Molecular cloning and characterization of three phenylalanine ammonia-lyase genes from *Schisandra chinensis*

FAN San-Peng, CHEN Wei, WEI Jiang-Chun, GAO Xiao-Xu, YANG Yong-Cheng, WANG An-Hua, HU Gao-Sheng*, JIA Jing-Ming*

School of Traditional Chinese Materia Medica, Shenyang Pharmaceutical University, Shenyang 110016, China

Available online 20 Jul., 2022

[ABSTRACT] Phenylalanine ammonia-lyase (PAL), which catalyzes the conversion from L-phenylalanine to trans-cinnamic acid, is a well-known key enzyme and a connecting step between primary and secondary metabolisms in the phenylpropanoid biosynthetic pathway of plants and microbes. *Schisandra chinensis*, a woody vine plant belonging to the family of Magnoliaceae, is a rich source of dibenzocyclooctadiene lignans exhibiting potent activity. However, the functional role of PAL in the biosynthesis of lignan is relatively limited, compared with those in lignin and flavonoids biosynthesis. Therefore, it is essential to clone and characterize the PAL genes from this valuable medicinal plant. In this study, molecular cloning and characterization of three PAL genes (ScPAL1−3) from *S. chinensis* was carried out. ScPALs were cloned using RACE PCR. The sequence analysis of the three ScPALs was carried out to give basic characteristics followed by docking analysis. In order to determine their catalytic activity, recombinant protein was obtained by heterologous expression in pCold-TF vector in *Escherichia coli* (BL21-DE3), followed by Ni-affinity purification. The catalytic product of the purified recombinant proteins was verified using RP-HPLC through comparing with standard compounds. The optimal temperature, pH value and effects of different metal ions were determined. *V_{max}, K_{cat}* and *K_{m}* values were determined under the optimal conditions. The expression of three ScPALs in different tissues was also determined. Our work provided essential information for the function of ScPALs.

(KEY WORDS) *Schisandra chinensis*; Phenylalanine ammonia-lyase; Molecular cloning and characterization; Kinetics


Introduction

*Schisandra chinensis* (Turcz.) Baill, a medicinal vine plant belonging to the family of Magnoliaceae and genus of *Schisandra*, is mainly distributed in Northeast China. The dried mature fruit of *S. chinensis* (SCF) is referred as Wuweizi in Chinese, which means that the fruit has five flavors, namely sour, bitter, sweet, spicy and salty. SCF is generally produced through scaled artificial cultivation and Liaoning Province is the most important major producing area of SCF. SCF was first recorded in Shennong’s *Classic of Materia Medica* and has been used for the treatment of chronic cough, wet dream, enuresis, chronic diarrhea, spontaneous sweating, night sweats, diabetes, palpitation and insomnia [1]. According to modern chemical and pharmacological studies, dibenzocyclooctadiene lignans are the main effective compounds in SCF and responsible for activities such as hepatoprotection [2], anti-depression [3], anti-cancer [4], anti-inflammation [5], anti-ulcer [6], anti-oxidation and detoxification [7]. Due to the chemical structure and pharmacological actions, dibenzocyclooctadiene lignans have been proved as a potent hepatoprotective agent, especially bicyclol which was screened from their derivates and developed as a hepatoprotective agent [8].

Dibenzozyoloctadiene lignan is a member of phenylpropanoid compounds, which is biosynthesized through the phenylpropanoid pathway [8], where L-phenylalanine was first deaminated by phenylalanine ammonia-lyase (PAL) to trans-cinnamic acid, followed by hydroxylation by CYP73A (C4H) to form p-coumaric acid (Fig. 1). p-Coumarate-CoA is then produced by p-coumarate-CoA ligase (4CL) and then hydroxylated to ferulyl-CoA by C3H and CCoAOMT. The produced ferulyl-CoA is converted to coniferyl alcohol by cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) [9]. After a series of catalysis, dibenzocyclooctadiene lignans are finally biosynthesized (Fig. 1). It has...
been reported in many plants and microbes that, PAL is a key and rate-limiting enzyme in the biosynthesis of phenylpropanoids [10]. Since Koukol and Conn discovered the first PAL gene from barley in 1961 [11], more and more PAL genes have been investigated in many higher plants, such as salvia [12], coffee [10], Ocimum basilicum [14], and Picrorhiza kurrooa [15], and even in some moss plants [16] and fungi [17]. In many plants, PAL protein is usually encoded by a multi-gene family [10]. For example, three PAL genes were found in salvia [18], and six were found in Populus trichocarpa [19] and Oriza sativa [20]. A total of 12 PAL genes were found in Citrullus lanatus [20].

