

# Cloning and Heterologous Expression of a New 3'-Hydroxylase Gene from *Lycoris radiata*

Ya-zhong Li<sup>a,b</sup>, Xiao-dan Li<sup>b</sup>, Yu-mei Jiang<sup>b</sup>, Ren Wang<sup>b</sup>, and Bing Xia<sup>b,\*</sup>

<sup>a</sup> College of Life Science, Nanjing Agricultural University, Nanjing 210095, China

<sup>b</sup> Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (Nanjing Botanical Garden, Mem. Sun Yat-sen), Nanjing 210014, China. E-mail: xiabingnb@sina.com

\* Author for correspondence and reprint requests

Z. Naturforsch. **64c**, 138–142 (2009); received March 31/May 30, 2008

A full-length cDNA (LC3'H) was obtained from a cDNA library of *Lycoris radiata* by DOP-PCR (degenerate oligonucleotide primer PCR), 3' race and 5' race methods. Compared with the other reported enzymes from different plants, the deduced amino acid sequence of LC3'H exhibits significant homologies to 3'-hydroxylases that are involved in the caffeic acid biosynthesis. These findings suggest that the new gene is closely related to the biosynthesis of caffeic acid, which is also an important step of the galanthamine biosynthesis in Amaryllidaceae plants.

**Key words:** Galanthamine Biosynthesis, 3'-Hydroxylase, Amaryllidaceae

## Introduction

The natural Amaryllidaceae alkaloid product galanthamine is an important drug for the treatment of Alzheimer's disease. The low yield of galanthamine 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 galanthamine in chemotherapy and its high cost have accelerated research towards the elucidation of its biosynthetic pathways leading to galanthamine and related alkaloids and definition of the responsible enzymes and genes. The biosynthesis of galanthamine from the plant primary metabolism is a very complex process. L-Phenylalanine (L-phe) and L-tyrosine (L-tyr) are considered as primary precursors of galanthamine and other Amaryllidaceae alkaloids. From L-phe to norbelladine, the common precursor of all Amaryllidaceae alkaloids, at least two hydroxylation steps are required (Bastida and Viladomat, 2002). Cinnamate 4-hydroxylase (C4H) introduces the first hydroxy group at the 4-position of the aromatic ring of cinnamic acid. The C4H activity is readily measured in plants, and was one of the first plant enzymes recognized to be a P450. The second hydroxy group is at the 3-position of the aromatic ring of cinnamic acid. The enzyme responsible for 3-hydroxylation of phenolic intermediates (forma-

tion of caffeic acid from *p*-coumaric acid) remained a long time uncharacterized. In recent years, the CYP98 family of cytochrome P450 was identified as the major 3-hydroxylase in the phenylpropanoid pathway (Schoch *et al.*, 2001; Nair *et al.*, 2002; Franke *et al.*, 2002).

Over the past several years, several hydroxylase genes responsible for 3-hydroxylation have been cloned and characterized as follows. CYP98A3 was confirmed as a *meta*-hydroxylase of the phenolic ring functioning as a 5-*O*-(4-coumaroyl) shikimate/quinic-3 hydroxylase (C3H) (Franke *et al.*, 2002). This enzyme is in particular responsible for the final step in the biosynthesis of chlorogenic acid, a widespread phenolic in plants. CYP98A35 and CYP98A36 genes were recently characterized from the coffee tree, which accumulates chlorogenic acid. CYP98A35 can be able to hydroxylate the chlorogenic acid precursor *p*-coumaroyl quinate (Mahesh *et al.*, 2007). In addition, other 4-coumaroyl esters or amides can be *meta*-hydroxylated by CYP98s in species which accumulate specific compounds such as rosmarinic acid or tyramine derivatives (Gang *et al.*, 2002; Morant *et al.*, 2007). All these genes belong to cytochrome P450-dependent mono-oxygenases (P450s) A-groups, which are playing an important role in the biosynthesis of secondary plant products (Bak *et al.*, 1998). However, up to date, there is scant report on CYP98 of Amaryllidaceae plants, which accumulate alkaloids.

