Supporting Information

Fungal Indole Alkaloid Biogenesis Through Evolution of a Bifunctional Reductase/Diels-Alderase

Qingyun Dan,[1,2,#] Sean A. Newmister,[1,#] Kimberly R. Klas,[3] Amy E. Fraley,[1,4] Timothy J. McAfoos,[3] Amber D. Somoza,[3] James D. Sunderhaus,[3] Ying Ye,[1] Vikram V. Shende,[1,5] Fengan Yu,[1] Jacob N. Sanders,[6] W. Clay Brown,[1] Le Zhao,[3] Robert S. Paton,[3] K. N. Houk,[6] Janet L. Smith,[1,2] David H. Sherman,[1,4,7,8,*] and Robert M. Williams[3,9,*]

[1]Life Sciences Institute, University of Michigan, Ann Arbor, Michigan 48109, USA; [2]Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA; [3]Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, USA; [4]Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA; [5]Program in Chemical Biology, University of Michigan, Ann Arbor, Michigan 48109, USA; [6]Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, USA; [7]Department of Microbiology & Immunology, University of Michigan, Ann Arbor, Michigan 48109; [8]Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109; and [9]University of Colorado Cancer Center, Aurora, Colorado 80045, USA.

[#]These authors contributed equally to this work.

*Robert M. Williams: robert.williams@colostate.edu; David H. Sherman: davidhs@umich.edu.

Online Methods

1. Materials and Strains

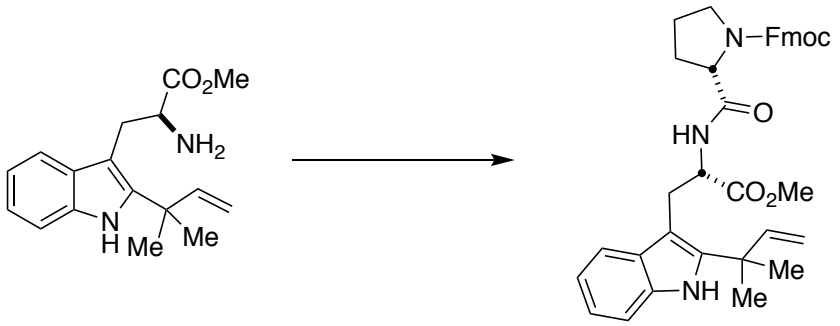
ATP (MilliporeSigma), dimethylallyl pyrophosphate (DMAPP, Isoprenoids) and NAD(P) (Roche) were purchased commercially. Optically pure (+)-premalbrancheamide and (+)-malbrancheamide were extracted from *Malbranchea aurantiaca* RRC1813A as previously described,[1,2] other chemical reagents used in this study were synthesized chemically. *E. coli* XL1-Blue cells were used for vector storage, *E. coli* DH10Bac (Invitrogen) cells were used for production of recombinant bacmids, *E. coli* BL21(DE3), pRare2-CDF[3], BAP1, pGro7 (Takara), BAP1-pG-KJE8 (Takara) and Insect High Five (BTI-TN-5B1-4, Invitrogen) cells were used for protein expression.[4]

2. General Chemical Procedures

^{1}H and ^{13}C spectra were obtained using 300 MHz, 400 MHz or 500 MHz spectrometers. The chemical shifts are given in parts per million (ppm) relative to residual $CDCl_3$ δ 7.26 ppm, CD_3OD δ 3.31 ppm, $(CD_3)_2CO$ δ 2.05 ppm or $(CD_3)_2SO$ δ 2.50 ppm for proton spectra and relative to $CDCl_3$ at δ 77.23 ppm, CD_3OD δ 49.00 ppm, $(CD_3)_2CO$ δ 29.84 ppm or $(CD_3)_2SO$ δ 39.52 ppm for carbon spectra. IR spectra were recorded on an FT-IR spectrometer as thin films. Mass spectra were obtained using a high/low resolution magnetic sector mass spectrometer. Flash column chromatography was performed with silica gel grade 60 (230-400 mesh). Preparative TLC was performed with silica gel 60 F_{254} 20 × 20 cm plates. Unless otherwise noted materials were obtained from commercially available sources and used without further purification. Dichloromethane (CH_2Cl_2), tetrahydrofuran (THF), N, N-dimethylformamide (DMF), acetonitrile (CH_3CN), triethylamine (Et_3N), and methanol (MeOH) were all degassed

with argon and passed through a solvent purification system containing alumina or molecular sieves in most cases.

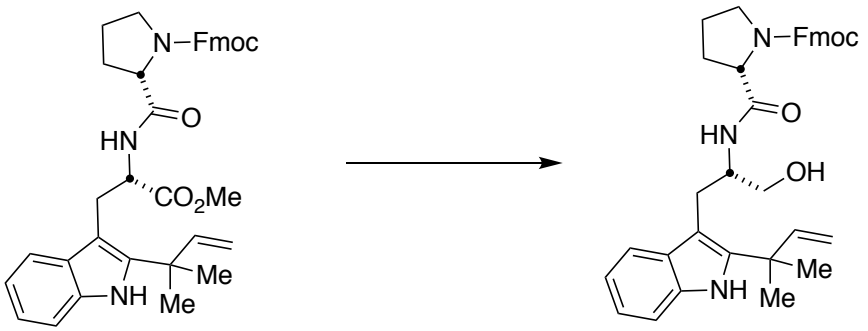
We attempted to coalesce rotomeric peaks by heating to 100 °C. In some cases it was successful, and others it was not. Reports show data of rotomeric compounds taken at 100 °C in DMSO.



13 **15**

Dipeptide 15. HATU (1.230 g, 3.23 mmol) and *i*-Pr_2NEt (1 mL, 742 mg, 5.74 mmol) were added to a solution of *N*-Fmoc-L-proline (1.120 g, 3.32 mmol) and tryptophan **13** (792 mg, 2.76 mmol) in CH_3CN (27 mL) and the reaction was stirred at room temperature for 2 hrs. The reaction was concentrated under reduced pressure, and the residue partitioned between Et_2O (50 mL) and 1 M HCl (50 mL). The layers were separated, and the organic phase was washed with saturated aqueous NaCl (50 mL), then dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 40% EtOAc/hexane to give 1.428 g (85%) of dipeptide **15** as an off-white foamy solid. ^{1}H NMR (300 MHz, DMSO, 100 °C) δ 10.12 (s, 1 H), 7.82 (d, *J* = 7.5 Hz, 2 H), 7.69-7.75 (m, 1 H), 7.61 (d, *J* = 7.7 Hz, 1 H), 7.60 (d, *J* = 7.7 Hz, 1 H), 7.35-7.45 (comp, 3 H), 7.23-7.40 (comp, 3 H), 6.93 (t, *J* = 7.5 Hz, 1 H), 6.84 (t, *J* = 7.5 Hz, 1 H), 6.10 (dd, *J* = 17.4, 10.6 Hz, 1 H), 4.99 (d, *J* = 17.4 Hz, 1 H), 4.97 (d, *J* = 10.6 Hz, 1 H), 4.60 (dd, *J* = 15.3, 7.6 Hz, 1 H), 4.14-4.30 (comp, 4 H), 3.74-3.87 (m, 1 H), 3.36 (s, 3 H), 3.23-3.33 (comp, 3 H), 3.09 (dd, *J* = 14.5, 7.0 Hz, 1 H), 1.98-2.08 (m, 1 H), 1.58-1.79 (comp, 3 H), 1.46 (s, 3 H), 1.45 (s, 3 H); ^{13}C NMR (75 MHz, DMSO, 100 °C) δ 171.6,

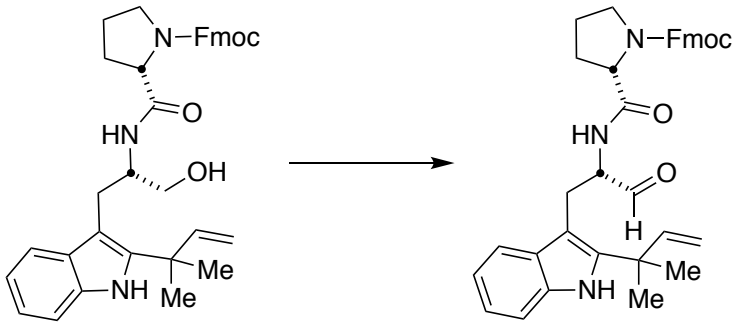
171.2, 153.7, 145.7, 143.5, 140.4, 140.3, 134.4, 128.8, 127.0, 126.5, 124.5, 119.9, 119.4, 117.7, 117.2, 110.6, 110.3, 104.5, 66.3, 59.3, 53.1, 50.8, 46.5, 46.3, 38.4, 29.7, 27.3, 27.2, 27.0, 22.7; IR (thin film) 3292, 1741, 1677, 1515, 1414, 1346, 1118, 911, 758 cm^{-1}; HRMS (ESI-APCI) *m/z* 606.2967 [$C_{37}H_{40}N_3O_5$ (M+H) requires 606.2968].



15 **16**

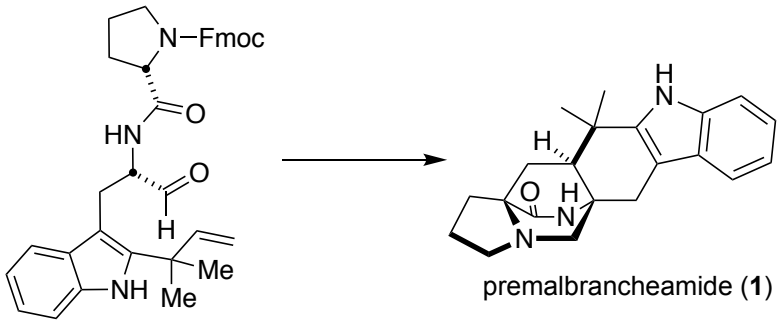
Alcohol 16. $NaBH_4$ (280 mg, 7.40 mmol) was added to a suspension of dipeptide **15** (1.428 g, 2.36 mmol) and LiCl (363 mg, 8.56 mmol) in THF (11 mL) and the reaction stirred for 5 min at room temperature. EtOH (11 mL) was then added and the reaction stirred for 7 hrs. The reaction was quenched with saturated NH_4Cl (25 mL) and extracted with EtOAc (50 mL). The layers were separated, and the organic phase was washed with saturated aqueous NaCl (25 mL), then dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 60-70% EtOAc/hexane to give 931 mg (82%) of alcohol **16** as a white foamy solid. ^{1}H NMR (300 MHz, DMSO, 100 °C) δ 10.02 (s, 1 H), 7.82 (d, *J* = 7.5 Hz, 2 H), 7.57-7.64 (comp, 3 H), 7.37 (t, *J* = 7.3 Hz, 2 H), 7.29 (t, *J* = 7.3 Hz, 2 H), 7.24 (d, *J* = 7.9 Hz, 1 H), 7.09 (d, *J* = 6.9 Hz, 1 H), 6.94 (t, *J* = 7.5 Hz, 1 H), 6.85 (t, *J* = 7.5 Hz, 1 H), 6.13 (dd, *J* = 17.4, 10.6 Hz, 1 H), 4.98 (d, *J* = 17.4 Hz, 1 H), 4.96 (d, *J* = 10.6 Hz, 1 H), 4.60 (dd, *J* = 15.3, 7.6 Hz, 1 H), 4.10-4.12-4.30 (comp, 5 H), 3.22-3.36 (comp, 4 H), 3.00 (dd, *J* = 14.5, 8.2 Hz, 1 H), 2.86 (dd, *J* = 14.5, 6.4 Hz, 1 H), 1.96-2.08 (m, 1 H), 1.55-1.77 (comp, 3 H), 1.48 (s, 3 H), 1.47 (s, 3 H); ^{13}C NMR (75 MHz, DMSO, 100 °C) δ 170.9, 153.9, 146.0, 143.5, 140.3, 140.0,

134.4, 129.2, 127.0, 126.5, 124.5, 119.7, 119.4, 117.8, 117.5, 110.4, 110.1, 106.3, 66.3, 62.1, 59.9, 52.2, 46.6, 46.3, 38.4, 29.8, 27.4, 27.3, 26.2, 22.7; IR (thin film) 3308, 1685, 1655, 1520, 1415, 1352, 1119, 910, 738 cm^{-1}; HRMS (ESI-APCI) *m/z* 578.3023 [$C_{36}H_{40}N_3O_4$ (M+H) requires 578.3019].

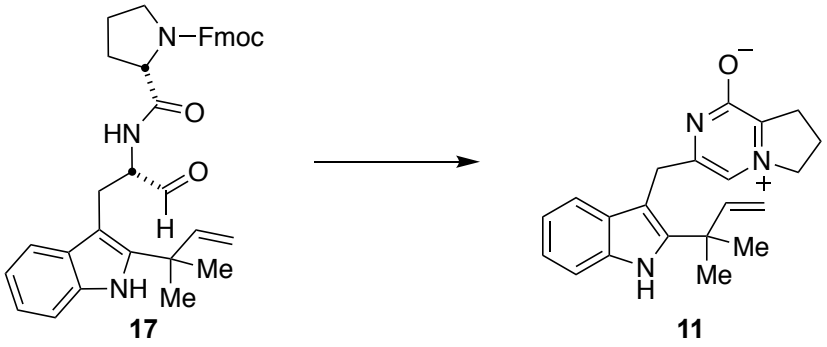


Aldehyde 17. $SO_3 \cdot Py$ (180 mg, 1.13 mmol) was added to a solution of alcohol **16** (160 mg, 0.28 mmol), Et_3N (0.2 mL, 145 mg, 1.43 mmol) and DMSO (1.5 mL) in CH_2Cl_2 (3 mL) and the reaction stirred for 2 hrs at room temperature. The reaction was partitioned between water (10 mL) and EtOAc (10 mL). The layers were separated, and the organic phase was washed with 1 M HCl (10 mL) and saturated aqueous NaCl (10 mL), and then dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 45% EtOAc/hexane to give 115 mg (72%) of aldehyde **17** as dark yellow solid. ^{1}H NMR (500 MHz, DMSO, 100 °C) δ 10.22 (bs, 1 H), 9.42 (s, 1 H), 7.90 (bs, 1 H), 7.85 (d, *J* = 7.55 Hz, 2 H), 7.61-7.65 (comp, 2 H), 7.46 (m, 1 H), 7.40 (t, *J* = 7.65 Hz, 2 H), 7.32 (m, 3 H), 7.00 (m, 1 H), 6.91 (m, 1 H), 6.18 (m, 1 H), 5.04 (m, 2 H), 4.44 (m, 1 H), 4.22-4.31 (comp, 4 H), 3.29-3.38 (comp, 3 H), 3.08 (m, 1 H), 2.00-2.10 (comp, 1 H), 1.70-1.80 (comp, 3 H), 1.51 (d, *J* = 3.45 Hz, 3 H) 1.50 (d, *J* = 2.05 Hz, 3H); ^{13}C NMR (75 MHz, DMSO, 100 °C) δ 199.67, 199.54, 171.72, 171.67, 153.80, 153.73, 145.85, 145.79, 143.50, 143.48, 143.41, 140.40, 140.37, 140.29, 134.45, 128.84, 127.05, 126.51, 124.56, 124.49, 119.98, 119.95, 119.41, 117.84, 117.82, 117.41, 117.28,

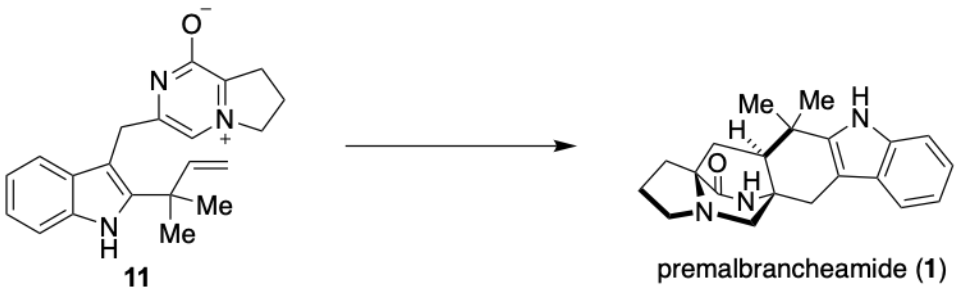
110.71, 110.69, 110.39, 110.36, 104.42, 104.36, 78.59, 66.27, 66.24, 59.47, 59.03, 58.95, 46.54, 46.51, 46.28, 38.36, 27.41, 27.40, 27.27, 27.26, 24.16, 24.07; IR (thin film) 3281, 1684, 1508, 1416, 1341, 1119, 912, 739 cm^{-1}; HRMS (ESI-APCI) *m/z* 576.2884 [$C_{36}H_{38}N_3O_4$ (M+H) requires 576.2862].



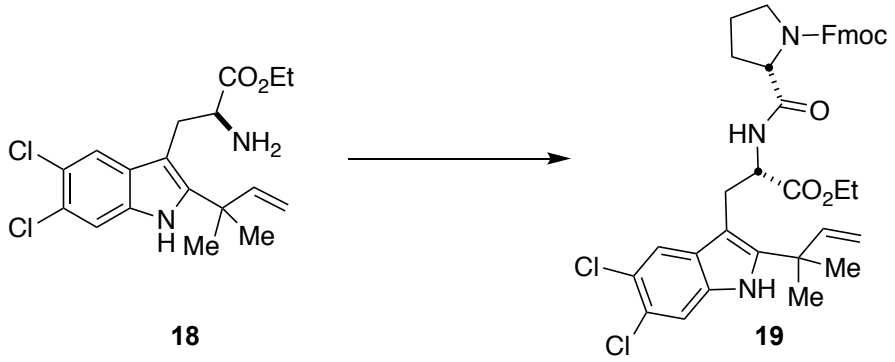
Premalbrancheamide 1. Et_2NH (0.5 mL) was added to CH_3CN (2.5 mL) and the resulting solution was sparged with argon for 15 min. The 5:1 CH_3CN:Et_2NH solution (2 mL) thus prepared was added to aldehyde **17** (54 mg, 0.094 mmol) and the reaction was stirred for 2 hrs at room temperature and the reaction was concentrated under reduced pressure. The residue was dissolved in the THF (2 mL) and TFA (0.02 mL, 30 mg, 0.26 mmol) was added. The reaction stirred for 24 hrs at room temperature. The reaction was quenched with saturated aqueous $NaHCO_3$ (10 mL) and the resulting mixture was extracted with CH_2Cl_2 (3 × 5 mL). The combined organic phases were dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 2% MeOH/CH_2Cl_2 to give 9 mg of impure premalbrancheamide as an off-white solid. Further purification by flash chromatography eluting with 50% EtOAc/hexane gave 5 mg (15%) of **1** as a white solid. All spectral data matched those previously reported.[5]



Prenylated zwitterion 11. Et_2NH (0.102 mL) was added to a solution of aldehyde **17** in MeCN (0.51 mL) and the reaction stirred at room temperature for 2 hrs. The reaction was concentrated under reduced pressure, the residue was taken up in CH_2Cl_2 (1.55 mL) and allowed to stand for 2 days. The resulting solution was concentrated under reduced pressure. The residue was purified by preparative thin layer chromatography eluting with 5% MeOH/CH_2Cl_2 to give 4.65 mg (45%) of **11** as a yellow solid. ^{1}H NMR (500 MHz, DMSO, 25 °C) δ 10.61 (s, 1 H), 7.33 (d, *J* = 7.95 Hz, 1 H), 7.28 (d, *J* = 7.75 Hz, 1 H), 7.01 (dd, *J* = 7.30, 7.75 Hz, 1 H), 6.89 (dd, *J* = 7.25, 7.65 Hz, 1 H), 6.66 (s, 1 H), 6.18 (dd, *J* = 10.50, 17.45 Hz, 1 H), 5.06 (d, *J* = 17.45 Hz, 1 H), 5.01 (d, *J* = 10.45 Hz, 1 H), 4.39 (t, *J* = 7.8 Hz, 2 H), 4.00 (s, 2 H), 2.95 (t, *J* = 7.45 Hz, 2 H), 2.11 (m, *J* = 7.5 Hz, 2 H), 1.48 (s, 6 H). ^{13}C NMR (500 MHz, DMSO, 25 °C) δ 173.63, 146.03, 141.64, 140.89, 136.49, 134.76, 129.05, 126.45, 120.50, 118.44, 117.79, 110.91, 105.25, 101.93, 64.08, 58.25, 34.64, 31.25, 29.61, 27.73; (ESI-M-TOFMS) *m/z* 334.1939 [$C_{21}H_{23}N_3O$ (M+H) requires 334.1919].

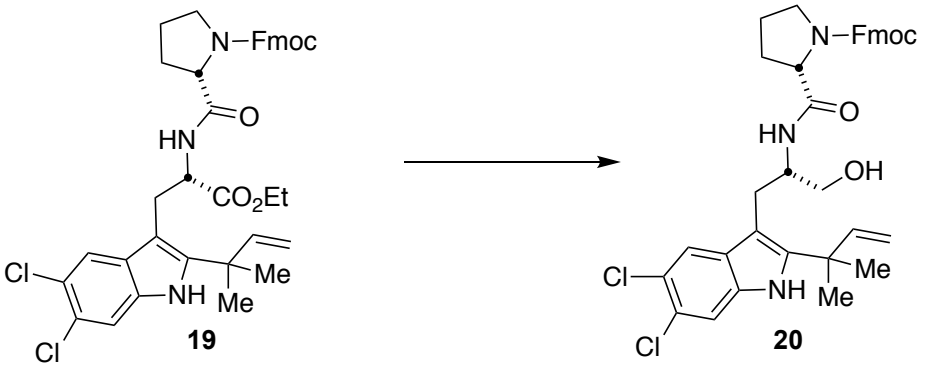


Premalbrancheamide 1. A solution of **11** (3.9 mg, 0.0117 mmol) in DMSO (0.65 mL) was added to a solution of NADH (89.36 mg, 0.117 mmol) in HEPES buffer (50 mM, 23.4 mL) and the resulting solution was sparged with argon for 15 min, then stirred for 18 hrs at room temperature. The resulting mixture was extracted with EtOAc (3 × 20 mL). The combined organic phases were washed with saturated aqueous NaCl (20 mL), dried (Na_2SO4), filtered and concentrated under reduced pressure. The residue was purified by preparative thin layer chromatography eluting with 80% EtOAc/Hex to give 3 mg (76.5%) of premalbrancheamide (**1**) as a white solid. All spectral data matched those previously reported.[4]



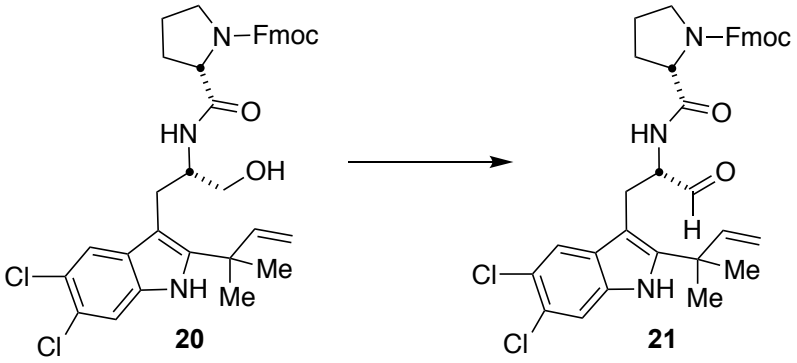
Dipeptide 19. HATU (31 mg, 0.081 mmol) and *i*-Pr_2NEt (0.04 mL, 30 mg, 0.216 mmol) were added to a solution of *N*-Fmoc-L-proline **14** (18.2 mg, 0.054 mmol) and tryptophan **18**[6] (20 mg, 0.054 mmol) in CH_3CN (1.5 mL) and the reaction was stirred at room temperature for 18 hrs. The reaction was concentrated under reduced pressure, and the residue was re-suspended in H_2O (10 mL) and extracted with ethyl acetate (30 mL). The organic phase was washed with saturated aqueous NaCl (30 mL), then dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 20% EtOAc/hexane (100 mL) and then 50% EtOAc/hexane (100 mL) to give 32.5 mg (87%) of dipeptide **19** as an off-white foamy solid. The product is a mixture of diastereomers and amide rotamers. ^{1}H NMR (300 MHz, DMSO, 100 °C) δ 10.49 (d, *J* = 8.7 Hz, 1 H), 7.84 (d, *J* = 7.5 Hz, 2 H), 7.61-7.64 (comp, 3 H),

7.46 (s, 1 H), 7.31-7.43 (comp, 4H), 6.12 (dd, *J* = 18.0, J = 10.5, 1 H), 5.06 (d, *J* = 11.7 Hz, 1 H), 5.01 (d, *J* = 18 Hz, 1 H), 4.67 (m, 2 H), 4.44 (m, 41H), 4.23-4.34 (comp, 4 H), 3.80-3.83 (m, 2 H), 3.37-3.52 (comp, 2 H), 1.90-1.98 (comp, 4 H), 1.50 (s, 3 H), 1.45 (s, 3 H), 0.88(t, *J* = 6.3 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$, 25°C) δ 172.0, 171.5, 156.3, 145.3, 144.1, 143.0, 141.5, 133.0, 130.0, 127.9, 127.2, 125.3, 123.5, 120.1, 119.4, 119.2, 112.9, 112.1, 105.8, 67.9, 61.6, 61.2, 53.7, 47.8, 47.3, 39.4, 31.3, 29.9, 28.3, 27.6, 24.8, 23.7; IR (thin film) 3326, 1735, 1681, 1514, 1417, 1352, 1264, 1118, 736 cm^{-1}; HRMS (ESI-APCI) *m/z* 688.2349 [$C_{38}H_{40}Cl_2N_3O_5$ (M+H) requires 688.2340].



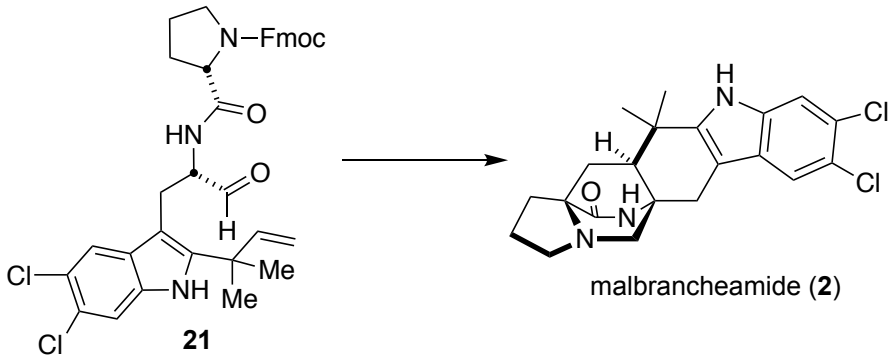
Alcohol 20. $NaBH_4$ (5.6 mg, 0.09 mmol) and LiCl (9.0 mg, 0.12 mmol) were added to a solution of dipeptide **19** (26.8 mg, 0.039 mmol) in CH_2Cl_2 (0.2 mL) and EtOH (0.2 mL). The reaction stirred for 18 hrs and was quenched with saturated NH_4Cl (3 mL) and extracted with EtOAc (9 mL). The layers were separated, and the organic phase was washed with saturated aqueous NaCl (10 mL), then dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue was purified by preparative TLC with 2% MeOH/CH_2Cl_2 to give 22.7 mg (90 %) of alcohol **20** as a white foamy solid. The product is a mixture of diastereomers and amide rotamers. ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C) δ 7.91 (s, 1 H), 7.77 (d, *J* = 6.8 Hz, 1 H), 7.76 (s, 1 H), 7.59 (comp, 2 H), 7.29-7.40 (comp, 5 H), 6.75 (br s, 1H), 6.50 (br s, 1H), 6.07 (dd, *J* = 17.6, 10.4 Hz, 1 H), 5.15 (d, *J* = 17.2 Hz, 2 H), 4.15-4.33-4.42 (comp, 5 H), 3.42-3.71 (comp, 4 H), 2.99 (s, 1

H), 2.30-2.09 (comp, 3 H), 1.75-2.05 (comp, 3 H), 1.53 (s, 3 H), 1.50 (s, 3 H); ^{13}C NMR (100 MHz, CD_3OD, 25 °C) δ 174.7, 156.6, 147.1, 145.2, 144.3, 142.5, 135.1, 131.2, 128.7, 128.1, 126.1, 125.0, 123.2, 120.8, 120.5, 112.9, 112.1, 108.0, 68.6, 64.2, 61.8, 54.3, 40.3, 32.3, 30.9, 28.1, 27.2, 25.1, 24.2; IR (thin film) 3309, 1672, 1532, 1450, 1262, 1119, 735 cm^{-1}; HRMS (ESI-APCI) *m/z* 668.2050 [$C_{36}H_{37}Cl_2N_3NaO_4$ (M+Na) requires 668.2053].

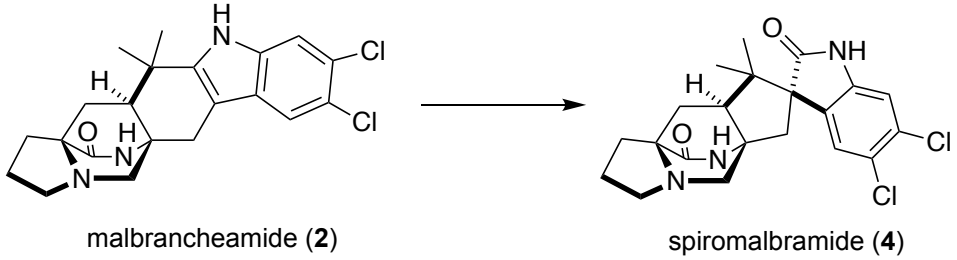


Aldehyde 21. SO_3·Py (44 mg, 0.28 mmol) was added to a solution of alcohol **20** (45 mg, 0.07 mmol), Et_3N (0.05 mL, 38 mg, 0.38 mmol) and DMSO (0.37 mL) in CH_2Cl_2 (0.75 mL) and the reaction stirred for 3 hrs at room temperature. The reaction was partitioned between water (5 mL) and EtOAc (5 mL). The layers were separated, and the organic phase was washed with 1 M HCl (7 mL) and saturated aqueous NaCl (7 mL), and then dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by preparative TLC with 2% MeOH/CH_2Cl_2 (3×) to give 40 mg (89%) of aldehyde **21** as an off-white foam. The product is a mixture of diastereomers and amide rotamers. ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C) δ 9.51 (s, 1 H), 7.96 (d. *J* = 4.6 Hz, 1H), 7.75 (comp, 2H), 7.29-7.41-7.63 (comp, 7 H), 6.46 (d, *J* = 8.8 Hz, 1 H), 6.10 (m, 1 H), 5.16 (m, 2 H), 4.62 (q, *J* = 6.8 Hz, 1 H), 4.04-4.29-4.35 (comp, 4 H), 2.99-3.49 (comp, 4 H), 1.84-2.09 (comp, 4 H), 1.56 (s, 3 H), 1.53 (s, 3 H); ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C) δ 199.8, 172.8, 156.5, 145.6, 145.2, 143.9, 142.9, 141.5, 133.1, 128.0, 127.3, 125.3, 125.2, 120.2, 119.7, 113.3, 112.3 96.7, 68.0, 61.5, 60.0, 56.1, 55.0, 51.1, 47.4, 39.3, 29.9, 28.3, 27.6, 26.8, 24.8,

23.9; IR (thin film) 3315, 1673, 1517, 1450, 1416, 1353, 1118, 737 cm^{-1}; HRMS (ESI-APCI) *m/z* 644.2080 [$C_{36}H_{36}Cl_2N_3O_4$ (M+H) requires 644.2080].

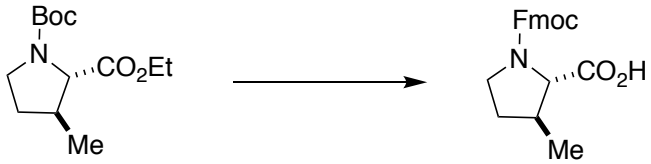


Malbrancheamide 2. Et_2NH (0.1 mL) was added to CH_3CN (0.4 mL) and the resulting solution was degassed. The 5:1 CH_3CN:Et_2NH solution (0.5 mL) was added to aldehyde **21** (13 mg, 0.020 mmol) and the reaction was stirred for 2 hrs at room temperature and the reaction was concentrated under reduced pressure. The residue was dissolved in a THF (0.5 mL) and TFA (0.005 mL, 7.45 mg, 0.06 mmol) degassed solution. The reaction stirred for 2 days at room temperature. The reaction was quenched with 1M NaOH (1 mL) and the resulting mixture was extracted with CH_2Cl_2 (3 × 5 mL). The combined organic phases were dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue was purified by preparative TLC with 3% MeOH/CH_2Cl_2 to give 3.1 mg (38%) of pure malbrancheamide as a white solid. All spectral data matched those previously reported.[6]



Spiromalbramide 4. NCS (5.0 mg, 0.037 mmol) was added to a solution of **2** (10.0 mg, 0.025 mmol) in DMF (0.5 mL) and the reaction stirred for 3 hrs at -15°C to 0 °C. To the solution was

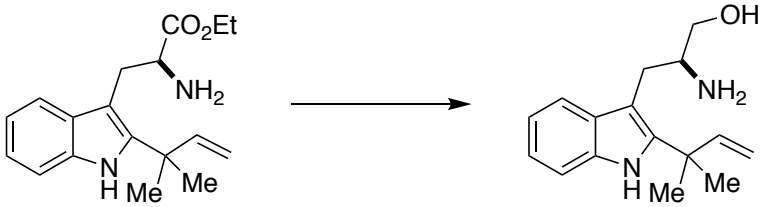
added pTsOH (60 mg, 0.35 mmol) and H_2O (0.2 mL) which stirred at 70 °C for 20 min and cooled to room temperature. The reaction was partitioned between 5% aqueous sodium carbonate (2 mL) and EtOAc. The layers were separated, and the organic phase was washed with saturated aqueous NaCl (6 mL), and then dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue was purified by preparative TLC with 3 × 4% MeOH/CH_2Cl_2 to give 5.2 mg (49.5% yield) of pure spiromalbramide as a white solid. All spectral and HRMS data matched those previously reported.[2]



40 **34**

***N*-Fmoc-*β*-methyl-L-proline 34.** NaOH (4.05 g, 15.75 mmol) was added to a solution of **40**[7] in MeOH (315 mL) and the reaction was heated to reflux for 18 hrs. The resulting solution was acidified (pH = 2) with 0.1 M HCl (900 mL) and washed with EtOAc (3 × 500 mL). The combined organic phases were dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue was recrystallized from EtOAc and Hexanes to yield 2.29 g (63.4%) of proline **34** as a white solid. TFA (9.92 mL) was added to a solution of *β*-MeProline **34** (1 g, 4.36 mmol) in DCM (9.92 mL) and the reaction stirred at 0 °C for 1 hr. The reaction was concentrated under reduced pressure, the residue was taken up in dioxane (21.8 mL) and Fmoc-Osu (1.67 g, 4.95 mmol) and K_2CO_3 (21.8 mL, 21.8 mmol) were added. The reaction was allowed to stir at room temperature for 18 hrs. The resulting solution was diluted with deionized H_2O (30 mL) and extracted with EtOAc (2 × 50 mL). The aqueous phase was acidified with 2 M HCl (25 mL) and washed with EtOAc (1 × 50 mL, 2 × 25 mL). The combined organic phases were washed with NaCl (2 × 75 mL), dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue

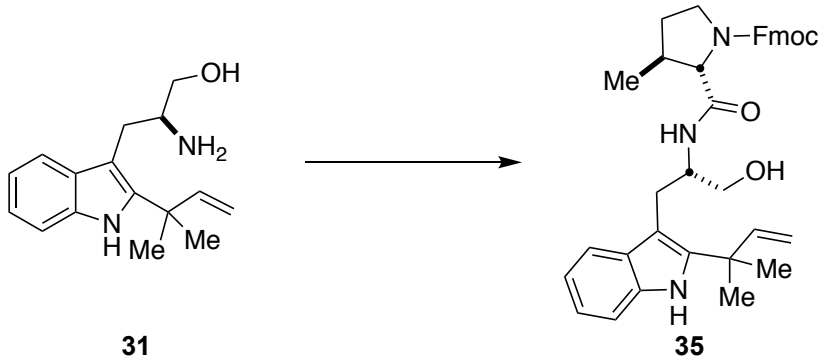
was purified by flash chromatography eluting with 20-50% EtOAc/Hexanes to yield 1.39 g (91.1%) of proline **34** as a foamy white solid. ^{1}H NMR (400 MHz, DMSO, 25 °C) δ 12.65 (bs, 1 H), 7.89 (t, *J* = 6.60 Hz, 2 H), 7.65 (m, 2 H), 7.42 (t, *J* = 7.42 Hz, 2 H), 7.33 (m, 2 H), 3.51 (m, 1 H), 3.37 (m, 1 H), 2.29 (m, 1 H), 1.99 (m, 1 H), 1.53 (m, 1 H), 1.11 (dd, *J* = 6.78, 17.10 Hz, 3H); ^{13}C NMR (400 MHz, DMSO, 25 °C) δ 173.66, 173.19, 153.90, 153.85, 143.83, 143.76, 143.68, 140.75, 140.66, 140.63, 127.69, 127.15, 127.14, 125.28, 125.21, 125.18, 125.10, 120.12, 66.94, 66.52, 65.74, 65.30, 46.74, 46.66, 45.92, 45.40, 37.88, 31.92, 30.87, 18.63, 18.28; Maxis Q-TOF (ESI) *m/z* 352.1547 [$C_{21}H_{21}NO_4$ (M+H) requires 352.1549]. Maxis Q-TOF (ESI) *m/z* 352.1547 [$C_{21}H_{21}NO_4$ (M+H) requires 352.1549].



41 **31**

Reduced Tryptophan 31. $NaBH_4$ (84 mg, 2.22 mmol) was added to a solution of tryptophan **41** in MeOH (1.16 mL) and the reaction was stirred at room temperature for 1 hr. The reaction was quenched with saturated NH_4Cl (6.5 mL) and washed with EtOAc (13 mL). The layers were separated, and the organic phase was washed with saturated aqueous NaCl (6.5 mL), then dried ($MgSO_4$), filtered and concentrated under reduced pressure to yield 139.2 mg (87.3 %) of alcohol **31** as a white solid. ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C) δ 7.98 (bs, 1 H), 7.55 (d, *J* = 7.8 Hz, 1 H), 7.29 (d, *J* = 7.88 Hz, 1 H), 7.14 (t, *J* = 6.94 Hz, 1 H), 7.07 (t, *J* = 7.44 Hz, 1 H), 6.13 (dd, *J* = 10.56, 17.36 Hz, 1 H), 5.18 (d, *J* = 6.08 Hz, 1 H), 5.14 (s, 1 H), 3.67 (dd, *J* = 2.96, 11.12 Hz, 1 H), 3.46 (t, *J* = 7.14 Hz, 1 H), 3.30 (bs, 1 H), 2.95 (dd, *J* = 5.6, 14.5 Hz, 1H), 2.86 (dd, *J* = 8.68, 14.52 Hz, 1H), 2.27 (s, 3 H), 1.54 (s, 6 H), 1.26; ^{13}C NMR (400 MHz, $CDCl_3$, 25 °C) δ 146.27,

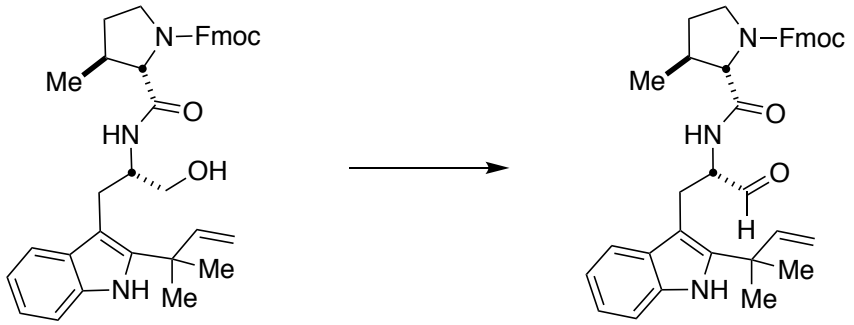
140.25, 134.31, 130.05, 121.62, 119.48, 118.73, 112.09, 110.55, 107.83, 66.46, 54.26, 39.25, 29.59, 28.01, 27.94; Maxis Q-TOF (ESI) *m/z* 259.1806 [$C_{16}H_{22}N_2O$ (M+H) requires 259.1810].



31 **35**

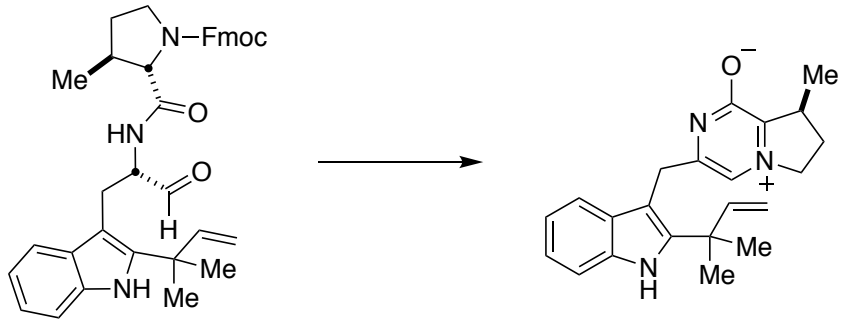
Alcohol 35. HATU (1.96 g, 5.16 mmol) and *i*-Pr_2NEt (3 mL, 2.22 g, 17.2 mmol) were added to a solution of *N*-Fmoc-*β*-methyl-L-proline **34** (1.59 g, 4.73 mmol) and reduced tryptophan **31** (1.11 g, 4.3 mmol) in CH_3CN (43 mL) and the reaction was stirred at room temperature for 4 hrs. The reaction was concentrated under reduced pressure, the residue was dissolved in EtOAc and partitioned between Et_2O (118 mL) and 1 M HCl (84 mL). The layers were separated, and the organic phase was washed with saturated aqueous NaCl (84 mL), then dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 60-80% EtOAc/hexane to give 2 g (81%) of alcohol **35** as a white foamy solid. ^{1}H NMR (500 MHz, DMSO, 100 °C) δ 10.05 (s, 1 H), 7.85 (d, *J* = 7.45 Hz, 2 H), 7.59-7.66 (comp, 3 H), 7.40 (t, *J* = 7.4 Hz, 2 H), 7.27-7.35 (comp, 3H), 6.95-6.99 (comp, 1 H), 6.87-6.93 (comp, 1 H), 6.14-6.26 (comp, 1 H), 4.97-5.09 (comp, 2 H), 4.22-4.33 (comp, 3 H), 4.16 (bs, 2 H), 3.73 (bs, 1 H), 3.34-3.44 (comp, 4 H), 3.02-3.07 (comp, 3 H), 2.86-2.96 (comp, 2 H), 2.06 (m, 1 H), 1.53 (s, 3 H), 1.51 (d, *J* = 3.95 Hz, 3 H), 0.99 (dd, *J* = 6.8, 22.3 Hz, 3 H) ^{13}C NMR (500 MHz, DMSO, 100 °C) δ 173.08, 146.24, 146.17, 142.35, 139.99, 139.97, 139.06, 137.09, 134.44, 129.28, 129.18, 128.31, 126.65, 120.67, 119.68, 119.66, 119.31, 117.96, 117.80, 117.46, 117.45, 110.34, 110.12, 110.06, 108.38, 106.37, 106.50, 78.58, 67.37, 67.28, 62.66, 62.30, 51.53, 51.30,

44.90, 44.59, 38.40, 38.17, 38.02, 33.98, 33.82, 27.51, 27.48, 27.37, 27.30, 26.32, 26.04, 18.88, 18.77; (ESI-M-TOFMS) *m/z* 592.3156 [$C_{37}H_{41}N_3O_4$ (M+H) requires 592.3175].

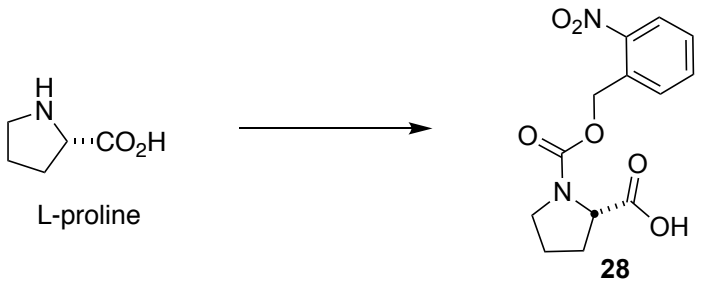


Aldehyde 36. $SO_3 \cdot Py$ (180 mg, 1.13 mmol) was added to a solution of alcohol **35** (160 mg, 0.28 mmol), Et_3N (0.2 mL, 145 mg, 1.43 mmol) and DMSO (1.5 mL) in CH_2Cl_2 (3 mL) at 0 °C and the reaction stirred for 2 hrs at the same temperature. The reaction was partitioned between water (10 mL) and EtOAc (10 mL). The layers were separated, and the organic phase was washed with 1 M HCl (10 mL) and saturated aqueous NaCl (10 mL), and then dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 45% EtOAc/hexane to give 115 mg (72%) of aldehyde **36** as dark yellow solid. ^{1}H NMR (500 MHz, DMSO, 100 °C) δ 10.21 (s, 1 H), 9.37 (s, 1 H), 7.85 (d, *J* = 7.55 Hz, 2 H), 7.61-7.64 (comp, 2 H), 7.38-7.48 (comp, 3 H), 7.29-7.33 (comp, 3 H), 6.97-7.01 (comp, 1 H), 6.88-6.93 (comp, 1 H), 6.17 (m, 1 H), 5.00-5.07 (comp, 2 H), 4.46 (bs, 1 H), 4.22-4.32 (comp, 4 H), 3.77 (d, *J* = 5 Hz, 1 H), 3.44 (bs, 1 H), 3.25-3.37 (comp, 2 H), 3.07 (m, 1 H), 2.05-2.14 (comp, 1 H), 1.51 (d, *J* = 4.05 Hz, 3 H), 1.49 (s, 3 H), 1.26 (d, *J* = 17.65 Hz, 3 H), 1.00 (d, *J* = 6.75 Hz, 2 H); ^{13}C NMR (500 MHz, DMSO, 100 °C) δ 199.68, 199.55, 171.32, 171.24, 153.80, 153.75, 145.82, 145.80, 145.78, 143.46, 143.41, 140.39, 140.36, 140.29, 134.45, 134.43, 128.84, 128.80, 127.03, 126.49, 126.47, 124.52, 124.47, 119.95, 119.39, 119.38, 117.80, 117.38, 117.28, 110.69, 110.66, 110.35, 104.41, 104.28, 78.58, 66.56, 66.19, 59.01, 58.80, 46.56, 46.53, 38.35,

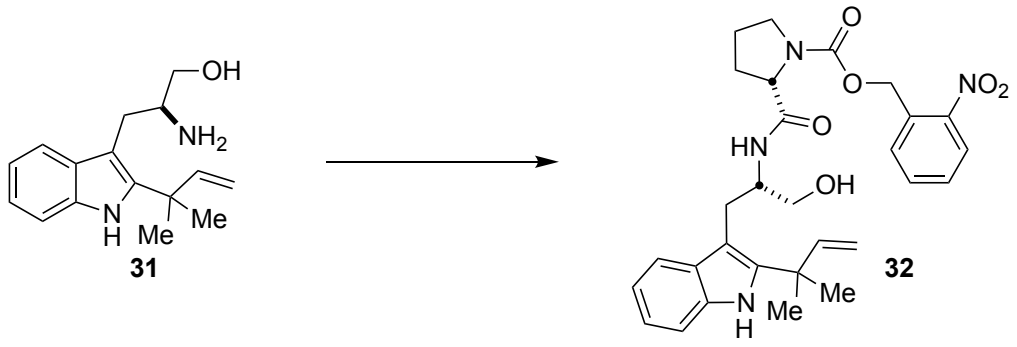
27.39, 27.24, 24.15, 18.02, 17.86; (ESI-M-TOFMS) *m/z* 590.3029 [$C_{37}H_{39}N_3O_4$ (M+H) requires 590.3019].



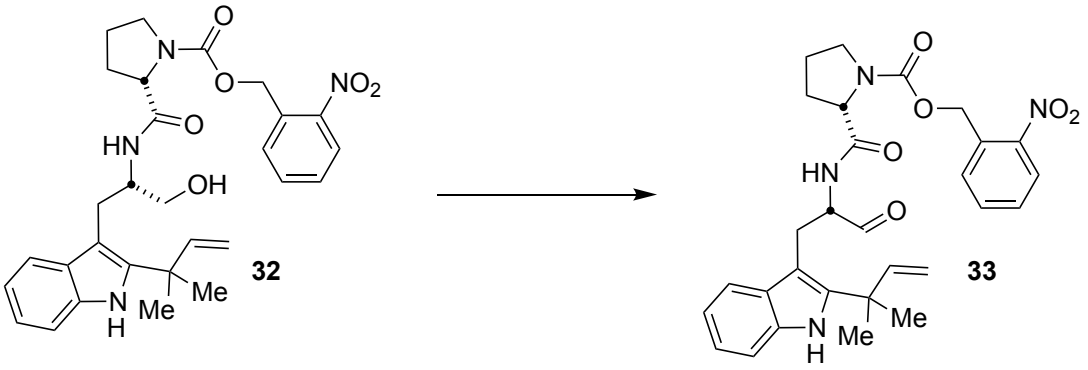
Prenylated zwitterion 38. Et_2NH (0.102 mL) was added to a solution of aldehyde **36** in MeCN (0.51 mL) and the reaction stirred at room temperature for 2 hrs. The reaction was concentrated under reduced pressure, the residue was taken up in CH_2Cl_2 (1.55 mL) and allowed to stand for 2 days. The resulting solution was concentrated under reduced pressure. The residue was purified by preparative thin layer chromatography eluting with 5% MeOH/CH_2Cl_2 to give 4.65 mg (45%) of **38** as a yellow solid. ^{1}H NMR (400 MHz, DMSO, 25 °C) δ 10.61 (s, 1 H), 7.33 (d, *J* = 8.00 Hz, 1 H), 7.30 (d, *J* = 7.92 Hz, 1 H), 7.01 (td, *J* = 1.10, 7.56 Hz, 1 H), 6.89 (td, *J* = 0.96, 7.92 Hz, 1 H), 6.65 (s, 1 H), 6.18 (dd, *J* = 10.52, 17.40 Hz, 1 H), 5.06 (dd, *J* = 1.16, 17.44 Hz, 1 H), 5.02 (dd, *J* = 1.20, 10.52 Hz, 1 H), 4.45 (m, 1 H), 4.30 (m, 1 H), 4.00 (s, 2 H), 3.37 (m, 1 H), 2.29-2.39 (comp, 1 H), 1.68-1.76 (comp, 1 H), 1.49 (d, *J* = 2.0 Hz, 6 H), 1.28 (d, *J* = 7.08 Hz, 3 H); ^{13}C NMR (400 MHz, DMSO, 25 °C) δ 164.31, 160.02, 146.17, 141.52, 141.00, 134.75, 129.06, 120.51, 118.45, 117.82, 110.98, 110.90, 108.59, 105.24, 57.14, 38.74, 37.31, 31.25, 27.74, 27.46, 16.40; HRMS (BTOF) *m/z* 348.2077 [$C_{22}H_{25}N_3O$ (M+H) requires 348.2070].