With respect to its importance in the phenylpropanoid biosynthesis pathway, PAL has attracted increasing attentions. In the current study, rapid amplification of cDNA ends (RACE)-PCR was used, while full length cDNA sequences of three PAL genes were cloned from S. chinensis. The bioinformatic analysis of the deduced amino acid sequence concerning the homology and phylogenetic relationship, amino acid sequence, structure and physical and chemical properties were performed. Through molecular docking, the active sites of three PAL genes with L-phenylalanine in S. chinensis were analyzed, and the optimized catalytic conditions and the catalytic kinetics were determined. The tissue specific expression in nine tissues of three ScPALs were also evaluated by RT-qPCR.

**Results and Discussion**

**Cloning and multiple alignment of ScPAL genes**

The full-length sequences of PALs from S. chinensis were designated as ScPAL1–3. Multiple sequence alignment results showed that the deduced ScPAL1–3 proteins shared the maximum identities (74.8%) with the PALs of other species, namely Arabidopsis thaliana (NP181241.1), Macleay a cordata (OVA08384.1), Nelumbo nucifera (XP010261982.1), Papaver somniferum (XP02644499.1), and Tetracentron sinense (KAF8393567.1) (Fig. 2). Furthermore, a conserved motif “GTITASDLVPLSY1AG” (Ala-Ser-Gly) was found in the three ScPALs shown in red rectangle. The presence of the Ala-Ser-Gly (ASG) motif in the active site residues is known for substrate binding and catalysis of MIO (3,5-dihydro-5-methylidine-4H-imidazole-4-one) autocatalytic domain formation [21]. Another conserved residue “FL” was also found in the three ScPALs shown in yellow rectangle. The “FL” residue is important for the substrate specificity of PAL enzyme [22, 23]. The motif tag “ASG” and “FL” residues presented in PAL protein are important for the consumption of L-phenylalanine as the sole substrate in ScPALs. Therefore, multiple alignment and conserved residue analysis revealed that all three ScPALs may possess the activity of catalyzing L-phenylalanine to form trans-cinnamic acid.

**Prediction of protein domains and physicochemical properties**

In the bioinformatic analysis of deduced ScPALs amino acid sequence, three ScPALs shared high sequence and structure conservation with known functional PAL proteins. The InterProscan analysis revealed that the three ScPALs protein structures contained three functional domains, namely furmarase/histidase (IPR024083), L-aspartase-like (IPR008948) and phenylalanine-ammonia-lyase (IPR023144). Furthermore, it was indicated that the three ScPALs possessed the activity catalyzing L-phenylalanine to form trans-cinnamic acid. We also analyzed the physicochemical properties of the three proteins on the bioinformatics website ExPASy. The results showed that the molecular weights of ScPAL1, ScPAL2, and ScPAL3 were 77.26, 84.15, and 78.61 kDa, respectively. The theoretical isoelectric point (pI) of these proteins ranged from 6.07 to 6.29. The instability index (II) ranged from 35.76 to 40.78 and the estimated half-life was 30 h, confirming that the deduced ScPALs were a stable protein. Furthermore, the average hydrophilicity was found to be −0.113 to −0.217, indicating that ScPAL proteins did not form transmembrane structure (Table S1).

