


ARTICLE

<https://doi.org/10.1038/s41467-019-11987-z>

OPEN

Tropane alkaloids biosynthesis involves an unusual type III polyketide synthase and non-enzymatic condensation

Jian-Ping Huang^{1,4}, Chengli Fang ^{2,4}, Xiaoyan Ma ^{1,3,4}, Li Wang ¹, Jing Yang¹, Jianying Luo¹, Yijun Yan¹, Yu Zhang ² & Sheng-Xiong Huang ¹

The skeleton of tropane alkaloids is derived from ornithine-derived *N*-methylpyrrolinium and two malonyl-CoA units. The enzymatic mechanism that connects *N*-methylpyrrolinium and malonyl-CoA units remains unknown. Here, we report the characterization of three pyrrolidine ketide synthases (PYKS), *Aa*PYKS, *Ds*PYKS, and *Ab*PYKS, from three different hyoscyamine- and scopolamine-producing plants. By examining the crystal structure and biochemical activity of *Aa*PYKS, we show that the reaction mechanism involves PYKS-mediated malonyl-CoA condensation to generate a 3-oxo-glutaric acid intermediate that can undergo non-enzymatic Mannich-like condensation with *N*-methylpyrrolinium to yield the racemic 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoic acid. This study therefore provides a long sought-after biosynthetic mechanism to explain condensation between *N*-methylpyrrolinium and acetate units and, more importantly, identifies an unusual plant type III polyketide synthase that can only catalyze one round of malonyl-CoA condensation.

¹State Key Laboratory of Phytochemistry and Plant Resources in West China, and CAS Center for Excellence in Molecular Plant Sciences, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China. ²The Key Laboratory of Synthetic Biology, and CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China. ³School of Chemistry and Chemical Engineering, Sichuan University of Science & Engineering, Zigong 643000, China. ⁴These authors contributed equally: Jian-Ping Huang, Chengli Fang, Xiaoyan Ma. Correspondence and requests for materials should be addressed to Y.Z. (email: y Zhang@sippe.ac.cn) or to S.-X.H. (email: sx Huang@mail.kib.ac.cn)

Tropane alkaloids (TA) are an important class of secondary metabolites with a characteristic 8-azabicyclo[3.2.1]octane ring in their chemical structures, and occur mainly in many members of the plant families Erythroxylaceae and Solanaceae^{1,2}. Among the >200 known TA, the anticholinergic drugs atropine (the racemic hyoscyamine), and scopolamine as well as the stimulant cocaine (Fig. 1a) are most prominent and have been used for a long time as medicinals worldwide^{2,3}.

The carbon atom source of TA is known from the amino acid ornithine^{4–7} and acetate units^{8–10} for a long time by elaborate isotope-labeling experiments. In recent 20 years, biochemical studies have demonstrated that the intermediate *N*-methylpyrrolinium cation (**1**) was biosynthesized from amino acid ornithine by three enzymes, the ornithine decarboxylase (ODC)¹¹, putrescine *N*-methyltransferase^{12,13}, and *N*-methylputrescine oxidase (MPO)^{14–16} (Fig. 1b). Very recently, tropinone biosynthesis has been investigated¹⁷, and it has been shown that the key intermediate 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoic acid (**2**)¹⁸ results from an atypical type III polyketide synthase (PKS) *AbPYKS* catalyzed condensation between **1** and malonyl-CoA, whereas tropinone is formed by a P450 enzyme *AbCYP82M3*-mediated oxidation and cyclization of racemic **2**. However, this very unusual PYKS-mediated condensation mechanism between **1** and acetate units still remains unknown. Herein, we report the characterization of three PYKSs, *AaPYKS*, *DsPYKS*, and *AbPYKS*, from three respective plants, *Anisodus acutangulus*, *Datura stramonium*, and *Atropa belladonna* as well as the condensation mechanism based on the crystal structures of *AaPYKS*.

Results

Identification of type III PKS genes from three solanaceous plants. The hyoscyamine and scopolamine are found in a variety of solanaceous plants including *A. acutangulus*, *A. belladonna*, and *D. stramonium*^{1,2,19}. Their biosynthesis occurs in the plant roots^{20,21}, which provides the advantage to study the biosynthesis using the plant hairy roots culture²². To facilitate gene discovery efforts, we constructed the hairy roots of *A. acutangulus*, *A. belladonna*, and *D. stramonium*, and confirmed the production of hyoscyamine and scopolamine by HPLC analysis (Supplementary Fig. 1). Previous labeling studies^{4–10,23–25} demonstrated that the second ring of TA originates from two acetate units indicating the involvement of PKS. De novo transcriptome assembly of three

established hairy roots was generated using the Hiseq platform. By a combination of direct screening of the transcriptome assembly annotations, BLAST searches and rapid amplification of cDNA ends, full-length transcripts of four, five, and two putative type III PKS unigenes were obtained from three hairy roots, respectively (Supplementary Table 1). Subsequent phylogenetic tree (Supplementary Fig. 2) and amino-acid sequence alignment (Supplementary Fig. 3) indicated that *AbPKS2*, *AaPKS4*, *AaPKS2*, *AbPKS3*, and *DsPKS1* from three different plants, were in a clade more divergent with others and the amino-acid sequence of *AbPKS3* is identical with the reported *AbPYKS*¹⁷. This information suggests that the five PKSs are presumably involved in TA biosynthesis.

Determining the enzyme activity of putative PYKS. In order to functionally assign these putative type III PKSs, the five proteins were heterologously expressed in *Escherichia coli* (*E. coli*) (Supplementary Fig. 4) and analyzed in vitro for their activity toward putative substrates, chemically synthesized **1** (Supplementary Fig. 5) and malonyl-CoA. Interestingly, liquid chromatography–mass spectrometry (LC-MS) analysis showed the appearance of a new peak at *m/z* 186 in the *AaPKS2*, *AbPYKS*, and *DsPKS1* catalyzed reactions that was not present in *AbPKS2* (Fig. 2 and Supplementary Fig. 6a) (*AaPKS4* forms inclusion bodies and was thus not characterized further). This new peak has the same MS and retention time as that of chemically synthesized **2** (Fig. 2 and Supplementary Fig. 7). Therefore, we established that *AaPKS2* and *DsPKS1* are involved in TA biosynthesis and named them as *AaPYKS* and *DsPYKS*, respectively¹⁷. However, the recent Barry and co-workers' work did not determine the absolute configuration of the enzymatic product **2** owing to its inherent instability^{17,18}. To solve this problem, we scaled up the enzymatic reactions and treated the product **2** with SOCl₂ either in methanol or benzyl alcohol to afford the expected esters (**3** and **4**) (Fig. 1c and Supplementary Fig. 8). Then, the chiral and racemic **3** and **4** were chemically synthesized as standards for comparisons (Supplementary Fig. 8). LC-MS analysis revealed both the methyl and the benzyl product of **2** matched the standards **3** and **4** (Supplementary Fig. 9). Whereas, the optical rotation data of enzymatic product **4** is around –4.5 comparing to the value of –31.2 for synthesized *S*-configuration **4** (Supplementary Fig. 7b), suggesting that the enzymatic product **2** is a racemate.

