Homolog of tocopherol C methyltransferases catalyzes N methylation in anticancer alkaloid biosynthesis

David K. Liscombe, Aimee R. Usera, and Sarah E. O’Connor

Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139

Edited* by Jerrold Meinwald, Cornell University, Ithaca, NY, and approved September 17, 2010 (received for review June 23, 2010)

Madagascar periwinkle (Catharanthus roseus) is the sole source of the anticancer drugs vinblastine and vincristine, bisindole alkaloids derived from the dimerization of the terpenoid indole alkaloids vindoline and catharanthine. Full elucidation of the biosynthetic pathways of these compounds is a prerequisite for metabolic engineering efforts that will improve production of these costly molecules. However, despite the medical and commercial importance of these natural products, the biosynthetic pathways remain poorly understood. Here we report the identification and characterization of a C. roseus cDNA encoding an S-adenosyl-L-methionine-dependent N methyltransferase that catalyzes a nitrogen methylation involved in vindoline biosynthesis. Recombinant enzyme produced in Escherichia coli is highly substrate specific, displaying a strict requirement for a 2,3-dihydro bond in the aspidosperma skeleton. The corresponding gene transcript is induced in methyl jasmonate-elicited seedlings, along with the other known vindoline biosynthetic transcripts. Intriguingly, this unique N methyltransferase is most similar at the amino acid level to the plastidic γ-tocopherol C methyltransferases of vitamin E biosynthesis, suggesting an evolutionary link between these two functionally disparate methyltransferases.

Madagascar periwinkle (Catharanthus roseus) plants produce a large family of monoterpenoid indole alkaloids derived from tryptamine and the iridoid terpene secologanin (1). These products include the bisindole alkaloids vinblastine 9a and vincristine 9b, which are derived through the dimerization of vindoline 7a, an aspidosperma-type alkaloid, and catharanthine 8 monomers (Fig. 1). Vinblastine 9a and vincristine 9b are microtubule disruptors that have been used extensively in the treatment of several types of cancer, including leukemia and lymphoma (2). Although yields of vincristine 9a and vinblastine 9b are approximately 0.0003% and 0.01%, respectively, from dried leaves of C. roseus (3), this plant nevertheless remains the only commercial source for these drugs. No other known plant species produces these compounds, and total syntheses of vinblastine 9a and vincristine 9b (for representative examples, see refs. 4 and 5) have not yet been adopted as industrial production methods.

Despite the medicinal and commercial importance of these compounds, several aspects of bisindole alkaloid biosynthesis are still not understood. Identification of the biosynthetic genes of this pathway is required for rational metabolic engineering efforts to improve production levels in the plant, explore heterologous production platforms, and to reprogram the pathway to make useful alkaloid derivatives. Although a number of the genes encoding vindoline 7a biosynthetic enzymes have been cloned and characterized (6–9), genes encoding the hydrating enzyme or N methyltransferase (NMT) involved in converting tabersonine 1 to vindoline 7a have not been identified (Fig. 1). Many of the validated biosynthetic enzymes are represented in the collection of approximately 20,000 ESTs for C. roseus available in GenBank (10, 11). Additional genes corresponding to unknown steps might also be represented in this dataset. Pioneering work by De Luca and colleagues demonstrated that the NMT of vindoline 7a biosynthesis is associated with thylakoid membranes of chloroplasts (12–14). Therefore, when we identified a number of candidate cDNAs that encode putative methyltransferases, we paid particular attention to a subset of these genes that also displayed sequence similarity to plastidic γ-tocopherol methyltransferases (γ-TMTs) (15–17). Functional expression of these genes confirmed this hypothesis.

Author contributions: D.K.L. and S.E.O. designed research; D.K.L. and A.R.U. performed research; A.R.U. contributed new reagents/analytic tools; D.K.L., A.R.U., and S.E.O. analyzed data; and D.K.L. and S.E.O. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession nos. Cr2276 (HM584929), Cr1196 (HM584930), and Cr2551 (HM584931)].

To whom correspondence should be addressed. E-mail: soc@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1009003107/-/DCSupplemental.
candidate cDNAs led to the identification of a γ-TMT homolog that methylates the aromatic nitrogen of a model aspidosperma substrate and can selectively methylate proposed native substrates present in seedling extracts.