**Three dimensional analysis of deduced ScPAL protein**

In order to clarify the structure and function of deduced ScPALs, a comparative analysis of 3D models of ScPALs was performed at ExPASy using SWISS-MODEL. The 3D structure of Petroselinum crispum PAL protein (PDB No.
was used as the template, while homology modeling was performed using SWISS-MODEL (http://www.swissmodel.expasy.org). Using the online tool PDB sum, it was found that the deduced ScPAL1 protein consisted of \( \alpha \)-helices (54.69\%), \( \beta \)-helices (6.57\%), \( \gamma \)-turns and random coils (31.19\%), and extended strands (7.55\%). The ScPAL2 protein consisted of \( \alpha \)-helices (53.65\%), \( \beta \)-helices (5.93\%), \( \gamma \)-turns and random coils (31.40\%), and extended strands (9.02\%). The ScPAL3 protein consisted of \( \alpha \)-helices (57.36\%), \( \beta \)-helices (5.83\%), \( \gamma \)-turns and random coils (30.00\%), and extended strands (6.81\%). Based on the earlier report on PAL protein of P. crispum [24], the structure of PALs were assumed to be ‘sea horse’, which was comprised of a precursor binding domain (4-methylidene-imidazole-5 one; MIO), a core domain and an inserted shielding domain. The docking results revealed that the deduced ScPAL1 protein possessed three amino acid residues (Gly-114, Asn-261, and Asn-385) connected by hydrogen bonds; ScPAL2 protein possessed four amino acid residues (Asn-264, Tyr-355, Arg-358, and Asn-388) connected by hydrogen bonds, and ScPAL3 protein possessed four amino acid residues (Asn-264, Tyr-355, Arg-358, and Asn-388) connected by hydrogen bonds surrounding L-phenylalanine (Fig. 3). These docking results were similar to the residues by which the model...
protein interacts with L-phenylalanine\(^{[24]}\). Overall, these results implied that the deduced ScPALs have similar functions as those found in the PAL proteins of other plant species.

**Phylogenetic analysis of ScPAL protein**

To investigate the evolutionary relationship between ScPALs and other PAL proteins, a set of PAL amino acid sequences from other plants were used, and a phylogenetic tree was constructed by the neighbor-joining method in MEGA 7.0 program (Fig. 4). In this phylogenetic tree, most plant PAL proteins were divided into four branches, namely dicots, monocots, gymnosperms and bryophytes. All the three ScPALs were clustered into the dicot group. ScPAL1 and ScPAL3 were clustered into one node. In this subgroup, ScPAL1 and ScPAL3 had a closer relationship with CsPAL from *Cucumis sativus*, which is a known functional PAL gene\(^{[25]}\). However, ScPAL2 was greatly different from the other two ScPALs.

**Heterologous expression in *E. coli* and activity of the recombinant ScPAL proteins**

Multiple sequence alignment and 3D model analysis showed that ScPALs had a high degree of sequence identity and structural similarity to other plant functional PAL proteins. In order to explore their functional activities, ScPALs were constructed into pColdIF vector and the recombinant

---

**Fig. 3** Molecular modeling of ScPAL1–3 and L-phenylalanine (A1, ScPAL1; A2, ScPAL2; and A3, ScPAL3)

**Fig. 4** Phylogenetic analysis of PALs from *S. chinensis* and other plant species by MEGA 7 from ClustalW alignments. The neighbor-joining method was used to construct the tree with p-distance. The number for each interior branch was the percent bootstraps value (1000 replicates)
proteins were heterologously expressed in *E. coli* BL21 (DE3) under the optimal incubation conditions at 16 °C for 24 h and 1 mmol·L⁻¹ IPTG and purified by Ni-affinity chromatography. The TF tag was fused to the N-terminus of ScPALs to increase the solubility of the recombinant protein. According to SDS-PAGE analysis of total crude protein, the recombinant ScPALs were mainly found in the supernatant, with the molecular weight (including 48 KD tags) of about 125 kDa (Fig. 5A), which was consistent with the theoretical value. In the enzyme activity assay, the characteristic trans-cinnamic acid peak was found at the retention time of around 24.5 min with ScPALs, which was be monitored in the control group (Fig. 5B). These results showed that the recombinant ScPALs are functional enzymes that convert L-phenylalanine to trans-cinnamic acid.

