High Yield Bacterial Expression and Purification of Active Cytochrome P450 ρ-coumarate-3-hydroxylase (C3H), the Arabidopsis Membrane Protein

Hee Jung Yang*, Wanyeon Kim*, Young Ju Yun, Ji Won Yoon, TaeWoo Kwon, HyeSook Youn* and BuHyun Youn*

College of Natural Sciences, Department of Biological Sciences, Pusan National University, Busan 609-735, Korea
1Department of Bioscience & Biotechnology/Institute of Bioscience, Sogang University, Seoul, 143-747, Korea

Received May 7, 2009 / Accepted May 28, 2009

The cytochrome P450s (P450s) metabolizing natural products are among the most versatile biological catalysts known in plants, but knowledge of the structural basis for their broad substrate specificity has been limited. The activity of ρ-coumarate 3-hydroxylase (C3H) is thought to be essential for the biosynthesis of lignin and many other phenylpropanoid pathway products in plants however, all attempts to express and purify the protein corresponding C3H gene have failed. As a result, no conditions suitable for the unambiguous assay of the enzyme are known. The detailed understanding of the mechanism and substrate-specificity of C3H demands a method for the production of active protein on the milligram scale. We have developed a bacterial expression and purification system for the plant C3H, which allows for the quick expression and purification of active wild-type C3H via introduction of combinational mutagenesis. The modified cytochrome P450 C3H (C3Hmut) could be purified in the absence of detergent using immobilized metal affinity chromatography and size exclusion chromatography following extraction from isolated membranes in a high salt buffer and catalytically activated. This method makes the use of isotopic labeling of C3H for NMR studies and X-ray crystallography practical, and is also applicable to other plant cytochrome P450 proteins.

Key words: Cytochrome P450s, C3H, phenylpropanoid, bacterial expression, mutagenesis

Introduction

The phenolic natural products are of pivotal importance in plant biology. Indeed, evolution of their biochemical pathways played a key role in the successful transition of plants to a land base. Furthermore, in addition to their roles in plants, many of these metabolites now find application and/or potential utilization in numerous health-related areas. For this reason, it is essential that each of these enzymatic steps is fully understood at the mechanistic level in order to exploit fully, for the benefit of humanity, the potential of the biotechnological applications that are now emerging [3]. Typically, the amounts of phenylpropanoid metabolites in vascular plants approach 30% of the biomass of which the lignins, lignans and (iso)flavonoids are the most abundant. Of these, the lignins are mainly employed for structural support of load-bearing cell walls and reinforcement of water/nutrient conducting elements. The lignans, stilbenes and isoflavonoids, by contrast, are mainly involved in plant defense due to their profound physiological properties, e.g., as cytotoxic compounds, biocides, antifeedants, antioxidants, and antifungal/antiviral agents, etc. In recent years, there has also been a most remarkable explosion of interest, both scientifically and in the public domain, in medicinally important and health-protecting plant phenolics [5].

Most of the genes encoding the enzymes of the phenylpropanoid pathway have been cloned over the last 20 years by standard biochemical approaches, followed by the isolation of an array of orthologues from various species. To our knowledge, there are currently no crystal structures of any plant cytochrome P450s, in spite of their abundance in the plant kingdom (e.g. Arabidopsis contains at least 272 P450s). However, their conversions are of enormous importance in plant biology (e.g. to the structural lignins, and to the various phenolic defense metabolites), as well as to various health-related biochemical pathways, e.g. to caffeic/chlorogenic acids, taxol, morfinhe, podophyllotoxin, daidzein/genistein and many other metabolic pathways to mention just a few. The known cytochrome P450-dependent
monooxygenases (P450s) in the pathway, p-coumarate 3-hydroxylase (C3H), cinnamate 4-hydroxylase (C4H) and ferulate 5-hydroxylase (F5H) were more difficult targets because the instability, low abundance, and membrane-bound nature of plant P450s makes conventional purification problematic [2]. Despite these difficulties, the gene encoding C4H and F5H were identified using the purification of the enzyme and a T-DNA tagged allele of the Arabidopsis fah1 mutant, respectively [8,14,15,23]. The biosynthesis of many phenylpropanoids requires at least two hydroxylation steps. C4H introduces the first hydroxyl group at the 4-position of the aromatic ring of cinnamic acid. C4H activity is readily measured in plants, and was one of the first plant enzymes recognized to be a P450 [17]. The next hydroxylation occurs at the 3-position of the ring, and p-coumarate 3-hydroxylase (C3H), the enzyme that catalyses this reaction, has not been unambiguously characterized. Over the past 30 years, many researchers have attempted to assay, characterize and purify C3H, and there is considerable disagreement in their published reports concerning the nature of the enzyme [1,6].