Results and Discussion

Identification of N-Methyltransferase Candidates. We relied on two distinguishing features of the NMT of vindoline 7a biosynthesis to guide the gene discovery process. First, this enzyme is associated with thylakoid membranes of chloroplasts (12, 13, 18), and, thus, we expected the translation product of the NMT gene to possess a transit peptide, or to resemble known plastid-localized proteins. It is also well established that vindoline 7a biosynthesis occurs primarily in the leaves (7, 11), so we focused on those candidates that were expressed in this biosynthetically active tissue. C. roseus EST assemblies on the PlantGDB server (http://plantgdb.org/cgi-bin/blast/PlantGDBblast) were mined using tBLASTn, and all annotated plant NMT sequences in GenBank were used as queries. This initial list of NMT candidates derived from BLAST results was manually edited, where contigs that were identical or nearly identical to methyltransferases of known function (i.e., protein methyltransferases) were removed, to provide the master candidate list presented in Table S1. A phylogenetic tree (Fig. 2) illustrates the relationship between C. roseus NMT candidates and selected, functionally characterized S-adenosyl-L-methionine (SAM)-dependent enzymes from plants and bacteria. Predicted translation products of Cr5804, Cr6424, Cr8458, and Cr3910022 were too short to be included in the tree. We noted that five of our candidates (Cr1196, Cr2270, Cr7756, Cr3710017, and Cr3910022) were homologous to γ-TMTs, enzymes that are known to reside in plastids, where they participate in tocopherol biosynthesis (15, 17). The other analyzed candidates fall within a clade representing functionally diverse type-I and -II methyltransferases (Fig. 2) (19), including tabersonine 16-O methyltransferase, O methyltransferases of benzyloquinoline and ipecac alkaloid biosynthesis, and an anthranilate N methyltransferase. The apparent full-length cDNAs of the latter candidate group do not encode putative chloroplast transit peptides, according to ChloroP (20), and, therefore, the corresponding polypeptides are not expected to be plastidic. The γ-TMT homologs Cr2270 and Cr7756 appeared to code for full-length proteins, lacking only the 5′ region that would be expected to encode transit peptides (Fig. S1). However, Cr2270 was the only apparent full-length γ-TMT homolog that is expressed in leaves (Table S1), and we were unable to amplify the Cr7756 ORF from leaf or seedling cDNA. ORFs encoding Cr2270, and Cr1196 (both lacking putative transit peptide region), as well as Cr6996, and Cr7789, were amplified from elicited seedling cDNA and expressed in E. coli. Recombinant Cr2270, Cr1196, and Cr6996 proteins were purified to homogeneity and used for NMT assays (Fig. S2). As previously reported (21), we also observed that recombinant Cr7789 was insoluble and could not be purified.

Recombinant Cr2270 Catalyzes Aromatic Nitrogen Methylation of 2,3-Dihydrotabersonine. The known enzymes of vindoline 7a biosynthesis are highly substrate specific (7, 22), which presumably ensures that the enzymes act in a strictly defined sequence. De Luca and coworkers previously reported that the NMT of vindoline 7a biosynthesis, partially purified from C. roseus leaves, only accepted aspidosperma substrates in which the 2,3 alkene was among the few substrates readily turned over by this enzyme (13, 18), and because this compound is synthetically accessible through sodium cyanoborohydride reduction of tabersonine I (Fig. 3A), we used this as a test substrate to screen for the desired NMT activity. The structure of the synthetic DHT substrate was confirmed by 1H-NMR (Fig. 3B; full spectrum in Fig. S3A), 13C-NMR, and COSY-NMR (Fig. S3 B and C). Analysis of enzyme assays by liquid chromatography-mass spectrometry (LC-MS) revealed that in the presence of SAM, Cr2270 converted DHT to a more hydrophobic compound.
that displayed the expected molecular weight of N-methyl-DHT 11 (Fig. 3C). Furthermore, an authentic standard of N-methyl-DHT 11 that was synthesized by treatment of DHT 10 with sodium hydride and iodomethane (Fig. 3A and B and Fig. S3 D–E) coeluted with the enzymatic product (Fig. 3C). In contrast, recombinant Cr1196 or Cr6996 was not capable of methylating DHT 10 when SAM was included in the assays. Steady-state kinetic analyses for Cr2270 showed that DHT 10 evidently serves as an excellent model substrate for Cr2270, exhibiting a $K_m$ of 8.8 μM and a $V_{max}$ of 2.4 s$^{-1}$ (Table 1 and Fig. S4 A and B). As expected, Cr2270 is sensitive to inhibition by S-adenosylhomocysteine (SAH) ($K_i = 7.3$ μM) (Fig. S4).

