Bull. Univ. of Tokyo For, 128, 121 − 137 (2013)

Original Article

A phenylalanine ammonia-lyase gene (*ErPAL1*) from *Eucalyptus robusta*: molecular cloning, expression and characterization.

Running title: Molecular cloning of *ErPAL1* from cultured eucalyptus cells

Shinya AKIMOTO*, Takahisa ISHINO*, Tamami TERADA*, Masahiro SAMEJIMA*, Shigehiro KAMODA**

Introduction

Phenylpropanoid pathway is one of the important secondary metabolism pathways and synthesizes a large number of biologically important secondary metabolites such as caffeic acid, flavonoids, stilbennoids, lignin, coumarins, salicylic acid and, chlorogenic acid (DIXON *et al*., 2002; JOOS and HAHLBROCK, 1992; LIU *et al*., 2006; PELLEGRINI *et al*., 1994). These compounds engage in and play important roles of the plant development, mechanical support and disease resistance (BARBER and MITCHELL, 1997; CHEN *et al*., 2007; HARAKAVA, 2005). Furthermore, salicylic acid, isoflavonoids, stilbenoids, and chlorogenic acids can act as signaling molecules or antagonistic ingredients (CHEN *et al*., 2007; DIXON and PAIVA, 1995).

Phenylalanine ammonia-lyase (PAL, EC 4.3.1.5) was first described in 1961 by Koukol and Conn (MACDONALD and D'CUNHA, 2007). This enzyme catalyzes the non-oxidative deamination of phenylalanine into *trans*-cinnamic acid and ammonia, which is the first step in the biosynthesis of phenolic compounds (HAHLBROCK and SCHEEL, 1989). Since PAL plays an important role in synthesizing various phenolics, it is one of the most extensively studied enzymes in plants (DIXON and PAIVA, 1995). In most plant species, such as *Arabidopsis thaliana* (WANNER *et al*., 1995), tobacco (*Nicotiana tabacum*) (REICHERT *et al*., 2009), and parsley (*Petroselinum crispum*) (APPERT *et al*., 1994), PAL is encoded by a multi-gene family.

Eucalyptus species being native to Australia and the adjacent areas is one of the most widely planted hardwood trees as its estimated plantation area covers approximately 20 million hectares worldwide according to GIT Forestry (http://www.git-forestry.com, September 8, 2008). The genus *Eucalyptus* comprises more than 700 species and hybrids and some of them have a key role as a source of raw material pulp. In the pulp production, lignin must be removed by chemical treatments, which is a costly process both to the mill and the environment (KAWAOKA *et al*., 2000). Lignin is derived from dehydrogenative polymerization of monolignols synthesized through the phenylpropanoid pathway (HIGUCHI, 1985).Therefore studying eucalyptus PAL, key enzyme of the pathway, is helpful in breeding the plant for pulp. There are a great variety of

* Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo
** Arboricultural Research Institute, The University of Tokyo Forests, Graduate School of Agricultural and Life Sciences, The University of Tokyo

phenolic extracts, such as ellagic and gallic acid derivatives, catechin, flavonoids and quercetin, produced by going through the phenylpropanoid pathway in *Eucalyptus* (CADAHÍA *et al*., 1997; CONDE *et al*., 1995; FECHTAL and RIEDL, 1991; VÁZQUEZ *et al*., 2012). In our previous report, it turned out that *Eucalyptus* had a quantity of 15% phenolics per dry weight in the leaf which has the greatest amount of phenolics (AKIMOTO *et al*., 2012). It was indicated that phenolic compounds performed crucial functions in eucalyputs. For these reasons, we have focused on its PAL.

The entire genome of *Eucalyptus camaldulensis* was published (HIRAKAWA *et al*., 2011), however, there is no investigation about characterization of eucalyptus PAL. In this report, to gain the knowledge about eucalyptus PAL, a full length PAL gene (*ErPAL1*) and its 5'-flanking region were cloned from *Eucalyptus robusta*. We examined its biochemical characterization of recombinant ErPAL1.

Materials and methods

Plant materials

Suspension cultured eucalyptus (*Eucalyptus robusta*) cells (YAMAGUCHI *et al*., 1986) were propagated in 100 ml of Linsmaier and Skoog (LINSMAIER and SKOOG, 1965) medium (pH 5.8-6.0) containing 100 mg l^{-1} of inositol, 1mg l^{-1} of thiamine hydrochloride, 3% sucrose, 0.5 mg l^{-1} of 2,4-dichlorophenoxyacetic acid (2,4-D), and 0.4 mg l^{-1} of kinetin in 500 ml Erlenmeyer flasks at 26.5° C on a rotary shaker in the dark. Fourteen-day-old cultured cells were transferred into fresh medium for subculture at inoculum size 7 mg ml^{-1} of fresh weight (TERADA *et al*., 1993). Dehydrated with dry filter paper, the fourteen-day-old cells were frozen in liquid nitrogen and stored at -80° C until DNA and RNA preparation.

Extraction of DNA and RNA, synthesis of cDNA

Genomic DNA was extracted from the fourteen-day-old cells using DNeasy Plant Mini Kit (QIAGEN). Total RNA was isolated from same cells with RNeasy Plant Mini Kit (QIAGEN). We performed both experiments according to the protocol suggested by the manufacturer. The quality and the concentration of genomic DNA and RNA were checked by 0.75% agarose gel electrophoresis and spectrometer analysis.

Total RNA was subject to extraction of mRNA using PolyATract mRNA Isolation System Ⅲ & Ⅳ (Promega) according to the protocol. Subsequently, to synthesize cDNA from mRNA, cDNA Synthesis Kit (M-MLV Version) (TaKaRa) was used with oligo $(dT)_{18}$ primer as directed.

Cloning of the full-length *ErPAL1* and its 5'-flanking region

In our previous study, an 805 bp *ErPAL1* fragment located near the 3' end of the coding region was obtained. Alignment analysis showed that the sequence of the fragment shared 97% identities with the partial sequence of *Eucalyptus globulus PAL* (GenBank accession No.AF167487) (AKIMOTO *et al*., 2012).

In order to obtain the full-length *ErPAL1* and its 5'-flanking region, GenomeWalker™

Universal Kit (Clontech) was used. The genomic DNA was completely digested with two blunt-end restriction enzymes (*Dra*I and *Hpa*I), DNA fragments were ligated to the GenomeWalker adaptor provided in the kit.

Cloning of the upstream and downstream fragments of the known 805 bp sequence consisted of four and one individual PCR amplifications, respectively. Each amplifications was composed by two PCR reactions. First PCR reaction used the outer adaptor primer AP1 provided in the kit and an outer gene-specific primer GSP1 which was designed from the known sequences. The reaction was performed in the PCR Thermal cycler Dice (TaKaRa) for 2 min at 94° C, then for 7 cycles with 25 sec at 94° C, 3 min at 72° C and then for 32 cycles with 25 sec at 94° C, 3 min at 67 ° C. After the final cycle, the amplification was extended for 7 min at 67° C. After 30-fold dilution, the first PCR mixture was used as a template for nested PCR with the nested adaptor primer AP2 provided in the kit and a nested gene-specific primer GSP2 designed from the known sequences. Reaction was performed as follows: 2 min at 94° C, then for 5 cycles with 25 sec at 94° C, 3 min at 72° C and then for 20 cycles with 25 sec at 94° C, 3 min at 67° C. After the final cycle, the amplification was extended for 7 min at 67° C. The PCR products were purified with SUPREC^{TM}-PCR (TaKaRa). Sequences of adaptor primers and gene specific primers used in the experiment were listed in Table 1.

Table 1. Primers used in cloning of the full length *ErPAL1*.

A

	Primer sequences
Adaptor primer	AP1 5'-GTAATACGACTCACTATAGGGC-3'
	AP2 5'-ACTATAGGGCACGCGTGGT-3'
First amplification	GPS1 5'- AGCTTCTGCATCAACGGGTAAGTGGCG-3'
	GPS2 5'- CTGCAGGGGTCATCAGCGTAGGTGAAC-3'
Second amplification	GPS1 5'- GATAGGCTGAATAGGACCGCTAACAGA-3'
	GPS2 5'- TTCAACTTTTTGTCTCGACTCGCTAGG-3'
Third amplification	GPS1 5'- ATAAGTTCCTTCTGCAAAGCACCGCC-3'
	GPS2 5'- TTGCTTCGTCCTCCGGTGAGAAGTGG-3'
Forth amplification	GPS1 5'- TGTACGAGGAGAGATGAGCGAACGAA-3'
	GPS2 5'- AAGAGAACTGGACGAGAGCGCTTTGC-3'

B

	Primer sequences
Adaptor primer	Same as Table 1a
First amplification	GPS1 5'- TGGATGGAAGCGCTTATGTGAAGGCTG-3'
	GPS2 5'- TTGCACGAGATGGATCCGCTCCAGAAG-3'

A: The upstream amplification primers
B: The downstream amplification primers

Sequence analysis of *ErPAL1*

The amino acid sequence of *ErPAL1* was deduced and analyzed with DNA to protein sequence converter (http://bioinformatics.picr.man.ac.uk/research/software/tools/sequenceconverter.html). BLAST was done at the NCBI server (http://www.ncbi.nlm.nih.gov/BLAST/). Multiple alignment analysis was performed by Clustal W (http://www.genome.jp/tools/clustalw/). Transcription start site was predicated by PlantProm (http://mendel.cs.rhul.ac.uk/mendel.php). Analysis for transcription response elements was done with PLACE (http://www.dna.affrc.go.jp/PLACE/signalscan.html).

