the contents of tubes containing 100 μl competent cells (section 8.1) were slowly melted on ice. The product of a ligation reaction (10 μl) (section 7.2) was pipetted to the cells and incubated for 30 min on ice. The transformation occurred during a heat shock (5 min in water bath of 42 ^{o}C), then the immediately transferred to ice (2 min incubation) and 0.5 ml LB medium were added to the cells. The cells were incubated for 45 min at 37 ^{o}C and were then transferred to the petri dishes containing LB agar plus ampicillin. The transformed bacteria were grown overnight at 37 ^{o}C, and single colony was used for further analysis.

9. Establishment of cDNA library in λ-phages

Since it was presumed that both plant polyketide synthases could occurred in both *C. alata* and *P. indica* roots, a poly A^{+} RNA encoding for plant polyketide synthase was isolated by screening *C. alata* and *P. indica* cDNA library under the low-stringency conditions. A ZAP-cDNA Synthesis Kit and ZAP-cDNA Gigapack III Gold Cloning Kit (Stratagene) were used in the preparation of *C. alata* and *P. indica* cDNA library.

9.1 Construction of λ-ZAP-cDNA Gigapack III library

Poly A^+ RNA was isolated from *C. alata* and *P. indica* roots as described in section 5.1. *C. alata* and *P. indica* cDNA library were synthesized following the Stratagene's protocol. The first strand cDNA was synthesized using Moloney Mouse Leukemia Virus (*M-MLV*) reverse transcriptase (37 ^{o}C, 1 hour). The inclusion of 5'-methyl dCTP instead of dCTP during the first strand cDNA synthesis allowed the protection of cDNA from the host bacteria's restriction enzymes in the subsequent steps.

The second strand cDNA was synthesized by DNA polymerase I. First, mRNA that was hybridized to the first strand cDNA, was nicked by Rnase H, resulting in a multitude of fragments that served as primers for the cDNA polymerase I to synthesize the second strand cDNA. During the second strand synthesis, the methyl-nucleotides were eliminated from the assay to allow the restriction with *Xho* I in a subsequent step. After the synthesis of second strand cDNA was completed, *Pfu* DNA polymerase was added to the assay to produce blunt 3'-and 5'-ends. The blunt 3'-and 5'-ends allowed a subsequent blunt-end ligation of *Eco*RI adapters.

The *Eco* RI adapter at the 5'-end was removed by *Xho* I restriction. This *Xho* I restriction resulted in a cDNA containing an *Xho* I sticky end at the 5'-end and an *Eco* RI sticky end at the 3'-end. The digested adapter as well as the free nucleotidess were removed after the size exclusion chromatography performed on a drip column (Separose CL-2B gel filtration medium). The size-fractionated cDNA was precipitated. The cDNA was ligated in a Uni-ZAP vector and the ligated vector was packaged with a Gigapack III Gold Packaging extract. The resulting phage was used to infect the host bacteria.

A single colony of *E. coli* strain XL1-Blue MRF' was used to inoculate 50 ml LB medium supplemented with 10 mM $MgSO_4$ and 0.2% maltose (w/v). The bacteria were grown on a rotary shaker (200 rpm) at 37 ^oC for 6 hours. The bacteria were harvested by centrifugation (500 x g, RT, 1 min) and the pellet was resuspended in 10 mM $MgSO_4$ to an OD_{600} of 0.5. To determine the size of the phage cDNA library, the λ-ZAP cDNA library was titered in the following way: To 200 μl of host bacteria, 1 μl of diluted λ-phage was added (phage dilutions: 1:10 and 1:100 in SM buffer) and incubated at 37 ^oC to allowed the attachment of λ-phage to the host bacteria. After 15 min incubation, melted NZY top agar (3 ml; 48 ^oC), IPTG (final conc.: 2.5 mM), and X-Gal (final conc.: 4 mg/ml) were added and mixed. This mixture was immediately plated onto pre-warmed (37 ^oC) NZY-agar plates. The plates were incubated overnight at 37 ^oC. Plaques were counted on the following day. The percentage of cDNA inserts in the cDNA library was determined by a blue-white screening.

9.2 Screening of λ-ZAP-cDNA Gigapack III Library

The phage cDNA library, plated for screening, was prepared in similar way as described in the previous section. To 200 μl of host bacteria resuspended in sterile 10 mM $MgSO_4$ to OD_{600} of 0.5, λ-phage was added and incubated at 37 ^oC for 15 min to allow phage to attach to the bacteria. Melted NZY top agar (6.5 ml for primary screening and 3 ml for subsequent screening; 48 ^oC) was added and mixed. This mixture was immediately plated on the pre-warmed (37 ^oC) petri dishes containing NZY agar. The petri dishes of 150 mm in diameter were used for primary cDNA screening and the petri dishes of 100 mm in diameter were used for the subsequent screening. The plaques developed after the overnight incubation at 37 ^oC.

The following component were used in cDNA library screening:

20 x SSC buffer :	0.3 M sodium citrate/HCl, pH 7.0
	3 M NaCl
SM-buffer :	50 mM Tris-HCl, pH 7.5
	100 mM NaCl
	8 mM $MgSO_4$
	5 ml 2 % (w/v) gelatin
Denaturation solution :	0.5 M NaOH
	1.5 M NaCl
Neutralization solution :	0.5 M Tris-HCl, pH 8.0
	1.5 M NaCl
Washing solution :	0.2 M Tris-HCl, pH 7.5
	2x SSC buffer

The screening was performed using nitrocellulose filters (ϕ132 mm; Satorius; ϕ 82 mm; Schleicher & Schuell). The filters were put on the agar plates and the following procedure was followed:

Transfer from the plate to filter	2 min
Denaturation of filter in denaturation buffer	2 min
Neutralization of filter in neutralization solution	1 min
Additional neutralization of filter in neutralization solution	4 min
Filter washing with washing solution	0.5 min

The filters were subsequently air-dried and the transferred DNA was fixed on the filters by baking at 80 ^{o}C for 2 hours. The filters with fixed DNA were hybridized with [α-^{32}P]-labeled probe as described in section 15. The phages of interest were localized on the NZY agar petri dishes using autoradiograms. The agar containing the plaques of interest was cut out and the phage was eluted from the agar in 500 μl SM buffer

supplemented with 20 μl $CHCl_3$. After the content of the tube was vortexed and centrifuged (500 x g, RT, 1 min), the phage-containing supernatant was diluted in SM buffer (ratio 1:100) and used for the subsequent screening steps or for *in vivo* excision.

9.3 *In vivo* excision

E. coli strain XL1-Blue MRF' and *E. coli* SOLR strain were grown over-night in LB medium supplemeted with 0.2 % (w/v) maltose and 10 mM $MgSO_4$ at 30 ^oC. The next day, cells were centrifuged (1000 x g, 1 min, RT) and resuspended in sterile 10 mM $MgSO_4$ to an OD_{600} of 1.0. The following components were transferred to a 10 ml polypropylene tube: 200 μl *E. coli* XLI-Blue MRF' cells in $MgSO_4$, 200 μl isolated phage and 1 μl ExAssist helper phage. After 15 min incubation at 37 ^oC (phage attaching to bacteria), 3 ml of LB broth were added to the tube. The bacteria were cultivated for 3 hours at 37 ^oC to enable the excission of pBluescript phagemid from the phage. A heating step (70 ^oC, 20 min) followed to inactivate the bacteria. The bacteria were centrifuged (1000 x g, 15 min, RT) and the supernatant containing the excised phagemid was stored at 4 ^oC.

The medium with the excised phagemid (1 μl) was combined with 200 μl *E. coli* SOLR cells, diluted in sterile $MgSO_4$. This strain contains an amber mutation that prevents the replication of phage genome, but allows the replication of phagemid. The bacteria were incubated for 15 min at 37 ^oC and plated on LB agar plates supplemented with ampicillin (50 mg/l). The bacteria were grown overnight (37 ^oC). The next day, single bacteria colonies were transferred to LB supplemented with ampicillin (50 mg/l). The phagemid were isolated from the bacteria using a Plasmid Midi Kit (Qiagen) and were analyzed for the occurrence of inserts by the restriction digestion with *Eco*RI and agrose gel electrophoresis.

10. DNA sequencing

The clones of *C. alat*a and *P. indica* from cDNA library that had expected insert size were selected for further DNA sequencing at MWG-Biotech (Ebersberg, Germany).

11. Comparison of genetic database

The sequence comparison was performed using BLAST search tool (Altschul *et al*., 1990). This software is available on Internet:http//www.ncbi.nlm.nih.gov:80/BLAST/

12. Amplification of partial cDNAs from *R. tataricum* using RACE-PCR

First-strand cDNA was synthesized from 5 μg of total RNA isolated from *R. tataricum* rhizomes using Superscript II reverse transcriptase (Life Technologies). One μl of the first-strand cDNA was used as a PCR template with degenerate primers 5'-AA(A/G)GC(C/T)AT(A/C)GAIGA(A/G)TGGGG-3 and 5'-CCACCIGG(A/G)TGI(A/G) CAATCC-3' based on those described by Helariutta *et al.* (1995) and *Taq* polymerase (Promega). Cycling conditions consisted of an initial denaturation at 94 ºC, 3 min; 30 cycles of 94 ºC, 30 sec; 48 ºC, 30 sec and 72 ºC, 1 min, followed by 10 min at 72 ºC. A PCR product of the expected length (575-600 bp) was purified by gel electrophoresis, ligated into pGEM-T Easy vector (Promega) and sequenced.

12.1 5' and 3' RACE PCR

A RACE PCR kit (SMART Technology; Clontech) was used to synthesize cDNA from total RNA isolated from *R. tataricum* rhizomes as described in section 5.1. 3' RACE was performed using the gene-specific primer 5'-CATAGACTCCATGGTAGG GCAAGC-3' and universal primer A mix (UPM) supplied by the manufacturer to amplify a 755 bp DNA fragment. Cycling conditions consisted of an initial denaturation at 94 º

C, 5 sec; 30 cycles of 94 ºC, 5 sec; 68 ºC, 10 sec and 72 ºC, 3 min, followed by 7 min at 72 ºC Advantage polymerase (Clontech). Similarly, 5' RACE used the gene-specific primer 5'-GCCGCACCGTCACCAAATATTGC-3' and UPM to amplify a 781 bp DNA fragment. Both 3' and 5' RACE PCR products were gel purified, ligated into pGEM T-Easy and sequenced.

13. Expression and purification of recombinant proteins

13.1 Construction of expression vectors

The expression of all investigated full-length genes were performed in the *E. coli.* For the protein purification, all of them were ligated into the pET-14b vectors (Novagen) that enabled fusion of a histidine tag at the N-terminal of the expressed protein (6 histidine residues; His Tag). The *CalPKS1*, *CalPKS2*, *CalPKS3*, *PinPKS* and *RtaPKS* full-length genes (each 1 ng/μl) were PCR amplified. The primers for this PCR amplification consisted of the nucleotides that were complementary to the 3'-and 5'-ends of the protein-coding region, and of the additional nucleotides that enabled subsequent digestion with restriction endonucleases and ligation into the pET-14b vectors (sticky ends ligation). The primer sequences were shown in **Table 1**.

CalPKS1, *CalPKS2*, *CalPKS3* and *PinPKS* were PCR-amplified by *Pfu* DNA polymerase (Promega) whereas *Rta*PKS was PCR-amplified by Taq DNA polymerase. in the PCR assay. The following PCR program was used for the DNA amplification:

Denaturing	3 min	94 ^{o}C	1 cycle
Denaturing	3 min	94 ^{o}C	25 cycles
Annealing	1 min	48 ^{oC}	
Extension	0.5 min	72 ^{o}C	
Final extension	7 min	72 ^{o}C	1 cycle

Table 1 Primers for the PCR amplification of protein-coding regions of *Cal*PKS1, *Cal*PKS*2*, *Cal*PKS*3*, *Pin*PKS and *Rt*PKS (start and stop). The restriction endonuclease recognition sites are underlined.

*Cal*PKS1	5' *Nde* I; 5'-TAGTCA TATGGTGAAGGTGGAAGA-3'
	5' *Xho* I; 5'-TTGGCTCGAGTTAAATAGCAATACTGT-3'
*Cal*PKS2	5' *Nde* I; 5'-ATTTTTCATATGGTGAGTGTTGAAGAG-3'
	5' *Bam* H I; 5'-TTTAAAGGATCCTTAGTTAACTGCCACACT-3'
*Cal*PKS3	5' *Nde* I; 5'- TAGTCATATGGTGAAGGTGGA AGA-3'
	5'*Bam* H I; 5'-TTGGCTCGAGTTAGACAGCCACACTAT-3'
*Pin*PKS	5' *Nde* I; 5'-ATTTTTCATATGGCACCAGCAGTTCAA-3'
	5' *Bam* H I; 5'-TTTAAAGGATCCTTAGTTAAGCGGCACACT-3'
*Rt*PKS	5' *Nde* I; 5'-AATAGTCATATGGCACCGGAGGAGTCG-3'
	5' *Bam* H I; 5'-TTTAAAGGATCCTCAGGTAATTAGCGGCAC-3'

The occurrence of PCR products was investigated by agarose gel electrophoresis. Single bands of the expected length (1200 bp) were detected. These bands were extracted from the agarose gel. The sticky ends of the amplified DNA and vectors were obtained by restriction endonuclease digestion using *Nde*I/*Xho*I for *Cal*PKS1, *Nde*I/ *Bam*H I for *Cal*PKS2, *Nde*I/ *Bam*H I for *Cal*PKS3, *Nde*I/ *Bam*H I for *Pin*PKS and *Nde*I/ *Bam*H I for *Rt*PKS, respectively. Each PCR product (4 μg) and pET-14b (1 μg) were digested overnight at 37 ^{o}C with specific endonucleases enzymes.

The short digested DNA fragments were removed from the restriction endonuclease assays by performing the agarose gel electrophoresis. The DNA-bands of expected length were extracted from the gel (1200 bp for each PKS genes and 4700 bp for pET-14b). The extracted DNA and vectors were combined in ligation reaction.

Fifty-ng each of PKS and 50 ng pET-14b were combined in the ligation reaction. All of ligation reactions were allowed to proceed at 16 ^{o}C overnight.

*Cal*PKS:pET-14b, *Pin*PKS:pET-14b and *Rt*PKS:pET-14b were subsequently transformed into a cloning *E. coli* strain (*E. coli* DH5α). The plasmids were isolated from bacteria and the occurrence of inserts in the plasmids was investigated by a restriction endonuclease digestion using *Nde*I/*Xho*I for *Cal*PKS1, *Nde*I/*Bam*HI for *Cal*PKS2, *Nde*I/*Bam*HI for *Cal*PKS3, *Nde*I/*Bam*HI for *Pin*PKS and *Nde*I/*Bam*HI for *Rt*PKS. The transformation of expression vector constructs in the expression host *E. coli* BL21 (DE3), the heterologous expression, and the preparation of crude bacterial extracts were performed as described below.

The following buffer was used during the expression and purification of recombinant proteins:

HLB buffer (lysis buffer) :	50 mM Tris-HCl, pH 7.0
	500 mM NaCl
	2.5 mM imidazole
	10 mM β-mercaptoethanol
	10% (v/v) glycerol
	1% (v/v) tween 20
	750 μg/ml lysozyme
HWB buffer (wash buffer) :	50 mM Tris-HCl, pH 7.0
	500 mM NaCl
	2.5 mM imidazole
	10 mM β-mercaptoethanol
	10 % (v/v) glycerol

HEB buffer (elution buffer) :	50 mM Tris-HCl, pH 7.0
	500 mM NaCl
	250 mM imidazole
	10 mM β-mercaptoethanol
	10 % (v/v) glycerol
HTB buffer (thrombin buffer) :	50 mM Tris-HCl, pH 7.0
	150 mM NaCl
	2.5 mM imidazole
	10 mM β-mercaptoethanol
	3 mM $CaCl_2$
	5 mM $MgCl_2$
	10 % (v/v) glycerol

13.2 Purification of recombinant proteins

DNA inserts were ligated into expression vectors (pET-14b, Novagen). *E. coli* strain BL21 (DE3) was transformed with the recombinant vectors and plated on petri dishes containing LB agar and ampicillin (100 mg/l). The bacteria were incubated at 37 ^{o}C overnight.

On the next day, 50 ml LB broth that contained the appropriate antibiotic were inoculated with a single bacterial colony. The bacteria were cultivated for six hours at 37 ^{o}C and were subsequently transferred to a two-liter flask containing 950 ml of pre-warmed LB broth (ampicillin (100 mg/l). The bacteria were cultivated on a rotary shaker (200 rpm) at 37 ^{o}C until the OD_{600} had reached 0.6. The flask was cooled on ice (5 min) and IPTG was added in required amounts to induce the protein expression. The final IPTG concentration for the bacteria transformed with pET-14b vector constructed was 0.4 mM. The expression of recombinant protein was allowed to proceed overnight at 28 ^{o}C (shaker 200 rpm).

On the next day, the bacteria culture was chilled on ice (10 min) and the bacteria were harvested by centrifugation (20 min, 4 ^{o}C, 6000 g). The supernatant was decanted and the pellet was resuspened in 20 ml HLB buffer. The cell suspension was incubated for 1 hour on ice to allow the lysis to proceed. Subsequently, the suspension was sonicated (pulse sonication; 3 x 30 sec) and centrifuged (20 min, 12000 g, 4 ^{o}C). The supernatant containing the heterologous protein was decanted into a new tube.

TALON resin was equilibrated with HWB buffer prior to the addition of supernatant according to the supplier's instructions. Two-ml of TALON resin was transferred to a 15 ml polypropylene tube, centrifuged (700 g, 2 min, 4 ^{o}C) and washed twice with 10 ml HWB buffer. The bacterial supernatant containing the recombinant protein was applied to the equilibrated TALON resin and was gently shaken for 20 min at room temperature to allow the binding of His-tagged recombinant protein to the resin. The resin was subsequently centrifuged (700 g, 5 min, 4 ^{o}C), resuspended in 10 ml washing buffer (HWB) and agitated again for 10 min at room temperature. The washing step was repeated once. Following centrifugation (700 g, 5 min, 4 ^{o}C), the resin was resuspended in 2 ml HWB buffer and transferred to an end-capped 2 ml gravity flow column (Clontech). The resin was allowed to settled. After the resin had settle, the end cap was removed to allow the buffer to drip from the column. The resin was washed with 10 ml HWB buffer. The protein that bound to the resin was eluted with 5 ml HEB buffer. One-ml fractions were collected. The protein content of individual fractions was determined by the Bradford assay.

To remove low molecular weight substances (especially imidazole and NaCl), the fractions containing recombinant protein were combined to a total volume of 2.5 ml and applied to a PD-10 column (Pharmacia) that had been previously equilibrated with 20 ml HTB buffer. After the protein fraction entered the PD-10 column, 3.5 ml HTB buffer were added to the column. A single 3.5 ml fraction containing recombinant protein was collected.

The histidine tag of the recombinant protein was removed by thrombin protease. Thrombin protease (20 U; Pharmacia) was added to the fraction that eluted from the PD-10 column. The thrombin digest was performed overnight at 4 ^{o}C. On the next day, the solution was applied to TALON resin in a gravity-flow column, previously equilibrated with 10 ml HTB buffer. When the resin settled, the end-cap was removed from the column. The eluent containing the protein without the histidine tag was collected in a single fraction, while His-tagged protein as well as the hydrolyzed histidine tag remained bound to the TALON resin. The purified protein was used for the determination of polyketide synthase activity as described in section 18.2.

14. Gel electrophoresis

14.1 Agarose gel electrophoresis

DNA or RNA mixtures were separated on agarose gels. Except for the Southern hybridizations that 0.8 % agarose gels were prepared, DNA analysis was performed using 1 % agarose gels containing 1 x TAE buffer:

1 x TAE buffer : 40 mM Tris
20 mM CH_3COOH
1 mM EDTA

The agarose gels (volume of 50 or 100 ml) were prepared as follows: Agarose was melted in 1x TAE buffer in a microwave oven (750 W, 2 min). Ethidium bromide (0.4 μg per 1 ml gel) was added to the gel when the solution had cooled down to about 65 ^{o}C. The gel was immediately poured into horizontal trays for the agarose gel electrophoresis. The hardened gel was transferred to the gel electrophoresis chamber containing 1x TAE buffer. DNA samples were diluted with 6x DNA Loading Solution (Sigma). The samples were pipetted into the gel pockets and run at 80 Volts for about 1-2 hours. Separated DNA was visualized by exposing to UV light and were photographed using the BioRad Gel Doc Equipment.

14.2 Agarose gel electrophoresis of RNA

For northern blot hybridization, RNA was resolved on 1.2% denatured agarose gels in 1x MOPS buffer, obtained by dilution of 10x MOPS buffer.

10 x MOPS buffer : 200 mM MOPS
50 mM CH_3COONa
10 mM EDTA

The agarose gels (volume of 100 ml) were prepared by combining 1.2 g agarose, 84 ml H_2O and 10 ml 10 x MOPS buffer. Following the microwave heating (750 W, 2 min) and cooling to about 60 ^oC, 5.5 ml of 37% formaldehyde were added. The gel solution was poured into the horizontal tray for agarose gel electrophoresis. The hardened gel was transferred to a gel electrophoresis chamber containing 1x MOPS buffer.

Prior to the electrophoresis, RNA-denaturing buffer (25 μl) was added to 10 μl samples containing RNA. Samples were heated at 65 ^oC for 20 min and subsequently cooled on ice. The agarose gel was loaded with the denatured samples and run in 1x MOPS buffer at 95 V. The separated RNA was visualized by exposing to UV light and photographed.

RNA-denaturing buffer : 26 mM MOPS
6.5 mM CH_3COONa
1.3 mM EDTA
8.5 % (v/v) formaldehyde
44 % (v/v) formamide
10 % (v/v) glycerol
0.05 % (w/v) bromophenol blue
0.05 % (w/v) xylenecyanol
0.002 % (w/v) ethidium bromide

14.3 Sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Following the enzyme purification, the purified protein and its molecular weight were investigated by the SDS-PAGE (Laemmli, 1970; Blackshear, 1984). Proteins become strongly negatively charged through the binding of SDS to the protein hydrophobic regions and the reduction of disulfide bonds. The total negative charge correlates with the molecular weight of protein. Therefore, SDS-PAGE analysis can be observed as a separation based on the protein size.

During this work, only 12 % SDS-PAGE gels were used. The acrylamide stock solutions (30% w/v, acrylamide with 0.8% w/v bisacrylamide) were obtained as ready solutions from Roth (Karlsruhe). The following components were needed to perform an SDS-PAGE electrophoresis:

12% SDS-PAGE (separation gel) :

4.0 ml H_2O

2.5 ml 1.5 M Tris-HCl, pH 8.8

3.3 ml acrylamide (Rotiphorese Gel30; Roth)

100 μl 10 % SDS

5 μl TEMED

50 μl 10 % APS

SDS-PAGE stacking gel :

2.3 ml H_2O

255 μl 1.5 M Tris-HCl, pH 8.8

390 μl acrylamide (Rotiphorese Gel 30; Roth)

30 μl 10 % SDS

3 μl TEMED

15 μl 10 % APS

SDS/glycine buffer : 15.0 g glycine; 6.0 g Tris; 1.5 g SDS

Dissolved in one liter H_2O

Sample buffer : 2 ml 0.5 M Tris-HCl, pH 8.8

2 ml 20 % (w/v) SDS

2 ml glycerol

1 ml β-mercaptoethanol

2.6 ml H_2O

0.4 ml 0.5 % (w/v) bromophenol blue

The separation gel was mixed and about 4 ml was poured between two glass plates (100 mm x 73 mm; 100 mm x 83 mm). A distance of 0.5 mm between plates was assured by appropriate spacers. The top of the gel was covered with *n*-butanol to allow the polymerization. After about 30 min, the *n*-butanol was removed and the stacking gel was added (about 1 ml). A comb was inserted to allow the formation of loading pockets. After the gel polymerized, it was transferred together with the glass plates to the electrophoresis apparatus that contained SDS/glycine buffer. For sample application, 20 μl sample buffer was added to 40 μl protein solution, mixed and boiled (95 ^{o}C, 5 min). Samples (20 μl) were pipetted into the gel pockets and run at 80 volts for about 2 hours.

14.4 Protein gel dying

Following the electrophoresis, SDS gels were fixed and dyed in a dying solution.

Dying solution **:** 0.07 % (w/v) Coomassie Brilliant Blue R-250

40 % (v/v) ethanol

10 % (v/v) CH_3COOH

Destaining solution : 40 % (v/v) ethanol

10 % (v/v) CH_3COOH

After at least two hours of staining, the gel background was destained for more than two hours in a destaining solution.

15. Radioactive labeling of nucleic acids

The labeling of nucleic acids was performed using [α-^{32}P] labeled dATP of RadPrime DNA Labeling System (Life Technologies). The supplier's protocol was precisely followed in the individual labeling procedures. The principle of the RadPrime DNA Labeling System is in the annealing of random primers (octamers) to the denatured DNA template and extended by Klenow fragment in the presence of [α-^{32}P] labeled dATP. Twenty-five-ng DNA in 21 μl H_2O was denatured by heating at 100 ^{o}C for 5 min. The solution was immediately chilled on ice and the following was added:

500 μM dCTP	1 μl
500 μM dGTP	1 μl
500 μM dTTP	1 μl
2.5x random primers solution	20 μl
[α-^{32}P] dATP	5 μl
Klenow fragment	1 μl

The reaction was incubated at 37 ^{o}C for 10 min then 5 μl of 0.8 M EDTA (pH 8.0) was added to stop the reaction and use for screening cDNA library (section 9.2) or northern blot analysis (section 16.1).

16. Blotting techniques

SSC buffer was used when northern blot and southern blot were performed.

20 x SSC buffer : 0.3 M sodium citrate/HCl, pH 7.0
3 M NaCl

16.1 Northern blot

The term "northern blot" describes the transfer of RNA to a nylon membrane. The hybridization of bound RNA with the single-stranded, labeled DNA enables the identification of gene transcripts in different tissue. This method is also used to identify the factors that may influence the expression level of the gene of interest.

Total RNA from various parts of *C. alata*, *P. indica*, and *R. tataricum*, were isolated as described in section 5.1. Twenty-microgram of the total RNA was dissolved in RNA denaturing buffer containing ethidium bromide. RNA agarose gel electrophoresis was performed as described above. Following the electrophoresis, the presence of equal amount of RNA on the gel was verified by the composition of ribosomal RNA bands of the individual samples (UV light exposure). The gel was agitated in 10x SSC buffer for 15-20 min. A nylon membrane (Hybond N(+) 0.45 mm, Amersham) was soaked in sterile H_2O for 10 min.

An electrophoresis chamber was used for the transfer stack assembly. The agarose gel was place on 1 layer Whatman 3MM paper that was connected to the reservoirs containing 10x SSC buffer. The nylon membrane was placed over the gel. Several layers of Whatman 3MM paper as well as common tissue paper were placed over the membrane and loaded with a weight to enable the capillary blotting overnight. All of these components were carefully placed one over the other to prevent air bubbles entering the system. The next morning, the nylon membrane was air dried and the the RNA was UV-cross-linked to membrane. The membranes were stored in a solution containing 2x SSC and 0.1% SDS until they were hybridized with labeled probes.

17. Hybridization techniques

The hybridization with [α-^{32}P]-dATP labeled probes was used to identify a plant polyketide synthase cDNA in phage library (section 9.2) and northern blot (section 15.1).

The following components were used during the hybridization:

20 x SSC buffer :	0.3 M sodium citrate/HCl, pH 7.0
	3 M NaCl
50x Denhardt's reagent :	10 g Ficoll (MW 400,000)
	5 g PVP (MW 360,000)
	5 g bovine serum albumin
	H_2O to 500 ml
20x SSPE buffer :	0.2 M Na_2HPO_4 pH 7.4
	3.6 M NaCl
	20 mM EDTA
Pre-hybridization solution :	5x SSPE buffer
	5x Denhardt's solution
	0.1 % (w/v) SDS
	50 % (v/v) formamide
	250 μg/ml salmon sperm DNA
	(Boiled and chilled on ice)
Hybridization solution :	same as pre-hybridization solution, but
	containing [α-^{32}P] labeled dATP

Round nitrocellulose filters with transferred phage cDNA were soaked in a solution containing 2x SSC buffer and 0.1% SDS for 10 min. The filters were subsequently transferred into a plastic tray containing 100 ml pre-hybridization solution. The tray was placed in a water bath (42 ^{o}C) for 3 hours. The pre-hybridization solution

was replaced by hybridization solution (100 ml) that contained [α-^{32}P]-dATP-labeled probe and was returned to the water bath (42 ^{o}C). The hybridization was allowed to proceed overnight.

The nylon membranes from northern blot and southern blotting experiments were hybridized in Duran glass tubes. The membranes were transferred in glass tubes and the pre-hybridization solution was added (10 ml). The tubes containing the membranes were incubated in a hybridization oven (65 ^{o}C) for 3 hours. The pre-hybridization solution was then replaced by the hybridization solution (10 ml) containing [α-^{32}P]-dATP-labeled probe. The hybridization was allowed to proceed overnight at 65 ^{o}C.

After the hybridization was completed, round nitrocellulose filters or nylon membranes obtained from northern and southern blotting experiments were washed twice for 10 min in a solution containing 5x SSPE buffer and 0.1 % SDS. The volume and the temperature of this solution was the same as the temperature of the solutions used during the hybridization (100 ml or 10 ml; 42 ^{o}C or 65 ^{o}C). The radioactivity of filters and membranes was verified with a Geiger counter and, when necessary, washed in subsequent step with a solution containing 1x SSPE and 0.1 % SDS. The filters and the membranes were air-dried. Phage cDNA library filters were exposed to the autoradiography film for 1 day at -80 ^{o}C and subsequently developed, while northern blot membranes were exposed to the phosphor imager screens for 1 day (room temperature) and were evaluated by a phosphor imager.

18. Analytical methods

18.1 Determination of protein

Protein concentration was determined using a Bradford assay (Bradford, 1976). This assay is based upon unspecific binding of Coomassie Brilliant Blue G-250 to proteins in acidic. A shift of absorbance-maximum from 465 nm to 595 nm can be photometrically determined. Bovine serum albumin (Sigma) was used as standard.

A sample with unknown protein concentration was dissolved in 800 µl H_2O. Two-hundred microliters of BioRad Protein Assay (BioRad, München) were added to the sample solution. After 5 min incubation at room temperature, the absorbance at 595 nm was determined. The protein concentration in the sample was determined using bovine serum albumin standard curve.

18.2 Determination of plant polyketide synthases

Polyketide synthase activity was measured by the conversion of various starter CoA molecules and [2-^{14}C]malonyl CoA into reaction products. The standard enzyme assay contained 100 mM Hepes buffer (pH 7.0), 20 µM starter CoA and 12,000 dpm malonyl CoA (specific activity 55 mCi/mmol) and 1.5 µg polyketide synthase in a 50 µl reaction volume. The assay mixture was incubated for 30 min at 30 °C. The reaction was stopped by the addition of 5 µl 10% (v/v) HCl and was extracted twice with 100 µl ethylacetate. The combined organic phase was evaporated to dryness, the residue dissolved in 10 µl ethylacetate and used for resolved by RP-C18 thin layer chromatography (Merck). The TLC plates were developed in methanol:H_2O:acetic acid (75:25:1). The radioactive products were quantified with a Rita Star TLC scanner (Raytest) or by phosphorimaging.

18.3 Optimum temperature of plant polyketide synthase activity

The optimum temperature for the CalCHS, PinPKS and RtSTS activity of purified recombinant enzymes were investigated in the experiments that followed. The standard enzyme assay contained 100 mM Tris-HCl buffer (pH 7.0), 20 μM *p*-coumaroyl-CoA, 12,000 dpm [2-^{14}C] malonyl-CoA and 1.5 μg purified enzyme in a 50 μl reaction volume. To determine the temperature optimum for the synthesis of naringenin, pyrone and resveratrol, respectively, several assays were performed at various temperatures. The standard conditions of the enzyme assays were as described in section 18.2.

18.4 Optimum pH of plant polyketide synthase activity

The assays for optimum pH were performed as described in section 19.3, except that 100 mM Tris-HCl, pH 7.0 was replaced by the following buffers: 100 mM sodium citrate buffers were used for the pH range 4.0 to 6.0, 100 mM potassium phosphate buffer were used for the pH at 6.5, 100 mM Tris-HCl buffer were used for the pH range 7.0 to 8.5, and 100 mM glycine buffers were used for the pH range 9.0 to 10.0. The standard enzyme assay contained 20 μM *p*-coumaroyl-CoA, 12,000 dpm [2-^{14}C] malonyl-CoA and 1.5 μg purified enzyme in a 50 μl reaction volume. Assays were incubated at 30 ^{o}C for 30 min. The standard conditions of the enzyme assays were as described in section 18.2.

19. Mass spectrometry for the analysis of enzymatic products.

To identify the enzymatic products, scaled up reactions containing 75 mM Tris-HCl (pH 7), 50 μM various starter CoA, 100 μM malonyl-CoA and 5 μg purified enzyme in a 200 μl reaction volume were used. The reaction proceeded for 1 hour at 30º before acidification and extraction with ethylacetate. After drying *in vacuo*, the residue was

dissolved in methanol and analysed by LC-MS. Positive and negative electrospray ionization (ESI) mass spectra were obtained with a Finnigan MAT TSQ 7000 instrument (electrospray voltage 4.5 kV; heated capillary temperature 220 ^{o}C; sheath and auxillary gas nitrogen) coupled to a Micro-Tech Ultra-Plus MicroLC system equipped with an RP18-column (5 μm, 100x1 mm, SepServ, Berlin). For all compounds, a gradient system was used that ranged from H_2O:CH_3CN 90:10 (each containing 0.2% acetic acid) to 10:90 over 15 min, followed by isocratic elution with a 10:90 mixture of both solvents for 10 min; flow rate 70 μl/min. The collision-induced dissociation mass spectra during an HPLC run were recorded with a collision energy of –20 or -25 eV for positive ions as well as +20 or +25 eV for negative ions, respectively (collision gas: argon, collision pressure: 1.8 x 10^{-3} Torr). These analyses were kindly performed by Dr.Jürgen Schmidt of Leibniz Institut für pflanzenbiochemie, Halle, Germany.

20. HPLC analysis

20.1 Flavonoid analysis

Plant tissues (leaves, stems and roots) of *C. alata* (1 g fresh weight) were frozen in liquid nitrogen and ground to a fine powder by a mortar and pestle. To the powder, 10 ml of 80% methanol with 3% concentrated HCl was added. The extract was heated at 100 ^{o}C for 60 min in a sealed tube. After cooling, the cell debris was removed from the extract by centrifugation at 10,000 x g for 10 min at room temperature. The extract was taken to dryness *in vacuo*, dissolved in 5 ml water and extracted twice with ethylacetate. The ethylacetate layer was removed and dried *in vacuo* and the residues dissolved in 50 μl of 50% (v/v) methanol and centrifuged at 10,000 x g for 10 min. For analysis of resveratrol, 10 μl of this solution was injected onto a Knauer Eurospher-100 C18 column (250x4mm, 5 μm) connected with Hewlett Packard Series 1100 HPLC.

The solvent system was (A): H_2O containing 2% CH_3CN and 0.2% H_3PO_4 and (B) CH_3CN containing 2% H_2O and 0.2% H_3PO_4, with a linear gradient of 20-45% B in 30 min, followed by 45-55 %B for another 5 min. Flow rate was 0.6 ml min^{-1}. The flavonoid content was monitored at 370 nm.

20.2 Resveratrol analysis

Plant tissues (leaves, petioles and rhizomes) of *R. tataricum* (1 g fresh weight) were frozen in liquid nitrogen and ground to a fine powder by a mortar and pestle. The preparation of resveratrol was as described in the previous section, except the HPLC analysis condition used linear gradient of 20-60% B in 30 min. Flow rate was 0.6 ml min^{-1}. The resveratrol content was monitored at 270 nm.

CHAPTER IV

RESULTS

1. Isolation of cDNA encoding polyketide synthases

1.1 Primer design

The homology comparison of amino acid sequences from known plant polyketide synthases (*Gerbera hybrida-2PS*, *Petroselinum crispum*-CHS, *Ruta graveolens*-ACS, *Vitis* spp.-STS) revealed that the residues with strictly conserved CHS active-site residues, Cys164, Phe215, Phe265, and His303 (numbering in Ferrer *et al*., 1999) are present in all CHS-superfamily as shown in the bold residues of **Figure 13.** However, the regions used for the primer design were based upon conserved regions in plant polyketide synthases as shown in the gray boxed **(Figure13)**. The upstream primer was designated "PK1.1 forward primer" and the downstream primer was designated "PK 2.3 reverse primer" based on those described by Helariutta *et al*., (1995). The sequences of degenerated primers were:

PK 1.1 forward primer : 5'-AA(A/G)GC(C/T)AT(A/C)GAIGA(A/G)TGGGG-3

PK 2.3 reverse primer : 5'-CCACCIGG(A/G)TGI(A/G)CAATCC-3'

1.2 cDNA Synthesis and PCR procedure

Total RNA was isolated from *C. alata*, *P. indica* and *R. tataricum* roots as described in section 5.1 of Materials and Methods, and the first strand cDNA were synthesized from total RNA as described in section 6.1 of Materials and Methods.

The PKS-like cDNA fragments were then obtained by PCR-amplification. The PCR assays including *C. alata* cDNA, *P. indica* cDNA and *R. tataricum* cDNA, an upstream primer (PK1.1 for), a downstream primer (PK2.3 rev), and the following PCR program, were prepared as described in section 6.2 of Materials and Methods.

Figure 13 Amino acid sequence comparison of plant polyketide synthases from *Ruta*-ACS, *Petroselinum*-CHS, *Gerbera*-2PS and *Vitis*-STS. The regions used for primer design are gray boxed.

After the PCR reactions were completed, the occurrence of PCR products was investigated by agarose gel electrophoresis. The results showed the appearance of a 584 bp PCR product from the first strand cDNA of each plant designated Cal1, Pin1, and Rt1 **(Figure 14)**. Each PCR product was then isolated and further analyzed.

