FEBS Letters 580 (2006) 1411–1416

Artemisia annua L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin

Keat H. Teoh, Devin R. Polichuk, Darwin W. Reed, Goska Nowak, Patrick S. Covello*

Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, Sask., Canada S7N OW9

Received 14 December 2005; revised 18 January 2006; accepted 23 January 2006

Available online 30 January 2006

Edited by Ulf-Ingo Flügge

Abstract **Artemisinin, a sesquiterpene lactone endoperoxide derived from the plant *Artemisia annua*, forms the basis of the most important treatments of malaria in use today. In an effort to elucidate the biosynthesis of artemisinin, an expressed sequence tag approach to identifying the relevant biosynthetic genes was undertaken using isolated glandular trichomes as a source of mRNA. A cDNA clone encoding a cytochrome P450 designated CYP71AV1 was characterized by expression in *Saccharomyces cerevisiae* and shown to catalyze the oxidation of the proposed biosynthetic intermediates amorpha-4,11-diene, artemisinic alcohol and artemisinic aldehyde. The identification of the CYP71AV1 gene should allow for the engineering of semi-synthetic production of artemisinin in appropriate plant or microbial hosts.**

Keywords: Artemisinin; Sesquiterpene; Cytochrome P450; Trichome; *Artemisia annua*

1. Introduction

A little over thirty years ago, as part of a search for new antimalarial drugs, Chinese scientists isolated a sesquiterpene lactone endoperoxide called qinghaosu or artemisinin (see Fig. 1) from *Artemisia annua* [1]. Artemisinin and its derivatives have since become the most important agents in the treatment of malaria, particularly in the form of artemisinin-based combination therapies (ACTs) [2].

As a member of the tribe Anthemideae in the Asteraceae, *A. annua* characteristically produces an essential oil rich in mono- and sesqui-terpenes which is sequestered in 10-celled biseriate glandular trichomes [3,4]. The overall composition of *A. annua* essential oil is dominated by monoterpenes including artemisia ketone, camphor, α-pinene and pinocarvone [4]. The sesquiterpene lactone artemisinin accumulates to levels of 0.01–1% of dry weight [5]. Recently, a reasonably clear picture of artemisinin biosynthesis has emerged as illustrated in Fig. 1 [6]. The identity of amorpha-4,11-diene as a biosynthetic intermediate was established, based on the presence of traces of amorpha-4,11-diene in *A. annua* extracts and the cloning and expression of cDNAs representing amorpha-4,11-diene synthase, a sesquiterpene cyclase [7,8]. Dihydroartemisinic acid is thought to be an intermediate produced by enzymatic oxidation at the C12 to carboxylate and reduction of the double bond at C11–C13 of amorpha-4,11-diene. The biochemical evidence reported recently [6] is consistent with cytochrome P450 (CYP) involvement for the initial hydroxylation of amorpha-4,11-diene and with distinct (soluble) enzyme activity required for subsequent oxidation to carboxylic acid.

In an effort to improve our understanding of the biosynthesis of artemisinin, and of sesquiterpene lactones in general, we undertook the molecular cloning of the enzyme involved in amorpha-4,11-diene oxidation. The expressed sequence tags (ESTs) generated from isolated trichomes of *A. annua* were used as a starting point for the molecular cloning and characterization of CYP71AV1, a multifunctional sesquiterpene oxidase.

2. Materials and methods

2.1. Chemicals and chemical analysis

(−)-α-Gurjunene, (+)-γ-gurjunene, (−)-alloisolongifolene, (+)-ledene and (+)-valencene were obtained from Sigma–Aldrich. Caryophyllene, limonene, α-pinene, β-pinene, pinocarveol and (+)-β-selinene were obtained from the Plant Biotechnology Institute terpene collection. (+)-β-Selinene (from celery seed) was confirmed by ^{1}H NMR and EI^{+}-GC/MS and was found to contain a 4:1 mixture of β- and α-selinene.

Artemisinic acid was isolated from *A. annua* leaf and flower bud material by sonication in dichloromethane. The acid was partitioned into 0.1 N KOH, acidified with HCl, extracted into dichloromethane and then recovered by reverse phase HPLC. The purified acid was used to synthesize artemisinic alcohol, artemisinic aldehyde and amorpha-4,11-diene according to the method described previously [8] with the exception that 1.5 equivalent of diisobutyl lithium aluminum hydride was used to reduce the acid and the resulting alcohol/aldehyde mixture was subsequently resolved. The products were purified by reverse phase HPLC and confirmed by ^{1}H NMR and EI^{+}-GC/MS [6,8].

