Missing enzymes in the biosynthesis of the anticancer drug vinblastine in Madagascar periwinkle

Lorenzo Caputi,1 Jakob Franke,1* Scott C. Farrow,1 Khoa Chung,1 Richard M. E. Payne,1† Trinh-Don Nguyen,1 Thu-Thuy T. Dang,1 Inês Soares Teto Carqueijeiro,2 Konstantinos Koudounas,2 Thomas Dugé de Bernonville,2 Belinda Ameyaw,1‡ D. Marc Jones,1 Ivo Jose Curcino Vieira,3 Vincent Courdavault,2§ Sarah E. O’Connor1§

Anticancer drugs vincristine 5 and vinblastine 6 were serendipitously discovered 60 years ago in Catharanthus roseus (Madagascar periwinkle). These compounds have been used for the treatment of several types of cancer, including Hodgkin’s lymphoma, lung and brain cancers. Much of the metabolic pathway (31 steps from geranyl pyrophosphate to vincblastine) has been elucidated (1–3). Here we report the genes encoding the missing enzymes that complete the vincblastine pathway. Two redox enzymes convert stemmadenine acetate 7 into an unstable molecule, likely dihydroprecondylocarpine acetate 11, which is then desacetylated by one of two hydrolases to generate, through Diels-Alder cyclizations, either tabersonine 2 or catharanthine 3, scaffolds that are ultimately dimerized to yield vincblastine and vincristine (Fig. 1A). In addition to serving as the precursors for vincristine 5 and vinblastine 6, tabersonine 2 and catharanthine 3 are also precursors for other biologically active alkaloids (4, 5) (Fig. 1B). The discovery of these two redox enzymes (precondylocarpine acetate synthase and dihydroprecondylocarpine acetate synthase), along with the characterization of two hydrolases (tabersonine and catharanthine synthase), provides insight into the mechanisms that plants use to create chemical diversity and also enables production of a variety of high value alkaloids.

Catharanthine 3 (iboga-type alkaloid) and tabersonine 2 (aspidosperma-type) scaffolds are likely generated by dehydration of the biosynthetic intermediate stemmadenine 1 to dehydrosecodine 9, which can then cyclize to either catharanthine 3 or tabersonine 2 via a net [4+2] cycloaddition reaction (Fig. 2) (6–9). We speculated that the missing components were an enzyme with dehydration and cyclization function. We hypothesized that the unstable nature of the dehydration product dehydrosecodine 9 (7) would preclude its diffusion out of the enzyme active site, and thus searched for an enzyme that could catalyze both dehydration and cyclization reactions.

Since the biosynthetic genes for vincristine 5 and vinblastine 6 are not clustered in the plant genome (10), we searched for gene candidates in RNA-seq data (10) from the vincristine/vinblastine producing plant C. roseus. Two genes annotated as alpha/beta hydrolases were identified by a shared expression profile with previously identified vinblastine biosynthetic enzymes (fig. S1A and data S1). A dehydratase could facilitate the isomerization of the 19,20-exo-cyclic double bond of stemmadenine 1 to form iso-stemmadenine 8 (Fig. 2), which would then allow dehydration to form dehydrosecodine 9 and, consequently, catharanthine/tabersonine (8). Virus-induced gene silencing (VIGS) of herein named Ta-bersonine Synthase (TS) and Catharanthine Synthase (CS) in C. roseus resulted in a marked reduction of tabersonine 2 (p = 0.0048) and catharanthine (p = 0.01), respectively. These silencing experiments implicate CS and TS in catharanthine 3 and tabersonine 2 biosynthesis in C. roseus. However, when CS and TS were heterologously
expressed in *E. coli* (fig. S4A) and tested for reactivity with stemmadenine 1 (fig. S5) or the acetylated form of stemmadenine 7 (fig. S6), in which spontaneous deformylation would be hindered (Fig. 2, figs. S5 and S6, and tables S4 and S5), no reaction was observed. Although TS and CS are known to be implicated in vinblastine biosynthesis (II), the substrates, and therefore the specific catalytic functions, remained elusive.