Expression of *ErPAL1* in *Escherichia coli* BL21

Primers PALFEcoRI (5'- CAGAATTCTATGGAGATGGAGAGCACCAC-3', *Eco*RI site was underlined) and PALRXbaI (5'-CATTCTAGACTAAGAGATGGGAAGAGGAG-3', *Xba*I site was underlined) were designed and synthesized to amplify the coding sequence of *ErPAL1*. Reaction was performed as follows: 94° C for 2 min, 35 cycles of amplification (94° C for 30 sec, 54° C for 30 sec, 72° C for 3min), then 72° C extension for 3 min. Digested with *Eco*RI and *Xba*I (TaKaRa), the PCR products were purified by QIAquick gel Extraction Kit (QIAGEN). After that, the products were ligated into expression vector pEcoli-Nterm 6 × HN (Clontech), which was pre-digested with the same restriction enzymes. Then recombinant plasmid, the vector-*ErPAL1*, was sequenced to confirm the correction of the ORF. Subsequently, the expression plasmid was transferred into *E.coli* strain BL21 (Promega) for protein expression.

A single colony of *E.coli* strain BL21 cells harboring the expression plasmid pEcoli-ErPAL1 was inoculated at 37° C in LB medium containing ampicillin (100 mg l^{-1}) with shaking until the optical density OD_{600} reached 0.6 ～ 0.8. Then the protein expression was induced by addition of isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 1 mM. Cells containing ErPAL1 were induced at 37° C for 4 h and harvested by centrifugation at 1,000 g for 15 min. Then supernatant was removed. In order to obtain enzyme solution, the cell pellets were resuspended in TALON xTractor Buffer (Clontech). Aliquot buffer was added to the pellet with 2 μl DNase I (TaKaRa) per 1 ml buffer. After incubation at room temperature for 30 min, the solution was centrifuged at 10,000 g for 20 min, and the supernatant containing recombinant ErPAL1 protein was obtained.

Purification and electrophoresis of ErPAL1 proteins

The recombinant ErPAL1 proteins with his-tag were purified on a HisTALON Superflow Cartridges which was charged with cobalt (Co^{2+}) (Clontech) according to the protocol suggested by the manufacturer.

The expression and purification level of the protein was assessed by 10% SDS-PAGE followed by EzStain Aqua staining (ATTO).

Activity assay and biochemical properties analysis

Using albumin standard curve, protein concentrations were determined by measuring the 595nm absorbance of purified ErPAL1 protein stained by Protein assay CBB solution (Nacalai tesque).

PAL activity was assayed by measuring *trans*-cinnnamic acid formation with the absorbance increase at 280 nm. *Trans*-cinnamic acid was used to generate the calibration curve. The reaction mixture contained 50 mM Tris-HCl, pH 9.0, 61 mM L-phenylalanine, and aliquot of purified ErPAL1 in a total volume of 50 µl. The reaction was carried out at 40° C for 1 h and terminated by addition of 200 µl of 1.2 N HCl.

In order to determine optimum pH, assays were performed at 40° C with citric acid buffer (50 mM; pH 5.0, 6.0), Tris-HCl buffer(50mM; pH 7.0, 7.5, 8.0, 8.5, 9.0) and Glycine-NaOH buffer (50 mM; pH 9.5, 10.0) as described above. For optimum temperature, incubations were carried out a constant pH of 9.0, with various temperatures (5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 and 75° C).

In order to determine the kinetic parameters, a range of concentrations of L-phenylalanine (0.01, 0.02, 0.04, 0.08, 0.1, 0.4, 0.8 and 1.0 mM) was used. Using DeltaGraph 6 software (Japan Poladegital), hyperbolic plots were obtained to calculate the K_m and k_{cat} values.

Results

Cloning and analysis of *ErPAL1*

Using the partial sequence of *ErPAL1* obtained from previous study (AKIMOTO *et al*., 2012), full-length DNA sequence was obtained by genome walking technology. It is designed for *ErPAL1* and its accession number at GenBank is AB698855. The full-length *ErPAL1* is 3,823 bp in size and consisted of two exons and an intron of 1,675 bp (Figure 1). Intron of *ErPAL1* had typical characteristics of plant introns, being rich in A + T (66.31 %) and having standard GT/AG splicing site (LORKOVIC *et al*., 2000).

The complete cDNA sequence was also obtained by primers designed from *ErPAL1*.The full-length cDNA sequences contain a 2,148 bp ORF encoding a 715-amino-acid protein. By using Expasy (http://web.expasy.org/compute_pi/), the theoretical isoelectric point (pI) and molecular weight (MW) of the deduced ErPAL1 protein were predicted to be 5.99 and 77.196 kDa respectively. According to BLAST in DDBJ (http://www.ddbj.nig.ac.jp/search/top-j.html) and multi-alignment analysis by Clustal W (http://www.genome.jp/tools/clustalw/), sequence of the protein was found to share 68-84 % identity with PALs from other plant species. For example, ErPAL1 protein had a similarity of 84 % identity to MsPAL (X58180), 82 % identity to DcPAL1 (D85850), 79 % identity to AtPAL1 (L33677), 75 % identity to BoPAL1 (AY450643) and 68 % identity to OsPAL (X16099) (Figure 2) (Table 2). Additionally, all PAL sequences contained some active site motifs conserved and marked by shadow box, and a Ala-Ser-Gly triad which can be converted into a 3,5-dihydro-5-methylidine-4H-imidazol-4-one (MIO) prothetic group (boxed) (RÖTHER *et al*., 2002; RÉTEY, 2003; MAHESH *et al*., 2006; MACDONALD and D'CUNHA, 2007).