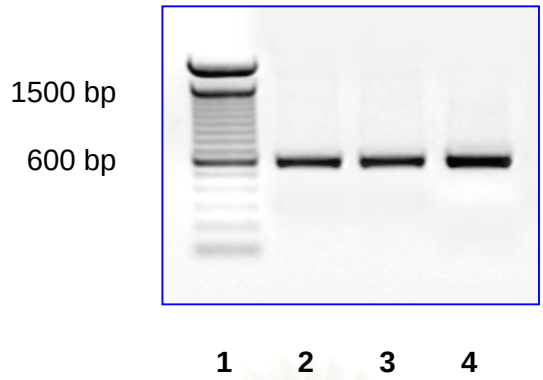


Figure 14 Agarose gel electrophoresis of PCR products, lane 1: 100 bp DNA marker, lane 2: Cal1 (*C. alata*), lane 3: Pin1 (*P. indica*), and lane 4: Rt1 (*R. tataricum*). Using the primer combination of PK1.1 forward and PK2.3 reverse, resulted in the amplification of approximately 584 bp PCR product.

1.3 Subcloning and sequence analysis of amplified PCR fragments

The PCR products (*Cal*1, *Pin*1 and *Rt*1) were isolated from agarose gel and ligated into pGEM-T Easy vectors (Promega). The *Cal*:pGEM-T Easy, *Pin*:pGEM-T Easy and *Rt*:pGEM-T Easy constructs were transformed into competent *E. coli* DH5α. Transformants was analyzed and the DNA inserts were subsequently sequenced at MWG, Ebersberg, Germany.

The sequence analysis revealed that the sequences of *Cal*1, *Pin*1and *Rt*1 were comparable to the sequences of plant PKS genes. These results suggested that *Cal*, *Pin* and *Rt* cDNA were related to a plant polyketide synthase superfamily.

1.4 Isolation of full-length plant polyketide synthase gene from *C. alata* and *P. indica*

cDNA Libraries of *C. alata* and *P. indica* were prepared as described in section 9 of Materials and Methods. The size of cDNA libraries of *C. alata* or *P. indica* cDNA were estimated to be 80,000 pfu (plaque forming units) with 98.5% pfu containing the Uni-ZAP vector with cDNA inserts (colorless plaques) and 1.5% pfu without the inserts (blue plaques).

1.4.1 Screening *C. alata* and *P. indica* cDNA-libraries

C. alata and *P. indica* cDNA libraries were screened using [α-^{32}P]-dATP labeled *Cal*1 and *Pin*1 fragments as probe, respectively. Plaque hybridization revealed 20 strong signals per each cDNA library. Each of the 20 corresponding phages was isolated and diluted in SM-buffer (1:100). The diluted phages were then reinfect host bacteria. The second cDNA library screening resulted in the identification of 20 strong signals (**Figure 15**). All of phages were isolated, diluted in SM-buffer (1:100), and were used for the third infection of host bacteria and screening. The third was performed and individual phages were isolated. The corresponding phages were designated *Cal*PKS1 to *Cal*PKS20 and *Pin*PKS1 to *Pin*PKS20 , respectively.

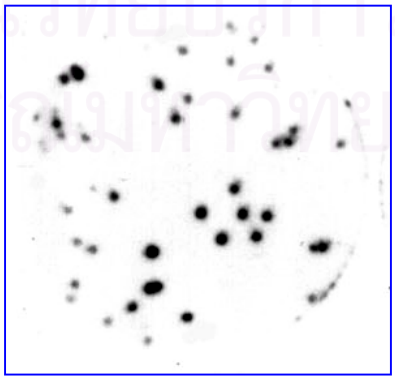


Figure 15 Positive clones in the second screening with strong signals.

Once each positive clone was identified, it was religated into pBluescript SK(-) with helper phage. The size of inserted cDNA was analyzed by digesting with suitable restriction endonucleases, followed by analysis on 1% agarose gel. The gel electrophoresis of digested phagemids demonstrated that the clones of *C. alat*a and *P. indica* from cDNA library that had expected insert size approximately 1,500 bp to 1,700 bp and then selected for further DNA sequencing at MWG-Biotech (Ebersberg, Germany).

The complete nucleotide sequences of *Cal*PKS1, *Cal*PKS2, *Cal*PKS3 and *Pin*PKS were determined and contained 1554, 1536, 1590 and 1622 bp encoding complete reading frames of 389, 390, 389 and 396 amino acid residues and shown in **Figure 16, 17, 18, 19,** and **20,** respectively. The predicted molecular weights for the encoded proteins are 42.6, 42.9, 42.6 and 43.0 kDa, respectively.

```
1      aattcggcacgaggatcatactagctagcacgcaccaagtggtacttccatctctttatt       60
61     attatttttgttagaacaagttcggggaaaatttaaaaATGGTGAAGGTGGAAGAGATC       120
                                            M  V  K  V  E  E  I   -    7
121    CGTAAGGCACAACGGGCAGAAGGTGCCGCGACGGTGATGGCCATCGGCACAGCCACTCCG       180
        R  K  A  Q  R  A  E  G  A  A  T  V  M  A  I  G  T  A  T  P   -   27
181    GCGAACTGTGTTGAACAAAGCACCTATCCAGATTACTATTTCCGTGTTACCAACAGTGAG       240
        A  N  C  V  E  Q  S  T  Y  P  D  Y  Y  F  R  V  T  N  S  E   -   47
241    CACATGACAGAGTTAAAAGAGAAATTTCAACGCATGTGTGATAAGTCAATGATCAAGAAG       300
        H  M  T  E  L  K  E  K  F  Q  R  M  C  D  K  S  M  I  K  K   -   67
301    AGATACATGCACTTGACAGAGGAGATCCTCAAGGAGAACCCTAACATGTGTGCTTACATG       360
        R  Y  M  H  L  T  E  E  I  L  K  E  N  P  N  M  C  A  Y  M   -   87
361    GCGCCTTCCATTGATGCTAGGCAAGACATAGTGGTTTTGGAAGTACCAAAGCTTGGAAAA       420
        A  P  S  I  D  A  R  Q  D  I  V  V  L  E  V  P  K  L  G  K   -   107
421    GAGGCTGCAACAAAGGCCATCAAGGAATGGGGTCAACCCAAATCCAAAATTACCCATTTG       480
        E  A  A  T  K  A  I  K  E  W  G  Q  P  K  S  K  I  T  H  L   -   127
481    ATCTTTTGCACCACAAGTGGTGTGGACATGCCCGGAGCTGATTATCAACTCACTAAGCTC       540
        I  F  C  T  T  S  G  V  D  M  P  G  A  D  Y  Q  L  T  K  L   -   147
541    TTGGGCCTTCGTCCCTCCGTCAAGCGATACATGATGTACCAACAAGGTTGCTTTGCCGGC       600
        L  G  L  R  P  S  V  K  R  Y  M  M  Y  Q  Q  G  C  F  A  G   -   167
601    GGCACGGTGCTCCGCTTGGCCAAGGACTTGGCCGAGAATAACAAAGGTGCACGTGTACTT       660
        G  T  V  L  R  L  A  K  D  L  A  E  N  N  K  G  A  R  V  L   -   187
661    GTGGTTTGTTCTGAGATTACTGCAGTCACATTCCGTGGGCCTAGTGATACCCATCTTGAC       720
        V  V  C  S  E  I  T  A  V  T  F  R  G  P  S  D  T  H  L  D   -   207
721    AGCCTTGTGGGCCAAGCATTGTTTGGTGATGGAGCAGCTGCTGTCATTGTTGGATCCGAC       780
        S  L  V  G  Q  A  L  F  G  D  G  A  A  A  V  I  V  G  S  D   -   227
781    CCAATCCCACAAGTTGAGACGCCCTTATTTGAATTAGTATGGACAGCCCAAACTATCCTC       840
        P  I  P  Q  V  E  T  P  L  F  E  L  V  W  T  A  Q  T  I  L   -   247
841    CCTGACAGTGAAGGAGCTATTGATGGACATCTTCGTGAAGTTGGGCTCACATTCCATCTT       900
        P  D  S  E  G  A  I  D  G  H  L  R  E  V  G  L  T  F  H  L   -   267
901    CTAAAAGACGTTCCTGGGCTCATTTCAAAGAACATTGAGAAAGCCTTGGTTGAAGCCTTC       960
        L  K  D  V  P  G  L  I  S  K  N  I  E  K  A  L  V  E  A  F   -   287
961    AATCCATTGGGAATATCTGACTACAACTCAATTTTCTGGATTGCTCACCCAGGTGGGCCA       1020
        N  P  L  G  I  S  D  Y  N  S  I  F  W  I  A  H  P  G  G  P   -   307
1021   GCAATTTTGGACCAAGTCGAGGCCAAATTGGGCTTGAAGCCCGAAAAGATGCAGGCCACT       1080
        A  I  L  D  Q  V  E  A  K  L  G  L  K  P  E  K  M  Q  A  T   -   327
1081   AGACATGTGCTTAGTGAGTATGGAAACATGTCCAGTGCATGCGTGTTATTCATTATGGAT       1140
        R  H  V  L  S  E  Y  G  N  M  S  S  A  C  V  L  F  I  M  D   -   347
1141   GAAATGAGGAGGAAATCAACAAAAGATGGGCTTGGCACAACAGGCGAAGGGCTTGAATGG       1200
        E  M  R  R  K  S  T  K  D  G  L  G  T  T  G  E  G  L  E  W   -   367
1201   GGTGTTTTATTTGGATTTGGGCCTGGCCTTACTGTCGAGACTGTGGTGCTCCACAGTATT       1260
        G  V  L  F  G  F  G  P  G  L  T  V  E  T  V  V  L  H  S  I   -   387
1261   GCTATTTAAgtcttcaaagcccaaagtattcttttaattgtctttgttattcatatttcg       1320
        A  I  .                                                        -   389
1321   catgttgtcttgaacactaaaataatgcttaatttgtattcaaattttgttatttaaatt       1400
1401   gttaaaagatgatgtatgagttatagctccggacaatttggggcaagtttgtaaaagggc       1480
1481   ttggcataacttatagtaagtcaccagtttgtgagtaaaaaaaaaaaaaaaaaa             1554
```

Figure 16 Nucleotide sequence and deduced amino acid sequence of *C. alata* cDNA clone *Cal*PKS1

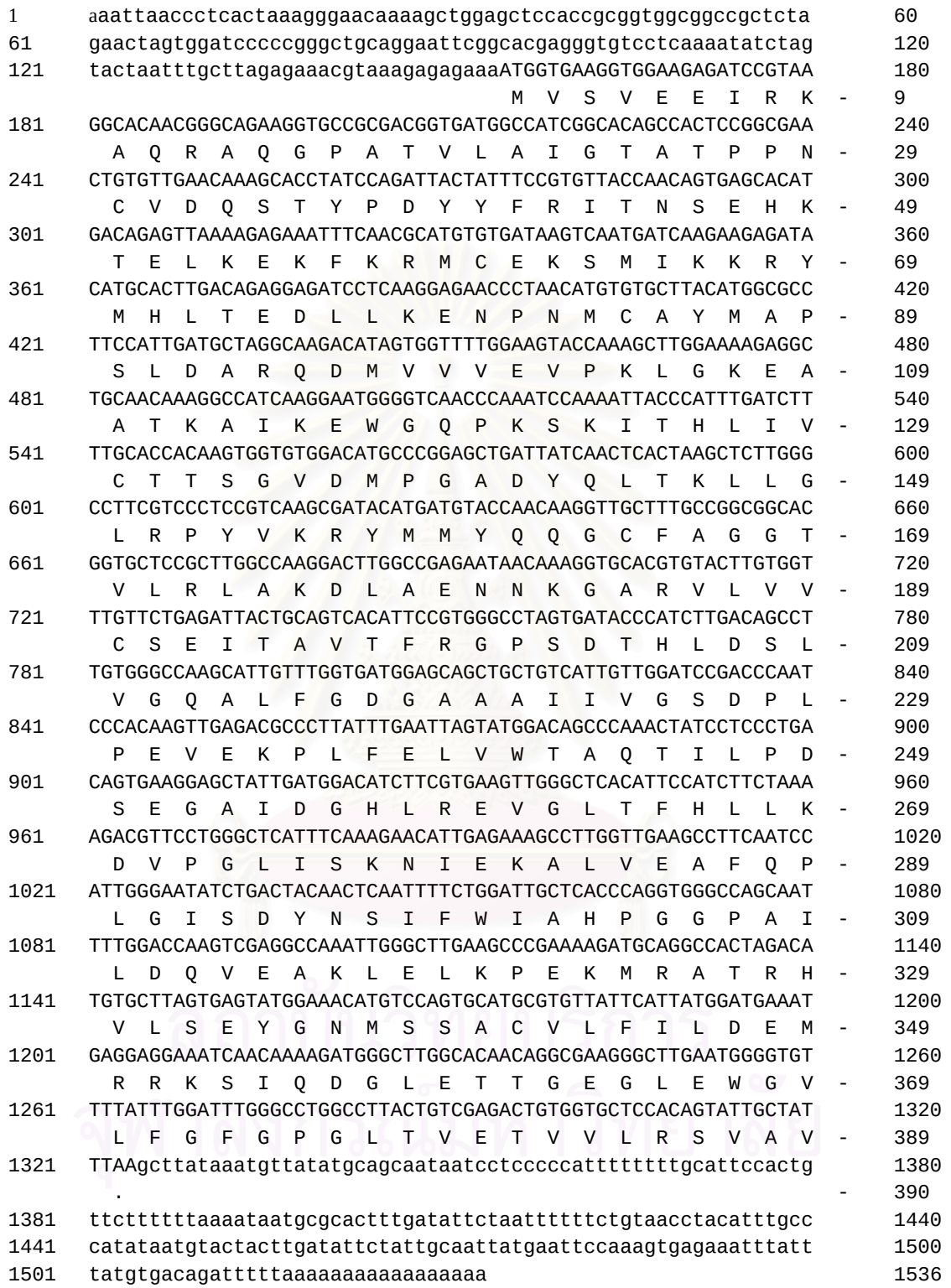


Figure 17 Nucleotide sequence and deduced amino acid sequence of *C. alata* cDNA clone *Cal*PKS2

```
1      aaagggaacaaaagctggagctccaccgcggtggcggccgctctagaactagtggatccc      60
61     ccgggctgcaggaattcgcacgaggtggtacttctcagttcaccacatcatcaacatcaa      120
121    cctttcccttagtagcctccccctctcagtttgtttgtcttcttgattagtatttttgt       180
181    tacaacaagtttaggggaaagttaaagatATGGTGAAGGTGGAAGAGATCCGTAAGGCAC      240
                                    M  V  K  V  E  E  I  R  K  A    -    10
241    AACGGGCGGAAGGCGCCGCCACGGTGATGGCGATTGGTACGGCAACTCCGGCTAACTGCG      300
       Q  R  A  E  G  A  A  T  V  M  A  I  G  T  A  T  P  A  N  C    -    30
301    TGGACCAAAGCACCTATCCTGACTACTACTTTCGTATCACAAATAGTGAGCACATGACGG      360
       V  D  Q  S  T  Y  P  D  Y  Y  F  R  I  T  N  S  E  H  M  T    -    50
361    AGTTGAAGGAAAAATTTCAGCGCATGTGTGATAAGTCAATGATCAAGAAAAGATACATGC      420
       E  L  K  E  K  F  Q  R  M  C  D  K  S  M  I  K  K  R  Y  M    -    70
421    ATTTGACAGAAGAGATCCTAAAGGAGAACCCTAACATGTGTGCTTACATGGCACCTTCCA      480
       H  L  T  E  E  I  L  K  E  N  P  N  M  C  A  Y  M  A  P  S    -    90
481    TTGATGCTAGGCAAGACATAGTGGTTTTGGAAGTACCAAAGCTAGGAAAAGAGGCTGCAA      540
       I  D  A  R  Q  D  I  V  V  L  E  V  P  K  L  G  K  E  A  A    -    110
541    CAAAGGCCATCAAGGAATGGGGTCAACCCAAATCCAAAATTACCCATTTGATCTTTTGCA       600
       T  K  A  I  K  E  W  G  Q  P  K  S  K  I  T  H  L  I  F  C    -    130
601    CCACAAGTGGTGTGGACATGCCTGGAGCTGATTATCAACTCACTAAGCTCTTGGGCCTTC      660
       T  T  S  G  V  D  M  P  G  A  D  Y  Q  L  T  K  L  L  G  L    -    150
661    GTCCCTCTGTCAAGCGATACATGATGTACCAACAAGGTTGCTTCGCCGGAGgcACGGTGC      720
       R  P  S  V  K  R  Y  M  M  Y  Q  Q  G  C  F  A  G  G  T  V    -    170
721    TCCGCTTGGCCAAGGACTTGGCCGAGAATAACAAAGGTGCACGTGTACTTGTGGTTTGTT      780
       L  R  L  A  K  D  L  A  E  N  N  K  G  A  R  V  L  V  V  C    -    190
781    CTGAGATTACTGCAGTCACATTCCGTGGGCCCAGTGATACCCATCTTGACAGCCTTGTGG      840
       S  E  I  T  A  V  T  F  R  G  P  S  D  T  H  L  D  S  L  V    -    210
841    GCCAAGCATTGTTTGGTGATGGAGCAGCTGCTGTCATTGTTGGATCCGACCCAATTCCAG      900
       G  Q  A  L  F  G  D  G  A  A  A  V  I  V  G  S  D  P  I  P    -    230
901    AAGTTGAGAAGCCCTTATTTGAATTAGTATGGACAGCCCAAACTATCCTCCCTGACAGTG      960
       E  V  E  K  P  L  F  E  L  V  W  T  A  Q  T  I  L  P  D  S    -    250
961    AAGGAGCTATTGATGGACATCTTCGTGAAGTTGGGCTTACATTCCATCTTCTGAAAGACG      1020
       E  G  A  I  D  G  H  L  R  E  V  G  L  T  F  H  L  L  K  D    -    270
1021   TTCCTGGGCTCATCTCAAAGAACATTGAGAAAGCCTTGGTTGAAGCTTTCAAACCATTGG      1080
       V  P  G  L  I  S  K  N  I  E  K  A  L  V  E  A  F  K  P  L    -    290
1081   GAATATCTGACTACAACTCAATTTTTTGGATTGCTCACCCAGGTGGGCCAGCAATTTTGG      1140
       G  I  S  D  Y  N  S  I  F  W  I  A  H  P  G  G  P  A  I  L    -    310
1141   ACCAAGTCGAGGCCAAATTGGGCTTGAAGCCCGAAAAGATGCAGGCCACTAGACATGTGC      1200
       D  Q  V  E  A  K  L  G  L  K  P  E  K  M  Q  A  T  R  H  V    -    330
1201   TTAGTGAGTATGGAAACATGTCCAGTGCATGCGTGTTATTCATTATGGATGAAATGAGGA      1260
       L  S  E  Y  G  N  M  S  S  A  C  V  L  F  I  M  D  E  M  R    -    350
1261   GGAAATCAACAAAAGATGGGCTTGGCACAACAGGTGAAGGGCTTGAGTGGGGTGTTTTAT      1320
       R  K  S  T  K  D  G  L  G  T  T  G  E  G  L  E  W  G  V  L    -    370
1321   TTGGATTTGGGCCTGGCCTTACTGTAGAGACTGTTGTGCTCCATAGTGTGGCTGTCTAAa      1380
       F  G  F  G  P  G  L  T  V  E  T  V  V  L  H  S  V  A  V  .    -    389
1381   tgcttataaatgttatatgcagcaataatcctcccccatttttttgcattccactgttc       1440
1441   ttttttaaaataatgcgcactttgatattctaatttttctgtaacctacatttgcccat       1500
1501   ataatgtactacttgatattctattgcaattatgaattccaaagtgagaaatttatttat      1560
1561   gtgacagattttaaaaaaaaaaaaaaaa                                      1590
```

Figure 18 Nucleotide sequence and deduced amino acid sequence of *C. alata* cDNA clone *Cal*PKS3

```
                                                          50
MVKVEEIRKAQRAEGAATVMAIGTATPANCVEQSTYPDYYFRVTNSEHMT   CalCHS1
--S----------Q-P---L-------P---D----------I-----K-   CalCHS2
------------------------------D----------I-------   CalCHS3
                                                         100
ELKEKFQRMCDKSMIKKRYMHLTEEILKENPNMCAYMAPSIDARQDIVVL   CalCHS1
------K---E------------DL--------------L-----M--V   CalCHS2
--------------------------------------------------   CalCHS3
                                                         150
EVPKLGKEAATKAIKEWGQPKSKITHLIFCTTSGVDMPGADYQLTKLLGL   CalCHS1
----------------------------V---------------------   CalCHS2
--------------------------------------------------   CalCHS3
                                                         200
RPSVKRYMMYQQGCFAGGTVLRLAKDLAENNKGARVLVVCSEITAVTFRG   CalCHS1
--Y-----------------------------------------------   CalCHS2
--------------------------------------------------   CalCHS3
                                                         250
PSDTHLDSLVGQALFGDGAAAVIVGSDPIPQVETPLFELVWTAQTILPDS   CalCHS1
---------------------I------L-E--K----------------   CalCHS2
------------------------------E--K----------------   CalCHS3
                                                         300
EGAIDGHLREVGLTFHLLKDVPGLISKNIEKALVEAFNPLGISDYNSIFW   CalCHS1
-------------------------------------Q------------   CalCHS2
-------------------------------------K------------   CalCHS3
                                                         350
IAHPGGPAILDQVEAKLGLKPEKMQATRHVLSEYGNMSSACVLFIMDEMR   CalCHS1
-----------------E------R-------------------L----   CalCHS2
--------------------------------------------------   CalCHS3

RKSTKDGLGTTGEGLEWGVLFGFGPGLTVETVVLHSIAI              CalCHS1
---IQ---E------------------------R-V-VN              CalCHS2
-----------------------------------V-V               CalCHS3
```

Figure 19 Comparison of translations of the nucleotide sequences of three polyketide synthase cDNAs isolated from root tissue of *C. alata*.

Comparison of those amino acid sequences of these cDNAs with the GenBank/EMBL databases indicated that the *Cassia* PKS genes are most similar to chalcone synthase from other members of the Leguminosae, especially, *Glycine max* of 92% identity (Akada *et al*. 1993) and *Phaseolus vulgaris* of 92% identity (Ryder *et al*. 1987) **(Table 2).**

Table 2 Amino acid similarity of *C. alata* root-specific type III polyketide synthases

	CalPKS1[a]	CalPKS2	CalPKS3	PsaCHS	GmaCHS
CalPKS1	100[b]	92	98	92	92
CalPKS2		100	93	92	93
CalPKS3			100	92	93
PsaCHS				100	92
GmaCHS					100

[a] *Cal*PKS1, *Cassia alata* type III polyketide synthase 1 (this work); *Psa*CHS, *Pisum sativum* chalcone synthase (Ichinose *et al*. 1992); *Gma*CHS, *Glycine max* chalcone synthase (Akada *et al*. 1991).

[b] % amino acid identity

In case of *Plumbago* PKS gene are most similar to chalcone synthase from *Gerbera hybrida* with 60% identity (Helariutta *et al*., 1995) and *Glycine max* with 60% identity (Akada *et al*., 1993). The complete nucleotide sequences of cDNAs *Pin*PKS was determined and contained 1622 bp encoding complete reading frames of 396 amino acid residues as shown in **Figure 20**.

```
1      ggcacgagggagctctctagcattcattatatattcctagagctaagctagtgcagttac      60
61     Tgcatatagagatctagagatcgaagagcagtactgctactaactagctagctagcaaat      120
121    acatttgaaATGGCACCAGCAGTTCAATCTCAATCTCACGGTGGAGCATACCGCAGCAAT      180
                M  A  P  A  V  Q  S  Q  S  H  G  G  A  Y  R  S  N   -    17
181    GGTGAGAGGTCAAAAGGGCCAGCGACCGTGCTAGCCATTGCTACTGCTGTGCCACCAAAT      240
        G  E  R  S  K  G  P  A  T  V  L  A  I  A  T  A  V  P  P  N   -    37
241    GTATACTATCAGGATGAATATGCCGACTTTTTCTTCCGCGTCACCAACAGCGAGCACAAG      300
        V  Y  Y  Q  D  E  Y  A  D  F  F  F  R  V  T  N  S  E  H  K   -    57
301    ACTGCGATCAAGGAGAAGTTTAACCGAGTTTGCGGTACCTCGATGATTAAGAAGAGGCAC      360
        T  A  I  K  E  K  F  N  R  V  C  G  T  S  M  I  K  K  R  H   -    77
361    ATGTACTTCACCGAGAAGATGCTTAACCAAAACAAAAACATGTGCACCTGGGATGATAAA      420
        M  Y  F  T  E  K  M  L  N  Q  N  K  N  M  C  T  W  D  D  K   -    97
421    TCCCTCAACGCCCGTCAGGACATGGTGATCCCAGCAGTCCCCGAGCTCGGCAAAGAAGCC      480
        S  L  N  A  R  Q  D  M  V  I  P  A  V  P  E  L  G  K  E  A   -    117
481    GCCTTGAAGGCCATCGAGGAGTGGGGAAAACCACTCTCTAACATCACCCACCTCATCTTC      540
        A  L  K  A  I  E  E  W  G  K  P  L  S  N  I  T  H  L  I  F   -    137
541    TGCACCACAGCCGGTAACGACGCCCCTGGAGCAGACTTCAGGCTAACCCAGCTCCTTGGA      600
        C  T  T  A  G  N  D  A  P  G  A  D  F  R  L  T  Q  L  L  G   -    157
601    CTGAACCCATCAGTGAACCGGTACATGATCTACCAGCAGGGATGCTTCGCTGGAGCCACC      660
        L  N  P  S  V  N  R  Y  M  I  Y  Q  Q  G  C  F  A  G  A  T   -    177
661    GCACTCCGCATAGCCAAGGACCTTGCTGAGAACAACAAGGGTGCTCGTGTGCTCATTGTA      720
        A  L  R  I  A  K  D  L  A  E  N  N  K  G  A  R  V  L  I  V   -    197
721    TGCTGTGAGATCTTTGCTTTTGCATTCCGTGGACCTCATGAGGACCACATGGACTCTTTG      780
        C  C  E  I  F  A  F  A  F  R  G  P  H  E  D  H  M  D  S  L   -    217
781    ATTTGCCAGCTGCTGTTTGGGGATGGTGCAGCTGCTGTCATTGTCGGTGGTGATCCTGAC      840
        I  C  Q  L  L  F  G  D  G  A  A  A  V  I  V  G  G  D  P  D   -    237
841    GAGACCGAGAATGCACTCTTTGAGCTCGAGTGGGCCAACTCAACCATCATACCACAATCA      900
        E  T  E  N  A  L  F  E  L  E  W  A  N  S  T  I  I  P  Q  S   -    257
901    GAAGAGGCCATCACCCTTAGAATGCGCGAAGAAGGTCTCATGATCGGTTTGTCCAAGGAA      960
        E  E  A  I  T  L  R  M  R  E  E  G  L  M  I  G  L  S  K  E   -    277
961    ATCCCAAGGCTCCTAGGCGAACAGATCGAAAGCATTTTGGTCGAGGCTTTCACACCCCTT      1020
        I  P  R  L  L  G  E  Q  I  E  S  I  L  V  E  A  F  T  P  L   -    297
1021   GGAATTACTGACTGGAGCTCACTCTTCTGGATTGCCCACCCAGGTGGTAAGGCCATCCTT      1080
        G  I  T  D  W  S  S  L  F  W  I  A  H  P  G  G  K  A  I  L   -    317
1081   GAGGCACTGGAGAAGAAAATCGGCGTTGAAGGTAAGTTGTGGGCTTCGTGGCACGTTCTT      1140
        E  A  L  E  K  K  I  G  V  E  G  K  L  W  A  S  W  H  V  L   -    337
1141   AAAGAATATGGAAACTTGACCAGTGCTTGTGTGCTGTTCGCCATGGACGAAATGAGGAAG      1200
        K  E  Y  G  N  L  T  S  A  C  V  L  F  A  M  D  E  M  R  K   -    357
1201   AGGTCCATTAAGGAAGGGAAGGCCACTACTGGAGACGGACACGAATATGGTGTTCTCTTC      1260
        R  S  I  K  E  G  K  A  T  T  G  D  G  H  E  Y  G  V  L  F   -    377
1261   GGTGTCGGCCCGGGTCTTACCGTCGAGACAGTTGTGCTAAAAAGTGTGCCGCTTAACTAA      1320
        G  V  G  P  G  L  T  V  E  T  V  V  L  K  S  V  P  L  N  .   -    396
1321   gaagaagttagttcgttggtccgtccgtccgaatcgaaagtagctatatatatacccatt      1380
1381   tttattaagttatttaattatcgttgctgccgctgctggtggctactgctgctctactgc      1440
1441   tgctgcatgagaataaagtcatatatatattatgacaccgagctacgtacgtgcttcttt      1500
1501   agagctgcttctgtttcaacctctcgtgagtgaatccgtcgttgacaatgctctctttg      1560
1561   tattgacgtgttctctttttgttttttttttttctttctttcttgcgctatcaaaaaaaa      1620
1621   aa                                                               1622
```

Figure 20 Nucleotide sequence and deduced amino acid sequence of *P. indica* cDNA clone *Pin*PKS

1.5 RACE PCR for *Rheum tataricum*

1.5.1 Cloning of core fragment

The total RNA was isolated from *R. tataricum* roots for first strand cDNA preparation. The core fragment was obtained by further with PK1.1 forward and PK2.3 reverse primers (results in section 1.1). The PCR product of approximately 584 bp was identified. Deduced amino acid sequence comparison revealed a high level of similarity to other plant PKSs. The result suggesting that the obtained fragment was a part of a gene for plant polyketide synthase superfamily.

1.5.2 3'-End Amplification

Rapid amplification of cDNA end (RACE) method was applied to obtain 3'-end of the cDNA. The specific primer (Rh1.1-3'-race; 5'- CATAGACTCCATG GTAGGGCAAGC-3') was designed in a sense direction based on the core fragment sequence. 3'-RACE with Rh1.1-3'-race primer for the PCR with a universal primer A mix (UPM) (Clontech) as an antisense primer gave an approximately 753-based pair product **(Figure 21)**. The nucleotide and deduced amino acid sequence of the 3'-end fragment showed perfect matching in overlapping sequence to the core fragment and showed a high degree of similarity to these of 3'-end of the relevant plant polyketide synthases.

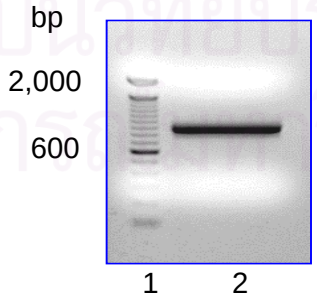


Figure 21 Agarose gel electrophoresis of amplified product 3'-RACE PCR, lane 1: 100 bp DNA marker, lane 2: 3'-RACE PCR product of *R. tataricum* cDNA using primer Rh1.1-3'-race and universal primer A mix (UPM)

1.5.3 5'-End Amplification

RACE method was also applied to obtain 5'-end of the cDNA. Specific antisense primer was designed based on the core fragment. The template cDNA was prepared with superscript II reverse transcriptase. 5'-RACE with Rh1.1 5'-race; 5'-GCCGCACCGTCACCAAATATTGC-3' and a universal primer A mix (UPM) (Clontech) were used as antisense primer to give a 781-based pair product. The nucleotide and deduced amino acid sequence of the 5'-end fragment showed perfect matching in overlapping sequence to the core fragment and showed a high degree of similarity to those of 5'-end of the relevant plant polyketide synthases.

1.5.4 Amplification of open reading frame (ORF) of *R. tataricum-PKS*

DNA fragment covering ORF was obtained by PCR with 5'-and 3'-flanking primers. The complete nucleotide sequences of *Rt*PKS1 was determined and contained 1,433 bp encoding complete reading frame of 391 amino acid residues and shown in **Figure 22**. The predicted molecular weights for the encoded proteins is 43.0 kDa.

Comparison of the translated nucleotide sequences of this cDNAs with the GenBank/EMBL databases indicated that the *Rheum* genes are most similar to benzalacetone synthase from *R. palmatum* with 82% identity (Abe *et al*., 2001) and chalcone synthase from *Petroselinum crispum* with 70% identity (Reimold *et al*.,1983).

```
1      gcgggactcgcacttgctcaacagacgctagctagagaatctaagcttaacatatctgct      60
61     agagATGGCACCGGAGGAGTCCAGGCATGCTGAAACTGCAGTTAACAGAGCCGCCACCGT      120
           M  A  P  E  E  S  R  H  A  E  T  A  V  N  R  A  A  T  V      19
121    CCTGGCCATCGGCACTGCCAACCCGCCAAACTGCTACTATCAAGCGGACTTTCCTGACTT      180
         L  A  I  G  T  A  N  P  P  N  C  Y  Y  Q  A  D  F  P  D  F      39
181    CTACTTCCGTGCCACCAACAGCGACCACCTCACGCACCTCAAGCAAAAATTTAAGCGCAT      240
         Y  F  R  A  T  N  S  D  H  L  T  H  L  K  Q  K  F  K  R  I      59
241    TTGTGAGAAATCGATGATTGAAAAACGTTATCTCCATTTGACGGAAGAAATTCTCAAGGA      300
         C  E  K  S  M  I  E  K  R  Y  L  H  L  T  E  E  I  L  K  E      79
301    GAATCCAAATATTGCTTCCTTCGAGGCGCCATCATTGGATGTAAGACATAACATTCAAGT      360
         N  P  N  I  A  S  F  E  A  P  S  L  D  V  R  H  N  I  Q  V      99
361    GAAAGAAGTGGTGCTGCTCGGAAAAGAGGCAGCTTTGAAGGCCATCAATGAGTGGGGCCA      420
         K  E  V  V  L  L  G  K  E  A  A  L  K  A  I  N  E  W  G  Q      119
421    ACCCAAGTCAAAGATCACGCGCCTCATTGTGTGTTGTATTGCCGGCGTTGACATGCCCGG      480
         P  K  S  K  I  T  R  L  I  V  C  C  I  A  G  V  D  M  P  G      139
481    CGCAGACTATCAACTCACTAAACTCCTTGGCTTACAACTTTCTGTTAAGCGATTTATGTT      540
         A  D  Y  Q  L  T  K  L  L  G  L  Q  L  S  V  K  R  F  M  F      159
541    TTACCACCTAGGATGCTATGCCGGTGGCACCGTCCTTCGCCTTGCGAAGGACATAGCAGA      600
         Y  H  L  G  C  Y  A  G  G  T  V  L  R  L  A  K  D  I  A  E      179
601    AAACAACAAGGAAGCTCGTGTTCTCATCGTTCGCTCTGAGATGACGCCAATCTGTTTCCG      660
         N  N  K  E  A  R  V  L  I  V  R  S  E  M  T  P  I  C  F  R      199
661    TGGGCCATCCGAAACCCACATAGACTCCATGGTAGGGCAAGCAATATTTGGTGACGGTGC      720
         G  P  S  E  T  H  I  D  S  M  V  G  Q  A  I  F  G  D  G  A      219
721    TGCGGCTGTTATAGTTGGTGCAAATCCCGACCTATCCATCGAAAGGCCGATTTTCGAGTT      780
         A  A  V  I  V  G  A  N  P  D  L  S  I  E  R  P  I  F  E  L      239
781    GATTTCTACATCCCAAACTATCATACCTGAATCCGATGGTGCGATTGAGGGACATTTGCT      840
         I  S  T  S  Q  T  I  I  P  E  S  D  G  A  I  E  G  H  L  L      259
841    TGAAGTTGGACTCAGTTTCCAACTCTACCAGACTGTTCCCTCATTAATCTCTAATTGTAT      900
         E  V  G  L  S  F  Q  L  Y  Q  T  V  P  S  L  I  S  N  C  I      279
901    CGAAACTTGTCTTTCAAAGGCTTTCACACCTCTTAACATTAGTGATTGGAACTCACTATT      960
         E  T  C  L  S  K  A  F  T  P  L  N  I  S  D  W  N  S  L  F      299
961    CTGGATTGCACACCCTGGTGGCCGTGCTATCCTTGACGATATCGAGGCTACTGTTGGTCT      1020
         W  I  A  H  P  G  G  R  A  I  L  D  D  I  E  A  T  V  G  L      319
1021   CAAGAAGGAGAAACTTAAGGCAACAAGACAAGTTTTGAACGACTATGGGAACATGTCAAG      1080
         K  K  E  K  L  K  A  T  R  Q  V  L  N  D  Y  G  N  M  S  S      339
1081   TGCTTGCGTATTTTTCATCATGGATGAGATGAGGAAGAAGTCGCTCGCAAACGGTCAAGT      1140
         A  C  V  F  F  I  M  D  E  M  R  K  K  S  L  A  N  G  Q  V      359
1141   AACCACTGGAGAAGGACTCAAGTGGGGTGTTCTTTTTGGGTTCGGGCCAGGTGTTACTGT      1200
         T  T  G  E  G  L  K  W  G  V  L  F  G  F  G  P  G  V  T  V      379
1201   GGAAACTGTGGTTCTAAGCAGTGTGCCGCTAATTACCTGAaaaccattgaataacttgta      1260
         E  T  V  V  L  S  S  V  P  L  I  T  .                         391
1261   ataagatacaagttactcatgttaagtyttgtgtgttgaaatgggttggtctgacgttgt      1320
1321   gatttcttgttgtgtgcttcagtgcttgtttatgtatttgaataattgtgagttaatccc      1380
1381   ctctacaatgggcgctctttatgcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa           1433
```

Figure 22 Nucleotide sequence and deduced amino acid sequence of *R. tataricum* cDNA clone *Rt*PKS.

