Structural Basis of Cycloaddition in Biosynthesis of Iboga and Aspidosperma Alkaloids

Lorenzo Caputi[#1], **Jakob Franke**[#2], **Kate Bussey**[3], **Scott C. Farrow**[3], **Ivo Jose Curcino Vieira**[4], **Clare E. M. Stevenson**[3], **David M. Lawson**[*,3], **Sarah E. O'Connor**[*,1]

[1]Max Planck Institute of Chemical Ecology, Department of Natural Product Biosynthesis, Hans-Knöll-Straße 8, 07745 Jena, Germany

[2]Leibniz University Hannover, Centre for Biomolecular Drug Research, Schneiderberg 38, 30167 Hannover, Germany

[3]John Innes Centre, Department of Biological Chemistry, Norwich Research Park, Norwich, NR4 7UH, UK

[4]Laboratorio de Ciencias Quimicas-UENF-Campos dos Goytacazes-RJ, 28013-602, Brazil

[#] These authors contributed equally to this work.

Abstract

Cycloaddition reactions generate chemical complexity in a single step. Here we report the crystal structures of three homologous plant-derived cyclases involved in the biosynthesis of iboga and aspidosperma alkaloids. These enzymes act on the same substrate, named angryline, to generate three distinct scaffolds. Mutational analysis reveals how these highly similar enzymes control regio- and stereo-selectivity.

Aspidosperma and iboga alkaloids are plant-derived natural products with diverse biological properties, including anti-cancer and anti-addiction activities (Supplementary Fig. 1)[1]. Despite dissimilar structural architectures, these molecules are synthesized through nearly identical pathways[2]. Extensive evidence suggests that the structural divergence occurs as a result of formal Diels-Alder reactions, where "Diels-Alderases" would act on the same substrate, dehydrosecodine, to generate the iboga and aspidosperma scaffolds[3–5]. The dihydropyridine group of dehydrosecodine could act as a diene and the methyl acrylate as a dienophile to form the iboga alkaloid (+)-catharanthine (**4**). Alternatively, the dihydropyridine could act as the dienophile and the vinyl indole could act as the diene to form the aspidosperma alkaloid (–)-tabersonine (**5**; Supplementary Fig. 2).

We recently demonstrated the enzymatic formation of both (+)-catharanthine (**4**) and (–)-tabersonine (**5**) by incubating the intermediate precondylocarpine acetate (**1**) with a reductase (DPAS) and either CS or TS from the plant *Catharanthus roseus* (Fig. 1) [2]. Furthermore, we identified a reductase (DPAS) and cyclase (CorS) pair (from the plant *Tabernanthe iboga*) that react with precondylocarpine acetate to form the iboga alkaloid (–)-coronaridine (**6**), which is similar in structure to (+)-catharanthine, but has the opposite configuration as well as a reduced oxidation state (Fig. 1)[6]. These three cyclases have high amino acid identity (Supplementary Fig. 3), yet distinct product profiles, providing a system in which to probe the basis for cycloaddition product selectivity.

We first identified the substrate of CS, TS and CorS, enzymes that had been previously characterized only in coupled assays with DPAS. Neither dihydroprecondylocarpine acetate (**2**) or dehydrosecodine (**3a**, **3b**) has been isolated or characterized due to instability[3–5, 7–8]. When precondylocarpine acetate (**1**) is reacted with DPAS, a compound at *m/z* 337.19 is observed but rapidly decomposes into a variety of products (Supplementary Figs. 4-5). After extensive optimization, the enzyme product was isolated and shown to be an intramolecularly protected form of dehydrosecodine that we colloquially named "angryline" (**3c**) due to its high reactivity (Fig. 1, Supplementary Note). CD spectral analysis indicated that angryline is racemic, suggesting that equilibration between achiral dehydrosecodine and chiral angryline is non-enzymatic (Supplementary Fig. 6). Angryline (**3c**) was incubated with CS or TS to yield (+)-catharanthine (**4**) or (–)-tabersonine (**5**), respectively (Supplementary Figs. 7-8) and with CorS and DPAS (*T. iboga*) to yield (–)-coronaridine (**6**)[6]. The isolation of angryline suggests that desacetoxylation of dihydroprecondylocarpine acetate (**2**) to form dehydrosecodine (**3a**) occurs either in solution, through a Grob-type fragmentation[9], or in the active site of DPAS, and that the sole function of CS, TS and CorS is to catalyse cyclization.

Isolated angryline (**3c**) did not undergo cyclization in the absence of a cyclase under optimal reaction conditions, indicating that these cycloaddition reactions are enzymatically catalysed (Supplementary Fig. 8). Angryline is stable in mild acidic conditions but under neutral/alkaline pH can undergo imine-enamine tautomerization (Supplementary Fig. 9) which, in the absence of a cyclase, leads to degradation (Supplementary Fig. 5, 8). The relatively slow equilibration of angryline (**3c**) to dehydrosecodine (**3b**) did not allow accurate measurement of kinetic constants.

