# Supplementary Information

## Fungal-derived Brevianamide assembly by a stereoselective semi-pinacolase

Ying Ye,[1#] Lei Du,[2,3#] Xingwang Zhang,[2,4] Sean A. Newmister,[1] Morgan McCauley,[5] Juan V. Alegre-Requena,[5] Wei Zhang,[2] Shuai Mu,[6] Atsushi Minami,[7] Amy E. Fraley,[1] Maria L. Adrover-Castellano,[1] Nolan A. Carney,[1] Vikram V. Shende,[1] Feifei Qi,[3] Hideaki Oikawa,[7] Hikaru Kato,[8] Sachiko Tsukamoto,[8] Robert S. Paton,[5,9] Robert M. Williams,[5,10*] David H. Sherman,[1,11*] and Shengying Li[2,3,4*]

[1]Life Sciences Institute, University of Michigan, Ann Arbor, Michigan 48109, USA
[2]State Key Laboratory of Microbial Technology, Shandong University, Qingdao, Shandong, 266237, China
[3]Shandong Provincial Key Laboratory of Synthetic Biology, CAS Key Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, Shandong, 266101, China
[4]Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao, Shandong 266237, China
[5]Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, USA
[6]Tianjin Institute of Pharmaceutical Research, Tianjin, 300193, China
[7]Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan
[8]Graduate School of Pharmaceutical Sciences, Kumamoto University, 5-1 Oe-honmachi, Kumamoto 862-0973, Japan
[9]Chemical Research Laboratory, University of Oxford, Mansfield Road, Oxford, OX1 3TA, UK
[10]University of Colorado Cancer Center, Aurora, CO 80045, USA
[11]Departments of Medicinal Chemistry, Chemistry and Microbiology & Immunology, University of Michigan, Ann Arbor, MI, USA

[#] These authors contributed equally.

[*]Shengying Li: lishengying@sdu.edu.cn
[*]David H. Sherman: davidhs@umich.edu
[*]Robert M. Williams: robert.williams@colostate.edu

**Supplementary Methods**

## Plasmid construction, DNA extraction and manipulation.

E.Z.N.A.$^{TM}$ Plasmid Miniprep Kit (Omega Biotek, Norcross, GA) was used for plasmids isolation. E.Z.N.A.$^{TM}$ Gel Extraction Kit (Omega Biotek, Norcross, GA) was used for DNA fragment purification. Vectors were constructed using T4 DNA ligase (Thermo Fisher Scientific, USA), recombinase (Pro Ligation-Free Cloning Kit, Applied Biological Materials Inc., Canada), In-Fusion HD Cloning Kit (Takara Bio USA), and NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, USA). Filamentous fungal genomic DNA was extracted using Rapid Fungi Genomic DNA Isolation Kit (Sangon Biotech, Shanghai, China). Restriction enzyme digestion, ligation, gel electrophoresis, vector transformation, and other standard techniques of molecular cloning were performed according to general methods[1].

## Determination of pH and temperature dependence of BvnE

To determine the pH and temperature dependency of BvnE, the *in vitro* assays as described above were conducted under the pH values varying from 5.0-9.0 with an interval of 0.5 or the temperatures ranging from 20-45 °C with an interval of 5 degrees. All reactions were quenched after an incubation of 20 min. Conversion rates were calculated based on HPLC analysis. All reactions were carried out in duplicate and the data were showed as means ± SD.

## Kinetic analysis of BvnE

In the kinetic assays, the concentrations of substrate **15** ranged from 0.05 to 4 mM. BvnE was diluted to a suitable concentration to ensure that the consumption of substrate was within the linear range during the reaction. Samples from each individual reaction were taken at 0, 5, 10 and 15 min. The substrate/product concentrations were calculated based on the integration of HPLC peak areas. The $K_m$ and $k_{cat}$ values were deduced by nonlinear regression fitting to the Michaelis-Menten equation. All reactions were carried out in duplicated and the data were shown as means ± SD.

## Analytical methods

For *Pb* fermentation, 6-d solid CD plates were cut into pieces and extracted with an ethyl acetate (EtOAc) and methanol (MeOH) mixture (EtOAc : MeOH = 85:15). Extract was centrifuged, dried and re-suspended with 1 mL MeOH. For *Ao* feeding experiments, spores of *Ao*-WT or mutants were inoculated into 10 mL of CMP medium containing appropriate nutrients in 50 mL Erlenmeyer flasks. Each precursor compound (50-100 μg in MeOH) was then added into the cultures. After an additional 3-d incubation at 30 °C, 200 rpm, 25 mL acetone was added into each flask for soaking extraction at room temperature overnight. After evaporating acetone, the remaining water phase was extracted with EtOAc. The organic layer was then concentrated *in vacuo*. The crude extracts were dissolved in MeOH and then analyzed by HPLC at 230 nm. For *in vitro* enzymatic assays, an equal volume of MeOH was added to quench reactions and precipitate proteins. After centrifugation at the maximum speed for 10 min, supernatants were used as samples for high performance liquid chromatography (HPLC) and high resolution mass spectrometry (HRMS). All samples were analyzed on an Agilent 1260 infinity HPLC system (Agilent Technologies, USA) with a photodiode array detector. Compounds were separated on a Triart C18 column (YMC Co., Ltd., Japan) using a linear mobile phase gradient ranging from 25% (v/v) acetonitrile in 0.1% (v/v) TFA aqueous solution to 50% (v/v) acetonitrile in 0.1% (v/v) TFA aqueous solution over 25 min. The flow rate was set to 1 mL/min and the injection volume was 10 μL. The detection wavelength was set to 230 nm. Structural assignments of metabolites were performed by HRMS, nuclear magnetic resonance (NMR, Bruker), ECD and mono-crystal x-ray diffraction analyses. HRMS data were recorded in the positive ionization mode on an LCQ Deca XP plus ion-trap mass spectrometer (Thermo-Finnigan, San Jose, CA, USA).

**Product quantification**

Products including **BA**, **BB**, **7**, **BX** and **BY** were quantified through calculations of integrated HPLC peak areas (230 nm for **7**, **BX** and **BY**, and 404 nm for **BA** and **BB**). The weighed purified authentic compounds were used as standards. Detailed peak areas used for product quantifications and final calculated ratios are shown in Supplementary Table 4.

**Scale-up biotransformation and purification**

Spores of *Ao* transformants were inoculated into 100 mL of CMP medium containing appropriate

nutrients in 500 mL Erlenmeyer flasks. A certain precursor (10 mg in the minimal volume of DMSO) was then administered to the culture medium. After an additional incubation at 30 °C for 3 d, the culture was treated with acetone (250 mL) at room temperature overnight. After filtration, the acetone in filtrate was evaporated *in vacuo*. The remaining aqueous phase was extracted with EtOAc, and the organic layer was concentrated *in vacuo*. The crude extracts were pre-treated by a C18 open column giving rise to a product mixture. Further purification was conducted by semi-preparative HPLC; compounds were separated on a SymmetryPrep C18 column (7 µm, 7.8 × 300 mm, Waters, USA) using a linear mobile phase gradient ranging from 5% (v/v) acetonitrile/$H_2O$ to 33.5% (v/v) acetonitrile/$H_2O$ over 20 min, and from 33.5% to 50% acetonitrile/$H_2O$ over 16.5 min. The flow rate was set to 3 mL/min.

