ORIGINAL ARTICLE

Two types of coumarins-specific enzymes complete the last missing steps in pyran- and furanocoumarins biosynthesis

Yucheng Zhao[a,†], Yuedong He[b,†], Liangliang Han[c,†], Libo Zhang[a], Yuanzheng Xia[c], Fucheng Yin[c], Xiaobing Wang[c], Deqing Zhao[d], Sheng Xu[e], Fei Qiao[d,*], Yibei Xiao[f,*], Lingyi Kong[c,*]

[a]Department of Resources Science of Traditional Chinese Medicines, School of Traditional Chinese Pharmacy, and State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing 210009, China
[b]College of Bioscience and Biotechnology, Hunan Agricultural University, Changsha 410128, China
[c]Jiangsu Key Laboratory of Bioactive Natural Product Research and State Key Laboratory of Natural Medicines, Department of Natural Medicinal Chemistry, China Pharmaceutical University, Nanjing 210009, China
[d]Key Laboratory of Crop Gene Resources and Germplasm Enhancement in Southern China, Ministry of Agriculture, Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou 517317, China
[e]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing 210014, China
[f]Department of Pharmacology, School of Pharmacy, and State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing 210009, China

Abstract Pyran- and furanocoumarins are key representatives of tetrahydropyrans and tetrahydrofurans, respectively, exhibiting diverse physiological and medical bioactivities. However, the biosynthetic mechanisms for their core structures remain poorly understood. Here we combined multiomics analyses of biosynthetic enzymes in *Peucedanum praeruptorum* and *in vitro* functional verification and identified two types of key enzymes critical for pyran and furan ring biosynthesis in plants. These included three distinct *P. praeruptorum* prenyltransferases (PpPT1−3) responsible for the prenylation of the simple coumarin skeleton **7** into linear or angular precursors, and two novel CYP450 cyclases (PpDC and PpOC) crucial for the cyclization of the linear/angular precursors into either tetrahydropyran or tetrahydrofuran scaffolds. Biochemical analyses of cyclases indicated that acid/base-assisted epoxide ring opening

contributed to the enzyme-catalyzed tetrahydropyran and tetrahydrofuran ring refactoring. The possible acid/base-assisted catalytic mechanisms of the identified cyclases were theoretically investigated and assessed using site-specific mutagenesis. We identified two possible acidic amino acids Glu303 in PpDC and Asp301 in PpOC as vital in the catalytic process. This study provides new enzymatic tools in the epoxide formation/epoxide-opening mediated cascade reaction and exemplifies how plants become chemically diverse in terms of enzyme function and catalytic process.

1. Introduction

Pyranocoumarins (PCs) and furanocoumarins (FCs) are widely distributed in the plant kingdom and exhibit diverse physiological and medical bioactivities[1−4]. For instance, FC compounds xanthotoxin and bergapten are commonly used clinical drugs approved by the FDA for skin diseases, and the structural diversity of these coumarins has enabled their development as broad-spectrum antifungals that can circumvent multidrug resistance[3]. However, there has been little progress in understanding the PCs and FCs biosynthetic mechanisms in plants, despite numerous chemosynthetic methods, including Prins cyclization, hetero-Diels−Alder reactions, and ring closing metatheses, being developed to synthesize pyran and furan rings[5].

Coumarins possess a common 1,2-benzopyrone skeleton and are generally subclassified as simple coumarins (SCs), linear FCs/PCs, and angular FCs/PCs. Linear/angular FCs and PCs all have an additional furan and pyran rings, respectively, and hence they also belong to complex coumarins[1,3]. Prenylation on coumarin skeleton **7** and subsequent cyclization are considered crucial steps among all catalytic steps involved in FC and PC biosynthesis as they determine the formation of the four types of FCs and PCs (Fig. 1)[3]. For instance, prenylation catalyzed by prenyltransferases (PTs) determine the linear (**7 → 8**) or angular (**7 → 9**) type of coumarins, which creates a key precursor/intermediate for the subsequent cyclization[6−9]. However, only a limited number of PTs have been functionally verified[3]. The cyclization step establishes the basic skeleton of furan and pyran rings, and it is an initial step in FC and PC biosynthesis[1]. Alongside PTs, at least four types of coumarin skeletons (**10−13**) can be generated. Nevertheless, to date, only one cyclase responsible for linear FC biosynthesis has been cloned in the Moraceae family[10], and genes responsible for linear PCs and angular FCs/PCs biosynthesis as well as mechanisms for the cyclization of different types of coumarins remain unknown. Furthermore, the formation of 6-*endo*-tet PCs is disfavored according to Baldwin's rules, the underlying mechanism of the formation of such six-membered rings in plants remains to be clarified[11].

Herein, we identified key enzymes in determining linear/angular FC and PC scaffolds and those completing the last missing steps in coumarin biosynthesis in plants. An acid/base-assisted enzymatic catalysis was evaluated and shown to influence the chemical distribution of PC and FC cyclic products. Furthermore, we found that the identified cyclases can catalyze the disfavored 6-*endo*-tet cyclization under acidic conditions. The findings of this study contribute to knowledge regarding the classical Baldwin's rules and provide useful enzymatic tools for PCs and FCs biosynthesis.

2. Materials and methods

2.1. Alignment and phylogenetic analyses of PT and CYP450 gene family

Amino acid sequences of the PT and CPY450 proteins were aligned using ClustalW with default options and then manually corrected in MEGA 7.0[12]. The resulting amino acid sequence alignments were used to guide the alignments of the nucleotide coding sequences. An unrooted maximum likelihood phylogenetic tree was constructed by IQ-TREE[13], the value of SH-aLRT test was 1000 (-alrt 1000), the value of Ultrasfast bootstrap was 1000 (-bb 1000), and the default values of other running parameter selection programs were used. The Newick format output file was used to visualize the phylogenetic tree online with the Figtree (http://tree.bio.ed.ac.uk/software/figtree) and iTOL (https://itol.embl.de/)[14]. In the end, the members of CYP71 Clan were extracted from the complete phylogenetic tree.

2.2. Correlation analysis

Correlation analysis refers to the analysis of two or more elements of variables that are correlated in order to measure the correlation closeness of different variables. The correlation between variables can be expressed by the simple correlation coefficient, the closer the correlation coefficient is to 1, the more correlated the two elements are, and the closer the correlation coefficient and 0 are, the more independent the two elements are. There are three main methods to measure the correlation of random variables: Pearson correlation coefficient, Spearman correlation coefficient, and Kendall correlation coefficient. We used Pearson correlation coefficient to analyze the correlation between target genes and candidate genes according to the previously report[15]. It is performed in R language using the cor function.

