Cloning and Sequencing of Hydroxylase Genes Involved in Taxol Biosynthesis

Jun Tu, Ping Zhu*, Ke-di Cheng*, and Chao Meng

Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union Medical College, 1 Xian Nong Tan Street, Beijing 100050, P. R.China. Fax: +86-10-63017757.
E-mail: zhuping@imm.ac.cn; chengkd@imm.ac.cn

* Authors for correspondence and reprint requests

Z. Naturforsch. **59 c**, 561–564 (2004); received February 20/March 19, 2004

Two full-length cDNAs (TCH1 and TCH2) were obtained from a cDNA library of *Taxus chinensis* mainly by the single specific-primer PCR (SSP-PCR) method. Compared with other reported enzymes from *Taxus* species, the deduced amino acid sequences of TCH1 and TCH2 exhibit significant homologies to hydroxylases that are involved in taxol biosynthesis. These findings imply that the two new genes are closely related to the biosynthesis of taxol/taxoids. Data Accession No: AF545833 and AY374652.

Key words: Taxol Biosynthesis, Hydroxylase, SSP-PCR

Introduction

The natural diterpenoid product taxol is an important antimitotic drug for the treatment of a variety of cancers (Baloglu and Kingston, 1999). The low yield of taxol from natural sources and the lack of a commercially viable total synthesis have prompted the development of alternative means of production such as hemisynthesis methods. Increasing applications of taxol in chemotherapy and its high cost have accelerated research towards elucidation of its biosynthetic pathways leading to taxol and related taxoids and definition of the responsible enzymes and genes (Jennewein and Croteau, 2001). The biosynthesis of taxol from plant primary metabolism is a very complex process, including at least 20 steps to construct its tetracyclic skeleton and the addition of the various hydroxyl and acyl functional groups. Over the past several years, several hydroxylase genes responsible for taxol biosynthesis have been cloned and characterized as follows: taxane 10β-hydroxylase (Schoendorf *et al.*, 2001), taxane 13α-hydroxylase (Jennewein *et al.*, 2001) and taxoid 14β-hydroxylase (Jennewein *et al.*, 2003). Many results from taxol biosynthesis pathways show that hydroxylations are important for the bioactivity of taxol. At present, taxane 10β-hydroxylase (CYP725A1), taxane 13α-hydroxylase (CYP725A2) and taxoid 14β-hydroxylase (CYP725A3) from *Taxus* species belong all to cytochrome P450-dependent monooxygenases (P450s) A-groups, which are playing an important role in the biosynthesis of secondary plant products (Bak *et al.*, 1998).

In this work, the SSP-PCR method was used in order to isolate the hydroxylase genes corresponding to taxol biosynthesis. Two new hydroxylase cDNAs were obtained. Compared with other reported enzymes from *Taxus* species, the deduced amino acid sequences of TCH1 and TCH2 exhibit significant homologies to hydroxylases that are involved in taxol biosynthesis.

Material and Methods

Strains and plasmids

E. coli DH5 α (TaKaRa, Dalian, China) was used as the host for plasmid amplification. Plasmid pMD18T (TaKaRa) was used as vectors for cloning.

Taxus chinensis cDNA library construction

To facilitate the isolation of clones encoding hydroxylases involved in taxol/taxoid biosynthesis, a cDNA library was constructed. Total RNA was extracted from *Taxus chinensis* callus cells using the single-step method with acid guanidinium thiocyanate/phenol/chloroform extraction (Chomczynski and Sacchi, 1987). The poly (A)+mRNA was purified by Magnesphere (Promega, Madison, USA) and a cDNA library of *T. chinensis* was constructed by the λ gt10 cDNA library Kit (Promega) using conditions recommended by the manufacturer.

