Published in final edited form as:

Tetrahedron Lett. 2010 August 18; 51(33): 4400–4402. doi:10.1016/j.tetlet.2010.06.075.

Biocatalytic asymmetric formation of tetrahydro- β -carboline

Peter Bernhardt^a, Aimee R. Usera^a, and Sarah E. O'Connor^{a,*}

^aDepartment of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Abstract

Strictosidine synthase triggers the formation of strictosidine from tryptamine and secologanin, thereby generating a carbon-carbon bond and a new stereogenic center. Strictosidine contains a tetrahydro- β -carboline moiety – an important N-heterocyclic framework found in a range of natural products and synthetic pharmaceuticals. Stereoselective methods to produce tetrahydro- β -carboline enantiomers are greatly valued. We report that strictosidine synthase from *Ophiorrhiza pumila* utilizes a range of simple achiral aldehydes and substituted tryptamines to form highly enantioenriched (ee >98%) tetrahydro- β -carbolines via a Pictet-Spengler reaction. This is the first example of aldehyde substrate promiscuity in the strictosidine synthase family of enzymes and represents a first step towards developing a general biocatalytic strategy to access chiral tetrahydro- β -carbolines.

The Pictet-Spengler reaction between tryptamine and an aldehyde yields the tetrahydro- β -carboline moiety, a heterocyclic framework found in a wide range of natural products and synthetic pharmaceuticals. Catalytic methodology has been developed to stereoselectively generate the C-C bond formed in this reaction.¹ Chiral tetrahydro- β -carbolines are also formed enzymatically in natural product biosynthetic pathways (Figure 1).² Although biocatalysis has proven useful for a number of synthetic transformations,³ natural product biosynthetic enzymes often have narrow substrate scopes that limits use of these enzymes as biocatalysts. Herein, we describe a Pictet-Spenglerase with a broad substrate range that asymmetrically generates tetrahydro- β -carbolines from tryptamine and a range of aldehydes.

Strictosidine synthase (STS), a Pictet-Spenglerase utilized in monoterpene and quinoline alkaloid biosynthesis, catalyzes the conversion of tryptamine **1** and secologanin **2** to the tetrahydro- β -carboline strictosidine **3** (Figure 1).² Two strictosidine synthase homologs from the medicinal plants *Rauvolfia serpentina* (RsSTS)^{4a} and *Catharanthus roseus* (CrSTS),^{4b} which share 82% amino acid sequence identity overall and near-identical active sites, have been studied in detail. Both of these homologs can turnover a variety of electron deficient and substituted tryptamine analogs.^{5,6} Site-directed mutagenesis of these enzymes has been successfully used to broaden the scope of strictosidine synthase for larger tryptamine derivatives,^{7a,b} However, only minor changes to the aldehyde substrate **2** are tolerated.^{5,7c} Therefore, any tetrahydro- β -carbolines generated by strictosidine synthase will have little application outside the context of alkaloid biosynthetic pathways that use secologanin **2**.

© 2010 Elsevier Ltd. All rights reserved.

* Corresponding author. Tel.: 617-324-0180;; soc@mit.edu (S. E. O'Connor)..

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Supporting Information (SI) **Available:** Experimental methods, modeling data and NMR spectra.

The recently cloned strictosidine synthase from *Ophiorrhiza pumila* (OpSTS)^{4c} has a lower sequence identity to both CrSTS (54%) and RsSTS (60%). This observation suggested that the active site of OpSTS could have different structural features, perhaps a result of the evolutionary distance between the Apocynaceae plant family from which CrSTS and RsSTS are derived, and the Rubiaceae family from which OpSTS is derived. Sequence alignments showed that OpSTS has a four-amino acid deletion that corresponds to $\Delta[\text{Gln282His283Gly284Arg285}]$ in CrSTS and $\Delta[\text{Met276His277Gly278Arg279}]$ in RsSTS. The crystal structure of RsSTS in complex with secologanin **2** (PDB: 2FPC)⁶ indicated that these amino acids are in the proximity of the secologanin binding site. Moreover, His283 (His277 in RsSTS), which is missing in OpSTS, is within hydrogen bonding distance to the glucose moiety of **2**. This observation prompted us to assess whether OpSTS has a broader, and therefore potentially more useful, aldehyde substrate scope compared to CrSTS and RsSTS.