2.3. Heterologous expression of candidate PTs in Escherichia coli and activity assays

All the PTs were amplified by PCR using their corresponding gene-specific primers and then cloned into the BamH I/EcoR I site of pETDuet-1 vector using the primers listed in Supporting Information Table S3. All the recombinant plasmids were inserted into *E. coli* competence cells, respectively, and cultured in Luria-Bertani medium at 37 °C, 220 rpm until the OD$_{600}$ reached 0.6−0.8. Then, the cultures were cooled down to 25 °C and afterwards, IPTG in a final concentration of 500 μmol/L was added for inducing protein expression overnight. The bacteria
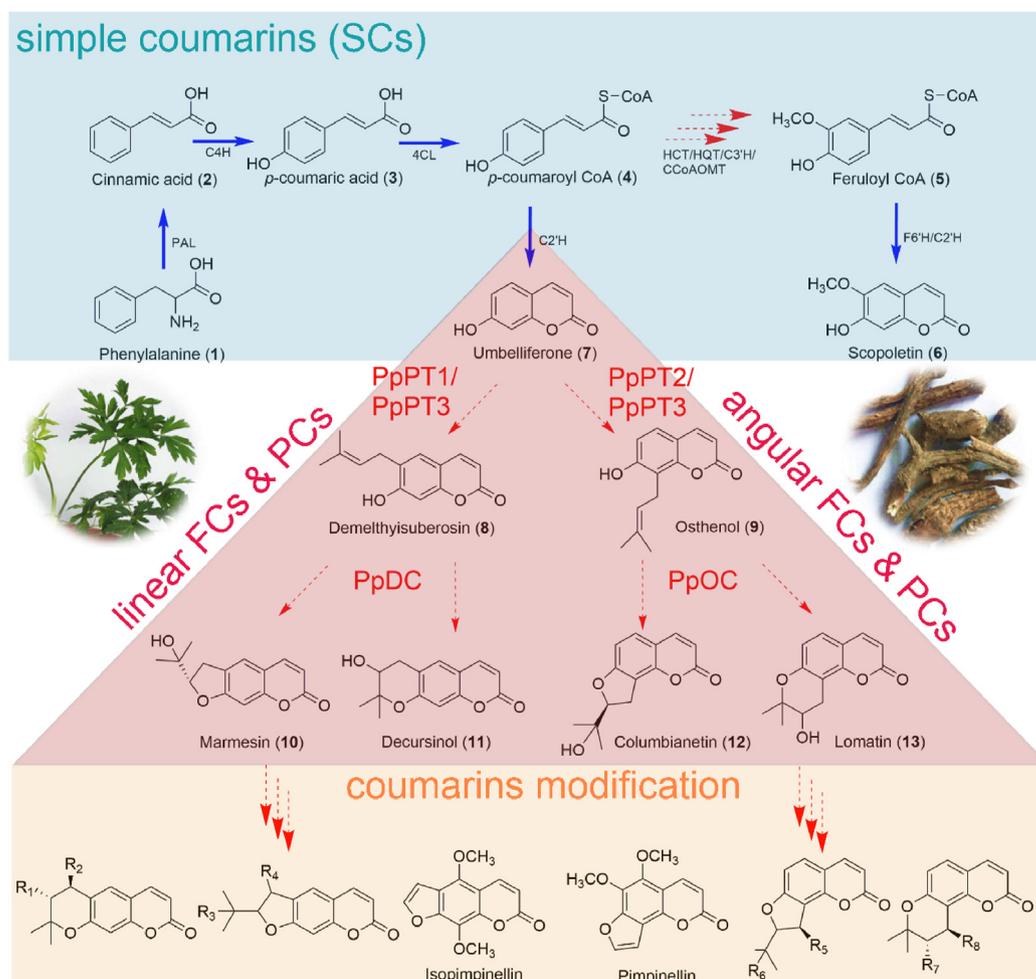
Figure 1 Summary of predicted coumarins biosynthetic pathway. SCs, linear/angular FCs and PCs and structurally modified coumarins were coated in different colors. Two pictures placed in two wings were to display the plant used in this study (*P. praeruptorum*) and its medicinal part (dry root). SCs, simple coumarins; FCs, furanocoumarins; PCs, pyranocoumarins; PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate: coenzyme A ligase; HCT/HQT, hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyl transferase; CCoAOMT, caffeoyl-CoA O-methyltransferase; C3′H, cinnamoyl ester 3′-hydroxylase; C2′H, *p*-coumaroyl CoA 2′-hydroxylase; F6′H, feruloyl-CoA 6′-hydroxylase; PpPT1, umbelliferone 6-prenyltransferase (U6PT); PpPT2, umbelliferone 8-prenyltransferase (U8PT); PpPT3, umbelliferone 6/8-prenyltransferase (U6,8PT); PpDC, demethyisuberosin cyclase; PpOC, osthenol cyclase. R1−R8 represent different substituent groups.

cells were collected by centrifugation (4000 rpm for 10 min), and resuspended in lysis buffer (20 mmol/L HEPES, 500 mmol/L NaCl, 10% *v/v* glycerol, pH 7.5). Suspended bacteria are crushed by ultrasonic on ice. Since the PTs are membrane-bound protein[6−9], it is difficult to obtain its purified protein, so we used the crude proteins to analyze their activities.

All the enzymatic activity assays were carried out on a shaking incubator (220 rpm) at 25 °C for 2 h. The reaction system contained 50 μL crude enzyme, 200 μmol/L umbelliferone **7**, 100 μmol/L DMAPP, and 200 μmol/L MgCl$_2$ in 200 μL of 100 mmol/L Tris-HCl, pH 7.5. To test the substrate specificity of different PTs, different substrates in a final concentration of 100 μmol/L were added to the reaction broth. The reaction temperature and pH, as well as the metal ions were also optimized by initiating the reactions at different temperatures (15−50 °C) and metal ions addition (Fe^{2+}, Fe^{3+}, Ca^{2+}, Mg^{2+}, Zn^{2+}, Cu^{2+}, Mn^{2+}, Co^{2+}, Ag$^+$, and Ni^{2+} in 10 μmol/L)[16]. Reactions were stopped by the addition of three times of methanol. Three independent reactions were conducted and the crude protein produced by

empty vector was used as the negative control. And then the supernatant was centrifuged at 15,000 rpm for 10 min, and 10 μL supernatant was analyzed by HPLC and MS.

2.4. Heterologous expression of CYPs in yeast and activity assays

Each candidate CYP (including the mutants) was cloned into the BamH I/EcoR I sites of yeast expression vector pYES2.0. Subsequently, the reconstructed plasmid pYES2-CYPs were transformed into *Saccharomyces cerevisiae* WAT11. The positive transformants that were screened on solid plates of SC-U (SC dropout medium without uracil) containing 20 g/L glucose were verified by microbial PCR and sequencing. Then, the positive clones were incubated with a shake at 30 °C until the OD$_{600}$ reached 2−3. After that, the cells were centrifuged at 4000 rpm for 10 min and washed at least three times with ddH$_2$O to remove glucose residue. The cell precipitates were resuspended in an SC-U medium containing 20 g/L galactose for 6−12 h at 30 °C to induce the expression of the target

protein. For preliminary activity screening, demethylsuberosin **8** and osthenol **9** were added into the cultures with a final concentration of 100 μmol/L. After incubation at 28 °C for 2 h, the cultures were collected and added three times of methanol to terminate the reaction, and then ultrasound twice for 30 min. The supernatants were collected for HPLC−MS analysis. For whole cell catalysis, the cultures were collected and resuspend with buffers in different pH values. For further activity assays, microsomes were prepared according to the previous report[17]. Without special statement, all the enzymatic activity assays were carried out on a shaking incubator (220 rpm) at 30 °C for 2 h and contained 500 μmol/L NADPH in 200 μL of 100 mmol/L Tris-HCl, pH 7.5. Reactions were stopped by the addition of three times of methanol and then used for HPLC−MS analysis. Three independent reactions were conducted and the empty pYES2.0 vector produced microsomal preparation was used as the negative control reaction. To test the substrate specificities of different CYPs, different substrates in a final concentration of 100 μmol/L were added to the reaction broth. The reaction temperature and pH, as well as the metal ions were also optimized the same as that in PTs.

2.5. HPLC−MS analysis

All the reactions are analyzed using 0.1% HCOOH aqueous solution−methanol in a 0.5 mL/min liner elution conditions: 0 min: 70:30 (*v/v*), 5 min: 35:65 (*v/v*), 12 min: 95:5 (*v/v*), 14 min: 95:5 (*v/v*), 16 min: 70:30 (*v/v*), 22 min: 70:30 (*v/v*). The detection wavelength for osthenol, demethylsuberosin, marmesin, decursinol, columbianetin, and lomatin is 335 nm, with Shimadzu LC-2010AT. For MS analysis, Agilent Poroshell 120 SB-Aq (3.0 mm × 150 mm, 2.7 μm) was used. It also using 0.1% HCOOH aqueous solution-methanol as mobile phase. The liner elution conditions are 0 min: 85:15 (*v/v*), 3 min: 85:15 (*v/v*), 8 min: 20:80 (*v/v*), 12 min: 5:95 (*v/v*), 16 min: 85:15 (*v/v*), 18 min: 85:15 (*v/v*). The conditions of the ESI source were as follows: drying gas (N$_2$) flow rate, 8.0 L/min; Collision energy, 35 eV; Spray voltage, 3.5 kV; Capillary temperature, 320 °C; Aux gas heater temp, 300 °C; Sheath gas flow rate, 35 arb; Aux gas flow rate, 15 arb; Sweep gas flow rate, 5 arb; aux gas, sheath gas, and sweep gas are all high purity N$_2$. All of the operations and analysis of data were positive ion mode.

2.6. Enzyme modeling, docking and site-directed mutagenesis

The structures of PpDC and PpOC were generated by colabalphafold2 which was released on the github by DeepMind and five structures with higher confidence level was selected[18]. Then, osthenol and demethylsuberosin, as well as heme were docked into the generated cyclase structures AutoDock[19]. Site-directed mutagenesis was conducted using a PCR method with KOD-plus-neo. Mutants were extracted from *E. coli* for sequencing and the positive plasmids were subsequently transferred to *S. cerevisiae* WAT11 for protein expression and activity test. Map drawing and molecular distances were generated using PyMOL and multiple sequence alignment was generated using MEGA.