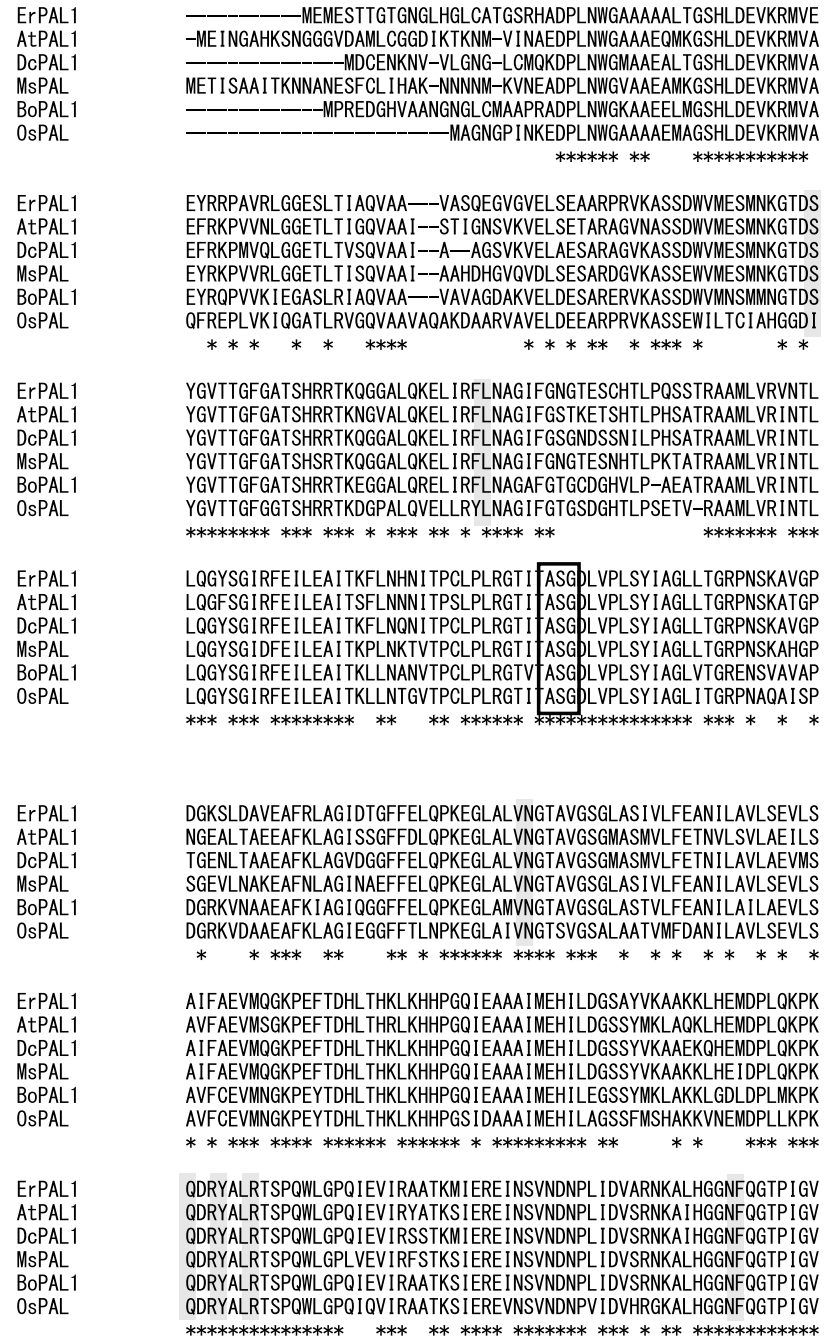
```
ATGGAGATGGAGAGCACCACCGGCACCGGCAACGGCCTTCACGGCCTCTGCGCCACCGGGAGCCGCCATGCCGACCCACTGAACTGGGGGGCGGCGGCAGCAGCCCTCACAGGGAGC
 M  E  M  E  S  T  T  G  T  G  N  G  L  H  G  L  C  A  T  G  S  R  H  A  D  P  L  N  W  G  A  A  A  A  A  A  L  T  G  S
CACCTCGACGAGGTGAAGCGGATGGTTGAGGAGTACCGGAGGCCGGCGGTGCCCTCGGCGGGGAGTCCCTCACGATAGCCCAGGTGGCGGCGGTGGCGAGTCAGGAGGGGGTAGGG
 H  L  D  E  V  K  R  M  V  E  E  Y  R  R  P  A  V  R  L  G  G  E  S  L  T  I  A  Q  V  A  A  V  A  S  Q  E  G  V  G
GTCGAGCTCTCGGAGGCGGCCCGTCCCAGGGTCAAGGCCAGCAGCGACTGGGTCATGGAGAGCATGAACAAGGGAACTGACAGCTACGGCGTCACCACCGGGTTCGGCGCCACTTCT
 V  E  L  S  E  A  A  R  P  R  V  K  A  S  S  D  W  V  M  E  S  M  N  K  G  T  D  S  Y  G  V  T  T  G  F  G  A  T  S
CACCGGAGGACGAAGCAAGGCGGTGCTTTGCAGAAGGAACTTATAAGGTAAACTTTGCACTACGTTAATGATTGTTAAATCACTCCATTTCGGCCATAATACCAATCATAGGTCTAA
 H  R  R  T  K  Q  G  G  A  L  Q  K  E  L  I  R  -  -  -  -  -  i  n  t  r  o  n  -  -  -  -  -  -  -  -  -  -  -  -  -
CAGATGATTGGCTTGTGATCTATATCTGTAATAATTAAAAGGTGGTAAGTACGTAAAAGGAGGTGGTAGGTGGGTTTTTGAGTGATCCACTTCACCGGTTTCTTCACAAATCGGTGA
AGGGTCATTAAAGCTTAGCTGGCCCTTTTTCATTCTTTCTTTGCTTCTTTATTTTAATTTTTAAGAAATTTGTGATTTCTTTTTTTTTCAAAATTCAATACCGTAATAAGCTTCGATG
AATTTTGCATGATATATTTTCCATAAGTATATATCATGTGAAGTCACTTATTGATGAGTATTCTTGTTCTTGGTGAGAAGTAATTTTATTTATTTATTATGATTAATTATTTCTCCA
TTTTATTTTAATCGGGAACTTAAATTTCATCAACTCTTTTCGTTTTTTTTTTTATACTAATTTCTGGATAATGACAAAAGGGCACCAGATAATTCAGAGTGAGGACTAATAATAACG
ATGAAAATAACTCGCTTTGTTATGCAAATTGGGATCATTTAGTGGTTGAGTAGGAACAACTATGGTCATAGTCATGTGACAAAAGGTGTCATATTTATTAGGCCTTGTTGAGATGTG
ATTGTGAGGACTACAAGTGGCCACAACTTTAGATGGGGGCAGGACACTGAATGATGTCGCCACCAAATTGGTTTGATTTAATGCTCCAATTGGAATTACTTTGCTGCTTGGAGGGAA
CATAATATATATTTTCAGTTCGATTTCCGGGTTTGAAACCTGACCGGACCCAAATTGAGACAATAGCAATACTTGAATTTCAAAGCTATTGCCAAAACTGAAATAGATCCAAACCCA
AAAAAACCCAAAACATTTTGATTCTGGTTGGTTGGACTTTTTGATAGCTCAACTAATTGGATAAACTGTCGGTACATGCTCTATTTTTGTTTTACAGAATATATATCTTTTCATTG
TAGAGATTCATCGTTTTCATCAACTCATGACTTCCAATTAAAAGGACAGATTGGACATCATAATTAGGCTCATCCTTAACTATAGTACCTGTCGATGTCTAGGCGATATACATGTTA
TTTATGAGGACCAATAGCTATTATACTTATCAATATTATTGTTAAAACTTGATTTACTCCATTTATATTTGATAAATTAGCATAATAAACACGTCCATCATGTTTAGTAGGTGCGGA
TGGCACCTTTTTCCTAATGGGTTGGCAGAAAGATAAGGATAAACTAGGTTTAAACGGATATTCCAACGTGCATCATGTATTTAAAATACCTATGTAGAATTTTTTTTTAAAATTTAA
TATCTTATATTTATTTTACAAGTCAGTGTGTGTTATAATTTCATCTCTATGGCTATCGTCGTACCTAGCGAGTCGAGACAAAAAGTTGAATAATTGAAATCTGACTTAGAAAAATCT
GTTAGCGGTCCTATTCAGCCTATCCTCGTTTGTCTTCTCTCCTGCATTTCAGTTCACACGGTTGCAATTGAAATGCAAAAGAGTCGCTTGATTACTACACTAAACAGAACTTTGAGC
GAGACCTTTGAACTGCAAAATACTTTACTCAGCAAGAACTAGGAGAGGTCCTTGTCTAATGGAAAGCTTGTGAATCTGGAACAGGTTCTTGAATGCCGGGATCTTCGGCAACGGCAC
- - - - - - - - i n t r o n - - - - - - - - - - - - - - - - F  L  N  A  G  I  F  G  N  G  T
GGAGTCGTGCCACACCCTGCCTCAATCCTCCACCCGAGCCGCCATGCTTGTCCGGGTCAACACCCTCCTCCAGGGCTACTCCGGCATCCGTTTTGAGATCCTCGAGGCCATCACCAA
 E  S  C  H  T  L  P  Q  S  S  T  R  A  A  M  L  V  R  V  N  T  L  L  Q  G  Y  S  G  I  R  F  E  I  L  E  A  I  T  K
GTTCCTCAACCACAACATCACCCCGTGCCTGCCCCTCAGGGGCACCATCACTGCCTCAGGCGACTTGGTCCCCCTCTCCTACATCGCCGGGCTCCTGACGGGCCGGCCCAACTCCAA