Pseudo steady-state kinetics with **1**, secologanin **2** and OpSTS ($k_{\text{cat}} = 1.1 \text{ s}^{-1}$, $K_{\text{m},2} = 21 \text{ }\mu\text{M}$) or CrSTS ($k_{\text{cat}} = 5.1 \text{ s}^{-1}$, $K_{\text{m},2} = 64 \text{ }\mu\text{M}$) verified that the heterologously expressed enzyme is active (Supporting Information, SI). Aldehydes **4-14** (1 mM) and tryptamine (1 mM) were incubated with OpSTS (0.2 mol% catalyst, 2 μM) or CrSTS (1.0 mol% catalyst, 10 μM) in aqueous buffer (pH 7.0). The resulting tetrahydro- β -carboline products were identified by mass spectrometry and by comparison with a conveniently prepared racemic authentic standard (SI).⁸ CrSTS turned over substrate **4** poorly, while no products formed with substrates **5-12** (Table 1). In contrast, OpSTS turned over aldehydes **4-10**. Although OpSTS displayed a preference for the natural substrate **2** by at least 1000-fold, this enzyme nevertheless catalyzed the conversion of a relatively broad range of substrates, including hydrogen bond accepting (**4**), aliphatic (**5-6** and **8**), and aromatic (**7**) aldehydes with similar rates ($k_{\text{obs}} = 0.013\text{-}0.048 \text{ min}^{-1}$). The k_{obs} values clearly indicated that OpSTS preferred aldehydes unsubstituted at the alpha position (compare **4-8** with **10-12**). A full kinetic analysis was not performed because of the high estimated K_{m} values for the aldehydes ($\geq 10 \text{ mM}$) and substrate solubility limitations.

The enantiomeric excess (*ee*) of OpSTS was >98% as evidenced by chiral HPLC (SI).⁹ The absolute configuration was determined for the enzymatic product derived from **1** and **8**. The reaction was carried out on milligram scale using OpSTS immobilized to a solid support (SI),¹⁰ and compared to an authentic 3(*S*)-standard obtained by an organocatalytic method. ^{1a} This confirmed that the absolute configuration of this enzymatic product is 3(*S*), which is the configuration observed with the natural substrate, secologanin **2** (SI).

Notably, the stereoselectivity observed for the reaction between **1** and **2** with RsSTS, CrSTS, and OpSTS was preserved with achiral aldehydes, indicating that stereoselectivity does not rely on the chirality of substrate **2**. The configuration at C3 is determined during the cyclization of iminium **13** (derived from condensation of **1** and **2**) to yield **14**, which is then deprotonated to form **3** (Figure 2A).^{11a,b}

To develop a model for stereoselectivity, computationally derived structures of **14**^{11a} were docked into the crystal structure of RsSTS (SI). The **14-2(R)3(R)** isomer could only be docked after considerable changes to the active site structure and substrate orientation, suggesting this intermediate is not formed in the active site (SI). Out of the three remaining isomers [**14-2(R)3(S)**, **14-2(S)3(S)**, and **14-2(S)3(R)**], only **14-2(R)3(S)** was appropriately positioned for deprotonation at C-2 by the catalytic base Glu309 (Figure 2B and SI). We propose that although alternate diastereomers of **14** may be formed within the active site, they cannot be deprotonated. Instead, these intermediates would undergo the reverse reaction, reforming iminium **13**, and cyclization would be repeated until the productive **14-2(R)3(S)** isomer is formed. This facial deprotonation model is consistent with the

reversible nature of the Pictet-Spengler reaction,^{11c} but should be interpreted cautiously, since it has not been experimentally validated.^{11d} Nevertheless, this model serves as a basis for understanding and engineering the stereoselectivity of Pictet-Spenglerases.