GC/MS analysis was performed using an Agilent 6890 GC equipped with an auto-injector split 30:1 onto a DB-5ms column (30 m × 0.25 mm i.d., J&W Scientific) which was temperature programmed from 125 to 300 °C at 5 °C/min. The column was connected to a mass selective detector (Agilent 5973) operating under standard EI^{+} conditions (70 eV).

2.2. Plant materials

Artemisia annua L. (seed source: Elixir Farm Botanicals, Brixey, MO, USA) were grown with 16 h/25 °C days and 8 h/20 °C nights to a height of approximately 1.2 m (about 3 months) then transferred to 12 h/25 °C days and 12 h/20 °C nights.

*Corresponding author. Fax: +1 306 975 4839.
E-mail address: Patrick.Covello@nrc-cnrc.gc.ca (P.S. Covello).

Abbreviations: AAFB, full-length flower bud cDNA library; AAGST, full-length glandular trichome cDNA library; CYP, cytochrome P450; DXP, 1-deoxy-D-xylulose-5-phosphate; EST, expressed sequence tag; GSTSUB, glandular-trichome-minus-flower-bud cDNA library

0014-5793/$32.00
doi:10.1016/j.febslet.2006.01.065

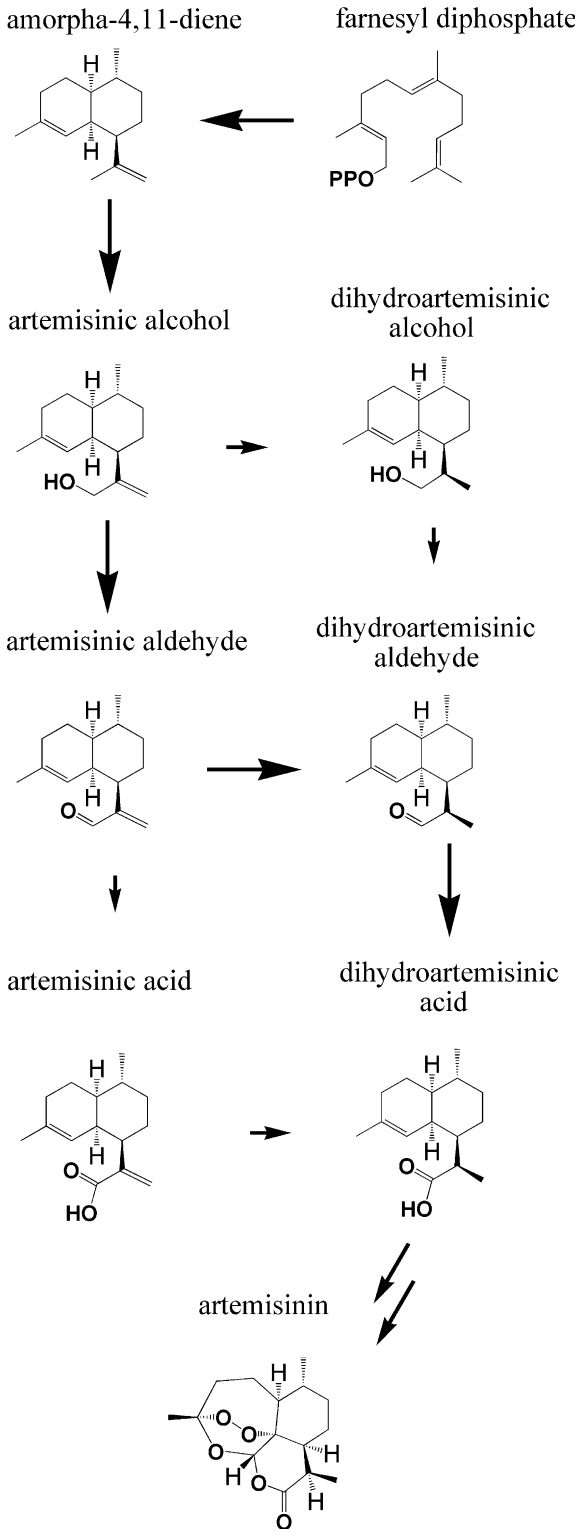

Fig. 1. Proposed pathway of artemisinin biosynthesis (redrawn from [6]).