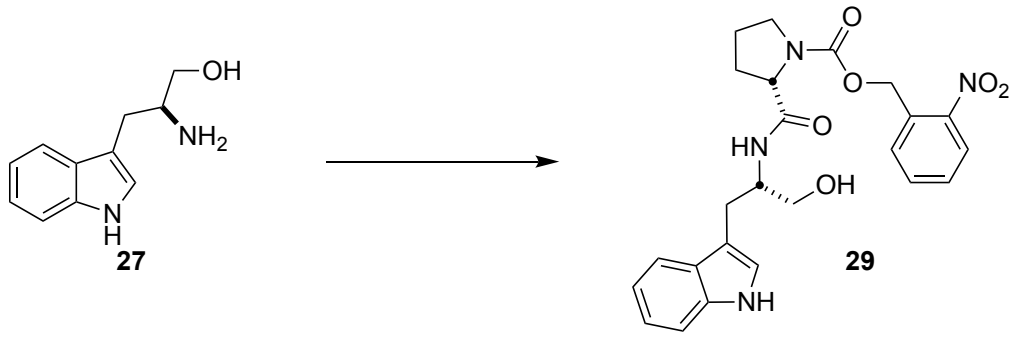
ONB-L-proline 28. A solution of previously prepared chloroformate[8] in DCM (8.6 mL) and dioxane (2.15 mL) at 0 °C was added in turn, alongside 2.37 mL NaOH (2 M), to a solution of L-proline in 2 M NaOH (2.15 mL) at 0 °C. The solution ran at the same temperature for 1 hr, then was allowed to warm to room temperature and run for another 18 hrs. The resulting organic phase was removed, and the aqueous phase was acidified (pH = 3-4) with 5 M HCl and washed with equal parts EtOAc. The combined organic phases were dried ($MgSO_4$), filtered and concentrated under reduced pressure. The crude reside was taken up in DCM and washed with equal parts 0.1 M HCl, dried over $MgSO_4$, filtered and concentrated. The residue was purified by flash chromatography eluting with 20% - 80% EtOAc/Hexanes to give 386 mg (30.6%) of proline **28** as a yellow oil. ^{1}H NMR (400 MHz, DMSO, 25 °C) δ 12.69 (bs, 1 H), 8.06-8.15 (comp, 1 H), 7.50-7.84 (comp, 3 H), 5.34-5.55 (comp, 2 H), 4.34 (m, 1 H), 4.34-3.53 (comp, 2 H), 2.14-2.36 (comp, 1 H), 1.76-2.03 (comp, 3 H); ^{13}C NMR (400 MHz, DMSO, 25 °C) δ 173.92, 173.47, 171.98, 171.62, 153.52, 153.41, 153.19, 152.98, 147.44, 147.36, 147.27, 146.96, 146.82, 134.27, 134.08, 134.02, 134.00, 132.84, 132.33, 132.25, 132.15, 131.03, 130.79, 130.66, 129.66, 129.51, 129.49, 129.44, 129.07, 129.02, 128.95, 128.78, 128.74, 128.56, 128.20, 124.96, 124.89, 124.85, 124.74, 63.15, 63.06, 62.89, 62.84, 59.01, 58.98, 58.45, 58.37, 46.89, 46.85, 46.20, 46.11, 30.46, 30.30, 29.37, 29.34, 23.95, 23.86, 23.03, 22.97.



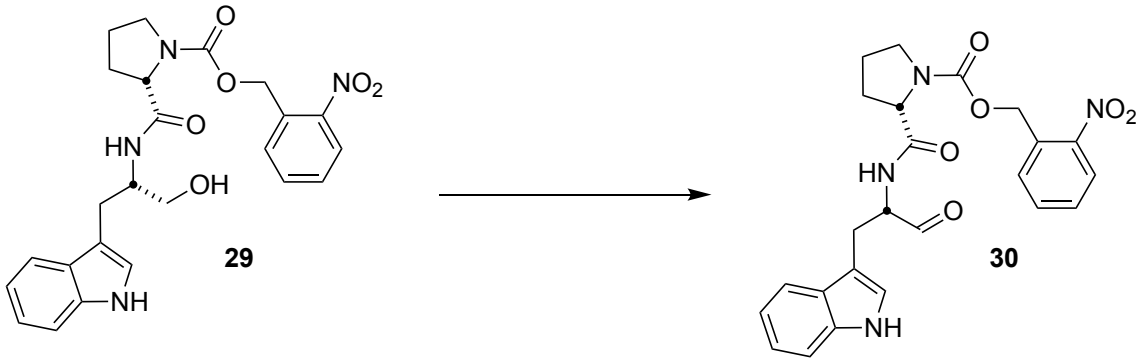
Alcohol 32. HATU (182.15 mg, 0.65 mmol) and *i*-Pr_2NEt (0.45 mL, 336.02 mg, 2.6 mmol) were added to a solution of **28** (229.4 mg, 0.78 mmol) and **31** (168 mg, 0.65 mmol) in CH_3CN (6.5 mL) and the reaction was stirred at room temperature for 2 hrs. The reaction was concentrated under reduced pressure, the residue was dissolved in 20 mL EtOAc and washed with 1 M HCl (20 mL), saturated aqueous NaCl (20 mL), dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 80% EtOAc/hexane to give 212.4 mg (61.2%) of alcohol **32** as a yellow solid. ^{1}H NMR (500 MHz, DMSO, 100 °C) δ 10.04 (s, 1 H), 8.05 (d, *J* = 8.05 Hz, 1 H), 7.62-7.74 (comp, 3 H), 7.56 (q, *J* = 7.8 Hz, 1 H), 7.28 (t, *J* = 8 Hz, 1 H), 7.17, (bs, 1 H), 6.95-6.99 (comp, 1 H), 6.88-6.93 (comp, 1 H), 6.20 (dq, *J* = 10.55, 6.8 Hz, 1 H), 5.36 (s, 2 H), 5.00-5.07 (comp, 2 H), 4.19-4.24 (comp, 2 H), 4.13 (bs, 1 H), 3.37-3.44 (comp, 4 H), 2.99-3.06 (m, *J* = 8.4 Hz, 2 H), 2.86 (m, *J* = 7.05 Hz, 1 H), 1.71-1.81 (comp, 2 H), 1.62-1.68 (comp, 1H), 1.54 (d, *J* = 1.5 Hz, 3 H), 1.51 (d, *J* = 2.75 Hz, 3 H); ^{13}C NMR (500 MHz, DMSO, 100 °C) δ 170.86, 170.79, 153.24, 153.12, 146.16, 146.14, 146.07, 139.94, 139.90, 134.40, 133.20, 133.17, 129.24, 129.21, 128.41, 128.36, 128.21, 128.17, 123.79, 123.78, 119.68, 119.65, 117.96, 117.81, 117.51, 117.46, 110.37, 110.34, 110.11, 110.08, 106.45, 106.36, 78.58, 62.20, 62.12, 59.82, 59.69, 52.23, 46.38, 38.40, 38.37, 37.70, 27.47, 27.46, 27.31, 26.19, 26.14; Maxis Q-TOF (ESI) *m/z* 535.2560 [$C_{29}H_{34}N_4O_6$ (M+H) requires 535.2557].



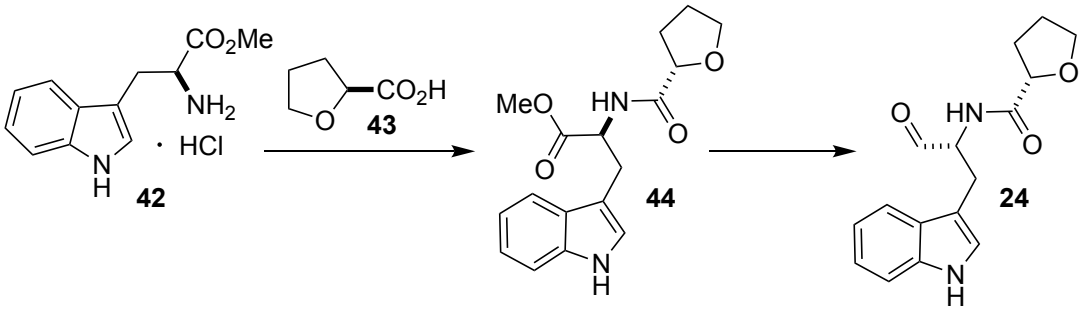
Aldehyde 33. $SO_3 \cdot Py$ (204.25 mg, 1.28 mmol) was added to a solution of alcohol **32** (171.4 mg, 0.321 mmol), Et_3N (0.23 mL, 165.7 mg, 1.64 mmol) and DMSO (1.72 mL) in CH_2Cl_2 (3.6 mL) and the reaction stirred for 3 hrs at room temperature. The reaction was partitioned between water (20 mL) and EtOAc (20 mL). The layers were separated, and the organic phase was washed with 1 M HCl (20 mL) and saturated aqueous NaCl (20 mL), and then dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 50% - 60% EtOAc/hexanes to give 75.6 mg (45%) of aldehyde **33** as yellow solid. ^{1}H NMR (500 MHz, DMSO, 100 °C) δ 10.22 (s, 1 H), 9.39 (s, 1 H), 8.05 (d, *J* = 8.15 Hz, 1 H), 7.94, (bs, 1 H), 7.68-7.71 (comp, 1 H), 7.62-7.65 (comp, 1 H), 7.57 (t, *J* = 7.6 Hz, 1 H), 6.44 (d, *J* = 7.95 Hz, 1 H), 7.30 (d, *J* = 7.9 Hz, 1 H), 7.00 (t, *J* = 7.55 Hz, 1 H), 6.92 (t, *J* = 7.35 Hz, 1 H), 6.17 (dd, *J* = 10.6, 17.45 Hz, 1 H), 5.31-5.39 (comp, 2 H), 5.02-5.09 (comp, 2 H), 4.3 (m, 1 H), 4.27-4.30 (comp, 1 H), 3.35-3.44 (comp, 2 H), 3.31 (q, *J* = 7.40 Hz, 1 H), 3.07 (m, 1 H), 2.10 (bs, 1 H), 1.68-1.81 (comp, 3 H), 1.50 (d, *J* = 1.80 Hz, 6 H); ^{13}C NMR (500 MHz, DMSO, 100 °C) δ 199.60, 171.62, 153.15, 145.81, 145.79, 140.36, 134.43, 133.15, 131.77, 128.83, 128.42, 128.27, 123.82, 119.94, 117.79, 117.28, 110.70, 110.36, 104.39, 62.19, 59.43, 58.98, 46.37, 38.35, 27.38, 27.26, 24.01, 22.70; Maxis Q-TOF (ESI) *m/z* 533.2398 [$C_{29}H_{32}N_4O_6$ (M+H) requires 533.2400].



Alcohol 29. HATU (144 mg, 0.513 mmol) and *i*-Pr_2NEt (0.3 mL, 222 mg, 1.72 mmol) were added to a solution of ONB-L-proline **28** (151 mg, 0.513 mmol) and previously prepared reduced tryptophan **27**[9] (81.34 g, 0.43 mmol) in CH_3CN (4.3 mL) and the reaction was stirred at room temperature for 2 hrs. The reaction was concentrated under reduced pressure, the residue was dissolved in EtOAc (10 mL) and washed with 1 M HCl (10 mL), saturated aqueous NaCl (10 mL), dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 20-80% EtOAc/hexane to give 124.4 mg (62.1%) of alcohol **29** as a yellow solid. ^{1}H NMR (500 MHz, DMSO, 100 °C) δ 10.44 (bs, 1 H), 8.03 (d, *J* = 8.1 Hz, 1 H), 7.64 (bs, 2 H), 7.57, (d, *J* = 7.9 Hz, 1 H), 7.52-7.55 (comp, 1 H), 7.48 (m, 1 H), 7.29 (d, *J* = 8.05 Hz, 1 H), 7.07 (bs, 1 H), 7.02 (t, *J* = 7.5 Hz, 1 H), 6.94 (t, *J* = 7.43 Hz, 1 H), 5.31-5.39 (comp, 2 H), 4.26 (dd, *J* = 3.2, 8.45 Hz, 1 H), 4.04 (m, 1 H), 3.42 (m, 4 H), 2.91 (m, 1 H), 2.82 (m, 1 H), 2.07-2.15 (comp, 1 H), 1.83-1.88 (comp, 1 H), 1.78 (m, 2 H); ^{13}C NMR (500 MHz, DMSO, 100 °C) δ 150.33, 135.87, 133.13, 128.36, 128.17, 128.07, 127.25, 123.76, 122.57, 120.17, 119.94, 117.81, 117.58, 110.89, 110.64, 75.21, 62.17, 62.13, 59.69, 56.13, 51.48, 46.39, 37.69, 25.91. HRMS (BTOF) *m/z* 467.19234 [$C_{24}H_{26}N_4O_6$ (M+H) requires 467.19251].



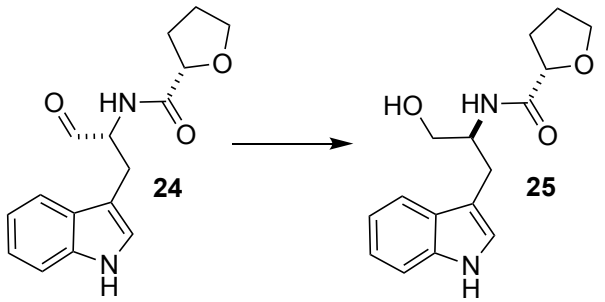
Aldehyde 30. $SO_3 \cdot Py$ (8.33 mg, 0.052 mmol) was added to a solution of alcohol **29** (6.1 mg, 0.0131 mmol), Et_3N (0.01 mL, 6.75 mg, 0.07 mmol) and DMSO (0.07 mL) in CH_2Cl_2 (0.15 mL) and the reaction stirred for 3 hrs at 0°C. The reaction was partitioned between water (2 mL) and EtOAc (2 mL). The layers were separated, and the organic phase was washed with 1 M HCl (2 mL) and saturated aqueous NaCl (2 mL), and then dried ($MgSO_4$), filtered and concentrated under reduced pressure. The residue was purified by preparative thin layer chromatography eluting with 3% MeOH/CH_2Cl_2 to give 2.2 mg (36.2%) of aldehyde **30** as dark yellow solid. ^{1}H NMR (500 MHz, DMSO, 100 °C) δ 10.56 (bs, 1 H), 9.51 (s, 1 H), 8.04 (d, *J* = 8.2 Hz, 1 H), 8.01 (bs, 1 H), 7.60-7.73 (comp, 2 H), 7.56 (m, 1 H), 7.50 (dd, *J* = 3.02, 7.77 Hz, 1 H), 7.32 (t, *J* = 7.00 Hz, 1 H), 7.10-7.12 (comp, 1 H), 7.03-7.07 (comp, 1 H), 6.95-6.99 (comp, 1 H), 5.36 (s, 2 H), 4.42 (q, *J* = 7.05 Hz, 1 H), 4.30-4.33 (comp, 1 H), 3.42-3.44 (comp, 2 H), 3.22 (m, 1 H), 3.04 (m, 2 H), 1.75-1.83 (comp, 3 H); ^{13}C NMR (500 MHz, DMSO, 100 °C) δ 199.72, 199.69, 171.73, 153.11, 135.89, 133.16, 133.13, 128.28, 128.26, 126.84, 123.81, 123.05, 120.41, 117.84, 117.54, 117.52, 110.81, 109.08, 62.21, 59.33, 58.39, 45.97, 40.42, 28.35, 23.63, 23.57, 8.16; HRMS (BTOF) *m/z* 465.17617 [$C_{24}H_{24}N_4O_6$ (M+H) requires 465.1774].



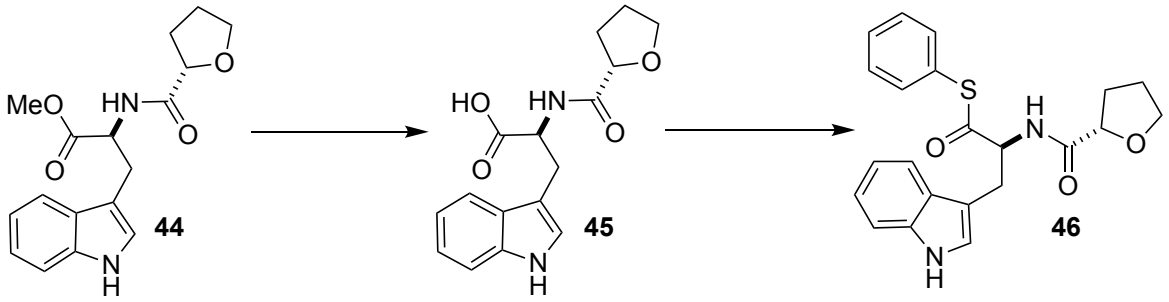
Aldehyde 24. HATU (6.730 g, 17.70 mmol) and *i*-Pr_2NEt (6.18 mL, 4.57 g, 35.40 mmol) were added to a solution of (S)-tetrahydrofuroic acid **43** (1.37 g, 11.80 mmol) and tryptophan **42** (3 g, 11.80 mmol) in CH_3CN (80 mL) and the reaction was stirred at room temperature for 20 hrs. The reaction was concentrated under reduced pressure, and the residue partitioned between equal parts EtOAc and saturated aqueous NH_4Cl. The layers were separated, and the organic phase was washed with equal parts saturated aqueous NaCl, then dried (Na_2SO_4), filtered and concentrated under reduced pressure. The crude methyl ester **44** was carried forward without further purification.

DIBAL-H (7.0 mL, 1M in toluene) was added to a solution of methyl ester **44** (800 mg, 2.50 mmol) in CH_2Cl_2 (25 mL) at -78 °C and the reaction stirred for 2 hrs at room temperature. The reaction was quenched with saturated aqueous Rochelle's Salt, allowed to warm to room temperature and stirred overnight. The layers were separated, and the aqueous phase was washed with equal parts CH_2Cl_2, dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 1:1 CH_2Cl_2/hexane to 1:2 CH_2Cl_2/EtOAc to give 1.69 g (50%), over two steps, of aldehyde **24**, as a mixture with alcohol **25**, as a brown foamy solid. ^1H NMR (400 MHz, $CDCl_3$, 25 °C) δ 9.65 (s, 1 H), 8.22 (bs, 1 H), 7.60 (d, *J* = 7.92, 1 H), 7.36 (d, *J* = 8.08 Hz, 1 H), 7.18-7.23 (comp, 1 H), 7.14 (q, *J* = 7.5 Hz, 1 H), 7.01 (d, *J* = 2.28 Hz, 1 H), 4.80 (m, 1 H), 4.35 (q, *J* = 5.76 Hz, 1 H), 3.84 (m, 1 H), 3.72 (m, 1 H), 3.57 (q, *J* = 7.0 Hz, 1 H), 3.23-3.42 (comp, 2 H), 2.17-2.29 (comp, 1 H), 2.03-2.11 (comp,

1 H), 1.86 (m, 2 H), 1.62 (m, 1 H); ^{13}C NMR (400 MHz, $CDCl_3$, 25 °C) δ 199.84, 199.69, 173.94, 136.38, 136.30, 127.64, 127.38, 123.01, 122.94, 122.58, 120.04, 119.95, 118.83, 118.73, 111.45, 109.97, 109.89, 78.50, 78.39, 69.58, 69.44, 58.65, 58.51, 30.38, 30.26, 25.53, 25.51, 25.29, 24.83; HRMS (ESI-APCI) *m/z* 287.1392 [$C_{16}H_{18}N_2O_3$ (M+H) requires 287.1396].



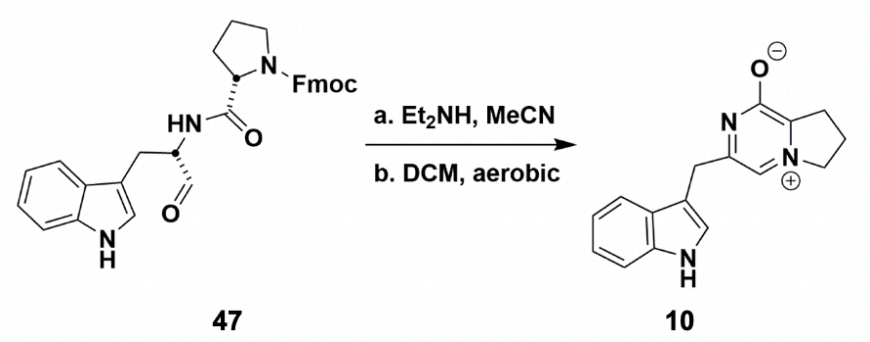
Alcohol 25. $NaBH_4$ (10 mg, 0.26 mmol) was added to a solution of aldehyde **24** (150 mg, 0.52 mmol) in MeOH (5 mL) and the reaction stirred for 7 hrs at room temperature. The reaction was quenched with saturated aqueous NH_4Cl and extracted with equal parts EtOAc, dried (Na_2SO_4), filtered and concentrated under reduced pressure. The residue was purified by flash chromatography eluting with 2:1 CH_2Cl_2/EtOAc to 30:1 CH_2Cl_2/MeOH to give 93 mg (61.6%) of alcohol **25** as a brown foamy solid. ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C) δ 8.71 (bs, 1 H), 7.62 (d, *J* = 7.84 Hz, 1 H), 7.33 (d, *J* = 8.12, 1 H), 7.16 (t, *J* = 7.42 Hz, 1 H), 7.09 (t, *J* = 7.46 Hz, 1 H), 6.99 (d, *J* = 2.08 Hz, 1 H), 6.96 (d, *J* = 7.92 Hz, 1 H), 4.28 (m, 2 H), 3.61-3.72 (comp, 3 H), 3.49 (q, *J* = 7.48, 1 H), 3.01 (m, 2 H), 2.07-2.16 (comp, 1 H), 1.79 (m, 1 H), 1.68 (m, 1 H), 1.45 (m, 1 H); ^{13}C NMR (400 MHz, $CDCl_3$, 25 °C) δ 194.26, 174.30, 136.36, 127.75, 123.02, 122.05, 119.45, 118.63, 111.39, 111.21, 78.31, 69.31, 64.68, 52.09, 30.15, 26.46, 25.25. HRMS (ESI-APCI) *m/z* 289.1549 [$C_{16}H_{20}N_2O_3$ (M+H) requires 289.1552].



Thioester 46. $LiOH \cdot H_2O$ (1.0 g, 23.80 mmol) was added to a solution of methyl ester **44** (1.5 g, 4.70 mmol) in THF (20 mL), MeOH (20 mL), and H_2O (10 mL) and the reaction stirred for 4 hrs at room temperature. The reaction was neutralized with 1 M HCl (22 mL). The organic phase was evaporated and the aqueous phase was washed with equal parts CH_2Cl_2, dried (Na_2SO_4), filtered and concentrated under reduced pressure. The crude carboxylic acid **45** was carried forward without further purification.

Thiophenol (0.5 mL, 550 mg, 5.0 mmol) and EDCI·HCl (725 mg, 3.75 mmol) were added to a solution of crude carboxylic acid **45** (750 mg, 2.5 mmol) and $HOBT \cdot H_2O$ (675 mg, 5.0 mmol) in EtOAc (25 mL) and the reaction was stirred at room temperature for 2 hrs. The reaction was quenched with saturated aqueous NH_4Cl. The layers were separated and the organic phase was washed with saturated aqueous NaCl, dried (Na_2SO_4), filtered and concentrated under reduced pressure. The reside was purified by flash chromatography eluting with 1:1 CH_2Cl_2/hexane to 15:1 CH_2Cl_2/EtOAc to give 740 mg (40%), over two steps, of thioester **46** as a yellow foamy solid. ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C) δ 8.52-8.60 (comp, 1 H), 7.58 (d, J = 7.84 Hz, 1 H), 7.32-7.42 (comp, 6 H), 7.21 (td, J = 1.0, 7.08 Hz, 1 H), 7.13 (td, J = 0.96, 7.92 Hz, 1 H), 6.99 (bs, 1 H), 5.14 (m, 1 H), 4.38 (t, J = 7.06 Hz, 1 H), 3.68 (q, J = 6.96 Hz, 1 H), 3.50 (q, J = 7.24, 1 H), 3.39 (d, J = 6.32 Hz, 2 H), 2.17 (m, 1 H), 1.85 (m, 1 H), 1.73 (m, 1 H), 1.51 (m, 1 H); (400 MHz, $CDCl_3$, 25 °C) δ 198.98, 198.93, 194.21, 173.75, 173.69, 136.22, 136.19, 134.70, 129.57, 127.13, 123.22, 123.15, 122.26, 119.66, 118.49, 111.52, 111.47, 109.47, 109.38, 78.27, 69.35, 59.07,

59.03, 29.99, 27.78, 25.31; HRMS (ESI-APCI) *m/z* 395.1416 [$C_{22}H_{22}N_2O_3S$ (M+H) requires 395.1429].



47 10

Zwitterion 10. Et_2NH (0.08 mL) was added to a solution of aldehyde **47** (12.1 mg, 0.024 mmol) in MeCN (0.40 mL) and the reaction stirred at room temperature for 2 hrs. The reaction was concentrated under reduced pressure, the residue was taken up in CH_2Cl_2 (1.20 mL) and allowed to stand for 2 days. The resulting solution was concentrated under reduced pressure. The residue was purified by preparative thin layer chromatography eluting with 5% MeOH/CH_2Cl_2 to give 3.0 mg (47.2%) of zwitterion **10** as a yellow solid. ^{1}H NMR (400 MHz, DMSO, 25 °C) δ 10.90 (s, 1 H), 7.53 (d, *J* = 7.84 Hz, 1 H), 7.33 (d, *J* = 8.08 Hz, 1 H), 7.21 (d, *J* = 2.24 Hz, 1 H), 7.14 (s, 1 H), 7.05 (t, *J* = 7.04 Hz, 1 H), 6.95 (t, *J* = 7.04 Hz, 1 H), 4.46 (t, *J* = 7.75 Hz, 2 H), 3.89 (s, 2 H), 2.95 (t, *J* = 7.58 Hz, 2 H), 2.13 (m, *J* = 7.6 Hz, 2 H). ^{13}C NMR (400 MHz, DMSO, 25 °C) δ 164.82, 159.04, 140.13, 136.24, 127.06, 123.75, 120.99, 118.67, 118.39, 111.38, 111.17, 109.30, 59.02, 31.69, 29.61, 19.09; (ESI-M-TOFMS) *m/z* 266.1234 [$C_{21}H_{23}N_3O$ (M+H) requires 266.1288].

Trp-SPh was prepared as previously reported[10], and ^{1}H and ^{13}C NMR spectra matched those of literature values.1 85% yield over 2 steps. ^{1}H NMR (599 MHz, Methanol-d4) δ 7.58 (d, J = 7.9 Hz, 1H), 7.49 – 7.46 (m, 1H), 7.43 (dd, J = 8.1, 6.7 Hz, 3H), 7.31 – 7.25 (m, 3H), 7.20 – 7.15 (m, 1H), 7.09 (t, J = 7.5 Hz, 1H), 4.59 (t, J = 7.1 Hz, 1H), 3.55 (dd, J = 14.8, 6.9 Hz, 1H), 3.41 (dd, J

= 14.8, 7.4 Hz, 1H). ^{13}C NMR (151 MHz, Methanol-d4) δ 195.32, 137.29, 134.80, 130.32, 129.55, 127.33, 125.65, 124.74, 122.00, 119.41, 118.04, 111.72, 106.33, 59.55, 28.08. HRMS (ESI-APCI) *m/z* 297.1069 [$C_{22}H_{22}N_2O_3S$ (M+H) requires 297.1056].

3. Construct Design of *malG* (*A_1-T_1, C, T_2, R*), *malE*, *malB*, *malC*, *malA*, *phqB R* and *phqE*

Coding sequences involved in this study were cloned from cDNA of *Malbranchea aurantiaca* RRC1813A and gDNA of *Penicillium fellutanum* ATCC20841. For cloning of *malG R*, *malE* and *malC*, PCR was used to amplify the cDNA template, followed by a ligation-independent cloning (LIC) procedure to insert the genes into the pMCSG7 vector.[11,12] For *malG A_1-T_1, malG C*, *malG T_2* and *phqB R*, the coding sequence was inserted into the pMCSG9 vector. For *phqE*, the pET28b vector was used. For *malB*, pH9 transfer vector was used.[13] The plasmids were transformed into *E. coli* XL1-Blue cells (pMCSG9-MalG A_1-T_1, pMCSG9-MalG C, pMCSG9-MalG T_2, pMCSG7-MalG R, pMCSG7-MalE, pMCSG7-MalC, pMCSG9-PhqB R and pET28b-PhqE) for storage and harvest. pH9-MalB was transformed into *E. coli* DH10Bac cells for production of recombinant bacmids as previously described.[13]

Cloning of *malA* and subsequent expression and MalA purification were as previously described.[1] Site-directed mutagenesis of *malC* and *phqE* was performed with the QuikChange kit (Agilent Technologies). Primers used for cloning are listed in Supplementary Table 1. All sequences were verified by Sanger sequencing at the University of Michigan Sequencing Core.

4. Protein Expression and Purification

For MalG C, MalG T_2, MalG R, MalE, PhqE or wild-type MalC, *E. coli* pRare2-CDF cells[3] were transformed with the corresponding plasmid and grown in 1 L Terrific Broth medium with 30

μg/mL ampicillin and 100 μg/mL spectinomycin at 37 °C to OD_{600} = ~0.8. For MalG T_2, a trace metals mix was added to ensure production of apo-T_2.[4] The culture was then shifted to 20 °C over 1 hr, induced with 0.4 mM IPTG, and incubated 18 - 20 hrs (20 °C, 225 rpm shake). Cells were harvested by centrifugation and stored at -20 °C. MalG A_1-T_1 was produced in *E. coli* BAP1-pG-KJE8 cells with the same protocol, excepting induction with 0.4 mM IPTG, 1 mg/mL L-arabinose and 4 ng/mL tetracycline. PhqB R was produced in *E. coli* pGro7 cells with the same protocol, excepting induction with 0.4 mM IPTG and 1 mg/mL L-arabinose. For production of selenomethionyl (SeMet) MalC, *E. coli* BL21(DE3) cells were transformed with pMCSG7-MalC and grown in SelenoMet medium (Molecular Dimensions), 30 μg/mL ampicillin and 50 μg/mL SeMet to OD_{600} = ~0.8 at 37 °C. The culture was shifted to 20 °C over 1 hr, induced with 0.4 mM IPTG, and incubated 18 - 20 hrs. Cells were harvested by centrifugation and stored at -20 °C. MalB was produced in High Five cells (Invitrogen). 1L of Insect X-press media (Lonza) in 2.8 L Fernbach flasks was seeded with High Five cells at 2×10^6 cells/mL. The cultures were infected at a multiplicity of infection (MOI) of 2 and incubated at 20 °C with shaking at 140 rpm for 72 hrs. The cells were harvested by centrifugation and stored at -80 °C.

For purification of MalG R, MalE, MalB, MalC, PhqE and PhqB R, the cell pellet was resuspended in lysis buffer (10% v/v glycerol, 500 mM NaCl, 20 mM Tris buffer pH 7.9, 20 mM imidazole pH 7.9, 0.1 mg/mL lysozyme, 0.05 mg/mL DNase and 1 mM $MgCl_2$), and mixed 30 min by vortex. Sonication and high speed centrifugation (16,000 rpm, 30 min) were applied to obtain the lysate soluble fraction. The soluble fraction was filtered and loaded on a Ni-NTA Histrap column and washed with 8 column volumes of Ni-NTA buffer (10% v/v glycerol, 500 mM NaCl, 20 mM imidazole pH 7.9, 20 mM Tris pH 7.9) at 3 mL/min. Proteins were eluted with an imidazole gradient (3 mL/min; 20 - 600 mM imidazole in 12 min). Fractions containing

the target protein were pooled and incubated with His-tagged tobacco etch virus (TEV) protease in a 1:50 w/w ratio at 20 °C for 4 hrs to remove the N-terminal His-tag or His-maltose binding protein (MBP)-tag. The tag-free protein was dialyzed overnight at 4 °C into 10% glycerol, 2 mM DTT, 500 mM NaCl, 20 mM Tris pH 7.9, and passed through the Ni-NTA Histrap column to remove TEV protease and any remaining tagged protein. Further homogeneity was achieved by size-exclusion chromatography with a GE Hiload 16/60 Superdex 200 prep grade column equilibrated with 10% v/v glycerol, 300 mM NaCl, 20 mM Tris pH 7.9 (1 mL/min). SDS-PAGE was used to assess protein homogeneity; all proteins were > 95% pure. MalG A_1-T_1, MalG C and MalG T_2 were purified with the same protocol described above, with different purification buffers (lysis buffer: 10% glycerol, 50 mM $(NH_4)_2SO_4$, 20 mM imidazole pH 7.0, 50 mM HEPES pH 7.0, 0.1 mg/mL lysozyme, 0.05 mg/mL DNase and 1 mM $MgCl_2$; Ni-NTA buffer: 10% glycerol, 50 mM $(NH_4)_2SO_4$, 20 mM imidazole pH 7.0, 50 mM HEPES pH 7.0; size-exclusion buffer: 10% glycerol, 50 mM $(NH_4)_2SO_4$, 50 mM HEPES pH 7.0). Sfp was prepared as previously described.[14]

5. MalG Substrate Loading

To load L-Pro onto MalG A_1-T_1, reaction of 150 μM MalG A_1-T_1 in 10% v/v glycerol, 50 mM NaCl, 50 mM HEPES pH 7, 5 mM ATP, 2 mM $MgCl_2$ was initiated by addition of 1 mM L-Pro. The reaction mix (50-100 μL) was incubated at 30 °C for 3 hrs, and dialyzed at 4 °C for 3 hrs into 10% v/v glycerol, 50 mM $(NH_4)_2SO_4$, 50 mM HEPES pH 7.0. To load L-Trp or dipeptidyl analog **23** onto MalG T_2, reaction of 150 μM MalG T_2 in 10% v/v glycerol, 50 mM NaCl, 50 mM HEPES pH 7, 5 μM Sfp, 20 mM $MgCl_2$ was initiated by addition of 1 mM Trp-CoA or **23**-CoA. CoA derivatives were prepared by transthioesterification of the phenol thioester of L-Trp

and **46** respectively.[15] The reaction mix (100-500 μL) was incubated at 30 °C for 3 hrs. For use in reconstitution assays, L-Trp-T_2 was dialyzed into 10% v/v glycerol, 50 mM $(NH_4)_2SO_4$, 50 mM HEPES pH 7.0. To analyze the efficiency of substrate loading, 5 μM loaded T domain in 10% v/v glycerol, 20 mM Tris pH 7.9 was analyzed by LC/MS (Phenomenex Aeris widepore C4 column (3.6 μm, 50 × 2.1 mm), buffer A: 0.2% v/v formic acid in water, buffer B: 0.2% v/v formic acid in acetonitrile. HPLC protocol: 5% buffer A for 2 min, 5 - 100% buffer B gradient for 4 min, 100% buffer B for 2 min. flow rate: 0.5 mL/min. Ionization parameters: fragmentor voltage, 225 V; skimmer voltage, 25 V; nozzle voltage, 1000 V; sheath gas temperature, 350 °C; drying gas temperature, 325 °C).

6. *In vitro* Malbrancheamide Pathway Reconstitution

Pathway reconstitution assays were performed with 150 μM L-Trp-T_2 and 10 μM of each enzyme in 10% v/v glycerol, 50 mM $(NH_4)_2SO_4$, 50 mM HEPES pH 7.0, 5 mM L-Pro, 5 mM ATP, 2 mM $MgCl_2$. Reactions were initiated by addition of cofactors: 5 mM NADPH or NADH, and 500 μM DMAPP for reactions including MalE or MalB. Reaction mixtures (100 μL) were incubated at 16°C with shaking at 300 rpm for 15 hrs, quenched with 50% v/v methanol, and cleared of denatured protein by centrifugation (13,000 rpm, 4 °C, 20 min). Products were analyzed by LC/MS (Phenomenex Kinetex C18 column (2.6 μm, 50 × 2.1 mm), buffer A: 0.2% v/v formic acid in water, buffer B: 0.2% v/v formic acid in acetonitrile. HPLC protocol: 5% buffer A for 2 min, 5 - 100% buffer B gradient for 8 min, 100% buffer B for 2 min. flow rate: 0.5 mL/min.). Chiral separations were performed using Phenomenex Lux cellulose-3 (5 μm, 250 × 4.6 mm) column (buffer A: water, buffer B: 95% acetonitrile; 19% acetonitrile for 3 min, 19 - 95%

acetonitrile gradient over 10 min, 95% acetonitrile for 2 min; flow rate 0.5 mL/min). All assays were performed in triplicate.

7. Aerobic Enzyme Assays

MalG R domain activity was assayed with **23**-T_2. Reaction of 150 μM **23**-T_2 and 20 μM MalG R in reaction buffer (10% v/v glycerol, 50 mM $(NH_4)_2SO_4$, 50 mM HEPES pH 7.0) was initiated by addition of 5 mM NADPH or NADH. The reaction mix (100 μL) was incubated at 25 °C with shaking at 300 rpm for 1 hr, quenched with 50% v/v methanol, and clarified by centrifugation. Products were analyzed by LC/MS.

MalE or MalB activity was assayed with free substrates (L-Trp or **10**) or with substrate-loaded T_2 (L-Trp-T_2 or **23**-T_2). Reaction of 150 μM substrate (L-Trp, **10**, L-Trp-T_2 or **23**-T_2) with 10 μM MalE or MalB in reaction buffer was initiated by addition of 500 μM DMAPP. Reaction mixtures (100 μL) were incubated 2 hrs at 25 °C with shaking (300 rpm). Reactions with free substrates were quenched with 90% v/v methanol, and cleared by centrifugation. Reactions with substrate-loaded T_2 were quenched with 1% formic acid.

MalC (wild type or mutant) was assayed in a 100 μL mixture containing 100 μM **11**, 10 μM MalC in reaction buffer. Reactions were initiated by addition of 1 mM NADPH or NADH, incubated at 25 °C with shaking at 300 rpm for 2 hrs, quenched with 90% v/v methanol, and clarified by centrifugation prior to product analysis by LC/MS. The effect of pH on the MalC activity was tested with the same reaction mix using buffers at five pHs (Bistris pH 6.0, HEPES pH 7.0/7.5/8.0, Tris pH 9.0). To determine kinetic constants for NADPH and NADH, reactions with 800 μM **11**, 10 μM MalC in reaction buffer were initiated with varying cofactor concentrations (NADPH: 0 μM, 5 μM, 10 μM, 20 μM, 30 μM, 50 μM, 100 μM; NADH: 0 μM,

25 μM, 50 μM, 100 μM, 200 μM, 350 μM, 500 μM). Each reaction mix (100 μL) was incubated at 25 °C with shaking at 300 rpm for 25 min, quenched with 90% v/v methanol, and clarified by centrifugation. Products were analyzed by LC/MS. Data were fit to the Michealis-Menten equation to calculate kinetic parameters.

The MalA assay with (+)-**1** produced from *in vitro* pathway reconstitution or chemically synthesized racemic **1** was performed as previously described.[1]

All assays were performed in triplicate.

8. Anaerobic Enzyme Assays

Assays were performed in an anaerobic chamber (25 °C, 0.8 ppm O_2). Prior to transfer to the chamber, the reaction buffer (10% v/v glycerol, 50 mM $(NH_4)_2SO_4$, 50 mM HEPES pH 7.0) was degassed with N_2, and solutions of individual reaction components were degassed with argon.

MalE activity was assayed with **30**. Caged compound **30** was photo-deprotected with UV light for 20 min to produce **8** immediately before adding the reaction components to reaction buffer in a mixture (100 μL) containing 150 μM **8**, 10 μM MalE and 500 μM DMAPP. Samples were removed from the anaerobic chamber after 2 hrs, immediately quenched with 90% v/v methanol, and re-gassed to convert unreacted **8** to **10**.

MalC activity was assayed with **33**. Caged compound **33** was photo-deprotected with UV light for 20 min to produce **9** immediately before adding the reaction components to reaction buffer in a mixture (100 μL) containing 100 μM **9**, 10 μM MalC and 5 mM $NADP^+$. A mixture without $NADP^+$ was used as a negative control. Samples were removed at 30-min to 18-hr time points, from the anaerobic chamber, immediately quenched with 90% v/v methanol, and re-gassed to convert unreacted **9** to **11**.

Denatured protein was removed by centrifugation and the products were analyzed by LC/MS. All assays were performed in triplicate.

9. Crystallization and Structure Determination

For crystallization of PhqB R domain, 10 mg/mL PhqB R (residues 2006 – 2429) was mixed with precipitant solution (10% PEG 8000, 200 mM $MgCl_2$, 100 mM Tris pH 7.0) in a 1:1 v/v ratio. For co-crystallization with NADPH, 10 mM NADPH was included in the precipitant solution. Crystals were grown at 4 °C within 24 - 48 hrs, harvested into precipitant solution with 20 - 25 % glycerol for cryo-protection, and flash cooled in liquid nitrogen. Diffraction data were collected at beamline 23-ID-D at the Advanced Photon Source (APS) using an X-ray wavelength of 1.033 Å (360° of data, 100 K, 0.2° image width). PhqB R crystals grew reproducibly, but had generally poor diffraction quality, d_{min} poorer than 4 Å for most crystals. The data used for processing were the best obtained from more than 400 crystals screened. Data were processed with XDS.[16] Attempts to solve the structure by molecular replacement with the similar bacterial NRPS R domain structures succeeded using the MR-ROSETTA[17,18] process in PHENIX.[19] Model building was carried out with Coot.[20] Refinement was carried out with PHENIX.refine.[21] For crystallization of wild-type MalC or SeMet MalC, 12 mg/mL protein stock was mixed with precipitant solution (32% PEG 2K MME, 0.1 M sodium acetate, 0.1 M MES pH 6.5) in a 1:1 v/v ratio. Crystals were grown at 20 °C within 24 - 48 hrs, harvested without additional cryo-protection and flash cooled in liquid nitrogen. Wild-type MalC data were collected at APS beamline 23-ID-D at an X-ray wavelength of 1.033 Å (360° of data, 100 K, 0.2° image width). SeMet data were collected at an X-ray wavelength of 0.979 Å. Data were processed with XDS, and the SeMet MalC crystal structure was solved by single-wavelength anomalous diffraction

(SAD) phasing with AutoSol.[22] Model building was carried out with Coot, and refinement was carried out with PHENIX.refine. For crystallization of PhqE, 10 mg/mL PhqE was mixed with precipitant solution (19% PEG 3350, 150 mM DL-malic acid, 2.5% ethylene glycol, 1 mM premalbrancheamide, 4 mM $NADP^+$, 1% DMSO) in a 1:1 v/v ratio. Crystals were grown at 20 °C for 7 days. A cryo-protectant solution (19% PEG 3350, 150 mM DL-malic acid, 22% ethylene glycol, 1 mM premalbrancheamide, 5 mM $NADP^+$, 1% DMSO, 10 mM HEPES pH 7.5, 50 mM NaCl) was added directly to the crystals prior to flash cooling in liquid nitrogen. For crystallization of PhqE D166N, 10 mg/mL PhqE D166N was mixed with precipitant solution (18% PEG 3350, 200 mM NaCl, 50 mM BisTris pH 6.75, 2.5% ethylene glycol, 1 mM **11**, 4 mM $NADP^+$, 1% DMSO) in a 1:1 v/v ratio. Crystals were grown at 20 °C for 7 days. A cryo-protectant solution (18% PEG 3350, 200 mM NaCl, 50 mM BisTris pH 6.75, 2.5% ethylene glycol, 1 mM **11**, 4 mM $NADP^+$, 1% DMSO, 10 mM HEPES pH 7.5, 50 mM NaCl) was added directly to the crystals prior to flash cooling in liquid nitrogen. Diffraction data were collected at APS beamline 23-ID-B at an X-ray wavelength of 1.033 Å (360° of data, 100 K, 0.2° image width). Data were processed with XDS. The structures were solved by molecular replacement using MalC as the search model. Final models were generated by alternating cycles of manual building in Coot and refinement in PHENIX.refine. The asymmetric unit contained 1.5 tetramers with the intact tetramer (chains A, B, C, D) well-ordered and the half tetramer (chains E and F) poorly packed (Fig. S20). These disordered chains were initially fitted using NCS-averaged maps and refined using NCS restraints to maintain geometry. Chains E and F were not used for structural interpretation and are responsible for the high R-factors of the refined PhqE complex structures. All structures were validated with MolProbity.[23] Multiple sequence alignments were generated from Clustal and Jalview.[24,25] Figures were prepared with PyMOL.[26]

10. Molecular Dynamics Simulations

Molecular dynamics simulations were prepared and equilibrated using the GPU code (*pmemd*)[27] of the AMBER 16 package.[28] Parameters for the ligands were generated within the *antechamber* module using the general AMBER force field (*gaff*),[29] with partial charges set to fit the electrostatic potential generated at the HF/6-31(d) level by the RESP model.[30] The partial charges were calculated according to the Merz–Singh–Kollman scheme[31,32] using the Gaussian 09 package.[33] Each protein was immersed in a pre-equilibrated cubic box with a 10 Å buffer of TIP3P[34] water molecules using the *leap* module, resulting in the addition of around 40,000 solvent molecules. The systems were neutralized by addition of explicit counter ions (Na^+ and Cl^-). All subsequent calculations were done using the Stony Brook modification of the Amber14 force field (*ff14sb*).[35] Water molecules were treated with the SHAKE algorithm such that the angle between the hydrogen atoms was kept fixed. For the heating and equilibration steps, long-range electrostatic effects were modeled using the particle-mesh-Ewald method.[36] An 8 Å cutoff was applied to Lennard–Jones and electrostatic interactions. First, a geometry optimization was performed on each system to minimize the positions of solvent molecules and ions while imposing positional restraints on the protein backbone and ligands using a harmonic potential with a force constant of 2 $kcal \cdot mol^{-1} \cdot Å^{-2}$. Second, each system was gently and continuously heated over 1 ns from 0 K to 300 K under constant-volume and periodic-boundary conditions. Harmonic restraints of 2 $kcal \cdot mol^{-1}$ were applied to the protein backbone and ligands, and the Andersen equilibration scheme was used to control and equalize the temperature. The time step was kept at 1 fs during the heating stages, allowing potential inhomogeneities to self-adjust. Third, each system was then equilibrated for a total of 4 ns at constant pressure of 1 atm with a

Berendson barostat with a 2 fs time step; harmonic restraints of 2 kcal·mol^{-1} were applied for the first 2 ns and harmonic restraints of 0.5 kcal·mol^{-1} were applied for the second 2 ns to the protein backbone and ligands. Finally, production trajectories without harmonic restraints were run on the Anton 2 supercomputer[37] for 1200 ns with a 2.5 fs time step at 300 K and 1 atm using the default NPT integrator and the default u-series treatment of electrostatic interactions.

11. Density Functional Theory Calculations

Density Functional Theory (DFT) calculations were performed with Gaussian 16.[38] Geometry optimizations were carried out with M06-2X functional and the 6-31+G(d,p) basis set.[39] An atom-pairwise density-independent D3 dispersion correction was applied with zero-damping throughout.[40] Experience has shown this level of theory to be well-suited to the study of Diels-Alder thermochemistry and kinetics, as well as non-covalent interactions: a comparison of this functional against benchmark values for the binding energies of 1744 non-covalent dimers gives an RMSD of 0.43 kcal/mol. A similar comparison using the barrier heights of 206 reactions, including pericyclic transformations, gives an RMSD of 2.57 kcal/mol.[41] Harmonic vibrational frequencies at the same level of theory were used to assign stationary points as either minima or transition structures (TSs) and to calculate the zero-point vibrational energy and thermal corrections. For these intramolecular reactions the choice of standard concentration is irrelevant. We obtained vibrational entropies using a quasi-harmonic approximation, treating vibrational modes below 100 cm^{-1} as free rotors and as rigid rotors above this cut-off using the GoodVibes program.[42,43] Molecular graphics were produced with PyMOL.[26]

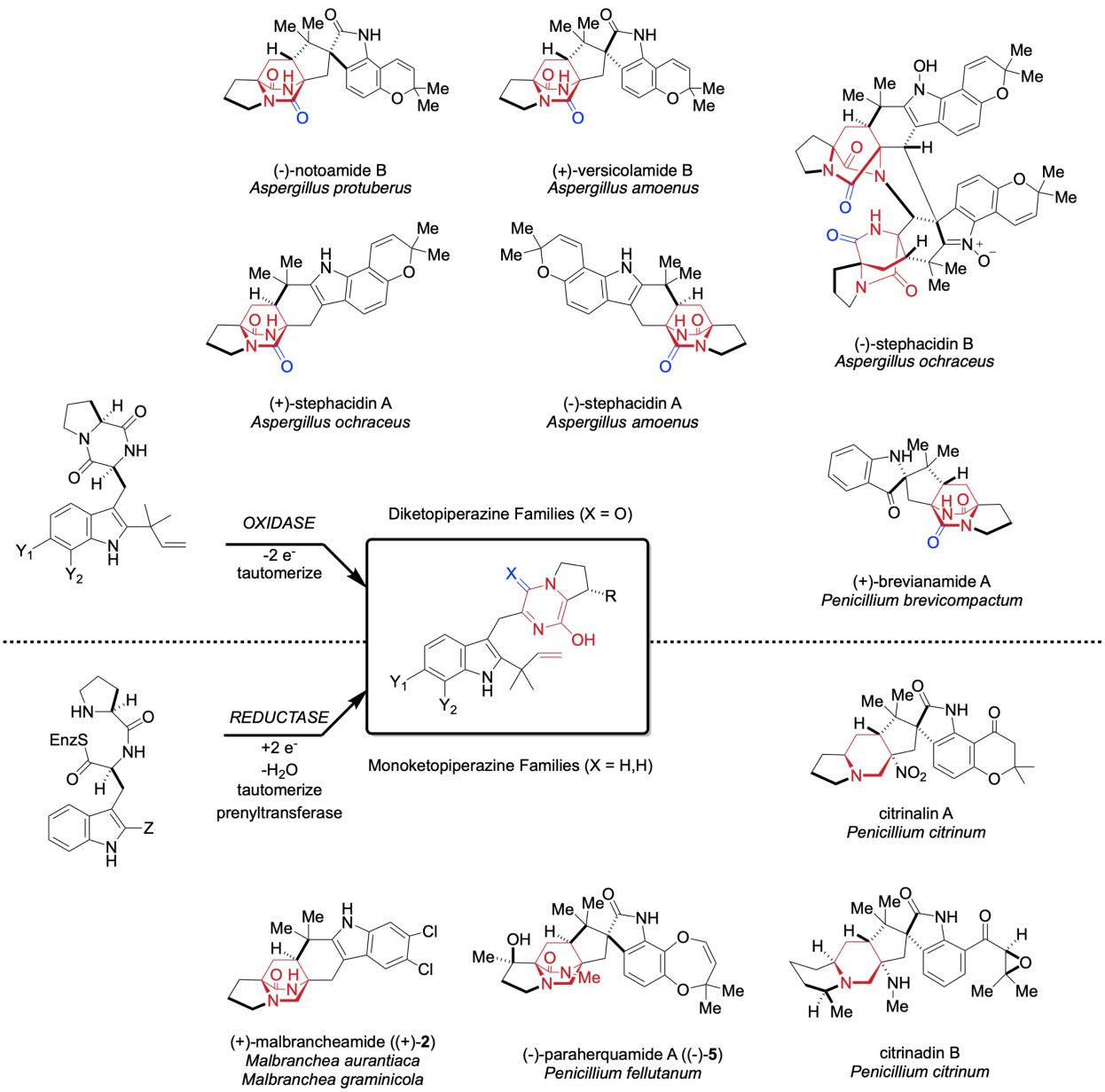
Notably, the conformation of the pyrrolidine ring in the most stable *syn*- and *anti*-TS differ for the monoketopiperazine substrate. This is not the case for the diketopiperazine substrate, since

the planarity of the azadiene is enforced by the sp^2-hybrization of all carbons in this heterocyclic ring. The introduction of a CH_2-group in the azadiene ring enables different ring-puckering modes of the pyrrolidine ring. The energetic preference for the *anti*-TS is enhanced for this substrate since unfavorable steric contacts between dienophile and the pyrrolidine can be minimized.