In an attempt to broaden the substrate scope of CrSTS, we mutated His283 for Gly/Leu/Phe and deleted Δ [282–285]. However, these modifications did not confer the substrate scope of OpSTS to CrSTS. His313 (His307 in RsSTS, His299 in OpSTS) is also within hydrogen bonding distance to the glucose moiety of secologanin **2**. However, neither His313Leu nor the double mutant His283Leu/His313Leu resulted in a broadened aldehyde substrate specificity. The factors that control the substrate specificity of **2** are complex, and mutations distant from the active site to impact substrate specificity are currently being explored.

The Rubiaceae family of plants contain chiral tetrahydro- β -carbolines different from those found in the Apocynaceae family.¹² The enzymes that catalyze formation of the precursor for these natural products are unknown, but it is tempting to speculate that they resemble OpSTS. Screening of Rubiaceae plants for new strictosidine synthase homologs and assay of the predicted substrates will experimentally test this possibility, and potentially broaden the base of biocatalysts available for use in this reaction. Although the catalytic rates using OpSTS are relatively low with simple aldehyde substrates, high stereoselectivity is maintained. Protein engineering may improve catalytic efficiency. This discovery allows the asymmetric biocatalytic formation of a wide range of tetrahydro- β -carbolines directly from tryptamine, and represents the first report of a Pictet-Spenglerase with broad aldehyde specificity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by the NIH (GM074820); ARU acknowledges NIH for a postdoctoral fellowship. We thank Omar Ahmad for assistance with chiral HPLC.

References and Notes

1. a Taylor MS, Jacobsen EN. *J. Am. Chem. Soc.* 2004; 126:10558–10559. [PubMed: 15327311] b Seayad J, Seayad AM, List B. *J. Am. Chem. Soc.* 2006; 128:1086–1087. [PubMed: 16433519] c Raheem IT, Thiara PS, Peterson EA, Jacobsen EN. *J. Am. Chem. Soc.* 2007; 129:13404–13405. [PubMed: 17941641] d Wanner MJ, van der Haas RN, de Cuba KR, van Maarseveen JH, Hiemstra H. *Angew. Chem. Int. Ed. Engl.* 2007; 46:7485–7487. [PubMed: 17702087] e Sewgobind NV, Wanner MJ, Ingemann S, de Gelder R, van Maarseveen JH, Hiemstra H. *J. Org. Chem.* 2008; 73:6405–6408. [PubMed: 18616320] f Klausen RS, Jacobsen EN. *Org. Lett.* 2009; 11:887–890. [PubMed: 19178157] g Bou-Hamdan FR, Leighton JL. *Angew. Chem. Int. Ed. Engl.* 2009; 48:2403–2406. [PubMed: 19222080]
2. O'Connor SE, Maresh JJ. *Nat. Prod. Rep.* 2006; 23:532–547. [PubMed: 16874388]
3. a Faber, K. *Biotransformations in Organic Synthesis*. 5th ed.. Springer; 2008. b Bornscheuer, UT.; Kazlauskas, RJ. *Hydrolases in Organic Synthesis*. 2nd ed.. Wiley-VCH; 2006.
4. a Kutchan TM, Hampp N, Lottspeich F, Beyreuther K, Zenk MH. *FEBS Lett.* 1988; 237:40–44. [PubMed: 3049153] b McKnight TD, Roessner CA, Devagupta R, Scott AI, Nessler CL. *Nucl. Acids Res.* 1990; 18:4939. [PubMed: 2395663] c Yamazaki Y, Sudo H, Yamazaki M, Aimi N, Saito K. *Plant Cell Physiol.* 2003; 44:395–403. [PubMed: 12721380]
5. a McCoy EA, Galan MC, O'Connor SE. *Bioorg. Med. Chem. Lett.* 2006; 16:2475–2478. [PubMed: 16481164] b Treimer JF, Zenk MH. *Eur. J. Biochem.* 1979; 101:225–233. [PubMed: 510306]
6. Ma X, Panjikar S, Koepke J, Loris E, Stöckigt J. *Plant Cell.* 2006; 23:532–547.

