Contents lists available at ScienceDirect

Tetrahedron Letters

journal homepage: www.elsevier.com/locate/tetlet

Biocatalytic asymmetric formation of tetrahydro-β-carbolines

Peter Bernhardt, Aimee R. Usera, Sarah E. O'Connor *

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, United States

ARTICLE INFO

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 toward developing a general biocatalytic strategy to access chiral tetrahydro-β-carbolines.

© 2010 Elsevier Ltd. All rights reserved.

The Pictet–Spengler reaction between tryptamine and an aldehyde yields 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 (Fig. 1).[2] Although biocatalysis has been proven useful for a number of synthetic transformations,[3] natural product biosynthetic enzymes often have narrow substrate scopes that limit the 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 (Fig. 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 these homologs can turnover a variety of electron-deficient and electron-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.

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 Δ[Gln282His283Gly284Arg285] in CrSTS and Δ[Met276His277-Gly278Arg279] 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_{cat} = 1.1 s^{-1}, $K_{m,2}$ = 21 μM) or CrSTS (k_{cat} = 5.1 s^{-1}, $K_{m,2}$ = 64 μM) verified that the heterologously expressed enzyme is active (Supplementary data). 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 (Supplementary data).[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_{obs} = 0.013–0.048 min^{-1}). The

---

* Corresponding author. Tel.: +1 617 324 0180.
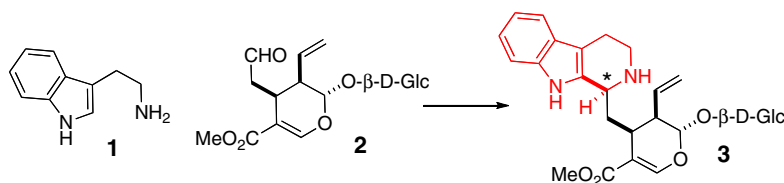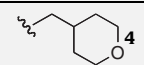E-mail address: soc@mit.edu (S.E. O'Connor).

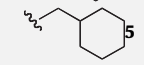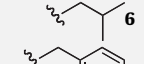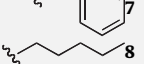**Figure 1.** Pictet–Spengler reaction catalyzed by strictosidine synthase. The tetrahydro-β-carboline moiety is highlighted in red, the C–C bond formed is shown in bold, and the stereogenic center (C3) is marked with an asterisk.

Table 1
Aldehyde scope for OpSTS and CrSTS

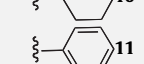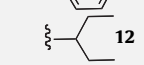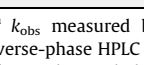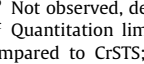| Starting material R =	k_{obs}[a] (min^{-1})	
	OpSTS	CrSTS
4	0.048	0.0002
5	0.037	Not obsd[b]
6	0.032	Not obsd[b]
7	0.022	Not obsd[b]
8	0.013	Not obsd[b]
9	<0.0001[c]	Not obsd[b]
10	<0.0001[c]	Not obsd[b]
11	Not obsd[b]	Not obsd[b]
12	Not obsd[b]	Not obsd[b]

[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 ~10^{-5} min^{-1}.

[c] Quantitation limit ~10^{-4} 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}.

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 ($\geqslant 10$ mM) and substrate solubility limitations.

The enantiomeric excess (ee) of OpSTS was >98% as evidenced by chiral HPLC (Supplementary data).[9] The absolute configuration was determined for the enzymatic product derived from **1** and **8**. The reaction was carried out on a milligram scale using OpSTS immobilized to a solid support (Supplementary data),[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** (Supplementary data).

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** (Fig. 2A).[11a,b]

To develop a model for stereoselectivity, computationally derived structures of **14**[11a] were docked into the crystal structure of RsSTS (Supplementary data). The **14**-2(R)3(R) isomer could only be docked after considerable changes to the active site structure and substrate orientation, suggesting that this intermediate is not formed in the active site (Supplementary data). 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 (Fig. 2B and Supplementary data). We propose that although alternate diastereomers of **14** may be
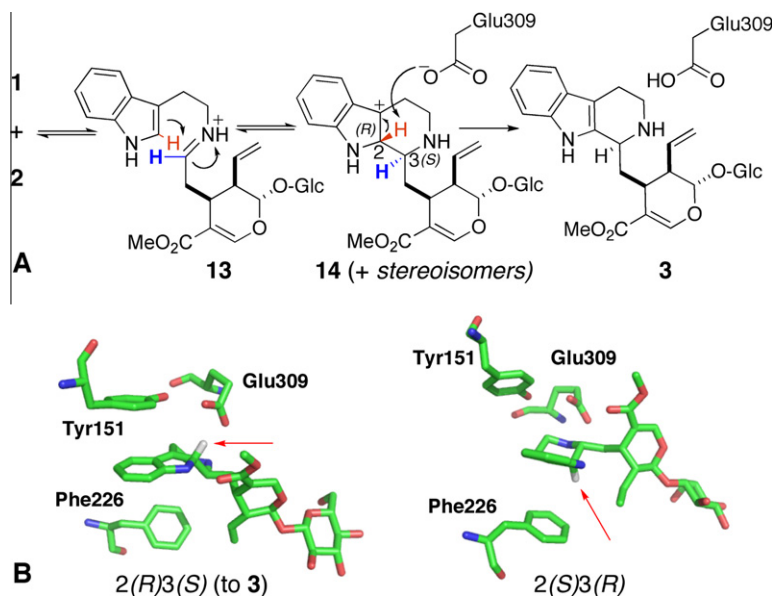

Figure 2. (A) Reaction sequence leading to **3**. (B) Computer-generated models of two diastereomers of **14** docked in RsSTS (RsSTS numbering).[6] Arrows highlight H-2 that is deprotonated to form **3**.

