doi.org/10.26434/chemrxiv.7505486.v1

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

Qingyun Dan, Sean A. Newmister, Kimberly R. Klas, Amy E. Fraley, Timothy J. McAfoos, Amber D. Somoza, James D. Sunderhaus, Ying Ye, Vikram V. Shende, Fengan Yu, Jacob N. Sanders, W. Clay Brown, Le Zhao, Robert S. Paton, K. N. Houk, Janet L. Smith, David H. Sherman, Robert M. Williams

Submitted date: 24/12/2018 • Posted date: 26/12/2018



Citation information: Dan, Qingyun; Newmister, Sean A.; Klas, Kimberly R.; Fraley, Amy E.; McAfoos, Timothy J.; Somoza, Amber D.; et al. (2018): Fungal Indole Alkaloid Biogenesis Through Evolution of a Bifunctional Reductase/Diels-Alderase. ChemRxiv. Preprint.

Prenylated indole alkaloids isolated from various fungi possess great structural diversity and pharmaceutical utility. Among them are the calmodulin inhibitory malbrancheamides and paraherquamides, used as anthelmintics in animal health. Herein, we report complete elucidation of the malbrancheamide biosynthetic pathway accomplished through complementary approaches. These include a biomimetic total synthesis to access the natural alkaloid and biosynthetic intermediates in racemic form, and in vitro enzymatic reconstitution that provides access to the natural antipode (+)-malbrancheamide. Reductive cleavage of a L-Pro-L-Trp dipeptide from the MalG nonribosomal peptide synthetase (NRPS) followed by reverse prenylation and a cascade of post-NRPS reactions culminates in an intramolecular [4+2] hetero-Diels-Alder (IMDA) cyclization to furnish the bicyclo[2.2.2]diazaoctane scaffold. Enzymatic assembly of optically pure (+)-premalbrancheamide involves an unexpected zwitterionic intermediate where MalC catalyzes enantioselective cycloaddition as a bifunctional NADPH-dependent reductase/Diels-Alderase. Crystal structures of substrate and product complexes together with site-directed mutagenesis and molecular dynamics simulations demonstrated how MalC and PhqE, its homolog from the paraherquamide pathway, catalyze diastereo- and enantioselective cyclization in the construction of this important class of secondary metabolites.

## File list (2)

# 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,*] and Robert M. Williams[3,7,*]

[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]Departments of Medicinal Chemistry, Microbiology & Immunology, and 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]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.

**Abstract**

Prenylated indole alkaloids isolated from various fungi possess great structural diversity and pharmaceutical utility. Among them are the calmodulin inhibitory malbrancheamides and paraherquamides, used as anthelmintics in animal health. Herein, we report complete elucidation of the malbrancheamide biosynthetic pathway accomplished through complementary approaches. These include a biomimetic total synthesis to access the natural alkaloid and biosynthetic intermediates in racemic form, and *in vitro* enzymatic reconstitution that provides access to the natural antipode (+)-malbrancheamide. Reductive cleavage of a L-Pro-L-Trp dipeptide from the MalG nonribosomal peptide synthetase (NRPS) followed by reverse prenylation and a cascade of post-NRPS reactions culminates in an intramolecular [4+2] hetero-Diels-Alder (IMDA) cyclization to furnish the bicyclo[2.2.2]diazaoctane scaffold. Enzymatic assembly of optically pure (+)-premalbrancheamide involves an unexpected zwitterionic intermediate where MalC catalyzes enantioselective cycloaddition as a bifunctional NADPH-dependent reductase/Diels-Alderase. Crystal structures of substrate and product complexes together with site-directed mutagenesis and molecular dynamics simulations demonstrated how MalC and PhqE, its homolog from the paraherquamide pathway, catalyze diastereo- and enantioselective cyclization in the construction of this important class of secondary metabolites.

Prenylated indole alkaloids comprised of the bicyclo[2.2.2]diazaoctane core have attracted considerable interest due to their wide spectrum of biological activities and offer compelling targets for chemical synthesis and biosynthetic studies.[1-3] Among them, 2-deoxy-paraherquamide A (derquantel) is a commercial therapeutic agent for treating parasitic nematodes in sheep.[4-6] It is now clear that in various genera of fungi, two distinct families containing a bicyclo[2.2.2]diazaoctane system have evolved: (1) the dioxopiperazines such as the anti-cancer stephacidins, insecticidal brevianamides and cytotoxic notoamides, and (2) the monooxopiperazines, including the anthelmintic paraherquamides, asperparalines and calmodulin-inhibitory malbrancheamides (Fig. 1a, **1-5**). In addition, the citrinadins represent another related series of alkaloids that are thought to be derived by deconstruction of monooxopiperazine progenitors containing the [2.2.2] ring system[7,8] (Fig. S1).

The bicyclo[2.2.2]diazaoctane core of these metabolites was first proposed in 1970 to arise from an intramolecular [4+2] Diels-Alder (IMDA) reaction.[9] A long-held hypothesis assumed that both the dioxopiperazine and monooxopiperazine families shared a common biogenesis, with the tryptophan carbonyl of the latter family involved in a net four-electron reduction subsequent to a putative Diels-Alder construction.[3,10] Based on initial genetic studies,[11] and experimental corroboration described in this report, we have discovered that Nature employs divergent biogenetic pathways and biochemical mechanisms to generate the bicyclo[2.2.2]diazaoctane nucleus in these two distinct families of alkaloids[11-15] (Fig. S1). Analysis of the malbrancheamide and paraherquamide biosynthetic gene clusters suggested that the bicyclo[2.2.2]diazaoctane ring system is directly produced in the monooxopiperazine oxidation state[11] (Fig. 1b). We reasoned that it proceeds by the cascade depicted in Figure 1b, **6-12**, following reductive cleavage of the tryptophan thiol ester by the nonribosomal peptide synthetase (NRPS) reductase domain.[11,15] The reduced Pro-Trp dipeptide intermediate is reverse prenylated, and we hypothesized that an intramolecular [4+2] Diels-Alder reaction follows, producing the bicyclo[2.2.2]diazaoctane ring system. However, annotation of the Mal and Phq gene clusters[11] failed to reveal a candidate enzyme for the IMDA reaction. The putative cycloaddition is stereospecific based on the *syn*- or *anti*- configuration of C12a (labeling in premalbrancheamide (**1**), Fig. 1a) and the relative position of the diene and the dienophile. Antipodal bicyclo[2.2.2]diazaoctanes have been isolated from different fungal strains producing the dioxopiperazine indole alkaloid family, while only (+)-malbrancheamide ((+)-**2**) has been isolated from *Malbranchea aurantiaca,*[16] indicating strict diastereo- and enantioselectivity of the biosynthetic IMDA. Thus, the identification and characterization of this presumed catalytic step is fundamental for understanding the formation of these structurally diverse molecules.

Reports of Diels-Alderases remain rare, with few examples over the past decade.[17-23] Among them, four crystal structures have been reported including, 1) the *S*-adenosyl-L-methionine (SAM)-dependent methyltransferase SpnF,[17,24] 2) the β-barrel protein PyrI4[25] and its homolog AbyU,[26] and 3) the flavin-dependent enzyme PyrE3.[27] Tang et al. recently reported functional studies on LepI,[21,28] a SAM-dependent enzyme involved in catalyzing a hetero-Diels-Alder reaction to form the Leporin family of natural products. A common theme in these Diels-

Alderases is their apparent evolution from divergent ancestors, with evident active site reconfiguration. Accordingly, in all reported cases these enzymes have lost ancestral function and the sole remaining activity facilitates a spontaneous [4+2] pericyclic reaction with regio- and stereoselectivity. Cofactors, if present, do not serve their canonical catalytic role for the Diels-Alder cycloaddition, but rather play a structural role in maintaining the active site in a catalytically productive conformation. Similarly, distinct catalytic residues that abolished the enzymatic function were not identified in any previously characterized Diels-Alderase, suggesting that catalysis is achieved primarily through substrate positioning in the protein scaffold. The malbrancheamide and paraherquamide gene clusters lack homologous genes that encode known Diels-Alderases, which indicated the existence of a novel class of biocatalysts. In this article, we reveal the molecular basis for stereocontrolled construction of the monooxopiperazine bicyclic core in the malbrancheamide and paraherquamide biosynthetic pathways. These genetically homologous systems proceed through a bifunctional reductase and Diels-Alderase that evolved from an ancestral short-chain dehydrogenase (SDR) and is also encoded in several other fungal natural product biosynthetic gene clusters.

## Biomimetic Synthesis of Premalbrancheamide, Malbrancheamide and Spiromalbramide

Early considerations regarding biogenesis of the bicyclo[2.2.2]diazaoctane core envisioned a biosynthetic Diels-Alder reaction. In order to chemically validate the sequence of events in malbrancheamide biosynthesis, we prepared the C2 reverse prenylated proposed biosynthetic intermediate, dipeptide aldehyde (**17**),[11] and found that this substance undergoes the cascade of ring closure, dehydration, tautomerization and intramolecular cycloaddition, upon deprotection to give premalbrancheamide (**1**) (Fig. 2). This strategy was applied to two additional natural products, malbrancheamide (**2**) and spiromalbramide (**4**) (Fig. S2), underscoring the utility of the biomimetic paradigm. The key, Fmoc-protected dipeptide aldehyde **17** was prepared through the peptide coupling of N-Fmoc proline (**14**) with the C2 reverse prenylated tryptophan methyl ester **13** through the agency of HATU in acetonitrile in 85% yield. Reduction of the methyl ester with sodium borohydride (82% to **15**) followed by a Parikh-Doering oxidation, furnished the N-Fmoc aldehyde **17** in 72% yield. Removal of the N-Fmoc residue with diethylamine under anaerobic conditions provided the di-enamine **9**, which could be isolated and characterized. Treatment of this substance with TFA in THF at temperatures between 0 °C and 50 °C, resulted in formation of (±)-**1**. The observed modest yield is possibly due to unfavorable tautomerization of **9** to **12**. Under aerobic conditions, **9** spontaneously and rapidly oxidized to aromatic zwitterion **11** (Fig. 1b), which we initially reasoned to be a non-physiological by-product. We later determined that **11** could be chemically reduced by NAD(P)H to **12**, resulting in the spontaneous formation of racemic premalbrancheamide (80 - 90% conversion). This discovery suggested two possible biosynthetic routes – aerobic vs. anaerobic – and the possibility that **11** is an authentic pathway intermediate depending on the intracellular conditions during fungal biosynthesis. Significantly, both routes lead to a single *syn*-diastereomer upon cyclization as the corresponding *anti*-isomer was not detected in even trace amounts from the cycloaddition reactions. This finding agrees

with density functional theory calculations that the azadiene **12** has a calculated relative transition state energy difference of about 2.6 kcal/mol favoring the *syn*-cycloadduct.[29] The *anti*-pathway experiences unfavorable steric interactions between the pyrrolidine ring and the prenyl group (Fig. S3). Our biomimetic synthesis of *syn*-malbrancheamides gave rise to the (+)- and (–)-enantiomers, raising the intriguing question regarding how optically pure (+)-premalbrancheamide is formed by *Malbranchea aurantiaca*. This further indicated the likely presence of an enzyme-directed cyclization *in vivo*, and motivated us to explore the biosynthetic origins of premalbrancheamide by *in vitro* pathway reconstitution.

## *In Vitro* Reconstitution of the Malbrancheamide Biosynthetic Pathway

We aimed to reconstitute the biosynthesis of malbrancheamide as a representative monooxopiperazine alkaloid in a multi-component *in vitro* reaction (Fig. S4). We hypothesized the first step of malbrancheamide biosynthesis involves coupling of L-proline and L-tryptophan by MalG, a dimodular NRPS containing six domains ($A_1$-$T_1$-C-$A_2$-$T_2$-R, Fig. 1b), to produce L-Pro-L-Trp aldehyde **6** through reductive off-loading. Since the full-length NRPS protein could not be produced in soluble form, we identified domain boundaries in MalG, developed expression constructs for the excised $A_1$-$T_1$, C, $T_2$ and R domains (Fig. S5), and loaded the putative amino acid substrates onto the MalG $T_1$ and $T_2$ domains (Fig. S6a, b). Phosphopantetheinylated MalG $A_1$-$T_1$ was loaded with L-proline in the presence of ATP and $Mg^{2+}$, consistent with our functional annotation. With no access to soluble MalG $A_2$, L-tryptophan was loaded onto MalG $T_2$ using Sfp,[30] a nonspecific 4'-phosphopantetheinyl transferase, and L-Trp-coenzyme A (CoA). L-Pro $A_1$-$T_1$ and L-Trp $T_2$ were incubated with the MalG C domain and R domain with the presumed NADPH cofactor. Product formation was determined by LC/MS and comparison with authentic standards. Instead of the proposed dipeptidyl aldehyde-derived product **8**, we identified aromatic zwitterion **10** as the main product (Fig. 1b and 3a). We hypothesized that **10** was produced from spontaneous oxidation of **8**. This was confirmed by chemical synthesis of **8**, which spontaneously and irreversibly converted to **10**. This transformation was suppressed under anaerobic conditions, leading to the conclusion that the malbrancheamide NRPS product rapidly cyclized and dehydrated to **8** and subsequently spontaneously oxidized to **10** under aerobic (i.e. physiological) conditions.

To further test the hypothesis that the MalG terminal R-domain catalyzes an NADPH-dependent two-electron reductive release to produce an aldehyde, we synthesized a dipeptidyl-CoA analog **23**, in which the prolyl-N-atom was replaced with an O-atom to prevent nucleophilic addition of the prolyl-N-atom to the CoA thioester, and loaded **23** onto MalG $T_2$ via Sfp (Fig. S6c). Product standards of aldehyde **24** and alcohol **25** were synthesized chemically. Compound **25** was nonreactive in methanol, while **24** epimerized and reacted to produce the hemiacetal **26** (Fig. S7). Assays with **23**-loaded $T_2$ and MalG R yielded product **26**, confirming that MalG generates an aldehyde product. NADPH was the preferred cofactor in this reaction (Fig. S7d, e). MalG R is an SDR reductase with catalytic Tyr and Lys amino acids, as demonstrated in the 2.6-Å crystal

structure of an NADPH complex of PhqB R, the MalG R homolog of paraherquamide biosynthesis (Figs. S8-10). The essential role of Tyr was confirmed with MalG R/Y2132F, which was incapable of reductive release (Fig. S10d).

We propose that the NRPS product **8** would undergo a reverse prenylation as the next biosynthetic step, thereby installing the dienophile for the IMDA reaction. Two genes, *malE* and *malB* (from the *mal* gene cluster) encode putative prenyltransferases. We incubated MalB or MalE with substrate-loaded MalG domains, NADPH and dimethylallyl pyrophosphate (DMAPP), the prenyl donor. MalE readily catalyzed a C2 reverse prenyl transfer reaction to produce zwitterion **11** (Fig. 3b), whereas MalB displayed modest activity, suggesting that *malB* may be a redundant gene in the pathway (Fig. 3c and S14). Because we could not distinguish in this assay whether **8** or **10** was the MalE substrate, synthetic **8** was produced under anaerobic conditions by UV irradiation of an *O*-nitrobenzyl (ONB) photo-protected dipeptide aldehyde **30** and subjected to the prenyltransferase assay (Fig. S11a). This substrate was rapidly prenylated by MalE in contrast to synthetic **10**, which showed low levels of conversion with the enzyme, indicating that **8** is the native substrate for MalE (Fig. S12). This raised the question regarding how MalE accesses substrate **8** prior to its rapid oxidation to **10**. Thus, we considered the possibility that C2 reverse prenylation occurs with the substrate tethered to the NRPS $T_2$ domain. To address this question, we tested whether MalE or MalB could prenylate L-Trp, L-Trp-loaded MalG $T_2$, or **23**-loaded MalG $T_2$ (Fig. S13). In all cases, no product was detected, confirming that the prenyl transfer reaction occurred on free substrate following the NRPS-catalyzed reaction.

We noticed low levels of premalbrancheamide in the reconstitution assays with MalG and MalE or MalB. Chiral LC/MS analysis revealed a 1:1 racemic mixture of (±)-**1** (Fig. 3f), in agreement with the biomimetic synthesis described above (Fig. 2). Further investigation using synthetic **11** revealed that racemic premalbrancheamide arose through non-enzymatic reduction of **11** by NADPH to azadiene **12**, which undergoes spontaneous cycloaddition in the reaction buffer, thereby explaining the background accumulation of the Diels-Alder products (±)-**1** from *in vitro* assays. From these studies we ascertained that MalG and MalE are the minimal components required for premalbrancheamide biosynthesis, albeit lacking stereocontrol in the IMDA reaction.

Premalbrancheamide isolated from *Malbranchea aurantiaca* is optically pure (+)-**1**, which strongly implicates enzymatic control in the IMDA reaction. Known Diels-Alderases have diverse origins, but the annotated *mal* and *phq* gene clusters did not contain an evident candidate biosynthetic enzyme. Nonetheless, we tested whether MalC, annotated as a short-chain dehydrogenase/reductase (SDR), could function as the presumed Diels-Alderase. When MalC was incubated with substrate-loaded MalG and MalE (NADPH and DMAPP included), neither aromatic zwitterion intermediate, **10** or **11**, was detected; instead the sole product was (+)-**1**, confirming that MalC functions as an intramolecular [4+2] Diels-Alderase (Fig. 3d). To our surprise, when MalC was added to the reaction mixture after significant amounts of **11** had accumulated, the oxidized intermediate was converted to (+)-**1**, indicating that MalC possessed the ability to reduce the zwitterion **11** to the reactive azadiene **12** prior to conducting the

diastereo- and enantio-controlled cycloaddition reaction. This unexpected reactivity of MalC was confirmed using synthetic **11** and NADPH (Fig. S15). To our knowledge, this is a unique example where reduction regenerates the biosynthetic substrate from an oxidized (aromatic) intermediate to provide a productive mode for cycloaddition. The fact that **11** is a MalC substrate indicates that it is an authentic pathway intermediate, and motivated us to address whether an aerobic or anaerobic biosynthetic route is operative *in vivo*. This question was interrogated in two ways; first by performing MalC assays under anaerobic conditions with synthetic **9**, which was generated by photo-deprotection of ONB prenyl dipeptidyl aldehyde **33** (Fig. S11b). Conversion to (+)-**1** was observed only in the presence of MalC and $NADP^+$ (Fig. 3g). However, the efficiency of this reaction was attenuated compared to the MalC-catalyzed conversion of **11** to (+)-**1** under aerobic conditions, indicating that the dienamine tautomer **9** is not optimally recognized by MalC. It is unknown whether MalC can play a role in tautomerization of **9**; notably, background conversion of **9** to racemic premalbrancheamide was not detected under these conditions. Second, gene disruption of the *malC* homolog *phqE* was conducted in the paraherquamide-producing strain *Penicillium simplicissimum* using a CRISPR-Cas9 system.[31] Extracts from the *phqE* mutant strain grown on CYA medium showed the presence of the expected (methyl-Pro-Trp prenyl) zwitterion intermediate **38** (Fig. S16) by LC/MS analysis and co-injection with a synthetic standard, confirming the accumulation of this oxidized metabolite *in vivo* (Fig. S17). Taken together, these data indicate that **11** is the native substrate for MalC en route to (+)-premalbrancheamide.

For the MalC-catalyzed reduction of **11**, NADH or NADPH are effective as the cofactor. However, NADPH is required for strict stereocontrol of the IMDA reaction, as MalC produced a 63:37 mixture of (+)-**1** and (–)-**1** when using NADH (Fig. 3h). This is consistent with the anaerobic experiment in which $NADP^+$ was required to generate (+)-**1**, and further indicates that NADPH plays an important role in the IMDA stereocontrol. The Michaelis-Menten kinetic constants for NADPH and NADH in reactions with MalC and **11** revealed a 10-fold greater catalytic efficiency ($k_{cat}/K_M$) with NADPH compared to NADH (Fig. S15d, e). A 6-fold greater $K_M$ with NADH also suggests that proper cofactor binding is a required component to achieve stereocontrol. Enzymatic rate enhancement of (+)-**1** formation is evident under both assay conditions: aerobically through substrate **11** (post-reduction), or anaerobically through substrate **9** (post-tautomerization) (Fig. 3g). The dramatic shift in enantiomeric excess for the enzymatic reaction (from 0% to 96%) is indicative of enzymatic catalysis for the IMDA reaction.

To complete the biosynthetic pathway, flavin-dependent halogenase MalA was employed to add chlorine atoms on C8 and C9 of premalbrancheamide (+)-**1** to provide malbrancheamide (+)-**2**.[32] We incubated MalA with its pathway partners (L-Pro and L-Trp MalG, MalE and MalC, NADPH, DMAPP, NaCl and $FADH_2$) and identified (+)-**2** as the final product (Fig. 3e). We also found that MalA is stereospecific: when incubated with racemic mixture of **1**, MalA chlorinated only the natural (+) enantiomer (Fig. S18).

## Probing the catalytic mechanism of the bifunctional Diels-Alderase

To gain further insight into the function of the SDR-derived Diels-Alderase, the crystal structure of ligand-free MalC was solved at 1.6-Å resolution, revealing a classical SDR fold with a nucleotide-binding subdomain that contains an invariant "$TGX_3GXG$" motif (P-loop), and a C-terminal substrate binding region that is less conserved and largely hydrophobic.[33,34] The closest structural homologs are a group of SDRs including RasADH (2.3 Å Cα R.M.S.D., 27% overall sequence identity),[35] which uses NAD(P) to catalyze reversible oxidation of secondary alcohols to aldehydes. Unexpectedly, MalC lacks the characteristic Tyr and Lys catalytic amino acids, and also the essential Asn and Ser residues of typical SDRs,[33,34] suggesting that the active site is reconfigured to fit its unique catalytic roles.

Neither cofactor nor substrate was captured in complex with MalC due to crystal lattice constraints. Thus, we turned to the homologous paraherquamide biosynthetic pathway. PhqE is a MalC homolog (54% identity) and catalyzed formation of (+)-premalbrancheamide using the zwitterionic prenylated Trp-Pro substrate **11** *in vitro* (Fig. S19). The 2.4-Å crystal structure of PhqE in complex with cofactor ($NADP^+$/$NAD^+$) and premalbrancheamide (Fig. S20) is highly similar to the MalC structure (1.0 Å Cα R.M.S.D.; Fig. 4b).

Consistent with our kinetic data, PhqE crystals grown with $NADP^+$ showed strong electron density for the cofactor (Fig. S21a) bound in a manner conserved with bacterial SDR homologs (Fig. S21c), while $NAD^+$ showed weak electron density (Fig. S21b). Lys50 accounts for preferential binding of $NADP^+$ through a salt bridge with the cofactor 2'-phosphate. Premalbrancheamide binds in a groove on the surface and is surrounded by hydrophobic residues. The bicyclo[2.2.2]diazaoctane ring system is buried against the nicotinamide and several amino acids including Arg131 (Fig. 4c). The indole lies in a pocket formed by Ala, Leu and Val side chains. The gem-dimethyl contacts Asp166 and Trp169, which are part of a conserved "PDPGW" motif (Fig. S24-25). Given the dual functions (reductase and Diels-Alderase) of MalC/PhqE, the product complex reveals PhqE to be well-adapted in its capacity as a stereo- and enantioselective biocatalyst due to the shape complementarity between the active site pocket and the product (Fig. 4f). Additionally, the short distance (~4.4 Å) between nicotinamide C4 and the deoxy C5 of premalbrancheamide suggests that reduction to the reactive azadiene also occurs in the same location of the active site and indicates that reduction and cycloaddition are highly coordinated.

We sought a non-reactive substrate complex with $NADP^+$ and **11**, and obtained strong electron density for **11** with indication of a flexible orientation in PhqE/D166N (Fig. 4d), whereas the wild-type PhqE yielded ambiguous density for **11**. The indole of **11** binds in the same pocket as premalbrancheamide while the pyrazinone is pushed towards the nicotinamide (Fig. 4e). Deoxy C5 of **11** lies 3.6 Å from the nicotinamide C4 consistent with hydride delivery to this position. In addition to hydride transfer, protonation of the pyrazinone alkoxide is required to form the reactive azadiene **12**. Curiously the corresponding oxygen atom is part of a hydrogen-bonding network involving the $NADP^+$ 2'-hydroxyl and Arg131, suggesting that the cofactor may play a role in proton-transfer during reduction. Superposition of the substrate and product complexes

revealed a high degree of pre-organization of **11** towards the Diels-Alder reaction and also affirmed that the reduction and IMDA reactions are spatially confined.

We used molecular dynamics (MD) simulations to explore the concepts of coordinated hydride delivery, proton transfer and pre-organization in the active site. First, we monitored the distance between the putative hydride acceptor (C5 of **11**) and nicotinamide C4 over a 1.2 μs simulation. The average distance between these carbon atoms was 4.2 Å, consistent with the crystal structures (Fig. S22). We next explored alkoxide protonation. In the crystal structure, the $NADP^+$ ribose 2'-hydroxyl is hydrogen-bonded to the alkoxide oxygen and Arg131. Within the first few ns of the simulation Arg131 displaces the ribose hydroxyl and interacts with the alkoxide for the remainder of the simulation, lending significance to the role of Arg131 in protonation of **11**. To assess facial selectivity in the cycloaddition reaction that forms the (+)- and (–)-premalbrancheamide enantiomers, we monitored the dihedral angle along N15-C5a-C12a-C13 (Fig. S23). Comparison of this dihedral angle in the constrained premalbrancheamide and unconstrained **11** revealed that the untethered diene explores only a single face of the pyrazinone ring corresponding to the natural (+)-enantiomer (Fig. S23). Interestingly, this pre-organization was lost when $NADP^+$ was omitted from the simulation, consistent with our observation that the cofactor is required for enantio-controlled cycloaddition. Together these results further support the conclusion that MalC/PhqE-catalyzed reduction and cycloaddition are coordinated and take place in the same active site pocket where the enzyme-cofactor complex provides stereocontrol by positioning the diene for [4+2] cycloaddition as the reactive azadiene is generated by reduction of **11**.

Based on this information, we probed the reaction mechanism by site-directed mutagenesis. MalC was chosen for this analysis to directly compare results with the reconstitution assay. All of the targeted amino acids are conserved in MalC and PhqE. With the *in vitro* reconstitution assay, MalC variants were assayed in the presence of the MalG NRPS and MalE prenyltransferase (Fig. 5a, b). Reductase activity was assessed by the levels of oxidized intermediate **11**: higher levels indicate less reductase activity. The effect on the IMDA reaction was determined by measuring levels of unnatural (–)-**1** as a percent of all premalbrancheamide: ~50% (–)-**1** formation indicates loss of enzymatic IMDA function. We identified five MalC substitutions that abolished reductase activity (D108A, R130A, D165A, D165N, W168L), and found that loss of function is highly correlated with loss of stereocontrol in the IMDA reaction. A single exception is MalC D165A, which produced mainly (+)-**1**, suggesting that Asp165 is required for reduction but not the IMDA reaction. The activity of MalC variants was also measured in assays with **11** (Fig. 5c). In agreement with the reconstitution assay, Asp108, Arg130 and Asp165 were required for reduction. Based on these data, we propose a mechanism for MalC in which Arg130 serves as a proton donor potentially in conjunction with 2'-OH of NADPH ribose (Fig. 5d). NADPH is the hydride donor, and Asp165 may stabilize the positive charge of the zwitterionic substrate **11** and facilitate formation of the reactive azadiene intermediate **12**. Stereocontrol of the IMDA reaction is primarily driven by shape

complementarity, with contacts between substrate and Trp168 and the cofactor playing a critical role in the IMDA process.

The MalC/PhqE Diels-Alderases clearly evolved from an ancestral SDR (Fig. S24). The SDR catalytic Tyr and Lys were replaced by shorter, non-polar residues (Ile and Cys) providing space to accommodate the substrate. The "PDPGW" motif positions the essential Asp165 3.0 Å closer to the substrate compared to the corresponding amino acid in canonical SDRs. The SDR hydrogen bonding network is partially maintained since the catalytic Arg side chain of MalC/PhqE occupies the position of the SDR catalytic Lys, providing a compelling example of protein evolution in molecular detail.

## Conclusions

Our comprehensive approach to studying the Diels-Alder mediated construction of bicyclo[2.2.2]diazaoctane indole alkaloids represents a culmination of conceptual, experimental and computational studies initiated almost a half-century ago by Birch.[36] The divergent biogenesis to create the monooxo- and dioxopiperazine-type molecules employed by diverse fungi was revealed through characterization of the respective biosynthetic gene clusters, which suggested a differential release mechanism from the functionally related bimodular NPRS systems (Fig. S1). This information was leveraged to design a biomimetic total synthesis of premalbrancheamide, providing a direct validation of the prenylated dipeptide azadiene intermediate and IMDA construction of the target natural product in racemic form. The basis for creating the (+)-antipodal form of premalbrancheamide via a presumed stereoselective Diels-Alderase motivated our search for the corresponding enzyme from Mal and Phq pathways, resulting in identification of a novel Diels-Alderase and a mechanistic understanding of enantio-induction. We have demonstrated that during biosynthetic assembly, the key step to produce the polycyclic core is catalyzed by a bifunctional reductase and intramolecular [4+2] Diels-Alderase, MalC/PhqE, providing exquisite diastereo- and enantiocontrol. Derived from SDR ancestors, the active site of MalC/PhqE evolved to accommodate an aromatic zwitterion substrate, and both the reduction and the IMDA steps are NADP(H)-dependent. In contradistinction to all other known putative Diels-Alderases[2] which are either redox-neutral cyclases or oxidases, we have discovered the first reductase-dependent Diels-Alderase. Our work reveals a distinct class of Diels-Alder enzymes and provides insights into the nature of IMDA catalysis as well as providing a bold evolutionary thesis. The availability of key intermediates provided from biomimetic synthesis enabled us to probe the molecular mechanism of this transformation in unprecedented detail. The MalC/PhqE-catalyzed reaction includes the remarkable step of “rescuing” an aromatic zwitterionic substrate **11** to create the bicyclic product **1** and avoid premature pathway termination. This brilliant evolutionary solution to protect the structural and stereochemical integrity of this architecturally unique family of alkaloids is, to the best of our knowledge, unprecedented and underscores the expanding plasticity and adaptability of secondary metabolite genes and enzymes. The Mal/Phq biosynthetic sequence represents a novel

"toolbox" for chemoenzymatic diversification of indole alkaloids with opportunities for facile access to improved calmodulin inhibitors, anthelmintics and other therapeutics to treat human and animal diseases.

**Acknowledgements** This work was supported by the National Institutes of Health (R01 CA070375 to R.M.W. and D.H.S.), R35 GM118101, the Hans W. Vahlteich Professorship (to D.H.S.), and R01 DK042303 and the Margaret J. Hunter Professorship (to J.L.S.). J.N.S. and K.N.H. acknowledge the support of the National Institute of General Medical Sciences of the National Institutes of Health under Award Numbers F32GM122218 (to J.N.S.) and R01GM124480 (to K.N.H.). Computational resources were provided by the UCLA Institute for Digital Research and Education (IDRE) and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the NSF (OCI-1053575). Anton 2 computer time was provided by the Pittsburgh Supercomputing Center (PSC) through Grant R01GM116961 from the National Institutes of Health. The Anton 2 machine at PSC was generously made available by D.E. Shaw Research. GM/CA@APS is supported by the National Institutes of Health, National Institute of General Medical Sciences (AGM-12006) and National Cancer Institute (ACB-12002). The Advanced Photon Source is a U.S. Department of Energy (DOE) Office of Science User Facility operated by Argonne National Laboratory under Contract No. DE-AC02- 06CH11357. We thank Prof. Stephen Ragsdale for his assistance with anaerobic enzyme assays.

**Author Contributions** Q.D., S.A.N., J.L.S., R.M.W. and D.H.S. contributed to the experimental design. Q.D., S.A.N., A.E.F. and W.C.B. performed molecular cloning, protein expression and purification. Q.D., S.A.N. and A.E.F. performed all enzymatic assays and LC/MS analysis. S.A.N. and Q.D. carried out all crystallographic experiments, structural analysis and structure-based site-directed mutagenesis. K.R.K., J.D.S., A.D.S., T.J.M., L.Z., S.A.N. and V.V.S. synthesized and validated all compounds described in this study. Y.Y. and F.Y. carried out the genetic knockout experiment, F.Y. and Q.D. performed genetic annotation. J.N.S. and S.A.N. performed MD simulations. R.S.P. performed DFT calculations. Q.D., S.A.N., K.N.H., J.L.S., R.M.W. and D.H.S. evaluated the data and prepared the manuscript.

**Author Information** The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.H.S. (davidhs@umich.edu) and R.M.W. (robert.williams@colostate.edu).

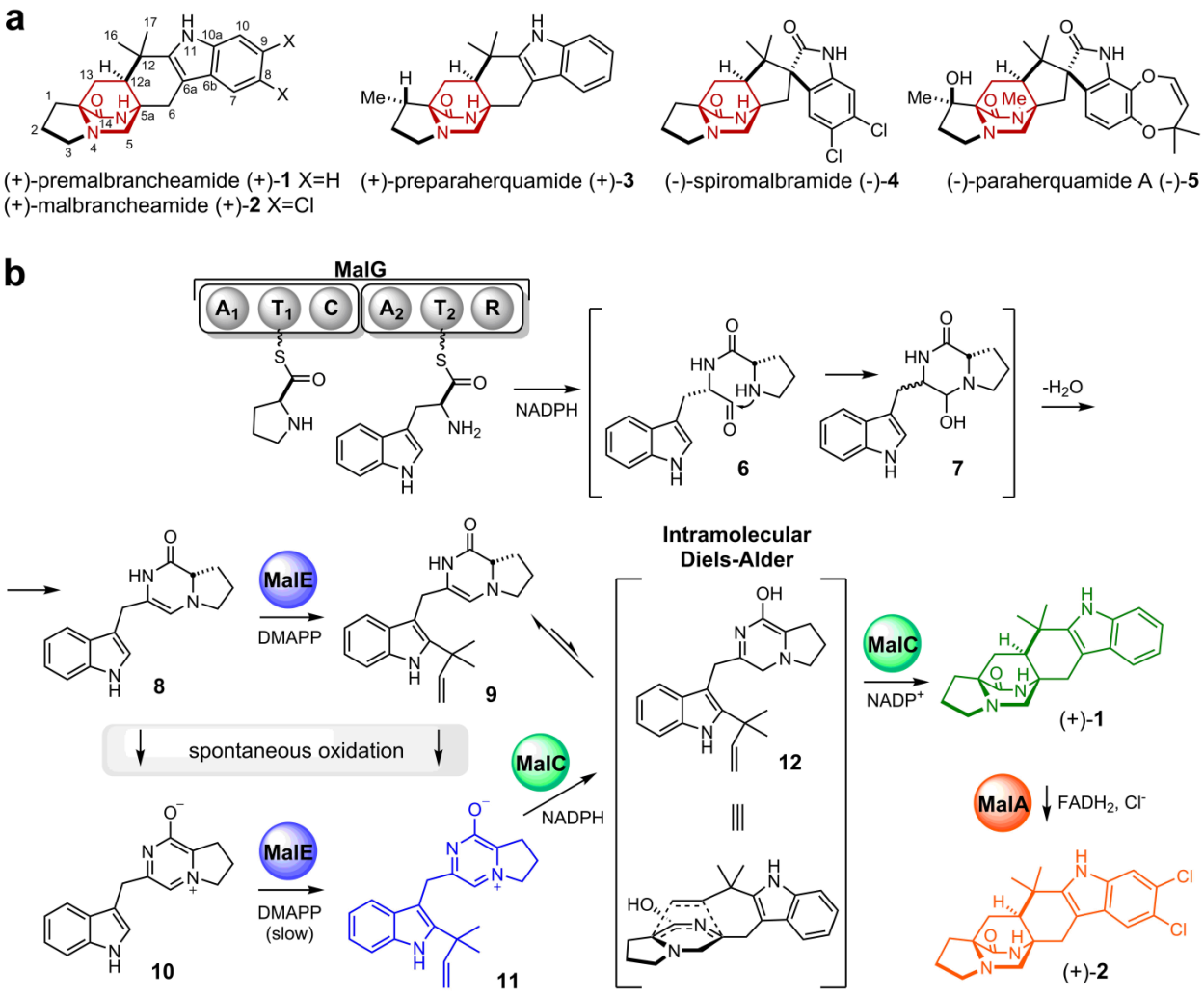



**Figure 1. Fungal bicyclo[2.2.2]diazaoctane indole alkaloids and biosynthesis.**

a. Representative natural products with the biocyclo[2.2.2]diazaoctane group colored in red. b. Scheme of malbrancheamide biosynthesis. The natural substrates are L-proline and L-tryptophan, and the final product is malbrancheamide (+)-**2**. The product of each biosynthetic step is colored differently. Proteins are indicated by spheres; MalG domains are adenylation ($A_1$ and $A_2$), thiolation ($T_1$ and $T_2$), condensation (C) and reductase (R).

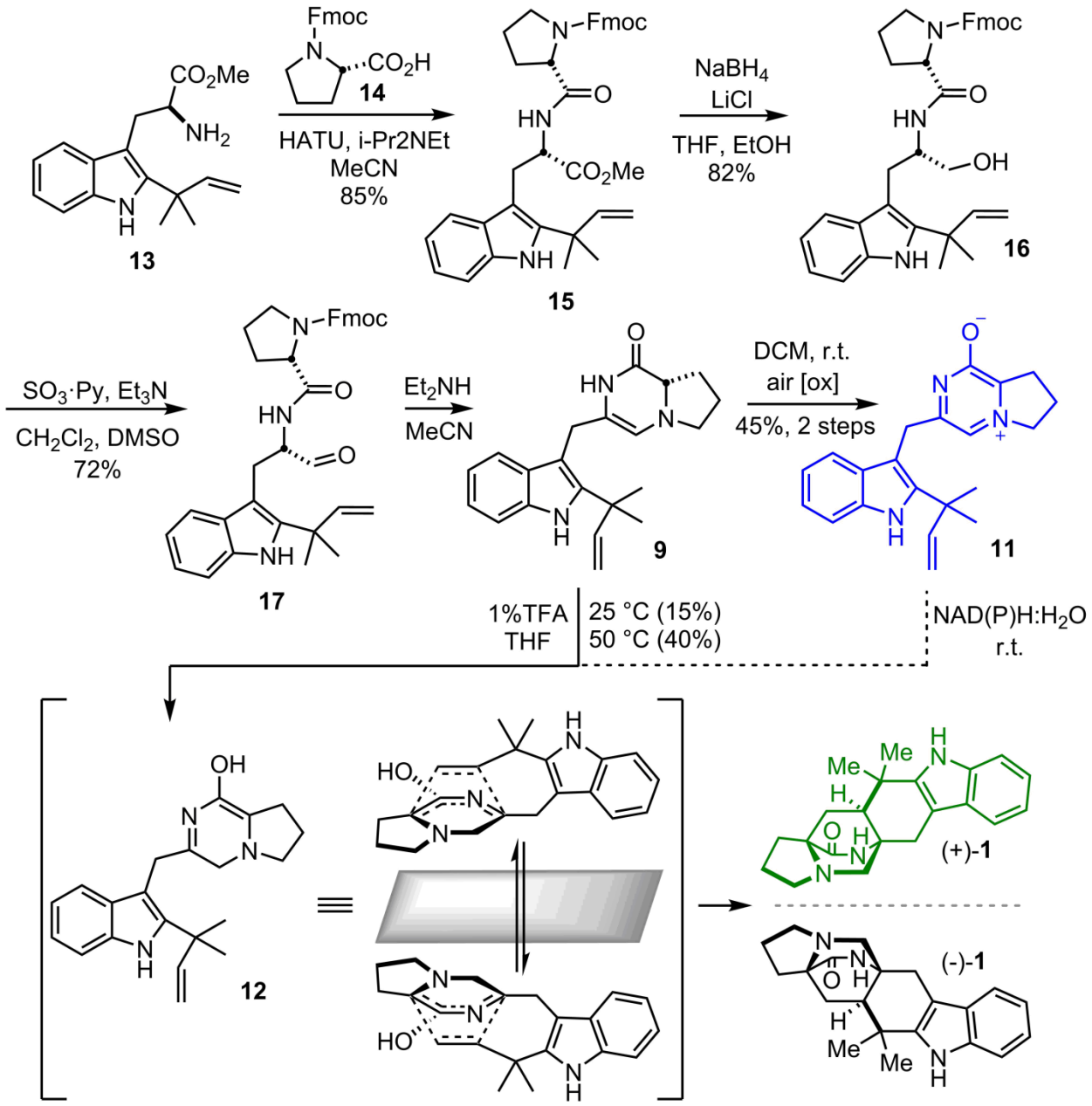



**Figure 2. Biomimetic synthesis of premalbrancheamide.**

The biomimetic synthesis proceeded through a spontaneous intramolecular [4+2] Diels-Alder reaction from a key azadiene intermediate **12** to produce a racemic mixture of *syn*-premalbrancheamides (**1**). Only optically pure (+)-**1** has been isolated from *Malbranchea aurantiaca.* See SI for complete methods.