In the present work, DOP-PCR (degenerate oligonucleotide primer polymerase chain reaction), 3' race and 5' race methods were used to isolate the hydroxylase gene corresponding to the galanthamine biosynthesis. A novel hydroxylase cDNA (LC3'H) was obtained from the Amaryllidaceae plant *Lycoris radiata*. Compared with the other reported enzymes from several species, the deduced amino acid sequence of LC3'H exhibits significant homologies to 3'-hydroxylases that are involved in the formation of caffeic acid from *p*-coumaric acid, which is also an important step of the galanthamine biosynthesis.

## Materials and Methods

### *Strains and plasmids*

*Escherichia coli* DH5 $\alpha$  (TaKaRa, Dalian, China) was used as the host for plasmid amplification. Plasmid pMD18-T (TaKaRa) was used as vectors for cloning. Plasmid pET29a and BL21(DE3) (Novagen, Madison, WI, USA) were used for heterologous expression

### *Generation of partial cDNA from Lycoris radiata*

Degenerate oligonucleotide primers: 5'-TTG GAT CCN GAR MMN TTY MAR CCN GAR MGR TT-3' (f1) and 5'-AAG GGC CCN GGR CAN DBY STY CTN CCN WYN CCR AA-3' (r1), were designed according to the highly homologous sites on both the amino acid and nucleotide levels in the plant cytochrome P450 mono-oxygenases, corresponding to the conserved domains C and D, respectively (Bozak *et al.*, 1990). PCR-generated partial cDNA encoding hydroxylase from *L. radiata* was produced by PCR using cDNA produced by reverse transcription of total RNA isolated from the leaves of *L. radiata*. DNA amplification was performed under the following conditions: 30 cycles at 94 °C for 30 s; 52 °C for 30 s; 72 °C for 30 s. At the end of 30 cycles, the reaction mixtures were incubated for an additional 5 min at 72 °C prior to cooling to 4 °C. The amplified DNA was then resolved by agarose gel electrophoresis; the bands of approx. the correct size were isolated, ligated into pMD18T vector, and transformed into *E. coli* DH5 $\alpha$  cells. These insert fragments were fully sequenced.

### *Isolation of full-length cDNA of hydroxylase*

3' Race and 5' race methods were used to isolate the full length cDNA of hydroxylase gene. The sequence between primers f1 and r1 was used to generate two gene-specific primers (3GSP1 and 3GSP2) for 3' race (TaKaRa, 3'-full race core set Ver. 2.0). PCR was then performed. A 387-bp fragment was gained. The amplified DNA was isolated, ligated, transformed and sequenced. Using this sequence, two gene-specific primers (5GSP1 and 5GSP2) were designed for 5' race (TaKaRa, 5'-full race kit). A 1311-bp fragment was amplified, cloned and sequenced. Full length of 1524-bp DNA was produced using two gene-specific primers: 5'-CATATGGCTCTCCCTCTCCTCCCTCTAGC-ACTAG3' and 5'-CTCGAGAGGGGTGTACA-TATTGGTAGGCAATCT3'. NdeI and XhoI restriction sites were introduced at 5' and 3', respectively. In order to get high expression, the upstream primer was reformed to 5'-CATATGGCTTTACCTTTATTACCTCTAGCACTAG3' in heterologous expression.

### *Nucleotide and protein sequence analysis*

The full length of the 3'-hydroxylase gene and the deduced amino acid sequence were compared to those available at the GenBank and were aligned by the BLAST program.

### *Heterologous expression of the protein*

Plasmid pET29a was used as vectors for heterologous expression. The LC3'H coding region was released from pMD18-T-LC3'H by digestion with NdeI and XhoI, and cloned into pET29a at the same sites, yielding plasmid pET29a-LC3'H. The correctness of the plasmid was confirmed by sequencing. Then it was introduced into *E. coli* strain BL21(DE3) cells. The bacteria carrying plasmid pET29a-LC3'H were inoculated in 5 ml LB medium (containing 50  $\mu$ g/ml kanamycin) and incubated overnight at 37 °C with shaking at 130 rpm. A drop of 0.5 ml of the overnight culture was inoculated in 50 ml TB medium (500 ml flask) and incubated at 37 °C with shaking at 210 rpm until the OD600 value was approx. 0.8. Isopropyl thiogalactoside (IPTG) (Sigma) was added to 1 mM and the culture was incubated for another 6 h at 30 °C. Cells were collected by centrifugation at 0, 2, 4, 6 h, respectively.

