Bacterial Cytochrome P450-Mediated Biosynthesis of Non-Symmetrical Diketopiperazine Dimers: Discovery, Catalysis, and Structure

by

Vikram V. Shende

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2019

Doctoral Committee:

Professor David H. Sherman, Chair
Professor Anna K. Mapp
Professor John Montgomery
Professor Corinna Schindler
Professor Emily Scott
Professor Janet L. Smith

Vikram V. Shende
vvshende@umich.edu
ORCID iD: 0000-0001-8396-6297

© Vikram V Shende 2019

Dedication

This dissertation is dedicated to my parents. My mother and father have always nurtured my curiosity and have provided a perfect template for the kind of person I want to be. None of this would be possible without their continued support.

Acknowledgements

Firstly, I must acknowledge and thank my advisor Professor David Sherman for his unwavering support throughout my Ph.D. I was extremely lucky to be able to join the Sherman lab with little experience or training in molecular biology, biochemistry or microbiology, but Pr. Sherman still accepted me and entrusted me to work on high priority projects with high the possibility for a large impact in the scientific community. As, if not more importantly, Pr. Sherman mentored me on being a good and complete human. Thank you, David.

I also must thank the members of my thesis committee; Pr. Corinna Schindler, Pr. Janet Smith, Pr. Anna Mapp, Pr. John Montgomery, and Pr. Emily Scott. All of you have helped to keep me “honest” throughout this entire processes, and also supported me through the difficult and trying times, offering me support and guidance. Thanks to you all.

I have also had the incredible luxury to work with an extraordinary group of students and post-docs in and around the Sherman lab, and have valued discussions with all of you. Without Dr. Yogan Khatri, Dr. Sean Newmister, Dr. Douglas Hansen, Dr. Aaron Koch, Pr. Alison Narayan and Dr. Matthew DeMars, teaching me the fundamentals I never would have been able to be successful at compiling this work. I owe you all many thanks and beers in hopefully the near future.

Finally, I have to thank my friends and family. Pursuing a graduate degree in the physical can be a lonesome and long road, and one that requires solitude with the unfortunate outcome of neglecting the relationships that mean the most to you, and with remain with you long after you’re done. While there are too many of you to name I love you all.

Table of Contents

List of Figures

List of Tables

Abstract

The simplest cyclic peptide is composed of two amino acids and forms a 2,5-diketopiperazine (DKP). Diketopiperazine containing natural products can be found from eukaryotic, fungal, and bacterial sources and are present in a diverse array of scaffolds. DKPs are most commonly biosynthesized by non-ribosomal peptide synthetases (NRPS) modules condensing two amino acids into a DKP upon offloading by a thioesterase. Recently a new class of enzymes, cyclodipeptide synthases (CDPS) have been identified as capable of synthesizing DKPs from tRNA charged amino acid precursors. Since this discovery, there has been a dramatic rise in the identification and characterization of CDPS containing gene clusters, and some classic natural products have been ascribed to these clusters, including albonoursin and bicyclomycin.

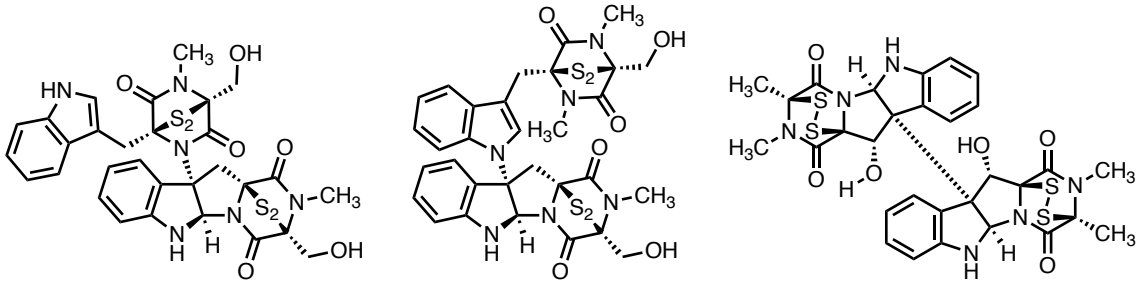
The subject of this thesis is the study of a cyclodipeptide synthase containing gene clusters which also harbor a cytochrome P450 which oxidatively dimerizes DKP monomers into topologically complex dimers with exquisite selectivity. These enzymes catalyze a C–H functionalization reaction and transform an sp^2 hybridized C–H bond into a C–C or C–N bond. We discovered a series of unique cytochromes, NznB, NascB, and homologues NascB-S1868, and NascB-F5053 all of which catalyze reactions novel site- stereo- and chemoselectivity. Chapter 1 discusses the synthetic and biosynthetic literature regarding diketopiperazine dimers. In the Chapters 2 and 3, these cytochromes are characterized biochemically using native and non-native substrates in *in vitro* and *in vivo* biocatalytic reactions demonstrating their synthetic utility, and Chapter 4 interrogates the structure and mechanism of the dimerization reaction of NascB-F5053. Chapter 5 discusses the future directions and applications of these cytochromes.

Chapter 1
Introduction

1.1 Diketopiperazine Dimers: History, Isolation and Biosynthetic Hypotheses

In the effort to discover of new molecules to inhibit bacterial growth, in 1944 Selman Waksman and Elizabeth Bugie and isolated a compound, (+)–chaetomin (named for the *Chaetomium cochliodes*)[1-2], with antibiotic activity similar to that of penicillin against a spectrum of gram-positive pathogenic bacteria including *Staphylococcus aureus.* The structure of chaetomin went unknown for over 30 years, despite the reconfirmation of its antibacterial activity as well as discovery of its ability to inhibit viral ribonucleic acid synthesis *in vivo* in a number of cell lines.[3] Elemental analysis and NMR spectroscopy lead to the initially proposed structure **1** from Safe and Taylor[3] in 1972, which was later corrected by McInnes, Taylor, Walter in 1976 to **2** through a series of shrewd ^{15}N NMR experiments using ^{15}N-enriched chaetomin isolated from feeding studies using $Na^{15}NO_3$.[4] Two years prior, the first structures of symmetrical dimeric epidithiodiketopiperazines (ETPs) (+)-chaetocin and (+)-verticillin were solved by Hauser[5] and Katagiri[6] who also reported antibiotic, and anti-cancer activities. Based on the structure of these compounds, as well as incorporation of radiolabelled amino acid precursors in the investigation of the biosynthesis of monomeric ETP, gliotoxin, these groups hypothesized that monomeric precursors of these dimers arise via the condensation of two amino acids into DKP monomers. They postulated that similar to the biosynthetic proposal put forth by Woodward[7] and Robinson[8] for the related polypyrroloindoline natural products, that these monomers undergo a single electron oxidation-dimerization cascade to give rise to the dimeric scaffold to undergo further thiolation.

A) First structures solved of dimeric DKP natural products

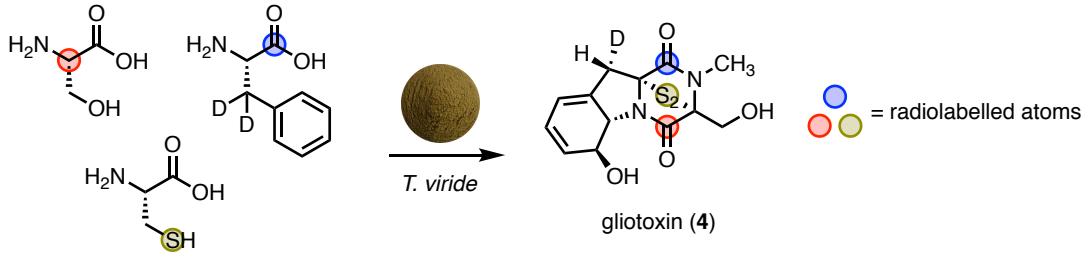


(+)-chaetomin (**1**)
initially proposed structure

(+)-chaetomin (**2**)
corrected structure

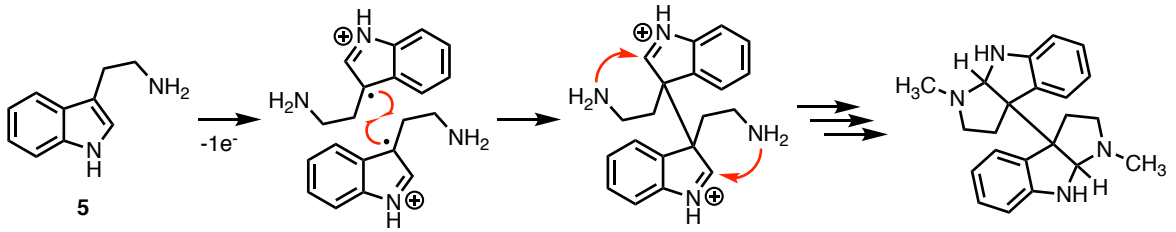
(+)-verticillin (**3**)

B) Incorporation of radiolabelled biosynthetic precursors into gliotoxin scaffold



gliotoxin (**4**)

C) Biosynthetic hypothesis put forth by Woodward and Robinson for related polypyyroloindole natural products



5

(±)-chimonanthine (**6**)
meso-chimonanthine (**7**)

Figure 1-1. A) Initially proposed structure of (+)-chaetomin (**1**), corrected structure of (+)-chaetomin (**2**), and (+)-verticillin (**3**). B) Incorporation of radiolabelled intermediates in the biosynthesis of gliotoxin (**4**). C) biosynthetic proposal of Woodward and Robinson.[9]

Since these seminal reports, the isolation of diketopiperazine dimers (DKPs) with a broad range of biological activites have been reported, including but not limited to antibacterial, anticancer, antiviral, antiparasitic, antifungal, antimalarial, immunosuppressive, immunomodulatory, phytotoxic, nematicidial and antiplatelet effects[10]. The broad range of biological activities as well as well as curiosity surrounding the biosynthesis of these molecules led to the discovery of over 100 members of this

diverse family[11], along with investigations into their access via total synthesis and as well as preliminary investigations into their biosynthesis.

1.2 Structural Diversity and Total Synthesis of Diketopiperazine Dimers And Their Putative Biosynthetic Congeners

The dimeric DKP natural products are composed of a wide variety of proteinogenic amino acids, possess a variety of oxidation states at both carbon and sulfur, (*R*)- and (*S*)- configurations at their axis of dimerization, and multiple connectivities between the two monomers (Figure 1-2). Given this multiplicity in constitutional variables, there are no general classes of DKP dimers. We chose to classify these compounds by their differences in connectivity; as the selectively in the biosynthetic dimerization of these compounds is unknown and likely represents a point of biosynthetic divergence.

Similarly, each connectivity also represents a point of synthetic divergence, as access to each connectivity is accompanied with a unique synthetic challenge. As such, a variety of elegant synthetic methods towards the total synthesis of these compounds have been developed in order to overcome each of these unique challenges. To exemplify both the variety of connectivities native to DKP dimers as well as the diversity of synthetic methods to assemble these compounds, representative examples from each connectivity (enumerated based on indole numbering) have been selected to display the history and variety of approaches to access these compounds. While a variety of non-indole linked dimers are also known (Figure 1-3**,** **15**-**19**), these compounds likely arise form an alternative biogenesis and as such will not be discussed.

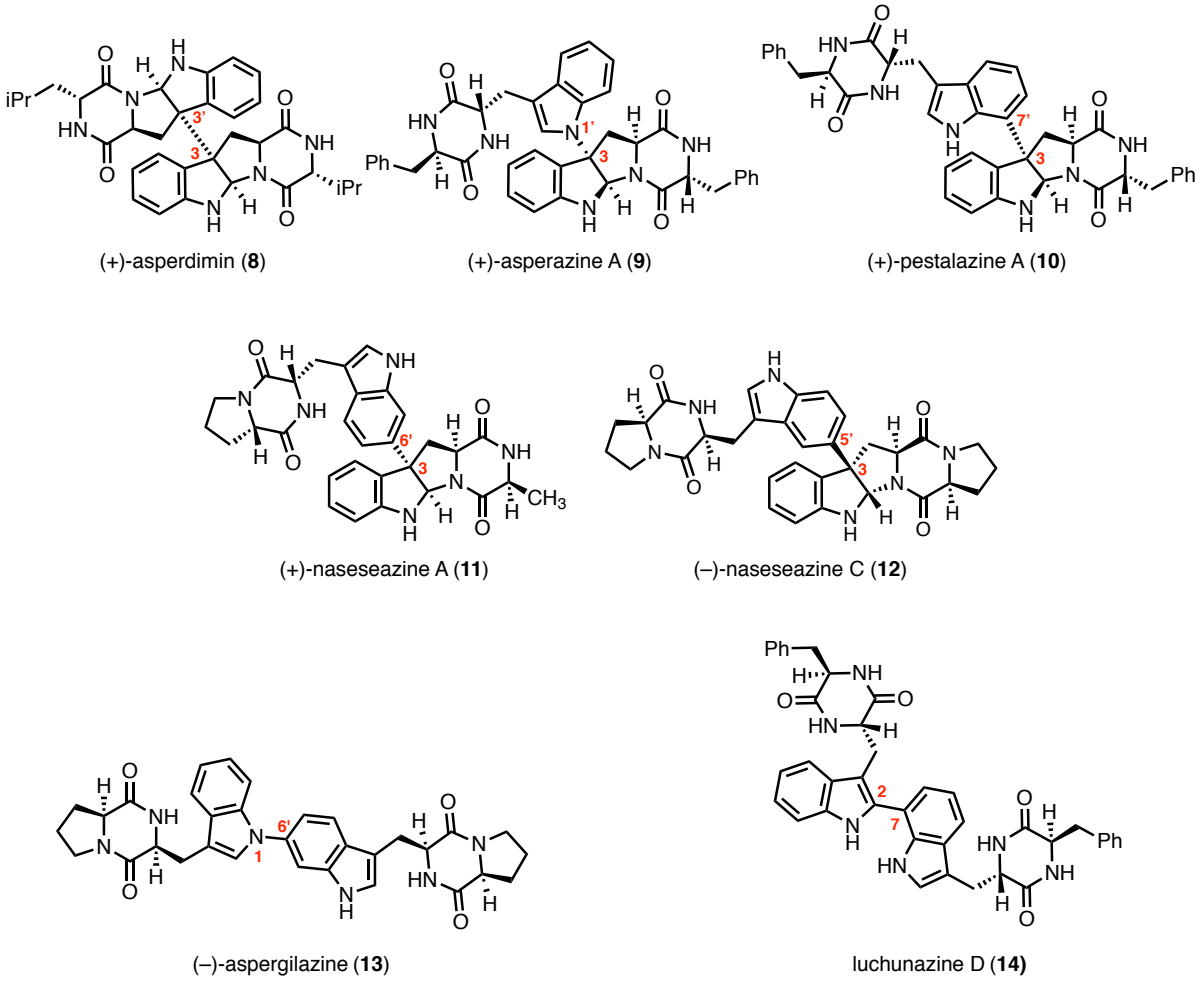


Figure 1-2. Representative dimers for all known tryptophan-tryptophan linked connectivities, including non-cyclized natural products (–)-aspergilazine (**13**) and lunchunazine D (**14**).[12-17]

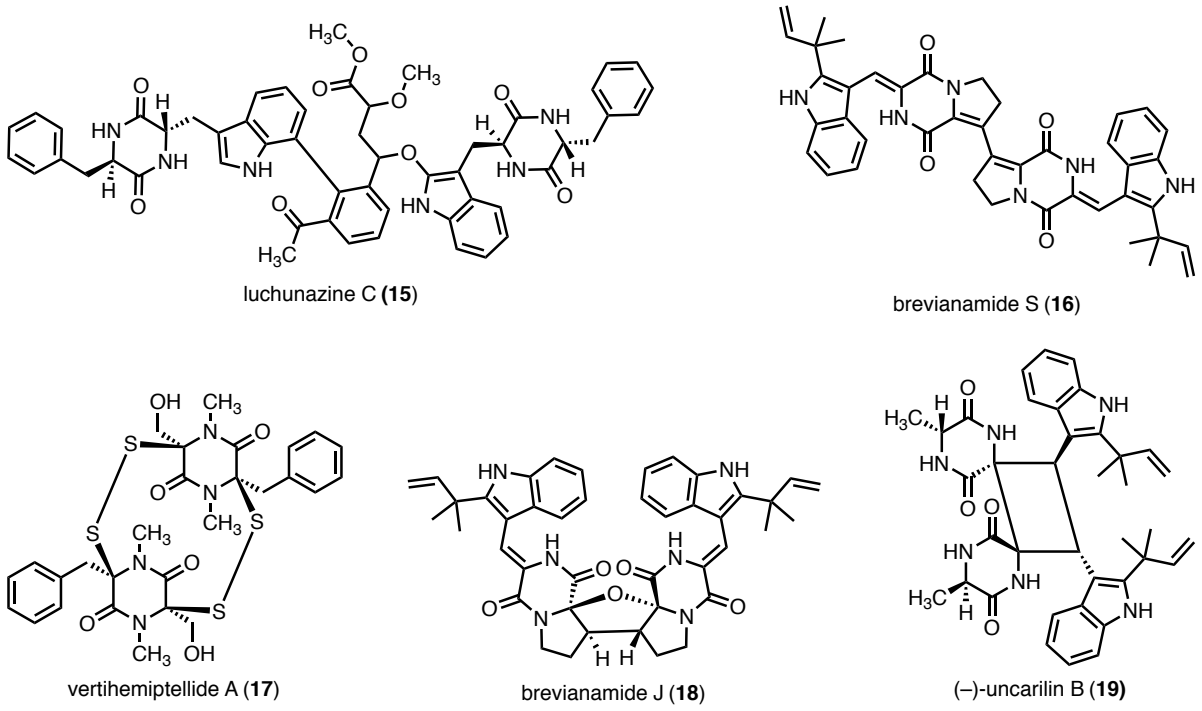


Figure 1-3. Representative examples of non-tryptophan linked DKP dimers. Notably all of these compounds have eluded total synthesis.[17-21]

1.3 Representative Total Syntheses of C3-C3' Linked Diketopiperazine Dimers

Biomimetic syntheses take inspiration from Nature's method of assembly of molecules to develop synthetic methods to perform a similar bond disconnection. The C3-C3' alkaloids are particularly well suited to biomimetic synthesis two simple building blocks can be oxidatively dimerized to assemble these symmetrical dimers, as shown by Hino an coworkers in their pioneering synthesis of (±)-folicanthine.[22] Based on the biosynthetic hypothesis put forth by Woodward[7] and Robinson[8] (Figure 1-4), Hino *et. al.* utilized a photooxidative dimerization to the total synthesis of (±)-folicanthine (**22**) from carboxymethyl protected tryptamine precursor (**20**).