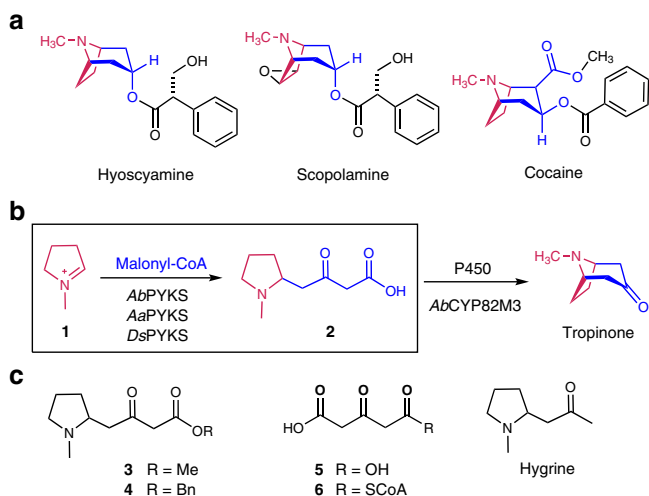


Fig. 1 The chemical structures mentioned in this study and the biosynthetic pathway of tropinone. **a** The tropane alkaloids. **b** The tropinone biosynthesis pathway with the current study highlighted in the box. **c** The chemical structures of other compounds

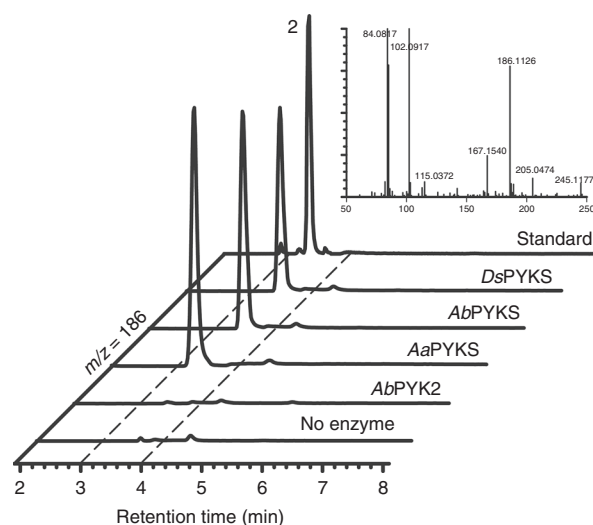


Fig. 2 LC-MS chromatograms at $[M + H]^+ = 186$ of product **2** in enzymatic reactions for *DsPYKS*, *AbPYKS*, *AaPYKS* or *AbPKS2* using **1** and malonyl-CoA as substrates

Table 1 Data collection and refinement statistics

	<i>AaPYKS-COB</i>	<i>AaPYKS-6</i>
Data collection		
Space group	<i>P</i> ₃ ₂ ₁	<i>P</i> ₃ ₂ ₁
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	108.1, 108.1, 191.1	108.4, 108.4, 191.5
α β γ (°)	90, 90, 120	90, 90, 120
Resolution (Å)	50.00–2.00 (2.03–2.00)	50.00–2.53 (2.57–2.53)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.104 (0.926)	0.092 (0.472)
<i>I</i> / σ <i>I</i>	24.9 (2.7)	34.1 (8.0)
Completeness (%)	100 (100)	100 (100)
Redundancy	11.0 (10.1)	16.4 (16.4)
Refinement		
Resolution (Å)	47.06–2.00	47.15–2.53
No. reflections	85,063	43,997
<i>R</i> _{work} / <i>R</i> _{free}	0.175/0.194	0.164/0.191
No. atoms	6466	6296
Protein	5805	5778
Ligand/ion	34	114
Water	627	404
B-factors		
Protein	27.3	32.9
Ligand/ion	39.8	63.9
Water	37.2	38.6
R.m.s. deviations		
Bond lengths (Å)	0.01	0.01
Bond angles (°)	1.14	1.45

*Values in parentheses are for highest-resolution shell

Structural basis for the PYKS catalytic mechanism. Plant type III PKSs are a family of enzymes known to catalyze the iterative decarboxylative condensation of malonyl-CoA upon CoA-tethered substrates and could generate a variety of natural products with aromatic ring^{26,27}. In this study, the type III PKSs were characterized to react onto the non-CoA-tethered substrate **1** to give a non-aromatic product, indicating a potential new family of type III PKS. To understand the structural basis for the catalytic mechanism, we determined a crystal structure of *AaPYKS* at 2.0 Å (*AaPYKS-COB*, Table 1; PDB ID: 6J1M). The overall structure of the dimeric *AaPYKS-COB* is highly homologous to those of previously reported type III PKSs (Fig. 3a and Supplementary Fig. 10)^{26,27}. The surface exposed characteristic CoA-binding tunnel reaches into the inner active center, where the conserved catalytic triad C166-H305-N338 resides as those in other type III PKSs. Surprisingly, we observed a large electron density extending from the thiol moiety of C166 (Fig. 3b). Inspired by previous structure of pentaketide chromone synthase (PCS) showing a co-purified CoASH in the active center²⁸, we suspect that the *AaPYKS-COB* prepared from *E. coli* cells probably utilized the endogenous malonyl-CoA to yield a reaction intermediate covalently bound to the thiol group. To confirm our hypothesis, we explored the possible covalent modification on C166 of purified *AaPYKS-COB* through mass spectrometry. The results in Supplementary Fig. 11 clearly show the presence of a C166-containing peptide with an additional mass of 128 Da, corresponding to exact mass of a 4-carboxy-3-oxobutanoyl (COB) moiety (Fig. 3c). The 4-carboxy-3-oxobutanoyl thioester could be readily fit into the electron density (Fig. 3b).