Some research has also suggested that the 3-hydroxylation reaction is catalysed by a P450 that uses p-coumaroyl quinate or p-coumaroyl shikimate as its substrate, or by a Zn$^{2+}$-dependent enzyme that converts p-coumaroyl CoA to caffeoyl CoA (Fig. 1) [9,11].

In spite of debates, C3H is probably the most common cytochrome P450 reactions in plant, and certainly the most involved in overall differential carbon allocation in vascular plants. Therefore, their structure determinations are thus arguably some of the most important plant P450 targets. Whereas biochemical studies might require only minute amounts of protein, biophysical and structural studies demand milligram amounts of homogenous protein of exceptional stability and purity. The low yields of soluble and active P450s using bacterial expression makes it difficult to produce sufficient amounts of protein for biophysical and structural studies. Currently, the most often used method for the expression of the active forms of plant P450s in the literature is by insect cell culture or by fusion with other soluble protein containing the modification of N-terminal membrane anchor of P450s [10,12,24]. Whereas insect cell culture can provide milligram amounts of protein, it is very demanding on time and cost, with the generation of a mutant protein typically in the range of 3-4 weeks. Furthermore, homogeneous isotopic labeling in cell culture is limited to 15N and 13C isotopes for NMR analysis, and is prohibitively expensive [22].

Here we present a method for bacterial expression of the Arabidopsis cytochrome P450 p-coumarate 3-hydroxylase (C3H) by introducing of combinational mutagenesis, which allows for the rapid generation, expression, and purification of C3H for structural studies.

Materials and Methods

Construction of expression vectors

The plasmid pET-24(a) (Novagen Inc.) was used as template for the modification of the Arabidopsis thaliana P450 C3H cDNA (CYP84A1) using the polymerase chain reaction (PCR) and for the expression of full-length P450 C3H-modified (C3H<sub>mod</sub>). The C3H<sub>mod</sub> (Fig. 2) differs from the wild-type protein by truncation of the N-terminal transmembrane domain (amino acids 1-22), addition of 7 amino acids in N-terminal (S'-MAKKTSS-3'), mutation of N-terminal residues (R24K, Y25G) and addition of 10 extra amino acids (AAALEE and 6His) from vector at the C-terminus. The primers which were used in mutagenesis are indicated in table 1. The site specific mutagenesis was carried out with a QuikChange XL site-directed mutagenesis kit (Stratagene, Inc.) following the manufacturer's instructions with PCR conditions as follows: initial denaturation at 95°C for 1 min, followed by 18 cycles of 95°C for 30 sec, 60°C for 50 sec and 68°C for 6 min, with 7 min at
68°C and an indefinite hold at 4°C. After PCR completion, the PCR product was treated with Dpn I for 1 hr at 30°C to digest the non-mutated parental DNA template, and was transformed into E. coli XL1-Blue (Stratagen Inc.). After selection on LB plates containing the site directed mutation, a positive clone was confirmed by sequencing to ensure that there were no other mutation(s) in the open reading frame as a result of the PCR.