7. a Bernhardt P, McCoy EA, O'Connor SE. Chem. Biol. 2007; 14:888–897. [PubMed: 17719488] b Loris EA, Panjikar S, Ruppert M, Barleben L, Unger M, Schubel H, Stockigt J. Chem. Biol. 2007; 14:979–85. [PubMed: 17884630] c Chen S, Galan MC, Coltharp C, O'Connor SE. Chem. Biol. 2006; 13:1137–1141. [PubMed: 17113995]
8. Racemic authentic standards were synthesized from tryptamine hydrochloride (0.5 M) and achiral aldehydes (0.5 M) in aqueous maleic acid buffer (10 mM, pH 2.0, total volume 2 ml) at 60 °C. Either during the reaction or upon standing, the tetrahydro- β -carboline product precipitated and could be filtered and washed with water (3×2 mL) to afford analytically pure product. See Supporting Information for NMR spectra and structures of products **15** through **21**. 1-((tetrahydro-2H-pyran-4-yl)methyl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (product **15** from **1** and **4**)- ¹H NMR (methanol-d₄): δ 7.48 (1H, dt, J = 0.9, 7.9), 7.36 (1H, dt, J = 0.9, 8.2), 7.16 (1H, td, J = 1.1, 8.2), 7.06 (1H, td, J = 0.9, 8.0), 4.86-4.82 (1H, m), 4.03-3.96 (2H, m), 3.75 (1H, dt, J = 1.3, 4.1, 12.4), 3.55-3.44 (3H, m), 3.12-3.07 (2H, m), 2.21-2.15 (1H, m), 1.95-1.83 (3H, m), 1.68 (1H, td, J = 1.9, 13.1), 1.50-1.37 (2H, m); ¹³C NMR (methanol-d₄): δ 137.61, 129.58, 126.73, 122.80, 119.91, 118.34, 111.63, 106.54, 68.03, 67.87, 51.08, 42.39, 39.94, 33.93, 32.40, 31.09, 18.80; ESI-MS(C₁₇H₂₃N₂O⁺) m/z calculated: 271.1810 [M+H]⁺, found: 271.1805 [M+H]⁺. 1-(cyclohexylmethyl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (product **16** from **1** and **5**)- ¹H NMR (methanol-d₄): δ 7.47-7.44 (1H, m), 7.36 (1H, td, J = 0.8, 8.2), 7.14 (1H, ddd, J = 1.1, 7.1, 8.2), 7.04 (1H, ddd, J = 1.0, 7.1, 8.0), 4.81-4.74 (1H, m), 3.71 (1H, ddd, J = 3.5, 5.5, 12.5), 3.41 (1H, ddd, J = 5.5, 9.9, 12.6), 3.09 (1H, dddd, J = 1.9, 5.5, 9.8, 15.4), 3.05-2.99 (1H, m), 2.12 (1H, ddd, J = 3.7, 10.2, 14.3), 2.08-2.00 (1H, m), 1.87-1.60 (6H, m), 1.49-1.34 (2H, m), 1.33-1.22 (1H, m), 1.19-1.02 (2H, m); ¹³C NMR (methanol-d₄): δ 137.59, 129.95, 126.75, 122.69, 119.84, 118.31, 111.64, 106.39, 51.57, 42.46, 40.53, 34.56, 33.59, 32.41, 26.74, 26.54, 26.16, 18.81; ESI-MS(C₁₈H₂₅N₂⁺) m/z calculated: 269.2012 [M+H]⁺, found: 269.2014 [M+H]⁺. 1-isobutyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (product **17** from **1** and **6**)- ¹H NMR (10:1, (methanol-d₄/DMSO-d₆) δ 7.50 (1H, td, J = 1.0, 7.8), 7.39 (1H, td, J = 0.9, 8.2), 7.17 (1H, ddd, J = 1.2, 7.1, 8.2), 7.08 (1H, ddd, J = 1.0, 7.1, 8.0), 4.82-4.79 (1H, m), 3.74 (1H, ddd, J = 3.9, 5.3, 12.6), 3.47 (1H, ddd, J = 5.9, 9.3, 12.6), 3.16-3.02 (2H, m), 2.08 (1H, ddd, J = 3.9, 10.2, 14.2), 1.99 (1H, dqd, J = 3.9, 6.5, 19.5), 1.85 (1H, ddd, J = 3.9, 10.3, 14.2), 1.15 (3H, d, J = 6.4), 1.10 (3H, d, J = 6.5); ¹³C NMR (DMSO-d₆): δ 137.22, 131.64, 126.98, 122.74, 120.07, 119.09, 112.46, 106.81, 51.51, 41.94, 24.67, 24.53, 22.51, 19.16; ESI-MS(C₁₅H₂₁N₂⁺) m/z calculated: 229.1699 [M+H]⁺, found: 229.1697 [M+H]⁺. 