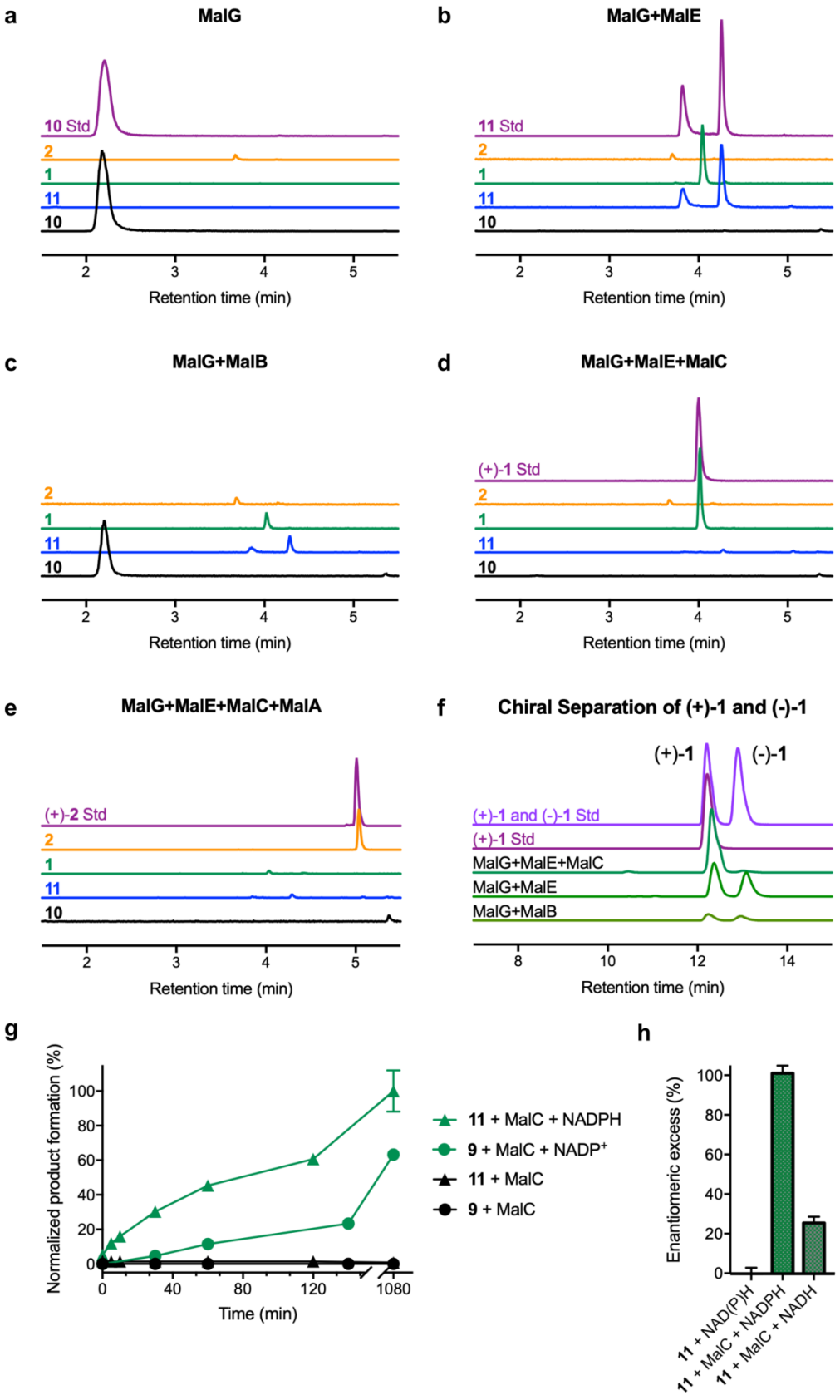

a
MalG
10 Std
2
1
11
10
Retention time (min)
b
MalG+MalE
11 Std
2
1
11
10
Retention time (min)
c
MalG+MalB
2
1
11
10
Retention time (min)
d
MalG+MalE+MalC
(+)-1 Std
2
1
11
10
Retention time (min)
e
MalG+MalE+MalC+MalA
(+)-2 Std
2
1
11
10
Retention time (min)
f
Chiral Separation of (+)-1 and (-)-1
(+)-1
(-)-1
(+)-1 and (-)-1 Std
(+)-1 Std
MalG+MalE+MalC
MalG+MalE
MalG+MalB
Retention time (min)
g
Normalized product formation (%)
Time (min)
11 + MalC + NADPH
9 + MalC + NADP+
11 + MalC
9 + MalC
h
Enantiomeric excess (%)
11 + NAD(P)H
11 + MalC + NADPH
11 + MalC + NADH

**Figure 3. *In vitro* enzymatic reconstitution of malbrancheamide biosynthesis.**
Reactions were monitored by LC/MS. Extracted ion counts (EIC) for key molecules in reaction mixtures are compared to authentic synthetic standards. a. MalG NRPS produced zwitterion **10** by spontaneous oxidation of **8**. b – c. Addition of MalE or MalB prenyltransferase formed three products: a prenylated zwitterion **11**, and (±)-**1**. d. MalC Diels-Alderase addition disabled formation of **11** and (–)-**1** (see panel f). e. Malbrancheamide **2**, the final pathway product, was produced by MalA halogenation of (+)-**1**. f. Chiral separation of (±)-**1** indicates that MalC is an intramolecular [4+2] Diels-Alderase, while MalE or MalB does not provide enantioselectivity for the spontaneous IMDA reaction. g. MalC-catalyzed reactions under aerobic (**11** + MalC) or anaerobic (**9** + MalC) conditions. The aerobic route with **11** as the pathway intermediate was more efficient than the anaerobic route from **9**. h. Effect of cofactor on the enantiomeric excess of the MalC-catalyzed Diels-Alder reaction. MalC provided limited enantioselectivity when NADH was used as cofactor. EIC traces are colored by compound as in Figure 1b, authentic standards are in purple or pink. For panel g and h, all data represent the average of triplicate measurements (error bars, SD; $n = 3$).

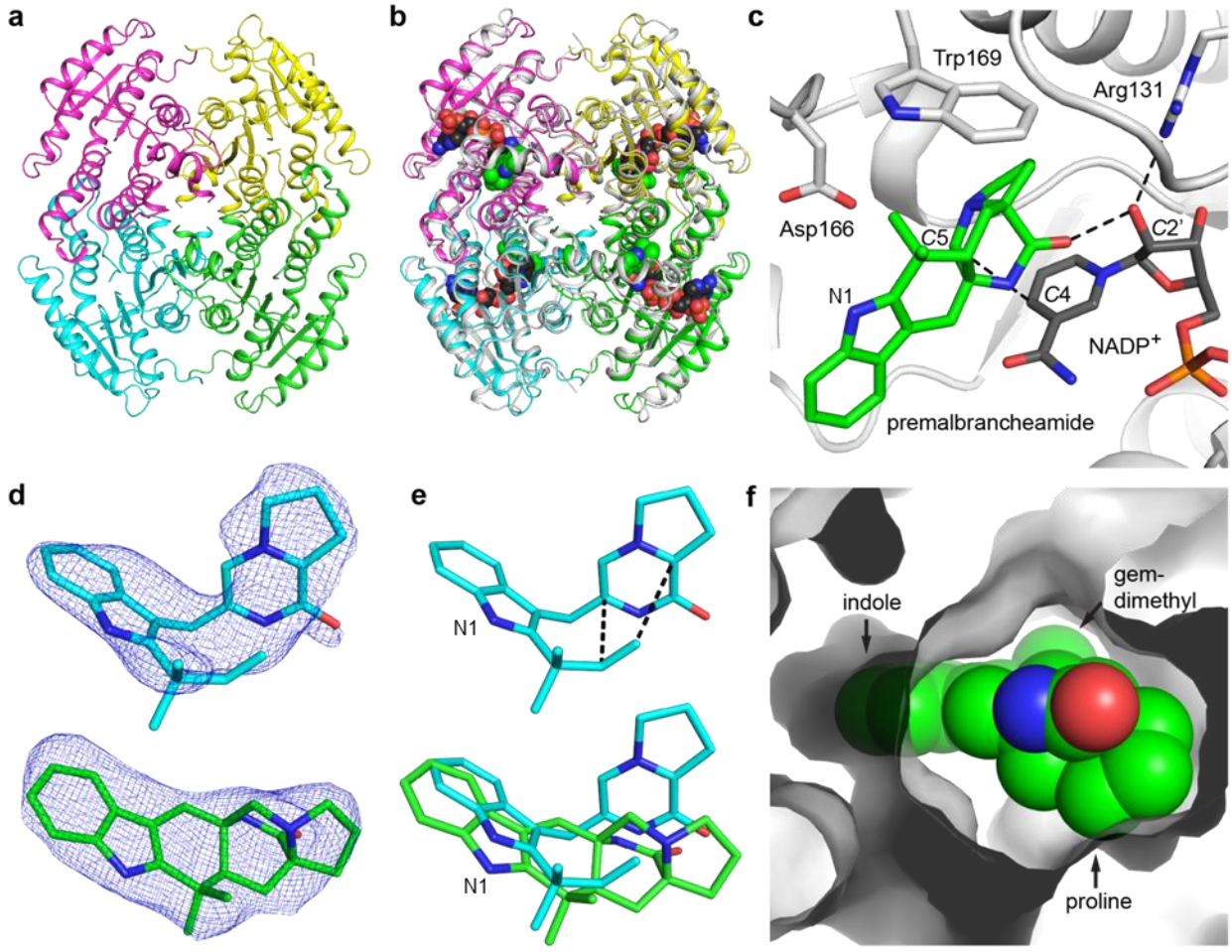



**Figure 4. Structures of MalC and PhqE.**

a. MalC tetramer colored by subunit. b. Superposition of MalC and PhqE product complex (gray); $NADP^+$ (black C) and premalbrancheamide (green C) are shown as spheres. c. PhqE active site showing close arrangement of the product and the $NADP^+$ cofactor. d. Omit electron density ($F_o$-$F_c$; contoured at 2.2 σ) for the substrate **11** (cyan) and the premalbrancheamide (+)-**1** product (green). e. Pre-organization for cycloaddition. Substrate **11** (upper) binds with the prenyl group poised for the IMDA (dashed lines), as seen in the overlay of **11** and premalbrancheamide (+)-**1** (lower). f. Surface representation of the product complex showing high shape complementarity between premalbrancheamide and PhqE.

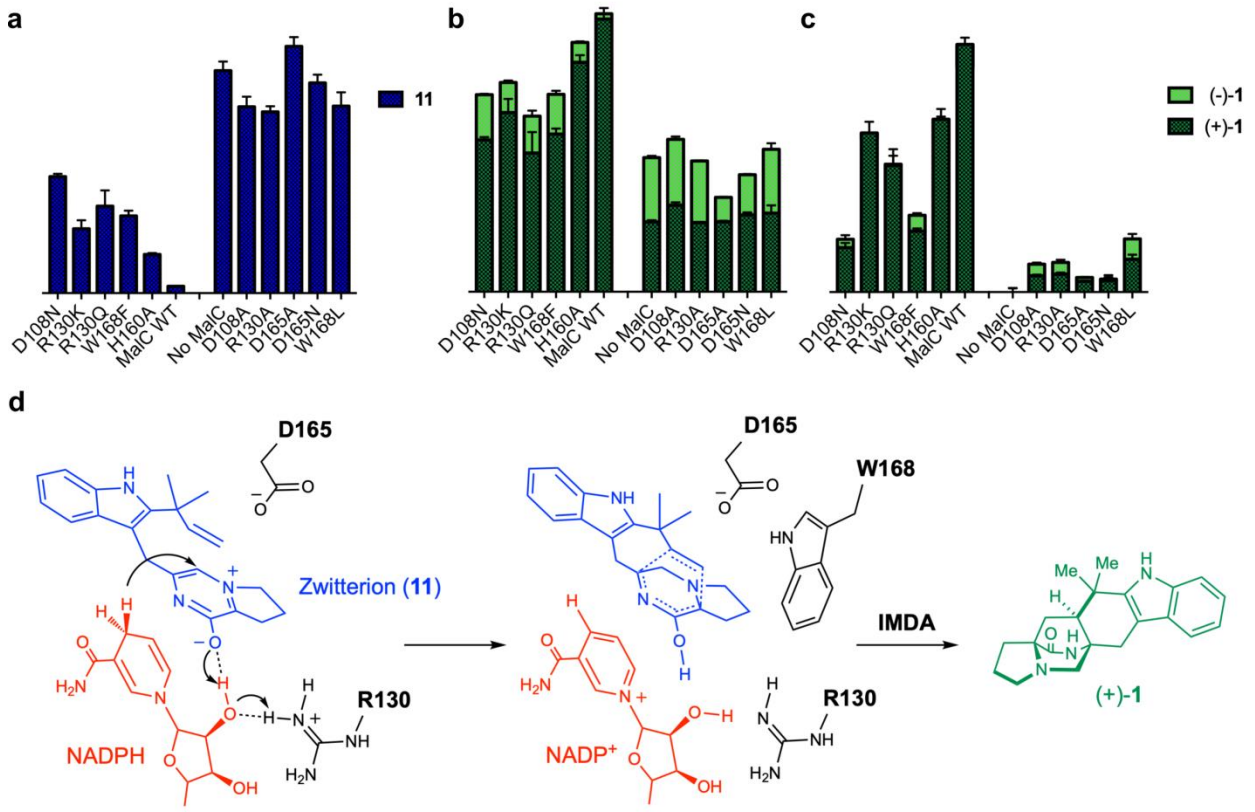



**Figure 5. Catalytic mechanism of the MalC/PhqE-catalyzed Diels-Alder reaction.**
a – b. Product profiles of **11** (blue), (+)-**1** (dark green) and (–)-**1** (light green) in the "MalG+MalE+MalC" reconstitution assay with MalC variants. c. MalC mutagenesis assessed by conversion of synthetic **11** to (+)-**1**. The results agree with those of the reconstitution assay in panel b. Product levels due to non-enzymatic conversion by NADPH were subtracted in all cases. All data represent the average of triplicate measurements (error bars, SD; *n* = 3). d. Proposed catalytic mechanism for MalC/PhqE, with residue numbers for MalC (PhqE residue number = MalC residue number + 1). Arg130 is the indirect proton donor, possibly mediated via the 2'-OH of NADPH ribose. Asp165 stabilizes the positive charge of **11**, and hydride transfer from NADPH completes the first reduction step, forming an unstable azadiene intermediate. The subsequent IMDA reaction is accelerated primarily via entropy trapping, with diastereo- and enantioselectivity achieved via close packing of the $NADP^+$ nicotinamide, the azadiene and MalC Trp168, which together restrain the conformations of both the diene ring and the dienophile to ensure a single cycloaddition mode.

## REFERENCES


1 Finefield, J. M., Frisvad, J. C., Sherman, D. H. & Williams, R. M. Fungal origins of the bicyclo[2.2.2]diazaoctane ring system of prenylated indole alkaloids. *J Nat Prod* **75**, 812-833 (2012).

2 Klas, K., Tsukamoto, S., Sherman, D. H. & Williams, R. M. Natural Diels-Alderases: Elusive and Irresistable. *J Org Chem* **80**, 11672-11685 (2015).

3 Klas, K. R. *et al.* Structural and stereochemical diversity in prenylated indole alkaloids containing the bicyclo[2.2.2]diazaoctane ring system from marine and terrestrial fungi. *Nat Prod Rep* **35**, 532-558 (2018).

4 Robertson, A. P. *et al.* Paraherquamide and 2-deoxy-paraherquamide distinguish cholinergic receptor subtypes in ascaris muscle. *J Pharmacol Exp Ther* **303**, 853-860 (2002).

5 Little, P. R. *et al.* Efficacy of a combined oral formulation of derquantel-abamectin against the adult and larval stages of nematodes in sheep, including anthelmintic-resistant strains. *Vet Parasitol* **181**, 180-193 (2011).

6 Buxton, S. K. *et al.* Investigation of Acetylcholine Receptor Diversity in a Nematode Parasite Leads to Characterization of Tribendimidine-and Derquantel-Sensitive nAChRs. *Plos Pathog* **10**, e1003870 (2014).

7 Mugishima, T. *et al.* Absolute stereochemistry of citrinadins a and B from marine-derived fungus. *J Org Chem* **70**, 9430-9435 (2005).

8 Mercado-Marin, E. V. *et al.* Total synthesis and isolation of citrinalin and cyclopiamine congeners. *Nature* **509**, 318-324 (2014).

9 Porter, A. E. A. & Sammes, P. G. A Diels-Alder Reaction of Possible Biosynthetic Importance. *J Chem Soc Chem Comm*, 1103 (1970).

10 Stocking, E. M. & Williams, R. M. Chemistry and biology of biosynthetic Diels-Alder reactions. *Angew Chem Int Ed Engl* **42**, 3078-3115 (2003).

11 Li, S. *et al.* Comparative analysis of the biosynthetic systems for fungal bicyclo[2.2.2]diazaoctane indole alkaloids: the (+)/(-)-notoamide, paraherquamide and malbrancheamide pathways. *Medchemcomm* **3**, 987-996 (2012).

12 Stocking, E. M., Sanz-Cervera, J. F. & Williams, R. M. Studies on the Biosynthesis of Paraherquamide: Synthesis and Incorporation of a Hexacyclic Indole Derivative as an Advanced Metabolite. *Angew Chem Int Ed Engl* **40**, 1296-1298 (2001).

13 Ding, Y. S. *et al.* Detection of VM55599 and preparaherquamide from *Aspergillus japonicus* and *Penicillium fellutanum*: Biosynthetic implications. *J Nat Prod* **71**, 1574-1578 (2008).

14 Ding, Y. S., 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).

15 Ding, Y. *et al.* Genome-based characterization of two prenylation steps in the assembly of the stephacidin and notoamide anticancer agents in a marine-derived *Aspergillus* sp. *J Am Chem Soc* **132**, 12733-12740 (2010).
16 Martinez-Luis, S. *et al.* Malbrancheamide, a new calmodulin inhibitor from the fungus *Malbranchea aurantiaca*. *Tetrahedron* **62**, 1817-1822 (2006).
17 Kim, H. J., Ruszczycky, M. W., Choi, S. H., Liu, Y. N. & Liu, H. W. Enzyme-catalysed [4+2] cycloaddition is a key step in the biosynthesis of spinosyn A. *Nature* **473**, 109-112 (2011).
18 Hudson, G. A., Zhang, Z. G., Tietz, J. I., Mitchell, D. A. & van der Donk, W. A. *In Vitro* Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin. *J Am Chem Soc* **137**, 16012-16015 (2015).
19 Wever, W. J. *et al.* Chemoenzymatic Synthesis of Thiazolyl Peptide Natural Products Featuring an Enzyme-Catalyzed Formal [4+2] Cycloaddition. *J Am Chem Soc* **137**, 3494-3497 (2015).
20 Tian, Z. H. *et al.* An enzymatic [4+2] cyclization cascade creates the pentacyclic core of pyrroindomycins. *Nat Chem Biol* **11**, 259-265 (2015).
21 Ohashi, M. *et al.* SAM-dependent enzyme-catalysed pericyclic reactions in natural product biosynthesis. *Nature* **549**, 502-506 (2017).
22 Li, L. *et al.* Genome Mining and Assembly-Line Biosynthesis of the UCS1025A Pyrrolizidinone Family of Fungal Alkaloids. *J Am Chem Soc* **140**, 2067-2071 (2018).
23 Kato, N. *et al.* Control of the Stereochemical Course of [4+2] Cycloaddition during trans-Decalin Formation by Fsa2-Family Enzymes. *Angew Chem Int Ed Engl* **57**, 9754-9758 (2018).
24 Fage, C. D. *et al.* The structure of SpnF, a standalone enzyme that catalyzes [4 + 2] cycloaddition. *Nat Chem Biol* **11**, 256-258 (2015).
25 Zheng, Q. *et al.* Enzyme-Dependent [4 + 2] Cycloaddition Depends on Lid-like Interaction of the N-Terminal Sequence with the Catalytic Core in PyrI4. *Cell Chem Biol* **23**, 352-360 (2016).
26 Byrne, M. J. *et al.* The Catalytic Mechanism of a Natural Diels-Alderase Revealed in Molecular Detail. *J Am Chem Soc* **138**, 6095-6098 (2016).
27 Zheng, Q. F. *et al.* Structural Insights into a Flavin-Dependent [4+2] Cyclase that Catalyzes trans-Decalin Formation in Pyrroindomycin Biosynthesis. *Cell Chem Biol* **25**, 718-728 (2018).
28 Cai, Y. *et al*. Structural Basis for Stereoselective Dehydration and Hydrogen-Bonding Catalysis by the SAM-Dependent Pericyclase LepI. *bioRxiv*, doi: 10.1101/491761 (2018).
29 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).

30 Quadri, L. E. N. *et al.* Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **37**, 1585-1595 (1998).
31 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).
32 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).
33 Filling, C. *et al.* Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J Biol Chem* **277**, 25677-25684 (2002).
34 Oppermann, U. *et al.* Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chem Biol Interact* **143-144**, 247-253 (2003).
35 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).
36 Birch, A. J. & Wright, J. J. Studies in relation to biosynthesis. XLII. The structural elucidation and some aspects of the biosynthesis of the brevianamides-A and -E. *Tetrahedron* **26**, 2329-2344 (1970).

## 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,*] and Robert M. Williams[3,7,*]

[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]Departments of Medicinal Chemistry, Microbiology & Immunology, and 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]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

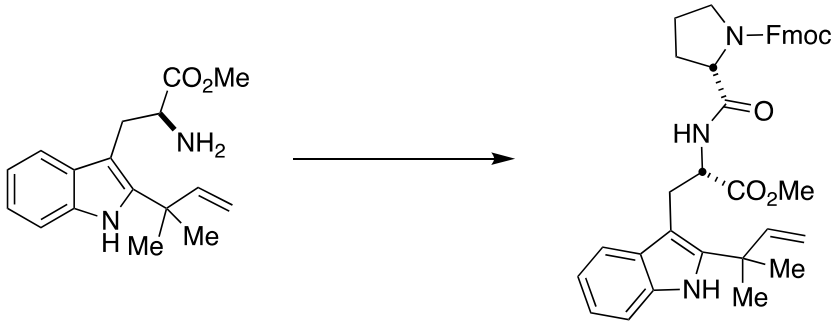
ATP (MilliporeSigma), dimethylallyl pyrophosphate (DMAPP, Isoprenoids) and NAD(P) (Roche) were purchased commercially. Optically pure (+)-premalbrancheamide and (+)-malbrancheamide were extracted from *Malbranchea aurantiaca* RRC1813A as previously described,[1] other chemical reagents used in this study 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,[2] BAP1, pGro7 (Takara), BAP1-pG-KJE8 (Takara) and Insect High Five (BTI-TN-5B1-4, Invitrogen) cells were used for protein expression.[3]

### 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.

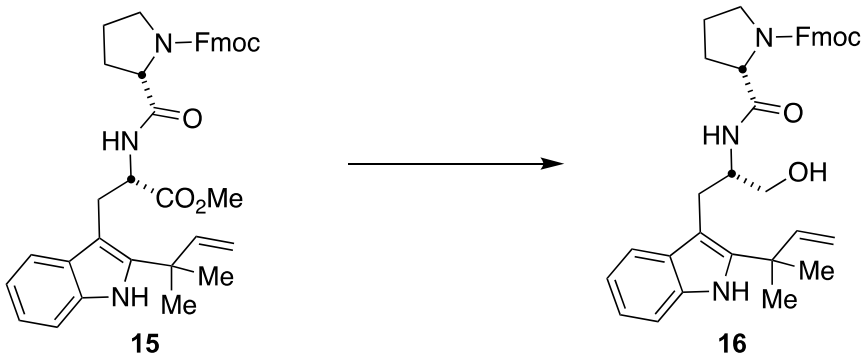
We attempted to coalesce rotomeric peaks by heating to 100 °C. In some cases it was successful, and others it was not. Reports show data of rotomeric compounds taken at 100 °C in DMSO.




**13** **15**

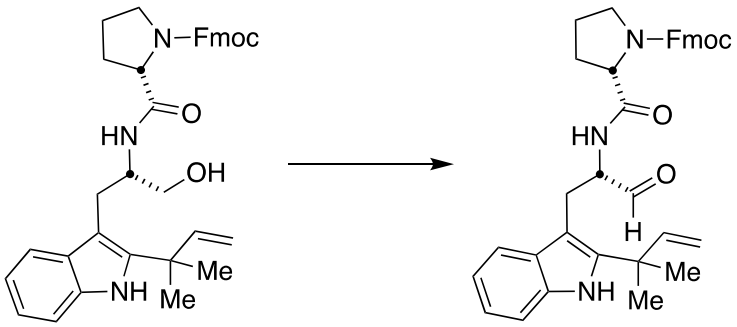
**Dipeptide 15.** HATU (1.230 g, 3.23 mmol) and *i*-$Pr_2NEt$ (1 mL, 742 mg, 5.74 mmol) were added to a solution of *N*-Fmoc-L-proline (1.120 g, 3.32 mmol) and tryptophan **13** (792 mg, 2.76 mmol) in $CH_3CN$ (27 mL) and the reaction was stirred at room temperature for 2 hrs. The reaction was concentrated 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].



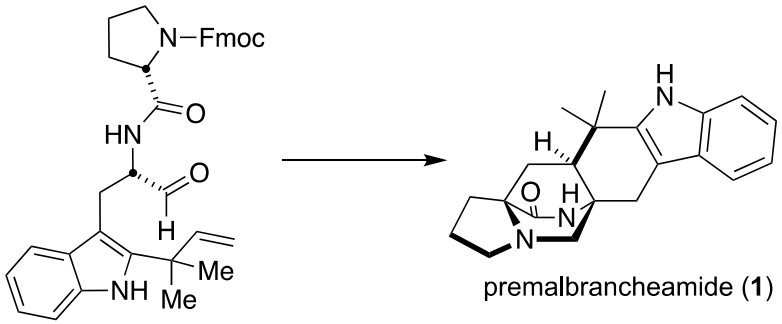

**Alcohol 16.** $NaBH_4$ (280 mg, 7.40 mmol) was added to a suspension of dipeptide **15** (1.428 g, 2.36 mmol) and LiCl (363 mg, 8.56 mmol) in THF (11 mL) and the reaction stirred for 5 min at room temperature. EtOH (11 mL) was then added and the reaction stirred for 7 hrs. The reaction was quenched 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].

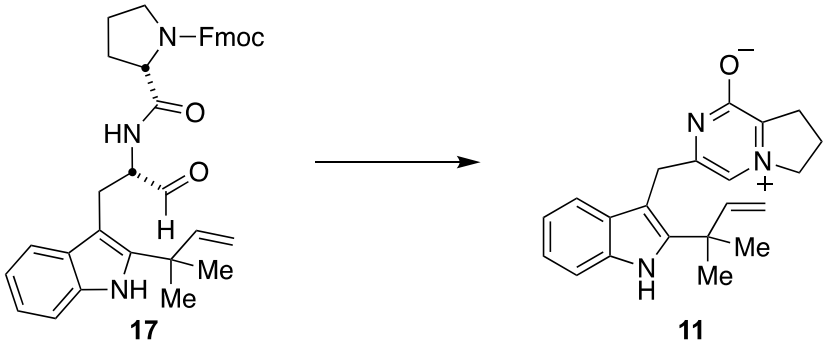



**Aldehyde 17.** $SO_3$·Py (180 mg, 1.13 mmol) was added to a solution of alcohol **16** (160 mg, 0.28 mmol), $Et_3N$ (0.2 mL, 145 mg, 1.43 mmol) and DMSO (1.5 mL) in $CH_2Cl_2$ (3 mL) and the reaction stirred for 2 hrs at room temperature. The reaction was partitioned between water (10 mL) and EtOAc (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].

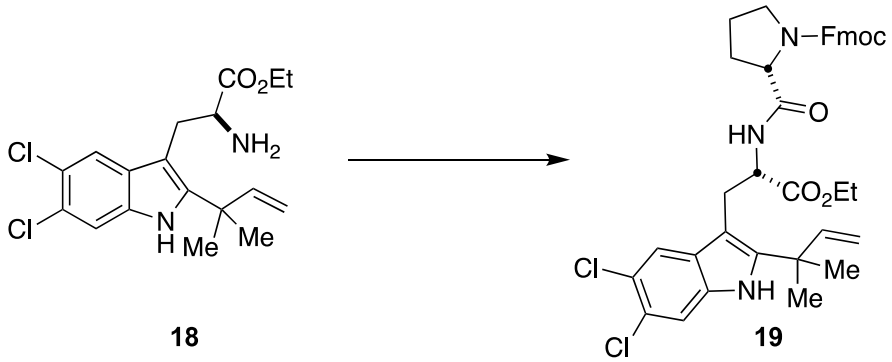



**Premalbrancheamide 1.** $Et_2NH$ (0.5 mL) was added to $CH_3CN$ (2.5 mL) and the resulting solution was sparged with argon for 15 min. The 5:1 $CH_3CN$:$Et_2NH$ solution (2 mL) thus prepared was added to aldehyde **17** (54 mg, 0.094 mmol) and the reaction was stirred for 2 hrs at room temperature 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.[4]



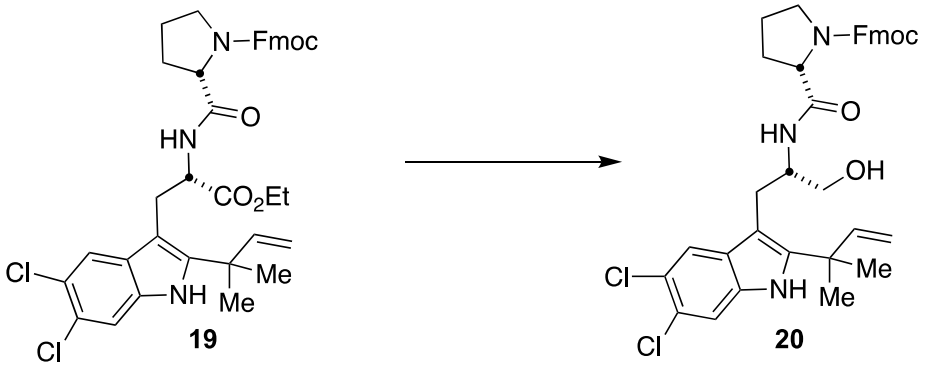

**17** **11**

**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].

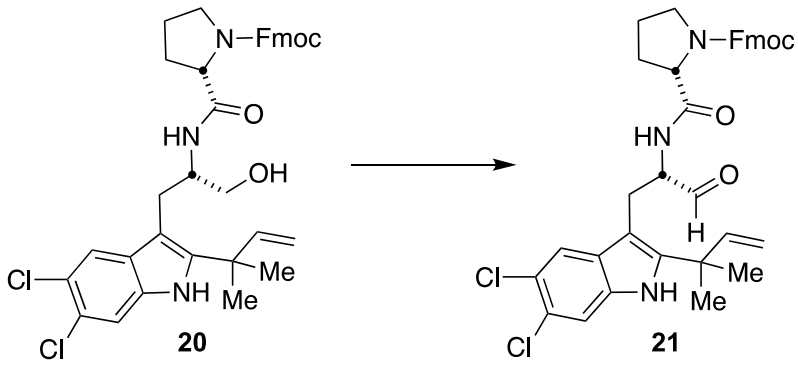



**18** **19**

**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**[5] (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].

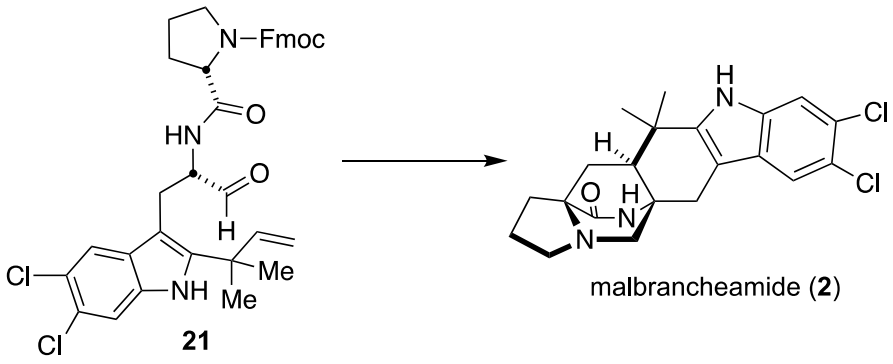


**Alcohol 20.** $NaBH_4$ (5.6 mg, 0.09 mmol) and LiCl (9.0 mg, 0.12 mmol) were added to a solution of dipeptide **19** (26.8 mg, 0.039 mmol) in $CH_2Cl_2$ (0.2 mL) and EtOH (0.2 mL). The reaction stirred for 18 hrs and was quenched with saturated $NH_4Cl$ (3 mL) and extracted with EtOAc (9 mL). The layers 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].



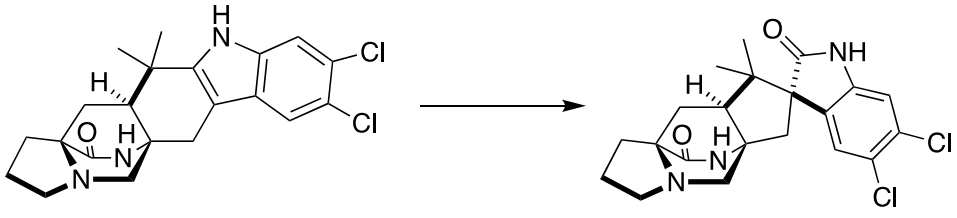

**Aldehyde 21.** $SO_3$·Py (44 mg, 0.28 mmol) was added to a solution of alcohol **20** (45 mg, 0.07 mmol), $Et_3N$ (0.05 mL, 38 mg, 0.38 mmol) and DMSO (0.37 mL) in $CH_2Cl_2$ (0.75 mL) and the reaction stirred for 3 hrs at room temperature. The reaction was partitioned between water (5 mL)

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].



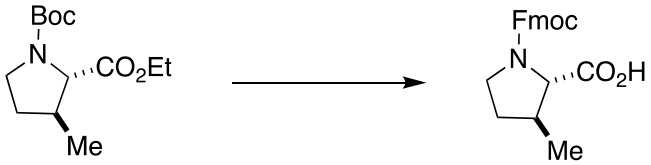

**Malbrancheamide 2.** $Et_2NH$ (0.1 mL) was added to $CH_3CN$ (0.4 mL) and the resulting solution was degassed. The 5:1 $CH_3CN$:$Et_2NH$ solution (0.5 mL) was added to aldehyde **21** (13 mg, 0.020 mmol) and the reaction was stirred for 2 hrs at room temperature and the reaction was concentrated under 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.[5]




malbrancheamide (**2**) spiromalbramide (**4**)

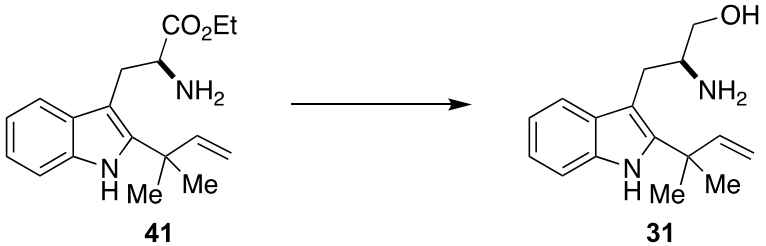
**Spiromalbramide 4.** NCS (5.0 mg, 0.037 mmol) was added to a solution of **2** (10.0 mg, 0.025 mmol) in DMF (0.5 mL) and the reaction stirred for 3 hrs at -15°C to 0 °C. To the solution was added pTsOH (60 mg, 0.35 mmol) and $H_2O$ (0.2 mL) which stirred at 70 °C for 20 min and cooled to room temperature. 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.[6]




**40** **34**

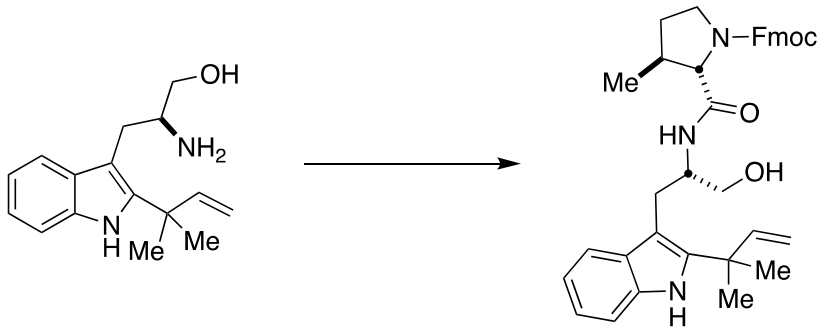
***N*-Fmoc-*β*-methyl-L-proline 34.** NaOH (4.05 g, 15.75 mmol) was added to a solution of **40**[7] in MeOH (315 mL) and the reaction was heated to reflux for 18 hrs. The resulting solution was acidified (pH = 2) with 0.1 M HCl (900 mL) and washed with EtOAc (3 × 500 mL). The

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**

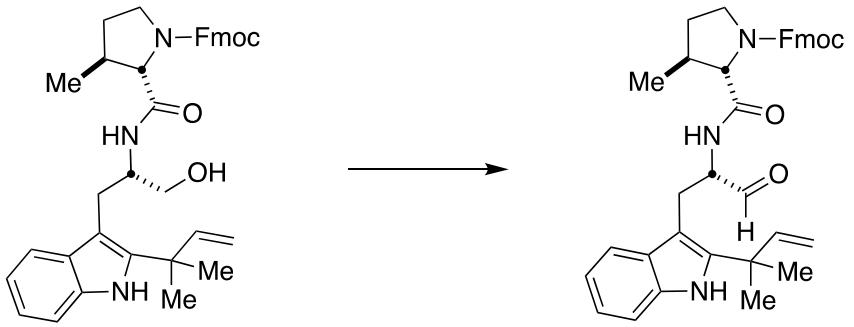
**Reduced Tryptophan 31.** $NaBH_4$ (84 mg, 2.22 mmol) was added to a solution of tryptophan **41** in MeOH (1.16 mL) and the reaction was stirred at room temperature for 1 hr. The reaction was quenched with saturated $NH_4Cl$ (6.5 mL) and washed with EtOAc (13 mL). The layers were separated, and the 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**

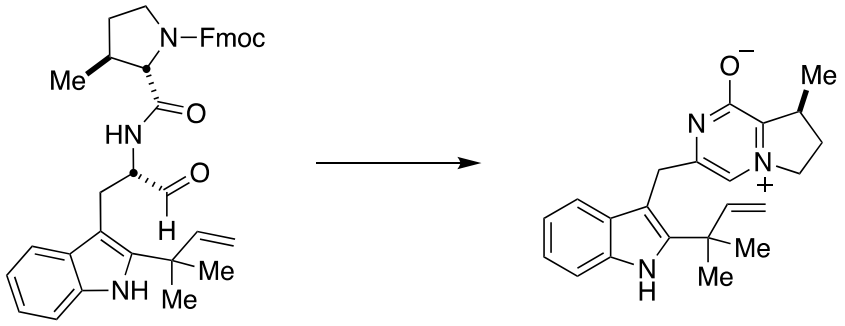
**Alcohol 35.** HATU (1.96 g, 5.16 mmol) and *i*-$Pr_2$NEt (3 mL, 2.22 g, 17.2 mmol) were added to a solution of *N*-Fmoc-*β*-methyl-L-proline **34** (1.59 g, 4.73 mmol) and reduced tryptophan **31** (1.11 g, 4.3 mmol) in $CH_3CN$ (43 mL) and the reaction was stirred at room temperature for 4 hrs. The 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.** $SO_3$·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

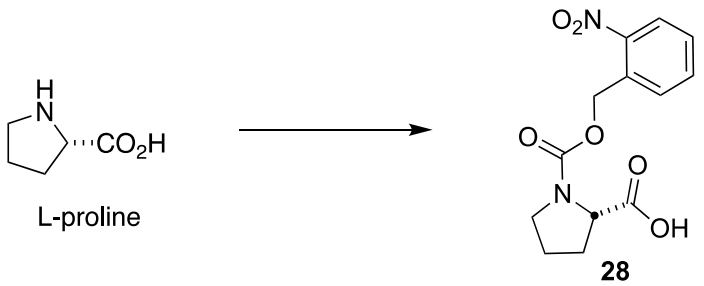
eluting with 45% EtOAc/hexane to give 115 mg (72%) of aldehyde **36** as dark yellow solid. $^{1}$H NMR (500 MHz, DMSO, 100 °C) δ 10.21 (s, 1 H), 9.37 (s, 1 H), 7.85 (d, *J* = 7.55 Hz, 2 H), 7.61-7.64 (comp, 2 H), 7.38-7.48 (comp, 3 H), 7.29-7.33 (comp, 3 H), 6.97-7.01 (comp, 1 H), 6.88-6.93 (comp, 1 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].




**36** **38**

**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,

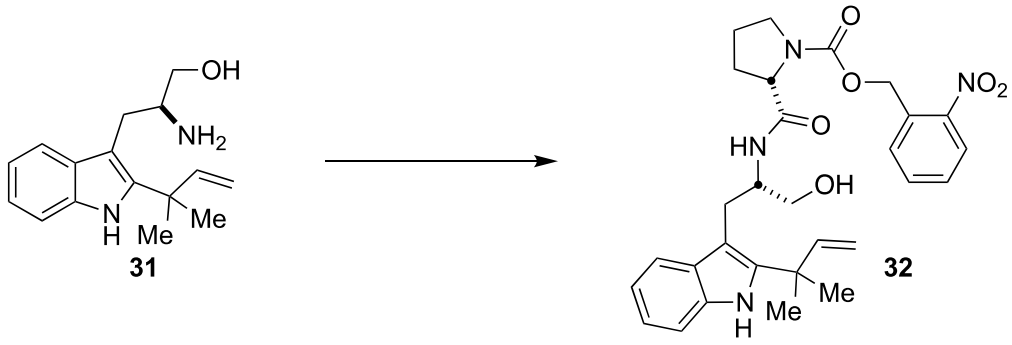
1 H), 7.30 (d, *J* = 7.92 Hz, 1 H), 7.01 (td, *J* = 1.10, 7.56 Hz, 1 H), 6.89 (td, *J* = 0.96, 7.92 Hz, 1 H), 6.65 (s, 1 H), 6.18 (dd, *J* = 10.52, 17.40 Hz, 1 H), 5.06 (dd, *J* = 1.16, 17.44 Hz, 1 H), 5.02 (dd, *J* = 1.20, 10.52 Hz, 1 H), 4.45 (m, 1 H), 4.30 (m, 1 H), 4.00 (s, 2 H), 3.37 (m, 1 H), 2.29-2.39 (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* 364.20162 [$C_{22}H_{25}N_3O$ (M+NH4) requires 364.22576].



**28**

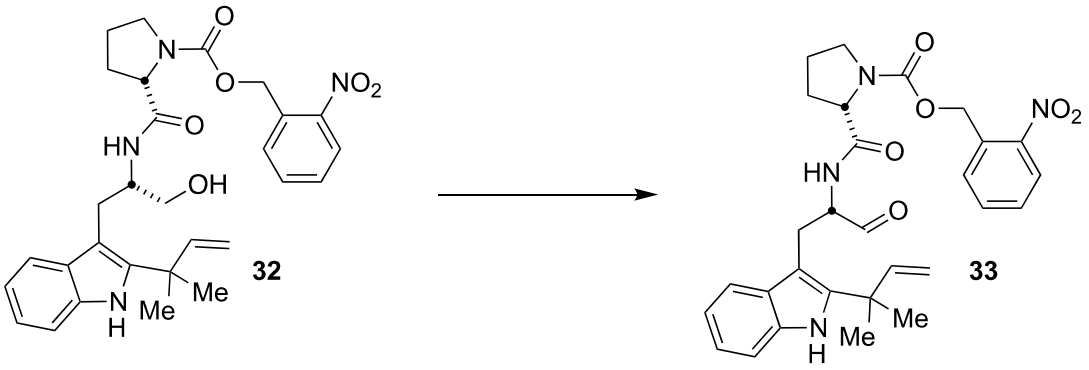
**ONB-L-proline 28.** A solution of previously prepared chloroformate[8] in DCM (8.6 mL) and dioxane (2.15 mL) at 0 °C was added in turn, alongside 2.37 mL NaOH (2 M), to a solution of L-proline in 2 M NaOH (2.15 mL) at 0 °C. The solution ran at the same temperature for 1 hr, then was allowed to warm 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.



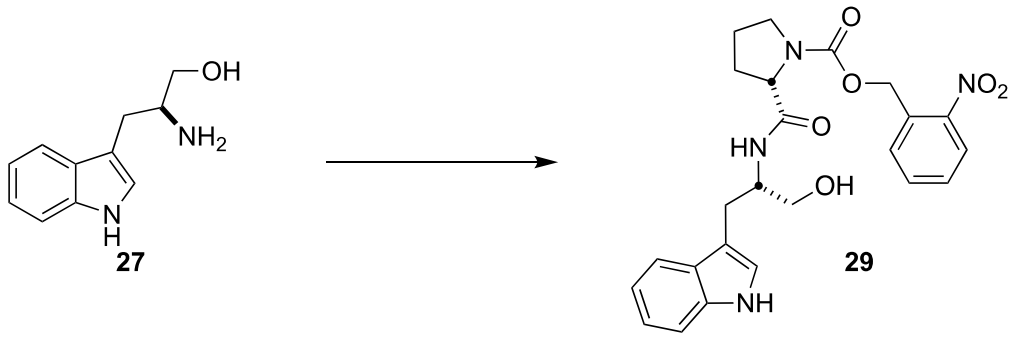

**Alcohol 32.** HATU (182.15 mg, 0.65 mmol) and *i*-$Pr_2NEt$ (0.45 mL, 336.02 mg, 2.6 mmol) were added to a solution of **28** (229.4 mg, 0.78 mmol) and **31** (168 mg, 0.65 mmol) in $CH_3CN$ (6.5 mL) and the reaction was stirred at room temperature for 2 hrs. The reaction was concentrated under reduced 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].