0939–5075/2004/0700–0561 $ 06.00

Screening Taxus chinensis cDNA library

A single-specific primer PCR (SSP-PCR) method was used to clone different hydroxylases. This PCR amplification strategy employs a 5′-forward primer/reverse primer directed against the heme binding region that is the most conserved amino acid sequence in P450s with a 3′-primer complementary to the cDNA λ gt10 vector sequence 5′-AGCAAGTTCAGCCTGGTTAAG-3′(B39) or 5′-CTTATGAGTATTTCTTCCAGGGTA-3′ (B40). For library screening, above primers were used to screen 10 secondary libraries in which 5×10^4 plagues grew in *E. coli* C_{600} *Hf1* (Promega) plate. cDNA amplification was performed using Taq and Pfu DNA polymerase (Invitrogen, Carlsbad, USA) under following conditions: 94 °C for 4 min, 30 cycles at 94 °C for 50 s, 56 °C for 1.5 min and 68 °C for 2.5 min and, finally, 68 °C for 7 min. Size-selected inserts were separated by agarose gel electrophoresis and these products were extracted from the gel, ligated into pMD18T vector, and transformed into *E. coli* DH5 α cells. These insert fragments were fully sequenced. Two full-length P450 cDNA fragments were successfully cloned by two-round SSP-PCR screening.

Nucleotide and protein sequence analysis

Two full-length hydroxylase insert fragment genes and deduced amino acid sequences were compared to those available at the GenBank and were aligned by the BLASTP and BLASTX program. Two nucleotide sequences reported in this work have been deposited at the GenBank database under the accession number AF545833 and AY374652.

Results and Discussion

Several P450-related genes were obtained from *T. chinensis* with the single-specific primer PCR (SSP-PCR) method. The SSP-PCR method was originally described for genome walking with one specific primer (Shyamada and Ames, 1989). It was used with a slight variation to clone different hydroxylases involved in taxol biosynthesis. This PCR amplification strategy employs a 5′-forward primer/reverse primer directed against the heme binding region that is the most conserved amino acid sequence in P450s with a 3′-primer complementary to λ gt10 vector universal arm sequence B39 or B40. Two full-length P450 cDNA fragments were successfully cloned through two-round SSP-PCR screening. In addition, several parts of hydroxylase cDNA sequences were also obtained by one-round SSP-PCR screening (data not shown).

One of the fragments was confirmed to be TCH1 by DNA sequencing. It contains a 1494-bp open-reading frame that encodes 497 amino acids for which a molecular weight of 56,470 Da was calculated. A theoretical isoelectric point (pI) of 9.42 was predicted by DNAMAN analysis software (Lynnon BioSoft, Vaudreuil, Canada). The analysis of the deduced amino acid sequence of TCH1 revealed several typical characteristics of cytochrome P450 monooxygenases including an *N*-terminal hydrophobic helix required for anchoring in the endoplasmic reticulum (ER) membrane, a proline-rich hinge region modulating folding and substrate access to the catalytic pocket, preceded by a cluster of basic residues (the stop-transfer signal) between the hydrophobic amino-terminal membrane anchoring segment and the globular part of the protein (Werck-Reichhart and Feyereisen, 2000). This genes showing three similar consensus sequences of these A-group genes are ES/TLR, PSRF and PFGGGXRXCXG. The deduced sequence of TCH1 resembles that of CYP725A2 (62% identity, 77% similarity) and CYP725A3 (57% identity, 72% similarity). Comparison with that of CYP725A1 from *T. cuspidata* revealed an even closer relationship (92% identity, 93% similarity). The minor differences observed are likely attributable to the species differences, so that the putative protein encoded by TCH1 should have the same function as that of CYP725A1, a taxane 10β-hydroxylase. However, alignment of the putative amino acid sequences of TCH1 with that of defined P450s from other plants, generally, shows a rather low 36% overall homology. The nomenclature of P450 genes is based on the amino acid identity among the proteins they encode; the P450s of the same family generally have > 40% identity of amino acid sequences, those of the subfamily generally have > 55% identity (Nelson, 1999). It is more likely to classify TCH1 into the CYP725A subfamily, so that TCH1 is considered as a possible hydroxylase gene involved in taxol biosynthesis.