2.7. Subcellular localization

Subcellular localization of cyclase was conducted using Arabidopsis protoplast as the method described in our previous report[20]. Firstly, the ORF of the target gene was cloned into pAN580 vector at Xba I/BamH I restriction sites for C-terminal GFP fusion. Then, we used Plant-mPLoc to predict subcellular locations of target

gene and the results indicated that the gene was located in Endoplasmic reticulum (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/#)[21]. Finally, the construct in combination with 35S: HDEL-mCherry was co-transformed into Arabidopsis leaf protoplasts using polyethylene glycol (PEG)-mediated transient method[22]. HDEL-mCherry was used as an Endoplasmic reticulum (ER) localization marker[22]. After transformation for about 16 h, the protoplasts were observed with a laser scanning confocal microscope (LSM 900; Carl Zeiss, Jena, Germany). The untransformed Arabidopsis protoplasts were used as negative control.

2.8. Molecular dynamics simulation

Molecular dynamics simulations of DC (and OC) in catalyzing compound **8** (and **9**) and its intermediate were carried out with YASARA Structure software using the standard YASARA macro md_run with AMBER14 force field. The cell was filled using the TIP3P water model at pH 7.4 and 298 K. An ionic concentration of 0.9% NaCl was used. p*K*a values were determined for each residue. The proteins were simulated in a periodic cuboid with a 10 Å boundary and an 8 Å PME cutoff value and 2.5 fs timesteps. The solvated and neutralized system was submitted to energy minimization and a production run of 300 ns at 298 K and a pressure of 1 atm (NPT ensemble), simulation snapshots were saved every 100 ps. According to the inferred mechanism, we constructed 10 complex systems in which (1) DC or OC binds to the initial substrate and forms a hydrogen bond with Glu303 or Asp301, respectively; (2) oxidative cyclization of the double bond between substrates C1 and C2 and forming a hydrogen bond with Glu303 or Asp301; (3) C5 hydroxyl protons of the substrates are transferred to Glu303 or Asp301 respectively; (4) In the catalytic sites of DC and OC, the conformation of epoxide substrates tend to generate pyran ring products; (5) the conformation of epoxide substrates tend to form furan ring construct.

The MD simulations were analyzed using the md_analyze macro in YASARA, which mainly includes interatomic distance and dihedral angle analysis. ChemDraw 16.0 has been used to draw 2D molecular schematic, and 3D structural figures were drawn using PyMOL (http://www.pymol.org).

2.9. Computational details

All structures in this work were optimized in water solvent and characterized by frequency analysis using Gaussian 16 program package and M062X density functional with the 6-31G(d,p) basis set[23]. Intrinsic reaction coordinate (IRC) was used to examine the transition state (TS) associating the corresponding minimum on the potential energy surfaces[24]. To obtain more accurate energy, single-point calculations were run at the M062X/6-311++G(d,p) (SMD,water) theoretical level. The optimized molecular structure was visualized by CYLview (2.0 version) software (CYLview20; Legault, C.Y., Université de Sherbrooke, 2020 (http://www.cylview.org)).

3. Results and discussion

3.1. Prenylation produces essential precursors for furan and pyran ring refactoring

To identify the genes involved in furan and pyran ring refactoring, we first focused on the prenylation reaction for the products essential for pyran and furan ring refactoring and that determines the formation of linear or angular pyran and furan rings[3,6−8]. We

use *P. praeruptorum* as plant material for it produces five types of coumarins and it is the primary coumarins resource in traditional Chinese medicine (Fig. 1 and Supporting Information Fig. S1)[25]. According to the genome information, seven PTs were classified as umbelliferone PTs. Remarkably, two genes (PpT1 and PpT2, corresponding to ASM_9.2469.3 and ASM_9.2434, respectively, Supporting Information Table S1) were distinctly classified as orthologs of the previously reported PTs (Fig. 2A). These two candidate genes exhibited high expression levels in all plant organs, especially in the roots (Supporting Information Fig. S2).

Such expression profiles largely corresponded to the accumulation of complex coumarins in different plant organs, implying a functional association (Fig. 2B). Consequently, we cloned these two PTs into *E. coli* to assess their activities. As a result, PpPT1 and PpPT2 were observed to display U6PT and U8PT activities, respectively, generating products equivalent to their corresponding standard compounds **8** and **9** (Fig. 2C and Supporting Information Fig. S3). PpPT1 displayed strict product-producing specificity, similar to the protein expressed by the PT gene in *Ficus carica*, whereas PpPT2 displayed more U8PT activity than U6PT activity
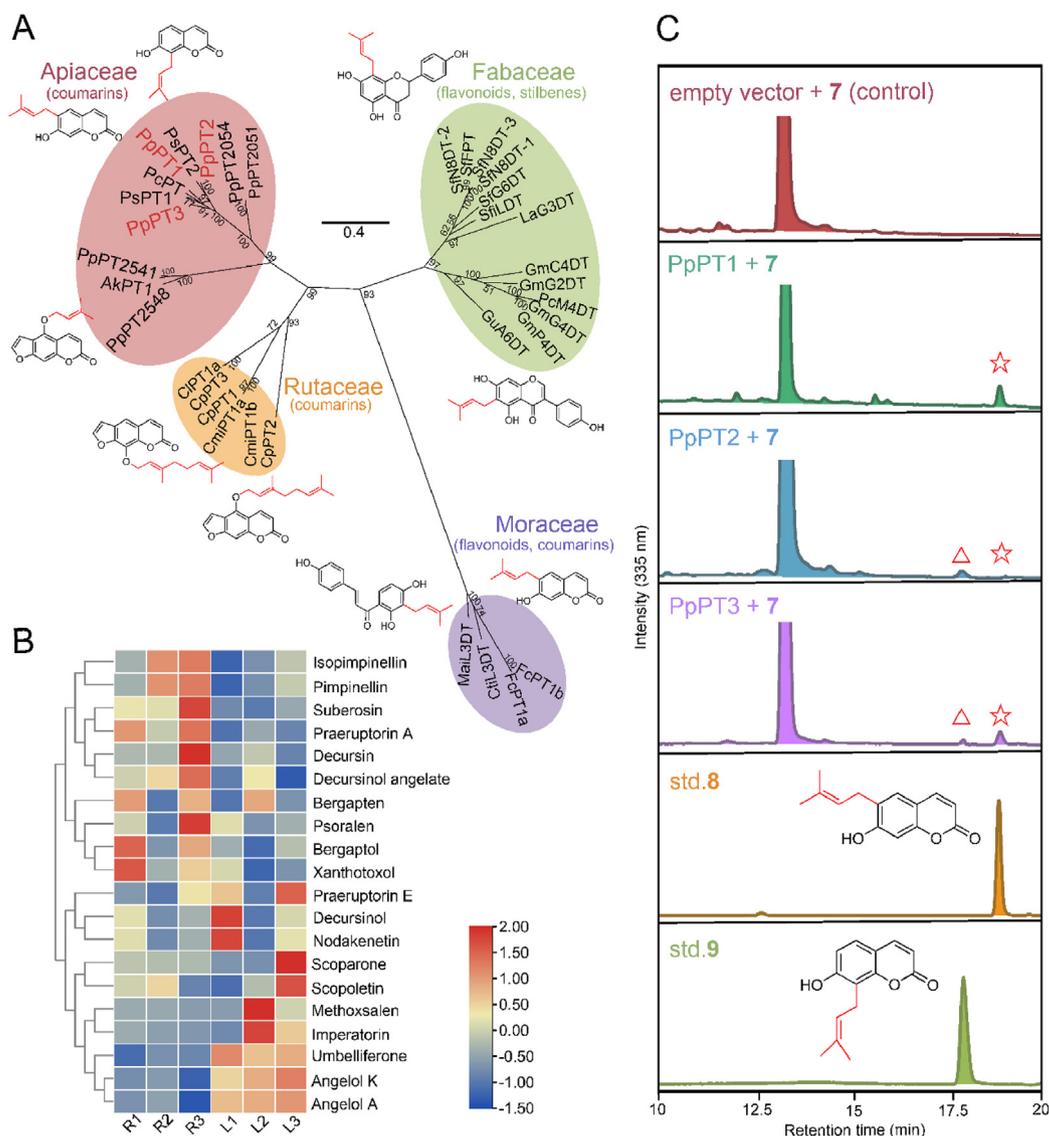

Figure 2 Discovery of *P. praeruptorum* prenyltransferases (PpPTs) involved in linear/angular coumarins biosynthesis. (A) Phylogenetic analysis of PpPTs and related prenyltransferases (PTs) identified from other species with similar functions. The products catalyzed by different PTs are depicted in the corresponding position of the PTs. PTs have been clustered into at least four groups, demonstrating their versatility. (B) Heatmap of the cluster analysis of the main coumarin metabolites derived from *P. praeruptorum*. SCs and FCs/PCs were largely clustered into two groups according to their abundance in the roots or leaves. SCs and FCs/PCs tend to accumulate in leaves and roots, respectively. (C) Enzymatic functions of three representative PTs detected *via* liquid chromatography at 335 nm (LC). The activities of PpPT1, PpPT2, and PpPT3 are arranged successively. All enzymatic activity assays were carried out on a shaking incubator (220 rpm) at 25 °C for 2 h. The reaction contained 50 μL crude enzyme, 200 μmol/L **7**, 100 μmol/L DMAPP, and 200 μmol/L MgCl₂ in 200 μL of 100 mmol/L Tris-HCl, pH 7.5. The LC map of the standards **8** and **9** as well as their chemical structures are depicted correspondingly. The retention times of standards **8** and **9** were 18.8 and 17.85 min, respectively.