The crystal structure of *AaPYKS-COB* also graphically explains the limited round of elongation *AaPYKS* could catalyze. In the structure, the 4-carboxy-3-oxobutanoyl moiety is stabilized by multiple polar interactions with R134, H305, N338, and S340 and hydrophobic interactions with M139, F217, L256, and L258 (Fig. 3c). Notably, R134 and S340 make a salt-bridge and

hydrogen bond interactions with the carboxy group of the thioester, respectively, which serves as a gate to prevent entry to the buried pocket near the active center, and consequently inhibits further chain elongation (Fig. 3d). Alanine, threonine, or serine substitution of R134 resulted in significant decrease of activity, highlighting the key role of R134 for *AaPYKS*-catalyzed one-round malonyl-CoA condensation (Supplementary Fig. 6b). To explore the necessity of hydrogen bonding between S340 and the carboxy group of the thioester, amino acids without hydroxyl group (leucine, glycine, and valine) or having more steric hindrance (leucine and valine) were selected for site mutation of S340. All the three mutations showed much decreased activity (Supplementary Fig. 6c), showing the importance of hydrogen bonding to the enzymatic activity. In addition, the presence of a leucine residue in the non-functional *AbPKS2* at the corresponding S340 position of *AaPYKS* also supports the key role for S340.

Our *in vitro* enzymatic experiment detects rapid conversion of malonyl-CoA to 3-oxo-glutaric acid (**5**) by *AaPYKS* in the absence of **1** (Supplementary Fig. 12), implicating that malonyl-CoA could help release the covalently bound intermediate and reactivates the enzyme. We therefore soaked the crystal of *AaPYKS-COB* with high concentration of malonyl-CoA and determined another structure at 2.5 Å (*AaPYKS-6*; PDB ID: 6J1N; Table 1 and Supplementary Fig. 10b). Unexpectedly, we observed an even larger electron density than that in *AaPYKS-COB* in the tunnel, the upper part of which is identical to that of COB moiety in the *AaPYKS-COB* structure and the lower part of which fills up the CoA tunnel (Fig. 3e, f). The expected intermediate 4-carboxy-3-oxobutanoyl-CoA (**6**) could be readily fitted into the density. In the structure of *AaPYKS-6*, the carboxyl moiety makes the same interactions as in structure of *AaPYKS-COB* and the sulfur of C166 is close to the C1 and C3 atoms of **6** (Fig. 3e, f). The structure suggests we trapped another intermediate **6** and the C166 is ready to attack the thiol ester carbonyl to recapture the COB moiety. To our knowledge, this is the first type III PKS structure trapped with elongated and CoA-tethered β -keto intermediate²⁹.

The crystal structures clearly suggest that the PYKSs take the solo malonyl-CoA as both starter and extender unit. This feature is similar to that of the PCS and octaketide synthase (OKS) from *Aloe arborescens*^{28,30,31}, which have been established as a novel class of plant type III PKSs because both enzymes accept the sole malonyl-CoA as both starter and extender unit to catalyze five- or eight-round malonyl-CoA condensations. Consistently, *AaPYKS*, *AbPYKS*, and *DsPYKS* all have the characteristic leucine residue at 258, 259, and 259, respectively, corresponding to Leu266 in both PCS and OKS (Supplementary Fig. 3), which was proposed to determine the selectivity of the starter molecule from traditional *p*-coumaroyl-CoA to the sole malonyl-CoA unit^{28,32}. Alanine substitution of *AaPYKS* L258 led to remarkable enzymatic activity reduction (Supplementary Fig. 6). This information also suggests that all the three PKSs catalyze the sole malonyl-CoA condensation to generate **5**, then condense with **1** to yield **2** rather than **1** reacts with malonyl-CoA first (Supplementary Fig. 13). This unique order of condensations is also preferred by the feeding experiments-based TA biosynthesis by Robins et al.²³.

Non-enzymatic condensation **1 with **5**.** To determine how the non-CoA-tethered substrate **1** interacts with the enzyme, high concentration of **1** was soaked into the crystal of *AaPYKS-COB*. However, the active center of the resulting crystal structure was highly similar to that of the structure *AaPYKS-COB*; and the C166 is still covalently bound with the COB moiety (Supplementary Fig. 14a). This result suggests that **1** may not be capable