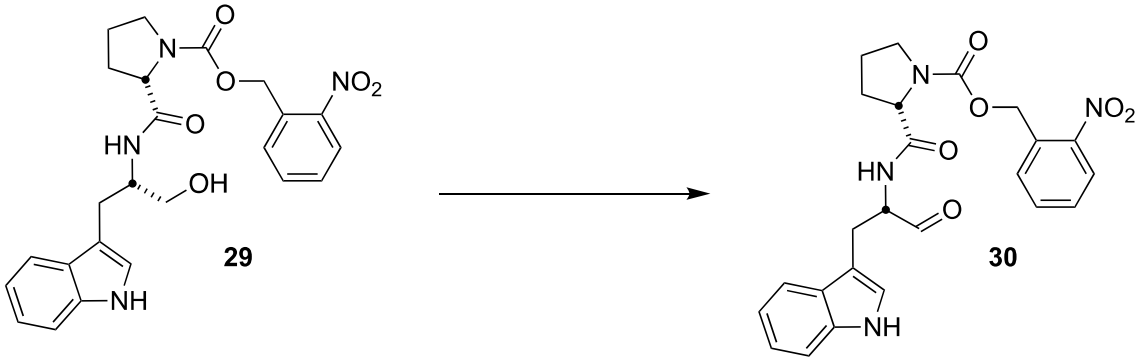



**Aldehyde 33.** $SO_3$·Py (204.25 mg, 1.28 mmol) was added to a solution of alcohol **32** (171.4 mg, 0.321 mmol), $Et_3N$ (0.23 mL, 165.7 mg, 1.64 mmol) and DMSO (1.72 mL) in $CH_2Cl_2$ (3.6 mL) and the reaction stirred for 3 hrs at room temperature. The reaction was partitioned between water (20 mL) 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* = ,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

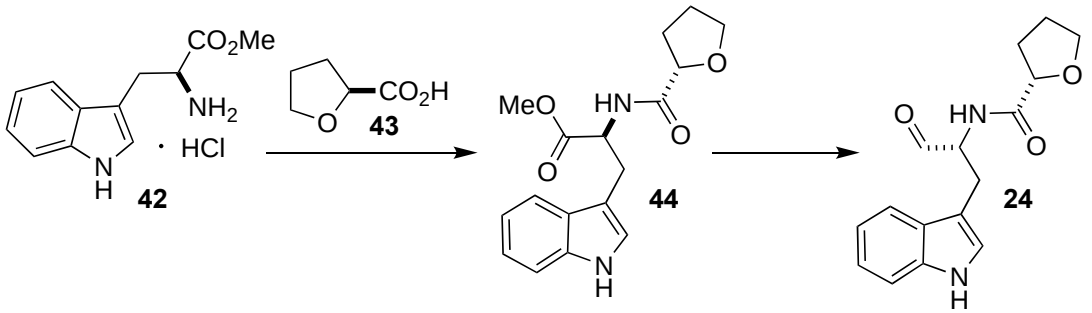
199.60, 171.62, 153.15, 145.81, 145.79, 140.36, 134.43, 133.15, 131.77, 128.83, 128.42, 128.27, 123.82, 119.94, 117.79, 117.28, 110.70, 110.36, 104.39, 62.19, 59.43, 58.98, 46.37, 38.35, 27.38, 27.26, 24.01, 22.70; Maxis Q-TOF (ESI) *m/z* 533.2398 [$C_{29}H_{32}N_4O_6$ (M+H) requires 533.2400].




**Alcohol 29.** HATU (144 mg, 0.513 mmol) and *i*-$Pr_2$NEt (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. The material was not analyzed by NMR due to the presence of rotomers. HRMS (BTOF) *m/z* 467.19234 [$C_{24}H_{26}N_4O_6$ (M+H) requires 467.19251].



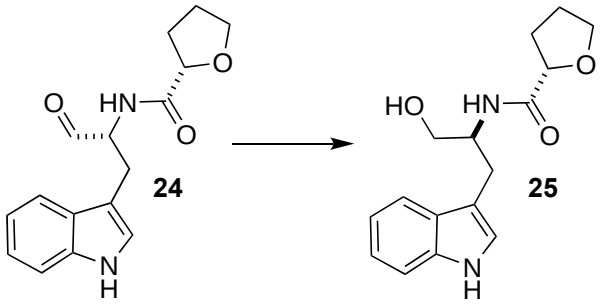
**Aldehyde 30.** $SO_3$·Py (8.33 mg, 0.052 mmol) was added to a solution of alcohol **29** (6.1 mg, 0.0131 mmol), $Et_3N$ (0.01 mL, 6.75 mg, 0.07 mmol) and DMSO (0.07 mL) in $CH_2Cl_2$ (0.15 mL) and the reaction stirred for 3 hrs at 0 °C. The reaction was partitioned between water (2 mL) and EtOAc (2 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. The material was not analyzed by NMR due to the presence of rotomers. 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*-$Pr_2$NEt (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.

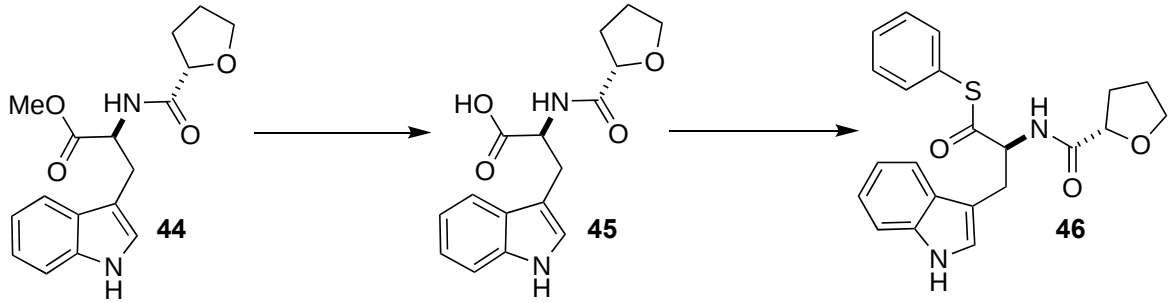
DIBAL-H (7.0 mL, 1M in toluene) was added to a solution of methyl ester **44** (800 mg, 2.50 mmol) in $CH_2Cl_2$ (25 mL) at -78 °C and the reaction stirred for 2 hrs at room temperature. The

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. $^1$H 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·$H_2O$ (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 N HCl (22mL). 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** (CAS# 1357586-64-1) 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·$H_2O$ (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].

### 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.[10,11] For *malG $A_1$-$T_1$*, *malG C*, *malG $T_2$*, *malG R* and *phqB R*, the coding sequence was inserted into the pMCSG9 vector. For *phqE*, the pET28b vector was used. For *malB*, pFastBac transfer vector (Invitrogen) was used. 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. pFastBac-MalB was transformed into *E. coli* DH10Bac cells for production of recombinant bacmids. To prepare bacmids, 80 μL SOC medium was added to 20 μL DH10 Bac cells (pFastBac-MalB) and incubated at 37 °C for 3 hrs. The sample was then plated on a Q-tray (48 wells) containing kanamycin, tetracycline, gentamycin, IPTG and bluo-gal. The tray was incubated for 48 hrs at 37 °C. Two white colonies

from each well were picked and patch streaked onto an indicator plate which was incubated overnight at 37 °C. Colonies that remained white were considered positive and were inoculated with 5 mL Lysogeny broth medium and cultured overnight at 300 rpm shake, 37 °C. Cells were pelleted and then subjected to alkaline lysis. Samples were spun in a microfuge for 10 min at 13,000 rpm. An 800 μL aliquot was then added to 800 μL of isopropanol, mixed by inversion and spun for 30 min at 13,000 rpm. The supernatant was decanted and the tubes were air-dried. The bacmid pellets were resuspended in 40 μL of sterile water and stored at -20 °C.

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[2] 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-$T_2$.[3] 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. 1L of Insect X-press media (Lonza) in 2.8 L Fernbach flasks was seeded with High Five cells at $2 \times 10^6$ 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, 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% 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 2 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, 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).

**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. 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 (Aeris widepore C4 column (3.6 μm, 50 × 2.10 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).

**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 Kinetix reverse-phase C18 column (40 mm × 2.1 mm, 2.6 μm), 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.). Chiral separations were performed using Phenomenex Lux cellulose-3 (250 × 4.6 mm, 5 μm) column (buffer A: water, buffer B: 95% acetonitrile; flow rate 0.5 mL/min; 19% acetonitrile for 3 min, 19 - 95% acetonitrile gradient over 10 min, 95% acetonitrile for 2 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). 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.[12] Attempts to solve the structure molecular replacement with the similar bacterial NRPS R domain structures succeeded using the MR-ROSETTA[13,14] process in PHENIX.[15] Model building was carried out with Coot.[16] Refinement was carried out with PHENIX.refine.[17] 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.[18] 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 well-ordered and the half tetramer poorly packed (Fig. S20). All structures were validated with MolProbity.[19] Multiple sequence alignments were generated from Clustal and Jalview.[20,21] Figures were prepared with PyMOL.[22]

**10. Molecular Dynamics Simulations**

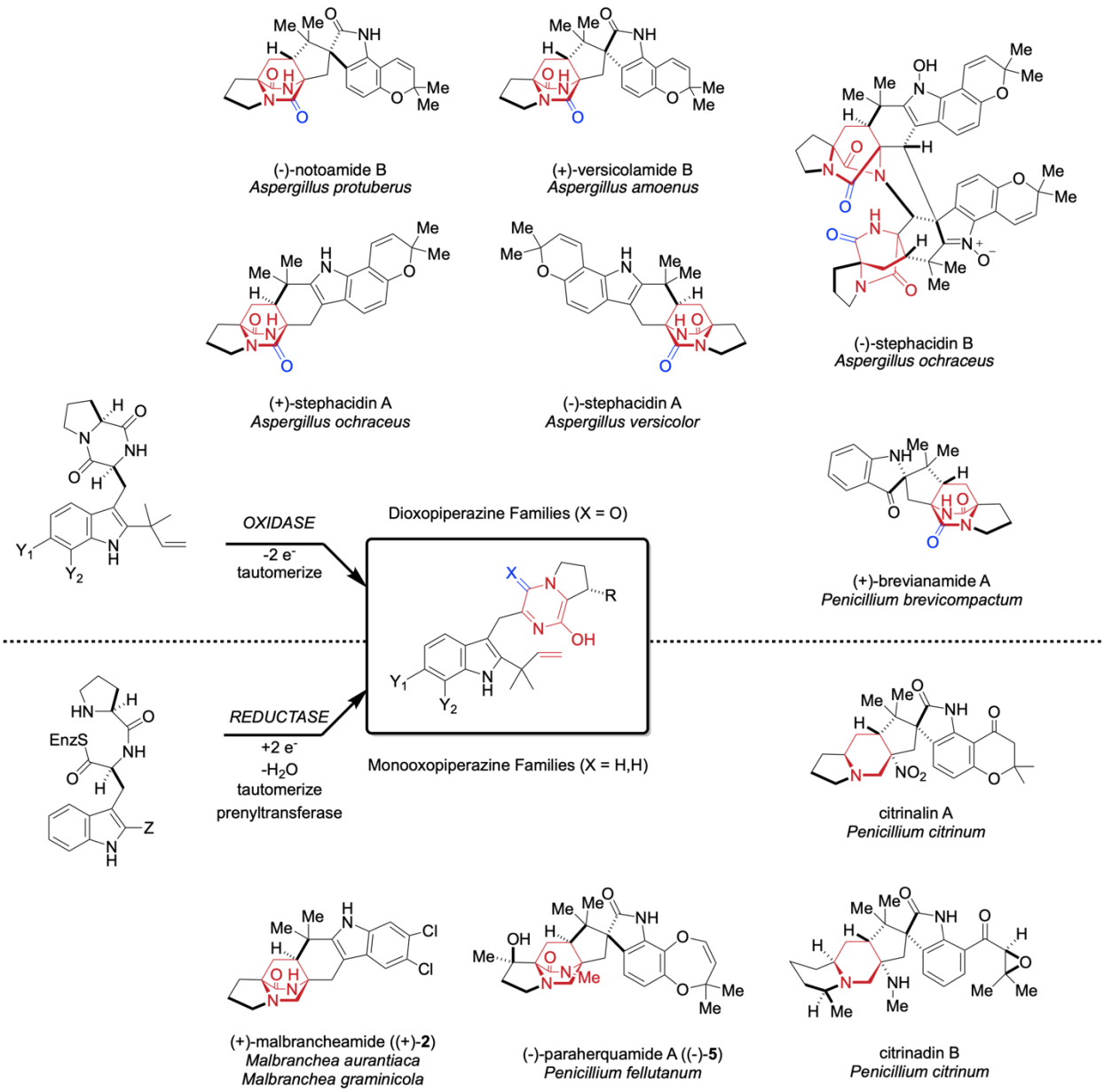
Molecular dynamics simulations were prepared and equilibrated using the GPU code (*pmemd*)[23] of the AMBER 16 package.[24] Parameters for the ligands were generated within the *antechamber* module using the general AMBER force field (*gaff*),[25] with partial charges set to fit the electrostatic potential generated at the HF/6-31(d) level by the RESP model.[26] The partial charges were calculated according to the Merz–Singh–Kollman scheme[27,28] using the Gaussian 09 package.[29] Each protein was immersed in a pre-equilibrated cubic box with a 10 Å buffer of TIP3P[30] 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*). [31] 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.[32] 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·mol$^{-1}$·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·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[33] 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.

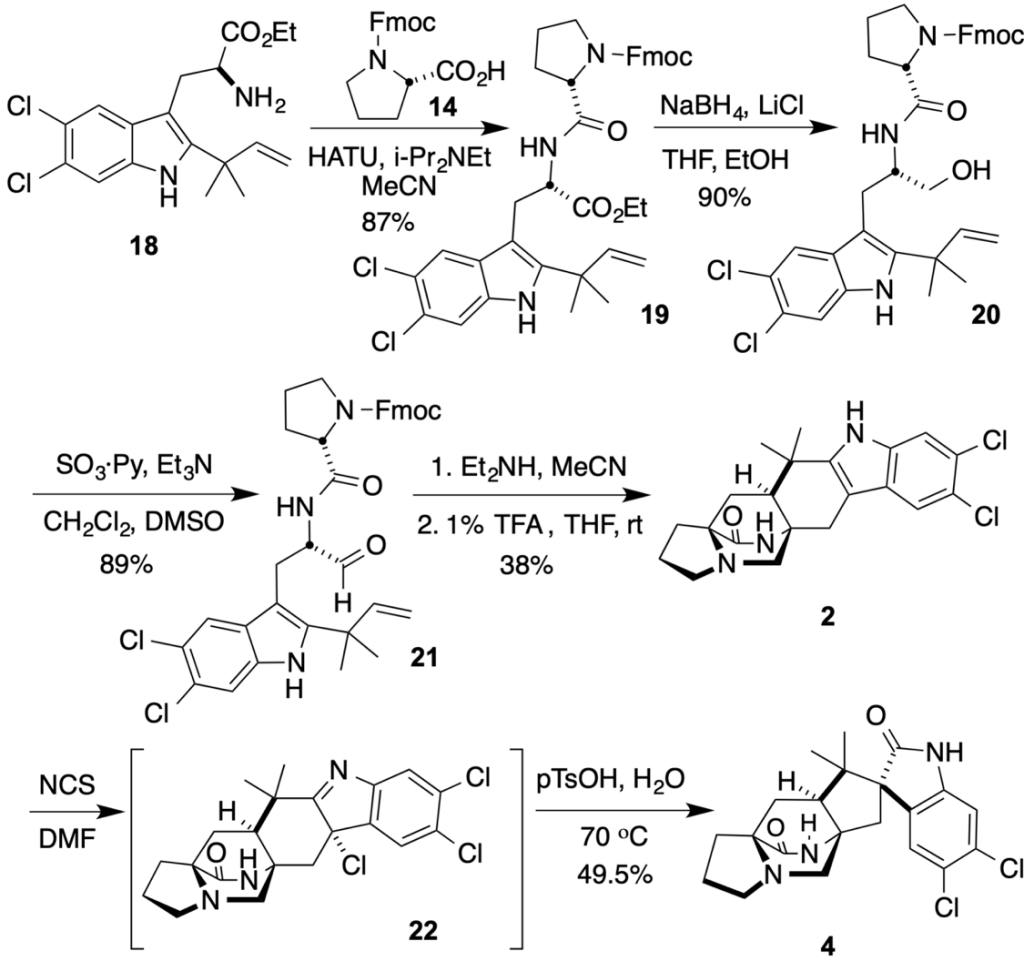
**11. Genetic Disruption of *phqE***

The gene disruption in *Penicillium simplicissimum* was performed by the CRISPR/Cas9 system for filamentous fungi.[34] 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

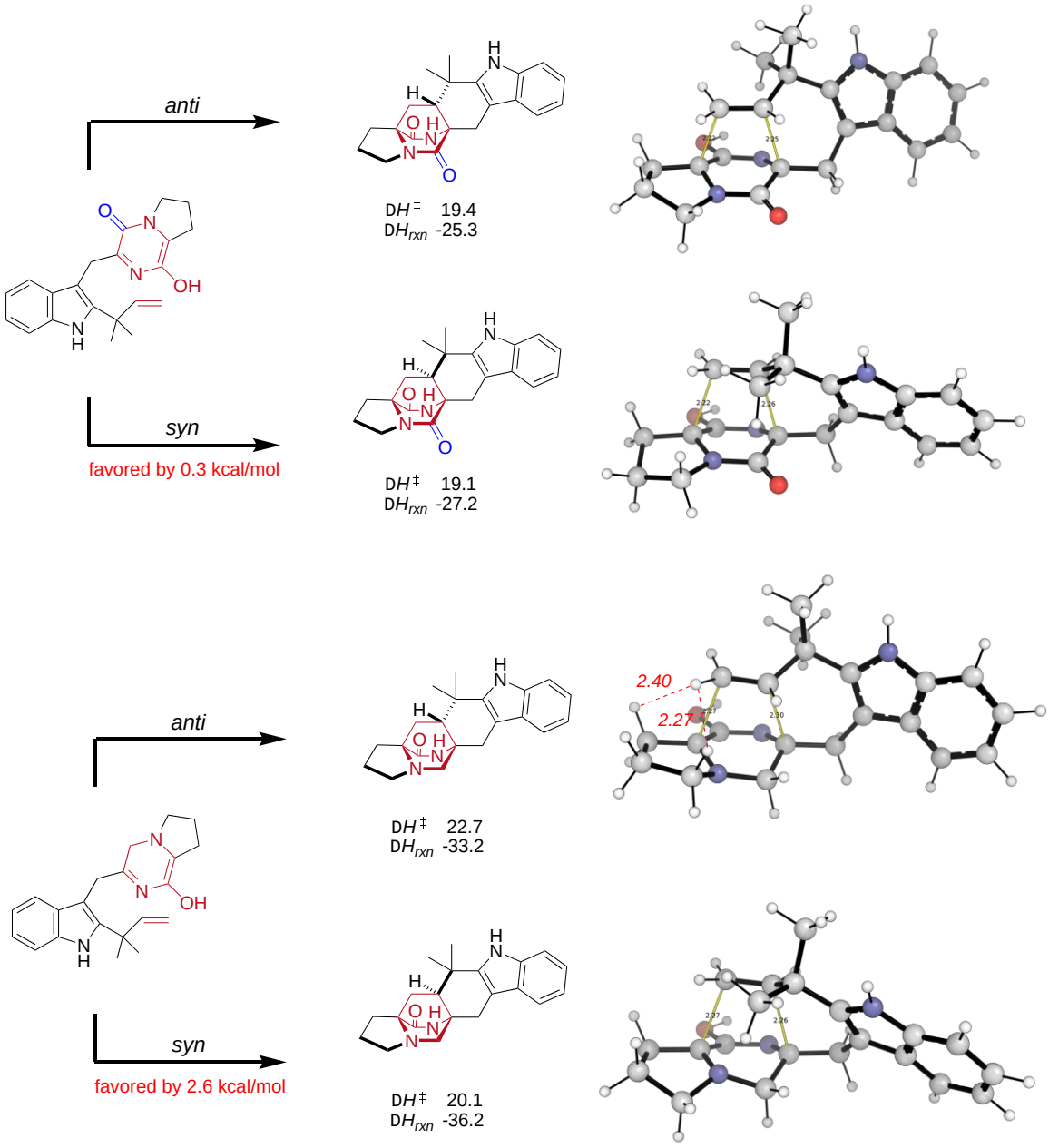
filtration and washed with STC solution (0.8 M sorbitol, 0.05 M Tris-HCl, 0.05 M $CaCl_2$, pH 8), and diluted to a concentration of $2 \times 10^8$ cells $mL^{-1}$. Then, the circular plasmid was added to the 200 μL protoplasts solution, and incubated on ice for 30 min, which was blended with 2 mL 30% 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.



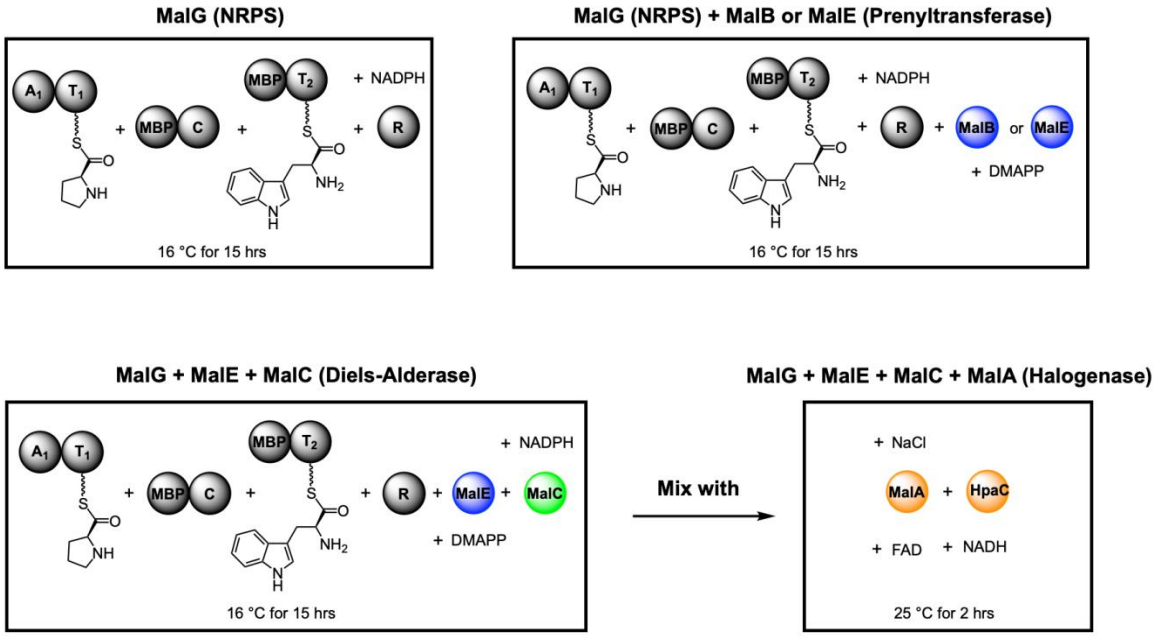

**Supplementary Figure 1.** Unified biogenesis of the dioxopiperazine and monooxopiperazine families of alkaloids possessing the bicyclo[2.2.2]diazaoctane core structures. The biocyclo[2.2.2]diazaoctane group is colored in red. The extra ketone group of dioxopiperazines are highlighted in blue.



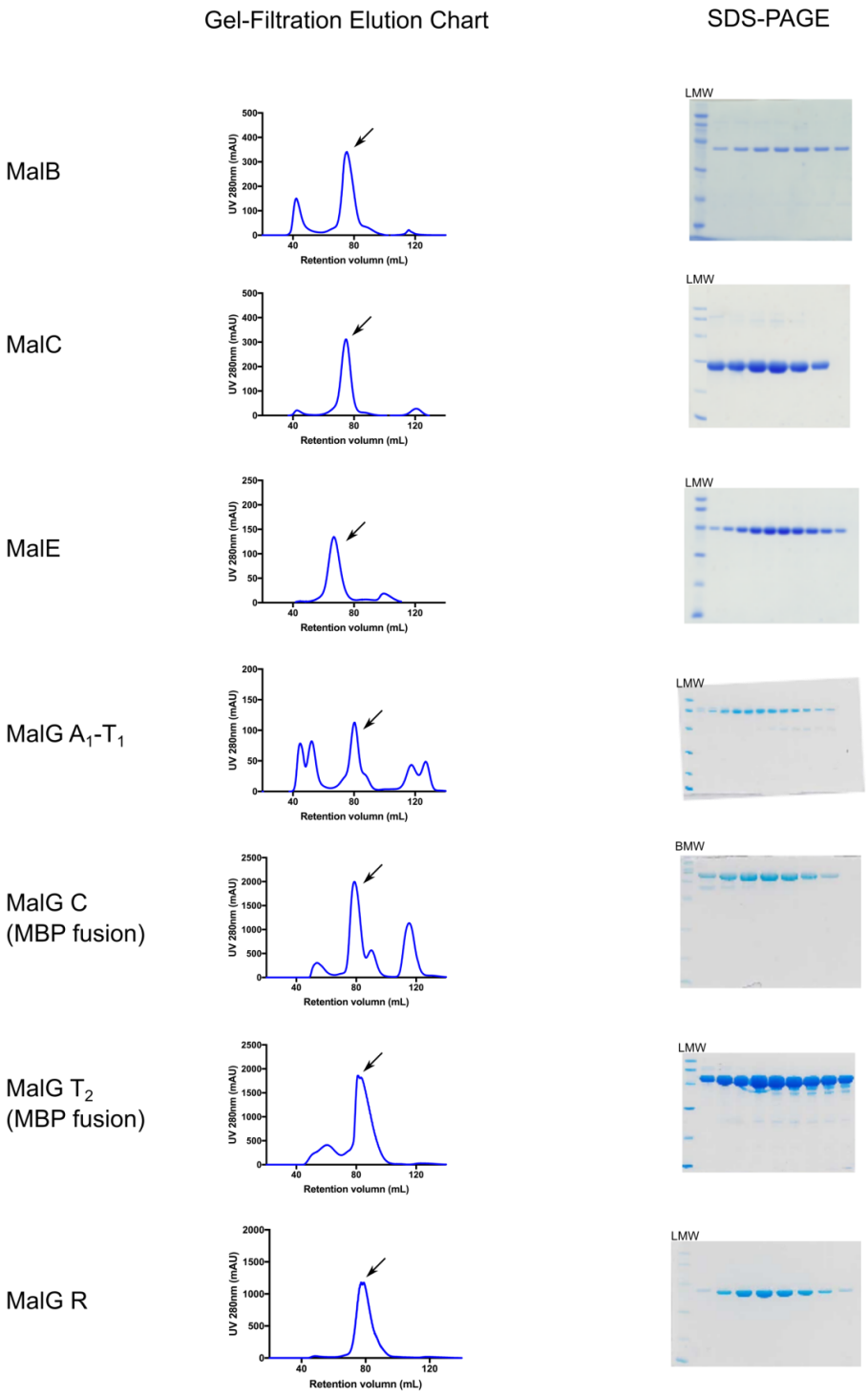

**Supplementary Figure 2.** Biomimetic synthesis of racemic malbrancheamide and spiromalbramide. We have applied an analogous strategy to two additional natural products, racemic (±)-malbrancheamide (**2**) and (±)-spiromalbramide (**4**) that underscores the utility of this new biomimetic paradigm. 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.



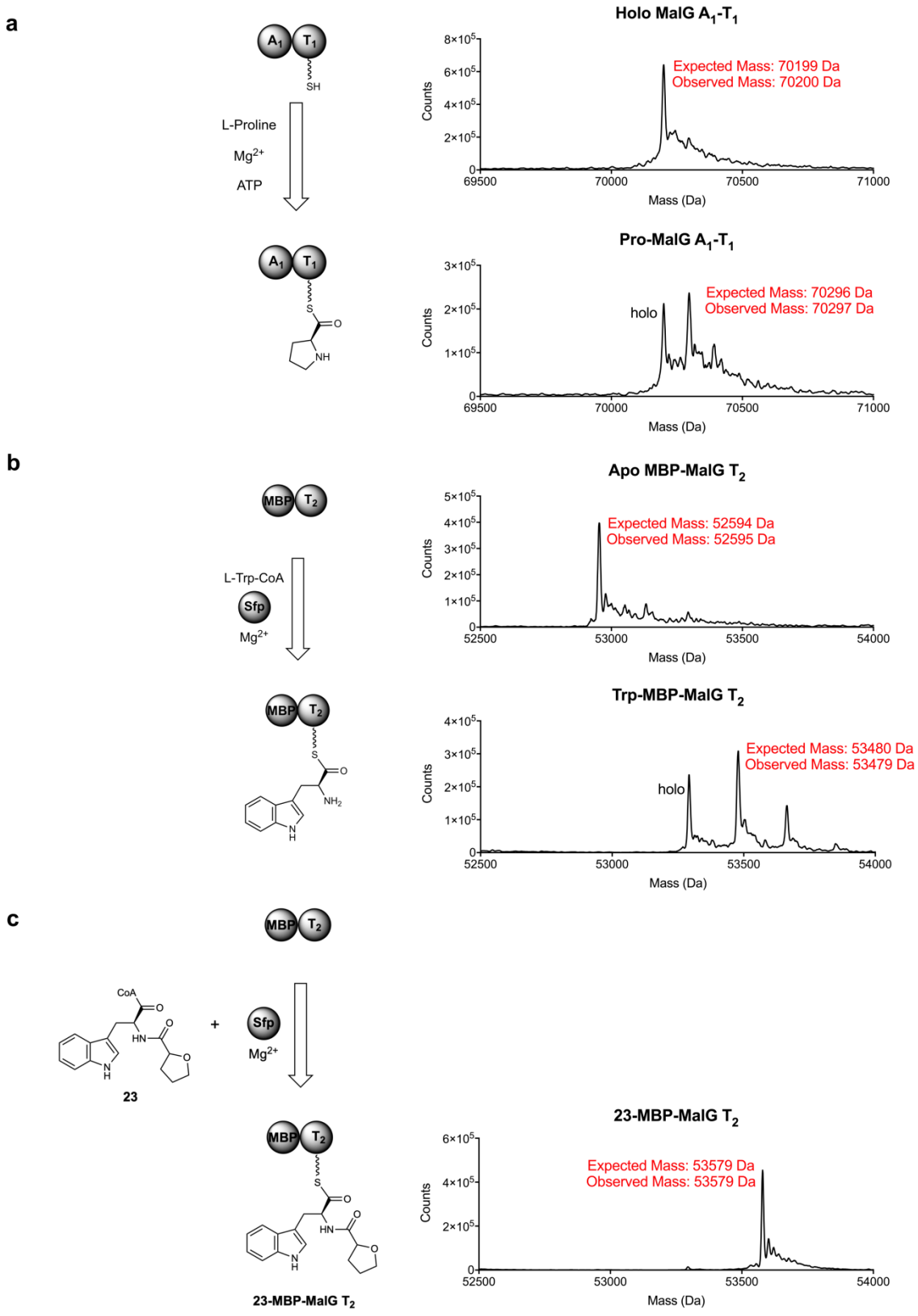

**Supplementary Figure 3.** M062X-D3/6-31+G(d,p) intramolecular Diels-Alder transition structures (TSs) for both oxidation states represented by azadiene, predict a relatively modest *syn*-: *anti*-diastereoselectivity for the more oxygenated azadiene species of 0.3 kcal/mol.[35] This value has been 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.



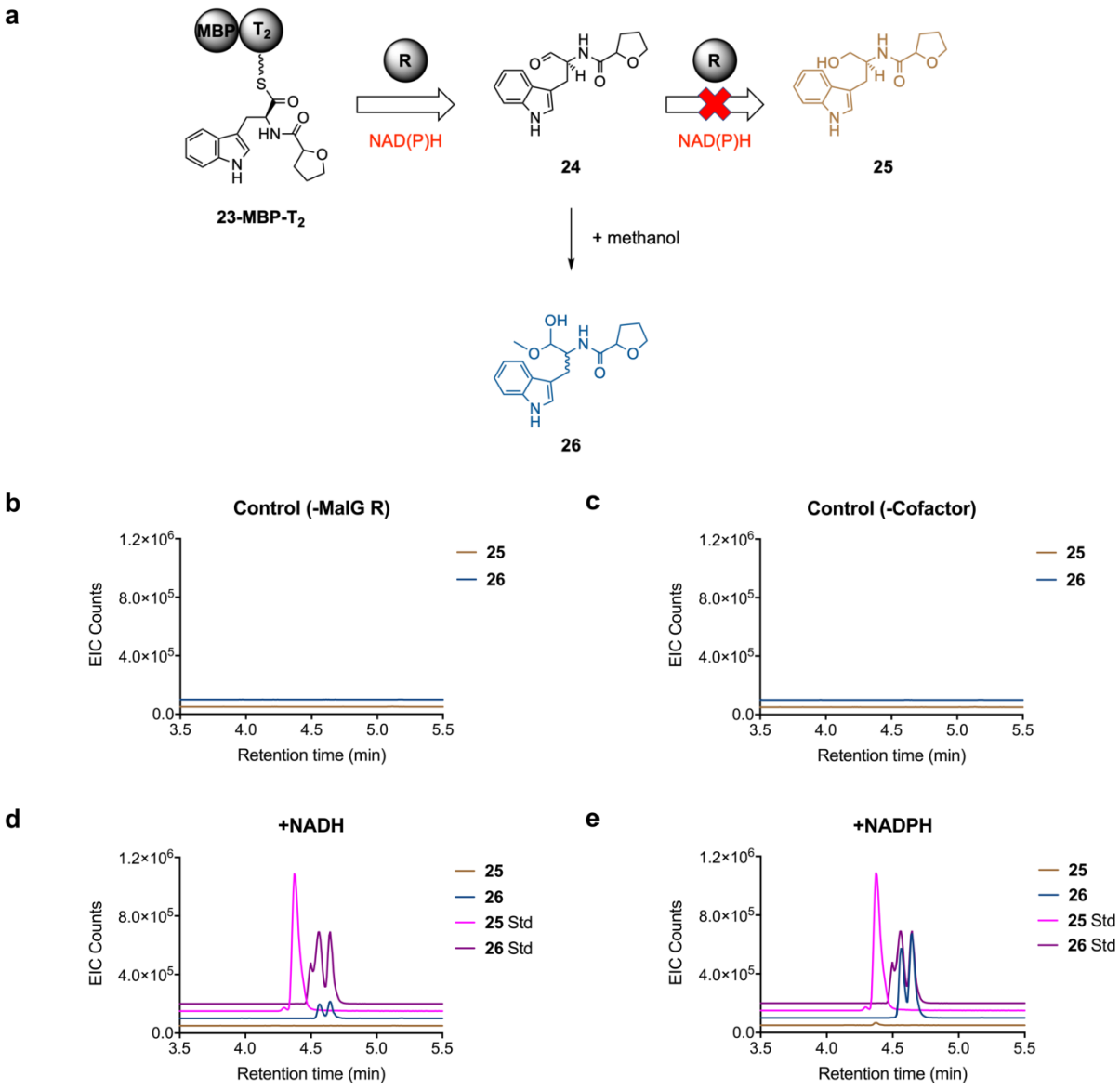

**Supplementary Figure 4.** Scheme of *in vitro* reconstitution assays. The flavin-dependent MalA halogenase requires a recycling system to reduce FAD to $FADH_2$ after each catalytic cycle, here HpaC reductase + NADH.[1] MBP = maltose binding protein.



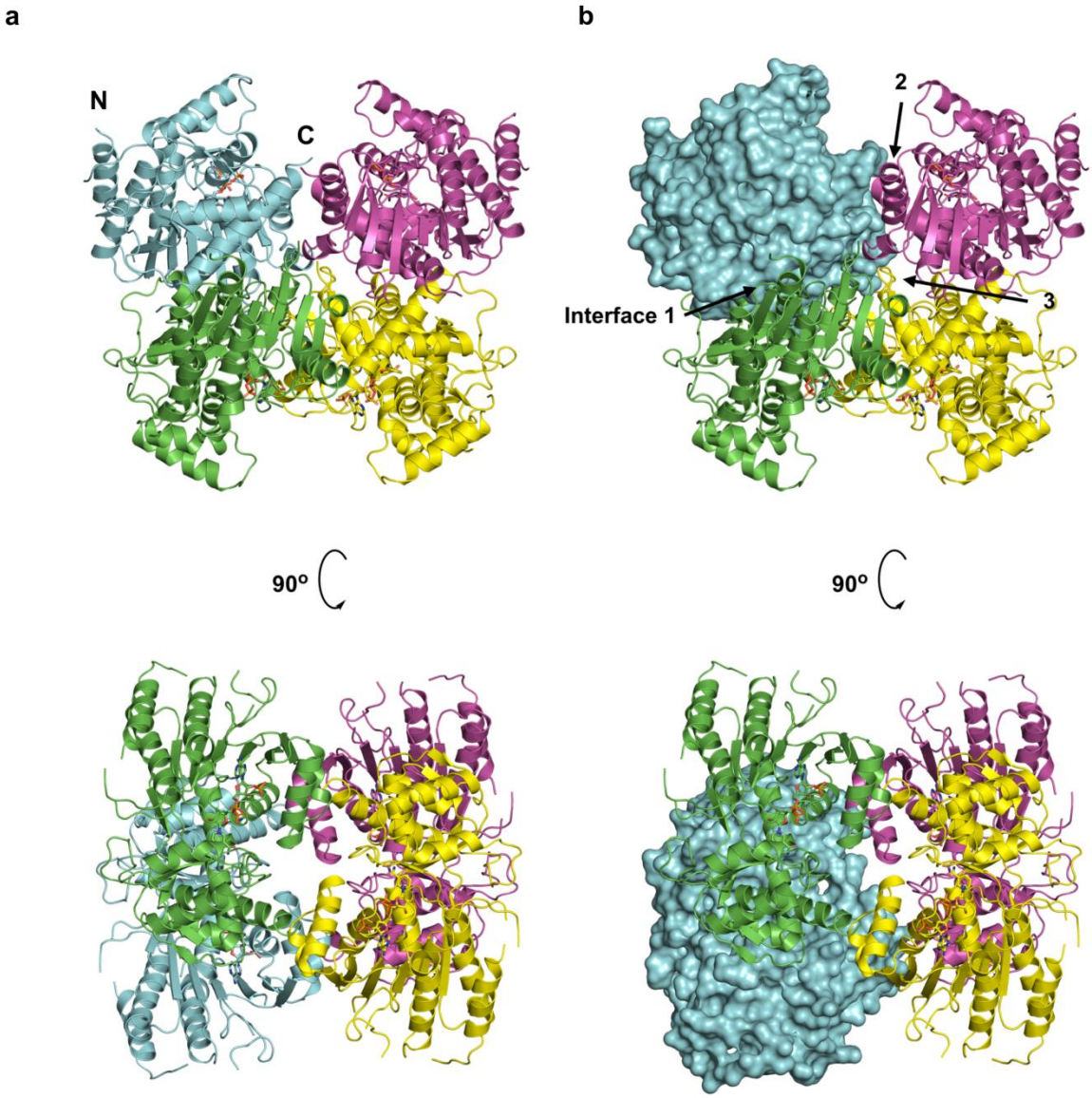

**Supplementary Figure 5.** Purification of enzymes and enzymatic domains involved in malbrancheamide biosynthesis with gel filtration profiles at left and SDS gels of the indicated peaks at right. Recombinant MalA was produced as previously described.[1] Protein molecular weight standards were LMW: 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.



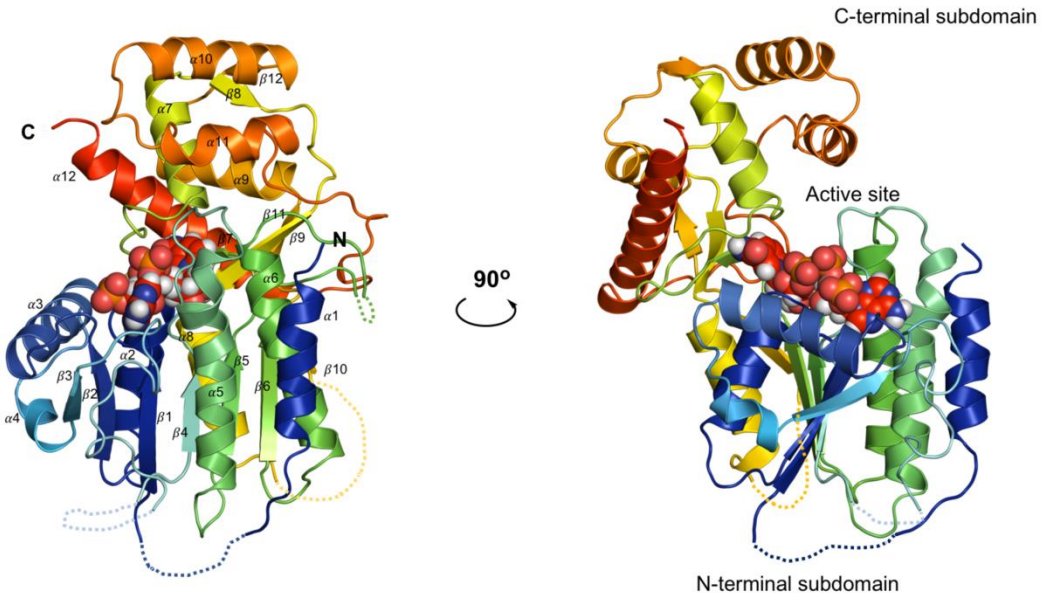

**Supplementary Figure 6.** Substrate loading of MalG $T_1$ and $T_2$ domains. Protein MS was applied to analyze efficiency of substrate loading, confirming successful loading in all cases.