Consistent with the previously reported substrate specificity for the NMT of vindoline 7a biosynthesis, Cr2270 does not methylate tabersonine 1 (13, 18). A variety of other alkaloids were assayed with recombinant Cr2270, but none were converted to any product as evidenced by monitoring the disappearance of starting material and formation of any product by LC-MS (Fig. S5). The failure of Cr2270 to turn over tabersonine 1 indicates that the 2,3-single bond of the aspidosperma skeleton is a critical recognition element for the enzyme. Without this functionality, the substrate is not recognized by the enzyme in vitro, presumably due to conformational differences between the substrates, or because the nitrogen of DHT 10 is more electron rich than that of tabersonine 1. The known alkaloid products of C. roseus—specifically the lack of N-methylated tabersonine—suggest that the substrate specificity is the same in vivo. In contrast, the 16-methoxy substitution on the indole ring does not appear to qualitatively affect the action of this NMT.

The substrate specificity of this NMT undoubtedly plays a role in determining the order of enzymatic transformations and defining the complement of alkaloids produced by C. roseus plants. The substrate specificity of subsequent steps is also consistent with the order of vindoline 7a biosynthesis (Fig. 1): D4H activity relies on the presence of the N-methyl group and reduced 2,3-bond (22), and deacetylvindoline acetyltransferase (DAT) clearly requires the 4-hydroxy group for acetylation. Although N deformylvincristine (N desmethylvinblastine) has been isolated, it is likely that this compound arises from deformylation of vincristine 9b in planta or during the isolation process (23), or via demethylation of vinblastine. The inherent substrate specificity of the biosynthetic enzymes has obvious metabolic engineering implications. Currently, all known N-methylated aspidosperma alkaloids in C. roseus lack the 2,3-double bond; understanding the molecular basis of substrate recognition by this enzyme may allow N methylation of a wider range of alkaloid intermediates for production of many previously undescribed alkaloid structures.

**Table 1. Apparent kinetic parameters for recombinant Cr2270**

<table>
<thead>
<tr>
<th>Substrates</th>
<th>DHT</th>
<th>SAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_m$ (μM)</td>
<td>8.8 ± 1.0</td>
<td>22.0 ± 1.8</td>
</tr>
<tr>
<td>$V_{max}$ (pmol s$^{-1}$ mg$^{-1}$ protein)</td>
<td>67.0 ± 1.8</td>
<td>151.6 ± 3.9</td>
</tr>
<tr>
<td>$k_{cat}$ (s$^{-1}$)</td>
<td>2.4 ± 0.1</td>
<td>5.4 ± 0.1</td>
</tr>
<tr>
<td>$k_{cat}/K_m$ (s$^{-1}$ μM$^{-1}$)</td>
<td>267,744</td>
<td>243,053</td>
</tr>
</tbody>
</table>

Recombinant Cr2270 Methylates the Proposed Native Substrates in Seedling Extracts. The precursors to vindoline 7a and vindorosine 7b (Fig. 1) are difficult to produce synthetically or isolate in sufficient quantities for kinetic studies. However, compounds with exact mass and formulas corresponding to these biosynthetic intermediates could be detected in crude extracts of C. roseus seedlings (Fig. 4). These extracts were therefore incubated with recombinant Cr2270 in the absence or presence of SAM to determine whether these intermediates are productive substrates for the enzyme. Notably, the compound with the exact mass matching 16-methoxy-2,3-dihydro-3-hydroxytabersonine 4a (Fig. 1) was consumed after addition of the enzyme, and a compound (desacetoxyvindoline 5a) with a mass corresponding to the addition of one methyl group accumulated (Fig. 4 A and C). Taken together with the results of the model substrate, this strongly suggests that Cr2270 also acts upon its predicted native substrate. A compound in crude extracts with an exact mass matching N desmethylvinblastine 4b also appeared to be methylated by this NMT to yield the corresponding intermediate vindoline 5b, a presumed precursor in vindorosine 7b biosynthesis (Fig. 4 B and C). Methylations of other putative vindoline 7a or vindorosine 7b intermediates in the extract was not detected. This targeted activity-based metabolic profiling approach (24) appears to be an excellent strategy to functionally characterize biosynthetic enzymes with native substrates that are challenging to isolate or synthesize.