We attempted to isolate the active substrate for the TS and CS enzymes from various aspidosperma- and iboga-alkaloid producing plants using enzyme-assay guided fractionation. We focused on *Tabernaemontana* plants that are known to accumulate more stemmadenine 1 intermediate relative to downstream alkaloids (I2). These experiments demonstrated that TS and CS were always active with the same fractions (fig. S7), consistent with previous hypotheses (6) that both enzymes utilize the same substrate. However, attempts to structurally characterize the substrate were complicated by its rapid decomposition, and the deformylated product tubotaiwine 12 (previously synthesized in reference (I3)) was the major compound detected in the isolated mixture (Fig. 2 and NMR data in fig. S8). Given the propensity for deformation in these structural systems (I4), we rationalized that tubotaiwine 12 could be the decomposition product of the actual substrate, which would correspond to *iso*-stemmadenine 8 (dihydroprecondylocarpine) or its protected form (dihydroprecondylocarpine acetate 11) (Fig. 2). We surmised that a coupled oxidation-reduction cascade could form a net isomerization to generate dihydroprecondylocarpine 8 (or dihydroprecondylocarpine acetate 11) from stemmadenine 1 (or stemmadenine acetate 7). This idea was initially proposed by Scott, who indicated that stemmadenine acetate 7 can be oxidized to precondylocarpine acetate 10, after which the 19,20-double bond can then be reduced to form dihydroprecondylocarpine acetate 11, which could then form traces of tabersonine 2 upon thermolysis (I5). Similar reactions with stemmadenine 1 resulted in deformation to form condylocarpine 13 (I6). Therefore, we re-examined the RNA-seq dataset for two redox enzymes that could convert stemmadenine acetate 7 to dihydroprecondylocarpine acetate 11.

We noted a gene annotated as reticuline oxidase that had low absolute expression levels, but a similar tissue expression pattern to the TS gene (fig. S1B). The chemistry of reticuline oxidase enzymes (I7) such as berberine bridge enzyme and dihydrobenzophenanthridine oxidase, suggests that these enzymes are capable of C-N bond oxidation, which is what would be required in this reaction sequence (Fig. 2) (I7). When this oxidase gene was silenced in *C. roseus*, a compound with a mass and "H NMR spectrum corresponding to semi-synthetically prepared stemmadenine acetate 7 (the proposed oxidase substrate) accumulated, suggesting that this gene encoded the correct oxidase. We named this enzyme precondylocarpine acetate synthase (PAS) (figs. S9 to S11). Similarly, silencing of a medium chain alcohol dehydrogenase, as part of an ongoing screen of alcohol dehydrogenases in *C. roseus* (I4, I8, I9), resulted in accumulation of a compound with a mass, retention time and fragmentation pattern consistent with a partially characterized synthetic standard of precondylocarpine acetate 10 (the proposed substrate of the reductase) (figs. S12 to S14). This standard could be synthesized from stemmadenine acetate 7 using Pt and O2 by established methods (8, I5, 20). With our small-scale reactions yields were low and variable, and the product decomposed during characterization. However, the limited 2D NMR data set was consistent with an assignment of precondylocarpine acetate 10. Thus, we renamed this alcohol dehydrogenase dihydroprecondylocarpine synthase (DPAS). Collectively, these data suggest that PAS and DPAS act in concert with CS or TS to generate catharanthine 3 and tabersonine 2.

To validate whether these enzymes produce catharanthine 3 and tabersonine 2, we transiently co-expressed PAS, DPAS and CS or TS in the presence of stemmadenine acetate 7 in *Nicotiana benthamiana*. These experiments illustrated the sequential activity of the newly discovered enzymes, whereby we observed formation of catharanthine 3 in plant tissue overexpressing PAS/DPAS/CS or tabersonine 2 in plant tissue overexpressing PAS/DPAS/TS, when the leaf was also co-infiltrated with stemmadenine acetate 7 (Fig. 3A). The presence of all proteins was validated by proteomics analysis (fig. S4B and data S2). Formation of precondylocarpine acetate 10 was observed when stemmadenine acetate 7 was infiltrated into *N. benthamiana* in the absence of any heterologous enzymes (Fig. 3A), suggesting that an endogenous redox enzyme(s) of *N. benthamiana* can oxidize stemmadenine acetate 7. Formation of 3 and 2 was validated by co-elution with commercial standards, and formation of 10 was validated by co-elution with the semi-synthetic compound.

Purified proteins were required to validate the biochemical steps of this reaction sequence in vitro. While CS, TS and DPAS all expressed in soluble form in *E. coli* (fig. S4A), the flavin-dependent enzyme PAS failed to express in standard expression hosts such as *E. coli* or *S. cerevisiae*. To overcome this obstacle, we expressed the native full-length PAS in *N. benthamiana* plants using a transient expression system (fig. S4B), in *Pichia pastoris* (fig. S4, C and D) and in Sf9 insect cells (fig. S4E). The presence of PAS was validated by proteomics data (data S3). Reaction of PAS from each of these expression hosts with stemmadenine acetate 7 produced a compound that had an identical mass and retention time to our semi-synthetic standard of precondylocarpine acetate 10 (figs. S15 and S16). The enzymatic assays with PAS protein...
derived from *P. pastoris* and Sf9 insect cells ensure that formation of the expected products is not the result of a protein contaminant found in the plant-expressed PAS protein. When the PAS proteins (from *N. benthamiana* and *P. pastoris*), along with stemmadenine acetate 7, were combined with heterologous DPAS and CS, catharanthine 3 was formed, and when combined with DPAS and TS, tabersonine 2 was observed (figs. S17 and S18).