We next solved the crystal structures (Supplementary Table 1) of CS (2.2 Å resolution, 6RT8), TS (1.3 Å resolution, 6RS4), and CorS (1.4 Å resolution, 6RJ8) using molecular replacement (PDB code 2O7R). These enzymes, which are all carboxylesterase homologues, were structurally similar (RMSD of the dimers 1.045-1.188 Å), showing an α/β-hydrolase fold (Supplementary Fig. 10). The hydrogen-bonding network of the canonical carboxylesterase catalytic triad Ser/Cys-His-Asp[10,11] is disrupted, most notably by replacement of His with Tyr in CS and TS, and with Phe in CorS (Fig. 2, Supplementary Fig. 11). Moreover, the characteristic Gly-Gly oxyanion hole is replaced by Ala-Gly in each of the cyclases (Supplementary Fig. 12). CS and TS fail to turn over model substrates that are typically hydrolyzed by this class of enzymes (4-methylumbelliferyl acetate and 4-methylumbelliferyl butyrate)[11], suggesting that ancestral esterase/hydrolase activity has

been lost (Supplementary Fig. 13). The attempt to reintroduce the hydrolysis function in TS by Y297H mutation failed, consistent with the observation that the entire active site is distorted compared to typical carboxylesterases (Supplementary Figs. 11, 13).

We attempted to co-crystallize CS, TS and CorS with a range of ligands. Electron density corresponding to a biologically relevant ligand was only observed in a structure of CS co-crystallized with (+)-catharanthine (**4**). The omit map suggested that the ligand was not (+)-catharanthine, but 16-carbomethoxycleaviminium (**7**; Fig. 2 and Supplementary Figs. 14, 15a). **7** has been indirectly shown to be formed under acidic conditions via a retro-Mannich opening of (+)-catharanthine (**4**) by the observation of 16-carbomethoxycleavamine (**8**) after chemical reduction[12–13]. A compound with the same mass and retention time was formed upon extended incubation of (+)-catharanthine with CS (Supplementary Fig. 16), indicating that CS is capable of catalysing this retro-Mannich reaction. This indirectly suggests that the forward reaction may also proceed via an ionic stepwise mechanism rather than a concerted Diels-Alder mechanism (Supplementary Fig. 15b). The negatively charged carboxyl group of a retro-Mannich intermediate (Supplementary Fig. 15b) could be stabilized by the oxyanion hole (Ala82-Gly83).

TS, in contrast, did not catalyze any detectable reaction with (–)-tabersonine (**5**), suggesting that the mechanisms of CS and TS are fundamentally different. We docked (–)-tabersonine into the active site of TS (Supplementary Fig. 12b). The enzyme active site may position the carbonyl of the methyl acrylate (Fig. 2) in plane with the indole, allowing the vinyl indole to function as the diene in a Diels-Alder mechanism. Additionally, due to a missing loop, the TS substrate is more exposed to the bulk water (Supplementary Fig. 17), which could facilitate protonation of the dihydropyridine amine. While tabersonine formation could also proceed via a stepwise mechanism, this protonated amine could serve as an electron withdrawing group to activate the dienophile in a Diels-Alder reaction (Supplementary Fig. 15c).

CorS reacts with angryline to form an unstable product that is reduced by DPAS to form (–)-coronaridine (**6**)[6]. We hypothesize that CorS converts angryline (**3c**) to (–)-coronaridine iminium (**9**) through an aza-Diels-Alder reaction or a stepwise mechanism[14] (Supplementary Fig. 15d), although this proposed iminium is too unstable to isolate. Furthermore, while (+)-catharanthine (**4**) and (–)-tabersonine (**5**) appear to form directly from dehydrosecodine enamine (**3b**), reaction in D_2O suggests that the formation of (–)-coronaridine (**6**) requires tautomerization of the dihydropyridine moiety of **3b** to form **3d** and **3e**[6] (Supplementary Fig. 15d). Therefore, CorS likely promotes tautomerization of dehydrosecodine in addition to cyclization. Thus, the switch in enantioselectivity between CS and CorS is caused by tautomerization to a different isomer of dehydrosecodine (**3e**), which would cyclize via a different mechanism. Docking of (–)-coronaridine in CorS (Fig. 2, Supplementary Fig. 12c) demonstrates that this active site is shaped to accommodate the (–)-iboga scaffold.

This study provides a structural basis for design and assay of 47 mutants (Supplementary Figs. 12, 14, 17). Five TS mutants (S170A, H78A, Y216D, Y216F, helix5 swap), five CS mutants (loop trim, loop trim+helix5 swap, loop trim+F218Y, helix5 swap, D273A) and three CorS mutants (S174A, S174C, P175T) showed qualitative changes in product profile

(Supplementary Fig. 18-24, Supplementary Table 2), providing the first clues to cyclization specificity.

The substrate conformation determines the cyclization selectivity between iboga and aspidoperma scaffolds, and the primary role of these enzymes is most likely to conformationally constrain dehydrosecodine. Most importantly, the methyl acrylate of dehydrosecodine must be in plane with the indole to form the aspidosperma scaffold, while these two moieties must be perpendicular to form the iboga scaffolds (Supplementary Fig. 15). Mutations of residues that appear to interact with the methyl acrylate or the indole had different effects in the three enzymes. Mutations S172A and S172C in CS, a residue that is close to the indole nitrogen and the ester group of dehydrosecodine, resulted in inactive enzyme. However, mutation S170C in TS did not affect activity, though S170A formed small amounts of (+)-catharanthine (**4**). When the same mutations were introduced into CorS, S174C produced (–)-tabersonine (**5**) but no (–)-coronaridine iminium (**9**), and CorS S174A produced a mixture of (–)-coronaridine iminium (**9**) and catharathine (**4**; Supplementary Figs. 20, 22, 24). Mutation of Pro175 in CorS, a residue that appears to contact the ester group of dehydrosecodine, resulted in formation of all three cyclization products, while mutation of the corresponding residue in CS (Pro173) and in TS (Thr171) only led to a decrease of activity.