 F  L  N  H  N  I  T  P  C  L  P  L  R  G  T  I  T  A  S  G  D  L  V  P  L  S  Y  I  A  G  L  L  T  G  R  P  N  S  K
GGCCGTCGGGCCTGATGGGAAGTCCCTGGACGCCGTCGAGGCCTTCCGGCTTGCCGGGATTGACACGGGATTCTTCGAGCTGCAGCCAAAGGAAGGGTTGGCGCTCGTGAACGGCAC
 A  V  G  P  D  G  K  S  L  D  A  V  E  A  F  R  L  A  G  I  D  T  G  F  F  E  L  Q  P  K  E  G  L  A  L  V  N  G  T
GGCAGTCGGGTCTGGCCTGGCTTCCATCGTCCTCTTCGAGGCCAACATACTCGCGGTCCTGTCCGAGGTCCTGTCAGCGATCTTCGCAGAGGTGATGCAGGGGAAGCCGGAGTTCAC
 A  V  G  S  G  L  A  S  I  V  L  F  E  A  N  I  L  A  V  L  S  E  V  L  S  A  I  F  A  E  V  M  Q  G  K  P  E  F  T
AGACCACTTGACGCATAAATTGAAGCACCATCCCGGGCAGATTGAGGCTGCGGCTATAATGGAGCACATTTTGGATGGAAGCGCTTACGTGAAGGCTGCTAAAAAGTTGCACGAGAT
 D  H  L  T  H  K  L  K  H  H  P  G  Q  I  E  A  A  A  I  M  E  H  I  L  D  G  S  A  Y  V  K  A  A  K  K  L  H  E  M
GGATCCGGCTCCAGAAGCCAAAGCAGGACAGGTACGCTCTCAGGACTTCTCCCCAGTGGCTAGGGCCCCAGATTGAGGTGATCCGAGCGGCGACCAAGATGATTGAGAGGGAAATCAA
 D  P  L  Q  K  P  K  Q  D  R  Y  A  L  R  T  S  P  Q  W  L  G  P  Q  I  E  V  I  R  A  A  T  K  M  I  E  R  E  I  N
TTCGGTCAACGACAACCCACTGATCGACGTCGCGAGGAACAAGGCCCTGCACGGTGGGAACTTCCAGGGGACCCCGATTGGTGTCTCCATGGACAACACTCGCCTGGCGATTGCGTC
 S  V  N  D  N  P  L  I  D  V  A  R  N  K  A  L  H  G  G  N  F  Q  G  T  P  I  G  V  S  M  D  N  T  R  L  A  I  A  S
CATAGGGAAGCTCATGTTCGCGCAGTTCTCAGAGCTTGTCAACGACTTCTACAACAACGGGCTGCCCTCAAATCTGTCGGGCGGCCGCAACCCTAGCCTGGACTACGGCTTCAAGGG
 I  G  K  L  M  F  A  Q  F  S  E  L  V  N  D  F  Y  N  N  G  L  P  S  N  L  S  G  G  R  N  P  S  L  D  Y  G  F  K  G
TGCAGAGATCGCCATGGCATCATACTGCTCGGAGCTCCAATTCCTCGCCAATCCCGTTACCAATCATGTCCAGAGCGCGGAGCAGCACAACCAGGATGTCAACTCCCTAGGCCTGAT
 A  E  I  A  M  A  S  Y  C  S  E  L  Q  F  L  A  N  P  V  T  N  H  V  Q  S  A  E  Q  H  N  Q  D  V  N  S  L  G  L  I
CTCATCGAGGAAGACTGCCGAGGCCATCGACATATTGAAGCTCATGTCTTCCACTTTCCTGGTTGCTCTGTGCCAGGCAGTTGACCTGAGGCATCTGGAGGAGAACCTGAAGAGCAC
 S  S  R  K  T  A  E  A  I  D  I  L  K  L  M  S  S  T  F  L  V  A  L  C  Q  A  V  D  L  R  H  L  E  E  N  L  K  S  T
GGTGAAGAACACGGTCGGCAGGGTGGCCAGGAAAGTCCTCATGGTGGGCGCTAATGGCGAGCTCCACCCGTCTCACTACTGTGAGAGGGACCTGCTCAAAGTGGTCGACCGCGAGCA
 V  K  N  T  V  G  R  V  A  R  K  V  L  M  V  G  A  N  G  E  L  H  P  S  H  Y  C  E  R  D  L  L  K  V  V  D  R  E  H
TGTGTTCACCTACGCTGATGACGCCTGCAGCGCCACCTACCCGCTGATGCAGAAGCTGAGGCAAGTCCTGGTCGACCAGGCACTGGTGAATGGCGAGAGCGAGCTGAACCCGAGCAC
 V  F  T  Y  A  D  D  A  C  S  A  T  Y  P  L  M  Q  K  L  R  Q  V  L  V  D  Q  A  L  V  N  G  E  S  E  L  N  P  S  T
TTCGATCTTCCAAAAGATCGTGGCCTTCGAGGAGGAGCTCAAGGCCCAGTTGCCGAAGGACGTTGAGGGCATTCGAGTCCAGTACGAGACAGGCAGCCTCGCCATCCCCAACCAGAT
 S  I  F  Q  K  I  V  A  F  E  E  E  L  K  A  Q  L  P  K  D  V  E  G  I  R  V  Q  Y  E  T  G  S  L  A  I  P  N  Q  I
CAAGGAATGCAGGTCCTATCCATTGTACAAGCTGGTGAGGGAGGAGCTGGGGACTGCCCTGCTCACGGGCGAGGGCGTGATATCCCCTGGCGAGGACTTTGACAAGGTCTTCACTGC
 K  E  C  R  S  Y  P  L  Y  K  L  V  R  E  E  L  G  T  A  L  L  T  G  E  G  V  I  S  P  G  E  D  F  D  K  V  F  T  A
GATCTGTGCTGGAAAACTGATTGATCCGCTGCTGGAGTGCCTAAGCGGTTGGAACGGTGCTCCTCTTCCCATCTCTTAG
 I  C  A  G  K  L  I  D  P  L  L  E  C  L  S  G  W  N  G  A  P  L  P  I  S
```