(Supporting Information Fig. S4 and Fig. 2C). Such functional differences between the two genes may be associated with the fact that different plants have different angular or linear coumarins[1].

Following the cloning of the other five PTs, functional experiments indicated that PpPT2051, PpPT2054 and PpPT2548 exhibited trace amounts of linear activity (Supporting Information Fig. S5). Interestingly, PpPT3 (ASM_9.2506-07, Table S1) demonstrated both linear and angular activities toward compound **7** (Fig. S3 and Fig. 2C). Generally, PpPT1 showed strict linear activity, while PpPT3 primarily catalyzed linear products and also displayed a considerable angular function (20%); PpPT2 generally demonstrated angular activity but exhibited remnant trace linear function (1%). We note that all the detected activities were very low, which is largely in accordance with the previously report for the low solubility of this kind of PTs (Supporting Information Fig. S6A)[6−8]. Despite the amount of crude enzyme and the yield of **8** is somewhat positive correlation in a specific reaction (Fig. S6B), only an extremely amount of protein could be detected using Western-blot (Fig. S6C). Hence, strategies to increase the amount of soluble protein will be meaningful for the synthetic biology of coumarins.

3.2. Identification of two CYP450 cyclases crucial for furan and pyran ring refactoring

The three PTs described above produce prenylated products as essential precursors for pyran and furan ring refactoring (Fig. 2). However, another cyclization reaction was required to complete the last step in the biosynthesis of the furan and pyran ring skeletons[1,3,10]. To identify the functional genes, we first screened all the CYP71 family genes in the *P. praeruptorum* genome, owing to speculation that the mechanism of isopentene group cyclization is similar to that of CYP71 menthofuran synthase derived from *Mentha piperita*[1,26,27]. In addition, psoralen/angelicin synthase and few hydroxylases that participate in coumarin downstream processes belong to the CYP71 family[2,25,28−31]. Collectively, 171 proteins were classified as CYP71 or CYP71-clan families according to gene annotation, including 57 CYP71, 31 CYP76, 20 CYP736, 10 CYP81, and 11 CYP82 proteins (Supporting Information Table S2 and Fig. S7). Metabolome-transcriptome correlation analysis was used for initially screening the candidate genes, although only a small number of CYP450s were excluded. Considering that genes within the same metabolic pathway usually exhibit similar levels of expression, we re-examined the 171 CYP71 genes and 26 CYP71s with high expression levels were selected (Supporting Information Fig. S8)[15]. Next, using PpPT1 and PpPT2 as the target expression levels, a co-expression analysis was conducted to find the candidate CYP450s with Pearson correlation coefficient[15]. The analysis revealed that ASM_0.35 exhibited a strong co-expression pattern with PpPT1 and that ASM_7.4765 was co-expressed with PpPT2 (Fig. 3A). Hence, these two genes were selected and cloned into yeast expression vectors for further functional verification. After induction, the strains containing ASM_7.4765 and ASM_0.35 were incubated with compound **8**, respectively. Fortunately, the strain containing ASM_7.4765 produced a clear peak (peak 1, Fig. 3B), the [M+H]$^+$ molecular weights of 247.09480 and 269.07654 and the retention time of peak 1 was similar to those of standard **10** (Supporting Information Fig. S9A). Additionally, two additional peaks (peaks 2 and 3) were also observed. Using standard calibration and high-resolution mass spectrometry (HRMS), we discovered that the target protein could also produce compound

11, a PC that initiates the biosynthesis of a key PCs (Fig. 3B and Fig. S9A). Peak 3 was not be confirmed at present, however, according to its *m/z* of [M+H]$^+$ = 265.10522 Da, we infer it maybe the by-product peusedanol. These results indicate that the identified cyclase mediated the biosynthesis of both FCs and PCs.

We also used compound **9** as a substrate to assess the functions of the candidate genes. Remarkably, another candidate, ASM_2436 (PpOC), demonstrated OC activity and produced two peaks (peaks 4 and 5), identified as compounds **12** and **13** using standard calibration and HR-MS (Fig. 3C and Fig. S9B). Considering that ASM_7.4765 and ASM_2436 specifically cyclized compounds **8** and **9** (Supporting Information Fig. S10), they were termed as PpDC and PpOC, respectively. These two CYP71s belong to the CYP736 family and are classified as CYP736A360 and CYP736A366, respectively, per the cytochrome P450 annotation committee (http://drnelson.utmem.edu/CytochromeP450.html). However, marmesin synthase (MS) obtained from Moraceae is a CYP76 protein, implying that furanocoumarin biosynthesis evolved independently in different species[10]. Furthermore, we functionally verified all other highly expressed genes in each CYP71 clan to avoid any omission in the gene co-expression analysis; however, none of them exhibited similar functions (Fig. S10). Hence, ASM_7.4765 and ASM_2436 were considered the functional DC and OC enzymes in *P. praeruptorum*, respectively. The expression profile indicated that DC and OC were more highly expressed in roots than in leaves (Supporting Information Fig. S11). This expression mode was similar to that of the PTs and agreed with the compound accumulation exhibited by these organs (Fig. 2B). The PpDC was localized in the endoplasmic reticulum, which was distributed across each organelle, potentially favoring the utilization of the compounds **8** and **9** produced by PTs (Supporting Information Fig. S12).