The residues of helix α5 also appear to interact with the methyl acrylate and dihydropyridine moieties, primarily through van der Waals interactions. Thus, this region of the active site may also be responsible for orienting the substrate. Swapping helix α5 (Asp217-Cys224) of CS with the corresponding TS sequence yielded a CS mutant that formed (–)-tabersonine (**5**) (Supplementary Fig. 20). When this grafting was combined with the removal of the extended loop, we observed formation of small amounts of (–)-coronaridine (**6**) in the presence of DPAS (Supplementary Fig. 25), suggesting that these residues may play a role in the tautomerization of the dihydropyridine ring required for (–)-coronaridine formation. However, the same swap in CorS led to a mutant that still formed small amounts of (–)-coronaridine, indicating that other residues are required for this function. When helix α5 from CS was introduced into TS, the enzyme lost (–)-tabersonine (**5**)-forming activity and produced small amounts of an additional unknown product. Swapping the helix from TS into CorS resulted only in loss of activity. Finally, numerous other residues in the binding pocket were mutated; mutation of Tyr216 and His78 in TS (Supplementary Fig. 22), and of Asp273 in CS (Supplementary Fig. 20) led to changes in product profiles in the respective enzymes, but these mutations had minimal qualitative effects in the other enzymes.

Given that no single mutation led to a complete switch in product selectivity, and since identical point mutations had different effects in the context of the three enzymes, all of the residues of the binding pockets likely work in concert to exert changes to the substrate conformation. Nevertheless, this study clearly demonstrates the plasticity of these enzymes with regard to product profile. More extensive analysis of CS, TS and CorS homologues found in other Apocynaceae plants, along with combinatorial mutagenesis, may be required to achieve complete changes in product selectivity. Although not characterized here, low levels of side products were observed (Supplementary Figs. 20, 22, 25), suggesting that future enzyme engineering efforts could enable unnatural iboga and aspidosperma scaffolds

to be accessed by additional variations of a cycloaddition reaction of dehydrosecodine isomers.

Here we report the enzymatic production and characterization of angryline (**3c**), a stabilized form of dehydrosecodine (**3a**, **3b**), an alkaloid cycloaddition substrate proposed 60 years ago[3]. When reacted with CS, TS and CorS, this substrate yields (+)-iboga, (–)-aspidosperma and (–)-iboga scaffolds. Structural data suggest that these enzymes control substrate conformation, namely the orientation of the methyl acrylate relative to the indole, to yield these different scaffolds. Tautomerization of the dihydropyridine ring, mediated in part by helix α5 of CorS, allows an additional cyclization mode. These enzymes generate chemical diversity from a single substrate via different cycloaddition reactions[15–18], and provide a foundation for enzyme engineering for catalysis of additional alkaloids.

Online Methods

Materials and molecular biology kits

All solvents used in this study were either of HPLC or MS grade, depending on the application. All were purchased from Fisher Scientific. Catharanthine **4** was purchased from Sigma Aldrich, whilst tabersonine **5** was obtained from Ava Chem Scientific. Stemmadenine acetate was synthesized as described before[2]. NADPH was from Roche. Carbenicillin was from Formedium. All genes and fragment amplifications were performed using Platinum Superfi polymerase (Thermo Fisher) whilst colony PCRs were performed using Phire II master mix (Thermo Fisher). PCR product purifications were performed using the Macherey-Nagel PCR clean-up kit. Plasmids purifications were performed using Promega Wizard minipreps. Sequencing of the clones was performed by Eurofins Genomics (Germany).

Cloning and mutagenesis

Cloning of CrPAS (MH213134), CrDPAS (KU865331), CS (MF770512) and TS (MF770513) was reported in Caputi *et al.*[2]. Cloning of DPAS and CorS from *Tabernanthe iboga* was reported in Farrow *et al.*[6]. CS, TS and CorS mutants were generated by overlap extension PCR. The codon(s) to be mutated was selected and two primers, one reverse and one forward (Supplementary Table 3), were designed to overlap and introduce the mutation. PCR products were gel purified, ligated into pOPINF expression vector[19] using the In-Fusion cloning kit (Clontech Takara) and transformed into competent *E. coli* Stellar cells (Clontech Takara) according to the manufacturer's instructions. Mutant constructs were sequenced to verify the mutant gene sequence and correct insertion.

Protein expression and purification

CrPAS and CrDPAS were expressed and purified as reported before[2]. All other proteins and mutants reported in this study were expressed in SoluBL21 *E. coli* cells (Amsbio). Chemically competent cells were transformed by heat shock at 42°C. Transformed cells were selected on LB agar plates supplemented with carbenicillin (100 µg/mL).