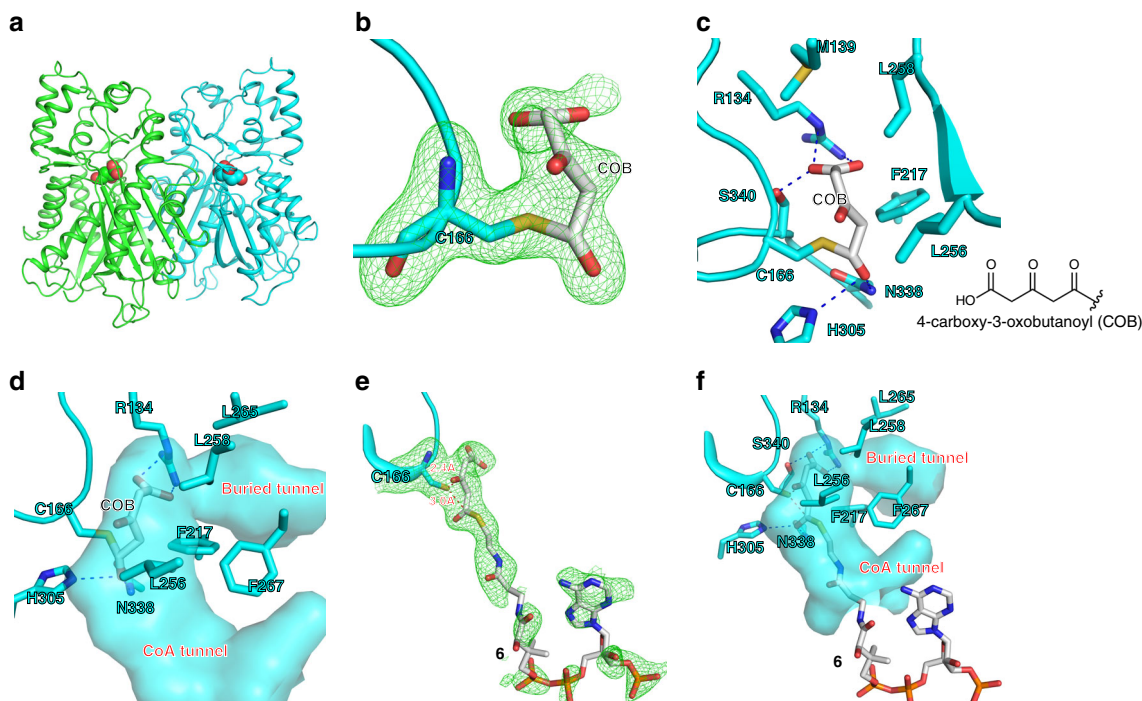


Fig. 3 Crystal structures of AaPYKS-COB (PDB ID: 6J1M) and AaPYKS-6 (PDB ID: 6J1N). **a** AaPYKS-COB is a dimer (cyan and green ribbon). The 4-carboxy-3-oxobutanoyl (COB) thioester is shown as spheres. **b** The $F_{O}-F_{C}$ electron density polder map contoured at 5σ of the catalytic residue C166 and the covalently bound COB. **c** The detailed interactions of COB with AaPYKS residues. **d** The surface presentation of AaPYKS shows the inner tunnels. **e** The $F_{O}-F_{C}$ electron density polder map contoured at 3σ of the catalytic residue C166 and **f** The surface presentation of AaPYKS-6 shows the inner tunnels. N, O, S, and P atoms are colored as blue, red, yellow, and orange, respectively. C atoms of protein residues and small molecules are colored as cyan and white, respectively

of releasing the COB moiety from C166 in the active center. Above information in conjunction with the production of racemic **2** prompts us to propose that the condensation **1** with **5** is non-enzymatic. We next sought to obtain evidence for the non-enzymatic reactions by performing *in vitro* reactions between the intermediate **5** and **1** in the conditions of with and without enzyme AaPYKS, respectively. As expected, both conditions did yield the same product **2**, and the kinetic analysis showed that there is no reaction velocity difference for the conditions of with or without enzyme (Supplementary Fig. 14b). These results clearly demonstrate that the racemic **2** is afforded in a non-enzymatic Mannich-like condensation. Hygrine (Fig. 1c) is a decarboxylative product of **2** biosynthetically^{17,18}. Coincidentally, all the isolated hygrine from different plants is racemic³³, corroborating the racemic feature of **2** from spontaneous condensation in TA biosynthesis.

Discussion

Based on the three structures and the *in vitro* reactions, we proposed a detailed mechanism for the identified PYKSs (Fig. 4). First, the AaPYKS runs one-round malonyl-CoA condensation using conserved Cys₁₆₆-His₃₀₅-Asn₃₃₈ catalytic triad to afford **6** in the active site (Fig. 3f), then the thiol group of Cys₁₆₆ attacks the thiol ester carbonyl of **6** (Fig. 3c) to recapture the COB moiety. Subsequently one molecule of water could come into the active site to form the hydrogen bonds with His₃₀₅ and the carbonyl as the case of the plant PKS BAS-mediated reaction³⁴. The nucleophilic attack to thiol ester by this water would yield **5**. Additional malonyl-CoA will come into the active site to push **5** out of the pocket to undergo the spontaneous condensation with **1** to yield **2**.

The mechanism of Mannich-like condensation between **1** and acetate units has intrigued chemists for two decades^{23–25}. Robert

Robinson finished the chemical synthesis of tropinone in one pot by the addition of succinaldehyde, methylamine, and salt of 3-oxo-glutaric acid (**5**) at physiological pH condition with the 42% yield in 1917 (Supplementary Fig. 15)³⁵, which has been widely applauded and has become the first synthesis classic. This classic synthesis route stimulated a proposal by Robinson, suggesting the biosynthesis of tropinone might occur via an analogous route involving the pyrrolidine ring moiety and **5** to furnish the tropane ring¹⁸. In this study, we clearly demonstrated that AaPYKS only takes one-round of malonyl-CoA condensation to form free acid **5**, and then undergo the non-enzymatic Mannich-like condensation with the substrate *N*-methylpyrrolinium (**1**) to afford **2** (Fig. 4). This biosynthetic route is analogous to the Robinson's proposal and suggests that the classic synthetic route reported > 100 years ago is biomimetic. Overall, our experiments have solved the longstanding question for reaction mechanism between **1** and acetate units in TA biosynthesis.

There is no reported type III PKS that takes one round of sole malonyl-CoA condensation in natural product biosynthesis, thereby, PYKS may represent an emerging class of plant type III PKSs. In addition, the key step of granatane alkaloids (GA)³⁶ and lycopodium alkaloids^{37,38} biosynthesis was proposed to be similar with that of TA¹⁷. Our findings suggest a unifying biosynthetic mechanism for a wide range of plant alkaloids including granatane and lycopodium families (Supplementary Fig. 16). Therefore, the mechanism depiction could provide a molecular blueprint for generating structure diversified TA, GA, and lycopodium alkaloids.

Methods

General experimental procedures. All moisture or oxygen-sensitive reactions were carried out under an argon or nitrogen atmosphere in oven or heat-dried flasks. The solvents used were purified by distillation over the drying agents