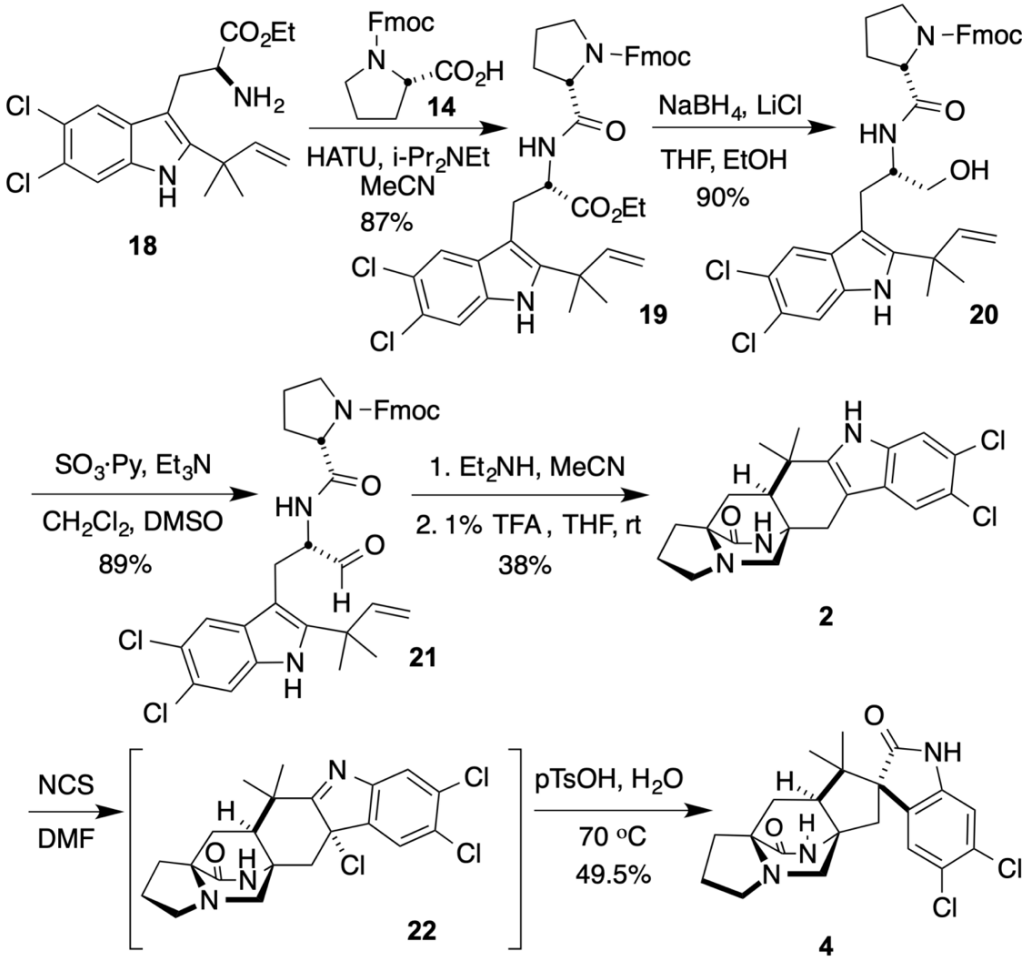
12. Genetic Disruption of *phqE*

The gene disruption in *Penicillium simplicissimum* was performed by the CRISPR/Cas9 system for filamentous fungi.[44] For the preparation of the *in vitro* transcriptional gRNA, the gRNA cassettes containing the T7 promoter, the protospacer sequence, and the synthetic gRNA scaffold for targeting genes were PCR amplified from the plasmid pFC333 as template, using the primers listed in Supplementary Table 1, and inserted into pFC332 to generate the plasmid pFC332-*phqE*. For transformation of *Penicillium simplicissimum*, the strain was inoculated into 100 mL YPD medium and cultivated at 28 °C, 200 rpm, for 2 days. The mycelia were collected and digested using vinoflow (64 mg/mL). The resulting protoplasts were then separated from mycelia by filtration and washed with STC solution (0.8 M sorbitol, 0.05 M Tris-HCl, 0.05 M $CaCl_2$, pH 8), and diluted to a concentration of 2×10^8 cells mL^{-1}. Then, the circular plasmid was added to the 200 μL protoplasts solution, and incubated on ice for 30 min, which was blended with 2 mL 30% PEG solution (40% PEG8000, 50 mM $CaCl_2 \cdot 2H_2O$, 10 mM Tris-HCl, pH 8.0) and incubated at the room temperature for 20 min. The resulting solution was then diluted with STC solution and distributed on selective PGA plates (PG broth, 1.2 M sorbitol, 100 μg/mL hygromycin B, 1.5% agar) The plates were incubated at 30 °C for 5 - 7 days. The colonies grown from these selective

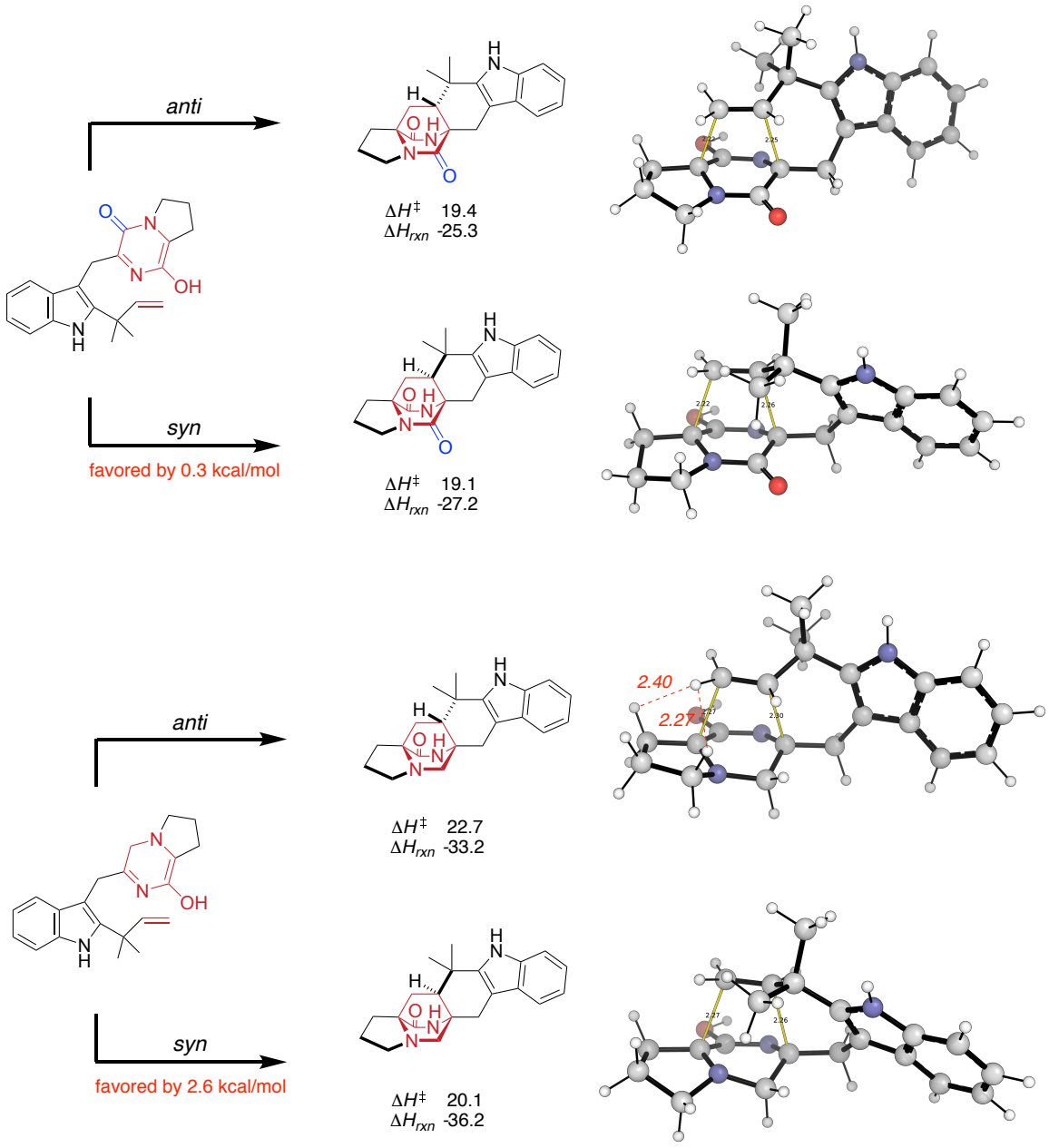
plates were cultured (stationary) in CYA medium (1L containing Difco Czapek-Dox 35 g, yeast extract 5 g, $CuSO_4 \cdot 5H_2O$ 5 mg, $ZnSO_4 \cdot 7H_2O$; pH 6.3) for 7 days and analyzed in TOF-MS.



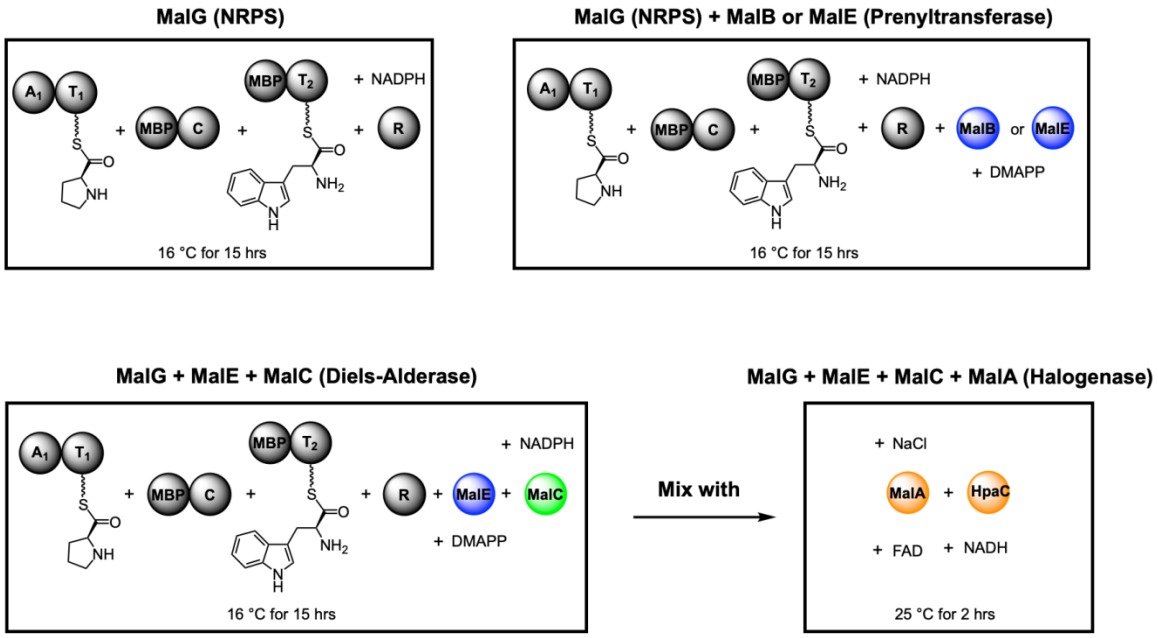
Supplementary Figure 1. Unified biogenesis of the diketopiperazine and monoketopiperazine families of alkaloids possessing the bicyclo[2.2.2]diazaoctane core structures. The biocyclo[2.2.2]diazaoctane group is colored in red. The second keto group of diketopiperazines is highlighted in blue, and arises from the terminal condensation domain of the biosynthetic NRPS that generates a diketopiperazine intermediate. In the monoketopiperazine biosynthetic pathways, a terminal reductive domain exists in the biosynthetic NRPS and generates a Pro-Trp aldehyde intermediate.



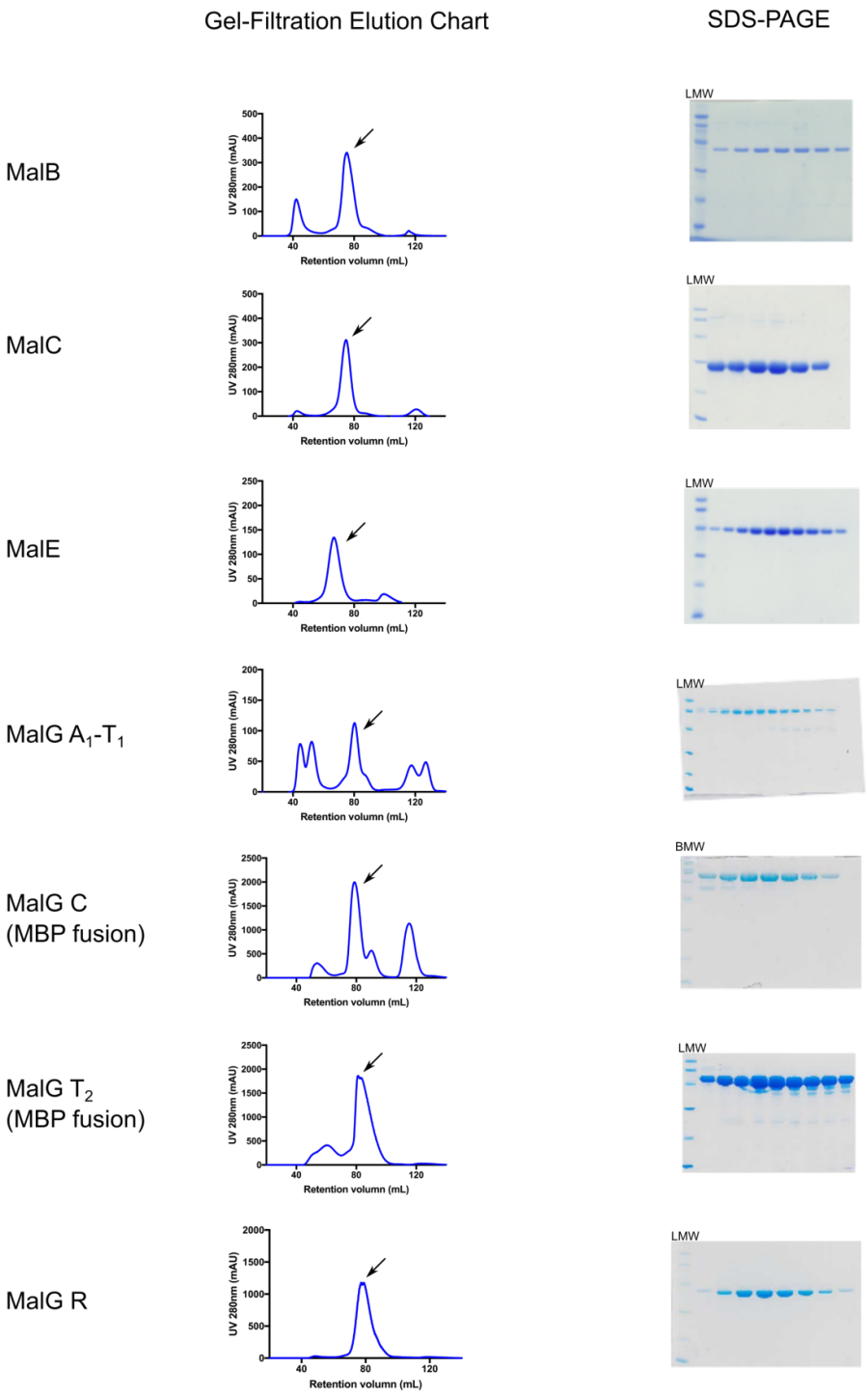
Supplementary Figure 2. Biomimetic synthesis of racemic malbrancheamide and spiromalbramide. We have applied an analogous strategy to two additional natural products, racemic (±)-malbrancheamide (**2**) and (±)-spiromalbramide (**4**) that underscores the utility of this new biomimetic paradigm. The key halogenated Fmoc-protected amino aldehyde (**21**) was prepared by peptide coupling of the reverse prenylated tryptophan methyl ester (**18**) with Fmoc-protected proline amino acid (**14**) using HATU (87% yield). The ethyl ester was reduced with sodium borohydride (**20**; 90%) and followed by a Parikh–Doering oxidation to provide the N-Fmoc aldehyde **21** in 89% yield. The Fmoc group was removed with diethylamine, and the crude product was directly treated with a degassed solution of 1% TFA in THF at room temperature to provide malbrancheamide (racemic) **2** in 38% yield. **2** was treated with N-chlorosuccinimide to form the incipient chloroindoline intermediate (**22**), which was directly hydrated under acidic conditions to undergo a pinacol-type rearrangement and form spiromalbramide (racemic) **4** in 49.5% yield. Including the four steps required to synthesize the reverse prenylated tryptophan species **18**, the synthesis of malbrancheamide was achieved in eight steps from commercially available materials and only four steps in the longest linear sequence; one additional transformation (two steps, one operation) being required to reach spiromalbramide.



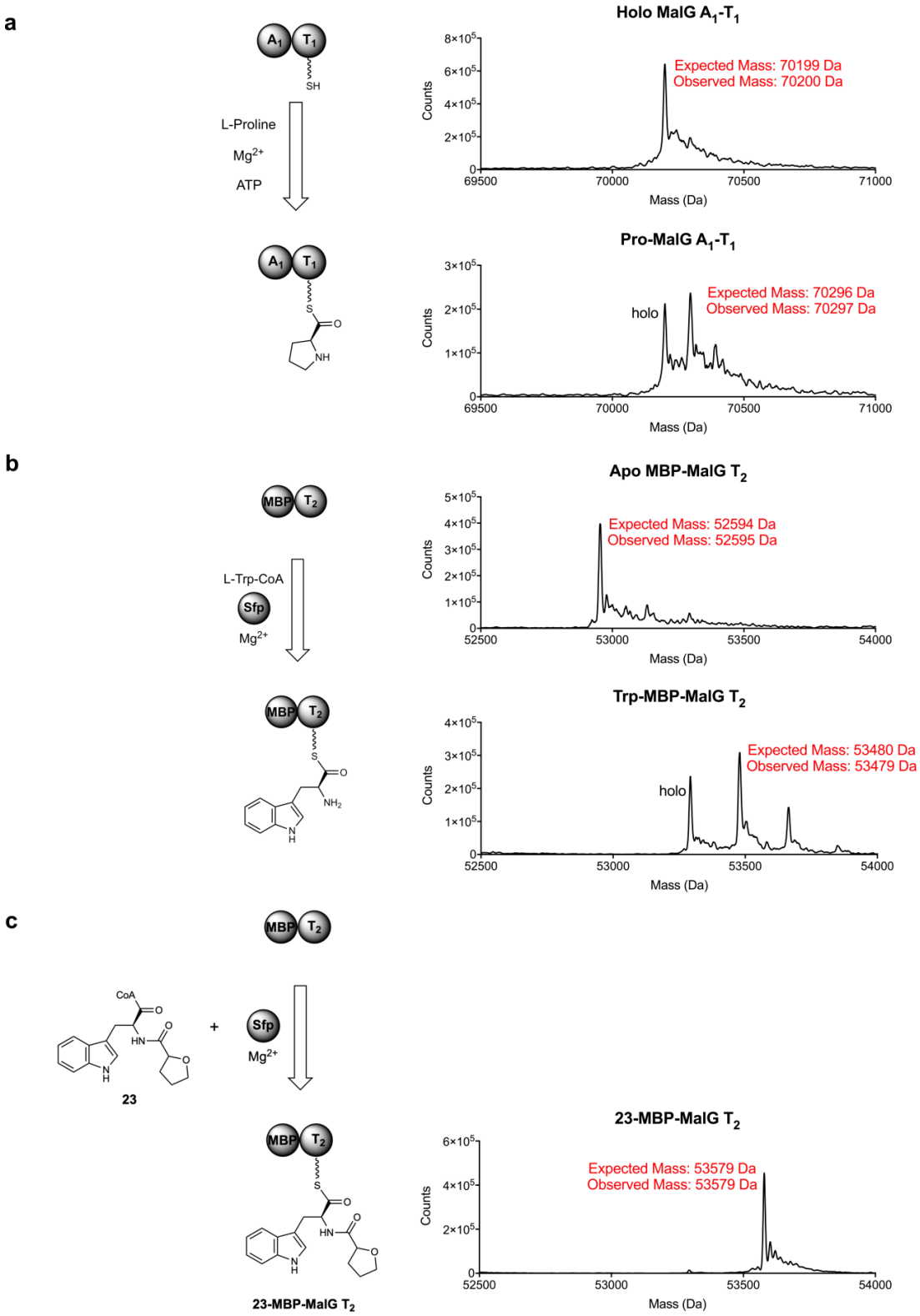
Supplementary Figure 3. M062X-D3/6-31+G(d,p) intramolecular Diels-Alder transition structures (TSs) for both oxidation states represented by the azadiene, predict a relatively modest *syn*-: *anti*-diastereoselectivity for the more oxygenated azadiene species of 0.3 kcal/mol.[45] This value has been corroborated experimentally in several systems where the *syn*- : *anti*-ratio is typically around 2.5:1. The reduced azadiene species has a more substantial TS difference of 2.6 kcal/mol favoring the *syn*-cycloadduct. The pyrrolidine ring adopts different conformations in these *syn*- and *anti*- TSs, puckering towards the dienophile in the less-favorable structure and resulting in short H---H contacts. Consistent with greater levels of selectivity, the only detectable diastereomeric cycloadducts were the *syn*-diastereomers by comparison with authentic, synthetic samples of the corresponding *anti*-diastereomers.



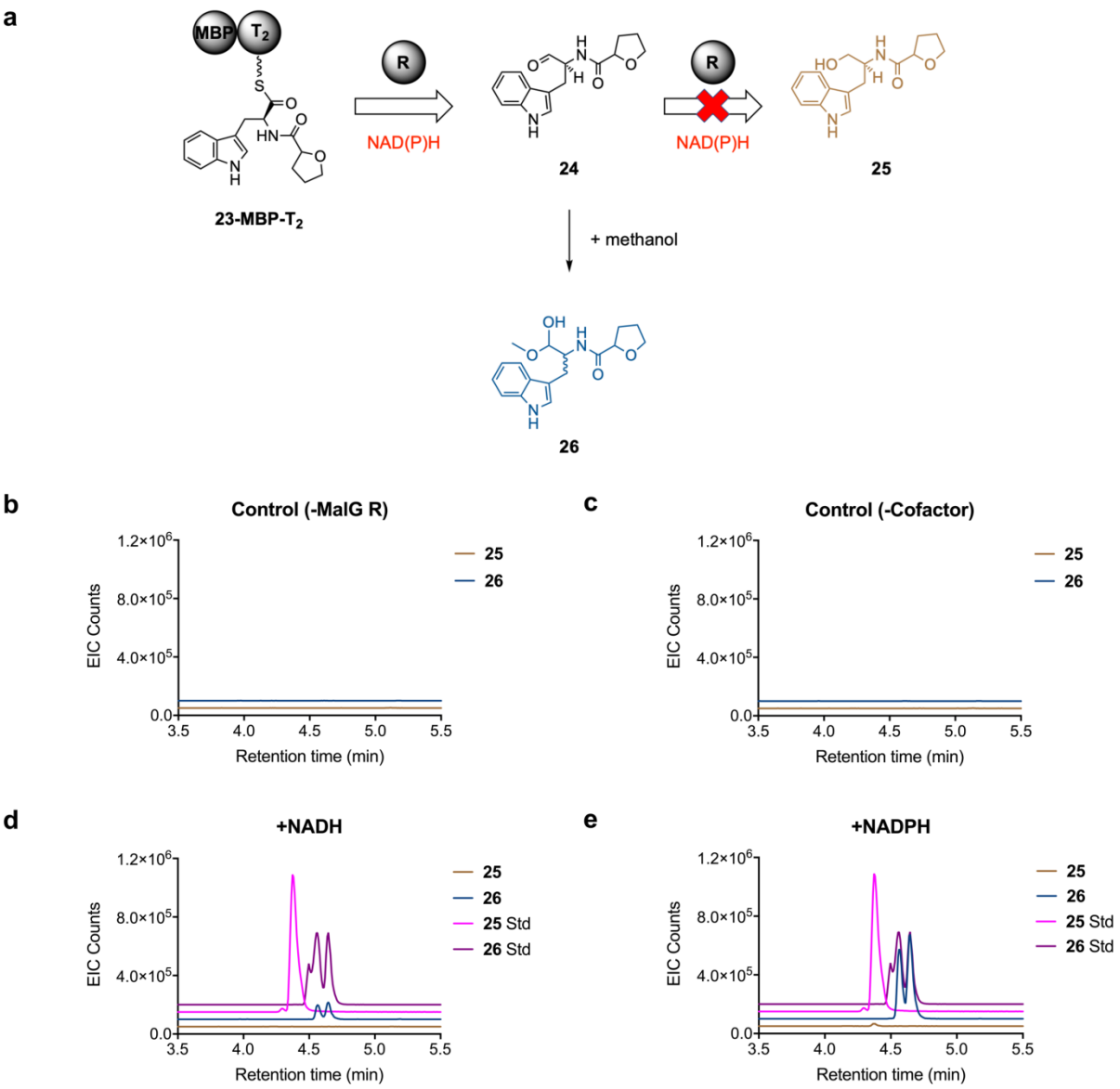
Supplementary Figure 4. Scheme of *in vitro* reconstitution assays. The flavin-dependent MalA halogenase requires a recycling system to reduce FAD to $FADH_2$ after each catalytic cycle, here HpaC reductase + NADH.[1,46] MBP = maltose binding protein. The domain boundaries for MalG used in the assays are: A_1-T_1 (198-838), C (846-1277, with N-terminal MBP fusion), T_2 (1841-1925, with N-terminal MBP fusion), and R (1932-2345).



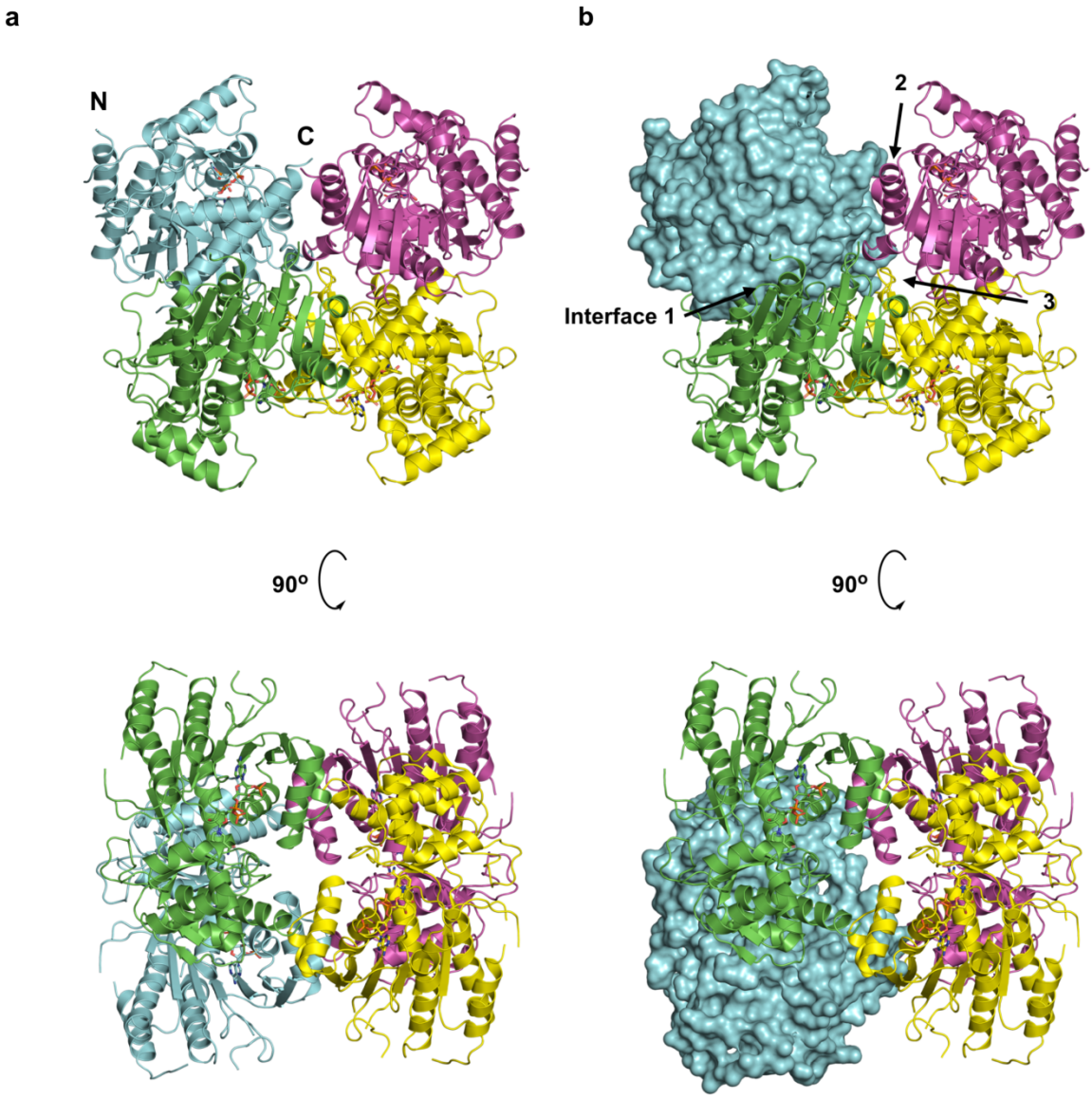
Supplementary Figure 5. Purification of enzymes and enzymatic domains involved in malbrancheamide biosynthesis with gel filtration profiles at left and SDS gels of the indicated peaks at right. Recombinant MalA was produced as previously described.[1] Protein molecular weight standards were LMW: 97.4, 66.2, 45.0, 31.0, 21.5, 14.4 kDa; and BMW: 200, 116.3, 97.4, 66.2, 45.0, 31.0, 21.5, 14.4, 6.5 kDa. These experiments were repeated independently with similar results for more than three times.



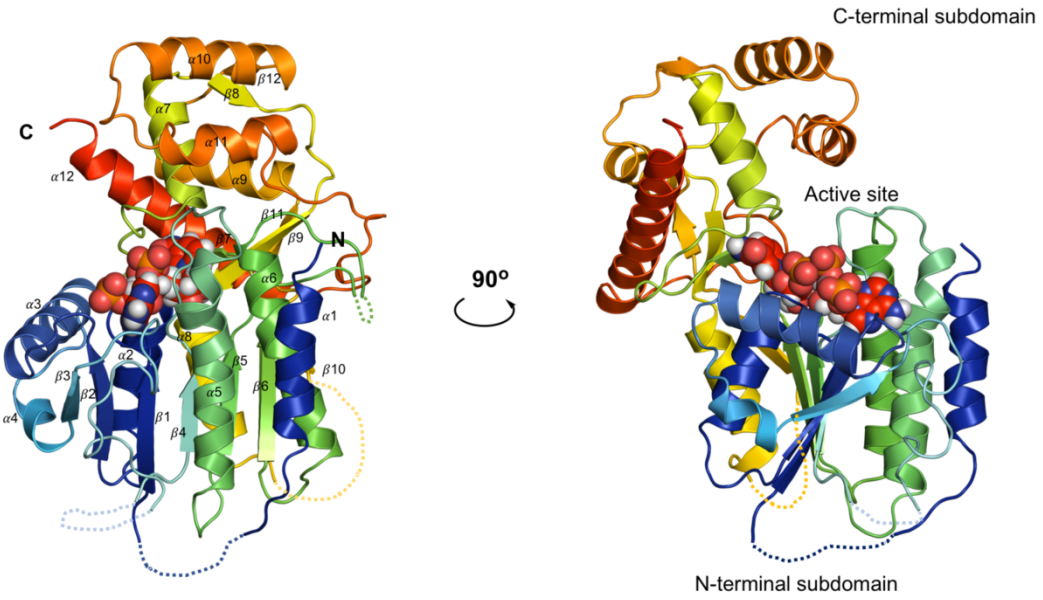
Supplementary Figure 6. Substrate loading of MalG T domains. Protein mass spectrometry was applied to analyze efficiency of substrate loading, confirming successful loading in all cases. These experiments were repeated independently with similar results for three times.



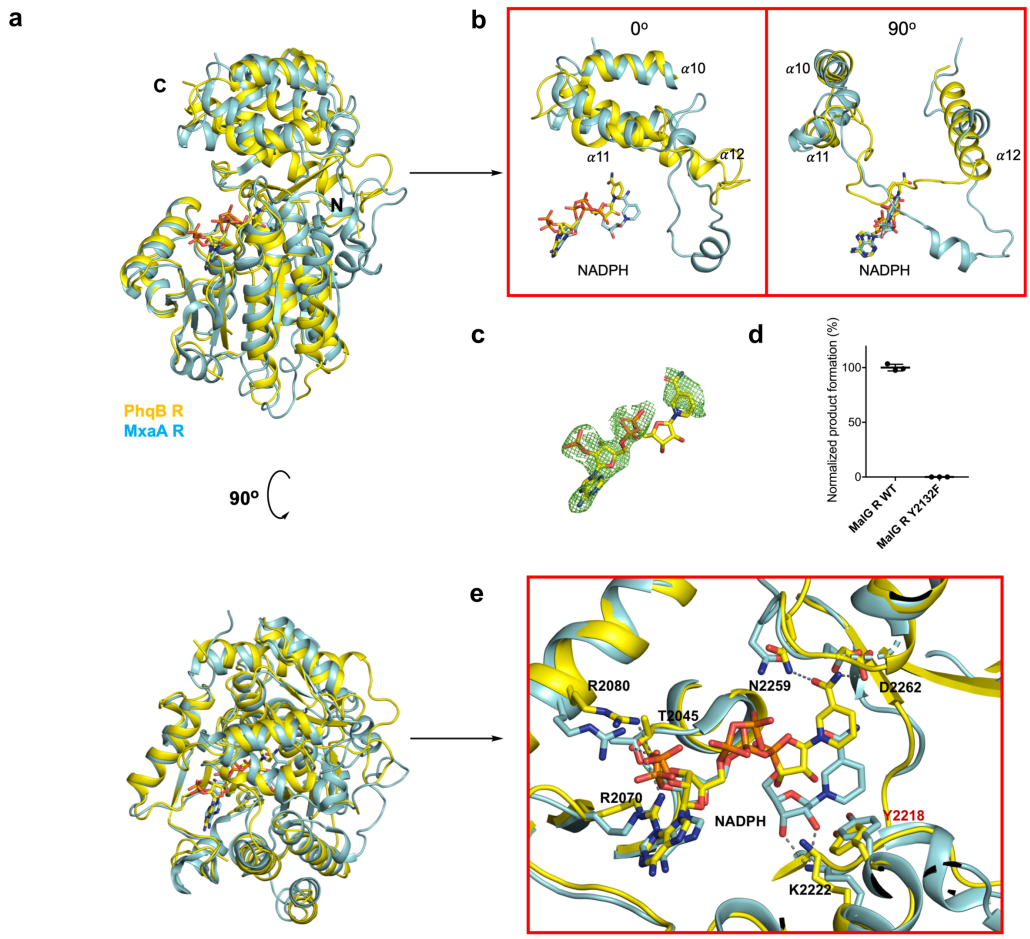
Supplementary Figure 7. MalG R catalyzes a 2-electron reductive release reaction. a. Reaction scheme of MalG R. b – c. EIC profiles of control experiments, with no enzyme (b) or no cofactor (c). d – e. EIC profiles of the MalG R-catalyzed reaction, using NADH (d) or NADPH (e) as cofactor. NADPH is the preferred cofactor. **25**, the product of a 4-electron reduction was not detected. Experiments in b – e were repeated independently with similar results for three times.



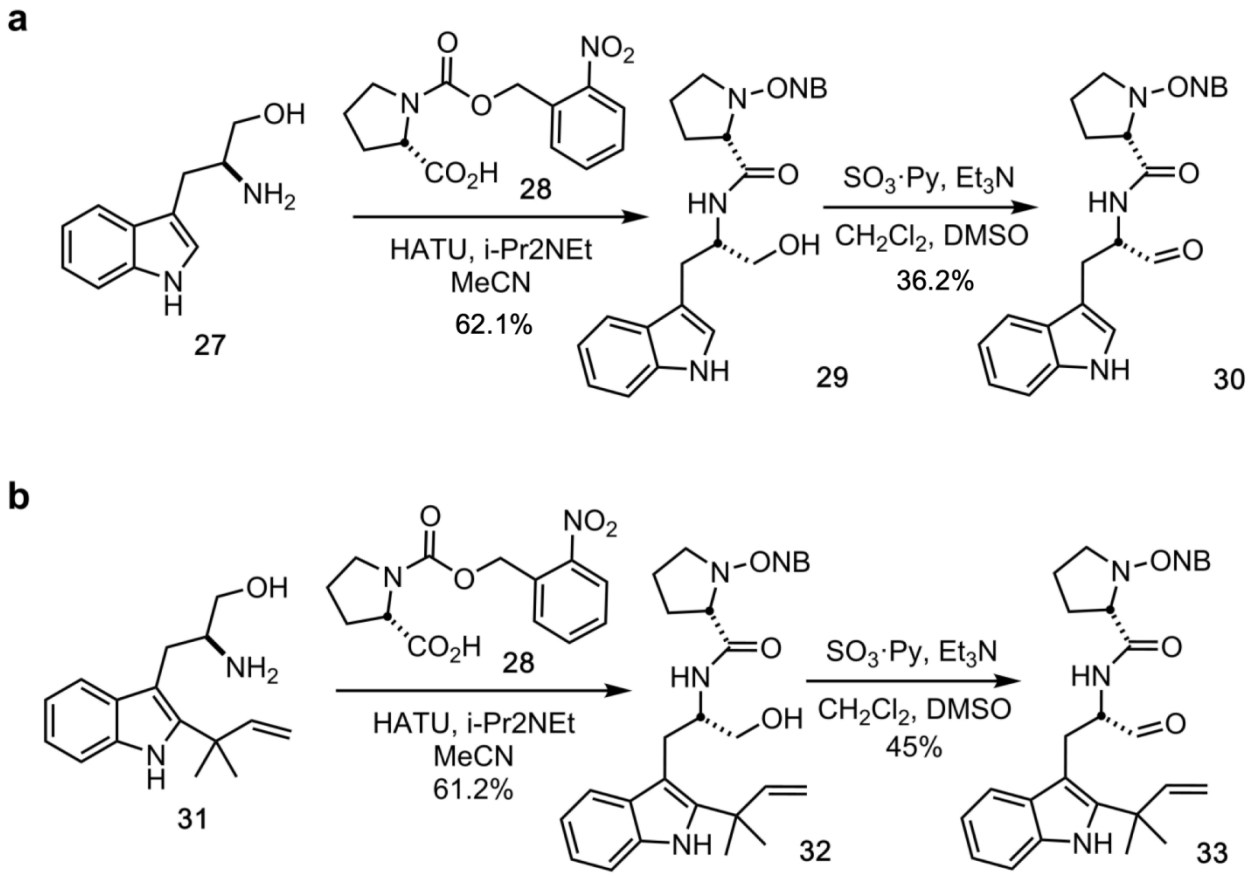
Supplementary Figure 8. Crystal structure of the PhqB R tetramer, with the four subunits shown in contrasting colors cartoon representation in (a), and the cyan subunit shown in surface representation in (b). The N- and C-termini are marked for the cyan subunit. The excised PhqB R tetramer has D2 point symmetry, and each subunit contacts all three other subunits. The tetrameric oligomer state, apparently inherited from short-chain dehydrogenase/reductase (SDR) ancestors, differs from the generally monomeric NRPS situation. However, the four N-termini that link to the rest of MalG are at the exterior of the tetramer and well separated from one another in an arrangement that would allow flexible tethering of a "monomeric" NRPS module.



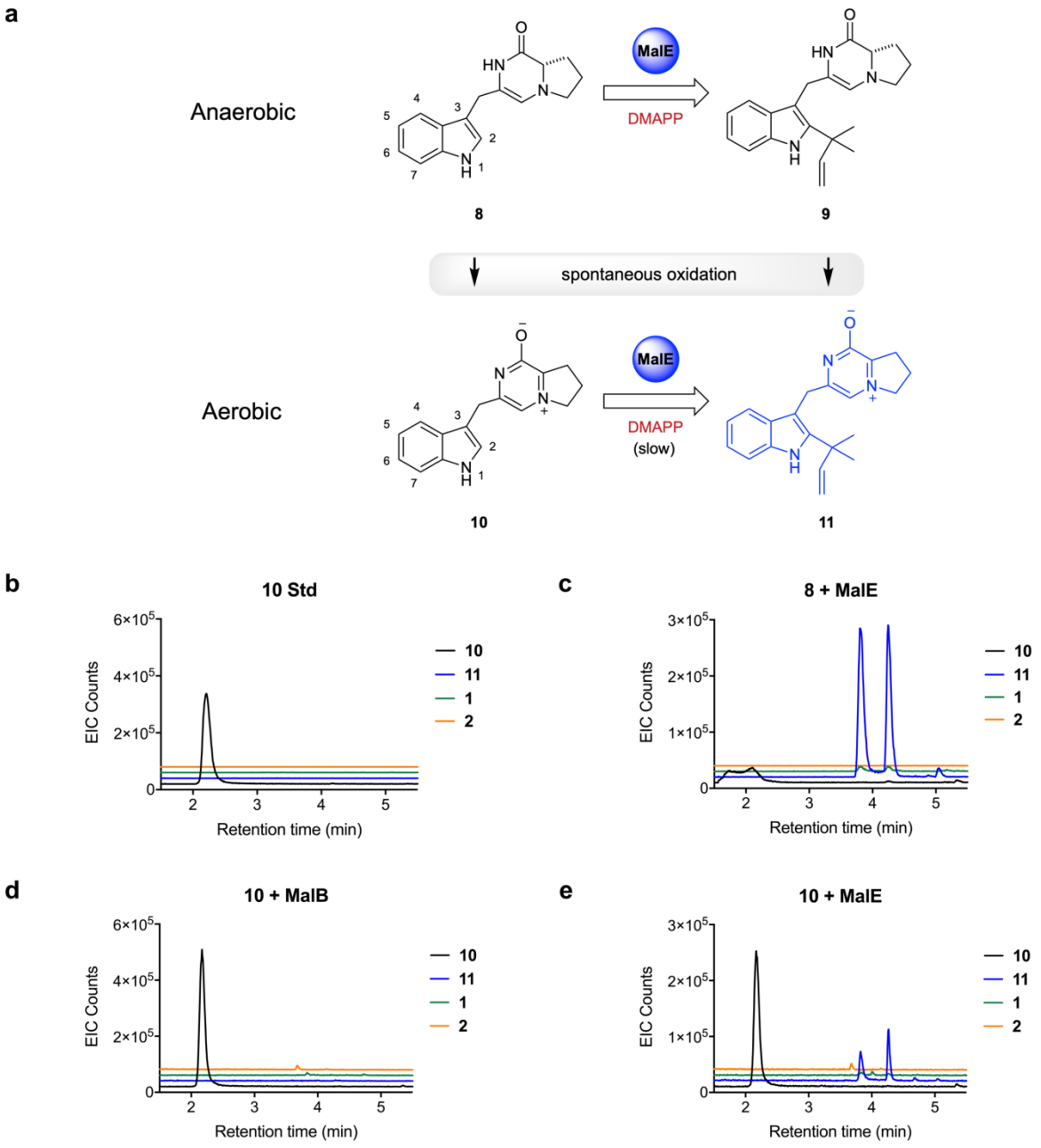
Supplementary Figure 9. Structure of the PhqB R subunit colored as a rainbow from blue N-terminus to red C-terminus. PhqB R consists of an N-terminal nucleotide-binding subdomain and a C-terminal substrate-binding subdomain, which recognizes Pro-Trp-T_2. The nucleotide-binding subdomain has a typical Rossmann fold, with a parallel β sheet (β1, β2, β3, β4, β5, β6 and β10) flanked by six α helices (α2, α3, α4, α5, α6 and α8) and an invariant "TGX_3GXG" motif (P-loop), as well as conserved Arg2070 and Arg2080, which coordinate the adenosine 2'-phosphate and account for the selectivity of NADPH (shown as spheres) over NADH. The C-terminal subdomain covers the active site, is unique to NRPS terminal reductases, and is composed of five α helices (α7, α9, α10, α11 and α12).



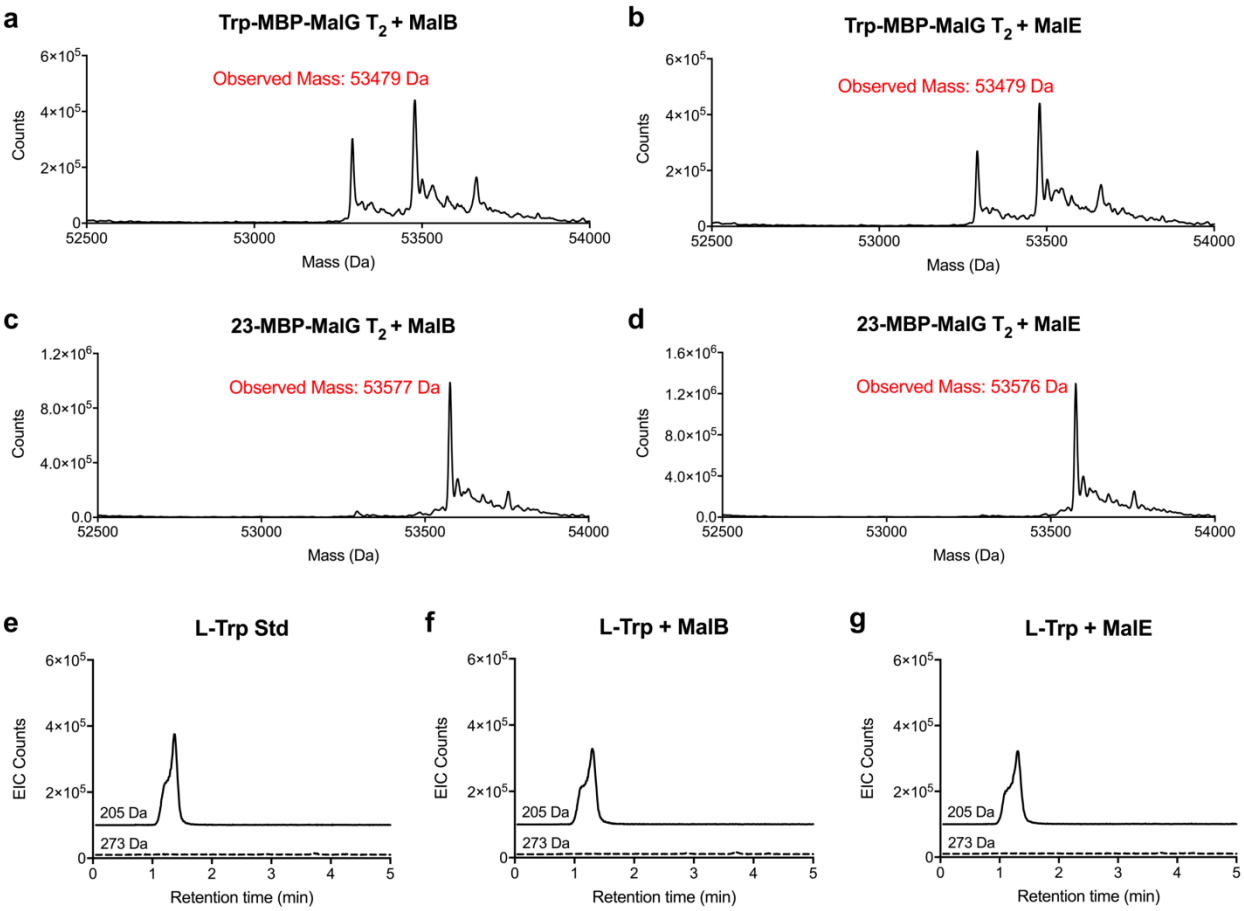
Supplementary Figure 10. Comparison of fungal and bacterial NRPS R domains. a. Superposition of the PhqB R subunit (yellow) and the bacterial MxaA R[47] (cyan, PDB ID: 4U7W). b. The major structural difference lies in the C-terminal subdomain. Relative to bacterial NRPS R domain structures, PhqB R α12 is tilted towards the core with a significantly shorter preceding loop that lacks a short helix. In the PhqB R tetramer this is the site of a subunit contact, which does not exist in the bacterial R domain. c. Electron density map for the PhqB NADPH cofactor (F_o-F_c omit contoured at 3σ). In the co-crystal structure of PhqB R and NADPH, the nicotinamide ring of NADPH is poorly resolved and partially occupies a non-catalytic position. This is in contrast to bacterial NRPS R domains and is correlated with strikingly different structures for α11-loop-α12 in the fungal and bacterial R domains. A poorly ordered nicotinamide also occurs in bacterial modular polyketide synthase (PKS) B-type ketoreductases.[48] Substrate may be required for optimal cofactor binding. d. Lack of detectable activity in MalG R/Y2132F (center values, mean; error bars, SD; $n = 3$). e. Active site detail. The active site contains conserved residues Tyr2218 and Lys2222, suggestive of a shared reaction mechanism with bacterial NRPS terminal R domains[47,49,50] and other Tyr-dependent SDRs[51,52] in which a catalytic Tyr serves as a proton donor and a catalytic Lys facilitates proton transfer. The catalytic Tyr is labeled in red. In the PhqB R, the nicotinamide is in a non-catalytic position away from Tyr2218 and Lys2222, but is hydrogen bonded to conserved Asn2259 and Asp2262.



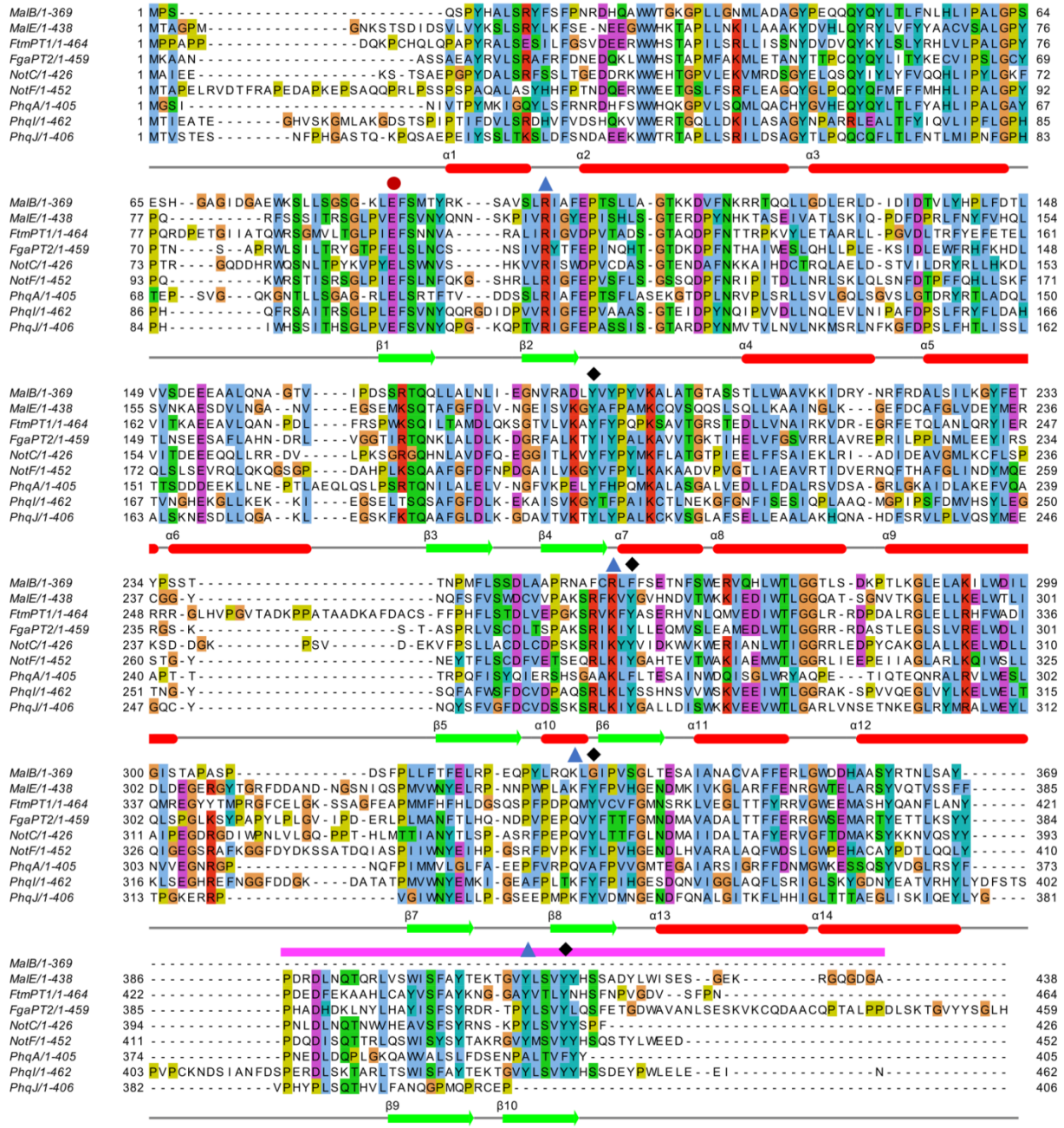
Supplementary Figure 11. Synthetic scheme for the ONB protected dipeptide aldehyde **30** (a) and ONB protected prenyl dipeptide **33** (b).