Semisynthetic precondylocarpine acetate 10 could be reacted with DPAS and TS/CS to yield tabersonine 2 or catharanthine 3, respectively (fig. S19). In addition, a crude preparation of what is proposed to be dihydroprecondylocarpine acetate 11, synthesized according to (11), was converted to catharanthine 3 and tabersonine 2 by the action of CS and TS, respectively (figs. S20). Reaction of PAS (purified from *N. benthamiana*) and DPAS with stemmadenine acetate 7 in the absence of CS or TS yielded a compound isomeric to tabersonine 2 and catharanthine 3, suggesting that cyclization can occur spontaneously under these reaction conditions (fig. S17). As observed during attempts to purify the CS/TS substrate from *Tabernaemontana* plants, dihydroprecondylocarpine acetate 11 can also deformylate to form tubotaiwine 12.

Solvant and reaction conditions likely determine how the reactive dihydroprecondylocarpine acetate 11 decomposes. PAS failed to react with stemmadenine 1, indicating the enzyme recognized the acetyl group (fig. S21). Oxidation of stemmadenine 1 produced a compound with a mass consistent with that of the shunt product condylocarpine 13. This identification was supported by comparison of the MS/MS spectrum to the related compound tubotaiwine 12 (fig. S22). The transformation of stemmadenine 1 to condylocarpine 13 is known (15, 21). We therefore suspect that the acetylation of stemmadenine 1 is necessary to slow spontaneous deformylation after oxidation, while also serving as a leaving group to allow formation of dehydrosecodine 9 (20). Acetylation also functions as a protecting group in the biosynthesis of noscapine in opium poppy (22).

The reactivity of the intermediates involved in the transformation of stemmadenine acetate 7 to catharanthine 3 or tabersonine 2 suggests that PAS, DPAS and CS/TS should be co-localized, since the unstable post precondylocarpine acetate 10 intermediates may not remain intact during transport between cell types or compartments. Using YFP-tagged proteins in *C. roseus* cell suspension culture, we showed that PAS is targeted to the vacuole through small vesicles budding from the endoplasmic reticulum (ER), as was previously observed for the PAS homolog, berberine bridge enzyme (23) (fig. S23). This localization suggests that stemmadenine acetate 7 oxidation occurs in the ER-lumen, ER-derived vacuole-targeted vesicles and/or vacuole. In contrast, co-localization of DPAS, CS and TS were confirmed in the cytosol (figs. S24 and S25). Bimolecular fluorescence complementation suggested preferential interactions between DPAS and TS (fig. 3B and figs. S26 and S27). Such interactions may not only prevent undesired reactions on the reactive dihydroprecondylocarpine acetate 11 intermediate but may control the flow of 11 into tabersonine 2.

Homologs of PAS are utilized throughout benzyliisoquino-line and pyridine alkaloid biosynthesis. Certain PAS mutations characterize the enzymes found in *aspidosperma* and iboga alkaloid producing plant clades (Fig. 3C). For instance, PAS lacks the His and Cys residues involved in covalent binding of the FAD cofactor (fig. S28). We anticipate that these aspidosperma-associated PAS homologs populate the metabolic pathways for the wide range of aspidosperma alkaloids found in nature. DPAS is a medium chain alcohol dehydrogenase, enzymes widely used in monoterpene indole alkaloid biosynthesis (14, 18, 19, 24). We hypothesize that CS and TS may retain the hydrolysis function of the putative ancestor hydrolase enzyme (25) to allow formation of dehydrosecodine 9 from dihydroprecondylocarpine acetate 11. In principle, the formation of tabersonine 2 and catharanthine 3 is formed via two different modes of cyclization, and dehydrosecodine 9 can undergo two distinct Diels-Alder reactions (26) to form either catharanthine 3 or tabersonine 2 (Fig. 2).

Here we report the discovery of two enzymes—PAS and DPAS—along with the discovery of the catalytic function of two other enzymes—CS and TS (11)—that convert stemmadenine acetate 7 to tabersonine 2 and catharanthine 3. Chemical investigations of this system (6–8, 15), coupled with plant DNA sequence data, enabled discovery of the last enzymes responsible for the construction of the tabersonine 2 or catharanthine 3 scaffolds. With the biosynthesis of stemmadenine acetate 7 (11), this completes the biosynthetic pathway for vindoline 4 and catharanthine 3, compounds that can be used to semi-synthetically prepare vinblastine. These discoveries allow the prospect of heterologous production of these expensive and valuable compounds in alternative host organisms, providing a new challenge for synthetic biology (27).

REFERENCES AND NOTES


5. G. K. Jana, S. Paul, S. Sinha, Progress in the synthesis of iboga-alkaloids and their
This work shows sequence data for PAS (MH213134), a patent on PAS and DPAS is currently being filed. The first release is on 3 May 2018.