3.3. Acid/base-assisted epoxide ring opening contributes to furan and pyran ring refactoring

The identification of DC and OC provided molecular insights into the biosynthesis of four different types of plant furan and pyran rings. This metabolic diversity aroused our interest, because a recently identified MS appeared to produce only one type of linear FCs[10]. We further purified the microsome proteins of DC and surveyed its biochemical characteristics for its high activity (Fig. 3B and 4A, and Supporting Information Figs. S13−S16). We found that excessive purification of protein abolishes its activity in PCs formation, only whole cell catalysis or the corresponding cell lysate of strain containing target gene could produces PCs **11** and the additional peak 3 (or peak 6, Fig. 3B, C and Fig. S16A). In addition, the products of DC/OC at different pH values were dissimilar: only one product (compound **10/12**) was observed under alkaline conditions and the formation of PCs needs an acidic or neutral environment (Fig. 4A). These findings imply that some intermediates or byproducts may be produced simultaneously, with different acid/base environments potentially changing the metabolic direction. To investigate the underlying mechanisms of furan and pyran ring refactoring, whole cell catalysis that could produce three peaks was employed. We first focused on the identification of peak 3 because it was largely epoxide intermediate **14** according to previous speculations and a recent publication (Supporting Information Fig. S17)[10,32−35]. However, the absence of a corresponding prospective charge-to-mass ratio of **14** in HR-MS and its high abundance convinced
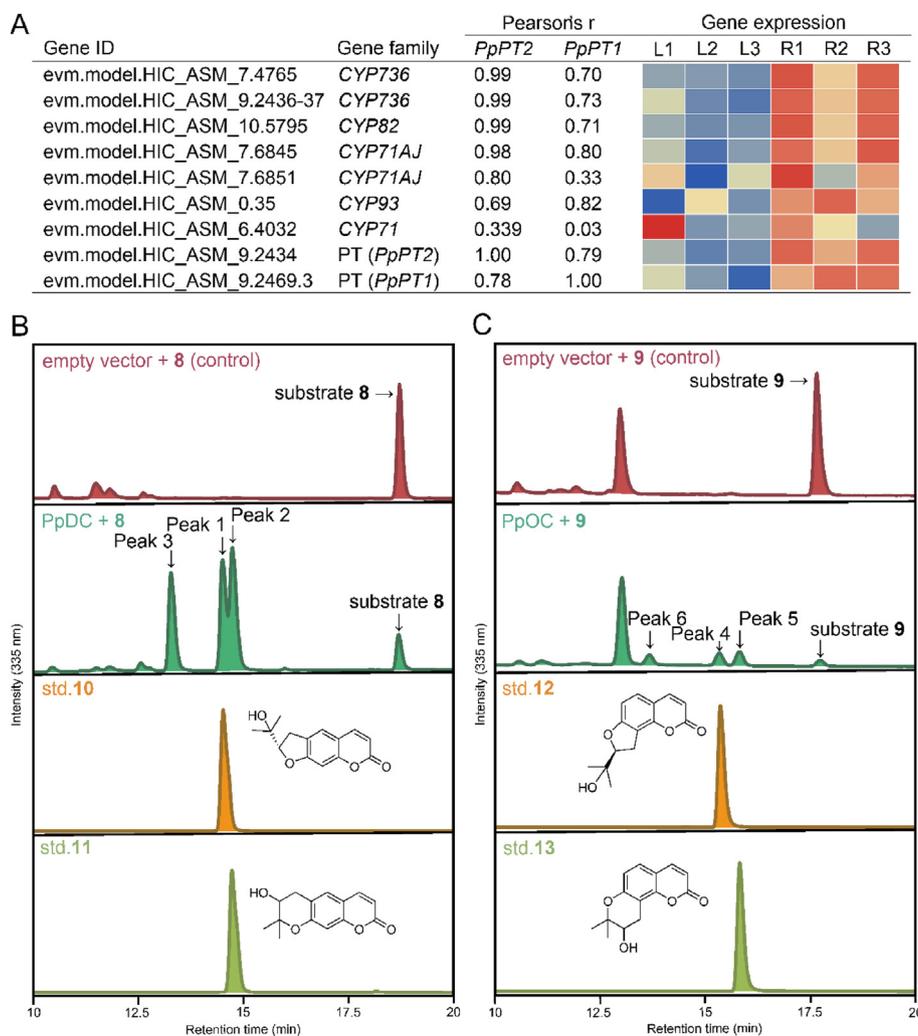
Figure 3 Discovery of cyclases involved in FCs and PCs biosynthesis. (A) Co-expression analysis of candidate cytochrome P450s using identified PpPTs as target. Pearson correlation coefficient was used to analyze the correlation between target genes and candidate genes. A strong co-expression with PsPT2 and PsPT1 was observed in some CYP736 family proteins and was subsequently evidenced to be involved in cyclization. Enzymatic functions of (B) PpDC and (C) PpOC were detected using liquid chromatography at 335 nm (LC). The enzymatic activity assays were carried out on a shaking incubator (220 rpm) at 30 °C for 2 h. The reaction broth contained 100 μmol/L demethylsuberosin or osthenol and 500 μmol/L NADPH in 200 μL crude enzyme resuspend in 100 mmol/L Tris-HCl, pH 7.5. The LC maps of standards **10**−**13** and their corresponding chemical structures are shown in the figure. Peak 1−3 indicate three specific products produced by PpDC compared with the sample using empty vector. The retention times of compounds **10**−**13** was 14.63, 14.86, 15.30, and 15.75 min, respectively.

us that an intermediate was absent. The HR-MS data were re-examined, and the $[M+H]^+ = 265.10522$ Da m/z of peak 3 indicated that it was a diol (peucedanol, **15**) instead of an epoxide. This was further confirmed using standard **15** and ultraviolet absorption (Supporting Information Figs. S18A, S18B, and S19). The identification of compound **15** and its associated derivatives from natural sources further confirmed the presence of diol[34]. The incorporation of water into compound **15** was also demonstrated *via* isotope labelling of the oxygen atom of one of the diols[35,36]. In addition, **15** could not be catalysed by PpDC; hence, hydroxylation followed by a cyclization mechanism similar to that of AurD to form furan/pyran ring seems impractical (Fig. S18B)[37]. All these results prove that **15** is not an intermediate but a by-product.

Although peak 3 was successfully identified as diol **15**, the reaction mechanism of the cyclase was more complex. Since we could not detect epoxides, the previous speculation that an epoxide

intermediate involved a stepwise reaction was challenged[10,32−35]. We then focused on whether dihydroxylation product **15** resulted from the overall syn- or anti-addition of the two hydroxyl groups to the double bond. The syn-addition strongly suggests the formation of a carbocation intermediate, whereas the latter is consistent with an oxonium ion or epoxide. Chiral phase chromatography analysis of enzymatic product **15** was conducted, as indicated in Supporting Information Fig. S18C and a specific chiral *o*-diol was obtained. However, if a carbocation intermediate is formed, the *o*-diol is racemic. Hence, we concluded that an epoxide intermediate was involved in the cyclase reaction mechanism. Although we could not detect the epoxide, identification of epoxide derivatives from other plants substantiated its existence (Supporting Information Fig. S20)[32−34]. In addition, when epoxide **14** was chemically synthesized and used as the substrate, it was easily converted to **10** and **11** (Supporting Information Figs. S18B and S21−S34). These
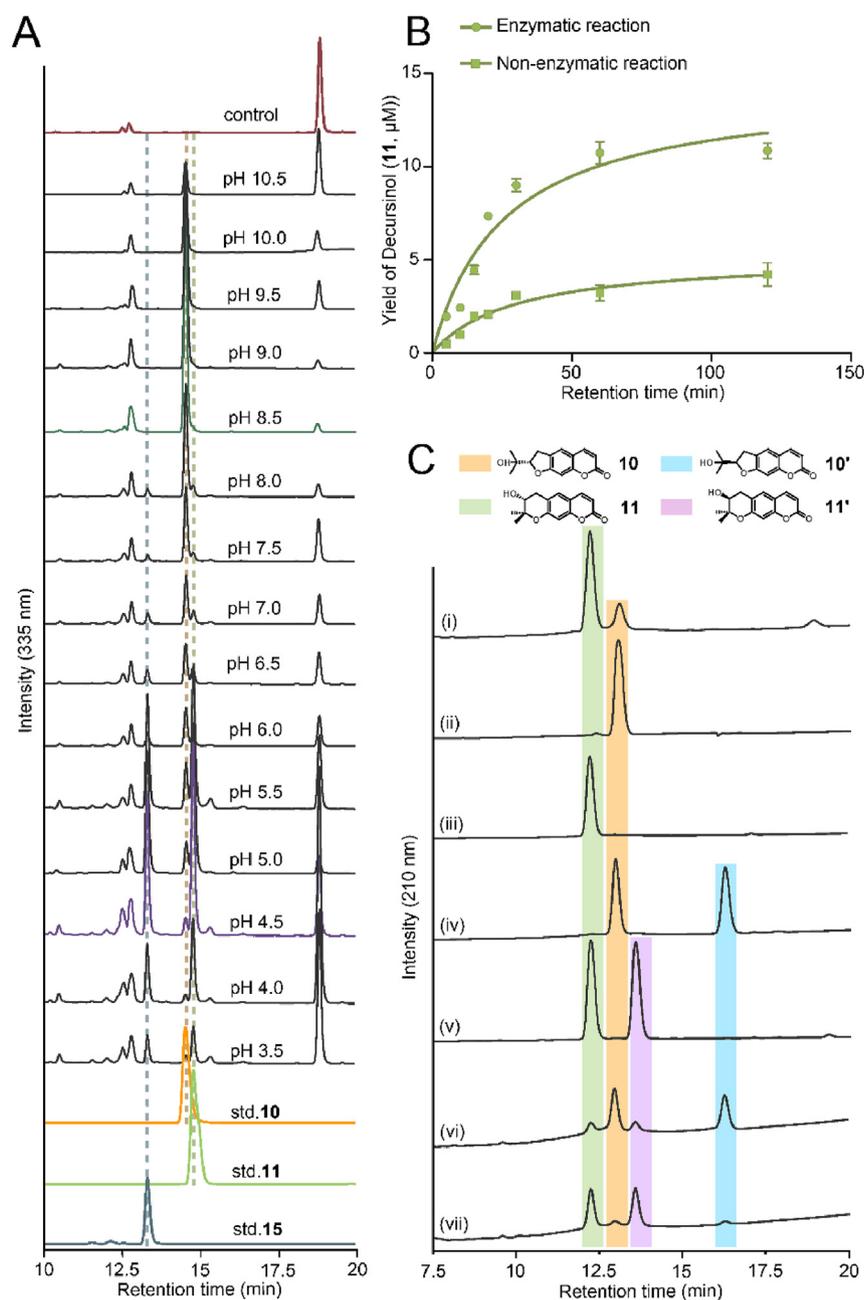
Figure 4 Acid/base-assisted epoxide ring opening with and without PpDC. (A) Reaction characteristics of the PpDC in different pH values. The buffer used in this work are as follows: pH 3.5−6.0, 50 mmol/L citric acid buffer; pH 6.5−8.5, 50 mmol/L Tris-HCl buffer; pH 9.0−10.5, 50 mmol/L glycine buffer. (B) Enzymatic and non-enzymatic reactions of **14** in pH 5.5. The reaction was conducted a final concentration of 250 μmol/L of epoxide **14** at 30 °C for 0−120 min. Only the yield of **11** was calculated for its high production in pH 5.5. (C) Chiral phase chromatography analysis of **10** and **11**. (i) represents the enzymatic products **10** and **11** using **8** as a substrate; the enzymatic products **10** (ii) and **11** (iii) were collected with preparative liquid phase and then detected with chiral phase chromatography, respectively; (iv) and (v) represent the chiral phase chromatography analysis of **10** and **11** produced by chemical synthesis. No chiral selectivity was observed in chemical synthesis; (vi) represents the chiral phase chromatography analysis of the reaction of **14** with PpDC; (vii) represents the chiral phase chromatography analysis of spontaneous reaction of **14**.

results support the epoxide intermediate involved in the step-wise cyclisation reaction of FCs/PCs[10,32−35], however, the role of cyclase in this kind of "step-wise" cyclization remains unclarified.