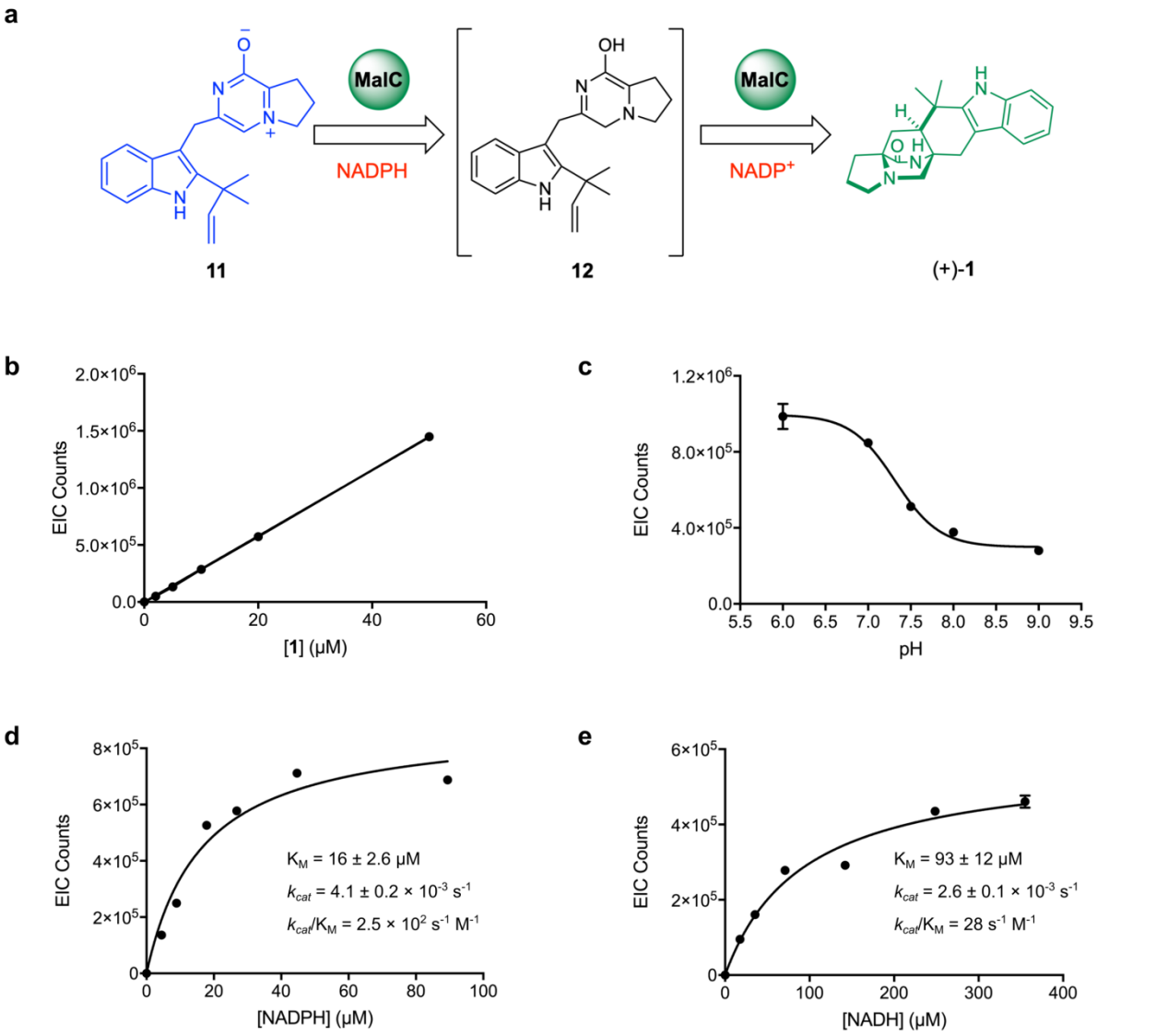
Supplementary Figure 12. Prenylation of **8** (anaerobic) and **10** (aerobic). **8** is the natural substrate of MalE. **10** can be prenylated by MalE but not MalB. a. C2 reverse prenyltransfer reaction scheme. b. EIC profile of **10** authentic standard. c. EIC profile of **8** prenylation by MalE in anaerobic conditions. Spontaneous oxidation of **9** to **11** occurred when the reaction mixture was subjected to LC/MS analysis. d. EIC profile of **10** prenylation by MalB in aerobic conditions. No prenylated product was detected. e. EIC profile of **10** prenylation by MalE, illustrating that oxidized **10** is less favored than **8**. Experiments in b – e were repeated independently with similar results for three times.



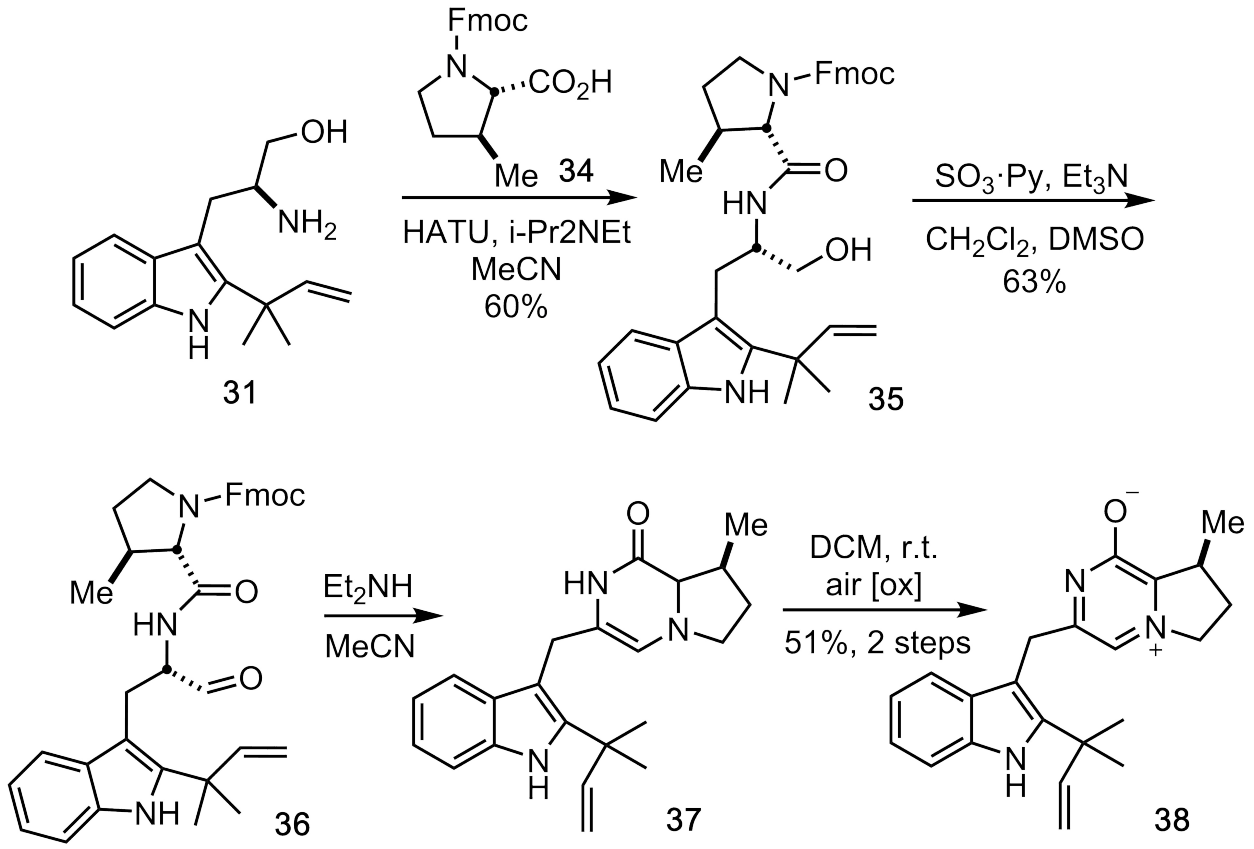
Supplementary Figure 13. Timing of prenylation reaction after NRPS reductive offloading. a. Protein MS profile of L-Trp-MBP-MalG T_2 after 2 hours of MalB and DMAPP incubation. b. Protein MS profile of L-Trp-MBP-MalG T_2 after 2 hours of MalE and DMAPP incubation. See Fig. S6b for protein MS profile of L-Trp-MBP-MalG T_2 control. c. Protein MS profile of **23**-MBP-MalG T_2 after 2 hours of MalB and DMAPP incubation. d. Protein MS profile of **23**-MBP-MalG T_2 after 2 hours of MalE and DMAPP incubation. See Fig. S6c for protein MS profile of **23**-MBP-MalG T_2 control. In all cases, no protein mass change was observed, showing no prenylation. e. EIC profile of L-Trp authentic standard. f. EIC profile of L-Trp after 2 hours of MalB and DMAPP incubation. g. EIC profile of L-Trp after 2 hours of MalE and DMAPP incubation. In all cases, no prenylation was detected (L-Trp $M+H^+$ *m/z* = 205; Prenylated L-Trp $M+H^+$ *m/z* = 273), demonstrating that MalG NRPS functions as the first enzyme in the malbrancheamide pathway. All experiments were repeated independently with similar results for three times.



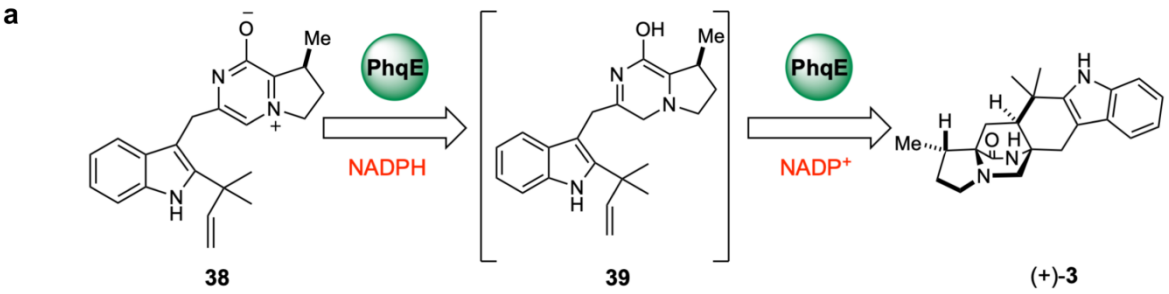
Supplementary Figure 14. Multiple sequence alignment of fungal indole prenyltransferases. The catalytic base Glu is highlighted with a red dot. Residues critical for coordinating the cofactor DMAPP are highlighted with blue triangles. Four Tyr that are expected to shield the active site are highlighted with black squares. MalE contains a full set of conserved residues, while MalB does not. The C-terminal sequence that MalB lacks (magenta) may include the last two β strands of the prenyltransferase barrel, possibly contributing to inefficiency of MalB catalysis.



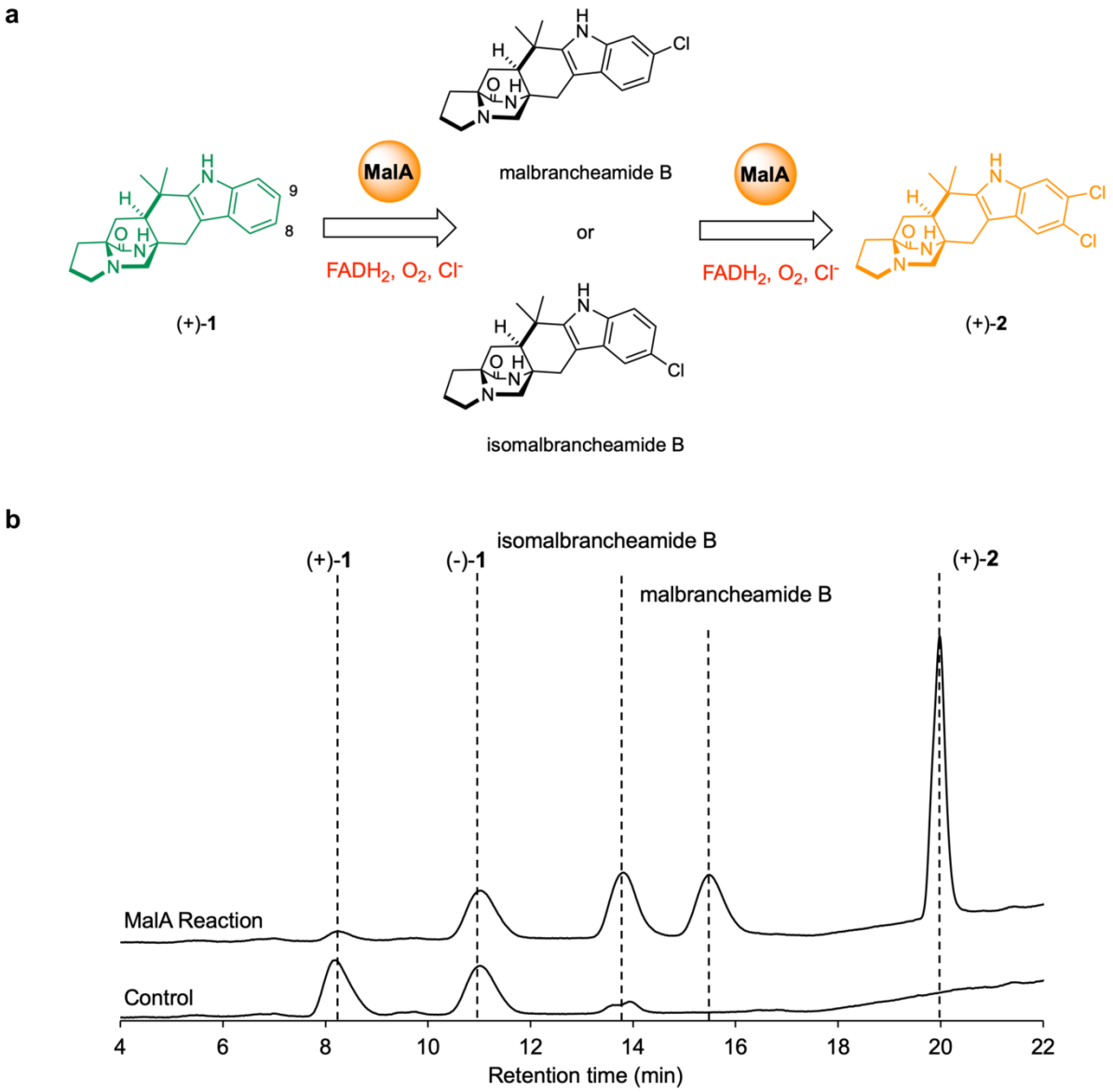
Supplementary Figure 15. a. MalC catalyzes a two-step reaction of **11** to (+)-**1**. b. Standard curve of (+)-**1**, presenting linear correlation of EIC counts to (+)-**1** concentration. c. pH profile of the MalC-catalyzed reaction, efficiency of which decreases beyond neutral pH. d – e. K_M measurement of NADPH (d) and NADH (e) for MalC catalysis. For all measurements in b – e, results were repeated three times (center values, mean; error bars, SD; $n = 3$).



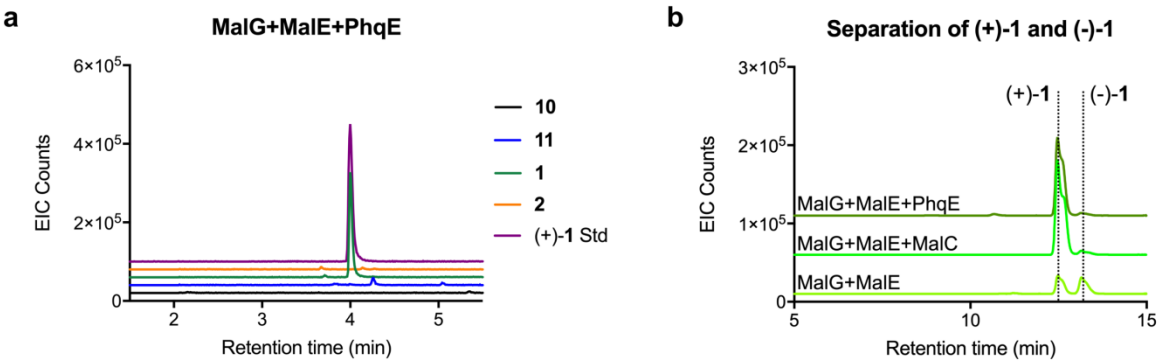
Supplementary Figure 16. Synthetic scheme of *β*-methyl prolyl prenyl zwitterion **38**, an intermediate in paraherquamide biosynthesis.



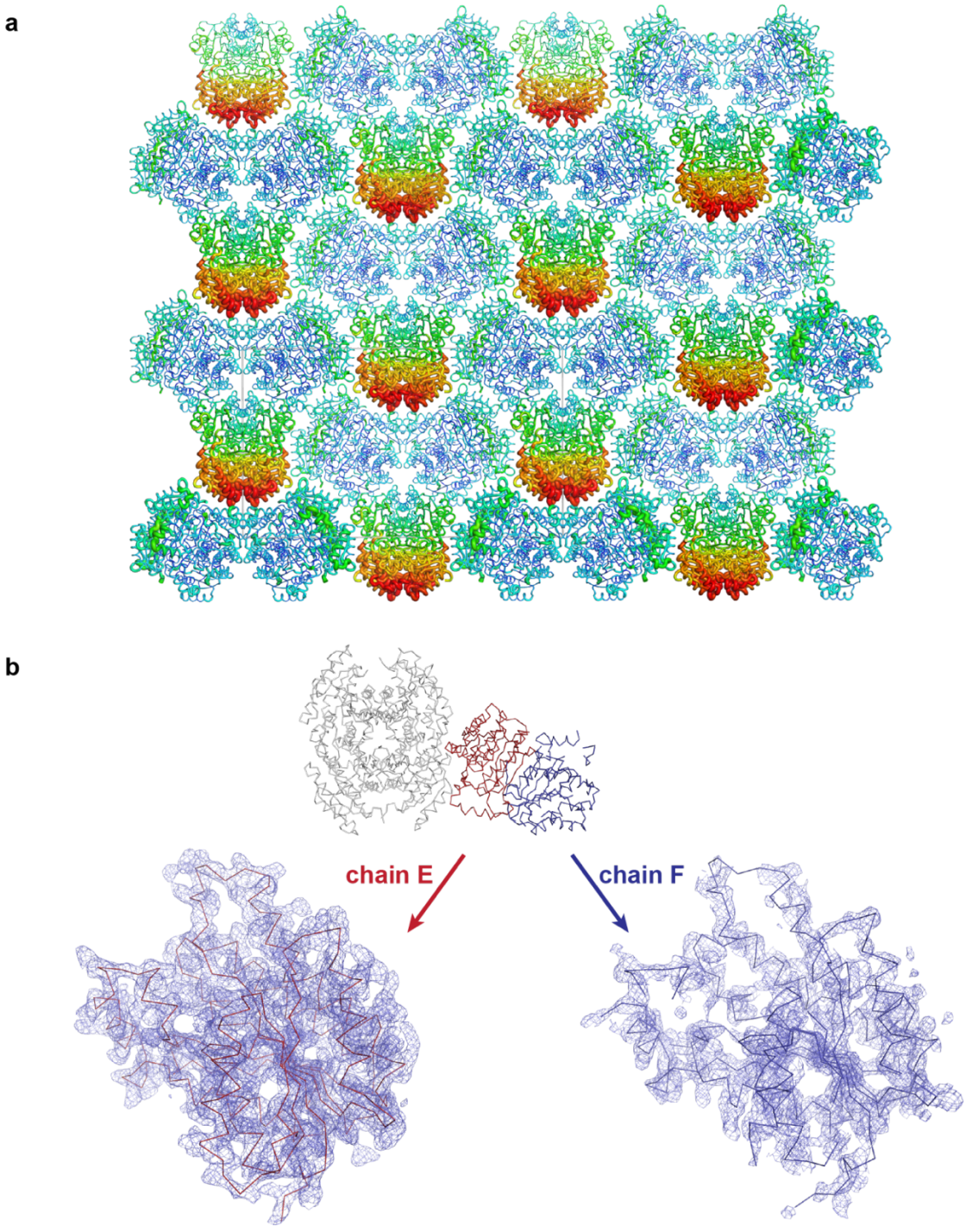
Supplementary Figure 17. Production analysis of **38** ($M+H^+$ m/z = 348) by TOF-MS from *Penicillium simplicissimum phqE* mutant. *phqE* is the homologous gene of *malC* in the paraherquamide biosynthetic pathway. a. PhqE catalyzes a two-step reaction of **38** to (+)-preparaherquamide **3**. b. *In vivo* production of **38** via *phqE* knockout. The EIC traces (from bottom to top) are: 1) *Penicillium simplicissimum* WT extracts; 2) *Penicillium simplicissimum phqE* knockout mutant extracts; 3) **38** authentic standard; 4) Co-injection of *phqE* mutant extracts and **38** standard. All peaks of **38** are highlighted with dashed lines. **38** was split into two or three peaks under our analysis condition, possibly due to tautomerization. All experiments were repeated independently with similar results for three times.



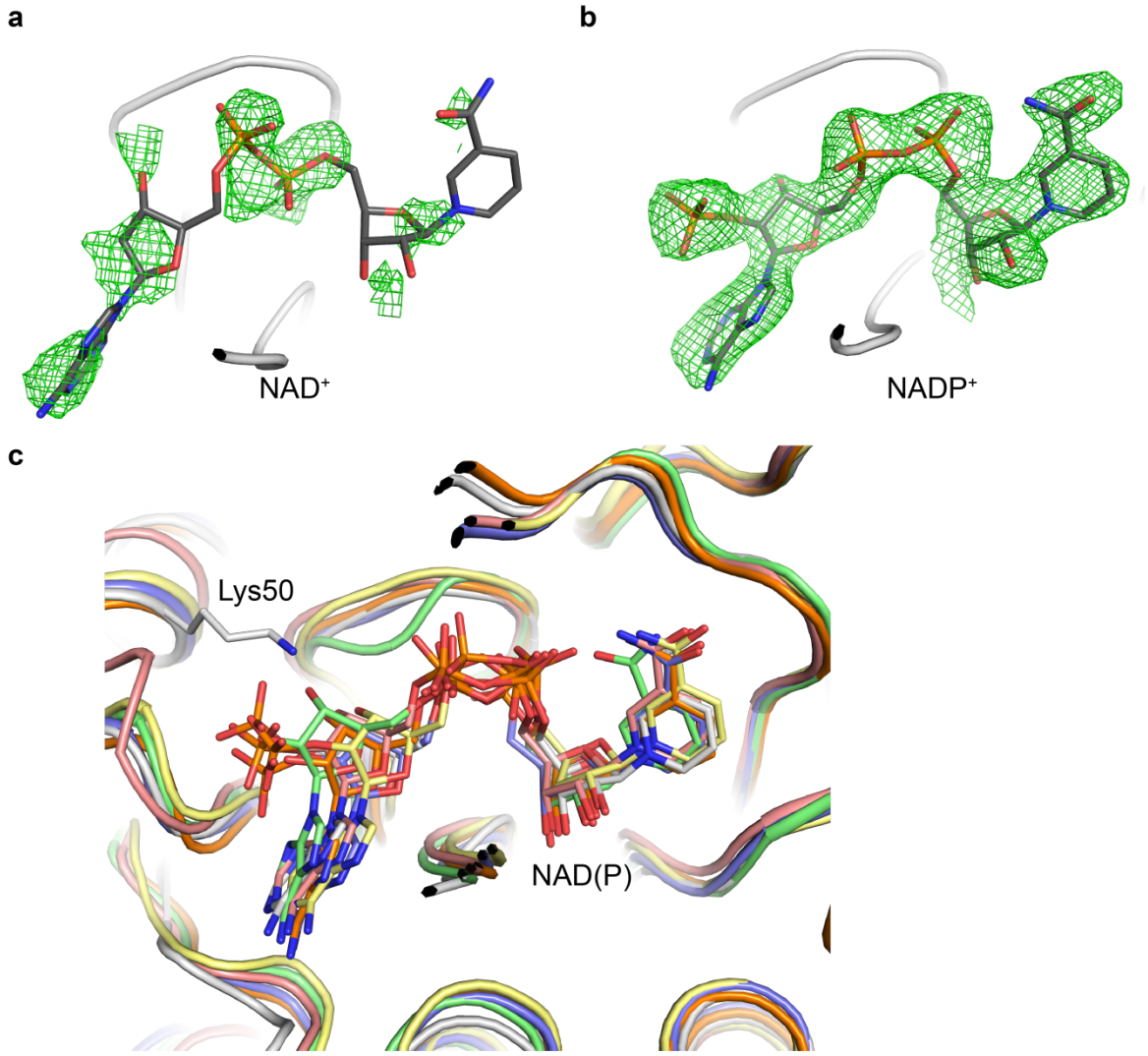
Supplementary Figure 18. a. MalA catalyzes an iterative dihalogenation reaction, converting (+)-**1** to (+)-**2**. b. MalA is stereospecific and does not react on (−)-**1**. The Y-axis is UV 240 nm absorbance. All experiments were repeated independently with similar results for three times.



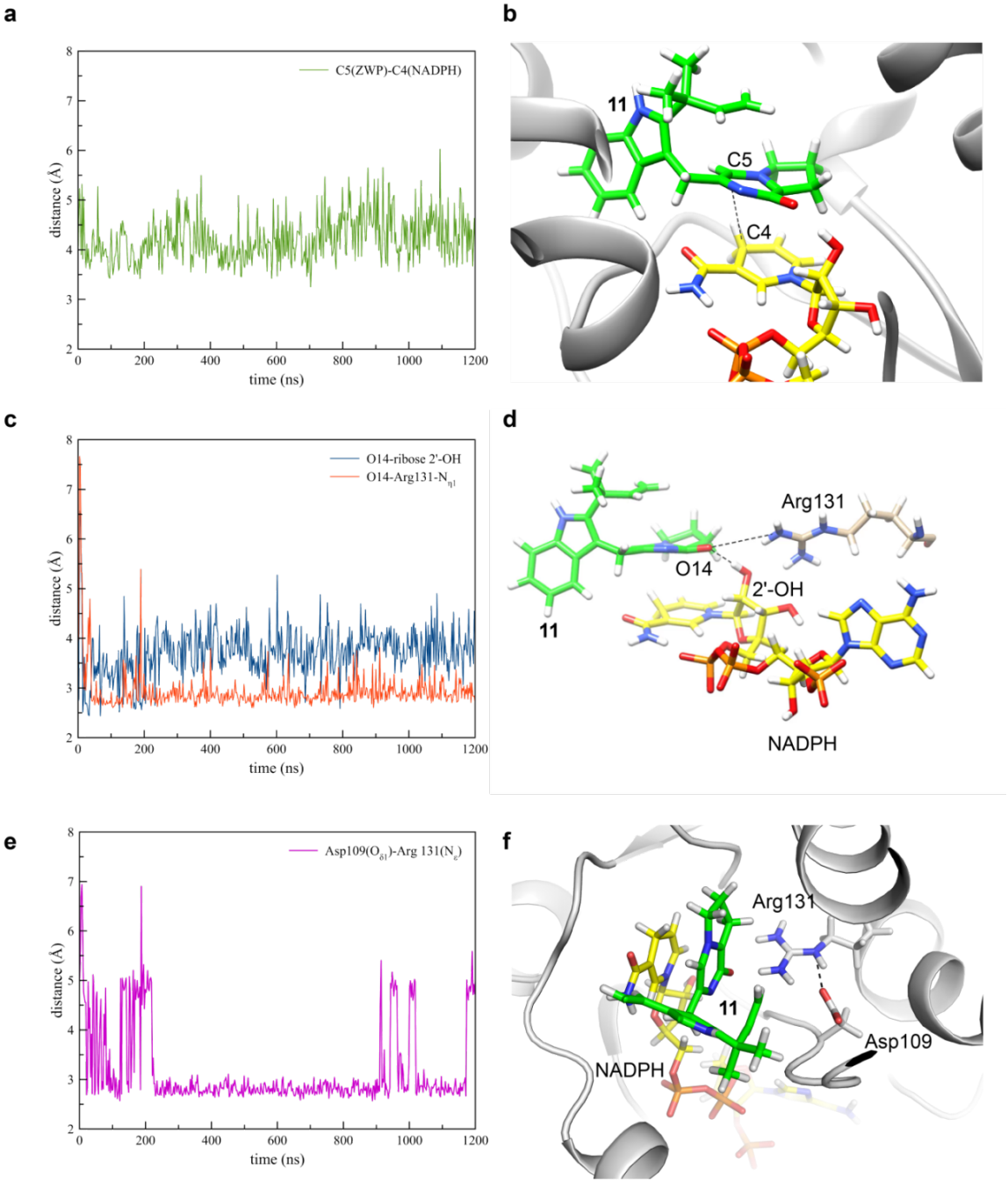
Supplementary Figure 19. PhqE is a bifunctional reductase and Diels-Alderase. a. EIC profile of *in vitro* malbrancheamide pathway reconstitution assay, with MalC replaced by PhqE. b. Chiral separation of (+)-**1** and (–)-**1**, indicating that PhqE is diastereo- and enantioselective. Reconstitution of "MalG+MalE" is shown as a negative control, and "MalG+MalE+MalC" is a positive control. All experiments were repeated independently with similar results for three times.



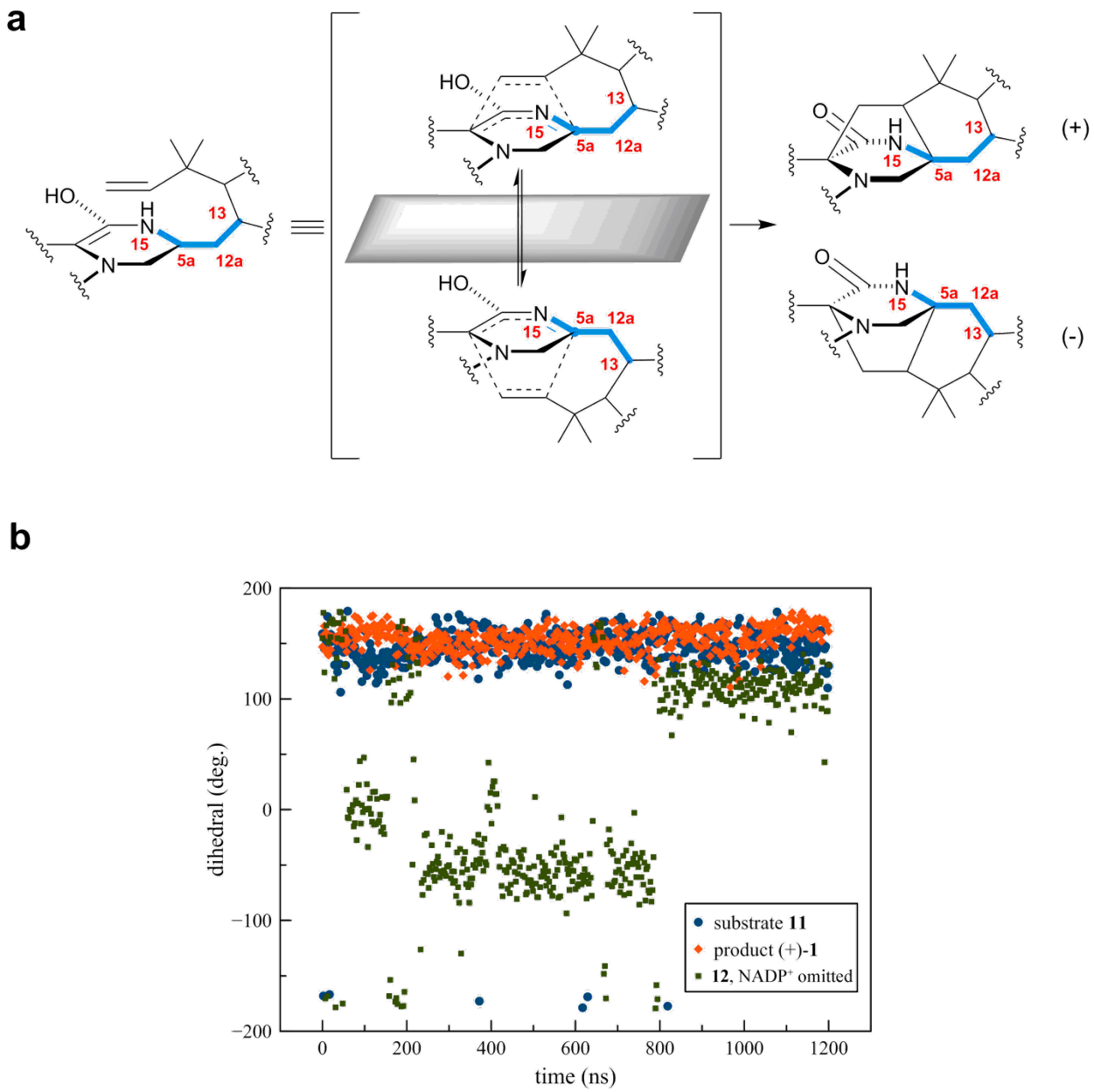
Supplementary Figure 20. PhqE crystal lattice in space group *C*2. a. Packing diagram colored by B factor from 20 Å^2 in blue to 50 Å^2 in red. The asymmetric unit contains 1.5 tetramers. One tetramer (chains A – D) is well-ordered, while in the half-tetramer (chains E and F) chain F is poorly packed along the crystallographic two-fold axis. b. Electron density (2Fo-Fc, contoured at 1σ) for the E-F half-tetramer showing the poor packing of chain F.



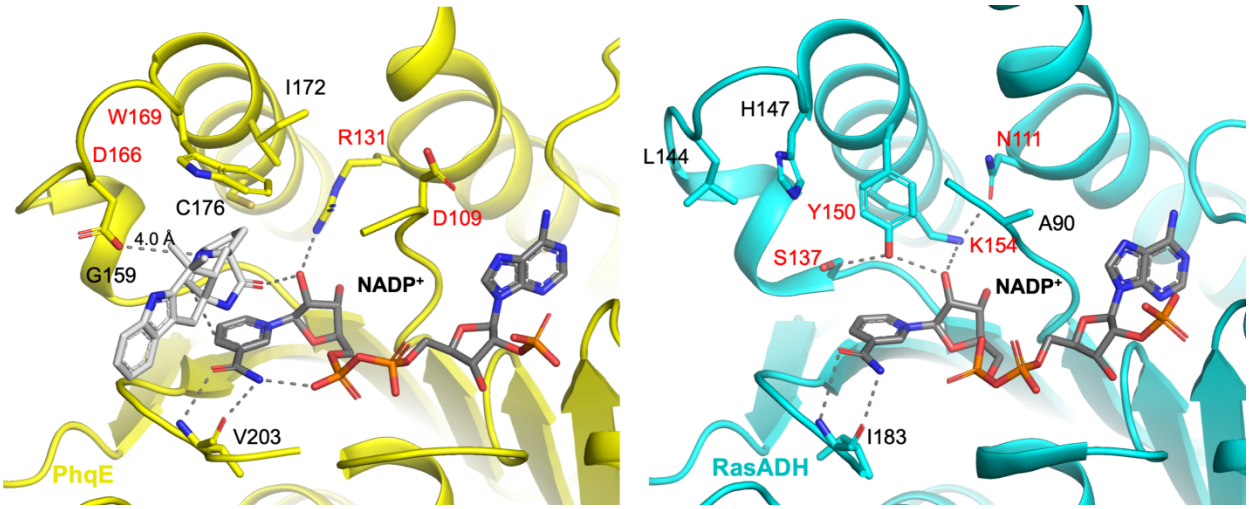
Supplementary Figure 21. Cofactor binding of PhqE. a – b. Poor omit density for NAD^+ in PhqE (contoured at 2σ) compared to well-ordered $NADP^+$ (contoured at 2σ). c. A selection of bacterial SDRs (*Burkholderia cenocepacia*, PDB 5U2W, pink; *Sinorhizobium meliloti*, PDB 3TOX, yellow; *Ralstonia sp.*, PDB 4BMS[53], blue; *Bacillus subtilis*, PDB 5ITV[54], orange; *Brucella melitensis*, PDB 5T5Q, green) superposed on PhqE (gray). The cofactor binding mode and loops surrounding the active site are remarkably similar.



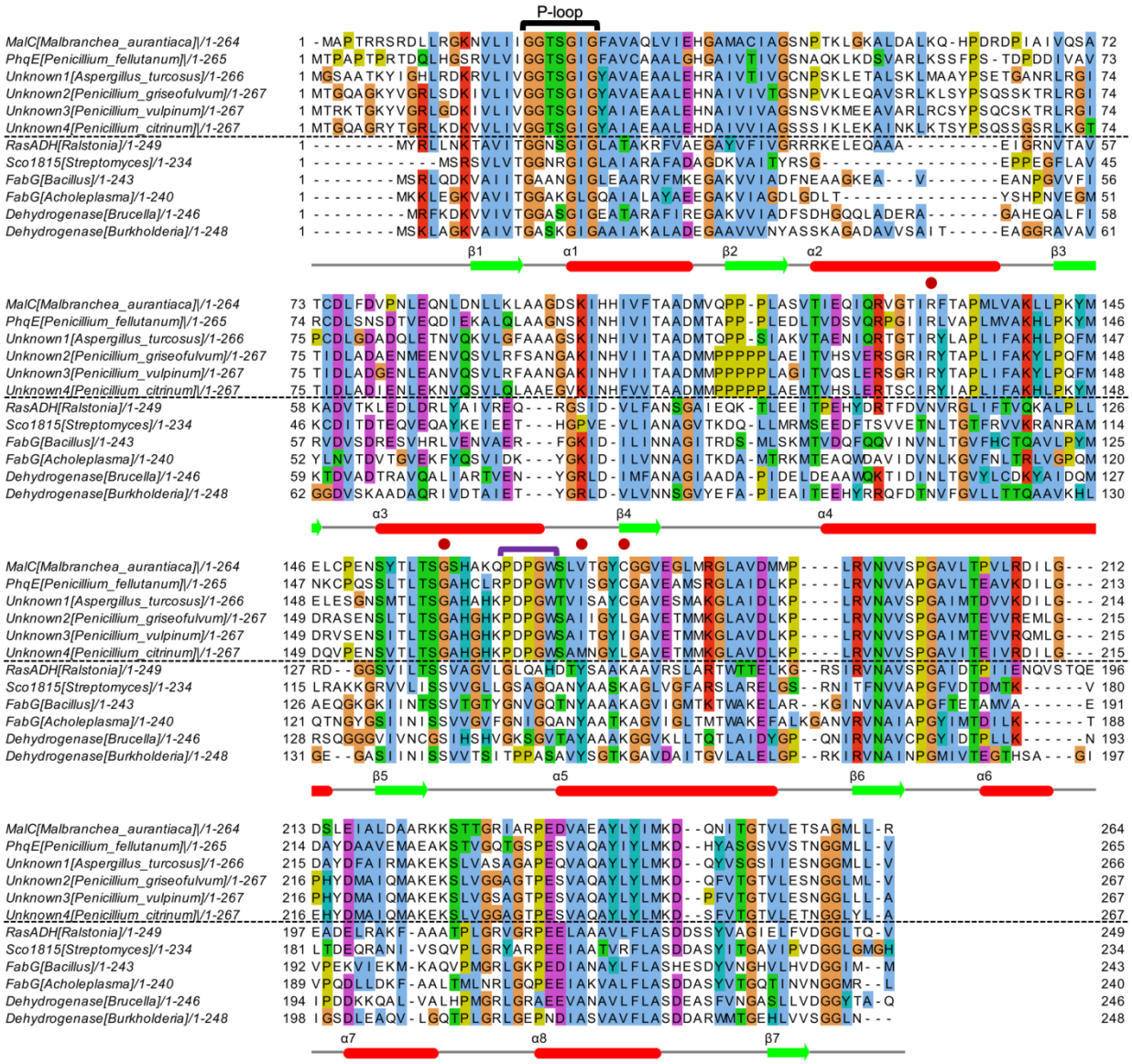
Supplementary Figure 22. a. Plot showing the distance between C5 of **11** and C4 of NADPH during 1200 ns simulation; the average distance is 4.4 Å. b. Snapshot depicting the atoms. c. Distance between the ribose hydroxyl and O14 (blue) and the distance between the Arg131 and O14 (red). Arg131 is expected to be protonated at physiological pH, based on the pKa prediction (PROPKA) of 11.10.[55,56] d. Snapshot from the beginning of the simulation. e. Arg131 also interacts with Asp109 during the simulation, which allows access to bulk solvent and may promote the ability of Arg131 to provide a proton source. f. Representative snapshot depicting this interaction.



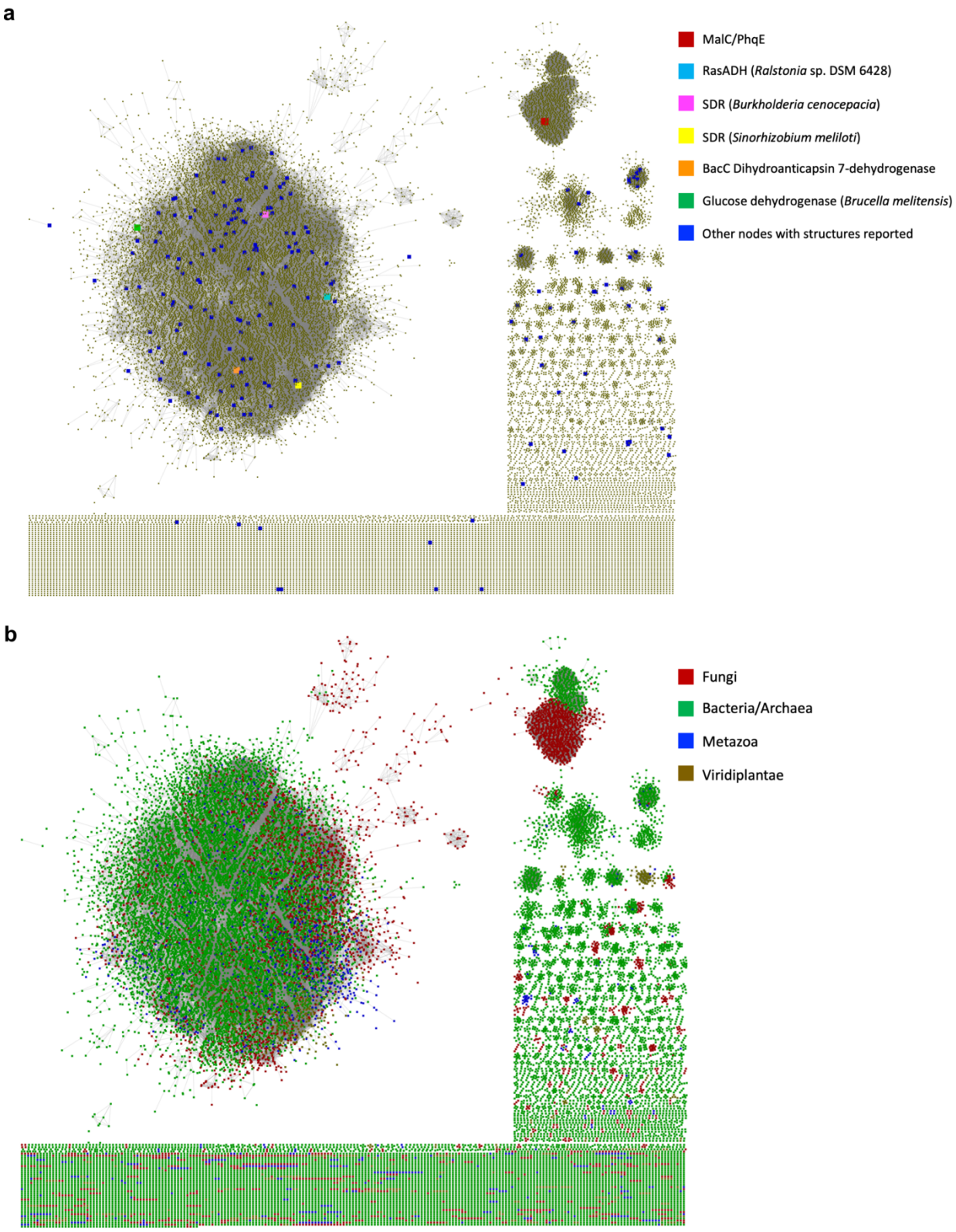
Supplementary Figure 23. a. Dihedral angles monitored to track facial selectivity in [4+2] cycloaddition reaction. b. 1200 ns MD simulation shows similar dihedrals throughout the simulation despite the fact that the substrate is unconstrained. $NADP^+$ cofactor is required in the simulation to maintain a restrained dihedral angle.



Supplementary Figure 24. Active site comparison of the PhqE product complex at left (yellow, product colored in white) and RasADH[53] at right (PDB ID: 4BMS, cyan). For each enzyme, the cofactor is colored in gray, amino acids essential to catalysis are labelled in red, hydrogen bonds are shown as gray dashed lines. Premalbrancheamide (**1**) is colored in white. In PhqE, positions of SDR catalytic residues (RasADH Asn111, Ser137, Tyr150, and Lys154) are occupied by Arg131, Gly159, Ile172, and Cys176, respectively. Given the striking difference in amino acids in the two active sites, the PhqE and RasADH backbones are remarkably similar.

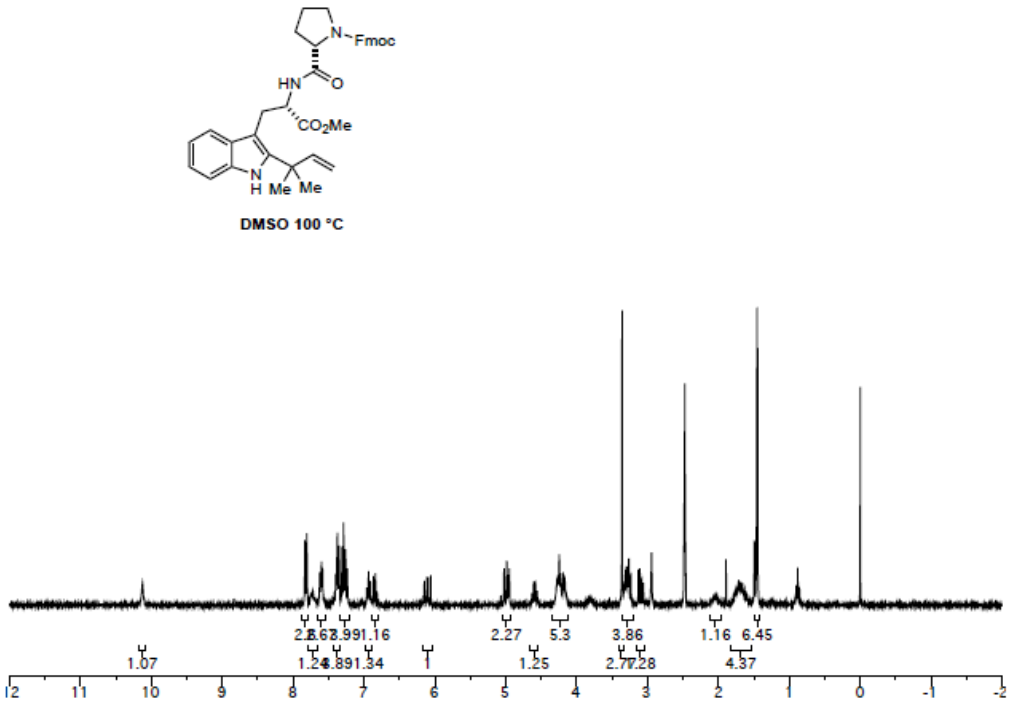


Supplementary Figure 25. Multiple sequence alignment of MalC and PhqE homologs. The P-loop critical for cofactor binding is highlighted with a black cap, the "PDPGW" motif is highlighted with a purple cap. SDR amino acids essential for catalysis (Asn-Ser-Tyr-Lys) are shown with red dots, all of which are different in the MalC/PhqE-type Diels-Alderases. In order to validate the reliability of the unknown sequences, we identified and annotated two fungal genomes (*Aspergillus turcosus*, GenBank accession number NIDN01000061; *Penicillium griseofulvum*, GenBank accession number LHQR01000065; Table S4)[57], confirming that both contain clustered homologs of *malG*, *malE* and *malC*, and revealing more potential pathways that produce the bicyclo[2.2.2]diazaoctane nucleus. Sequences below the dashed line are conventional SDRs of known structure.

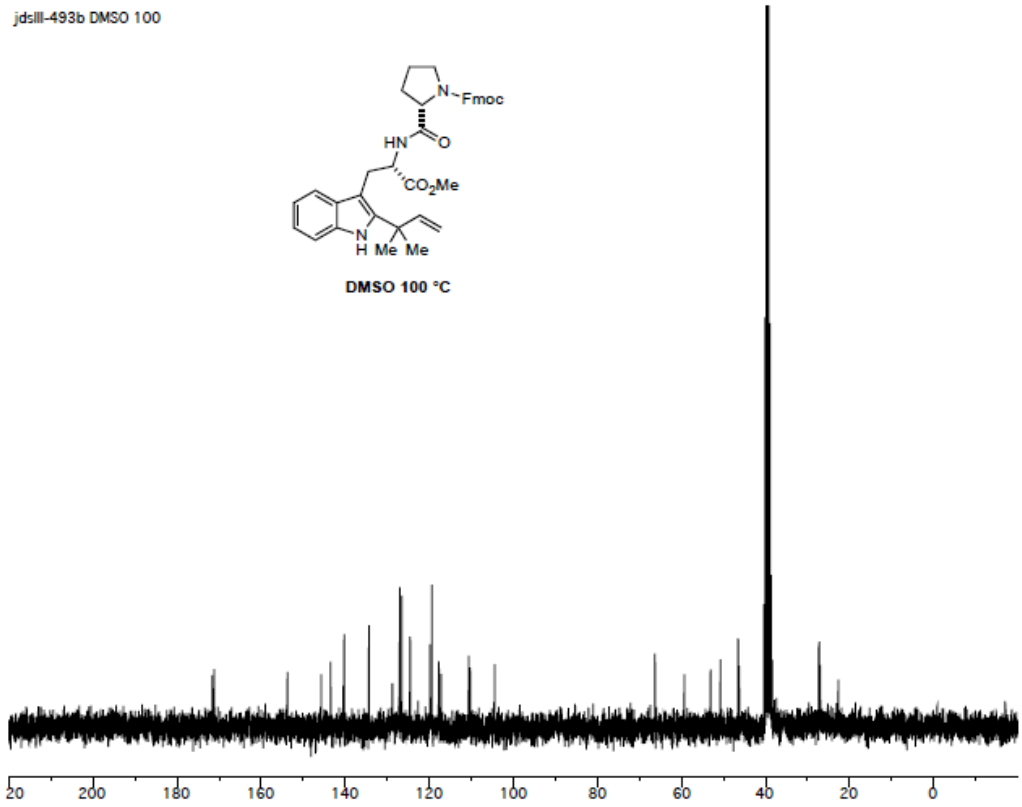


Supplementary Figure 26. Sequence similarity network (SSN) of Pfam family PF13561 (adh_short_C2), a subset of short-chain dehydrogenase/reductases including the MalC/PhqE

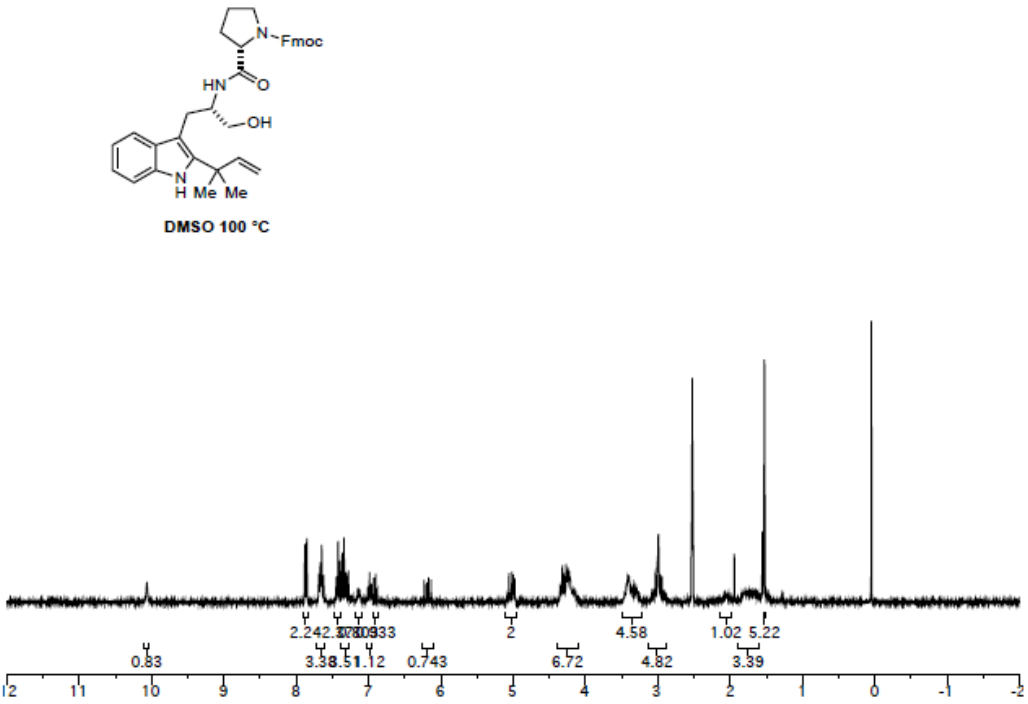
Diels-Alderases. The SSN was generated with the Enzyme Function Initiative web tool (https://efi.igb.illinois.edu/efi-est/)[58] using the UniRef50 databases, in which sequences with ≥ 50% sequence identity over 80% of the sequence length are grouped together and represented by a single node. A total of 28,659 nodes are included in this SSN. a. PF13561 SSN with highlights for the MalC/PhqE Diels-Alderases (red), PDB entries (blue) and selected enzymes. The MalC/PhqE Diels-Alderases reside in an evolutionary island (upper right) and are also the first crystal structures within the island. b. PF13561 SSN colored by biological (super)kingdom. The island of the MalC/PhqE Diels-Alderases is nearly exclusively in the fungal kingdom.