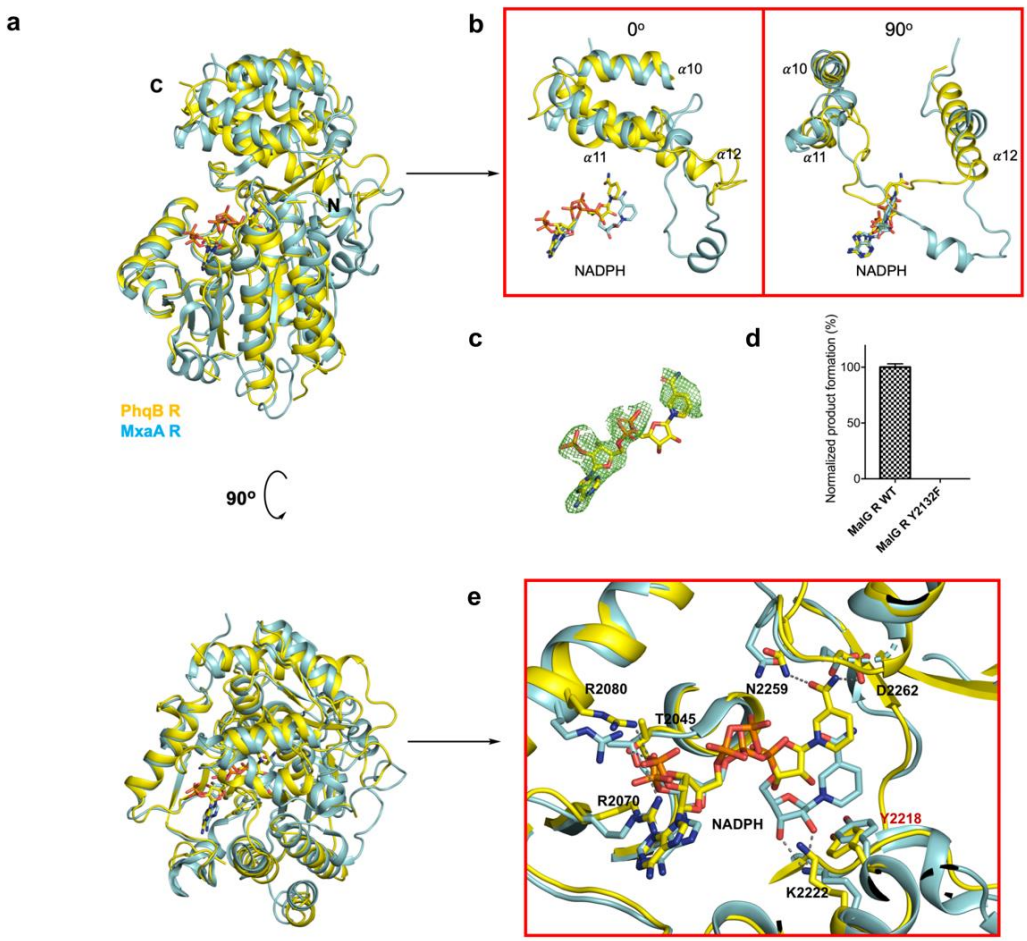

**Supplementary Figure 7.** MalG R catalyzes a 2-electron reductive release reaction. a. Reaction scheme of MalG R. b – c. EIC profiles of control experiments, with no enzyme (b) or no cofactor (c). d – e. EIC profiles of the MalG R-catalyzed reaction, using NADH (d) or NADPH (e) as cofactor. NADPH is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



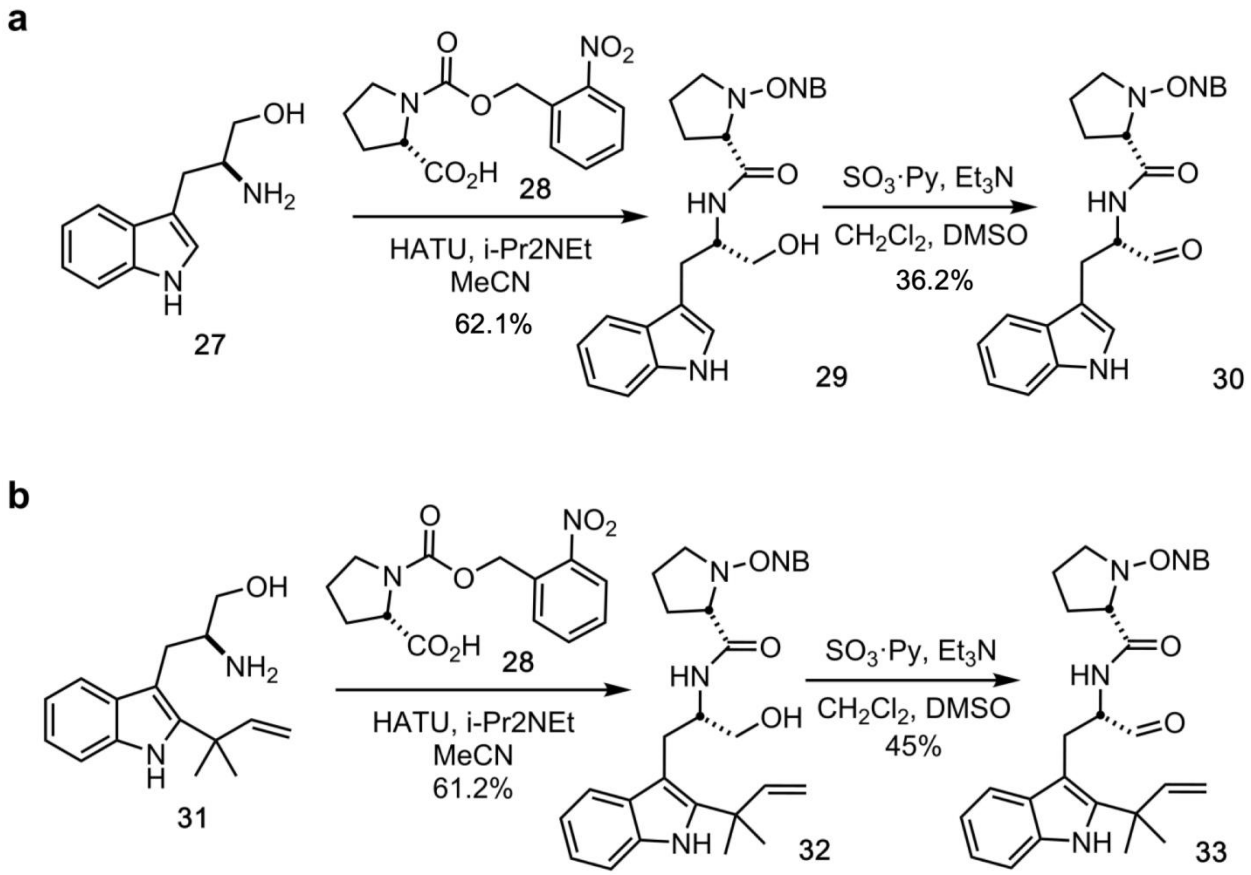

**Supplementary Figure 8.** Crystal structure of the PhqB R tetramer, with the four subunits shown in contrasting colors cartoon representation in (a), and the cyan subunit shown in surface representation in (b). The N- and C-termini are marked for the cyan subunit. The excised PhqB R tetramer has D2 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, 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.



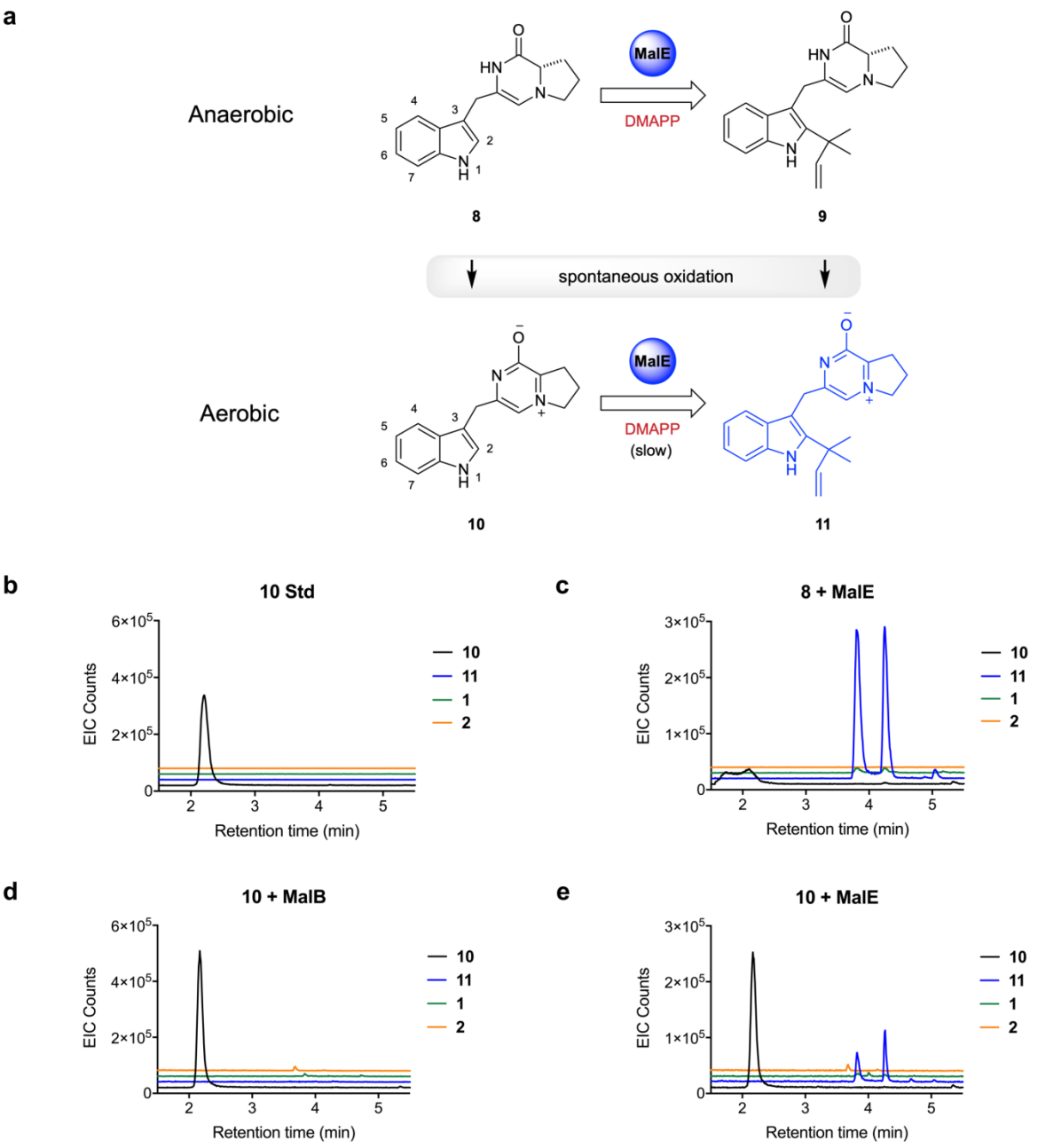

**Supplementary Figure 9.** Structure of the PhqB R subunit colored as a rainbow from blue N-terminus to red C-terminus. PhqB R consists of an N-terminal nucleotide-binding subdomain and a C-terminal substrate-binding subdomain, which recognizes Pro-Trp-$T_2$. The nucleotide-binding subdomain has a typical 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).



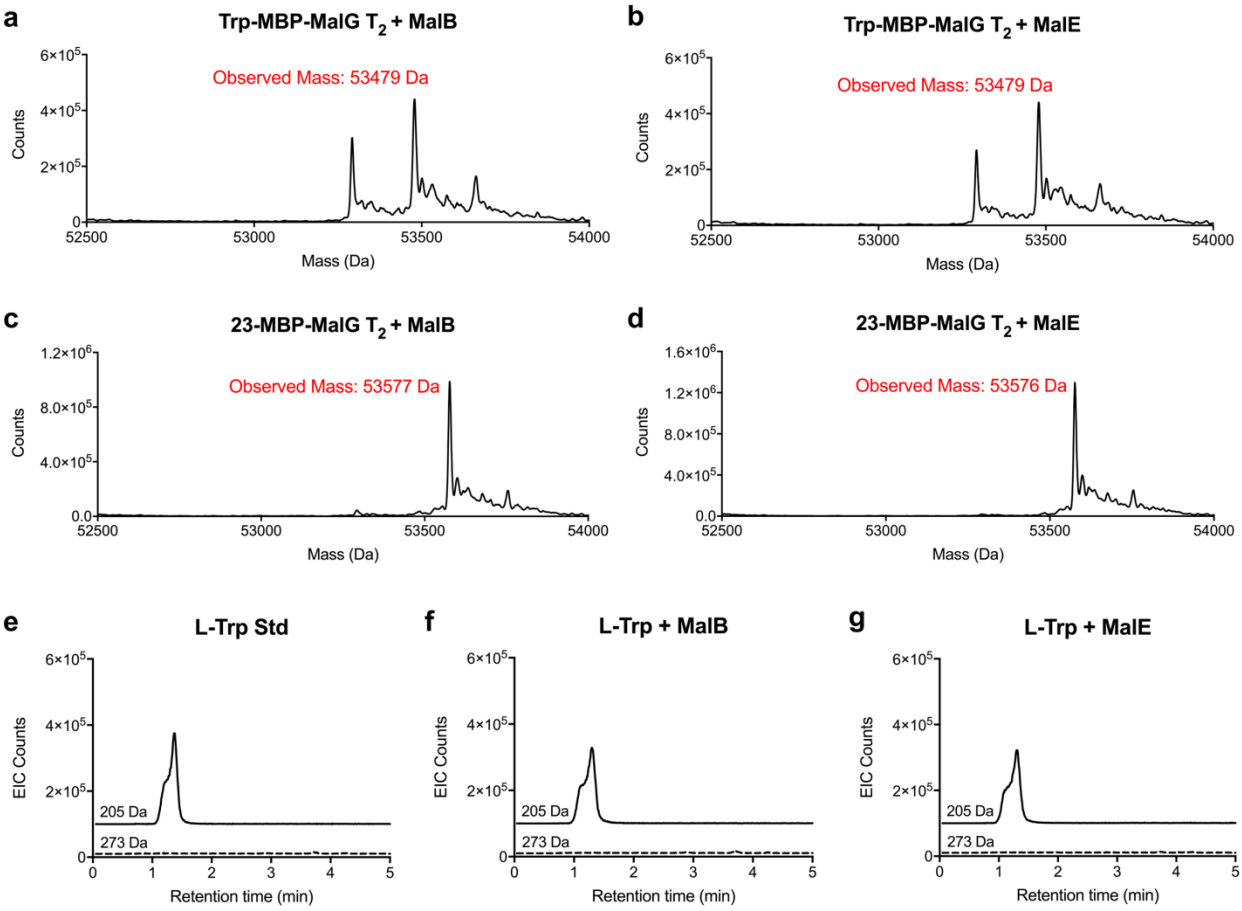

**Supplementary Figure 10.** Comparison of fungal and bacterial NRPS R domains. a. Superposition of the PhqB R subunit (yellow) and the bacterial MxaA R[36] (cyan, PDB ID: 4U7W). b. The major structural difference lies in the C-terminal subdomain. Relative to bacterial NRPS R domain structures, PhqB 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 keto-reductases.[37] Substrate may be required for optimal cofactor binding d. Lack of detectable activity in MalG R/Y2132F (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[36,38,39] and other Tyr-dependent SDRs[40,41] 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.



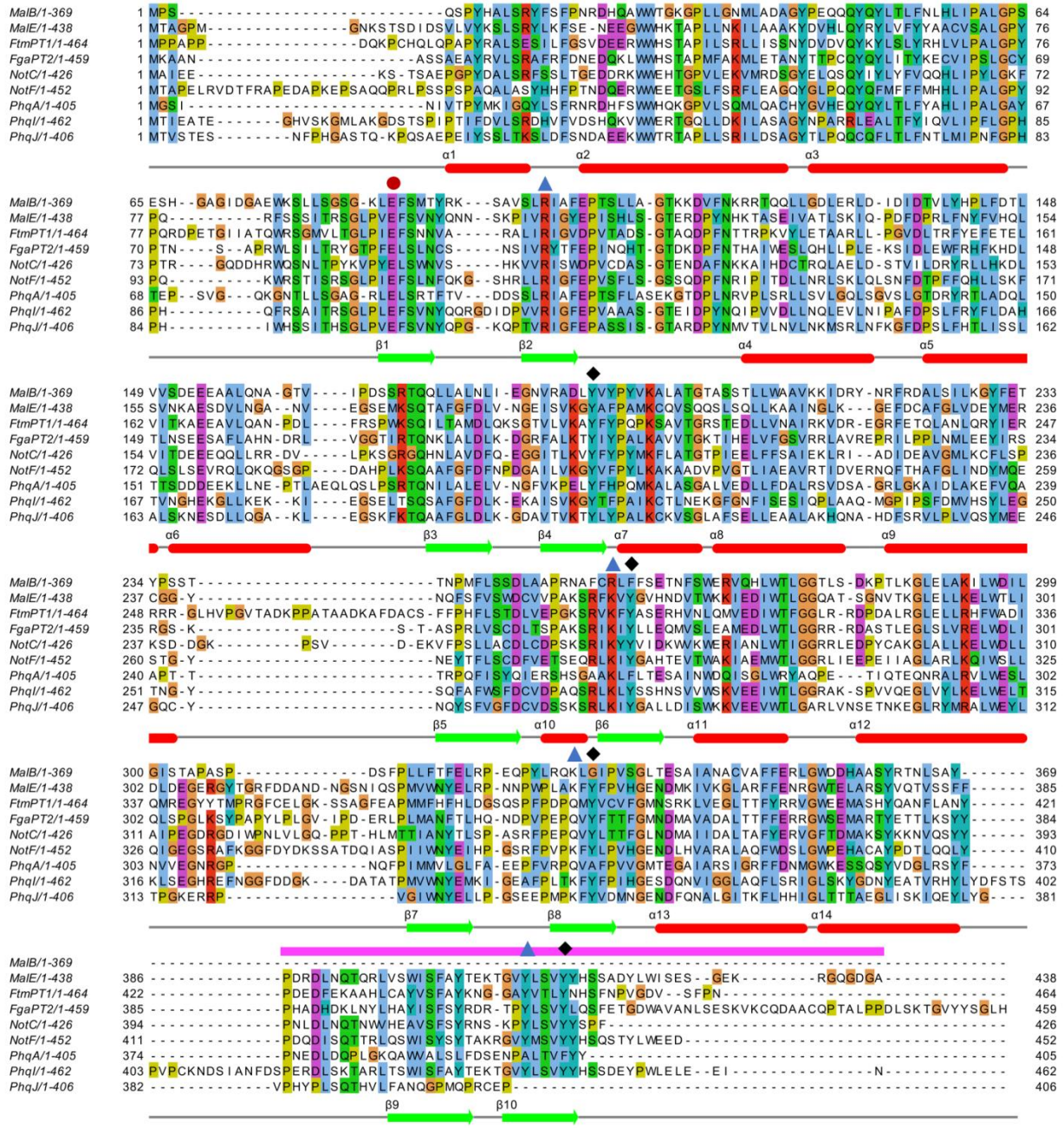

**Supplementary Figure 11.** Synthetic scheme for the ONB protected dipeptide aldehyde **30** (a) and ONB protected prenyl dipeptide **33** (b).



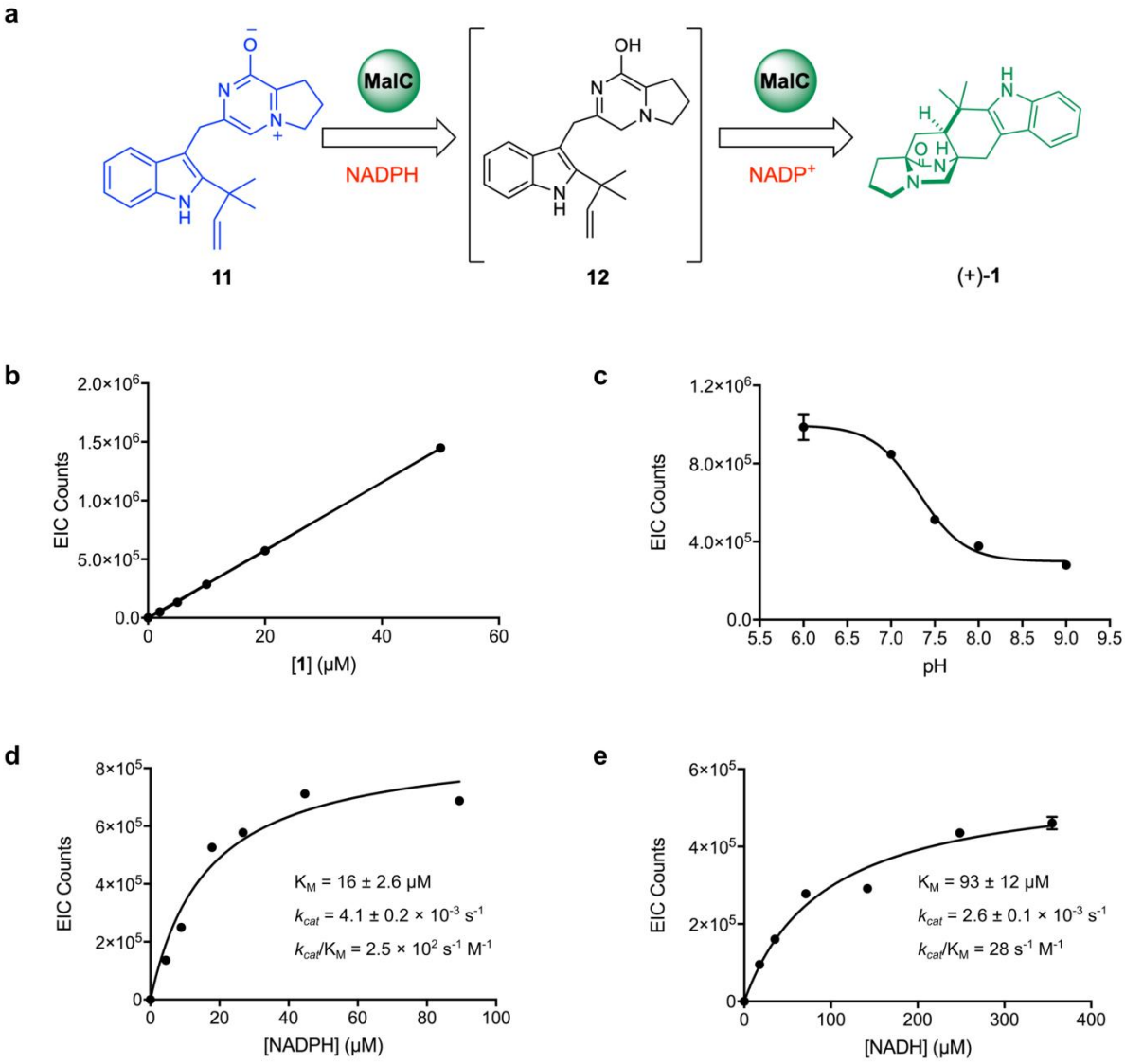

**Supplementary Figure 12.** Prenylation of **8** (anaerobic) and **10** (aerobic). **8** is the natural substrate of MalE. **10** can be prenylated by MalE but not MalB. a. C2 reverse prenyltransfer reaction scheme. b. EIC profile of **10** authentic standard. c. EIC profile of **8** prenylation by 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 anaerobic conditions. No prenylated product was detected. e. EIC profile of **10** prenylation by MalE, illustrating that oxidized **10** is less favored than **8**.



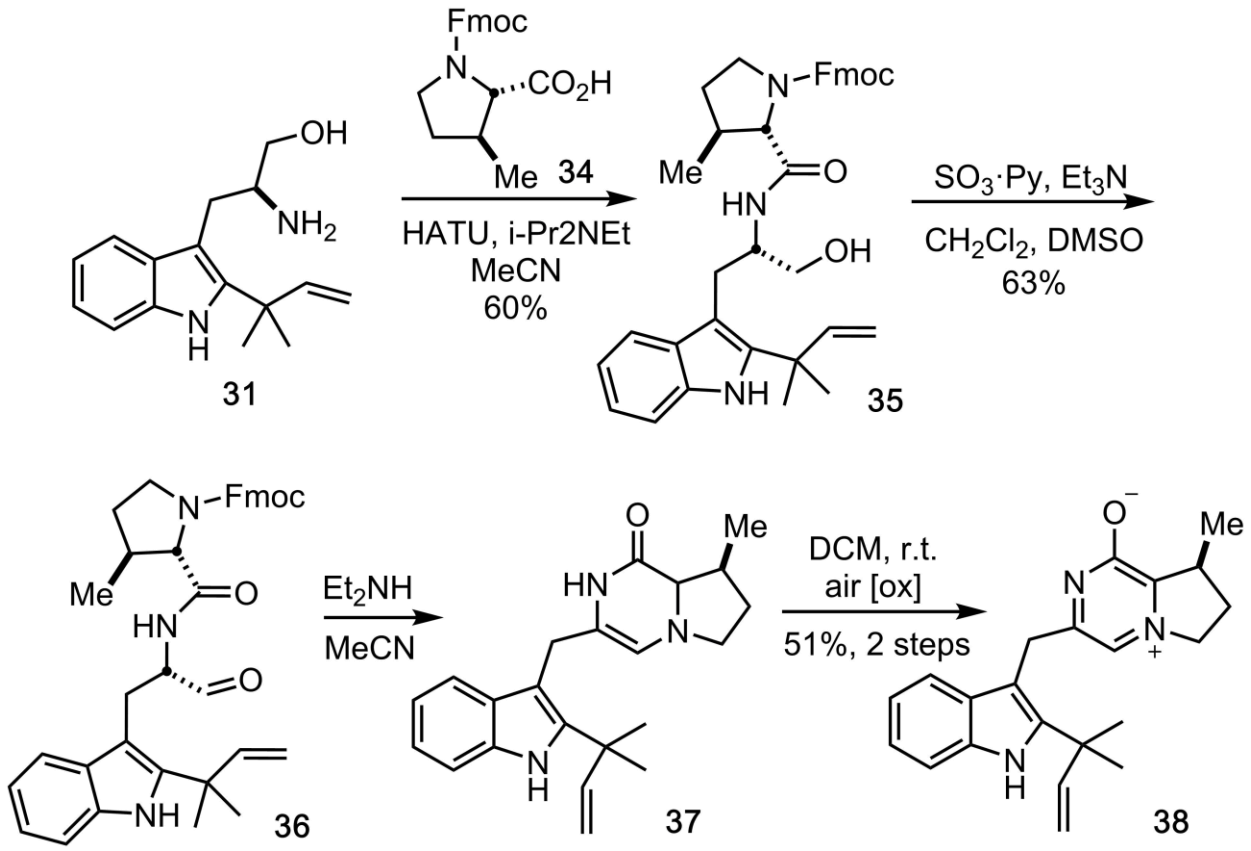

**Supplementary Figure 13.** Timing of prenylation reaction after NRPS reductive offloading. a. Protein MS profile of L-Trp-MBP-MalG $T_2$ after 2 hours of MalB and DMAPP incubation. b. Protein MS profile of L-Trp-MBP-MalG $T_2$ after 2 hours of MalE and DMAPP incubation. c. Protein MS profile of **23**-MBP-MalG $T_2$ 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. 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.



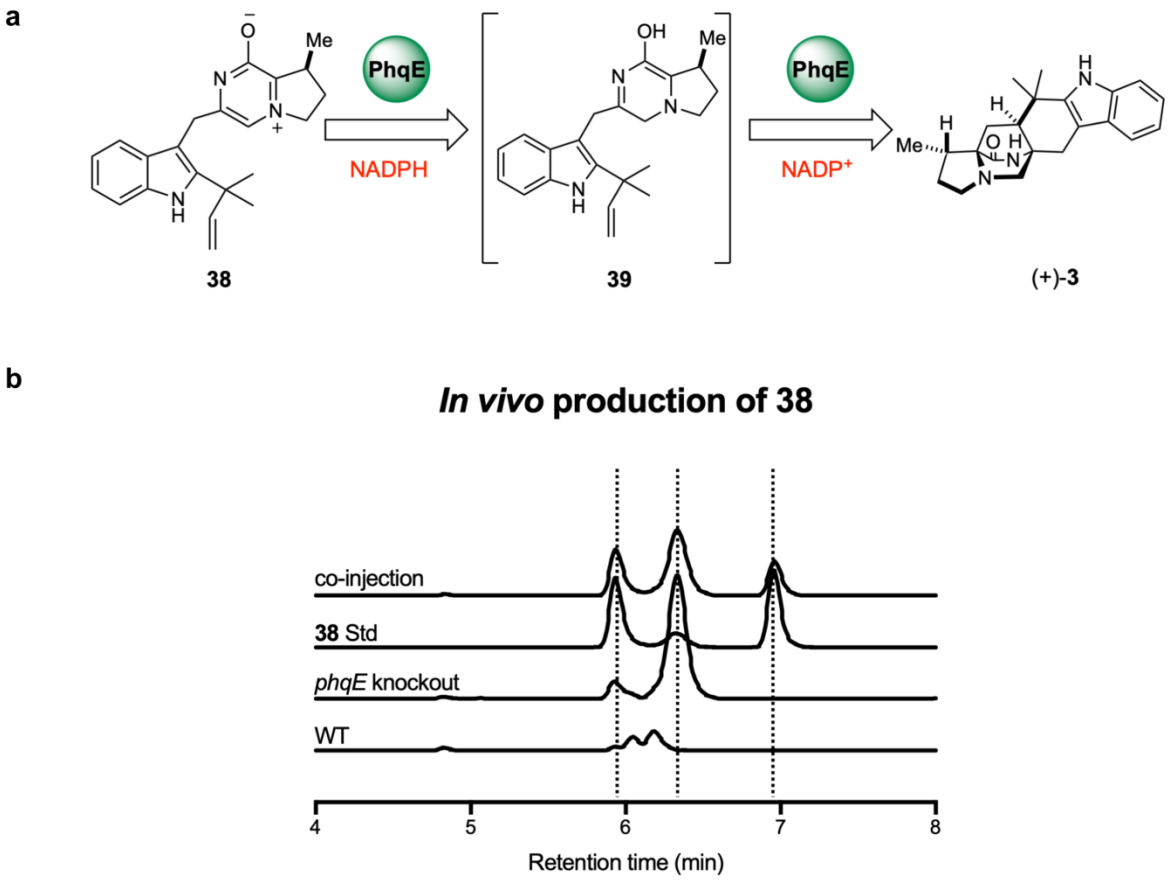

**Supplementary Figure 14.** Multiple sequence alignment of fungal indole prenyltransferases. The catalytic base Glu is highlighted with a red dot. Residues critical for coordinating the cofactor DMAPP are highlighted with blue triangles. Four Tyr that are expected to shield the active site are highlighted 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.



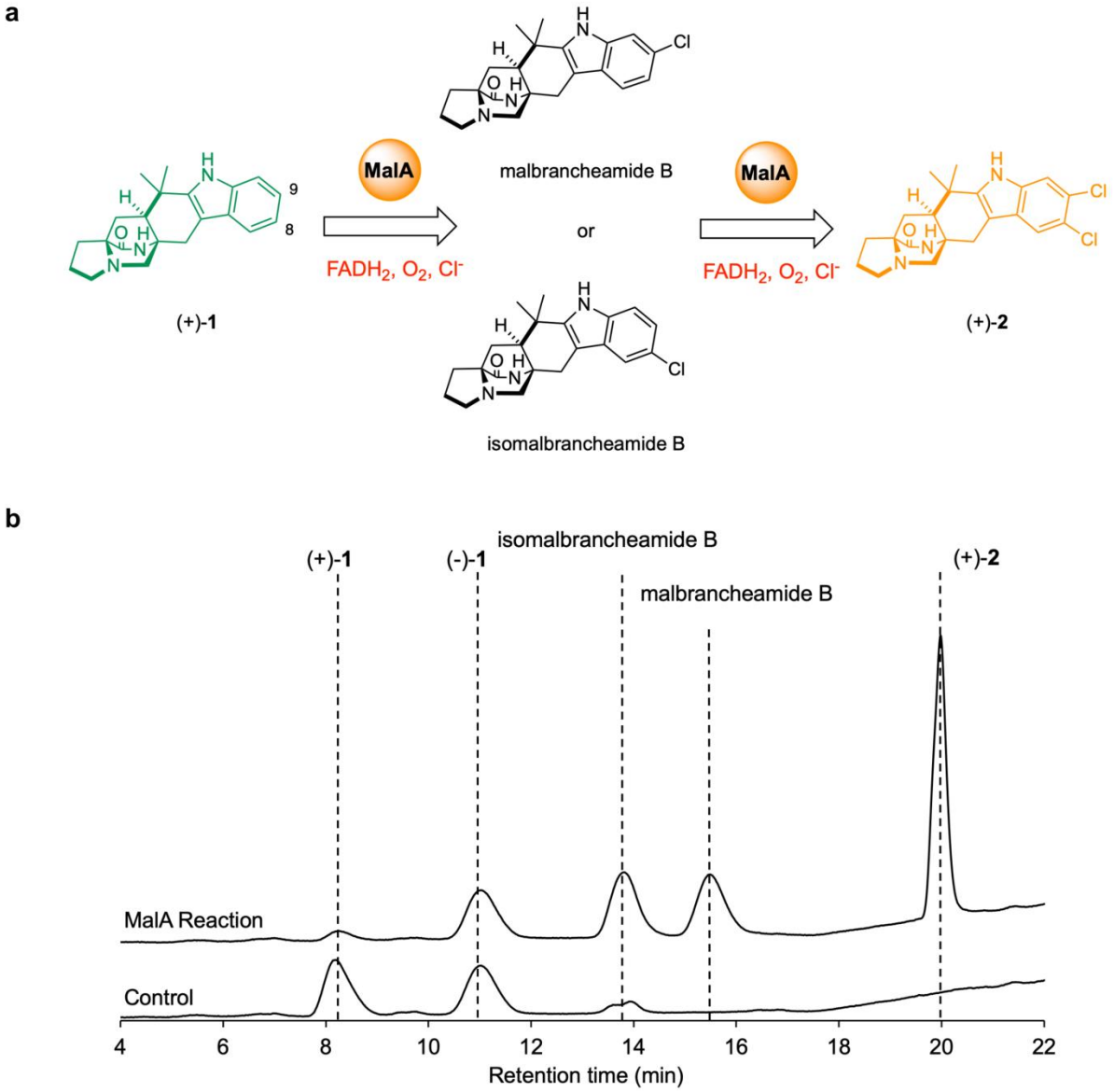

**Supplementary Figure 15.** a. MalC catalyzes a two-step reaction of **11** to (+)-**1**. b. Standard curve of (+)-**1**, presenting linear correlation of EIC counts to (+)-**1** concentration. c. pH profile of the MalC-catalyzed reaction, efficiency of which decreases drastically beyond neutral pH.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



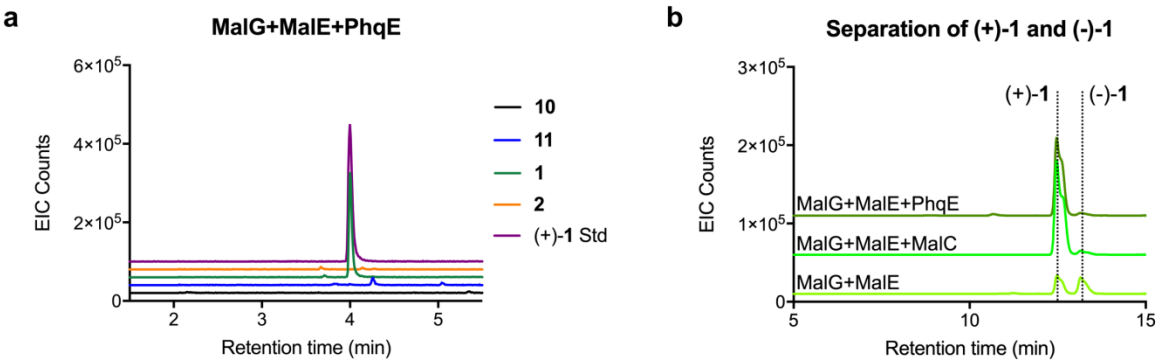

**Supplementary Figure 16.** Synthetic scheme of *β*-methyl prolyl prenyl zwitterion **38**, an intermediate in paraherquamide biosynthesis.



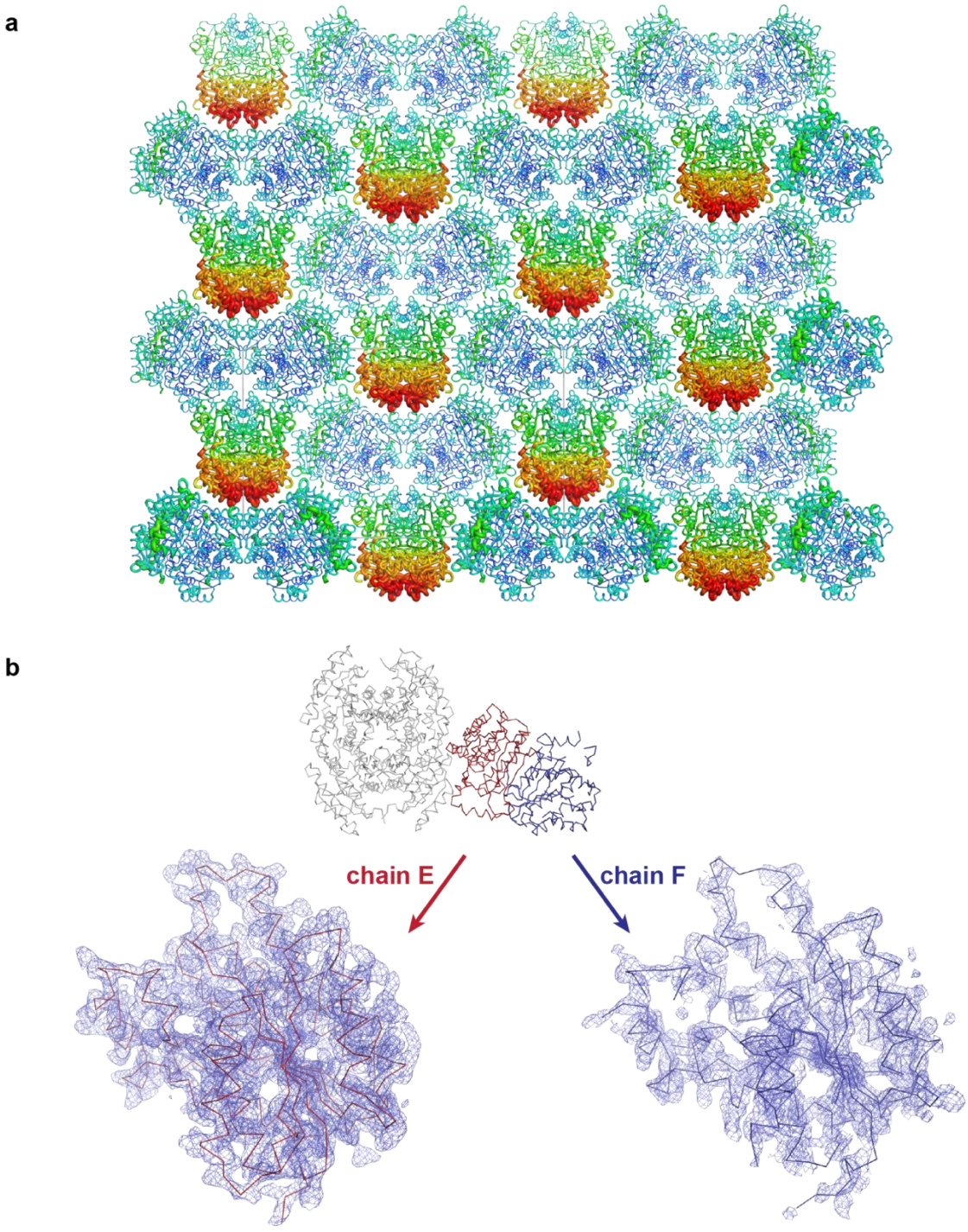

**Supplementary Figure 17.** Production analysis of **38** (M+H+ m/z = 348) by TOF-MS from *Penicillium simplicissimum phqE* mutant. *phqE* is the homologous gene of *malC* in the paraherquamide biosynthetic pathway. a. PhqE catalyzes a two-step reaction of **38** to (+)-preparaherquamide **3**. b. *In 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.



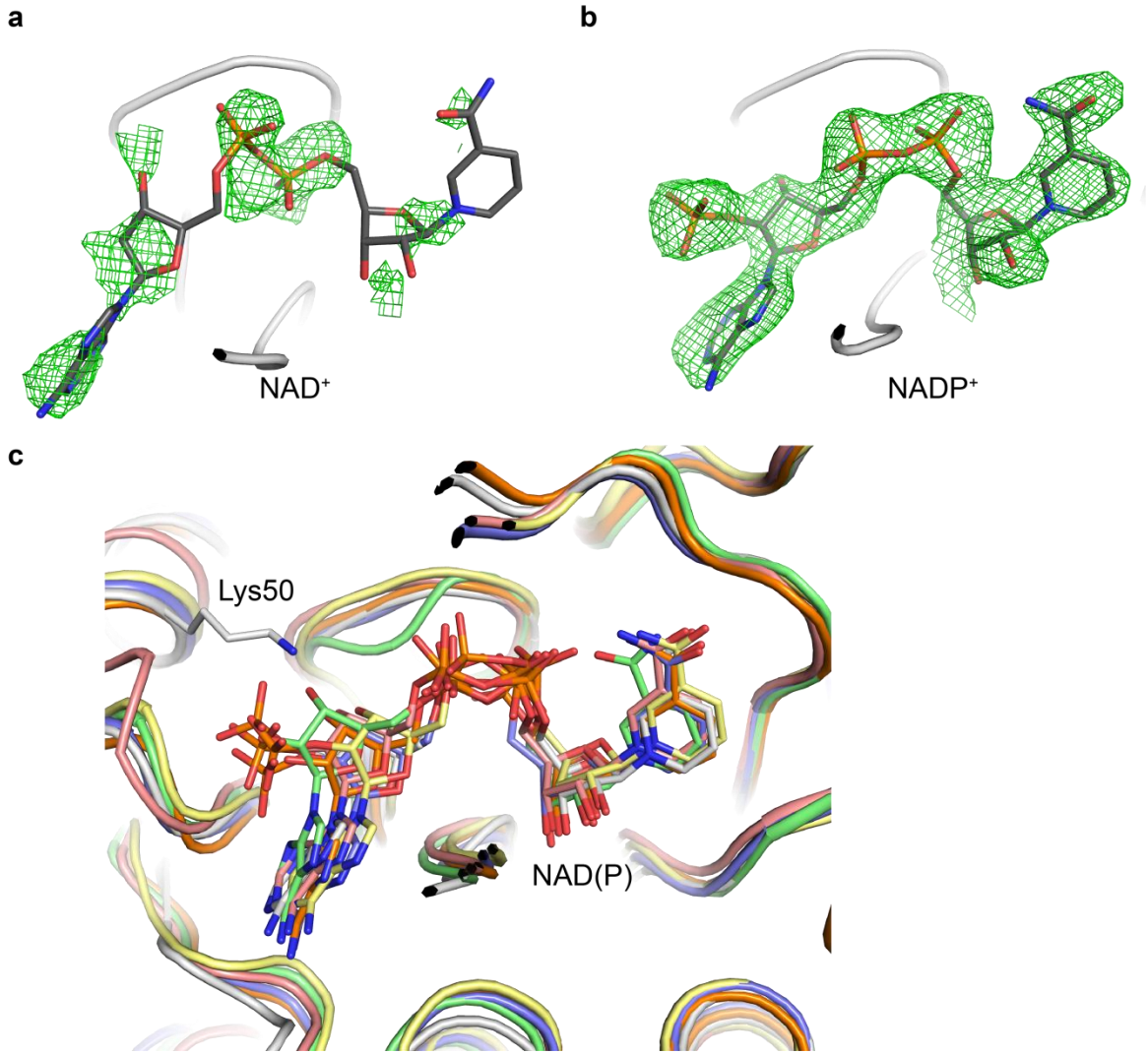

**Supplementary Figure 18.** a. MalA catalyzes an iterative dihalogenation reaction, converting (+)-**1** to (+)-**2**. b. MalA is stereospecific and does not react on (–)-**1**. The Y-axis is UV 240 nm absorbance.