**ECD measurement for compound 7**

The experiments were performed as previously described[2]. The ECD spectrum of **7** was recorded on a JASCO J-820 spectropolarimeter in $CH_3CN$. ECD calculations were performed at the B3LYP/TZVP level, and the wavelength was corrected (–24 nm) to match the experimental and calculated maxima of the UV spectra.

**Single-crystal x-ray diffraction analysis**

Colorless needles of **BX** and **BY** were grown from an acetonitrile/water solution upon spontaneous evaporation at 4 °C.

A **BY** crystal of dimensions 0.25 × 0.08 × 0.08 mm was mounted on a Rigaku AFC10K Saturn 944+ CCD-based X-ray diffractometer equipped with a low temperature device and Micromax-007HF Cu-target micro-focus rotating anode ($\lambda$ = 1.54187 Å) operated at 1.2 kW power (40 kV, 30 mA). The X-ray intensities were measured at 85(2) K with the detector placed at a distance of 42.00 mm from the crystal. A total of 2028 images were collected with an oscillation width of 1.0° in ω. The exposure times were 1 s for the low angle images, 4 s for high angle. Rigaku d*trek images were exported to CrysAlisPro (v1.171.38.41, Rigaku Oxford Diffraction, 2015) for processing and corrected for absorption. The integration of the data yielded a total of 27067 reflections to a maximum 2θ value of 138.43° of which 3304 were independent and 3283

were greater than 2σ(I). The final cell constants were based on the xyz centroids of 23493 reflections above 10σ(I). Analysis of the data showed negligible decay during data collection. The structure was solved and refined with the Bruker SHELXTL (version 2018/3) software package, using the space group P2(1)2(1)2(1) with Z = 4 for the formula $C_{21}H_{23}N_3O_3$. All non-hydrogen atoms were refined anisotropically with the hydrogen atoms placed in a combination of idealized and refined positions. Full matrix least-squares refinement based on $F^2$ converged at R1 = 0.0274 and wR2 = 0.0715 [based on I > 2σ(I)], R1 = 0.0277 and wR2 = 0.0719 for all data[3]. Additional details are presented in Supplementary Table 3.

The X-ray diffraction analysis of **BX** was conducted in a similar way compared to **BY**. The dimension of **BX** crystal was 0.11 × 0.10 × 0.07 mm. The exposure times were 3 s for the low angle images, 20 s for high angle. The integration of the data yielded a total of 28204 reflections to a maximum 2θ value of 138.50° of which 6550 were independent and 6377 were greater than 2σ(I). The final cell constants were based on the xyz centroids of 22108 reflections above 10σ(I). The structure was solved and refined with the Bruker SHELXTL (version 2018/3) software package, using the space group P2(1) with Z = 2 for the formula $2(C_{21}H_{24}N_3O_3)(H_2O)$. Full matrix least-squares refinement based on $F^2$ converged at R1 = 0.0337 and wR2 = 0.0872 [based on I > 2σ(I)], R1 = 0.0351 and wR2 = 0.0901 for all data[3]. Additional details are presented in Supplementary Table 3.

Colorless needles of compound **15** were grown from an acetonitrile/water solution upon spontaneous evaporation at 16 °C. A crystal with dimensions of 0.984 mm × 0.304 mm × 0.048 mm was chosen for structure determination. Data were obtained at 99.99(10) K with a Rigaku Oxford XtaLAB Synergy-DW diffractometer with monochromated Mo Kα radiation ($\lambda$ = 1.54187 Å). The crystal-to-detector distances were set as 50 mm. The CrysAlisPro software was used for data reduction and integration. The structures were established by direct methods and refined through full-matrix least-squares fitting on $F^2$ using OLEX2. All atoms were refined using full matrix least-squares techniques, and final least-squares refinement was on $F_o^2$ with data having $F_o^2 \geq 2\sigma(F_o^2)$. Numerical absorption corrections were carried out using the SCALE program for

the area detector. The structures were solved with the use of Shel-XT to determine the atomic coordinates of the cations.

The CIF entries for structures were modified to "_atom_sites_solution_primary" "dual-space" to reflect the method of phase determination. Hooft parameters were calculated with PLATON (v1.19).

**Chemical properties of key products**

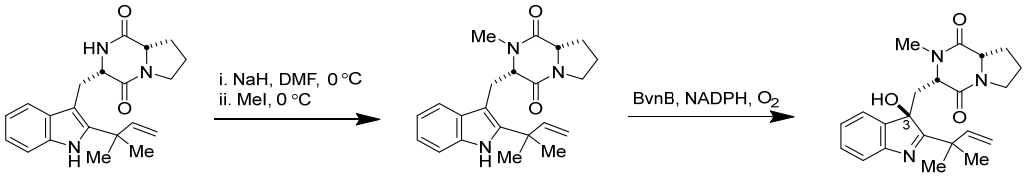
**BA**: yellow powder, $[\alpha]^{26}_{D}$ +226.0426 (*c* 0.16, MeOH), HRMS *m/z* 366.1815 $[M + H]^+$ (*calc.* 366.1812) (Supplementary Fig. 4). $^1H$ and $^{13}C$ NMR data, see Supplementary Table 1; NMR spectrum, see Supplementary Figs. 2-3.

Compound **7**: colorless powder, $[\alpha]^{26}_{D}$ -274.1998 (*c* 0.47, MeOH), HRMS *m/z* 366.1813 $[M + H]^+$ (*calc.* 366.1812) (Supplementary Fig. 4). $^1H$ and $^{13}C$ NMR data, see Supplementary Table 2; NMR spectrum, see Supplementary Figs. 5-10.

**BX**: white powder, $[\alpha]^{25}_{D}$ +92.2667 (*c* 0.13, MeOH), HRMS *m/z* 366.1810 $[M + H]^+$ (*calc.* for 366.1812) (Supplementary Fig. 4). $^1H$ and $^{13}C$ NMR data, see Supplementary Table 1; NMR spectrum, see Supplementary Figs. 11-16.

**BY**: white powder, $[\alpha]^{26}_{D}$ +27.1053 (*saturated**, MeOH), $[\alpha]^{25}_{D}$ +255.1634 (*c* 0.37, DMSO), HRMS *m/z* 366.1813 $[M + H]^+$ (*calc.* 366.1812) (Supplementary Fig. 4). $^1H$ and $^{13}C$ NMR data, see Supplementary Table 1; NMR spectrum, see Supplementary Figs. 17-22. *literature[4]: $[\alpha]^{25}_{D}$ +11.5 (*c* 0.2, MeOH), at which concentration **BY** could not be dissolved in our case, so here the MeOH saturated solution was used for measurement.