We, therefore, proposed two hypotheses regarding the "step-wise" cyclisation: (1) PpDC is only used for epoxide intermediate **14** formation, and then, **10** and **11** are cyclized by spontaneous

reaction; (2) PpDC is not only indispensable for **14** formations, but also irreplaceable for **10** and/or **11** cyclization. The dissimilar products of PpDC at different pH values implied that the products may be produced by non-enzymatic reaction (Fig. 4A). Hence, the cyclization mechanism may originate from the instability of the epoxide intermediates. To prove hypothesis (1), epoxide **14** was

used as a substrate to test its reaction rate in enzymatic and non-enzymatic reaction processes (Fig. 4B). Consistent with our expectations, non-enzymatic reaction also occurred when **14** was used as substrate. However, at the same time, the addition of enzyme could significantly improve the reaction velocity of the epoxide, and the enzymatic activities is 1.2−2.7 times more active when compared with non-enzymatic reactions. This challenges opinion (1) that only spontaneous reactions lead to cyclization. Hence, except for **14** formation, PpDC may also be involved in cyclization similar to XimE in xiamenmycin biosynthesis[35]. To prove hypothesis (2), we analysed the enantioselectivity of the cyclic products for the PpDC-medicatd enzymatic reaction involving the S_N2 reaction on the epoxide and the expected stereoselectivity. We then analyzed the enzymatic products of **8** in Fig. 4A, as depicted in Fig. 4C(i-iii) of chiral phase chromatography, and two strictly chiral selective products were produced. Based on the standards **10/11** produced by chemical synthesis (Fig. 4C(iv-v)) and X-ray diffraction (Supporting Information Fig. S35), the absolute configuration of FCs is (*S*)-type, while the PCs one is (*R*)-type. Hence, PpDC was not only used for **14** formations but was also involved in the follow-up cyclization.

To further verify hypothesis (2), nonenzymatic and enzymatic transformations were performed using chemically synthesized **14** as substrates. Considering that **11** was produced in an alkaline environment and the yield of PpDC was low in a neutral environment (Fig. 4A), citric acid buffer (pH 5.5) was used as the reaction buffer. Under these conditions, non-enzymatic reactions tended to produce furan-favoring compounds with the ratio of FCs (**10** + **10′**) to PCs (**11** + **11′**) of 4.43:1 (Fig. 4C(vi)). However, when PpDC was added into the buffer, the production of PCs is predominant, in which the ratio of PCs (**11** + **11′**) to FCs (**10** + **10′**) is 9.2:1 (Fig. 4C(vii)). This indicates that both enzymatic and non-enzymatic cyclization occurred regardless of whether DC was added. The enzymatic reaction leads to excessive pyran-formation of the end products, such as minor (*S*)-configuration FCs and major (*R*)-configuration PCs. Because the formation of 6-*endo*-tet PCs is disfavored according to Baldwin rules[11], the higher ratio of **11** may be evidence of the enzymatic cyclization mechanism. Hence, we propose a stepwise reaction in which epoxides are essential (Supporting Information Fig. S36). In contrast to the reactions of XimD and XimE, which require two enzymes to complete the architecture of 6-*endo*-tet PCs[35], we found that an individual enzyme, DC, could realize this process in plants (Supporting Information Fig. S36). However, we could not know the exact mechanism of the relationship between the products and pH in the process. Maybe, the different pH in different growing environments of *P. praeruptorum* lead to the formation of the geotropism. Otherwise, another enzyme, that is needed for epoxide ring opening and it does not rely on acid or base, remain to be identified in planta. In addition, why further purification of DC/OC abolishes its activity of PCs formation remains unclear and need to be explicated in the future (Fig. S16).

3.4. Possible mechanisms of furan and pyran ring formation in coumarin biosynthesis

Oxidative cyclization is a powerful method of forming variable skeletons *via* radical or nonradical cyclization, epoxidation, halogenation, and ring rearrangement[38,39]. According to the above findings and those of previous studies, we speculated that this form of pyran and furan cyclization occurs *via* epoxide ring-opening

reactions and obeys Baldwin's "rules for ring closure", despite no reports of this in plants[11,40,41]. However, the formation of disfavored 6-*endo*-tet cyclization and individual furan products under alkaline conditions makes it challenging to meet the classical Baldwin's rules[11]. Recently identified enzymes involved in polyether and xiamenmycin biosynthesis from streptomyces have demonstrated the mechanism of pyran and furan ring formation in bacteria, although the reaction required an additional epoxide hydrolase (Lsd19B and XimE) to reinforce pyran ring formation[35,42,43]. Thus, the novel cyclase identified herein may suggest certain modifications and/or extensions to the classical Baldwin's rules[41].