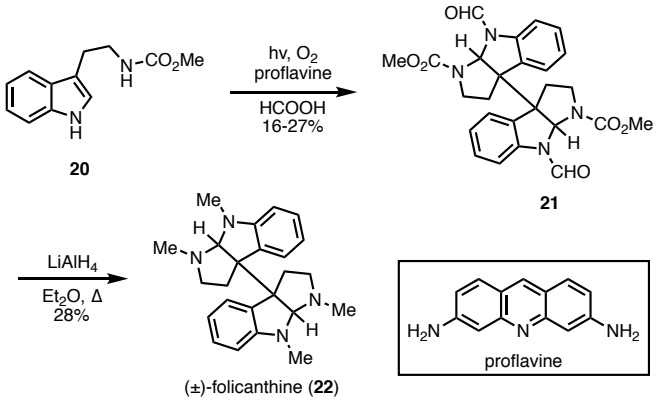


Figure 1-4. Biomimetic total synthesis of (±)-folicanthine via photooxidation.[22]

Given the shared skeleton between (±)-folicanthine and **25**, they attempted the same oxidative strategy. However, this resulted in oxidative cyclization of diketopiperazine precursor **23** rather than dimerization, giving hydroxypyrroloindole **24**, as the major product in 20% yield as a mixture of diastereomers. After screening a variety of oxidants, it was found the thallium (III) trifluoroacetate was successful in dimerizing **23** into desired product **25**, albeit in a modest 3% yield. Despite the low yield of the dimerization, this pioneering synthesis demonstrated the value of pursuing biomimetic approaches to symmetrical DKP dimers.[23]

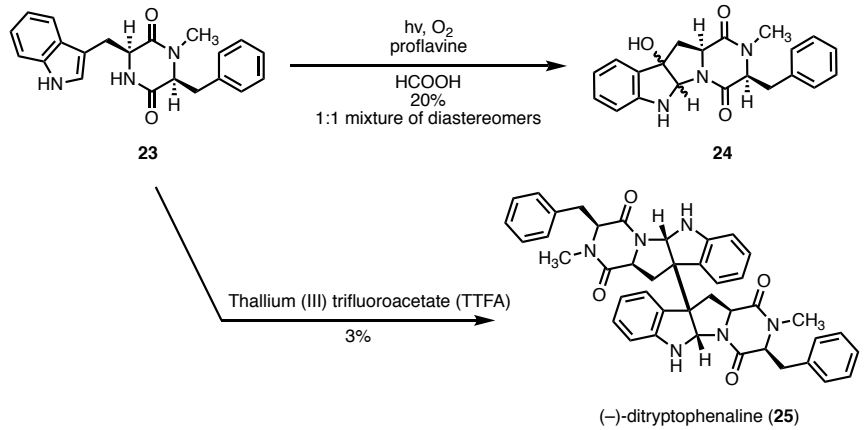


Figure 1-5. Biomimetic synthesis of (–)-ditryptophenaline via TTFA oxidation.

Movassaghi and coworkers further exemplified the value of biomimetic synthetic methodology in their stereospecific total synthesis of (+)-WIN 64821 and (–)-ditryptophenaline (Figure 1-6). Bromocyclization of diketopiperazine **26**, generated a

near (1:1) mixture of *exo*- and *endo*- cyclized tertiary bromides **27** and **28** respectively. After methylation of **28** to give **29**, these two monomers were carried forward in parallel and under cobalt-mediated solvent caged reductive coupling conditions similar to those developed by Baldwin *et. al.* in their synthesis of (+)-biatractylolide[24-25], **27** and **29** were dimerized into symmetrical dimers **30** and **31** respectively, which were then deprotected to give **25** and **32**. The parallel nature of this synthesis takes advantage of the (1:1) diastereoselectivity of bromocyclization access both natural products, and pseudo-diastereomers **25** and **32**. While this strategy is extremely efficient, the outcome is substrate-controlled, as the configuration of the tertiary bromide sets the stereochemistry of the dimerization.[26]

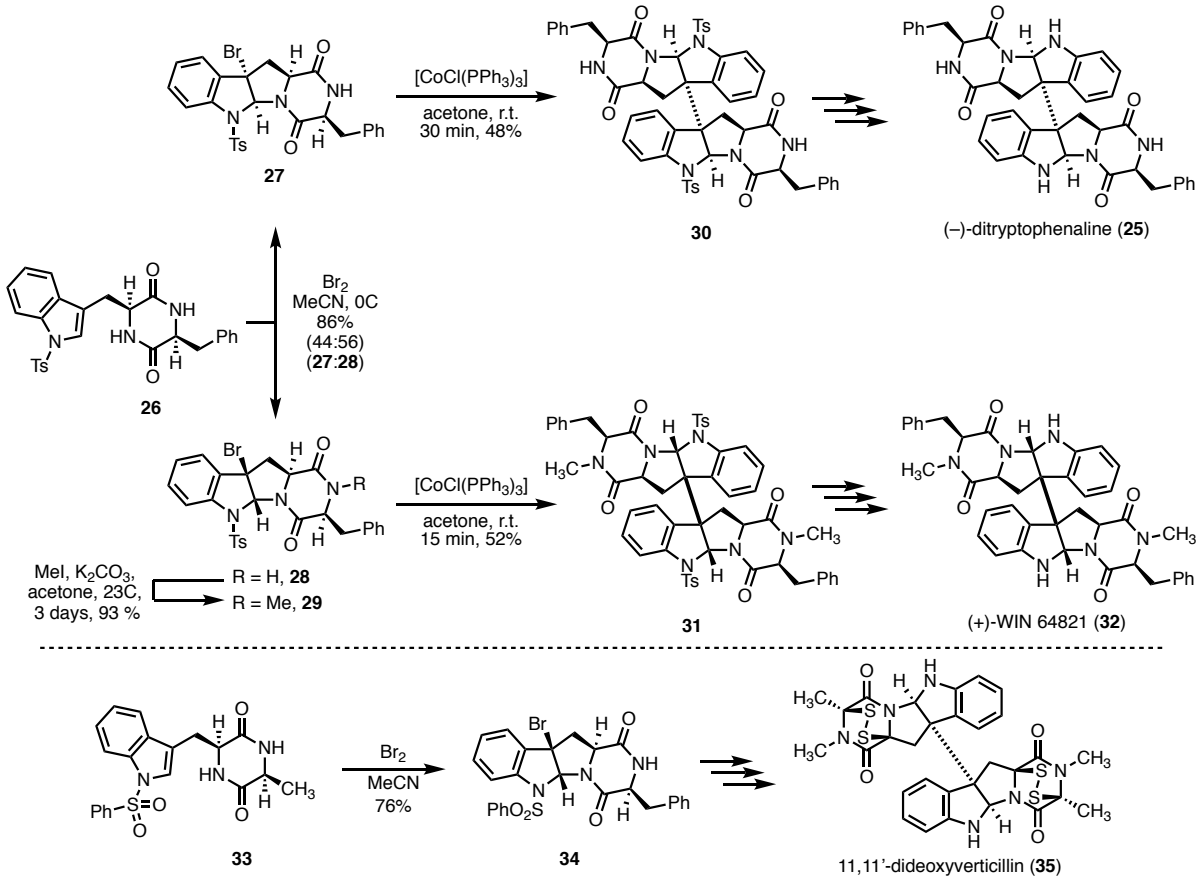


Figure 1-6. Movassaghi's bioinspired parallel total synthesis of **25** and **32** via solvent caged reductive radical coupling, and the general scheme for total synthesis of **35**.[26]

Further building on the biomimetic approach, Xia *et. al* developed a copper catalyzed oxidative cyclization/dimerization reaction to forge the dimeric structure of **32** and **25** (Figure 1-7). While this approach is again extremely efficient, it requires the use of a protecting group for the indole nitrogen, and as such the starting material is a protected tryptophan, rather than the diketopiperazine itself, and as such the diketopiperazine must be formed after dimerization. And although also sets the dimerization stereochemistry as well as forms the dimeric scaffold all at once, the reaction is again under substrate control, as selection of protecting group of the indole nitrogen steers the stereochemistry of the cyclization cascade.[27]

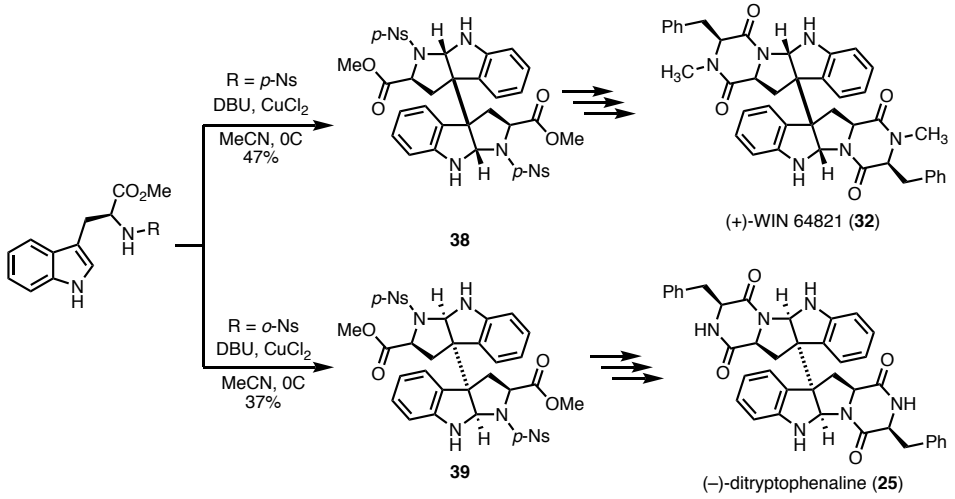


Figure 1-7. Xia's biomimetic total synthesis of **32** and **25** via copper-catalyzed oxidative dimerization of protected tryptophans.[27]

1.4 Representative Total Syntheses of C3-N1' Linked Diketopiperazine Dimers

While there have been biomimetic approaches the C3-N1' connectivity, they are based on methodology developed for the C3-C6' connectivity and will be discussed in that section. Rather, the approaches for the Liao *et. al.* utilized in the total synthesis of C3-N1' linked dimer, pestalazine C (**44**),[28] employed a chiral phosphoric acid (**46**) to set the stereochemistry of the dimerization in the the C–N bond forming reaction (Figure 1-8). Diketopiperazine **40** was subjected to the optimized azo-coupling conditions with chiral phosphoric acid **46**, which gave azo-compound **41** in greater than 20:1 d.r. in 79% yield. The N–N bond was reductively cleaved in good yield to give amine **42**. **42** was then employed in arylation conditions using diazonium nucleophile **45**, and the dimeric

scaffold was completed as the second DKP core was appended via Pd-catalyzed Larock annulation with alkyne **47**, as seen in the synthesis of C3-C6' linked dimers (+)-naseseazines A as well as the synthesis of C6-N1' dimer (–)-aspergilazine A (**13**) by Reisman *et. al*.[29]

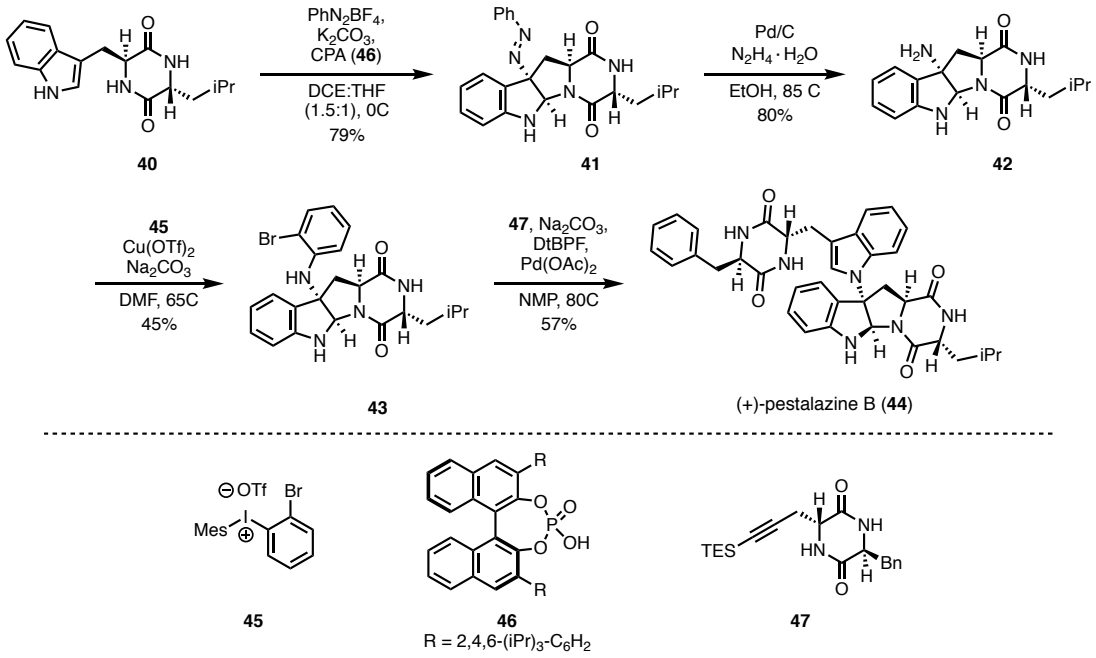


Figure 1-8. Liao's total synthesis of (+)-pestalazine B via azo-coupling, amination, and indole synthesis.[28]

Highlights of this synthesis include the development and utilization of robust chemistry which requires no protecting group manipulations of the reactive amine and amide functional groups, and generation of the C–N dimerization axis in high diastereoselectivity. However, this methodology is exclusively applicable to the formation of C–N bond-linked DKP dimers, and as the northern tryptophan indole is synthesized at a late stage via alkynyl DKP **47**, cannot take advantage of low-cost, and commercially available protected proteinogenic amino acids as chiral pool starting materials. Biomimetic syntheses that utilize commercially available amino acids as starting materials to construct the diketopiperazine and form the C–N bond at a late-stage are known[30]. However these syntheses require protecting and steering groups to guide selective C–N bond formation.

1.5 Representative Total Synthesis of C3-C7' Linked Diketopiperazine Dimers

The recent development of cross-coupling methodology to forge sp^3-sp^2 carbon-carbon bonds has been applied to the synthesis of the C3-C7' linked family of diketopiperazine dimers.

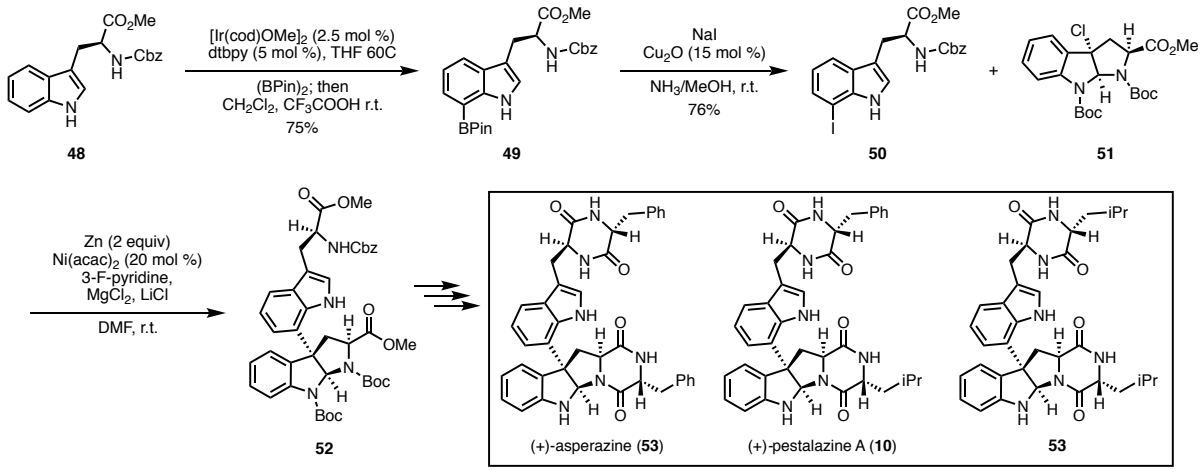


Figure 1-9. Collective total synthesis of C3-C7' linked dimers **53**, **10** and **54**, by Gong *et. al.* [31]

The route developed by Gong *et. al.* (Figure 1-9) is extremely efficient, as it takes advantage of amino acid chiral pool starting materials, and utilizes contemporary C–H functionalization methods to install the C7 boronic ester, and nickel-catalyzed sp^2-sp^3 bond cross-electrophile coupling to generate the C3-C7'dimeric axis.[31] However, similar to other dimerization methods, the stereochemistry of the dimerization is set early in the halocyclization and resulting stereochemistry of tertiary chloride **51**. Starting from amino acid derivatives **48** and **51**, allows for diversification of dimeric core **52**, however this strategy requires the use of multiple orthogonal protecting groups, and cannot be employed in dimerizing diketopiperazines directly. Biomimetic approaches which directly couple diketopiperazines and form the C3-C7' carbon-carbon bond at a late stage are known,[32] however these syntheses also suffer from the use of protecting groups and the stereochemical outcome of dimerization is substrate–controlled.

