Planta (2001) 212: 460–465

Planta

Amorpha-4,11-diene synthase: cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin

T. Eelco Wallaart[1], Harro J. Bouwmeester[2], Jacques Hille[3], Lucas Poppinga[1], Niels C. A. Maijers[1]

[1]GenoClipp biotechnology B.V., Meditech Center, L.J. Zielstraweg 1, 9713 GX Groningen, The Netherlands
[2]Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands
[3]Department Molecular Biology of Plants, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands

Received: 8 June 2000 / Accepted: 21 August 2000

Abstract. The sesquiterpenoid artemisinin, isolated from the plant *Artemisia annua* L., and its semi-synthetic derivatives are a new and very effective group of antimalarial drugs. A branch point in the biosynthesis of this compound is the cyclisation of the ubiquitous precursor farnesyl diphosphate into the first specific precursor of artemisinin, namely amorpha-4,11-diene. Here we describe the isolation of a cDNA clone encoding amorpha-4,11-diene synthase. The deduced amino acid sequence exhibits the highest identity (50%) with a putative sesquiterpene cyclase of *A. annua*. When expressed in *Escherichia coli*, the recombinant enzyme catalyses the formation of amorpha-4,11-diene from farnesyl diphosphate. Introduction of the gene into tobacco (*Nicotiana tabacum* L.) resulted in the expression of an active enzyme and the accumulation of amorpha-4,11-diene ranging from 0.2 to 1.7 ng per g fresh weight.

Key words: *Artemisia* – Artemisinin – Antimalarial drug – Amorpha-4,11-diene synthase

Introduction

Artemisinin, a new and very potent antimalarial drug, is produced by the plant *Artemisia annua* L. in relatively small amounts ranging from 0.01% to 0.5% on a dry-weight basis (Wallaart et al. 2000). This makes artemisinin an expensive drug. Several studies describe the chemical synthesis of this sesquiterpene, but none of these are an economical alternative for isolation of artemisinin from the plant (Webster and Lehnert 1994; Van Geldre et al. 1997). Therefore, a higher artemisinin concentration in the plant is desirable to make artemisinin available as a relatively cheap antimalaria drug. Knowledge of the biosynthetic pathway of this sesquiterpene may enable us to influence its formation in a direct way, for example by metabolic pathway engineering.

It has been postulated that the regulatory step in the biosynthesis of sesquiterpenes is represented by the cyclisation of the ubiquitous precursor farnesyl diphosphate (FDP) to the highly specific olefinic sesquiterpene skeletons (McCarvey and Croteau 1995). Furthermore, terpene synthases, like most enzymes involved in the biosynthesis of secondary metabolites, are known to occur in very low intracellular concentrations and to catalyse reactions rather slowly (Croteau and Cane 1985). The accumulation of artemisinic and dihydroartemisinic acids (Fig. 1) in the plant, and the absence of any intermediates en route from FDP to these two compounds, indicates that the terpene synthase-catalysed step is indeed rate limiting and may represent a regulatory point. Hence, cloning of the gene encoding amorpha-4,11-diene synthase and overexpression of this gene in *A. annua* seems to be the most rational approach to elevate the artemisinin level. To prove the functional expression of recombinant amorpha-4,11-diene synthase in a plant, a non-amorphadiene-producing plant should preferably be used for transformation. Tobacco is free of any amorpha-4,11-diene synthase and, because of its good transformation characteristics, is also a good candidate.

The cloning of terpene synthases can be achieved using a method based on the homology within these enzymes (Colby et al. 1993). The polymerase chain reaction (PCR) primers can be designed and used in a PCR with cDNA of *A. annua* as a template. The PCR product can subsequently be used as a probe for screening of an *A. annua* cDNA library. Because there are several terpene synthases present in *A. annua* (Bouwmeester et al. 1999), such a PCR may yield a

The cDNA sequence of amorpha-4,11-diene synthase has been submitted to the GeneBank/NCBI data bank with accession number AY006482 and to the European Patent Office under application number 98202854.0 (Wallaart and Bouwmeester 1999)

Abbreviations: DHAA = dihydroartemisinic acid; FDP = farnesyl diphosphate; PCR = polymerase chain reaction

Correspondence to: T. E. Wallaart;
E-mail: mail@genoclipp.com; Fax: +31-50-3176790

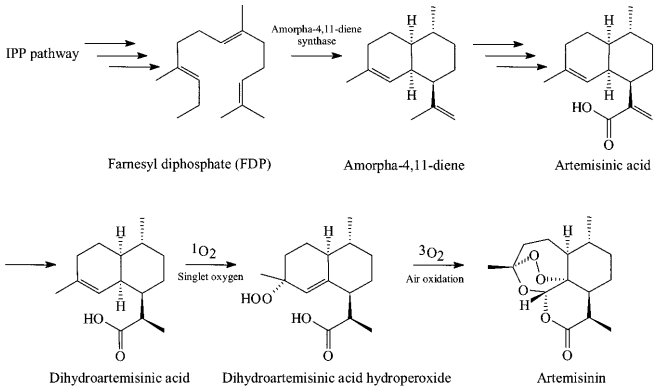

Fig. 1. Proposed biosynthetic pathway of artemisinin in *Artemisia annua*, starting with the amorpha-4,11-diene synthase-catalysed conversion of FDP into amorpha-4,11-diene. The conversion of dihydroartemisinic acid into artemisinin consists of non-enzymatically catalysed reactions starting with the photooxidation (singlet oxygen, $^1O_2$) of dihydroartemisinic acid, which proceeds through a classical ene reaction with rearrangement of the endocyclic double bond, yielding dihydroartemisinic acid hydroperoxide. Ring cleavage, induced by air oxidation (triplet oxygen, $^3O_2$), and re-formation then complete the ring structure with artemisinin as the end-product. *IPP* Isopentenyl pyrophosphate (=isopentenyl diphosphate)