1.6 Phylogenetic comparison of the deduced polypeptide sequence of *Cal*CHS, *Pin*PS and *Rt*STS

A phylogenetic tree was constructed from the genetic distance matrix using Neighbor-Joining (NJ) method (Saito and Nei, 1987). The GenBank accession numbers of the CHS superfamily members used to construct the phylogenetic tree are *Phalaenopsis* BBS (P53416), *Ruta graveolens* ACS (S60241), *R. graveolens* CHS (CAC14059), *Gerbera hybrida* 2-PS (CAA86219), *G. hybrida* CHS (S56699), *Rheum palmatum* BAS (AAK82824), *Hydrangea macrophylla* CTAS (BAA32733), *H. macrophylla* CHS (BAA32732), *Humulus lupulus* VPS (BAA29039), *H. lupulus* CHS (BAB47196), *Vitis vinifera* STS (CAA54221), *V. vinifera* CHS (CAA53583), *Petroselinum crispum* CHS (CAA24779), *Oryza sativa* CHS (CAA61955), *Arachis hypogaea* STS (P20178), *Medicago sativa* CHS (P30074), *Pinus densiflora* CHS (BAA94594), *P. densiflora* STS (BAA94593), *Pinus sylvestris* CHS (CAA43166), *P. sylvestris* STS (CAA43165), *Psilotum nudum* CHS (BAA87922), *P. nudum* VPS (BAA87923) and *P. nudum* STS (BAA87925) In this study, *C. alata* CHS1 (AF358430), *C. alata* CHS2 (AF358431), *C. alata* CHS3 (AF358432), *P. indica* PS (no accession number) and *R. tataricum* STS (AF508150), respectively.

The relationships of *Cal*CHS, *Pin*PS, *Rt*STS and other chalcone synthase superfamily members **Figure 23** along with bootstrap values greater than 70%. *Cal*CHS to be closely related to the same cluster CHS from other members of the Leguminosae, for example, *M. sataiva* CHS and other CHS, such as *V. vinifera* CHS, *H. macrophylla* CHS.

*Rt*STS to be closely related to the same cluster of STS of *V. venifera* STS, *A. hypogaea* STS and also included *R. graveolens* ACS.

*Pin*PS seem to be closely related to *P. nodum* valerophenone synthase (VPS) and forming a group with functionally diverse enzyme such as *R. palmatum*

benzalacetone synthase (BAS), *G. hybrida* 2-pyrone synthase (2-PS). This group appears to be distinct from groups composed of angiosperm CHSs (*P. crispum*, *V. vinifera*, *O. sativa*, *R. graveolens*, *H. macrophylla*) and from PKSs from the Leguminosae (*C. alata, M. sativa*).

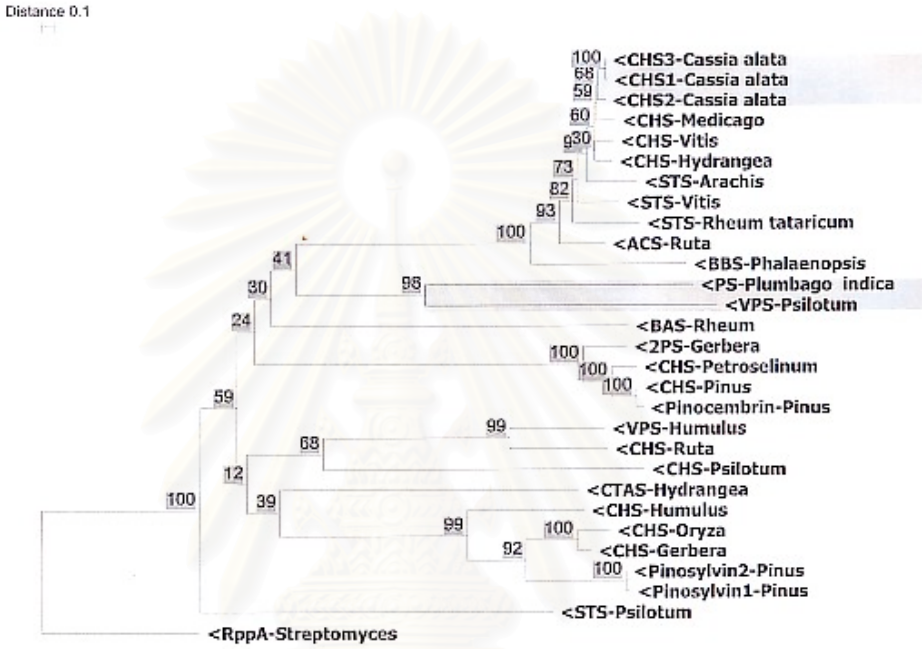


Figure 23 Phylogenetic comparison of the deduced polypeptide sequence of *Cal*CHS, *Pin*PS and *Rt*STS with members of the CHS superfamily. The phylogenetic tree was constructed using TREECON (Van de Peer and Wachter, 1994). *Streptomyces griseus* RppA was used as an outgroup.

Alignment of the deduced amino acid sequences with other PKS indicates that the gene products from *Cassia, Plumbago and Rheum*, were highly similar to CHS and CHS-like proteins from other plant species. Almost all of strongly conserved residues were found in the deduced protein sequences. For example, the four residues which define the active site (Cys164, Phe215, Phe265, His303, and Asn336; Ferrer *et al*., 1999) that are conserved in all known CHS and CHS-like enzymes as shown in **Figure 24**.

Figure 24 Amino acid sequence aligment of *Cal*PKS*, Pin*PKS, *Rta*PKS and plant polyketide synthases ; *Ruta*-ACS, *Petroselinum*-CHS, *Gerbera*-2PS and *Vitis*-STS, respectively. The different amino acid sequences are shaded black. The sequence alignment was done using the CLUSTAL algorithm.

2. Heterologous expression in *E. coli* of plant polyketide synthases from *Cal*PKS, *Pin*PKS and *Rt*PKS.

Although the amino acid sequence comparison between *Cal*PKS, *Pin*PKS and *RtPKS* did not show any obvious differrences, we decided to heterologously express and characterize the protein encoded by these *Cal*PKS, *Pin*PKS and *Rt*PKS genes.

2.1 Construction of expression vectors

The expression of all investigated full-length genes were performed in the *E. coli* pET expression system (Novagen). For the reason of protein purification, all of them were ligated into the pET-14b vectors that enabled fusion of a histidine tag at the N-terminal of the expressed protein (6 histidine residues; His Tag). The transformation of expression vector constructs in the expression host *E. coli* BL21 (DE3), the heterologous expression, and the preparation of crude bacterial extracts were performed as described in section 13.2 of Materials and Methods.

2.2 Purification of recombinant *CalPKS, PinPKS* and *RtPKS*

A bacteria expression system has the advantage of producing large quantities of functional proteins within a relatively short time span. Besides the high levels of PKS which can be produced in *E. coli*, they can also be purified by relatively simple methods. For example, addition of histidine tags at the N-terminus or C-terminus makes purification of heterologously expressed PKS easy. Based on these data, we constructed 6xHis-*Cal*PKS, 6xHis-*Pin*PKS and 6xHis*Rt*PKS and expressed them in *E. coli*.

Based on this method, the *Cal*PKS, *Pin*PKS and *Rt*PKS proteins were each purified from 1-liter each of *E. coli* BL21 (DE3) culture. After isopropyl-β-D-thiogalactopyranoside (IPTG) induction, the crude protein extract containing heterologously expressed His-tagged fusion were purified by TALON resin (Clontech). The protein SDS-PAGE gel of the individual purification steps of *Cal*PKS, *Pin*PKS and *Rt*PKS were shown in **Figure 25, 26** and **27**, respectively, purified of each enzyme had a molecular mass of approximately 42, 43, and 43 kDa, respectively.

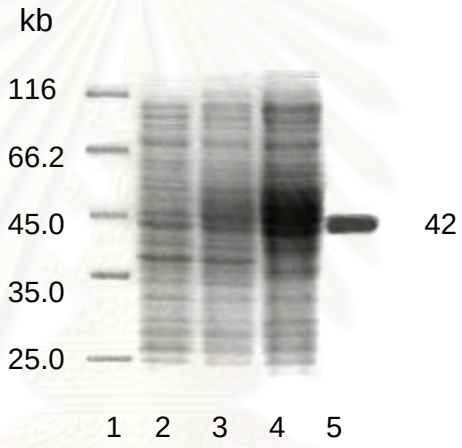


Figure 25 SDS-PAGE of the individual *Cal*PKS purification steps. Five microgram protein from each individual purification steps was applied on the 12% SDS gel. The gel was stained with Coomassie brilliant blue solution. Lane 1, molecular mass markers; Lane 2, uninduced cells; Lane 3, induced cell; Lane 4, total cell lysate and Lane 5, *Cal*PKS purified by affinity chromatography with a His-bind resin.

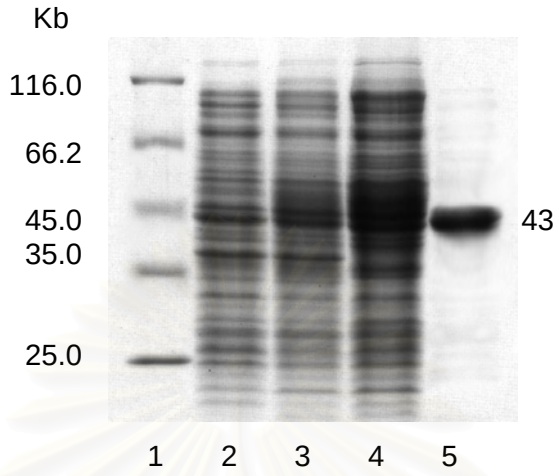


Figure 26 SDS-PAGE of the individual *Pin*PKS purification steps. Five microgram protein from each individual purification steps was applied on the 12% SDS gel. The gel was stained with Coomassie brilliant blue solution. Lane 1, molecular mass markers; Lane 2, uninduced cells; Lane 3, induced cell; Lane 4, total cell lysate and Lane 5, *Pin*PKS purified by affinity chromatography with a His-bind resin.

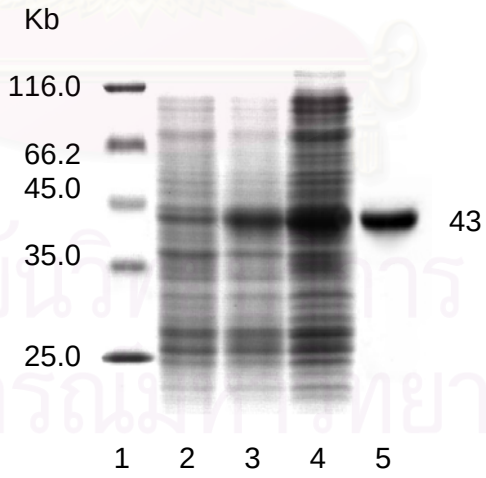


Figure 27 SDS-PAGE of the individual *Rt*PKS purification steps. Five microgram protein from each individual purification steps was applied on the 12% SDS gel. The gel was stained with Coomassie brilliant blue solution. Lane 1, molecular mass markers; Lane 2, uninduced cells; Lane 3, induced cell; Lane 4, total cell lysate and Lane 5, *Rt*PKS purified by affinity chromatography with a His-bind resin.

3. Enzyme assay and mass spectrometric analysis of enzymatic products

3.1 Enzymatic products of *Cal*PKS

The assay for the enzyme activity of *Cal*PKS was performed according to the method described in section 18.2 of Materials and Methods. In order to determine the catalytic activity of *Cal*PKS, recombinant enzyme was incubated with [2-^{14}C] malonyl-CoA and the potential starter CoA esters: acetyl-CoA, *n*-butyryl-CoA, isovaleryl-CoA, *n*-hexanoyl-CoA, benzoyl-CoA, cinnamoyl-CoA, and *p*-coumaroyl-CoA. The reaction mixtures were acidified before product extraction to increase recovery of pyrone (Yamaguchi *et al*., 1999) and to prevent cyclization of chalcones to their corresponding flavanones. The resulting radioactive products were resolved by reversed-phase thin layer chromatography. As shown in **Figure 28**, *Cal*PKS accepted both aliphatic and aromatic CoA esters as primers for polyketide synthesis and produced multiple products from each. The major product formed in the *p*-coumaroyl-CoA assay co-chromatographed with naringenin, suggesting that the *Cassia* PKS encoded a CHS.

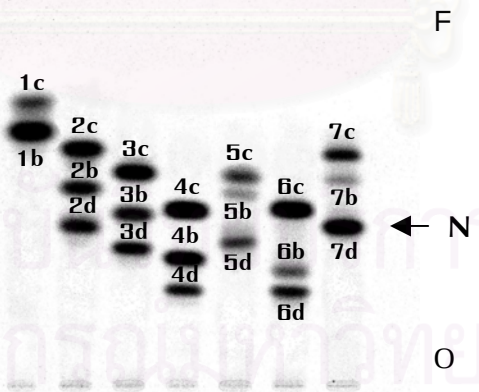


Figure 28 Reversed-phase TLC analysis of *Cal*PKS reaction products. The reactions were performed with 1.5 μg of purified enzyme, [2-^{14}C] malonyl-CoA and acetyl-CoA (lane 1), *n*-butyryl-CoA (lane 2), isovaleryl-CoA (lane 3), n-hexanoyl-CoA (lane 4), benzoyl-CoA (lane 5), cinnamoyl-CoA (lane 6), p-coumaroyl-CoA (lane 7). O= TLC origin, F= solvent front, N= position of naringenin standard.

In order to thoroughly characterize the heterologously expressed *Cassia* cDNAs, the reactions were repeated with unlabeled substrates on a larger scale so that the reaction products could be analyzed by mass spectrometry. The chemical structures of the compounds identified by liquid chromatography electrospray ionization mass spectrometry (LC/ESI-MS) are summarized in **Table 3** and the corresponding mass spectral data are shown in **Table 4.** With the exception of the assay in which acetyl-CoA is used as the starter molecule, all substrates yielded a phloroglucinol derivative resulting from three sequential condensations of acetate units derived from three malonyl-CoA decarboxylations. The tetraketide intermediate thus formed cyclizes by an intramolecular Claisen condensation into the hydroxylated aromatic ring system. Every substrate tested in the assay also produced two pyrone derivatives, one resulting from two acetate unit condensations (a bis-noryangonin-type pyrone derailment product) and one resulting from three acetate unit condensations (a 4-coumaroyltriacetic acid lactone-type pyrone derailment product).

Table 3 Enzymatic product of *Cal*CHS with different starter CoA esters

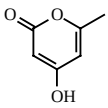
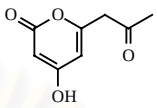
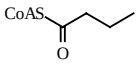
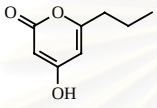
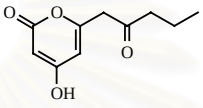
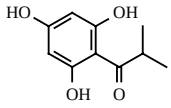
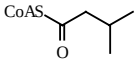
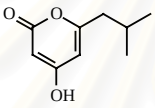
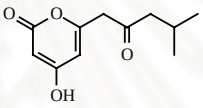
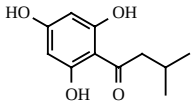
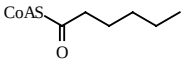
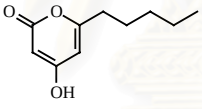
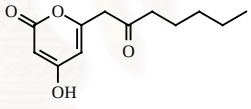
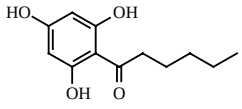
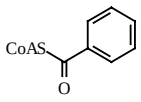
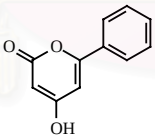
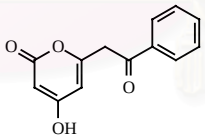
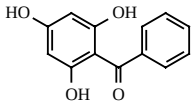
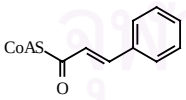
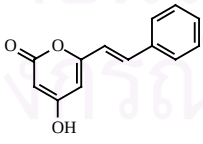
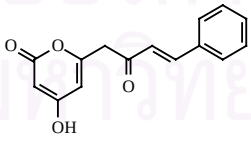
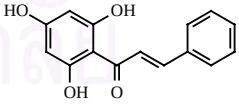
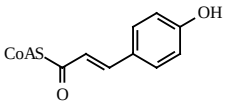
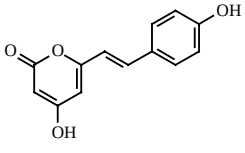
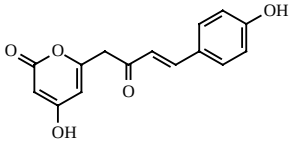
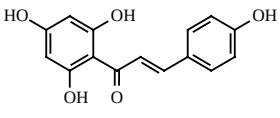
Starter-CoA	Enzymatic product		
	BNY-type	**CTAL-type**	**Chalcone-type**
 acetyl- CoA **(1a)**	 *m/z* 127 $[M+H]^+$ **(1b)**	 *m/z* 169 $[M+H]^+$ **(1c)**	
 *n*-butyryl-CoA **(2a)**	 *m/z* 155 $[M+H]^+$ **(2b)**	 *m/z* 197 $[M+H]^+$ **(2c)**	 *m/z* 197 $[M+H]^+$ **(2d)**
 *Iso*valeroyl-CoA **(3a)**	 *m/z* 169 $[M+H]^+$ **(3b)**	 *m/z* 211 $[M+H]^+$ **(3c)**	 *m/z* 211 $[M+H]^+$ **(3d)**
 Hexanoyl-CoA **(4a)**	 *m/z* 183 $[M+H]^+$ **(4b)**	 *m/z* 225 $[M+H]^+$ **(4c)**	 *m/z* 225 $[M+H]^+$ **(4d)**
 Benzoyl-CoA **(5a)**	 *m/z* 189 $[M+H]^+$ **(5b)**	 *m/z* 231 $[M+H]^+$ **(5c)**	 *m/z* 231 $[M+H]^+$ **(5d)**
 Cinnamoyl-CoA **(6a)**	 *m/z* 215 $[M+H]^+$ **(6b)**	 *m/z* 257 $[M+H]^+$ **(6c)**	 *m/z* 257 $[M+H]^+$ **(6d)**
 *p*-Coumaroyl-CoA **(7a)**	 *m/z* 231 $[M+H]^+$ **(7b)**	 *m/z* 273 $[M+H]^+$ **(7c)**	 *m/z* 273 $[M+H]^+$ **(7d)**

Table 4. Collision-induced dissociation (CID) mass spectra of enzymatically formed products from *Cal*CHS, compound 1-7 obtained from the $[M+H]^+$ ions by LC/ESI-MS

Substrate	Product	RT (min)[a]	$[M+H]^+$ (m/z)	Ions in the CID mass spectra (m/z, rel.int.)	CE (eV)
Acetyl CoA **(1a)**	**1c**	10.20	169[b]	169, 127 (100), 85 (50), 43 (23)	-20
	1b	11.10	127	127 (47), 109 (27), 99 (18), 87 (22), 71 (21), 69 (68), 67 (18), 65 (39), 53 (20), 43 (100)	-25
n-Butyryl CoA **(2a)**	**2c**	14.62	197	197 (1), 127 (100), 113 (8), 71 (46), 43 (25)	-20
	2b	15.14	155	155 (100) , 137 (12), 126 (68), 109 (15), 91 (40), 87 (41), 81 (55), 79 (45), 69 (70), 65 (18), 57 (18)	-25
	2d	17.60	197	197 (74), 179 (95), 155 (93), 139 (33), 123 (67), 113 (100), 95 (42), 43 (16)	-20
Isovaleryl CoA **(3a)**	**3c**	15.65	211	211 (2), 127 (100), 111 (14), 85 (55), 57 (57)	-20
	3b	16.00	169	169 (45), 126 (100), 105 (16), 87 (28), 69 (22), 67 (26), 43 (25)	-25
	3d	17.98	211	211 (53), 193 (20), 169(14), 165 (8), 155 (100), 151 (63), 139 (15), 129 (9), 123 (30), 57 (5), 43 (4)	-20
Hexanoyl CoA **(4a)**	**4c**	17.46	225	225 (10), 141 (6), 127 (100), 99 (27), 71 (13)	-15
	4b	17.81	183	183 (100), 155 (12), 141 (5), 126 (68), 109 (33), 99 (25), 95 (28), 87 (27), 81 (45), 71 (32), 69 (44), 67 (35), 57 (28)	-25
	4d	19.26	225	225 (35), 207 (100), 165 (15), 155 (52), 151 (10), 143 (12), 139 (61), 123 (10), 101 (35), 83 (17)	-20
Benzoyl CoA **(5a)**	**5c**	15.76	231	231 (5), 189 (3), 157 (4), 147 (22), 105 (100), 71 (4)	-20
	5b	16.20	189	189 (100), 171 (95), 161 (28), 147 (10), 129 (55), 127 (75), 115 (63), 105 (72), 103 (43), 77 (17), 69 (87)	-25
	5d	16.61	231	231 (4), 153 (100), 105 (28)	-20
Cinnamoyl CoA **(6a)**	**6c**	17.13	257	257 (5), 221 (2), 173 (8), 131 (100), 103 (3), 69 (2)	-20
	6b	17.93	215	215 (28), 197 (18), 187 (15), 179 (7), 169 (45), 155 (20), 141 (80), 131 (100), 129 (15), 103 (15), 69 (28)	-25
	6d	19.53	257	257 (25), 153 (100), 131 (55), 123 (7), 103 (11), 77 (6), 69 (5)	-25
p-Coumaroyl CoA **(7a)**	**7c**	15.59	273	273 (2), 189 (5), 147 (100), 119 (12)	-25
	7b	16.26	231	231 (35), 213 (12), 203 (33) 171 (12), 157 (25), 147 (100), 121 (15), 110 (5), 69 (4)	-20
	7d	17.84	273	273 (7), 255 (9), 153 (100), 147 (23), 123 (4), 85 (5)	-25

CE = collision energy.

[a]Retention times (RT) were obtained from the corresponding ion chromatogrames in the HPLC.

[b]compound **1c** was identified by selected reaction monitoring (SRM), the relative abundances were calculated from the peak areas of the three detected reactions m/z 169 m/z ⟶ 127, m/z 169 ⟶ m/z 85, and m/z 169 ⟶ m/z 43.

3.2 Enzymatic products of *Pin*PKS

The assay for the enzyme activity of *Pin*PKS was performed according to the method described in section18.2 of Materials and Methods. As shown in **Figure 29**, *Pin*PKS accepted both aliphatic and aromatic CoA esters as primers for polyketide synthesis and produced multiple products from each.

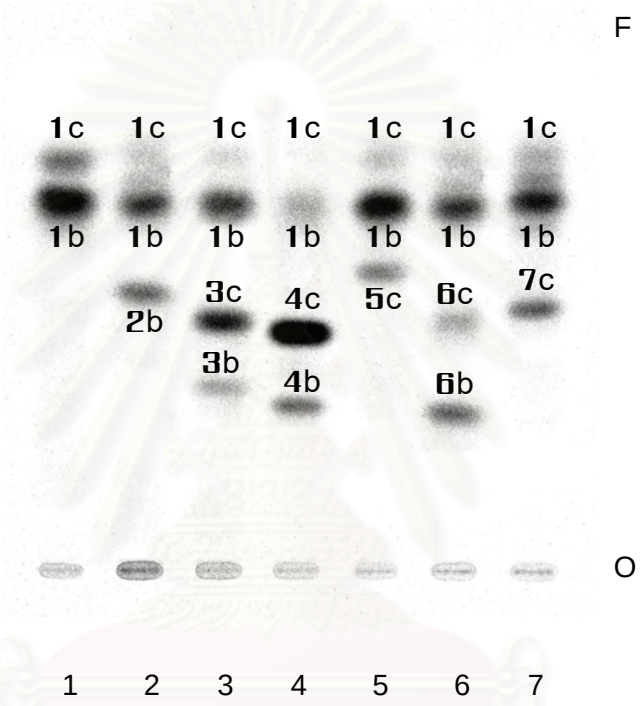


Figure 29 Reversed-phase TLC analysis of *Pin*PKS reaction products. The reactions were performed with 1.5 μg of purified enzyme, [2-^{14}C] malonyl-CoA and acetyl-CoA (lane 1), *n*-butyryl-CoA (lane 2), isovaleryl-CoA (lane 3), *n*-hexanoyl-CoA (lane 4), benzoyl-CoA (lane 5), cinnamoyl-CoA (lane 6), *p*-coumaroyl-CoA (lane 7). O= TLC origin, F= solvent front.

Similar to *Cal*PKS, substrate CoA (1a, 2a, 3a, 4a, 5a, 6a and 7a) and [2-^{14}C] malonyl-CoA were incubated with the recombinant *Pin*PKS overexpressed in *E. coli,* and their products were isolated by radio TLC. It is interesting that the enzyme of *Pin*PKS is the novel plant polyketide synthase which has the different characteristics from the other plant polyketide synthases.

It was found that the recombinant PinPKS displayed broad substrate specificity for the nonphysiological substrates and efficiently afforded the unnatural natural product. 6-Methyl-4-hydroxy-2-pyrone (triacetic acid lactone) **(1b)** (*m/z* 127 $[M+H]^+$) was formed as the main product. In addition, 6-acetonyl-4-hydroxy-2-pyrone (tetracetic acid lactone) **(1c)** (*m/z* 169 $[M+H]^+$) which was also produced when *Pin*PKS was incubated with acetyl-CoA as starter-CoA. However, it shows that the enzyme of *Pin*PKS was capable of producing the *in vitro* products of BNY-type **(2b, 3b, 4b, 5b. 6b and 7b)**, CTAL-type **(3c, 4c, 5c, 6c and 7c)** compounds were also produced as minor products *in vitro* assay. The enzyme of *Pin*PKS did not cause the formation of the chalcone-type compound. The chemical structures of the compounds identified by liquid chromatography electrospray ionization mass spectrometry (LC/ESI/MS) were summarized in **Table 5** and the corresponding mass spectral data were shown in **Table 6**.

Table 5 Enzymatic product of *Pin*PKS with different starter CoA esters.

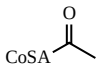
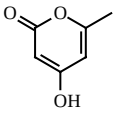
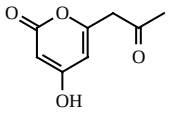
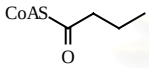
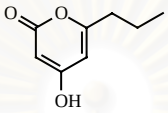
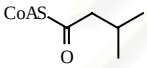
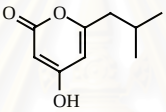
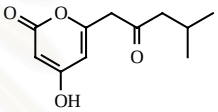
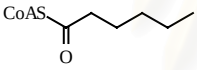
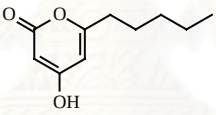
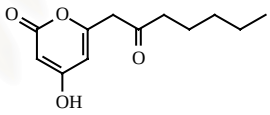
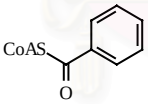
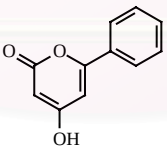
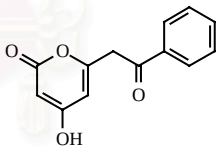
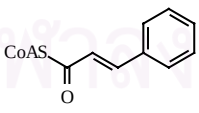
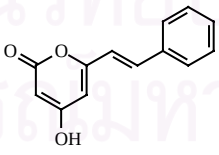
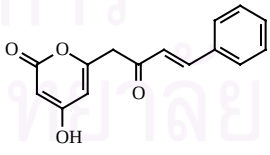
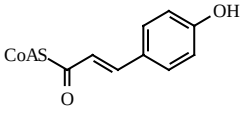
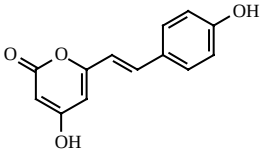
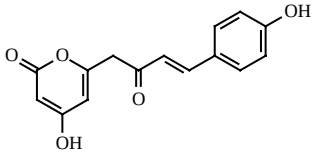
Starter-CoA	Enzymatic product	
	BNY-type	CTAL-type
 acetyl- CoA **(1a)**	 *m/z* 127 $[M+H]^+$ **(1b)**	 *m/z* 169 $[M+H]^+$ **(1c)**
 *n*-butyryl-CoA **(2a)**	 *m/z* 155 $[M+H]^+$ **(2b)**	
 *Iso*valeroyl-CoA **(3a)**	 *m/z* 169 $[M+H]^+$ **(3b)**	 *m/z* 211 $[M+H]^+$ **(3c)**
 Hexanoyl-CoA **(4a)**	 *m/z* 183 $[M+H]^+$ **(4b)**	 *m/z* 225 $[M+H]^+$ **(4c)**
 Benzoyl-CoA **(5a)**	 *m/z* 189 $[M+H]^+$ **(5b)**	 *m/z* 231 $[M+H]^+$ **(5c)**
 Cinnamoyl-CoA **(6a)**	 *m/z* 215 $[M+H]^+$ **(6b)**	 *m/z* 257 $[M+H]^+$ **(6c)**
 *p*-Coumaroyl-CoA **(7a)**	 *m/z* 231 $[M+H]^+$ **(7b)**	 *m/z* 273 $[M+H]^+$ **(7c)**

Table 6 Collision-induced dissociation (CID) mass spectra of enzymatically formed products from *Pin*PKS, compound 1-7 obtained from the $[M+H]^+$ ions by LC/ESI-MS

Substrate	Product	RT (min)[a]	$[M+H]^+$ (m/z)	Ions in the CID mass spectra (m/z, rel.int.)	CE (eV)
Acetyl-CoA	**1c**	9.37	169[b]	169, 127 (100), 85 (50), 43 (23)	-20
(1a)	**1b**	10.27	127	127 (72), 109 (18), 99 (27), 87 (57), 71 (16), 69 (68), 67 (33), 65 (51), 53 (23), 43 (100)	-25
n-Butyryl-CoA **(2a)**	**2b**	14.46	155	155 (45) , 137 (22), 126 (100), 109 (10), 91 (55), 87 (32), 81 (77), 79 (25), 69 (23), 65 (3), 57 (27)	-25
Isovaleryl-CoA	**3c**	15.85	211	211 (2), 127 (100), 111 (13), 85 (56), 57 (58)	-20
(3a)	**3b**	16.21	169	169 (27), 126 (100), 105 (6), 87 (47), 69 (22), 67 (26), 43 (25)	-25
Hexanoyl-CoA	**4c**	16.80	225	225 (2), 141 (15), 127 (100), 99 (26), 71 (37), (69 (2)	-20
(4a)	**4b**	17.23	183	183 (100), 155 (22), 141 (3), 126 (86), 109 (41), 99 (15), 95 (58), 87 (7), 81 (40), 71 (18), 69 (26)	-25
Benzoyl-CoA	**5c**	15.76	231	231 (2), 189 (3), 157 (4), 147 (18), 105 (100), 71 (5)	-20
(5a)	**5b**	16.20	189	189 (72), 171 (61), 161 (28), 147 (10), 129 (55), 127 (75), 115 (63), 105 (100), 103 (11), 77 (37), 69 (25)	-25
Cinnamoyl-CoA	**6c**	17.01	257	257 (5), 221 (2), 173 (13), 131 (100), 103 (3), 69 (2)	-20
(6a)	**6b**	17.81	215	215 (26), 197 (17), 187 (14), 179 (7), 173 (9), 169 (40), 155 (18), 141 (77), 131 (100), 129 (15), 103 (16), 69 (28)	-25
p-Coumaroyl-CoA	**7c**	15.25	273	273 (10), 147 (100), 119 (18), 91 (2), 69 (2)	-25
(7a)	**7b**	15.78	231	231 (4), 213 (18), 185 (60), 171 (20), 157 (60), 147 (100), 121 (35), 69 (63)	-25

CE = collision energy.

[a]Retention times (RT) were obtained from the corresponding ion chromatogrames in the HPLC.

[b]compound **1c** was identified by selected reaction monitoring (SRM), the relative abundances were calculated from the peak areas of the three detected reactions

m/z 169. m/z ⟶ 127, m/z 169 ⟶ m/z 85, and m/z 169 ⟶ m/z 43.

3.3 Enzymatic products of *Rt*PKS

*Rt*PKS was found to accepted both aliphatic and aromatic CoA esters as primers for polyketide synthesis and produced multiple products from each. The major product formed in the *p*-coumaroyl-CoA assay was co-chromatographed with resveratrol as shown in **Figure 30,** suggesting that the *Rheum* PKS encoded a STS.

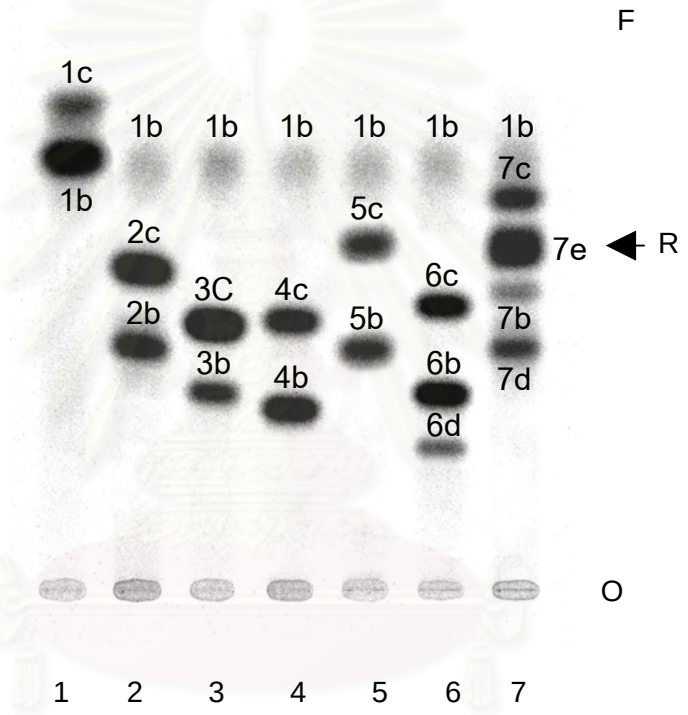


Figure 30 Reversed-phase TLC analysis of *Rt*STS reaction products. The reactions were performed with 1.5 μg of purified enzyme, [2-^{14}C] malonyl-CoA and acetyl-CoA (lane 1), *n*-butyryl-CoA (lane 2), isovaleryl-CoA (lane 3), *n*-hexanoyl-CoA (lane 4), benzoyl-CoA (lane 5), cinnamoyl-CoA (lane 6), *p*-coumaroyl-CoA (lane 7). O= TLC origin, F= solvent front, R= position of resveratrol standard.

To conclusively identify the polyketide products, *Rt*STS was incubated with unlabelled malonyl-CoA and the starter CoA esters in scaled up *in vitro* reactions and the resulting product mixtures were analyzed by LC-MS and LC-MS/MS. Products were identified by their parent ions and collision induced decomposition (CID) spectra. The chemical structures of the compounds identified by liquid chromatography electrospray ionization mass spectrometry (LC/ESI-MS) were summarized in **Table 7** and the corresponding mass spectral data were shown in **Table 8.**

*Rt*STS formed aromatic products only with *p*-coumaroyl-CoA and cinnamoyl-CoA. Incubation with *p*-coumaroyl-CoA gave resveratrol (**7e**) (*m/z* 229, $[M+H]^+$ and *m/z* 227, $[M-H]^-$) as the major product, demonstrating that *Rt*STS is an STS. Resveratrol was identified by its electrospray CID spectrum of the $[M-H]^-$ ion (*m/z* 227) displaying key ions at *m/z* 185 ($[M-H-CH_2CO]^-$) and *m/z* 143 ($[M-H-2CH_2CO]^-$) carparable with an authentic sample. Smaller amounts of BNY-type (**7b**) and CTAL-type (**7c**) pyrones, and naringenin chalcone (**7d**), were also detected in the *p*-coumaroyl-CoA assay. This spectrum of four products, including the naringenin chalcone formed as a cross-reaction product, has been previously described for *Arachis hypogaea* STS (Yamaguchi *et al*., 1999). These resulted suggesting that the *Rheum* PKS encoded a STS.