**Chemoenzymatic synthesis of compound 14 and 15**

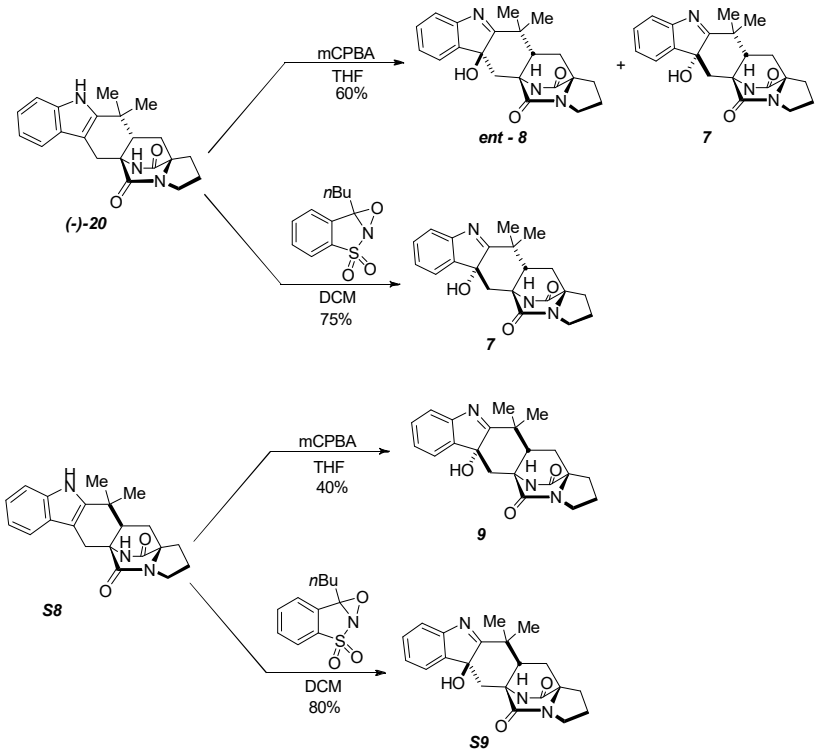



Compound **14** was chemically synthesized according to the method described by Li[5] with some modifications. NaH (60%, 24 mg, 0.6 mmol) was added to a solution of **DE** (140 mg, 0.4 mmol) in DMF (4 mL) at 0 ºC and the reaction was stirred for 30 min at 0 ºC. MeI (30 µL, 68 mg, 0.48 mmol) was added and the reaction was stirred for 4 h at 0 ºC. The reaction was partitioned between $CH_2Cl_2$ (20 mL) and saturated aqueous $NH_4Cl$ (20 mL). The organic phase was collected, and the aqueous phase was extracted with 20 mL $CH_2Cl_2$ twice. The combined organic phases were then concentrated under reduced pressure. The residue was re-suspended with 4 mL DMSO and purified through preparative C18 reversed-phase HPLC using a linear mobile phase gradient ranging from 30% (v/v) acetonitrile in 0.1% (v/v) TFA aqueous solution to 70% (v/v) acetonitrile in 0.1% (v/v) TFA aqueous solution over 30 min. The flow rate was set to 10 mL/min and the injection volume was 1 mL. Detection wavelength was set to 230 nm. Structure of synthetic **14** was confirmed by NMR (Supplementary Figs. 36-37). $^1$H NMR (600 MHz, $CDCl_3$), δ7.99 (s, 1 H), 7.50 (d, *J* = 7.8 Hz, 1 H), 7.29 (d, *J* = 8.4 Hz, 1 H), 7.16 (t, *J* = 7.2 Hz, 1 H), 7.10 (t, *J* = 7.8, 1 H), 6.13 (dd, *J* = 17.4, 10.8 Hz, 1 H), 5.17 (dd, *J* =21.0, 17.4 Hz, 2 H), 4.56 (dd, *J* = 9.0, 3.6 Hz, 1 H), 4.04 (dd, *J* = 11.4, 6.0 Hz, 1 H), 3.87 (dt, *J* = 12.0, 7.8 Hz, 1 H), 3.72 (dd, *J* = 15.0, 3.6 Hz, 1 H), 3.34 (dd, *J* = 15.6, 9.0 Hz, 1 H), 3.30 (m, 1H), 2.71 (s, 3 H), 2.35 (m, 1 H), 1.86 (m, 1 H),1.55 (overlap, 1 H), 1.55 (s, 6 H); $^{13}$C NMR (600 MHz, $CDCl_3$): δ 166.8, 165.8, 146.0, 140.6, 134.3, 129.1, 122.1, 120.0, 118.6, 112.3, 110.8, 106.1, 63.1, 59.7, 45.1, 39.4, 34.3, 31.6, 29.3, 27.9, 27.8, 22.0.

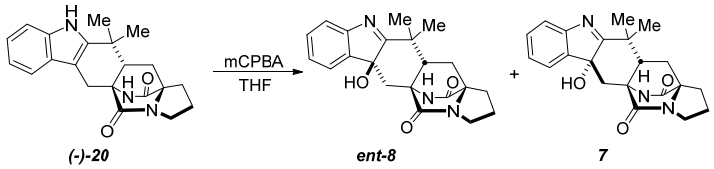
Compound **15** was obtained through large scale enzymatic conversion with BvnE as described above except that the reaction was conducted in a 20 mL scale. Its structure was confirmed by UV (Supplementary Fig. 24), HRMS (Supplementary Fig. 35), NMR (Supplementary Table 5, Supplementary Figs. 38-43) spectra and single-crystal X-ray diffraction (Extended Data Fig. 4).

**Chemical synthesis of pathway metabolites: hydroxyindolenines**

**Overview scheme:**



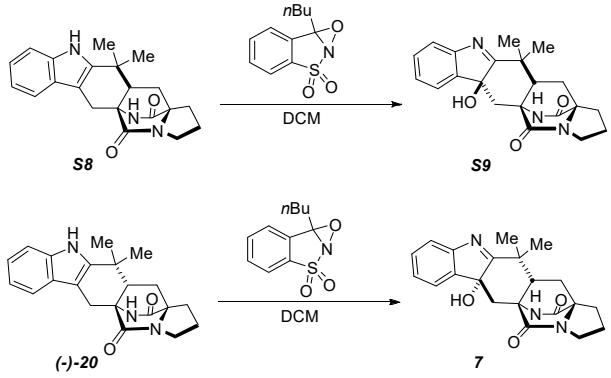

**Chemical synthesis of compound *ent*-8 and 7**




To a solution of compound (-)-**20** (30.0 mg, 087 mmols) in THF (3 mL) at rt was added mCPBA (19.2 mg, 0.111 mmols). The reaction was stirred vigorously for 2 hours or until TLC showed completion (6% MeOH/$CDCl_3$). The reaction was quenched with $Me_2S$ and concentrated under reduced pressure. The diastereomers were separated and purified by preparative TLC (6% MeOH/$CDCl_3$) to yield pale yellow solids (70%). *ent*-**8:** $^1$H NMR (400 MHz, Chloroform-*d*) δ 7.53 – 7.43 (m, 2H), 7.33 (td, *J* = 7.6, 1.3 Hz, 1H), 7.28 (d, *J* = 2.5 Hz, 1H), 7.21 (td, *J* = 7.4, 1.1 Hz, 1H), 3.69 (d, *J* = 15.3 Hz, 1H), 3.39 – 3.28 (m, 2H), 3.26 (ddd, *J* = 11.3, 7.2, 6.0 Hz, 1H), 2.64 (dt, *J* = 13.4, 6.8 Hz, 1H), 2.07 (dd, *J* = 13.2, 10.2 Hz, 1H), 2.02 – 1.72 (m, 5H), 1.46 (d, *J* = 15.3 Hz, 1H), 1.35 (s, 3H), 1.30 (s, 3H).$^{13}$C NMR (101 MHz, Chloroform-*d*) δ 189.78, 173.62, 170.55, 153.56, 139.89, 129.84, 126.48, 122.55, 120.69, 82.00, 67.14, 60.27, 44.38, 43.32, 39.03, 34.48,