The other fragment named TCH2 is a new sequence compared with reported genes. TCH2 cDNA contains a 1482-bp open-reading frame, which encodes 493 amino acids with a calculated molecular weight of 55,238 Da and an estimated pI point of 9.74. Same as TCH1, TCH2 shows

some typical characteristics of P450s and possesses several conserved domains: GTLR, PSRF, PFGGGRRACPG and ASYDTT. In the putative amino acid sequence of TCH2, the arginine in the very conserved E-X-X-R motif has been replaced by glycine. This may imply that the TCH2 protein has some special changes in its second and tertiary structure that might be related with its special function. The deduced amino acid sequence of TCH2 resembles those of other cloned hydroxylases (53–54% identity, 70–71% similarity) involved in taxol biosynthesis. Multiple alignment of the putative amino acid sequences of TCH1 and TCH2 with those amino acid sequences of other cloned hydroxylases, CYP725A1, CYP725A2 and CYP725A3, are shown in Fig. 1.

The application of the SSP-PCR method facilitates cloning of genes encoding isolated proteins based on primers designed with available amino acid sequence. In this work, the single specify primer is designed according to the heme binding region PFGXG because it is the most conserved amino acid sequence in plant P450s. Considering codon degeneration, we can design a set of single specific primers with PFGXG to clone hydroxylase involved in taxol biosynthesis. In this study, we present two new hydroxylase sequence data by a two-round SSP-PCR screening. This method presented in this work also be applicable for cloning of cytochrome P450 families/subfamilies genes from various sources.