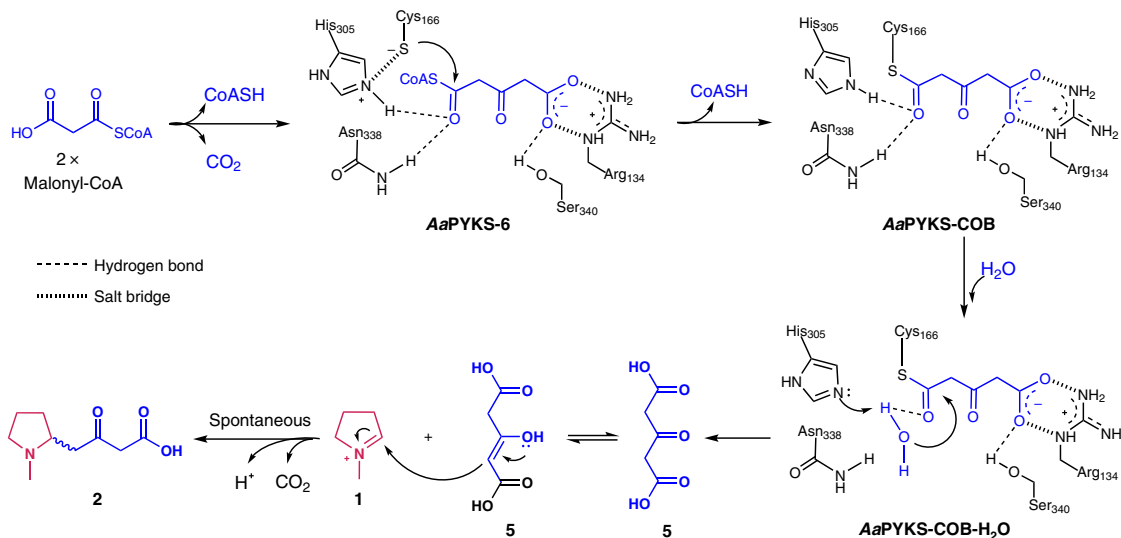


Fig. 4 Proposed mechanism for *AaPYKS*-catalyzed reactions

indicated and were transferred under argon: THF (Na), CH_2Cl_2 (CaH_2), MeOH (Mg), NEt_3 (CaH_2). All reactions were monitored by thin-layer chromatography on silica gel F254 plates using UV light as visualizing agent (if applicable), and a solution of phosphomolybdic acid (50 g L^{-1}) in EtOH followed by heating as developing agents. The products were purified by flash column chromatography on silica gel (200–300 meshes from the Anhui Liangchen Silicon Material Company in China). ^1H NMR and ^{13}C NMR spectra were recorded in D_2O , CDCl_3 or acetone- d_6 solution on a Bruker AM 400 MHz instrument. Chemical shifts were denoted in ppm (δ), and calibrated by using residual undeuterated solvent D_2O (4.79 ppm), CDCl_3 (7.27 ppm), acetone- d_6 (2.05 ppm) or tetramethylsilane (0.00 ppm) as internal reference for ^1H NMR and the deuterated solvent CDCl_3 (77.00 ppm), acetone- d_6 (29.8 ppm) or tetramethylsilane (0.00 ppm) as internal standard for ^{13}C NMR. The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, br = broad, brs = broad singlet, m = multiplet. Optical rotation data was measured on a Rudolph Autopol VI Automatic Polarimeter. The high-resolution mass spectral analysis (HRMSIMS) data were measured on Agilent G6230 Q-TOF mass instrument (Agilent Corp., USA) by means of the ESI technique. LC-MS analysis was conducted on AGILENT 1290/6530 system. HPLC analysis was conducted on a HITACHI Chromaster system equipped with a DAD detector, a YMC-Triart C_{18} column (250 mm \times 4.6 mm i.d., 5 μm), and a flow rate of 1.0 mL min^{-1} at a column temperature of 25 $^\circ\text{C}$.

Plant materials. The seeds of *A. acutangulus*, *A. belladonna*, and *D. stramonium* were collected from Yunnan province, Hunan province, and Heilongjiang province of China, respectively. The hairy root induction was performed using young plant tissues³⁹. In brief, leaf disks and stem segments prepared from 2–6-week-old sterilized plant seedlings were inoculated with *Agrobacterium tumefaciens* strain C58C1 (pRiA4) suspended in liquid Murashige and Skoog (MS) medium containing 30 g L^{-1} sucrose and $100\text{ }\mu\text{M}$ acetosyringone, incubated in darkness on co-cultivation solid $\frac{1}{2}$ MS medium containing 30 g L^{-1} sucrose and $100\text{ }\mu\text{M}$ acetosyringone, disinfected on solid $\frac{1}{2}$ MS medium containing 30 g L^{-1} sucrose and 0.5 g L^{-1} cefotaxime, and sub-cultured in solid or liquid $\frac{1}{2}$ MS medium with 30 g L^{-1} sucrose.

Scopolamine and hyoscyamine analysis. Scopolamine and hyoscyamine were extracted from 40 $^\circ\text{C}$ dried hairy roots⁴⁰. An extraction solvent of chloroform:methanol:25% ammonia (15:5:1, v/v/v; 100 mL g^{-1} sample) was added to the weighed sample (200 mg), vortexed, and sonicated for 30 min, and then centrifuged for 30 min at room temperature (rt). The supernatant was evaporated to dryness at 40 $^\circ\text{C}$, and the resulting sample was dissolved in 5 mL of 0.5 M sulfuric acid and 15 mL chloroform. The sulfuric acid phase was collected and adjusted to pH ~ 10 with ice-bathed 25% ammonia, and then alkaloids were extracted from the sulfuric acid phase with 5 mL of chloroform thrice. The combined chloroform phase was evaporated to dryness at 40 $^\circ\text{C}$, and the resulting sample was dissolved in 1 mL methanol for HPLC analysis. The mobile phase was made up from 57% or 59% HPLC gradient methanol and 43% or 41% HPLC gradient water (containing 50 mM ammonium acetate and adjust pH to 4.6 by acetic acid). The chromatogram was monitored by detecting the absorbance at 215 nm. Standards for hyoscyamine and scopolamine (Yuanye, Shanghai, China) were prepared in methanol at a final concentration of 1 mg mL^{-1} . Hyoscyamine and scopolamine peak in samples were identified by comparing their retention time and UV spectra with those of standards and confirmed by MS spectra. Quantitative analysis was performed using standard curves of hyoscyamine and scopolamine based on the peak area.

Time-course culture of hairy roots and de novo transcriptome sequencing.

Fresh hairy roots (0.2 g) were transferred to 100 mL liquid $\frac{1}{2}$ MS medium with 30 g L^{-1} sucrose and maintained in darkness at 25 $^\circ\text{C}$ on a rotary shaker (110 rpm.). Samples of cultured hairy roots were collected every 5 days, and the fresh weights, dry weights, and the contents of TA were determined in triplicates. To maximize recovery of genes involved in TA biosynthesis, the 5d-old hairy roots showing the highest increase and the 15d-old hairy roots achieving maximum alkaloid content were selected as transcriptome study materials based on the time course of alkaloid production. Total RNAs of the representative 5d-old and 15d-old hairy roots were extracted, and cDNA libraries were generated and sequenced on a HiSeq platform (Personalbio, Shanghai, China). De novo assembled transcriptomes of *A. acutangulus* hairy roots, *A. belladonna* hairy roots, and *D. stramonium* hairy roots were generated using Trinity (r20140717, K-mer 25 bp), and 229,401 unigenes, 163,301 unigenes, and 103,470 unigenes were obtained, respectively. Databases including NR, GO, KEGG, eggNOG, and SwissProt were used for unigene function annotation.