Table 7 Enzymatic product of *Rt*STS with different starter CoA esters

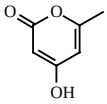
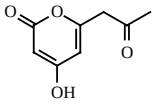
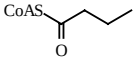
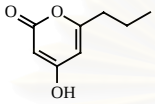
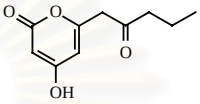
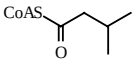
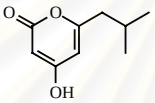
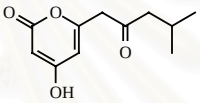
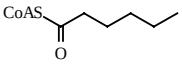
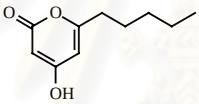
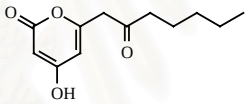
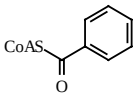
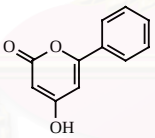
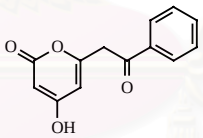
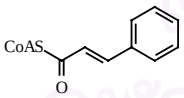
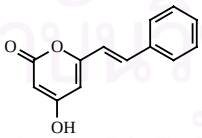
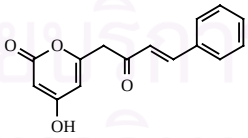
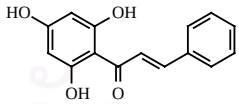
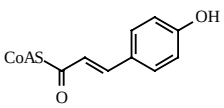
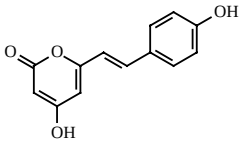
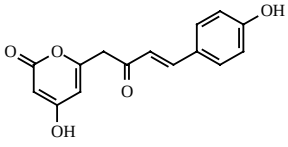
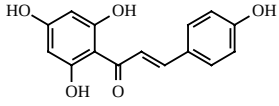
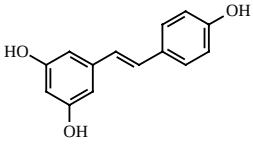
Starter-CoA	Enzymatic product		
	BNY-type	CTAL-type	Chalcone-type
 acetyl- CoA **(1a)**	 *m/z* 127 $[M+H]^+$ **(1b)**	 *m/z* 169 $[M+H]^+$ **(1c)**	
 *n*-butyryl-CoA **(2a)**	 *m/z* 155 $[M+H]^+$ **(2b)**	 *m/z* 197 $[M+H]^+$ **(2c)**	
 *Iso*valeroyl-CoA **(3a)**	 *m/z* 169 $[M+H]^+$ **(3b)**	 *m/z* 211 $[M+H]^+$ **(3c)**	
 Hexanoyl-CoA **(4a)**	 *m/z* 183 $[M+H]^+$ **(4b)**	 *m/z* 225 $[M+H]^+$ **(4c)**	
 Benzoyl-CoA **(5a)**	 *m/z* 189 $[M+H]^+$ **(5b)**	 *m/z* 231 $[M+H]^+$ **(5c)**	
 Cinnamoyl-CoA **(6a)**	 *m/z* 215 $[M+H]^+$ **(6b)**	 *m/z* 257 $[M+H]^+$ **(6c)**	 *m/z* 257 $[M+H]^+$ **(6d)**
 *p*-Coumaroyl-CoA **(7a)**	 *m/z* 231 $[M+H]^+$ **(7b)**	 *m/z* 273 $[M+H]^+$ **(7c)**	 *m/z* 273 $[M+H]^+$ **(7d)**
	 *m/z* 229 $[M+H]^+$ **(7e)**		

Table 8 Collision-induced dissociation (CID) mass spectra of enzymatically formed products from *Rt*STS, compound 1-7 obtained from the $[M+H]^+$ ions by LC/ESI-MS

Substrate	Product	RT (min)[a]	$[M+H]^+$ (m/z)	Ions in the CID mass spectra (m/z, rel.int.)	CE (eV)
Acetyl-CoA	**1c**	10.02	169[b]	169, 127 (100), 85 (50), 43 (23)	-20
(1a)	**1b**	11.07	127	127 (47), 109 (27), 99 (18), 87 (22), 71 (21), 69 (68), 67 (18), 65 (39), 53 (20), 43 (100)	-25
n-Butyryl-CoA	**2c**	14.38	197	197 (1), 127 (100), 113 (8), 71 (46), 43 (25)	-20
(2a)	**2b**	14.78	155	155 (100) , 137 (12), 126 (68), 109 (15), 91 (40), 87 (41), 81 (55), 79 (45), 69 (70), 65 (18)	-25
Isovaleryl-CoA	**3c**	15.50	211	211 (2), 127 (100), 111 (14), 85 (55), 57 (57)	-20
(3a)	**3b**	15.81	169	169 (54), 126 (100), 105 (19), 87 (31), 69 (23), 67 (19), 43 (25)	-25
Hexanoyl-CoA	**4c**	16.77	225	225 (14), 141 (6), 127 (100), 99 (27), 71 (12)	-15
(4a)	**4b**	17.19	183	183 (100), 155 (9), 141 (4), 126 (48), 109 (16), 99 (14), 95 (25), 87 (26), 81 (42), 71 (28), 69 (35), 67 (30), 57 (10)	-25
Benzoyl-CoA	**5c**	15.13	231	231 (2), 147 (18), 111 (4), 105 (100)	-20
(5a)	**5b**	15.57	189	189 (100), 171 (17), 161 (10), 127 (4), 115 (8), 105 (24), 103 (4), 69 (11)	-20
Cinnamoyl-CoA	**6c**	16.44	257	257 (5), 221 (2), 173 (8), 131 (100), 103 (3), 69 (2)	-20
(6a)	**6b**	17.23	215	215 (100), 197 (13), 187 (40), 179 (3), 169 (35), 155 (12), 141 (22), 131 (60), 129 (6), 69 (7)	-20
	6d	18.64	257	257 (25), 153 (100), 131 (55), 123 (7), 103 (11), 77 (6), 69 (5)	-25
p-Coumaroyl-CoA	**7c**	15.21	273	273 (2), 189 (5), 147 (100), 119 (15)	-25
	7e	15.41	227[c]	227 (63), 185 (87), 183 (68), 159 (25), 143 (100)	+25
(7a)	**7b**	15.74	231	231 (35), 213 (12), 203 (18) 171 (12), 157 (32), 147 (100), 121 (10), 110 (5), 69 (4)	-20
	7d	17.21	273	273 (15), 255 (9), 153 (100), 147 (7), 123 (4), 85 (5)	-25

CE = collision energy.

[a]Retention times (RT) were obtained from the corresponding ion chromatogrames in the HPLC.

[b]compound **1c** was identified by selected reaction monitoring (SRM), the relative abundances were calculated from the peak areas of the three detected reactions m/z 169 ⟶. m/z 127, m/z 169 ⟶ m/z 85, and m/z 169 ⟶ m/z 43.

[c]compound **7e** was identified by CID of the $[M-H]^-$ at the m/z 227.

4. Optimum temperature of the expressed PKS

The optimum temperature optimum for the *Cal*CHS, *Pin*PKS and *Rt*STS activities of purified recombinant enzymes were investigated. The results were shown in **Figure 31**. *Cal*CHS has a optimum temperature at 35 ^{o}C for the synthesis of naringenin. *Pin*PS has a optimum temperature at 30 ^{o}C for the synthesis of pyrone product. *Rt*STS has a optimum temperature at 35 ^{o}C for the synthesis of resveratrol.

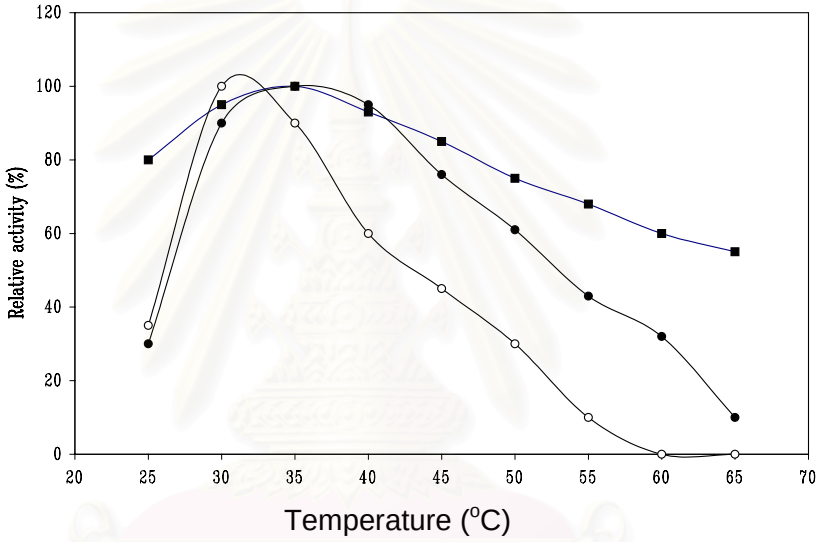


Figure 31 Optimum temperature for polyketide assay. A, *Cal*CHS (■), B, *Pin*PS (○) and C. *Rt*STS (●) .

5. Optimum pH of the expressed PKS

Optimum pH of each purified enzyme was shown in **Figure 32**. *Cal*CHS has a optimum pH for naringenin synthesis at 7.0. *Pin*PS has a optimum pH for pyrone synthesis at 8.0. *Rt*STS has a optimum pH for resveratrol synthesis at 7.0.

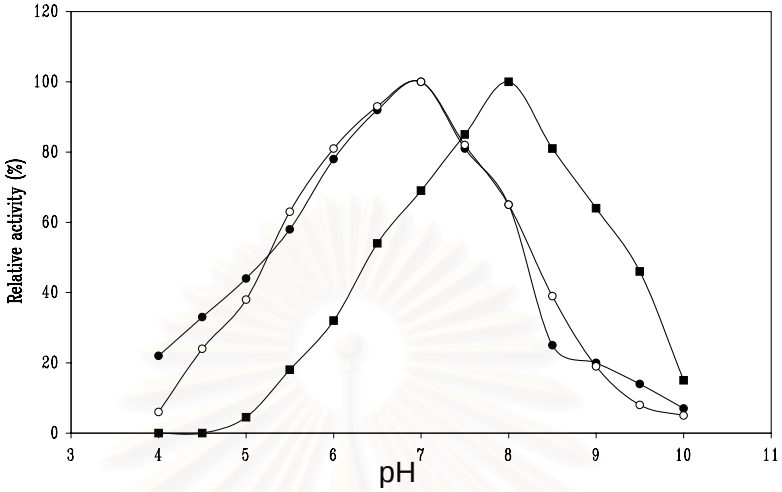


Figure 32 Optimum pH for polyketide assay. A, *Cal*CHS (■), B, *Pin*PS (○) and C, *Rt*STS (●) .

6. Tissue-specific expression of *Calchs, Pinps* and *Rtsts*

Total RNAs isolated from the leaves, stems and roots of greenhouse-grown *C. alata* were all subjected to RNA gel blot analysis. The full-length gene of *Cal-chs*1-3 was used as hybridization probe. RNA gel blot analysis of gene transcript accumulation in various organs of mature *C. alata* plant indicated that *Calchs* transcript is present predominantly in young and old root, and at higher levels in young leaf than in stem. However, in old leaf present at lower level **(Figure 33).** This corresponds to high accumulation of the flavonoids, naringenin, kampferol, quercetin in roots, as shown in **Figure 34**.

The flavonoid content in leaves and roots of *C. alata* was analyzed as described in section 20.1 of Materials and Methods. The elution patterns are shown in **Figure 34,** The retention times of the peaks in **Figure 34 B** and **34 C** were compared with those of known authentic flavonoids **Figure 34 A**. The retention times of peaks at

13.6, 20.9 and 24.9 min were identical to quercetin, naringenin and kaempferol, respectively. These data indicated that the flavonoids were accumulated mostly in the leaves and roots of *C. alata,* corresponding to the tissue specific expression of all *Calchs* which were highly expressed in the roots and moderately in the leaves as evidenced in northern blot analysis **(Figure 33)**.

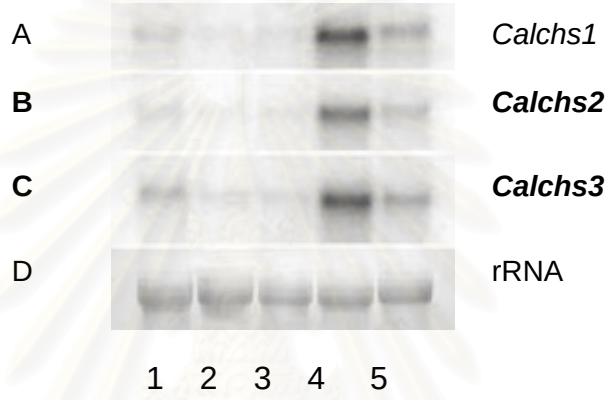


Figure 33 Expression of *Calchs* in leaves, stems and roots of *C. alata*. Twenty μg of total RNA were hybridized with [α-^{32}P]-dATP-labelled full-length *Calchs*1-3. The application of equal amounts of RNA was achieved by the comparison of ribosomal RNA from the individual samples (lower lane). Lane1; young leaves; Lane 2, mature leaves; Lane 3, stem; Lane 4, young roots and Lane 5, mature roots.

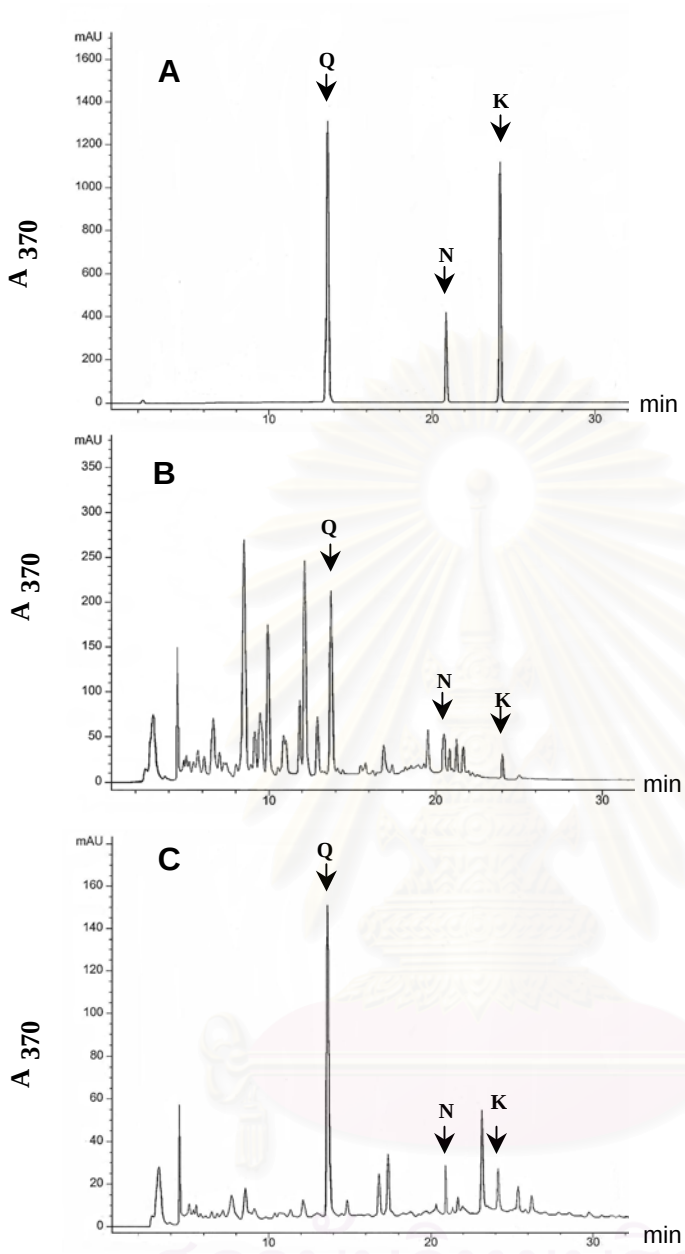


Figure 34 Analytical reversed-phase HPLC profiles of flavonoids in *C. alata* plants. Panels show chromatograms of A, authentic standard (Q, quercetin, R_t = 13.6 min; N, naringenin, R_t = 20.9 min; K, kaempferol, R_t = 24.9 min); B, leaves extract, and C, root extract.

The accumulation of *Calchs* mRNA during the seedling development was then monitored. In doing this, RNA was isolated from developing seedling at 0.5, 1, 2, 3, 4, and 5 cm-length after germination to analyze the temporal accumulation of *Calchs* gene transcript. The transcript of *Calchs* was present at a low level in the seedling at 0.5 and 1.0 cm-length after germination, and then increases during the growth up to 5 cm-length and in mature root **(Figure 35)**.

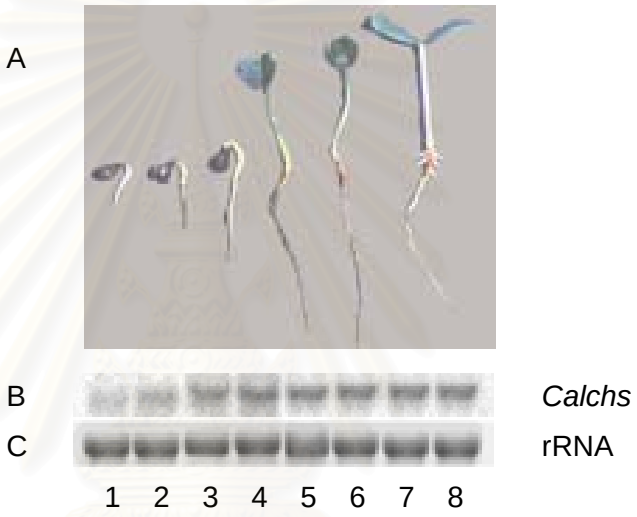


Figure 35 Accumulation of *Calchs* gene in *C. alata* seedling. A, developing *C. alata* seedling length, Lane 1-6 was 0.5, 1, 2, 3, 4 and 5 cm, respectively, Lane 7-8, mature root; B*, Calchs* transcription and C, rRNA control.

For tissue-specific expression of *Pinps,* the highest levels of *Pinps* transcripts were found in young roots and moderately in old roots. Low levels of *Pinps* transcript were present in young leaves, old leaves and stems, respectively, as shown in **Figure 36**. This result indicates the roots as the source of mRNA of *P. indica Pinps*.

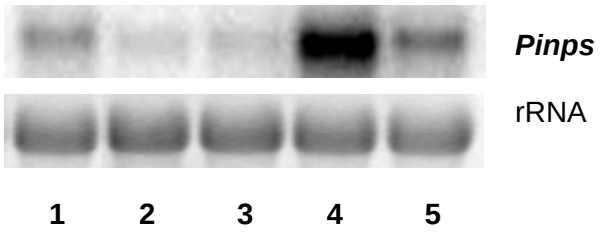


Figure 36 Expression of *Pinps* in leaves, stems and roots of *P. indica*. Twenty μg of total RNA were hybridized with [α-^{32}P]-dATP-labelled full-length *Pinps*. The application of equal amounts of RNA was achieved by comparison of ribosomal RNA from individual samples (lower lane). Lane1, young leaves; Lane 2, old leaves; Lane 3, stem; Lane 4, young roots and Lane 5, old roots.

For tissue-specific expression of *Rtsts*, the highest levels of *Rtsts* transcripts were found in roots. Low levels of *Rtsts* transcript were present in leaves and petioles, as shown in **Figure 37**. This corresponds to the high accumulation of stilbene (resveratrol) in roots, as shown in **Figure 38**.

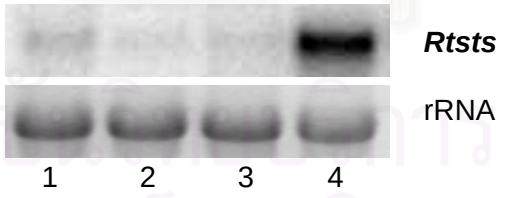


Figure 37 Expression of *Rtsts* in leaves, petiole and roots of *R. tataricum*. Twenty μg of total RNA were hybridized with [α-^{32}P]-dATP-labelled full-length of *Rtsts*. The application of equal amounts of RNA was achieved by comparison of ribosomal RNA from individual samples (lower lane). Lane1, young leaves; Lane 2, old leaves; Lane 3, petiole and Lane 4, roots.

Analysis of resveratrol as a component in the rhizomes of *R. tataricum* was performed as described in section 20.1 of Materials and Methods. The elution patterns were monitored as shown in **Figure 38**. The retention time of peak at 17.8 min was identical to that of resveratrol. This was confirmed by a co-injection experiment with authentic standard. This data indicated that the highest level of resveratrol accumulated in the rhizomes of *R. tataricum,* which corresponded to the tissue specific expression of *Rtsts* as the highest expressed in roots as shown in northern blot analysis **(Figure 37)**.

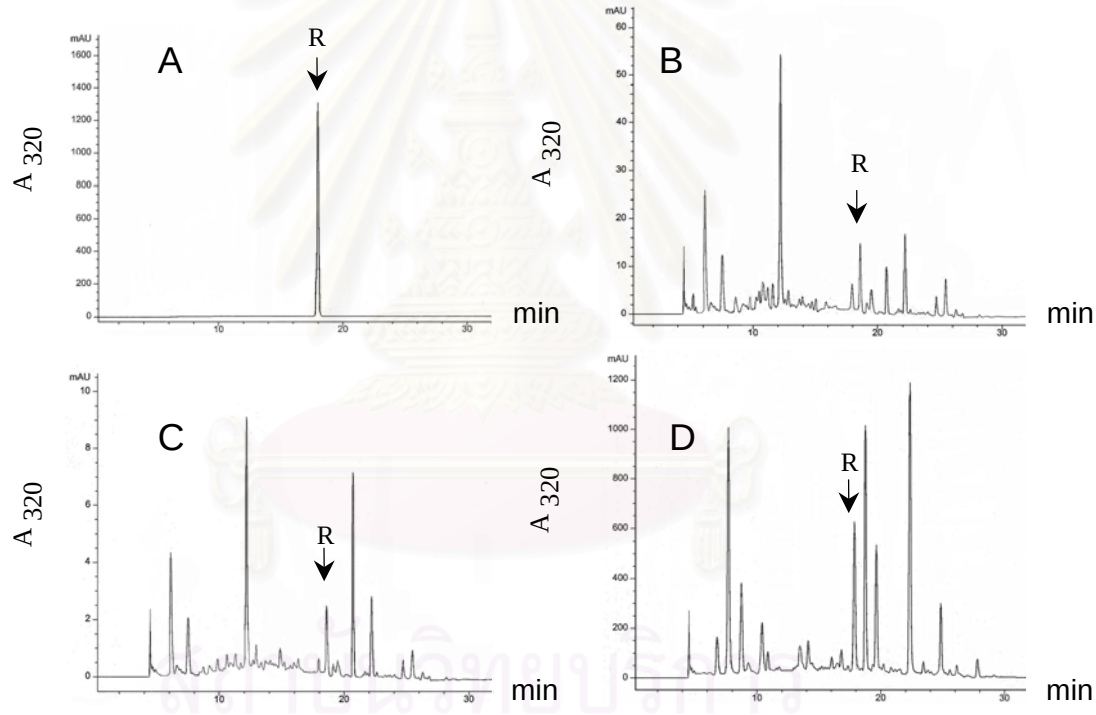


Figure 38 Analytical reversed-phase HPLC separation profiles of resveratrol in *R.tataricum* plants. Panels show chromatograms of A, authentic standard (R, resveratrol, R_t = 17.8 min); B, leaves extract; C, stems extract; and D, roots extract.

CHAPTER V

DISCUSSION

Plants contain a group of polyketide synthases (PKS) that appear to be unique in several aspects and they are often called PKSIII because they do not really fit the definitions of PKSI and PKSII. The basic mechanisms of their condensing reactions, however, follow the general model developed for all condensing enzymes (Schroder, 2000). In this investigation of PKSs type III that occur in Thai medicinal plants. We have isolated three cDNAs that encode chalcone synthase (CHS) from *C. alata*, one cDNA encoding stilbene synthase (STS) from *R. tataricum* and one cDNA encoding pyrone synthase from *P. indica*. CHS and STS are the most well-known members of this plant PKS superfamily. CHS catalyzes the first committed step in the biosynthesis of a large number of biologically important substances such as flavonoids (flower colour) and phytoalexins (defense against pathogens). STS forms the backbone of the stilbene phytoalexins; these enzymes are rare in higher plants. CHS and STS are plant-specific polyketide synthases and we have found in this study that the gene transcripts of both CalCHS and RtSTS accumulate in roots (Figure 33 and 37). This expression pattern correlates with the occurance of flavonoids quercetin, naringenin and kaempferol in *C. alata* roots (Figure 34) and of resveratrol in *R. tataricum* rhizomes (Figure 38). Structurally, the sequences of the two enzymes are closely related (70% homology). It has been proposed that STS in higher plants evolved from CHS (Tropf *et al*., 1994).

For substrate utilization, our heterologously expressed cDNAs *Cal*CHS and *Rt*STS show broad substrate specificity for non-physiological substrates and produce multiple aromatic and pyrone derailment products from each substarte. So far, CHS and STS have been only the well-known members of a superfamily of related proteins that use the same type of condensing reaction, but with widely varying substrates (e.g. phenylpropanoid-CoA esters, benzoyl-CoA derivatives, linear CoA-esters) and serving quite different pathways in natural product biosynthesis (For review see Schroder, 1999). There are also CHS/STS-type enzymes programmed for only one or two condensation reactions (Schroder, 1997). Taken together, the use of different starter substrates, modifications of intermediates, and programming for one, two, or three condensing reactions indicates a large variety of products that can be synthesized by proteins of the CHS/STS-type enzyme superfamily (Schroder, 1999; 2000).

We have isolated one cDNA encoding pyrone synthase (PS) from *P. indica*. PS is the member of this plant PKS superfamily. PS catalyzes the condensation of acetyl-CoA with two molecules of malonyl-CoA in the formation of gerberin and parasorboside in *G. hybrida* that has been previous described in Eckermann *et al*., 1998. We have found in this study that the gene transcripts of *Pin*PS accumulate in *P. indica* roots (Figure 36). Structurally, the sequences of this enzyme is closely related to CHSs (60% homology). It has been proposed that PS in higher plants evolved from CHS (Eckermann *et al*., 1998).

For substrate utilization, our heterologously expressed cDNA *Pin*PS show a broad substrate specificity for non-physiological substrates and produce 6-methyl-4-hydroxy-2-pyrone (triacetic acid lactone) **(1b)** (*m/z* 127 $[M+H]^+$) as the main product, in addition to 6-acetonyl-4-hydroxy-2-pyrone (tetracetic acid lactone) **(1c)** (*m/z* 169 $[M+H]^+$) also produced, when *Pin*PS was incubated with acetyl-CoA as starter-CoA and malonyl Co-A as extender unit.

1. Isolation of cDNAs encoding *C. alata*, *P. indica* and *R. tataricum* polyketide synthases

A homology-based approach was used successfully in this study to amplify gene fragments from all the three species of *C. alata*, *P. indica*, and *R. tataricum* by RT-PCR. The roots of these plants were targeted because the polyketides of interest, including anthraquinones, are accumulated in the tissue of each plant. The resulting PCR products showed similarity to CHS superfamily at the nucleotide level. For *C. alata* and *P. indica*, their partial clones were used as a hybridization probe to screen cDNA libraries in λ-ZAP II prepared from root mRNAs, whereas for *R. tataricum*, the cDNA regions 5' and 3' from the core fragment were obtained using RACE PCR. In this manner, three clones that encoded complete reading frames were isolated from *Cassia* cDNA, one clone from *Plumbago* cDNA and one full-length *Rt*STS cDNA were obtained. Translation of the nucleotide sequences of the three *Cassia* cDNAs indicated that they were highly similar, but not-identical **(Figure 19)**.

A comparison of the translation of the nucleotide sequences of these cDNAs with those present in the GenBank/EMBL databases indicated that the *Cassia* genes are most similar to chalcone synthase from other members of the Leguminosae, especially *G. max* (Akada *et al*., 1993) and *P. vulgaris* (Ryder *et al*., 1987) **(Table 2).** The complete nucleotide sequences of the three cDNAs *Cal*PKS1, *Cal*PKS2 and *Cal*PKS3 were determined as containing 1,554, 1,536 and 1,590 bp encoding complete reading frames of 389, 390 and 389 amino acid residues, respectively. The predicted molecular masses for the encoded proteins are 42.6, 42.9 and 42.6 kDa, respectively. In the *Plumbago* gene is most similar to chalcone synthase from *G. hybrida* with 60% identity (Helariutta *et al*., 1995) and *G. max* with 60% identity (Akada *et al*., 1993). The complete nucleotide sequences of the three cDNAs *Pin*PKS was determined as containing 1,622 bp encoding complete reading frames of 396 amino acid residues.

For *R. tataricum*, a full-length *Rt*STS cDNA was 1,433 bp encoding a protein of 391 amino acids with a predicted molecular weight of 43.0 kDa. The 59 bp 5' untranslated region contained a stop codon in-frame with the presumed start codon indicating the cDNA was full length. A comparison of the translation of the nucleotide sequences of these cDNAs with those present in the GenBank/EMBL databases indicated that the *R. tataricum* genes are most similar to *R. palmatum* BAS and *R. graveolens* acridone synthase (ACS).

2. Functional expression and characterization of *C. alata, P. indica and R. tataricum* polyketide synthases

The open reading frames of the putative polyketide synthases from *Cassia, Plumbago and Rheum,* were ligated into a T7 expression vector containing an amino terminal hexahistidine fusion tag. The proteins were then purified from *E. coli* BL21(DE3) in two steps (sonication/centrifugation, cobalt affinity-chromatography) to yield electrophoretically homogeneous enzyme of approximately 42-43 kDa. The purified enzyme was used in enzyme assays to determine the substrate/product specificity of the putative polyketide synthases.

A series of seven coenzyme A esters were tested together with [2-^{14}C]malonyl-CoA for their capacity to serve as polyketide synthase "starter" molecules. The compounds used were acetyl-, *n*-butyryl-, isovaleryl-, *n*-hexanoyl-, benzoyl-, cinnamoyl-, and *p*-coumaroyl-CoA. In each case, multiple products were formed as determined by phosphorimagery following resolution of the reaction mixtures by TLC. This broad product spectrum had been reported previously for several of the CoA-esters with chalcone synthase, for example, from *Pinus sylvestris*, *P. strobus* and *Sinapis alba* (Zuurbier *et al*., 1998) *Scutellaria baicalensis* (Morita *et al*., 2000) and from raspberry *Rubus idaeus* (Zheng *et al*., 2001). For STS has been reported previously from *A. hypogaea* (Yamaguchi *et al*., 1999; Morita *et al*., 2001).

In order to thoroughly characterize the heterologously expressed *Cassia, Plumbago and Rheum* cDNAs, the reactions were repeated with unlabeled substrates on a larger scale so that the reaction products could be analyzed by mass spectrometry. The chemical structures of the compounds identified by liquid chromatography electrospray ionization mass spectrometry (LC/ESI-MS) and the corresponding mass spectral data are summarized in **section 3.1, 3.2 and 3.3** of Results, respectively.

*Cal*CHS and *Rt*STS formed aromatic products only with *p*-coumaroyl-CoA and cinnamoyl-CoA. When *Cal*CHS incubation with *p*-coumaroyl-CoA gave naringenin chalcone **(7d)** (*m/z* 273, $[M+H]^+$)as the major product, demonstrating that *Cal*CHS is a CHS. Naringenin chalcone was identified by its electrospray CID spectrum of the $[M-H]^+$ ion (*m/z* 273) displaying key ions at *m/z* 153 ($[M+H-CH_2CO]^+$) and *m/z* 147 ($[M+H-CO]^+$) comparable with an authentic sample. Smaller amounts of BNY-type **(7b)** and CTAL-type **(7c)** pyrones were also detected in the *p*-coumaroyl-CoA assay. This spectrum of a reaction product, has been previously described for *S. baicalensis* CHS (Morita *et al*., 2000).

For cinnamoyl-CoA, *Cal*CHS synthesized BNY-type **(6b)** *m/z* 215, $[M+H]^+$, CTAL-type **(6c)** *m/z* 257, $[M+H]^+$, as well as a small amount of pinocembrin chalcone **(6d)** *m/z* 257, $[M+H]^+$ from cinnamoyl-CoA. Pinocembrin chalcone was clearly identified by its parent ion (*m/z* 257 $[M+H]^+$) and the typical collision-induced fragments at *m/z* 153 (trihydroxybenzoyl ion) and *m/z* 131 (cinnamoyl ion).

For *Rt*STS, when incubation with *p*-coumaroyl-CoA gave resveratrol **(7e)** (*m/z* 229, $[M+H]^+$ and *m/z* 227, $[M-H]^-$) as the major product, demonstrating that *Rt*STS is a STS. Resveratrol was identified by its electrospray CID spectrum of the $[M-H]^-$ ion (*m/z* 227) displaying key ions at *m/z* 185 ($[M-H-CH_2CO]^-$) and *m/z* 143 ($[M-H-2CH_2CO]^-$) in

comparison with an authentic sample. Smaller amounts of BNY-type **(7b)** and CTAL-type **(7c)** pyrones, and naringenin chalcone **(7d),** were also detected in the *p*-coumaroyl-CoA assay. This spectrum of four products, including the naringenin chalcone formed as a cross-reaction product, has been reported previously for *A. hypogaea* STS (Yamaguchi *et al*., 1999).

*Rt*STS synthesized BNY-type **(6b)** and CTAL-type **(6c)** pyrones as well as a small amount of pinocembrin chalcone **(6d)** and pinosylvin **(6e)** from cinnamoyl-CoA. Pinocembrin chalcone was clearly identified by its parent ion (*m/z* 257 $[M+H]^+$) and the typical collision-induced fragments at *m/z* 153 (trihydroxybenzoyl ion) and *m/z* 131 (cinnamoyl ion), while pinosylvin **(6e)** was detected by the LC-MS analysis showing a peak with masses of *m/z* 213 ($[M+H]^+$) and *m/z* 211 ($[M-H]^-$). Only three products were visible in the ^{14}C-labeled assay with cinnamoyl CoA , the least polar (lower R_f value) of which likely corresponds to an unresolved mixture of pinocembrin chalcone and pinosylvin.

It is interesting that when incubated *Cal*CHS with substarte starter-CoA, *n*-butyryl- **(2a),** isovaleryl- **(3a),** *n*-hexanoyl-**(4a),** benzoyl-**(5a),** cinnamoyl-**(6a),** and *p*-coumaroyl-CoA **(7a)** with malonyl-CoA as extender molecule. It can be produced small amount of chalcone-type compound, **2d, 3d, 4d, 5d, 6d and 7d,** respectively, as shown in **section 3.1** of Results. This results as the same has been reported previously for *S. baicalensis* CHS (Morita *et al*., 2000). However, when incubated all the substrate CoA and malonyl-CoA with *Pin*PS and *Rt*STS, it can not formed the chalcone-type compound *in vitro* reaction as shown in **section 3.2 and 3.3** of Results.

It should be note that when *Cal*CHS, *Pin*PS and *Rt*STS were incubated with all the substrate starter Co-A and malonyl-CoA, the two type of derailment pyrone product can be formed, one is capable to perform the minor *in vitro* product were BNY-type **(1b, 2b, 3b, 4b, 5b. 6b and 7b)**, suggesting the structure of the BNY-type compound, indicating the reaction had terminated after two condensation reaction of malonyl-CoA and CTAL-type **(1c, 2c, 3c, 4c, 5c, 6c and 7c)** compounds, suggesting the structure of the CTAL-type compound, indicating the reaction had terminated after three condensation reaction of malonyl-CoA respectively, The MS/MS and the fragment showed in the **Table 4,6** and **8**, confirming the presence of α-pyrone ring. The enzyme reaction was therefor terminate without aromatic ring formation. For CHS and STS activity has been reported previously (Morita *et al*., 2001; Yamaguchi *et al*., 1999).

It was remarked for the first time that the recombinant *Cal*CHS, *Pin*PS and *Rt*STS also show the formation of 6-acetonyl-4-hydroxy-2-pyrone (tetracetic acid lactone) **(1c)** (*m/z* 169 $[M+H]^+$), in addition to 6-methyl-4-hydroxy-2-pyrone (triacetic acid lactone) **(1b)** (*m/z* 127 $[M+H]^+$), when *Cal*CHS, *Pin*PS and *Rt*STS, was incubated with acetyl-CoA is notable. Compound **(1b)** has been previously described as the main product formed from the use of acetyl-CoA as a starter CoA ester by *G. hybrida* 2-pyrone synthase (Eckermann *et al*., 1998) and CHS from *P. crispum* (Schuz *et al*., 1983). 6-Acetonyl-4-hydroxy-2-pyrone (**1c**) is a new enzymatic product from a plant PKS. Its synthesis is analogous to the formation of *p*-coumaroyl triacetic lactone and likely involves the condensation of acetyl-CoA with three C_2 units derived from malonyl-CoA followed by a pyrone-type cyclization. This compound has not been reported previously from plants, although it has been isolated as a polyketide metabolite of *Penicillium stipatatum* (Bentley and Zwitkowits, 1967).

With some notable exceptions, the exclusive formation of pyrones by plant PKSs seems to be an indicator that the starter CoA ester in the reaction is not the physiologically relevant substrate for the enzyme under assay. This phenomenon was observed in the assays of *Cal*CHS, *Pin*PS and *Rt*STS with aliphatic and benzoyl CoA substrates. Pyrones also form as byproducts in reactions with correct substrates but the amounts are generally minor compared to the main aromatic product. In a limited number of cases, pyrone formation appears to be the true catalytic activity of plant PKS enzymes. 2-Pyrone synthase from *G. hybrida* forms 6-methyl-4-hydroxy-2-pyrone as its major product *in vitro*, with the involvement of this enzyme in pyrone metabolism supported by antisense “knockout” of 2-pyrone synthase in *G. hybrida* (Eckermann *et al*., 1998). A *p*-coumaroyl triacetic acid synthase from *H. macrophylla* may also represent an example of a pyrone-forming PKS, although this enzyme has been suggested to yield a linear tetraketide *in vitro* (Akiyama *et al*., 1999)

3. Tissue-specific expression of *Calchs, Pinps* and *Rtsts*

The accumulation of the *Cassia* polyketide synthase transcript was investigated in developing seedlings and plants of *C. alata*. Since the nucleotide sequences of the three polyketide synthase clones were more than 90% identical, no attempt was made to differentially detect the transcripts. RNA was isolated from seedlings of 0.5-5 cm in length. RNA gel blot analysis indicated that transcript was already present in 0.5 cm long seedlings. As the seedlings grow, the level of polyketide synthase transcript increases. A 5 cm long seedlings contains approximately the same level of transcript as mature root.

In the mature *C. alata* plant, the highest level of polyketide synthase transcript was found in young root, whereas much less accumulated in leaf and stem. It was surprising to find a root-specific polyketide synthase, with a substrate and product spectrum similar to that of chalcone synthase, in a plant known to accumulate anthraquinones in both aerial and underground tissue. To investigate the flavonoid composition of leaf and root tissue, an acidic methanol extract was prepared and was analyzed by reversed-phase HPLC. Quercetin was found to be a major component of the root and was also present as part of a more complex mixture in the leaf. Given the distribution of gene transcript and of flavonoids, it is reasonable to assume that the *C. alata* root-specific polyketide synthases are chalcone synthases that demonstrate broad substrate specificity *in vitro*.

For *Pin*PS, the expression pattern of the *Pinps* transcript was determined by northern blot analysis of total RNA isolated from young leaves, mature leaves, young roots and mature roots. Young leaves showed low expression levels, and the transcript was virtually absent from older leaves and mature roots. High expression level of *Pinps* was detected in roots where plumbagin levels were also highest (data not shown). Given the distribution of gene transcript, it is reasonable to assume that the *P. indica* root-specific polyketide synthases are pyrone synthases that demonstrate broad substrate specificity *in vitro*.