Figure 1. Nucleotide sequence and the deduced amino acid sequence of *ErPAL1*. The start codon (ATG) and the stop codon (TAG) are boxed.
GT/AG splicing site are shaded.
The active site consensus (Ala-Ser-Gly) sequence is underlined.

Analysis of the 5'-flanking region of *ErPAL1*

The 5'-flanking sequence (913 bp from start codon ATG) was obtained from *E. robusta* DNA by using genome walking technology. By Plant Prom database, transcription start site of *ErPAL1* was predicted and described as +1. It is located at 95 bp upstream from the start codon ATG (Figure 3).

The sequence analysis of the putative *cis*-acting elements was performed by PLACE. Location


ErPAL1 ---------MEMESTTGTGNGLHGLCATGSRHADPLNWGAAAAALTGSHLDEVKRMVE
AtPAL1 -MEINGAHKSNGGGVDAMLCGGDIKTKNM-VINAEDPLNWGAAAEQMKGSHLDEVKRMVA
DcPAL1 ----------MDCENKNV-VLGNG-LCMQKDPLNWGMAAEALTGSHLDEVKRMVA
MsPAL METISAAITKNNANESFCLIHAK-NNNNM-KVNEADPLNWGVAAEAMKGSHLDEVKRMVA
BoPAL1 ----------MPREDGHVAANGNGLCMAAPRADPLNWGKAAEELMGSHLDEVKRMVA
OsPAL ----------MAGNGPINKEDPLNWGAAAAEMAGSHLDEVKRMVA
EYRRPAVRLGGESLTIAQVAA---VASQEGVGVELSEAARPRVKASSDWVMESMNKGTDS
EFRKPVVNLGGETLTIGQVAAI--STIGNSVKVELSETARAGVNASSDWVMESMNKGTDS
EFRKPMVQLGGETLTVSQVAAI--A--AGSVKVELAESARAGVKASSDWVMESMNKGTDS
EYRKPVVRLGGETLTISQVAAI--AAHDHGVQVDLSESARDGVKASSEWVMESMNKGTDS
EYRQPVVKIEGASLRIAQVAA---VAVAGDAKVELDESARERVKASSDWVMNSMMNGTDS
QFREPLVKIQGATLRVGQVAAVAQAKDAARVAVELDEEARPRVKASSEWILTCIAHGGDI
YGVTTGFGATSHRRTKQGGALQKELIRFLNAGIFGNGTESCHTLPQSSTRAAMLVRVNTL
YGVTTGFGATSHRRTKNGVALQKELIRFLNAGIFGSTKETSHTLPHSATRAAMLVRINTL
YGVTTGFGATSHRRTKQGGALQKELIRFLNAGIFGSGNDSSNILPHSATRAAMLVRINTL
YGVTTGFGATSHSRTKQGGALQKELIRFLNAGIFGNGTESNHTLPKTATRAAMLVRINTL
YGVTTGFGATSHRRTKEGGALQRELIRFLNAGAFGTGCDGHVLP-AEATRAAMLVRINTL
YGVTTGFGGTSHRRTKDGPALQVELLRYLNAGIFGTGSDGHTLPSETV-RAAMLVRINTL
LQGYSGIRFEILEAITKFLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGP
LQGFSGIRFEILEAITSFLNNNITPSLPLRGTITASGDLVPLSYIAGLLTGRPNSKATGP
LQGYSGIRFEILEAITKFLNQNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGP
LQGYSGIDFEILEAITKPLNKTVTPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAHGP
LQGYSGIRFEILEAITKLLNANVTPCLPLRGTVTASGDLVPLSYIAGLVTGRENSVAVAP
LQGYSGIRFEILEAITKLLNTGVTPCLPLRGTITASGDLVPLSYIAGLITGRPNAQAISP
DGKSLDAVEAFRLAGIDTGFFELQPKEGLALVNGTAVGSGLASIVLFEANILAVLSEVLS
NGEALTAEEAFKLAGISSGFFDLQPKEGLALVNGTAVGSGMASMVLFETNVLSVLAEILS
TGENLTAAEAFKLAGVDGGFFELQPKEGLALVNGTAVGSGMASMVLFETNILAVLAEVMS
SGEVLNAKEAFNLAGINAEFFELQPKEGLALVNGTAVGSGLASIVLFEANILAVLSEVLS
DGRKVNAAEAFKIAGIQGGFFELQPKEGLAMVNGTAVGSGLASTVLFEANILAILAEVLS
DGRKVDAAEAFKLAGIEGGFFTLNPKEGLAIVNGTSVGSALAATVMFDANILAVLSEVLS
AIFAEVMQGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSAYVKAAKKLHEMDPLQKPK
AVFAEVMSGKPEFTDHLTHRLKHHPGQIEAAAIMEHILDGSSYMKLAQKLHEMDPLQKPK
AIFAEVMQGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSSYVKAAEKQHEMDPLQKPK
AIFAEVMQGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSSYVKAAKKLHEIDPLQKPK
AVFCEVMNGKPEYTDHLTHKLKHHPGQIEAAAIMEHILEGSSYMKLAKKLGDLDPLMKPK
AVFCEVMNGKPEYTDHLTHKLKHHPGSIDAAAIMEHILAGSSFMSHAKKVNEMDPLLKPK
QDRYALRTSPQWLGPQIEVIRAATKMIEREINSVNDNPLIDVARNKALHGGNFQGTPIGV
QDRYALRTSPQWLGPQIEVIRYATKSIEREINSVNDNPLIDVSRNKAIHGGNFQGTPIGV
QDRYALRTSPQWLGPQIEVIRSSTKMIEREINSVNDNPLIDVSRNKAIHGGNFQGTPIGV
QDRYALRTSPQWLGPLVEVIRFSTKSIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGV
QDRYALRTSPQWLGPQIEVIRAATKSIEREINSVNDNPLIDVSRNKALHGGNFQGTPIGV
QDRYALRTSPQWLGPQIQVIRAATKSIEREVNSVNDNPVIDVHRGKALHGGNFQGTPIGV

```
ErPAL1          SMDNTRLAIASIGKLMFAQFSELVNDFYNNGLPSNLSGGRNPSLDYGFKGAEIAMASYCS
AtPAL1          SMDNTRLAIAAIGKLMFAQFSELVNDFYNNGLPSNLTASRNPSLDYGFKGAEIAMASYCS
DcPAL1          SMDNTRLAIAAIGKLMFAQFSELVNDFYNNGLPSNLSGGRNPSLDYGFKGAEIAMASYCS
MsPAL           SMDNTRLALASIGKLMFAQFSELVNDFYNNGLPSNLSASRNPSLDYGFKGAEIAMASYCS
BoPAL1          SMDNTRLAIAAIGKLMFAQFSELVNDFYNNGLPSNLSGGRNPSLDYGFKGAEIAMASYCS
OsPAL           SMDNARLAIANIGKLMFAQFSELVNEFYNNGLTSNLAGSRNPSLDYGFKGTEIAMASYSS
                **** *** * ************** ****** ***   *********** ******* *

ErPAL1          ELQFLANPVTNHVQSAEQHNQDVNSLGLISSRKTAEAIDILKLMSSTFLVALCQAVDLRH
AtPAL1          ELQYLANPVTSHVQSAEQHNQDVNSLGLISSRKTSEAVDILKLMSTTFLVAICQAVDLRH
DcPAL1          ELQFLGNPVTNHVQSAEQHNQDVNSLGLISSRKTAEAVEILKLMSTTFLVGLCQAVDLRH
MsPAL           ELQYLANPVTTHVQSAEQHNQDVNSLGLISARKTNEAIEILQLMSSTFLIALCQAIDLRH
BoPAL1          ELQFLGNPVTNHVQSAEQHNQDVNSLGLISSRKTAEAIDILKIMSSTFLVALCQAIDLRH
OsPAL           ELQYLANPITNHVQSAEQHNQDVNSLGLVSARKTLEAVDILKLMTSTYIVALCQAVDLRH
                *** * ** * ********************* * *** **  **  *  *    *** ****

ErPAL1          LEENLKSTVKNTVGRVARKVLMVGANGELHPSHYCERDLLKVVDREHVFTYADDACSATY
AtPAL1          LEENLRQTVKNTVSQVAKKVLTTGVNGELHPSRFCEKDLLKVVDREQVYTYADDPCSATY
DcPAL1          LEENLKSTVKNTVSQVAKKVLTMGVNGELHPSRFCELDLLRVVDREYIFAYIDDPCSATY
MsPAL           LEENLKNSVKNTVSQVAKKTLTMGVNGELHPSRFCEKDLLKVVDREHVFAYIDDPCSATY
BoPAL1          IEENVKSAVKSCVMTVAKKTPSTNSTGDLHVARFCEKDLLKEIDREAVFAYADDPCSPNY
OsPAL           LEENIKSSVKNCVTQVAKKVLTMNPTGDLSSARFSEKNLLTAIDREAVFSYADDPCSANY
                 ***    **  *  ** *       * *      *  **   ***    * ** **  *

ErPAL1          PLMQKLRQVLVDQALVNGESELNPSTSIFQKIVAFEEELKAQLPKDVEGIRVQYETGSLA
AtPAL1          PLIQKLRQVIVDHALINGESEKNAVTSIFHKIGAFEEELKAVLPKEVEAARAAYDNGTSA
DcPAL1          PLMQKLRQVLVEHALKNGETEKNLSTSIFQKIAAFEDELKALLPKEVESARAVVESGNPA
MsPAL           PLSQKLRQVLVDHALVNGESEKNFNTSIFQKIATFEEELKTLLPKEVESARTAYESGNPT
BoPAL1          PLMKKLRSVLVESALANGMAEFNAETSIFARVALFEEELRAALPRAVEAARASVENGTAA
OsPAL           PLMQKLRAVLVEHALTSGDRRARGLRVLQDHQVRGGAPLCAAPGDRGRPRRRRQRTAPVA
                **  *** * *  **  *                   *          *

ErPAL1          IPNQIKECRSYPLYKLVREELGTALLTGEGVISPGEDFDKVFTAICAGKLIDPLLECLSG
AtPAL1          IPNRIKECRSYPLYRFVREELGTELLTGEKVTSPGEEFDKVFTAICEGKIIDPMMECLNE
DcPAL1          IPNRIKECRSYPLYKFIREELGTVYLTGEKVTSPGEEFDKVFTAMSKGEIIDPLLACLES
MsPAL           IPNKINGCRSYPLYKFVREELGTGLLTGENVISPGEECDKLFSAMCQGKIIDPLLECLGE
BoPAL1          APNRITECRSYPLYRFVREELGTEYLTGEKTRSPGEELNKVLLAINQGKHIDPLLECLKE
OsPAL           ——NRIVESRSFPLYRFVREELGCVFLTGEKLKSPGEECNKVFLGISQGKLIDPMLDCLKE
                  * *   ** ***   *****   ****   ****   *      *  ***   **

ErPAL1          WNGAPLPIS
AtPAL1          WNGAPIPIC
DcPAL1          WNGAPLPIA
MsPAL           WNGAPLPIC
BoPAL1          WNGEPLPIC
OsPAL           WNGEPLPIN
                *** * **
```

Figure 2. Multiple alignment analysis of PAL protein sequences by Clustal W.
There are 8 conserved active site motif (shaded).
Ala-Ser-Gly triad is boxed.
Abbreviations of plant names are the same as Table 2.

and functions of the elements were indicated in Figure 3 and Table 3. A putative TATA box of *ErPAL1* was found to be located at position -85. Its sequence (TATTTAA) was the same as that of *OsPAL* (ZHU *et al*., 2002). There are about 7 CAAT boxes, the most common elements, and two of them are marked in figure 3. A significant *cis*-element for the promoters of phenylpropanoid pathway, Box L (YCYYACCWACC) (SONG and WANG, 2009) was located at position -280. Furthermore, Box P (YYMMCMAMCMMC) and Box A (MCGTCY) (LOGEMANN *et al*., 1995;

Table 2. Amino-acid size of some plant PALs and their homology with ErPAL1 protein.

	Homology with ErPAL1	Amino-acid size	Accession number
ErPAL1	100%	715	AB698855
AtPAL1	79%	725	L33677
BoPAL1	75%	712	AY450643
DcPAL1	81%	708	AB041361
MsPAL	84%	665	X58180
NtPAL1	80%	715	M84466
PcPAL1	80%	716	X15473
PtPAL1	81%	711	AF480619
OsPAL	68%	701	X16099
SmPAL1	81%	711	EF462460

Er, *Eucalyptus robusta*; At, *Arabidopsis thaliana*; Bo, *Bambusa oldhamii*; Dc, *Daucus carota*; Ms, *Medicago sativa*; Nt, *Nicotiana tabacum*; Pc, *Petroselium crispum*; Pt, *Populus tremuloides*; Sm, *Salvia miltiorrhiza*

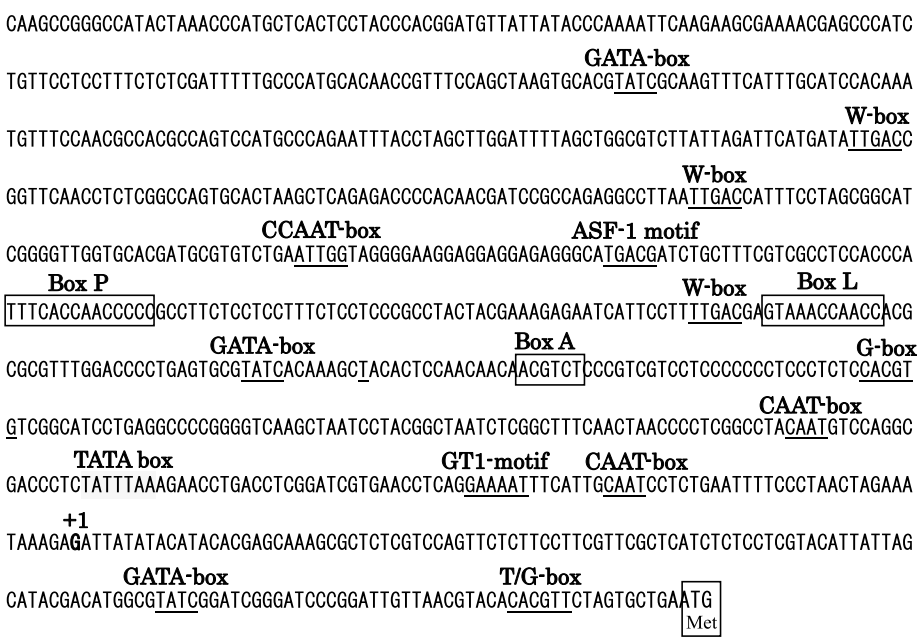


Figure 3. The 5'-flanking region of *ErPAL1*.
TATA box is shaded, and Box P, A, L are boxed.
Other *cis*-elements are underlined.
The putative transcription start by Plant Prom is marked with +1.