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 contains chiral tetrahydro-β-carbolines different from those found in the Apocynaceae family.[12] The enzymes that catalyze the 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.

Acknowledgments

Supplementary data

Supplementary data (experimental methods, modeling data and NMR spectra) associated with this article can be found, in the online version, at doi:10.1016/j.tetlet.2010.06.075.

References and notes

1. (a) Taylor, M. S.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2004**, *126*, 10558–10559; (b) Seayad, J.; Seayad, A. M.; List, B. *J. Am. Chem. Soc.* **2006**, *128*, 1086–1087; (c) Raheem, I. T.; Thiara, P. S.; Peterson, E. A.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2007**, *129*, 13404–13405; (d) Wanner, M. J.; van der Haas, R. N.; de Cuba, K. R.; van Maarseveen, J. H.; Hiemstra, H. *Angew. Chem., Int. Ed.* **2007**, *46*, 7485–7487; (e) Sewgobind, N. V.; Wanner, M. J.; Ingemann, S.; de Gelder, R.; van Maarseveen, J. H.; Hiemstra, H. *J. Org. Chem.* **2008**, *73*, 6405–6408; (f) Klausen, R. S.; Jacobsen, E. N. *Org. Lett.* **2009**, *11*, 887–890; (g) Bou-Hamdan, F. R.; Leighton, J. L. *Angew. Chem., Int. Ed.* **2009**, *48*, 2403–2404.
2. O'Connor, S. E.; Maresh, J. J. *Nat. Prod. Rep.* **2006**, *23*, 532–547.
3. (a) Faber, K. *Biotransformations in Organic Synthesis*, 5th ed.; Springer, 2008; (b) Bornscheuer, U. T.; Kazlauskas, R. J. *Hydrolases in Organic Synthesis*, 2nd ed.; Wiley-VCH, 2006.
4. (a) Kutchan, T. M.; Hampp, N.; Lottspeich, F.; Beyreuther, K.; Zenk, M. H. *FEBS Lett.* **1988**, *237*, 40–44; (b) McKnight, T. D.; Roessner, C. A.; Devagupta, R.; Scott, A. I.; Nessler, C. L. *Nucleic Acids Res.* **1990**, *18*, 4939; (c) Yamazaki, Y.; Sudo, H.; Yamazaki, M.; Aimi, N.; Saito, K. *Plant Cell Physiol.* **2003**, *44*, 395–403.
5. (a) McCoy, E. A.; Galan, M. C.; O'Connor, S. E. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2475–2478; (b) Treimer, J. F.; Zenk, M. H. *Eur. J. Biochem.* **1979**, *101*, 225–233.
6. Ma, X.; Panjikar, S.; Koepke, J.; Loris, E.; Stöckigt, J. *Plant Cell* **2006**, *23*, 532–547.
7. (a) Bernhardt, P.; McCoy, E. A.; O'Connor, S. E. *Chem. Biol.* **2007**, *14*, 888–897; (b) Loris, E. A.; Panjikar, S.; Ruppert, M.; Barleben, L.; Unger, M.; Schubel, H.; Stöckigt, J. *Chem. Biol.* **2007**, *14*, 979–985; (c) Chen, S.; Galan, M. C.; Coltharp, C.; O'Connor, S. E. *Chem. Biol.* **2006**, *13*, 1137–1141.
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 Supplementary data 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**): ^{1}H NMR (methanol-d_4): 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); ^{13}C NMR (methanol-d_4): 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 ($C_{17}H_{23}N_2O^+$) *m/z* calcd: 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**): ^{1}H NMR (methanol-d_4): 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); ^{13}C NMR (methanol-d_4): 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 ($C_{18}H_{25}N_2^+$) *m/z* calcd: 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**): ^{1}H NMR (10:1, (methanol-d_4/DMSO-d_6) 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); ^{13}C NMR (DMSO-d_6): 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 ($C_{15}H_{21}N_2^+$) *m/z* calcd: 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**): ^{1}H NMR (methanol-d_4): 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); ^{13}C NMR (methanol-d_4): 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 ($C_{18}H_{19}N_2^+$) *m/z* calcd: 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**): ^{1}H NMR (methanol-d_4): 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); ^{13}C NMR (methanol-d_4): 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 ($C_{16}H_{23}N_2^+$) *m/z* calcd: 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**): ^{1}H NMR (methanol-d_4): 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, 3.8, 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); ^{13}C NMR (methanol-d_4): 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 ($C_{13}H_{17}N_2^+$) *m/z* calcd: 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**): ^{1}H NMR (methanol-d_4): 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); ^{13}C NMR (methanol-d_4): 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 ($C_{17}H_{23}N_2^+$) *m/z* calcd: 255.1856 [M+H]$^+$, found: 255.1858 [M+H]$^+$.
9. Enantiomeric products derived from **4**, **5**, **7**, and **8** were successfully separated (Supplementary data), but enantiomeric products derived from **6** could not be separated despite testing many chromatography conditions.
10. Pfitzner, U.; Zenk, M. H. *Methods Enzymol.* **1987**, *136*, 342–350.
11. (a) Maresh, J. J.; Giddings, L. A.; Friedrich, A.; Loris, E. A.; Panjikar, S.; Trout, B. L.; Stöckigt, J.; Peters, B.; O'Connor, S. E. *J. Am. Chem. Soc.* **2008**, *130*, 710–723; (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, G. L.; Kerber, V. A.; Farias, F. M.; Konrath, E. L.; Moreno, P.; Sobral, M. E.; Zuanazzi, J. A. S.; Henriques, A. T. *Biochem. Syst. Ecol.* **2004**, *32*, 1187–1195.