**Enzyme properties and kinetic constant determination**

The effect of pH value on the ScPAL activity with L-phenylalanine was investigated between 3.0 and 11.0. As shown in Fig. 6, A1–A3, the specific activity of ScPALs notably increased with the increasing of pH value from 3.0 to 8.0, and decreased from pH 9.0 to 11.0. All three ScPALs possessed higher catalytic activity under alkaline conditions (pH 7–11) than acidic conditions (pH 3–6). When the pH value increased from 6 to 8, the catalytic ability of all three ScPALs sharply increased; when the pH value further increased, the catalytic ability decreased. The optimal pH for ScPALs ranged 7.5–8.0. It was the same as the optimal pH for two recently reported PALs from *Pseudezyma antarctica* [28]. The effects of metal ions on the ammonia-lyase activity of ScPALs were determined. As shown in Fig. 6, B1–B3, Cu²⁺, Ag⁺, and Fe²⁺ exerted strong inhibitory effects on the catalytic activity of ScPALs, but Mn²⁺ and Mg²⁺ exhibited auxiliary effects. These results indicated that the ScPALs activity were influenced by the presence of certain metal ions. The effect of temperature (15–65 °C) on the catalytic reaction was evaluated using the purified recombinant ScPALs (Figs. 6C1–6C3). The optimal temperature for the catalytic activity of these ScPALs were 45, 45 and 50 °C respectively.

In order to further understand the catalytic ability of ScPALs, the catalytic parameters of reported PALs in different plants are listed in Table 1. The kinetics of ammonia elimination by ScPAL was characterized using L-phenylalanine as the substrate (Figs. 7A1–7A3). The $K_m$ value of L-phenylalanine with ScPAL1 (0.17 mmol·L⁻¹), ScPAL2 (0.25 mmol·L⁻¹) and ScPAL3 (0.21 mmol·L⁻¹) were similar to those found with the closely related EsPAL1 (0.15 mmol·L⁻¹) from *Ephedra sinica* [25] and CdPAL1 (0.10 mmol·L⁻¹) from *Cistanche deserticola* [26], higher than those with PsPAL1 (0.017 mmol·L⁻¹) from *P. crispum* [29] and AtPAL1 (0.068 mmol·L⁻¹) from *A. thaliana* [30], but lower than those with TcPAL (1.10 mmol·L⁻¹) from *Taxus chinensis* [31] and BoPAL4 (2.07 mmol·L⁻¹) from *Bambusa oldhamii* [32].

**Tissue-specific expression analysis of ScPALs**

In order to examine the expression of ScPALs in different tissues, *GAPDH* gene was cloned using RACE PCR and used as an internal reference gene, while qRT-PCR was performed with gene-specific primers (Table S2). Quantitative PCR analysis was performed using ScPAL genes in nine different tissues of *S. chinensis*, including leaf, petiole, stem xylem, stem phloem, root xylem, root phloem, pericarps, fruit stem, and seed (Fig. 8). The expression of ScPAL genes in different tissues were obviously different, and each gene has a tissue-specific pattern. Among them, all three PAL genes had the highest expression in steam xylem, with the lowest expression in pericarps. As we known, PAL is the first enzyme in the phenylpropanoid biosynthetic pathway and the first enzyme to synthesize lignin, and PAL had the highest expression in tissues rich in lignin. The expression of ScPAL genes were higher in leaves and petioles. Among three ScPAL genes, the expression of ScPAL1 in the nine tissues was relatively low, and the expression of ScPAL2 and ScPAL3 were 2.5 and 2.3 times that of ScPAL1, respectively.
Conclusion

In this study, molecular cloning and characterization of three full-length PAL genes from *S. chinensis* were carried out. Multiple sequence alignment, 3D modeling, and phylogenetic tree construction were performed, and the results demonstrated that ScPAL1–3 showed possible PAL activity. The recombinant proteins of ScPAL1–3 were heterologously expressed and purified. The catalytic product was identified by RP-HPLC and the effects of temperature, pH and metal ions were investigated. Finally, the transcription levels of *ScPAL*1–3 were determined in different tissues. Our results demonstrated that ScPAL1–3 can catalyze the reaction from *L*-phenylalanine to *trans*-cinnamic acid. The optimal temperature of recombinant ScPALS was 45–50 °C and the optimal pH was 7.5–8.0. Cu$^{2+}$, Ag$^{+}$, and Fe$^{3+}$ significantly inhibited the ScPAL activity, while Mg$^{2+}$ improved the ScPAL activity. The $K_{m}$ value of *L*-phenylalanine with ScPAL1 (0.17 mmol·L$^{-1}$) was smaller than ScPAL2 (0.25 mmol·L$^{-1}$) and ScPAL3 (0.21 mmol·L$^{-1}$). The expression of *ScPAL* genes was the highest in steam xylem and lowest in pericarps.