For large scale expression of proteins for crystallography studies, single colonies were used to inoculate starter cultures in 50 mL of 2 x YT medium supplemented with carbenicillin (100 μg/mL) that were grown overnight at 37°C. Starter culture (10 mL) was used to inoculate 1 L of 2 x YT medium containing the antibiotic. The cultures were incubated at 37°C until OD600 reached 0.6 and then transferred to 18°C for 30 min before induction of protein expression by addition of IPTG (0.2 mM). Protein expression was carried out for 16 h. Cells were harvested by centrifugation and re-suspended in 50 mL of Buffer A (50 mM Tris-HCl pH 8, 50 mM glycine, 500 mM NaCl, 5% glycerol, 20 mM imidazole,) with EDTA-free protease inhibitors (Roche Diagnostics Ltd.). Cells were lysed by sonication for 4 minutes on ice. Cell debris was pelleted by centrifugation at 35,000 g for 20 min. His$_6$-tagged enzymes were purified on an AKTA Pure system (GE Healthcare) using a HisTrap HP 5 mL column (GE Healthcare) equilibrated with Buffer A. Samples were loaded at a flow rate of 2 mL/minute and step-eluted using Buffer B (50 mM Tris-HCl pH 8, 50 mM glycine, 500 mM NaCl, 5% glycerol, 500 mM imidazole). Eluted proteins were subjected to further purification on a Superdex Hiload 16/60 S200 gel filtration column (GE Healthcare) at a flow rate of 1 mL/minute using Buffer C (20 mM HEPES pH 7.5, 150 mM NaCl) and collected in 1.5 mL fractions. Fractions containing the protein of interest were collected and treated with His$_6$-tagged 3C protease overnight at 4°C to cleave the His$_6$-tag. After filtration through low protein binding filters, the protein samples were passed through a HisTrap HP 1 mL column (GE Healthcare) equilibrated with Buffer C to remove the protease and the uncut protein. The proteins were concentrated using centrifugal filters.

For small scale expression of proteins for enzymatic activity, single colonies were used to inoculate starter cultures in 10 mL of 2 x YT medium supplemented with carbenicillin (100 μg/mL) that were grown overnight at 37°C. Starter culture (1 mL) was used to inoculate 100 mL of 2 x YT medium containing the antibiotic. The cultures were incubated at 37°C until OD600 reached 0.6 and then transferred to 18°C for 30 min before induction of protein expression by addition of IPTG (0.2 mM). Protein expression was carried out for 16 h. Cells were harvested by centrifugation and re-suspended in 10 mL of Buffer A with EDTA-free protease inhibitors (Roche Diagnostics Ltd.). Cells were lysed by sonication for 1 minute on ice. Cell debris was pelleted by centrifugation at 35,000 g for 20 min. 250 μL of Ni-NTA slurry (Qiagen) were added to the lysates and incubated for 1 h at 4°C on a shaker. The Ni-NTA beads were harvested by centrifugation at 1000 g for 1 min and washed three times with buffer A. Proteins were eluted by addition of 2 x 300 μL of buffer B. After filtration through low protein binding filters, the proteins were dialyzed in buffer C using centrifugal filters.

Protein Crystallization

Crystallization screens were conducted by sitting-drop vapour diffusion in MRC2 96-well crystallization plates (Swissci) with a mixture of 0.3 μl well solution from the PEGs (Qiagen) and PACT (Qiagen) screens and 0.3 μl of protein solution. Protein concentrations were adjusted to 20–30 mg/mL for all proteins. Ligands (catharanthine **4**, tabersonine **5** and stemmadenine acetate) were prepared at a concentration of 30 mM in methanol and added to a final concentration of 1 mM (1μl in 30 μl of protein solution). Solutions were dispensed by an Oryx8 robot (Douglas Instruments).

Crystals were obtained in several screening conditions and tested for diffraction on the beamline. The best diffracting CS crystals were obtained from 0.2 M Na thiocyanate and 20% (w/v) PEG3350; the best TS crystals were obtained from 25% (w/v) PEG 1500 in 0.1 M MIB buffer pH 4.0 (MIB = Malonate dibasic monohydrate, Imidazole, Boric acid; Molecular Dimensions); and the best CorS crystals were obtained from 25% (w/v) PEG 1500 in 0.1 M SPG buffer pH 6.0 (SPG = Succinic acid, sodium dihydrogen Phosphate and Glycine; Molecular Dimensions).

All crystals were cryoprotected by soaking in crystallization solution containing 25% (v/v) ethylene glycol before flash-cooling in liquid nitrogen.

X-ray data collection, processing and structure solution

All X-ray data were recorded on beamline I03 at the Diamond Light Source (Oxfordshire, UK) at a wavelength of either 0.976 Å (CS and CorS) or 0.980 Å (TS) using a Pilatus3 6M hybrid photon counting detector (Dectris) with the crystal maintained at 100 K by a Cryojet cryocooler (Oxford Instruments). Diffraction data were integrated and scaled using either DIALS[20] or XDS[21] via the XIA2 expert system[22] then merged using AIMLESS[23]. Data collection statistics are summarized in Supplementary Table 1.