1.6 Representative Total Syntheses of C3-C6' Linked Diketopiperazine Dimers

The (+)-naseseazines A (**11**) and B (**48**) are the first diketopiperazine dimers isolated from bacterial sources.[15] In their retrobiosynthetic analysis, Movassaghi *et. al* proposed that these monomers are likely assembled via nucleophilic addition to an enzymatically generated carbocation (Figure 1-10).[33]

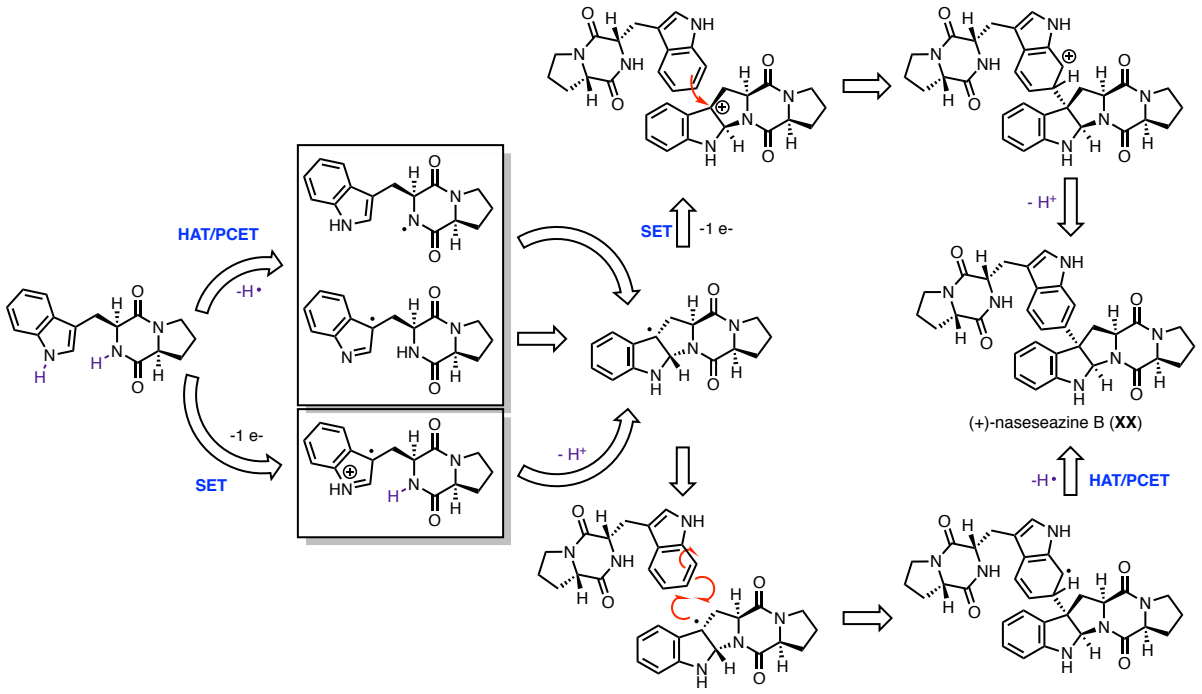


Figure 1-10. Possible mechanisms of biosynthetic dimerization described by Movassaghi *et. al*.[33]

To mimic this proposed biosynthetic hypothesis, access these non-symmetrical dimers, Movassaghi *et. al.* developed a convergent approach using a silver-mediated ionization of Cbz-protected tertiary bromide **55** in the presence of Cbz- protected diketopiperazine **54,** resulting in a near 1:1 mixture of constitutional isomers **56** and **57** in 53% yield. **56** and **57** could be deprotected to give (+)-naseseazine B (**58**) as well as (+)-*iso*-naseseazine B (**59**). This synthesis was applied to heterodimer (+)-naseseazine A (**11**) as well (Figure 1-11).

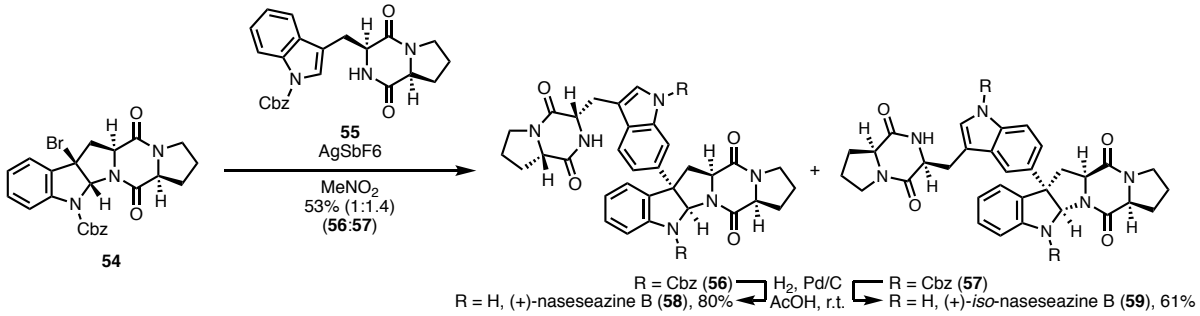


Figure 1-11. First generation biomimetic synthesis of (+)-naseseazine B (**58**) and (+)-*iso*-naseseazine B (**59**)[33]

To improve the selectivity site-selectivity, a second generation synthesis employed electrofugal trifluoroborate **60,** to build up charge density at the desired nucleophilic carbon to direct C–C bond formation. This resulted in complete control of site selectivity during dimerization (50% yield), and dimers could again be deprotected to give desired natural products **58** and **11** (Figure 1-12).

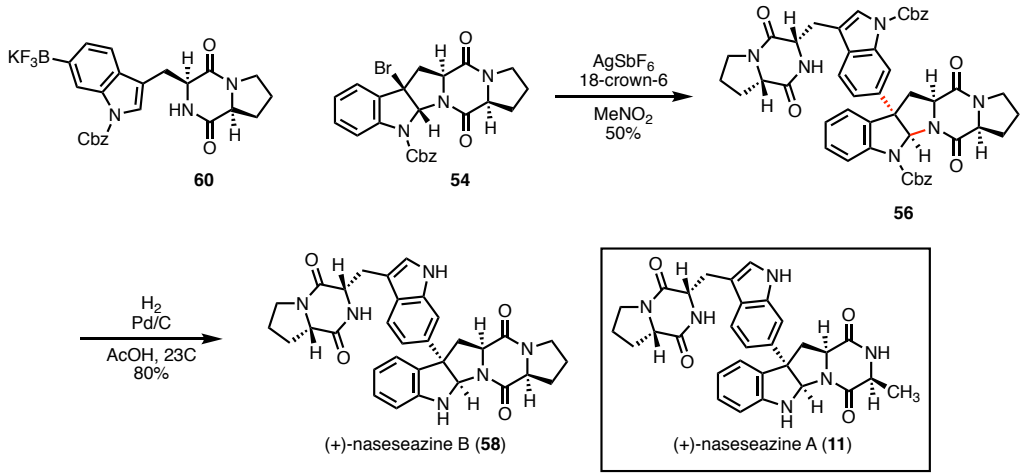


Figure 1-12. Second-generation, site-selective biomimetic syntheses of (+)-naseseazines A (**11**) and B (**58**).[33]

The use of trifluoroborates to direct C–C bond formation enabled complete selectivity in the dimerization reaction. However, this strategy requires pre-functionalization of a single monomer which then predetermines the connectivity of dimerization, and as in previous syntheses the stereochemistry of dimerization is preset by configuration of tertiary bromide **54.**

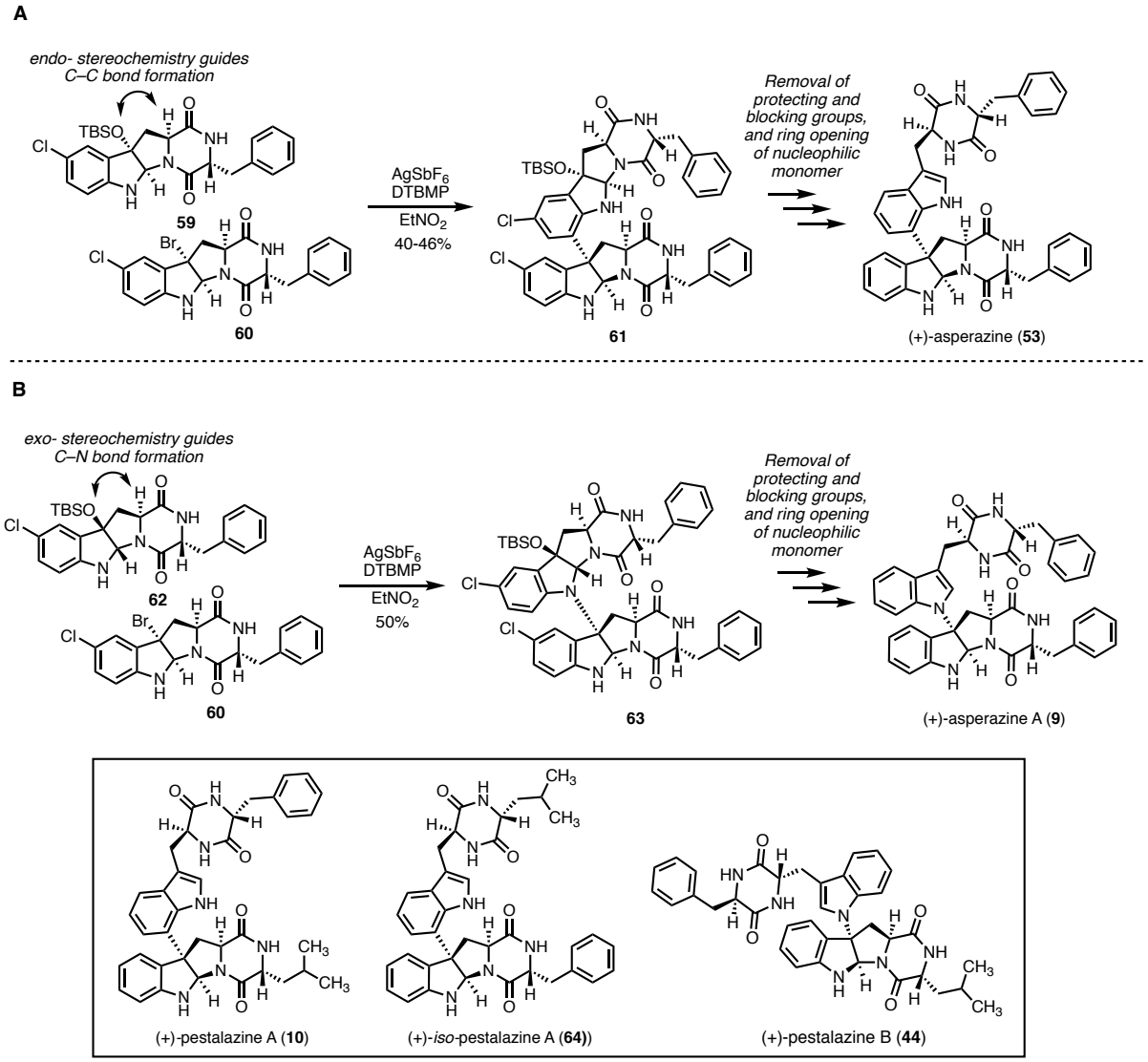


Figure 1-13. Biomimetic synthesis of C3-C7' (+)-asperazine (**53**), (+)-*iso*-pestalazine A (**64**) (+)-pestalazine A (**10**), and C3-N1' linked (+)-asperazine A (**9**) and (+)-pestalazine B (**44**) by Movassaghi *et. al.*[30, 32]

As referenced, this biomimetic Friedel-Crafts approach was also utilized in the synthesis of unsymmetrical dimers C3-C7' linked (+)-asperazine (**53**), (+)-*iso*-pestalazine A (**64**), and (+)-pestalazine A (**10**) and C3-N1' and (+)-asperazine A (**9**) and (+)-pestalazine B (**44**) (Figure 1-13). However, the use of trifluoroborates did not prove to be a general strategy for guiding bond formation, and the transient use of blocking and steering functionalities were required for selective bond formation. Transiently cyclized of nucleophilic monomers **59** and **60** were utilized to direct bond formation, and

as in the synthesis of the (+)-naseseazines, the configuration of dimerization was set by stereochemistry of tertiary bromide **60**.

1.7 Total Syntheses of C6-N1' Linked Diketopiperazine Dimer, (–)-Aspergilazine A (13)

Unique among indole linked DKP dimers, (–)-aspergilazine A (**13**) is unique as neither monomer is cyclized, and the axis of dimerization is at two sp^2 hybridized atoms (C6–N1'). Reisman *et. al*. were able to rapidly assemble this dimer via leveraging the innate pseudo-symmetry of this homodimer, and a utilization of a single diketopiperazine precursor (**68**) (Figure 1-14)[29]. Selective Pd-catalyzed amination of 2-bromoiodobenzene **65**, with bromoaniline **66**, gave diarylamine **67**. Larock indole synthesis utilized with 3 equivalents of **68**, and *in situ* desilation post annulation gave (–)-aspergilazine A (**13)** in 45% over two steps.

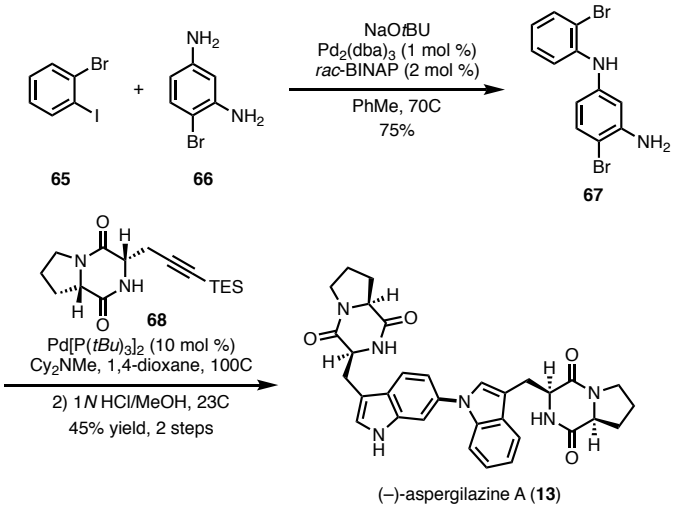


Figure 1-14. Total synthesis of (–)-aspergilazine A via selective amination and Larock annulation to install both DKP cores by Reisman *et. al.*[29]

1.8 Total Synthesis of Putative Oxidatively-Tailored Indole Linked Dimer (+)-gliocladin C (77)

In the isolation of C3-C3' linked, C2-symmetrical dimers such as the (+)-verticillins ((+)-verticillin, **3**) and (+)-leptosins, structurally similar molecules such as gliocladine C (**68**) are often co-isolated.[34] In these isolation reports, treatment of **3** with

strong base, generates **68** in an unspecified yield. This led to the hypothesis that **3** and **68** may be biosynthetically related, and **3** may be generated via a retro-aldol type reaction, cleaving off a diketopiperazine core and re-aromatizing the northern indole (Figure 1-15, A). Recently an example of a retro-aldol reaction in the biosynthesis of diketopiperazine-containing fungal metabolite, aspirochlorine (**72**) has been demonstrated, lending credence to this biosynthetic proposal (Figure 1-15, B).[35-36]

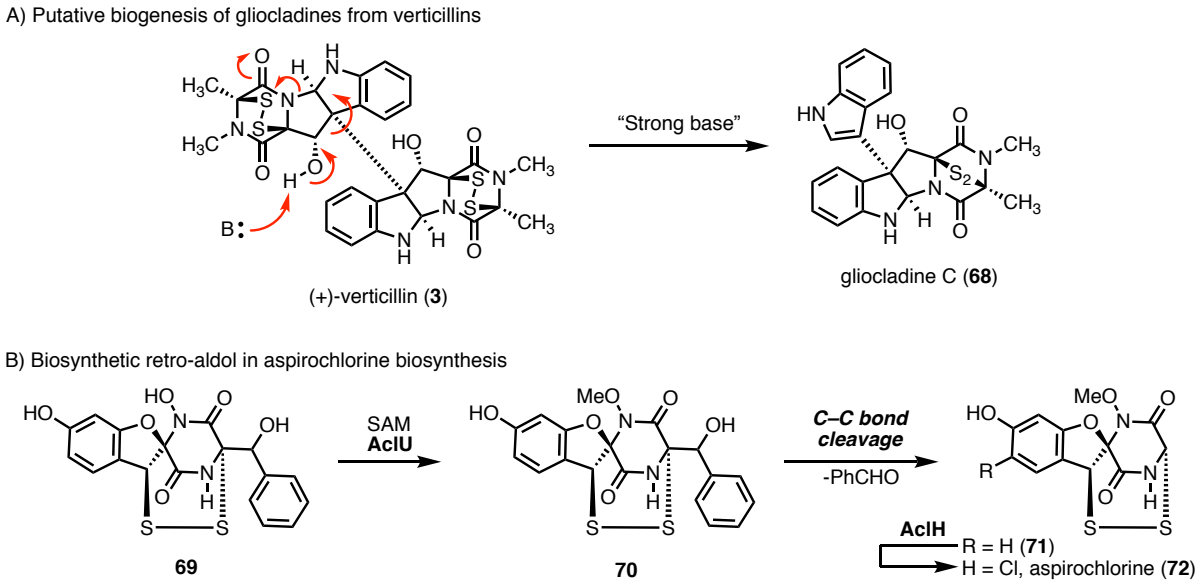


Figure 1-15. A) Chemical degradation of **3** to **68**, suggesting the biosynthetic relationship between these two co-isolates, B) Recently discovered retro-aldol in biosynthesis of fungal metabolite aspirochlorine (**72**). [34]

Stephenson *et. al.* used visible light photocatalysis to assemble the core structure of related compound, (+)-gliocladin C (**77**) in a biomimetic radical arylation.[37] Boc-protection and bromocyclization of protected tryptophan **73**, gave tertiary bromide **74** in 91% yield. After amide formation with methylamine (87% yield), under conditions developed for radical dehalogenation and arylation of alkyl halides,[38] **75** was coupled to indole-2-carboxaldehyde in high yield (82%). This core structure was then further elaborated into (+)-gliocladin C.