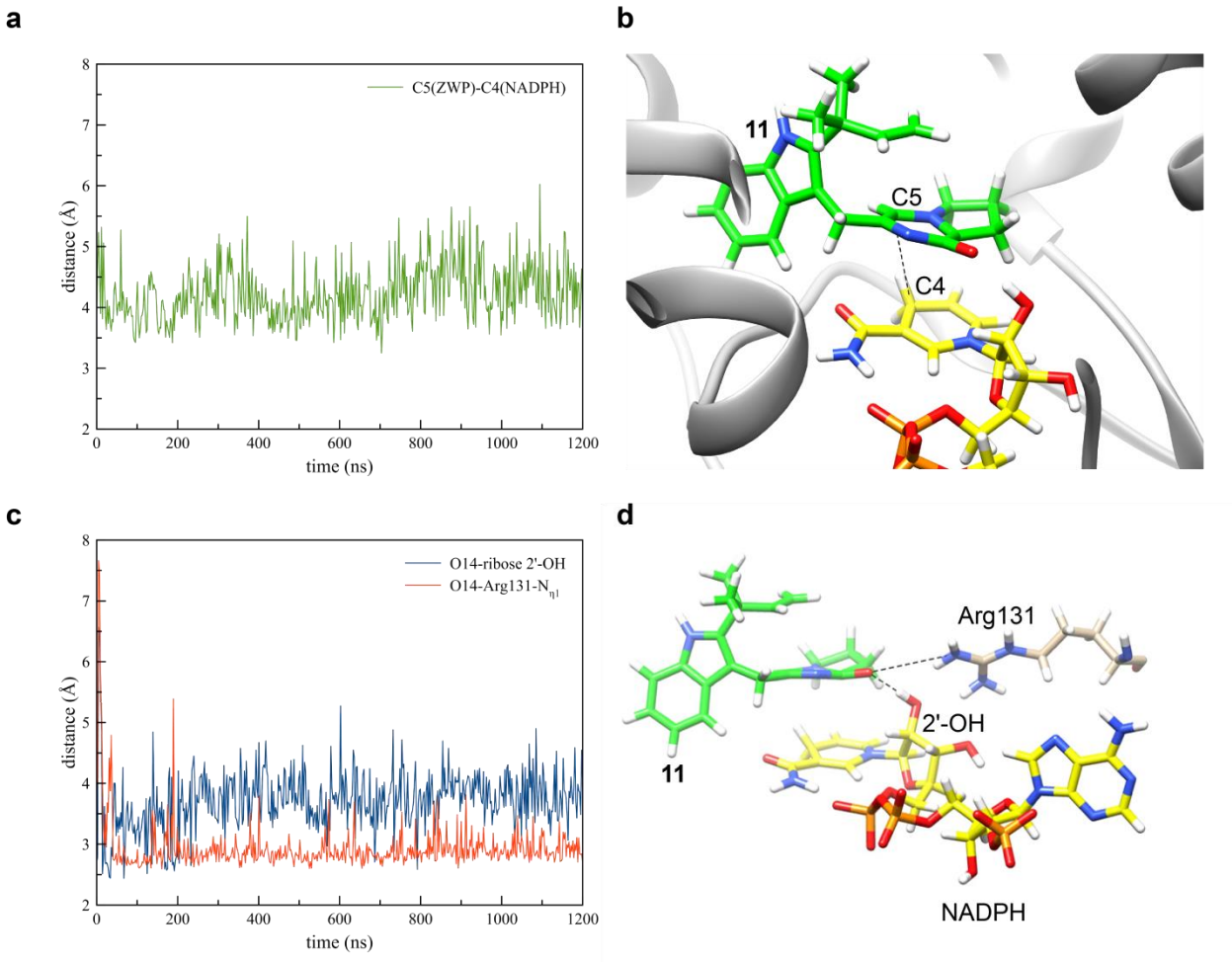

**Supplementary Figure 19.** PhqE is a bifunctional reductase and Diels-Alderase. a. EIC profile of *in vitro* malbrancheamide pathway reconstitution assay, with MalC replaced by PhqE. b. Chiral separation of (+)-**1** and (–)-**1**, indicating that PhqE is diastereo- and enantioselective. Reconstitution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



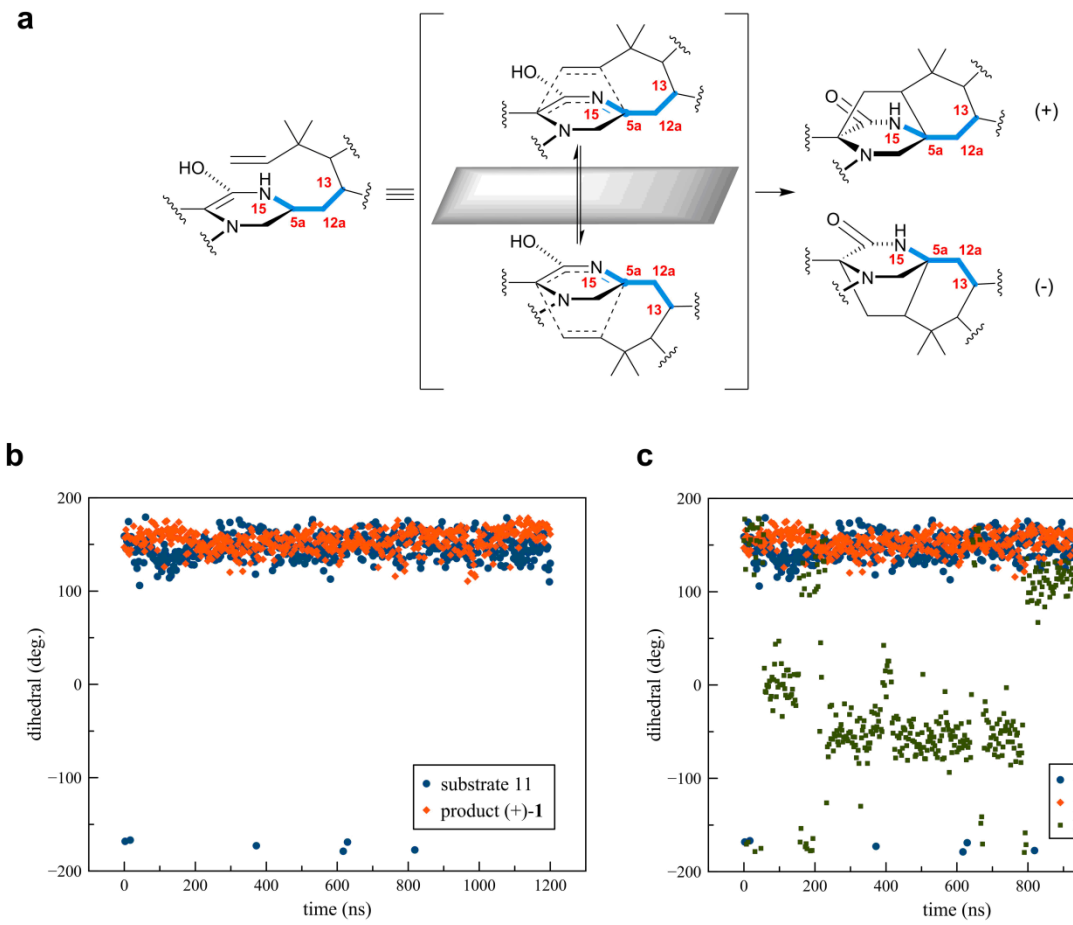

**Supplementary Figure 20.** PhqE crystal lattice in space group *C*2. a. Packing diagram colored by B factor from 20 Å$^2$ in blue to 50 Å$^2$ in red. The asymmetric unit contains 1.5 tetramers. One tetramer (chains A – D) is well-ordered, while in the half-tetramer (chains E and F) chain F is poorly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



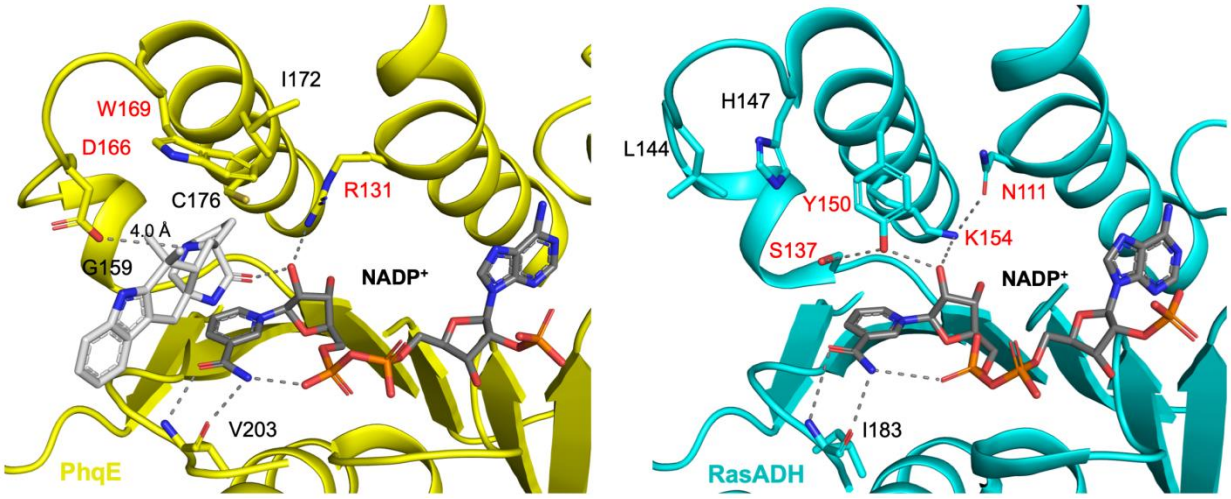

**Supplementary Figure 21.** Cofactor binding of PhqE. a – b. Poor omit density for $NAD^+$ in PhqE (contoured at 2σ) compared to well-ordered $NADP^+$ (contoured at 2σ). c. A selection of bacterial SDRs (*Burkholderia cenocepacia*, PDB 5U2W, pink; *Sinorhizobium meliloti*, PDB 3TOX, yellow; *Ralstonia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



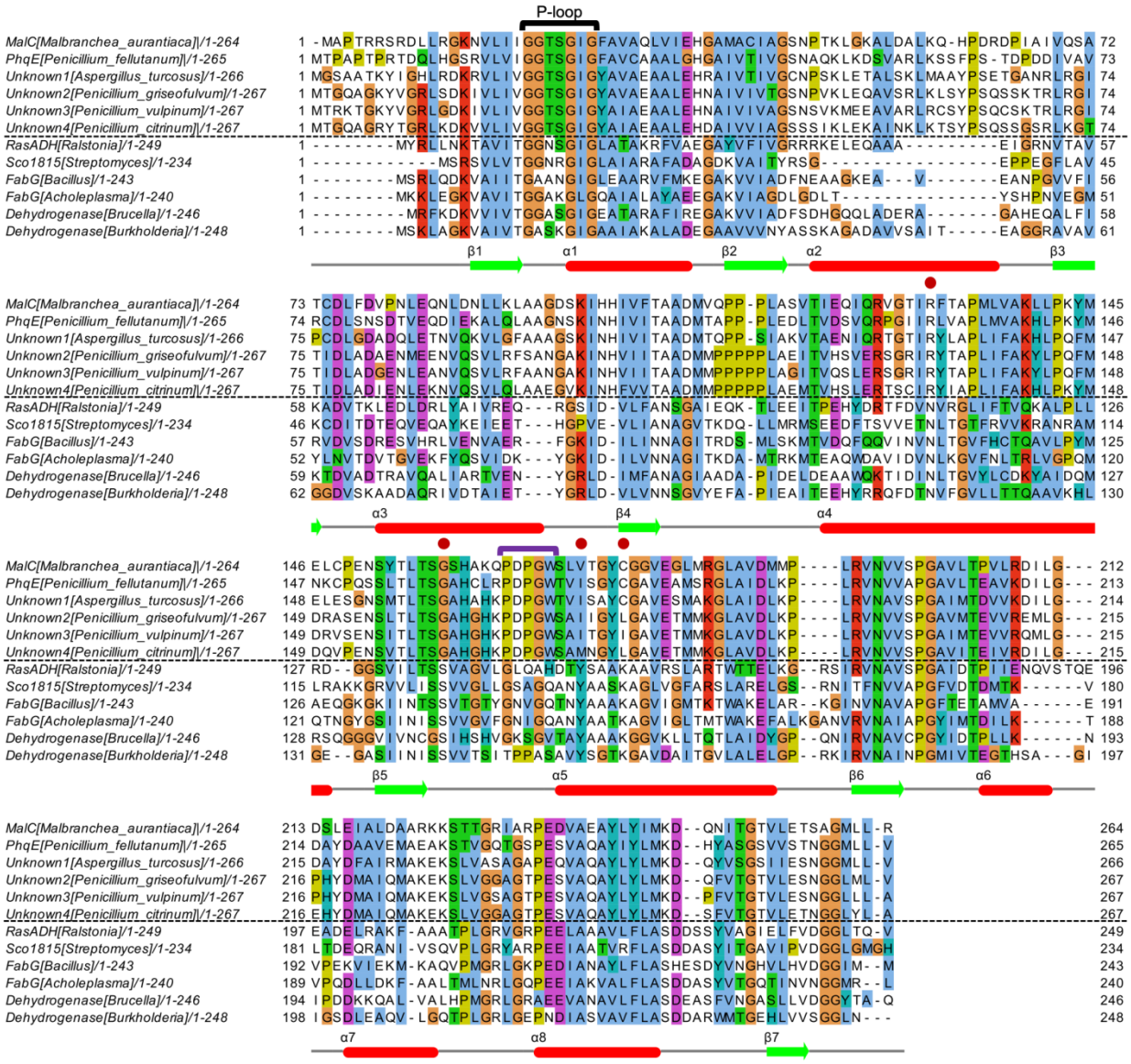

**Supplementary Figure 22.** a. Plot showing the distance between C5 of **11** and C4 of NADPH during 1200 ns simulation; the average distance is 4.4 Å. b. Snapshot depicting the atoms. c. Plot showing the distance between the ribose hydroxyl and O14 (blue) and the distance between the Arg131 and O14 (red). Arg131 is expected to be protonated at physiological pH, based on the pKa prediction (PROPKA) of 11.10. Arg131 also interacts with Asp109 during the simulation allowing access to bulk solvent. d. Snapshot from the beginning of the simulation.



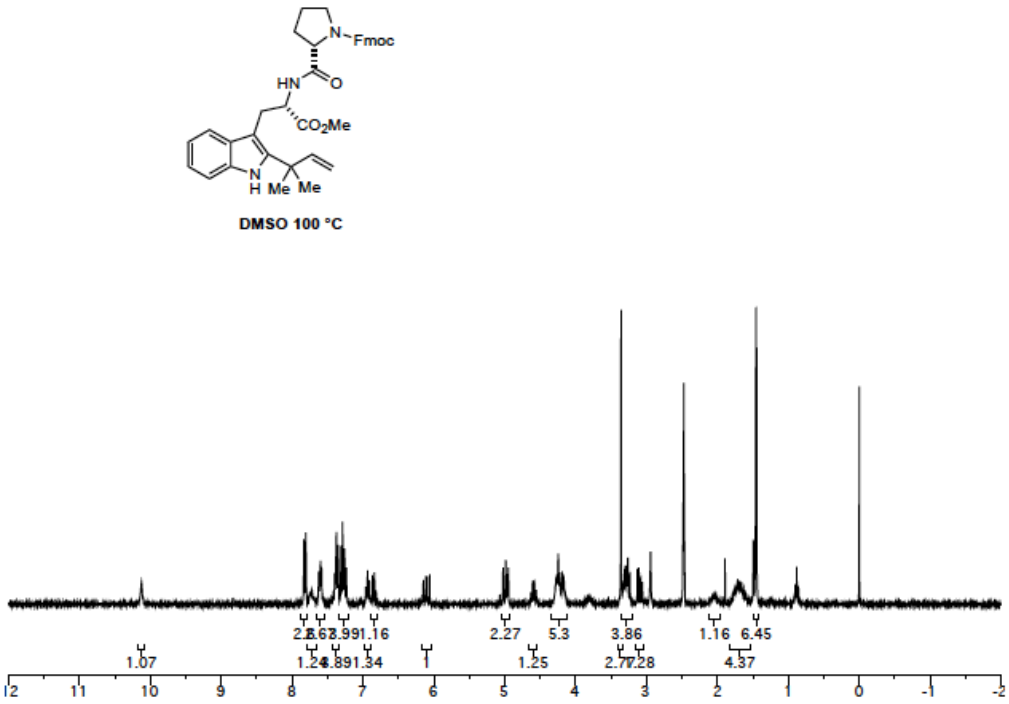

**Supplementary Figure 23.** a. Dihedral angles monitored to track facial selectivity in [4+2] cycloaddition reaction. b. 1200 ns MD simulation shows similar dihedrals throughout the simulation despite the fact that the substrate is unconstrained. c. $NADP^+$ cofactor is required in the simulation to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



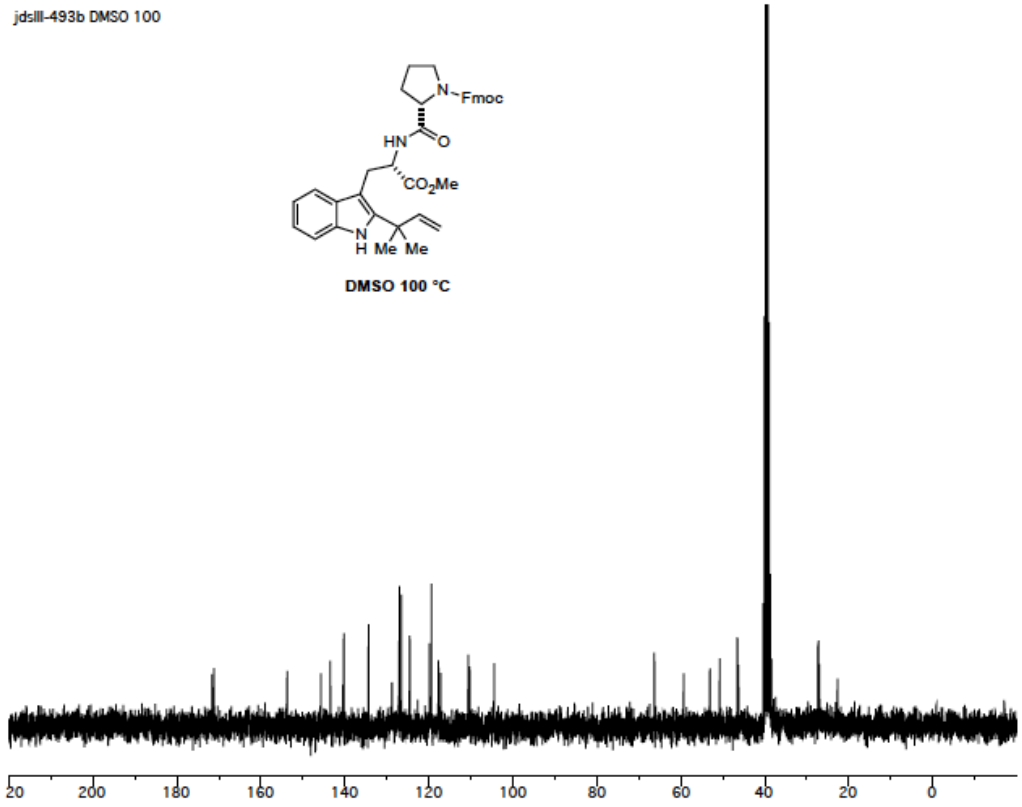

**Supplementary Figure 24.** Active site comparison of the PhqE product complex at left (yellow, product colored in white) and RasADH[42] at right (PDB ID: 4BMS, cyan). For each enzyme, the cofactor is colored in grey, amino acids essential to catalysis are labelled in red, hydrogen bonds are shown as grey dashed lines. Given the striking difference in amino acids in the two active sites, the PhqE and RasADH backbones are remarkably similar.



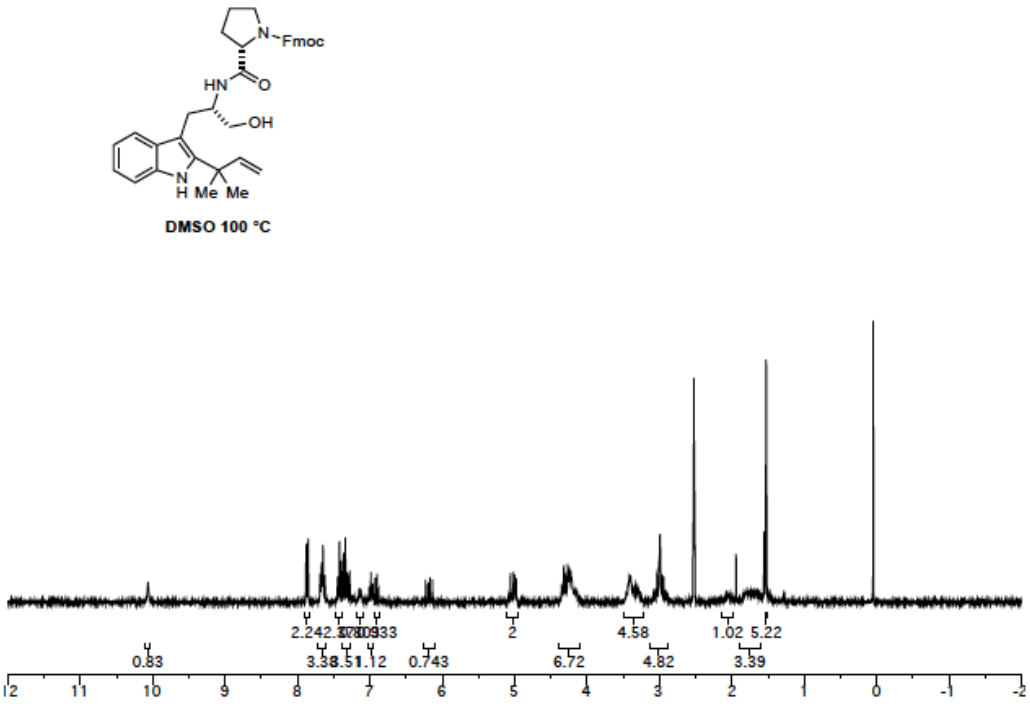

**Supplementary Figure 25.** Multiple sequence alignment of MalC and PhqE homologs. The P-loop critical for cofactor binding is highlighted with a black cap, the "PDPGW" motif is highlighted with a purple cap. SDR amino acids essential for catalysis (Asn-Ser-Tyr-Lys) are shown with red dots, all of which 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 S3[43]), 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.