Heterologous expression, fractionation and purification of C3Hmod in E.coli

E.coli XL1-Blue (Stratagen Inc.) was used as a host for the expression of the C3Hmod enzyme. Cultures (5 ml) were grown overnight in Super Broth [18] supplemented with kanamycin (30 mg/liter). Then, 0.5 liter of Terrific Broth [18] supplemented with kanamycin (30 mg/liter) and contained in a 1-liter flask was inoculated with an aliquot of the overnight culture to yield a density (A600) between 0.05 and 0.1. The culture was incubated at 37°C until A600 reached 0.4 then δ-aminolevulinic acid (ALA) was added to a final concentration of 80 mg/liter. Following an incubation for 15 min at 30°C with gentle shaking (200 rpm), the expression of C3Hmod enzyme was induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside. The cells were harvested 48 hr later by centrifugation at 5,000× g for 10 min at 4°C. The cell pellet from a 0.5-liter culture of E.coli cells was resuspended in 50 ml of 20 mM KPi (pH 7.4), containing 20% (v/v) glycerol and 1 mM PMSF. Lysozyme was added to a final concentration of 0.2 mg/ml, and the suspension was incubated for 30 min on ice. An equal volume of H2O was added, and the mixture was sonicated on ice with an Ultrasonics cell disruptor using three 45-s pulses with 60-s intervals at 80% of maximum power. Cell debris and unbroken cells were removed by centrifugation at 5,000× g for 10 min at 4°C, and the membranes were pelleted by centrifugation at >123,000× g for 90 min at 4°C using Beckman Ti 502 rotor. In the absence of EDTA, almost all of the C3Hmod was found in the membrane fraction. Membranes were resuspended in 100 ml of 10 mM Buffer A (KPi, pH 7.4, containing 20 % glycerol, 0.2 mM PMSF, and 10 mM β-mercaptoethanol). The C3Hmod was extracted from the membranes by combining the suspension with an equal volume of 1M Buffer A to raise the ionic strength. After stirring for 30 min, the suspension was centrifuged as described above to pellet membranes. After centrifugation, the supernatant was incubated with nickel-nitrilotriacetic acid agarose (Qiagen Co.) for 1 hr in an overhead shaker at 4°C. After washing with 20 column volumes of wash buffer (100 mM buffer A containing 100 mM NaCl and 50 mM glycine), the fusion protein was eluted stepwise with Buffer B (50 mM KPi, pH 8.0, 300 mM NaCl, 100-500 mM imidazole). Fractions containing C3Hmod were identified by SDS-PAGE, pooled, and purified to homogeneity by using a Superdex G100 size exclusion column (Pharmacia 2×100 cm) equilibrated with 10 mM Buffer A. The purified C3Hmod were con

Table 1. Primers used for multiple mutagenesis of C3H for construction of C3Hmod

<table>
<thead>
<tr>
<th>Mutation</th>
<th>5' primer</th>
<th>3' primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutation 1</td>
<td>5'-GTCGTACATACATAGCGACAAAAAAGCAGCATCATATCGTA</td>
<td>AGGCAGCGGTGG-3'</td>
</tr>
<tr>
<td>Mutation 2</td>
<td>5'-GGCAGCGATACATAGCGACAAAAAAGCAGCATCATATCGTA</td>
<td>AGGCAGCGGTGG-3'</td>
</tr>
<tr>
<td>Mutation 3</td>
<td>5'-GTCGTACATACATAGCGACAAAAAAGCAGCATCATATCGTA</td>
<td>AGGCAGCGGTGG-3'</td>
</tr>
</tbody>
</table>

Mutation 1: 5'-MSWFLIVAVATIAVAVSVYKLIQR-3' → 5'-MAKKTTSS-3'; Mutation 2: R24→K24; Mutation 3: Y25→G25.
centrated with a final purity of >99% as estimated by SDS-PAGE (Coomassie Blue staining) (Fig. 3), and the concentrations were determined by using the Bradford method. C3Hmod was concentrated to >0.60 mM by using a centrifugal device.

Dynamic Light Scattering

The radius and molecular weight of C3Hmod were estimated using a DynaPro-Titan (Wyatt Technology) instrument at 25°C. Purified C3Hmod (2 mg/mL) was freshly prepared buffer was filtered through polyvinylidene difluoride filter (0.2 mm, Millipore). Scattering data were acquired through accumulation (5 times) of 10 scans with 10 s/scan, with the laser intensity set to a range of 50-60% (30-36 mW). The corresponding molecular weight and radius were calculated using the software package DYNAMICS V6 supplied with the instrument.