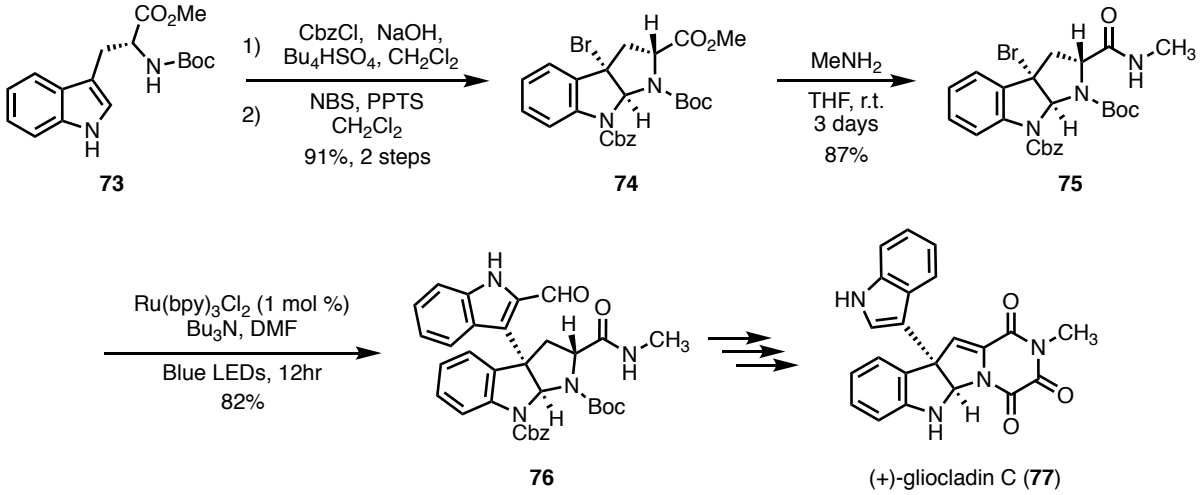


Figure 1-16. Total synthesis of (+)-gliocladin C (**77**) enabled by photoredox catalysis by Stephenson *et. al*.[37]

1.9 Investigations In Fungal Biosynthesis of DKP Dimers

In 2013, Walsh *et. al.* published the first report disclosing the biosynthetic gene cluster for unsymmetrical ETP dimer (+)-chaetocin.[39] Genetic disruption and accumulation of intermediates in the gliotoxin biosynthetic pathway identified key genes implicated in the assembly of the DKP core (GliP, non-ribosomal peptide synthetase (NRPS)).[40] Genome assembly from Illumina sequencing of total nucleic acids isolated from (+)-chaetocin producer, *Chaetomium virescens* (ATCC 26417), was queried using GliP as a probe gene to identify putative (+)-chaetocin biosynthetic gene clusters. This resulted in discovery of putative NRPS, ChaP and primers were designed to screen a cosmid library prepared from *C. viriscens* genomic DNA, and lead to the identification of a 38 kb cluster of 14 genes putatively involved in (+)-chaetocin biosynthesis (Table 1-1).

Gene	Accession # of closest homologue	Proposed function	Similarity (E value)
chaM	XM_001258084	methyltransferase	6×10^{-118}
chaE	XM_001933514	cytochrome P450 ; dimerization	6×10^{-72}
chaT	XM_001226009	thioredoxin reductase ; dithiol formation	3×10^{-91}
chaZ	XM_002375564	regulation of chaetocin gene cluster expression	9×10^{-13}
chaB	XM_002842656	cytochrome P450 ; pyrroloindole formation	2×10^{-40}
chaA	XM_003066250	ABC transporter ; chaetocin efflux	0
chaG	AY838877	glutathione S-transferase	3×10^{-29}

chaJ	XM_001258081	dipeptidase	4×10^{-100}
chaK	XM_001258086	g-glutamylcyclotransferase	3×10^{-52}
chaP	XP_750855	nonribosomal peptide synthase	0
chaC	AY838877	cytochrome P450	9×10^{-99}
chaI	AY838877	PLP-dependent carbon–sulfur bond lyase	5×10^{-77}
chaN	XM_002846230	methyltransferase	5×10^{-48}

Table 1-1. Putative (+)-chaetocin biosynthetic gene cluster from *Chaetomium virescens* (ATCC 26417).[39]

Comparison of this putative cluster with known gliotoxin cluster revealed an additional cytochrome P450, ChaE, which the authors attributed to the dimerization of DKP monomers into the (+)-chaetocin dimeric scaffold however were unable to heterologously express and characterize this enzyme, and no genetic system for *Chaetomium* fungi had been described at that point.

In 2014, Watanabe *et. al.* were also able to uncover the biosynthetic gene cluster for C_2-symmetrical dimer (–)-ditryptophenaline (**25**) from *Aspergillus flavus* through disruption of the NRPS, DtpA.[41] This gene cluster consisted of an NRPS (DtpA), N-methyltransferase (DtpB), and cytochrome P450 (DtpC). Watanabe *et. al.* established the role each of these genes through disruption *in vivo*, as well as reconstitution of the native enzymatic activity of DtpC *in vitro* through incubation of biosynthetic DKP precursor **78** with microsomal fractions of S. cerevisiae expressing DtpC to yield (–)-ditryptophenaline, **25**. Wild-type DtpC displayed modest biosynthetic flexibility, and when DtpC containing microsomal fractions were incubated with non-cognate diketopiperazines, known natural products **80** and **81** were obtained and the authors were able to use this approach to generate milligram quantities of these dimers (5.6mg of **25**, 2.0mg of **80** and 1.6mg of **81**).

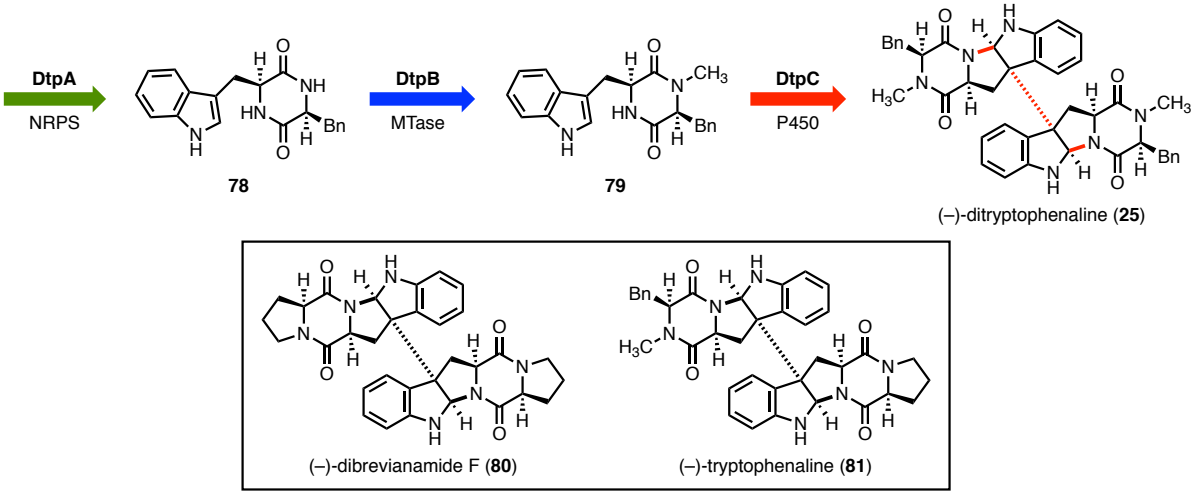


Figure 1-17. (–)-Ditryptophenaline **25** gene cluster and compounds generated through microsomal reactions. [41]

Recently, the biosynthetic gene cluster for related C_2-symmetric dimer (+)-verticillin A (**3**) as also discovered from *Clonostachys rogersoniana* and disruption of putative NRPS, *verA,* abolished production of (+)-verticillin (**3**) *in vivo*.[42] In a survey of biosynthetic gene clusters from *Aspergillus tubingensis* G131, a known producer of (+)-asperazine (**53**), a putative gene cluster for non-symmetrical was also disclosed.[43]

1.10 Specific Aims

The structural diversity and biological activity of dimeric DKPs are very diverse and dependent on stereochemistry and site of dimerization. In spite of the pharmaceutical potential of these compounds, the majority of in-depth biological study is are from dimeric DKPS containing the epidisulfide bridge from the epipolythiodiketopiperazine (ETP) family, likely do to the high titers from fermentation and commercial availability of members of this class. However, as noted in isolations of the non-thiolated dimers, the lack of their availability of these materials inhibits a full examination of their pharmaceutical potential.[13] Development of methods for total synthesis of these molecules has forwarded the ability assess this potential, as well as structure activity relationships, and although there are >20 total syntheses for dimeric diketopiperazines, all of these syntheses are substrate controlled, and to date there exists no method which allows for selective construction of two isomers of different connectivity and stereochemistry from the same monomeric starting materials.

The ultimate goal of this biosynthetic study is to uncover the biosynthetic machinery involved in the bacterial biosynthesis of diketopiperazine dimers, and to utilize these enzymes as biocatalysts for the synthesis and biological evaluation of known diketopiperazine dimer and novel derivatives. The focus of this dissertation is the site- and stereoselectivity of enzymes involved in the selective dimerization of diketopiperazines. Bioinformatic analysis of multiple *Streptomyces* revealed that in the biogenesis of unsymmetrical diketopiperazines dimers are derived from a diketopiperazine generated by a cyclodipeptide synthase (CDPS) which are oxidatively dimerized by a cytochrome P450 (P450). Thus, through heterologous gene expression and *in vitro* enzymatic assays, we propose to: a) identify key enzymes involved in site - and stereoselective assembly of non-symmetrical diketopiperazine dimers and; b) to probe the catalytic mechanism and selectivity of these enzymes through X-ray crystallography and site-directed mutagenesis.

1.11 References

1. Waksman, S. A.; Bugie, E., *Journal of Bacteriology* **1944,** *48* (5), 527-530.
2. Geiger, W. B.; Conn, J. E.; Waksman, S. A., *Journal of Bacteriology* **1944,** *48* (5), 531-536.
3. Safe, S.; Taylor, A., *J Chem Soc Perkin 1* **1972,** *4*, 472-9.
4. McInnes, A. G.; Taylor, A.; Walter, J. A., *J Am Chem Soc* **1976,** *98* (21), 6741.
5. Hauser, D.; Weber, H. P.; Sigg, H. P., *Helv. Chim. Acta* **1970,** *53*, 1061-1073.
6. Katagiri, K.; Sato, K.; Hayakawa, S.; Matsushi.T; Minato, H., *J Antibiot* **1970,** *23* (8), 420-&.
7. Woodward, R. B.; Yang, N. C.; Katz, T. J., *P Chem Soc London* **1960,** (2), 76-78.
8. Robinson, R.; Teuber, H. J., *Chem Ind-London* **1954,** (27), 783-784.
9. Kim, J.; Movassaghi, M., *Acc. Chem. Res.* **2015,** *48* (4), 1159-71.
10. Jiang, C. S.; Guo, Y. W., *Mini-Rev Med Chem* **2011,** *11* (9), 728-745.
11. Borthwick, A. D., *Chem. Rev.* **2012,** *112* (7), 3641-716.
12. Ovenden, S. P. B.; Sberna, G.; Tait, R. M.; Wildman, H. G.; Patel, R.; Li, B.; Steffy, K.; Nguyen, N.; Meurer-Grimes, B. X., *Journal of Natural Products* **2004,** *67* (12), 2093-2095.
13. Varoglu, M.; Corbett, T. H.; Valeriote, F. A.; Crews, P., *J Org Chem* **1997,** *62* (21), 7078-7079.
14. Ding, G.; Jiang, L.; Guo, L.; Chen, X.; Zhang, H.; Che, Y., *J. Nat. Prod.* **2008,** *71* (11), 1861-5.
15. Raju, R.; Piggott, A. M.; Conte, M.; Aalbersberg, W. G.; Feussner, K.; Capon, R. J., *Org. Lett.* **2009,** *11* (17), 3862-5.
16. Cai, S.; Kong, X.; Wang, W.; Zhou, H.; Zhu, T.; Li, D.; Gu, Q., *Tet. Lett.* **2012,** *53* (21), 2615-2617.
17. Sun, S. S.; Ma, K.; Tao, Q. Q.; Han, J. J.; Bao, L.; Liu, L.; Liu, H. W., *Fitoterapia* **2018,** *125*, 266-272.
18. Song, F. H.; Liu, X. R.; Guo, H.; Ren, B.; Chen, C. X.; Piggott, A. M.; Yu, K.; Gao, H.; Wang, Q.; Liu, M.; Liu, X. T.; Dai, H. Q.; Zhang, L. X.; Capon, R. J., *Organic Letters* **2012,** *14* (18), 4770-4773.
19. Isaka, M.; Palasarn, S.; Rachtawee, P.; Vimuttipong, S.; Kongsaeree, P., *Organic Letters* **2005,** *7* (11), 2257-2260.
20. Li, G. Y.; Yang, T.; Luo, Y. G.; Chen, X. Z.; Fang, D. M.; Zhang, G. L., *Org. Lett.* **2009,** *11* (16), 3714-7.
21. Geng, C. A.; Huang, X. Y.; Ma, Y. B.; Hou, B.; Li, T. Z.; Zhang, X. M.; Chen, J. J., *Journal of Natural Products* **2017,** *80* (4), 959-964.
22. Hino, T.; Kodako, S.; Takahashi, K.; Yamaguchi, H.; Masako, N., *Tet. Lett.* **1978,** *19* (49), 4913-4916.
23. Nakagawa, M.; Sugumi, H.; Kodato, S.; Hino, T., *Tet. Lett.* **1981,** *22* (52), 5323-5326.
24. Bagal, S. K.; Adlington, R. M.; Baldwin, J. E.; Marquez, R.; Cowley, A., *Org. Lett.* **2003,** *5* (17), 3049-52.
25. Bagal, S. K.; Adlington, R. M.; Baldwin, J. E.; Marquez, R., *J Org Chem* **2004,** *69* (26), 9100-8.

26. Movassaghi, M.; Schmidt, M. A.; Ashenhurst, J. A., *Angew Chem Int Ed Engl* **2008,** *47* (8), 1485-7.
27. Liang, K.; Deng, X.; Tong, X.; Li, D.; Ding, M.; Zhou, A.; Xia, C., *Org. Lett.* **2015,** *17* (2), 206-9.
28. Li, Q.; Xia, T.; Yao, L.; Deng, H.; Liao, X., *Chem Sci* **2015,** *6* (6), 3599-3605.
29. Chuang, K. V.; Kieffer, M. E.; Reisman, S. E., *Org. Lett.* **2016,** *18* (18), 4750-3.
30. Nelson, B. M.; Loach, R. P.; Schiesser, S.; Movassaghi, M., *Org. Biomol. Chem.* **2018,** *16* (2), 202-207.
31. Luo, L.; Zhai, X. Y.; Wang, Y. W.; Peng, Y.; Gong, H., *Chemistry* **2018**.
32. Loach, R. P.; Fenton, O. S.; Movassaghi, M., *J .Am. Chem. Soc.* **2016,** *138* (3), 1057-64.
33. Kim, J.; Movassaghi, M., *J Am Chem Soc* **2011,** *133* (38), 14940-3.
34. Katagiri, K.; Sato, K.; Hayakawa, S.; Matsushima, T.; Minato, H., *J Antibiot (Tokyo)* **1970,** *23* (8), 420-2.
35. Chankhamjon, P.; Boettger-Schmidt, D.; Scherlach, K.; Urbansky, B.; Lackner, G.; Kalb, D.; Dahse, H. M.; Hoffmeister, D.; Hertweck, C., *Angew Chem Int Ed Engl* **2014,** *53* (49), 13409-13.
36. Tsunematsu, Y.; Maeda, N.; Yokoyama, M.; Chankhamjon, P.; Watanabe, K.; Scherlach, K.; Hertweck, C., *Angew Chem Int Ed Engl* **2018,** *57* (43), 14051-14054.
37. Furst, L.; Narayanam, J. M.; Stephenson, C. R., *Angew Chem Int Ed Engl* **2011,** *50* (41), 9655-9.
38. Staveness, D.; Bosque, I.; Stephenson, C. R., *Acc Chem Res* **2016,** *49* (10), 2295-2306.
39. Gerken, T.; Walsh, C. T., *Chembiochem* **2013,** *14* (17), 2256-8.
40. Balibar, C. J.; Walsh, C. T., *Biochemistry* **2006,** *45* (50), 15029-38.
41. Saruwatari, T.; Yagishita, F.; Mino, T.; Noguchi, H.; Hotta, K.; Watanabe, K., *Chembiochem* **2014,** *15* (5), 656-9.
42. Wang, Y.; Hu, P.; Pan, Y.; Zhu, Y.; Liu, X.; Che, Y.; Liu, G., *Fungal Genet. Biol.* **2017,** *103*, 25-33.
43. Choque, E.; Klopp, C.; Valiere, S.; Raynal, J.; Mathieu, F., *BMC Genomics* **2018,** *19* (1), 200.

Chapter 2

Discovery and Characterization of NascB and NznB, Site- and Stereodivergent Diketopiperazine Dimerizing Cytochromes P450

2.1. Abstract

In this chapter, we describe the identification of two biosynthetic gene clusters from *Streptomyces sp.* CMB-MQ030 encoding the biosynthesis of the (+)-naseseazines, *nzn* gene cluster) as well as regio- and stereoisomeric cyptic natural products, the (–)-naseseazines (*nasc* gene cluster). Both gene clusters are composed of a cyclodipeptide synthase (CDPS) and cytochrome P450. The products of the cyclodipeptide synthases were characterized through *in vivo* expression and isolation of DKPs from the media. The P450s from these clusters were expressed and characterized *in vitro* as the first cytochrome P450s characterized which catalyze site- and stereodivergent C–H functionalization in a carbon-carbon and carbon-nitrogen bond forming cascade. This work was enabled through great help from Dr. Yogan Khatri who assisted in optimizing expression of the cytochromes described in this chapter, as well as Dr. Fengan Yu who initially annotated the gene clusters.