variety of terpene synthase representing PCR products with the same length. This problem can be circumvented by specific enhancing of the expression level of amorpha-4,11-diene synthase. Recently, we reported that dihydroartemisinic acid (DHAA; Fig. 1), a biosynthetic derivative of amorpha-4,11-diene, acts as a scavenger of singlet oxygen ($^1O_2$; Wallaart et al. 1999a, b). The DHAA, and consequently also its precursor amorpha-4,11-diene, are produced in response to the release of reactive oxygen (Wallaart et al. 1999b). Thus, exposure of *A. annua* to stress conditions in which reactive oxygen species (e.g. $^1O_2$) are produced, will preferentially induce the transcription of the gene encoding amorpha-4,11-diene synthase (Knox and Dodge 1985; Wallaart et al. 1999b). Here we describe the isolation of a cDNA encoding amorpha-4,11-diene synthase and its functional expression in *Escherichia coli* and *N. tabacum*.

Materials and methods

Plant material and growth conditions

Artemisia annua L. plants originating from Hanoi, Vietnam (taxonomically verified specimen is deposited at our company; deposit no. Aa002) were grown under sterile conditions on Murashige and Skoog medium (Duchefa, Haarlem, The Netherlands) supplemented with 3% (w/v) sucrose, pH 5.9 and solidified with 0.8% (w/v) micro agar (Duchefa), at 23 °C, >95% moisture, under a 16 h light (ca. 3,000 lx) and 8 h dark cycle. Tobacco (*Nicotiana tabacum* L., cv. Petite Havana SR1) plants were grown on the same medium and under the same conditions (Malaga et al. 1973).

Total RNA isolation and cDNA synthesis

Artemisia annua plants were stressed by placing them at approximately 30% relative humidity (drought stress) and 6,000 lx (photo stress) for 30 min. After a recovery period of 1 h, total RNA was isolated from young leaves (shoots) by using a DNaesy RNA isolation kit according to the manufacturer's instructions (Qiagen). Impurities (such as chromosomal DNA) were removed from the RNA by incubation with DNase I. (RNase free; Gibco BRL). Copy DNA was synthesised using SuperScript II reverse transcriptase (Gibco BRL), according to the manufacturer's instructions.

Polymerase chain reaction-based probe generation

Based on comparison of the deduced amino acid sequences of spearmint [(-)-limonene synthase], tobacco (5-*epi*-aristolochene synthase) and castor bean (casbene synthase) two highly conserved domains were selected (Back et al. 1994) for which a set of consensus inosine-containing, degenerated primers (primers A and B) was synthesised. Primer A was designed based on the highly conserved domain of the proposed active site of the terpene synthases with the characteristic DDXXD sequence motif, in detail: DD(T/I)(I/Y/F)D(A/V)Y(A/G); primer B was designed to the highly conserved domain (D/N)(S/E)(D/E/N)G(K/E)FKE. The sequences of the anti-sense primer A and the sense primer B were respectively: 5′-C RTA IGC RTC RAA IGT RTC RTC-3′ and 5′-GAY GAR AAY GGI AAR TTY AAR GA-3′.The letters R (A or G), and Y (C or T) designate IUB codes for variable nucleotide sites, and I denotes inosine.

The PCR was performed using 0.5 μM of each primer, 0.25 μg of template cDNA, and Super*Taq* DNA polymerase (HT Biotechnology, Cambridge, UK) according to the manufacturer's instructions. Amplification was performed in a thermal cycler (Mastercycler personal; Eppendorf) as follows: 2 min at 94 °C; 40 cycles of 1 min at 94 °C, 2 min at 40 °C, and 1 min 15 s at 72 °C, followed by 10 min at 72 °C. The amplification products of the expected size (ca. 0.5 kb) were made blunt by the DNA polymerase I large (Klenow) fragment, ligated into *Sma*I-digested plasmid pGEM-7Zf(+) (Promega) and used to transform *E. coli* DH5α (Gibco BRL).

Cloning of cDNA

The cDNA was made double stranded by using the RiboClone cDNA synthesis system (Promega). After ligation with *Eco*RI (*Not* I) adapters (Gibco BRL) the double-stranded DNA was ligated into a λ ExCell *Eco*RI/CIAP vector (Amersham-Pharmacia Biotech). For packaging and plating of the cDNA library, the Ready-To-Go Lambda Packaging Kit (Amersham-Pharmacia Biotech) was used. For library screening, 200 ng of the PCR-amplified probe was gel-purified, randomly labelled with α-[^{32}P]dCTP, according to the manufacturer's recommendation (Random Primed DNA Labeling Kit; Boehringer Mannheim) and used to screen the cDNA library according to Sambrook et al. (1989). The positive clones were sequenced with an automatic LI-COR Biotechnology sequencer. Both isoelectric point and molecular weight were calculated from the deduced amino acid sequence using the ExPASy ftp server.

Expression of the amorpha-4,11-diene synthase cDNA in E. coli

The full-length open reading frame of amorpha-4,11-diene synthase, amplified by PCR (under conditions as described above) using primers with *Nco*I and *Bam*HI sites at the ends [forward primer C (5′-GT CGA CAA ACC ATG GCA CTT ACA GAA G-3′), reverse primer D (5′- GGATGGATCC_TCA TAT ACT CAT AGG ATA AAC G-3′)], was inserted into a *Nco*I/*Bam*HI-digested pET 11d expression vector (Stratagene). For expression, this gene construct was used to transform *E. coli* BL21(DE3) (Stratagene). Clones were grown overnight on Luria broth (LB; Duchefa,