The majority of the downstream analysis was performed through the CCP4i2 graphical user interface[24]. Both TS and CorS were solved by molecular replacement using the structure of *Actinidia eriantha* carboxylesterase (AeCXE1) (PDB entry 2O7R)[11] as a template, with which they share 29% and 32% sequence identity, respectively. Templates were prepared with reference to the appropriate sequences using SCULPTOR[25] and then all side-chains were truncated to Cβ. The structures were solved using PHASER[26] which located two copies of the template in the $P2_1$ asymmetric unit (ASU) of TS (with a corresponding solvent content of 53%), and a single copy of the template in the $P6_522$ ASU of CorS (with a corresponding solvent content of 63%). After editing these solutions to remove poorly fitting regions in COOT[27] and subsequent refinement with REFMAC5[28] the resultant models were completely rebuilt with BUCCANEER[29] to give much-improved models. These models were finalised through several iterations of manual editing in COOT and further refinement with REFMAC5. Anisotropic temperature factors were used for the latter in both cases.

The CS structure was solved in a similar fashion, but instead starting from a template derived from TS with which it shares 78% sequence identity. PHASER was successful in finding eight copies of the subunit template in the ASU, arranged as four homodimers (with a corresponding solvent content of 53%). From the initial electron density maps there was clear evidence for a bound ligand in each of the eight CS active sites. However, it could not be reconciled with the structure of catharanthine, which had been added to the crystallisation. It was only after the rest of the model was essentially complete that a cleavaminium intermediate could be confidently assigned to the density. For the final refinement cycles, TLS refinement was used with a single TLS domain defined for each of the eight CS subunits.

The statistics of the final refined models are shown in Supplementary Table 4. The Ramachandran statistics (favoured/allowed/outlier expressed as percentages), evaluated

using MolProbity[30], were as follows: CS - 95.2/4.2/0.6; TS - 98.2/1.5/0.3; CorS - 96.1/3.3/0.6. All structural figures were prepared using CCP4mg[31].

Docking simulations

Ligands were docked into the active sites of the TS and CorS crystal structures using AutoDock Vina[32]. Ligand coordinate files were generated using the Lidia function within COOT[27]. To prevent the appearance of extended, non-productive, conformations for the non-cyclized ligands, the three rotatable bonds linking the indole to the dihydropyridine ring were manually adjusted to place the latter alongside the methyl acrylate and then they were fixed. Otherwise all the remaining rotatable bonds were allowed to rotate during the simulations. Both the protein and ligand coordinates were prepared for the docking calculations using AutoDockTools[33]. The side-chains of several residues lining the active site pocket were allowed to rotate during the simulations, which are listed below for each protein. For the final simulations, an exhaustiveness value of 128 was used and, given the largely hydrophobic character of the active site pockets, the hydrogen bonding strength was increased (weight_hydrogen value of -1.2). Otherwise the default AutoDock Vina parameters were used. In assessing the results, we reasoned that the indole moiety would probably occupy roughly the same region of the active site pocket throughout the catalytic cycle and that the carbonyl oxygen of the methyl acrylate would most likely interact with the oxyanion hole, and possibly the catalytic Ser. Moreover, we favoured results that were broadly consistent in placement and orientation with the 16-carbomethoxycleaviminium **7** seen in the CS crystal structure. We thus selected plausible poses that satisfied these criteria for each of the ligands, which were not always the lowest energy solutions. These are displayed in Supplementary Fig. 10.

For TS, the best results were obtained using the protein chain B, possibly because it adopts a slightly more open conformation. During the simulations, the side-chains of Tyr14, Ser170, Thr171, Tyr216, Tyr224 and Tyr297 were allowed to rotate, and a search space of $20 \times 20 \times 20$ Å encompassing the active site cavity was used.

For CorS the side-chains of Tyr18, Ser174, Tyr205, Tyr219, Tyr224, Tyr227, Phe300, Phe301 and Phe304 were allowed to rotate, and a search space of $22 \times 22 \times 20$ Å encompassing the active site cavity was used. However, no suitable docking poses were obtained. Closer inspection of these in the context of a molecular surface indicated that the side-chain of Phe300 was mostly directed into the active site and likely to sterically inhibit ligand binding. Thus, for subsequent runs, Phe300 was manually rotated out of the active site and fixed.

NMR and CD spectroscopy

NMR spectra (1D and 2D NMR) were acquired using a Bruker Neo 600 MHz NMR spectrometer equipped with a TCI cryoprobe. The solvent residual peaks of CD_3OD (δ 3.31 and 49.0, respectively), CD_3CN (δ 1.94 and 1.32, respectively) and $CDCl_3$ (δ 7.26 and 77.16, respectively) were used as internal standards in ^1H and ^{13}C NMR. The number of scans depended on sample concentration and are indicated in SI Figures and Tables accordingly.

CD spectra were recorded in 1 nm steps with a 0.5 s averaging time on a Chirascan Plus spectropolarimeter (Applied Photophysics) at 20°C in a 1 mm cuvette. Measurements were collected in triplicates, averaged and background subtracted with 100 mM MES buffer pH 6.0.

Activity assays

Activity assays of CS and TS for pH optimum were performed in the following buffers, all at a concentration of 100 mM: MES (pH 5.5, pH 6.0 and pH 6.5), HEPES (pH 7.0 and pH 7.5), TRIS-HCl (pH 8.0 and pH 8.5) and CHES (pH 9.0, pH 9.5 and pH 10.0). The total volume of the reactions was 100 μL and they contained 100 nM enzyme and 1 μM angryline **3c**. Reactions were started by addition of the substrate and incubated at 37°C for 20 min. 10 μL of the reaction mixtures were collected at time 0 and after 20 min and quenched in 90 μL of 90:9:1 MeOH:H_2O:FA.