2.2. Introduction

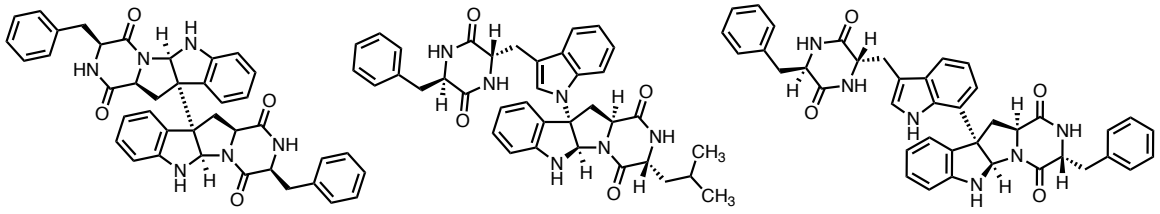
Dimeric diketopiperazines are structurally complex natural products whose biological activities are intimately linked to the connectivity of the two diketopiperazine (DKP) monomers.[1-3] While the isolation of DKP dimers from filamentous fungi dates back nearly fifty years,[4] their corresponding biosynthetic gene clusters have only been recently identified through genome sequencing, gene disruption, and heterologous expression.[5-7] Preliminary characterization of these fungal pathways indicates that biosynthetic cytochromes P450 mediate the oxidative convergence of two DKP monomers to generate C_2 symmetrical dimers, validating long-standing hypotheses

proposed in isolation, characterization, and total syntheses of these molecules.[7] However, direct biochemical studies of cytochromes P450s that catalyze formation of microbial-derived DKP dimers has been stymied due to the lack of suitable protein expression systems for their direct study *in vitro*.

Biological activity and exigent structural features native to C_2 symmetrical dimeric DKPs, such as the sterically encumbered vicinal chiral quaternary carbons about the C_2 symmetry axis, elicited development of elegant synthetic methodologies towards the total synthesis of these compounds. Among them, biomimetic strategies have enabled efficient access to these natural products as two monomers may be identically functionalized and coupled at a late stage resulting in highly convergent syntheses. The selective chemical synthesis of the exotic non-symmetrical DKP dimers pose a comparable synthetic challenge as they require independent synthesis of two dissimilar subunits, each with appropriate functionalization to guide site- and stereospecific formation of the desired C–C or C–N bond upon dimerization. [8-10]

These non-symmetrical dimers were previously exclusively isolated from fungi,[11-13] however marine Actinobacteria have recently emerged as a new source for these rare non-symmetrical dimers, and notably all DKP dimers isolated from bacterial sources to date have lacked symmetry elements (Figure 2-1).[14-16]

Representative Fungal Diketopiperazine Dimers

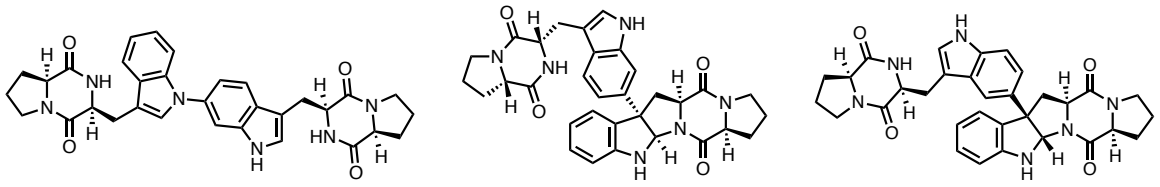


(+)-WIN 64821 (**5**)
Aspergillus sp. SC319

(+)-pestalazine B (**6**)
Pestalotiopsis theae (W148)

(+)-asperazine (**7**)
Aspergillus niger

Bacterial Diketopiperazine Dimers Isolated to Date



(–)-aspergilazine (**3**)
Aspergillus taichungensis (ZHN-7-07)
Streptomyces sp. SMA-1

(+)-naseseazine B (**1**)
Streptomyces sp. CMB-MQ030

(–)-naseseazine C (**4**)
Streptomyces sp. USC-636

Figure 2-1. Representative examples of fungal diketopiperazine dimers and bacterial diketopiperazine dimers (include co-isolation of dimers).

Furthermore, there are multiple reports of co-isolation of constitutional isomers of a given DKP dimer from the same organism, giving rise to two possible biosynthetic hypotheses: 1) a single oxidase has flexibility in dimerization and generates multiple products as a result of active site plasticity, or 2) each constitutional outcome is the result of a discrete oxidase with a high degree of selectivity. Identification of the bacterial biosynthetic enzymes responsible for non-symmetrical DKP dimer biogenesis may provide ready access to synthetically useful biocatalysts for selective assembly of this challenging class of natural products, and offers new opportunities to study late-stage enzymatic C–H functionalization and C–C bond formation.

2.3 Research and Discussion

Streptomyces sp. CMB-MQ030 was the first bacterial strain from which the unsymmetrical DKP dimers (+)-naseseazines A (**1**) and B (**2**) were isolated.[14] This strain was obtained, sequenced and mined for non-ribosomal peptide synthetase (NRPS) containing gene clusters based on known fungal biosynthetic systems.[7] This effort failed to reveal any relevant NRPS-driven systems containing cytochrome P450s, however we identified three cyclodipeptide synthase (CDPS) driven candidate biosynthetic gene clusters (BGCs), each containing a cytochrome P450 from which the (+)-naseseazines A (**2**) and B (**1**) may be derived from (Figure 2-2). Based on these biosynthetic gene clusters, as well as structural analysis of **1** and **2**, we proposed the diketopiperazine monomers are biosynthesized by a CDPS followed by dimerization by a downstream cytochrome (Figure 2-3).

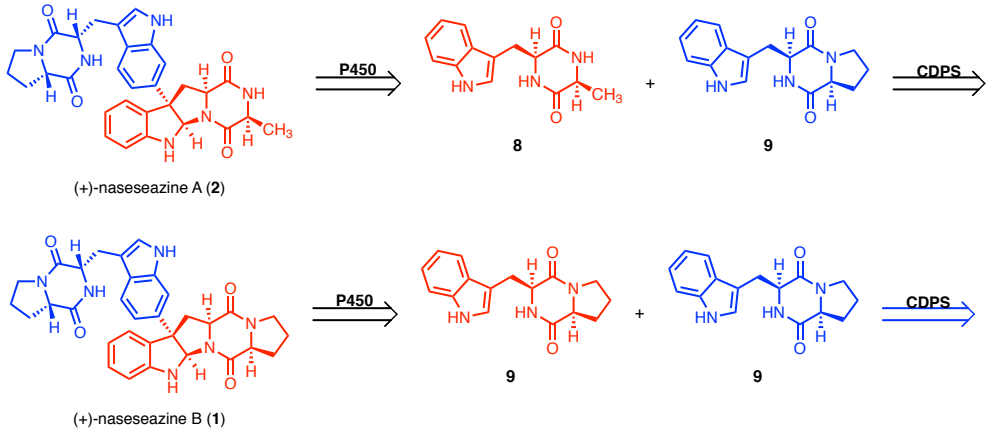
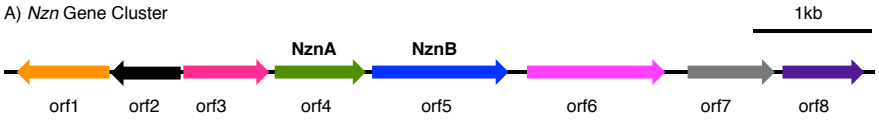


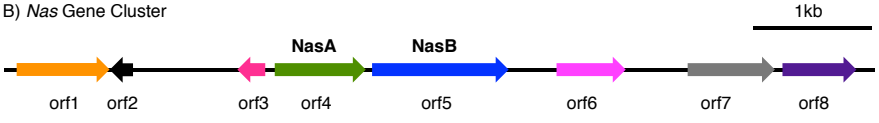
Figure 2-2. Retrobiosynthetic hypothesis for (+)-naseseazines A (**2**) and B (**1**)

A) *Nzn* Gene Cluster



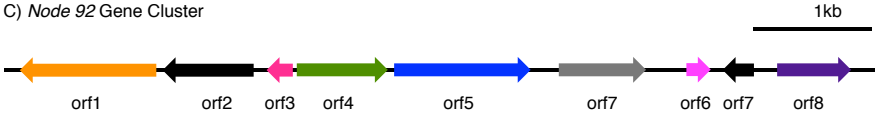
Frame	NCBI Homologue	Identity/Similarity
orf1	Dehydrogenase [*S. marina*]	(76/84)
orf2	Short-chain dehydrogenase [*S. griseoaurantiacus*]	(82/89)
orf3	XRE family transcriptional regulator [*N. brasiliensis*]	(45/58)
orf4 (NznA)	AlbC [*Nocardiopsis sp.* CNS639]	(35/48)
orf5 (NznB)	Cytochrome P450 [*A. mirum*]	(36/49)
orf6	Multi-drug efflux MFS permease [*R. ruber*]	(48/66)
orf7	F420-dependent oxidoreductase [*A. marinum*]	(52/65)
orf8	GntR family transcriptional regulator [*S. albulus*]	(83/91)

B) *Nas* Gene Cluster



Frame	NCBI Homologue	Identity/Similarity
orf1	Serine/threonine protein kinase [S. acidiscabies]	(37/47)
orf2	Hypothetical protein	(82/89)
orf3	Alpha amylase [*S. cryophillis*]	(45/58)
orf4 (NasA)	AlbC [*Nocardiopsis sp.* CNS639]	(31/43)
orf5 (NasB)	Cytochrome P450 CYP121 [*S. rimosus*]	(34/49)
orf6	LuxR transcription regulator [*S. flavogriseus*]	(83/89)
orf7	Histidine kinase [*S. griseoaurantiacus*]	(52/65)
orf8	CAAX amino terminal protease [*S. griseoaurantiacus]*	(66/78)

C) *Node 92* Gene Cluster



Frame	NCBI Homologue	Identity/Similarity
orf1	D-Ala-D-Ala carboxypeptidase [*A. vancoresmycina*]	(53/67)
orf2	LuxR regulator [*S. rimosus*]	(88/92)
orf3	Histidine kinase [*Streptomyces* sp. TOR3209]	(45/58)
orf4	AlbC [*Nocardiopsis sp.* CNS639]	(39/55)
orf5	Cytochrome P450 CYP121 [*S. rimosus*]	(36/51)
orf6	N-methyl-transferase [*S. alboniger*]	(37/57)
orf7	Nitroreductase [*Nocardiopsis halotolerans*]	(51/67)
orf8	LacI-transcription regulator [S. venezuelae]	(66/75)

Figure 2-3. Cyclodipeptide synthase (CDPS) containing gene clusters.

Annotation of the surrounding genes in each CDPS gene cluster revealed that one cluster (Node 92) contained cyclic dipeptide oxidases (CDOs) upstream of the

CDPS, and as the (+)-naseseazines contain no exocyclic unsaturation,[17] we hypothesized that this gene cluster is likely not responsible for the biosynthesis of the naseseazines, leaving two gene clusters, *nasc* and *nzn*, both containing a CDPS and P450 within the same operon. Alignment of these two clusters showed high sequence homology for both the CDPS (Figure 2-4) and P450s (Figure 2-5) and an overall lack of homology to characterized homologues in the NCBI. In order to determine whether these gene clusters were responsible for the biosynthesis of the (+)-naseseazines the two CDPS, NascA and NznA, and respective P450s, NascB and NznB, were selecting for heterologous expression and characterization.

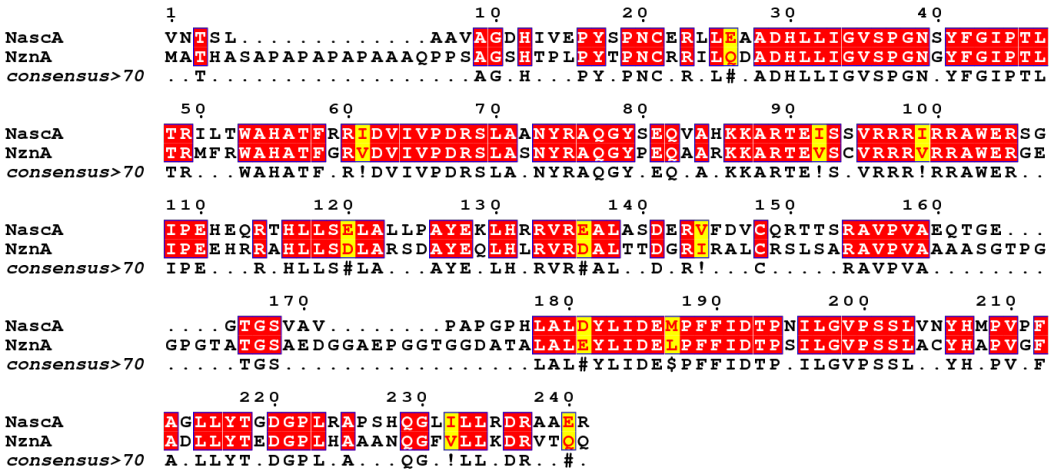


Figure 2-4. Sequence alignment for NascA and NznA CDPSs. Identical residues are highlighted in red, yellow are similar, and consensus sequence below.

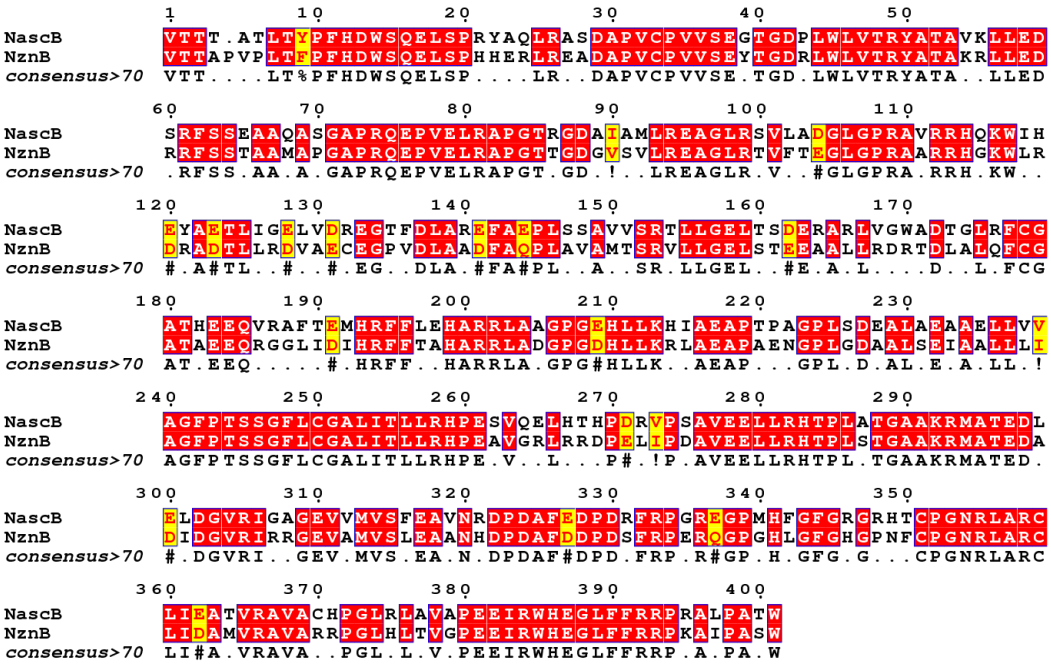


Figure 2-5. Sequence alignment of NascB and NznB P450s.

2.3.1 Identification of CDPS Products Through *In Vivo* Analysis

When heterologously expressed in *E. coli*, CDPSs utilize endogenous aminoacyl-tRNAs as substrates to generate DKPs which are secreted into the media via an unknown mechanism. Upon expression of NascA, **8** was identified as the sole metabolite of the fermentation culture. Based on our biosynthetic hypothesis, **8** is a precursor to both **1** and **2**, and as such we hypothesized that the *nasc* BCG may be responsible for the biosynthesis of the (+)-naseseazines. Expression of NznA, **8** and **9** were identified as secreted metabolites, presenting all the required precursors required for construction of **1** and **2**. However, while we found the P450 from the *nasc* gene cluster to be readily obtained through heterologous expression in *E. coli*, the cytochrome from the *nzn* cluster, NznB, was intransigent to this approach regardless of promoter, cell line, codon optimization, or co-expression with chaperones and characterization of the *nasc* gene cluster was pursued.

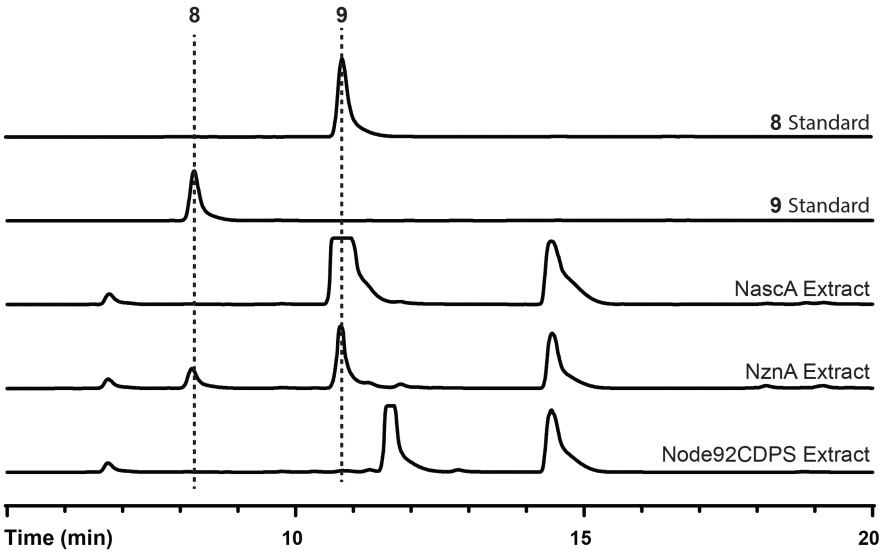


Figure 2-6. HPLC traces from extraction of CDPS expression media from NascA, NznA and Node 92 CDPS.