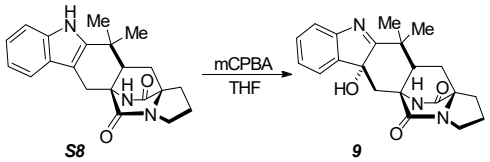
30.70, 29.03, 28.32, 24.65, 23.25. **7:** $^{1}$H NMR (400 MHz, Chloroform-*d*) δ 7.53 (dt, *J* = 7.7, 0.8 Hz, 1H), 7.42 (ddd, *J* = 7.3, 1.3, 0.6 Hz, 1H), 7.36 (td, *J* = 7.6, 1.3 Hz, 1H), 7.23 (td, *J* = 7.5, 1.0 Hz, 1H), 7.13 (s, 1H), 3.67 – 3.62 (m, 1H), 3.41 (t, *J* = 6.8 Hz, 2H), 2.81 (d, *J* = 15.7 Hz, 1H), 2.74 (dt, *J* = 13.4, 6.8 Hz, 1H), 2.12 – 1.92 (m, 6H), 1.83 (dt, *J* = 13.1, 7.2 Hz, 1H), 1.44 (s, 3H), 1.32 (s, 3H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 188.21, 172.55, 168.36, 152.11, 140.58, 130.30, 126.71, 122.41, 121.24, 82.50, 67.29, 62.10, 50.49, 44.21, 40.54, 38.04, 32.68, 29.15, 27.45, 24.51, 20.17.

**Chemical synthesis of compounds 7 and S9**




Davis Oxaziridine (66 mg, 0.28 mmols) was added at rt to a solution of cycloadduct **S8** (32 mg, 0.092 mmols) in DCM (4 mL). The reaction mixture was stirred for 18 hrs, then concentrated. The crude material was placed on a column and eluted with 5 % MeOH/DCM. The material was further purified by preparative TLC using 6% MeOH/DCM to yield hydroxyindolenine **S9** (75%). Compound **7** was synthesized in the same manner from *(-)*-**20**. Data for compound **7** is given above alongside data for compound ***ent*-8**. **S9:** $^{1}$H NMR (400 MHz, Chloroform-*d*) δ 7.54 (dt, *J* = 7.7, 0.9 Hz, 1H), 7.47 (ddd, *J* = 7.4, 1.3, 0.6 Hz, 1H), 7.39 (td, *J* = 7.7, 1.3 Hz, 1H), 7.32 – 7.24 (m, 1H), 6.91 (s, 1H), 3.86 (dd, *J* = 10.2, 6.9 Hz, 1H), 3.55 – 3.47 (m, 2H), 3.27 (d, *J* = 16.8 Hz, 1H), 2.81 (ddd, *J* = 13.0, 7.1, 5.8 Hz, 1H), 2.60 (d, *J* = 16.8 Hz, 1H), 2.27 (dd, *J* = 13.1, 10.2 Hz, 1H), 2.11 – 1.95 (m, 2H), 1.95 – 1.83 (m, 2H), 1.46 (s, 3H), 1.17 (s, 3H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 190.04, 173.79, 168.04, 153.39, 139.63, 130.24, 126.59, 122.25, 120.95, 83.23, 67.10, 60.23, 48.37, 44.54, 38.67, 34.20, 31.50, 29.55, 28.75, 24.61, 21.15.

**Chemical synthesis of compound 9**

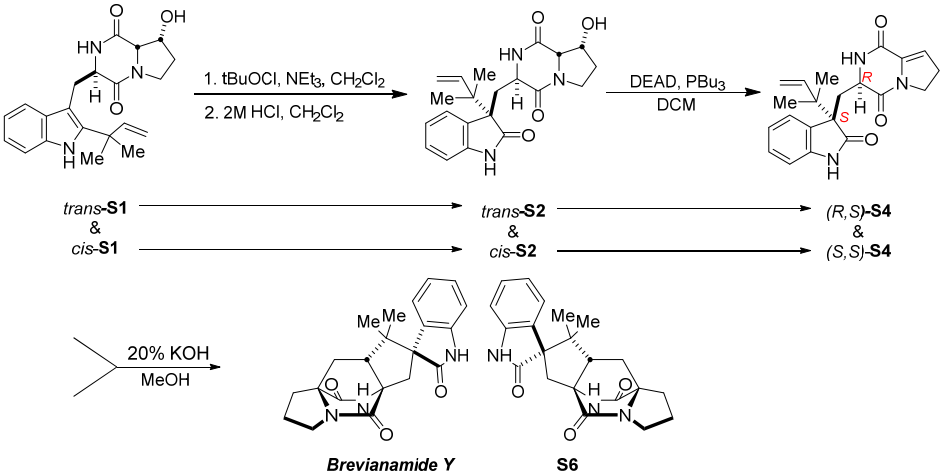



*S8* *9*

To a solution of compound **S8** (20.0 mg, 0435 mmols) in THF (2 mL) at rt was added mCPBA (9.6 mg, 0.056 mmols). The reaction was stirred vigorously for 4 hours or until TLC showed completion (6% MeOH/$CDCl_3$), quenched with $Me_2S$, and concentrated under reduced pressure. The diastereomers were separated and purified by preparative TLC (6% MeOH/$CDCl_3$) to yield a pale yellow solid (40%). **9:** $^1$H NMR (400 MHz, Chloroform-*d*) δ 7.48 (d, *J* = 7.6 Hz, 1H), 7.44 – 7.31 (m, 2H), 7.21 (t, *J* = 7.5 Hz, 1H), 6.77 (s, 1H), 3.73 (dd, *J* = 10.0, 7.2 Hz, 1H), 3.56 – 3.39 (m, 2H), 2.90 (d, *J* = 15.9 Hz, 1H), 2.82 – 2.71 (m, 1H), 2.34 (d, *J* = 16.0 Hz, 1H), 2.18 (dd, *J* = 12.9, 10.1 Hz, 1H), 2.01 (tp, *J* = 12.9, 6.7, 6.1 Hz, 2H), 1.91 – 1.79 (m, 2H), 1.34 (s, 3H), 1.14 (s, 3H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 189.81, 173.51, 167.85, 153.35, 139.41, 130.19, 126.46, 122.10, 120.89, 83.12, 66.99, 60.10, 48.30, 44.41, 38.52, 34.15, 31.39, 29.43, 28.65, 24.48, 21.03.