SCF is a famous traditional Chinese medicinal, with significant pharmacological activity and prominent application prospects. However, the current research mainly focuses on the characterization of chemical components and pharmacological activity, with few reports concerning its biosynthesis. The study of PAL in *S. chinensis* will not only assist to further understand its biosynthetic pathway, but also act as a stepping stone to improve the pharmaceutical and biotechnological applications of this valuable medicinal plant in the future.

Material and methods

**Plant materials**

Plant samples were collected from Fengcheng County, Liaoning Province, and authenticated as *Schisandra chinensis* by Prof. JIA Jing-Ming in School of Traditional Chinese Material Medica, Shenyang Pharmaceutical University (Shengyang, China). Plant specimen were preserved in School of Traditional Chinese Material Medica, with specimen number: P-20200921. The collected plant materials were separated into different tissue groups, including leaf, petiole, stem xylem, stem phloem, root xylem, root phloem, fruit stem, pericarp and seed, immediately frozen in liquid nitrogen, and then stored at −80 °C until later use.

**Cloning of ScPALS and GAPDH cDNA**

Total RNA was extracted from frozen tissue using Fruitmate (Takara, DaLian, China) and RNAiso Plus (Takara, DaLian, China) according to the manufacturer’s instructions. First strand cDNA synthesis was performed with the PrimeScript<sup>II</sup> Master Mix (Takara, DaLian, China). The full length cDNA sequences of *ScPAL* genes were amplified by
Table 1  The catalytic parameters of PALs in different plants

<table>
<thead>
<tr>
<th>Plant</th>
<th>Gene name</th>
<th>$K_m$/(μmol·L$^{-1}$)</th>
<th>$V_{max}$/(μmol·mg$^{-1}$·min$^{-1}$)</th>
<th>$K_{cat}$/(s$^{-1}$)</th>
<th>$K_{cat}/K_m$ (s$^{-1}$·mmol·L$^{-1}$)</th>
<th>Optimal pH</th>
<th>Optimal $T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schisandra chinensis</td>
<td>ScPAL1</td>
<td>174.1</td>
<td>0.66</td>
<td>0.85</td>
<td>4.87</td>
<td>8</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>ScPAL2</td>
<td>253.8</td>
<td>0.52</td>
<td>0.67</td>
<td>2.65</td>
<td>7.5</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>ScPAL3</td>
<td>212.6</td>
<td>0.77</td>
<td>0.99</td>
<td>4.65</td>
<td>8</td>
<td>50</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>AtPAL1</td>
<td>68</td>
<td>1.80</td>
<td>26.47</td>
<td>8.4–9.2</td>
<td>46–48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AtPAL2</td>
<td>64</td>
<td>3.20</td>
<td>50.00</td>
<td>8.4–8.9</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AtPAL3</td>
<td>2560</td>
<td>0.10</td>
<td>0.04</td>
<td>8.7–8.9</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AtPAL4</td>
<td>71</td>
<td>3.00</td>
<td>42.25</td>
<td>8.4–8.8</td>
<td>46–48</td>
<td></td>
</tr>
<tr>
<td>Nicotiana tabacum</td>
<td>NtPAL1</td>
<td>59.8</td>
<td>1.09</td>
<td>18.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NtPAL2</td>
<td>39.5</td>
<td>1.14</td>
<td>28.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NtPAL3</td>
<td>36.4</td>
<td>0.78</td>
<td>21.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NtPAL4</td>
<td>52.4</td>
<td>1.53</td>
<td>29.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Petroselinum crispum</td>
<td>PsPAL1</td>
<td>17.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PsPAL2</td>
<td>16.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PsPAL3</td>
<td>24.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PsPAL4</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taxus chinensis</td>
<td>TcPAL</td>
<td>1100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bambusa oldhamii</td>
<td>BoPAL1</td>
<td>230</td>
<td>16.29</td>
<td>70.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BoPAL2</td>
<td>993</td>
<td>21.30</td>
<td>21.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BoPAL4</td>
<td>2072</td>
<td>16.32</td>
<td>7.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lycoris radiata</td>
<td>LrPAL3</td>
<td>510</td>
<td></td>
<td></td>
<td></td>
<td>8.5–9.5</td>
<td>35–40</td>
</tr>
<tr>
<td>Cistanche deserticola</td>
<td>CdPAL1</td>
<td>101.3</td>
<td>3.36</td>
<td>33.17</td>
<td>8.5</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EsPAL1</td>
<td>152</td>
<td>0.59</td>
<td>3.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ephedra sinica</td>
<td>EsPAL2</td>
<td>144</td>
<td>0.46</td>
<td>3.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EsPAL3</td>
<td>112</td>
<td>0.36</td>
<td>3.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EsPAL4</td>
<td>180</td>
<td>0.52</td>
<td>2.87</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 7  Enzyme kinetics of ScPAL1 (A1), ScPAL2 (A2), and ScPAL3 (A3) calculated from Michaelis–Menten equation. Data are expressed as mean ± SD, and error bars represent the average value of three parallel experiments.
(Vazyme, Nanjing, China), before sequencing (GENEWIZ, TianJin, China). Gene specific primers were designed based on the new sequences and the full length sequences of ScPALs were amplified followed by T vector construction, transformation and sequence verification. All primers were designed with Primer Premier 5.0 software (Premier, BC, Canada) and the sequences of ScPAL1–3 were submitted to Genbank under the accession numbers ScPAL1 (OK357444), ScPAL2 (OK357443), and ScPAL3 (OK357445). ScGAPDH gene was cloned following the procedure mentioned above. The multiple sequence alignment of GAPDH from other plants is shown in Fig. S2. Primers are listed in Table S2 and the gene sequence was submitted to Genbank under the accession number OK467969.