2.3.2 Characterization of Cytochrome P450, NascB

For purposes of scale, **9** was synthesized as previously described,[3, 8] and the binding of **9** to NascB was assessed via spectrophotometry. The interaction showed a clear low-spin to high-spin transition, characteristic of direct binding of the substrate to the enzyme active site affirming that this cytochrome P450 may .[18]

Incubation of **9** with NascB, ferredoxin and ferredoxin reductase from *Spinacia oleracea*[19-20], and NADPH led to formation of a single product whose mass corresponded to the predicted DKP dimer, (+)-naseseazine B (**1**), as assessed by LC-MS (Figure 2-7). To determine the structure of this dimeric product, a 10mg scale NascB-catalyzed dimerization was performed and crude products purified by preparatory HPLC for NMR analysis.

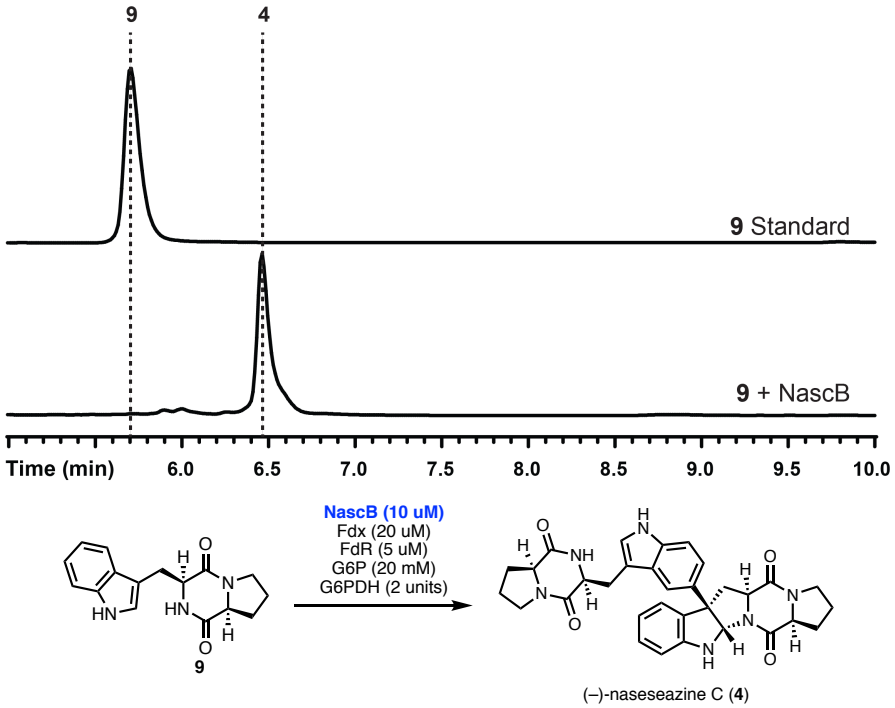


Figure 2-7. HPLC trace from NascB reaction with **9**.

Detailed evaluation of the ^{1}H NMR spectrum of the NasB catalyzed dimerization product of DKP **9** revealed significant differences with spectroscopic data for (+)-naseseazine B (**1**). These differences were also evident in the ^{13}C NMR and subsequent 2D NMR spectra. With mass spectrometry supporting the formation of a dimeric compound, we compared the NMR data against spectra from known dimers of **9** dimers.[7, 15-16, 21] The ^{1}H NMR data from the product of the NascB reaction aligned favorably with that reported for "new *iso*-naseseazine B",[15,22] the C3(*R*)-diastereomer of (+)-naseseazine B (**11**). However, the structure of "new *iso*-naseseazine B"[22] had been disputed in a later isolation report, which reassigned the dimerization axis from C3–C6' to the unprecedented C3–C5' linkage[23] and renamed the molecule naseseazine C (**4**) (Figure 2-8).[16] However, the use of different solvents in the reports describing the structure of the "new *iso*-naseseazine B (**11**)"[22] and naseseazine C (**4**) prevented their direct comparison and an unambiguous correlation to our NascB catalyzed dimerization product.

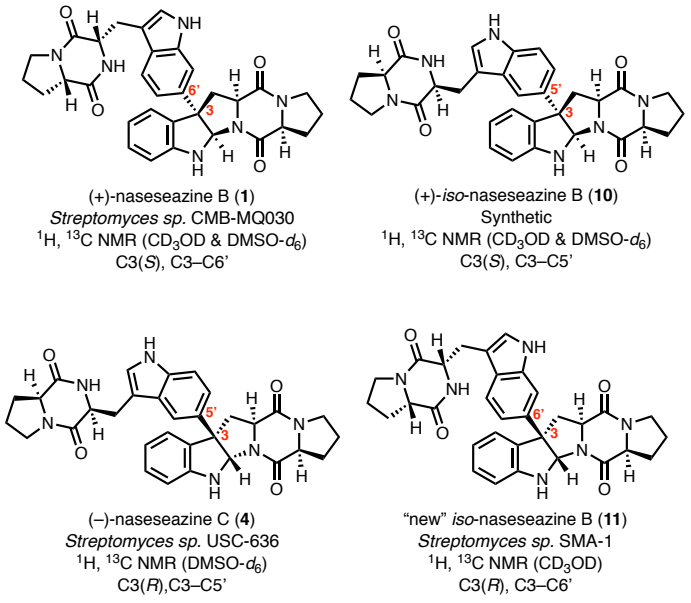


Figure 2-8. Structures of (+)-naseseazine B (**1**), synthetic (+)-*iso*-naseseazine B (**10**) proposed structure of "new *iso*-naseseazine B (**11**), and naseseazine C (**4**)

2.3.3 Total Synthesis of (–)-naseseazine C (4) and (–)-*iso*-naseseazine C (10)

To validate the structure of the NasB dimerization product, we opted to access both potential regioisomers recently reported as new naseseazines[15-16] through chemical total synthesis in collaboration with the Movassaghi lab (MIT). This enabled full structural assignment of these alkaloids in both methanol-d_4 and DMSO-d_6, and direct comparison with the new product **4** isolated from the NascB mediated dimerization of DKP **9**.

Exposure of (–)-**12** to silver hexafluoroantimonate in the presence of **13** led to a 1:1.4 (C5':C6') regioisomeric mixture of dimers (–)-**14** and (–)-**15** in 40% combined yield (Scheme 1). [24] Notably, the same ratio of regioisomers has been observed in the first generation synthesis of dimeric DKP alkaloids, (+)-naseseazine A (**2**) and (+)-naseseazine B (**1**).[8] Hydrogenolytic removal of the carboxybenzyl groups afforded both C5' and C6' constitutional isomers in 80% and 84% yield, respectively. All spectroscopic data for the C5' isomer **4** were consistent with the literature data for both (–)-naseseazine C (**4**) and "new iso-naseseazine B (**11**)."[22]

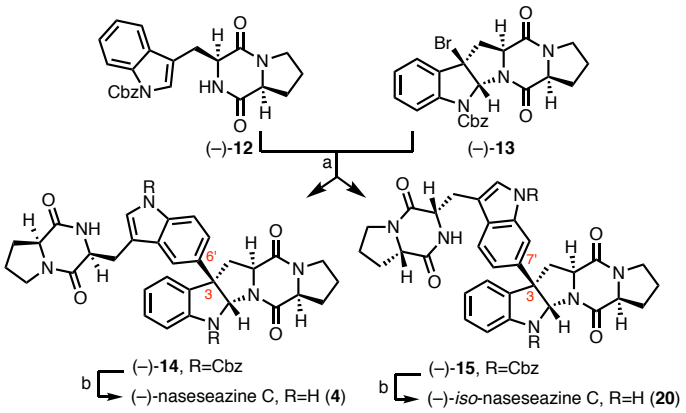


Figure 2-9. Total synthesis of (–)-naseseazine C (4) and (–)-isonaseseazine C (16)
(a), $AgSbF_6$, $EtNO_2$, (–)-**14**:(–)-**15** = 1:1.4, 40%. (b) H_2/Pd/C, EtOH, (80% for (–)-**4** and 84% for (–)-**16**.

Furthermore, the spectroscopic data for the C6' isomer termed *iso*-naseseazine C (**16**)[25] was not consistent with data reported for either (–)-naseseazine C or "new *iso*-naseseazine B". Based on these results, we confirm the relative and absolute stereochemistry of (–)-naseseazine C (**7**) as the C3(*R*)-DKP dimer with C3–C6' linkage. Additionally, we were able to revise the optical rotation for (–)-naseseazine C (observed: $[\alpha]_D^{23}$ = –180 (*c* = 0.16, MeOH); lit.: $[\alpha]_D^{23}$ = –18.6 (*c* = 0.004 MeOH)[16]). Notably the magnitude of the observed optical rotation for our synthetic sample of (–)-naseseazine C (**7**) is consistent with other members of this family of alkaloids.[9-10, 26] Importantly, with both regioisomers in hand and with confirmation of the structure of (–)-naseseazine C, we were able to validate its assignment as the direct product of the dimerization event of brevianamide F catalyzed by NasB.

Although (–)-naseseazine C (**4**) was not detected under the reported fermentation conditions from which (+)-naseseazines A (**1**) and B (**2**) were obtained, the isolation of isomeric diketopiperazine dimers from a single organism is not uncommon (Figure 2-10).[12-13] Based on this literature precedent, and the products of the CDPS, NznA, we hypothesized that the (+)-naseseazines may be produced by the *nzn* biosynthetic gene cluster and attempted to solubilize the encoded cytochrome P450, NznB.

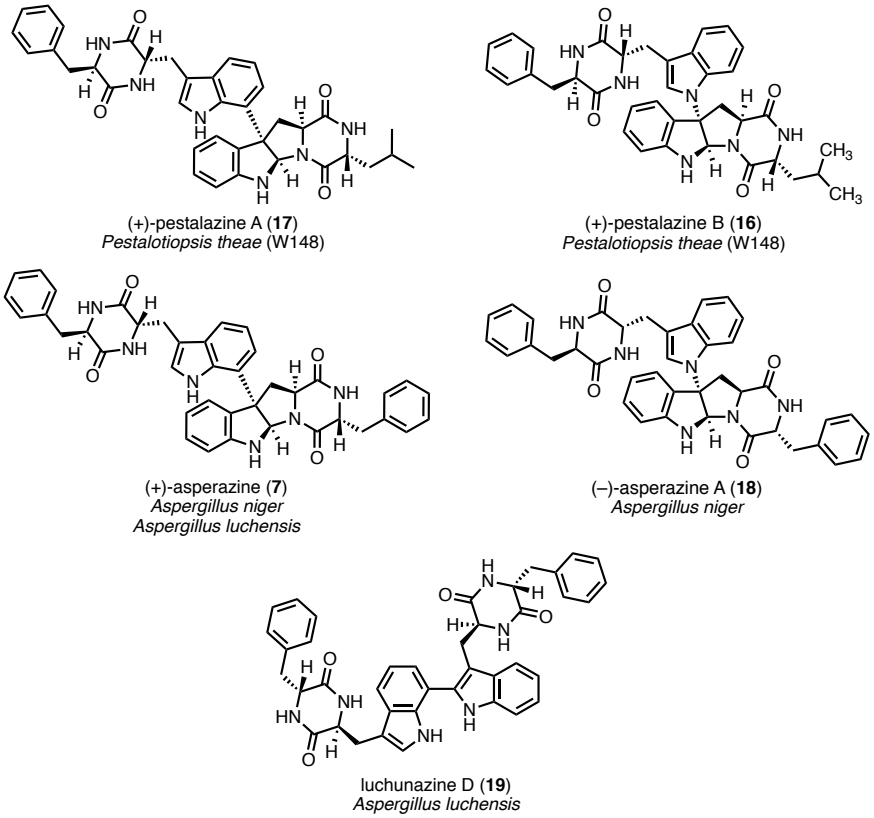


Figure 2-10. Examples of isomeric DKP dimers isolated from a single organism.

2.3.4 Expression and Characterization of Cytochrome P450, NznB

Alternative start sites for biosynthetic genes are well known in *Streptomyces*[27] and the start site for this NznB was annotated through a sequence alignment using NascB as a template. Upon further analysis, we identified two in-frame putative alternative start sites at -30nt and -60nt of the originally annotated NznB start site. Extension of a codon-optimized gene to generate constructs with start sites at -30nt and -60nt and co-expression of chaperones via co-transformation with pGro7 plasmid enabled soluble expression of this NznB, with the -30 site garnering a dramatically increased expression yield (3-4mg/L) and enabled *in vitro* characterization of the P450 product.

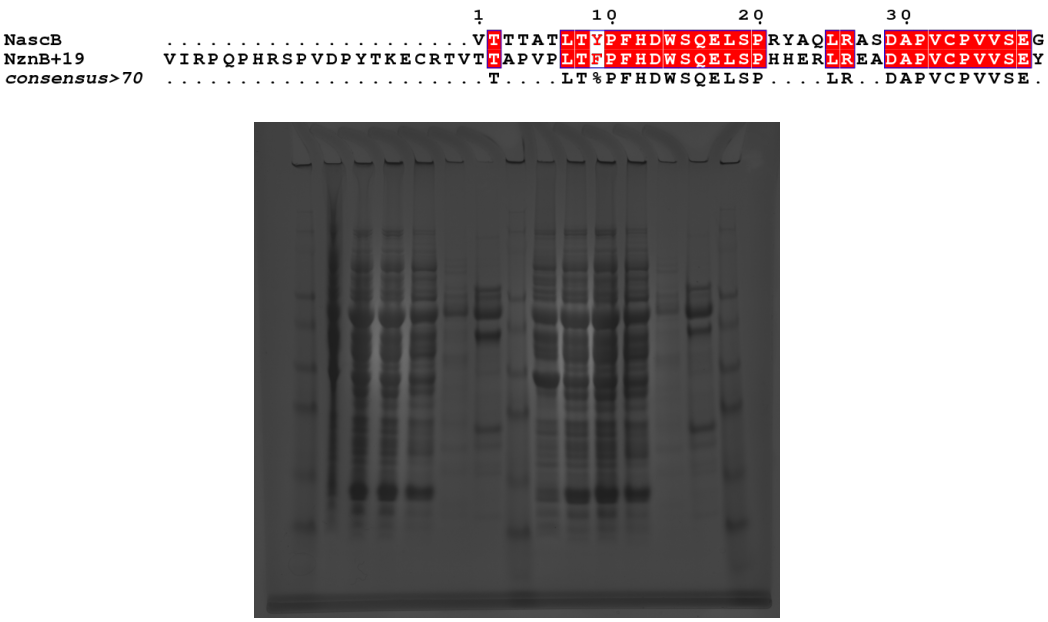


Figure 2-11. Top: Sequence alignments of NascB and NznB P450s and alternative start sites highlighted in red. Bottom: SDS-Page gel from expression of NznB+09 (-30nt, left) and NznB+19 (-60nt, right). Lanes for both expressions from left to right: 1) Ladder, 2) Pellet, 3) supernatant, 4) 25mM imidazole wash, 5) 50mM imidazole Wash, 6) 200mM imidazole elution.

Binding of CDPS products **8** and **9** to NznB was assessed via spectrophotometry and both induced a clear low-spin to high-spin transition. Incubation of **9** with NznB, ferredoxin (Fdx) and ferredoxin reductase (FdR) from *Spinacia oleracea*[19-20], and NADPH led to formation of a single product whose mass corresponded to homodimeric (+)-naseseazine B (**1**) as assessed by LC-MS, and upon scale up and isolation was confirmed by NMR to be (+)-naseseazine B (Figure 2-11).[8] (+)-naseseazine A is a heterodimer composed as **8** and **9,** and when combined together in equimolar quantities with NznB, Fdx, and FdR the major product of the reaction has a mass corresponding to (+)-naseseazine A (**2**). While the homodimer of **9**, (+)-naseseazine B (**1**), and the heterodimer of **8** and **9**, (+)-naseseazine A (**2**), are both reported in the original isolation report, the homodimer of **8** was absent. However a peak who's mass corresponded to this product was observed in minute quantities from the heterodimerization reaction of **8** and **9** with NznB. Upon incubation of only **8** with NznB, Fdx, and FdR, a single product was once again formed in high yield and high conversion whose mass corresponded to a homodimer of **8**. Scale up and NMR characterization of these compounds are currently underway.