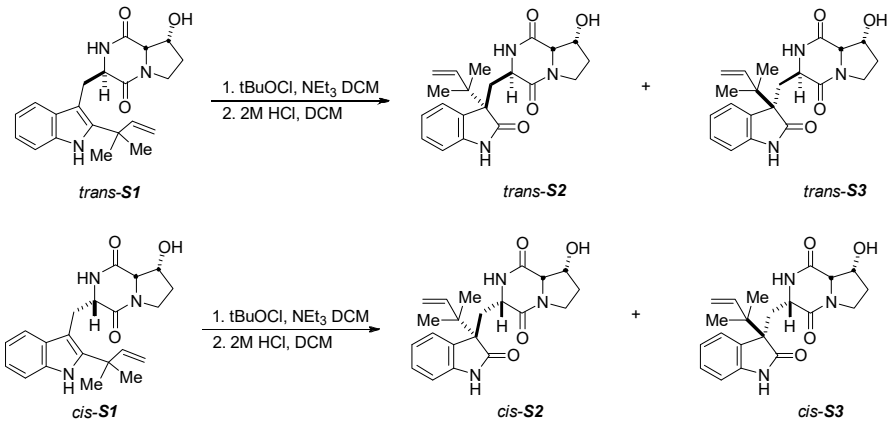
**Chemical synthesis of compound BY**

**Overview scheme:**




***Brevianamide Y*** **S6**

**Chemical synthesis of compounds S2 and S3** (*the *cis*- and *trans*-diastereomeric nomenclature refers to the relative stereochemistry of the proline- and tryptophan-derived α-methine protons)




tBuOCl (227 µL, 2.18 mmols) was added to a solution of each individual diastereomer of *trans*-**S1** and *cis*-**S1** (400 mg, 1.09 mmols) and $NEt_3$ (334 µL, 2.40 mmols) in DCM (20 mL) at 0ºC and stirred for 3 hours. The reaction mixture then was immediately concentrated before being taken up in 2M HCl (40 mL) and DCM (40 mL), and left to stir for two days. The reaction mixture was neutralized with $NH_4Cl$, extracted with DCM (×2), washed with brine, dried over $Na_2SO_4$, filtered, and concentrated. The crude yellow solid was purified by flash column chromatography eluting with 3-6% MeOH/DCM to yield pale yellow solids. Each diastereomer was run separately and their resultant diastereomers were separated as mentioned above*: trans*-**S2** to *trans*-**S3** (70%; 3:1) and *cis*-**S2** to *cis*-**S3** (60%; 1:1).

**From *trans*-S1: *trans*-S2 (*top* diastereomer):** $^1$H NMR (400 MHz, Chloroform-*d*) δ 8.67 (s, 1H), 7.30-7.23 (m, 2H), 7.04 (t, *J* = 7.6 Hz, 1H), 6.90 (d, *J* = 7.7 Hz, 1H), 6.09 (dd, *J* = 17.4, 10.8 Hz, 1H), 6.01 (s, 1H), 5.13 (d, *J* = 10.8 Hz, 1H), 5.03 (d, *J* = 17.4 Hz, 1H), 4.52 – 4.42 (m, 1H), 4.05 (s, 1H), 3.90 (t, *J* = 2.6 Hz, 1H), 3.55 – 3.42 (m, 3H), 3.01 (dd, *J* = 15.0, 4.3 Hz, 1H), 2.60 – 2.50 (m, 2H), 2.01 – 1.81 (m, 2H), 1.73 (d, *J* = 14.9 Hz, 1H), 1.14 (s, 3H), 1.05 (s, 3H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 180.21, 167.70, 164.87, 142.57, 129.52, 129.23, 126.89, 122.06, 114.74, 110.20, 70.72, 63.91, 56.41, 53.94, 44.13, 43.04, 31.51, 30.24, 29.86, 22.43, 21.60.

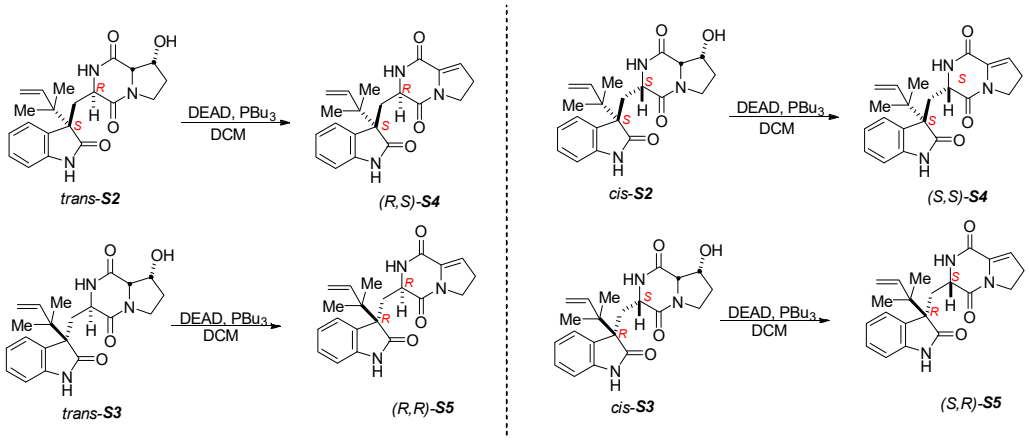
***trans*-S3 (*bottom* diastereomer):** $^1$H NMR (400 MHz, Chloroform-*d*) δ 9.29 (s, 1H), 7.29 (s, 1H), 7.23 (t, *J* = 7.7 Hz, 1H), 7.01 (t, *J* = 7.6 Hz, 1H), 6.90 (d, *J* = 7.8 Hz, 1H), 6.33 (s, 1H), 6.10 (q, *J* = 17.4, 10.8 Hz, 1H), 5.12 (d, *J* = 10.9 Hz, 1H), 5.01 (d, *J* = 17.4 Hz, 1H), 4.46 (s, 1H), 4.07 (s, 1H), 3.90 (s, 1H), 3.50 – 3.45 (m, 2H), 2.97 (dd, *J* = 15.0, 4.7 Hz, 1H), 2.68-2.60 (m, 2H), 1.90 (m,

2H), 1.13 (s, 3H), 1.04 (s, 3H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 181.77, 167.86, 165.22, 143.08, 142.27, 129.43, 129.24, 127.78, 121.81, 114.81, 110.76, 70.94, 64.37, 56.78, 54.35, 44.29, 43.30, 31.99, 30.63, 22.74, 21.93.