Sequencing, phylogenetic analysis and molecular modeling of ScPALs

The assembled full-length cDNA sequence of ScPAL genes were analyzed using ORF Finder tool (http://www.ncbi.nlm.nih.gov/gorf/gorf) to generate the complete coding regions of ScPAL proteins. The deduced ScPALs were aligned with other PAL sequences retrieved from NCBI database for multiple sequence alignment with ClustalW (http://www.ebi.ac.uk/Tools/clustalw/index). A phylogenetic tree was constructed using the MEGA 7.0 with the neighbor-joining (NJ) method. The SWISS-MODEL (http://swissmodel.expasy.org) was performed to predict the 3D structure of ScPALs and the L-phenylalanine ligand using the phenylalanine ammonia-lyase from P. crispum (PDB: 1w27) as a template. To predict the binding site and binding energy between ScPALs and L-phenylalanine, docking was also performed with Schrodinger 11.1 program. Protein structure was analyzed by PDBsum program (https://www.ebi.ac.uk/pdbsum). The physicochemical properties, protein functional site, and hydrophobicity index of ScPALs protein were analyzed with online tools including ProtParam (https://www.expasy.org/tools/protparam), InterProscan (https://www.ebi.ac.uk/tools/pfa/iprscan) and ProtScale (https://www.expasy.ch/tools/protscale).

Heterologous expression and purification of ScPALs in E. coli BL21 (DE3)

The coding sequence of ScPALs were amplified by homologous primers (Table S2) and inserted into the pCold-TF expression vector (Novagen, Madison, WI, USA) digested by Kpn I and BamH I using the ClonExpress Ultra One Step Cloning Kit (Vazyme, NanJing, China). The positive recombinant plasmids, pCold-TF-ScPALs were confirmed by sequencing and then transformed into E. coli BL21 (DE3) strain for protein expression. The transformants were inoculated at 37 °C in LB medium until OD600 reaching about 0.6. Then, a final concentration of 1 mmol·L⁻¹ isopropyl β-D-1-thiogalactopyranoside (IPTG) was added into the cultures to induce protein expression at 16 °C for 24 h. The cells were collected by centrifugation at 5000 g at 4 °C for 5 min, then resuspended in PBS (50 mmol·L⁻¹ phosphate buffered saline, pH 7.4), and lysed by ultrasound in an ice bath for 5 min. The total protein was subjected on 10% SDS-PAGE followed by Coomassie Brilliant Blue R250 staining, so as to evaluate the expression of the induced recombinant proteins. Affinity chromatography was used to purify recombinant ScPAL proteins with 6 × His-tag fused to the N-terminus of ScPALs. Specifically, a 1 mL Ni column (GE Healthcare, Pittsburgh USA) was equilibrated with 20 column volumes of PBS, followed by lysate loading. The column was washed with 20 column volume of PBS, and the target protein with eluted by washing buffer (PBS, 500 mmol·L⁻¹ imidazole, pH 7.4). The purity of the purified recombinant was determined using SDS-PAGE and the concentration of the purified recombinant proteins was determined using Bradford method following the manufacturer’s instructions (Solarbio, Beijing, China).