To investigate the possible catalytic mechanisms of this cyclase, we first constructed a three-dimensional structural model using *Salvia miltiorrhiza* CYP76AH1 as a template for its high identity with PpDC (33.82%)[44]. Combined with the primary structure alignment of PpDC and PpOC, we identify some amino acid residues located around the substrate binding and catalytic regions (Supporting Information Figs. S37−S39). Based on the site-specific mutagenesis-approved docking results and biochemical analyses, we proposed a possible catalytic mechanism (Supporting Information Figs. S38−S41), which was further supported by molecular dynamics simulation (MD) and density functional theory (DFT) calculations (Supporting Information Tables S4−S6, Supporting Information Figs. S42 and S43)[33,39]. First, the C1−C2 double bond of the substrate was cyclized by the superoxide of the iron porphyrin at the catalytic site to form an epoxide (Figs. S40 and S41). This cyclization was evidenced by the suitable distance (5−6 Å) between the iron porphyrin ring and C1/C2 in a 300 ns steady state and the low-energy barrier of 9.1 kcal/mol, which gave a possible oxidation path to the epoxide intermediate (Figs. S42 and S43, and Tables S3 and S4). Meanwhile, a hydrogen bond was formed between the C5−OH of compound **8** and the carboxyl group of Glu303 (<2 Å) (Fig. S42). The mutation of Glu303 to alanine, methionine, or glycine completely abolished its activity implying the importance of the identified amino acids (Fig. S38). The same situation also occurs in PpOC (Fig. S39). Furthermore, we found that compound **14** exists in both acidic and alkaline environments (Fig. 5), theoretically, proving that the cyclization requires epoxide ring opening. Nevertheless, as an intermediate, the fate of the epoxide was controlled by the concentration of Brønsted−Lowry acids/bases in the reaction solution (Figs. S40 and S41). In an acidic environment, free protons combine with the oxygen lone pair of compound **14**. Under these conditions, the hydroxyl proton at C5 is transferred to PpDC-Glu303, and subsequently, the deprotonated oxygen adopts S_N2 nucleophilic attack at C1 or C2 to form 6-*endo*-tet pyran or 5-*exo*-tet furan ring products. In addition, the oxygen in H_2O undergoes a nucleophilic attack at C1 or C2 *via* a spontaneous S_N1-like reaction to form the byproducts **15** or **15′**[37]. DFT calculations revealed that the energy barrier for 5-*exo*-tet formation was significantly higher than that for 6-*endo*-tet and byproduct formation (10.3 *vs* 1.2 *vs* 0.0 kcal/mol), indicating that pyran and water addition to the epoxide ring openings were the major pathways occurring in acidic conditions (Fig. 5A, and Tables S3 and S5). However, when the intermediate interacted with the proton in the solution, the former preferentially formed a pyran ring or the oxygen atom on C5 performed a nucleophilic attack at C1, which agrees with the result presented in Fig. 4A. Under this circumstance, the energy barrier of 5-*exo*-tet was lower than that of 6-*endo*-tet by 2.2 kcal/mol, and the major furan ring product **10** was subsequently produced under alkaline conditions.
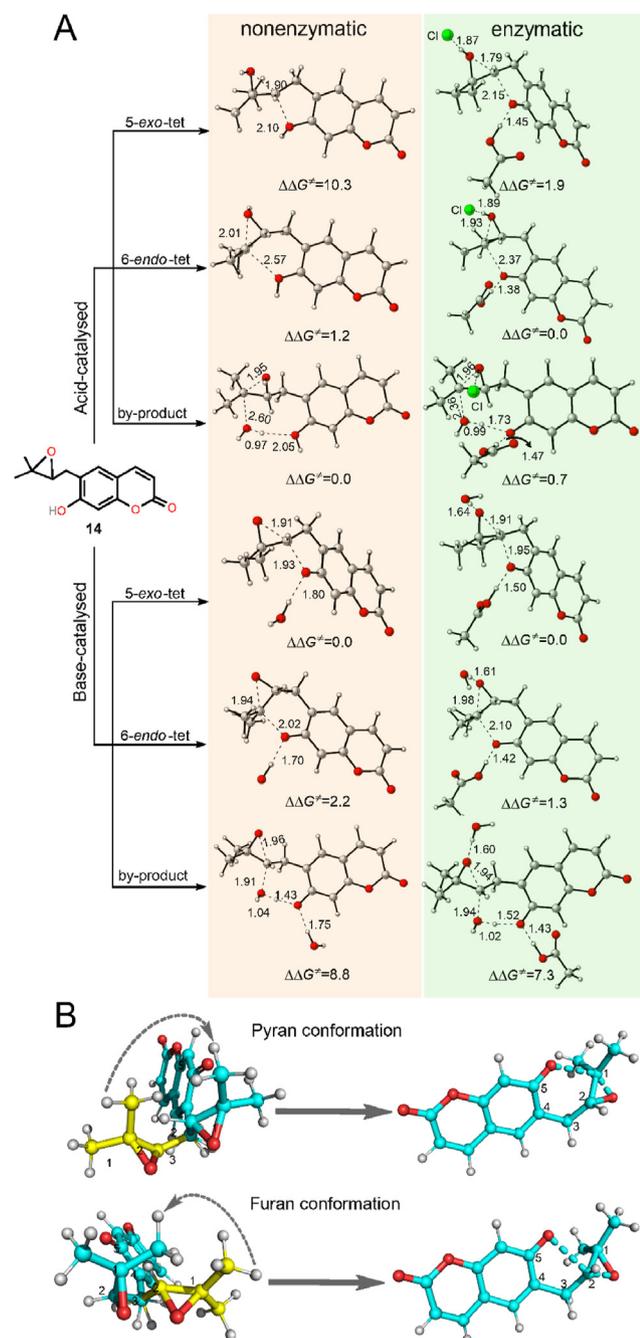
Figure 5 Possible mechanisms of furan and pyran ring formation in coumarins biosynthesis. (A) Acid/base-catalyzed cyclization of FCs via 5-exo or 6-endo cyclization with and without PpDC at different pH values. Gibbs free energies (kcal/mol) obtained at the M062X/6-311++G(d,p) (SMD, water)//M062X/6-31G(d,p) (SMD, water) level and forming/breaking bond distances (Å) are given. (B) The changes of dihedral angles formed by C5−C4−C3−C2 and C4−C3−C2−C1 in two conformations of the epoxide **14** intermediate during the formation of pyran and furan rings.

We found that the coexisting conformations of compounds **14** and **16** and their corresponding conformations in different pH values contribute to product diversity (Supporting Information Figs. S44 and S45). In other words, in the conformational

characteristics of compounds **14** to **10**, the dihedral angle of C5−C4−C3−C2 was maintained as obtuse and that of C4−C3−C2−C1 was maintained as acute (Fig. 5B and Supporting Information Figs. S44 and S46). According to previous studies, the dihedral angle of C5−C4−C3−C2 can reach a low-energy transition state via clockwise rotation, which is beneficial for the formation of the 6-endo pyran ring structure[35,37,42,43]. In the conformational characteristics of compounds **14** to **11**, C5−C4−C3−C2 was maintained as an acute dihedral angle, whereas C4−C3−C2−C1 was maintained as an obtuse dihedral angle (Fig. 5B, and Supporting Information Figs. S44, S46−S48). In a more alkaline environment, free protons gradually decrease, and hence, no protons are available to form electrostatic mutual bonds with the epoxide (Fig. 4A and Fig. S48). Accordingly, the deprotonated oxygen on C5 directly conducts an S_N2 nucleophilic attack at C2 to generate a furan product without generating any byproducts (Figs. 4 and 5 and Fig. S40). As per the DFT calculations, this kind of pH-dependent transformation largely confirms the hypothetical synergistic mechanism wherein a protonated pathway prefers to form a pyran ring, whereas under deprotonated conditions, a furan ring is preferentially generated (Fig. 5A)[35].

Although the reaction can be theoretically processed in a nonenzymatic environment, the involvement of additional enzymes could considerably reduce the energy barrier (Fig. 5A and Table S5). MD and DFT calculations indicated that Glu303 in PpDC or Asp301 in PpOC are important for forming hydrogen bonds with the hydroxyl groups of the substrates to enhance the stability of the enzyme−substrate complex and benefit electron transfer (Figs. 4A and 5A and Fig. S43). The mutation of Glu303 to aspartic acid retained its activity, implying that the latter was also functional during proton transfer (Fig. S38). This is largely in accordance with a previous study that reported that during the formation of the XimE-catalyzed pyran ring, the protonation and stabilization of epoxides are primarily realized by catalyzing acidic residues Tyr46 and Glu136 and the basic residue His102[35]. This type of general acid/base catalysis is also observed in microbes[35,37,42], whereas the formation of "anti-Baldwin" 6-endo pyran ring requires two enzymes: a flavin-dependent monooxygenase to produce epoxide and an epoxide hydrolase for cyclization[35,37]. Hence, the cyclase identified herein may broaden the cognition of the classical Baldwin's rules[11]. We also identified a key amino acid residue (Gly308) in determining the substrate specificity of PpOC, which is Glu303 in PpDC. When we mutate Gly308 in PpOC into glutamic acid (E), G308E, it endows PpOC possesses the DC activity (Supporting Information Fig. S49). However, the mechanism by which this cyclase evolved such specific functions and their strict substrate specificity remains unclear, especially because they exhibit a high similarity with other CYP450 proteins that lack this function[2,25]. In addition, due to the lack of the crystal structures of PpDC and PpOC, many key catalytic residues maybe remained undiscovered.

3.5. Data availability

All reported data in this study are available via database or by request from the corresponding author. Coding DNA sequences for the genes characterized and assessed in this study deposited in the National Center for Biotechnology Information (NCBI) database with the following accessions: PpPT2051 (ON934682), PpPT2054 (ON934683), PpPT2548 (ON934684), PpPT2469.3

(ON934685), PpPT2434 (ON934686), PpPT2506-07 (ON934687), PpDC (ON934691), PpOC (ON934692). The transcriptome sequences of tissues have been submitted to the NCBI Sequence Read Archive (SRA) database under the BioProject accession PRJNA847169.

4. Conclusions

We identified two types of enzymes that mediate prenylation and cyclization reactions, respectively. These newly discovered enzymes complete the last missing steps in linear/angular PC and FC biosynthesis. The acid/base-assisted epoxide ring opening mediated the chemical diversity and the formation of disfavored 6-endo-tet cyclization products further expanding our knowledge regarding the classical Baldwin's rules[11,40,41]. This study provides a suitable explanation for the generation of chemical diversity in plants. Further efforts to probe the specific and precise acid/base-assisted cyclization mechanisms at the protein structural level are ongoing.

Author contributions

Conceptualization: Yucheng Zhao, Lingyi Kong, Fei Qiao, Yibei Xiao; Methodology and experiments: Yucheng Zhao, Yuedong He, Liangliang Han, Libo Zhang; Investigation: Yuanzheng Xia, Fucheng Yin, Xiaobing Wang; Visualization: Deqing Zhao, Sheng Xu; Funding acquisition: Yucheng Zhao, Fei Qiao, Yibei Xiao; Project administration: Lingyi Kong, Yucheng Zhao; Supervision: Yucheng Zhao, Fei Qiao, Yibei Xiao; Writing−original draft: Yucheng Zhao, Writing−review & editing: Yucheng Zhao.

Conflicts of interest

The authors declare no conflicts of interest.

Appendix A.   Supporting information

Supporting data to this article can be found online at https://doi.org/10.1016/j.apsb.2023.10.016.

References

1. Bourgaud F, Hehn A, Larbat R, Doerper S, Gontier E, Kellner S, et al. Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes. *Phytochemistry Rev* 2006;**5**:293−308.