Table 3. Function of *cis*-elements in *ErPAL1* 5'-flanking region.

Cis-element	Function
TATA-box	DNA sequence that indicates the point at which a genetic sequence can be read
CAAT-box	A highly conserved DNA sequence found about 75 bp 5' to the site of transcription
Box-P	One of putative *cis*-acting elements of phenylalanine ammonia-lyase genes
Box-A	One of putative *cis*-acting elements of phenylalanine ammonia-lyase genes
Box-L	One of putative *cis*-acting elements of phenylalanine ammonia-lyase genes
GATA-box	One of c*is*-acting regulatory element involved in light responsiveness
G-box	One of c*is*-acting regulatory element involved in light responsiveness
GT1-motif	One of c*is*-acting regulatory element involved in light responsiveness
W-box	Wounding and pathogen responsive element
ASF-1 motif	*Cis*-acting element involved in auxin and salicylic acid treatments
CCAAT-box	Heat shock responsive element
T/G-box	Element involved in jasmonate induction

SCHMIDT *et al*., 2004), which were reported to exist in the promoter region of many plant species, were found at position -361 and -228, respectively. Three types of *cis*-acting elements related to light responsiveness (GATA-box, G-box, GT1-motif) were found in the promoter region (LAM and CHUA, 1989; WILLIAMS *et al*., 1992; TERZAGHI and CASHMORE, 1995). In addition, W-box (TTGAC) (YU *et al*., 2001), ASF-1 motif (TGACG) (DESPRES *et al*., 2003), CCAAT-box (ATTGG) (WENKEL *et al*., 2006) and T/G-box (CACGTT) (BOTER *et al*., 2004) were found.

Expression of *ErPAl1* in *E. coli* BL21

The recombinant $(HN)_6$-tagged ErPAL1 protein was purified by the HisTALON Superflow Catridge. Purified ErPAL1 was confirmed by SDS-PAGE (Figure 4). A major band located between molecular mass of 97.2 and 66.4 kDa stained by CBB agreed with the putative mass of ErPAL1 (77.2 kDa). Furthermore, it was found that purified ErPAL1 was active (data not shown).

Estimating from single band of lane 2 in figure 4, it seemed that ErPAL1 eluted from the cartridge was nearly homogenous. Thus, approximately 7mg of recombinant ErPAL1 was purified from 100 ml of culture. High level of expression of the recombinant protein made it possible to do further biochemical and kinetic studies of ErPAL1.

With purified protein, the optima temperature and pH of recombinant ErPAL1 were investigated. It revealed that the optima temperature and pH of the protein was 50° C (Figure 5A) and pH 9.0 (Figure 5B), respectively. The kinetic parameters of ErPAL1 were calculated by Delta Graph 6 software (Figure 6). The values of K_m and k_{cat} were 203 μM and 5.26 s^{-1}, each.

Discussion

In this analysis, genomic sequence and cDNA sequence of *ErPAL1* were cloned. The full length of genomic *ErPAL1* was 3,823 bp and composed of two exons (first: 398 bp, second: 1,750 bp) and an intron (1,675 bp). The size of second exon of *ErPAL1*, 1,750 bp, was the same size with

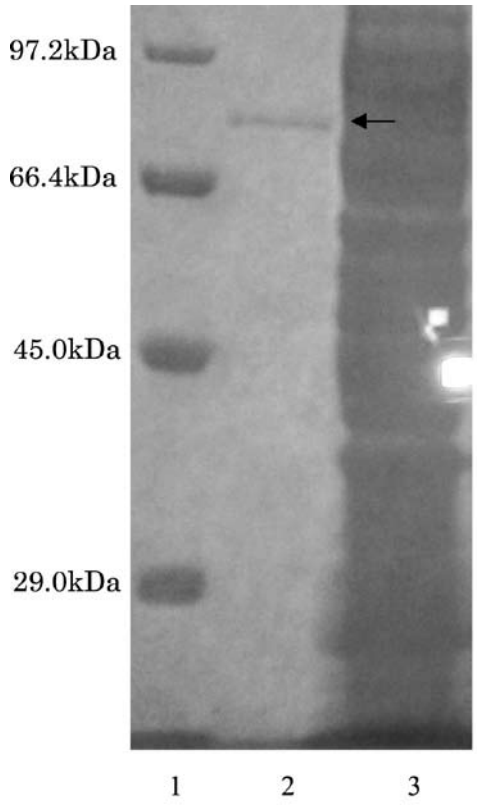


Figure 4. The 10 % SDS-PAGE of the recombinant ErPAL1 expressed in *E. coli* BL21.
Lane 1: Marker;
Lane 2: Purified ErPAL1;
Lane 3: Pre-purified supernatant containing recombinant ErPAL1 and E. *coli*-derived proteins.
The black arrow indicates the target ErPAL1 protein (77.2KD).

that of some other plant *PALs*. The genomic organization of *ErPAL1* was the most common structure of PAL genes in angiosperm plants characterized to date. For instance, *AtPAL1* and *AtPAL2* in *Arabidopsis thaliana* (Wanner *et al*., 1995), *Nicotiana tabacum* (Reichert *et al*., 2009), *Dioscorea bulbifera* (Terauchi and Kahl, 2000), *Salvia miltiorrhiza* (Song and Wang, 2009), and *Oryza sativa* (Minami *et al*., 1989) possesed two exons and an intron respectively. Exceptions were *BoPAL1* (Hsieh *et al*., 2010) and *GbPAL* (Xu *et al*., 2008), which had no intron, and *AtPAL3* and *AtPAL4* in *Arabidopsis thaliana* which consisted of three exons and two introns (Huang *et al*., 2010).

ErPAL1 encoded a 715 amino-acid protein. Its length was similar to PALs of other plants in previous reports (Table 2). To determine the function of the protein, *ErPAL1* was expressed in *E. coli* BL21. It turned out that the recombinant protein possessed PAL activity, which catalyzed the

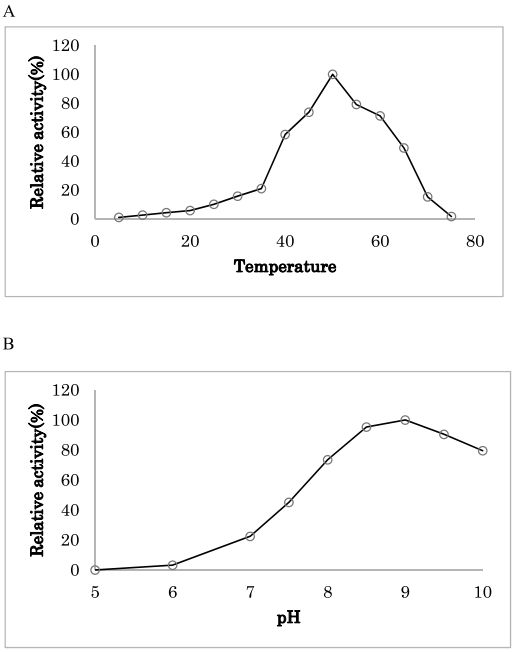


Figure 5. Character of recombinant ErPAL1 protein expressed in *E. coli* BL21.
A: The optima temperature analysis of recombinant ErPAL1.
B: The optima pH analysis of recombinant ErPAL1

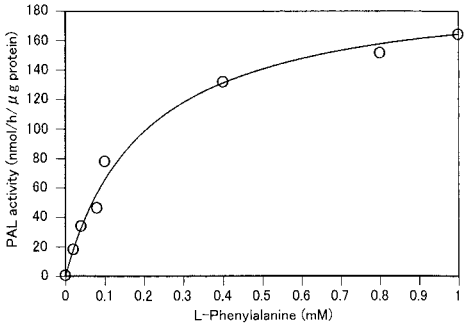


Figure 6. Substrate saturation plot of recombinant ErPAL1.
It was determined as the change in the absorbance of 280 nm for 1 h incubation at various concentrations of L-phenylalanine. Approximated curve was shaped by DeltaGraph 6.
The K_m and k_{cat} values were calculated as 203 μM and 5.26 s^{-1}, respectively.

deamination of phenylalanine to *trans*-cinnamic acid. This suggests that ErPAL1 encoded a functional PAL.