Gene cloning of plant type III PKSs. Total RNAs were extracted from hairy roots with Total RNA isolation kit (Promega), and quantified using NanoDrop 2000C (Thermo Scientific). In all, 1 μg of total RNA was used to prepare cDNA using SMARTer RACE 5'/3' Kit (Clontech Laboratories, Inc.). Full-length *AaPKSs*, *AbPKSs*, and *DsPKSs* were obtained using 5'-RACE and 3'-RACE PCR with primers listed in Supplementary Table 2.

Multiple sequence alignment and phylogenetic analysis. Sequences of *AaPKSs*, *AbPKSs*, *DsPKSs*, and representative plant and bacterial type III PKSs were aligned by MUSCLE implemented in MEGA version 7.0.14⁴¹, and visualized by GeneDoc version 2.7. The phylogenetic tree was generated with MEGA version 7.0.14 using the Maximum Likelihood method, a bootstrap test of 2000 replicates, and the Jones-Taylor-Thornton model using default parameters.

Protein expression and purification. The full-length sequences of *AaPYKS*, *AbPYKS*, *AbPKS2*, *DsPYKS* were amplified from hairy root cDNA respectively using the primers listed in Supplementary Table 2. The PCR products were confirmed by agarose gel electrophoresis, purified by gel extraction kit (omega, D2500-02), digested by *EcoRI/Sal I* (*AaPYKS*) or *BamHI/SalI* (*AbPYKS*, *AbPKS2*, and *DsPYKS*), ligated to pET28a vector digested by the corresponding restriction enzymes using T4 ligase, and transformed into chemically competent *E. coli* cells DH5 α . Recombinant colonies were selected on LB agar plates supplemented with kanamycin ($50\text{ }\mu\text{g mL}^{-1}$). Positive clones were identified by colony PCR and the corresponding plasmids were isolated and confirmed by sequencing and restriction enzyme digestion. Site mutated pET28a-*AaPYKSs* were obtained by fusion PCR using wild pET28a-*AbPYKS* as template and the primers are listed in Supplementary Table 2.

The *E. coli* Rosetta (DE3) cells containing pET28a-*AaPYKS*, *-AbPYKS*, *-AbPKS2*, *-DsPYKS* or site mutated *AaPYKS* were cultured in LB liquid medium supplemented with kanamycin ($50\text{ }\mu\text{g mL}^{-1}$) and chloramphenicol ($25\text{ }\mu\text{g mL}^{-1}$) at 37 $^\circ\text{C}$ in a shaker at 200 rpm. The protein expression was induced with IPTG (0.1 mM for *AbPYKS*, *AbPKS2*, and *DsPYKS*; 0.5 mM for wild *AaPYKS* and site mutated *AaPYKS*) for 18–20 h at 16 $^\circ\text{C}$ when OD_{600} reached 0.6.

Cell pellets were collected at 4 $^\circ\text{C}$ by centrifugation and re-suspended in 50 mL ice-cold Buffer A (15 mM imidazole, 50 mM Tris, 300 mM NaCl, 10% glycerol,

The large-scale enzyme reaction was performed as described in the Synthesis of chemical compounds **4** and **2** part in the presence of $6\ \mu\text{M}$ AaPYKS. The benzyl modification product of **2** was purified by HPLC using the same gradient elution program as that for the LC-MS analysis, and its stereochemistry was studied. $[\alpha]_{\text{D}}^{37} = -4.5$ (*c* 0.40, H_2O).

Crystallization and structure determination of AaPYKS. For preparation of AaPYKS, *E. coli* Rosetta (DE3) cells carrying pET28a-AaPYKS were cultured in LB at 37°C , and the expression of N-terminal 6xHis-tagged AaPYKS was induced at 18°C for 16 h with $0.5\ \text{mM}$ IPTG at OD₆₀₀ of 0.8. Cells were harvested by centrifugation ($8000\ \text{g}$, 4°C), re-suspended in lysis buffer ($50\ \text{mM}$ Tris-HCl pH 8.0, $0.3\ \text{M}$ NaCl, 5% (*v/v*) glycerol, $5\ \text{mM}$ β -mercaptoethanol, protease inhibitor cocktail (bimake.cn)) and lysed using an Avestin EmulsiFlex-C3 cell disrupter (Avestin, Inc.). The lysate was centrifuged ($16,000\ \text{g}$; $45\ \text{min}$; 4°C) and the supernatant was loaded on to a $2\ \text{mL}$ column packed with Ni-NTA agarose (SMART, Inc.). The protein was washed by lysis buffer containing 20 , $40\ \text{mM}$ imidazole and eluted with lysis buffer containing $500\ \text{mM}$ imidazole. The eluted fractions were loaded on a HiLoad 16/60 Superdex S200 column (GE Healthcare, Inc.) equilibrated in $10\ \text{mM}$ Tris-HCl pH 8.0, $0.1\ \text{M}$ NaCl, 1% (*v/v*) glycerol, $1\ \text{mM}$ DTT. Fractions containing AaPYKS were collected, concentrated to $15\ \text{mg mL}^{-1}$, and stored at -80°C .

Crystals of AaPYKS-COB were grown by vapor diffusion at 4°C in $2\ \mu\text{L}$ drops containing a 1:1 mixture of $15\ \text{mg mL}^{-1}$ protein and crystallization buffer (1% w/v Tryptone, $0.05\ \text{M}$ HEPES sodium 7.0 , 12% PEG3,350 (*w/v*), $1\ \text{M}$ Sodium azide). AaPYKS-COB crystals grown 7 days were harvested for X-ray diffraction data collection. AaPYKS-6 crystals were obtained by soaking the AaPYKS-COB crystals in crystallization buffer containing $10\ \text{mM}$ malonyl-CoA. Crystals were transferred and stabilized in crystallization buffer containing 21% ethylene glycol and cooled in liquid nitrogen.