Activity assays of the mutant proteins were performed in 100 mM TRIS-HCl buffer pH 8.5. Total volume of the reactions was 100 μL. Reactions contained 100 nM enzyme and 1 μM angryline **3c**. Reactions involving the coupling of CorS and DPAS were performed as described in Farrow et al.[6]. Assays were performed at 37°C for 20 min. 10 μL of the reaction mixtures were collected at time 0 and after 20 min and quenched in 90 μL of 90:9:1 MeOH:H_2O:FA. All reactions were performed in triplicates.

UPLC/MS and UPLC/MS-MS methods

For high resolution MS analysis and MS/MS studies, angryline **3c** was infused at 5-10 μL/min using a Harvard Apparatus syringe pump onto a Synapt G2 HDMS mass spectrometer (Waters) calibrated using a sodium formate solution. Samples were analyzed for 1 minute with a scan time of 1 sec in the mass range of 50-1200 *m/z*. Capillary voltage was 3.5 kV, cone voltage 40 V, source temperature 120°C, desolvation temperature 350°C, desolvation gas flow 800 L/h. Leu-enkephaline peptide (1 ng/μL) was used to generate a dual lock-mass calibration with [M+H]$^+$ = 556.2766 and *m/z* = 278.1135 measured every 10 s. Spectra were generated in MassLynx 4.1 by combining a number of scans and peaks were centered using automatic peak detection with lock mass correction.

UPLC/QqQ-MS analysis was carried out on a UPLC (Waters) equipped with an Acquity BEH C18 1.7 μm (2.1 x 50 mm) column connected to Xevo TQS triple quadrupole (Waters). Chromatographic separation was performed using 0.1% FA as mobile phase A and methanol as mobile phase B. A linear gradient from 30% to 35% B in 3 min was performed for separation of the compounds followed by 0.5 min isocratic at 35%. The column was then re-equilibrated at 30% B for 1.5 min. The column was kept at 35°C throughout the analysis and the flow rate was 0.6 mL/min. MS detection was performed in positive ESI. Capillary voltage was 3.0 kV; the source was kept at 150°C; desolvation temperature was 500°C; cone gas flow, 50 L/h; and desolvation gas flow, 800 L/h. Unit resolution was applied to each quadrupole. The MRM transitions used to monitor the elution of the alkaloids of interest are reported in Supplementary Table 4.

Formation of 16-carbomethoxycleavamine from catharanthine

Chemical synthesis of 16-carbomethoxycleavamine **8** from catharanthine **4** was performed using a modified method derived from Andriamialisoa *et al.*[12]. 2.5 mg of catharanthine were dissolved in 100 μL of TFA and stirred at room temperature for 3 h. Formation of 16-carbomethoxycleaviminium **7** intermediate (m/z = 337.19) was monitored every hour by UPLC/QqQ-MS analysis. 1 μL aliquotes of the reaction mixture were diluted to 1 mL with 90:9:1 MeOH:H$_2$O:FA and injected. Before appearance of by products, the iminium intermediate was reduced by dropwise addition of an excess of sodium cyanoborohydride in MeOH (5 mg in 200 μL of MeOH). After stirring at room temperature for 30 min, the final products, 16-carbomethoxycleavamine **8** isomers (m/z = 339.19) were analyzed by UPLC/QqQ-MS.

Enzymatic synthesis of 16-carbomethoxycleavamine **8** from catharanthine **4** using CS as catalyst was performed in 50 mM HEPES buffer pH 7.5. 3 mg of catharanthine **4** were dissolved in 20 μL of DMSO and added to the reaction together with 1.4 mg of enzyme. Reactions were incubated at 37°C for 3 h and progress of the reaction and formation of 16-carbomethoxycleaviminium **7** intermediate was monitored by UPLC/QqQ-MS analysis. When the reactions had generated enough product, the enzyme was removed by SPE on an OASIS BHL cartridge (30 mg) and the product was eluted with 200 μL of 90:9:1 MeOH:H$_2$O:FA. 500 μL of material were reacted with an excess of sodium cyanoborohydride in MeOH (1 mg in 200 μL of MeOH). After stirring at room temperature for 30 min, the final products were analyzed by UPLC/QqQ-MS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. O'Connor SE, Maresh JJ. Nat Prod Rep. 2006; 23:532–472. [PubMed: 16874388]