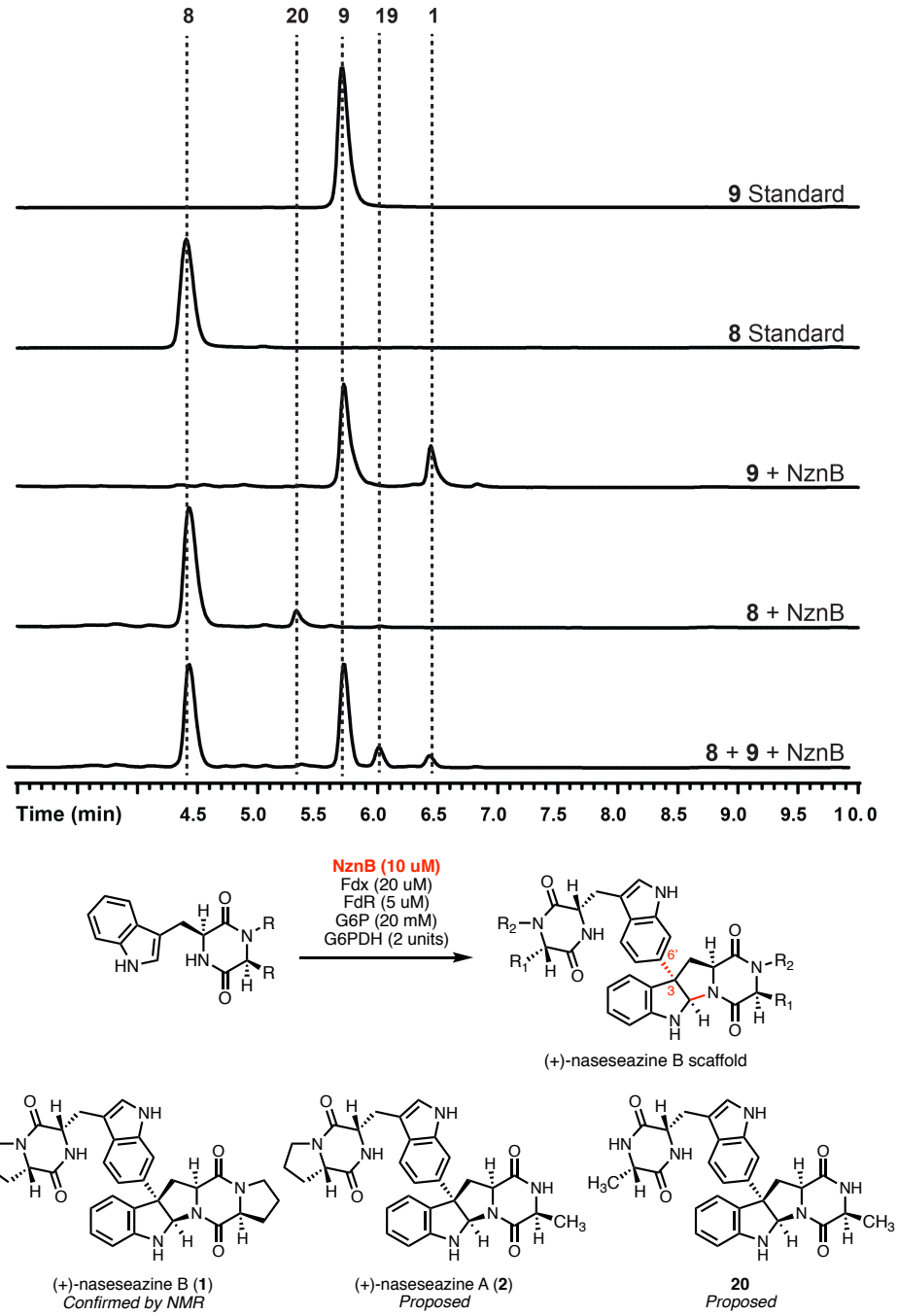


Figure 2-12. Reactions of **8** and **9** with NznB and structure of confirmed product **1**, and proposed structures of **2** and **20**.

2.3.5 *In vitro* Simulation of Putative Biosynthetic Cross-Talk

Tailoring of DKP natural products from oxidative enzymes via "biosynthetic crosstalk" has been observed in the biosynthesis of both bacterial, and fungal secondary metabolites, such as the erythrochelin[28] and spirotryprostatins,[29] respectively. Given the high sequence homology between both the *nasc* and *nzn* BGCs, we hypothesized that DKP **8** generated by NznA, and not NascA, may be accepted as a dimerization substrate by NascB via such pathway crosstalk. To simulate the possibility for this crosstalk *in vitro*, co-incubation of **8** and **9** resulted in a product who's mass corresponds to a single heterodimer (Figure 2-13). Curiously, incubation of **8** alone with NascB results in an ~1:1 mixture of two products whose masses correspond to formation of a homodimer. We hypothesize the homodimerization of **8** resulting in two constitutional isomers is due to the loss of steric bulk on the substrate (alanine relative to proline), and therefore decreased molecular recognition by NascB leading to loss of selectivity and the formation of multiple products. The formation of a single DKP heterodimer corroborates this hypothesis, as the larger proline side chain "rescues" the selectivity of NascB and yields (–)-naseseazine C exclusively.

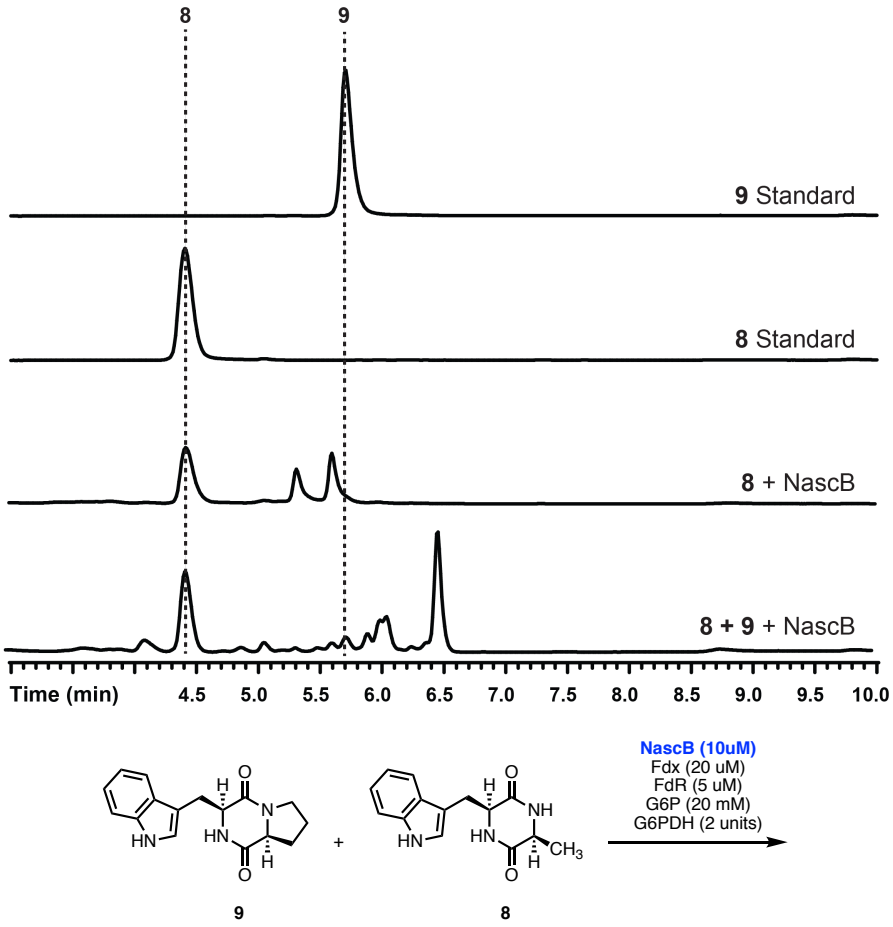
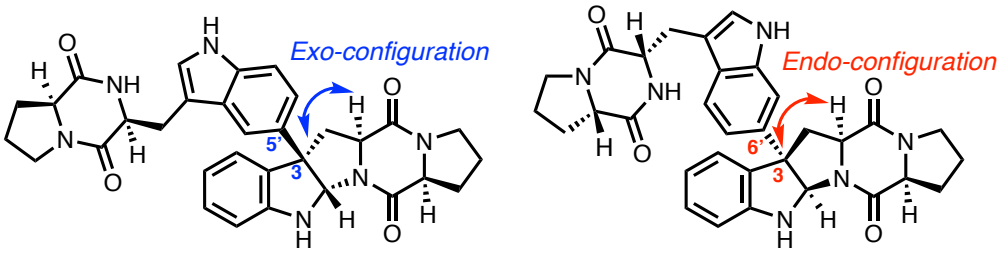


Figure 2-13. HPLC traces from reactions of **8** and **9** with NascB.

2.4 Conclusion

We identified the biosynthetic gene clusters for the biosynthesis of the (+)-naseseazines as well as the cryptic (–)-naseseazines and validated our biosynthetic hypothesis wherein cyclodipeptide synthases NascA and NznA generated diketopiperazine precursors **8** and **9** which were then dimerized to (+)-naseseazine or (–)-naseseazine scaffolds by NznB and NascB respectively. *In vitro* characterization of the cytochromes demonstrated that NascB and NznB are site- and stereodivergent biocatalysts, with NascB forming the C3-C5' linked diketopiperazine dimers with exo-stereochemistry such as (–)-naseseazine C (**4**), and NznB forming C3-C6' linked dimers with endo chemistry such as (–)-naseseazine B (**1**). Identification and initial characterization of these gene clusters may enable the use of these biocatalysts in the assembly of non-native dimers and their derivatives.



(–)-naseseazine C (**4**) (+)-naseseazine B (**1**)

Figure 2-14. Comparison of (–)-naseseazine C (**4**) and (+)-naseseazine B (**1**) structures.

We also generated previously unreported dimer **20** (structure under assignment) and simulated biosynthetic crosstalk between these two clusters demonstrating DKP **8**, produced only in the *nzn* cluster by NznA, can act as a substrate in dimerization reactions with NascB generating a suite of new, and previously uncharacterized dimers.

2.5 References

1. Anthoni, U.; Christophersen, C.; Nielsen, P. H., In *Alkaloids: Chemical and Biological Perspectives*, Pelletier, S. W., Ed. Pergamon: London, 1999; Vol. 13, pp 163-236.
2. Hino, T.; Nakagawa, M., In *The Alkaloids: Chemistry and Pharmacology*, Brossi, A., Ed. Academic Press: New York, 1989; Vol. 34, pp 1-75.
3. Borthwick, A. D., *Chem. Rev.* **2012,** *112* (7), 3641-716.
4. Hauser, D.; Weber, H. P.; Sigg, H. P., *Helv. Chim. Acta* **1970,** *53*, 1061-1073.
5. Gerken, T.; Walsh, C. T., *Chembiochem* **2013,** *14* (17), 2256-8.
6. Wang, Y.; Hu, P.; Pan, Y.; Zhu, Y.; Liu, X.; Che, Y.; Liu, G., *Fungal Genet. Biol.* **2017,** *103*, 25-33.
7. Saruwatari, T.; Yagishita, F.; Mino, T.; Noguchi, H.; Hotta, K.; Watanabe, K., *Chembiochem* **2014,** *15* (5), 656-9.
8. Kim, J.; Movassaghi, M., *J Am Chem Soc* **2011,** *133* (38), 14940-3.
9. Loach, R. P.; Fenton, O. S.; Movassaghi, M., *J .Am. Chem. Soc.* **2016,** *138* (3), 1057-64.
10. Nelson, B. M.; Loach, R. P.; Schiesser, S.; Movassaghi, M., *Org. Biomol. Chem.* **2018,** *16* (2), 202-207.
11. Varoglu, M.; Corbett, T. H.; Valeriote, F. A.; Crews, P., *J Org Chem* **1997,** *62* (21), 7078-7079.
12. Ding, G.; Jiang, L.; Guo, L.; Chen, X.; Zhang, H.; Che, Y., *J. Nat. Prod.* **2008,** *71* (11), 1861-5.
13. Li, X. B.; Li, Y. L.; Zhou, J. C.; Yuan, H. Q.; Wang, X. N.; Lou, H. X., *J Asian Nat. Prod. Res.* **2015,** *17* (2), 182-7.
14. Raju, R.; Piggott, A. M.; Conte, M.; Aalbersberg, W. G.; Feussner, K.; Capon, R. J., *Org. Lett.* **2009,** *11* (17), 3862-5.
15. Xiong, Z. Q.; Liu, Q. X.; Pan, Z. L.; Zhao, N.; Feng, Z. X.; Wang, Y., *Arch. Microbiol.* **2015,** *197* (2), 299-309.
16. Buedenbender, L.; Grkovic, T.; Duffy, S.; Kurtböke, D. I.; Avery, V. M.; Carroll, A. R., *Tet. Lett.* **2016,** *57* (52), 5893-5895.
17. Lautru, S.; Gondry, M.; Genet, R.; Pernodet, J. L., *Chem Biol* **2002,** *9* (12), 1355-64.
18. Noble, M. A.; Miles, C. S.; Chapman, S. K.; Lysek, D. A.; Mackay, A. C.; Reid, G. A.; Hanzlik, R. P.; Munro, A. W., *Biochem J* **1999,** *339*, 371-379.
19. Aliverti, A.; Jansen, T.; Zanetti, G.; Ronchi, S.; Herrmann, R. G.; Curti, B., *Eur. J. Biochem.* **1990,** *191* (3), 551-5.
20. Piubelli, L.; Aliverti, A.; Bellintani, F.; Zanetti, G., *Prot. Expr. Purif.* **1995,** *6* (3), 298-304.
21. Cai, S.; Kong, X.; Wang, W.; Zhou, H.; Zhu, T.; Li, D.; Gu, Q., *Tet. Lett.* **2012,** *53* (21), 2615-2617.
22. We refer to this isolated natural product as “new iso-naseseazine B” for clarity since the term iso-naseseazine B was originally used to refer to a regioisomer of naseasazine B in ref. 8.
23. Wyche, T. P.; Ruzzini, A. C.; Schwab, L.; Currie, C. R.; Clardy, J., *J Am Chem Soc* **2017,** *139* (37), 12899-12902.

24. For synthesis of (–)-**14** and (–)-**15** please see Experimental.
25. This structure is that proposed for “new iso-naseseazine B” but as described above that assignment is incorrect. Thus, this compound is being first reported here and has been denominated as iso-naseseazine C.
26. Movassaghi, M.; Schmidt, M. A.; Ashenhurst, J. A., *Angew Chem Int Ed Engl* **2008,** *47* (8), 1485-7.
27. Bibb, M.; Hesketh, A., *Methods Enzymol* **2009,** *458*, 93-116.
28. Lazos, O.; Tosin, M.; Slusarczyk, A. L.; Boakes, S.; Cortes, J.; Sidebottom, P. J.; Leadlay, P. F., *Chemistry & Biology* **2010,** *17* (2), 160-173.
29. Tsunematsu, Y.; Ishikawa, N.; Wakana, D.; Goda, Y.; Noguchi, H.; Moriya, H.; Hotta, K.; Watanabe, K., *Nat. Chem. Biol.* **2013,** *9* (12), 818-25.

2.5 Experimental

Instrumentation

All UV-Visible spectra were acquired using a single beam Molecular Devices Spectra Max M5 spectrophotometer, with a 1 cm quartz cuvette. Analytical HPLC data was acquired using a Shimadzu HPLC system comprised of two LC-20AD$_{XR}$ pumps, a SIL-20AC$_{XR}$ autosampler, and a SPD-M20A diode array detector. Preparatory HPLC was performed using a Beckman Coulter stack comprised of a System Gold 125 solvent module, 168 detector, and SC100 fraction collector. Isothermal titration calorimetry experiments were conducted in a Nano-ITC Low volume (TA Instruments). ^{1}H NMR spectra were recorded on Varian 600 MHz spectrometers and are reported relative to residual solvent peak (CD_2HOD: δ 3.31). ^{13}C NMR spectra were reported relative to residual solvent peaks (CD_2HOD: δ 49.0).

Streptomyces sp. CMBMQ-030 genomic DNA sequencing

The genomic DNA of *Streptomyces* sp. CMB-MQ030 was extracted and purified as prescribed using the Promega Wizard® Genomic DNA Purification Kit. The genomic DNA (2x250 bp paired-end) was sequenced using Illumina HiSeq 2500 sequencing system at University of Michigan DNA Sequencing Core. Briefly, the genomic DNA sample was first sheared to approximately 500 nucleotide average fragment size, then Illumina-compatible sequencing libraries were prepared from those fragments on an Apollo 324 robotic workstation (WaferGen Biosystems), using the Kapa HTP Library Preparation Kit (KAPABiosystem) according to the manufacturer's protocols. Subsequent libraries were sequenced on an Illumina HiSeq 2000, obtaining paired-end sequence data with 100 nucleotide reads at each end, as per recommended protocols from Illumina, Inc. The generated next-generation sequencing data were then *de novo* assembled using Velvet 1.2.10.[1]

[1] Zerbino, D. R.; Birney, E., *Gen. Res.* **2008,** *18* (5), 821-829.

Supplementary Table 2-1. Summary of genome assembly dataset and parameters

Genome	Number of reads	k-mer	Coverage cutoff	N50	Number of contigs
Streptomyces sp. (CMB-MQ030)	33,263,004	51	10	66,868	591

Cloning of NascA, NascB, and NznA

The PCR primers were designed to introduce an *Nde*I restriction site at the 5′ end of the fragment and a *Hind*III restriction site 3′ end, respectively. The coding sequences were amplified from the genomic DNA of *Streptomyces sp.* CMB-MQ030 using primers for the associated target gene (NasA, primer 1 &2, NasB primer 3 & 4). The expression constructs were designed in pET21b vector via *Nde*I and *Hind*III restriction sites. The nucleotide sequences were confirmed by automated sequencing (University of Michigan DNA Sequencing Core).

Supplementary Table 2-2. Primers used in cloning NascA, NasB, NznA, NznB, and extending codon optimized sequence of NznB

Primer #	Gene	Primer	Rest. Site	5'→3'
1	NascA	Forward	*Nde*I	CCAACCcatatgGTGAACACTTCCCTCGCTGCGGTGGCCGGC
2	NascA	Reverse	*Hind*III	CCAACCaagcttGCGTTCGGCCGCCCGGTCCCGCAGCAGGAT
3	NascB	Forward	*Nde*I	CCAACCcatatgGTGACCACCACCGCCACGCTGACCTACCCC
4	NascB	Reverse	*Hind*III	CCAACCaagcttCCAGGTGGCGGGAAGCGCCCGCGGACGGCG
5	NznA	Forward	*Nde*I	CCAACCcatatgATGGCCACACACGCCTCCGCACCCGCACCC
6	NznA	Reverse	*Hind*III	CCAACCaagcttCTGCTGCGTCACGCGGTCCTTGAGGAG
7	NznB	Forward	*Nde*I	CCAACCcatatg GTGACCACCGCACCGGTCCCCCTCACCTTC
8	NznB	Reverse	*Hind*III	CCAACCaagcttCCAGGAAGCGGGGATCGCCTT

				CGGGCGGCG
9	NznB (CO) +09	Forward		GTGGATCCGTATACCAAAGAATGCCGTACCGT GACCACCGCGCCGGTTCCGCT
10	NznB (CO) +19	Forward		GTGATTCGTCCGCAGCCGCATCGTAGCCCGGT GGATCCGTATACCAAAGAATGCCGTACCGTGA CCACCGCGCCGGTTCCGCT

Note: Lowercase sequences represent restriction sites.