1-benzyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (product **18** from **1** and **7**)- ¹H NMR (methanol-d₄): δ 7.55-7.37 (m, 7H), 7.21 (1H, ddd, J = 1.1, 7.1, 8.2), 7.11 (1H, dd, J = 1.0, 7.1, 8.0), 5.06-5.01 (1H, m), 3.79-3.73 (1H, m), 3.62 (1H, ddd, J = 3.6, 5.3, 12.5), 3.36 (1H, ddd, J = 5.7, 9.6, 12.6), 3.21-3.01 (3H, m); ¹³C NMR (methanol-d₄): δ 136.68, 135.82, 130.10, 129.87, 127.71, 126.26, 122.36, 119.54, 118.59, 111.90, 106.72, 54.20, 41.90, 37.76, 18.45; ESI-MS(C₁₈H₁₉N₂⁺) m/z calculated: 263.1543 [M+H]⁺, found: 263.1540 [M+H]⁺. 1-pentyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (product **19** from **1** and **8**)- ¹H NMR (methanol-d₄): δ 7.48 (1H, dt, J = 0.9, 7.9), 7.36 (1H, dt, J = 0.9, 8.2), 7.15 (1H, td, J = 1.1, 8.2), 7.06 (1H, td, J = 0.9, 8.0), 4.69 (1H, dd, J = 4.1, 8.9), 3.73 (1H, ddd, J = 3.8, 5.5, 12.5), 3.44 (1H, ddd, J = 5.6, 9.6, 12.6), 3.11 (1H, dddd, J = 1.9, 5.6, 9.5, 15.1), 3.05 (1H, dddd, J = 1.3, 3.8, 5.3, 16.3), 2.31-2.23 (1H, m), 1.97-1.89 (1H, m), 1.64-1.54 (2H, m), 1.53-1.37 (4H, m), 0.97 (3H, t, J = 7.1); ¹³C NMR (methanol-d₄): δ 138.36, 130.32, 127.46, 123.48, 120.59, 119.10, 112.39, 107.18, 55.17, 43.19, 33.38, 32.79, 25.92, 23.48, 19.54, 14.37; ESI-MS(C₁₆H₂₃N₂⁺) m/z calculated: 243.1856 [M+H]⁺, found: 243.1851 [M+H]⁺. 1-ethyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (product **20** from **1** and **9**)- ¹H NMR (methanol-d₄): δ 7.47 (1H, td, J = 0.9, 7.9), 7.37 (1H, td, J = 0.9, 8.2), 7.15 (1H, ddd, J = 1.1, 7.1, 8.2), 7.06 (1H, ddd, J = 1.0, 7.1, 8.0), 4.63-4.57 (1H, m), 3.71 (1H, ddd, J = 3.7, 5.6, 12.5), 3.39 (1H, ddd, J = 5.5, 9.7, 12.6), 3.10 (1H, dddd, J = 2.0, 5.6, 9.7, 15.3), 3.01 (1H, dddd, J = 1.4, 3.7, 5.3, 16.3), 2.33 (1H, dqd, J = 4.2, 7.6, 15.1), 2.05-1.93 (1H, m), 1.19 (3H, t, J = 7.5); ¹³C NMR (methanol-d₄): δ 138.30, 130.14, 127.42, 123.46, 120.57, 119.13, 112.38, 107.25, 56.19, 43.15, 26.33, 19.52, 9.93; ESI-MS(C₁₃H₁₇N₂⁺) m/z calculated: 201.1386 [M+H]⁺, found: [M+H]⁺ 201.1385. 1-cyclohexyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (product **21** from **1** and **10**)- ¹H NMR (methanol-d₄): δ 7.45 (1H, d, J = 7.9), 7.38 (1H, d, J = 8.2), 7.14 (1H, ddd, J = 1.1, 7.2, 8.2), 7.04 (1H, ddd, J = 1.0, 7.2, 8.0), 4.60-4.57 (1H, m), 3.71 (1H, ddd, J = 3.2, 5.5, 12.4), 3.38 (1H, ddd, J = 5.4, 10.1, 12.5), 3.10 (1H, dddd, J = 2.0, 5.6, 10.1, 15.8), 3.04-2.95 (1H, m), 2.35-2.22 (1H, m), 1.98-1.68 (4H, m), 1.57-1.10 (6H, m); ¹³C NMR (methanol-d₄): δ 137.64, 128.27, 126.71, 122.76,