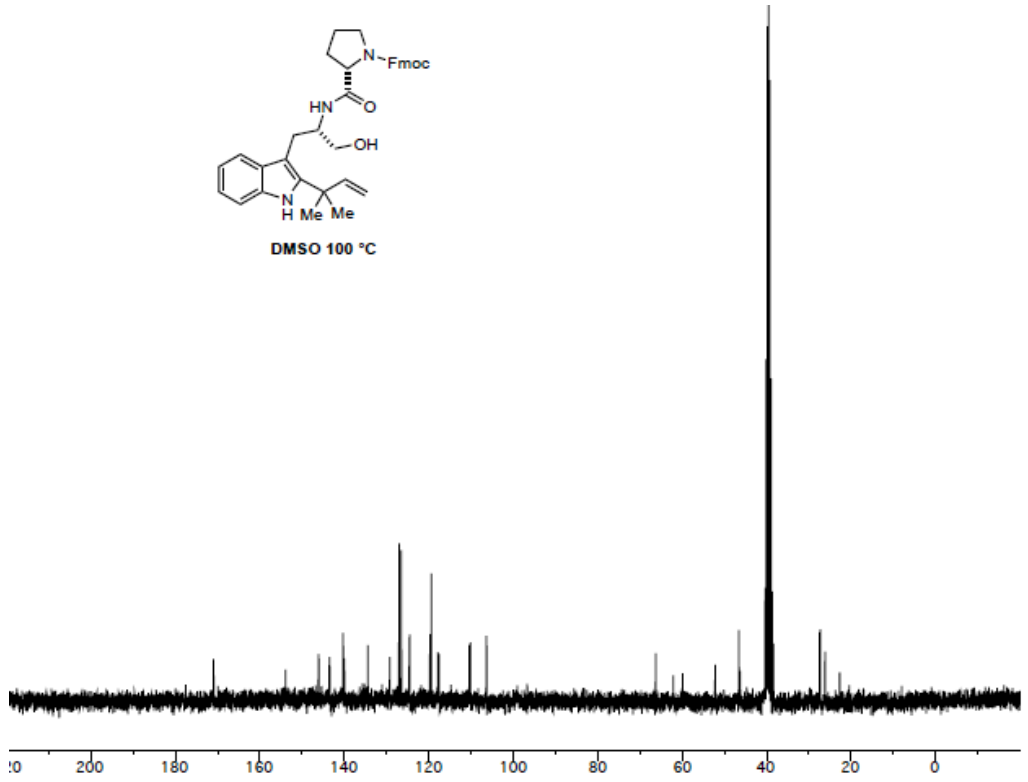
Supplementary Figure 27. ^{1}H NMR spectrum of **15**.



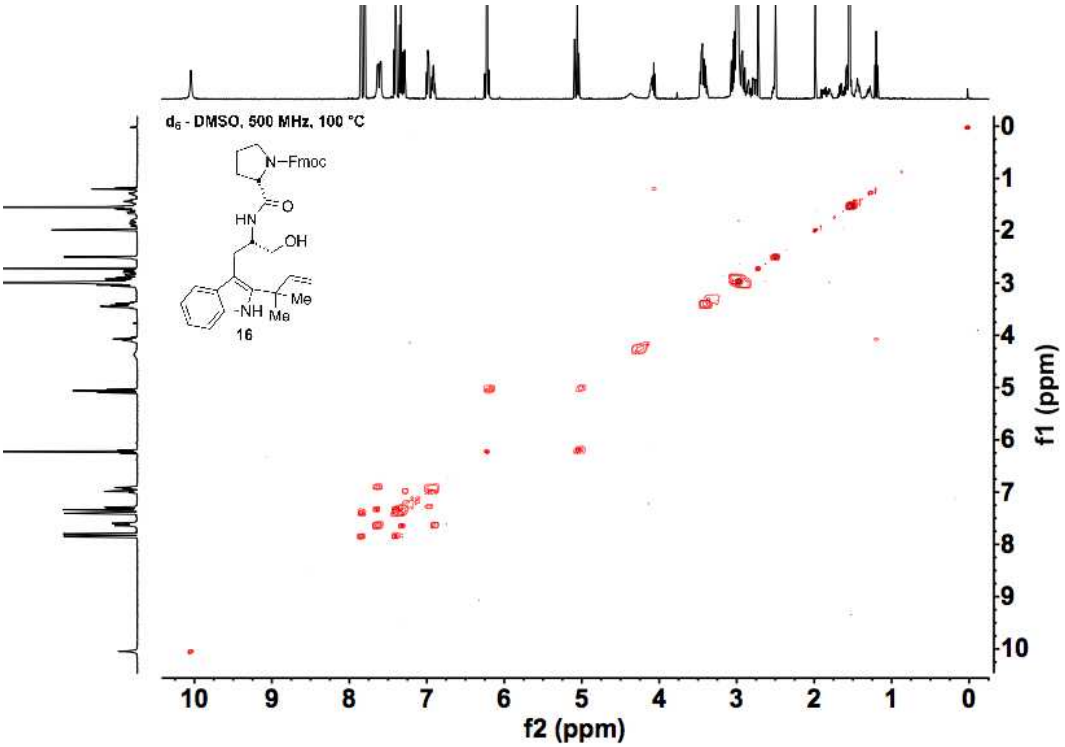
Supplementary Figure 28. ^{13}C NMR spectrum of **15**.



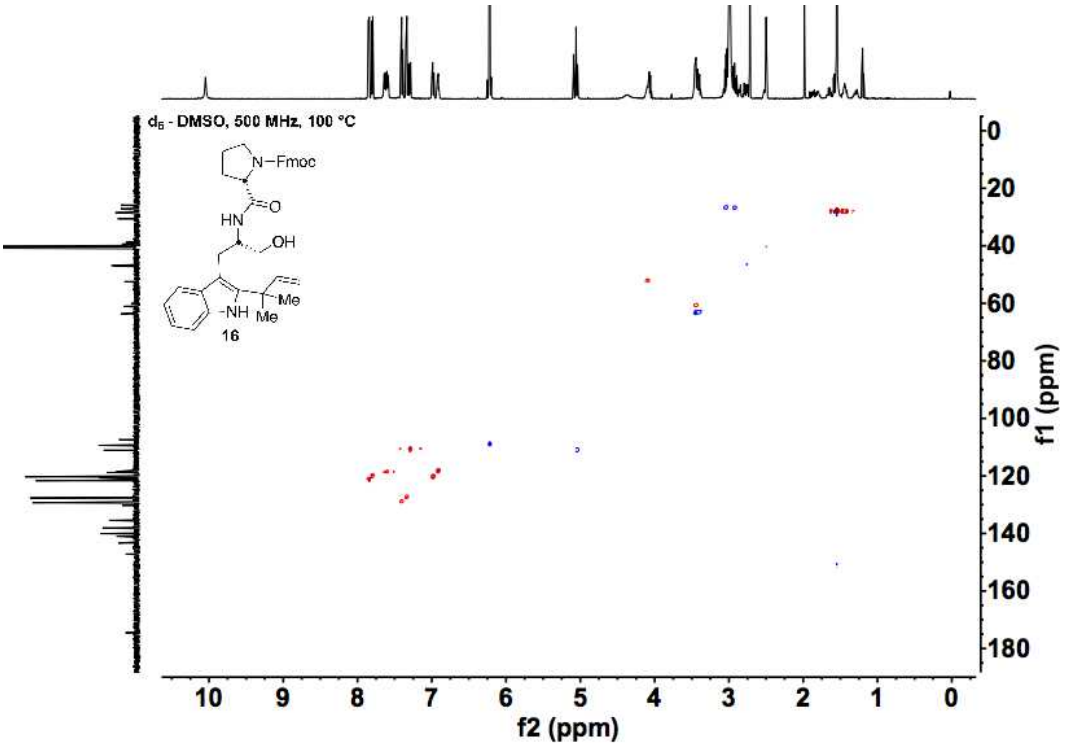
Supplementary Figure 29. ^{1}H NMR spectrum of **16**.



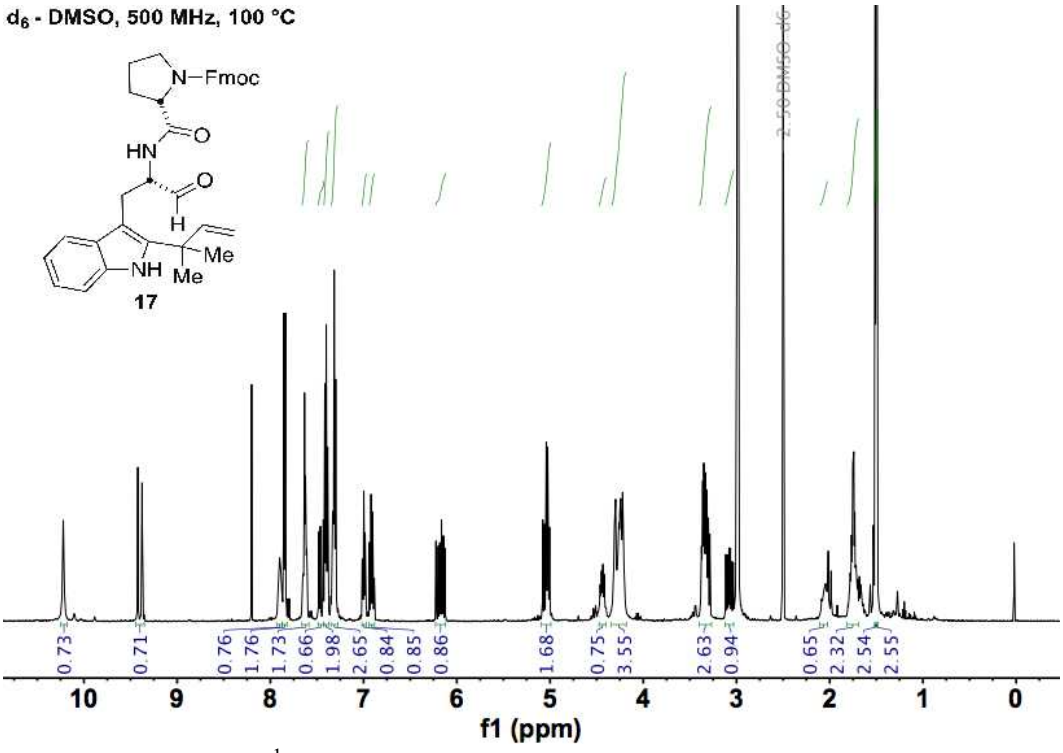
Supplementary Figure 30. ^{13}C NMR spectrum of **16**.



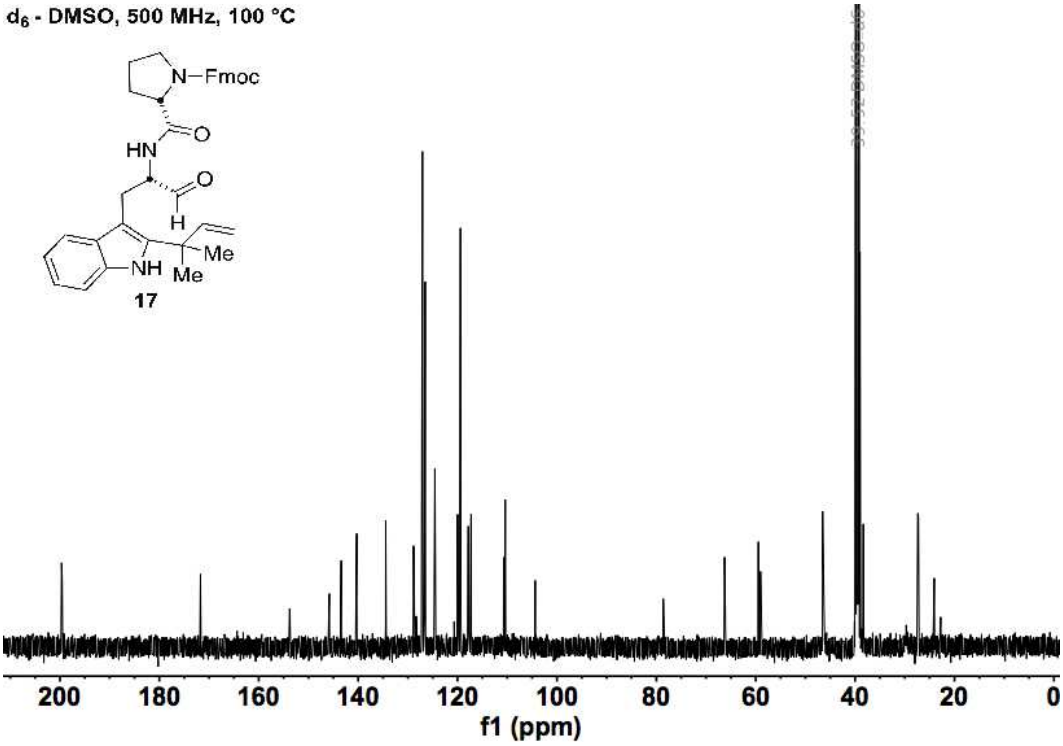
Supplementary Figure 31. ^{1}H-^{1}H COSY spectrum of **16**.



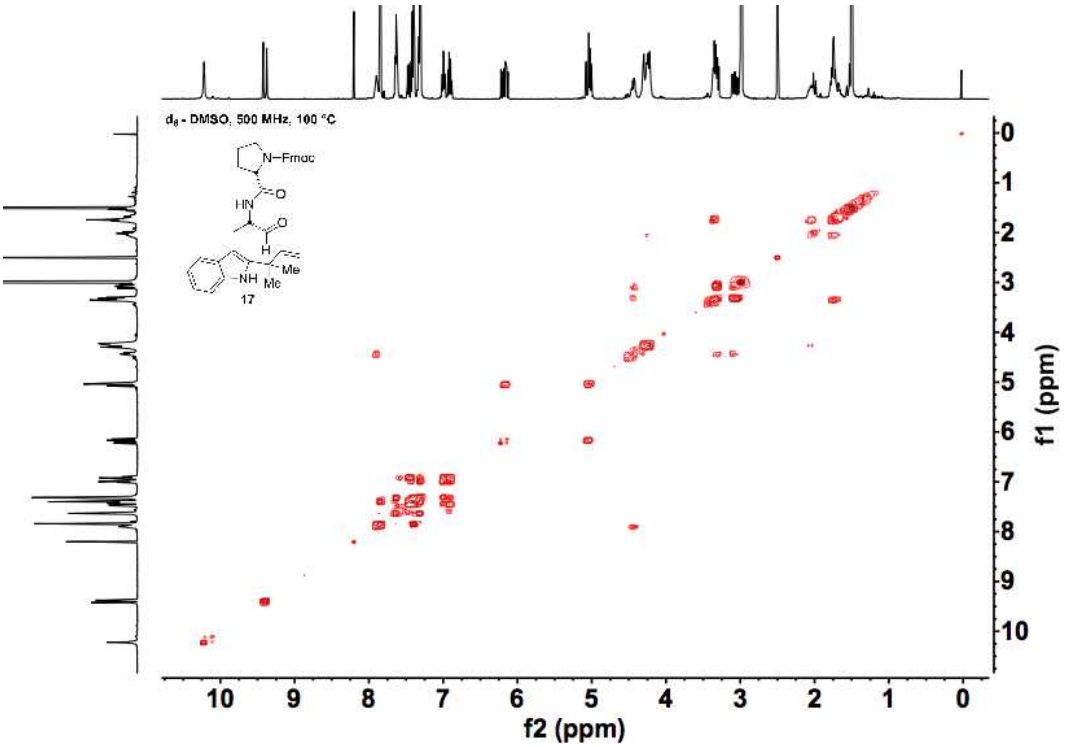
Supplementary Figure 32. ^{1}H-^{13}C HSQC spectrum of **16**.



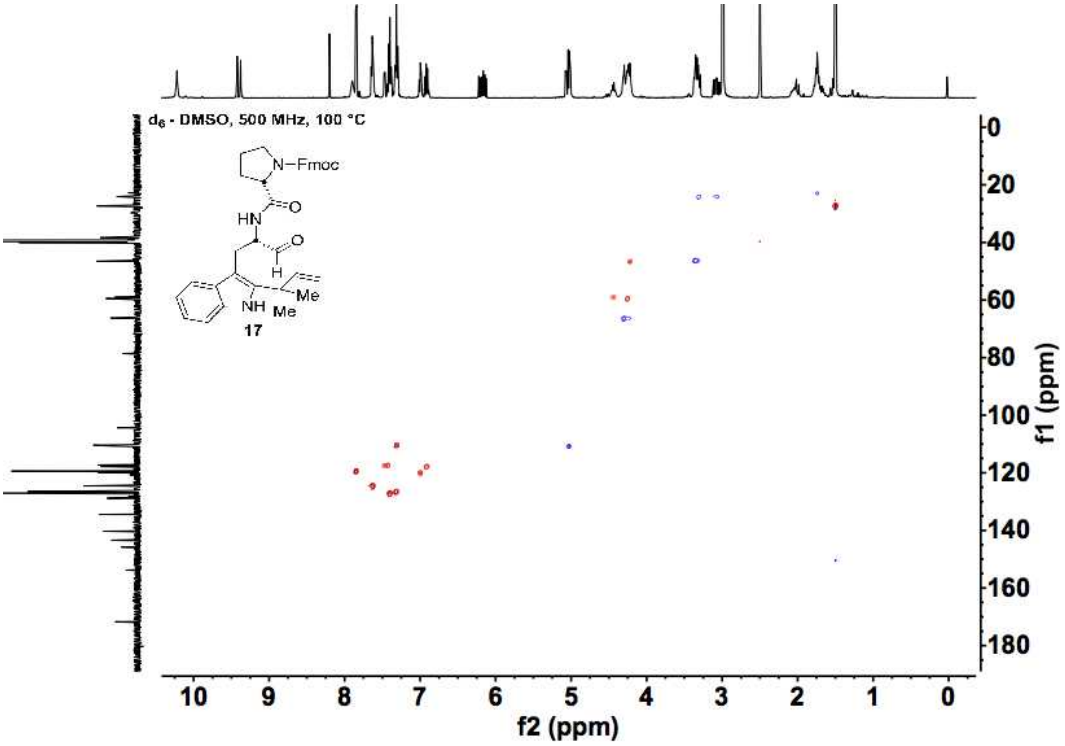
Supplementary Figure 33. ^{1}H NMR spectrum of **17**.



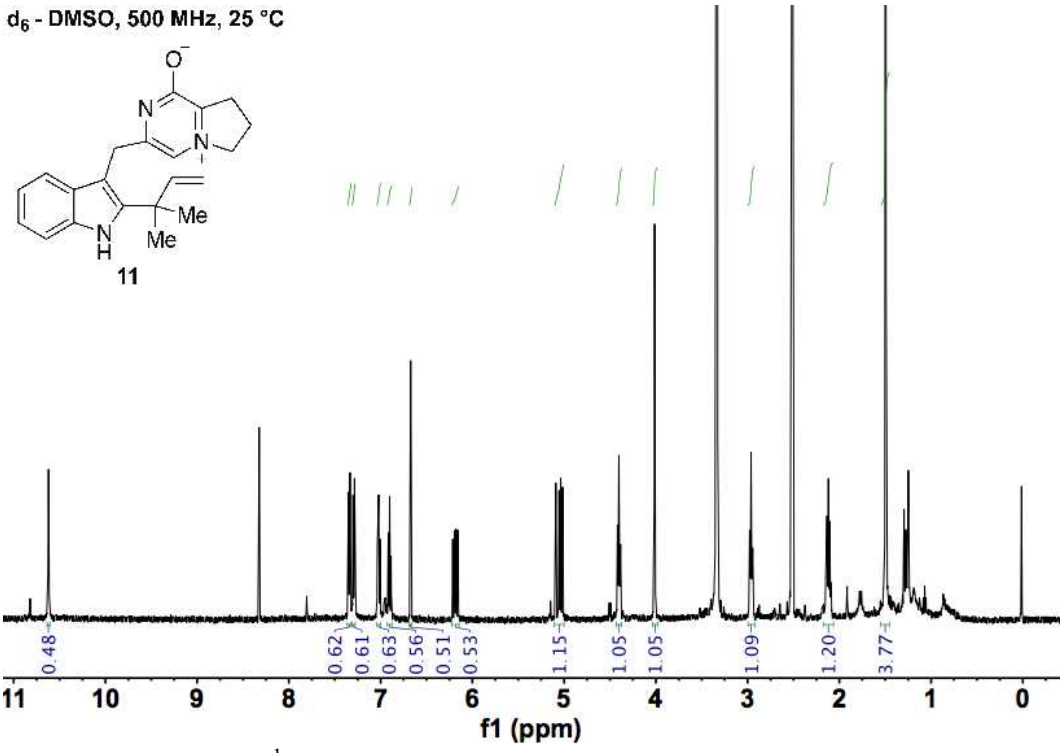
Supplementary Figure 34. ^{13}C NMR spectrum of **17**.



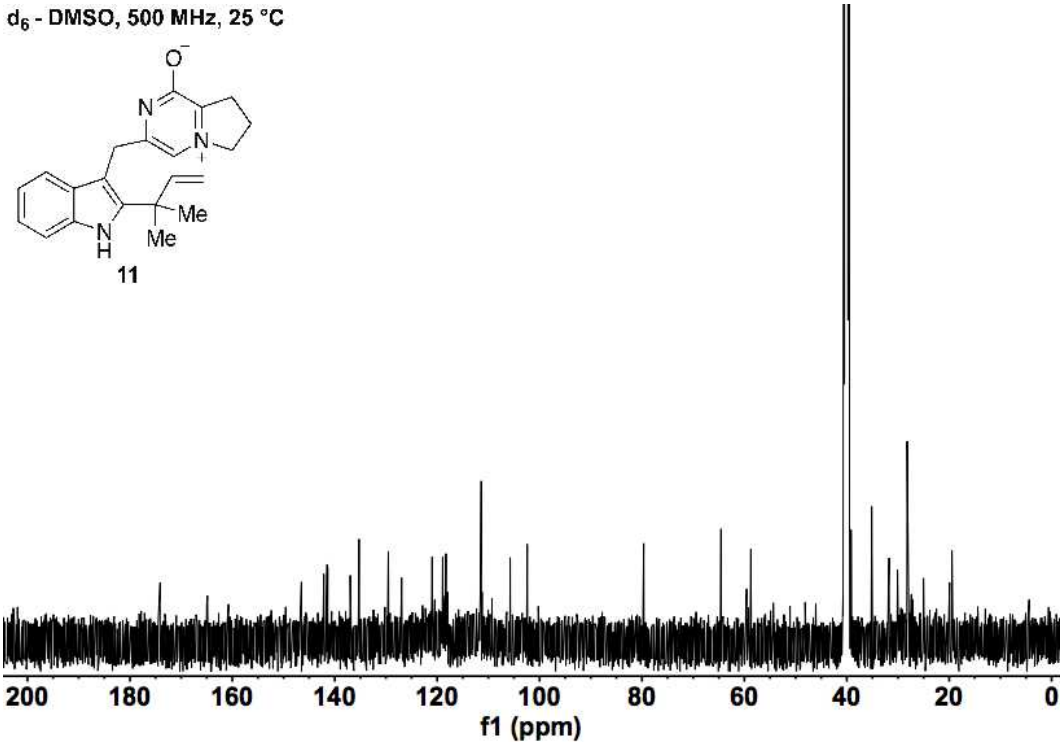
Supplementary Figure 35. ^{1}H-^{1}H COSY spectrum of **17**.



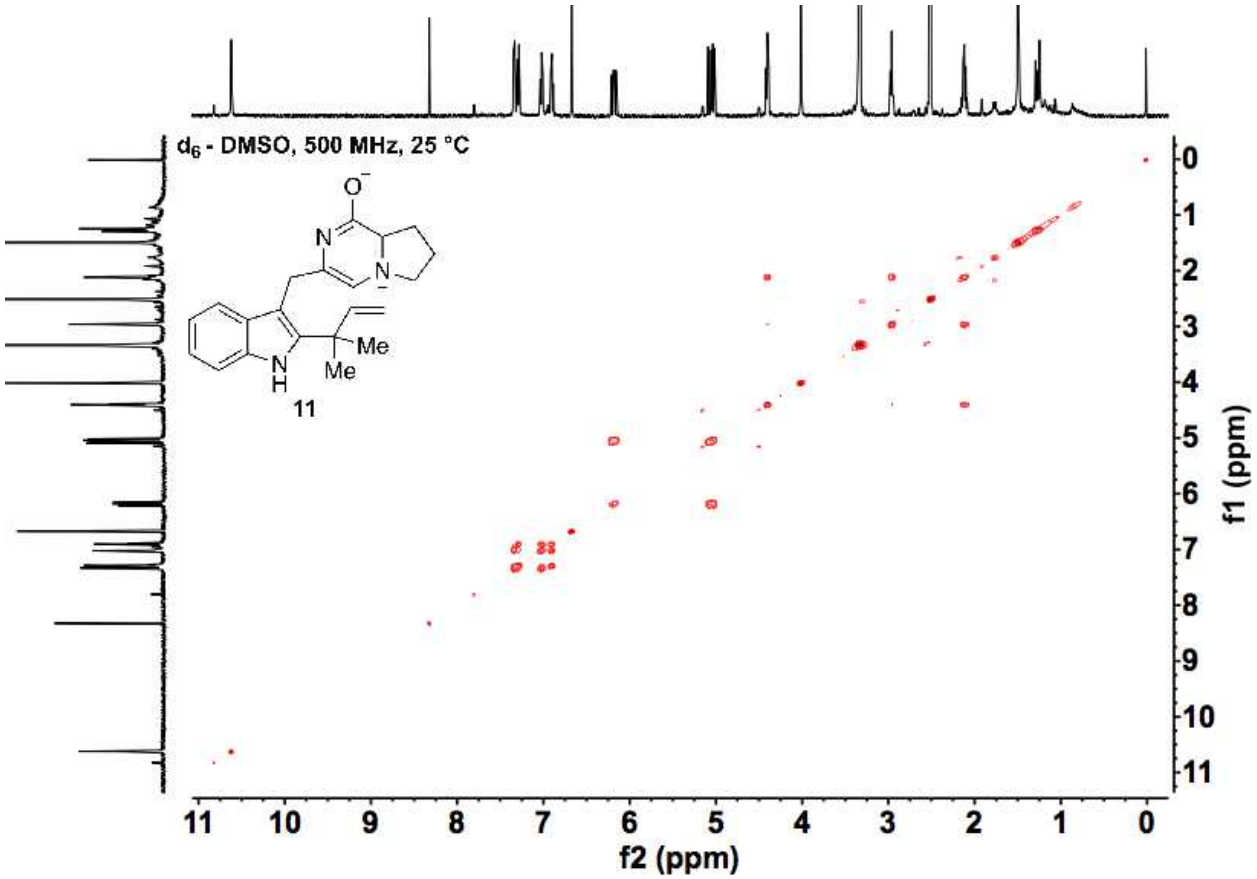
Supplementary Figure 36. ^{1}H-^{13}C HSQC spectrum of **17**.



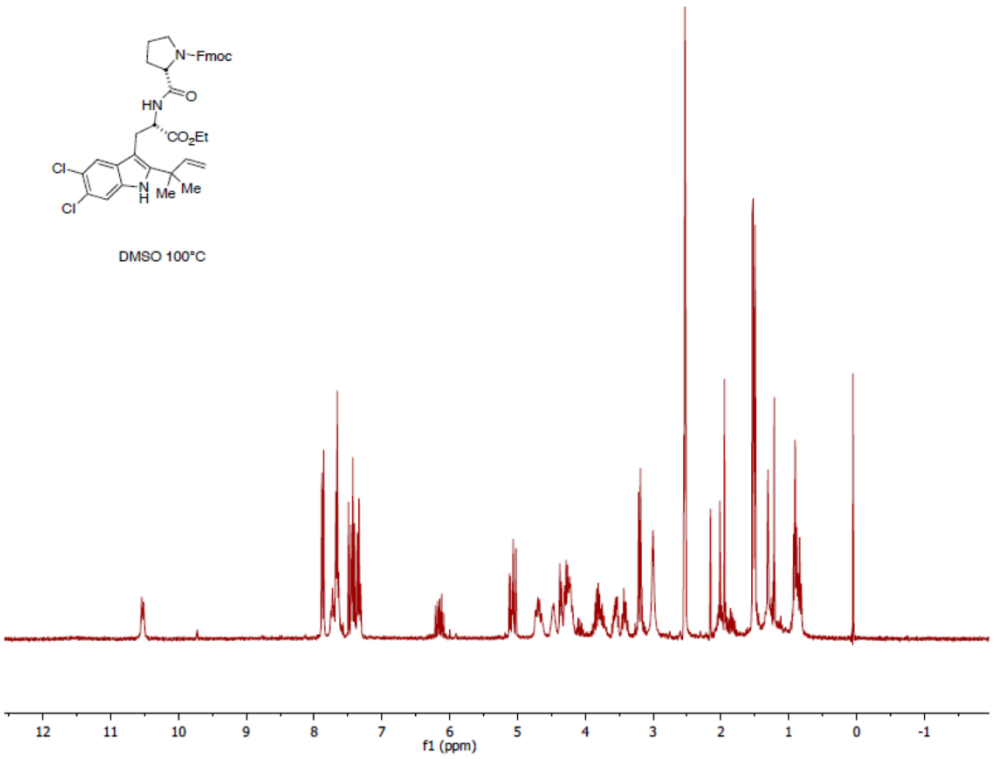
Supplementary Figure 37. ^{1}H NMR spectrum of **11**.



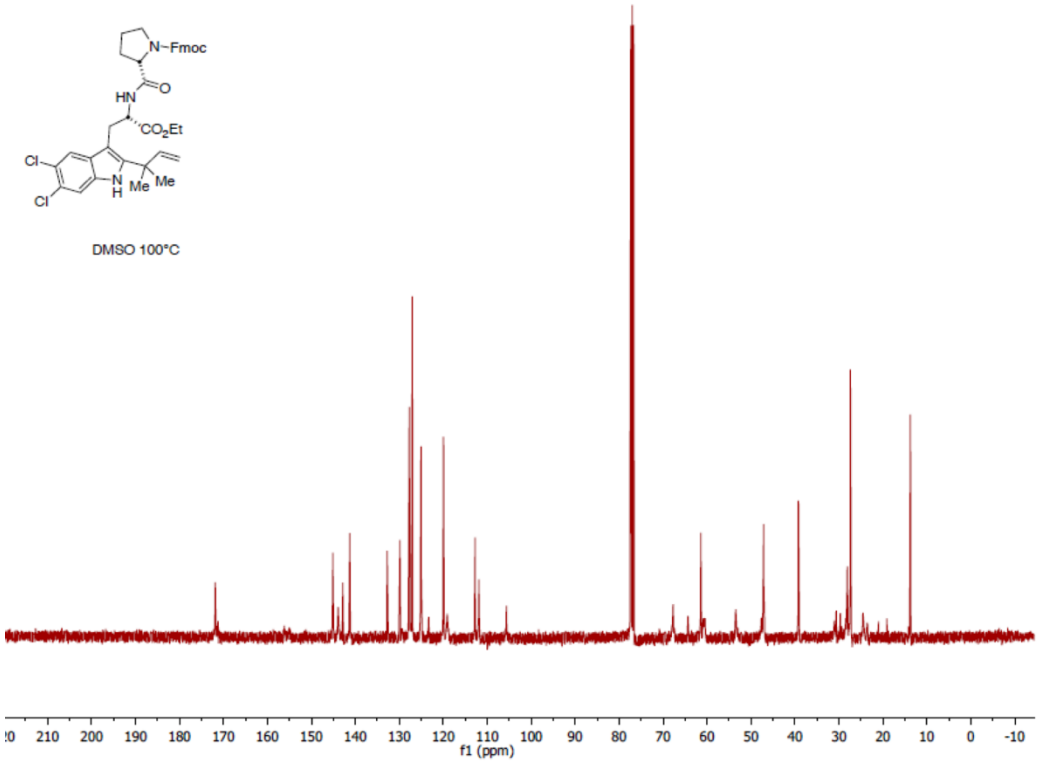
Supplementary Figure 38. ^{13}C NMR spectrum of **11**.



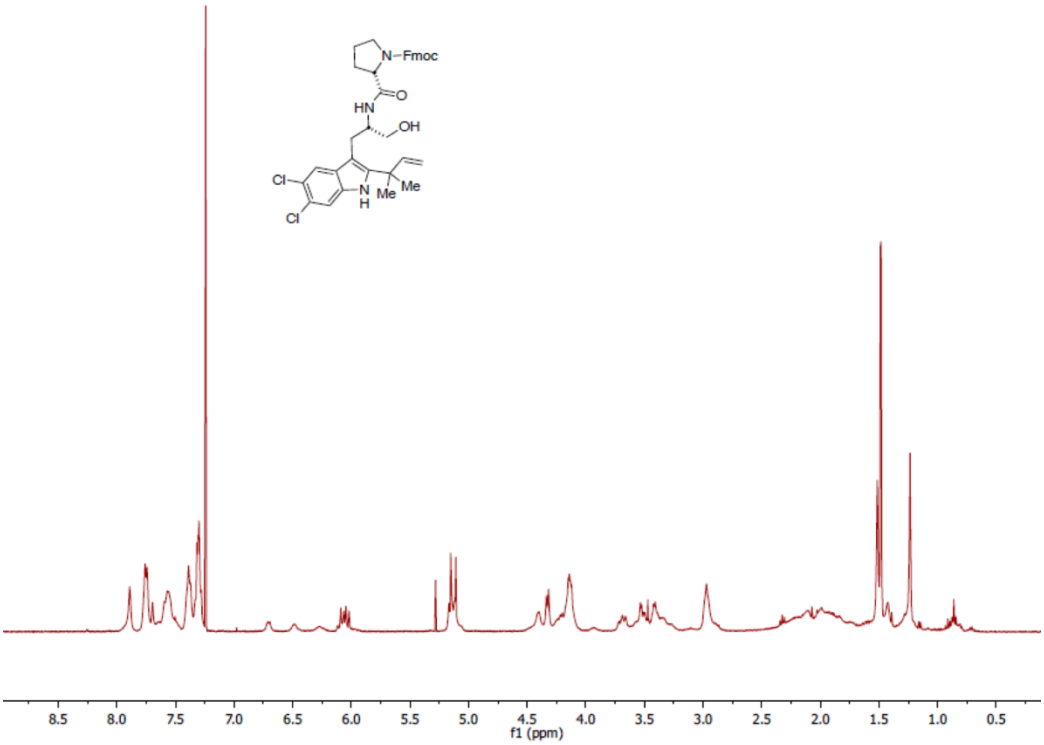
Supplementary Figure 39. ^{1}H-^{1}H COSY spectrum of **11**.



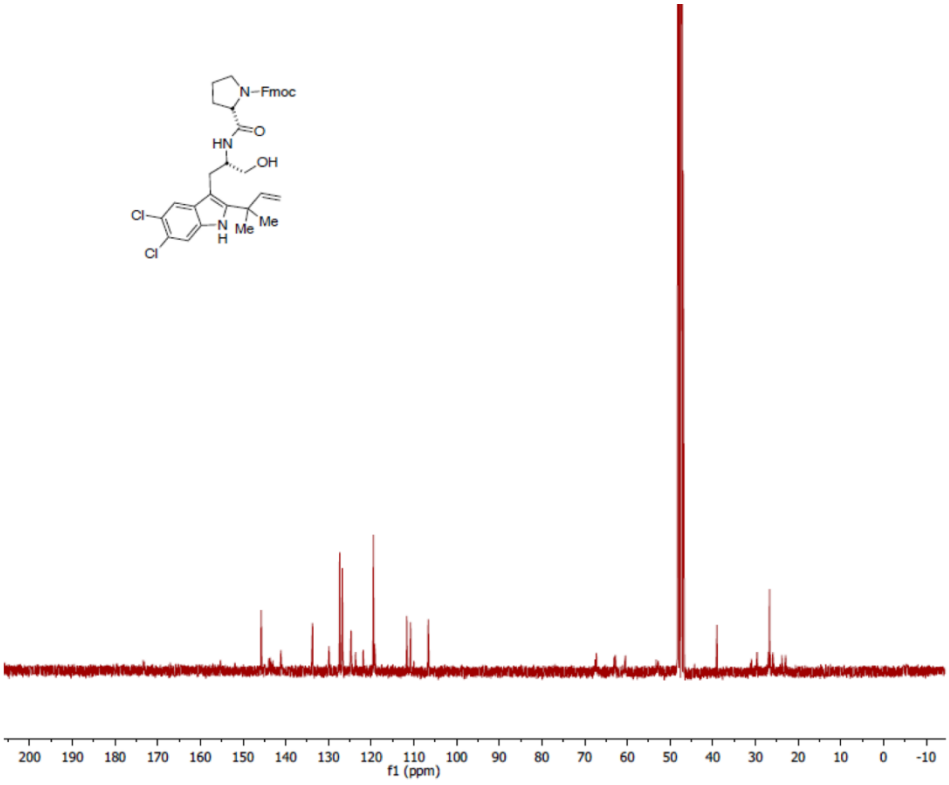
Supplementary Figure 40. ^{1}H NMR spectrum of **19**.



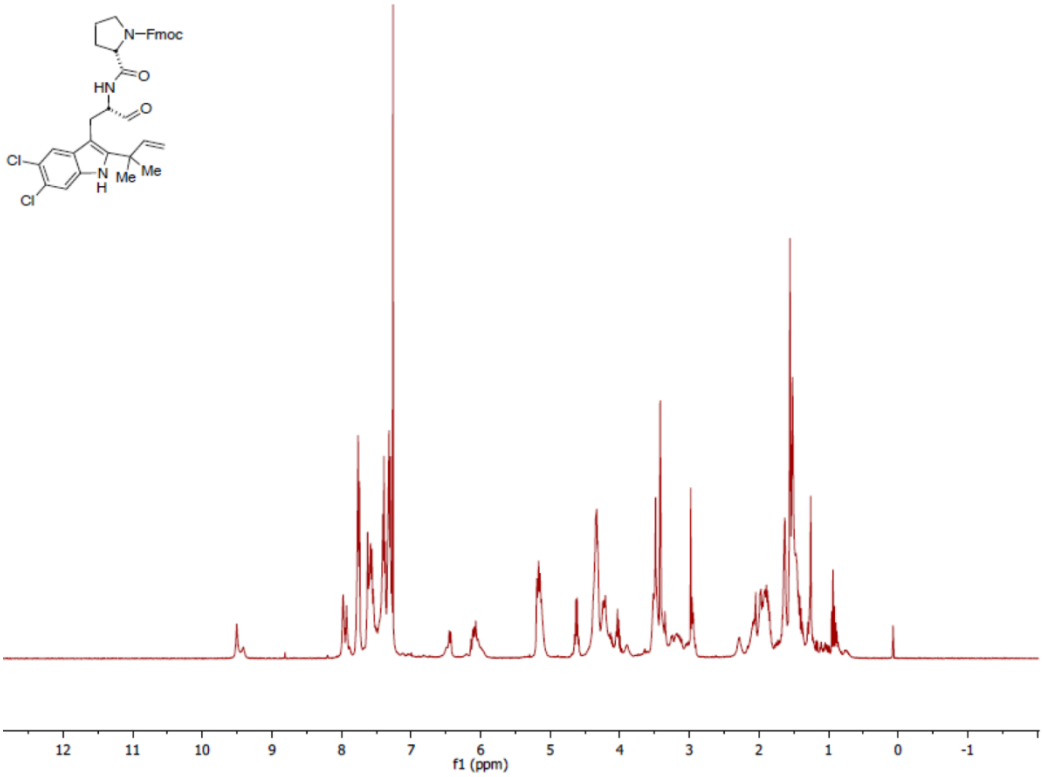
Supplementary Figure 41. ^{13}C NMR spectrum of **19**.



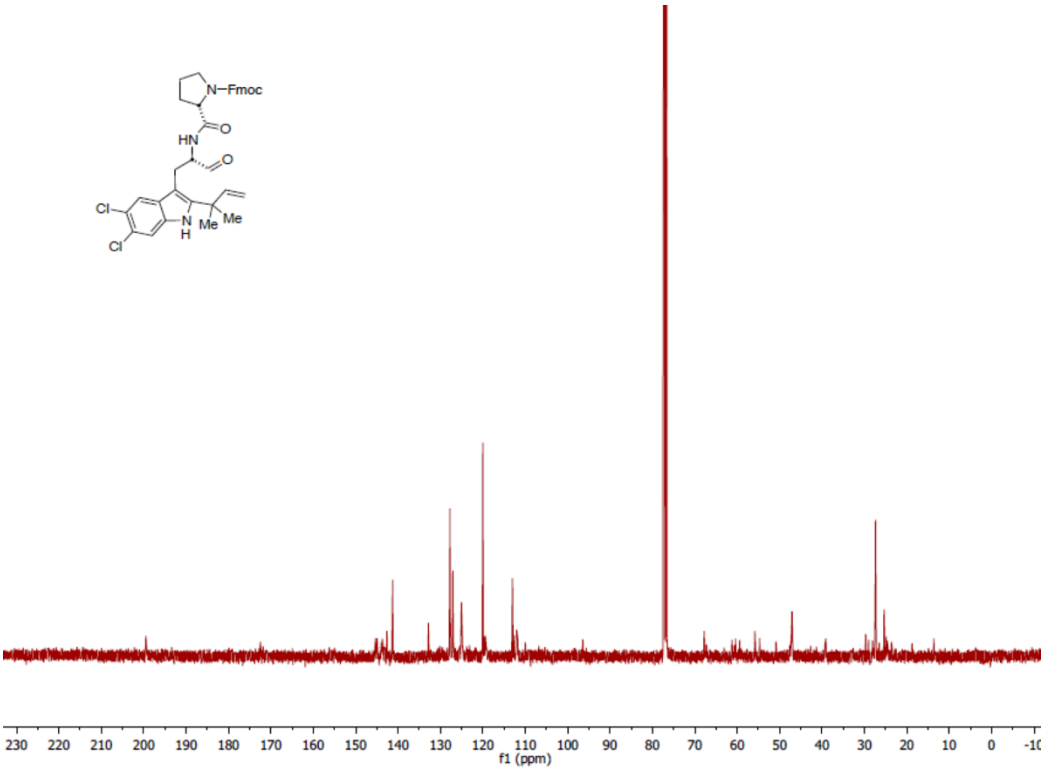
Supplementary Figure 42. ^{1}H NMR spectrum of **20**.



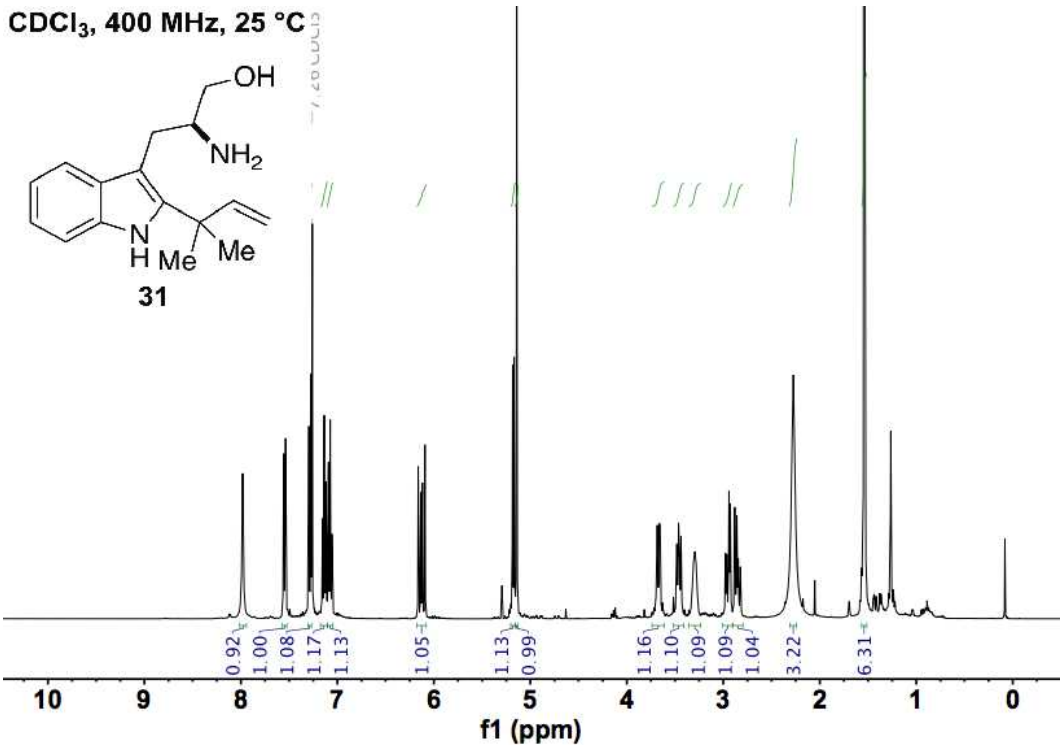
Supplementary Figure 43. ^{13}C NMR spectrum of **20**.



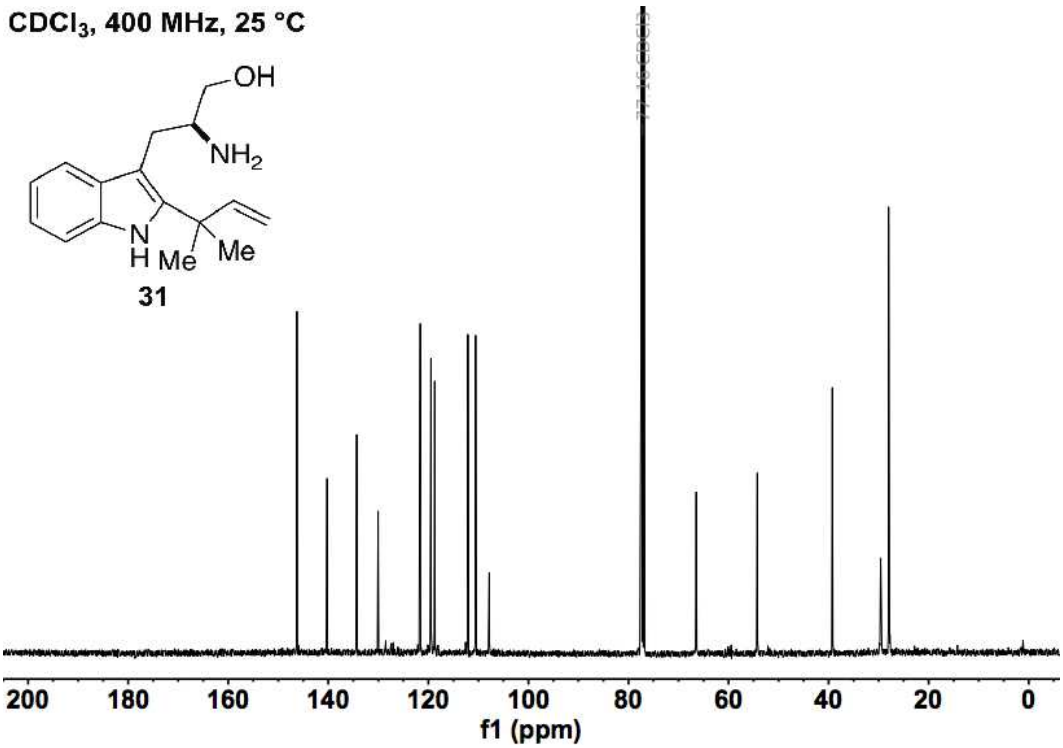
Supplementary Figure 44. ^{1}H NMR spectrum of **21**.



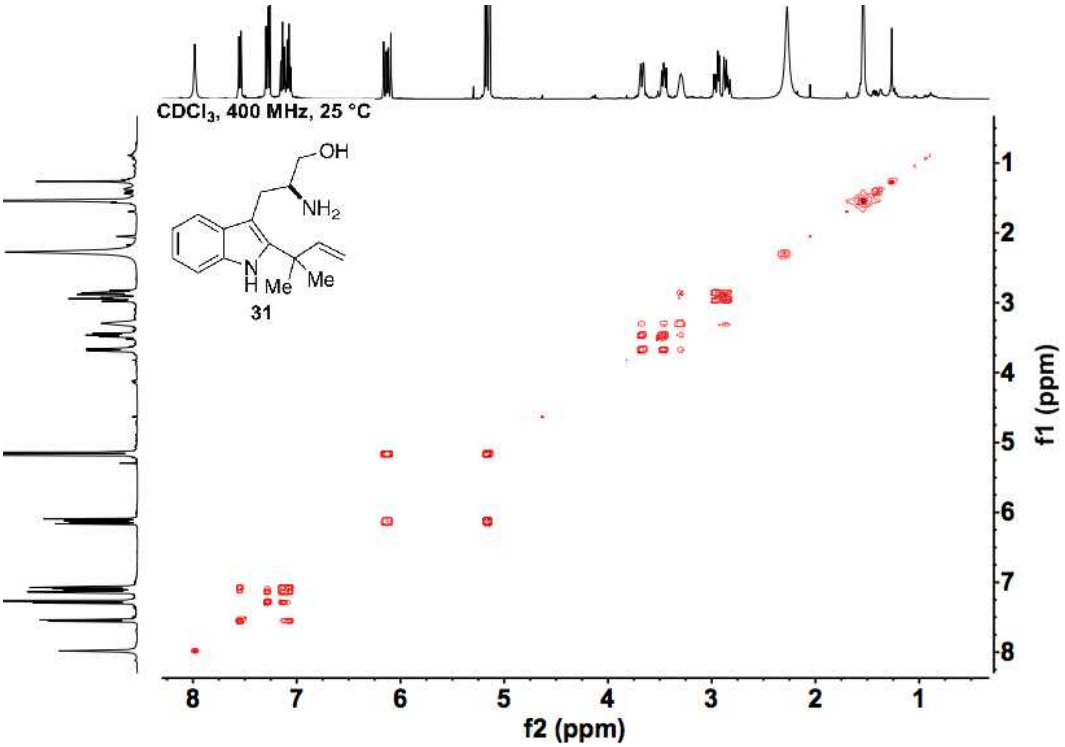
Supplementary Figure 45. ^{13}C NMR spectrum of **21**.