**From *cis*-S1: *cis*-S2 (*top* diastereomer):** $^{1}$H NMR (400 MHz, Chloroform-*d*) δ 10.80 (s, 1H), 8.48 (s, 1H), 7.23 (t, *J* = 7.5 Hz, 2H), 7.04 (t, *J* = 7.5 Hz, 1H), 6.90 (d, *J* = 7.8 Hz, 1H), 6.06 (q, *J* = 17.4, 10.8 Hz, 1H), 5.14 (d, *J* = 10.8 Hz, 1H), 5.03 (d, *J* = 17.4 Hz, 1H), 4.54 (t, *J* = 3.7 Hz, 1H), 4.21 (s, 1H), 3.70 – 3.60 (m, 2H), 3.42 (m, 1H), 2.82 (s, 1H), 2.54 – 2.43 (m, 1H), 2.40 (dd, *J* = 14.1, 3.7 Hz, 1H), 2.17 – 2.07 (m, 1H), 2.06 – 1.93 (m, 1H), 1.15 (s, 3H), 1.05 (s, 3H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 183.85, 170.32, 167.33, 143.01, 142.68, 128.44, 128.07, 126.59, 122.30, 114.97, 112.17, 71.17, 63.52, 56.92, 56.16, 44.66, 42.84, 34.28, 30.40, 22.39, 22.19.

***cis*-S3 (*bottom* diastereomer):** $^{1}$H NMR (400 MHz, Chloroform-*d*) δ 10.70 (s, 1H), 8.73 (s, 1H), 7.19 (m, 2H), 6.94 (t, *J* = 7.2 Hz, 2H), 6.02 (q, *J* = 17.3, 10.8 Hz, 1H), 5.10 (d, *J* = 10.8 Hz, 1H), 4.97 (d, *J* = 17.4 Hz, 1H), 4.27 (t, *J* = 4.0 Hz, 1H), 4.21 (d, *J* = 7.4 Hz, 1H), 3.70 (m, 1H), 3.12 (t, *J* = 11.1, 10.6 Hz, 1H), 2.93 (d, *J* = 14.9 Hz, 1H), 2.76 (dd, *J* = 15.0, 7.5 Hz, 1H), 2.38 (s, 1H), 1.90 – 1.68 (comp, 3H), 1.11 (s, 3H), 0.97 (s, 3H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 183.02, 167.03, 165.33, 143.53, 142.61, 128.65, 128.63, 128.34, 120.52, 114.52, 110.38, 71.08, 62.27, 56.30, 55.42, 43.27, 42.89, 33.26, 29.41, 22.02, 21.62.

**Chemical synthesis of compounds S4 and S5**

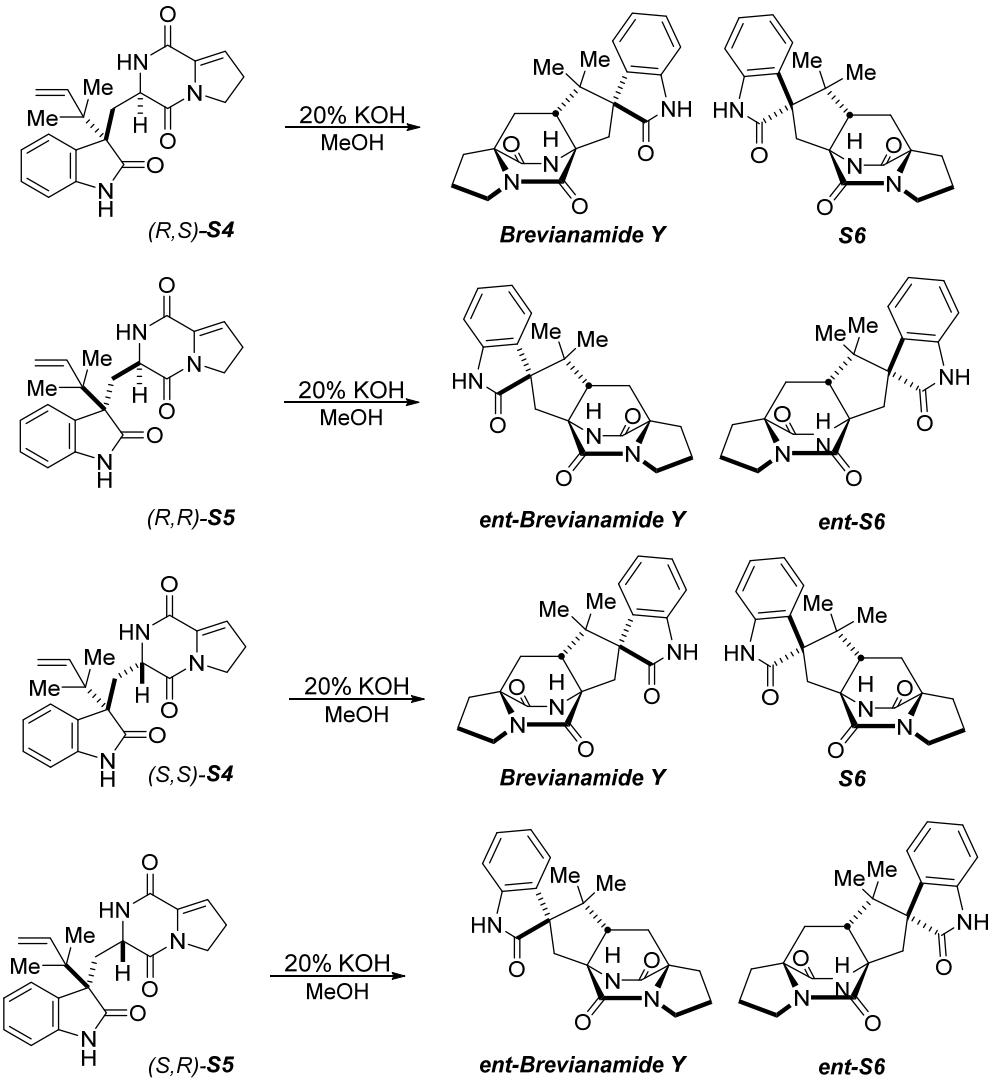



DEAD (527 µL, 1.16 mmols) was added to a solution of the individual diastereomers *trans*-**S2** or

***cis*-S3**. **From *trans*-S2:** (148 mg, 0.386 mmols) in DCM (7.80 mL) at room temperature. After running for 15 minutes, $PBu_3$ (290 μL, 1.16 mmols) was added. The reaction was run until starting material was consumed (~ 3 hrs) and concentrated. The crude material was purified by flash column chromatography eluting with 3-10% MeOH/DCM to yield a pale yellow solid (70%). Each diastereomer was carried forward separately. *(R,S)*-**S4:** $^1$H NMR (400 MHz, Chloroform-*d*) δ 9.72 (s, 1H), 7.28-7.26 (m, *J* = 7.3 Hz, 2H), 7.10 – 7.02 (m, 2H), 6.97 (d, *J* = 8.2 Hz, 1H), 6.04 (dd, *J* = 17.4, 10.8 Hz, 1H), 5.98 (t, *J* = 3.0 Hz, 1H), 5.15 (t, *J* = 10.6 Hz, 1H), 5.02 (d, *J* = 17.4 Hz, 1H), 4.02 (m, 1H), 3.85 (m, 1H), 3.59 – 3.51 (m, 1H), 2.86 – 2.64 (comp, 2H), 2.41 (dd, *J* = 14.1, 11.1 Hz, 1H), 1.71-1.49 (m, 1H), 1.15 (s, 3H), 1.05 (s, 3H).$^{13}$C NMR (101 MHz, Chloroform-*d*) δ 178.29, 164.61, 158.39, 142.62, 140.75, 129.12, 124.82, 123.94, 121.34, 115.52, 80.89, 68.32, 67.99, 65.48, 63.71, 43.82, 43.08, 28.12, 22.76, 21.78, 21.40. Diastereomer *cis*-**S3** provides the enantiomeric product (*S,R*)-**S5** with identical $^1$H NMR and $^{13}$C NMR spectra.