For *Rt*STS, to confirm that stilbenes were present in the tissue of *R. tataricum.* Acid hydrolysis of extracts was performed to cleave sugar residues and release stilbenes as their aglycones. Resveratrol was detected in all tissues, with higher amounts occurring in the rhizomes. To the best of our knowledge, this is the first report of resveratrol from *R. tataricum*. The presence of this compound provides further

evidence that the *in vivo* function of *Rt*STS is the biosynthesis of resveratrol. The distribution of resveratrol correlated with the expression pattern of the *Rtsts* transcript as determined by northern blot analysis of total RNA isolated from young leaves, mature leaves, petioles and rhizomes. Young leaves showed low expression levels, and the transcript was virtually absent from older leaves and petioles. High expression level was detected in rhizomes where resveratrol levels were also highest.

4. Phylogenetic comparisons of the deduced polypeptide sequence of *Cal*CHS, *Pin*PS and *Rt*STS

A comparison with a CHS consensus from CHS superfamily with established enzyme activity as shown in **Figure 23 and 24**, indicated that the *Cassia, Plumbago and Rheum* proteins confirmed to the consensus, with only a few exceptions. However, Jez *et al*. (2000a) compared the crystal structure of *M. sativa* chalcone synthase to that of *G. hybrida* pyrone synthase and found that three catalytic residues are present in the active sites of both type III plant polyketide synthases (Cys 164, His 303, Asn 336 in chalcone synthase and Cys 169, His 308, Asn 341 in pyrone synthase). The catalytic mechanism for polyketide formation is expected to be equivalent in both enzymes. Pyrone synthase, however, has a smaller initiation/elongation cavity that cannot accommodate the larger starter molecule *p*-coumaroyl CoA that is required for naringenin formation. Four amino acid residues have been identified as important to cavity size (Thr 197, Ile 254, Gly 256, Ser 338 in chalcone synthase and Leu 202, Met 259, Leu 261, Ile 343 in pyrone synthase). The *Cal*CHS cDNAs contain catalytic and initiation/elongation cavity residues identical to those of *M. sativa* chalcone synthase. Jez *et al*. (2000a) has been reported that the methylpyrone formation by the *M. sativa* chalcone synthase, which is consistent with results reported herein. From sequence comparison the initiation/elongation cavity of

*Cal*CHS is expected to be as large as that of the *M. sativa* chalcone synthase. It is therefore not surprising that starter molecules that are smaller than *p*-coumaroyl CoA can be accommodated by the enzyme with varying degrees of efficiency. The relative amount of products should depend on the assay conditions employed. Formation of multiple products from multiple starter molecules *in vitro* is likely a result of the large initiation/elongation cavity of *Cal*CHS and does not necessarily suggest *in vivo* activity. Additional three dimensional structure determinations for type III polyketide synthases with varying substrate and product specificities will certainly provide insight into the efficiency of various starter molecules and the formation of derailment products for this important class of enzymes.

CONCLUSION

From this research work of “Cloning and expression of plant polyketide synthase genes from *Cassia alata*, *Plumbago indica*, and *Rheum tataricum*.” The following conclusion can be drawn :

1. As part of our investigation of type III polyketide synthases that occur in Thai medicinal plants, we have isolated three cDNAs that encode chalcone synthase from roots of *C. alata* . The gene transcripts were found predominantly in roots, which also accumulate the flavonoids quercetin, naringenin and kaempferol. The heterologously expressed cDNAs show broad substrate specificity for non-physiological substrates and produce multiple aromatic and pyrone derailment products from each substrate. The enzymes can accept CoA-esters that are not derived from phenylpropanoid metabolism and convert them to compounds that are not known to occur in the *Cassia* plant.

2. One full-length cDNA clone from roots of *P. indica* encoding PS. The gene transcripts were found predominantly in roots, which also accumulate the naphthoquinone, plumbagin. The heterologously expressed cDNAs show broad substrate specificity in *in vitro* assay and produce pyrone products from each substrate.

3. One full-length cDNA clone from roots of *R. tataricum* encoding STS. The gene transcripts were found predominantly in rhizomes, which also high accumulate the stilbene, resveratrol. The heterologously expressed cDNAs show broad substrate specificity for non-physiological substrates and produce multiple aromatic and pyrone derailment products from each substrate. *Rt*STS synthesized resveratrol and a trace amount of naringenin chalcone from *p*-coumaroyl-CoA, supporting the enzyme's identification as a resveratrol-type stilbene synthase. Bis-noryangonin and *p*-coumaroyl triacetic acid lactone (CTAL)-type pyrones were observed in minor amounts in the reaction with *p*-coumaroyl-CoA. The enzymes can accept CoA-esters that are not derived from phenylpropanoid metabolism and convert them to compounds that are not known to occur in the *Rheum* plant.

REFERENCES

Abe, I., Morita, H., Nomura, A., and Noguchi, H. 2000. Substrate specificity of chalcone synthase: enzymatic formation of unnatural polyketides by chalcone synthase. **J. Am. Chem. Soc**. 122: 11242-11243.

Abe, I., Takahashi, Y., Morita, H., and Noguchi, H. 2001. Benzalacetone synthase. A novel polyketide synthase that plays a crucial role in the biosynthesis of phenylbutanones in *Rheum palmatum*. **Eur. J. Biochem**. 268: 3354-9.

Akada, S, Kung, S.D., and Dube, S.K. 1991. The nucleotide sequence of gene 1 of the soybean chalcone synthase multigene family. **Plant Mol Biol.** 16: 751-752

Akada, S, Kung, S.D., and Dube, S.K. 1993. Nucleotide sequence and putative regulatory elements of gene 2 of the soybean (*Glycine max*) chalcone synthase multigene family. **Plant Physiol.** 102: 317-319

Akashi, T., Furuno, T., Futami, K., Honda, M., Takahashi, T., Welle, R., and Ayabe, S.I. 1996. A cDNA for polyketide reductase (accession no. D83718) that catalyzes the formation of 6'-deoxychalcone from cultured *Glycyrhiza echinata* L. cells. **Plant Physiol.** 111: 247

Akiyama, T., Shibata, M., Liu, H.M., and Ebizuka, Y. 1999. *p*-coumaroyltriacetic acid synthase, a new homologue of chalcone synthase, from *Hydrangea macrophylla* var. *thunbergii*. **Eur. J. Biochem**. 263: 834-839.

Altschul S.F. and Lipman D.J. 1990. Protein database searches for multiple alignments. **Proc Natl Acad Sci USA**. 87(14):5509-5513.

Atanassove, I., Russinova, E., Antonov, L., and Atanassov, A. 1998. Expression of an anther-specific chalcone synthase-like gene is correlated with uninucleate microspore development in Nicotina sylvestris. **Plant Mol. Biol**. 38: 1169-1178.

Ayabe, S.I., Udagawa, A., and Furuya, T. 1988. NAD(P)H-dependent 6'-deoxychalcone synthase activity in *Glycyrhiza echinata* cells induces by yeast extract**.** **Arch. Biochem. Biophys**. 261: 458-462.

Balance, G.M., and Dixon, R.A., 1995. *Medicago sativa* cDNA encoding chalcone reductase. **Plant physiol**. 107: 1027-1028.

Bangera, M.G., and Thomashow, L.S. 1996. Characterization of a genomic locus required for synthesis of the antibiotic 2,4-diacetylphloroglucinol by the biological control agent *Pseudomonas fluorescens* Q2-87. **Mol. Plant-Microbe Interact.** 9: 83-90.

Baumert, A., Maier, W., Groger, D., and Deutzmann, R. 1994. Purification and properties of acridone synthase from cell suspension cultures of *Ruta graveolens*. **Z. Naturforsch**. 49c: 26-32.

Baumert, A., Porzel. A., Schmidt, J., and Groger, D. 1992. Formation of 1,3-dihydroxy-N-methylacridone from N-methylanthraniloyl-CoA and malonyl-CoA by cell-free extracts of *Ruta graveolens*. **Z. Naturforsch**. 47c: 365-368.

Beckert, C., Horn, C., Schnitzler, J.-P., Lehning, A., Heller, W., and Veit, M. 1997. Styrylpyrone biosynthesis in *Equisetum arvense* L. **Phytochemistry.** 44: 275-283.

Beerhues, L. 1996. Benzophenone synthase from cultured cells of *Centarium erythraea*. **FEBS Lett**. 383: 264-266.

Beerhues, L., and Berger, U. 1994. Xanthones in cell suspension cultures of two *Centauricum* species. **Phytochemistry.** 35: 1227-1231.

Beerhues, L., and Berger, U. 1995. Differential accumulation of xanthones in methyl-jasmonate-and yeast-extract-treated cell cultures *of Centaurium erythrea* and *Centaurium littorale*. **Planta**. 197: 608-612.

Belguendouz, L., Fremont, L., and Linard, A. 1997. Resveratrol inhibits metal ion-dependent and independent peroxidation of porcine low-density lipoprotein. **Biochem. Pharmacol**. 53: 1347-1355.

Bennett, G.J., and Lee, H.H. 1989. Xanthones from Guttiferae. **Phytochemistry** 28: 967-998.

Bentley, R. 1975. **Biosynthesis.** 3: 181-246.

Bentley, R. and Campbell, I.M. 1974. In **"The Chemistry of Quinonoid Compounds"** (Patai, S., ed.), 683-736. Wiley, New York.

Bentley, R., and Zwitkowits, P. M., 1967. Biosynthesis of tropolones in *Penicillium stipitatum*. VII. The formation of polyketide lactones and other nontropolone compounds as a result of ethionine inhibition. **J. Am. Chem. Soc**. 89: 676-80.

Bhatia, B.B. and Lal, S. 1933. The pharmacological action of *Plumbago zeylanica* and its active principle (plumbagin). **Ind. J. Med. Res**. 20: 777-788.

Bingham, S.A., Atkinson, C., Liggins, J., Bluck, L., and Coward, A. 1998. Phyto-estrogens: Where are we now? **Br. J. Nutr**. 79: 393-406.

Bisang, C., Long, P.F., Cortes, J., Westcott, J., Crosby, J., Matharu, A.L., Cox, R.J., Simpson, T.J., Staunton, J., and Leadlay, P.F. 1999. A chain initiation factor common to both modular and aromatic polyketide synthases. **Nature** (London). 401: 502-505.

Borejsza-Wysocki, W., and Hrazdina, G. 1994. Biosynthesis of *p*-hydroxyphenylbutan-2-one in raspberry fruits and tissue cultures. **Phytochemistry.** 35: 623-628.

Borejsza, Wysocki, W., and Hrazdina, G. 1996. Aromatic polyketide synthases. Purification, characterization, and antibody development to benzalacetone synthase from raspberry fruits. **Plant Physiol**. 110: 791-799.

Borejsza-Wysocki, W., Lester, C., Attygalle, A.B., and Hrazdina, G. 1999. Elicited cell suspension cultures of apple (*Malus x domestica*) cv. Liberty produce biphenyl phytoalexins. **Phytochemistry.** 50: 231-235.

Brinker, A.M., and Seigler, D.S. 1991. Isolation and identification of piceatannol as a phytoalexin from sugarcane. **Phytochemistry.** 30: 3229-3232.

Burkhardt, G., Schild, W., Becker, H., and Grubert, M. 1992. Biphenyls and xanthones from the Podostemaceae. **Phytochemistry.** 31: 543-548.

Carbo, N., Costelli, P., Baccino, F.M., Lopez-Soriano, F.J., and Argiles, J.M. 1999. Resveratrol. A natural product present in wine, decrease tumor growth in a rat tumour model. **Biochem. Biophys. Res. Commun**. 254: 739-743.

Capuano, V., Galleron, N., Pujic, P., Sorokin, A., and Ehrlich, S.D. 1996. Organization of the *Bacillus subtilis* 168 chromosome between *kdg* and the attachment site of the sp-beta prophage: use of long accurate PCR and yeast artificial chromosomes for sequencing. **Microbiol**. 142: 3005-3015.

Christensen, A.B., Gregersen, P.L., Schroder, J., and Collinge, D.B. 1998. A chalcone synthase with an unusual substrate perference is expressed in barley leaves in response to UV-light and pathogen attack**. Plant Mol. Biol**. 37: 849-857.

Clement, M.Y., Hirpara, J.L., Chawdhury, S.H., and Pervaiz, S. 1998. Chemopreventive agent resveratrol, a natural product derived from grapes, triggers CD95 signalling-dependent apoptosis in human tumor cells. **Blood.** 92: 996-1002.

Cohn, J., Day, R.B., and Stacey, G. 1998. Legume nodule organogenesis. **Trends Plant Sci.** 3: 105-110.

Crockett, C.O., Guede-Guina, F., Pugh, D., Vangah-Manda, M., Robinson, T.J., Olubadewo, J.O., Ochillo, R.F. 1992. *Cassia alata* and the preclinical search for therapeutic agents for the treatment of opportunistic infections in aids patients. **Cell Mol Biol.** 38: 799-802

De Cooman, L., Everaert, E., Fache, P., Vande Casteele, K., and Van Sumere, C.F. 1993. Flavonoid biosynthesis in petals of *Rhododendron simsii*. **Phytochemistry.** 33: 1419-1426.

Davies, K.M., Bloor, S.J., Spiller, G.B., and Deroles, S.C. 1998. Production of yellow colour in flowers: redirection of flavonoid biosynthesis in Petunia. **Plant J.** 13: 259-266.

Dellaporta, S. L., Wood, J., and Hicks, J. B., 1983. A plant DNA minipreparation. **Plant Mol. Biol. Rep.** 1: 19-21.

Dewick, P.M. 1997. **Medicinal Natural Products - A Biosynthetic Approach**. John Wiley & Sons, Chichester.

Dixon, R.A. 1999. **Isoflavonoids: biochemistry, molecular biology, and biological functions.** pp. 773-823 *in* Polyketides and Other Secondary Metabolites Including Fatty Acids and Their Derivatives (U. Sankawa, ed.), Elsevier, Amsterdam.

Dixon, R.A., and Steele, C.L. 1999. Flavonoids and isoflavonoids-a gold mine for metabolic engineering. **Trends Plant Sci**. 4: 394-400.

Drawert, F. and Beier, J. 1976. Monoprenyllierte acylphloroglucine. **Phytochemistry**. 15: 1965.

Durand, R. and Zenk, M.H. 1971. Biosynthesis of plumbagin (5-hydroxy-2-methyl-1,4-naphthoquinone) via the acetate pathway in higher plants. **Tetrahedron Letters**. 32: 3009-3012.

Eckermann, S., Schroder, G., Schmidt, J., Strack, D., Edrada, R.A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpelainen, I., Proksch, P., Teeri, T.H., and Schroder, J. 1998. New pathway to polyketide in plants. **Nature** (London). 396: 387-390.

Elomaa, P., Helariutta, Y., Kotilainen, M., and Teeri, T.H. 1996. Transformation of anitisense constructs of the chalcone synthase gene superfamily into *Gerbera hybrida*: differential effect on the expression of family members. **Mol. Breed**. 2: 41-50.

Elujoba, A.A., Ajulo, O.O., Iweibo, G.O. 1989. Chemical and biological analyses of Nigerian *Cassia* species for laxative activity. **J Pharm Biomed Anal.** 7: 1453-1457

Erdtman, H. 1963. **Some aspects of chemotaxonomy**. pp. 89-125 in Chemical Plant Taxonomy(T. Swain, ed.), Acadamic Press, London.

Ferrer, J.L., Jez, J.M., Bowman, M.E., Dixon, R.A., and Noel, J.P. 1999. Structure of chalcone synthases and the molecular basis of plant polyketide biosynthesis. **Nature Struct. Biol**. 6: 775-784.

Fischer, R., and Hain, R. 1994. Plant disease resistance resulting from the expression of foreign in a phytoalexin. **Curr. Opin. Biotechnol**. 5: 125-130.

Fischer, R., Budde, I., and Hain, R. 1997. Stilbene synthase gene expression causes changes in flower colour and male sterility in tobacco. **Plant J**. 11: 489-498.

Fliegmann, J., Schroder, G., Schanz, S., Britsch, L., and Schroder, J., 1992. Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and differential regulation of these and related enzyme activities in stressed plants. **Plant Mol. Biol**. 18: 489-503.

Forkmanm, G., and Heller, W. 1999. **Biosynthesis of flavonoids**. pp. 713-748 *in* Polyketides and Other Secondary Metabolites Including Fatty Acids and Their Derivatives (U. Sankawa, ed.), Elsevier, Amsterdam.

Foust, C. M., 1992. **Rhubarb: The Wondrous Drug**. Princeton, Princeton University Press.

Fremont, L., Belguendouz, L., and Delpal, S. 1999. Antioxidant activity of resveratrol and alcohol-free wine polyphenols related to LDL oxidation and polyunsaturated fatty acids. **Life Sci**. 64: 2511-2521.

Fujii, I., Mori, Y., Watanabe, A., Kubo, Y., Tsuji, G., and Ebizuka, Y. 1999. Heterologous expression and product identification of *Collectotrichum lagenarium* polyketide synthase encoded by the *PKS1* gene involved in melanin biosynthesis. **Biosci. Biotechnol. Biochem**. 63: 1445-1452.

Funa, N., Ohnishi, Y., Fujii, I., Shibuya, M., Ebizuka, Y., and Horinouchi, S. 1999. A new pathway for polyketide synthesis in microorganisms. **Nature** (London). 400: 897-899.

Fung, S.Y., Brussee, J., Van Der Hoeven, R.A.M., Niessen, W.M.A., Scheffer, J.J.C., and Verpoorte, R. 1994. Analysis of proposed aromatic precursors of hop bitter acids. **J. Nat. Prod**. 57: 452-459.

Graham, T.L. 1998. Flavonoid and Flavonol glycoside metabolism in Arabidopsis. **Plant Physiol. Biochem**. 36: 135-144.

Grayer, R.J., Harborne, J.B. 1994. A survey of antifungal compounds from higher plants, 1982-1993. **Phytochemistry.** 37: 19-42.

Goormachtig, S., Lievens, S., Herman, S., Van Montagu, M., and Holsters, M. 1999. Chalcone reductase-homologous transcripts accumula during development of stem-borne nodules on the tropical legume *Sesbania rostrata*. **Planta.** 209: 45-52.

Goormachtig, S., Valerio-Lepiniec, M., Szczyglowski, K., Van Montagu, M., Holsters, M., and De Bruijn, F.N. 1995. Use of differential display to identify novel *Sesbania rostrata* genes enhanced by *Azorhizobium caulinodans* infection. **Mol. Plant-Microbe Interact**. 8: 816-824.

Gorham, J. 1995. **The Biochemistry of the Stilbenoids**. Chapmann Hall, London.

Greenhagen, B., and Chappell, J., 2001. Molecular scaffolds for chemical wizardry: learning nature's rules for terpene cyclases. **Proc. Natl. Acad. Sci**. USA 98, 13479-81.

Hagen, S.E., Prasad, J.V., Boyer, F.E., Domagala, J.M., Ellsworth, E.L., Gajda, C., Hamilton, H.W., Markoski, L.J., Steinbaugh, B.A., Tait, B.D., Lunney, E.A., Tummino, P.J., Ferguson, D., Hupe, D.J., Nouhan, C., Gracheck, S.J., Saunders, J.M., and Van Der Roest, S., 1997. Synthesis of 5,6-dihydro-4-hydroxy-2-pyrones as HIV-1 protease inhibitors: the profound effect of polarity on antiviral activity. **J. Med. Chem**. 40: 3707-3711.

Hain, R., Bieseler, B., Kindl, H., Schroder, G., and Stocker, R. 1990. Expression of a stilbene synthase gene in *Nicotiana tabacum* results in synthesis of the phytoalexin resveratrol. **Plant Mol. Biol**. 15: 325-335.

Hain, R., Reif, H.-J., Karuse, E., Langebartels, R., Kindl, H., Vornam, B., Wiese, W., Schmelzer, E., Schreier, P.H., Stocker, R., and Stenzel, K. 1993. Disease resistance results from foreign phytoalexin expression in a novel plant. **Nature** (London). 361: 153-156.

Hakamatsuka, T., Noguchi, H., Ebizuka, Y., and Sankawa, U. 1988. Deoxychalcone synthase from cell suspension cultures of *Pueraria lobata*. **Chem. Phar. Bull.** (Tokyo) 36: 4225-4228.

Harano, K., Okada, N., Furuno, T., Takahashi, T., Ayabe, S.I., and Welle, R. 1993. Enzymatic synthesis of 6'-deoxychalcone in cultured *Glycyrrhiza echinata* cells. **Plant Cell Rep.** 12: 66-69.

Harborne, J.B., Greenham, J., Williams, C.A., Eagles, J., and Markham, K.R. 1993. Ten isoprenylated and C-methylated flavonoids from the leaves of three *Vellozia* species. **Phytochemistry.** 34: 219-226.

Hayashi, H., Murayama, K., Hiroaka, N., and Ikeshiro, Y. 1996. Nucleotide sequences of cDNA (accession no. D86558 and D86559) for *Glycyrhiza glabra* polyketide reductase. **Plant Physiol.** 112: 864.

Helariutta, Y., Elomaa, P., Kotilainen, M., Griesbach, R.J., Schroder, J., and Teeri, T.H. 1995. Chalcone synthase-like genes active during corolla development are differentially expressed and encode enzymes with different catalytic properties in *Gerbera hybrida* (Asteraceae). **Plant Mol. Biol**. 28: 47-60.

Helariutta, Y., Kotilainen, M., Elomaa, P., Kalkkinen, N., Bremer, K., Teeri, T.H., and Albert, V.A. 1996. Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: evolution with substarte change and catalytic simplification. **Proc. Natl. Acad. Sci. USA.** 93: 9033-9038.

Heller, W., and Forkmann, G. 1994. **Biosynthesis of flavonoids**. pp. 499-535. *in* The Flavonoids: Advances in Research Since 1986 (J.B. Harborne, ed.), Chapman & Hall, London.

Heller, W., and Hahlbrock, K. 1980. Highly purified 'flavanone synthase' from parsley catalyzes the formation of naringenin chalcone. **Arch Biochem Biophys**. 200: 617-619

Herderich, M., Beckert, C., and Veit, M. 1997. Establishing styrylpyrone synthase activity in cell free extracts obatined from gametophytes of *Equisetum arvense* L. by high performance liquid chromatography-tandem mass spectrometry. **Phytochem. Anal**. 8: 194-197.

Hrazdina, G., Kreuzaler, F., Hahlbrock, K., and Grisebach, H. 1976. Substrate specificity of flavanone synthase from cell suspension cultures of parsley and structure of release products *in vitro*. **Arch. Biochem. Biophys**. 175: 392-399.

Huang, C.S., Ma, W.Y., Goranson, A., and Dong, Z.G. 1999. Resveratrol suppresses cell transformation and induces apoptosis through a p53-dependent pathway. **Carcinogenesis**. 20: 237-242.

Ichinose, Y, Kawamata, S, Yamada, T, An, C, Kajiwara, T, Shiraishi, T, and Oku, H 1992. Molecular cloning of chalcone synthase cDNAs from *Pisum sativum*. **Plant Mol Biol.** 18: 1009-1012

Jang, M., Cai, L., Udeani, G.O., Slowing, K.V., Thomas, C.F., Beecher, C.W.W., Fong, H.H.S., Farnsworth, N.R., Kinghorn, A.D., Mehta, R.G., Moon, R.C., and Pezzuto, J.M. 1997. Cancer chemopreventive activity of resveratrol, a natural product derived from grapes. **Science**. 275: 218-220.

Jayaraman, K.S. 1987. India seeks scientific basis of traditional remedies. **Nature** (London) 326: 323.

Jez, J.M., Austin, M.B., Ferrer, J.L., Bowman, M.E., Schroder, J., and Noel, J.P. 2000a. Structural control of polyketide fromation in plant-specific polyketide synthases. **Chem Biol.** 7: 919-930

Jez, J. M., Bowman, M. E., and Noel, J. P., 2002. Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity**. Proc. Natl. Acad. Sci**. **USA**. 99: 5319-5324.

Jez, J.M., Ferrer, J.L., Bowman, M.E., Dixon, R.A., and Noel, J.P. 2000b. Dissection of malonyl-Coenzyme A decarboxylation from Polyketide formation in the reaction mechanism of a plant polyketide synthase**. Biochemistry.** 39: 890-902

Jez, J.M., Flynn, T.G., and Penning, T.M. 1997. A new nomenclature for the aldo-keto reductase superfamily. **Biochem. Pharmacol**. 54: 639-647.

Jez, J. M., and Noel, J. P., 2000. Mechanism of chalcone synthase. pKa of the catalytic cysteine and the role of the conserved histidine in a plant polyketide synthase. **J. Biol. Chem**. 275: 39640-6.

Jiang, Z. Y., Swem, L.R. Rushing, B.G., Devanathan, S., Tollin, G., and Bauer, C.E. 1999. Bacteria photoreceptor with similarity to photoactive yellow protein and plant phytochromes. **Science.** 285: 406-409.

Junghanns, K.T., Kneusel, R.E., Baumert, A., Maier, W., Groger, D., and Matern, U. 1995. Molecular cloning and recombinant expression of acridone synthase from elicited *Ruta graveolens* L. cell suspension cultures. **Plant Mol. Biol**. 27: 681-692.

Kamal, G.M., Gunaherath, B., Lesie Gunatilaka, A.A., Uvasis, M., Sultanbawa, S. and Balsubramanian, S. 1983. 1,2(3)-Tetrahydro-3,3'-Biplumbagin: A naphthalenone and other constituents from *Plumbago zeylanica*. **Phytochemistry**. 22: 1245-1247.

Kashiwada, Y., Nonaka, G., Nishioka, I., Nishizawa, M. and Yamagishi, T. 1988. Studies on rhubarb (*Rhei rhizoma*). XIV. Stilbene glucosides. **Chem. Pharm. Bull.** 36 : 1545-1549.

Kelly, T.R., Ma, Z., Xu, W. 1994. Revision of the structure of alatinone to emodin. **Phytochemistry.** 36: 253-254

Koes, R.E., Quattrocchio, F., and Mol, J.N.M. 1994. The flavonoid biosynthetic pathway in plants: function and evolution. **BioEssays.** 16: 123-132.

Kokubun, T., and Harborne, J.B. 1994. A survey of phytoalexin induction in leaves of the Rosaceae by copper ions. **Z. Naturforsch**. 49c: 628-634.

Kokubun, T., and Harborne, J.B. 1995. Phytoalexin induction in the sapwood of plants of the Maloideae (Rosaceae): biphenyls or dibenzofurans. **Phytochemistry.** 40: 1649-1654.

Kokubun, T., Harborne, J.B., Eagles, J., and Waterman, P.G. 1995. Antifungal biphenyl compounds are the phytoalexins of the sapwood of *Sorbus aucuparia*. **Phytochemistry.** 40: 57-59.

Kreuzaler, F. 1979. Flavanone synthase from Petroselinum hortense, Molecular weight, subunit composition, size of messenger RNA, and absence of pantetheinyl residue. **Eur. J. Biochem**. 99: 89-96.

Kreuzaler, F., and Hahlbrock, K. 1972. Enzymatic synthesis of aromatic compounds in higher plants: formation of naringenin (5,7,4'-trihydroxy-flavanone) from p-coumaroyl-coenzyme A. **FEBS Lett**. 28: 69-72.

Kreuzaler, F, and Hahlbrock, K. 1975. Enzymatic synthesis of aromatic compounds in higher plants. Formation of bis-noryangonin (4-hydroxy-6[4-hydroxystyryl]2-pyrone) from *p*-coumaroyl-CoA and malonyl-CoA. **Arch Biochem Biophys.** 169: 84-90

Kreuzaler, F., Light, R.J. and Hahlbrock, K. 1978. Flavanone synthase catalyzes CO2 exchange and decarboxylation of malonyl-CoA. **FEBS Lett**. 94: 175-178.

Krishnaswamy, M. and Purusottaman, K.K. 1980. Plumbagin, a study of its anticancer, antibacterial and antifungal properties. **Ind. J. Expt. Biol**. 18: 876-877.

Kubitzki, K. 1987. Phenylpropanoid metabolism in relation to land plant origin and diversification. **J. Plant Physiol**. 131: 17-24.

Lans, C., Harper, T., Georges, K., Bridgewater, E. 2000. Medicinal plants used for dogs in Trinidad and Tobago. **Prev Vet Med**. 45: 201-220

Lanz, T., Tropf, S., Marner, F.-J., Schroder, J., and Schroder, G. 1991. The role of cysteins in polyketide synthases: site-directed mutagenesis of resveratrol and chalcone synthases, two key enzymes in different plant-specific pathways. **J. Biol. Chem**. 266: 9971

Larsen, M., Poli, L., Callesen, O., and Lewis, M. 1991. Relations between the aroma compounds and the sensory evaluation of 10 raspberry varieties (*Rubus idaeus*) L. **Acta Agric. Scand**. 41: 447-454.

Leckband, G., and Lorz, H. 1998. Transformation and expression of a stilbene synthase gene of *Vitis vinifera* L. in barley and wheat for increased fungal resistance. **Theor. Appl. Genet**. 96: 1004-1012.

Leete, E., Muir, A., and Towers, G.H.N. 1982. Biosynthesis of psilotin from [2'-3'-$^{13}C_2$-^{14}C,4-^{3}H]phenylalanine studied with ^{13}C-NMR. **Tetrahedron Lett.** 23: 2635-2638.

Leistner, E. 1980. In **"Pigments in Plants"** Czygan, I.C., ed) 352-369. G.Fischer-Verlag, Stuttgart and New York.

Leistner, E. 1981. **Biosynthesis of plant quinones**. The Biochemistry of Plants, Vol.7. Academic Press, Inc.

Lopidot, T., Harel, S., Akiri, B., Granit, R., and Kanner, J. 1999. pH-dependent forms of red wine anthocyanins as antioxidant. **J. Agri. Food Chem**.47: 67-70.

Lukacin, R., Springob, K., Urbanke, C., Ernwein, C., Schroder, G., Schroder, J., and Matern, U. 1999. Native acridone synthases I and II from *Ruta graviolens* L. form homodimers. **FEBS Lett**. 448: 135-140.

Maier, W., Baumert, A., Schumann, B., Furukawa, H., and Groger, D. 1993. Synthesis of 1,3-dihydroxy-N-methylacridone and its conversion to rutacridone by cell-free extracts of *Ruta graveolens* cell cultures. **Phytochemistry.** 32: 691-698.

Manitto, P., and Sammes, P.G. 1981. **Biosynthesis of natural products**. New york: Halsted press.

Manning, K. 1998. Isolation of a set of ripening-related genes from stawberry: their identification and possible relationship to fruit quality traits. **Planta**. 205: 622-631.

Markham, K.R. 1988. **Distribution of flavonoids in the lower plants and its evolutionary significance**. pp. 427-468 *in* The Flavonoids (J.B. Harborne, ed.) Chapman and Hall, London.

Markham, K.R. 1990. **Bryophyte flavonoids, their structures, distribution, and evolutionary significance**. pp. 143-161 *in* Bryophytes Their Chemistry and Chemical Taxonomy (H.D. Zinmeister and R. Mues, eds.), Clarendon Press, Oxford.

Martin, C.R. 1993. **Structure, function, and regulation of the chalcone synthase**. pp. 233-284. *in* International Reviw of Cytology: A Survey of Cell Biology (K. Jeon and J. Jarvik, eds.), Acadamic Press, New York.

Matsuda, H., Morikawa, T., Toguchida, I., Park, J. Y., Harima, S., Yoshikawa, M., 2001. Antioxidantconstituents from rhubarb: structural requirements of stilbenes for the activity and structures of two new anthraquinone glucosides. **Bioorg. Med. Chem**. 9, 41-50.

McInnes, A.G., Yoshida, S., and Towers, G.H.N. 1965. A phenolic gylcoside from *Psilotum nudum* (L) Griseb. **Tetrahedron.** 21: 2939-2946.

Mckhann, H.I., Paiva, N.L., Dixon, R.A., and Hirsch, A.M. 1998. Expression of genes for enzymes of the flavonoid biosynthetic pathway in the early stages of the *Rhizobium*-legume symbiosis. **Adv. Exp. Med. Biol**. 439: 45-54.

Melchior, F., and Kindl, H., 1990. Grapevine stilbene synthase cDNA only slightly differing from chalcone synthase cDNA is expressed in *Escherichia coli* into catalytically active enzyme. **FEBS Lett**. 268, 17-20.

Miyamoto, M., Imai, S., Shinohara, M., Fujioka, S., Goto, M. Matsuoka, T., Fujimura, H. 1967. Studies on purgative substances I. Isolation of sennoside A. one of the most active principles from rhubarb. J. **Pharm. Soc**. 87: 1040-1043.

Mol, J.N.M., Grotewold, E., and Koes, R.E. 1998. How gene paint flowers and seeds. **Trends Plant Sci**. 3: 212-217.

Morita, H., Noguchi, H., Schroder, J., and Abe, I., 2001. Novel polyketides synthesized with a higher plant stilbene synthase. **Eur. J. Biochem**. 268: 3759-3766.

Morita, H., Takahashi, Y., Noguchi, H., and Abe, I., 2000. Enzymatic formation of unnatural aromatic polyketides by chalcone synthase. **Biochem. Biophys. Res. Commun**. 279: 190-195.

Nagumo, S., Toyonaga, T., Inoue, T., and Nagai, M. 1989. New glucosides of a-4-hydroxy-5-methylcoumarin and a dihydron-a-pyrone from *Gerbera Jamesonii hybrida*. **Chem. Pharm. Bull**. (Tokyo) 37: 2621-2623.

Nonaka, G. and Nishioka, I. 1983. Tannins and related compounds. X. Rhubarb (2). **Chem. Pharm. Bull**. 31: 1652-1658.

Nonaka, G., Nishioka, I., Nagasawa, T. and Oura, H. 1981. Tannins and related compounds. I. Rhubarb (1). **Chem. Pharm. Bull**. 29: 2862-2870.

Okabe, H., Matsuo, K. and Nishioka, I. 1973. Studies on rhubarb (*Rhei rhizoma*). II. Anthraquinone glycosides. **Chem. Pharm. Bull**. 21: 1254-1260.

Oshio, H., Imai, S., Fujioka, S., Sugawara, T., Miyamoto, M., and Tsukui, M. 1974. Investigation of rhubarbs III. New purgative constituents, Sennosides E and F. **Chem Pharm Bull**. 22: 823-831.

Pabst, A., Barron, D., Adda, J., and Schreier, P. 1990. Phenylbutan-2-one b-D-glucosids from raspberry fruit. **Phytochemistry.** 29: 3855-3858.

Pace-Asciak, C.R., Hahn, S., Diamandis, E.P., Soleas, G.J., and Goldberg, D.M. 1995. The red wine phenolics trans-resveratrol and quercetin block human platelet aggregation and eicosanoid synthesis: implications for protection against coronary heart disease. **Clin. Chim. Acta.** 235: 207-219.

Paganga, G., Miller, N., and Rice-evans, C.A. 1999. The polyphenolic content of fruit and vegetables and their antioxidant activities. What does a serving constitute? **Free Radc. Res**. 30: 153-162.

Paniego, N.B., Zuurbier, K.W.M., Fung, S.-Y., Van Der Heijden, R., Scheffer, J.J.C., and Verpoorte, R. 1999. Phlorisovalerophenone synthase, a novel polyketide synthase from hop (*Humulus lupulus* L.) cones. **Eur. J. Biochem**. 262: 612-616.

Perrin, P.W., and Towers, G.H.N. 1973. Hispidin biosynthesis in cultures of *Polyporus hispidus*. **Phytochemistry.** 12: 589-592.

Peters, S., Schmidt, W., and Beerhues, L. 1998. Regioselective oxidative phenol coupling of 2,3',4,6-tetrahydroxybenzophenone in cell cultures of *Centaurium erythraea* RAFN and *Hypericum androsaemum* L. **Planta.** 204: 64-69.

Powell, R.G., Tepaske, M.R., Plattner, R.D., White, J.F., and Clement, S.L. 1994. Isolation of resveratrol from *Festuca versuta* and evidence for the widespread occurrence of this stilbene in the Poaceae. **Phytochemistry.** 35: 335-338.

Preisig-Muller, R. Gehlert, R., Melchior, F., Stietz, U., and Kindl, H. 1997. Plant polyketide synthase leading to stilbenoids have a domain catalyzing malonyl-CoA:CO_2 exchange malonyl-CoA decarboxylation, and covalent enzyme modification and a site for chain lengthening. **Biochemistry.** 36: 8349-8358.

Reimold, U., Kroeger, M., Kreuzaler, F., and Hahlbrock, K. 1983. Coding and 3' noncoding nucleotide sequence of chalcone synthase messenger RNA and assignment of amino acid sequence of the enzyme. **EMBO J.** 2: 1801-1806.

Rice-Evans, C.A., and Miller, N.J. 1996. Antioxidant activities of flavonoids as bioactive components of food. **Biochem. Soc. Transaction.** 24: 790-794.

Rice-Evans, C.A., Miller, N.J., and Paganga, G. 1997. Antioxidant properties of phenolic components. **Trends Plant Sci**. 2: 152-159.

Rimm, E.B., Klatsky, A., Grobbee, D., and Stampfer, M.J. 1996. Review of moderate alcohol consumption and reduced risk of coronary heart disease-is the effect due to beer, wine, or spirits. **British med. J**. 312: 731-736.

Rupprich, N., and Kindl, H. 1978. Stilbene synthases and stilbenecarboxylate synthases. I. Enzymatic synthesis of 3,5,4x-trihydroxystilbene from *p*-coumaroyl coenzyme A and malonyl coenzyme A. **Hoppe-Seyler's Z. Physiol. Chem**. 359: 165-172.

Ryder, T.B., Hedrick, S.A., Bell, J.N., Liang, X.W., Clouse, S.D., and Lamb, C.J. 1987. Organization and differential activation of a gene family encoding the plant defense enzyme chalcone synthase in *Phaseolus vulgaris*. **Mol Gen Genet** . 210: 219-233

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

Saleh, N.A.M., Fritsch, H., Kreuzaler, F., and Grisebach, H. 1978. Flavanone synthase from cell suspension cultures of *Haplopappus gracilis* and comparision with the synthase from parsley. **Phytochemistry.** 17: 183-186.