Data were collected at Shanghai Synchrotron Radiation Facility beamlines 17U and 19U1, processed using HKL2000⁴⁵. The structure was solved by molecular replacement with Phaser MR⁴⁶ using the structure of chalcone synthase (PDB: 1B15) as the search model. Cycles of iterative model building and refinement were performed in Coot⁴⁷ and Phenix⁴⁸.

LC-MS/MS analysis of AaPYKS. AaPYKS protein ($50\ \mu\text{g}$) was loaded onto a PD Minitrap G-25 column (GE Healthcare, USA) equilibrated with water to remove salt contamination. The AaPYKS fractions were collected and digested with $1\ \mu\text{g}$ of trypsin (Promega) in $50\ \mu\text{L}$ of $50\ \text{mM}$ NH_4HCO_3 for 16 h at 37°C . The samples were subsequently evaporated to dryness. Peptides were then analyzed in an Orbitrap Fusion Lumos Mass spectrometer (Thermo Scientific, USA). Peptide fragmentation was performed via higher-energy collision dissociation. Data processing was performed using Proteome Discoverer 2.1 software (Thermo Scientific, USA) and peptide sequences were determined by matching protein databases with the acquired fragmentation pattern by SEQUEST HT algorithm. The precursor mass tolerance was set to $10\ \text{ppm}$ and fragment ion mass tolerance to $0.05\ \text{Da}$.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

RNA-Seq data that support the findings of this study have been deposited in National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) with accession numbers SRR9888534, SRR9888536, and SRR9888538. The GenBank accession numbers for AaPYKS, DsPYKS, and AbPKS2 are MN025472, MN025473, and MN025474, respectively. The PDB IDs for AaPYKS-COB and AaPYKS-6 are 6J1M and 6J1N, respectively. The authors declare that all other relevant data supporting the findings of this study are available within the article, its Supplementary Information, and Source Data files.

Received: 2 June 2019 Accepted: 14 August 2019

Published online: 06 September 2019

References

- Griffin, W. J. & Lin, G. D. Chemotaxonomy and geographical distribution of tropane alkaloids. *Phytochemistry* **53**, 623–637 (2000).
- Fodor, G. & Dharanipragada, R. Tropane alkaloids. *Nat. Prod. Rep.* **7**, 539–548 (1990).
- Gryniewicz, G. & Gadzikowska, M. Tropane alkaloids as medicinally useful natural products and their synthetic derivatives as new drugs. *Pharmacol. Rep.* **60**, 439–463 (2008).
- Leete, E., Marion, L. & Spenser, I. D. Biogenesis of hyoscyamine. *Nature* **174**, 650–651 (1954).
- Leete, E. The stereospecific incorporation of ornithine into the tropane moiety of hyoscyamine. *J. Am. Chem. Soc.* **84**, 55–57 (1962).
- Leete, E. Biosynthesis of the pyrrolidine rings of cocaine and cuscohygrine from [$5\text{-}^{14}\text{C}$]-ornithine via a symmetrical intermediate. *J. Am. Chem. Soc.* **104**, 1403–1408 (1982).
- Hashimoto, T., Yamada, Y. & Leete, E. Species-dependent biosynthesis of hyoscyamine. *J. Am. Chem. Soc.* **111**, 1141–1142 (1989).
- Leete, E. & Kim, S. H. A revision of the generally accepted hypothesis for the biosynthesis of the tropane moiety of cocaine. *J. Am. Chem. Soc.* **110**, 2976–2978 (1988).
- Leete, E., Bjorklund, J. A., Couladis, M. M. & Kim, S. H. Late intermediates in the biosynthesis of cocaine: 4-(1-Methyl-2-pyrrolidinyl)-3-oxobutanoate and methyl ecgonine. *J. Am. Chem. Soc.* **113**, 9286–9292 (1991).
- Hemscheidt, T. & Spenser, I. D. Biosynthesis of β -hydroxytropane in *Datura stramonium*: nonregiospecific incorporation of [$1,2\text{-}^{13}\text{C}_2$]acetate. *J. Am. Chem. Soc.* **114**, 5472–5473 (1992).
- Docimo, T. et al. The first step in the biosynthesis of cocaine in *Erythroxylum coca*: The characterization of arginine and ornithine decarboxylases. *Plant Mol. Biol.* **78**, 599–615 (2012).
- Hashimoto, T., Yun, D. J. & Yamada, Y. Putrescine and putrescine *N*-methyltransferase in the biosynthesis of tropane alkaloids in cultured roots of *Hyoscyamus albus*. *Planta* **178**, 123–130 (1989).
- Biastoff, S., Brandt, W., Dräger, B. & Putrescine *N*-methyltransferase—the start for alkaloids. *Phytochemistry* **70**, 1708–1718 (2009).
- Heim, W. G. et al. Cloning and characterization of a *Nicotiana tabacum* methylputrescine oxidase transcript. *Phytochemistry* **68**, 454–463 (2007).
- Katoh, A., Shoji, T. & Hashimoto, T. Molecular cloning of *N*-methylputrescine oxidase from tobacco. *Plant Cell Physiol.* **48**, 550–554 (2007).
- Friesen, J. B. & Leete, E. Nicotine synthase – an enzyme from *nicotiana* species which catalyzes the formation of (*S*)-nicotine from nicotinic acid and 1-methyl- δ -pyrrolinium chloride. *Tetrahedron Lett.* **31**, 6295–6298 (1990).
- Bedewitz, M. A., Jones, A. D., D'Auria, J. C. & Barry, C. S. Tropinone synthesis via an atypical polyketide synthase and P450-mediated cyclization. *Nat. Commun.* **9**, 5281 (2018).
- Humphrey, A. J. & O'Hagan, D. Tropane alkaloid biosynthesis. A century old problem unresolved. *Nat. Prod. Rep.* **18**, 494–502 (2001).
- Zheng, G., He, J. & Wang, S. The influence of callus differentiation of *Anisodus acutangulus* on the contents of hyoscyamine and scopolamine. *Acta Phytophysiol. Sin.* **6**, 377–385 (1980).
- Hashimoto, T. et al. Hyoscyamine β -hydroxylase, an enzyme involved in tropane alkaloid biosynthesis, is localized at the pericycle of the root. *J. Biol. Chem.* **266**, 4648–4653 (1991).
- Nakajima, K. & Hashimoto, T. Two tropinone reductases, that catalyze opposite stereospecific reductions in tropane alkaloid biosynthesis, are localized in plant root with different cell-specific patterns. *Plant Cell Physiol.* **40**, 1099–1107 (1999).
- Georgiev, M. I., Pavlov, A. I. & Bley, T. Hairy root type plant in vitro systems as sources of bioactive substances. *Appl. Microbiol. Biotechnol.* **74**, 1175–1185 (2007).
- Robins, R. J., Abraham, T. W., Parr, A. J., Eagles, J. & Walton, N. J. The biosynthesis of tropane alkaloids in *Datura stramonium*: The identity of the intermediates between *N*-methylpyrrolinium salt and tropinone. *J. Am. Chem. Soc.* **119**, 10929–10934 (1997).
- Abraham, T. W. & Leete, E. New intermediate in the biosynthesis of the tropane alkaloids in *Datura innoxia*. *J. Am. Chem. Soc.* **117**, 8100–8105 (1995).
- Newquist, M. L., Abraham, T. W. & Leete, E. Biosynthetic incorporation of ethyl (*RS*) [$2,3\text{-}^{13}\text{C}_2,3\text{-}^{14}\text{C}$]-4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoate into cuscohygrine in *Erythroxylum coca*. *Phytochemistry* **33**, 1437–1440 (1993).
- Austin, M. B. & Noel, J. P. The chalcone synthase superfamily of type III polyketide synthases. *Nat. Prod. Rep.* **20**, 79–110 (2003).
- Abe, I. & Morita, H. Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. *Nat. Prod. Rep.* **27**, 809–838 (2010).
- Morita, H. et al. Structural insight into chain-length control and product specificity of pentaketide chromone synthase from *Aloe arborescens*. *Chem. Biol.* **14**, 359–369 (2007).
- Shimizu, Y., Ogata, H. & Goto, S. Type III polyketide synthases: functional classification and phylogenomics. *ChemBioChem* **18**, 50–65 (2017).
- Abe, I. et al. A plant type III polyketide synthase that produces pentaketide chromone. *J. Am. Chem. Soc.* **127**, 1362–1363 (2005).
- Abe, I., Oguro, S., Utsumi, Y., Sano, Y. & Noguchi, H. Engineered biosynthesis of plant polyketides: chain length control in an octaketide-producing plant type III polyketide synthase. *J. Am. Chem. Soc.* **127**, 12709–12716 (2005).
- Abe, I., Watanabe, T., Morita, H., Kohno, T. & Noguchi, H. Engineered biosynthesis of plant polyketides: manipulation of chalcone synthase. *Org. Lett.* **8**, 499–502 (2006).
- McGaw, B. A. & Woolley, J. G. Stereochemistry of tropane alkaloid formation in *Datura*. *Phytochemistry* **17**, 257–259 (1978).