**From *trans*-S3 or cis-S2: *(R,R)*-S5 (from trans-S3):** $^1$H NMR (400 MHz, Chloroform-*d*) δ 10.85 (s, 1H), 8.67 (s, 1H), 7.18 (d, *J* = 7.5 Hz, 1H), 7.12 (t, *J* = 7.6 Hz, 1H), 6.94 – 6.81 (m, 2H), 6.08 (dd, *J* = 17.4, 10.8 Hz, 1H), 5.55 (t, *J* = 3.0 Hz, 1H), 5.10 (d, *J* = 10.8 Hz, 1H), 4.97 (d, *J* = 17.4 Hz, 1H), 4.31 (d, *J* = 7.2 Hz, 1H), 3.71 (t, *J* = 9.1 Hz, 2H), 3.03 (d, *J* = 14.7 Hz, 1H), 2.75 (dd, *J* = 14.8, 7.4 Hz, 1H), 1.48 – 1.37 (m, 1H), 1.32 – 1.23 (m, 1H), 1.11 (s, 3H), 0.97 (s, 3H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 183.12, 162.60, 157.60, 143.04, 132.11, 128.37, 128.05, 120.49, 118.28, 114.18, 111.09, 56.26, 55.79, 45.25, 42.62, 33.73, 27.51, 24.52, 24.38, 22.16, 21.61. Diastereomer *cis*-**S2** provides the enantiomeric product (*S,S*)-**S4** with identical $^1$H NMR and $^{13}$C NMR spectra.

**Chemical synthesis of compounds BY, *ent*-BY, S6, and *ent*-S6**



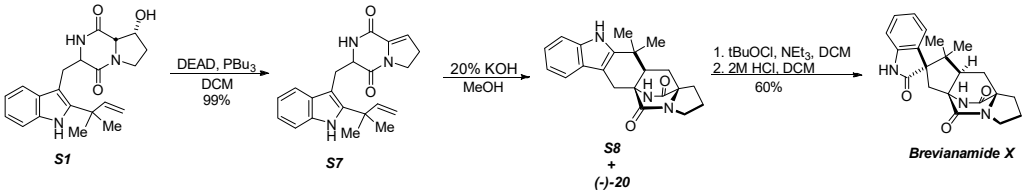

A 20% aqueous solution of KOH (7.53 mL) was added to a solution of *(R,S)*-**S4**, (*S,S*)-**S4** or *(R,R)*-**S5**, (*S,R*)-**S5** (110 mg, 0.30 mmols; each diastereomer was carried forward separately) in MeOH (25 mL) at 0 ºC, and slowly warmed to room temperature over the course of an hour. The reactions were run for 8 hours. The reaction was quenched with $NH_4Cl$ and extracted with DCM (×3). The combined organic layers were dried over $Na_2SO_4$, filtered, and concentrated. The crude oil was purified by flash column chromatography with 3-6% MeOH/DCM to yield a mixture of diastereomers (1:0.85, 80 mg, 73%). A small amount of each diastereomer was isolated by trituration. **BY** was isolated by trituration with chloroform as a white solid. The other diastereomer, **S6**, was partially isolated by trituration with acetone as a white solid. Data collected for **BY** matched previously published data.

**BY and *ent*-BY:** $^{1}$H NMR (400 MHz, DMSO-$d_6$) δ 10.30 (s, 1H), 8.79 (s, 1H), 7.42 (dd, $J$ = 7.6, 1.2 Hz, 1H), 7.19 (td, $J$ = 7.7, 1.2 Hz, 1H), 6.99 (td, $J$ = 7.6, 1.1 Hz, 1H), 6.81 (dd, $J$ = 7.8, 1.1 Hz, 1H), 3.28 (s, 2H), 3.17 (dd, $J$ = 10.4, 7.2 Hz, 1H), 2.82 (d, $J$ = 15.1 Hz, 1H), 2.48 – 2.43 (m, 1H), 2.13 (d, $J$ = 15.1 Hz, 1H), 2.05 – 1.88 (m, 2H), 1.87 – 1.74 (m, 2H), 1.65 (dd, $J$ = 13.0, 7.2 Hz, 1H), 0.99 (s, 3H), 0.69 (s, 3H). $^{13}$C NMR (101 MHz, DMSO-$d_6$) δ 181.87, 172.53, 169.08, 142.42, 129.78, 128.15, 126.31, 120.85, 109.06, 68.59, 67.18, 62.11, 50.05, 46.86, 43.26, 33.63, 28.45, 27.97, 24.47, 23.00, 20.42.

**S6 and *ent*-S6:** $^{1}$H NMR (400 MHz, DMSO-$d_6$) δ 10.30 (s, 1H), 8.71 (s, 1H), 7.18 (td, $J$ = 7.7, 1.2 Hz, 1H), 7.07 (d, $J$ = 7.2 Hz, 1H), 6.94 (td, $J$ = 7.5, 1.1 Hz, 1H), 6.83 (dd, $J$ = 7.8, 1.1 Hz, 1H), 4.16 – 4.04 (m, 1H), 3.18 (d, $J$ = 5.2 Hz, 1H), 2.68 – 2.55 (m, 2H), 2.51 – 2.42 (m, 2H), 2.08 – 1.89 (m, 2H), 1.92 – 1.73 (m, 2H), 1.66 (dd, $J$ = 13.0, 7.4 Hz, 1H), 1.06 (s, 3H), 0.42 (s, 3H). $^{13}$C NMR (126 MHz, DMSO-d$_6$) δ 178.68 172.32, 169.50, 141.52, 134.34, 127.88, 124.04, 121.07, 109.15, 68.53, 67.54, 62.44, 53.07, 47.79, 43.38, 33.12, 28.43, 27.31, 24.43, 23.18, 19.51.

**Chemical synthesis of compound BX**

**Overview scheme:**




tBuOCl (129.4 µL, 1.14 mmols) was added to a solution of **S8** (100 mg, 0.286 mmols) and $NEt_3$ (176 µL, 1.26 mmols) in DCM (5.10 mL) at 0 ºC and stirred for 3 hours. The reaction mixture was immediately concentrated before being retaken up in 2M HCl (10.5 mL) and DCM (10.5 mL), and left to stir for two days. The reaction mixture was neutralized with $NH_4Cl$, extracted with DCM (×2), washed with brine, dried over $Na_2SO_4$, filtered, and concentrated. The crude yellow solid was purified by flash column chromatography eluting with 3-6% MeOH/DCM to yield **BX**.