Glandular trichomes were isolated from flower buds that developed after 19–21 12-h days as described previously with some modifications [9]. Flower buds were chilled in ice-cold water (1 h), then abraded using a cell disrupter (Bead Beater, Biospec Products, Bartlesville, OK, USA). The 350 ml chamber was filled with 20–30 g of plant material, 80–100 g of glass beads (0.5 mm diameter), XAD-4 resin (1 g/g plant material), and isolation buffer (25 mM MOPSO, pH 6.6, 200 mM sorbitol, 10 mM sucrose, 5 mM thiourea, 2 mM dithiothreitol, 5 mM $MgCl_2$, 0.5 mM sodium phosphate, 0.6% (w/v) methylcellulose and 1% (w/v) polyvinylpyrrolidone (M_r 40000)). The flower buds were abraded in an ice-chilled chamber by 3 pulses of operation of 1 min each separated by 1 min. Following abrasion, the contents of the chamber were filtered through 350 μm nylon mesh (Small Parts Inc., Miami Lake, FL, USA). The residual plant material and beads were scraped from the mesh and rinsed twice with rinse buffer (isolation buffer without polyvinylpyrrolidone and methylcellulose) that was also passed through the 350 μm mesh. The 350 μm filtrate was successively passed through a 105, 40 and 30 μm nylon mesh. Glandular trichomes were recovered from the 30 μm mesh.

2.3. cDNA libraries and ESTs

Total RNA used to construct the glandular trichome and flower bud EST libraries was extracted as described previously [10]. cDNA synthesis and construction of two full-length EST libraries (AAGST, trichome; AAFB, flower bud) were carried out with a Creator SMART cDNA Library Construction Kit (Clontech) using Long-Distance PCR following the manufacturer's instructions and using the vector pDNR-LIB. A subtracted cDNA library (GSTSUB) was constructed with a PCR-Select cDNA Subtraction Kit (Clontech) according to the manufacturer's instructions using poly A^+ RNA isolated from glandular trichomes and flower buds (PolyATract mRNA Isolation System, Promega). Double stranded cDNAs of glandular trichomes and flower buds were used as tester and driver, respectively. Sequencing of randomly picked clones was performed on an AB13700 DNA sequencer using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and the M13 reverse primer. DNA sequence traces were interpreted, and vector and low quality sequences were eliminated using PHRED [11] and LUCY [12]. Clustering of the resulting EST dataset was done using STACKPACK [13] and sequence similarity was identified by BLAST [14].

2.4. Expression and characterization of CYP71AV1 in Yeast

The open reading frame (ORF) of a CYP designated CYP71AV1, encoded by the AAGST clone pKT101, was obtained through PCR using gene-specific primers 5′-CACCATGGCACTCTCACTGACCAC-3′ and 5′-CTAGAAACTTGGAACGAGTAACAAC-3′ and Vent polymerase (New England BioLabs, Cambridge, MA, USA). The resulting PCR product was cloned via the Gateway entry vector, pENTR/D-TOPO (Invitrogen) into a Gateway yeast expression vector, pYES-DEST52 (Invitrogen) to generate a yeast expression clone pKT011. The plasmid pKT011 was introduced into the *Saccharomyces cerevisiae* strain WAT11 [15] using S.c. EasyComp Transformation Kit (Invitrogen) to give the yeast strain WAT11/pKT011. Transformants were selected on synthetic complete medium lacking uracil (SC-U) containing 2% raffinose and grown at 30 °C for 24–48 h. Yeast cultures were initially grown in SC-U liquid medium containing 2% raffinose at 30 °C to an OD_{600} of 2–3. Cells were centrifuged and resuspended in induction medium (SC-U containing 1% raffinose and 2% galactose) and grown at 30 °C for 12–16 h. Yeast microsomes were prepared according to Katavic et al. [16] and the microsomal pellet was resuspended in storage buffer containing 50 mM sodium phosphate buffer, pH 7.4, 1 mM EDTA and 20% (v/v) glycerol.

Microsomes of yeast cells transformed with pKT011 were assayed with amorpha-4,11-diene and other isoprenoid substrates, followed by GC/MS analysis. Enzyme reactions were initiated by adding substrate to a final concentration of 100 μM to 500 μl Tris–HCl buffer (25 mM, pH 8.0) containing 1 mM NADPH, 5 mM glucose 6-phosphate, 0.5 U of glucose 6-phosphate dehydrogenase, 300 μg of microsomal proteins, 2 mM DTT, 5 μM FAD, 5 μM FMN, 1 mM ascorbic acid and 10 μg/ml octadecane as an internal standard. Reactions were allowed to proceed for 30 min at 30 °C with shaking and immediately stopped by extracting twice with 500 μl diethyl ether. The ether extracts were pooled, dried down and dissolved in 20 μl of a 1:1 mixture of pyridine and *N*,*O*-bis(trimethylsilyl)acetamide (Sigma) followed by GC/MS analysis.

2.5. Analysis of CYP71AV1 expression in A. annua

For gene expression analysis total RNA was isolated from root tissue (RNeasy Plant RNA Isolation Kit; Qiagen), leaf and flower bud tissue (Trizol, Invitrogen), and isolated trichomes (guanidine chloride