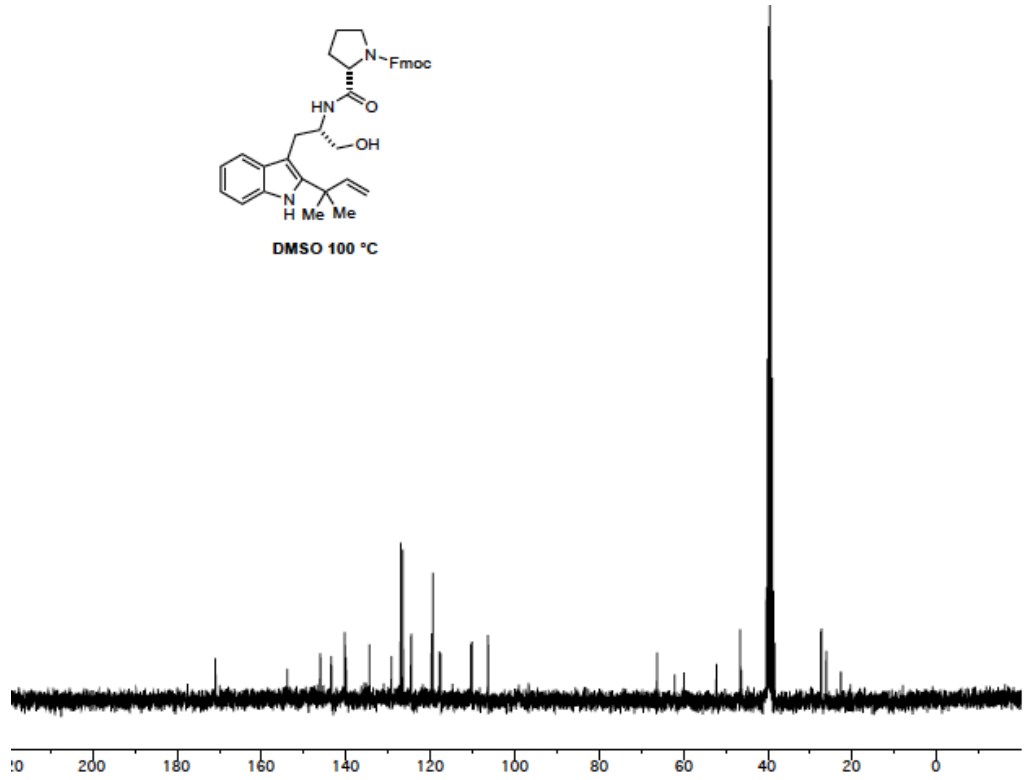

**Supplementary Figure 26.** $^{1}$H NMR spectrum of **15**.



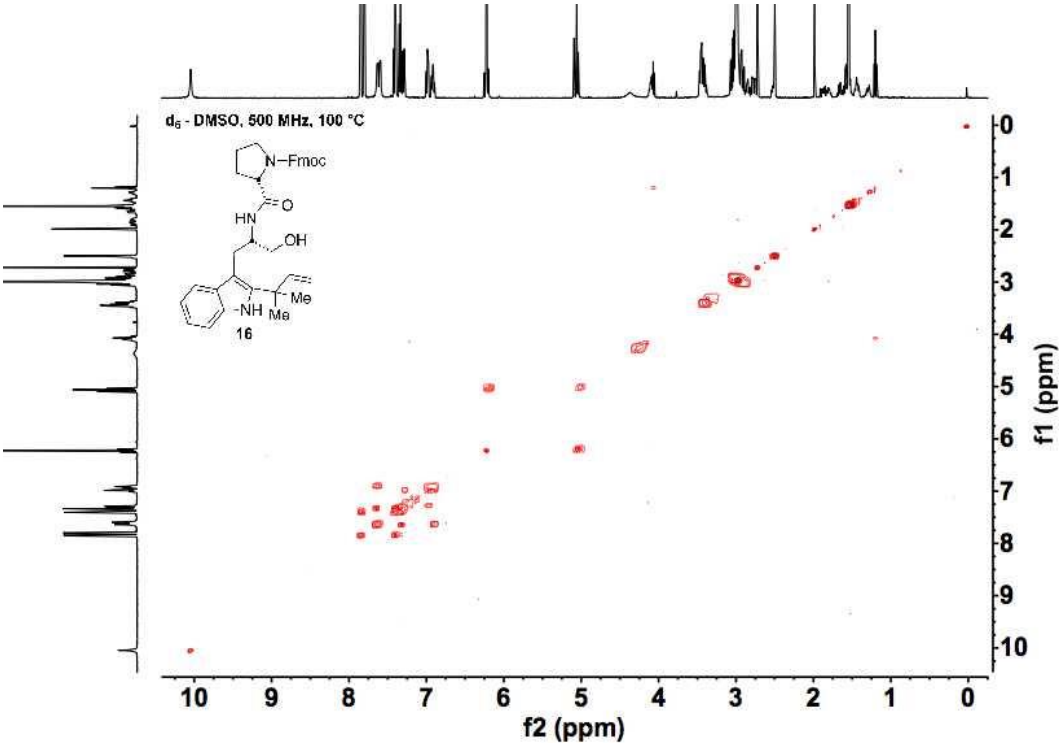

**Supplementary Figure 27.** $^{13}$C NMR spectrum of **15**.



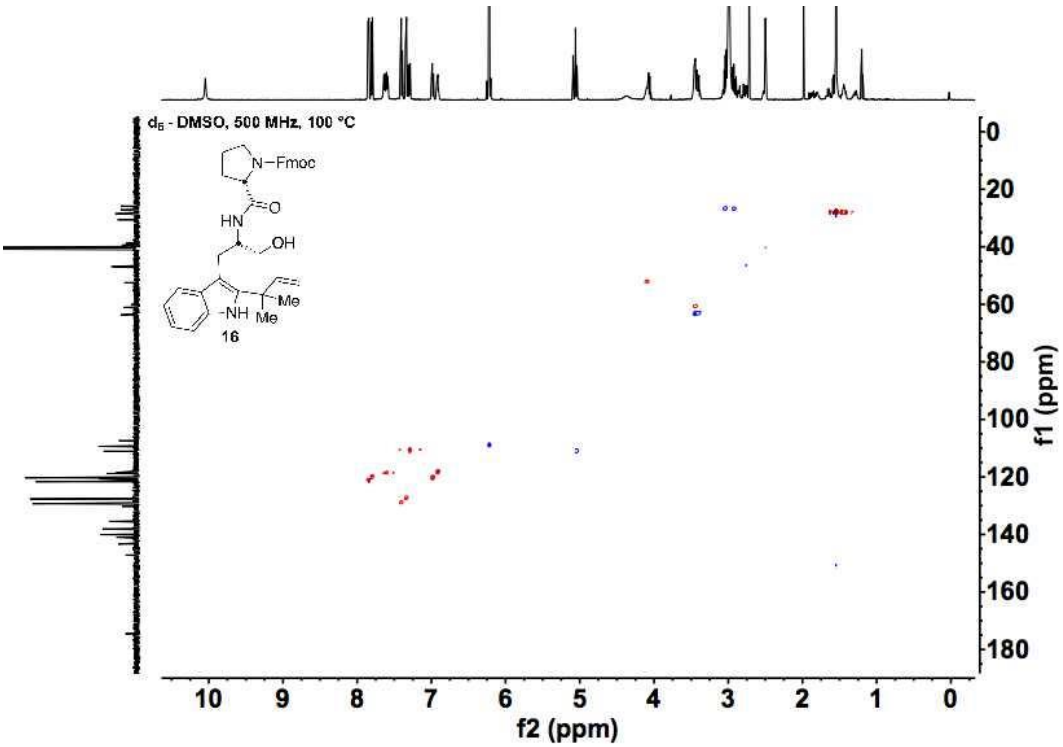

**Supplementary Figure 28.** $^{1}$H NMR spectrum of **16**.



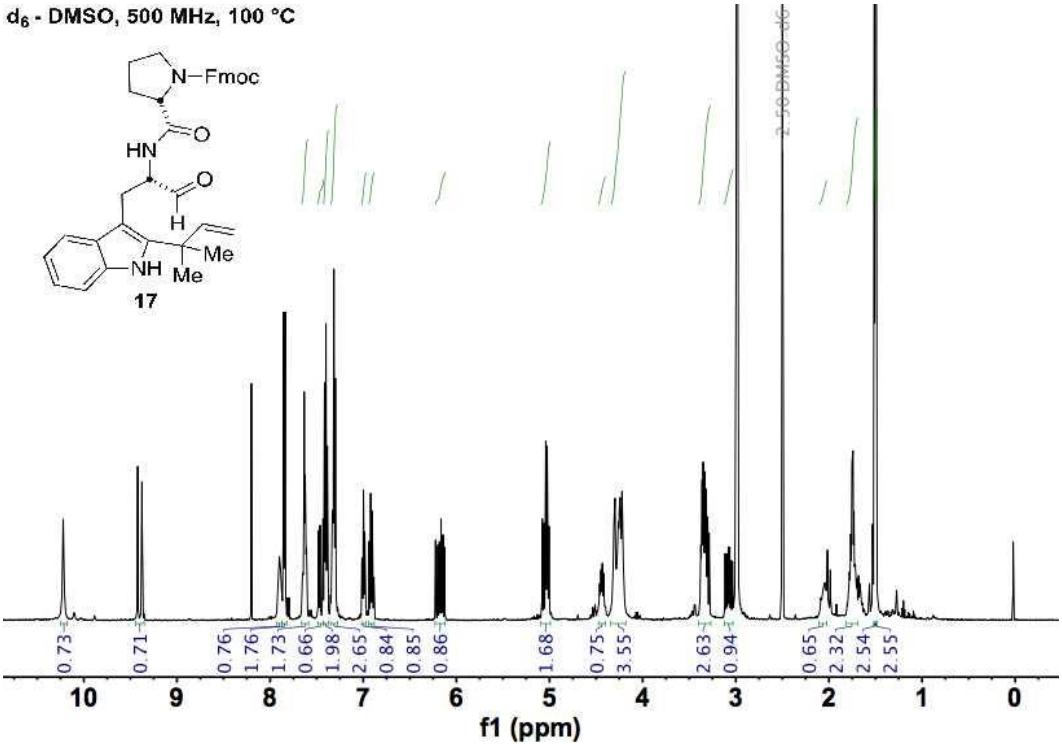

**Supplementary Figure 29.** $^{13}$C NMR spectrum of **16**.



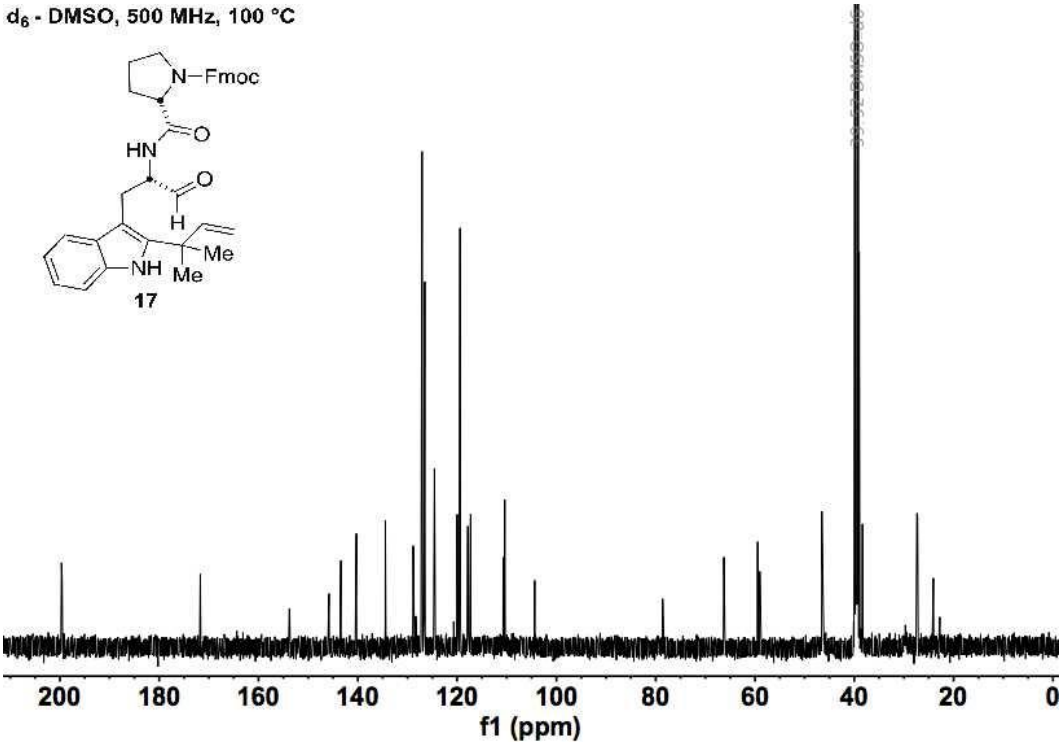

**Supplementary Figure 30.** $^{1}$H-$^{1}$H COSY spectrum of **16**.



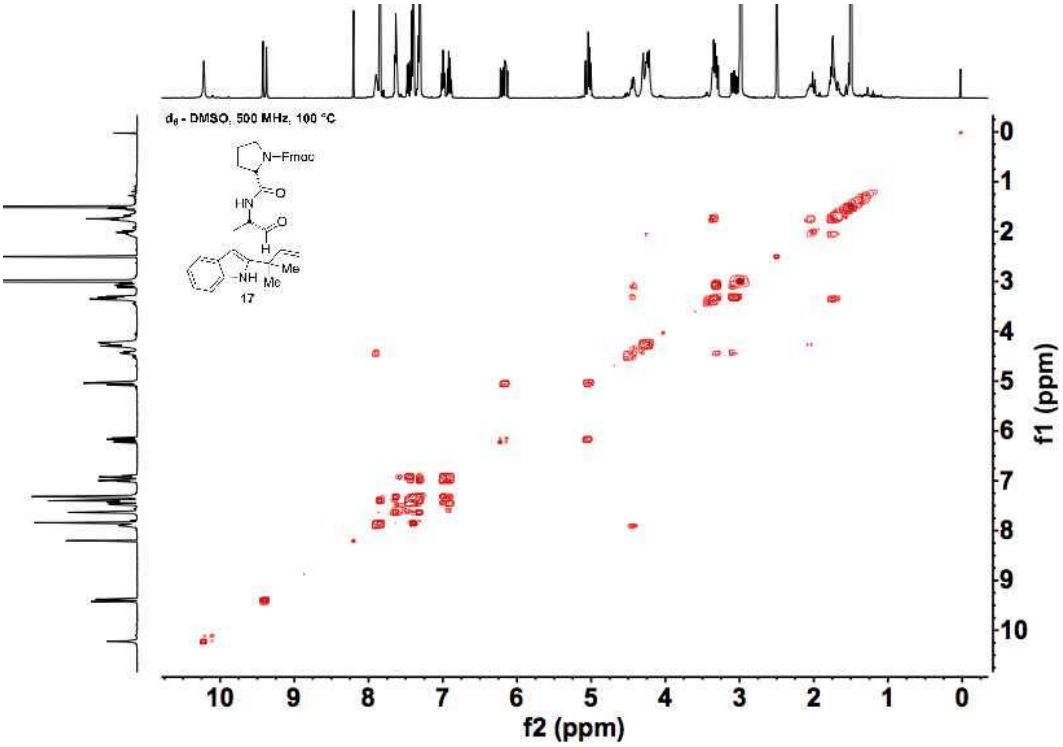

**Supplementary Figure 31.** $^{1}$H-$^{13}$C HSQC spectrum of **16**.



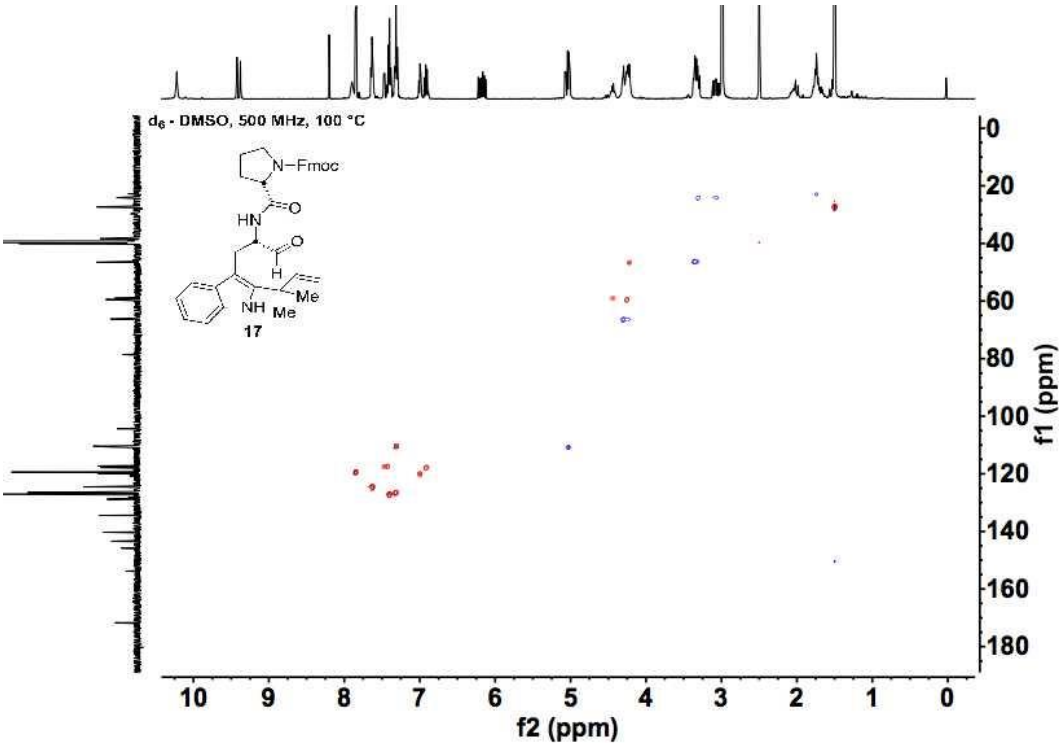

**Supplementary Figure 32.** $^{1}$H NMR spectrum of **17**.



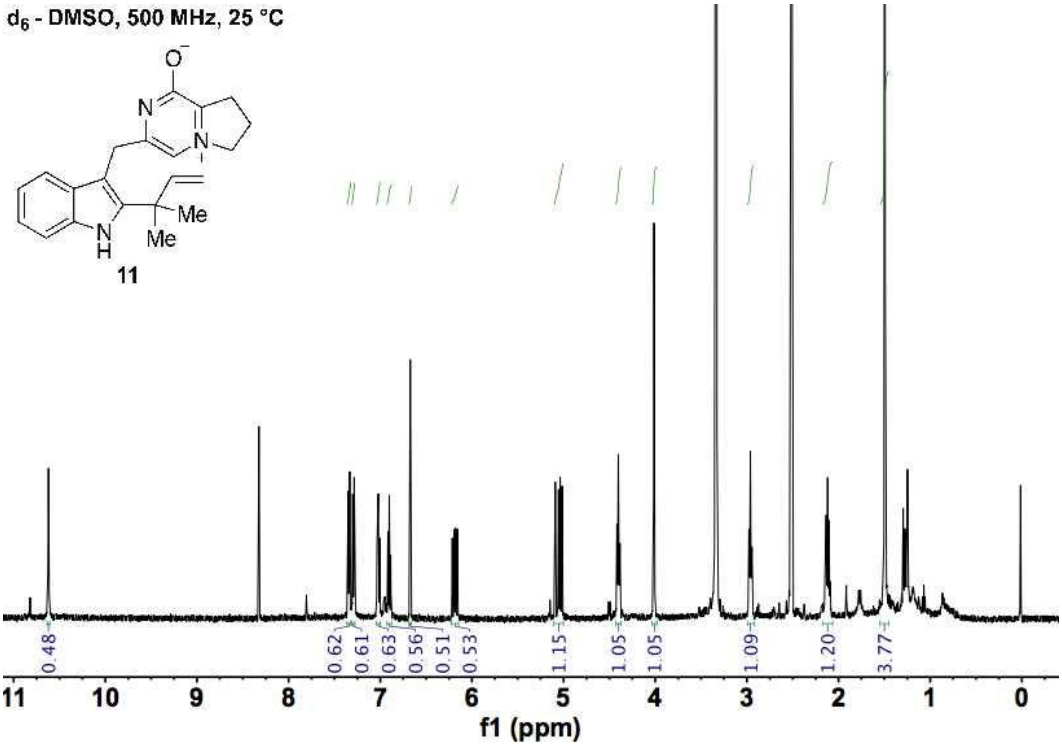

**Supplementary Figure 33.** $^{13}$C NMR spectrum of **17**.



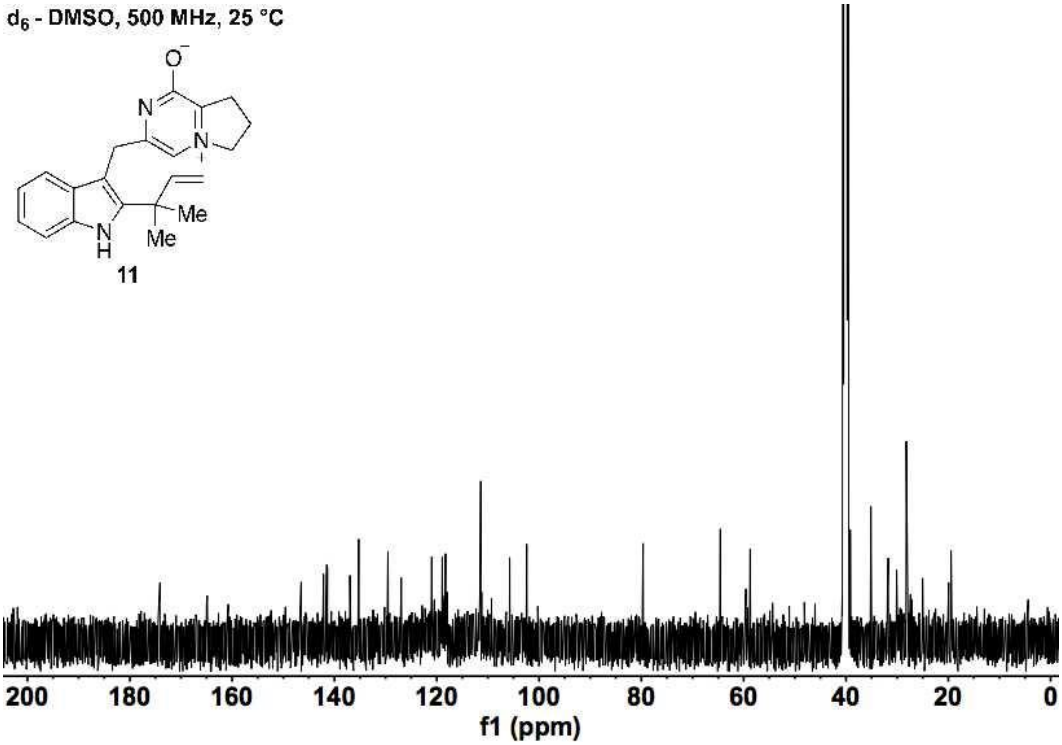

**Supplementary Figure 34.** $^{1}$H-$^{1}$H COSY spectrum of **17**.



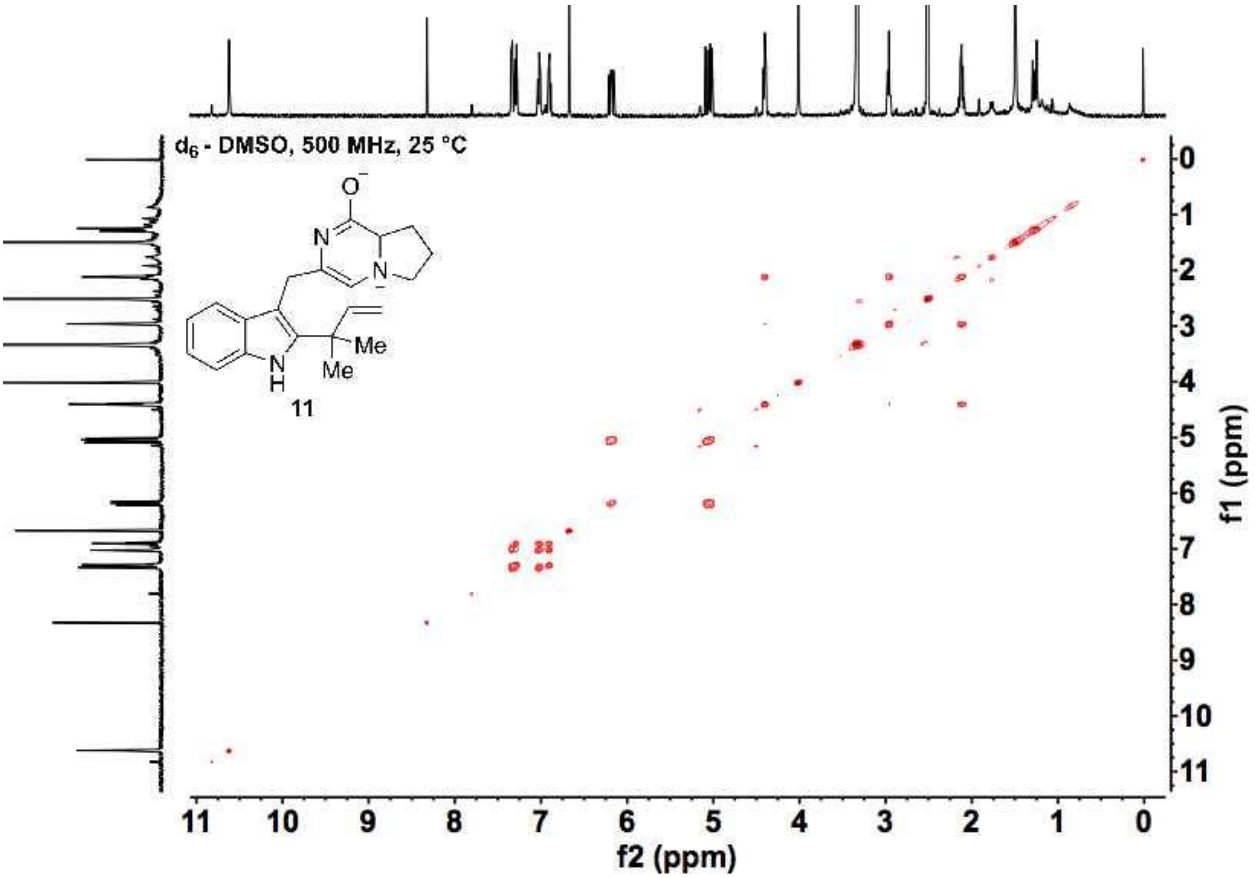

**Supplementary Figure 35.** $^{1}$H-$^{13}$C HSQC spectrum of **17**.



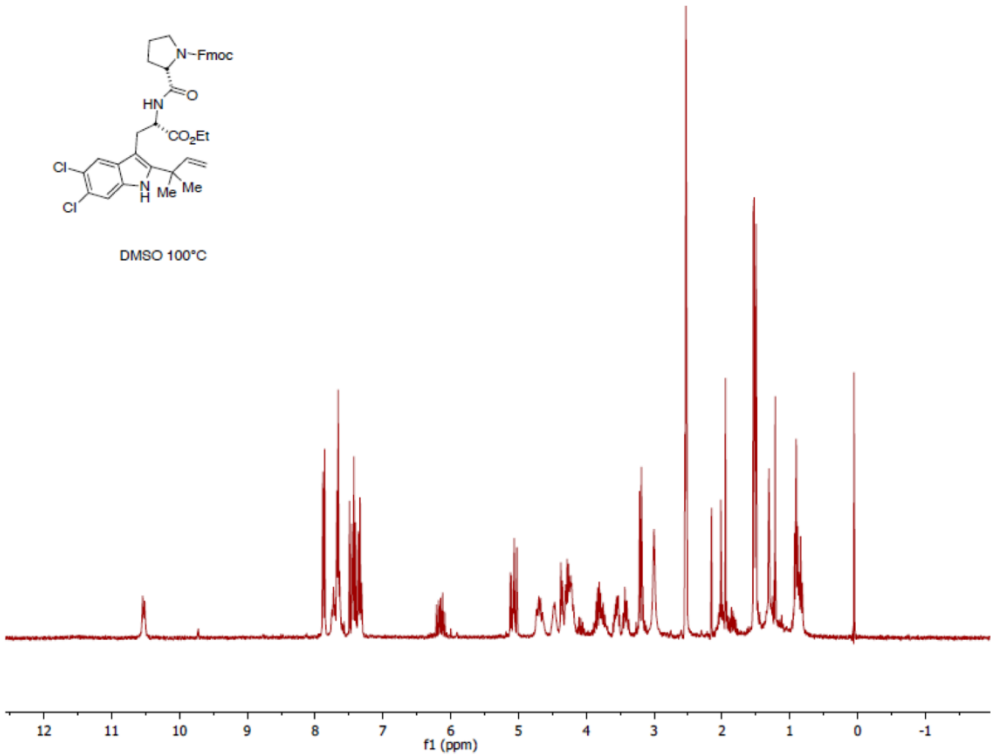

**Supplementary Figure 36.** $^{1}$H NMR spectrum of **11**.



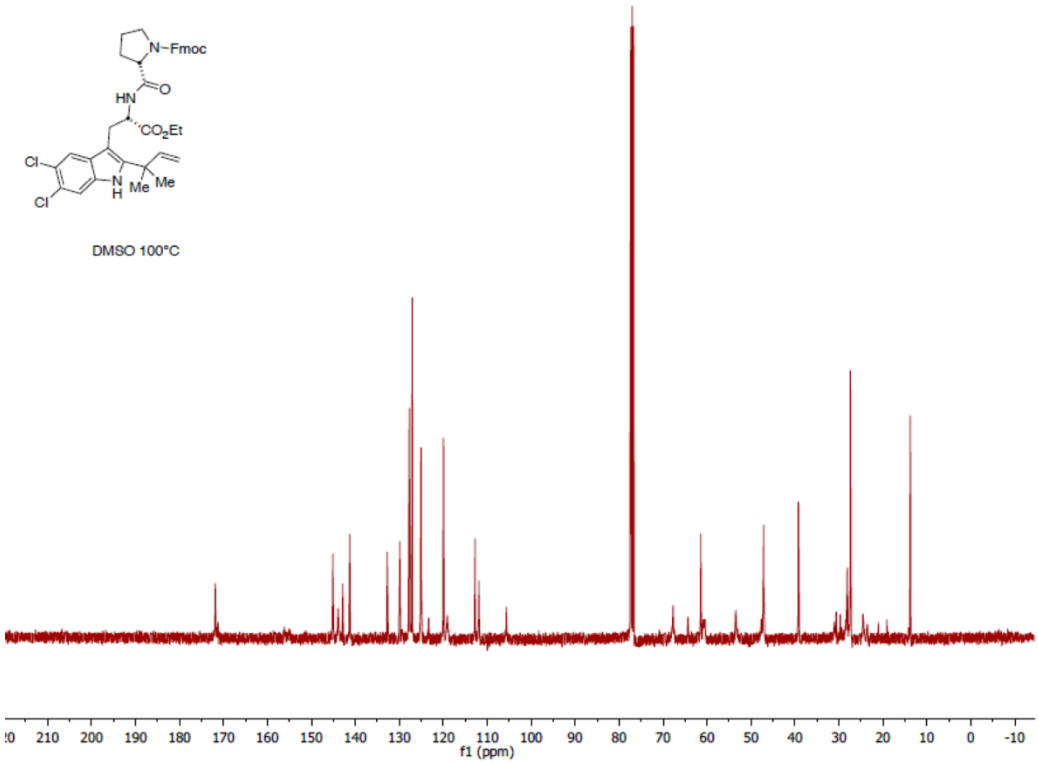

**Supplementary Figure 37.** $^{13}$C NMR spectrum of **11**.



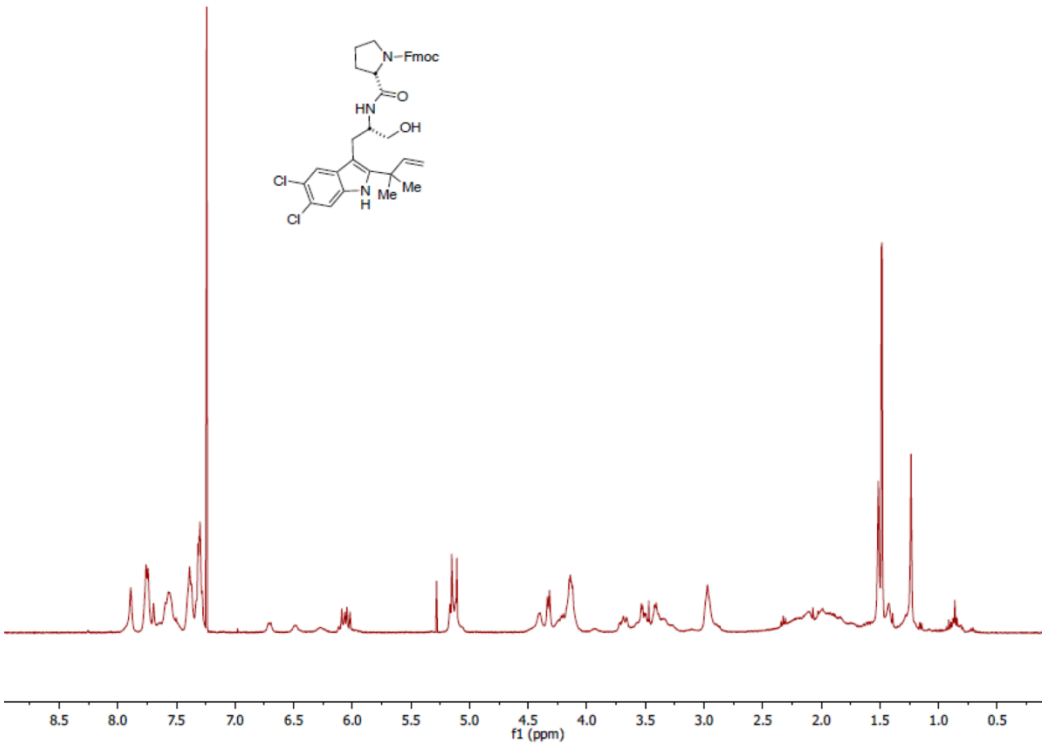

**Supplementary Figure 38.** $^{1}$H-$^{1}$H COSY spectrum of **11**.



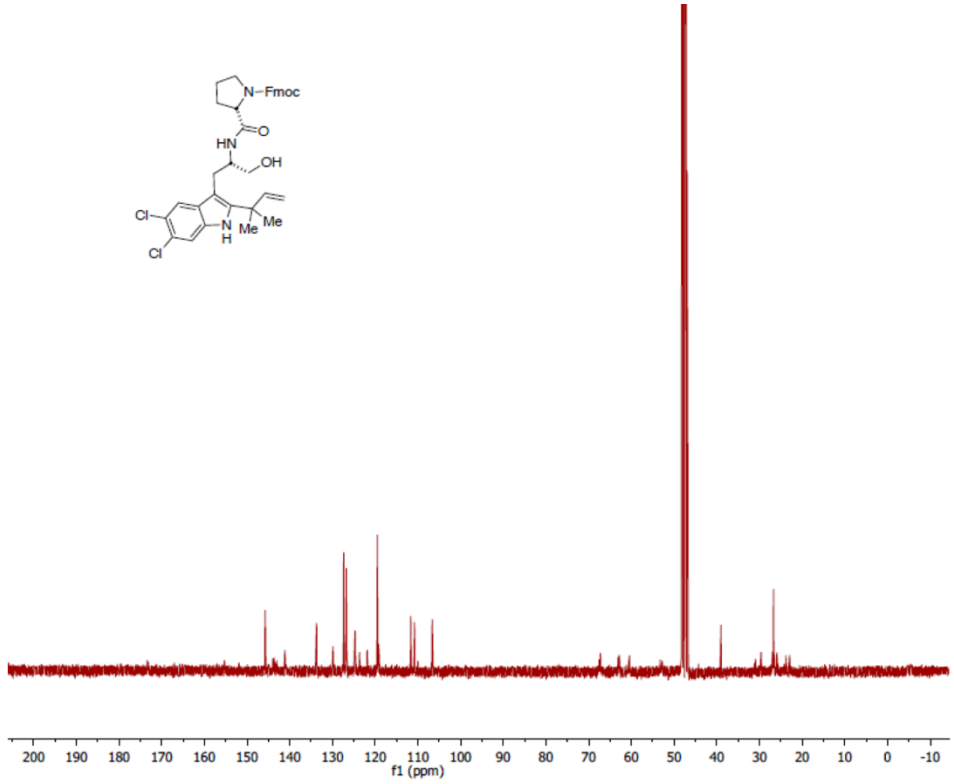

**Supplementary Figure 39.** $^{1}$H NMR spectrum of **19**.



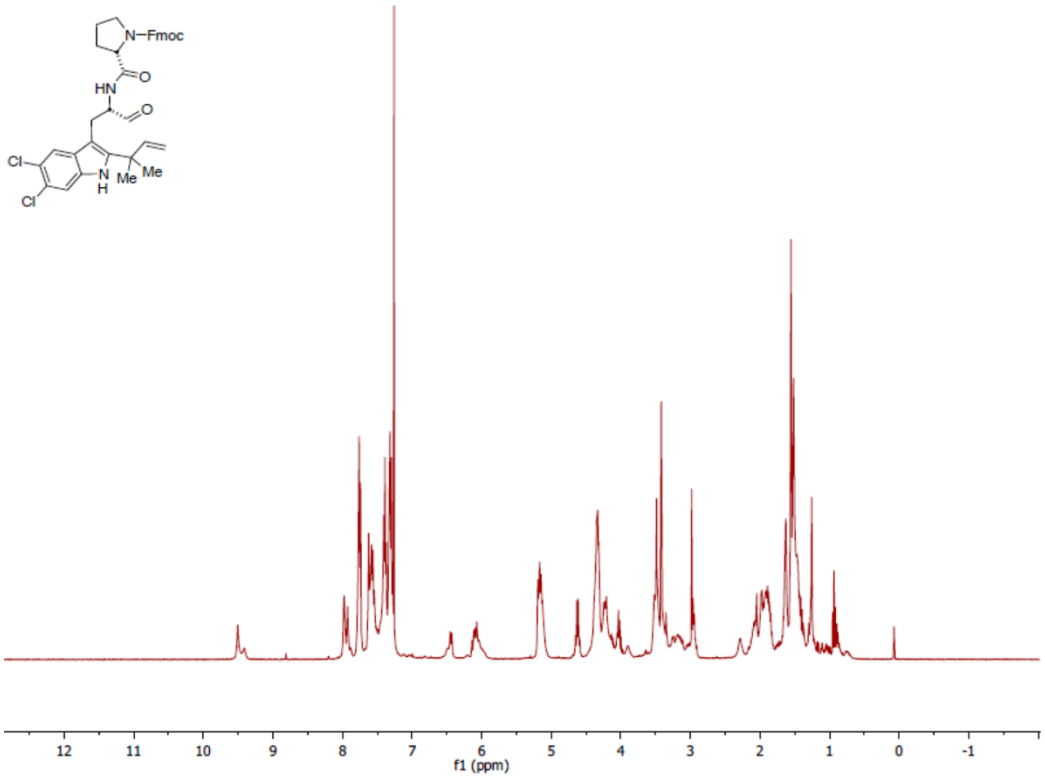

**Supplementary Figure 40.** $^{13}$C NMR spectrum of **19**.



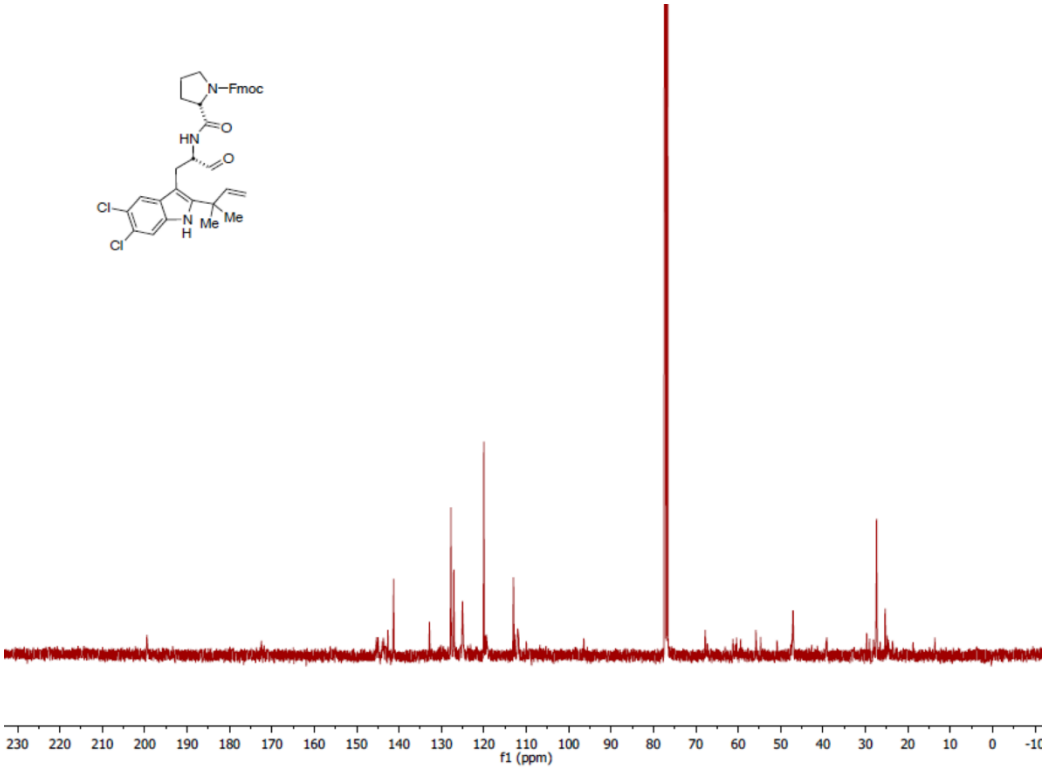

**Supplementary Figure 41.** $^{1}$H NMR spectrum of **20**.



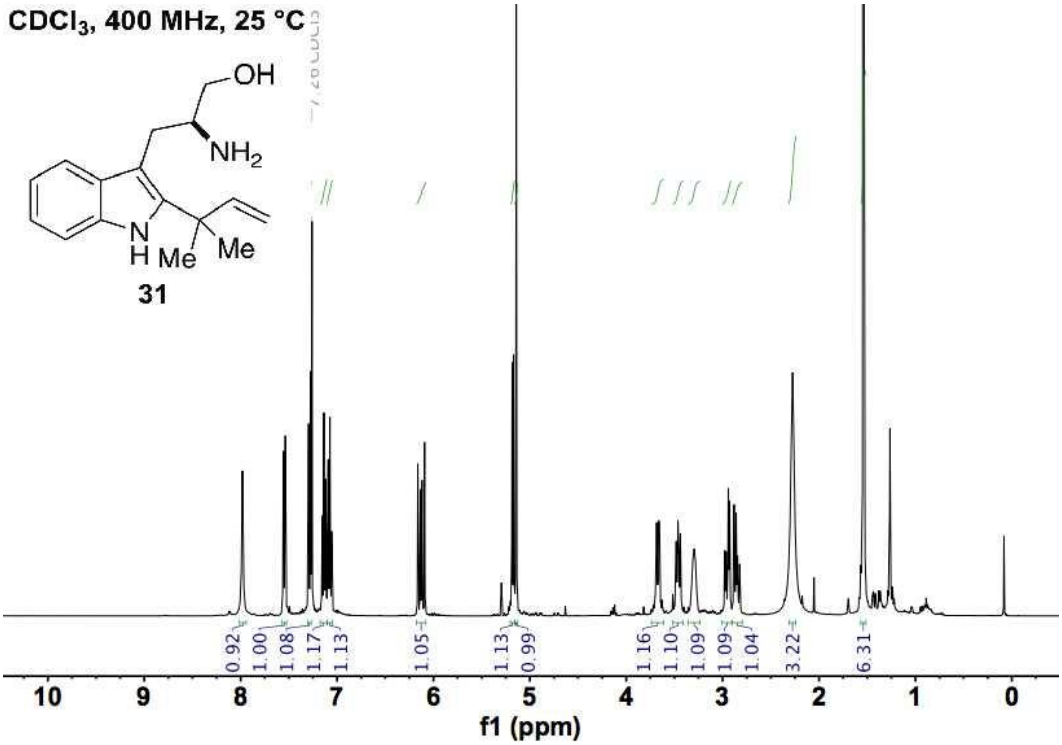

**Supplementary Figure 42.** $^{13}$C NMR spectrum of **20**.



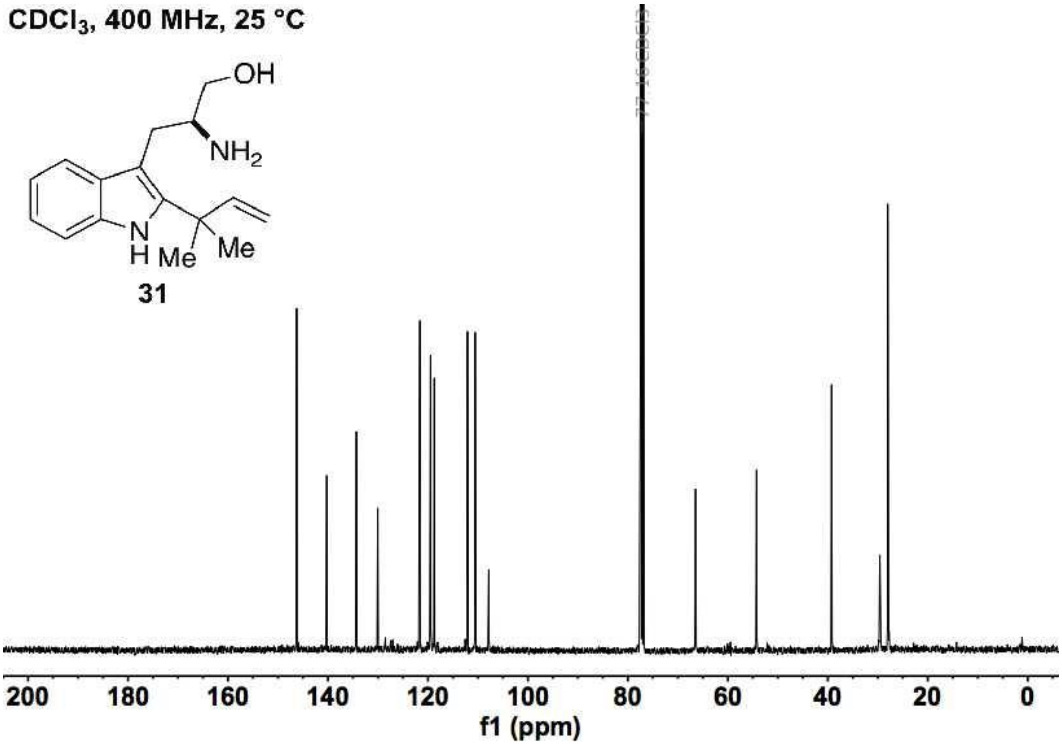

**Supplementary Figure 43.** $^{1}$H NMR spectrum of **21**.



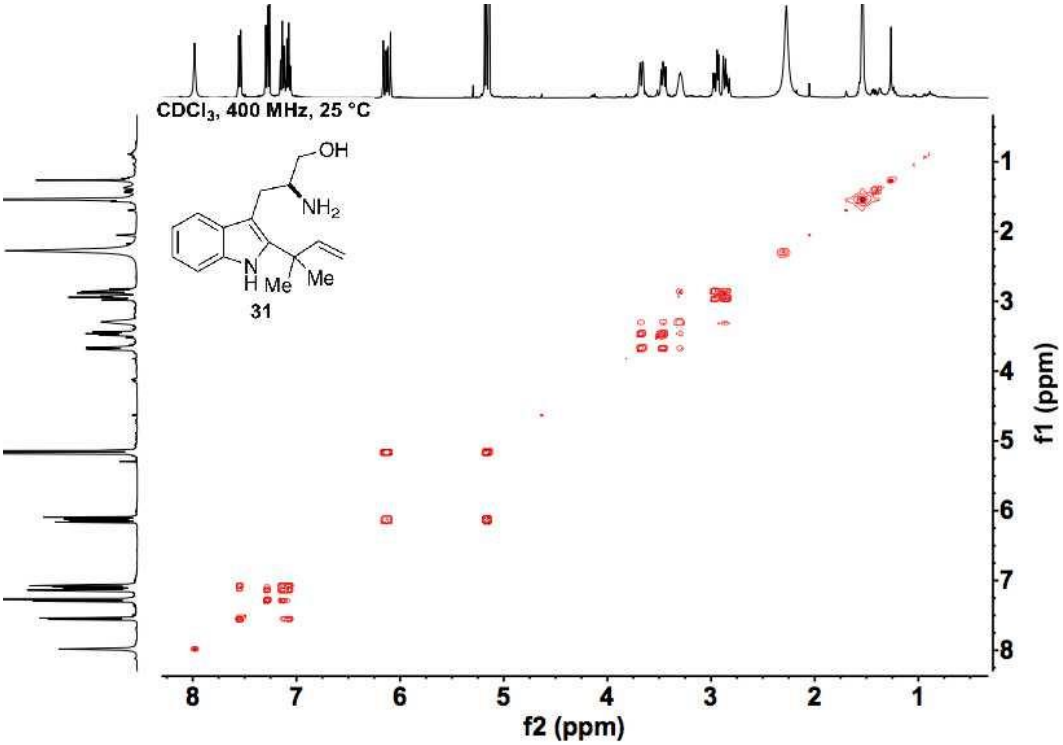

**Supplementary Figure 44.** $^{13}$C NMR spectrum of **21**.



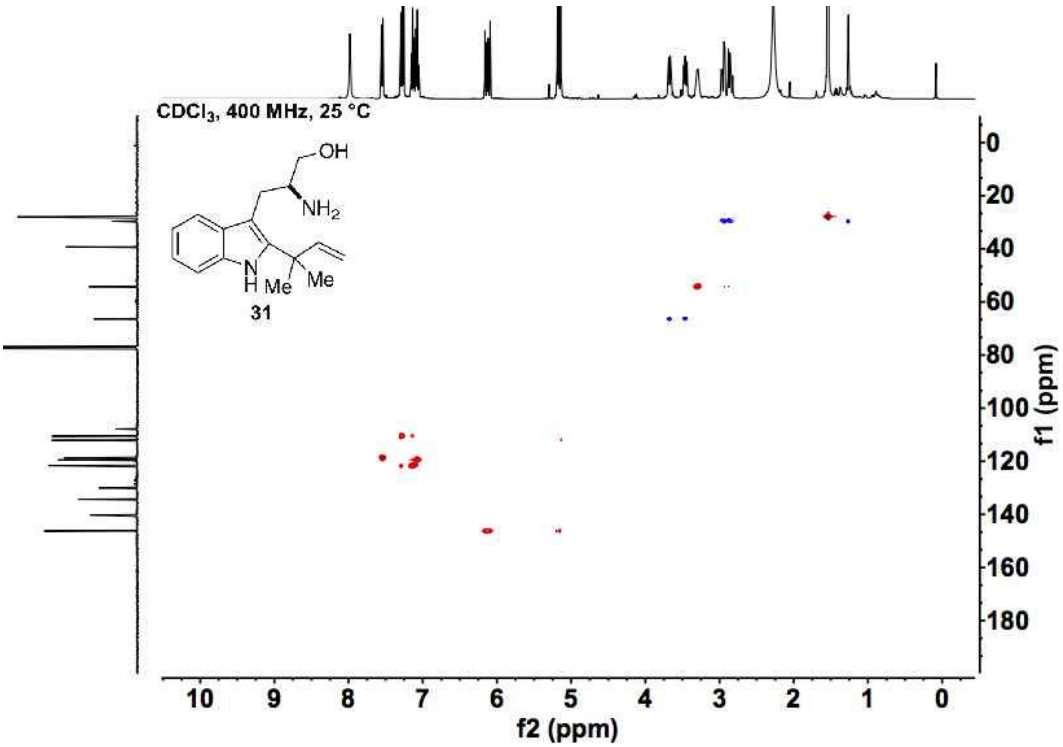

**Supplementary Figure 45.** $^{1}$H NMR spectrum of **31**.



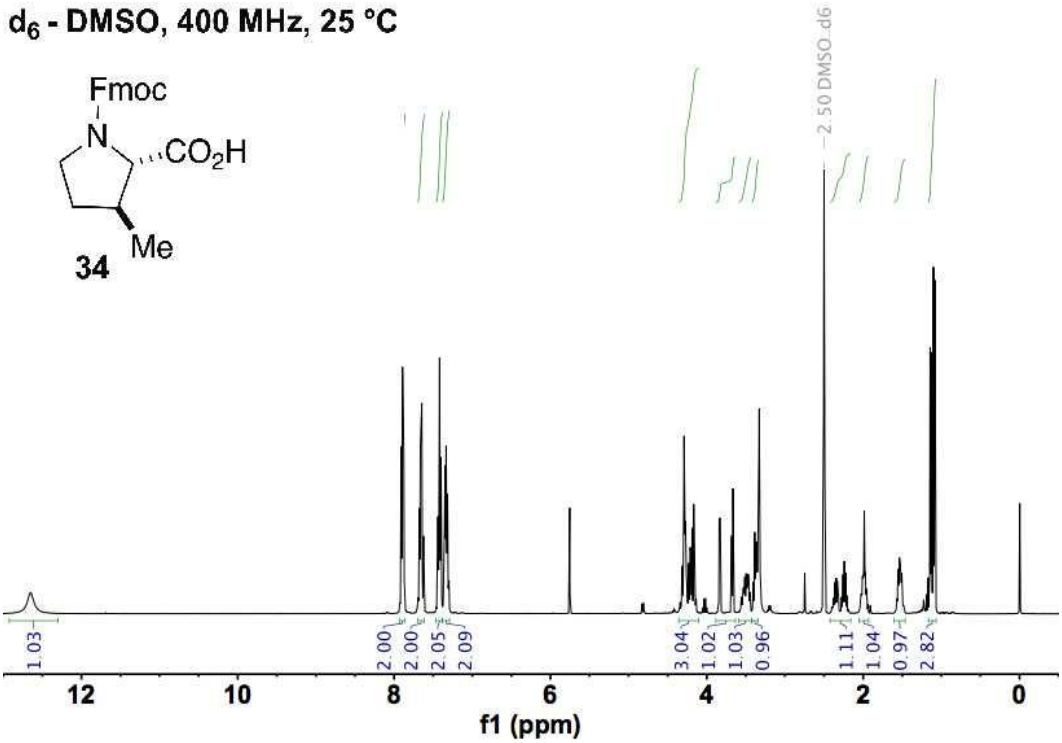

**Supplementary Figure 46.** $^{13}$C NMR spectrum of **31**.



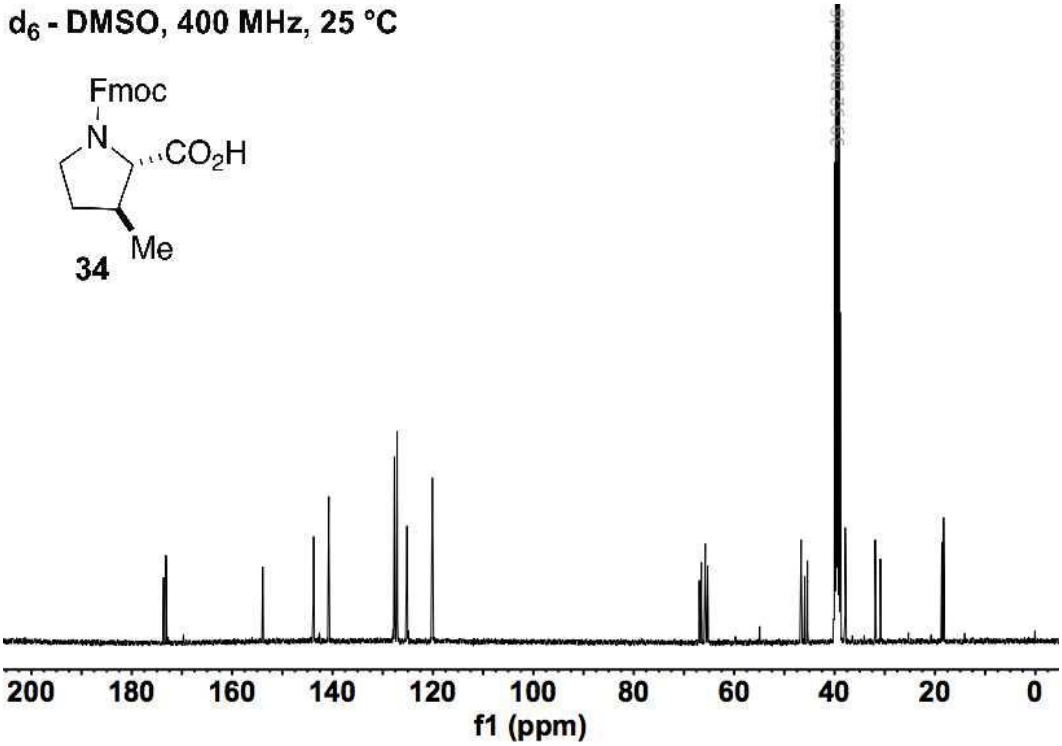

**Supplementary Figure 47.** $^{1}$H-$^{1}$H COSY spectrum of **31**.



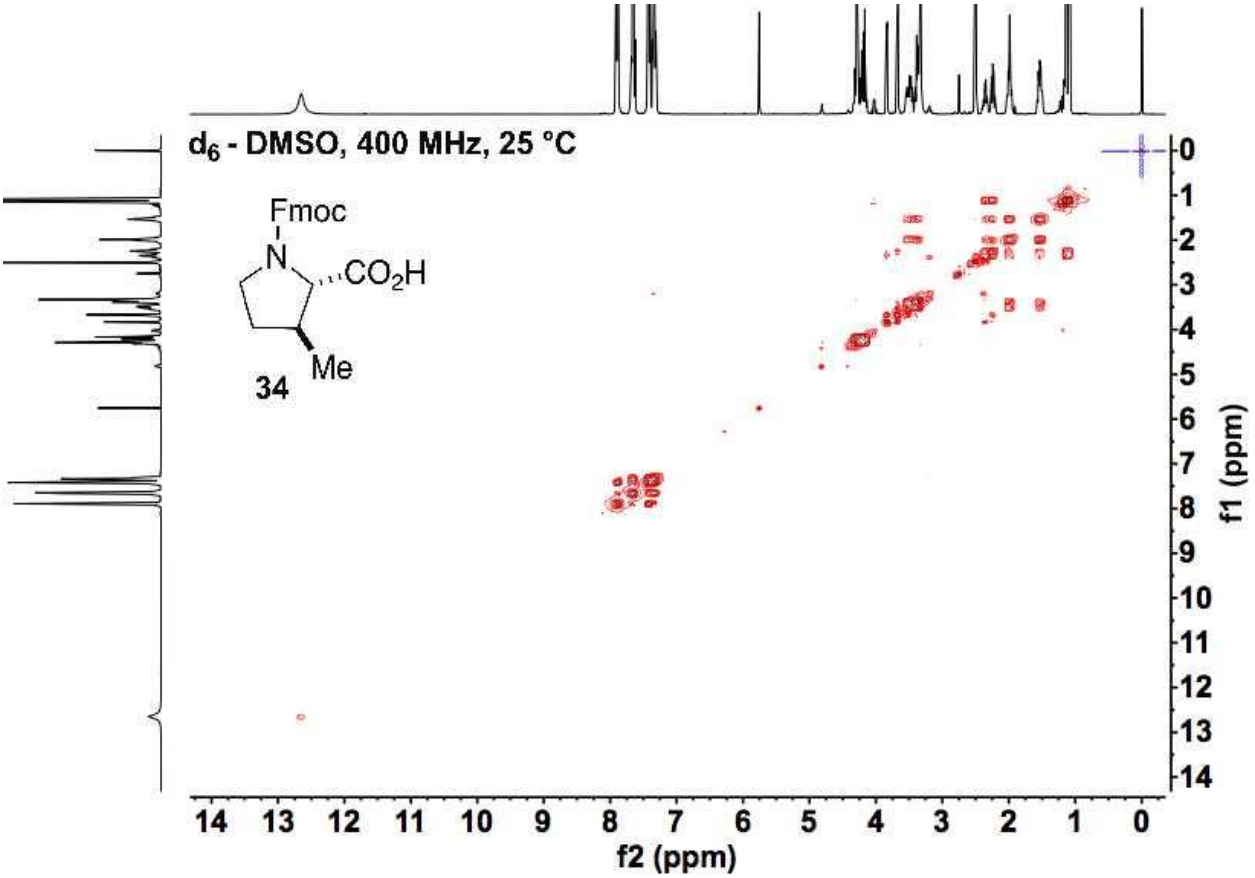

**Supplementary Figure 48.** $^{1}$H-$^{13}$C HSQC spectrum of **31**.



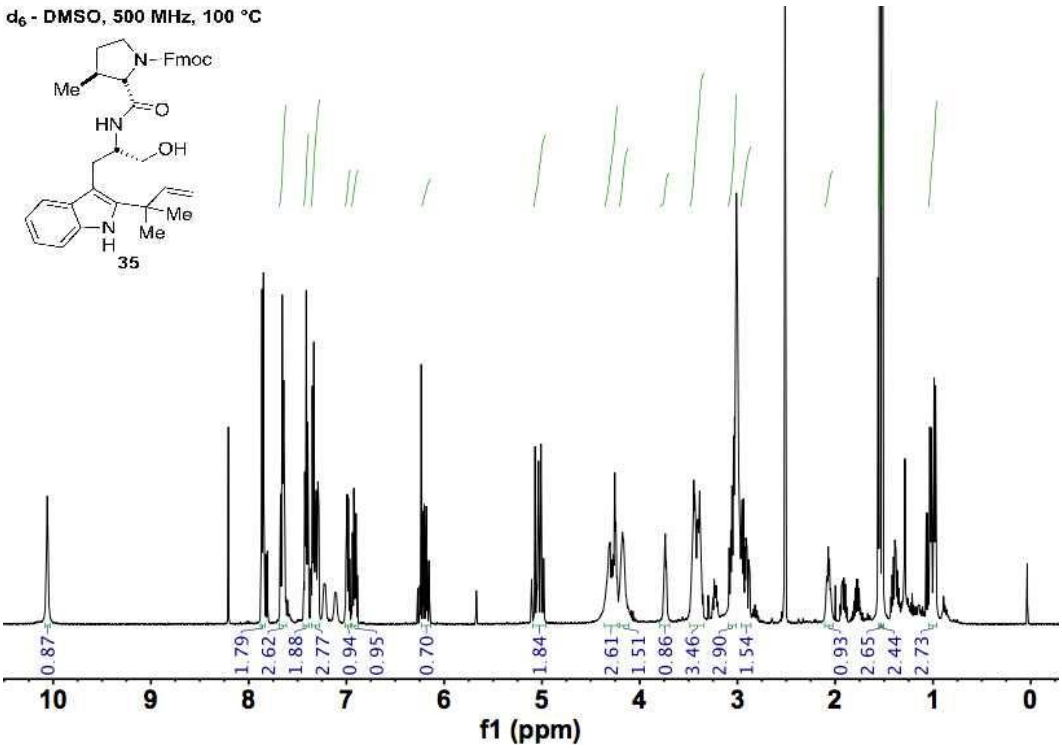

**Supplementary Figure 49.** $^{1}$H NMR spectrum of **34**.



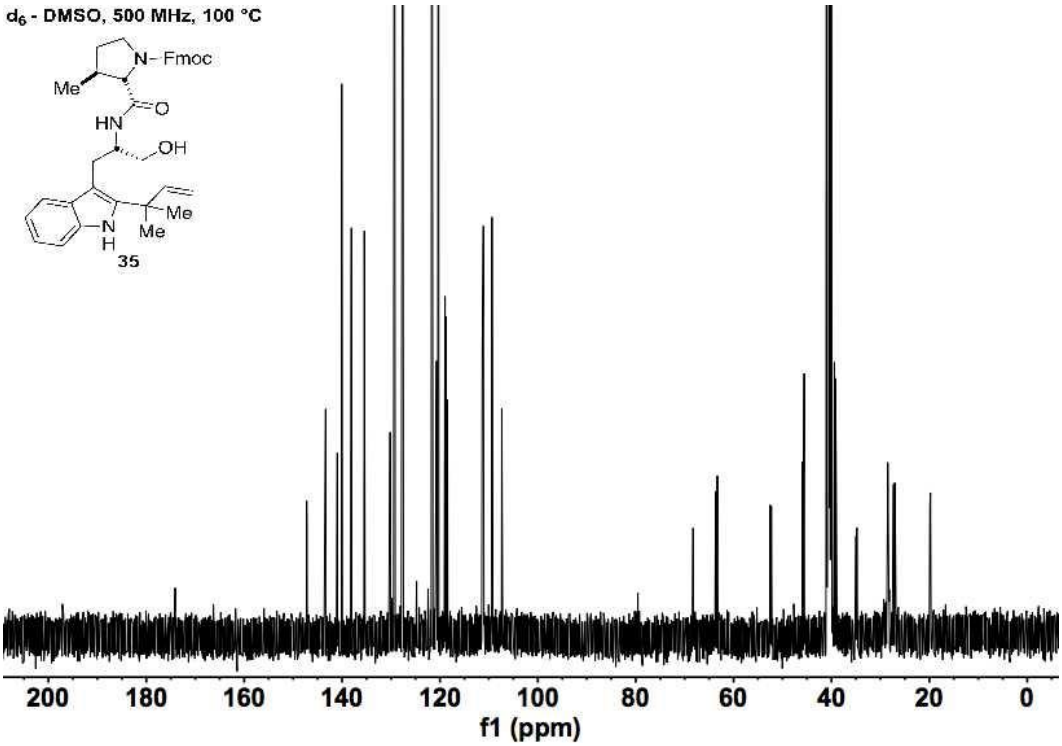

**Supplementary Figure 50.** $^{13}$C NMR spectrum of **34**.



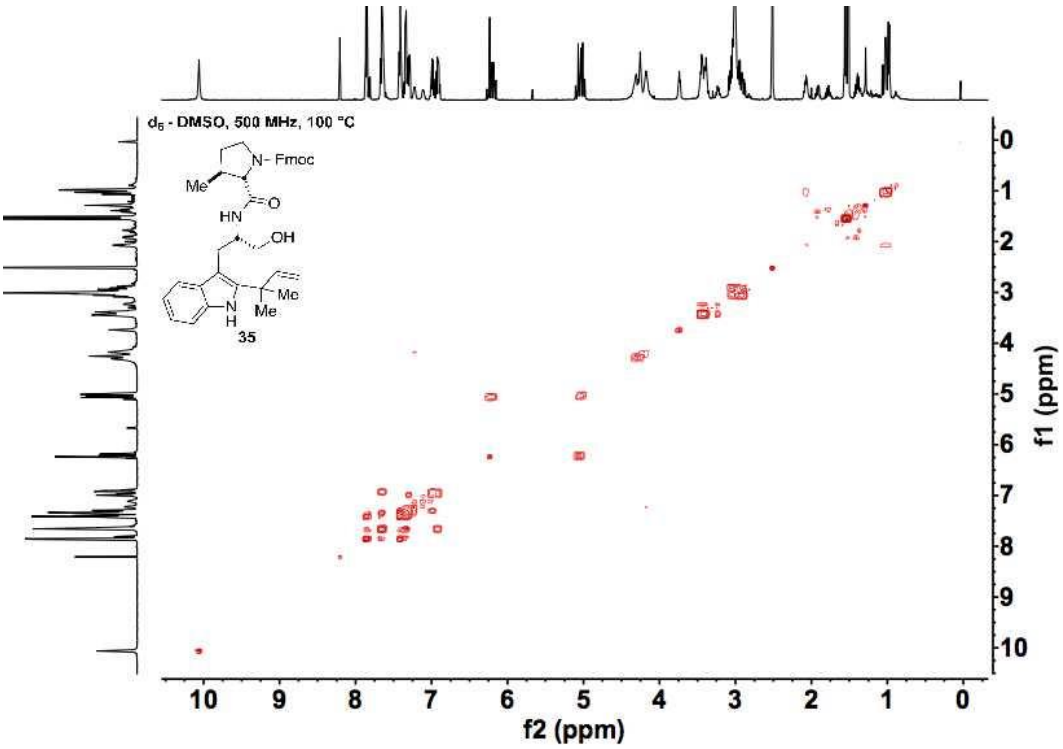

**Supplementary Figure 51.** $^{1}$H-$^{1}$H COSY spectrum of **34**.



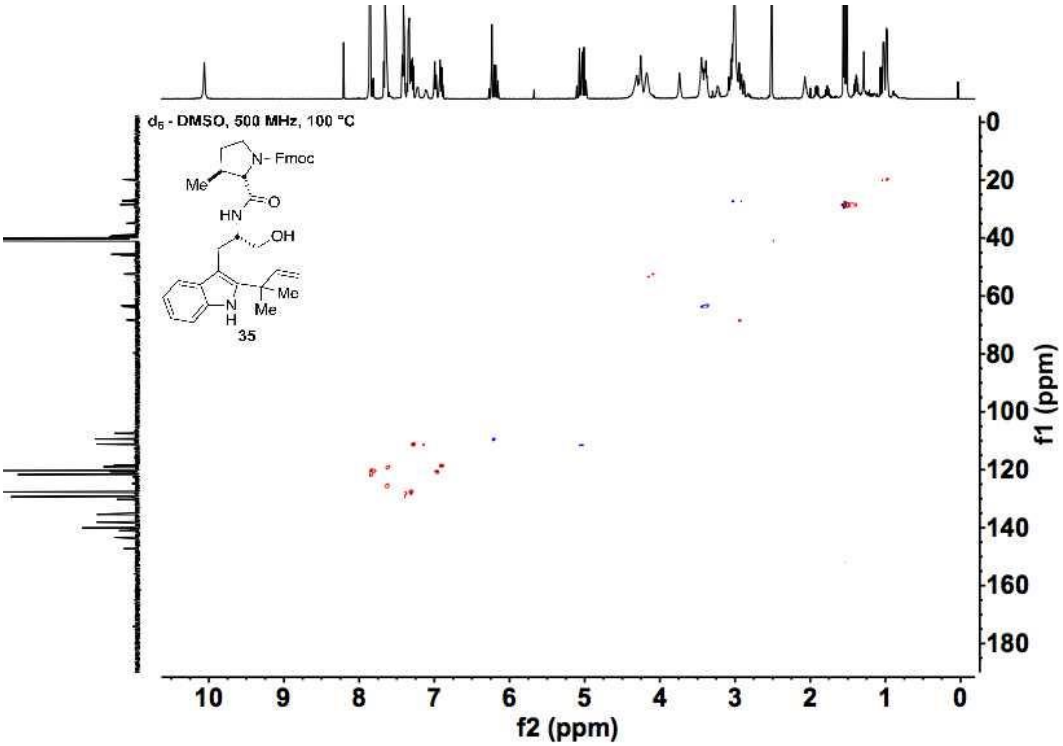

**Supplementary Figure 52.** $^{1}$H NMR spectrum of **35**.



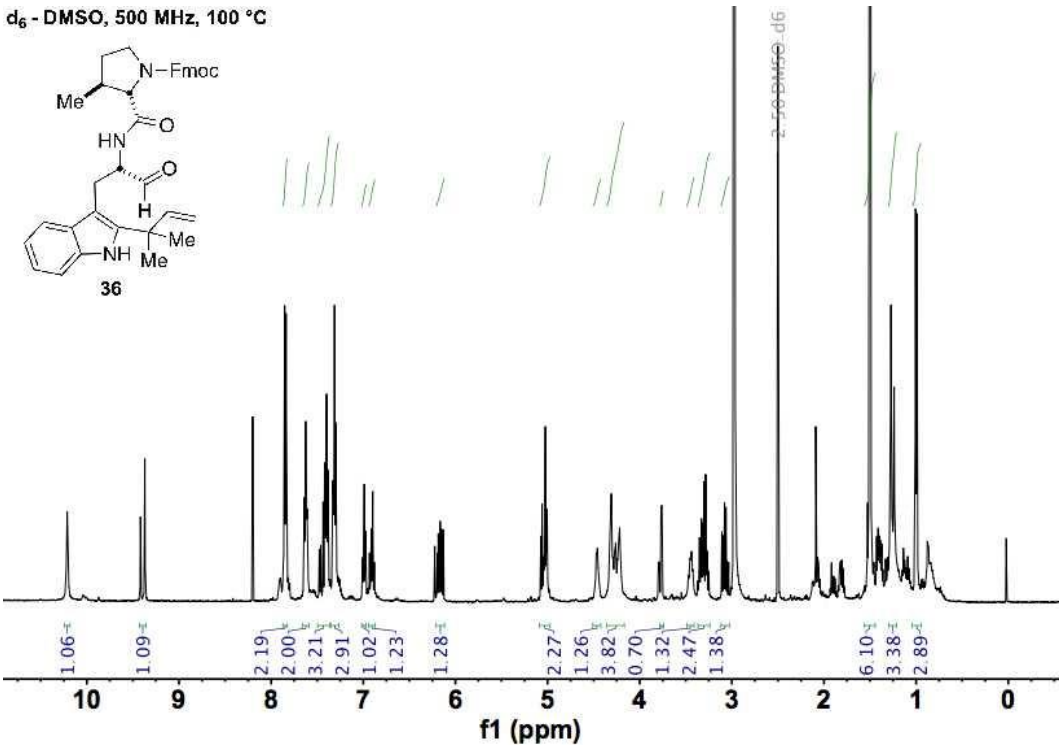

**Supplementary Figure 53.** $^{13}$C NMR spectrum of **35**.



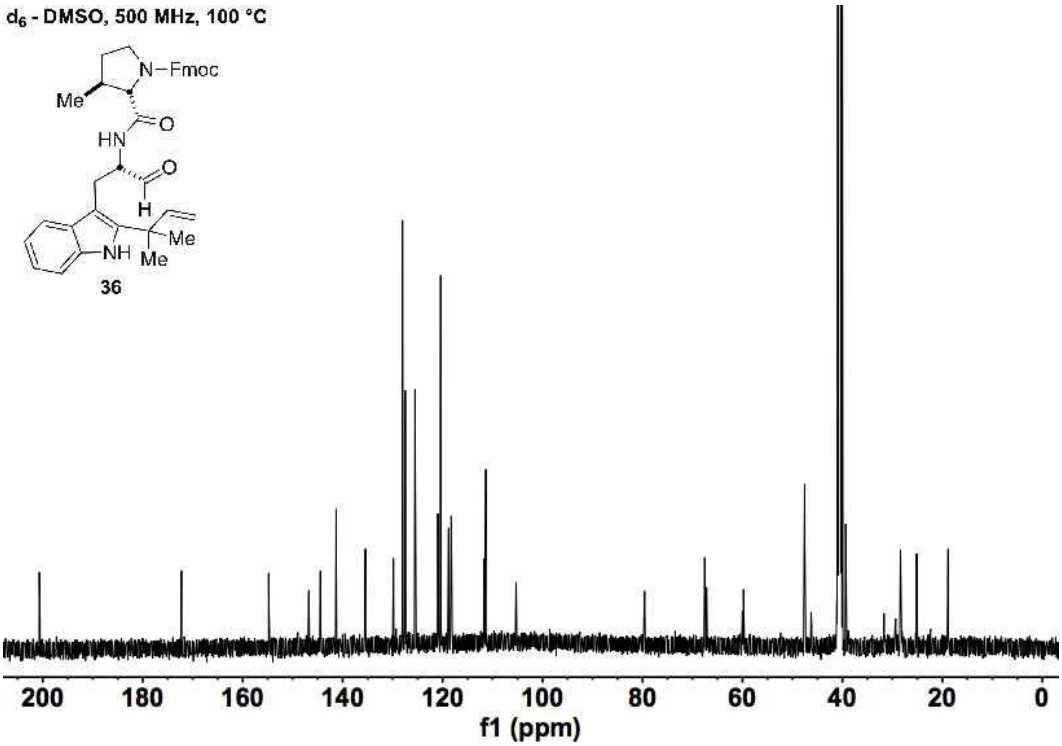

**Supplementary Figure 54.** $^{1}$H-$^{1}$H COSY spectrum of **35**.



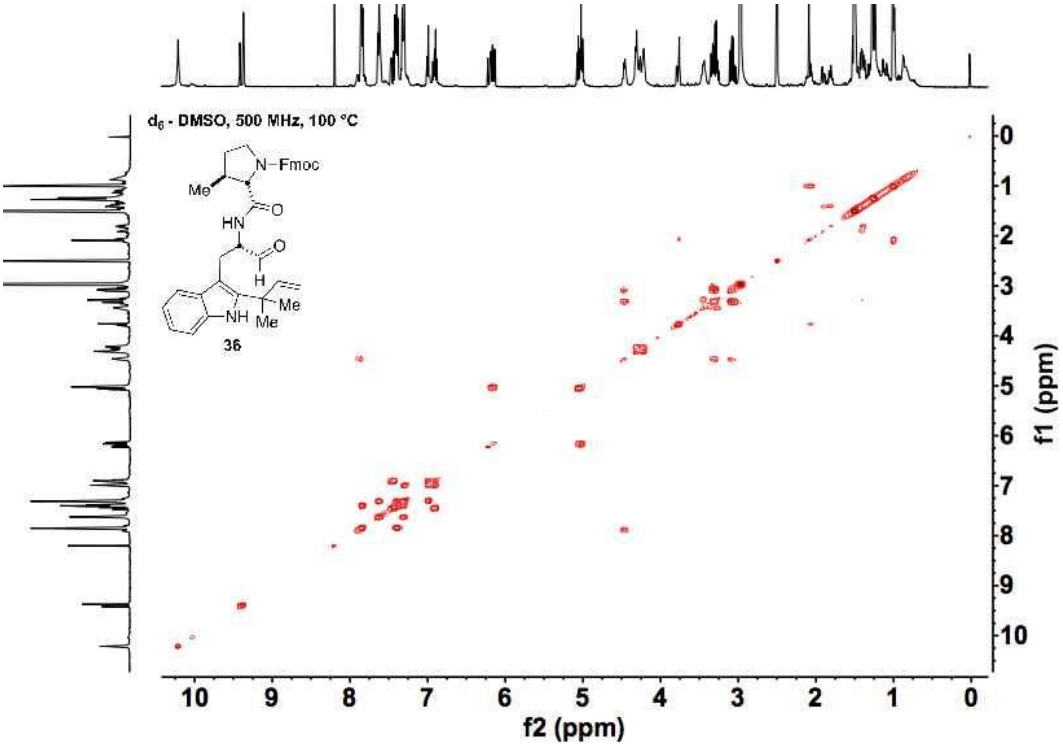

**Supplementary Figure 55.** $^{1}$H-$^{13}$C HSQC spectrum of **35**.



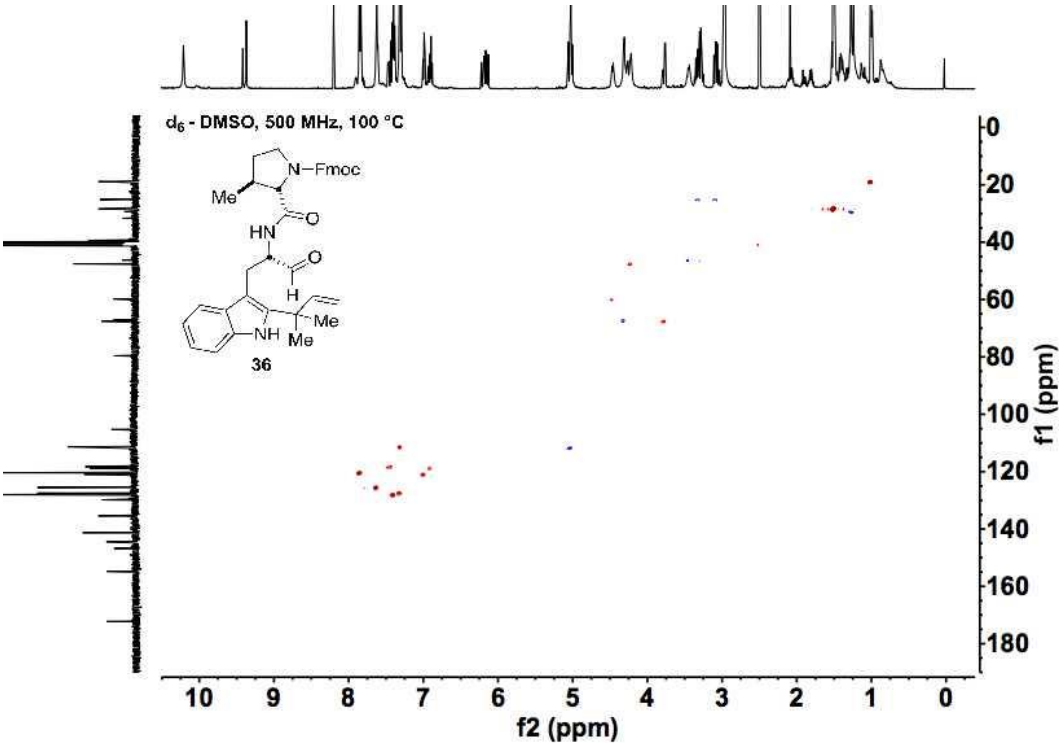

**Supplementary Figure 56.** $^{1}$H NMR spectrum of **36**.



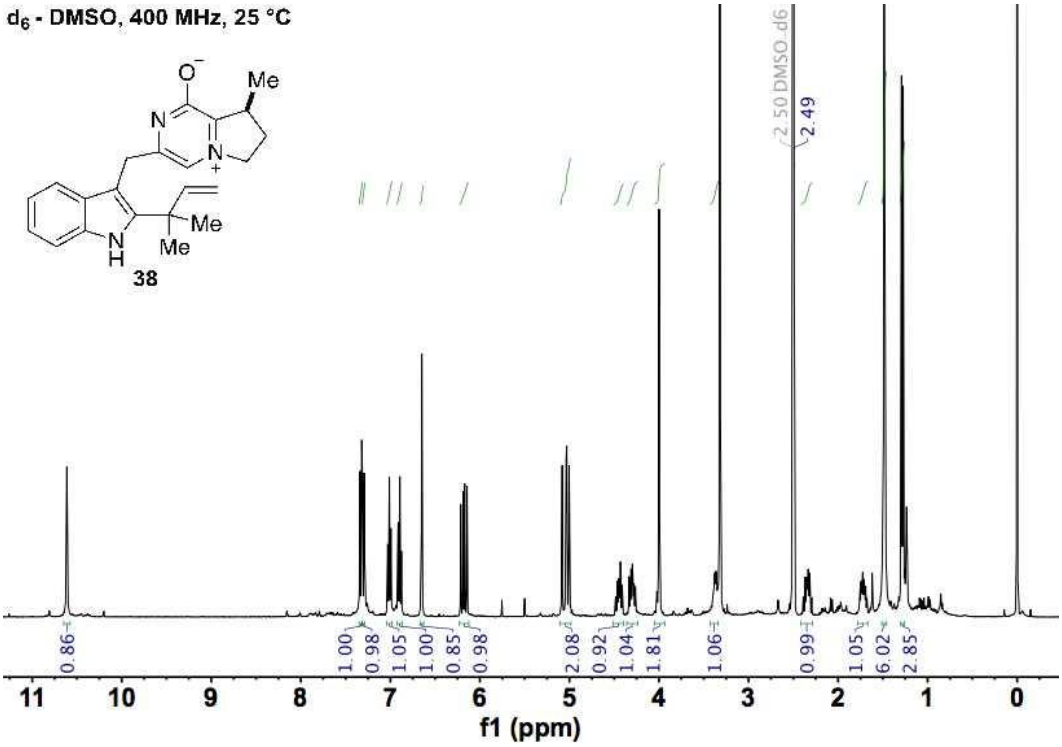

**Supplementary Figure 57.** $^{13}$C NMR spectrum of **36**.



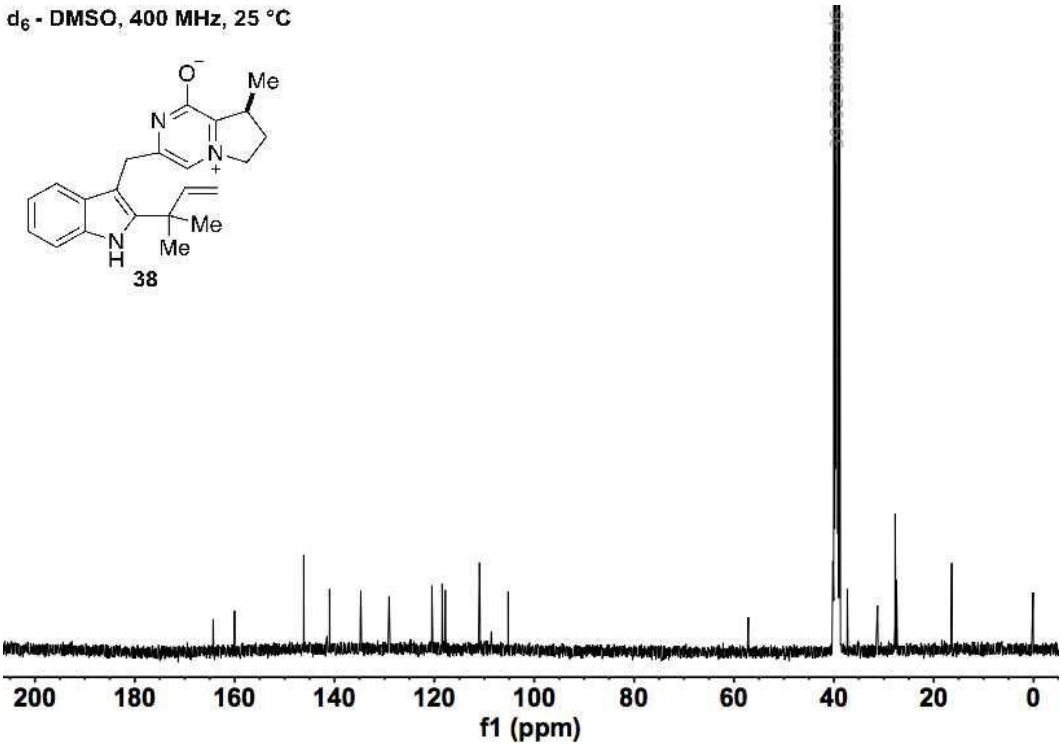

**Supplementary Figure 58.** $^{1}$H-$^{1}$H COSY spectrum of **36**.



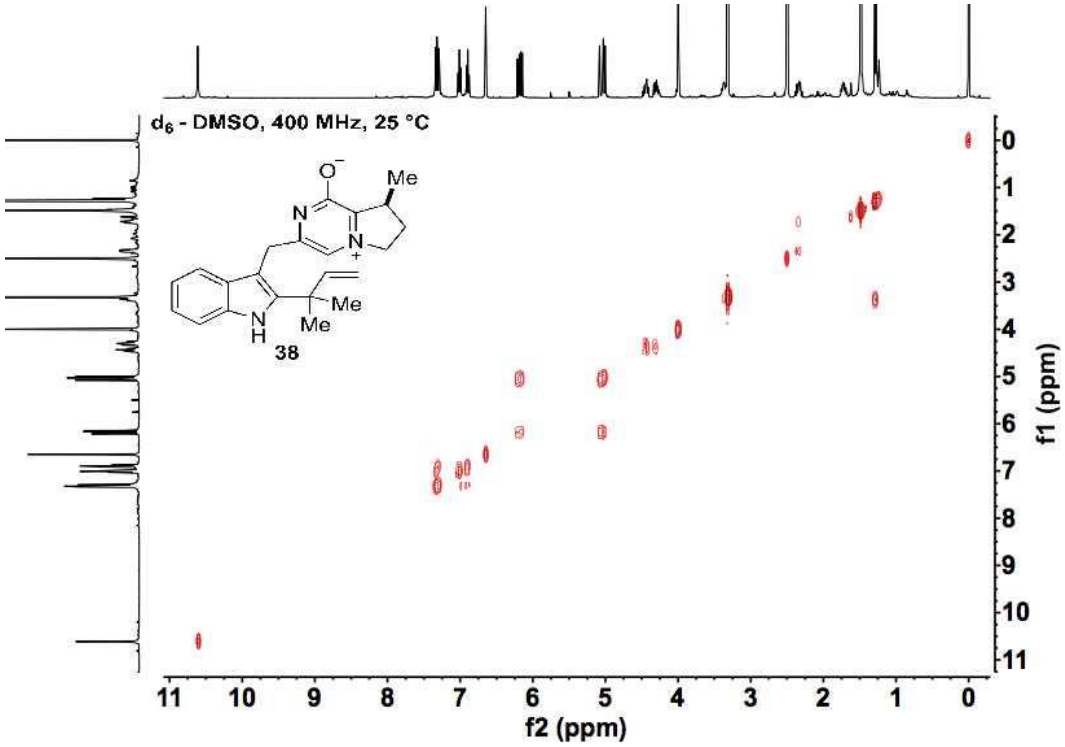

**Supplementary Figure 59.** $^{1}$H-$^{13}$C HSQC spectrum of **36**.



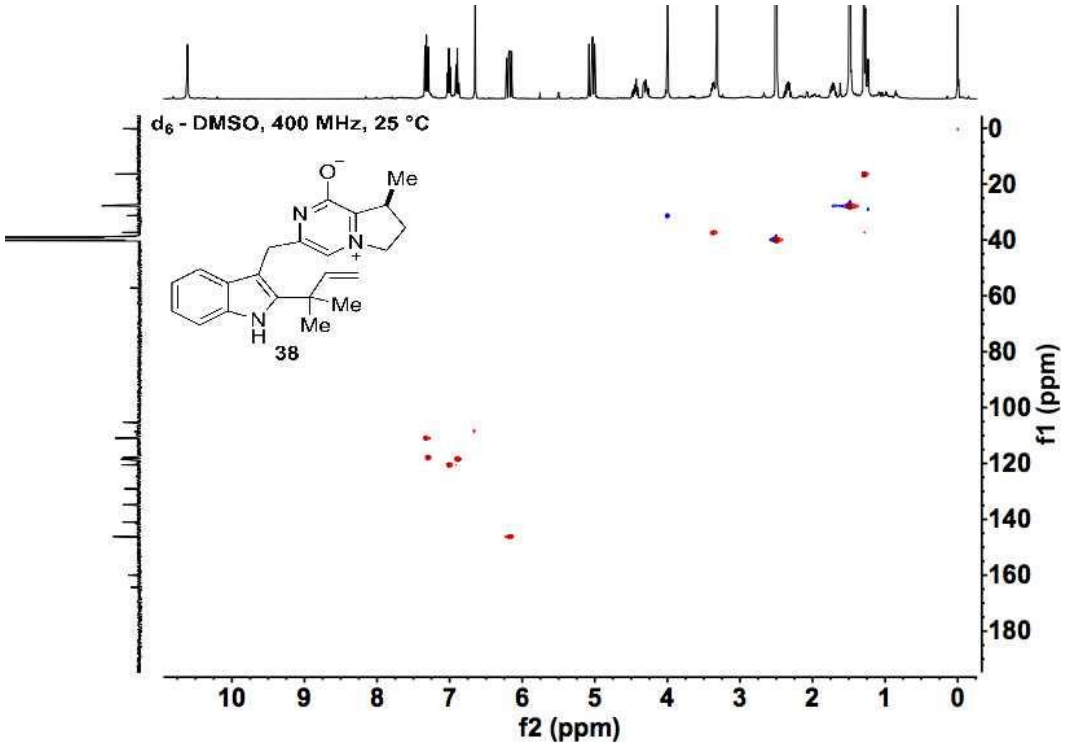