Moreover, biochemical features of recombinant ErPAL1 were examined. The optimum pH of

ErPAL1 was 9.0. It was similar to those of A. thaliana (8.4-9.2) (COCHRANE *et al*., 2004), *S. miltiorrhiza* (8.7) (SONG and WANG, 2009) and *B. oldhamii* (9.0) (HSIEH *et al*., 2010; 2011). The optimum temperature of ErPAL1 was 50° C. It was equal to those of BoPAL1 and BoPAL2 (50° C) (HSIEH *et al*., 2010; 2011), higher than that of *A. thaliana* (31-45° C), and lower than those of *S.miltiorrhizaI* (60° C) and *Petroselium crispum* (58° C) (APPERT *et al*., 1994). Kinetic parameters were also investigated. The K_m value was calculated as 203 μM which was comparable level to BoPAL2 (330 μM) (HSIEH *et al*., 2010) and ZmPAL (250-290 μl) in *Zea mays* (RÖSLER *et al*., 1997). The k_{cat} value was 5.26 s^{-1}. It was similar to that of AtPAL2 (3.2 s^{-1}), and lower than those of BoPAL (10.11 and 16.04 s^{-1}) and PcPAL (22 s^{-1}) in *P. crispum* (APPERT *et al*., 1994). The biochemical properties of the recombinant ErPAL1 were in the same ranges of the attributes of other plant PALs.

Cis-acting elements are the important components of the promoters in plants. The typical elements for the promoters of phenylpropanoid pathway, Box P, Box A, and Box L were found at the region. These three elements were existing in 5'-flanking region of some plant PALs (LOGEMANN *et al*., 1995; SCHMIDT *et al*., 2004; SONG and WANG, 2009). It was reported that Box L was also existing in the promoter region of parsley 4CL (4-coumarate:coenzyme A ligase) (LOGEMANN *et al*., 1995). Additionally, WAKO *et al*. (2010) identified the MYB protein which could bind Box L of *DcPAL3* in *Daucus carota*. On the other hand, there was no specific report about Box P and Box A to date. A lot of light responsive *cis*-elements were present in the 5' -flanking region, suggesting that light could play an important role for regulation of *ErPAL1* expression. There were some W-boxes which binded WRKY transcription factors and functioned as wounding and pathogenic responsiveness. This indicates that the expression of *ErPAL1* may be related to wounding and pathogenic infection. In addition, salicylic acid (ASF-1 motif) and jasmonic acid (T/G-box) treatments, and heat shock (CCAAT-box) were involved in the expression of *ErPAL1*.

In this paper, the full length DNA and cDNA of *ErPAL1* were cloned. The recombinant ErPAL1 protein was biochemically characterized for the first time in eucalyptus. The protein was active and its biochemical properties were similar to those of plant PALs in previous reports. It allowed to investigate the relationship between eucalyptus PAL and the biosynthesis of various phenolics by using the information of *ErPAL1*.

Summary

Phenylalanine ammonia-lyase (PAL) is the first enzyme of phenylpropanoid pathway. It produces precursors of compounds which play important roles during plant development and defense. Using genome walking technology, a PAL gene designated as *ErPAL1* was cloned from *Eucalyptus robusta*. The full length of *ErPAL1* was 3,823 bp. It consists of an intron and two exons encoding a 715 amino-acid polypeptide. Sequence alignment indicated that *ErPAL1* shared 70-80 % identity with the plant PAL sequences in some previous reports. We also obtained the 5'-flanking sequence and found some putative *cis*-acting elements such as TATA box, CAAT box, GT1-motif and G-box. *ErPAL1* was expressed in *Escherichia coli* strain BL21 with pEcoli vector.

The recombinant protein indicated PAL activity which could catalyze the deamination of L-phenylalanine into *trans*-cinnamic acid. The optimum temperature and pH for ErPAL1 activity was 50° C and pH9.0, respectively. The values of K_m and k_{cat} for L-phenylalanine were 203 μM and 5.26 s^{-1} each. The recombinant protein had similar biochemical properties to other plant PALs.

Keywords: *Eucalyptus robusta*, expression, molecular cloning, phenylalanine ammonia-lyase (PAL)

References

AKIMOTO, S., ISHINO T., TERADA, T., KAMODA, S., and SAMEJIMA, M. (2012) Phenolics and Phenylalanine ammonia-lyase gene in *Eucalyptus*. Bull. Tokyo Univ. For. 127: 35-44 (in Japanese).

APPERT, C., LOGEMMAN, E., HAHLBROCK, K., SCHMID, J., and AMRHEIN, N. (1994) Structural and catalytic properties of the four phenylalanine ammonia-lyase isozymes from parsley (*Petroselinum crispum* Nym.). Eur. J. Biochem. 225: 491-499.

BARBER, MS., and MITCHELL, HJ. (1997) Regulation of phenylpropanoid metabolism in relation to lignin biosynthesis in plants. Int. Rev. Cytol. 172: 243-293.

BOTER, M., RUIZ-RIVERO, O., ABDEEN, A., and PRAT, S. (2004) Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and *Arabidopsis*. Genes Dev. 18: 1577-1591.

CADAHÍA, E., CONDE, E., DESIMON, BF., and GARCÍA-VALLEJO, M. (1997) Tannin composition of *Eucalyptus camaldulensis*, *E.globulus* and *E.rudis*. Part I. Bark. Holzforschung 51: 125-129.

CHEN ,AH., CHAI, YR., LI, JN., and CHEN ,L. (2007) Molecular cloning of two genes encoding cinnamate 4-hydroxylase (C4H) from oilseed rape (*Brassica napus*). J. Biochem. Mol. Biol. 40: 247-260.

COCHRANE, FC., DAVIN, LB., and LEWIS, NG. (2004) The *Arabidopsis* phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. Phytochemistry 65: 1557-1564.

CONDE, E., CADAHÍA, E., GARCÍA, E., GARCÍA-VALLEJO, M., and TOMÁS-BARBERÁ, F. (1995) Low molecular weight polyphenols in wood and bark of *Eucalyptus globulus*. Wood Fiber Sci. 27: 379-383.

DESPRES, C., CHUBAK, C., ROCHON, A., CLARK, R., BETHUNE, T., DESVEAUX, D., and FOBERT, PR. (2003) The *Arabidopsis* NPR1 disease resistance protein is a novel cofactor that confers redox regulation of DNA binding activity to the basic domain/leucine zipper transcription factor TGA1. Plant Cell 15: 2181-2191.

DIXON, RA., ACHNINE, L., KOTA, P., LIU, CJ., REDDY, MSS., and WANG, L. (2002) The phenylpropanoid pathway and plant defense-a genomics perspective. Mol. Plant Pathol .3: 371-390.

DIXON, RA., and PAIVA, NL. (1995) Stress-induced phenylpropanoid metabolism. Plant Cell 7: 1085-1097.

FECHTAL, M., and RIEDL, B. (1991) Analyse des ectraits tannants des écorces des *Eucalyptus* après hydrolyse acide par la chromatograhie en phase gazeuse couplée avec la spectrométrie de masse (GC-MS). Holzforschung 45: 269-273.

HAHLBROCK, K., and SCHEEL, D. (1989) Physiology and molecular biology of phenyl-propanoid metabolism. Annu. Rev. Plant Physiol. Plant Mol. Biol. 400: 347-369

HARAKAVA, R. (2005) Genes encoding enzymes of the lignin biosynthesis pathway in *Eucalyptus*. Genet. Mol. Biol. 28: 601-607.

HIGUCHI, T. (1985) Biosynthesis of lignin. In: Biosynthesis and Biodegradation of Wood Components. HIGUCHI, T. (eds.) , Academic Press, New York, 141-160.

HIRAKAWA, H., NAKAMURA, Y., KANEKO, T., ISOBE, S., SAKAI, H., KATO, T., HIBINO, T., SASAMOTO, S., WATANABE, A., YAMADA, M., NAKAYAMA, S., FIJISHIRO, T., KISHIDA, Y., KOHARA, M., TABATA, S., and SATO, S. (2011) Survey of the genetic information carried in the genome of *Eucalyptus camaldulensis*. Plant Biol 28: 471-480.

HSIEH, LS., HSIEH, YL., YEH, CS., CHENG, CY., YANG, CC., and LEE, PD. (2011) Molecular characterization

of a phenylalanine ammonia lyase gene (*BoPAL1*) from *Bambusa oldhamii*. Mol. Biol. Rep. 38: 283-290.

HSIEH, LS., YEH, CS., PAN, HC., CHENG, CY., YANG, CC., and LEE, PD. (2010) Cloning and expression of a phenylalanine ammonia-lyase gene (*BoPAL2*) from *Bambusa oldhamii* in *Escherichia coli* and *Pichia pastoris*. Protein Expr. Purif. 71: 224-230.