34. Morita, H. et al. A structure-based mechanism for benzalacetone synthase from *Rheum palmatum*. *Proc. Natl. Acad. Sci. USA* **107**, 669–673 (2010).
35. Robinson, R. A synthesis of tropinone. *J. Chem. Soc. Trans.* **111**, 762–768 (1917).
36. Kim, N., Estrada, O., Chavez, B., Stewart, C. Jr & D'Auria, J. C. Tropane and granatane alkaloid biosynthesis: a systematic analysis. *Molecules* **21**, 1510 (2016).
37. Hemscheidt, T. & Spenser, I. D. Biosynthesis of lycopodine: incorporation of acetate via an intermediate with C_{2v} symmetry. *J. Am. Chem. Soc.* **115**, 3020–3021 (1993).
38. Hemscheidt, T. & Spenser, I. D. A classical paradigm of alkaloid biogenesis revisited: acetonedicarboxylic acid as a biosynthetic precursor of lycopodine. *J. Am. Chem. Soc.* **118**, 1799–1800 (1996).
39. Qin, B. et al. Effects of acetylsalicylic acid and UV-B on gene expression and tropane alkaloid biosynthesis in hairy root cultures of *Anisodus luridus*. *Plant Cell Tiss. Org. Cult.* **117**, 483–490 (2014).
40. Kamada, H., Okamura, N., Satake, M., Harada, H. & Shimomura, K. Alkaloid production by hairy root cultures in *Atropa belladonna*. *Plant Cell Rep.* **5**, 239–242 (1986).
41. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
42. Nielsen, L., Lindsay, K. B., Faber, J., Nielsen, N. C. & Skrydstrup, T. Sterecontrolled synthesis of methyl silanediol peptide mimics. *J. Org. Chem.* **72**, 10035–10044 (2007).
43. Hillgren, J. M., Oberg, C. T. & Eloffson, M. Syntheses of pseudoceramines A-D and a new synthesis of spermatinamine, bromotyrosine natural products from marine sponges. *Org. Biomol. Chem.* **10**, 1246–1254 (2012).
44. Bulman Page, P. C. et al. A short and versatile route to a key intermediate for lactacystin synthesis. *Org. Lett.* **5**, 353–355 (2003).
45. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
46. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
47. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
48. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

Acknowledgements

This research was supported by the National Key R&D Program of China (2018YFA0900600), the Strategic Priority Research Program of the CAS (No. XDB27020205), the National Natural Science Foundation of China (Nos. 81522044, 31822001 and U1702285), the Key Research Program of Frontier Sciences of the CAS

(No. QYZDB-SSW-SMC051), and the Research Foundations of Yunnan Province (Nos. 2019ZF011–2, 2013HA022, and 2018HC012).

Author contributions

S.-X.H. designed the research, and J.-P.H., C.F., Y.Z., and S.-X.H. wrote the paper. J.-P.H. cloned all the genes and mutants, and carried out the in vitro chemical reactions and as well as HPLC analysis. C.F. and Y.Z. performed the protein crystallization and structure analysis. X.M. synthesized enzymatic reaction substrates and standards. X.M., L.W., J.Y., J.L., and Y.Y. analyzed the NMR data and performed the structures determination of synthesized compounds.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-019-11987-z>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Peer review information: *Nature Communications* thanks Peter Facchini, Rajan Sankaranarayanan and other anonymous reviewer(s) for their contribution to the peer review of this work.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019