2. Caputi L, et al. Science. 2018; 60:1235–1239.

3. Wenkert E. J Am Chem Soc. 1962; 84:98–102.

4. Scott AI, Cherry PC, Qureshi AA. J Am Chem Soc. 1969; 91:4932. [PubMed: 5798098]

5. Kutney JP, Ehret C, Nelson VR, Wigfield DC. J Am Chem Soc. 1968; 90:5929. [PubMed: 5679179]

6. Farrow SC, et al. JACS. 2019; 141:12979–12983.

7. Scott AI. Acc Chem Res. 1970; 3:151.

8. Qureshi AA, Scott AI. J Chem Soc Chem Commun. 1968; 948

9. Kuehne ME, Roland DM, Haften RJ. Org Chem. 1978; 43:3705–3710.

10. Marshall SD, Putterill JJ, Plummer KM, Newcomb RD. J Mol Evol. 2003; 57:487–500. [PubMed: 14738307]

11. Ileperuma NR, et al. Biochemistry. 2007; 46:1851–1859. [PubMed: 17256879]

12. Andriamialisoa RZ, Langlois N, Langlois Y. Heterocycles. 1981; 15:245–250.

13. Langlois N, Guéritte F, Langlois Y, Potier P. J Am Chem Soc. 1976; 98:7017–7024. [PubMed: 965661]

14. Buonora P, Olsen J-C, Oh T. Tetrahedron. 2001; 57:6099–6138.

15. Lichman BR, O'Connor SE, Kries H. Eur J Chem. 2019; 25:6864–6877.

16. Tan D, et al. J Am Chem Soc. 2019; 141:769–773. [PubMed: 30609896]

17. Jeon BS, et al. Proc Natl Acad Sci. 2017; 114:10408–10413. [PubMed: 28874588]

18. Zhang B, et al. Nature. 2019; 568:122–126. [PubMed: 30867595]

19. Berrow NS, et al. Nucleic Acids Res. 2007; 35:e45. [PubMed: 17317681]

20. Winter G, et al. Acta Crystallogr D Biol Crystallogr. 2018; 74:85–97.

21. Kabsch W. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–132. [PubMed: 20124692]

22. Winter G. J Appl Crystallogr. 2010; 43:186–190.

23. Evans PR, Murshudov GN. Acta Crystallogr D Biol Crystallogr. 2013; 69:1204–1214. [PubMed: 23793146]

24. Potterton L, et al. Acta Crystallogr D Biol. 2018; 74:68–84.

25. Bunkoczi G, Read RJ. Acta Crystallogr D Biol Crystallogr. 2011; 67:303–312. [PubMed: 21460448]

26. McCoy AJ, et al. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]

27. Emsley P, Cowtan K. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

28. Murshudov GN, Vagin AA, Dodson EJ. Acta Crystallogr D Biol Crystallogr. 1997; 53:240–255. [PubMed: 15299926]

29. Cowtan K. Acta Crystallogr D Biol Crystallogr. 2006; 62:1002–1011. [PubMed: 16929101]

30. Davis IW, et al. Nuc Acids Res. 2007; 35:W375–W383.

31. McNicholas S, Potterton E, Wilson KS, Noble ME. Acta Crystallogr D Biol Crystallogr. 2011; 67:386–394. [PubMed: 21460457]