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) [27] was performed at 25°C using a VP-ITC instrument (MicroCal Inc.). The pure substrate p-coumaroyl shikimate was first individually dissolved in MeOH (0.5 cm³, 48 mM), with buffer C (20 mM Tris, pH 8.0, 1 mM DTT, 1 mM EDTA, and 10% glycerol) next added such that the MeOH content was 1% and the final concentration of p-coumaroyl shikimate was 0.48 mM. NADPH (0.48 mM) was dissolved in buffer C containing 1% MeOH. Purified C3Hmod was dialyzed extensively against buffer C in an Amicon stirred cell at 4°C, after which MeOH was added to give a 1% final concentration. The final enzyme concentration (30 µM) was determined by the Bradford method using BSA as standard. Both the enzyme and substrate/cofactor solutions were individually degassed in vacuo for 5 min before ITC. Titration experiments were performed as follows: aliquots (10 mm³) of the ligand solution (p-coumaroyl shikimate or NADPH) were injected into the reaction cell containing C3Hmod (2 cm³) with stirring set at 300 rpm. Twenty nine injections for each assay condition were performed with an equilibration interval of 400 s between each. Binary complex titrations were individually carried out by adding p-coumaroyl shikimate or NADPH solution to the C3Hmod solution in the reaction cell. Ternary complex titrations were carried out by mixing C3Hmod and NADPH in a 1:1 ratio (30 µM each) before adding the solution to the reaction cell, with titrations individually performed using solutions of pure p-coumaroyl shikimate. Heats of dilution were determined by titration of each ligand individually in Buffer C. The experimental data were fitted to an n-equivalent binding site model using the nonlinear least-squares regression from the Origin software package (OriginLab Corp Inc.). Since the number of binding sites, n, converged to values between 0.9 and 1.1 in the initial regression analyses, a single-site binding model (1.0) was used for the final analyses. This yielded the following thermodynamic parameters: binding constant (Kb), binding enthalpy (ΔH) and the entropy change (ΔS), respectively.

Results

Expression and purification of C3Hmod

When we initiated a project to establish a bacterial expression system for the C3H, plant cytochrome P450 enzyme, we found that the several mammalian P450s were successfully overexpressed, purified and used for structural studies. Surprisingly, C3H shows highest sequence similarity to the human cytochrome P450, such as P450-2C8 and -2C9, rather than to fungal and bacterial P450 with similarities ranging 58-62% [13]. From the high sequence similarity and characteristics of mammalian P450s and plant C3H, we thought that the multiple mutagenesis method which was broadly applied to mammalian P450s can be the good strategy for plant C3H overexpression. Consequently, C3H is expressed at high levels in E. coli as a modified protein, C3Hmod, where the native N-terminal sequence is replaced with that employed for the heterologous expression of mammalian P450s in E. coli such as P450-2A6, -2B4, -2C5, -2C9, and -2D6 [14]. The expression and purification of C3Hmod in bacteria is shown in Fig. 3A. The results indicate that the expression and purification of C3Hmod is similar to that of the mammalian P450s. From these observations, we chose E. coli for the expression and purification of our enzyme of interest.
-2C8 and -2C9 [20,21,25,26,28] (Fig. 2). These changes (deletion, addition and site-directed mutagenesis) to the N-terminus of CHI (Fig. 2) retain the hydrophobic character of the putative membrane-spanning segment while incorporating codons that facilitate expression in E. coli [16]. In the present study, a hydrophobic segment in the N-terminus of CHI that is thought to function as a membrane-spanning domain was removed to generate CHI_{mod} with the aim of preventing the incorporation of the cytochrome into E. coli membranes following its expression. Additionally, six histidines at the C-terminus were incorporated from expression vector to permit the isolation of protein from high salt extracts by affinity chromatography on an immobilized nickel ion affinity resin. This modification did not significantly alter the distribution of enzyme and kinetic properties which was shown in ITC experiments. In expression of CHI_{mod}, supplementation of the growth medium with the heme precursor ALA increased the total expression level approximately two-fold as reported in P450-2C3 [24].