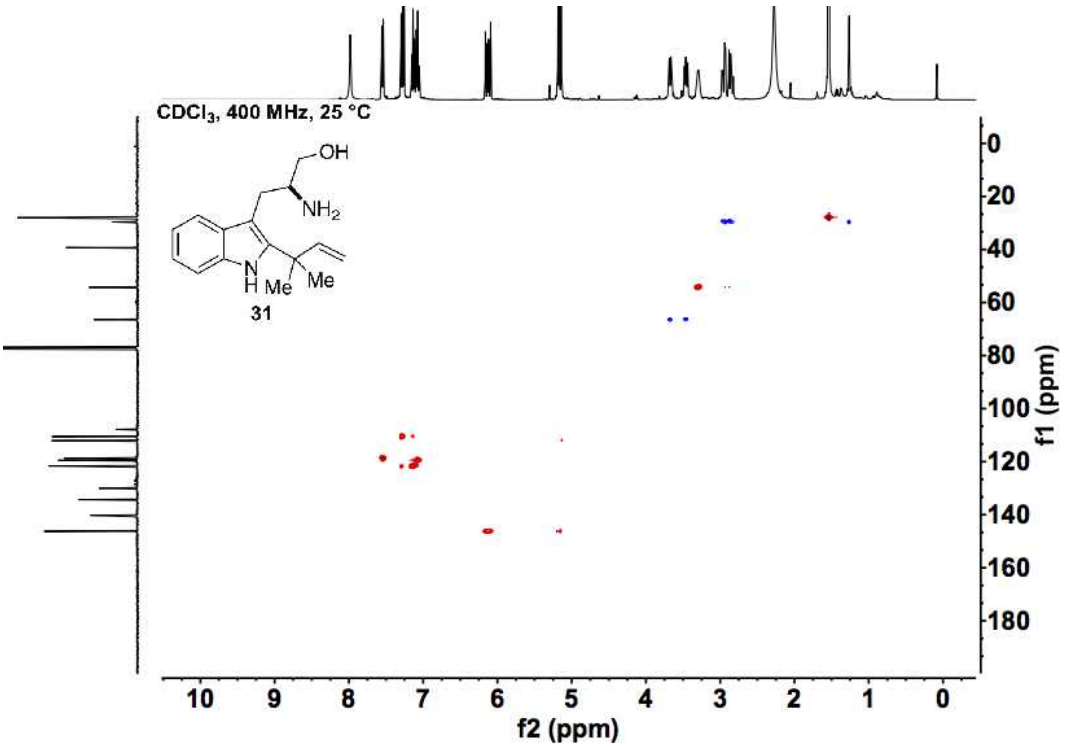
Supplementary Figure 46. ^{1}H NMR spectrum of **31**.



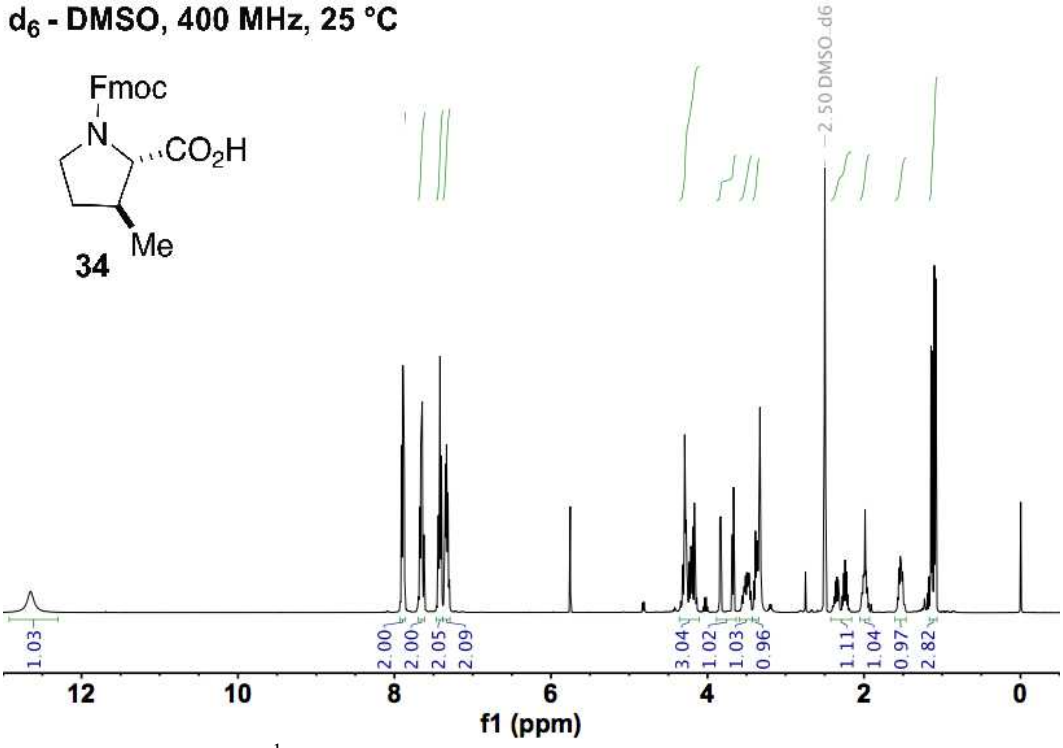
Supplementary Figure 47. ^{13}C NMR spectrum of **31**.



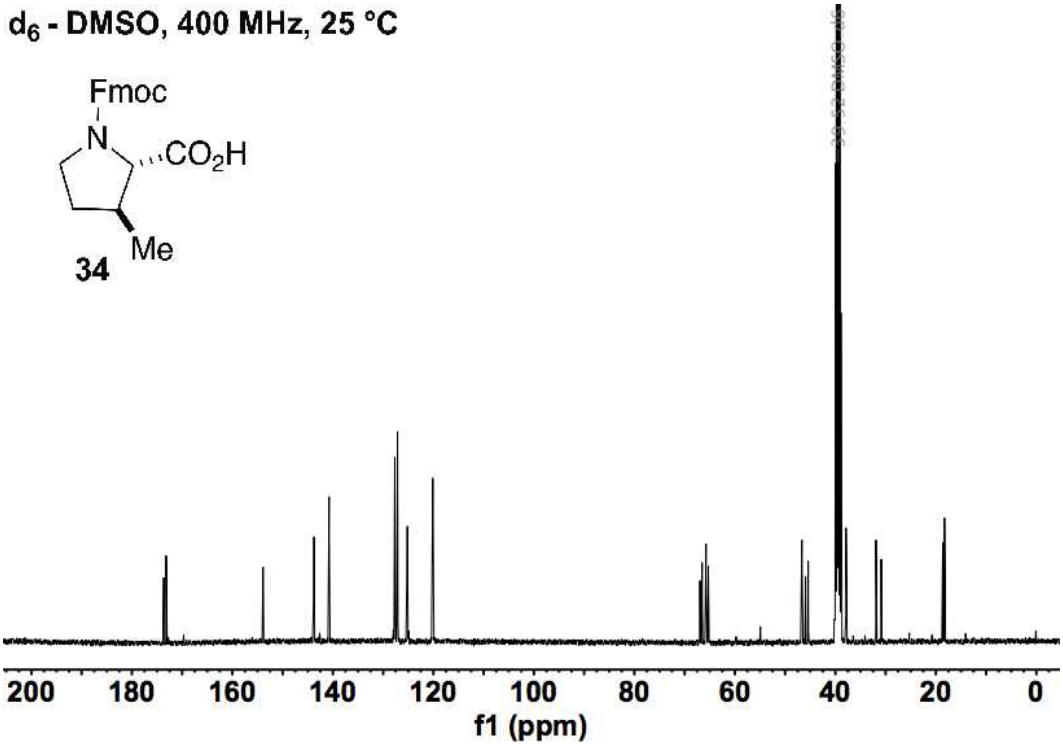
Supplementary Figure 48. ^{1}H-^{1}H COSY spectrum of **31**.



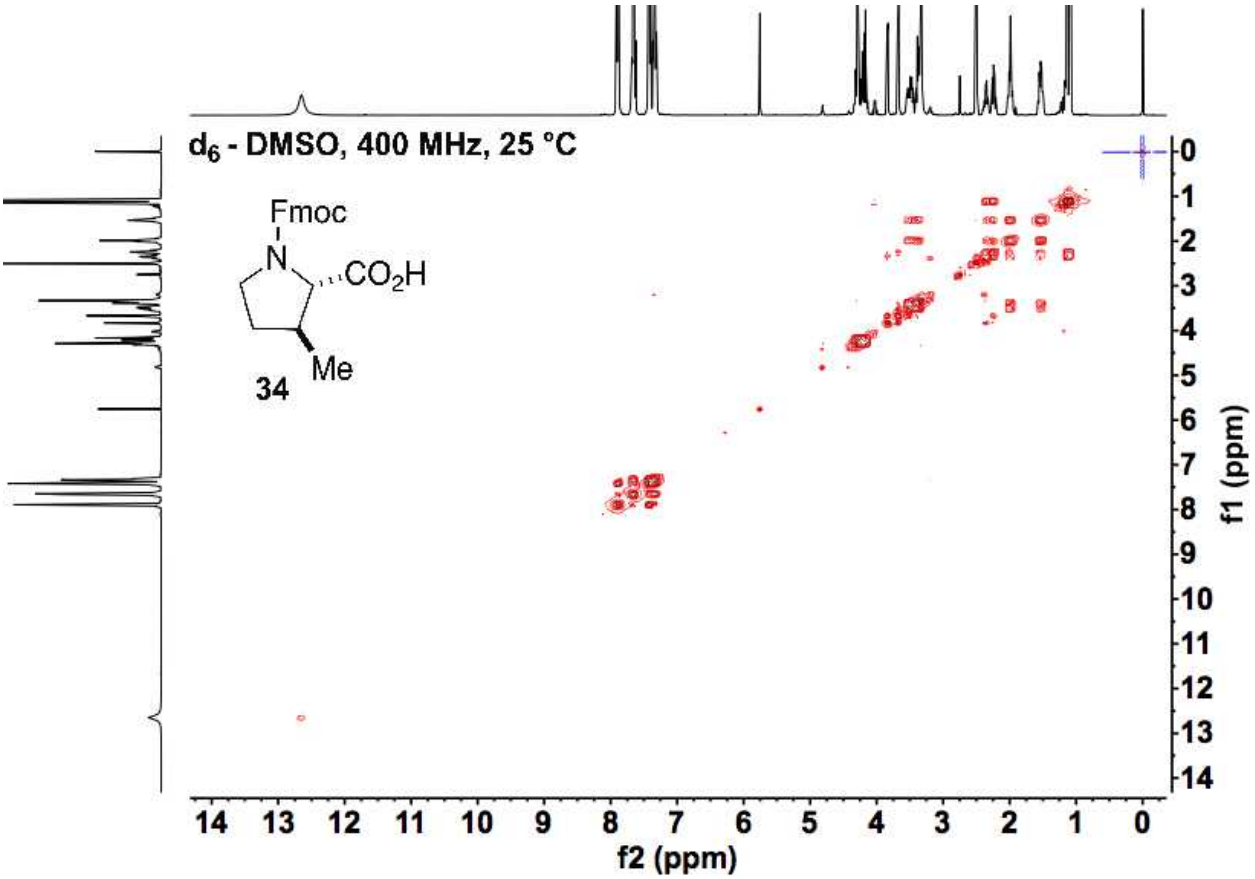
Supplementary Figure 49. ^{1}H-^{13}C HSQC spectrum of **31**.



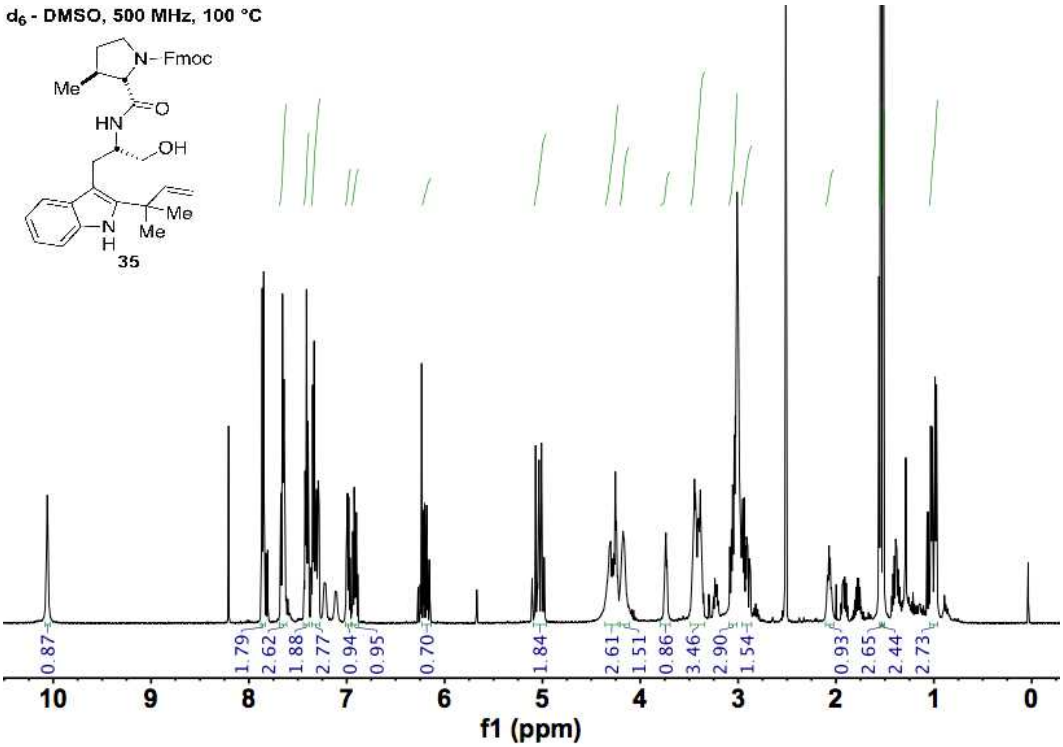
Supplementary Figure 50. ^{1}H NMR spectrum of **34**.



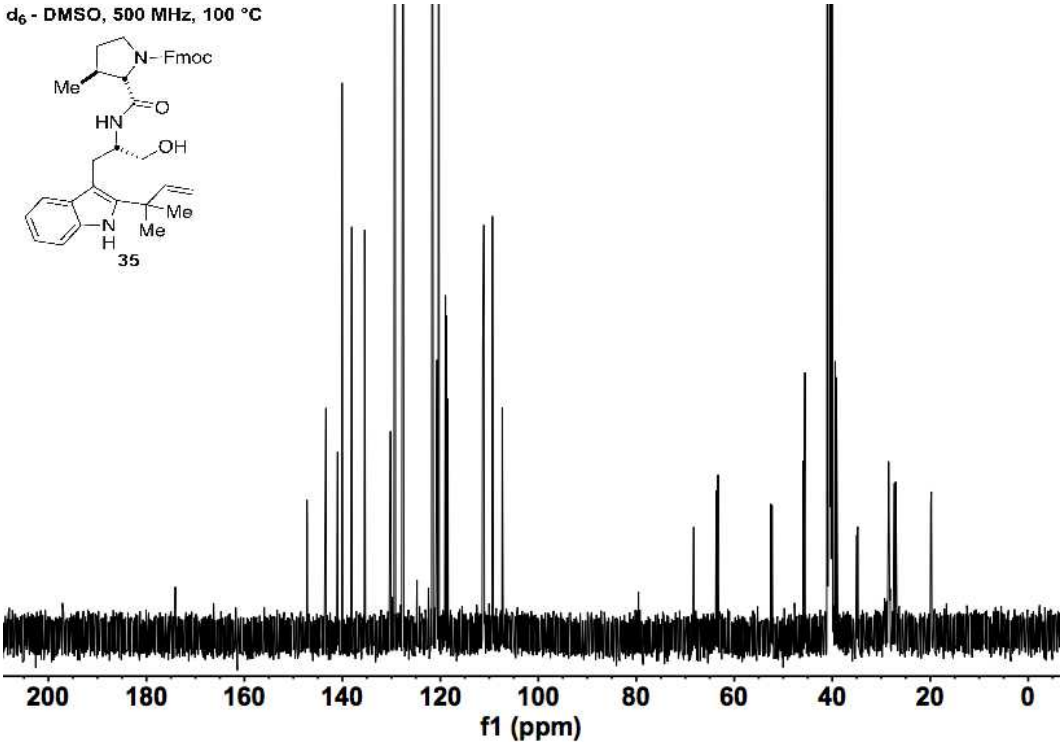
Supplementary Figure 51. ^{13}C NMR spectrum of **34**.



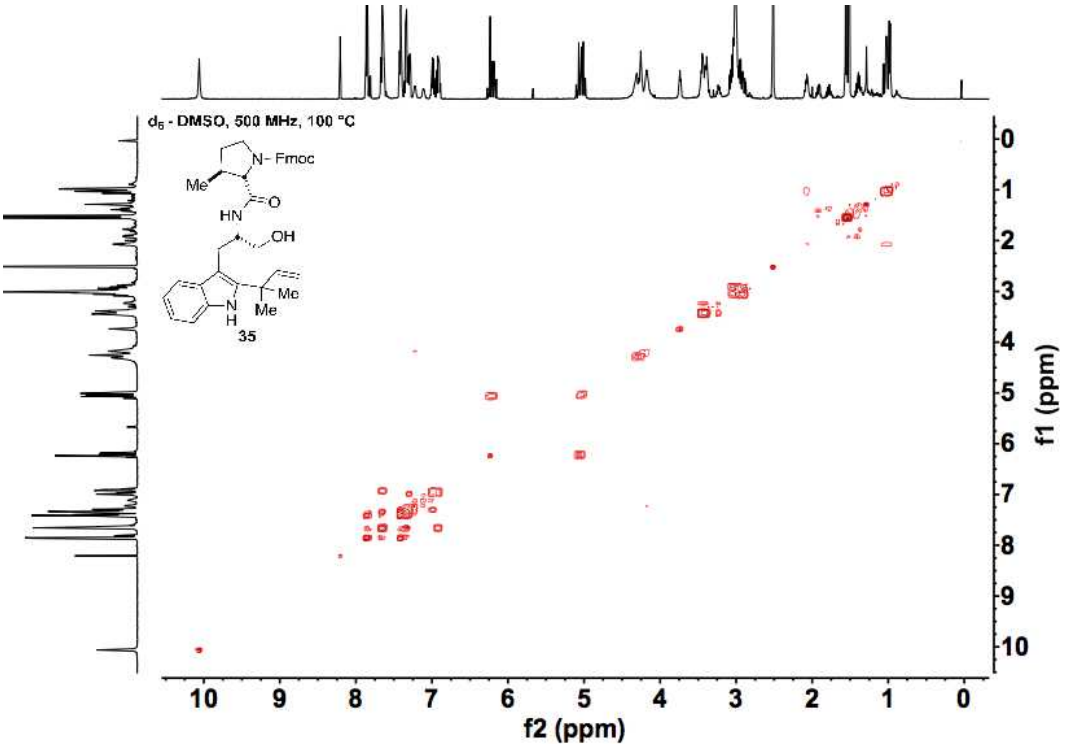
Supplementary Figure 52. ^{1}H-^{1}H COSY spectrum of **34**.



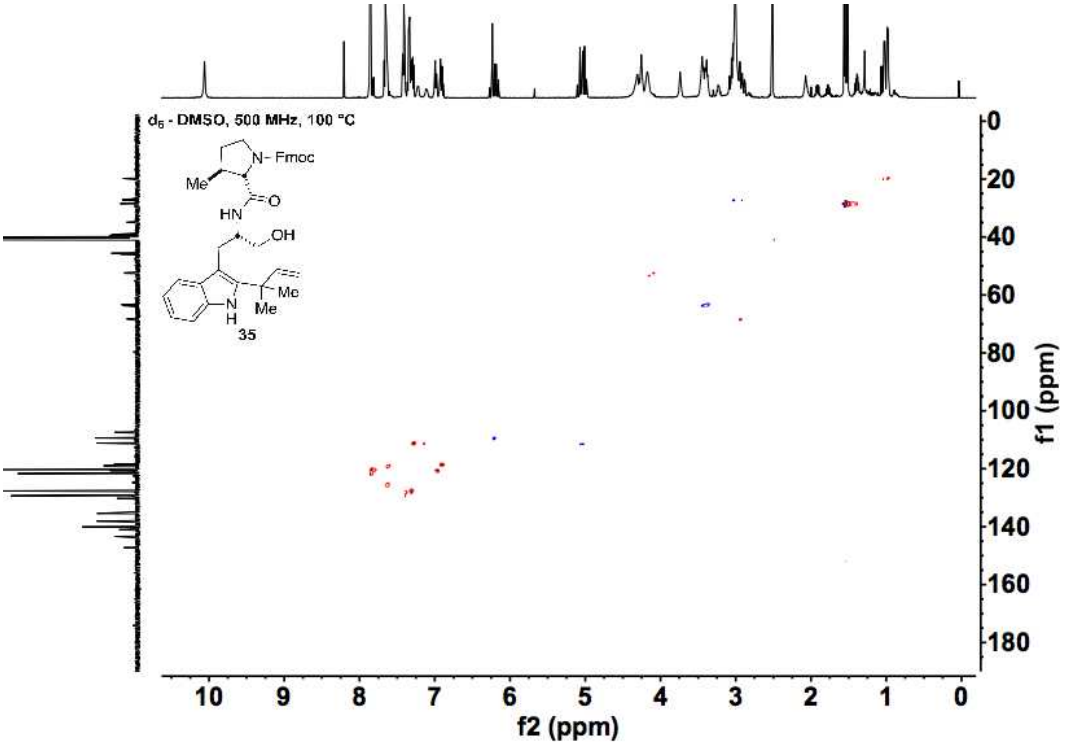
Supplementary Figure 53. ^{1}H NMR spectrum of **35**.



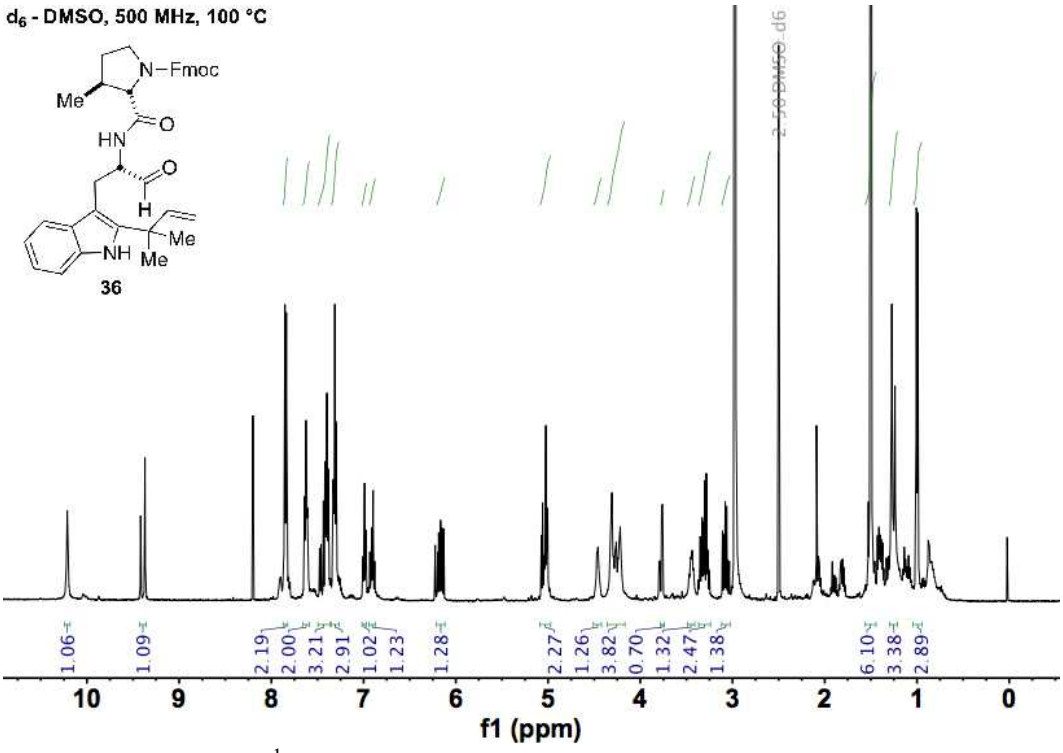
Supplementary Figure 54. ^{13}C NMR spectrum of **35**.



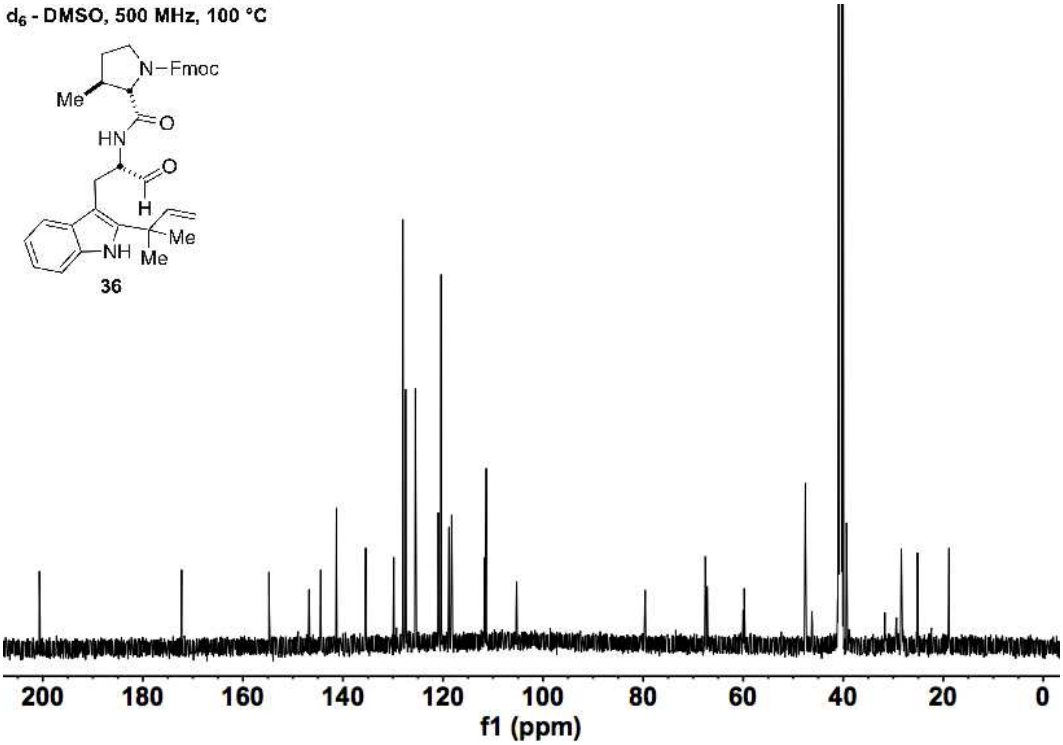
Supplementary Figure 55. ^{1}H-^{1}H COSY spectrum of **35**.



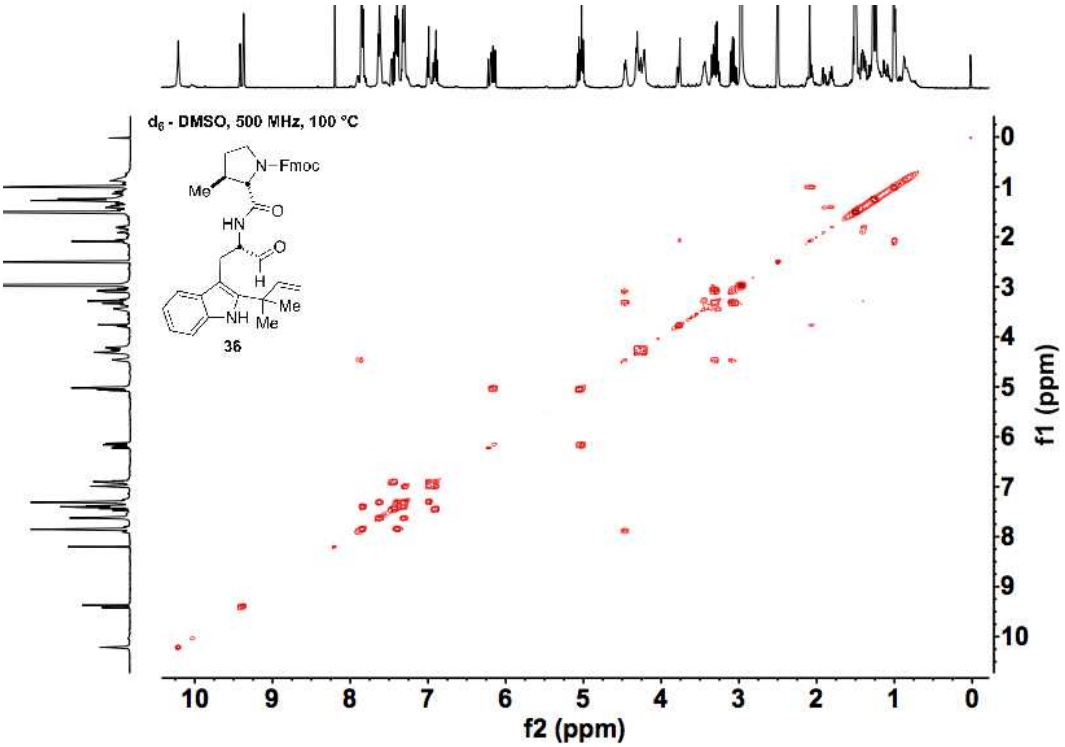
Supplementary Figure 56. ^{1}H-^{13}C HSQC spectrum of **35**.



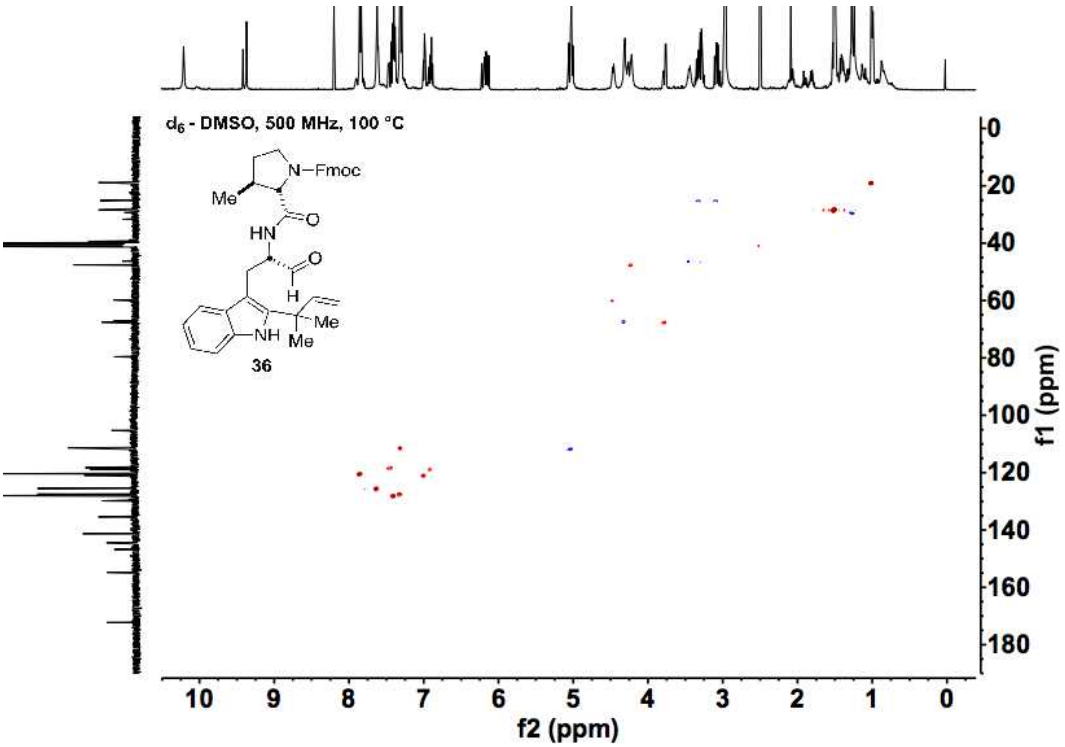
Supplementary Figure 57. ^{1}H NMR spectrum of **36**.



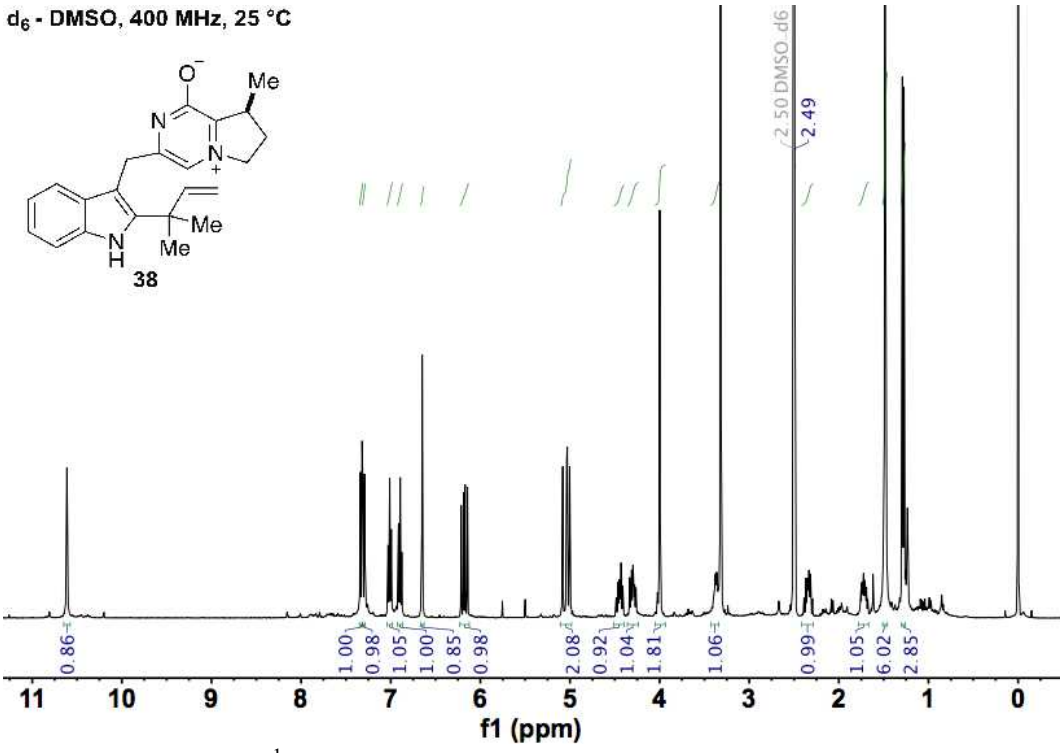
Supplementary Figure 58. ^{13}C NMR spectrum of **36**.



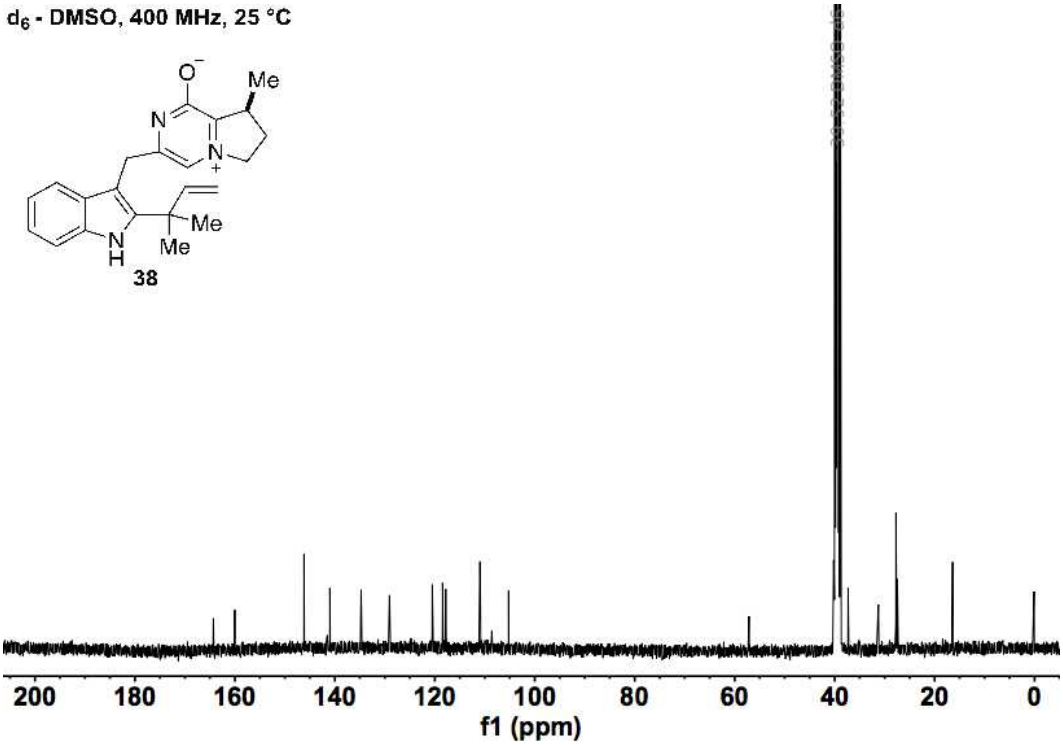
Supplementary Figure 59. ^{1}H-^{1}H COSY spectrum of **36**.



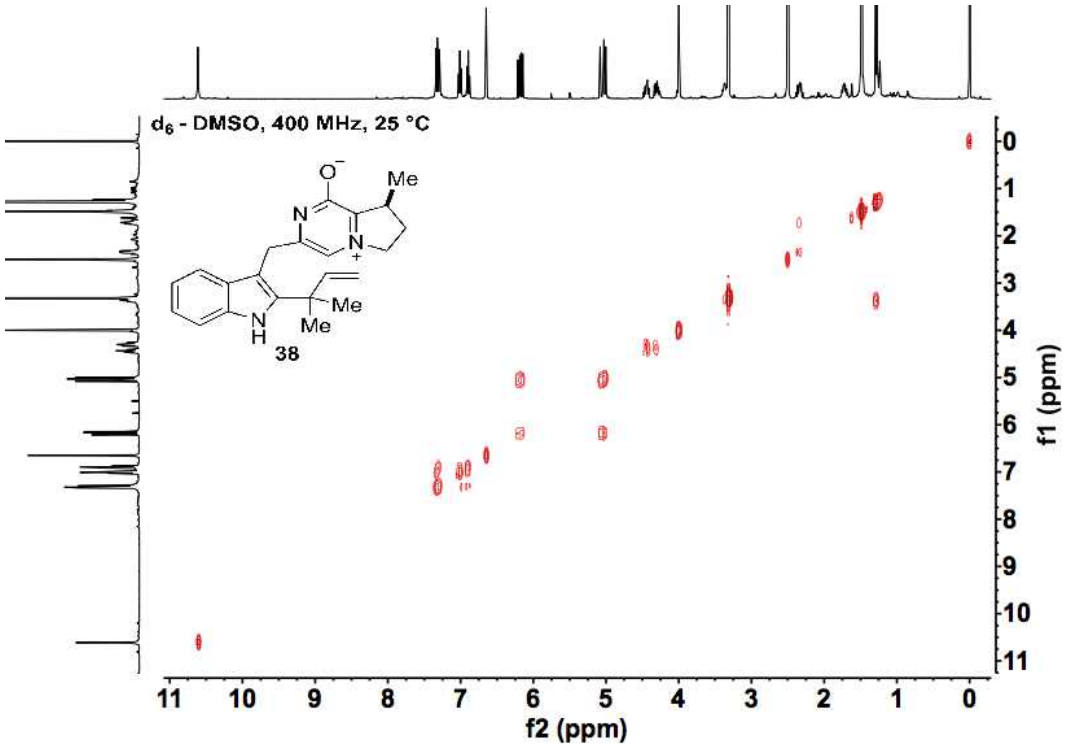
Supplementary Figure 60. ^{1}H-^{13}C HSQC spectrum of **36**.



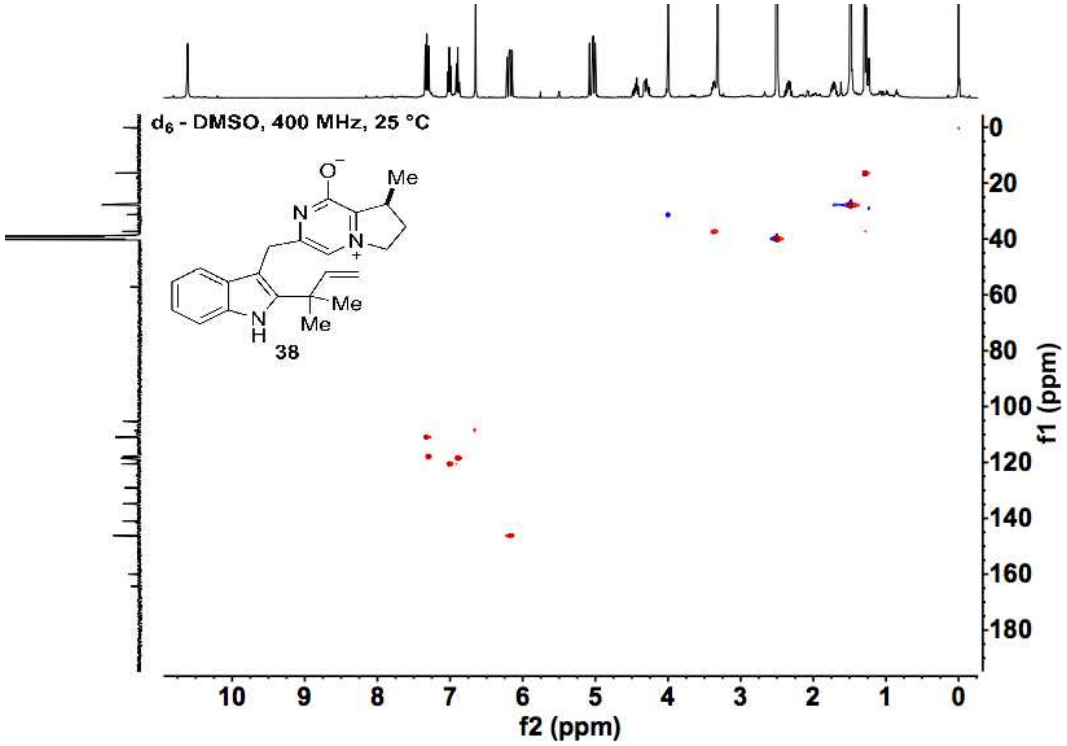
Supplementary Figure 61. ^{1}H NMR spectrum of **38**.



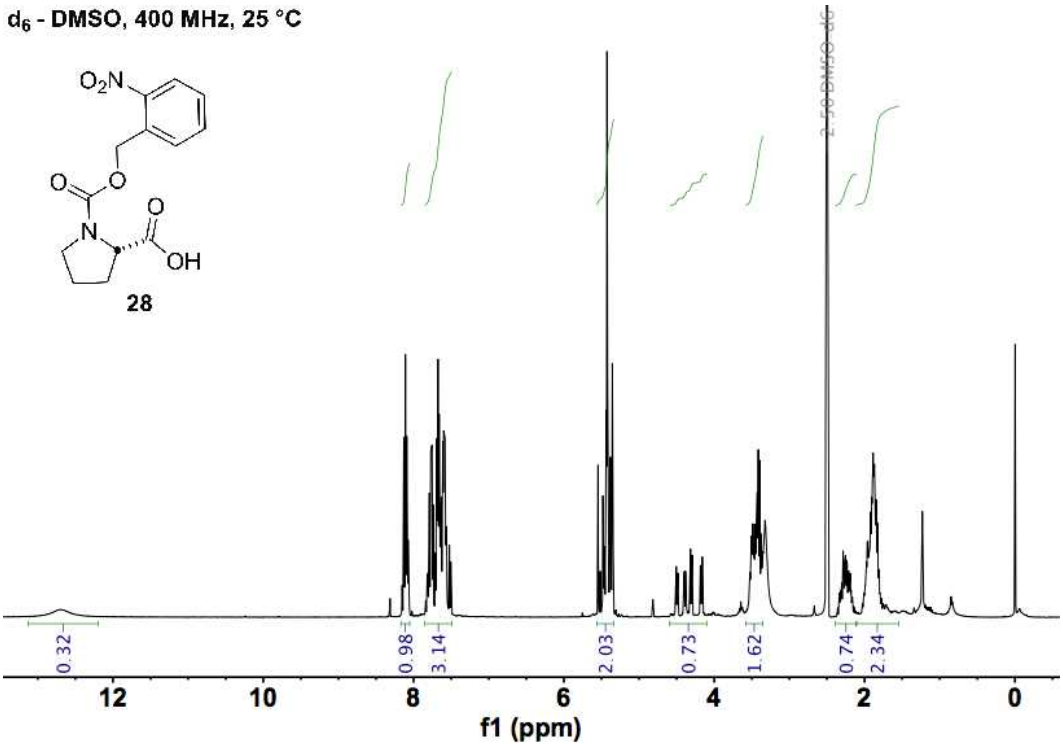
Supplementary Figure 62. ^{13}C NMR spectrum of **38**.



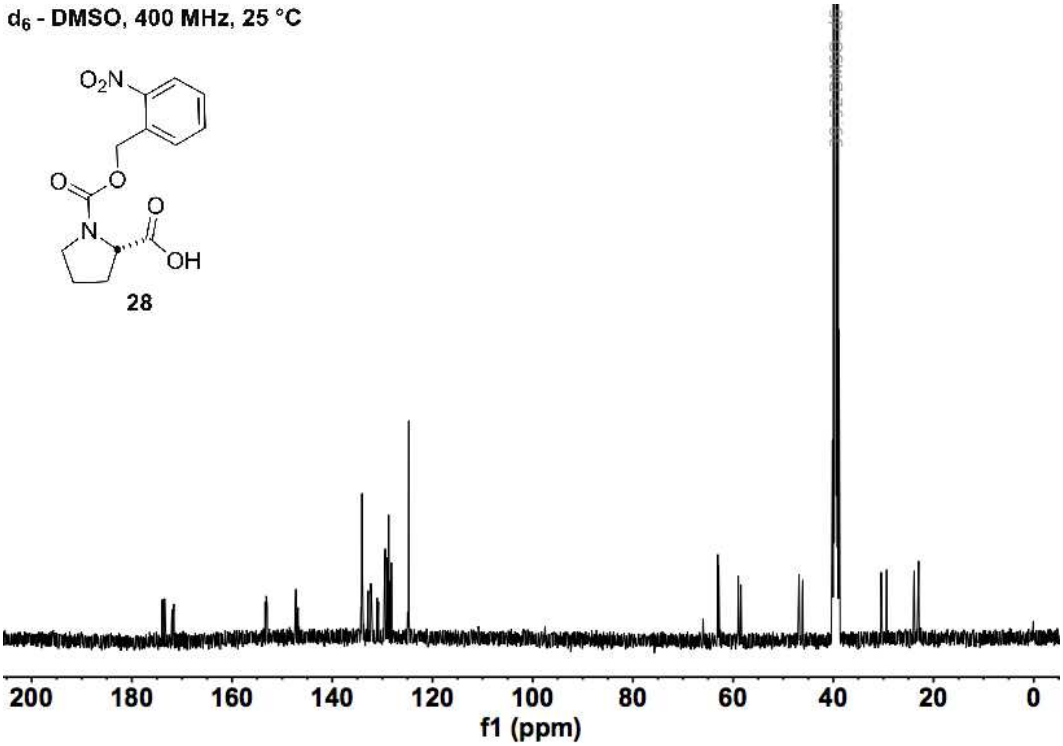
Supplementary Figure 63. ^{1}H-^{1}H COSY spectrum of **38**.



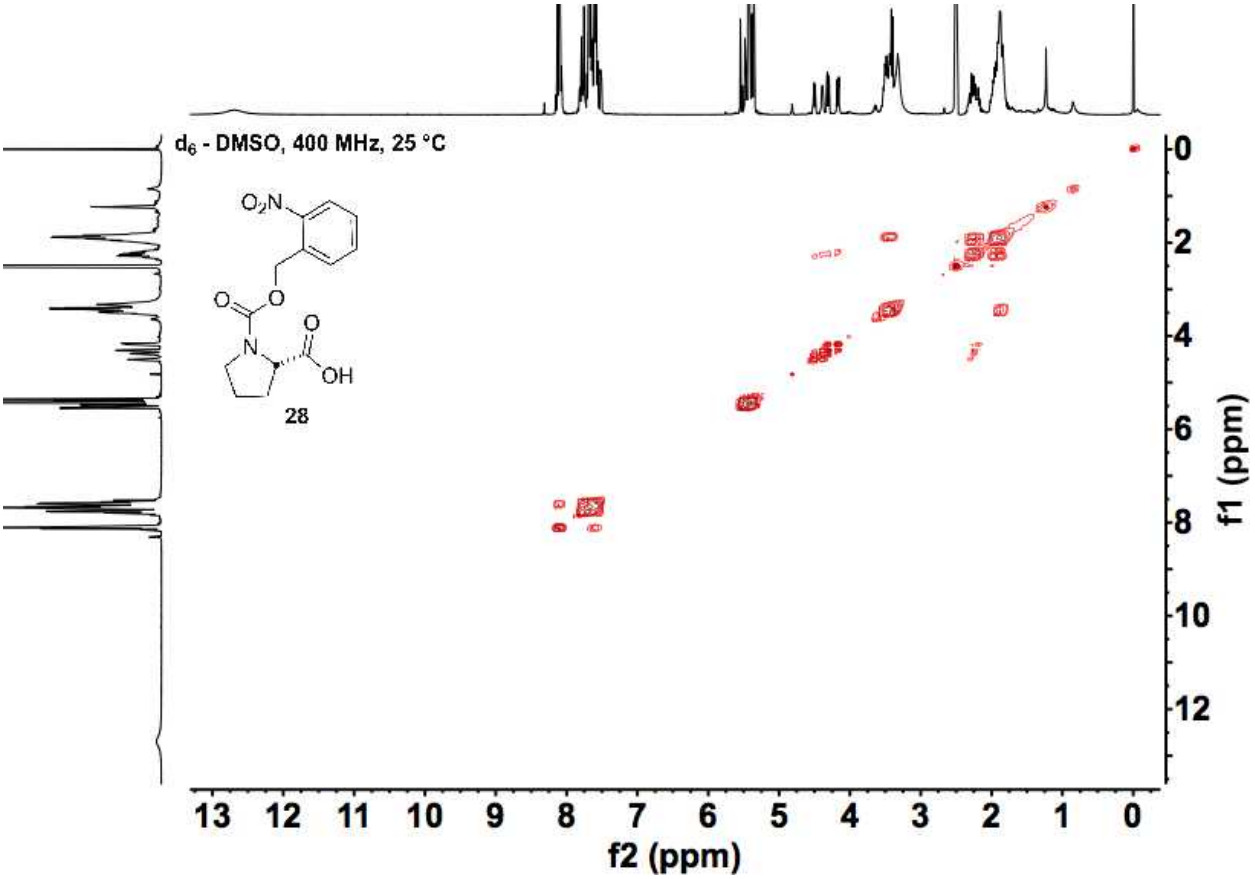
Supplementary Figure 64. ^{1}H-^{13}C HSQC spectrum of **38**.



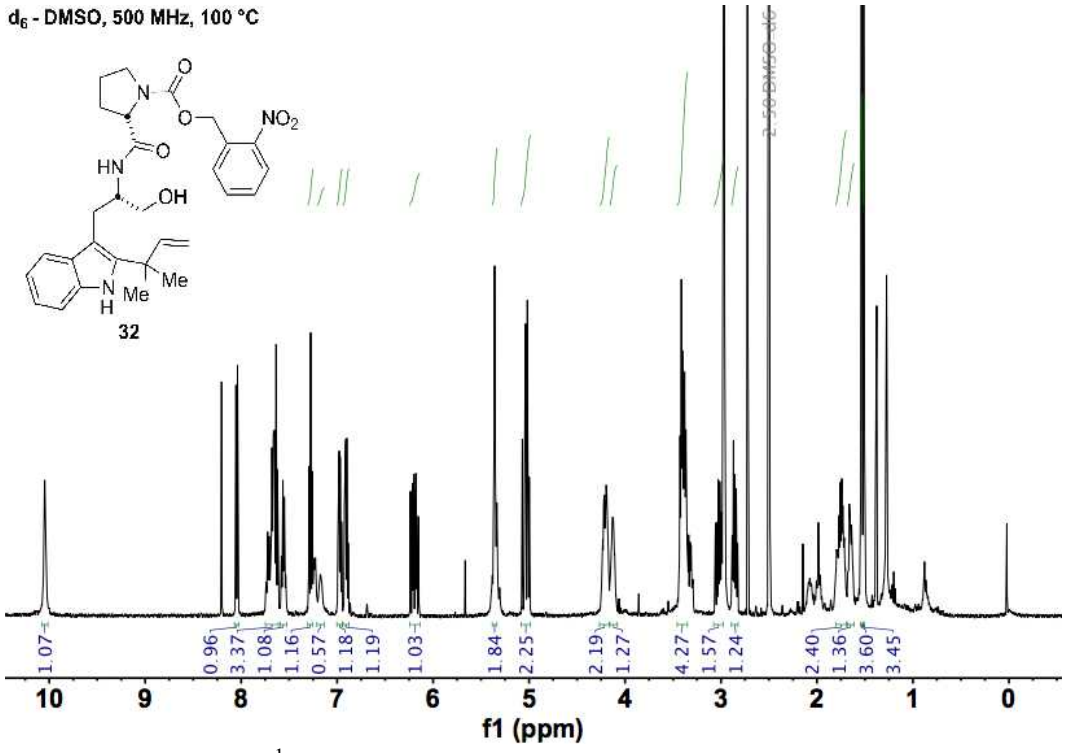
Supplementary Figure 65. ^{1}H NMR spectrum of **28**.