**Supplementary Figure 60.** $^{1}$H NMR spectrum of **38**.



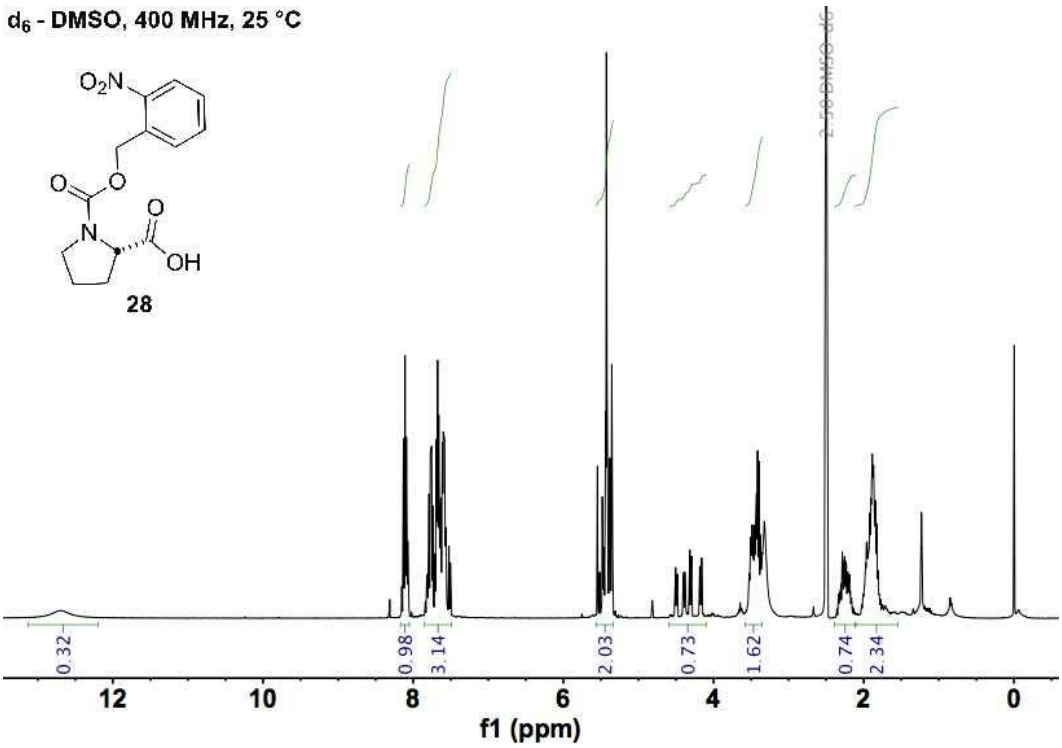

**Supplementary Figure 61.** $^{13}$C NMR spectrum of **38**.



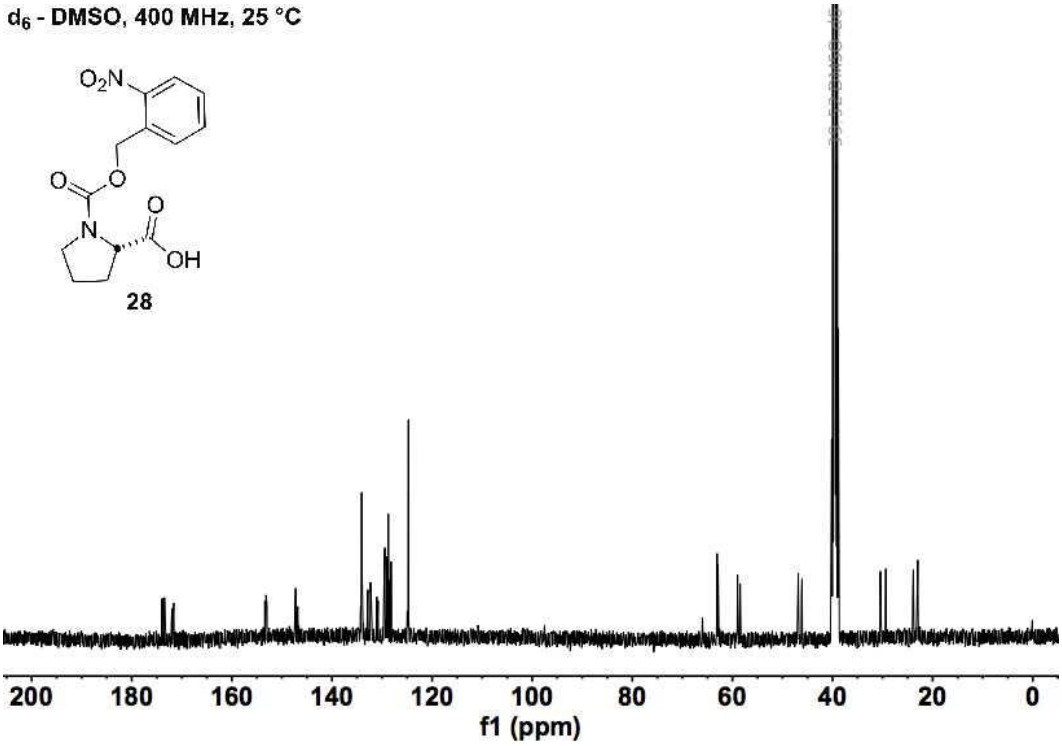

**Supplementary Figure 62.** $^{1}$H-$^{1}$H COSY spectrum of **38**.



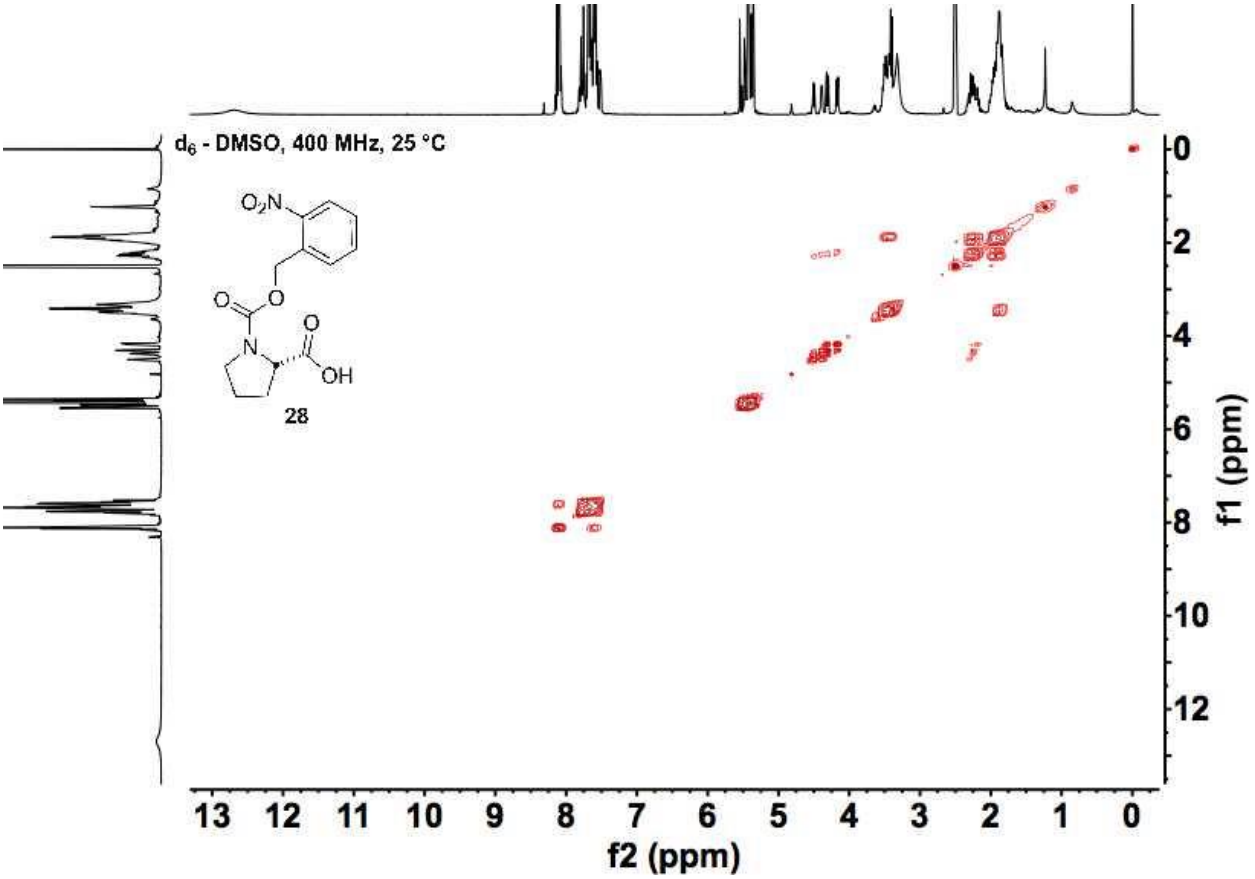

**Supplementary Figure 63.** $^{1}$H-$^{13}$C HSQC spectrum of **38**.



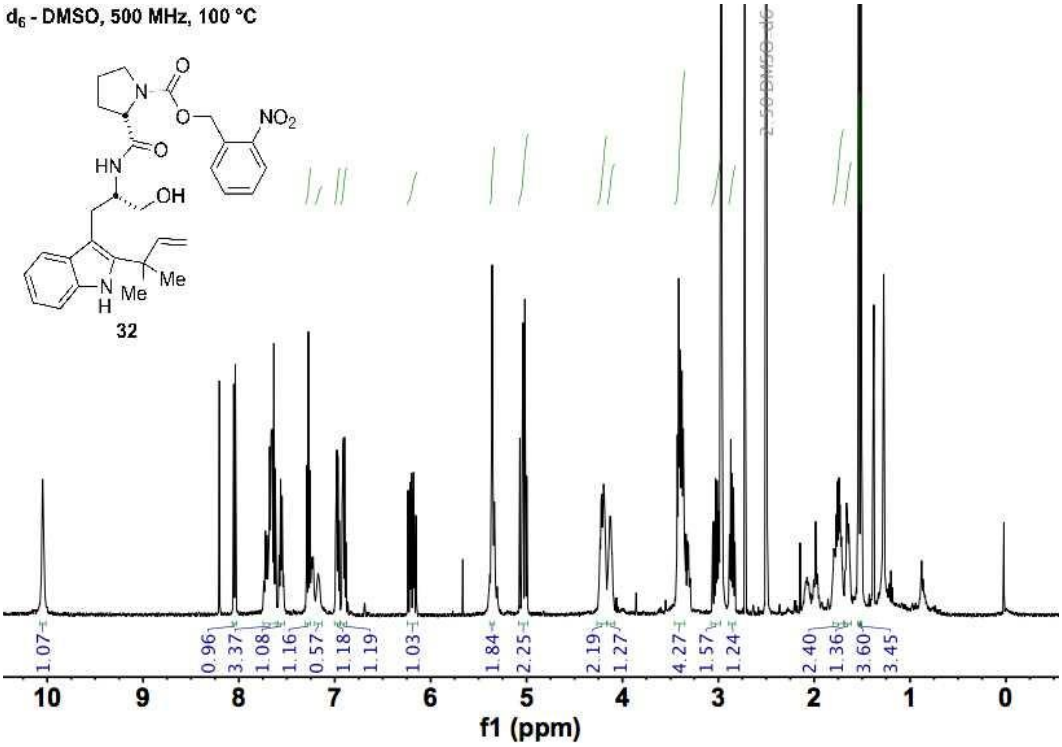

**Supplementary Figure 64.** $^{1}$H NMR spectrum of **28**.



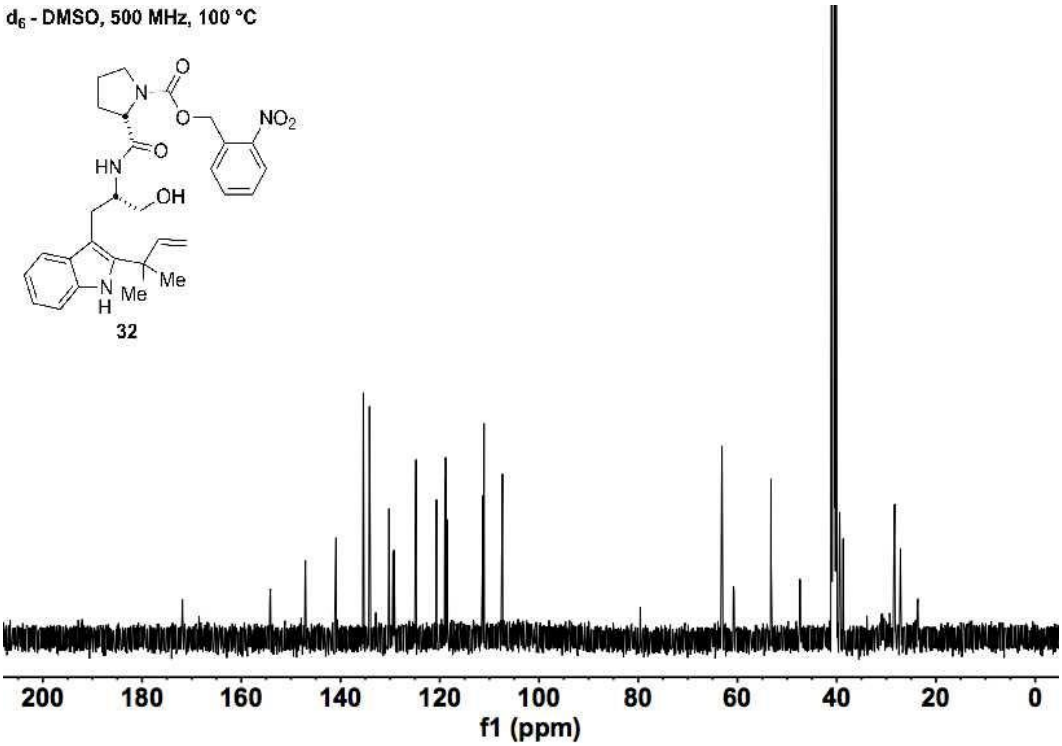

**Supplementary Figure 65.** $^{13}$C NMR spectrum of **28**.



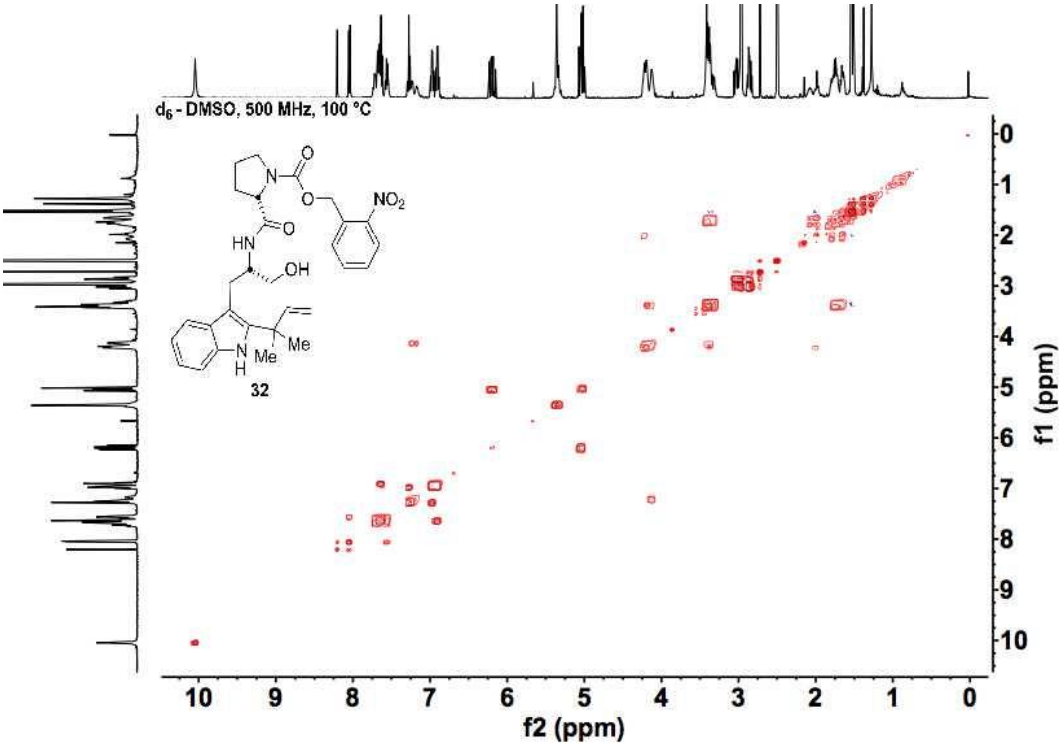

**Supplementary Figure 66.** $^{1}$H-$^{1}$H COSY spectrum of **28**.



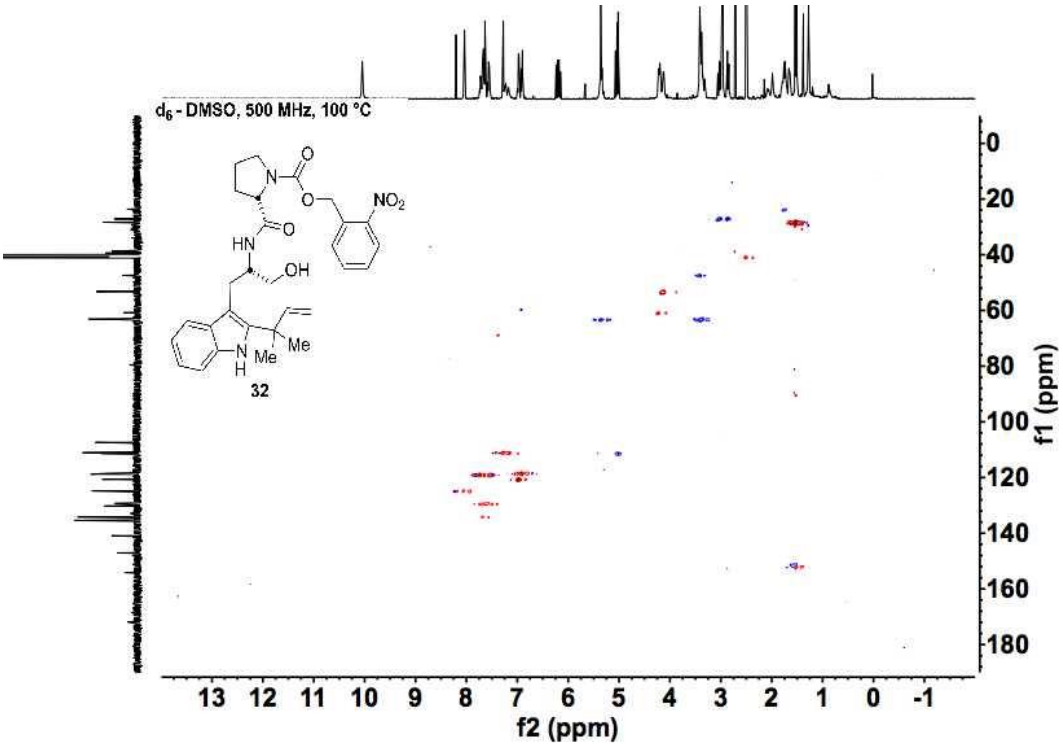

**Supplementary Figure 67.** $^{1}$H NMR spectrum of **32**.



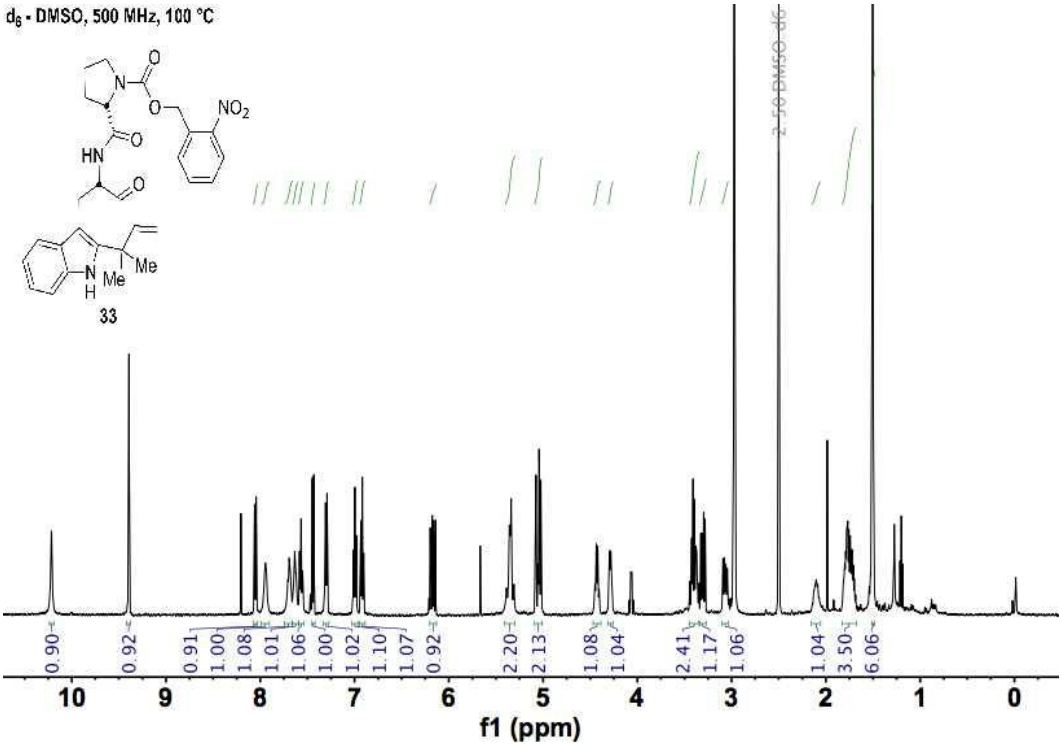

**Supplementary Figure 68.** $^{13}$C NMR spectrum of **32**.



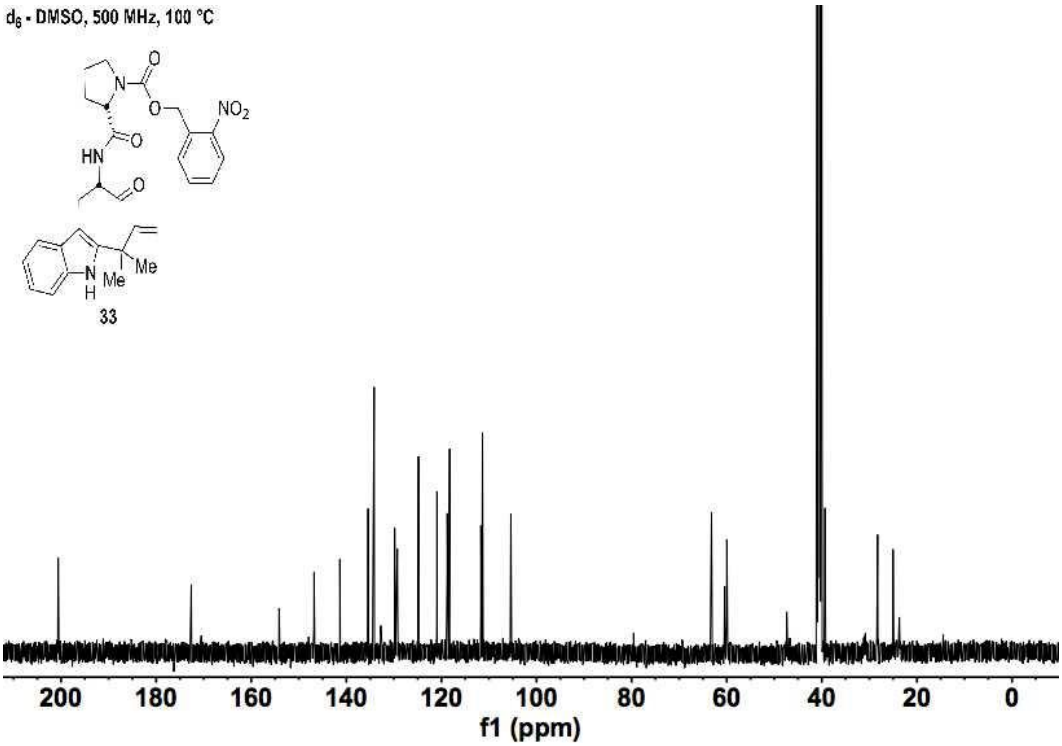

**Supplementary Figure 69.** $^{1}$H-$^{1}$H COSY spectrum of **32**.



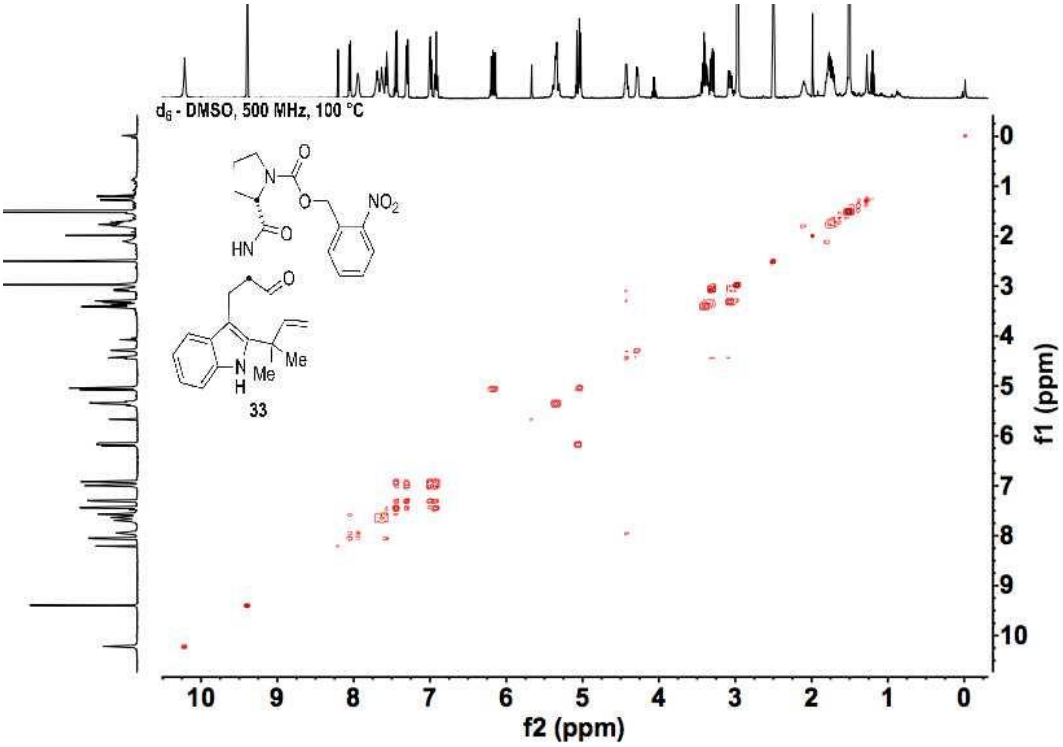

**Supplementary Figure 70.** $^{1}$H-$^{13}$C HSQC spectrum of **32**.



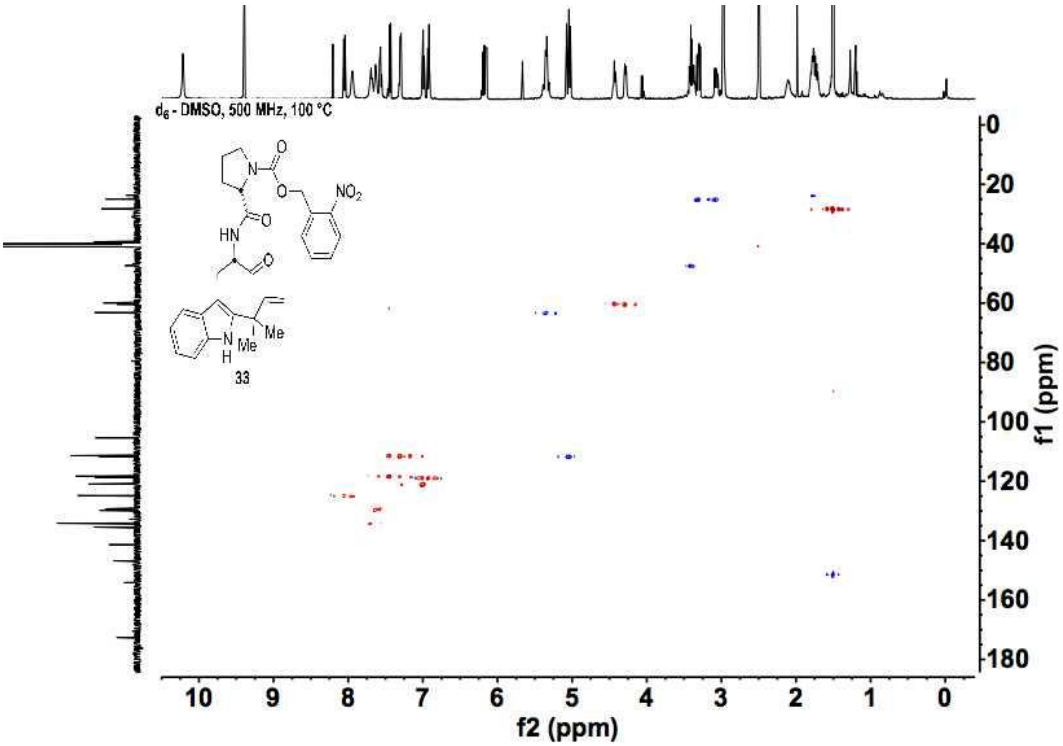

**Supplementary Figure 71.** $^{1}$H NMR spectrum of **33**.



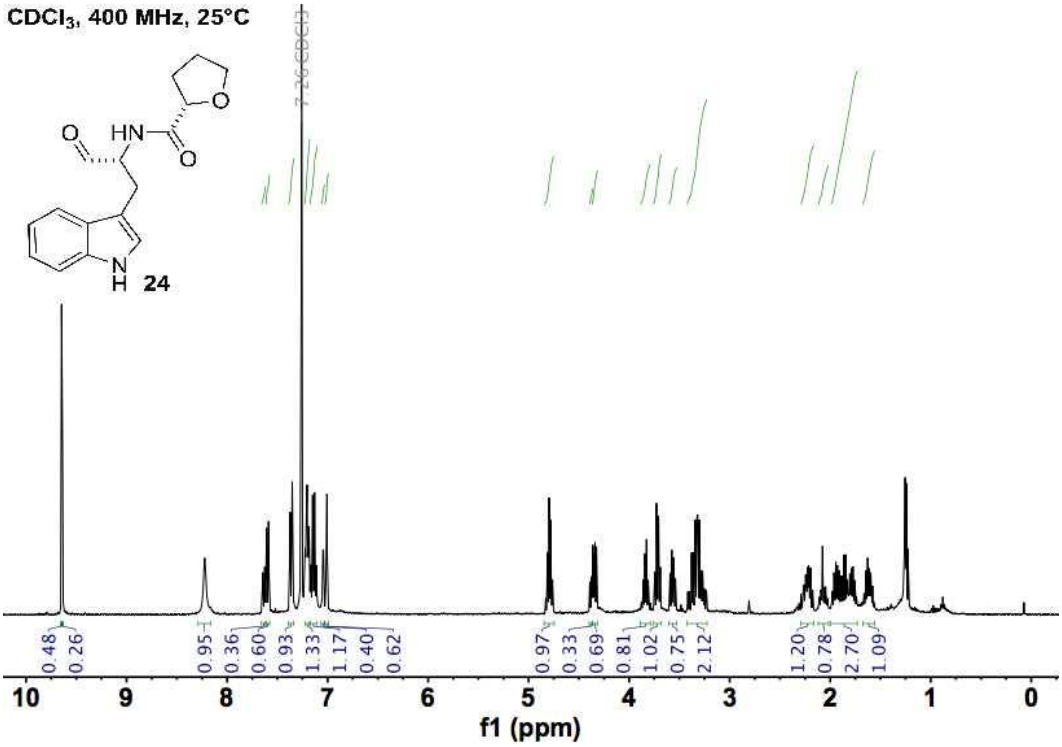

**Supplementary Figure 72.** $^{13}$C NMR spectrum of **33**.



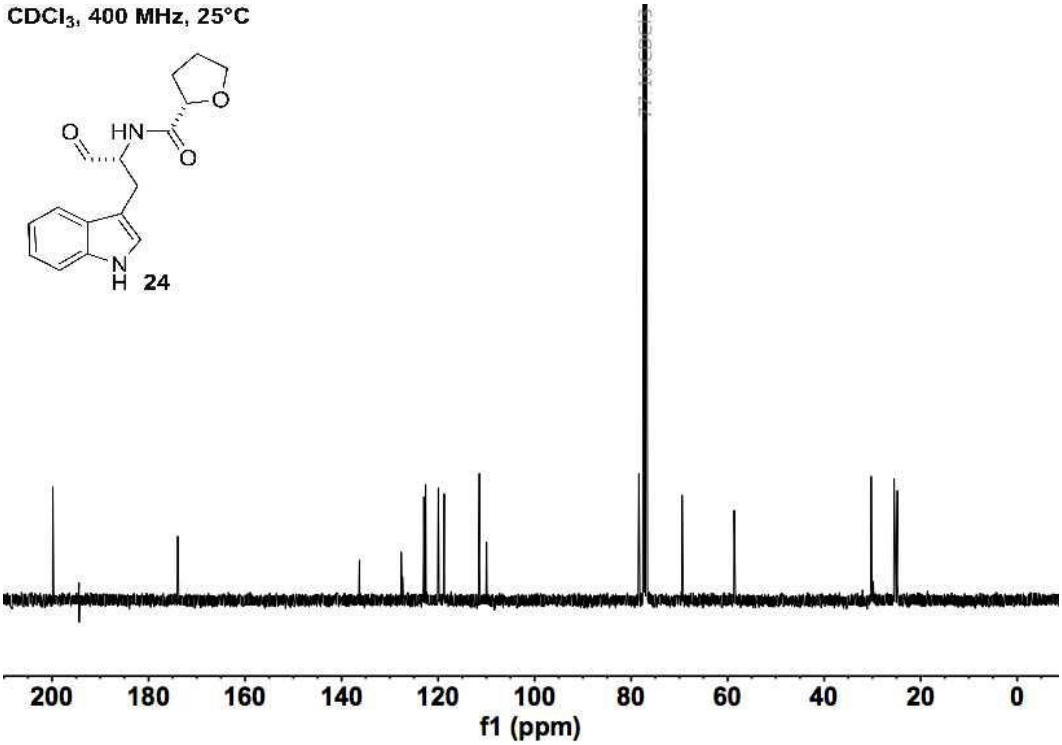

**Supplementary Figure 73.** $^{1}$H-$^{1}$H COSY spectrum of **33**.



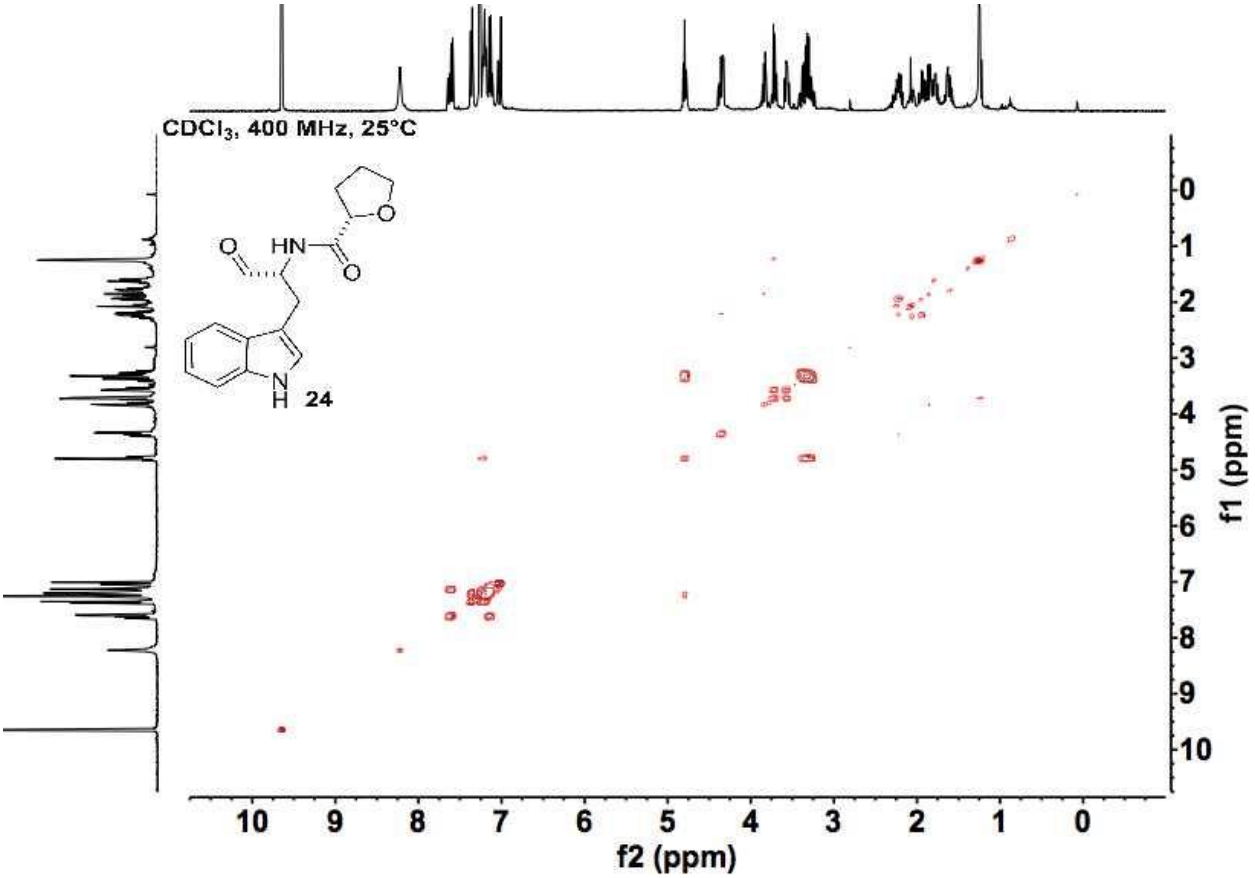

**Supplementary Figure 74.** $^{1}$H-$^{13}$C HSQC spectrum of **33**.



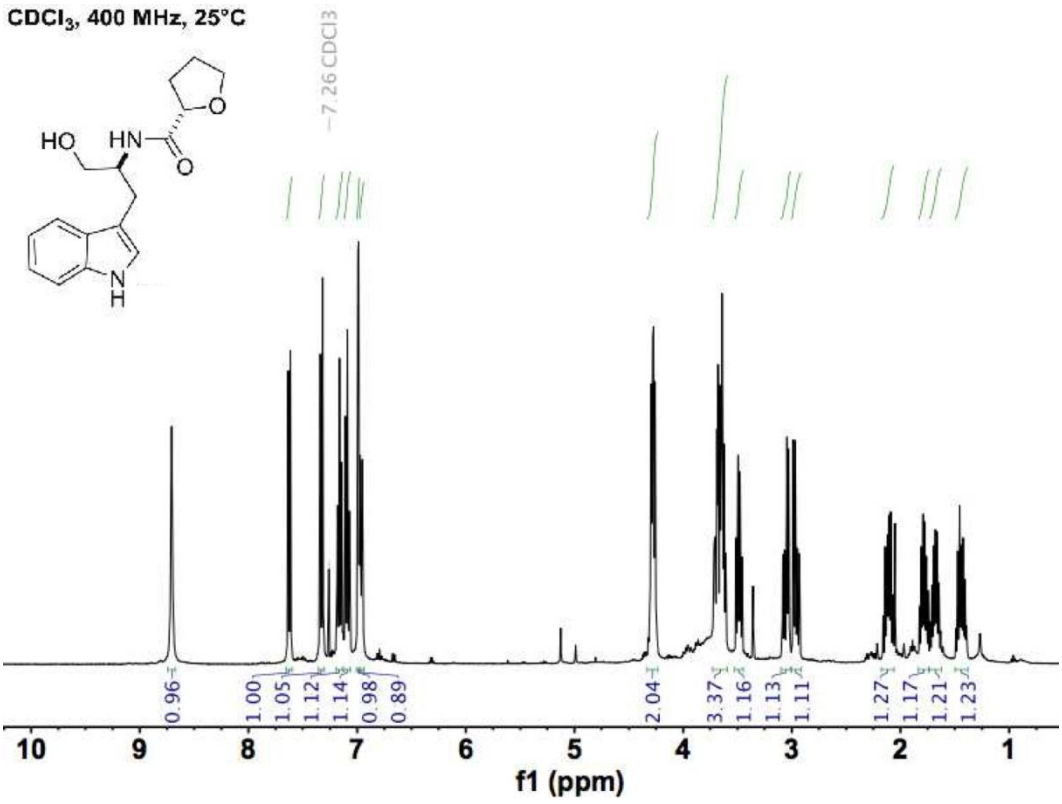

**Supplementary Figure 75.** $^{1}$H NMR spectrum of **24**.



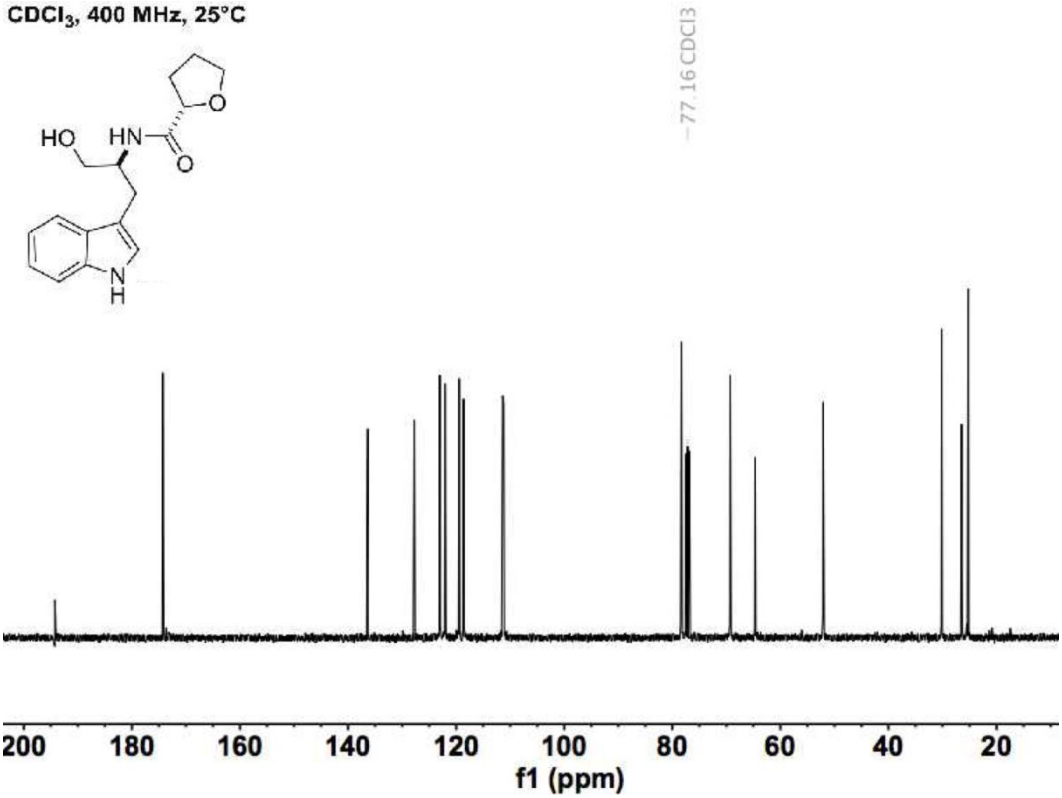

**Supplementary Figure 76.** $^{13}$C NMR spectrum of **24**.



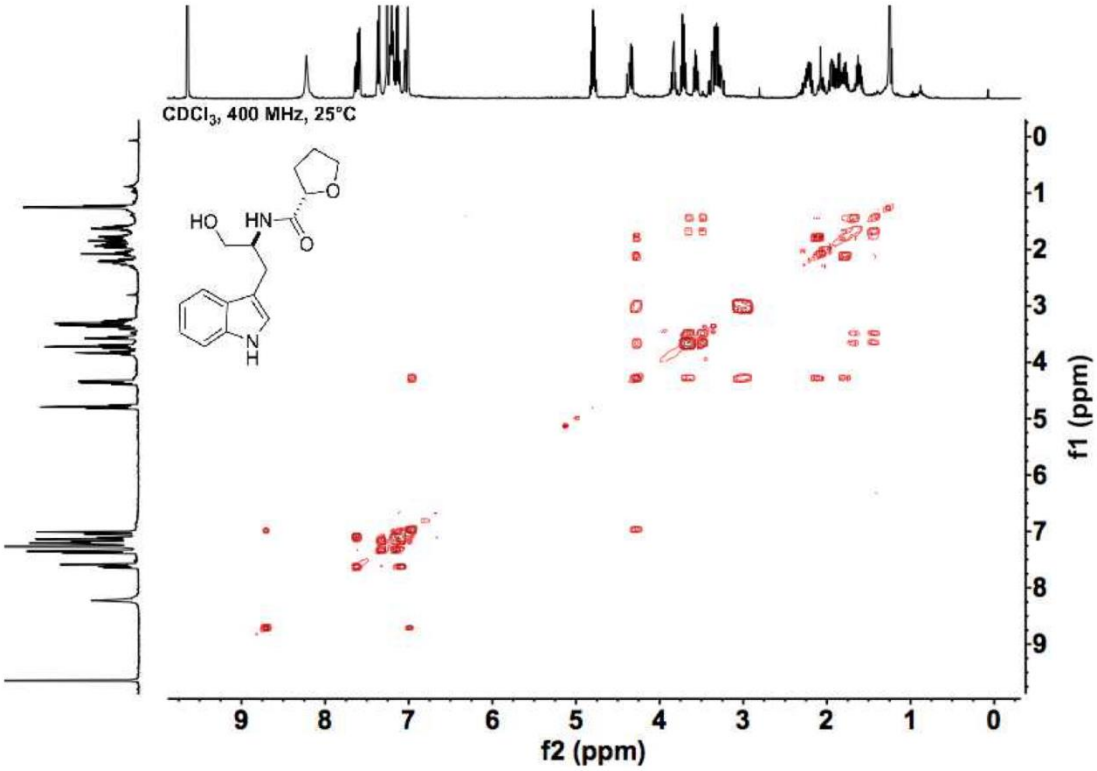

**Supplementary Figure 77.** $^{1}$H-$^{1}$H COSY spectrum of **24**.



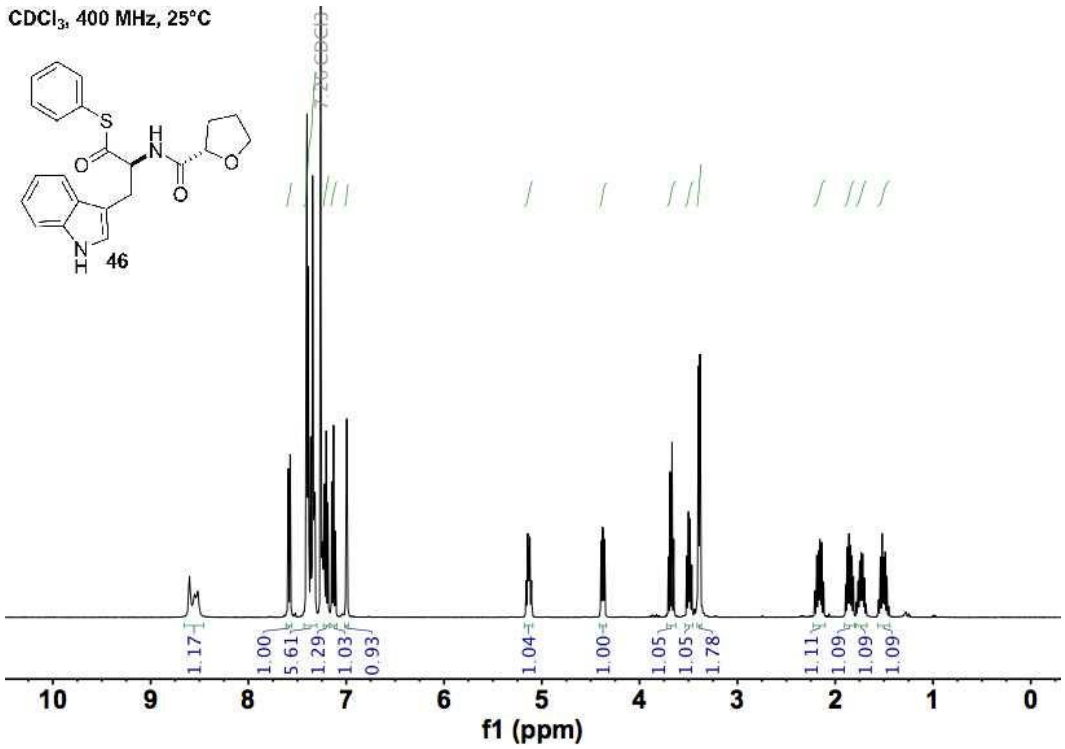

**Supplementary Figure 78.** $^{1}$H NMR spectrum of **25**.



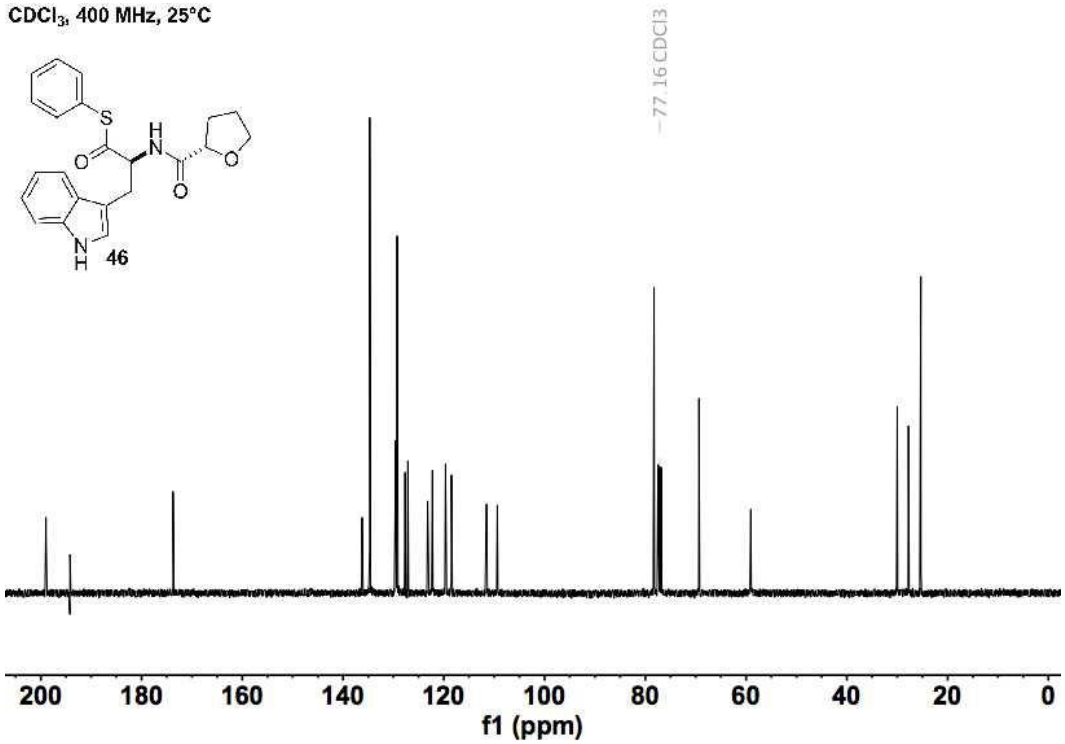

**Supplementary Figure 79.** $^{13}$C NMR spectrum of **25**.



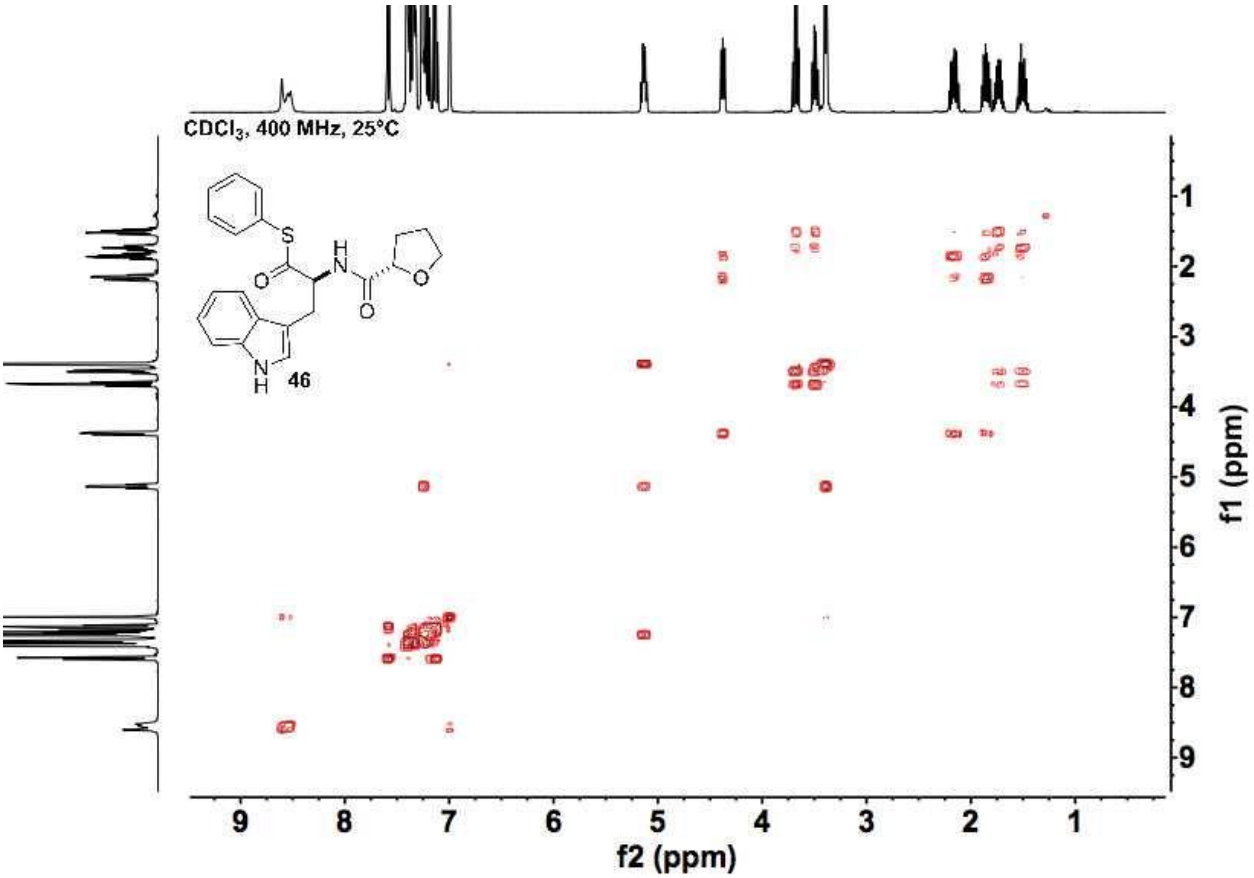

**Supplementary Figure 80.** $^{1}$H-$^{1}$H COSY spectrum of **25**.




**Supplementary Figure 81.** $^{1}$H NMR spectrum of **46**.




**Supplementary Figure 82.** $^{13}$C NMR spectrum of **46**.




**Supplementary Figure 83.** $^{1}$H-$^{1}$H COSY spectrum of **46**.

## 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’ |

## Supplementary Table 2. Crystallographic table

| | PhqB R · NADPH | MalC | PhqE D166N · **11** · $NADP^+$ | PhqE · **1** · $NADP^+$ |
|---|---|---|---|---|
| **Data collection** | | | | |
| Space group | *I 222* | *P 42* | *C 2* | *C 2* |
| Unit cell parameters | | | | |
| *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 |
| Wavelength (Å) | 1.033 | 1.033 | 1.033 | 1.033 |
| 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) |
| Completeness (%) | 99.8 (99.1) | 99.2 (92.1) | 97.2 (91.3) | 99.3 (97.4) |
| Multiplicity | 13.3 (12.7) | 6.6 (5.0) | 6.9 (6.1) | 6.5 (6.9) |
| Mean $I/\sigma$ | 20.5 (1.1) | 15.4 (1.4) | 12.8 (0.7) | 8.7 (1.6) |
| $R_{meas}$ | 0.077 (2.57) | 0.079 (1.21) | 0.077 (2.14) | 0.180 (1.07) |
| CC1/2 | 1 (0.64) | 1 (0.45) | 1 (0.81) | 1 (0.88) |
| **Refinements** | | | | |
| Resolution (Å) | 45.82 – 2.60 | 42.97 – 1.60 | 48.76 – 1.89 | 46.38 – 2.29 |
| No. Reflections | 14737 | 107859 | 112785 | 65748 |
| $R_{work}$ | 0.27 | 0.17 | 0.28 | 0.29 |
| $R_{free}$ | 0.31 | 0.20 | 0.31 | 0.34 |
| No. of Atoms | 2755 | 8566 | 11696 | 11735 |
| Protein | 2707 | 7735 | 11196 | 11190 |
| Ligands | 48 | -- | 313 | 438 |