32. Trott O, Olson AJ. J Comput Chem. 2010; 31:455–461. [PubMed: 19499576]

33. Morris GM, et al. J Comput Chem. 2009; 30:2785–2791. [PubMed: 19399780]
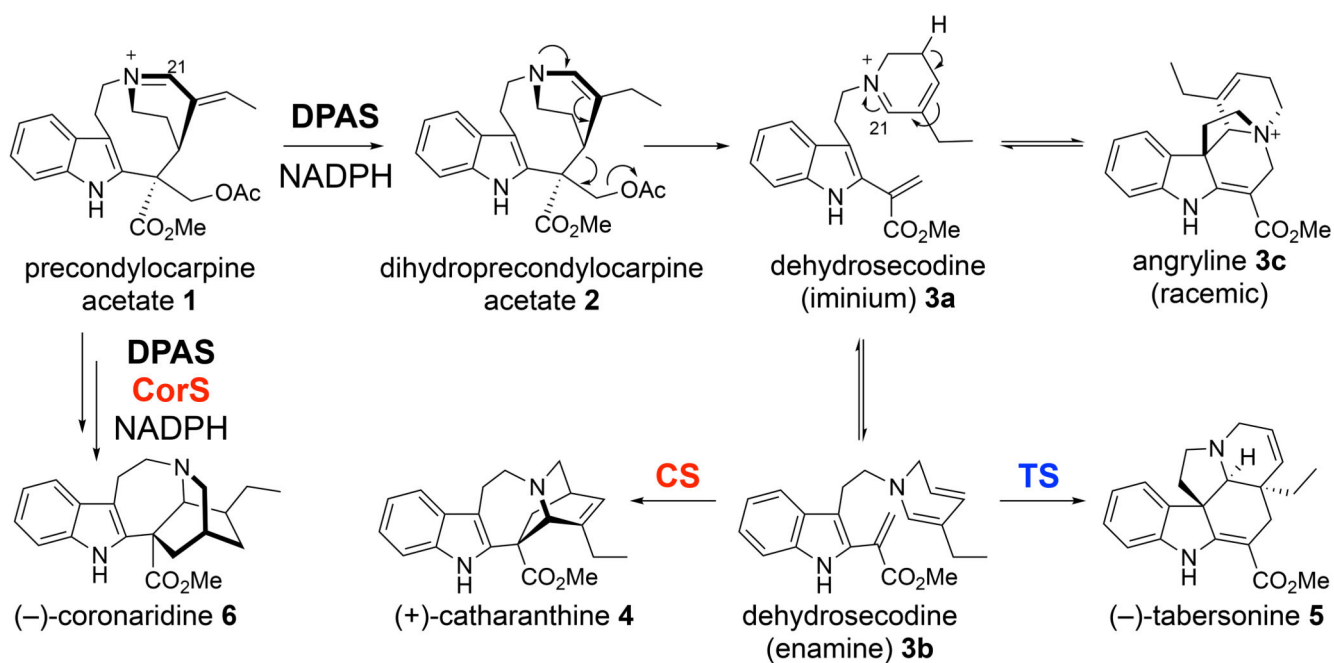
Figure 1. Biosynthesis of aspidosperma and iboga alkaloids.
Biosynthesis begins with reduction of precondylocarpine acetate (**1**) to generate dehydrosecodine (**3b**), which can undergo one of two formal Diels-Alder reactions to form either (+)-catharanthine (**4**) or (−)-tabersonine (**5**). The biosynthesis of (−)-coronaridine **6** also starts from precondylocarpine acetate **1**, but involves an additional reduction step.
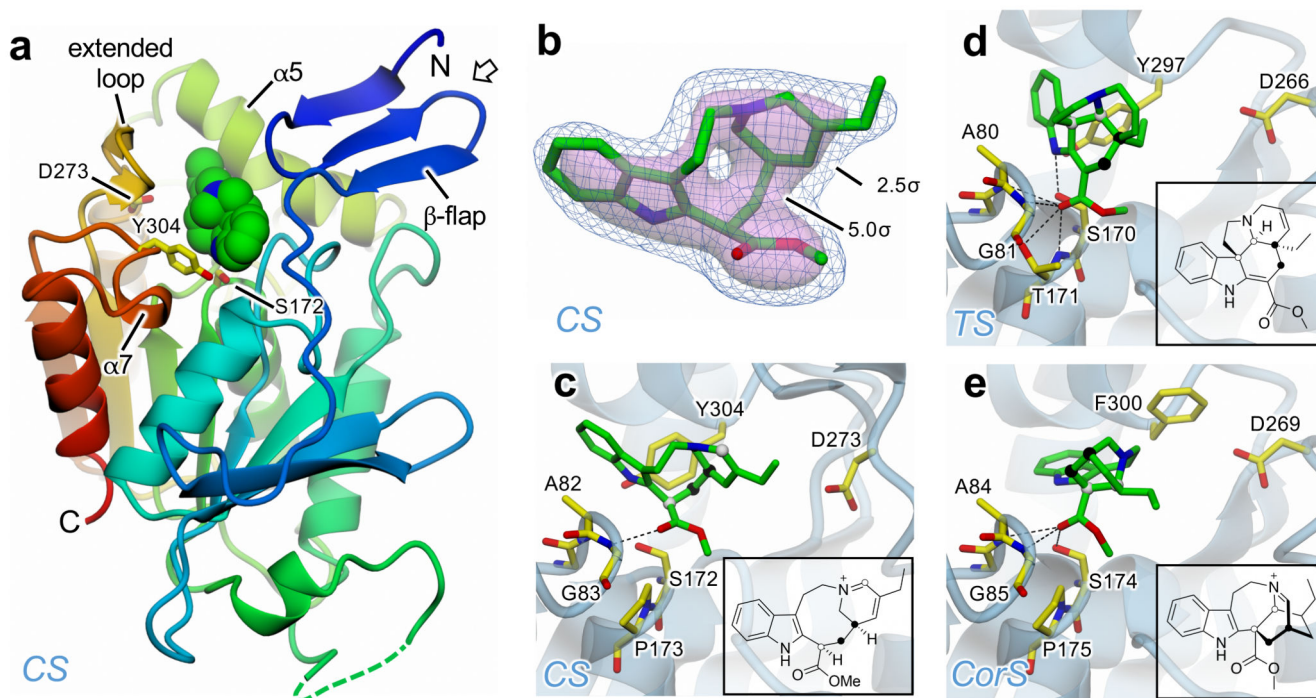
Figure 2. Crystal structures of CS, TS and CorS.

a. A single CS subunit depicted in cartoon representation with rainbow coloration from blue at the N-terminus through to red at the C-terminus. Also shown as van der Waals spheres with green carbons is the 16-carbomethoxycleaviminium (**7**) intermediate bound in the active site; the catalytic triad residues are in stick mode with yellow carbons. The sides of the active site cavity are delineated largely by helices α5 and α7, the N-terminal β-flap and, unique to CS, the extended loop; the base of the cavity is formed by the "nucleophilic elbow" which bears the catalytic Ser, and the loop providing the oxyanion hole residues. The open arrow shows the approximate direction of view for panels **b**–**e**. **b.** Omit electron density at 2.2 Å resolution for the cleaviminium intermediate **7** bound to CS shown at two contour levels: 2.5σ in blue mesh, and 5.0σ as a semi-transparent pink surface. **c.** Detail of CS active site, where residues from the catalytic triad and oxyanion hole are displayed with yellow carbons and the backbone trace is depicted as a semi-transparent pale blue cartoon. The cleaviminium intermediate **7** is shown with green carbons and also depicted as a 2D representation in the inset. In both representations, the two C-C bonds that form as a result of catalysis are indicated by spheres of the same colour for each bond (i.e. black or white). **d.** Detail of TS active site with docked (−)-tabersonine (**5**) displayed as for CS in panel **c**. **e.** Detail of CorS active site with docked (−)-coronaridine iminium (**9**) displayed as for CS in panel **c**. In **c**–**e**, only hydrogen bonds between ligands and the active site residues are shown (dashed lines).