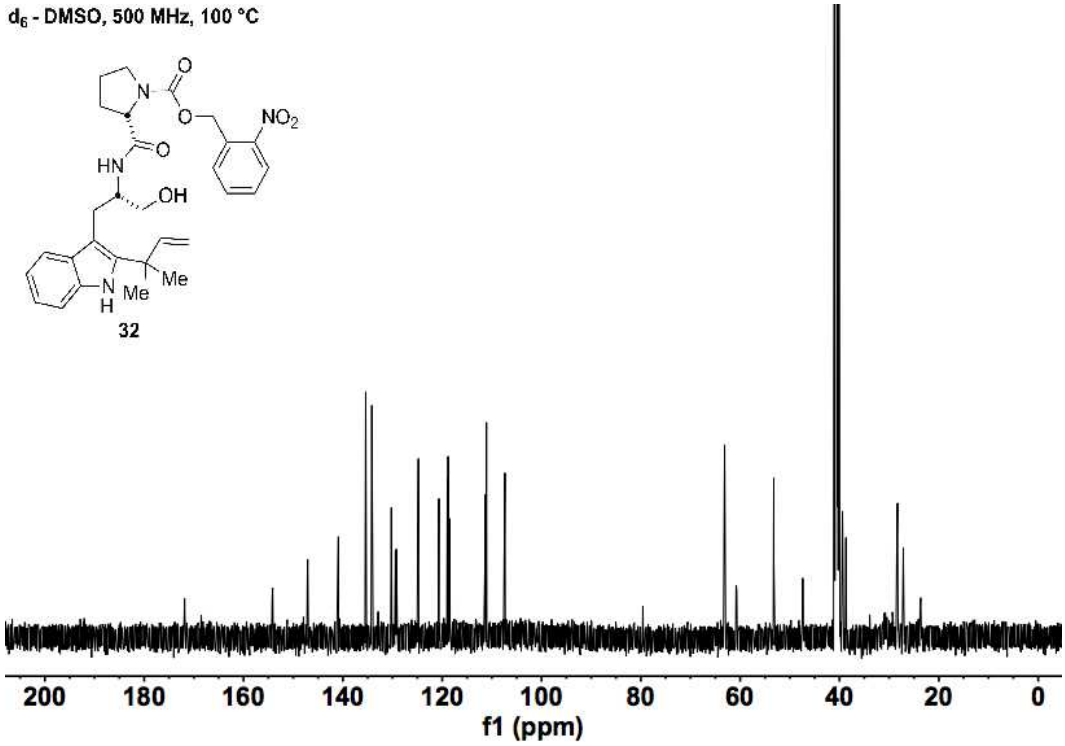
Supplementary Figure 66. ^{13}C NMR spectrum of **28**.



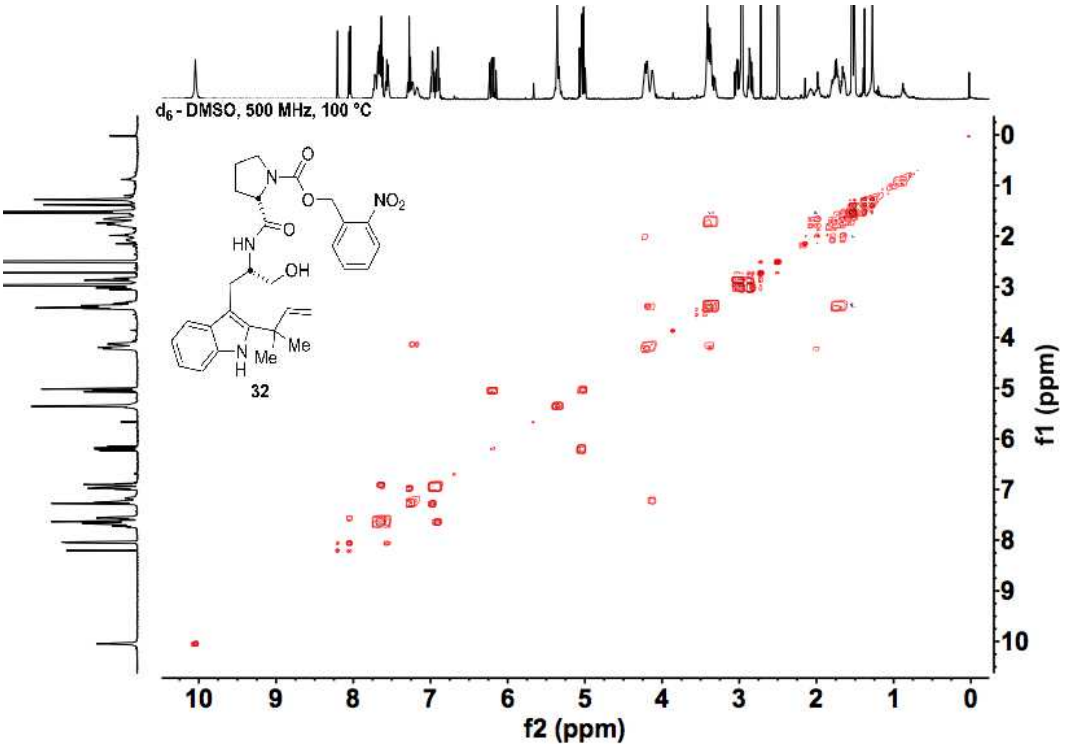
Supplementary Figure 67. ^{1}H-^{1}H COSY spectrum of **28**.



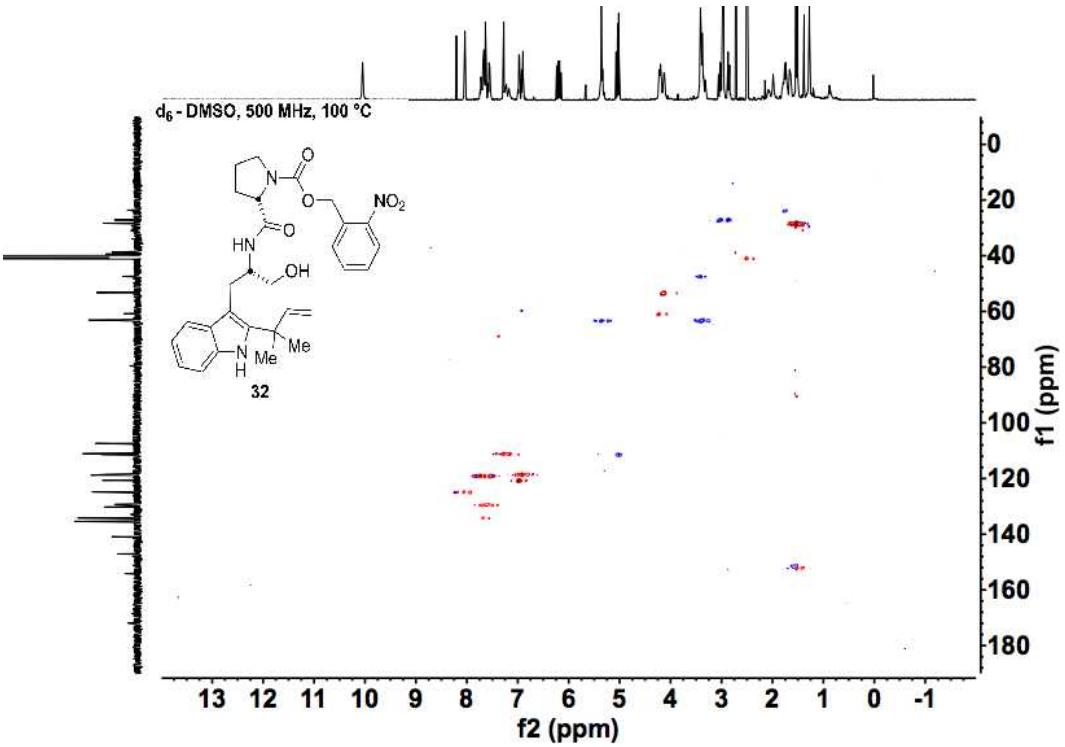
Supplementary Figure 68. ^{1}H NMR spectrum of **32**.



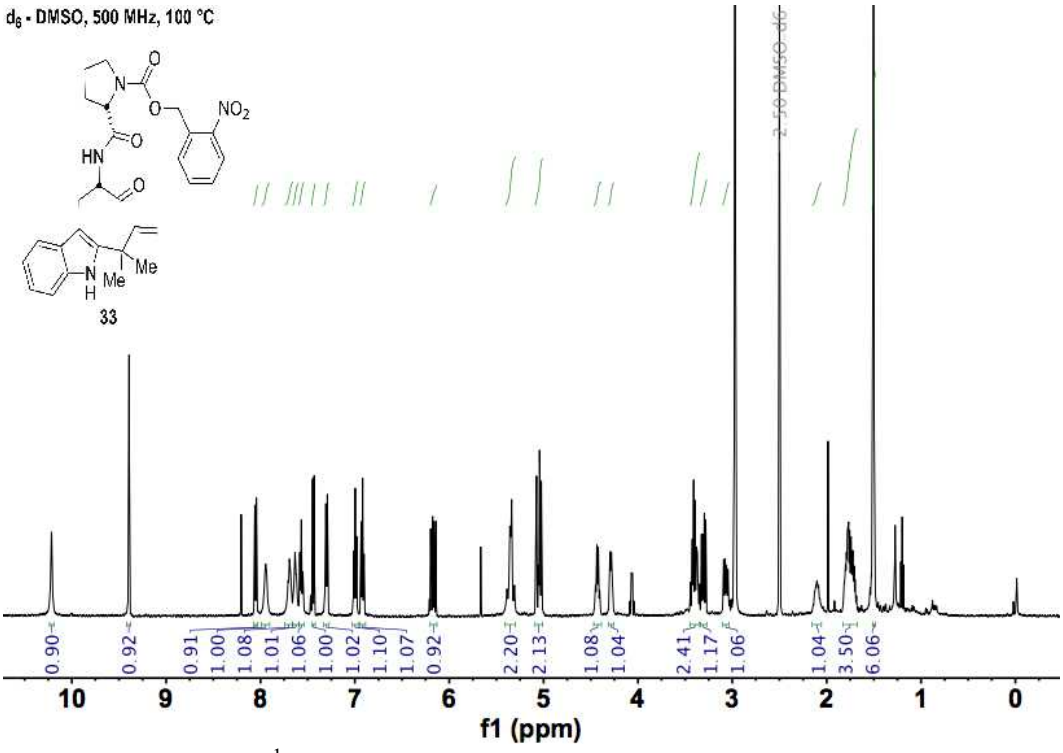
Supplementary Figure 69. ^{13}C NMR spectrum of **32**.



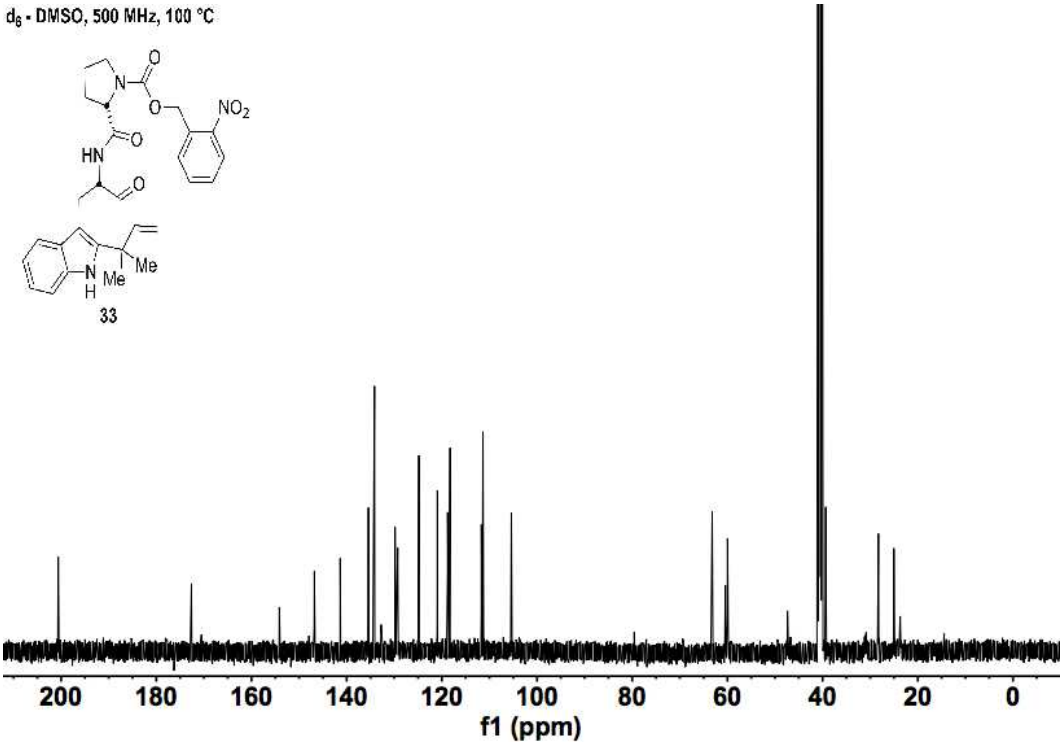
Supplementary Figure 70. ^{1}H-^{1}H COSY spectrum of **32**.



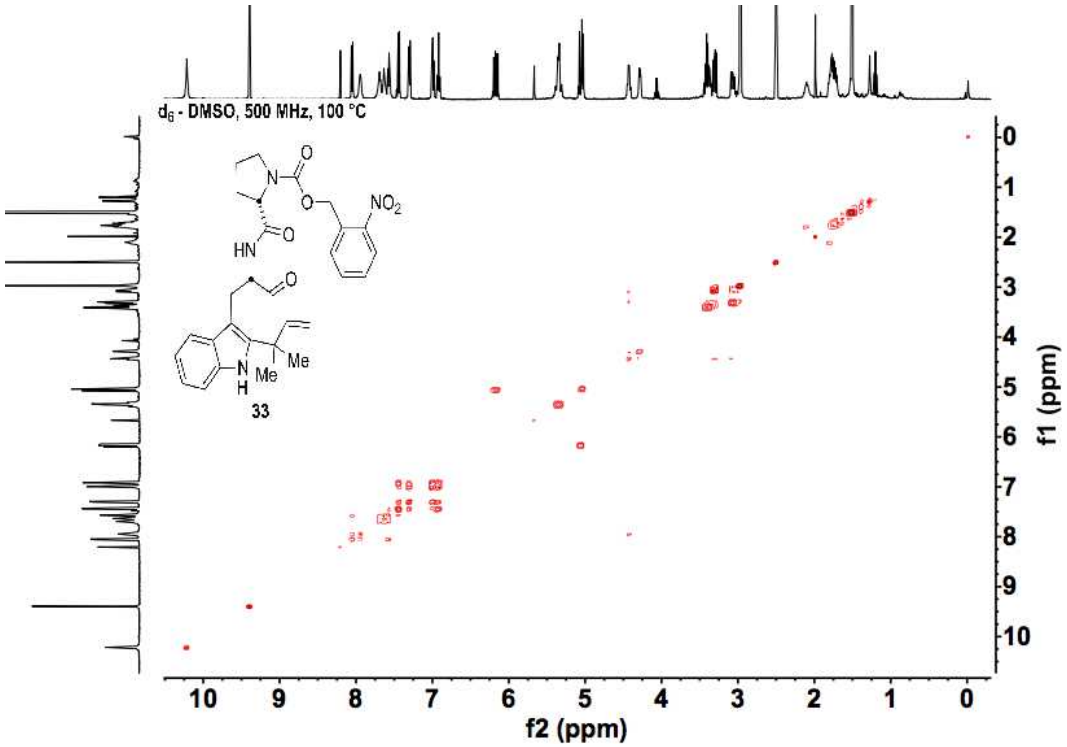
Supplementary Figure 71. ^{1}H-^{13}C HSQC spectrum of **32**.



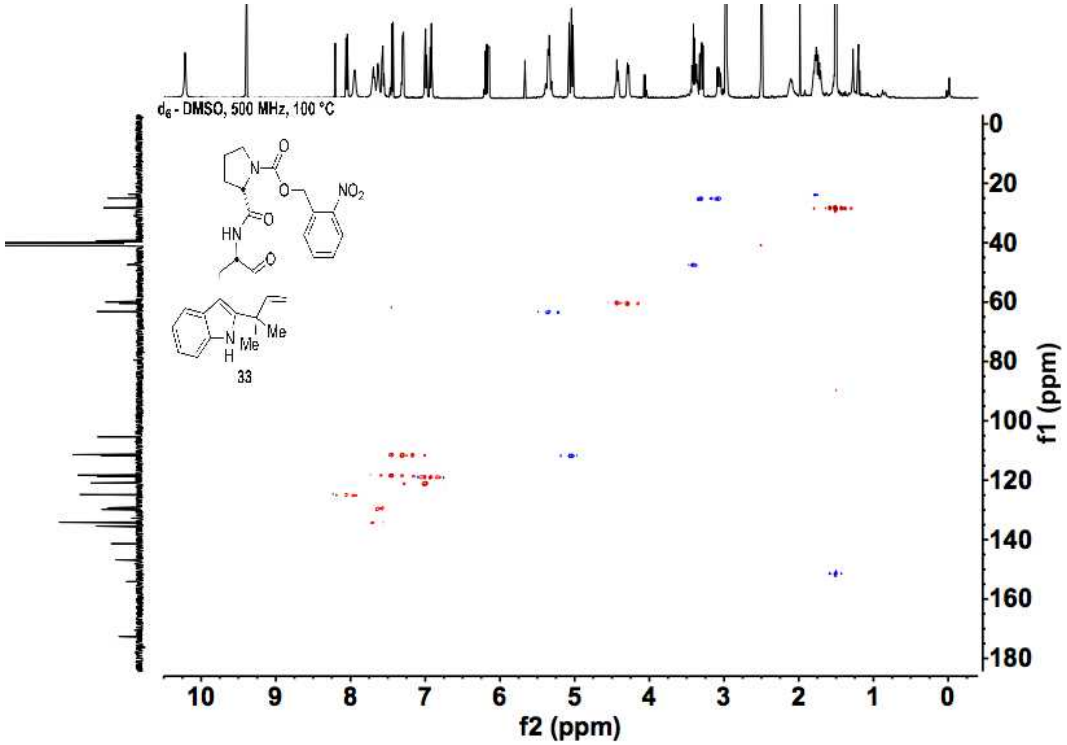
Supplementary Figure 72. ^{1}H NMR spectrum of **33**.



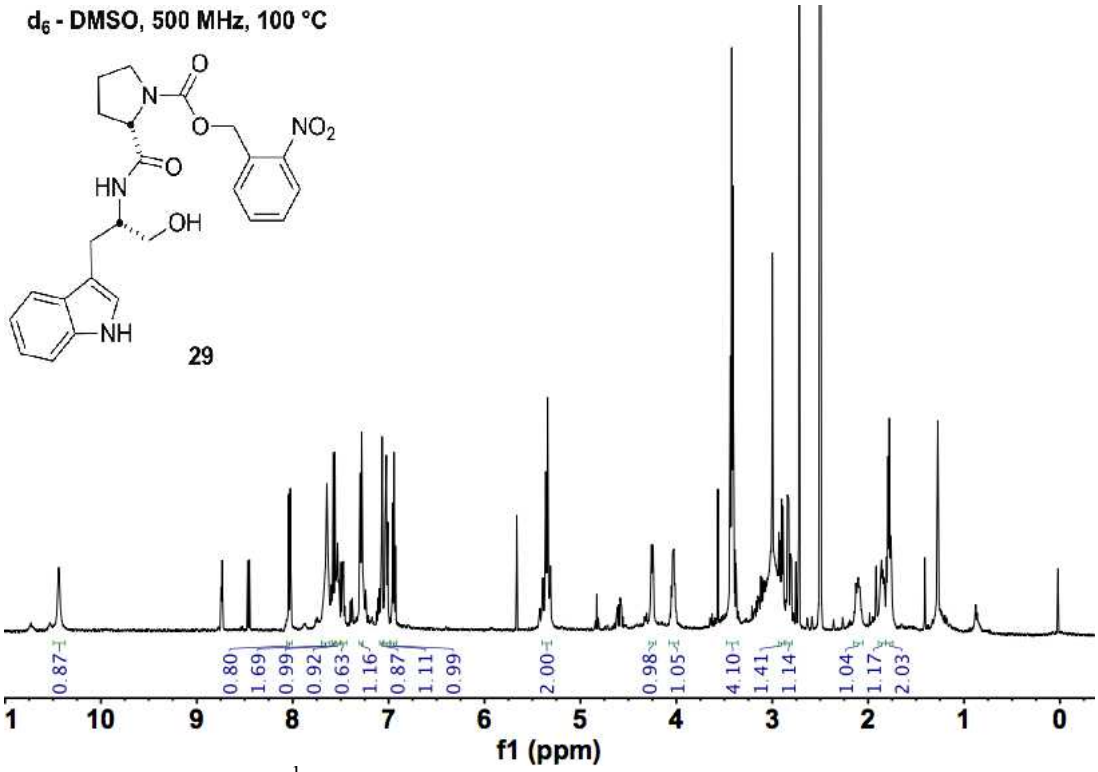
Supplementary Figure 73. ^{13}C NMR spectrum of **33**.



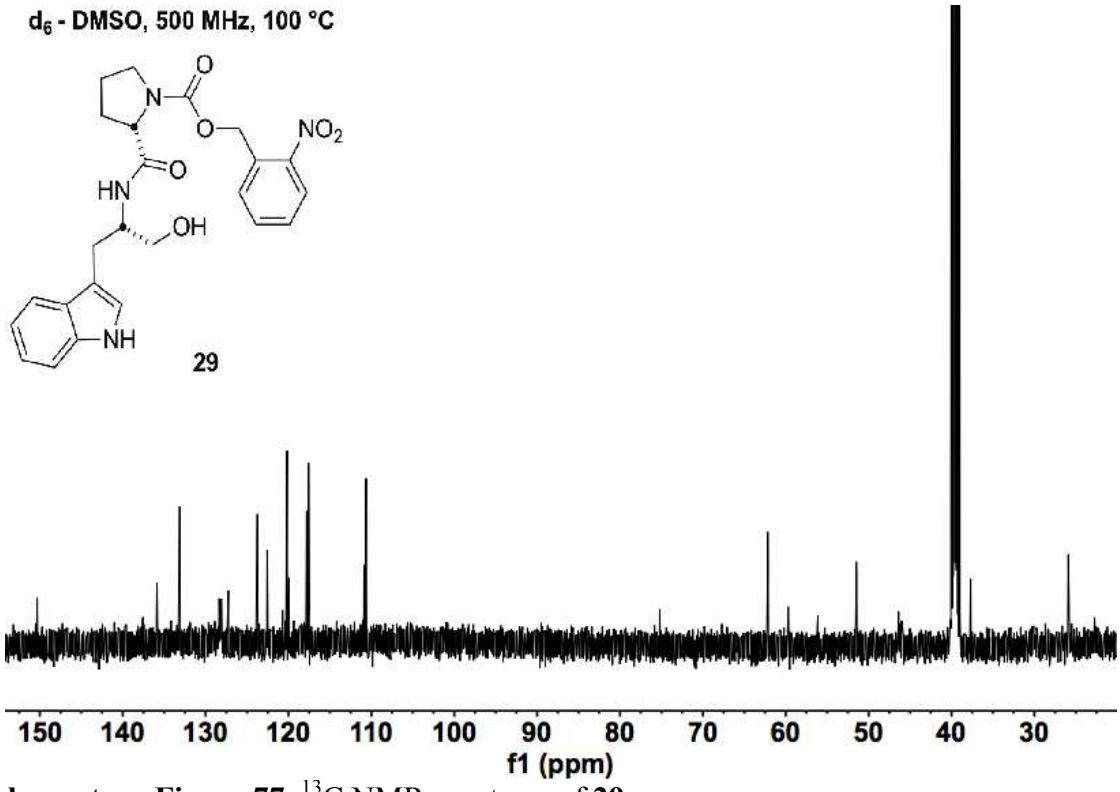
Supplementary Figure 74. ^{1}H-^{1}H COSY spectrum of **33**.



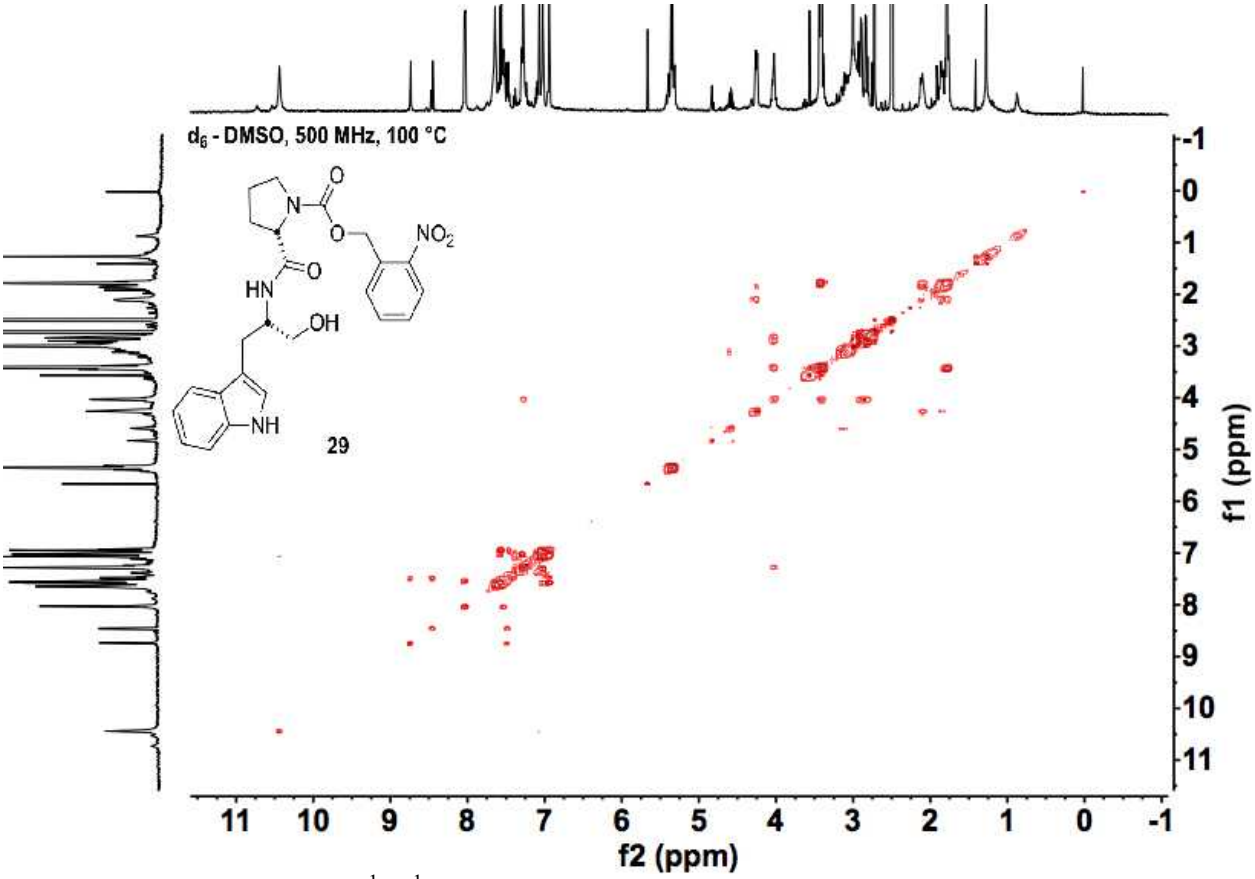
Supplementary Figure 75. ^{1}H-^{13}C HSQC spectrum of **33**.



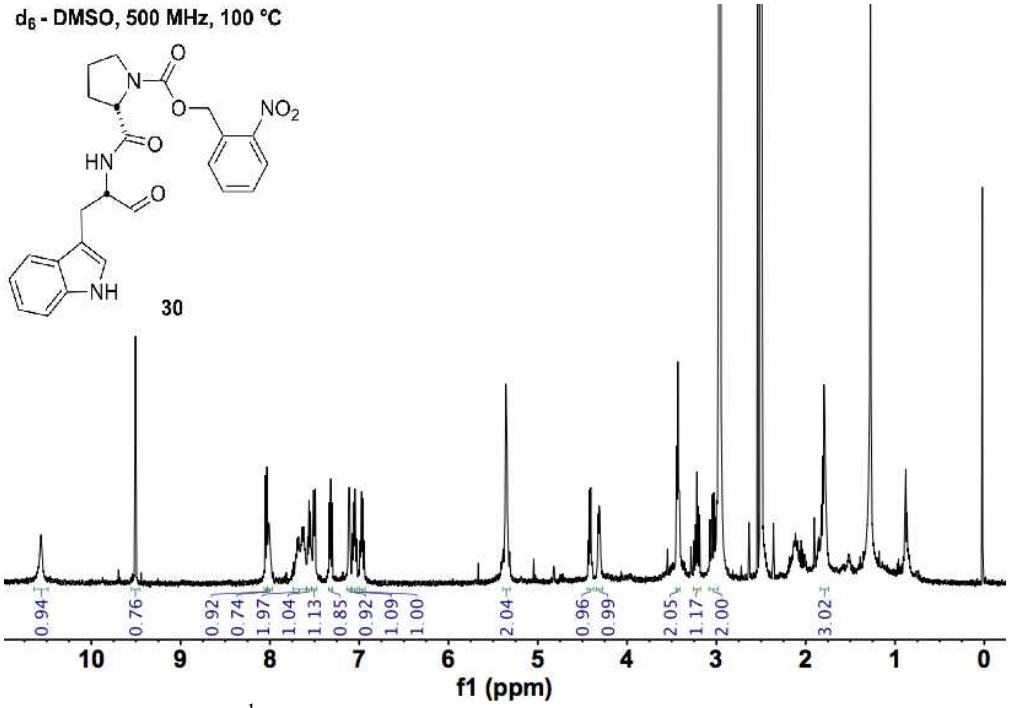
Supplementary Figure 76. ^{1}H NMR spectrum of **29**.



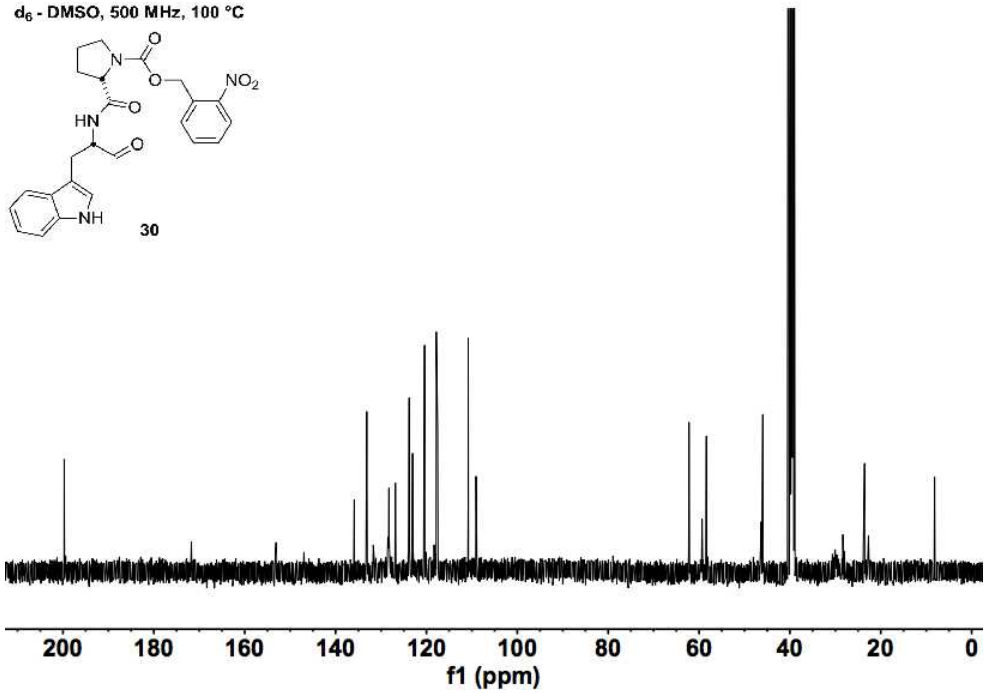
Supplementary Figure 77. ^{13}C NMR spectrum of **29**.



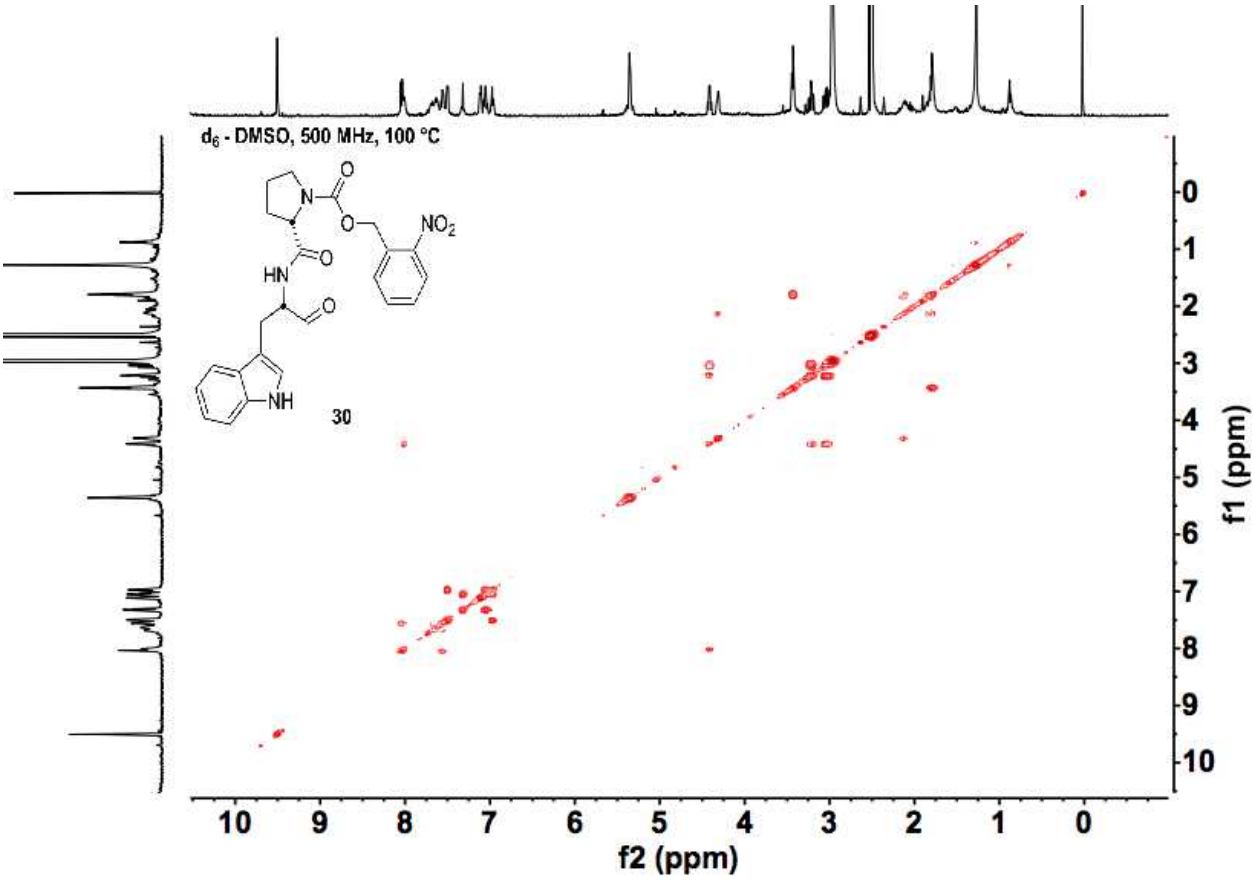
Supplementary Figure 78. ^{1}H-^{1}H COSY spectrum of **29**.



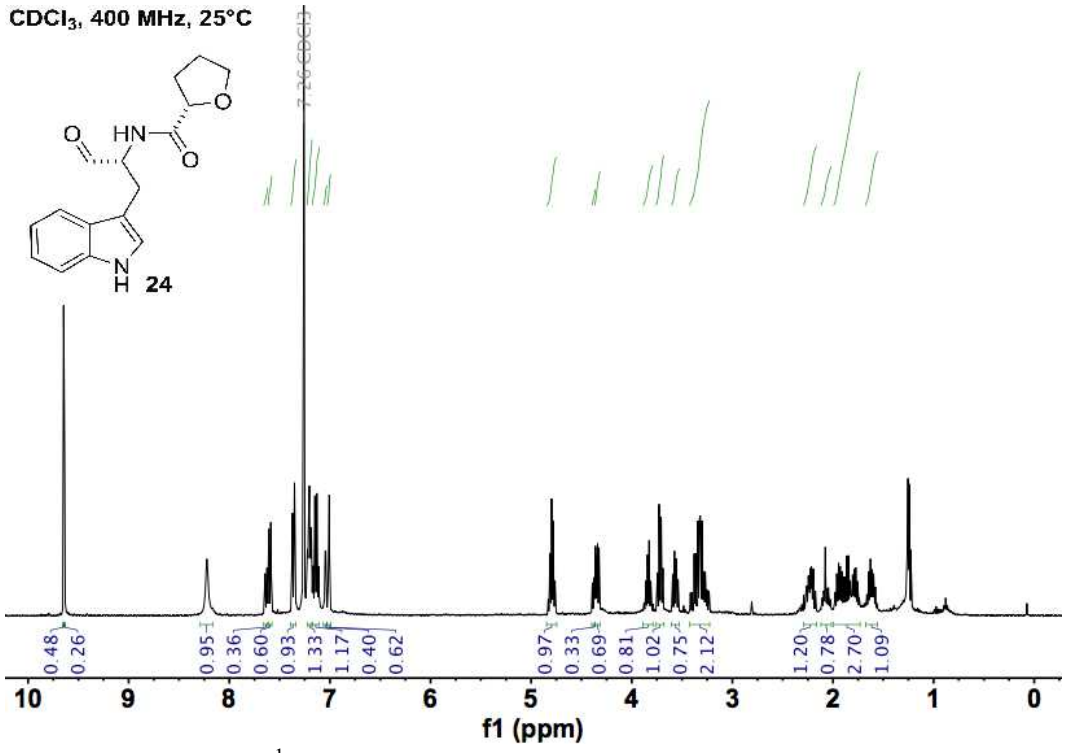
Supplementary Figure 79. ^{1}H NMR spectrum of **30**.



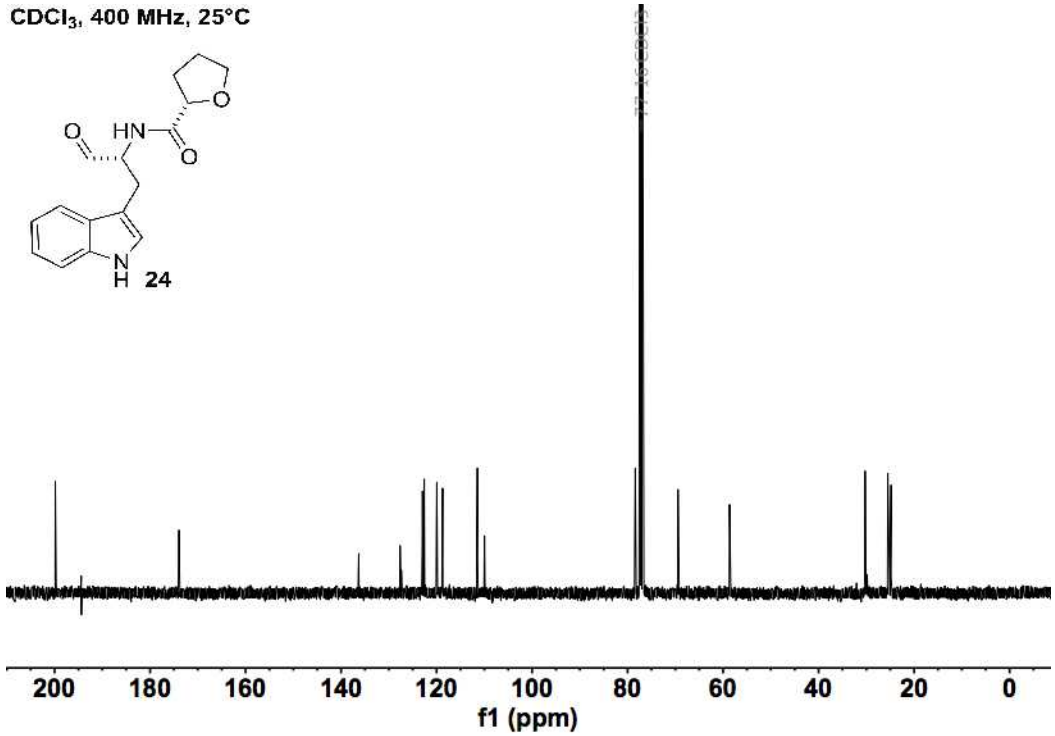
Supplementary Figure 80. ^{13}C NMR spectrum of **30**.



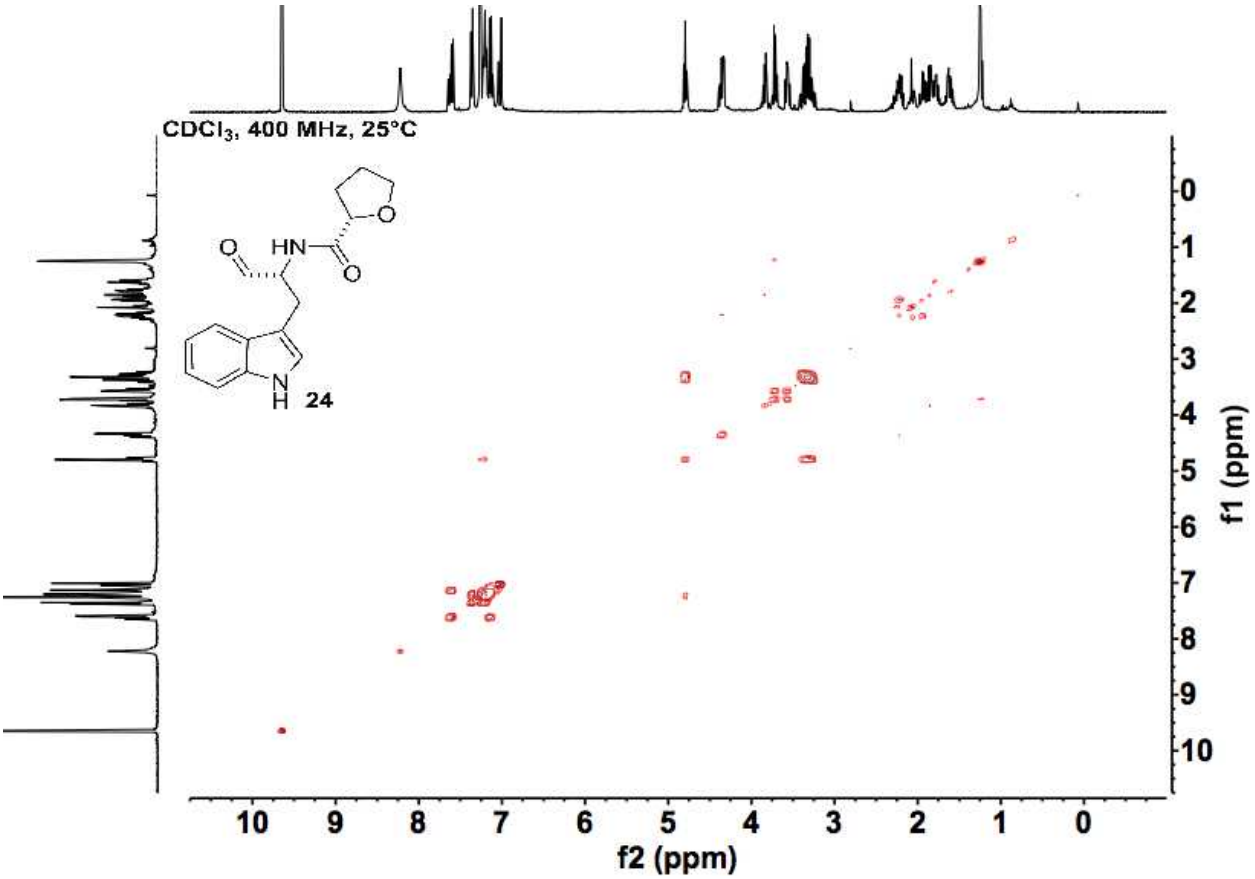
Supplementary Figure 81. ^{1}H-^{1}H COSY spectrum of **30**.



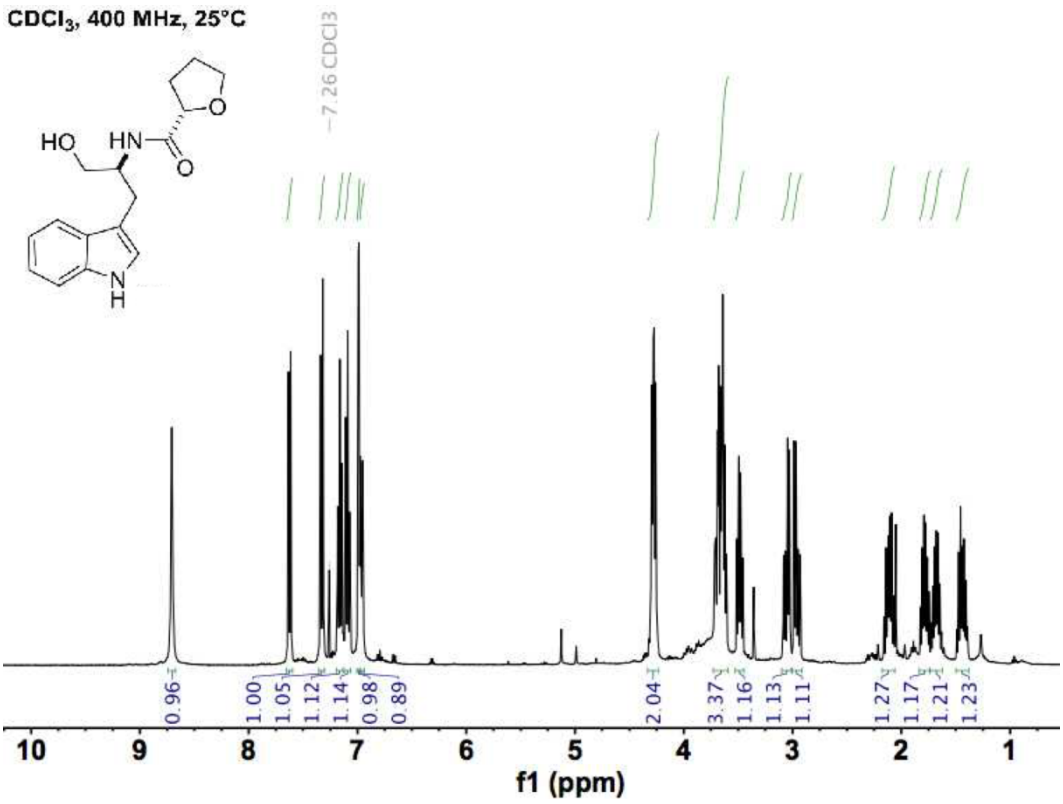
Supplementary Figure 82. ^{1}H NMR spectrum of **24**.



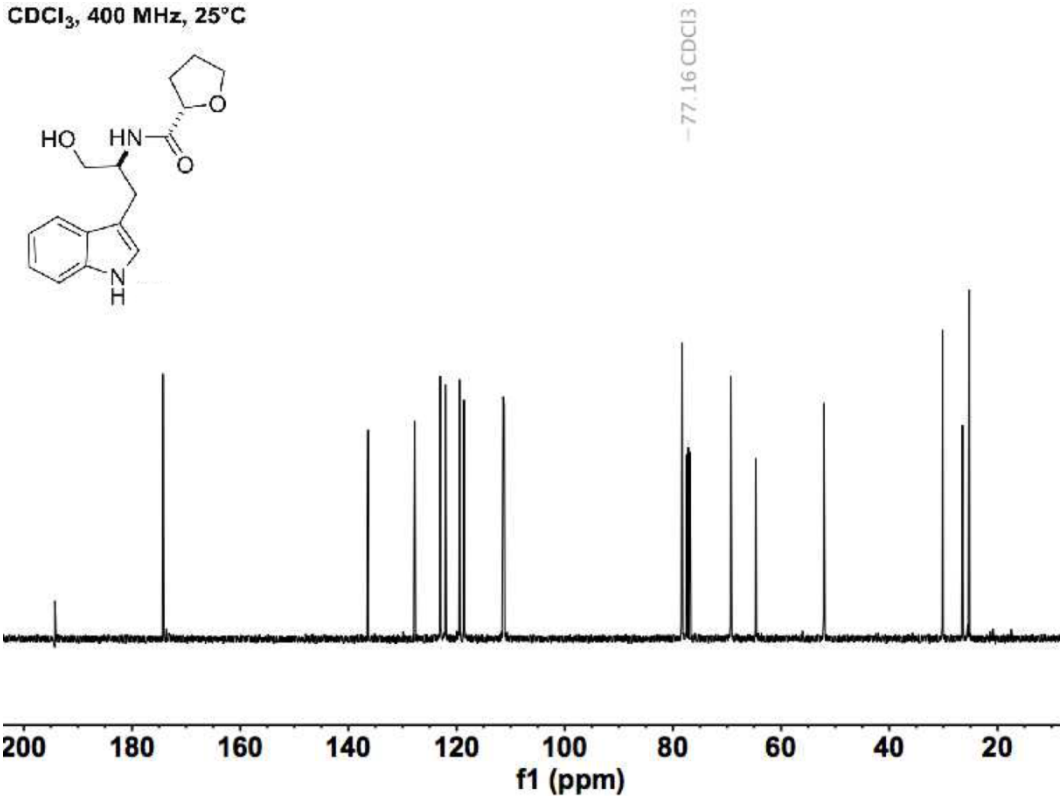
Supplementary Figure 83. ^{13}C NMR spectrum of **24**.



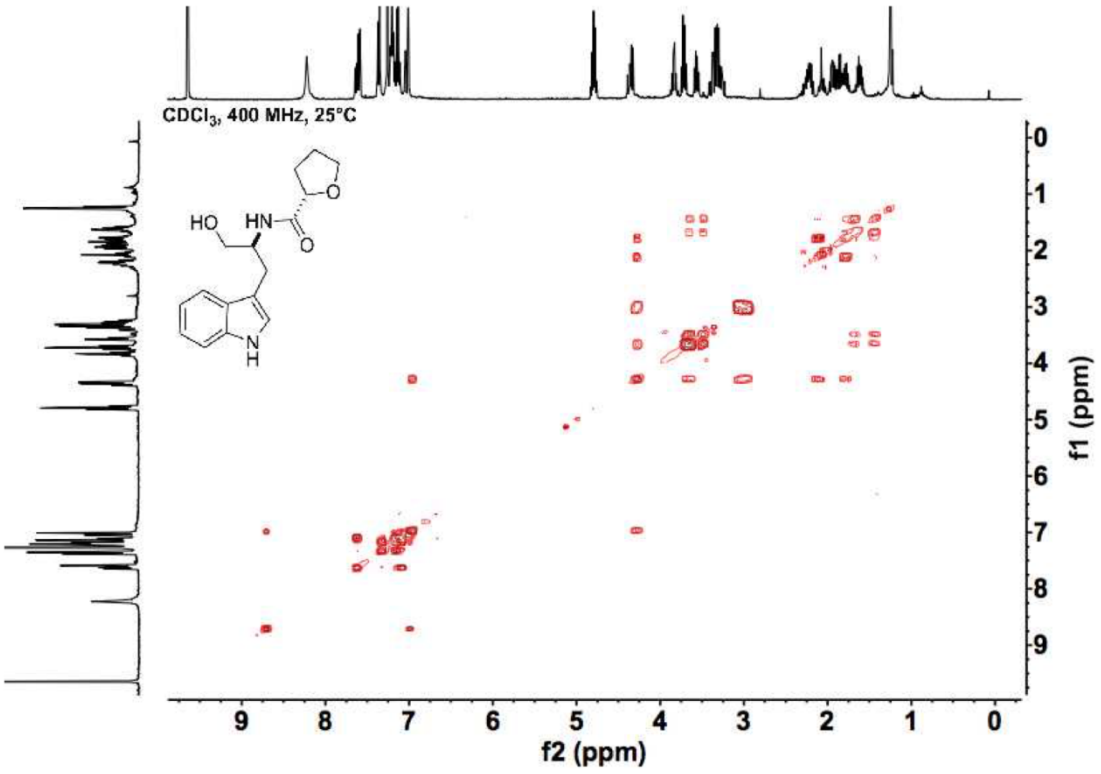
Supplementary Figure 84. ^{1}H-^{1}H COSY spectrum of **24**.