**BX:** $^{1}$H NMR (400 MHz, Chloroform-$d$) δ 10.34 (s, 1H), 9.10 (s, 1H), 7.26 – 7.14 (m, 2H), 6.97 (t, $J$ = 7.5 Hz, 1H), 6.82 (d, $J$ = 7.6 Hz, 1H), 3.46 – 3.34 (m, 2H), 3.27 – 3.18 (m, 1H), 2.85 (d, $J$ = 14.2

Hz, 1H), 2.57 – 2.47 (m, 2H), 2.19 (d, *J* = 14.2 Hz, 1H), 2.07 – 1.89 (m, 2H), 1.88 – 1.70 (m, 2H), 0.74 (s, 3H), 0.71 (s, 3H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 182.30, 173.06, 169.38, 142.31, 130.56, 128.05, 125.97, 121.06, 109.17, 68.06, 65.63, 61.41, 55.42, 45.09, 43.38, 33.12, 29.41, 28.99, 24.33, 23.25, 19.71.

*Note*: NMR spectra of synthetic hydroxyindolenines, **BY**, **BX** and related intermediates are shown in Supplementary Figs. 66-91.

**Computational methods**

The range-separated dispersion-corrected M06-2X density functional[6] and the 6-31+G(d,p)[7-11] basis set were used with Grimme's D3 correction using the original D3 damping function[12] to optimize the geometries of all stationary points in gas phase. Previous studies have demonstrated that the combination of M06-2X with D3 dispersion is robust, reliable and accurate[13]. To these optimized structures, single-point energy corrections were applied. The single-point energy calculations were obtained using correlated wavefunction theory (WFT) (DLPNO-CCSD(T)/cc-pV(DT)Z)[14-21] and the integral equation formalism variant of the polarizable continuum model (IEF-PCM)[22-26] with the SMD solvation model (solvent=water)[27] was used to account for solvent effects. For the bond dissociation energy (BDE) calculations, no solvent was included.

Domain-based local pair-natural orbital coupled cluster with perturbative triple excitations (DLPNO-CCSD(T)) calculations were carried out with *ORCA 4.0*[28,29]. DLPNO-CCSD(T) energies with a normal truncation threshold ("normalPNO") show an accuracy of 1 kcal mol$^{-1}$ or better compared to CCSD(T)[30], which is generally recognized as a "gold standard" in computational chemistry[31]. Extrapolation to the basis set limit was carried out employing cc-pVDZ and cc-pVTZ energies, treating the convergence of SCF and correlation energies separately:

(a) The convergence of the HF energy to the basis set limit is calculated as:

$$E_{SCF}^{(X)} = E_{SCF}^{(\infty)} + Ae^{(-\alpha\sqrt{X})} \qquad (1)$$

where $E_{SCF}^{X}$ is the SCF energy calculated with the basis set having highest angular momentum X, $E_{SCF}^{\infty}$ is the basis set limit SCF energy, α = 4.42 (empirically optimized for cc-pV(DT)Z), and A is a parameter to be determined.

(b) The correlation energy is assumed to converge as:

$$E_{corr}^{(\infty)} = \frac{X^{\beta} E_{corr}^{(X)} - Y^{\beta} E_{corr}^{(Y)}}{X^{\beta} - Y^{\beta}} \qquad (2)$$

where $E_{corr}^{(\infty)}$ is the correlation energy calculated with the basis sets having successive highest angular momentums X and Y (in our case 3 and 2), and β = 2.46 (empirically optimized for cc-pV(DT)Z)[32].

A python script was used in the conformational sampling (carrying out bond rotations) of ground state structures. For transition states, manual conformational searches were performed. In this process, the different parts of the molecules were rotated and six- and five-membered rings were represented with diverse conformations. Also, the amide and imidate tautomeric forms of dioxopiperazine rings were taken into account. During the conformational search, we found a great number of conformers for each reaction step (see attached *Thermochemistry.dat* file for more information).

We performed vibrational frequency calculations to verify that stationary points were either minima or first-order saddle points on the potential energy surface, and to calculate thermal corrections to Gibbs free energies at 298.15 K (25˚C). Moreover, intrinsic reaction coordinate (IRC) calculations[33] were performed to ensure that the intermediates (**Int**) connected to their corresponding transition structure (**TS**). The computed thermochemistry data were corrected following Grimme's quasi-harmonic (QHA) model for entropy[34] with a frequency cut-off value of 100.0 $cm^{-1}$ using the *GoodVibes* program[35]. Also, *GoodVibes* applied a 1 M concentration correction to account for reactions in solution and a vibrational scaling factor of 0.967 for M062X/6-31+G(d,p) level of theory as previously suggested[36].

*Gaussian 16*[37] was employed for all density functional theory (DFT) calculations, using an "ultrafine" pruned (99,590) grid for numerical integration of the exchange-correlation functional

and its derivatives. Natural atomic charges were computed using natural population analysis (NPA) with *NBO 6.0*[38], interfaced to *Gaussian 16*. Our display settings have been made openly accessible (https://gist.github.com/bobbypaton).

Configurational/conformational analysis was performed for all ground and transition states as explained above. To enumerate the different conformers, we used different letters at the end of each name (*i.e.* a, b, c, *etc.*). Also, to differentiate the different media used in the calculations, the structures with one $H^+$ equivalent contain a "-H" suffix after the name of their corresponding reaction step (*i.e.* **Int-I-H**, **TS-I-H**, *etc.*).

Representations in the main text and supporting information refer to the most stable rotameric conformation found for each step. G values of all the energy profiles correspond to the Boltzmann weighted G of all the conformers found in each step ($G_{av}$) (see section *Thermochemical data and absolute energy values*). The different reaction steps that were not included within the main text were named differently following (*i*) the order in which they appear along the pathways, (*ii*) the pathway containing the steps and (*iii*) their protonation state.

Boltzmann weighted G ($G_{av}$) were calculated with *GoodVibes* as:

$$G_{av} = \sum_i G_i \times p_i \qquad (3)$$

where $G_i$ is the relative Gibbs free energy of the corresponding conformers of a certain reaction step and pi is the probability of each conformer calculated as:

$$p_i = \frac{e^{\frac{-G_i}{RT}}}{\sum_i \left(e^{\frac{-G_i}{RT}}\right)} \qquad (4)$$

Small persistent imaginary frequencies (< 50*i* $cm^{-1}$) were inverted to their respective positive values before the QHA entropy and enthalpy corrections were computed as seen in previous examples[39].

In each of the different types of mechanisms studied, its corresponding most stable reaction step was used as the reference with $G_{av}$ = 0 kcal/mol. The thermochemical data of the most stable conformers of all the systems studied including absolute energies, zero-point energies (ZPE) and T·S, among

other parameters, at the M06-2X-D3/6-31+G(d,p) level, as well as the absolute energies, corrected final G and relative G obtained with DLPNO-CCSD(T)/cc-pV(DT)Z (SMD=water), were generated in an automated manner using *GoodVibes* and tabulated in a separate file of the ESI called *Thermochemistry.dat*. This automated process for creating G profiles provides a powerful method to avoid errors related to human manipulation of the data. The command lines used by *GoodVibes* are included in the same file. Additionally, we tested that no input mistakes were made in the calculations by using the --check function of *GoodVibes*.

The xyz files containing all the geometries studied are provided separately along with the ESI. The generation of xyz files was automated with *GoodVibes* using the --xyz option.