## Results and Discussion

A new P450-related gene was obtained from *L. radiata* by DOP-PCR, 3' race and 5' race methods. In the present work, one of the degenerate oligonucleotide primers was designed according to the heme-binding region PFGXG because it is the most conserved amino acid sequence in plant P450s. Another degenerate oligonucleotide primer was corresponding to PERF, another conserved domain of A-group genes. Considering codon degeneration, we designed a set of degenerate oligonucleotide primers. Using the degenerate oligonucleotide primers, several parts of cDNA sequences were obtained by PCR screening. One of the fragments was confirmed by DNA sequencing to be the 3'-hydroxylase gene. Then, using 3' race and 5' race methods, a full-length cDNA of LC3'H was obtained. It contains a 1515-bp open-reading frame (ORF) that encodes 505 amino acids for which a molecular weight of 57,586 Da was calculated. A theoretical isoelectric point (pI) of 8.73 was predicted by DNAMAN analysis software (Lynnon BioSoft, Vaudreuil, Canada).

The analysis of the deduced amino acid sequence of LC3'H revealed several typical characteristics of cytochrome P450 mono-oxygenases 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 gene showed three similar consensus sequences of these A-group genes: E-X-X-R-R, PERF and PFGGGXRXCXG. The deduced amino acid sequence of LC3'H resembles that of CYP98A3 (65% identity) and CYP98A36 (65% identity). Multiple alignment of the putative amino acid sequences of LC3'H with those amino acid sequences of other cloned hydroxylases (CYP98A3, NP850337; CYP98A36, DQ269127) is shown in Fig. 1. The differences observed are likely attributable to the species differences, so that the putative protein encoded by LC3'H should have the same function as that of CYP98A3, a *p*-coumarate 3-hydroxylase.

However, alignment of the putative amino acid sequences of LC3'H with that of defined P450s from other plants, generally, shows a rather low 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 LC3'H into the CYP98A subfamily, so that LC3'H is considered as a possible 3'-hydroxylase gene involved in the formation of caffeic acid from *p*-coumaric acid, which is also an important step of the galanthamine biosynthesis in Amaryllidaceae plants.

The analysis of sequence homology and secondary structure of the putative protein product of LC3'H strongly suggest that this putative protein may have function in the *p*-coumaric acid *meta*-hydroxylate biosynthesis. The 58 kDa of putative protein is in agreement with those of corresponding expression protein bands observed by SDS-PAGE (Fig. 2). Further studies on functional expression and characterization of this hydroxylase gene will facilitate the detailed understanding of the *p*-coumaric acid *meta*-hydroxylate biosynthesis step, which is also an important step in the galanthamine biosynthesis. It will make it possible to provide genes, methods and routes for the galanthamine combinatorial biosynthesis.

The analysis of sequence homology and secondary structure of the putative protein product of LC3'H strongly suggest that this putative protein may have function in the *p*-coumaric acid *meta*-hydroxylate biosynthesis. The 58 kDa of putative protein is in agreement with those of corresponding expression protein bands observed by SDS-PAGE (Fig. 2). Further studies on functional expression and characterization of this hydroxylase gene will facilitate the detailed understanding of the *p*-coumaric acid *meta*-hydroxylate biosynthesis step, which is also an important step in the galanthamine biosynthesis. It will make it possible to provide genes, methods and routes for the galanthamine combinatorial biosynthesis.

## Acknowledgements

This study was supported by the National Natural Science Foundation of China (grant no. 30700057) and the Opening Fund of Jiangsu Province Key Laboratory for Plant Ex-situ Conservation (KF07001).

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.