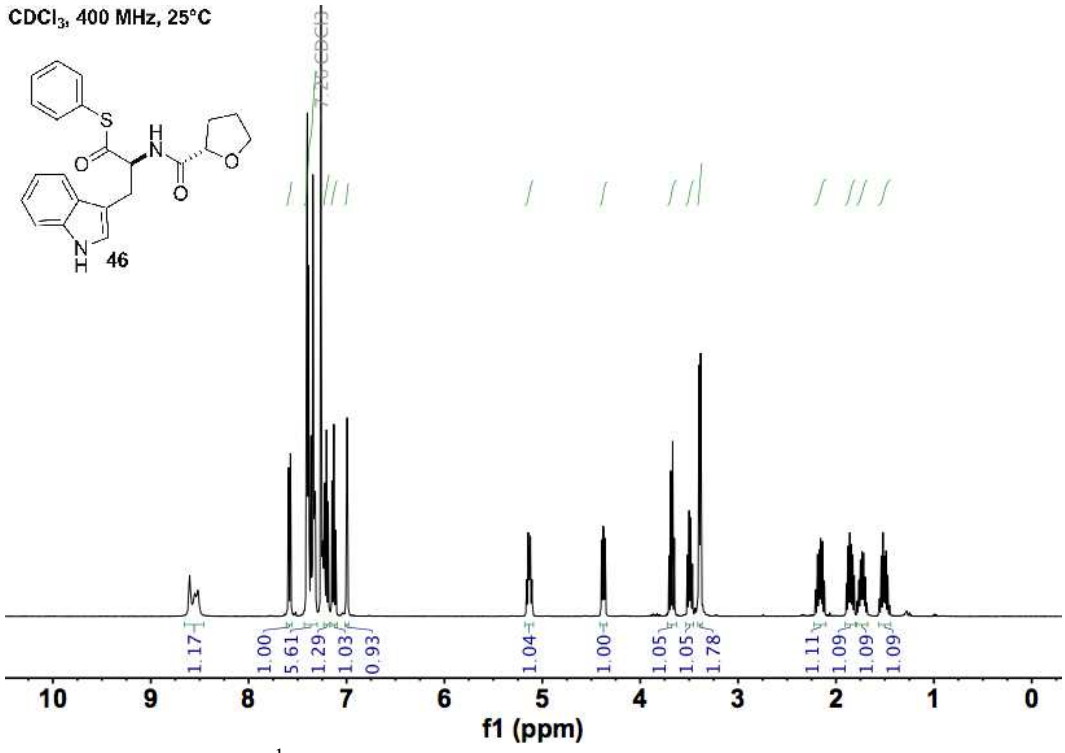
Supplementary Figure 85. ^{1}H NMR spectrum of **25**.



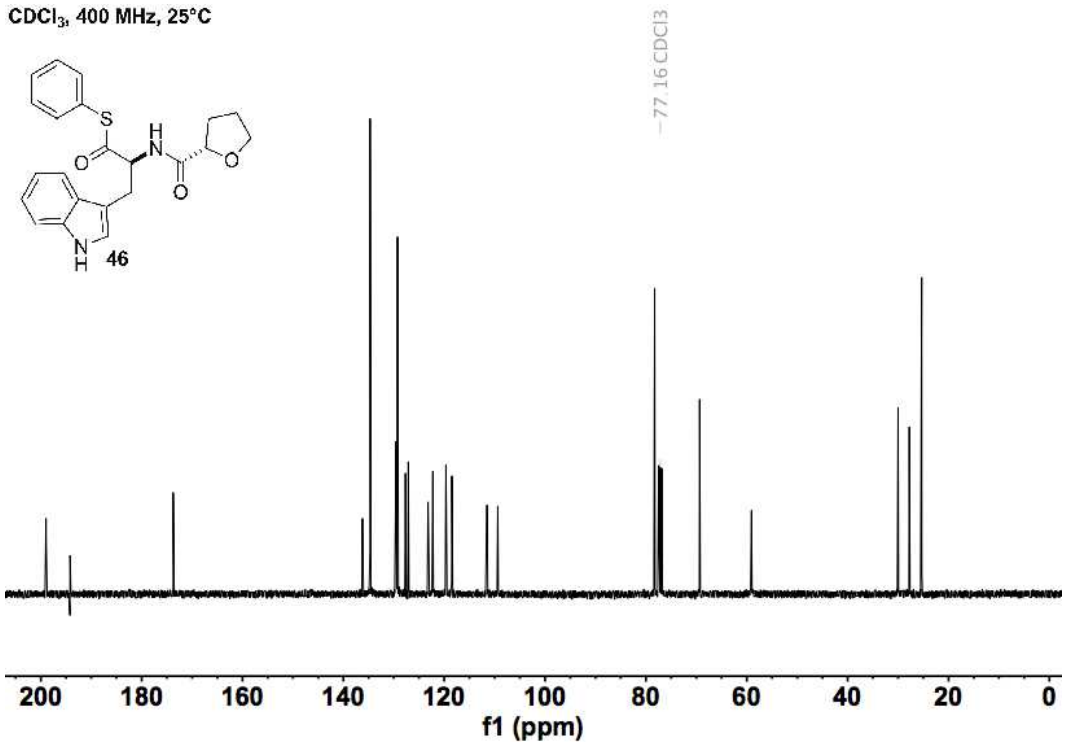
Supplementary Figure 86. ^{13}C NMR spectrum of **25**.



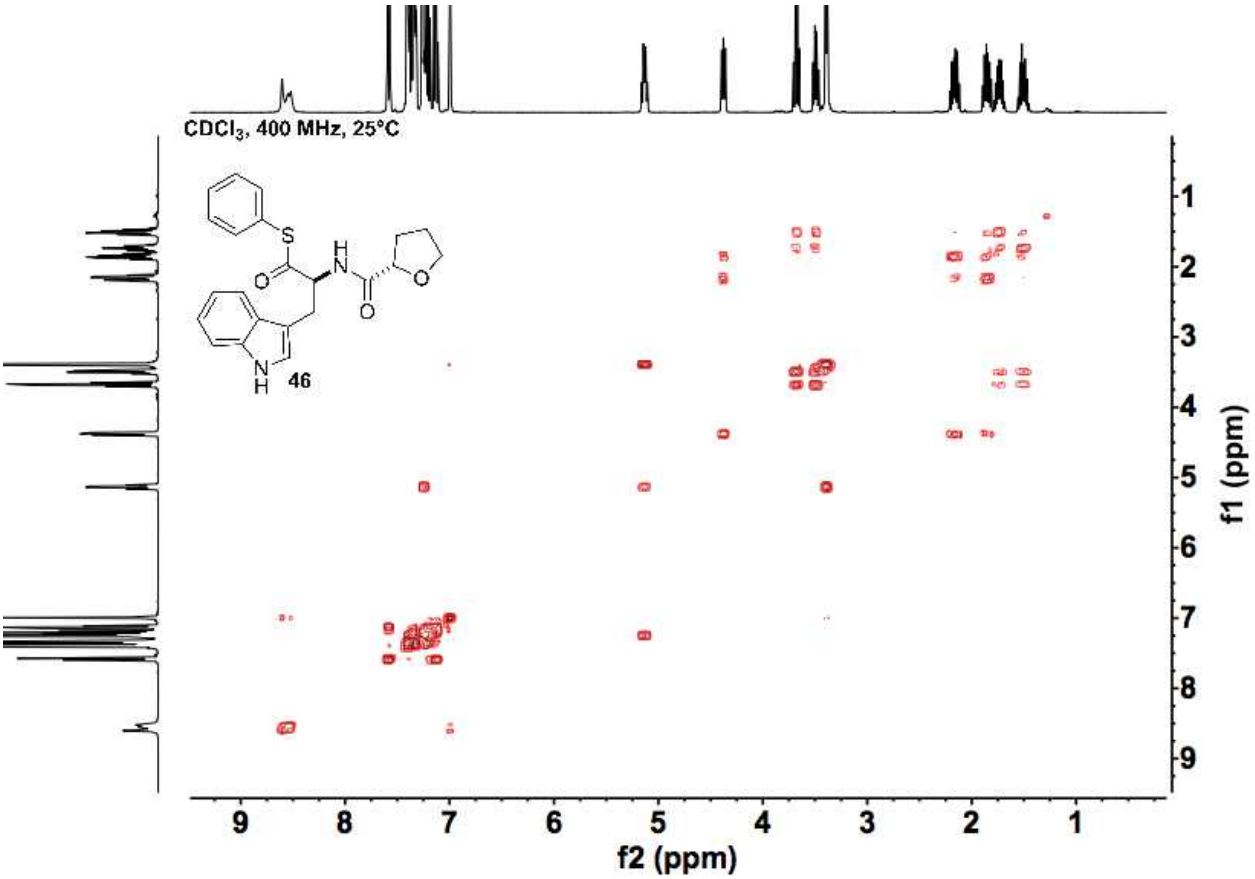
Supplementary Figure 87. ^{1}H-^{1}H COSY spectrum of **25**.



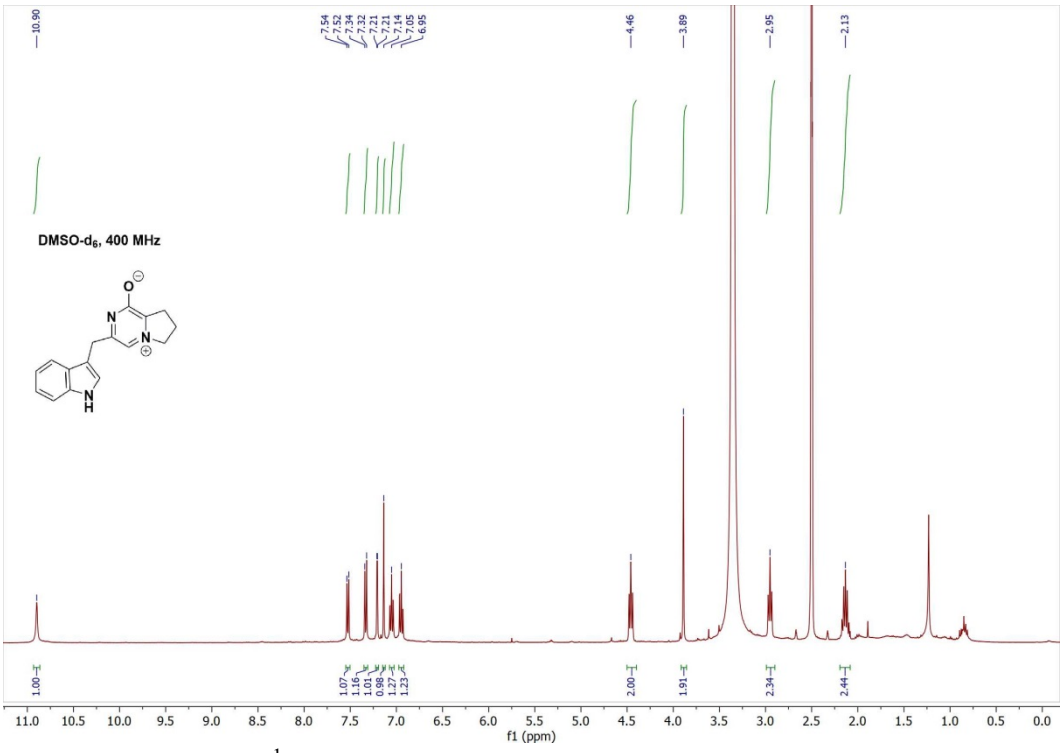
Supplementary Figure 88. ^{1}H NMR spectrum of **46**.



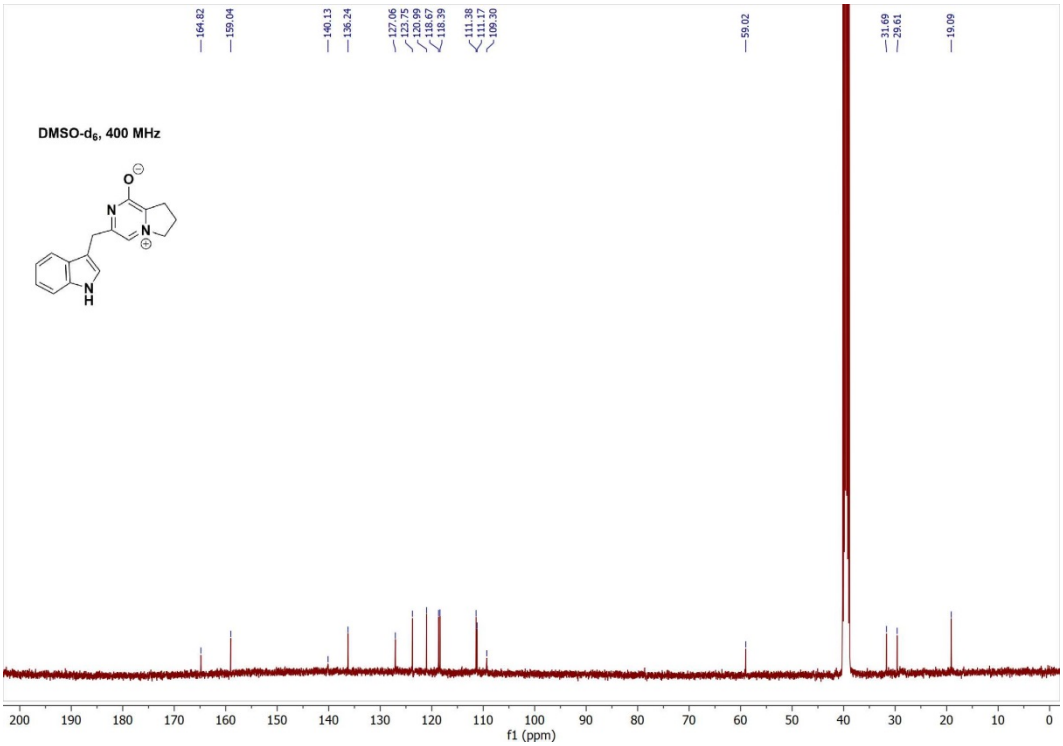
Supplementary Figure 89. ^{13}C NMR spectrum of **46**.



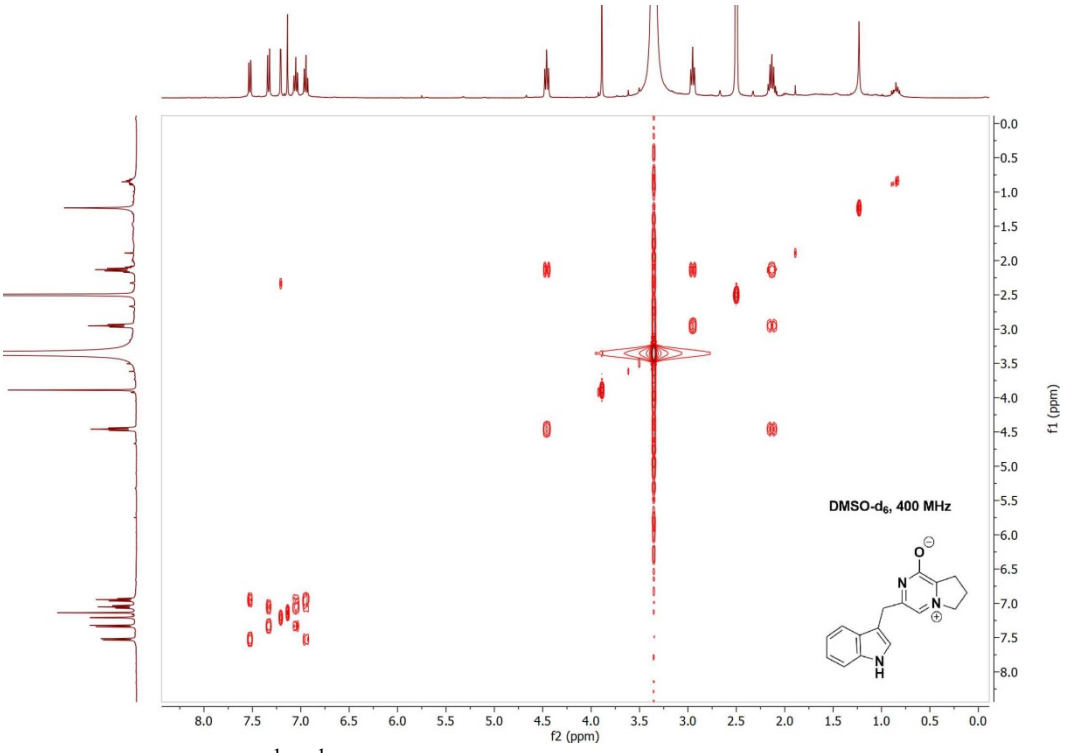
Supplementary Figure 90. ^{1}H-^{1}H COSY spectrum of **46**.



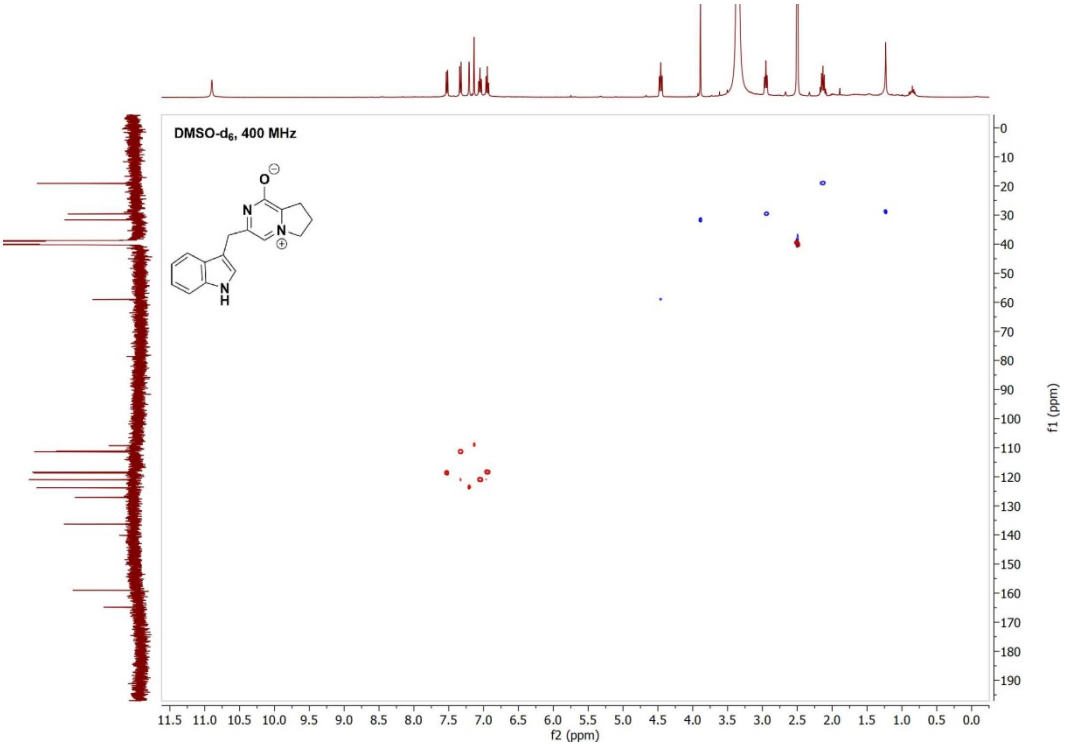
Supplementary Figure 91. ^{1}H NMR spectrum of **10**.



Supplementary Figure 92. ^{13}C NMR spectrum of **10**.



Supplementary Figure 93. ^{1}H-^{1}H COSY spectrum of **10**.



Supplementary Figure 94. ^{1}H-^{13}C HSQC spectrum of **10**.

Supplementary Table 1. Oligonucleotides used in this study

Gene	Primer Direction	Primer Sequence
malG A_1-T_1 (198 – 838)	Forward	5'-TACTTCCAATCCAATGCCTTGATGTGTGAGTCCGATATCGAA-3'
	Reverse	5'-TTATCCACTTCCAATGCTAAGCAGAGATCATGGTTCCAGCA-3'
malG C (846 – 1277)	Forward	5'-TACTTCCAATCCAATGCCTCAAAGACATCATTCGCCATTAACAAAT-3'
	Reverse	5'-TTATCCACTTCCAATGCTATTGTCGGTATGGGTCAACCTC-3'
malG T_2 (1841 – 1925)	Forward	5'-TACTTCCAATCCAATGCCACACTTCAACCTCACGAAAGCAC-3'
	Reverse	5'-TTATCCACTTCCAATGCTAAACCCCTTCAATGAGCCTGG-3'
malG R (1932 – 2345)	Forward	5'-TACTTCCAATCCAATGCCTCACAATTCGATCTCTATGCCAAGTA-3'
	Reverse	5'-TTATCCACTTCCAATGCTATCACAGGACGCGTCTAAAAATACG-3'
malE	Forward	5'-TACTTCCAATCCAATGCCATGACAGCAGGTCCGATGG-3'
	Reverse	5'-TTATCCACTTCCAATGCTATCAAGCACCATCTCCTTGACC-3'
malB	Forward	5'-TACTTCCAATCCAATGCCATGCCTTCACAAAGCCCATATCAT-3'
	Reverse	5'-TTATCCACTTCCAATGCTACTAGTAAGCTGACAAGTTGGTTCG-3'
malC	Forward	5'-TACTTCCAATCCAATGCCATGGCACCTACCAGGAGATC-3'
	Reverse	5'-TTATCCACTTCCAATGCTATCAGCGCAAAAGCATCCCC-3'
phqB R (2006 – 2449)	Forward	5'-TACTTCCAATCCAATGCCTGGTGGGAGAGGGTGCAA-3'
	Reverse	5'-TTATCCACTTCCAATGCTATTAAGAGTTGATAAGACCATTCCC-3'
phqB R (2006 – 2429)	Forward	5'-TACTTCCAATCCAATGCCTGGTGGGAGAGGGTGCAA-3'
	Reverse	5'-TTATCCACTTCCAATGCTAGGCAGCCAGGTCCTCAAG-3'
phqE	Forward	5'-GATCCAGCTAGCATGACACCCGCTCCGACACCAC-3'
	Reverse	5'-ATCAGACTCGAGTTAGACAAGAAGCATGCCACCGTTTG-3'
phqE D166N	Forward	5'-CGAGCGGTGCACACTGTTTGAGACCAAACCCTGGCTGGACCGTTATCTCGGGATATTG-3'
	Reverse	5'-CAATATCCCGAGATAACGGTCCAGCCAGGGTTTGGTCTCAAACAGTGTGCACCGCTCG-3'
malG R Y2132F	Forward	5'-CGTATTTGGTCTGACTGAAGCCATCATCGGAGGTT-3'
	Reverse	5'-AACCTCCGATGATGGCTTCAGTCAGACCAAATACG-3'
malC D108A	Forward	5'-GGCTGCACCATGGCTGCGGCGGTGA-3'
	Reverse	5'-TCACCGCCGCAGCCATGGTGCAGCC-3'
malC D108N	Forward	5'-GCTGCACCATGTTTGCGGCGGTGAAGAC-3'
	Reverse	5'-GTCTTCACCGCCGCAAACATGGTGCAGC-3'
malC R130A	Forward	5'-CATCGGTGCTGTAAATGCAATCGTGCCAACGCGC-3'
	Reverse	5'-GCGCGTTGGCACGATTGCATTTACAGCACCGATG-3'
malC R130K	Forward	5'-AGCATCGGTGCTGTAAATTTAATCGTGCCAACGCGCTG-3'
	Reverse	5'-CAGCGCGTTGGCACGATTAAATTTACAGCACCGATGCT-3'
malC R130Q	Forward	5'-CATCGGTGCTGTAAATTGAATCGTGCCAACGCG-3'
	Reverse	5'-CGCGTTGGCACGATTCAATTTACAGCACCGATG-3'
malC H160A	Forward	5'-ATCGGGCTGTTTTGCAGCCGATCCGCTGGTCAAG-3'
	Reverse	5'-CTTGACCAGCGGATCGGCTGCAAAACAGCCCGAT-3'
malC D165A	Forward	5'-GACTCCATCCCGGAGCGGGCTGTTTTGCA-3'
	Reverse	5'-TGCAAAACAGCCCGCTCCGGGATGGAGTC-3'
malC D165N	Forward	5'-GACTCCATCCCGGATTGGGCTGTTTTGCATG-3'
	Reverse	5'-CATGCAAAACAGCCCAATCCGGGATGGAGTC-3'
malC W168F	Forward	5'-CCCGTAACAAGACTGAATCCCGGATCGGGCTG-3'
	Reverse	5'-CAGCCCGATCCGGGATTCAGTCTTGTTACGGG-3'
malC W168L	Forward	5'-CCGTAACAAGACTCAATCCCGGATCGGGCT-3'
	Reverse	5'-AGCCCGATCCGGGATTGAGTCTTGTTACGG-3'
phqE-Cas9	Forward 1	5'-TCATAGCTGTTTCCGCTGA-3'
	Reverse 1	5'-CAGACAGGTCTTGCTGGAGAGACGAGCTTACTCGTTTCGTC CTCACGGACTCATCAGTCTCCACGGTGATGTCTGCTCAAG-3'
	Forward 2	5'-TCGTCTCTCCAGCAAGACCTGTCTGGTTTTAGAGCTAGAAATAGCAAGTTAAA-3'
	Reverse 2	5'-ATTCTGCTGTCTCGGCTG-3'

Supplementary Table 2. Absolute energies for M06-2X/6-31+G(d,p) computed species (Figure S3). Figures are given in atomic units.

Species	E_{el}	ZPE	H	qh-G
Diketopiperazine:				
Reactant	-1128.389725	0.409576	-1127.956095	-1128.028628
TS-*syn*	-1128.358397	0.410449	-1127.925660	-1127.994007
TS-*anti*	-1128.358040	0.410412	-1127.925174	-1127.994130
Adduct-*syn*	-1128.436341	0.415344	-1127.999384	-1128.066515
Adduct-*anti*	-1128.433281	0.415118	-1127.996351	-1128.064000
Monoketopiperazine:				
Reactant	-1054.350306	0.427763	-1053.899225	-1053.970483
TS-*syn*	-1054.316793	0.427655	-1053.867196	-1053.934773
TS-*anti*	-1054.313190	0.428129	-1053.863097	-1053.930872
Adduct-*syn*	-1054.411919	0.434114	-1053.956863	-1054.022661
Adduct-*anti*	-1054.406867	0.433676	-1053.952161	-1054.018121

Cartesian Coordinates.

Diketopiperazine TS-*syn*
Imaginary Freq: -449.75 cm^{-1}

C	-0.728	-1.180	-0.360
C	-0.810	0.894	-1.257
C	0.278	1.718	-0.573
C	1.512	0.871	-0.348
C	1.700	-0.488	-0.352
C	0.701	-1.589	-0.558
H	0.790	-2.030	-1.557
C	3.703	0.541	0.108
C	5.063	0.680	0.399
C	5.811	-0.483	0.520
C	5.221	-1.752	0.356
C	3.871	-1.880	0.065
C	3.092	-0.720	-0.064
H	5.518	1.657	0.529
H	6.870	-0.413	0.748
H	5.836	-2.640	0.461
H	3.425	-2.863	-0.058
N	2.724	1.489	-0.070
H	2.866	2.483	-0.002
H	0.922	-2.381	0.171
C	-0.168	2.358	0.755
H	-1.021	3.027	0.601
H	0.648	2.964	1.161
H	-0.413	1.604	1.505
C	0.624	2.854	-1.567
H	1.389	3.526	-1.163
H	-0.271	3.452	-1.766

H	0.992	2.448	-2.514
C	-3.219	-0.345	-0.066
N	-2.415	-0.353	1.061
N	-1.657	-1.718	-1.184
C	-2.887	-1.294	-1.046
O	-3.831	-1.635	-1.955
C	-2.158	1.226	-1.217
H	-2.503	1.991	-0.526
H	-2.787	1.074	-2.088
C	-4.581	0.188	0.304
H	-5.377	-0.443	-0.098
H	-4.720	1.195	-0.104
C	-4.536	0.203	1.849
H	-4.919	-0.745	2.237
H	-5.127	1.012	2.281
C	-3.044	0.325	2.190
H	-2.708	1.370	2.224
H	-2.746	-0.162	3.120
H	-0.454	0.407	-2.165
H	-3.390	-2.163	-2.636
C	-1.100	-0.735	1.002
O	-0.350	-0.696	1.974

Diketopiperazine TS-*anti*
Imaginary Freq: -470.04 cm^{-1}

C	0.697	-0.988	0.217
C	0.822	0.894	-1.016
C	-0.402	1.712	-0.591
C	-1.619	0.845	-0.364
C	-1.750	-0.495	-0.115
C	-0.677	-1.535	-0.052
H	-0.910	-2.242	0.757
C	-3.833	0.472	-0.080
C	-5.223	0.576	0.035
C	-5.929	-0.588	0.303
C	-5.272	-1.825	0.453
C	-3.893	-1.921	0.335
C	-3.154	-0.759	0.066
H	-5.731	1.529	-0.080
H	-7.009	-0.544	0.400
H	-5.857	-2.715	0.661
H	-3.394	-2.879	0.446
N	-2.881	1.429	-0.337
H	-3.075	2.402	-0.506
H	-0.636	-2.119	-0.977
C	-0.680	2.674	-1.774
H	0.211	3.274	-1.977
H	-1.498	3.367	-1.546
H	-0.943	2.118	-2.678
C	-0.160	2.560	0.673
H	-1.081	3.085	0.943
H	0.613	3.316	0.511
H	0.122	1.926	1.516

C	3.121	-0.024	0.623
N	3.016	-1.130	-0.206
N	0.937	-0.419	1.421
C	2.129	0.086	1.608
O	2.391	0.847	2.700
C	2.108	1.365	-0.775
H	2.892	1.128	-1.487
H	2.271	2.267	-0.193
C	4.565	0.388	0.694
H	5.022	-0.045	1.591
H	4.678	1.473	0.758
C	5.156	-0.215	-0.599
H	6.198	-0.516	-0.476
H	5.120	0.521	-1.405
C	4.250	-1.416	-0.938
H	4.022	-1.518	-2.002
H	4.661	-2.368	-0.590
H	0.660	0.299	-1.915
H	1.566	0.916	3.203
C	1.791	-1.623	-0.560
O	1.647	-2.463	-1.447

Monoketopiperazine TS-*syn*
Imaginary Freq: -488.80 cm^{-1}

C	-0.741	-1.226	-0.342
C	-0.883	0.809	-1.251
C	0.258	1.677	-0.711
C	1.477	0.843	-0.380
C	1.675	-0.515	-0.368
C	0.679	-1.595	-0.670
H	0.715	-1.899	-1.723
C	3.640	0.529	0.206
C	4.978	0.680	0.582
C	5.726	-0.476	0.754
C	5.158	-1.751	0.557
C	3.831	-1.891	0.183
C	3.050	-0.739	0.003
H	5.417	1.661	0.735
H	6.769	-0.397	1.045
H	5.774	-2.633	0.699
H	3.403	-2.878	0.029
N	2.666	1.469	-0.032
H	2.801	2.465	0.027
H	0.945	-2.487	-0.079
C	-0.132	2.532	0.506
H	-0.960	3.204	0.261
H	0.711	3.157	0.818
H	-0.418	1.920	1.363
C	-1.006	-0.913	1.130
H	-0.810	-1.825	1.724
H	-0.305	-0.150	1.487
C	0.633	2.644	-1.863
H	1.437	3.329	-1.569

H	-0.240	3.241	-2.140
H	0.968	2.086	-2.742
C	-3.232	-0.515	0.207
N	-2.390	-0.493	1.301
N	-1.698	-1.768	-1.092
C	-2.943	-1.389	-0.833
O	-3.926	-1.686	-1.732
C	-2.224	1.130	-1.093
H	-2.530	1.878	-0.366
H	-2.940	0.935	-1.885
C	-4.554	0.105	0.580
H	-5.373	-0.606	0.449
H	-4.768	0.966	-0.062
C	-4.346	0.509	2.059
H	-4.778	-0.251	2.715
H	-4.802	1.470	2.304
C	-2.820	0.526	2.239
H	-2.405	1.516	1.985
H	-2.503	0.273	3.255
H	-0.611	0.329	-2.189
H	-3.488	-2.105	-2.486

Monoketopiperazine TS-*anti*
Imaginary Freq: -462.44 cm^{-1}

C	0.757	-0.819	0.828
C	0.817	0.942	-0.650
C	-0.320	1.813	-0.110
C	-1.570	0.964	-0.075
C	-1.722	-0.283	0.471
C	-0.684	-1.037	1.248
H	-0.745	-0.774	2.311
C	-3.721	0.385	-0.435
C	-5.064	0.333	-0.819
C	-5.780	-0.808	-0.487
C	-5.175	-1.872	0.212
C	-3.842	-1.815	0.588
C	-3.090	-0.675	0.259
H	-5.531	1.155	-1.354
H	-6.826	-0.881	-0.767
H	-5.767	-2.747	0.459
H	-3.387	-2.639	1.130
N	-2.777	1.370	-0.617
H	-2.921	2.222	-1.135
H	-0.903	-2.114	1.181
C	-0.501	2.994	-1.085
H	0.419	3.584	-1.112
H	-1.312	3.660	-0.762
H	-0.705	2.646	-2.103
C	1.281	-1.812	-0.200
H	1.251	-2.829	0.232
H	0.629	-1.821	-1.085
C	-0.039	2.384	1.291
H	-0.873	3.021	1.602

H	0.870	2.996	1.283
H	0.091	1.587	2.025
C	3.300	-0.476	0.141
N	2.647	-1.465	-0.569
N	1.581	-0.344	1.760
C	2.847	-0.157	1.414
O	3.677	0.521	2.261
C	2.147	1.322	-0.636
H	2.763	1.135	-1.511
H	2.507	2.118	0.014
C	4.655	-0.258	-0.467
H	5.416	-0.763	0.135
H	4.913	0.804	-0.491
C	4.525	-0.896	-1.877
H	5.261	-1.692	-2.005
H	4.693	-0.165	-2.671
C	3.091	-1.471	-1.951
H	2.429	-0.844	-2.574
H	3.062	-2.487	-2.356
H	0.498	0.347	-1.509
H	3.126	0.835	2.990

Supplementary Table 3. Data collection and refinement statistics

	PhqB R · NADPH	MalC	PhqE D166N · **11** · $NADP^+$	PhqE · **1** · $NADP^+$
Data collection				
Space group	*I* 222	*P* 4_2	*C* 2	*C* 2
Cell dimensions				
a, *b*, *c* (Å)	81.6, 91.6, 124.6	79.4, 79.4, 133.6	209.5, 117.2, 63.7	209.6, 117.2, 64.8
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 107.4, 90	90, 107.9, 90
Resolution (Å)	2.60 (2.69 – 2.60)*	1.60 (1.66 – 1.60)	1.89 (1.96 – 1.89)	2.29 (2.38 – 2.29)
R_{meas}	0.077 (2.57)	0.079 (1.21)	0.077 (2.14)	0.180 (1.07)
$I / \sigma I$	20.5 (1.1)	15.4 (1.4)	12.8 (0.7)	8.7 (1.6)
Completeness (%)	99.8 (99.9)	99.2 (92.1)	97.2 (91.3)	99.3 (97.4)
Redundancy	13.3 (12.7)	6.6 (5.0)	6.9 (6.1)	6.5 (6.9)
Refinement				
Resolution (Å)	45.82 – 2.60	42.97 – 1.60	48.76 – 1.89	46.38 – 2.29
No. Reflections	14746	107863	112785	65748
R_{work} / R_{free}	0.28 / 0.31	0.17 / 0.20	0.28 / 0.31	0.29 / 0.34
No. atoms	2755	8475	11696	11735
Protein	2707	7735	11196	11190
Ligand/ion	48	--	313	438
Water	--	740	187	107
B-factors (Å^2)	127.5	29.2	68.4	75.4
Protein	126.8	29.0	69.0	75.9
Ligand/ion	164.9	--	61.2	68.6
Water	--	31.9	52.1	49.8
R.m.s. deviations				
Bond lengths (Å)	0.010	0.006	0.010	0.010
Bond angles (°)	1.25	0.86	1.50	1.60

*Values in parentheses are for highest-resolution shell.

Supplementary Table 4. Gene cluster annotation of *mal*/*phq* homologous pathways

Aspergillus turcosus (GenBank accession number NIDN01000061)				
ORF	Size (aa)	Putative Function	Relative identity/similarity (%)	Accession No.
1	266	short-chain dehydrogenase	[*Penicillium fellutanum*] (58/75); *phqE*	AGA37272.1
2	773	P-loop containing nucleoside triphosphate hydrolase protein	[*Aspergillus steynii* IBT 23096], multi-drug resistance	PLB53566.1
3	599	L-amino-acid oxidase	[*Madurella mycetomatis*] (55/70)	KXX80598.1
4	1080	NRPS	[*Aspergillus oryzae*] (39/57)	OOO14897.1
5	455	cytochrome P450	[*Penicillium griseofulvum*] (44/60)	KXG49078.1
6	445	FAD monooxygenase	[*Penicillium oxalicum*] (37/59); *phqK*	AOC84388.1
7	330	cytochrome P450	[*Penicillium griseofulvum*] (61/76)	KXG49078.1
8	411	prenyltransferase	[*Malbranchea aurantiaca*] (56/74); *malE*	AGA37265.1
9	308	negative regulator	[*Penicillium fellutanum*] (65/76); *phqG*	AGA37274.1
10	364	prenyltransferase	[*Malbranchea aurantiaca*] (41/59); *malE*	AGA37265.1
11	1048	hypothetical protein CFD26_02683	[*Aspergillus turcosus*] (89/90)	OXN18465.1
12	323	2OG-Fe(II)-oxygenase	[*Penicillium fellutanum*] (41/57); *phqC*	AGA37270.1 AFT91382.1
13	2324	NRPS	[*Malbranchea aurantiaca*] (41/58); *malG*	AGA37267.1
14	502	cytochrome P450	[*Aspergillus ruber* CBS 135680] (48/64)	EYE91288.1
15	420	P450 monooxygenase	[*Penicillium fellutanum*] (38/57); *phqM*	AGA37280.1
16	295	methyltransferase	[*Aspergillus ochraceoroseus* IBT 24754] (33/50)	PLB24695.1
17	2553	Type I Iterative Polyketide synthase (PKS)	[*Pseudogymnoascus* sp. 23342-1-I1] (41/59)	OBT66706.1
18	363	cytochrome P450	[*Aspergillus oryzae*] (42/62)	OOO07737.1
19	327	Phytanoyl-CoA dioxygenase	[*Penicillium expansum*] (36/54)	XP_016600816.1
20	350	putative Proline utilization protein PrnX	[*Aspergillus calidoustus*] (57/73)	CEL10788.1
21	492	transcriptional regulator	[*Quercus suber*]	XP_023878682.1
22	312	Phytanoyl-CoA dioxygenase	[*Penicillium griseofulvum*] (38/59)	KXG48658.1
23	620	oxidoreductase	[*Malbranchea aurantiaca*] (49/65); *malF*	AGA37266.1
24	293	NmrA-like transcriptional regulator	[*Penicillium roqueforti* FM164] (71/82)	CDM28291.1
25	246	short-chain dehydrogenase	[*Penicillium occitanis*] (60/68)	PCG98875.1
26	73	hypothetical protein CFD26_02699	[*Aspergillus turcosus*] (100/100)	OXN18438.1
27	247	NUDIX family hydrolase, putative	[*Aspergillus fischeri* NRRL 181] (80/89)	XP_001261565.1
28	208	endoglucanase-1	[*Aspergillus lentulus*] (83/89)	GAQ05884.1

29	852	glycosyl hydrolase, putative	[*Aspergillus fischeri* NRRL 181] (91/94)	XP_001261562.1
30	406	ankyrin repeat domain-containing protein 50	[*Aspergillus udagawae*] (67/82)	GAO86765.1
Penicillium griseofulvum (GenBank accession number LHQR01000065)				
ORF	Size (aa)	Putative Function	Relative identity/similarity (%)	Accession No.
1	711	glycogen/starch/alpha-glucan phosphorylase	[*Penicillium griseofulvum*] (96/96)	KXG49065.1
2	356	fungal G-protein, alpha subunit	[*Penicillium griseofulvum*]	KXG49066.1
3	368	MAP kinase SakA	[*Penicillium digitatum* PHI26] (98/99)	EKV06178.1
4	809	late secretory pathway protein AVL9	[*Penicillium griseofulvum*] (100/100)	KXG49068.1
5	785	Cullin homology	[*Penicillium griseofulvum*] (100/100)	KXG49069.1
6	2422	NRPS	[*Penicillium fellutanum*] (37/54); *phqB*	AGA37269.1
7	302	NmrA-like family protein	[*Aspergillus niger*]	GAQ40480.1
8	381	O-methyltransferase	[*Coccidioides posadasii* str. Silveira] (33/47)	EFW19547.1
9	452	cytochrome P450	[*Penicillium griseofulvum*] (96/96)	KXG49073.1
10	470	monooxygenase, FAD-binding	[*Penicillium griseofulvum*] (100/100)	KXG49074.1 AOC84388.1 AGC83573.1
11	394	P450 monooxygenase	[*Penicillium fellutanum*] (44/63); *phqM*	AGA37280.1 KXG49075.1
12	452	P450 monooxygenase	[*Penicillium fellutanum*] (38/55); *phqL*	AGA37279.1 KXG49076.1
13	387	cytochrome P450	[*Penicillium griseofulvum*] (100/100)	KXG49078.1
14	336	cytochrome P450	[*Penicillium griseofulvum*] (100/100)	KXG49078.1
15	618	oxidoreductase	[*Penicillium fellutanum*] (68/78); *phqH*	AGA37275.1
16	383	prenyltransferase	[*Penicillium fellutanum*] (44/62)	AGA37277.1 KXG49080.1
17-1	383	short-chain dehydrogenase	[*Malbranchea aurantiaca*] (51/73); *malC*	AGA37263.1
17-2	462	prenyltransferase	[*Penicillium fellutanum*] (82/87); *phqI*	AGA37276.1
18	178	Hp	[*Penicillium griseofulvum*] (100/100)	KXG49083.1
19	369	Hp	[*Penicillium griseofulvum*] (100/100)	KXG49084.1
20	274	Hp	[*Penicillium griseofulvum*] (94/93)	KXG49085.1
21	404	Calcium-binding EF-hand	[*Penicillium griseofulvum*] (100/100)	KXG49086.1
22	214	pectate lyase, catalytic	[*Penicillium griseofulvum*] (89/89)	KXG49087.1
23	884	SNF2-related protein	[*Penicillium griseofulvum*] (98/97)	KXG49088.1

SUPPLEMENTARY REFERENCES

1 Fraley, A. E. *et al.* Function and structure of MalA/MalA', iterative halogenases for late-stage C-H functionalization of indole alkaloids. *J Am Chem Soc* **139**, 12060-12068 (2017).

2 Watts, K. R. *et al.* Utilizing DART mass spectrometry to pinpoint halogenated metabolites from a marine invertebrate-derived fungus. *J Org Chem* **76**, 6201-6208 (2011).

3 Whicher, J. R. *et al.* Cyanobacterial polyketide synthase docking domains: a tool for engineering natural product biosynthesis. *Chem Biol* **20**, 1340-1351 (2013).

4 Skiba, M. A. *et al.* PKS-NRPS Enzymology and Structural Biology: considerations in protein production. *Methods Enzymol* **604**, 45-88 (2018).

5 Ding, Y., Greshock, T. J., Miller, K. A., Sherman, D. H. & Williams, R. M. Premalbrancheamide: synthesis, isotopic labeling, biosynthetic incorporation, and detection in cultures of *Malbranchea aurantiaca*. *Org Lett* **10**, 4863-4866 (2008).

6 Miller, K. A. *et al.* Biomimetic total synthesis of malbrancheamide and malbrancheamide B. *J Org Chem* **73**, 3116-3119 (2008).

7 Sommer, K. & Williams, R. M. Studies towards paraherquamides E & F and related C-labeled putative biosynthetic intermediates: stereocontrolled synthesis of the alpha-alkyl-beta-methylproline ring system. *Tetrahedron* **64**, 7106-7111 (2008).

8 Hu, L. C., Yonamine, Y., Lee, S. H., van der Veer, W. E. & Shea, K. J. Light-triggered charge reversal of organic-silica hybrid nanoparticles. *J Am Chem Soc* **134**, 11072-11075 (2012).

9 Chong, H. S. *et al.* Efficient synthesis of functionalized aziridinium salts. *J Org Chem* **75**, 219-221 (2010).

10 Nagorny, P., Sane, N., Fasching, B., Aussedat, B. & Danishefsky, S. J. Probing the frontiers of glycoprotein synthesis: the fully elaborated beta-subunit of the human follicle-stimulating hormone. *Angew Chem Int Ed Engl* **51**, 975-979 (2012).

11 Aslanidis, C. & de Jong, P. J. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res* **18**, 6069-6074 (1990).

12 Stols, L. *et al.* A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif* **25**, 8-15 (2002).

13 Brown, W. C. *et al.* New ligation-independent cloning vectors compatible with a high-throughput platform for parallel construct expression evaluation using baculovirus-infected insect cells. *Protein Expr Purif* **77**, 34-45 (2011).

14 Quadri, L. E. *et al.* Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **37**, 1585-1595 (1998).

15 Woithe, K. *et al.* Oxidative phenol coupling reactions catalyzed by OxyB: a cytochrome P450 from the vancomycin producing organism. implications for vancomycin biosynthesis. *J Am Chem Soc* **129**, 6887-6895 (2007).

16 Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132 (2010).

17 DiMaio, F. *et al.* Improved molecular replacement by density- and energy-guided protein structure optimization. *Nature* **473**, 540-543 (2011).

18 Terwilliger, T. C. *et al.* phenix.mr_rosetta: molecular replacement and model rebuilding with Phenix and Rosetta. *J Struct Funct Genomics* **13**, 81-90 (2012).

19 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
20 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
21 Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352-367 (2012).
22 Terwilliger, T. C. *et al.* Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65**, 582-601 (2009).
23 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).
24 Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
25 Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191 (2009).
26 The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.
27 Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. explicit solvent particle mesh Ewald. *J Chem Theory Comput* **9**, 3878-3888 (2013).
28 Case, D. A.; Cerutti, D. S.; Cheatham, III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Greene, D.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017, University of California, San Francisco.
29 Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-1174 (2004).
30 Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges - the RESP model. *J Phys Chem* **97**, 10269-10280 (1993).
31 Singh, U. C. & Kollman, P. A. An approach to computing electrostatic charges for molecules. *J Comput Chem* **5**, 129-145 (1984).
32 Besler, B. H., Merz, K. M. & Kollman, P. A. Atomic charges derived from semiempirical methods. *J Comput Chem* **11**, 431-439 (1990).
33 Gaussian 09, Revision D.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.;

Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2013.

34 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926-935 (1983).

35 Maier, J. A. *et al.* ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput* **11**, 3696-3713 (2015).

36 Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J Chem Phys* **98**, 10089 (1993).

37 Shaw, D. E.; Grossman, J. P.; Bank, J. A.; Batson, B.; Butts, J. A.; Chao, J. C.; Deneroff, M. M.; Dror, R. O.; Even, A.; Fenton, C. H.; Forte, A.; Gagliardo, J.; Gill, G.; Greskamp, B.; Ho, C. R.; Ierardi, D. J.; Iserovich, L.; Kuskin, J. S.; Larson, R. H.; Layman, T.; Lee, L.-S.; Lerer, A. K.; Li, C.; Killebrew, D.; Mackenzie, K. M.; Mok, S. Y.-H.; Moraes, M. A.; Mueller, R.; Nociolo, L. J.; Peticolas, J. L.; Quan, T.; Ramot, D.; Salmon, J. K.; Scarpazza, D. P.; Schafer, U. B.; Siddique, N.; Snyder, C. W.; Spengler, J.; Tang, P. T. P.; Theobald, M.; Toma, H.; Towles, B.; Vitale, B.; Wang, S. C.; Young C. Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. *IEEE SC* **2014**, 41–53.

38 Gaussian 16, Revision B.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2016.

39 Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor Chem Acc* **120**, 215-241 (2008).

40 Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J Chem Phys* **132**, 154104 (2010).

41 Mardirossian, N. & Head-Gordon, M. How accurate are the Minnesota density functionals for noncovalent interactions, isomerization energies, thermochemistry, and barrier heights involving molecules composed of main-group elements? *J Chem Theory Comput* **12**, 4303-4325 (2016).

42 Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem Eur J* **18**, 9955-9964 (2012).

43 GoodVibes V 2.0.2, Paton, R. S. DOI: 10.5281/zenodo.595246.

44 Nodvig, C. S., Nielsen, J. B., Kogle, M. E. & Mortensen, U. H. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. *PLoS One* **10**, e0133085 (2015).
45 Domingo, L. R., Zaragoza, R. J. & Williams, R. M. Studies on the biosynthesis of paraherquamide A and VM99955. A theoretical study of intramolecular Diels-Alder cycloaddition. *J Org Chem* **68**, 2895-2902 (2003).
46 Chakraborty, S., Ortiz-Maldonado, M., Entsch, B. & Ballou, D. P. Studies on the mechanism of p-hydroxyphenylacetate 3-hydroxylase from *Pseudomonas aeruginosa*: a system composed of a small flavin reductase and a large flavin-dependent oxygenase. *Biochemistry* **49**, 372-385 (2010).
47 Barajas, J. F. *et al.* Comprehensive structural and biochemical analysis of the terminal myxalamid reductase domain for the engineered production of primary alcohols. *Chem Biol* **22**, 1018-1029 (2015).
48 Bonnett, S. A. *et al.* Structural and stereochemical analysis of a modular polyketide synthase ketoreductase domain required for the generation of a cis-alkene. *Chem Biol* **20**, 772-783 (2013).
49 Chhabra, A. *et al.* Nonprocessive [2 + 2]e- off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. *Proc Natl Acad Sci U S A* **109**, 5681-5686 (2012).
50 Wilson, D. J., Shi, C., Teitelbaum, A. M., Gulick, A. M. & Aldrich, C. C. Characterization of AusA: a dimodular nonribosomal peptide synthetase responsible for the production of aureusimine pyrazinones. *Biochemistry* **52**, 926-937 (2013).
51 Filling, C. *et al.* Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J Biol Chem* **277**, 25677-25684 (2002).
52 Oppermann, U. *et al.* Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chem Biol Interact* **143-144**, 247-253 (2003).
53 Man, H. *et al.* Structures of alcohol dehydrogenases from *Ralstonia* and *Sphingobium* spp. reveal the molecular basis for their recognition of 'Bulky-Bulky' ketones. *Top Catal* **57**, 356-365 (2014).
54 Perinbam, K., Balaram, H., Guru Row, T. N. & Gopal, B. Probing the influence of non-covalent contact networks identified by charge density analysis on the oxidoreductase BacC. *Protein Eng Des Sel* **30**, 265-272 (2017).
55 Sondergaard, C. R., Olsson, M. H., Rostkowski, M. & Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *J Chem Theory Comput* **7**, 2284-2295 (2011).
56 Olsson, M. H., Sondergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J Chem Theory Comput* **7**, 525-537 (2011).
57 Banani, H. *et al.* Genome sequencing and secondary metabolism of the postharvest pathogen *Penicillium griseofulvum*. *BMC Genomics* **17**, 19 (2016).
58 Gerlt, J. A. Genomic enzymology: web tools for leveraging protein family sequence-function space and genome context to discover novel functions. *Biochemistry* **56**, 4293-4308 (2017).