Sallaud, C., El-Turk, J., Bigarre, L., Sevin, H., Welle, R., and Esnault, R. 1995. Nucelotide sequences of three chalcone reductase genes from alfafa. **Plant Physiol**. 108: 869-870.

Sallaud, C., El-Turk, J., Breda, C., Buffard, D., and De Kozak, I., Esnault, R., Kondorosi, A. 1995. Differential expression of cDNA coding for chalcone reductase, a key enzyme of the 5'-deoxyflavonoid pathway, under various stress conditions in *Medicago sativa*. **Plant Sci**. 109: 179-190.

Salzman, R. A., Fujita, T., Zhu-Salzman, K., Hasagawa, P. M., and Bressan, R. A., 1999. An improved RNA isolation method for plant tissue containing high levels of phenolic compounds or carbohydrates. **Plant Mol. Biol. Rep**. 17: 11-17.

Samappito, S., Page, J., Schmidt, J., De-Eknamkul, W., and Kutchan, T. M., 2002. Molecular characterization of root-specific chalcone synthases from *Cassia alata* . **Planta**. (Accepted).

Sambrock, J., Fritsch, E.F., and Maniatis, T. 1989. **Molecular Cloning. A Laboratory Manual.** 2nd ed., Cold Spring Harbor Laboratory Press.

Sankawa, U., Hakamatsuka, T., Shinkai, T., Yoshida, M., Park, H.H., and Ebizuka, Y. 1995. **Changes of secondary metabolism by elicitor treatment in *Pueraria lobata* cell cultures**. pp. 595-604 *in* Current Issues in Plant Molecular and Cellular Biology (M. Terzi, R. Cella and A. Falavigna, eds.) Kluwer Academic Publishers Dordrecht.

Santha Kumari, G., Rathinam, K. and Seshadri, C. 1978. Anticoagulant activity of plumbagin (from *Plumbago zeylanica*). **Ind. J. Expt. Biol**. 16: 485-487.

Schmidt, W., and Beerhues, L. 1997. Alternative pathway of xanthone biosynthesis in cell cultures of *Hypericum androsaemum* L. **FEBS Lett**. 420:143-146.

Schoppner, A. and Kindl, H. 1984. Purification and properties from Scot pine (*Pinus sylvestris*). **FEBs Lett**. 313: 71-74

Schroder, J. 1997. A family of plant-specific polyketide synthases: facts and predictions. **Trends Plant Sci**. 2: 373-378.

Schroder, J. 1999. Probing plant polyketide synthesis. **Nature Struct. Biol**. 6: 714-716.

Schroder, J. 1999. **The chalcone/stilbene synthase-type family of condensing enzyme**. pp. 749-771 *in* Polyketides and Other Secondary Metabolites Including Fatty Acids and Their Derivatives (U. Sankawa, ed.), Elsevier, Amsterdam.

Schroder, J. 2000. **The family of chalcone synthase-related proteins: functional diversity and evolution.** Pp.55-89 in Evolution of Metabolic Pathways (J.T. Romeo et al. Ed), Elsevier Science Ltd.

Schroder, G., Brown, J.W.S., and Schroder, J., 1988. Molecular analysis of resveratrol synthase: cDNA, genomic clones and relationship with chalcone synthase. **Eur. J. Biochem**. 172: 161-169.

Schroder, J. Raiber, S., Berger, T., Schmidt, A., Schmidt, J., Soares-Sello, A.M., Bardshiri, E., Strack, D., Simpson, T.J., Veit, M., and Schroder, G. 1998. Plant polyketide synthases: a chalcone synthase-type enzyme which performs a condensation reaction with methymalonyl-CoA in the biosynthesis of C-methylated chalcones. **Biochemistry.** 37: 8417-8425.

Schroder, J. and Schafer, E. 1980. Radio-iodinated antibodies, a tool in studies on the presence and role of inactive enzyme forms: regulation of chalcone synthase in parsley cell suspension cultures. **Arch. Biochem. Biophys**. 203, 800-808.

Schuz, R., Heller, W., and Hahlbrock, K. 1983. Substrate specificity of chalcone synthase from *Petroselinum hortense*. Formation of phloroglucinol derivatives from aliphatic substrate. **J. Biol. Chem**. 258: 6730-6734.

Soleas, G.J., Diamandis, E.P., and Goldberg, D.M. 1997. Resvertrol: a molecule whose time has come? And gone? **Clin. Biochem**. 30: 91-113.

Soleas, G.J., Diamandis, E.P., and Goldberg, D.M. 1997. Wine as a biological fluid: history, production, and role in disease prevention. **J. Clin. Lab. Analysis.** 11: 287-313.

Sparvoli, F., Martin, C., Scienza, A., Gavazzi, G., and Tonelli, C. 1994. Cloning and Molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). **Plant Mol Biol.** 24: 743-755

Springob, K., Lukacin, R., Ernwein, C., Gröning, I., and Matern, U. 2000. Specificities of functionally expressed chalcone and acridone synthases from *Ruta graveolens*. **Eur J Biochem.** 267: 6552-6559

Stafford, H.A. 1991. Flavonoid evolution: an enzymic approach. **Plant Physiol**. 96: 680-685.

Stafford, H.A. 1997. Roles of flavonoids in symbiotic and defense functions in legume roots. **Bot. Rev**. 63: 27-39.

Stark-Lorenzen, P., Nelke, B., Hanssler, G., Muhlbach, H.P., and Thomzik, J.E. 1997. Transfer of a grapevine stilbene synthase gene to rice (*Oryza sativa* L.). **Plant Cell Rep.** 16: 668-673.

Stoll, A., Becker, B., Helfenstein, A. 1950. Die Konstitution der Sennoside. **Helv Chim Acta**. 33 : 313-336.

Suh, D.Y., Fukuma, K., Kagami, J., Yamazaki, Y., Shibuya, M., and Ebizuka, Y. 2000. Identification of amino acid residues important in the cyclization reactions of chalcone and stilbene synthases. **Biochem J.** 350: 229-235

Sultanbawa, M.U.S. 1980. Xanthonoids of tropical plants. Tetrahedron 36: 1465-1506.

Surh, Y.J. 1999. Molecular mechanisms of chemopreventive effects of selected dietary and medicinal phenolic substances. **Mutat. Res. Fund. Mol. Mech. Mutagen**. 428: 305-327.

Surh, Y.J., Hurh, Y.J., Kang, J.Y., Lee, E., Kong, G., and Lee, S.J. 1999. Resvertrol, an antioxidant present in red wine induces apoptosis in human promyelocytic leukemia (HL-60) cells. **Cancer Lett**. 140: 1-10.

Takahashi, S., Nakamura, F., Sahashi, N., Ohmoto, T., Mizushima, U., Sankawa, U., and Towers, G.H.N. 1990. Chemical markers of the Psilotaceae**. Biochem. Syst. Ecol.** 18: 11-12.

Takano, Y., Kubo, Y., Shimizu, K., Mise, K., Okuno, T., and Furusawa, I. 1995. Structural analysis of *PKS1*, a polyketide synthase gene involved in melanin biosynthesis in *Collectotrichum lagenarium*. **Mol. Gen. Genet**. 249: 162-167.

Thomson, R.H. 1971. **Naturally Occurring Quinones**. Acadmic Press, New York.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucl. Acids Res**. 25, 4876-82.

Thomzik, J.E., Stenzel, K., Stocker, R., Schreier, P.H., Hain, R., and Stahl, D.J. 1997. Synthesis of a grapevine phytoalexin in transgenic tomatoes (*Lycopersicon esculentum* Mill.) conditions resistance against *Phytophthora infestans*. **Phyiol. Mol. Plant Pathol**. 51: 265-278.

Torssell, K.B.G. 1997. **Natural Product Chemistry : A mechanistic, biosynthetic and ecological approach.** Apotekasocieteten, Swedish Pharmaceutical Society.

Tropf, S., Lanz, T., Rensing, S.A., Schroder, J., and Schroder, G. 1994. Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. **J. Mol. Biol**. 38: 610-618.

Tropf, S., Karcher, B., Schroder, G., and Schroder, J. 1995. Reaction mechanisms of homodimeric plant polyketide synthases (stilbene and chalcone synthase): a single active site for the condensing reaction is suffient for synthesis of stilbenes, chalcones, and 6'-deoxychalcones. **J. Biol. Chem**. 270: 7922-7928.

Tsuboi, M., Minami, M., Nanoka, G. and Nishioka, I. 1977. Studies on rhubarb (*Rhei rhizoma*). IV. Naphtalene glycoside. **Chem. Pharm. Bull**. 25 : 2708-2712.

Ueda, K., Kim, K.-M., Beppu, T., and Horinouchi, S. 1995. Overexpression of a gene cluster encoding a chalcone synthase-like protein confers redbrown pigment production in Streptomyces griseus. **J. Antibiotics.** 48: 638-646.

Van de Peer, Y. and De Wachter, R. 1994. TREECON for windows a software package for the construction and drawing of the evolutionary trees for the microsoft windows environment. **Comput. Appl. Biosci**. 10: 569-570

Van den Berg, A.J.J. and Labadie, R.P. 1989. **Quinones.** Methods in Plant Biochemistry. Vol.1. Academic Press Limited.

Van der Vijver, L.M. 1972. Distribution of plumbagin in the Plumbaginaceae. **Phytochemistry** 11: 3247-3248.

Veit, M., Geiger, H., Wray, V., Abou-Mandour, A., Rozdzinski, W., Witte, L., Strack, D., and Czygan, F.C. 1993. Equisetumpyrone, a stylpyrone glucoside in gametophytes from *Equisetum arvense*. **Phytochemistry.** 32: 1029-1032.

Wang, H.K., Xia, Y., Yang, Z.Y., Natschke, S.L.M., and Lee, K.H. 1998. Recent advances in the discovery and development of flavonoids and their analogues as antitumor and anti-HIV agents. **Adv. Exp. Med. Biol**. 439: 191-225.

Wat, C.K., and Towers, G.H.N., 1979. **Metabolism of the aromatic amino acids by fungi**. pp. 371-432 *in* Biochemistry of Plant Phenolics (T. Swain, J.B. Harborne, and C.F. Van Sumere, eds.) Plenum Press, New York.

Welle, R., and Grisebach, H. 1989. Phytoalexin synthesis in soybean cells: elicitor induction of reductase involved in biosynthesis of 6'-deoxychalcone**. Arch. Biochem. Biophys**. 272: 97-102.

Welle, R., Schroder, G., Schiltz, E., Grisebach, H., and Schroder, J. 1991. Induced plant responses to pathogen attack: analysis and heterologous expression of the key enzyme in the biosynthesis of phytoalexins in soybean (*Glycine max* L. Merr.cv. Harosoy 63). **Eur. J. Biochem**. 196: 423-430.

Widyastuti, S.M., Nonaka, F., Watanabe, K., Sako, N., and Tanaka, K. 1992. Isolation and characterization of two aucuparin-related phytoalexins from *Photinia glabra* Maxim. **Ann. Phytopath. Soc. Japan.** 58: 228-233.

Yadav, S.K., Kalidhar, S.B. 1994. Alquinone: an anthraquinone from *Cassia alata*. **Planta Med.** 60: 601

Yamaguchi, T., Kurosaki, F., Suh, D.Y., Sankawa, U., Nishioka, M., Akiyama, T., Shibuya, M., and Ebizuka, Y. 1999. Cross-reaction of chalcone synthase and stilbene synthase overexpressed in *Escherichia coli*. **FEBS Lett.** 460: 457-461

Yamazaki, Y., Suh, D. Y., Sitthithaworn, W., Ishiguro, K., Kobayashi, Y., Shibuya, M., Ebizuka, Y., and Sankawa, U., 2001. Diverse chalcone synthase superfamily enzymes from the most primitive vascular plant, *Psilotum nudum*. **Planta.** 214: 75-84.

Yoshikawa, M., Ueda, T., Matsuda, H., Yamahara, J., and Murakami, N. 1994. Absolute stereostructures of hydramacrolides A and B, new bioactive secoiridoid glucoside complexes from the leaves of *Hydrangea macrophylla* Sering var. *thunbergii* Makino. **Chem Pham. Bull**. (Tokyo) 42: 1691-1693.

Zheng, D., Schroder, G., Schroder, J., and Hrazdina, G. 2001. Molecular and biochemical characterization of three aromatic polyketide synthase genes from *Rubus idaeus*. **Plant Mol Biol.** 46: 1-15

Zuurbier, K.W.M., Fung, S.-Y., Scheffer, J.J.C., and Verpoorte, R. 1995. Formation of aromatic intermediates in the biosynthesis of bitter acids in *Humulus lupulus*. **Phytochemistry.** 38: 77-82.

Zuubier, K.W.M., Leser, J., Berger, T., Hoete, A.J.P., Schroder, G. Verpoorte, R., and Schroder, J. 1998. 4-hydroxy-2-pyrone formation by chalcone and stilbene synthase with nonphysiological substrates. **Phytochemistry.** 49: 1945-1951.

APPENDIX I

Original manuscript I

The title “Molecular characterization of root-specific chalcone synthases from *Cassia alata*” has been accepted by Planta.

Dedicated to Nikolaus Amrhein, Zurich, on the occasion of his 60th birthday.

Molecular characterization of root-specific chalcone synthases from *Cassia alata*

Supachai Samappito[a,b,*] Jonathan Page[a] Jürgen Schmidt[a]
Wanchai De-Eknamkul[b] Toni M. Kutchan[a]

[a]*Leibniz-Institut für Pflanzenbiochemie, Weinberg 3, 06120 Halle/Saale, Germany*
[b]*Department of Pharmacognosy, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand*

The nucleotide sequences reported in this paper have been submitted to the GenBank™/EBI Data Bank with accession numbers AF358430 (*CalCHS1*), AF358431 (*CalCHS2*), AF358432 (*CalCHS3*).

Corresponding author. Tel.: +49 345 5582 1200; fax: +49 345 5581 1209
E-mail address: kutch@ipb-halle.de (T.M. Kutchan)
[*]Present address: Department of Biotechnology, Faculty of Technology, Mahasarakham University, Mahasarakham 44001, Thailand

Abstract Three cDNAs encoding very similar but unique isoforms of chalcone synthase [EC 2.3.1.74] were isolated from a cDNA library prepared from RNA from root tissue of the Thai medicinal plant *Cassia alata* L. (ringworm bush, Leguminosae). Gene transcript for these three type III polyketide synthases was found to accumulate predominantly in root. The heterologously expressed enzymes accepted acetyl-, *n*-butyryl-, isovaleryl-, *n*-hexanoyl-, benzoyl-, cinnamoyl-, and *p*-coumaroyl-CoA as starter molecules and together with the co-substrate malonyl-CoA, formed multiple products. With the exception of the assay in which acetyl-CoA was used as the starter molecule, all substrates yielded a phloroglucinol derivative resulting from three sequential condensations of acetate units derived from three malonyl-CoA decarboxylations. Every substrate tested also produced two pyrone derivatives, one resulting from two acetate unit condensations (a bis-noryangonin-type pyrone derailment product) and one resulting from three acetate unit condensations (a 4-coumaroyltriacetic acid lactone-type pyrone derailment). *C. alata* accumulates the flavonoids quercetin, naringenin and kaempferol in roots,

suggesting that the *in planta* function of these enzymes is the biosynthesis of root flavonoids.

Keywords *Cassia,* Leguminosae, Ringworm Bush, Polyketide Synthase, Chalcone Synthase, Flavonoids

Abbreviations CE: collision energy, CHS: chalcone synthase, CID-MS: collision-induced dissociation mass spectrometry, CoA: coenzyme A, DTT: dithiothreitol, HEPES: *N*-2-hydroxypiperazine-*N*-2-ethane sulfonic acid, LC/ESI-MS: liquid chromatography/electrospray ionization-mass spectrometry, PCR: polymerase chain reaction, 2-PS: 2-pyrone synthase, RT: retention time, RT-PCR: reverse transcription-polymerase chain reaction, PKS: polyketide synthase, SRM: selected reaction monitoring, TLC: thin layer chromatography, Tris: tri(hydroxymethyl)methylamine

Introduction

Plant polyketide synthases are encoded by a multigene family that has chalcone synthase as a prototype. Gene family evolution in plants appears to occur by gene duplication followed by nucleotide substitution that can lead to biochemical diversity. Plant polyketide synthases are presumably derived from a common ancestor that diverged to perform different reactions. This is clearly demonstrated by the highly similar genes that encode chalcone synthase and resveratrol synthase in a single species, such as found in *Vitis vinifera* (Melchior and Kindl 1990; Sparvoli et al. 1994), or by the occurrence of the closely related, but enzymatically unique chalcone and acridone synthases in *Ruta graveolens* (Springob et al. 2000).

Detailed crystal structure and site-directed mutagenesis studies have provided insight into the structural control of polyketide formation in plant homodimeric iterative (type III) polyketide synthases (Jez et al. 2000a; 2000b; Suh et al. 2000). The active site cavity can serve as a size-based filter to sterically exclude bulkier starter molecules (Jez et al. 2000a). This explains why pyrone synthase cannot accept *p*-coumaroyl CoA as a starter molecule, whereas chalcone synthase can use acetyl CoA as a starter molecule, albeit inefficiently, and forms a small amount of methylpyrone (Schüz et al. 1983). Modification of as few as three amino acids can change chalcone synthase into a pyrone synthase (Jez et al. 2000a). This demonstrates how readily plant type III polyketide synthases can be modified by nucleotide substitution to produce new natural products. Plants produce a variety of polyketide-derived secondary metabolites such as naphtha- and anthraquinones that could potentially be biosynthesized from type III

polyketide synthases, such as has been demonstrated for the formation of tetrahydroxynaphthalene in the microorganism *Streptomyces griseus* (Funa et al. 1999).

Cassia alata L. (ringworm bush, Leguminosae), distributed mainly in the tropics and subtropics, is used to treat superficial mycosis of the skin (Crockett et al. 1992). Anthraquinones accumulate in aerial plant parts as well as in roots (Kelly et al. 1994; Yadav and Kalidhar 1994) and are the principal laxative constituents in *Cassia* species used as purgatives (Elujoba et al. 1989). *C. alata* also finds use in preventative veterinary medicine and is used to deworm dogs (Lans et al. 2000). In an effort to characterize polyketide synthase genes from medicinal plants native to Thailand and to search for polyketide synthases catalyzing anthraquinone formation, we have isolated and functionally expressed three cDNAs encoding isoforms of a root-specific polyketide synthase from *C. alata*. We report herein on the molecular characterization of these three polyketide synthase isoforms.

Materials and methods

Plant material

Seeds of *Cassia alata* L. were obtained from plants growing in an open field of The Walai Rukhavej Botanical Research Institute, Mahasarakham University, Mahasarakham, Thailand. They were sown in the greenhouse at the Institute of Plant Biochemistry, Halle, Germany, at 24 °C with 18 h of light and 50% humidity. Young plant material was used for the following experiments.

cDNA isolation

Partial cDNA clones encoding putative polyketide synthases were generated by PCR using cDNA produced by reverse transcription of total RNA. Total RNA (10 g) was reverse transcribed using a cDNA synthesis kit (Life Technologies) as described by the manufacturer. DNA amplification using degenerate primers, (5'-AA(AG)GC(CT)AT(AC)GAIGA(AG)TGGGG-3') as the forward primer and (5'-CCACCIGG(AG)TGI(AG)CAATCC-3') as the reverse primer, was performed under the following conditions: 3 min at 94 °C, 3 min for 1 cycle, then 30 cycles of 94 °C for 30 s, 48 °C for 30 s, 72 °C for 1 min. The last step was followed by an additional elongation step at 72 °C for 7 min. The amplified DNA was resolved by agarose gel electrophoresis. The band of approximately the desired size (584 bp) was eluted from the gel using a gel extraction kit (Qiagen). The DNA was

subsequently ligated into pGEM-T Easy (Promega) and was transformed into *E. coli* DH5α. Sequencing of several clones confirmed that the PCR products showed similarity to polyketide synthases.

Poly(A)$^+$RNA was used as the template for first and second strand cDNA synthesis, using a ZAP-cDNA synthesis kit (Stratagene). The cDNA was extended with *Eco* RI and *Xho* I adaptors prior to ligation into the Uni-ZAP XR vector (Stratagene), which was then ligated to λ-ZAPII. The phagemids were packaged with a Gigapack III Gold Packaging Extract (Stratagene).

The partial clones encoding polyketide synthases that were generated by RT-PCR, as described above, were labeled with ^{32}P and used to screen a *C. alata* cDNA library of 100,000 phages for full-length clones. The library was plated and the plaque DNA was transferred to nitrocellulose membranes (Schleicher & Schuell). Filters were hybridized with the individual hybridization probes in a solution of 5x SSPE containing 50% formamide, 5% Denhardt's reagent, 0.5% SDS and 100 mg/ml salmon sperm DNA at 42 °C overnight. The hybridized filters were washed three times in 0.1x SSPE and 0.1% SDS for 20 min at 42 °C. Radioactivity was visualized on X-ray film using an intensifying screen. In this manner, three clones (*CalPKS1*, *CalPKS2* and *CalPKS3*) were identified and taken through several rounds of screening until the viruses were pure enough to convert to plasmids by excision. To ascertain the identity of the cDNAs as encoding polyketide synthases, the nucleotide sequences were determined on both strands and the translations were compared to those sequences available in the GenBank/EMBL databases.

Functional expression and enzyme assays

The CalPKS reading frames were amplified by PCR, CalPKS1 using 5'-TAGTCATATG
GTGAAGGTGGAAGAG-3' as the forward primer (*Nde* I site is underlined and the translation start site is bold) and 5'-TTGGCTCGAGTTAAATAGCAATACTGT-3' as the reverse primer (*Xho* I site is underlined and stop codon is bold); CalPKS2 using 5'-ATTTTTCATATGGTGAGTGTTGAAGAG-3' as the forward primer (*Nde* I site is underlined and the translation start site is bold) and 5'-TTTAAAGGATCCTTAGTTAAC
TGCCACACT-3' as the reverse primer (*Bam*H I site is underlined and stop codon is bold); CalPKS3 using 5'-TAGTCATATGGTGAAGGTGGAAGAG -3' as the

forward primer (*Nde* I site is underlined and the translation start site is bold), and 5'-TTGGCTC GAGTTAGACAGCCACACTAT-3' as the reverse primer (*Xho* I site is underlined and stop codon is bold) with the cDNAs in pBluescript as template. The 1.2 kb PCR products were digested with *Nde* I and *Xho* I, gel purified and ligated into *Nde*I/*Xho*I-digested pET-14b (Novagen) to generate the expression plasmids pET*Ca1PKS1*, pET*CalPKS2* and pET*CalPKS3*.

pET*Ca1PKS1*, pET*CalPKS2* and pET*CalPKS3* were each transformed into *E. coli* BL21(DE3). Transformed *E. coli* were grown at 37 °C in Luria-Bertani medium containing 50 µg/ml ampicillin until $A_{600\ nm}$ = 0.6. After induction with 0.4 mM isopropyl-1-thio- -D-galactopyranoside, the cultures were grown at 28 °C for an additional 12 hrs. Cells were collected by centrifugation (10,000 x *g*, 10 min, 4 °C) and resuspended in a buffered solution containing 50 mM Tris-HCl (pH 7.0), 500 mM NaCl, 2.5 mM imidazole, 10 mM -mercaptoethanol, 10% (v/v) glycerol, 1% (v/v) Tween-20 and 0.75 g/ml lysozyme (47,000 units/mg solid). After sonication and centrifugation (10,000 x *g*, 10 min, 4 °C), the supernatant was passed over a Talon (Clontech) metal affinity resin column. The resin was washed with 10 bed volumes of a buffered solution containing 50 mM Tris-HCl (pH 7.0), 500 mM NaCl, 2.5 mM imidazole, 10 mM β-mercaptoethanol, 10% (v/v) glycerol, and the His-tagged protein was eluted with a buffered solution containing 50 mM Tris-HCl (pH 7.0), 500 mM NaCl, 10 mM β-mercaptoethanol, 10% (v/v) glycerol and 250 mM imidazole. The eluted protein solution was desalted on a PD-10 column (Pharmacia) equilibrated with 50 mM HEPES (pH 7.0), 5 mM DTT, 10% glycerol and was stored at 4 °C.

Polyketide synthase activity was measured by the conversion of various "starter" CoA molecules and [2-^{14}C]malonyl CoA into reaction products. The standard enzyme assay contained 100 mM HEPES buffer (pH 7.0), 20 µM "starter" CoA and 12,000 dpm malonyl CoA (2 GBq/mmol) (Biotrend Chemikalien, Cologne) and 1.5 µg polyketide synthase in a 50 µl reaction volume. The assay mixture was typically incubated for 30 min at 30 °C. The reaction was stopped by addition of 5 µl 10% (v/v) HCl and was extracted twice with 100 µl ethylacetate. The combined organic phase was evaporated to dryness, the residue dissolved in 10 µl ethylacetate and this entire volume was resolved by RP-C18 thin layer chromatography (Merck). The TLC plates were

developed in methanol:H_2O:acetic acid (75:25:1). The radioactive products were quantified with a Rita Star TLC scanner (Raytest) or by phosphorimagery.

Mass spectrometric analysis

The positive and negative electrospray ionization (ESI) mass spectra were obtained with a Finnigan MAT TSQ 7000 instrument (electrospray voltage 4.5 kV; heated capillary temperature 220 ^oC; sheath and auxillary gas: nitrogen) coupled to a Micro-Tech Ultra-Plus MicroLC system equipped with an RP18-column (5 μm, 1x100 mm, SepServ, Berlin). For all compounds (**1**-**7**), a gradient system was used that ranged from H_2O:CH_3CN 90:10 (each containing 0.2% HOAc) to 10:90 over 15 min, followed by isocratic elution with a 10:90 mixture of both solvents for 10 min; flow rate 70 μl min^{-1}. The collision-induced dissociation (CID) mass spectra during an HPLC run were recorded with a collision energy of –20 or -25 eV for positive ions as well as +20 or +25 eV for negative ions, respectively (see Table 2); collision gas: argon, collision pressure: 1.8 x 10^{-3} Torr.

Flavonoid analysis

Plant tissue (1 g fresh weight) was frozen in liquid nitrogen and ground to a fine powder with a mortar and pestle. To the powder was added 10 ml 80% methanol, 3% concentrated HCl. The suspension was heated to 100 °C for 60 min in a sealed tube. After cooling, the cell debris was removed from the extract by centrifugation at 10,000 x g for 10 min at room temperature. The extract was taken to dryness *in vacuo*, dissolved in 5 ml water and extracted twice with 2 volumes of ethylacetate. The ethylacetate was removed *in vacuo* and the residue dissolved in 50 µl 50% methanol. Isolubles were separated by centrifugation (10,000 x g, 10 min, room temperature). For resolution of the natural products, 10 µl of the supernatant was injected onto a Hewlett Packard Series 1100 HPLC equipped with a Knauer Eurospher-100 C18 column (5 µm, 250 x 4 mm). The solvent system was 97.8% (v/v H_2O), 2% CH_3CN, 0.2% H_3PO_4 (A) and 97.8% (v/v CH_3CN), 2% H_2O, 0.2% H_3PO_4 (B) with a gradient of 0-30 min 20-45% B, 30-35 min 45-55% B at a flow rate of 0.6 ml min^{-1}. The flavonoid content was monitored at 370 nm.

General methods

Young leaves and roots were harvested, frozen immediately in liquid nitrogen and stored at -80 °C. The isolation of total RNA was carried out as described by Salzman, et al. (1999). Poly(A)$^+$ RNA was isolated with Oligotex beads (Qiagen) according to the manufacture's instructions.

For RNA gel blot analysis, total RNA (10 μg) was resolved by electrophoresis on 1% agarose gels containing formaldehyde. The RNA was transferred to a nylon membrane (Hybond N, Amersham) by capillary blotting. The three *Cassia* polyketide synthase cDNAs were labeled with ^{32}P using a random primer DNA labeling kit (Life Technologies) and were then used as hybridization probes. Hybridization was performed at 65 °C overnight and the membranes were then washed three times for 15 min with 2x SSC and 0.1% SDS at 65 °C. Radioactivity was visualized by phosphorimagery.

Genomic DNA was isolated from young leaves according to Dellaporta et al. (1983). The DNA (10 μg) was incubated with each of the following restriction endonucleases: *Apo*I, *Bcl* I and *Bst*X I (one restriction recognition site within the reading frame), and *Dra* III, *Nde* I and *Xba* I (no restriction recognition site within the reading frame) and was then resolved by electrophoresis on a 0.8% agarose gel, and capillary blotted onto a Hybond N nylon membrane. The three *Cassia* polyketide synthase cDNAs were labeled with ^{32}P using a random primer DNA labeling kit and were then used as hybridization probes. Hybridization was performed at 65 °C overnight and the membranes were then washed three times for 15 min with 2x SSC and 0.1% SDS at 65 °C. Radioactivity was visualized by phosphorimagery.

Results and discussion

Isolation of cDNAs encoding C. alata polyketide synthases

Using oligodeoxynucleotide primers that were based upon conserved regions in plant polyketide synthases, RT-PCR was carried out using *C. alata* root total RNA as template. The PCR products were resolved by agarose gel electrophoresis and the band of approximately the correct size (584 bp) was purified and ligated into pGEM-T Easy. Nucleotide sequence analysis revealed that the PCR products were highly similar to plant chalcone synthases. These partial clones were then used as a

hybridization probe with which to screen a cDNA library in λZAP II prepared from *C. alata* root mRNA. In this manner, three clones that encoded complete reading frames were isolated. Translation of the nucleotide sequences of the three cDNAs indicated that they were highly similar, but non-identical (Fig. 1). A comparison of the translation of the nucleotide sequences of these cDNAs with those present in the GenBank/EMBL databases indicated that the *Cassia* genes are most similar to chalcone synthase from other members of the Leguminosae, specifically *Glycine max* (Akada et al. 1993) and *Phaseolus vulgaris* (Ryder et al. 1987) (Table 1).

The complete nucleotide sequences of the three cDNAs *CalPKS1*, *CalPKS2* and *CalPKS3* were determined and contained 1554, 1536 and 1590 bp encoding complete reading frames of 389, 390 and 389 amino acid residues, respectively. The predicted molecular masses for the encoded proteins are 42.6, 42.9 and 42.6 kDa, respectively.

Functional expression and characterization

The open reading frames of the putative polyketide synthases from *Cassia* were ligated into a T7 expression vector containing an amino terminal hexahistidine fusion tag. The proteins were then purified from *Escherichia coli* BL21(DE3) in two steps (sonication / centrifugation, cobalt affinity-chromatography) to yield electrophoretically homogeneous enzyme of approximately 42 kDa (Fig. 2). Per liter, the bacterial culture typically produced 1.1 mg of recombinant enzyme. Electrophoretically homogeneous enzyme was used in enzyme assays to determine the substrate / product specificity of the putative polyketide synthases.

A series of seven coenzyme A esters were tested together with [2-^{14}C]malonyl CoA for their capacity to serve as polyketide synthase “starter” molecules. The compounds used were acetyl-, *n*-butyryl-, isovaleryl-, *n*-hexanoyl-, benzoyl-, cinnamoyl-, and *p*-coumaroyl-CoA. In each case, multiple products were formed as determined by phosphorimagery following resolution of the reaction mixtures by TLC (Fig. 3a). This broad product spectrum had been reported previously for several of the CoA-esters with chalcone synthase, for example, from *Pinus sylvestris*, *P. strobus* and *Sinapis alba* (Zuurbier et al. 1998) *Scutellaria baicalensis* (Morita et al. 2000) and from raspberry *Rubus idaeus* (Zheng et al. 2001). In order to thoroughly characterize the heterologously expressed *Cassia* cDNAs, the reactions were repeated with unlabeled substrates on a larger scale so that the reaction products could be analyzed by mass

spectrometry. The chemical structures of the compounds identified by liquid chromatography electrospray ionization mass spectrometry (LC/ESI-MS) are summarized in Fig. 4 and the corresponding mass spectral data are shown in Table 2. With the exception of the assay in which acetyl-CoA is used as the starter molecule, all substrates yielded a phloroglucinol derivative resulting from three sequential condensations of acetate units derived from three malonyl-CoA decarboxylations. The tetraketide intermediate thus formed cyclizes by an intramolecular Claisen condensation into the hydroxylated aromatic ring system. Every substrate tested in the assay also produced two pyrone derivatives, one resulting from two acetate unit condensations (a bis-noryangonin-type pyrone derailment product) and one resulting from three acetate unit condensations (a 4-coumaroyltriacetic acid lactone-type pyrone derailment). Unlike the complete reaction that leads to the formation of naringenin chalcone, these condensations do not result in the formation of an aromatic ring.

All three *Cassia* polyketide synthases produced the same spectrum of enzymatic products. This capacity to build multiple polyketides from a variety of CoA-substrates raised the question as to the physiological reaction catalyzed by these polyketide synthases *in planta*.

Tissue-specific expression

The accumulation of the *Cassia* polyketide synthase transcript was investigated in developing seedlings and plants of *C. alata*. Since the nucleotide sequences of the three polyketide synthase clones were more than 90% identical, no attempt was made to differentially detect the transcripts. RNA was isolated from seedlings of 0.5-5cm in length (Fig. 5a). RNA gel blot analysis indicated that transcript was present already in 0.5 cm long seedlings (Fig. 5b, Lanes 1-6). As the seedlings grow, the level of polyketide synthase transcript increases. A 5 cm long seedling contains approximately the same level of transcript as mature root (Fig. 5b, Lanes 7-8).

In the mature *C. alata* plant, the highest level of polyketide synthase transcript was found in young root, whereas much less accumulated in leaf and stem (Fig. 6a-c). It was surprising to find a root-specific polyketide synthase, with a substrate and product spectrum similar to that of chalcone synthase, in a plant known to accumulate anthraquinones in both aerial and underground tissue. To investigate the phenylpropanoid composition of leaf and root tissue, an acidic methanol extract was

prepared and was analyzed by reversed-phase HPLC. Quercetin was found to be a major component of root (Fig. 7c) and was also present as part of a more complex mixture in leaf (Fig. 7b). Given the distribution of gene transcript and of flavonoids, it is reasonable to assume that the *C. alata* root-specific polyketide synthases are chalcone synthases that demonstrate broad substrate specificity *in vitro*. *CalPKS1*, *CalPKS2* and *CalPKS3* were thus renamed *CalCHS1*, *CalCHS2* and *CalCHS3*, respectively.

Conclusions

As part of our investigation of type III polyketide synthases that occur in Thai medicinal plants, we have isolated three cDNAs that encode chalcone synthase. The gene transcripts were found predominantly in roots, which also accumulate the flavonoids quercetin, naringenin and kaempferol. The heterologously expressed cDNAs show broad substrate specificity for non-physiological substrates and produce multiple aromatic and pyrone derailment products from each substrate. The enzymes can accept CoA-esters that are not derived from phenylpropanoid metabolism and convert them to compounds that are not known to occur in the *Cassia* plant. Jez et al. (2000a) compare the crystal structure of *Medicago sativa* chalcone synthase to that of *Gerbera hybrida* pyrone synthase. From that study, three catalytic residues are present in the active sites of both type III plant polyketide synthases (Cys 164, His 303, Asn 336 in chalcone synthase and Cys 169, His 308, Asn 341 in pyrone synthase). The catalytic mechanism for polyketide formation is expected to be equivalent in both enzymes. Pyrone synthase, however, has a smaller initiation/elongation cavity that cannot accommodate the larger starter molecule *p*-coumaroyl CoA that is requisite to naringenin formation. Four amino acid residues have been identified as important to cavity size (Thr 197, Ile 254, Gly 256, Ser 338 in chalcone synthase and Leu 202, Met 259, Leu 261, Ile 343 in pyrone synthase). The *Cal*CHS cDNAs contain catalytic and initiation/elongation cavity residues identical to those of *M. sativa* chalcone synthase. Jez et al. (2000a) report methylpyrone formation by the *M. sativa* chalcone synthase, which is consistent with results reported herein. The initiation/elongation cavity of *Cal*CHS is, by sequence comparison, expected to be as large as that of the *M. sativa* chalcone synthase. It is therefore not surprising that starter molecules that are smaller than *p*-coumaroyl CoA can be accommodated by the enzyme with varying degrees of efficiency. The relative amount of products should depend on the assay conditions employed. Formation of multiple products from multiple starter molecules *in vitro* is likely a result of the large initiation/elongation cavity of *Cal*CHS and does not

necessarily suggest *in vivo* activity. Additional three dimensional structure determinations for type III polyketide synthases with varying substrate and product specificities will certainly provide insight into the efficiency of various starter molecules and the formation of derailment products for this important class of enzymes.

Roots of *Cassia* were also found to accumulate the potentially polyketide-derived anthraquinones chrysophanol, emodin and rhein as well as the anthracenone germichrysone (data not shown). Since the only cDNAs that were isolated by the RT-PCR approach used herein clearly encoded chalcone synthase, the nature of the enzyme that catalyzes formation of the anthraquinone skeleton remains to be identified.

Acknowledgements

This work was supported by the Thai Royal Golden Jubilee Fund, Bangkok, the Deutsche Akademische Austausch Dienst, Bonn, the Alexander von Humboldt Foundation, Bonn and by Fonds der Chemischen Industrie, Frankfurt.