HUANG, J., GU, M., LAI, Z., FAN, B., SHI, K., ZHOU, YH., YU, JQ., and CHEN, Z. (2010) Functional analysis of the Arabidopsis *PAL* gene family in plant growth, development, and response to environmental stress. Plant Physiol. 153: 1526-1538.

JOOS, HJ., and HAHLBROCK, K. (1992) Phenylalanine ammonia-lyase in potato (*Solanum tuverosum* L.). Genomic complexity, structural comparison of two selected genes and modes of expression. Eur. J. Biochem. 204: 621-629.

KAWAOKA, A., KAOTHIEN, P., YOSHIDA, K., ENDO, S., YAMADA, K., and EBINUMA, H. (2000) Functional analysis of tobacco LIM protein Ntlim1 involved in lignin biosynthesis. Plant J. 22: 289-301.

LAM, E., and CHUA, NH. (1989) ASF-2: a factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in Cab promoters. Plant Cell 1: 1147-1156.

LINSMAIER, EM., and SKOOG, F. (1965) Organic growth factors requirements of tobacco tissue culture. Physiol. Plant. 18: 100-127.

LIU, RR., XU, SH., LI, JL., HU, YL., and LIN, ZP. (2006) Expression profile of a PAL gene from *Astragalus membranaceus* var. *Mongholicus* and its crucial role in flux into flavonoid biosynthesis. Plant Cell Rep. 25: 705-710.

LOGEMANN, E., PARNISKE, M., and HAHLBROCK, K. (1995) Modes of expression and common structural features of the complete phenylalanine ammonia lyase gene family in parsley. Proc. Natl. Acad. Sci. USA 92:5905-5909.

LORKOVIC, ZJ., WIECZOREK, KDA., LAMBERMON, MH., and FILIPOWICZ, W. (2000) Pre-mRNA splicing in higher plants. Trends Plant Sci. 5:160-167.

MACDONALD, MJ., and D'CUNHA, GB. (2007) A modern view of phenylalanine ammonia lyase. Biochem. Cell Biol. 85: 273-282.

MAHESH, V., RAKOTOMALALA, JJ., LE GAL, L., VIGNE, H., de KOCHKO, A., HAMON, S., NOIROT, M., and CAMPA, C. (2006) Isolation and genetic mapping of a *Coffea canephora* phenylalanine ammonia-lyase gene (*CcPAL1*) and its involvement in the accumulation of caffeoyl quinic acids. Plant Cell Rep. 25: 986-992.

MINAMI, E., OZEKI, Y., MATSUOKA, M., KOIZUKA, N., and TANAKA, Y. (1989) Structure and some characterization of the gene for phenylalanine ammonia-lyase from rice plants. Eur. J. Biochem. 185: 19-25.

PELLEGRINI, L., ROHFRITSCH, O., FRITIG, B., and LEGRAND, M. (1994) Phenylalanine ammonia-lyase in tobacco: molecular cloning and gene expression during the hypersensitive reaction to tobacco mosaic virus and the response to a fungal elicitor. Plant Physiol. 106: 877-886.

REICHERT, AI., HE, XZ., and DIXON, RA. (2009) Phenylalanine ammonia-lyase (PAL) from tobacco (*Nicotiana tabacum*): characterization of the four tobacco *PAL* genes and active heterotetrameric enzymes. Biochem. J. 424: 233-242.

RÉTEY, J. (2003) Discovery and role of methylidene imidazolone, a highly electrophilic group. Biochim. Biohpys. Acta 1647: 179-184.

RÖSLER, J., KREKEL, F., AMRHEIN, N., and SCHMID, J. (1997) Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. Plant Physiol. 113: 175-179.

RÖTHER, D., POPPE, L., MORLOCK, G., VIERGUTZ, S., and RÉTEY, J. (2002) An active site homology model of phenylalanine ammonia-lyase from *Petroselinum crispum*. Eur. J. Biochem. 269: 3065-3075.

SCHMIDT, K., HEBERLE, B., KURRASCH, J., NEHLS, R., and STAHL, DJ. (2004) Suppression of phenylalanine ammonia lyase expression in sugar beet by the fungal pathogen *Cercospora beticola* is mediated at the core promoter of the gene. Plant Mol. Biol. 55: 835-852.

SONG, J., and WANG, Z. (2009) Molecular cloning, expression and characterization of a phenylalanine ammonia-lyase gene (*SmPAL1*) from *Salvia miltiorrhiza*. Mol. Biol. Rep. 36: 939-952.

TERADA, T., YAMAGUCHI, T., and SABURI, Y. (1993) L-phenylalanine ammonia-lyase activities in cell suspension cultures of *Eucalyptus polybractea* I. effects of inoculums densities and conditioned medium

on L-phenylalanine ammonia-lyase activities. J. Wood Sci. 39:958-962.

TERAUCHI, R., and KAHL, G. (2000) Rapid isolation of promoter sequences by TAIL-PCR: the 5'-flanking regions of *Pal* and *Pgi* genes from yams (*Dioscorea*). Mol. Gen. Genet. 263: 554-560.

TERZAGHI, WB., and CASHMORE, AR. (1995) Light-regulated transcription. Annu. Rev. Plant Physiol. Plant Mol. Biol. 46: 445-474.

VÁZQUEZ, G., SANTOS, J., FREIRE, M., ANTORRENA, G., and GONZÁLEZ-ÁLVAREZ, J. (2012) Extraction of antioxidants from eucalyptus (*Eucalyptus globulus*) bark. Wood Sci. Technol. 46: 443-457.

WAKO, T., KIMURA, S., CHIKAGAWA, Y., and OZEKI, Y. (2010) Characterization of MYB proteins acting as transcriptional regulatory factores for carrot *phenylalanine ammonia-lyase* gene (*DcPAL3*). Plant Biotech. 27: 131-139.

WANNER, LA., LI, G., WARE, D., SOMSSICH, IE., and DAVIS, KR. (1995) The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*. Plant Mol. Biol. 27: 327-338.

WENKEL, S., TURCK, F., SINGER, K., GISSOT, L., LE GOURRIEREC, J., SAMACH, A., and COUPLAND, G. (2006) CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of *Arabidopsis*. Plant Cell 18: 2971-2984.

WILLIAMS, ME., FOSTER, R., and CHUA, NH. (1992) Sequences flanking the hexameric G-BOX core CACGTG affect the specificity of protein binding. Plant Cell 4: 485-496.

XU, F., CAI, R., CHENG, S., DU, H., WANT, Y., and CHENG, S. (2008) Molecular cloning, characterization and expression of phenylalanine ammonia-lyase gene from *Ginkgo biloba*. Afr. J. Biotechnol. 7: 721-729.

YAMAGUCHI, T., FUKUZUMI, T., and YOSHIMOTO, T., (1986) Phenolics of tissue cultures from *Eucalyptus* spp. I. Mokuzai Gakkaishi 32:203-208.

YU, D., CHEN, C., and CHEN, Z. (2001) Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression. Plant Cell 13: 1527-1540.

ZHU, Q., ORDIZ, MI., DABI, T., BEACHY, RN., and LAMB C. (2002) Rice TATA binding protein interacts functionally with transcription factor IIB and the RF2a bZIP transcriptional activator in an enhanced plant in vitro transcription system. Plant Cell 14: 795-803.

(Received Jun. 29, 2012)
(Accepted Sep. 19, 2012)

Eucalyptus robusta 由来のフェニルアラニンアンモニアリアーゼ遺伝子のクローニングと解析

秋元真也 *，石野貴久 *，寺田珠実 *，鮫島正浩 *，鴨田重裕 **

* 東京大学大学院農学生命科学研究科生物材料科学専攻
** 東京大学大学院農学生命科学研究科附属演習林樹芸研究所

要　旨

フェニルアラニンアンモニアリアーゼ（PAL）はフェニルプロパノイド経路の初発酵素である。この酵素反応により，植物の成長過程や防御作用において重要な役割を担っていると考えられる化合物が生成されている。ユーカリ属樹木 *Eucalyptus robusta* から，ゲノムウオーキングという手法を用いて PAL ゲノム遺伝子（*ErPAL1*）ならびにその上流配列をクローニングした。*ErPAL1* は全長 3,823bp で 1 つのイントロンと 2 つのエクソン（715 アミノ酸をコード）から成り立っており，これまでに報告されている他の植物の PAL 遺伝子配列と 70-80％の相同性があった。5’ フランキング配列には TATA ボックス，CAAT ボックス，GT1 モチーフや G ボックスなど多くのシス因子に相当する配列が見られた。pEcoli ベクターと大腸菌（BL21）システムを用いて発現させた ErPAL1 は PAL 活性を有していた。さらに ErPAL1 の酵素学的特性を解析したところ，最適温度 50℃，最適 pH 9.0 であり，フェニルアラニンに対する Km 値 203 μ M，Kcat 5.26s^{-1} であった。これらの数値は他の植物から得られた PAL の値と同様であった。

キーワード：*Eucalyptus robusta*，発現，クローニング，フェニルアラニンアンモニアリアーゼ（PAL）