2. Limones-Mendez M, Dugrand-Judek A, Villard C, Coqueret V, Froelicher Y, Bourgaud F, et al. Convergent evolution leading to the appearance of furanocoumarins in citrus plants. *Plant Sci* 2020;**292**: 110392.

3. Rodrigues JL, Rodrigues LR. Biosynthesis and heterologous production of furanocoumarins: perspectives and current challenges. *Nat Prod Rep* 2020;**38**:869−79.

4. Robe K, Izquierdo E, Vignols F, Rouached H, Dubos C. The coumarins: secondary metabolites playing a primary role in plant nutrition and health. *Trends Plant Sci* 2021;**26**:248−59.

5. Vetica F, Chauhan P, Dochain S, Enders D. Asymmetric organocatalytic methods for the synthesis of tetrahydropyrans and their application in total synthesis. *Chem Soc Rev* 2017;**46**:1661−74.

6. Karamat F, Olry A, Munakata R, Koeduka T, Sugiyama A, Paris C, et al. A coumarin-specific prenyltransferase catalyzes the crucial biosynthetic reaction for furanocoumarin formation in parsley. *Plant J* 2014;**77**:627−38.

7. Munakata R, Olry A, Karamat F, Courdavault V, Sugiyama A, Date Y, et al. Molecular evolution of parsnip (*Pastinaca sativa*) membrane-bound prenyltransferases for linear and/or angular furanocoumarin biosynthesis. *New Phytol* 2016;**211**:332−44.

8. Zou Y, Zhan Z, Li D, Tang M, Cacho RA, Watanabe K, et al. Tandem prenyltransferases catalyze isoprenoid elongation and complexity generation in biosynthesis of quinolone alkaloids. *J Am Chem Soc* 2015;**137**:4980−3.

9. Munakata R, Olry A, Takemura T, Tatsumi K, Ichino T, Villard C, et al. Parallel evolution of UbiA superfamily proteins into aromatic *O*-prenyltransferases in plants. *Proc Natl Acad Sci U S A* 2021;**118**: e2022294118.

10. Villard C, Munakata R, Kitajima S, van Velzen R, Schranz ME, Larbat R, et al. A new P450 involved in the furanocoumarin pathway underlies a recent case of convergent evolution. *New Phytol* 2021;**231**:1923−39.

11. Baldwin JE. Rules for ring closure. *Ciba Found Symp* 1978;**18**:85−99.

12. Thompson JD, Higgins DG, Gibson TJ. Clustal W. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;**22**:4673−80.

13. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;**32**:268−74.

14. Letunic I, Bork P. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021; **49**:W293−6.

15. Nett RS, Lau W, Sattely ES. Discovery and engineering of colchicine alkaloid biosynthesis. *Nature* 2020;**584**:148−53.

16. Han BY, Wang ZL, Li JH, Jin Q, Wang HT, Chen K, et al. A highly selective *C*-rhamnosyltransferase from *Viola tricolor* and insights into its mechanisms. *Acta Pharm Sin B* 2023;**13**:3535−44.

17. Pompon D, Louerat B, Bronine A, Urban P. Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol* 1996;**272**:51−64.

18. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;**596**:583−9.

19. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 2009;**30**: 2785−91.

20. Zhao Y, Wang N, Zeng Z, Xu S, Huang C, Wang W, et al. Cloning, functional characterization, and catalytic mechanism of a bergaptol *O*-methyltransferase from *Peucedanum praeruptorum* Dunn. *Front Plant Sci* 2016;**7**:722.

21. Chou KC, Shen HB. Cell-PLoc: a package of Web servers for predicting subcellular localization of proteins in various organisms. *Nat Protoc* 2008;**3**:153−62.

22. Nelson BK, Cai X, Nebenführ A. A multicolored set of *in vivo* organelle markers for co-localization studies in Arabidopsis and other plants. *Plant J* 2007;**51**:1126−36.

23. Zhao Y, Truhlar DG. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor Chem Acc* 2008;**119**:525.

24. Gonzalez C, Schlegel HB. An improved algorithm for reaction path following. *J Chem Phys* 1989;**90**:2154−61.

25. Jian X, Zhao Y, Wang Z, Li S, Li L, Luo J, et al. Two CYP71AJ enzymes function as psoralen synthase and angelicin synthase in the biosynthesis of furanocoumarins in *Peucedanum praeruptorum* Dunn. *Plant Mol Biol* 2020;**104**:327−37.

26. Bertea CM, Schalk M, Karp F, Maffei M, Croteau R. Demonstration that menthofuran synthase of mint (Mentha) is a cytochrome P450 monooxygenase: cloning, functional expression, and characterization of the responsible gene. *Arch Biochem Biophys* 2001;**390**:279−86.

27. Croteau RB, Davis EM, Ringer KL, Wildung MR. (−)-Menthol biosynthesis and molecular genetics. *Naturwissenschaften* 2005;**92**:562−77.

28. Larbat R, Kellner S, Specker S, Hehn A, Gontier E, Hans J, et al. Molecular cloning and functional characterization of psoralen synthase, the first committed monooxygenase of furanocoumarin biosynthesis. *J Biol Chem* 2007;**282**:542−54.

29. Larbat R, Hehn A, Hans J, Schneider S, Jugdé H, Schneider B, et al. Isolation and functional characterization of CYP71AJ4 encoding for the first P450 monooxygenase of angular furanocoumarin biosynthesis. *J Biol Chem* 2009;**284**:4776−85.

30. Dueholm B, Krieger C, Drew D, Olry A, Kamo T, Taboureau O, et al. Evolution of substrate recognition sites (SRSs) in cytochromes P450 from Apiaceae exemplified by the CYP71AJ subfamily. *BMC Evol Biol* 2015;**15**:122.

31. Krieger C, Roselli S, Kellner-Thielmann S, Galati G, Schneider B, Grosjean J, et al. The CYP71AZ P450 Subfamily: a driving factor for the diversification of coumarin biosynthesis in Apiaceous plants. *Front Plant Sci* 2018;**9**:820.

32. Hamerski D, Matern U. Elicitor-induced biosynthesis of psoralens in *Ammi majus* L. suspension cultures. *Eur J Biochem* 1988;**171**:369−75.

33. Brown SA, el-Dakhakhny M, Steck W. Biosynthesis of linear furanocoumarins. *Can J Biochem* 1970;**48**:863−71.

34. Steck W, Brown SA. Biosynthesis of angular furanocoumarins. *Can J Biochem* 1970;**48**:872−80.

35. He BB, Zhou T, Bu XL, Weng JY, Xu J, Lin S, et al. Enzymatic pyran formation involved in Xiamenmycin biosynthesis. *ACS Catal* 2019;**9**:5391−9.

36. Steyn PS, Vleggaar R. Mechanistic studies on the biosynthesis of the aurovertins using ^{18}O-labelled precursors. *J Chem Soc Chem Commun* 1985;**24**:1796−8.

37. Mao XM, Zhan ZJ, Grayson MN, Tang MC, Xu W, Li YQ, et al. Efficient biosynthesis of fungal polyketides containing the dioxabicyclo-octane ring system. *J Am Chem Soc* 2015;**137**:11904−7.

38. Tang MC, Zou Y, Watanabe K, Walsh CT, Tang Y. Oxidative cyclization in natural product biosynthesis. *Chem Rev* 2017;**117**:5226−333.

39. Walsh CT, Tang Y. Recent advances in enzymatic complexity generation: cyclization reactions. *Biochemistry* 2018;**57**:3087−104.

40. Alabugin IV, Gilmore K. Finding the right path: Baldwin "rules for ring closure" and stereoelectronic control of cyclizations. *Chem Commun* 2013;**49**:11246−50.

41. Gilmore K, Mohamed RK, Alabugin IV. The Baldwin rules: revised and extended. *Interdiscip Rev Comput Mol Sci* 2016;**6**:487−514.

42. Hotta K, Chen X, Paton RS, Minami A, Li H, Swaminathan K, et al. Enzymatic catalysis of anti-Baldwin ring closure in polyether biosynthesis. *Nature* 2012;**483**:355−8.

43. Jiang C, He BB, Zhao RL, Xu MJ, Zhao YL. Computational exploration of how enzyme XimE converts natural *S*-Epoxide to pyran and *R*-Epoxide to furan. *ACS Catal* 2021;**11**:7928−42.

44. Gu M, Wang M, Guo J, Shi C, Deng J, Huang L, et al. Crystal structure of CYP76AH1 in 4-PI-bound state from *Salvia miltiorrhiza*. *Biochem Biophys Res Commun* 2019;**511**:813−9.