119.84, 118.31, 111.65, 107.40, 59.19, 43.15, 40.32, 30.12, 27.07, 26.82, 26.43, 26.41, 18.70; ESI-MS(C₁₇H₂₃N₂⁺) m/z calculated: 255.1856 [M+H]⁺, found: 255.1858 [M+H]⁺.

9. Enantiomeric products derived from **4**, **5**, **7**, **8** were successfully separated (SI), but enantiomeric products derived from **6** could not be separated despite testing many chromatography conditions.
10. Pfitzner U, Zenk MH. *Methods Enzymology*. 1987; 136:342–350.
11. Maresh JJ, Giddings LA, Friedrich A, Loris EA, Panjikar S, Trout BL, Stöckigt J, Peters B, O'Connor SE. *J. Am. Chem. Soc.* 2008; 130:710–723. [PubMed: 18081287] b) Computational studies (11a) suggest the spiroindolenine intermediate is not a productive intermediate; attack from indole C2 to directly form **14** is the productive mechanism. c) Kinetic isotope effects suggest the Pictet-Spengler reaction is reversible prior to the final deprotonation step (11a). d) Point mutation of Phe226 (Figure 1B) to numerous ionizable residues did not impact the stereoselectivity of the reaction. More extensive mutagenesis of STS is underway to experimentally validate this hypothesis.
12. Lopes S, von Poser GL, Kerber VA, Farias FM, Konrath EL, Moreno P, Sobral ME, Zuanazzi JAS, Henriques AT. *Biochem. Syst. Ecol.* 2004; 32:1187–1195.

Figure 1. Pictet-Spengler reaction catalyzed by strictosidine synthase. The tetrahydro-β-carboline moiety is highlighted in red, the C-C bond formed is shown in bold, and the stereogenic center (C3) is marked with an asterisk.

Figure 2.
A) Reaction sequence leading to 3. B) Computer generated models of two diastereomers of 14 docked in RsSTS (RsSTS numbering).⁶ Arrows highlight H-2 that is deprotonated to form 3.

Table 1

Aldehyde scope for OpSTS and CrSTS.

Starting material R =		k_{obs} [min^{-1}] ^a	
		OpSTS	CrSTS
	4	0.048	0.0002
	5	0.037	not obsvd. ^b
	6	0.032	not obsvd. ^b
	7	0.022	not obsvd. ^b
	8	0.013	not obsvd. ^b
	9	<0.0001 ^c	not obsvd. ^b
	10	<0.0001 ^c	not obsvd. ^b
	11	not obsvd. ^b	not obsvd. ^b
	12	not obsvd. ^b	not obsvd. ^b

^a k_{obs} measured by monitoring product formation with **1** at 280 nm using reverse-phase HPLC with a PDA detector^b not observed, detection limit $\sim 10^{-5} \text{ min}^{-1}$ ^c quantitation limit $\sim 10^{-4} \text{ min}^{-1}$. k_{obs} of **4** was 230-fold improved for OpSTS compared to CrSTS; k_{obs} of **5-10** were at least 3700-fold improved, given the detection limit of 10^{-5} min^{-1} .