References

Akada S, Kung SD, Dube SK (1991) The nucleotide sequence of gene 1 of the soybean chalcone synthase multigene family. Plant Mol Biol 16: 751-752

Akada S, Kung SD, Dube SK (1993) Nucleotide sequence and putative regulatory elements of gene 2 of the soybean (*Glycine max*) chalcone synthase multigene family. Plant Physiol 102: 317-319

Akiyama T, Shibuya M, Liu,H-M, Ebizuka Y (1999) *p*-Coumaroyltriacetic acid synthase, a new homologue of chalcone synthase, from *Hydrangea macrophylla* var. *thunbergii*. Eur J Biochem 263: 834-839

Crockett CO, Guede-Guina F, Pugh D, Vangah-Manda M, Robinson TJ, Olubadewo JO, Ochillo RF (1992) *Cassia alata* and the preclinical search for therapeutic agents for the treatment of opportunistic infections in aids patients. Cell Mol Biol 38: 799-802

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation. Plant Mol Biol Rep 1: 19-21

Elujoba AA, Ajulo OO, Iweibo GO (1989) Chemical and biological analyses of Nigerian

Cassia species for laxative activity. J Pharm Biomed Anal 7: 1453-1457

Funa N, Ohnishi Y, Fujii I, Shibuya M, Ebizuka Y, Horinouchi S (1999) A new pathway for polyketide synthesis in microorganisms. Nature 400: 897-899

Ichinose Y, Kawamata S, Yamada T, An C, Kajiwara T, Shiraishi T, Oku H (1992) Molecular cloning of chalcone synthase cDNAs from *Pisum sativum*. Plant Mol Biol 18: 1009-1012

Jez JM, Austin MB, Ferrer J-L, Bowman ME, Schröder J, Noel JP (2000a) Structural control of polyketide fromation in plant-specific polyketide synthases. Chem Biol 7: 919-930

Jez JM, Ferrer J-L, Bowman ME, Dixon RA, Noel JP (2000b) Dissection of malonyl-Coenzyme A decarboxylation from Polyketide formation in the reaction mechanism of a plant polyketide synthase. Biochemistry 39: 890-902

Kelly TR, Ma Z, Xu W (1994) Revision of the structure of alatinone to emodin. Phytochemistry 36: 253-254

Kreuzaler F, Hahlbrock K (1975) Enzymatic synthesis of aromatic compounds in higher plants. Formation of bis-noryangonin (4-hydroxy-6[4-hydroxystyryl]2-pyrone) from *p*-coumaroyl-CoA and malonyl-CoA. Arch Biochem Biophys 169: 84-90

Lans C, Harper T, Georges K, Bridgewater E (2000) Medicinal plants used for dogs in Trinidad and Tobago. Prev Vet Med 45: 201-220

Melchior F, Kindl H (1990) Grapevine stilbene synthase cDNA only slightly differing from chalcone synthase cDNA is expressed in *Escherichia coli* into a catalytically active enzyme. FEBS Lett 268: 17-20

Morita H, Takahashi Y, Noguchi H, Abe I (2000) Enzymatic formation of unnatural aromatic polyketides by chalcone synthase. Biochem Biophys Res Commun 279: 190-195

Ryder TB, Hedrick SA, Bell JN, Liang XW, Clouse SD, Lamb CJ (1987) Organization and differential activation of a gene family encoding the plant defense enzyme chalcone synthase in *Phaseolus vulgaris*. Mol Gen Genet 210: 219-233

Salzman RA, Fujita T, Zhu-Salzman K, Hasagawa PM, Bressan RA (1999) An improved RNA isolation method for plant tissue containing high levels of phenolic compounds or carbohydrates. Plant Mol Biol Rep 17: 11-17

Schüz R, Heller W, Hahlbrock K (1983) Substrate specificity of chalcone synthase from *Petroselinum hortense*. Formation of phloroglucinol derivatives from aliphatic substrates. J Biol Chem 258: 6730-6734

Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C (1994) Cloning and molecular

analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). Plant Mol Biol 24: 743-755

Springob K, Lukacin R, Ernwein C, Gröning I, Matern U (2000) Specificities of functionally expressed chalcone and acridone synthases from *Ruta graveolens*. Eur J Biochem 267: 6552-6559

Suh D-Y, Fukuma K, Kagami J, Yamazaki Y, Shibuya M, Ebizuka Y (2000) Identification of amino acid residues important in the cyclization reactions of chalcone and stilbene synthases. Biochem J 350: 229-235

Yadav SK, Kalidhar SB (1994) Alquinone: an anthraquinone from *Cassia alata*. Planta Med 60: 601

Yamaguchi T, Kurosaki F, Suh D-Y, Sankawa U, Nishioka M, Akiyama T, Shibuya, M, Ebizuka Y (1999) Cross-reaction of chalcone synthase and stilbene synthase overexpressed in *Escherichia coli*. FEBS Lett 460: 457-461

Zheng D, Schröder G, Schröder J, Hrazdina G (2001) Molecular and biochemical characterization of three aromatic polyketide synthase genes from Rubus idaeus. Plant Mol Biol 46: 1-15

Zuurbier KWM, Leser J, Berger T, Hofte AJP, Schröder G, Verpoorte R, Schröder J (1998) 4-Hydroxy-2-pyrone formation by chalcone and stilbene synthase with nonphysiological substrates. Phytochemistry 49: 1945-1951

Table 1 Amino acid similarity of *C. alata* root-specific type III polyketide synthases

	*Cal*PKS1[a]	*Cal*PKS2	*Cal*PKS3	*Psa*CHS	*Gma*CHS
*Cal*PKS1	100[b]	92	98	92	92
*Cal*PKS2		100	93	92	93
*Cal*PKS3			100	92	93
*Psa*CHS				100	92
*Gma*CHS					100

[a] *Cal*PKS1, *Cassia alata* type III polyketide synthase 1 (this work); *Psa*CHS, *Pisum sativum* chalcone synthase (Ichinose et al. 1992); *Gma*CHS, *Glycine max* chalcone synthase (Akada et al. 1991).

[b] % amino acid identity

```
                                                    50
MVKVEEIRKAQRAEGAATVMAIGTATPANCVEQSTYPDYYFRVTNSEHMT   CalCHS1
--S----------Q-P---L-------P---D----------I-----K-   CalCHS2
------------------------------D----------I-------   CalCHS3
                                                   100
ELKEKFQRMCDKSMIKKRYMHLTEEILKENPNMCAYMAPSIDARQDIVVL   CalCHS1
------K---E-------------DL--------------L-----M--V   CalCHS2
--------------------------------------------------   CalCHS3
                                                   150
EVPKLGKEAATKAIKEWGQPKSKITHLIFCTTSGVDMPGADYQLTKLLGL   CalCHS1
----------------------------V---------------------   CalCHS2
--------------------------------------------------   CalCHS3
                                                   200
RPSVKRYMMYQQGCFAGGTVLRLAKDLAENNKGARVLVVCSEITAVTFRG   CalCHS1
--Y-----------------------------------------------   CalCHS2
--------------------------------------------------   CalCHS3
                                                   250
PSDTHLDSLVGQALFGDGAAAVIVGSDPIPQVETPLFELVWTAQTILPDS   CalCHS1
--------------------I------L-E--K----------------   CalCHS2
------------------------------E--K----------------   CalCHS3
                                                   300
EGAIDGHLREVGLTFHLLKDVPGLISKNIEKALVEAFNPLGISDYNSIFW   CalCHS1
-----------------------------------Q------------   CalCHS2
-----------------------------------K------------   CalCHS3
                                                   350
IAHPGGPAILDQVEAKLGLKPEKMQATRHVLSEYGNMSSACVLFIMDEMR   CalCHS1
----------------E------R--------------------L----   CalCHS2
--------------------------------------------------   CalCHS3

RKSTKDGLGTTGEGLEWGVLFGFGPGLTVETVVLHSIAI              CalCHS1
---IQ---E------------------------R-V-VN              CalCHS2
-----------------------------------V-V               CalCHS3
```

Fig. 1 Comparison of translations of the nucleotide sequences of three polyketide synthase cDNAs isolated from root tissue of *C. alata*.

Table 2. Collision-induced dissociation (CID) mass spectra of enzymatically formed products from *Cal*CHS, compound 1-7 obtained from the $[M+H]^+$ ions by LC/ESI-MS

Substrate	Product	RT (min)[a]	$[M+H]^+$ (m/z)	Ions in the CID mass spectra (m/z, rel.int.)	CE (eV)
Acetyl CoA **(1a)**	**1c**	10.20	169[b]	169, 127 (100), 85 (50), 43 (23)	-20
	1b	11.10	127	127 (47), 109 (27), 99 (18), 87 (22), 71 (21), 69 (68), 67 (18), 65 (39), 53 (20), 43 (100)	-25
n-Butyryl CoA **(2a)**	**2c**	14.62	197	197 (1), 127 (100), 113 (8), 71 (46), 43 (25)	-20
	2b	15.14	155	155 (100) , 137 (12), 126 (68), 109 (15), 91 (40), 87 (41), 81 (55), 79 (45), 69 (70), 65 (18), 57 (18)	-25
	2d	17.60	197	197 (74), 179 (95), 155 (93), 139 (33), 123 (67), 113 (100), 95 (42), 43 (16)	-20
Isovaleryl CoA **(3a)**	**3c**	15.65	211	211 (2), 127 (100), 111 (14), 85 (55), 57 (57)	-20
	3b	16.00	169	169 (45), 126 (100), 105 (16), 87 (28), 69 (22), 67 (26), 43 (25)	-25
	3d	17.98	211	211 (53), 193 (20), 169(14), 165 (8), 155 (100), 151 (63), 139 (15), 129 (9), 123 (30), 57 (5), 43 (4)	-20
Hexanoyl CoA **(4a)**	**4c**	17.46	225	225 (10), 141 (6), 127 (100), 99 (27), 71 (13)	-15
	4b	17.81	183	183 (100), 155 (12), 141 (5), 126 (68), 109 (33), 99 (25), 95 (28), 87 (27), 81 (45), 71 (32), 69 (44), 67 (35), 57 (28)	-25
	4d	19.26	225	225 (35), 207 (100), 165 (15), 155 (52), 151 (10), 143 (12), 139 (61), 123 (10), 101 (35), 83 (17)	-20
Benzoyl CoA **(5a)**	**5c**	15.76	231	231 (5), 189 (3), 157 (4), 147 (22), 105 (100), 71 (4)	-20
	5b	16.20	189	189 (100), 171 (95), 161 (28), 147 (10), 129 (55), 127 (75), 115 (63), 105 (72), 103 (43), 77 (17), 69 (87)	-25
	5d	16.61	231	231 (4), 153 (100), 105 (28)	-20
Cinnamoyl CoA **(6a)**	**6c**	17.13	257	257 (5), 221 (2), 173 (8), 131 (100), 103 (3), 69 (2)	-20
	6b	17.93	215	215 (28), 197 (18), 187 (15), 179 (7), 169 (45), 155 (20), 141 (80), 131 (100), 129 (15), 103 (15), 69 (28)	-25
	6d	19.53	257	257 (25), 153 (100), 131 (55), 123 (7), 103 (11), 77 (6), 69 (5)	-25
p-Coumaroyl CoA **(7a)**	**7c**	15.59	273	273 (2), 189 (5), 147 (100), 119 (12)	-25
	7b	16.26	231	231 (35), 213 (12), 203 (33) 171 (12), 157 (25), 147 (100), 121 (15), 110 (5), 69 (4)	-20
	7d	17.84	273	273 (7), 255 (9), 153 (100), 147 (23), 123 (4), 85 (5)	-25

CE = collision energy.

[a]Retention times (RT) were obtained from the corresponding ion chromatogrames in the HPLC.

[b]compound **1c** was identified by selected reaction monitoring (SRM), the relative abundances were calculated from the peak areas of the three detected reactions m/z 169 m/z ⟶ 127, m/z 169 ⟶ m/z 85, and m/z 169 ⟶ m/z 43.

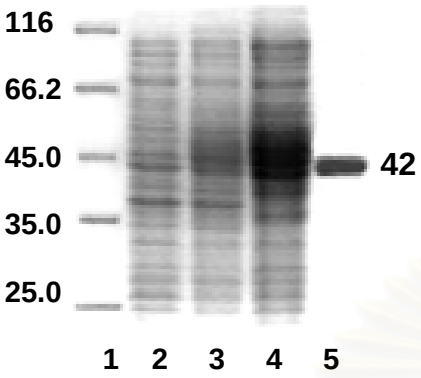


Fig. 2 SDS-PAGE analysis of the purification of recombinant *Cal*PKS1 visualized with Coomassie Brilliant Blue R250. Lane 1, molecular mass markers (kDa); Lane 2, un-induced *E. coli* cells; Lane 3, induced cells; Lane 4, total cellular lysate and Lane 5, *Cal*PKS1 purified by affinity chromatography on a cobalt resin.

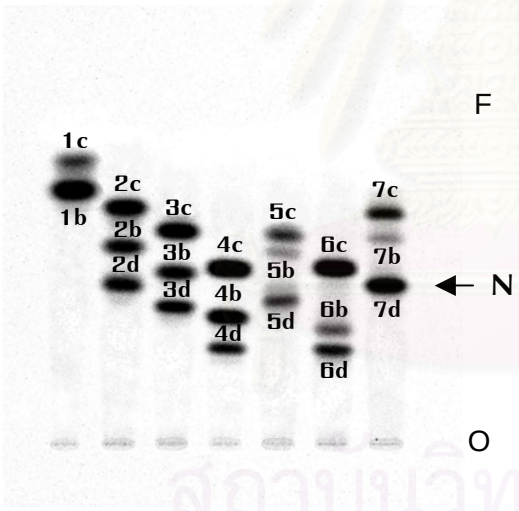
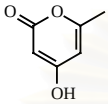
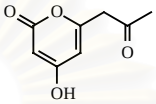
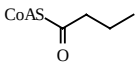
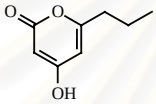
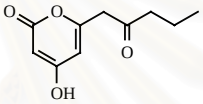
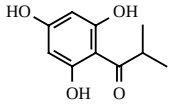
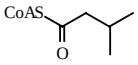
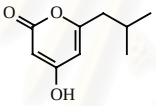
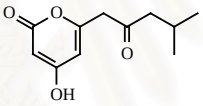
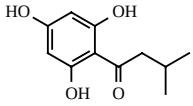
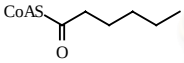
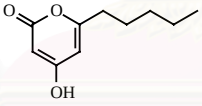
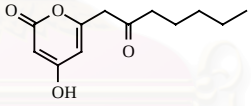
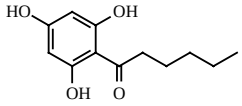
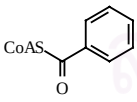
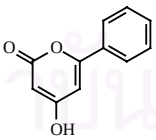
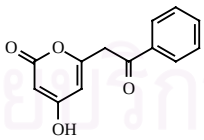
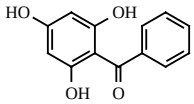
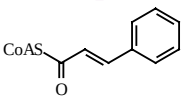
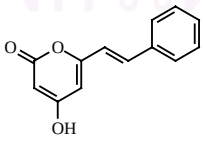
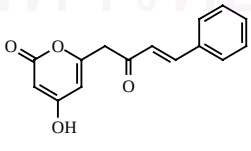
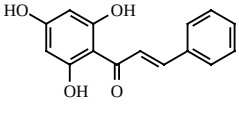
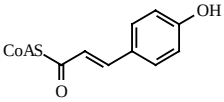
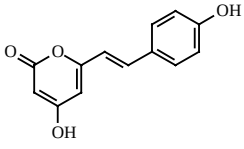
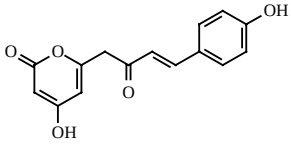
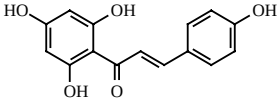


Fig. 3a,b a) Radiogram of *Cal*PKS1 reaction products resolved by reverse-phase (C18) TLC. The products were tentatively identified by comparison to authentic standards. The substrates used in the enzyme assays were: Lane 1, acetyl-CoA + [2-^{14}C]malonyl-CoA; Lane 2, *n*-butyryl-CoA + [2-^{14}C]malonyl-CoA; Lane 3, isovaleryl-CoA + [2-^{14}C]malonyl-CoA; Lane 4, *n*-hexanoyl-CoA + [2-^{14}C]malonyl-CoA; Lane 5, benzoyl-CoA + [2-^{14}C]malonyl-CoA; Lane 6, cinnamoyl-CoA + [2-^{14}C]malonyl-CoA; Lane 7, *p*-coumaroyl-CoA + [2-^{14}C]malonyl-CoA. O, TLC origin and F, solvent front. b) Assignment of the spots on the TLC plate in a) to the corresponding compounds given in Figure 4 and Table 2.

Fig. 4 Structures of the "starter" molecule substrates and of the reaction products resulting from incubation with *Cal*PKS1. The chemical structures of the reaction products were determined by LC-ESI mass spectrometry.

Starter-CoA	**Enzymatic product**		
	BNY-type	**CTAL-type**	**Chalcone-type**
 acetyl- CoA **(1a)**	 *m/z* 127 $[M+H]^+$ **(1b)**	 *m/z* 169 $[M+H]^+$ **(1c)**	
 *n*-butyryl-CoA **(2a)**	 *m/z* 155 $[M+H]^+$ **(2b)**	 *m/z* 197 $[M+H]^+$ **(2c)**	 *m/z* 197 $[M+H]^+$ **(2d)**
 *Iso*valeroyl-CoA **(3a)**	 *m/z* 169 $[M+H]^+$ **(3b)**	 *m/z* 211 $[M+H]^+$ **(3c)**	 *m/z* 211 $[M+H]^+$ **(3d)**
 Hexanoyl-CoA **(4a)**	 *m/z* 183 $[M+H]^+$ **(4b)**	 *m/z* 225 $[M+H]^+$ **(4c)**	 *m/z* 225 $[M+H]^+$ **(4d)**
 Benzoyl-CoA **(5a)**	 *m/z* 189 $[M+H]^+$ **(5b)**	 *m/z* 231 $[M+H]^+$ **(5c)**	 *m/z* 231 $[M+H]^+$ **(5d)**
 Cinnamoyl-CoA **(6a)**	 *m/z* 215 $[M+H]^+$ **(6b)**	 *m/z* 257 $[M+H]^+$ **(6c)**	 *m/z* 257 $[M+H]^+$ **(6d)**
 *p*-Coumaroyl-CoA **(7a)**	 *m/z* 231 $[M+H]^+$ **(7b)**	 *m/z* 273 $[M+H]^+$ **(7c)**	 *m/z* 273 $[M+H]^+$ **(7d)**

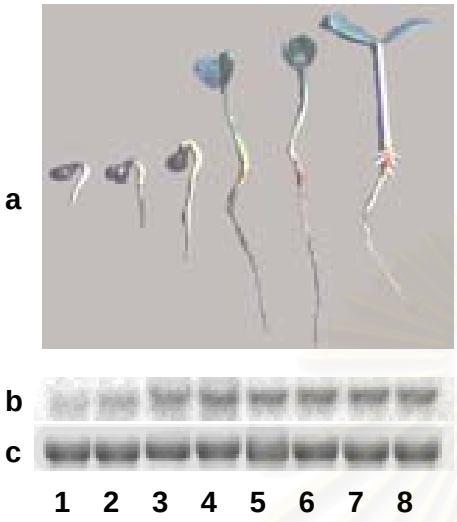


Fig. 5a-c RNA gel blot analysis of *Cal*PKS1 expressed in developing seedlings of *C. alata*. a) seedling development from 0.5 to 5 cm; b) result obtained after RNA blotting and hybridization to the *Cal*PKS1 full-length cDNA visualized by phosphorimagery; c) negative image of the rRNA loading control. Lane 1, 0.5 cm seedling; Lane 2, 1 cm seedling; Lane 3, 2 cm seedling; Lane 4, 3 cm seedling; Lane 5, 4 cm seedling; Lane 6, 5 cm seedling; Lanes 7-8, mature root.

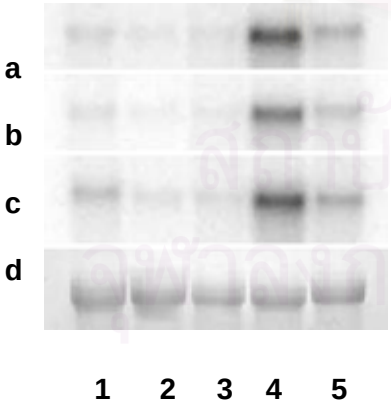


Fig. 6a-d RNA gel blot analysis of *Cal*PKS1*, 2* and *3* expressed in mature *C. alata* plants. a) blot hybridized to *Cal*PKS1; b) *Cal*PKS2; c) *Cal*PKS3; d) negative image of the rRNA loading control. Lane 1, young leaf; Lane 2, mature leaf; Lane 3, stem; Lane 4, young root; Lane 5, mature root.

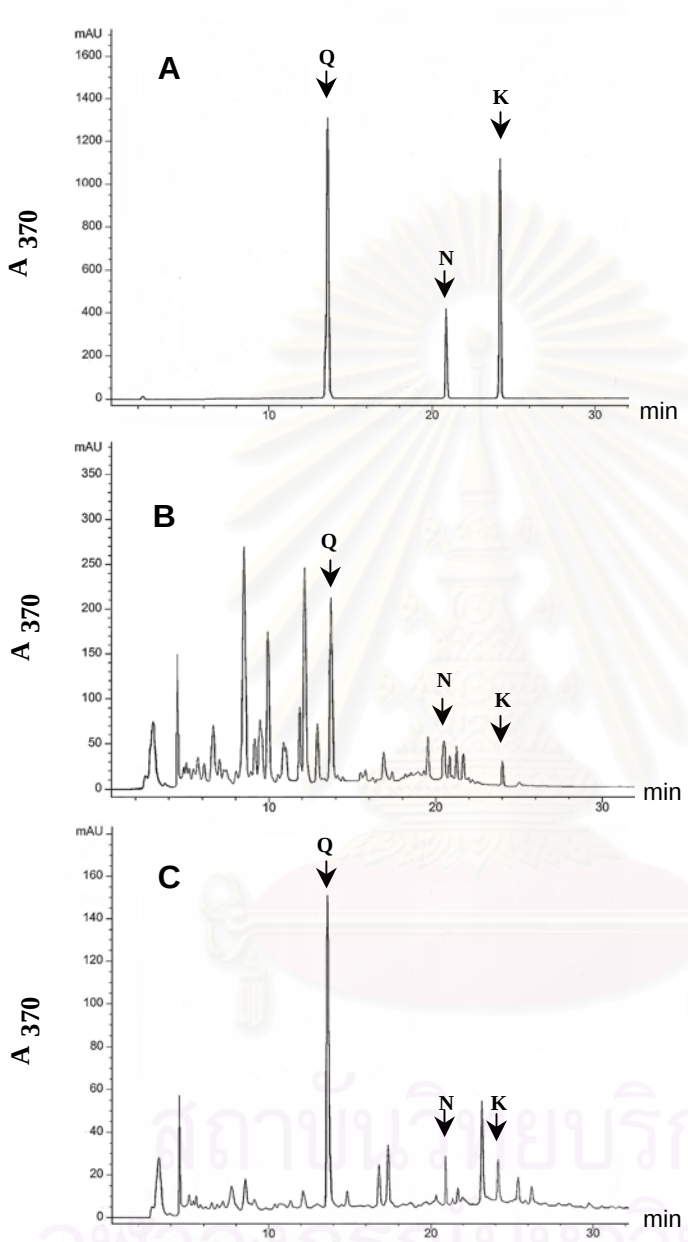


Fig. 7a-c Reversed-phase HPLC chromatogram of an acidic methanol extract of mature *C*. *alata* plant parts. a) authentic flavonoid standards; b) leaf extract; c) root extract. Q, quercetin (13.6 min); N, naringenin (20.9 min); K, kaempferol (24.9 min).

APPENDIX II
Original manuscript II

The title “Aromatic and pyrone polyketides synthesized by a stilbene synthase from *Rheum tataricum*” has been accepted by Phytochemistry.

Dedicated to Meinhart H. Zenk on the occasion of his 70th birthday.

Aromatic and pyrone polyketides synthesized by a stilbene synthase from *Rheum tataricum*

Supachai Samappito[a,b,c], Jonathan E. Page[a,*], Jürgen Schmidt[a],
Wanchai De-Eknamkul[b], Toni M. Kutchan[a]

[a]Leibniz-Institut für Pflanzenbiochemie, Weinberg 3, D-06120 Halle (Saale), Germany
[b]Department of Pharmacognosy, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand

The nucleotide sequence reported in this paper has been submitted to the GenBank/EMBL/DDBJ with the accession number AF508150.

[*]Corresponding author. Tel.: +49-345-5582-1220; fax: +49-345-5581-1209
E-mail address: jpage@ipb-halle.de
[c]Present address: Department of Biotechnology, Faculty of Technology, Mahasarakham University, Mahasarakham 44001, Thailand

Abstract

A cDNA encoding a stilbene synthase, RtSTS, was isolated from the rhizomes of Tatar rhubarb, *Rheum tataricum* L. (Polygonaceae), a medicinal plant containing stilbenes and other polyketides. Recombinant RtSTS was expressed in *E. coli* and assayed with acetyl-coenzyme A (CoA), *n*-butyryl-CoA, isovaleryl-CoA, *n*-hexanoyl-CoA, cinnamoyl-CoA and *p*-coumaroyl-CoA as primers of polyketide synthesis. RtSTS synthesized resveratrol and a trace amount of naringenin chalcone from *p*-coumaroyl-CoA, supporting the enzyme's identification as a resveratrol-type stilbene synthase (EC 2.3.1.95). Bis-noryangonin and *p*-coumaroyl triacetic acid lactone (CTAL)-type pyrones were observed in minor amounts in the reaction with *p*-coumaroyl-CoA and as major products with cinnamoyl CoA. As well, such pyrones, and not aromatic polyketides, were identified as the only products in assays with aliphatic and benzoyl CoA esters. Acetonyl-4-hydroxy-2-pyrone, a pyrone synthesized from acetyl-CoA, was identified as a new product of a stilbene synthase. Using Northern blot analysis, *RtSTS* transcript was found to be highly expressed in *R. tataricum* rhizomes, with low transcript levels

also present in young leaves. This expression pattern correlated with the occurrence of resveratrol, which was detected in higher amounts in *R. tataricum* rhizomes compared with leaves and petioles using HPLC. Few stilbene synthases have been found in plants, and the identification of *RtSTS* provides additional sequence and catalytic information with which to study the evolution of plant polyketide synthases.

Keywords: *Rheum tataricum* L., Polygonaceae, Tatar rhubarb, polyketide synthase, stilbene synthase, chalcone synthase, resveratrol, pyrone.

Abbreviations:
ACS, acridone synthase,
BAS, benzalacetone synthase
BBS, bibenzyl synthase
BNY, bis-noryangonin
CoA, coenzyme A
CHS, chalcone synthase
CID, collision-induced decomposition
CTAL, *p*-coumaroyl triacetic acid lactone
CTAS, *p*-coumaroyl triacetic acid synthase
IPTG, isopropyl-β-D-thiogalactopyranoside
PKS, polyketide synthase
2-PS, 2-pyrone synthase
RACE, rapid amplification of cDNA ends
RtSTS, *Rheum tataricum* stilbene synthase
SDS, sodium dodecyl sulfate
SSC, saline sodium citrate
STS, stilbene synthase
UPM, universal primer mix
VPS, valerophenone synthase

Introduction

The chalcone synthase superfamily of polyketide synthases (PKSs) are key enzymes in the biosynthesis of many plant natural products including the ubiquitous flavonoids and related anthocyanins, and stilbenes, pyrones, phloroacylphenones and acridone alkaloids of more restricted distribution (Schröder, 2000). The enzymes responsible are

homodimeric type III PKSs that, using aromatic or aliphatic CoA esters as reaction primers, carry out sequential decarboxylation and condensation reactions with one, two or three C_2 units derived from malonyl-CoA. The prototype enzyme of the superfamily is chalcone synthase (CHS) (EC 2.3.1.74), which has been extensively investigated at the biochemical (Jez and Noel, 2000) and structural level (Ferrer et al., 1999). CHS uses *p*-coumaroyl-CoA to form naringenin chalcone via a Claisen-type cyclization, with this compound further cyclizing to give naringenin. A closely related plant PKS, stilbene synthase (STS) (EC 2.3.1.-), also uses phenylpropanoid CoA starter esters but cyclizes the tetraketide intermediate in an aldol-type reaction to yield resveratrol, from *p*-coumaroyl-CoA (EC 2.3.1.95), or pinosylvin, from cinnamoyl-CoA (EC 2.3.1.146).

Plant PKSs are catalytically flexible, usually forming several products *in vitro*, and accepting a variety of physiological and non-physiological substrates. Both chalcone and stilbene synthases convert *p*-coumaroyl-CoA to the major products naringenin chalcone or resveratrol, respectively, as well as to the byproduct pyrones bis-noryangonin (BNY) and *p*-coumaroyl triacetic acid lactone (CTAL) (Yamaguchi et al., 1999). The latter are formed by derailment after two or three condensation reactions with malonyl-CoA, respectively. Additionally, cross reactivity between CHS and STS has been demonstrated with CHS forming resveratrol and STS forming naringenin in assays with *p*-coumaroyl-CoA, albeit in small amounts (Yamaguchi et al., 1999). CHSs and STSs accept aliphatic CoA esters as reaction primers. For example, *Scutellaria baicalensis* CHS converts isovaleryl-CoA to phloroacylphenones (via a correct Claisen cyclization) (Morita et al., 2000), while *Arachis hypogaea* STS forms BNY-type pyrones rather than aromatic products with this substrate (Morita et al., 2001). CHS and STS also utilize thiophene, furan and halogenated analogues of *p*-coumaroyl-CoA to form unnatural polyketides (Abe et al., 2000; Morita et al., 2001). The ability of plant PKSs to accept different CoA primers and carry out various condensation and cyclization reactions (Jez et al., 2002) makes this enzyme class, along with terpene cyclases (Greenhagen and Chappell, 2001), one of the major generators of carbon skeleton diversity in natural products.

As part of a study of the biosynthesis of polyketides present in medicinal plants, we investigated the CHS superfamily members present in *Rheum tataricum* L. (Polygonaceae), or Tatar rhubarb. *Rheum* species are a rich source of polyketides including phenylbutanoids, anthraquinones, naphthalenes and stilbenes (Kashiwada et

al., 1988). Members of this genus are also used as medicinal plants (Foust, 1992), particularly in Asian traditional medicine, and stilbenes have been reported to mediate the antioxidant activity of *Rheum* extracts (Matsuda et al., 2001). It is interesting to note that the first *in vitro* evidence for STS activity was shown with protein extracts of *Rheum rhaponticum* rhizomes (Rupprich and Kindl, 1978), although the enzyme responsible was never isolated using biochemical or molecular approaches. Abe et al. (2001) recently reported the isolation of a cDNA encoding benzalacetone synthase (BAS), a PKS involved in phenylbutanoid biosynthesis, from the leaves of *Rheum palmatum*. In this paper, we describe the cloning of *RtSTS*, a resveratrol-type STS from *Rheum tataricum* rhizomes and the characterization of its *in vitro* reaction products.

Results and Discussion

A homology-based approach with degenerate oligonucleotide primers that correspond to conserved regions of plant PKSs was used to amplify gene fragments from *Rheum tataricum* rhizome RNA by RT-PCR. Rhizomes were targeted because many of the polyketides of interest, including stilbenes and anthraquinones, are synthesized in this tissue. In this manner, two 584 bp fragments, *RtPKS1* and *RtPKS2*, that showed similarity to CHS superfamily members were isolated. Since these fragments showed 98.5% similarity at the nucleotide level, only *RtPKS1* was selected for further analysis. The cDNA regions 5' and 3' from this fragment were obtained using RACE PCR. The full-length *RtSTS* cDNA was 1429 bp encoding a protein of 391 amino acids with a predicted molecular weight of 43.0 kDa. The 59 bp 5' untranslated region contained a stop codon in-frame with the presumed start codon indicating the cDNA was full length. The deduced amino acid sequence of *RtSTS* is shown in Fig. 1, as a multiple sequence alignment with other plant CHS superfamily members. Southern blot analysis (data not shown) found that three copies of *RtSTS* are present in the *R. tataricum* genome.

*Rt*STS was expressed in *E. coli* and purified by immobilized metal affinity chromatography on a cobalt resin to give homogeneous recombinant enzyme (Fig. 2). In order to determine the catalytic activity of RtSTS, recombinant enzyme was incubated with ^{14}C-malonyl-CoA and the potential starter CoA esters, acetyl-CoA, *n*-butyryl-CoA, isovaleryl-CoA, *n*-hexanoyl-CoA, cinnamoyl-CoA and *p*-coumaroyl-CoA. The resulting radioactive products were resolved by reversed-phase thin layer chromatography. We acidified enzyme assays before extraction to increase recovery of

pyrones (Yamaguchi et al., 1999) and prevent cyclization of chalcones to their corresponding flavanones. As shown in Fig. 3, RtSTS accepted both aliphatic and aromatic CoA esters as primers for polyketide synthesis and produced multiple products from each. The major product formed in the *p*-coumaroyl-CoA assay co-chromatographed with resveratrol, suggesting that the *Rheum* PKS encoded an STS.

To conclusively identify the polyketide products, RtSTS was incubated with unlabelled malonyl-CoA and the starter CoA esters in scaled up *in vitro* reactions and the resulting product mixtures analyzed by LC-MS and LC-MS/MS. Products were identified by their parent ions and collision-induced decomposition (CID) spectra. We have previously published the results from the LC-MS analysis of a series of polyketides formed enzymatically from CHS (Samappito et al. 2002); therefore, only a graphical summary of the enzymatic products of RtSTS is presented here (Table 1). RtSTS formed aromatic products only with *p*-coumaroyl-CoA and cinnamoyl-CoA. Incubation with *p*-coumaroyl-CoA gave resveratrol (**7e**) (*m/z* 229, $[M+H]^+$ and *m/z* 227, $[M-H]^-$) as the major product, demonstrating that RtSTS is an STS. Resveratrol was identified by its electrospray CID spectrum of the $[M-H]^-$ ion (*m/z* 227) displaying key ions at *m/z* 185 ($[M-H-CH_2CO]^-$) and *m/z* 143 ($[M-H-2CH_2CO]^-$) in comparison with an authentic sample. Smaller amounts of BNY-type (**7b**) and CTAL-type (**7c**) pyrones, and naringenin chalcone (**7d**), were also detected in the *p*-coumaroyl-CoA assay. This spectrum of four products, including the naringenin chalcone formed as a cross-reaction product, has been previously described for *Arachis hypogaea* STS (Yamaguchi et al., 1999). RtSTS synthesized BNY-type (**6b**) and CTAL-type (**6c**) pyrones as well as a small amount of pinocembrin chalcone (**6d**) and pinosylvin (**6e**) from cinnamoyl-CoA. Pinocembrin chalcone was clearly identified by its parent ion (*m/z* 257 $[M+H]^+$) and the typical collision-induced fragments at *m/z* 153 (trihydroxybenzoyl ion) and *m/z* 131 (cinnamoyl ion), while pinosylvin (**6e**) was detected by the LC-MS analysis showing a peak with masses of *m/z* 213 ($[M+H]^+$) and *m/z* 211 ($[M-H]^-$). Only three products were visible in the ^{14}C-labeled assay with cinnamoyl CoA (Fig. 3), the least polar (lower R_f value) of which likely corresponds to an unresolved mixture of pinocembrin chalcone and pinosylvin.

With benzoyl-CoA and the aliphatic substrates tested, RtSTS catalyzed only the formation of BNY-type (**1b**, **2b**, **3b**, **4b**, **5b**) and CTAL-type (**1c**, **2c**, **3c**, **4c**, **5c**) pyrones as a result of derailment after two and three condensation reactions. With some notable exceptions, the exclusive formation of pyrones by plant PKSs seems to be an indicator that the starter CoA ester in the reaction is not the physiologically relevant substrate for the enzyme under assay. This phenomena was observed in the assays of RtSTS with aliphatic and benzoyl CoA substrates. Pyrones also form as byproducts in reactions with correct substrates but the amounts are generally minor compared to the main aromatic product. In a limited number of cases, pyrone formation appears to be the true catalytic activity of plant PKS enzymes. 2-Pyrone synthase from *Gerbera hybrida* forms 6-methyl-4-hydroxy-2-pyrone as its major product *in vitro*, with the involvement of this enzyme in pyrone metabolism supported by antisense “knockout” of 2-pyrone synthase in *G. hybrida* (Eckermann et al., 1998). A *p*-coumaroyl triacetic acid synthase from *Hydrangea macrophylla* may also represent an example of a pyrone-forming PKS, although this enzyme has been suggested to yield a linear tetraketide *in vitro* (Akiyama et al., 1999)

The formation of 6-acetonyl-4-hydroxy-2-pyrone (tetracetic acid lactone) (**1c**) (*m/z* 169 $[M+H]^+$), in addition to 6-methyl-4-hydroxy-2-pyrone (triacetic acid lactone) (**1b**) (*m/z* 127 $[M+H]^+$), when RtSTS was incubated with acetyl-CoA is notable. Compound **1b** has been previously described as the main product formed from the use of acetyl-CoA as a starter CoA ester by *Gerbera hybrida* 2-pyrone synthase (Eckermann et al., 1998) and CHS from *Petroselinum crispum* (Schuz et al., 1983). 6-Acetonyl-4-hydroxy-2-pyrone (**1c**) is a new enzymatic product from a plant PKS, although we have also observed it as an enzymatic product from *Senna alata* CHS (Samappito et al., 2002). Its synthesis is analogous to the formation of CTAL and likely involves the condensation of acetyl-CoA with three C_2 units derived from malonyl-CoA followed by a pyrone-type cyclization. This compound has not been reported previously from plants, although it has been isolated as a polyketide metabolite of *Penicillium stipatatum* (Bentley and Zwitkowits, 1967).

To confirm that stilbenes were present in the tissue from which the *RtSTS* cDNA was isolated, we analyzed leaf, petiole and rhizome extracts of *Rheum tataricum* by HPLC. Acid hydrolysis of extracts was used to cleave sugar residues and release stilbenes as their aglycones. Resveratrol was detected in all tissues, with higher amounts occurring

in the rhizomes (Fig. 4). To the best of our knowledge, this is the first report of resveratrol from *R. tataricum*. The presence of this compound provides further evidence that the *in vivo* function of RtSTS is the biosynthesis of resveratrol. The distribution of resveratrol correlated with the expression pattern of the *RtSTS* transcript as determined by Northern blot analysis of total RNA isolated from young leaves, mature leaves, petioles and rhizomes (Fig. 5). Young leaves showed low expression levels, and the transcript was virtually absent from older leaves and petioles. A high expression level was detected in rhizomes where resveratrol levels were also highest.