CHI_{mod} could be purified in an active form without using any detergents. However, it was found that when the purified CHI_{mod} was highly concentrated (>200 μM) and stored over several days at 4°C, a slow time-dependent aggregation occurred leading to visible precipitation. If octylglucoside (12%), sodium cholate (0.5%), or Chaps (0.5%) was included in one of the wash buffers used during the initial purification step when the enzyme was bound to the Ni-NTA resin, the problem with aggregation in concentrated solutions of purified CHI_{mod} was greatly diminished. Additional purification step employed chromatography of protein on Superdex G100 size exclusion column in detergent-free buffers. The purified protein preparations appeared to be nearly homogeneous as analyzed by SDS-PAGE (Fig. 3), and they exhibited specific contents of 6 mg/ml protein.

**Multimeric status of CHI_{mod}**

CHI_{mod} was studied by light scattering to determine its tendency to form oligomer in different salt concentration buffer. Dynamic light scattering experiments revealed single peak and an estimated relative mass of 140,000 which was close to that expected for a dimer nature of CHI_{mod} in both buffer solutions at 2 mg/ml (Fig. 4). In the presence of zwitterionic detergent Chaps at a concentration of 0.5% (w/v), CHI_{mod} eluted at an estimated mass corresponding to a monomer (M_{r} 68,000) (Fig. 4).

![Molecular mass of CHI_{mod} in solution. Elution profile of CHI_{mod} monitored by Dynamic light scattering (2 mg/ml). The calculated molecular radius and molecular weight are 4.37 nm and 140 KDa (CHI_{mod} in low salt buffer, solid line), 4.39 nm and 139 KDa (CHI_{mod} in high salt buffer, bold dotted-line), respectively, indicating its dimeric nature. The 0.5% Chaps-treated CHI_{mod} showed monomeric nature (3.85 nm and 68 KDa, respectively, thin dotted-line). The inset shows an SDS-PAGE gel of the purified CHI_{mod} protein.](http://example.com)

**Catalytic properties of CHI_{mod}**

Isothermal titration calorimetry (ITC) was next employed to study CHI_{mod} in terms of its NADPH cofactor-binding characteristics and differential ligand affinity between the substrate p-coumaroyl shikimate only and NADPH/ p-coumaroyl shikimate complex. In every case, the calorimetric data revealed that heat was released when pure substrate p-coumaroyl shikimate and cofactor were individually associated with CHI_{mod}, indicating that those interactions had significant enthalpic contributions in binding (Table 2). It also revealed slightly unfavorable entropic contributions for each, possibly indicating that the enzyme was slightly stabilized upon binding. This effect is especially noticeable for NADH, as possibly contributed by the significant reduction of B-values of the binding pocket upon formation of the cofactor-substrate binary complex [29]. Isothermal titration calorimetry shows that the CHI_{mod} binds substrates with a stoichiometry of 1:1, indicating that the protein concentration was determined reasonably accurately and that the protein is properly folded. As shown in Fig. 5 and Table 2, substrate p-coumaroyl shikimate showed comparable binding affinities to CHI_{mod} in the absence of NADPH. However, when the latter was present, there was about 35 fold decrease in \( K_d \) values of the p-coumaroyl shikimate (\( K_d =9.8 \) μM) over that of same substrates without NADPH (\( K_d =35.1 \) μM).
Table 2. C3Hmod binding parameters determined by isothermal titration calorimetry

<table>
<thead>
<tr>
<th>Compound(s)</th>
<th>$K_d$/uM</th>
<th>$ΔH$/kcal mol$^{-1}$</th>
<th>$ΔS$/cal mol$^{-1}$ degree$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$-coumaroyl shikimate</td>
<td>35.1±1.4</td>
<td>224±1.0</td>
<td>-63.2</td>
</tr>
<tr>
<td>NADPH</td>
<td>20.6±0.4</td>
<td>-30.7±0.7</td>
<td>-80.7</td>
</tr>
<tr>
<td>$p$-coumaroyl shikimate / NADPH</td>
<td>9.8±0.3</td>
<td>-14.7±0.2</td>
<td>-179</td>
</tr>
</tbody>
</table>