Expression of Cr2270 Transcripts Is Coordinatey Regulated with Known Alkaloid Biosynthetic Genes. C. roseus seedlings produce alkaloids as part of their normal developmental program, and this accumulation can be enhanced by methyl jasmonate (25). As previously demonstrated, we noted the amount of vindoline 7a found in seedlings increased over the course of a 10-d methyl jasmonate
elicitation when compared to nonelicited controls (Fig. 5). We therefore used qPCR to track the expression levels of Cr2270, in addition to the four previously characterized vindoline 7a biosynthetic genes. We noted that all known vindoline 7a biosynthetic genes and the NMT Cr2270 gene were up-regulated relative to the nonelicited seedlings (Fig. S6). The expression profile for Cr2270 exhibits the same pattern as 16OMT, suggesting that expression of these genes is coordinately regulated (Fig. 5B). In contrast, expression of the putative methytransferase gene Cr1196, which appears to have no role in alkaloid biosynthesis, was not up-regulated by methyl jasmonate (Fig. 5B). The coordinate regulation of Cr2270 expression with other vindoline 7a biosynthetic genes in response to methyl jasmonate, along with the lack of transcriptional induction of a highly similar methyltransferase that is not involved in alkaloid biosynthesis (Cr1196; Fig. 2), provides further support for the role of Cr2270 as the 16-methoxy-2,3-dihydro-3-hydroxytabersonine N methyltransferase in vindoline 7a biosynthesis.

**On the Evolution of N-Methyltransferase Function.** Sequence analysis of Cr2270 and related proteins (Fig. 2 and Fig. S1) suggest that Cr2270 represents a previously undescribed class of NMT that likely arose from γ-TMTs. Arabidopsis thaliana possesses a single gene that encodes a functional γ-TMT (17), whereas Brassica napus (26) and C. roseus both have at least four γ-TMT-like genes. In the latter case, we now know that members of this gene family encode enzymes that are functionally distinct. We speculate that Cr2270 arose from γ-TMT gene duplication events, which would explain why these gene products reside in the same subcellular compartment. Moreover, although DHT methylation by Cr2270 is inhibited by γ-tocopherol (Fig. S7A), γ-tocopherol is not turned over by this enzyme (Fig. S7B). Perhaps the evolution of substrate specificity was facilitated by the presence of nitrogen-containing γ-tocopherol derivatives, such as 5-nitro-γ-tocopherol, which is generated in diverse plant systems (27). Rubisco large subunit methyltransferase (RubLSMT) is known to modulate γ-TMT activity through methylation of a C-terminal lysine residue (28). Interestingly, the RubLSMT consensus motif (28) is conserved in γ-TMTs, and Cr1196, but has been lost in Cr2270, Cr7756 (29) and Cr3710017 (Fig. S1).

**Conclusion**

The discovery of this NMT further closes the gap in the tabersonine to vindoline branch of monoterpene indole alkaloid biosynthesis, with only one uncharacterized class of NMT that likely arose from γ-TMTs. This pathway will facilitate metabolic engineering efforts for both enhanced production of bisindole alkaloids in plants or in heterologous expression systems. Moreover, these data will allow a detailed study of the substrate specificity of this enzyme, which controls the production of the distinct alkaloid structures produced naturally. Furthermore, the amino acid sequence of this enzyme provides a suggestion as to how one
piece of this specialized metabolic pathway evolved from an entirely distinct metabolic pathway.

Materials and Methods

Plant Growth. Catharanthus roseus cv. ‘Little Bright Eyes’ seeds were purchased from B & T World Seeds (http://b-and-t-world-seeds.com). Sterile seedlings were cultured and treated with methyl jasmonate according to Aerts et al. (25).

Chemicals. Tabersonine was generously provided by Viresh Rawal (University of Chicago, Chicago). Unless otherwise noted, all other chemicals were obtained from Sigma-Aldrich.

Synthesis of (2R,3S)-2,3-Dihydrotabersonine. Excess sodium cyanoborohydride (0.027 g, 0.430 mmol) was added to a solution of tabersonine (0.024 g, 0.071 mmol) in 2.4 mL of methanol-water (5:1) acidified with acetic acid (pH 4). This mixture was stirred under argon for 16 h at room temperature. Additional sodium cyanoborohydride (0.008 mg, 0.127 mmol) was then added, and the reaction was allowed to stir at room temperature for another 4 h. The mixture was then evaporated to dryness, resuspended in 500 µL of water, and then extracted 4 times with ethyl acetate (500 µL). The organic fractions were collected, evaporated to dryness, and purified by flash silica chromatography (100% MeCN) in quantitative yield to give (2R,3S)-2,3-dihydrotabersonine (0.024 g, 0.071 mmol). 1H-NMR (CD3OD, 400 MHz): δ 7.07 (d, 1H, J = 7.5 Hz), 6.95 (dd, 1H, J = 8.0, 7.5, 1.0 Hz), 6.60 (dd, 1H, J = 7.5, 7.5, 1.0 Hz), 5.61 (d, 1H, 8.0 Hz), 5.79–5.75 (m, 1H), 5.37–5.35 (m, 1H), 4.02 (d, 1H, J = 2.5 Hz), 3.77 (t, 3H, δ 3.84–3.60 (m, 1H), 3.34 (d, 1H, J = 5.5 Hz), 3.22 (dt, 1H, J = 9.0, 3.0 Hz), 2.62 (d, 1H, J = 15.5 Hz), 2.40–2.34 (m, 1H), 2.30 (s, 1H), 2.10 (t, 1H), 1.77 (s, 1H), 1.74–1.61 (m, 2H), 1.3–15.09 (m, 2H), 1.05–0.97 (m, 2H), 0.56 (t, 3H, J = 7.5 Hz).