Acknowledgments

We gratefully acknowledge Dr. Gerhard Saalbach of the John Innes Proteomics Centre for proteomics work. We thank Dr. Franziska Kellner and Dr. Fernando Geo Flores for early silencing results with CS/TS, and we thank Dr. Benjamin Lichman for helpful discussions. We thank Dr. Richard Hughes and Dr. Marina Franceschetti for preparing the modified TRBO vector. We also thank MRCPPU Reagents and Services for PAS expression in insect cells. Funding: This work was supported by grants from the European Research Council (311363), BBSRC (BB/J004561/1) and from the Région Centre-Val de Loire, France (CatharSIS grant and BiophOPHRM project ARD2020) (V.C.). J.F. gratefully acknowledges DFG postdoctoral funding (FR 3720/1-1) and T.T.T.D. (ALTF 239-2015) and S.F. (ALTF 046-2016) an EMBO Long Term Fellowship. K.K. gratefully acknowledges the French Embassy in Greece for a SSSH grant. R.E.M.P. gratefully acknowledges the support of BBSRC for a PhD studentship. Author contributions: L.C. characterized all protein, purified substrates, coordinated all experiments; J.F. performed experiments with PAS and developed chemical mechanism and synthetic strategy; S.F. helped with candidate selection, VIGS and performed the qPCR experiments; K.C. performed all semi-synthetic transformations; R.M.E.P. performed initial cloning and characterization of PAS; D.N. expressed PAS in P. pastoris; T.T.T.D. performed reconstitution in N. benthamiana; B.A. performed cloning and silencing experiments; M.J. helped identify PAS and designed initial silencing constructs; I.J.C.V purified stemmadine alkaloid from Tabernanontana heyneana. Magn. Reson. Chem. 48, 474–476 (2002). doi:10.1002/marc.1030


SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.aat4100/DC1

Data and materials availability: Sequence data for PAS (MH213134), DPAS (KU865331), CS (MF770512) and TS (MF770513) are deposited in GenBank and are also provided in the supplementary materials. Transcription data used in this study has already been published as referenced in the manuscript (see also, http://medicinalplantgenomics.msu.edu/). Cell lines used in this manuscript will be made available upon request. Data supporting the findings of this study are available within the article and the supplementary materials files.

www.R-project.org/
Fig. 1. Vincristine and vinblastine biosynthesis. (A) Vincristine 5 and vinblastine 6 are formed by dimerization from the monomers catharanthine 3 and vindoline 4 by a peroxidase (28) or chemical methods (29). The genes that convert tabersonine 2 to vindoline 4 have been identified (24). (B) Representative bioactive alkaloids derived from stemmadenine.
Fig. 2. Biosynthesis of catharanthine and tabersonine scaffolds. Stemmadenine acetate 7 (generated from stemmadenine 1 using conditions i. (Ac2O (excess) pyridine (excess) r.t., 4 hours, >99%) undergoes an oxidation to form precondylocarpine acetate 10. This is catalyzed enzymatically by the reticuline oxidase homolog PAS, or alternatively can be generated synthetically using conditions as reported by Scott and co-workers (8) ii. (Pt (from 7.5 eq. PtO2), EtOAc, O2 atm., r.t., 10 hours, yields varied). Next, precondylocarpine acetate 10 is reduced by the alcohol dehydrogenase DPAS. This reduced intermediate could not be isolated due to its lability, but it is assumed to be dihydroprecondylocarpine acetate 11 based on the degradation product tubotaiwine 12. Dihydroprecondylocarpine acetate 11, in the open form, can form dehyrosecodine 9 through the action of CS or TS to form catharanthine 3 and tabersonine 2, respectively.
Fig. 3. Biosynthesis of tabersonine 2 and catharanthine 3 from stemmadenine acetate 7 starting substrate. (A) Reconstitution of tabersonine 2 and catharanthine 3 in N. benthamiana from stemmadenine acetate 7. Extracted ion chromatograms for ions m/z 397.19 (stemmadenine acetate 7), m/z 395.19 (precondylocarpine acetate 10) and m/z 337.19 (catharanthine at RT 4.0 and tabersonine at RT 4.4 min) are shown. (B) Interaction of CS and TS with DPAS by bimolecular fluorescence complementation (BiFC) in C. roseus cells. Efficiency of BiFC complex reformation reflected by YFP fluorescence intensity highlighted that CS and DPAS performed weak interactions (i–iii) while TS and DPAS strongly interacted (iv–vi). No interactions with loganic acid methyltransferase (LAMT) were observed (vii–ix). (C) Phylogenetic relationship of PAS with other functionally characterized berberine bridge enzymes.
Missing enzymes in the biosynthesis of the anticancer drug vinblastine in Madagascar periwinkle


published online May 3, 2018