Discussion

Removal of the putative membrane-spanning domain from the N-terminus of C3H successfully prevented the integration of the modified proteins into E.coli membrane as judged by the relatively complete and facile release of each protein from the membrane by increase of each protein from the membranes by increasing the ionic strength of the buffer. This permitted the subsequent isolation of the modified C3Hmod without the necessity of detergent and C3Hmod exhibited catalytic properties. In addition, C3Hmod exists predominantly as a dimer when characterized by light scattering method. These results suggest that the hydrophobicity of the N-terminus contributes to the formation of the larger oligomer as commonly shown in other mammalian P450s [19,24]. The relatively facile conditions that released C3Hmod from E.coli membranes indicate that the association of this modified protein with E.coli membranes is peripheral. The observation that specific proteolysis at an engineered site near N-terminus of the yeast microsomal P450 52A3 also releases the catalytic domain from the membrane provides additional evidence that the N-terminus constitutes the only membrane-spanning portion [19].

Isothermal titration calorimetry (ITC) data showed catalytic ability of C3Hmod is dependent on the NADPH binding and formation of binary complex in $p$-coumaroyl shikimate (Fig. 5). This is indicative of a significant involvement of NADPH to facilitate binding interactions of the substrate, even though only substrates are processed catalytically. Therefore, it is reasonable to speculate that NADPH binds first and NADP$^+$ leaves last, indicating that the kinetic mechanism of C3Hmod could be an ordered Bi-Bi or Theorell-Chance mechanism, as in the cases of 12-HD/PGR and AOR [47]. The dissociation of the NADP$^+$ could be rate-limiting.

The changes incorporated into C3H provide a means for isolating large quantities of highly purified protein in a relatively uniform and low aggregation state that can easily be concentrated to greater than 5 mg protein/ml without the use of detergents. It is hoped that the benefits derived from this modification will facilitate their structural studies such as NMR and X-ray crystallography. The striking effect of multiple mutageneses on the bacterial expression of C3H may be adaptable to other plant membrane proteins in P450 family.

Acknowledgment

This work was supported for two years by Pusan National University Research Grant.

References

3. Davin, L. B., M. Jourdes, A. M. Patten, K. W. Kim, D. G. Vassao, and N. G. Lewis. 2008. Dissection of lignin macromolecular configuration and assembly: comparison to re-
초록: 대장균 시스템을 이용한 Arabidopsis 막 단백질 cytochrome P450 p-coumarate-3-hydroxylase (C3H) 활성형의 과발현 및 분리정제

양희정† · 김완연† · 윤지휘 · 권태우 · 윤혜숙† · 윤부현* (부산대학교 자연과학대학 생명과학과, 1세종대학교 생명과학대학 생명공학부 생명공학과)

다양한 천연물의 합성대사에 관여하는 식물 cytochrome P450 (P450s)은 그 기능적 다양성에도 불구하고, 이들 효소의 정량과 기질 특이성을 설명해 줄 수 있는 구조분석에 대해서는 충분한 연구가 이루어지지 못하고 있는 실정이다. 식물 p-coumarate 3-hydroxylase (C3H)에 의해 매개되는 효소 반응은 lignin과 다양한 phenylpropanoid 부산물들의 생합성에 매우 중요한 것으로 여겨지지만, 단 단백질인 C3H의 발현 및 정제가 효과적으로 이루어지지 못하여, 활성을 측정하기 위한 분석법이 제계화 되지 못하고 있다. C3H의 작용기작과 기질독이성에 대해 폭넓은 이해를 위한 구조분석의 선형단계는 활성을 갖는 C3H를 밀리그램 단위로 분리, 정제하는 실험적 방법을 확립하는 것이라 할 수 있다. 이를 위해, 본 연구에서는 다양한 돌연변이 방법을 도입하여 식물 막단백질 C3H를 대장균 시스템에서 효과적으로 발현 및 정제할 수 있는 시스템을 사용하였다. 변형된 cytochrome P450 C3H (C3Hmod)을 세 포막으로부터 고농도의 염화춘용액을 이용하여 개질활성화 없이 추출하였으며, 2단계 chromatography를 통해 활성을 유지한 상태로 분리할 수 있었다. 이러한 실험적 기법은 NMR 및 X-ray crystallography와 같은 구조분석을 통한 C3H의 효과적인 분석에 적용될 수 있을 것이며, 또한 다른 식물 cytochrome P450 단백질의 효과적인 분석에도 적용 될 수 있을 것이다.