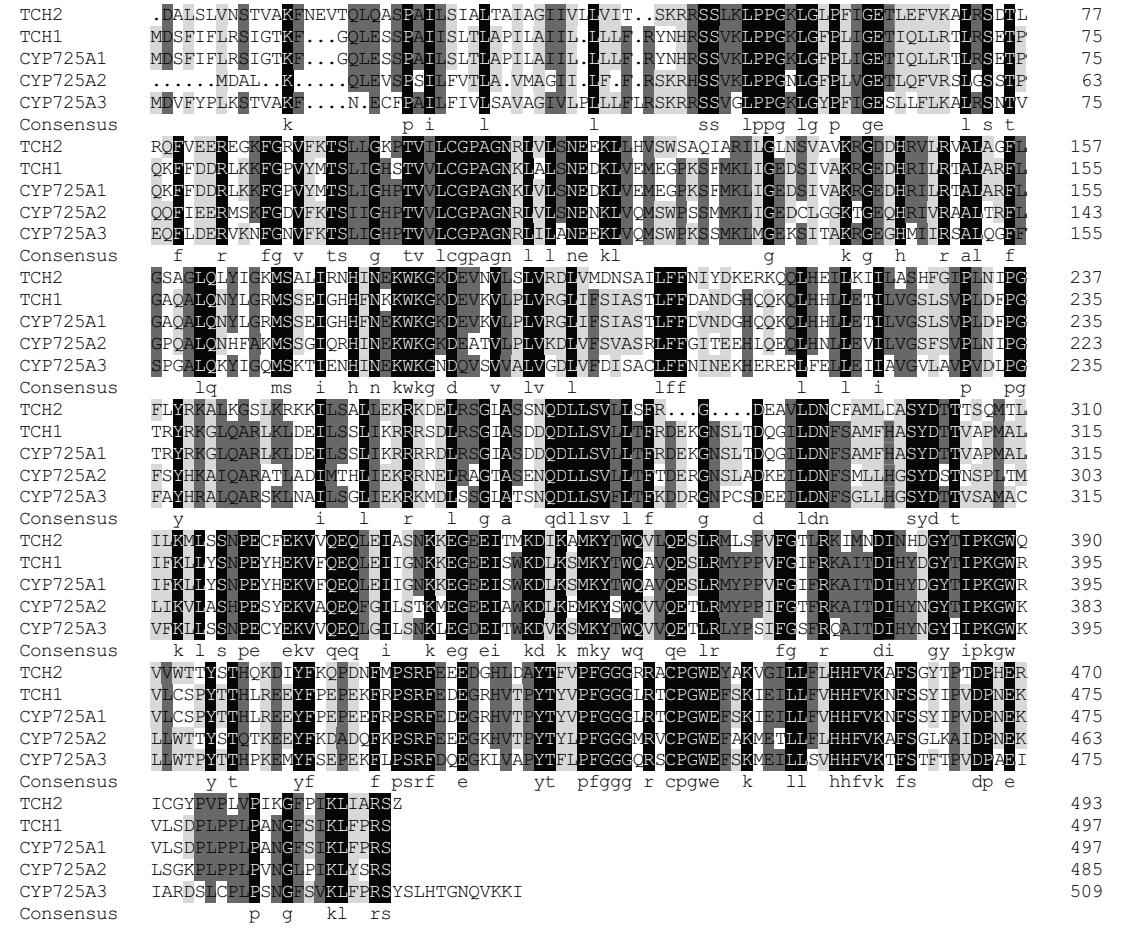

Fig. 1. Multiple alignment of the putative amino acid sequences of TCH1 and TCH2 with those of other cloned hydroxylases involved in taxol biosynthesis. Black boxes denote identical residues while grey boxes denote conservative substitutions. Numbering at the side refers to the entire protein sequences: TCH1 (AF545833); TCH2 (AY374652); CYP725A1 (AF318211); CYP725A2 (AY056019); CYP725A3 (AY188177).

Sequence homology analysis and secondary structure of putative protein products of TCH1 and TCH2 strongly suggest that these putative proteins may function in taxol biosynthesis. The sizes of putative enzyme proteins are in agreement with those of corresponding expression protein bands observed by SDS-PAGE (data not shown). Further studies on functional expression and characterization of these hydroxylase genes will facilitate the detailed understanding of taxol biosynthesis pathways. It will make it possible to provide genes, methods and routes for taxol combinatorial biosynthesis.

Acknowledgements

This work was supported by grants of the National 863 High Technology Programs (grant number 2001AA234021) and the National Science Foundation of China (grant number 20172071).

Bak S., Kahn R. A., Nielsen H. L., Moller B. L., and Halker B. A. (1998), Cloning of three A-type cytochrome P450, CYP71E1, CYP98, and CYP99 from *Sorghum bicolor* (L.) Moench by a PCR approach and identification by expression in *Escherichia coli* of CYP71E1 as a multifunctional cytochrome P450 in the biosynthesis of cyanogenic glucoside dhurrin. Plant Mol. Biol. **36**, 393–405.

Baloglu E. and Kingston D. G. I. (1999), The taxane diterpenoids. J. Nat. Prod. **62**, 1448–1472.

Chanczynski P. and Sacchi P. (1987), Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**, 236–241.

Jennewein S. and Croteau R. (2001), Taxol: biosynthesis, molecular genetics, and biotechnological applications. Appl. Microbiol. Biotech. **57**, 13–19.

Jennewein S., Rithner C. D., William R. M., and Croteau R. (2001), Taxol biosynthesis: Taxane 13α-hydroxylase is a cytochrome P450-dependent monooxygenase. Proc. Natl. Acad. Sci. USA **98**, 13595–13600.

Jennewein S., Rithner C. D., William R. M., and Croteau R. (2003), Taxoid metabolism: Taxoid 14β-hydroxylase is a cytochrome P450-dependent monooxygenase. Arch. Biochem. Biophys. **413**, 262–270.

Nelson D. R. (1999), Cytochrome P450 and the individuality of species. Arch. Biochem. Biophys. **369**, 1–10

Schoendorf A., Rithner C. D., Williams R. M., and Croteau R. (2001) Molecular cloning of a cytochrome P450 taxane 10β-hydroxylase cDNA from *Taxus* and functional expression in yeast. Proc. Natl. Acad. Sci. USA. **98**, 1501–1506.

Shyamada, V. and Ames G. F.-L. (1989), Genome walking by single-specific-primer polymerase chain reaction: SSP-PCR. Gene **84**, 1–8.

Werck-Reichhart D. and Feyereisen R. (2000) Cytochromes P450: a success story. Genome Biol. **1**, reviews 3003.1.