Enzyme Assays. NMT assays contained 1 µg of purified protein, 0.78–200 µM 2,3-dihydrotabersonine, 1.13–1000 µM S-adenosyl-L-methionine, and 10 µM ajamiline (internal standard), in buffer C (100 mM NaPO4, pH 7.4; 20 mM L-mercaptoethanol, 1 mM EDTA), in a total volume of 50 µL. Assays were incubated at 30 °C for 30 or 60 min and quenched with the addition of 500 µL methanol. Reactions were then centrifuged at 17,000 × g for 10 min to pellet precipitated protein. Assays were further diluted 1:150, and 7 µL were analyzed by LC-MS.

Kinetic Analysis. Kinetic assays were performed under optimal conditions (30 °C, pH 7.4). For DHT and SAH kinetics, SAM was fixed at 200 µM, and for SAM kinetics DHT was fixed at 100 µM. Kinetic data were fitted by nonlinear regression and analyzed using GraphPad Prism version 4 for Macintosh (GraphPad Software).

Production and Purification of Recombinant Protein. Escherichia coli Rosetta 2 (DE3) pLysS cells (EMD) harboring the pET28-Cr2270, pET28-Cr6996, or pRSET-Cr1196 expression constructs were grown in 1 L of Luria–Bertani medium supplemented with 50 µg mL−1 kanamycin (pET28) or 100 µg mL−1 ampicillin (pRSET), 34 µg mL−1 chloramphenicol and 3% (vol/vol) ethanol for Cr2270 at 37 °C with shaking (220 rpm) to an OD600 of 0.6, and then induced with 1 mM IPTG. After 3 h of growth at 37 °C, cells were collected by centrifugation at 5,000 × g, flash-frozen in liquid N2, and stored at −80 °C until further use. Pelleted cells were resuspended in buffer A [100 mM Tris-HCl (pH 7.5), 100 mM KCl, 10% glycerol, 20 mM L-mercaptoethanol] (14) and lysed by sonication. For Cr2270 and Cr1196, Tergitol NP-10 (0.1%, w/v) was added to the supernatant, which was subsequently incubated for 2 h at 4 °C with gentle agitation. Cell debris was removed by centrifugation (10,000 × g) for 30 min at 4 °C. The supernatant was bound to Talon cobalt affinity resin (Clontech), which was then washed twice with 4 mL of buffer A and eluted in a stepwise manner with 750 µL of buffer A containing increasing concentrations (10, 50, 100, and 200 mM) of imidazole. To remove imidazole, these fractions were exchanged into buffer B [100 mM NaPO4 (pH 7.4), 25% glycerol, 20 mM L-mercaptoethanol] using a Zeba desalt spin column (Fisher).

Quantitative Real-Time PCR (qPCR). Gene expression analysis was performed as described by Runguphan et al. (29) using the qPCR primers listed in Table S2. In addition to C. roseus rpsb, EF1α was used as a second endogenous reference transcript (90). The ΔΔCT method was used for relative gene expression analysis (31).

Phylogenetic Analysis. Neighbor-joining phylogeny was generated, and bootstrap analysis was performed with TREECON (33). GenBank accession numbers for the sequences used are provided in Table S3.

ACKNOWLEDGMENTS. The authors are grateful to Dr. Viresh Rawal (University of Chicago, Chicago) for the gift of tabersonine, and Dr. Li Li of the DCF (Massachusetts Institute of Technology, Cambridge, MA) for obtaining exact mass data. We gratefully acknowledge support from GM074820. We also acknowledge additional support from National Science Foundation and the American Cancer Society. D.K.L. is the recipient of a Natural Sciences and Engineering Research Council of Canada postdoctoral fellowship. A.R.U. is the recipient of a National Institutes of Health postdoctoral fellowship.