| Water | -- | 831 | 187 | 107 |
| Avg *B*-values ($Å^2$) | 127.51 | 29.77 | 68.42 | 75.38 |
| Protein | 126.85 | 29.32 | 68.90 | 75.89 |
| Ligands | 164.86 | -- | 61.17 | 68.58 |
| Water | -- | 33.89 | 52.10 | 49.76 |
| Ramachandran plot: favored/allowed/outlier (%) | 95.6/4.4/0 | 98.2/1.8/0 | 97.8/2.2/0 | 96.7/3.3/0 |
| RMSD bonds (Å) | 0.010 | 0.008 | 0.010 | 0.010 |
| RMSD angles (°) | 1.25 | 0.97 | 1.50 | 1.60 |

## Supplementary Table 3. 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<br>AOC84388.1<br>AGC83573.1 |
| 11 | 394 | P450 monooxygenase | [*Penicillium fellutanum*] (44/63); *phqM* | AGA37280.1<br>KXG49075.1 |
| 12 | 452 | P450 monooxygenase | [*Penicillium fellutanum*] (38/55); *phqL* | AGA37279.1<br>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<br>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 Whicher, J. R. *et al.* Cyanobacterial polyketide synthase docking domains: a tool for engineering natural product biosynthesis. *Chem Biol* **20**, 1340-1351 (2013).

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

4 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).

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

6 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).

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

11 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).

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

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

14 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).

15 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).

16 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).

17 Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352-367 (2012).

18 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).

19 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).

20 Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
21 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).
22 The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.
23 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).
24 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.
25 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).
26 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).
27 Singh, U. C. & Kollman, P. A. An approach to computing electrostatic charges for molecules. *J Comput Chem* **5**, 129-145 (1984).
28 Besler, B. H., Merz, K. M. & Kollman, P. A. Atomic charges derived from semiempirical methods. *J Comput Chem* **11**, 431-439 (1990).
29 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.
30 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).
31 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).
32 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).

33 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.

34 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).

35 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).

36 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).

37 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).

38 Chhabra, A. *et al.* Nonprocessive [2 + 2]e- off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. *Proc Natl Acad Sci* **109**, 5681-5686 (2012).

39 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).

40 Filling, C. *et al.* Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J Biol Chem* **277**, 25677-25684 (2002).

41 Oppermann, U. *et al.* Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chem Biol Interact* **143-144**, 247-253 (2003).

42 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).

43 Banani, H. *et al.* Genome sequencing and secondary metabolism of the postharvest pathogen *Penicillium griseofulvum*. *BMC Genomics* **17**, 19 (2016).