A phylogenetic comparison of the relationships between *RtSTS* and other chalcone synthase superfamily members (Fig. 6) shows *RtSTS* to be closely related to *Rheum palmatum* BAS (82% amino acid identity) and forming a group with functionally diverse enzymes such as *Ruta graveolens* acridone synthase (ACS), *Gerbera hybrida* 2-pyrone synthase (2-PS) and *Vitis vinifera* STS. This group appears to be distinct from groups composed of angiosperm CHSs, and from PKSs from the Leguminosae (*Arachis* and *Medicago*), *Pinus* and *Psilotum*. Abe et al. (2001) analyzed the phylogenetic relationships of *Rheum palmatum* BAS and characterized the group to which this enzyme belongs as more primitive than flavonoid forming CHSs. While the sequences and methods used for phylogenetic analysis may differ from those used in this study, the close relationship of *R. palmatum* BAS with *RtSTS* and *V. vinifera* STS and *Humulus lupulus* CHS suggests that no evolutionary conclusions as to the primitive or advanced nature of PKSs can be drawn on the basis of a simple comparison of sequence and enzyme function. The close relationship of *Rheum palmatum* BAS and *R. tataricum* STS supports a trend, first noted by Tropf et al. (1994), that functionally distinct enzymes that occur in the same plant taxa are more similar to each other than they are to functionally identical proteins that occur in more distantly related taxa.

STSs are not common in the plant kingdom. In comparison to the more than 100 genes encoding CHS that have been isolated from different species of higher plants (as reported in Swiss-Prot (2002)), STSs have been found in only four plant genera: *Arachis* (Schröder et al., 1988), *Vitis* (Melchior and Kindl, 1990), *Pinus* (Fliegmann et al., 1992) and *Psilotum* (Yamazaki et al., 2001). This is due both to the restricted distribution of stilbenes and to the inability to separate stilbene from CHSs based on sequence analysis; some sequences annotated as CHSs in sequence databases may in fact encode STSs. The cloning and characterization of a resveratrol-type STS from

the medicinal plant *Rheum tataricum* provides a new example of a plant PKS participating in the biosynthesis of the pharmacologically and agriculturally important stilbene, resveratrol. It will be interesting to obtain the sequence of CHSs from *Rheum* and *Arachis*, so that a pairwise phylogenetic comparison of co-occurring CHS and STSs from *Vitis*, *Pinus*, *Arachis*, *Psilotum* and *Rheum* may be undertaken.7

Experimental

Chemicals

Acetyl-CoA, malonyl-CoA, *n*-butyryl-CoA, isovaleryl-CoA, *n*-hexanoyl-CoA, benzoyl-CoA, naringenin and resveratrol were purchased from Sigma. [2-^{14}C]-malonyl-CoA was from Biotrend Chemikalien (Cologne, Germany). *p*-Coumaroyl-CoA and cinnamoyl-CoA were kindly provided by D. Knöfel of the Abteilung Sekundärstoffwechsel, Leibniz-Institut für Pflanzenbiochemie (IPB), Halle.

General methods

Rheum tataricum was grown from seed in the greenhouse at the IPB, Halle. Total RNA was isolated as described by Salzman et al. (1999). Poly(A)$^+$ RNA was isolated with Oligotex beads (Qiagen) according to the manufacturer's instructions. For Northern blot analysis, 10 μg total RNA was resolved by electrophoresis and transferred to a nylon membrane (Hybond N, Amersham Pharmacia Biotech). The full-length cDNA was excised by restriction enzyme digest and labeled with ^{32}P using random primers. Hybridization was performed at 65° overnight and the membranes were then washed three times for 15 min with 2x SSC and 0.1% SDS at 65°. Radioactivity was visualized by phosphorimaging.

For Southern blot analysis, genomic DNA was isolated from young leaves according to Dellaporta et al. (1983). DNA (10 μg) was digested with *Apo*I, *Bcl*I and *Bst*XI (each cutting once within the reading frame), resolved by electrophoresis on a 0.8% agarose gel, and capillary blotted onto a Hybond N nylon membrane. The cDNA labeled as above was used as a hybridization probe as described for Northern blot analysis. Radioactivity was visualized by phosphorimaging.

Amplification of partial cDNAs from R. tataricum

First-strand cDNA was synthesized from 5 μg of total RNA isolated from *R. tataricum* rhizomes using Superscript II reverse transcriptase (Life Technologies). One μl of the first-strand cDNA was used as a PCR template with degenerate primers 5'-AA(A/G)GC (C/T)AT(A/C)GAIGA(A/G)TGGGG-3 and 5'-CCACCIGG(A/G)TGI(A/G)CAATCC-3' based on those described by Helariutta et al. (1995) and *Taq* polymerase (Promega). Cycling conditions consisted of an initial denaturation at 94º, 3 min; 30 cycles of 94º, 30 sec; 48º, 30 sec and 72º, 1 min, followed by 10 min at 72º. A PCR product of the expected length (575-600 bp) was purified by gel electrophoresis, ligated into pGEM-T Easy (Promega) and sequenced. Comparison of several fragments showed that two sequences of 98.5% similarity were amplified from *R. tataricum* rhizomes.

5' and 3' RACE PCR

A RACE PCR kit (SMART technology; Clontech) was used to synthesize cDNA from total RNA isolated from *R. tataricum* rhizomes. 3' RACE was performed using the gene-specific primer 5'-CATAGACTCCATGGTAGGGCAAGC-3' and universal primer A mix (UPM) supplied by the manufacturer to amplify a 755 bp DNA fragment. Cycling conditions consisted of an initial denaturation at 94º, 5 sec; 30 cycles of 94º, 5 sec; 68º, 10 sec and 72º, 3 min, followed by 7 min at 72º Advantage polymerase (Clontech). Similarly, 5' RACE used the gene-specific primer 5'-GCCGCACCGTCACC AAATATTGC-3' and UPM to amplify a 781 bp DNA fragment. Both 3' and 5' RACE PCR products were gel purified, ligated into pGEM T-Easy and sequenced.

Expression in *E. coli*

To express the cDNA in *E. coli*, the open reading frame was cloned into pET14b (Novagen), which contains a hexahistidine N-terminal fusion tag. Amplification was performed using primers 5'-AATAGTCATATGGCACCGGAGGAGTCG-3' (the NdeI site is underlined) and 5'-TTTAAAGGATCCTCAGGTAATTAGCGGCAC-3' (the *Bam*HI site is underlined) with the 5'-RACE cDNA as template and *Pfu* polymerase (Promega). The 1.2 kb PCR product was digested with *Nde*I and *Bam*HI, gel purified and ligated into *Nde*I/*Bam*HI-digested pET-14b. Cloning of the expression construct was confirmed by sequencing. The plasmid was transformed into *E. coli* BL21(DE3) cells, 1 l cultures in LB broth containing 50 g/ml ampicillin were grown at 37° until OD_{600} ~0.6 and protein expression induced with 1mM IPTG. Induced cultures were grown at 28° for 12 hours, harvested by centrifugation and frozen at -80°.

Purification of recombinant RtSTS

Frozen *E. coli* cells were resuspended in buffer containing 50 mM Tris-HCl (pH 7), 500 mM NaCl, 2.5 mM imidazole, 10% (v/v) glycerol, 10 mM β-mercaptoethanol, 1% Tween-20 and 750 μg/ml lysozyme. Cells were incubated for 30 min on ice followed by a brief sonication. The lysate was clarified by centrifugation and the recombinant protein bound to Talon resin (Clontech). After washing with lysis buffer lacking detergent and lysozyme, recombinant protein was eluted with buffer containing 50 mM Tris-HCl (pH 7), 500 mM NaCl, 250 mM imidazole, 10% (v/v) glycerol, 10 mM β-mercaptoethanol. The protein-containing fractions were pooled and the buffer exchanged to 50 mM HEPES (pH 7), 10 mM β-mercaptoethanol and 10% (v/v) glycerol using a gel filtration column (PD10, Amersham Pharmacia Biotech). Recombinant RtSTS was determined to be pure by electrophoresis on a denaturing polyacrylamide gel (Laemmli, 1970). Purified enzyme was stored at -20° before use.

Enzyme assays

Polyketide synthase activity was measured by the conversion of starter CoA esters and [2-^{14}C]malonyl-CoA into reaction products. The standard enzyme assay contained 100 mM HEPES buffer (pH 7), 20 μM starter CoA and 12,000 dpm [2-^{14}C]-malonyl-CoA (55 mCi/mmol) and 1.5 μg RtSTS in a 50 μl reaction volume. The assay mixture was incubated for 30 min at 30°. The reaction was stopped by addition of 5 μl 10% (v/v) HCl and was extracted twice with 100 μl ethylacetate. The combined organic phase was evaporated to dryness and the products separated by thin layer chromatography (RP18, Merck) developed in MeOH-H_2O-acetic acid (70:30:1). Selected reactions were co-chromatographed with a resveratrol standard. The ^{14}C-labelled products were visualized by phosphorimaging.

Mass spectrometric analysis of enzymatic products

To identify the enzymatic products, scaled up reactions containing 75 mM Tris-HCl (pH 7), 50 μM starter CoA, 100 μM malonyl-CoA and 5 μg purified enzyme in a 200 μl reaction volume were used. The reaction proceeded for 1 hour at 30º before acidification and extraction with ethyl acetate. After drying *in vacuo*, the residue was dissolved in methanol and analysed by LC-MS. Positive and negative electrospray ionization (ESI) mass spectra were obtained with a Finnigan MAT TSQ 7000 instrument (electrospray voltage 4.5 kV; heated capillary temperature 220 ^{o}C; sheath

and auxillary gas nitrogen) coupled to a Micro-Tech Ultra-Plus MicroLC system equipped with an RP18-column (5 μm, 100x1 mm, SepServ, Berlin). For all compounds, a gradient system was used that ranged from H_2O:CH_3CN 90:10 (each containing 0.2% (v/v) acetic acid) to 10:90 over 15 min, followed by isocratic elution with a 10:90 mixture of both solvents for 10 min; flow rate 70 μl/min. The collision-induced dissociation mass spectra during an HPLC run were recorded with a collision energy of –20 or -25 eV for positive ions as well as +20 or +25 eV for negative ions, respectively; collision gas: argon, collision pressure: 1.8×10^{-3} Torr.

HPLC analysis of R. tataricum rhizomes

To measure resveratrol in the leaves, petioles and rhizomes of *R. tataricum*, 1 g fresh plant material was frozen in liquid nitrogen and ground to a fine powder in a mortar and pestle. Ten-ml of 80% (v/v) MeOH containing 3% (v/v) HCl was added and the extract was hydrolysed at 100º for 60 min in a sealed tube. After centrifugation (10,000 x g, 10 min) the clarified extract was evaporated, dissolved in water, and extracted with ethylacetate. The organic layer was dried *in vacuo*, the residue dissolved in 50 μl 50% (v/v) methanol and 10 μl were analysed by HPLC. The HPLC system consisted of a Hewlett Packard Series 1100 instrument with a Eurospher-100 RP18 column (5 μm, 250x4mm, Knauer). Chromatographic separation was performed using a solvent sysem of (A) H_2O containing 2% CH_3CN and 0.2% phosphoric acid and (B) CH_3CN containing 2%) H_2O and 0.2% phosphoric acid (B) with a linear gradient of 20-60% B over 30 min Flow rate was 0.6 ml/min with detection at 320 nm. All HPLC solvent percentages are expressed v/v.

Accession numbers

The GenBank accession numbers of the CHS superfamily members used to construct the phylogenetic tree are *Phalaenopsis* BBS (P53416), *Ruta graveolens* ACS (S60241), *R. graveolens* CHS (CAC14059), *Gerbera hybrida* 2-PS (CAA86219), *G. hybrida* CHS (S56699), *Rheum palmatum* BAS (AAK82824), *Rheum tataricum* STS (AF508150), *Hydrangea macrophylla* CTAS (BAA32733), *H. macrophylla* CHS (BAA32732), *Humulus lupulus* VPS (BAA29039), *H. lupulus* CHS (BAB47196), *Vitis vinifera* STS (CAA54221), *V. vinifera* CHS (CAA53583), *Petroselinum crispum* CHS (CAA24779), *Oryza sativa* CHS (CAA61955), *Arabidopsis thaliana* CHS (BAB11121), *Arachis hypogaea* STS (P20178), *Medicago sativa* CHS (P30074), *Pinus densiflora* CHS (BAA94594), *P. densiflora* STS (BAA94593), *Pinus sylvestris* CHS (CAA43166),

P. sylvestris STS (CAA43165), *Psilotum nudum* CHS (BAA87922), *P. nudum* VPS (BAA87923) and *P. nudum* STS (BAA87925).

Acknowledgements

This work was supported by the Thai Royal Golden Jubilee Fund, Bangkok, the Deutsche Akademische Austausch Dienst, Bonn, the Alexander von Humboldt Foundation, Bonn and by Fonds der Chemischen Industrie, Frankfurt. We thank Dagmar Knöfel of the Abteilung Sekundärstoffwechsel, IPB, Halle, for preparing *p*-coumaroyl-CoA and cinnamoyl-CoA as well as Christine Kuhnt, Abteilung Natur- und Wirkstoffchemie, IPB, Halle, for LCMS measurements.

References

Abe, I., Morita, H., Nomura, A., Noguchi, H., 2000. Substrate specificity of chalcone synthase: enzymatic formation of unnatural polyketides by chalcone synthase. J. Am. Chem. Soc. 122, 11242-11243.

Abe, I., Takahashi, Y., Morita, H., Noguchi, H., 2001. Benzalacetone synthase. A novel polyketide synthase that plays a crucial role in the biosynthesis of phenylbutanones in *Rheum palmatum*. Eur. J. Biochem. 268, 3354-9.

Akiyama, T., Shibuya, M., Liu, H. M., Ebizuka, Y., 1999. *p*-Coumaroyltriacetic acid synthase, a new homologue of chalcone synthase, from *Hydrangea macrophylla* var. *thunbergii*. Eur. J. Biochem. 263, 834-9.

Bentley, R., Zwitkowits, P. M., 1967. Biosynthesis of tropolones in *Penicillium stipitatum*. VII. The formation of polyketide lactones and other nontropolone compounds as a result of ethionine inhibition. J. Am. Chem. Soc. 89, 676-80.

Dellaporta, S. L., Wood, J., Hicks, J. B., 1983. A plant DNA minipreparation. Plant Mol. Biol. Rep. 1, 19-21.

Eckermann, S., Schröder, G., Schmidt, J., Strack, D., Edrada, R. A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpeläinen, I., Proksch, P., Teeri, T. H., Schröder, J., 1998. New pathway to polyketides in plants. Nature 396, 387-390.

Ferrer, J. L., Jez, J. M., Bowman, M. E., Dixon, R. A., Noel, J. P., 1999. Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nat. Struct. Biol. 6, 775-84.

Fliegmann, J., Schröder, G., Schanz, S., Britsch, L., Schröder, J., 1992. Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and differential regulation of these and related enzyme activities in stressed plants. Plant Mol. Biol. 18, 489-503.

Foust, C. M., 1992. Rhubarb: The Wondrous Drug. Princeton, Princeton University Press.

Greenhagen, B., Chappell, J., 2001. Molecular scaffolds for chemical wizardry: learning nature's rules for terpene cyclases. Proc. Natl. Acad. Sci. USA 98, 13479-81.

Helariutta, Y., Elomaa, P., Kotilainen, M., Griesbach, R. J., Schröder, J., Teeri, T. H., 1995. Chalcone synthase-like genes active during corolla development are differentially expressed and encode enzymes with different catalytic properties in *Gerbera hybrida* (Asteraceae). Plant Mol. Biol. 28, 47-60.

Jez, J. M., Bowman, M. E., Noel, J. P., 2002. Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity. Proc. Natl. Acad. Sci. USA 99, 5319-24.

Jez, J. M., Noel, J. P., 2000. Mechanism of chalcone synthase. pKa of the catalytic cysteine and the role of the conserved histidine in a plant polyketide synthase. J. Biol. Chem. 275, 39640-6.

Kashiwada, Y., Nonaka, G., Nishioka, I., Nishizawa, M., Yamagishi, T., 1988. Studies on Rhubarb (Rhei Rhizoma). XIV. Isolation and characterization of stilbene glucosides from Chinese rhubarb. Chem. Pharm. Bull. 36, 1545-1549.

Laemmli, U. K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-5.

Matsuda, H., Morikawa, T., Toguchida, I., Park, J. Y., Harima, S., Yoshikawa, M., 2001. Antioxidant constituents from rhubarb: structural requirements of stilbenes for the activity and structures of two new anthraquinone glucosides. Bioorg. Med. Chem. 9, 41-50.

Melchior, F., Kindl, H., 1990. Grapevine stilbene synthase cDNA only slightly differing from chalcone synthase cDNA is expressed in *Escherichia coli* into catalytically active enzyme. FEBS Lett. 268, 17-20.

Morita, H., Noguchi, H., Schröder, J., Abe, I., 2001. Novel polyketides synthesized with a higher plant stilbene synthase. Eur. J. Biochem. 268, 3759-66.

Morita, H., Takahashi, Y., Noguchi, H., Abe, I., 2000. Enzymatic formation of unnatural aromatic polyketides by chalcone synthase. Biochem. Biophys. Res. Commun. 279, 190-5.

Rupprich, N., Kindl, H., 1978. Stilbene synthases and stilbenecarboxylate synthases: enzymatic synthesis of 3,4,5-trihydroxystilbene from *p*-coumaroyl coenzyme A and malonyl coenzyme A**.** Hoppe-Seyler's Z. Physiol. Chem. 359, 165.

Salzman, R. A., Fujita, T., Zhu-Salzman, K., Hasagawa, P. M., Bressan, R. A., 1999. An improved RNA isolation method for plant tissue containing high levels of phenolic compounds or carbohydrates. Plant Mol. Biol. Rep. 17, 11-17.

Samappito, S., Page, J., Schmidt, J., De-Eknamkul, W., Kutchan, T. M., 2002. Molecular characterization of root-specific chalcone synthases from *Cassia alata*. (Accepted).

Schröder, G., Brown, J.W.S., Schröder, J., 1988. Molecular analysis of resveratrol synthase: cDNA, genomic clones and relationship with chalcone synthase. Eur. J. Biochem. 172, 161-169.

Schröder, J., 1999. The chalcone/stilbene synthase-type family of condensing enzymes. Comprehensive Natural Products Chemistry, Vol. 1: Polyketides and Other Secondary Metabolites Including Fatty Acids and Their Derivatives. U. Sankawa. Oxford, Elsevier Science.

Schröder, J., 2000. The family of chalcone synthase-related proteins: functional diversity and evolution. Rec. Adv. Phytochem. 34, 55-89.

Schuz, R., Heller, W., Hahlbrock, K., 1983. Substrate specificity of chalcone synthase from *Petroselinum hortense*. Formation of phloroglucinol derivatives from aliphatic substrates. J. Biol. Chem. 258, 6730-4.

Swiss-Prot Database, 2002. http://www.expasy.ch/sprot/

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., Higgins, D. G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl. Acids Res. 25, 4876-82.

Tropf, S., Lanz, T., Rensing, S. A., Schröder, J. and Schröder, G., 1994. Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. J. Mol. Evol. 38, 610-8.

Yamaguchi, T., Kurosaki, F., Suh, D. Y., Sankawa, U., Nishioka, M., Akiyama, T., Shibuya, M., Ebizuka, Y., 1999. Cross-reaction of chalcone synthase and stilbene synthase overexpressed in *Escherichia coli*. FEBS Lett. 460, 457-61.

Yamazaki, Y., Suh, D. Y., Sitthithaworn, W., Ishiguro, K., Kobayashi, Y., Shibuya, M., Ebizuka, Y., Sankawa, U., 2001. Diverse chalcone synthase superfamily enzymes from the most primitive vascular plant, *Psilotum nudum*. Planta 214, 75-84.

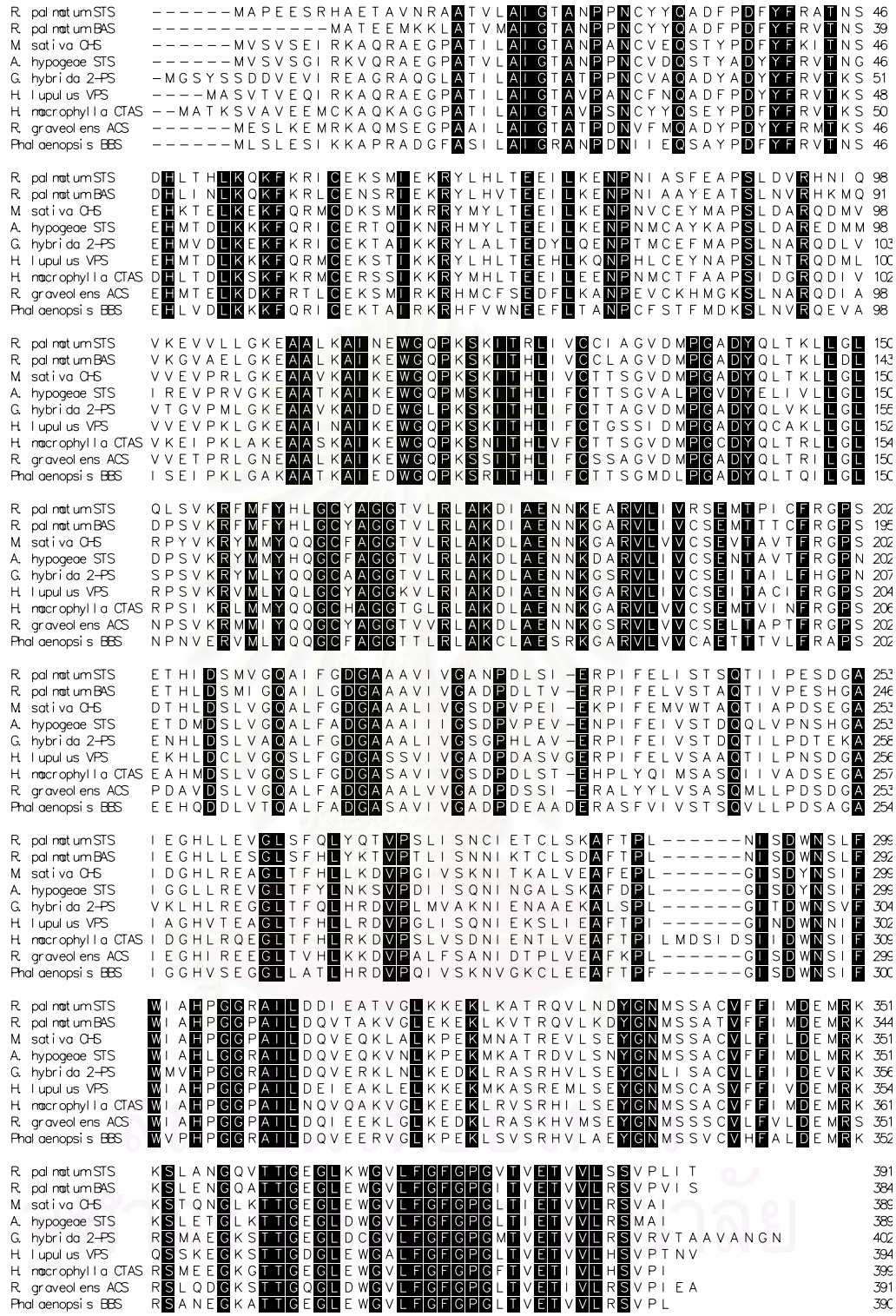


Fig. 1. Alignment of deduced amino acid sequence of *RtSTS* from *Rheum tataricum* with the plant CHS superfamily members *Rheum palmatum* BAS, *Medicago sativa* CHS, *Arachis hypogaea* STS, *Gerbera hybrida* 2-PS, *Humulus lupulus* VPS, *Hydrangea macrophylla* CTAS, *Ruta graveolens* ACS and *Phalaenopsis* BBS. Conserved motifs to which degenerate PCR primers were directed are shown with arrows. Shaded letters indicate conserved amino acid residues. Hyphens denote gaps introduced to maximize similarity. Accession numbers of proteins are found in the Experimental.

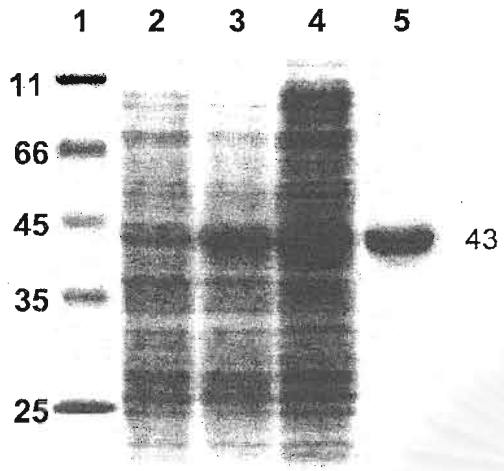


Fig. 2. SDS-PAGE analysis of recombinant RtSTS stained with Coomassie blue. Lane 1, molecular mass markers with masses in kDa indicated; Lane 2, uninduced cells; Lane 3, cells grown for 12 hours at 28° after induction with IPTG; Lane 4, crude cell lysate and Lane 5, RtSTS (2 μg) purified by affinity chromatography with a Talon cobalt resin.

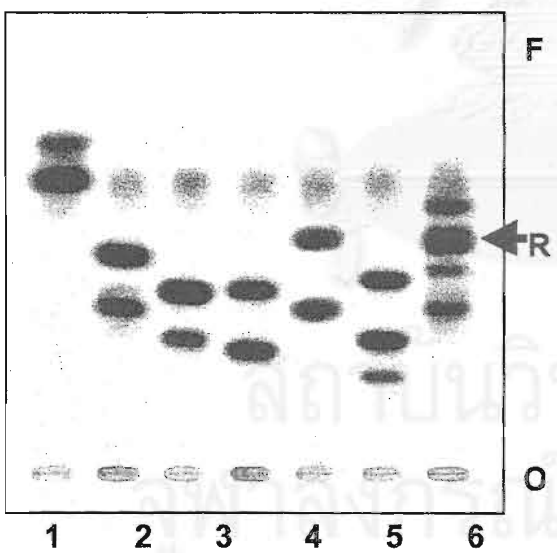


Fig. 3. Reversed-phase TLC analysis of RtSTS reaction products. The reactions were performed with 1.5 μg of purified enzyme, [2-^{14}C]malonyl-CoA and acetyl-CoA (lane 1), *n*-butyryl-CoA (lane 2), isovaleroyl-CoA (lane 3), *n*-hexanoyl-CoA (lane 4), benzoyl-CoA (lane 5), cinnamoyl-CoA (lane 6), *p*-coumaroyl-CoA (lane 7). O= TLC origin, F= solvent front, R= position of resveratrol standard.

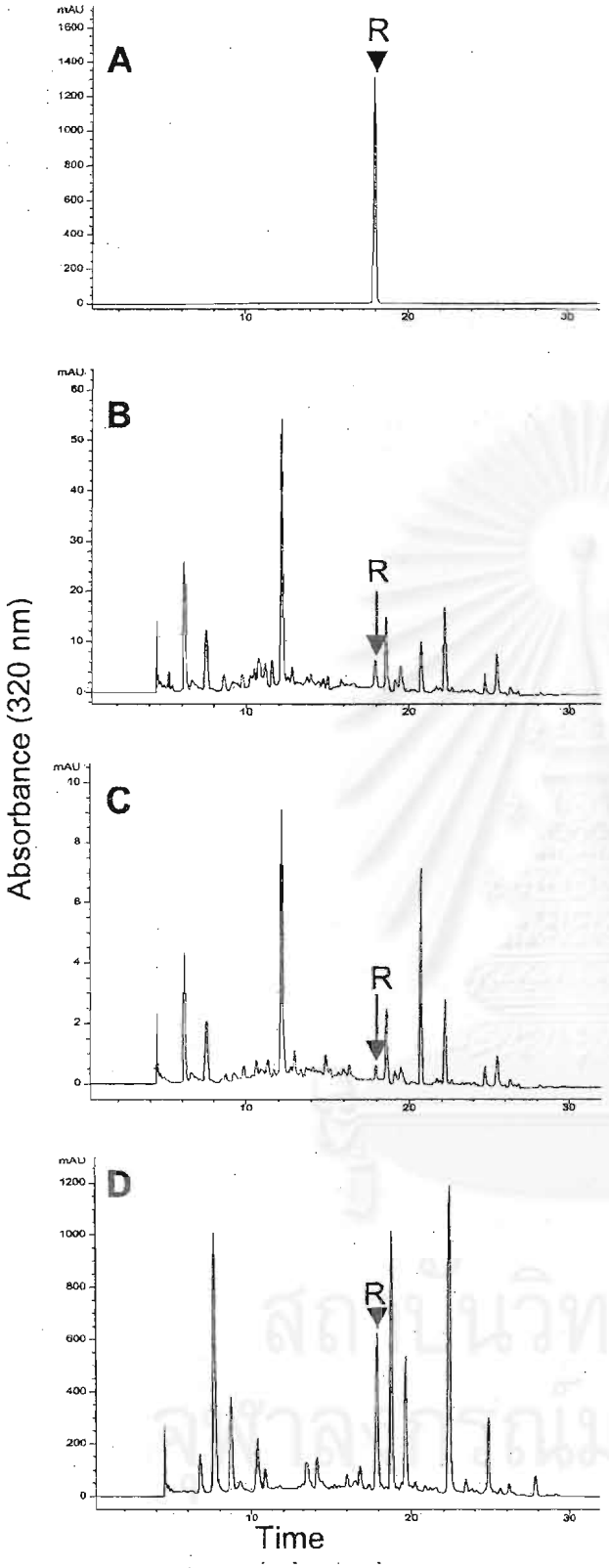


Fig. 4. Reversed-phase HPLC analysis of resveratrol in different tissues of *Rheum tataricum*. Methanolic extracts of each tissue were acid hydrolyzed before analysis to cleave sugar residues. Chromatograms represent (A) resveratrol standard, (B) leaf extract, (C) petiole extract and (D) rhizome extract. Note that the scale of the absorbance (y) axis differs in each of the panels due to the large variation in resveratrol content between the tissues analyzed. Chromatograms were obtained at 320 nm by photodiode array detection.

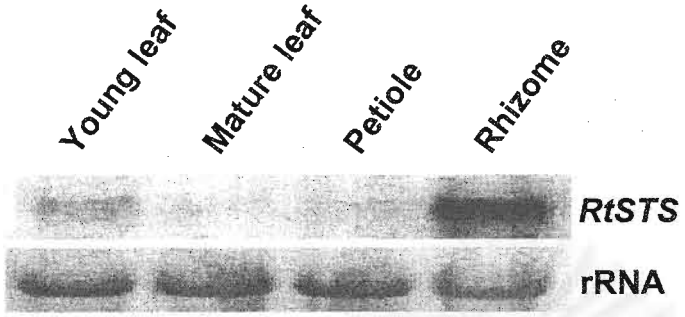


Fig. 5. Northern blot analysis of *RtSTS* expression in different tissues of *Rheum tataricum*. The lower panel shows equal loading of RNA as determined by ethidium bromide-staining of ribosomal RNA.

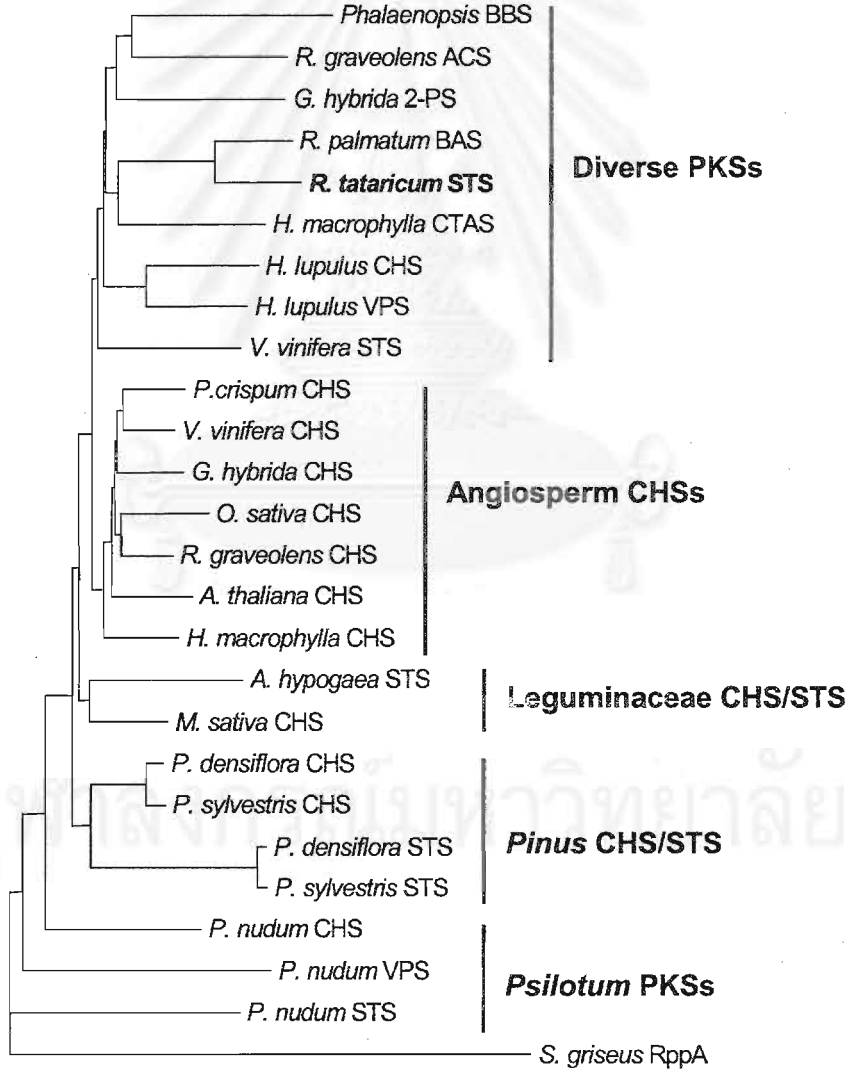
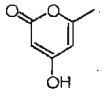
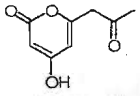
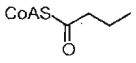
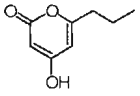
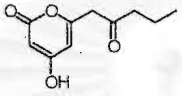
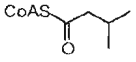
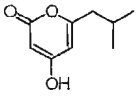
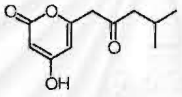
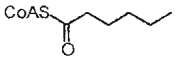
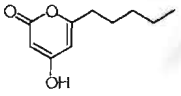
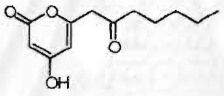
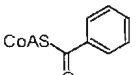
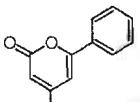
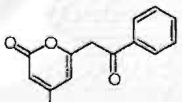
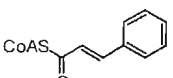
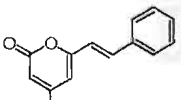
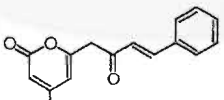
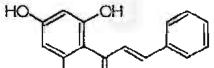
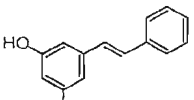
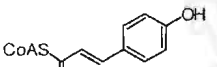
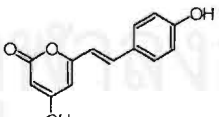
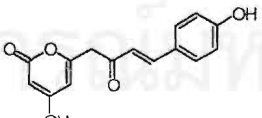
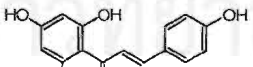
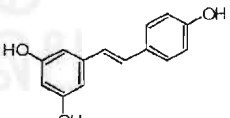


Fig. 6. Phylogenetic comparison of the deduced polypeptide sequence of *RtSTS* with members of the CHS superfamily. The phylogenetic tree was constructed using ClustalX (Thompson et al., 1997). *Streptomyces griseus* RppA was used as an outgroup. Accession numbers of proteins used to construct the phylogeny are found in the Experimental.

Table 1. Enzymatic products of RtSTS with different starter CoA esters

Substrate	Products			
 Acetyl-CoA (**1a**)	 *m/z* 127 $[M+H]^+$ (**1b**)	 *m/z* 169 $[M+H]^+$ (**1c**)		
 *n*-Butyryl-CoA (**2a**)	 *m/z* 155 $[M+H]^+$ (**2b**)	 *m/z* 197 $[M+H]^+$ (**2c**)		
 Isovaleryl-CoA (**3a**)	 *m/z* 169 $[M+H]^+$ (**3b**)	 *m/z* 211 $[M+H]^+$ (**3c**)		
 *n*-Hexanoyl-CoA (**4a**)	 *m/z* 183 $[M+H]^+$ (**4b**)	 *m/z* 225 $[M+H]^+$ (**4c**)		
 Benzoyl-CoA (**5a**)	 *m/z* 189 $[M+H]^+$ (**5b**)	 *m/z* 231 $[M+H]^+$ (**5c**)		
 Cinnamoyl-CoA (**6a**)	 *m/z* 215 $[M+H]^+$ (**6b**)	 *m/z* 257 $[M+H]^+$ (**6c**)	 Pinocembrin chalcone *m/z* 257 $[M+H]^+$ (**6d**)	 Pinosylvin *m/z* 213 $[M+H]^+$ (**6e**)
 *p*-Coumaroyl-CoA (**7a**)	 *m/z* 231 $[M+H]^+$ (**7b**)	 *m/z* 273 $[M+H]^+$ (**7c**)	 Naringenin chalcone *m/z* 273 $[M+H]^+$ (**7d**)	 Resveratrol *m/z* 227 $[M-H]^-$ (**7e**)

VITA

Mr. Supachai Samappito was born on August, 8 1968 in Saraburi, Thailand. He received his Bachelor degree of Science, program in Biology in 1991, from Srinakarinwirot University. After that, he received his Master degree of Science, program in Biotechnology in 1994, from Chulalongkorn University. At present, he is an instructor staff of Department of Biotechonoloy, Faculty of Technology, Mahasarakham University, Thailand.