Bastida J. and Viladomat F. (2002), Alkaloids of *Narcissus*. In: *Medicinal and Aromatic Plants—Industrial Profiles: The Genus Narcissus* (Hanks G., ed.). Taylor and Francis, London and New York, pp. 141–214.



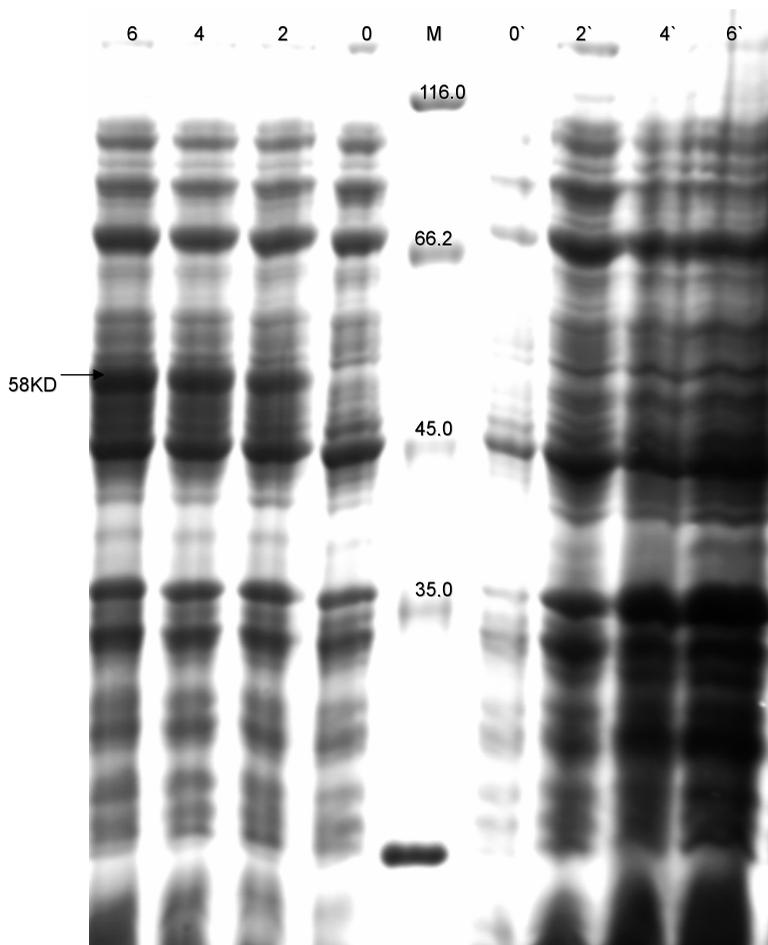


Fig. 2. SDS-PAGE profiles of protein expression 0–6 h after induction of pET29a-LC3'H/BL21(DE3) at 30 °C. Lane 0, pET29a-LC3'H/BL21(DE3) were not induced with IPTG; lanes 2–6, pET29a-LC3'H/BL21(DE3) were induced with IPTG 2, 4, 6 h, respectively; lanes 0'–6': vacancy pET29a/BL21(DE3) were induced with IPTG 0, 2, 4, 6 h; lane M, protein marker.

- Reichhart D. (2007), Catalytic activity, duplication and evolution of the CYP98 cytochrome P450 family in wheat. *Plant Mol. Biol.* **63**, 1–19.
- Nair R. B., Xia Q., Kartha C. J., Kurylo E., Hirji R. N., Datla R., and Selvaraj G. (2002), *Arabidopsis* CYP98A3 mediating aromatic 3-hydroxylation. Developmental regulation of the gene, and expression in yeast. *Plant Physiol.* **130**, 210–220.
- Nelson D. R. (1999), Cytochrome P450 and the individuality of species. *Arch. Biochem. Biophys.* **369**, 1–10.
- Schoch G., Goepfer T. S., Morant M., Hehn A., Meyer D., Ullmann P., and Werck-Reichhart D. (2001), CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J. Biol. Chem.* **276**, 36566–36574.
- Werck-Reichhart D. and Feyereisen R. (2000), Cytochromes P450: A success story. *Genome Biol.* **1**, 3003.1–3003.9.