Enzyme activity assay

The catalyzing activity was determined using a previously reported method [29] with modifications. Specifically, the reaction system contained 1 mL 0.01 mol·L⁻¹ PBS (pH 7.5), 100 μL crude protein extract and 0.02 mol·L⁻¹ L-phenylalanine, while crude enzyme solution was inactivated in a boiling water bath for 5 min and set as the control. The reactions were incubated at 30 °C for 30 min and an equal volume of ice acetonitrile was added to terminate the reaction. The reaction product was filtered through a filter with 0.45 μm diameter micro pores. RP-high-performance liquid chromatography (an Agilent 1260 system equipped with a photo-diode array, Palo Alto, CA, USA) was used to analyze the reaction product. The sample (10 μL) was eluted through column chromatography (C18, 5 μm, 4.6 mm × 250 mm; Agi-
lent, Palo Alto, CA, USA) in the following gradients: methanol (solvent A) and 0.1% formic acid in water (solvent B): 0 min 10% A; 30 min, 90% A; 30–40 min, 90% A; and 40–50 min, 10% A. The flow rate was 0.8 mL min⁻¹ and the detection wavelength was 220 nm. The retention times of reaction product and substrate were compared with those of the standard compounds of trans-cinnamic acid and L-phenylalanine.

**Enzyme properties and kinetic constants of ScPALs**

The effects of temperature on enzyme activity were investigated by measuring ScPAL activity in PBS buffer in a temperature range of 15–65 °C. Three buffer systems (50 mmol·L⁻¹ citric acid-sodium citrate pH 3.0–6.0, 50 mmol·L⁻¹ PBS pH 6.0–8.0, and 50 mmol·L⁻¹ sodium carbonate-sodium hydroxide pH 9–11) were used to identify the optimal pH value for enzyme activity. To assess the effects of metal ions (Na⁺, K⁺, Cu²⁺, Zn²⁺, Al³⁺, Fe²⁺, Fe³⁺, Cu²⁺, Ag⁺, and Mn²⁺) on ScPAL activity, the ScPALs solution was incubated with these metal ions (1 mmol·L⁻¹) at 4 °C for 24 h. Then, L-phenylalanine was added to the reaction system at a final concentration of 0.1 mmol·L⁻¹. In order to measure the kinetic constants, the catalytic product was detected at 290 nm using a spectrophotometer (Varioskan Flash, Thermo Fisher Scientific, Massachusetts, USA) with the substrate L-trans-cinnamic acid and L-phenylalanine.

**Quantitative real-time PCR (qRT-PCR)**

In order to analyze the transcription level of ScPALs in various tissues of *S. chinensis*, the primers of individual ScPALs were designed based on the corresponding genes using Primer5 software. Quantitative primers and internal reference gene primers used for the study are listed in Table S2. qRT-PCR amplification assay was performed using a real-time PCR system (Bio-Rad, CF96), in an ultimate 20 μL reaction system containing 8 μL EvaGreen Master Mix (Gene-wiz, NanJing, China), 1.5 μL of diluted cDNA, 0.25 μmol·L⁻¹ of each primer before the addition of double distilled water. qPCR was done with a temperature profile of 5 min at 95 °C, 35 cycles of 15 s at 96 °C, 20 s at the specific annealing temperature for each primer, and 20 s at 72 °C. The 2⁻ΔΔCt method was used to analyze the relative transcription level after PCR was completed. For each group, three replicates were performed.

**Supplementary Materials**

All the supporting information of this paper can be requested by sending E-mails to the corresponding authors.

**References**