Codon optimized sequence for NznB

ACCACCGCGCCGGTTCCGCTGACCTTCCCGTTTCACGATTGGAGCCAAGAGCTGA
GCCCGCACCATGAGCGTCTGCGTGAAGCGGATGCGCCGGTGTGCCCGGTGGTTA
GCGAGTACACCGGTGACCGTCTGTGGCTGGTTACCCGTTATGCGACCGCGAAGC
GTCTGCTGGAAGATCGTCGTTTCAGCAGCACCGCGGCGATGGCGCCGGGTGCGC
CGCGTCAGGAGCCGGTGGAACTGCGTGCGCCGGGCACCACCGGTGATGGTGTGA
GCGTTCTGCGTGAGGCGGGTCTGCGTACCGTTTTTACCGAAGGTCTGGGTCCGCG
TGCGGCGCGTCGTCACGGTAAATGGCTGCGTGATCGTGCGGACACCCTGCTGCG
TGATGTGGCGGAGTGCGAAGGTCCGGTTGATCTGGCGGCGGACTTTGCGCAGCC
GCTGGCGGTGGCGATGACCAGCCGTGTTCTGCTGGGTGAACTGAGCACCGAGGA
AGCGGCGCTGCTGCGTGATCGTACCGATCTGGCGCTGCAGTTTTGCGGTGCGAC
CGCGGAGGAACAACGTGGTGGCCTGATCGATATTCACCGTTTCTTTACCGCGCAT
GCGCGTCGTCTGGCGGATGGTCCGGGTGACCACCTGCTGAAGCGTCTGGCGGAG
GCGCCGGCGGAAAACGGTCCGCTGGGTGATGCGGCGCTGAGCGAAATTGCGGC
GCTGCTGCTGATTGCGGGTTTCCCGACCAGCAGCGGCTTTCTGTGCGGTGCGCTG
ATCACCCTGCTGCGTCACCCGGAGGCGGTGGGCCGTCTGCGTCGTGATCCGGAA
CTGATTCCGGACGCGGTTGAGGAACTGCTGCGTCATACCCCGCTGAGCACCGGTG
CGGCGAAACGTATGGCGACCGAGGACGCGGATATCGACGGCGTGCGTATTCGTC
GTGGTGAGGTGGCGATGGTTAGCCTGGAAGCGGCGAACCACGATCCGGACGCGT
TCGACGATCCGGATAGCTTTCGTCCGGAACGTCAAGGTCCGGGTCACCTGGGTTT
TGGCCACGGTCCGAACTTTTGCCCGGGTAACCGTCTGGCGCGTTGCCTGATTGAT
GCGATGGTGCGTGCGGTTGCGCGTCGTCCGGGTCTGCACCTGACCGTTGGTCCG
GAGGAAATCCGTTGGCACGAAGGCCTGTTCTTTCGTCGTCCGAAGGCGATTCCGG
CGAGCTGG

Overexpression and analysis of NasA and NznA.

The expression of pET21b-NasA was performed by transforming the plasmid into the competent cell *E. coli* strain C41(DE3) and selected on Luria-Bertani (LB) medium plates containing 100 µg/mL ampicillin. A single colony was grown overnight in 5 mL LB broth containing the same concentration of the antibiotics. The main culture was prepared by inoculating 1% with overnight culture into a 2 L baffled flask containing 1 L of LB containing 100 µg/mL ampicillin. The cultures were incubated at 37°C and 250 rpm for 3-4 h until absorbance A_{590nm} = 0.6. The expression of NasA was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cultures were incubated further at 18°C for 18 h shaking at 180 rpm. Cells were harvested by centrifugation at 5000×g at 4°C for 15 min and a 1 mL aliquot of the supernatant was extracted with 1 mL $CHCl_3$ (3x). The organic extracts were dried and solvents were removed *en vacuo*. Residue was resuspended in 100 uL of HPLC grade methanol for HPLC analysis in which the samples was resolved using a linear gradient of 5-100% acetonitrile: water (0.1% formic acid) over 30 min (1.5 mL/min flow rate) on a Phenomenex Luna 5µ C-18(2) 100A, 250 × 4.60mm 5 micron column.

Overexpression, purification and spectral characterization of cytochrome P450 NascB

The expression of pET21b-NasB was performed by transforming the plasmid into the competent cell *E. coli* strain C41(DE3) and selected on Luria-Bertani (LB) medium plates containing100 µg/mL ampicillin. A single colony was grown overnight in LB broth containing the same concentration of the antibiotics. The main culture was prepared by inoculating 1% of each overnight culture into a 2 L baffled flask containing 500 ml of Terrific Broth (TB) containing 100 µg/mL ampicillin. The cultures were incubated at 37°C and 250 rpm for 3-4 h unless the absorbance A_{590nm} = 1. 0.5 mM delta-5-aminolevulinic acid (w/v) was added and expression of NascB was induced by the addition of 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cultures were incubated further for

36 h, 28°C at 160 rpm. Cells were harvested by centrifugation at 5000×*g* at 4°C for 10 min and the cell pellet was stored at –80°C until purification of the protein.

The cell pellet was resuspended in 5% culture volume of lysis buffer (50 mM Tris, pH 7.4 containing 50 mM NaCl, 2% glycerol, 0.5 mM EDTA, 10 mM beta-mercaptoethanol and 1mM phenylmethane sulfonyl fluoride (PMSF) and disrupted by sonication. The cell lysate was centrifugated (35,000 rpm for 35 min) and filtered (0.2 µM Millipore filter). The soluble His_6-tagged NasB was purified by affinity chromatography using Ni-NTA (Qiagen) column and the collected fractions were analyzed by SDS-PAGE. The suitable red fractions were pooled and dialyzed at 4°C three times using 50 mM Tris, pH 7.4 containing 10% glycerol, 50 mM NaCl and 1 mM dithiothreitol (DTT) against a total of 6 L buffer.

Overexpression, purification and spectral characterization of cytochrome P450 NznB

The expression of pET28b-NznB+09 was performed by transforming the plasmid into the competent cell *E. coli* strain C41 (DE3) and selected on Luria-Bertani (LB) medium plates containing100 µg/mL kanamycin. A single colony was grown overnight in LB broth containing the same concentration of the antibiotics. The main culture was prepared by inoculating with 1% overnight culture into a 2 L baffled flask containing 500 ml of Terrific Broth (TB) containing 100 µg/mL kanamycin. The cultures were incubated at 37°C and 250 rpm for 3-4 h unless the absorbance A_{590nm} = 1. 0.5 mM delta-5-aminolevulinic acid (w/v) as well as 1mM thiamine were then added and the expression of NznB was induced by the addition of 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cultures were incubated further for 30 h, 18°C at 160 rpm. Cells were harvested by centrifugation at 5000×*g* at 4°C for 10 min and the cell pellet was stored at –80°C until purification of the protein.

The cell pellet was resuspended in 5% culture volume of lysis buffer (50 mM kPi, pH 8.0 containing 200 mM $(NH_4)_2SO_4$, 20% glycerol, 0.5 mM EDTA, 10 mM beta-mercaptoethanol and 1mM phenylmethane sulfonyl fluoride (PMSF) and disrupted by

sonication. The cell lysate was centrifugated (35,000 rpm for 35 min) and filtered (0.2 µM Millipore filter). The soluble His_8-tagged NznB was purified by affinity chromatography using Ni-NTA (Qiagen) column and the collected fractions were analyzed by SDS-PAGE. The suitable red fractions were pooled and dialyzed at 4°C three times using 50 mM Tris, pH 7.4 containing 50 mM kPi, pH 8.0 containing 200 mM $(NH_4)_2SO_4$, 20% glycerol, 0.5 mM EDTA, and 1 mM dithiothreitol (DTT) against a total of 6 L buffer.

UV–Vis spectra for the purified cytochomes were recorded at room temperature in buffer (50 mM Tris-HCl buffer, pH 7.4 containing 5% glycerol) was used for the spectral measurements of the oxidized and reduced form. NasB was reduced by the addition of a small amount of sodium dithionite. The concentration of the P450s was estimated by CO-difference spectra, assuming ε(459–490) = 91 mM−1 cm−1 according to the method of Omura and Sato.

General reaction conditions for analytical scale P450 reactions

The conversion of DKPs by P450s were carried out with the heterologous redox electron partners Fdx and FdR from *S. olerecea*. A protein ratio of P450: Fdx: FdR of 1: 20: 3 was used. The *in vitro* reaction mixture included 50 mM Tris-HCl, pH 7.4, 0.5 µM P450, 10 µM Fdx and 1.5 µM FdR at the end volume of 250 µL. Without exceeding 1% of the total reaction volume, 200 µM of DKP dissolved in DMSO was added. The reaction was initiated by adding NADPH (1 mM). Two of the control reactions including all the contents except for NADPH or P450 were used as controls. The reactions were incubated at 30°C for 1 h agitating at 600 rpm in a thermoshaker (Multi-thermoshaker, Benchmark). The conversion was stopped by adding an equal volume of chloroform and was extracted with chloroform for 3 times. The samples were dried under a stream of nitrogen and the residue was resuspended in 100 µL of HPLC grade methanol and analyzed by HPLC in which the reaction mixture was resolved using a linear gradient of 5-100% acetonitrile: water (0.1% formic acid) over 30 min (1.5 mL/min flow rate) on a Phenomenex Luna 5µ C-18(2) 100A, 250 × 4.60mm 5 micron column.

General reaction conditions for preparatory scale P450 reactions

Large scale reaction were performed in 84 identical reactions as described above. The reactions were combined after 2 hours, diluted with 1 volume of brine and quenched with 100 mL of $CHCl_3$. The mixture was centrifuged at 5,000 X g and poured into a separatory funnel. The organic layer was partitioned and the aqueous layer was extracted 2 more times (100 mL each). The organic layers were combined, dried and solvent was removed *en vacuo*.

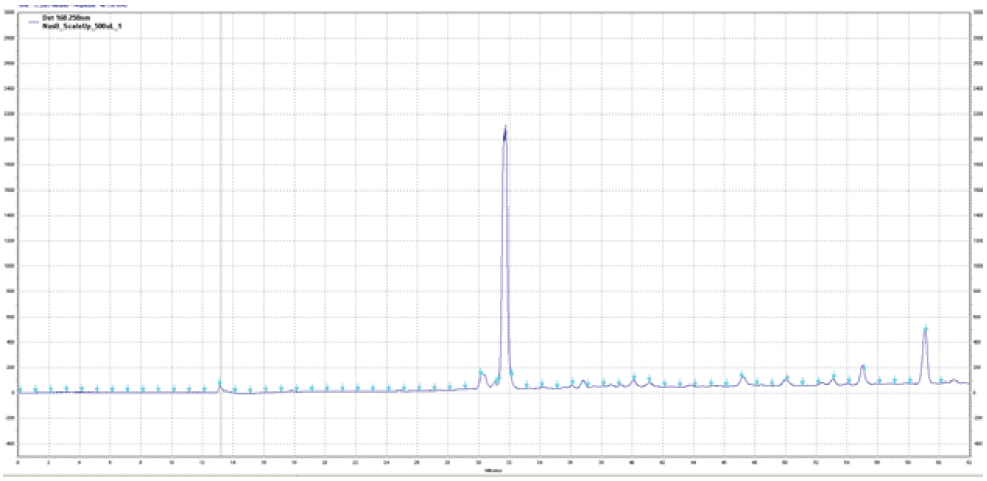
Prep-Scale Synthesis of (–)-naseseazine C (4)

Procedure above was followed with DKP **9**, as a substrate and NascB as P450. Residue was diluted in 8 mL HPLC grade methanol, sonicated to solvate, and filtered through a 0.2 µm syringe filter. The filtered solution was then purified via preparative HPLC (Luna preparative HPLC column, C18, 5 µm, 21.2 × 250 mm;) using gradient of 5% to 100% acetonitrile in water containing 0.1% formic acid at the flow of 7.0 mL/min for 60 min. The retention time (t_R) of **7** was 31.2 min. The product fractions were pooled, solvent was removed *en vacuo*, to give 8 mg of (–)-naseseazine C which was used for detailed NMR characterization without any further purification.

Prep-Scale Synthesis of (+)-naseseazine B (1)

Procedure above was followed with DKP **9**, as a substrate and NznB as P450. Residue was diluted in 8 mL HPLC grade methanol, sonicated to solvate, and filtered through a 0.2 µm syringe filter. The filtered solution was then purified via preparative HPLC (Luna preparative HPLC column, C18, 5 µm, 21.2 × 250 mm;) using gradient of 5% to 100% acetonitrile in water containing 0.1% formic acid at the flow of 7.0 mL/min for 60 min. The retention time (t_R) of **7** was 34.1 min. The product fractions were pooled, solvent was removed *en vacuo*, to give 11 mg of (+)-naseseazine B which was used for detailed NMR characterization without any further purification.

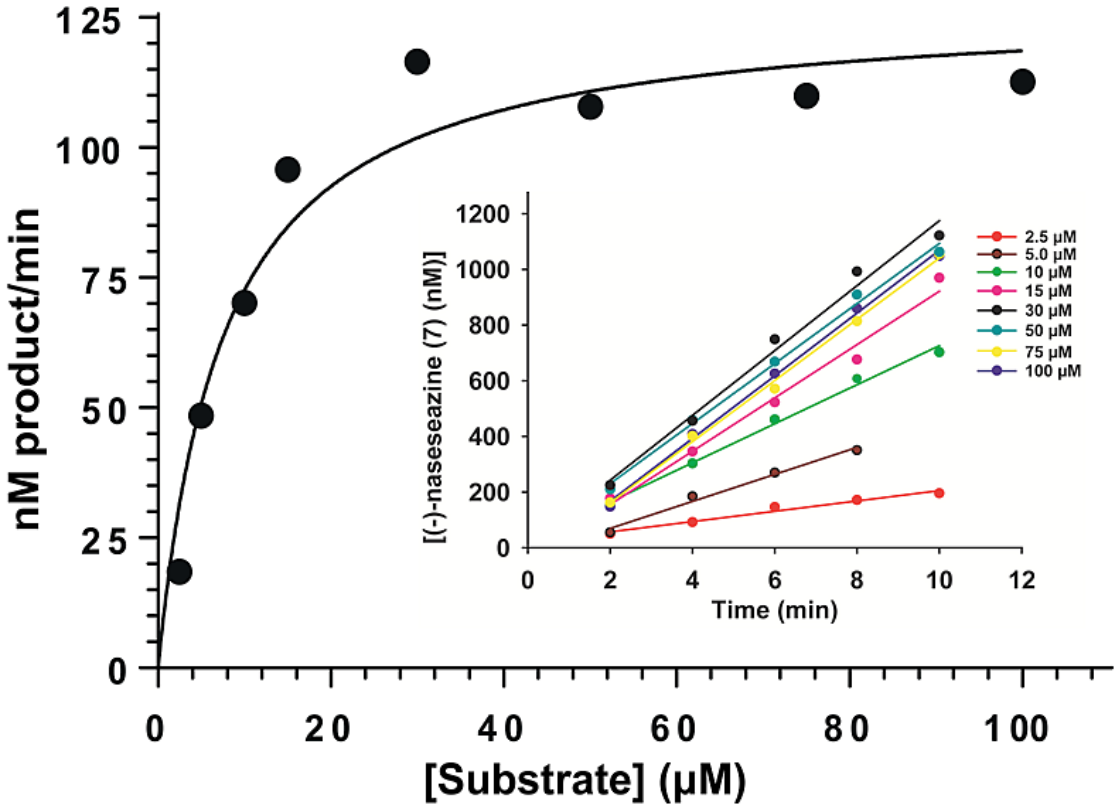
Supplementary Figure 2-1. Sample Prep-HPLC trace from purification of 4 from large scale conversion of 9 by NascB.



Steady state kinetics of conversion of brevianamide F (9) to (–)-naseseazine C (4) by NasB

In order to capture the steady state, the concentration of NascB and **9** had to be significantly lower than in previous enzymatic reactions. As such to visualize product formation, reaction volume was scaled up to 2 mL containg NascB (30 nM), Fdx (600 nM), and FdR (90 nM), and the substrate concentrations were varied from 2.5-100 μM. Reactions were initiated with the addition of NADPH (1mM final concentration), and quenched at 2 min intervals with the addition of 1 mL of $CHCl_3$ and vortexing at high speeds for 30s. An additional 500 uL of $CHCl_3$ was added and 800 uL of organic layer was removed, and this was repeated 2 more times. Solvents were removed *en vacuo*, and residue was resuspended in 100 uL of HPLC grade methanol, and analyzed by HPLC in which the reaction mixture was resolved using a linear gradient of 20-55% acetonitrile : water (0.1% formic acid) for 12 min (1.5 mL/min flow rate) on a Phenomenex Luna 5 μ C-18(2) 100A, 250 × 4.60 mm 5 micron column. Initial velocities were determined based on the formation of **7** and quantified using a standard curve. Initial velocities were plotted against substrate concentration and fit by non-linear regression to the Michaelis-Menten equation using GraphPad Prism Version 6.01 to determine the kinetic constants k_{cat} and K_M.

Supplementary Figure 2-2. Steady state kinetics of 9 conversion by NascB. The Michaelis-Menten hyperbolic fit for [naseseazine C (**9**, nM)] /min vs [brevianamide F (**4**, nM)] is shown. The inset showed the capture of steady state product formation (nM) using varying concentration (2.5 μM–100 μM) of **9** at varying time interval (2 min to 10 min). The data represents the mean of 3 experiments.



Supplementary Table 2-3. Steady state kinetics of 6 conversion by NasB.

Enzyme	Substrate	Kcat (s^{-1})	K_M (μM)	kcat/K_M (M^{-1} s^{-1})
NascB	Brevianamide F (**9**)	0.071	7.6	9.4×10^3

Chemistry

General Procedures. All reactions were performed in oven-dried or flame-dried round-bottom flasks. The flasks were fitted with rubber septa and reactions were conducted

under a positive pressure of argon. Gas-tight syringes with stainless steel needles or cannulae were used to transfer air- and moisture-sensitive liquids. Where necessary (so noted), reactions were performed in Schlenk tubes fitted with a PTFE stopcock. Flash column chromatography was performed as described by Still et al. using granular silica gel (60 Å pore size, 40–63 μm, 4–6% H_2O content, Zeochem).[2] Analytical thin layer chromatography (TLC) was performed using glass plates pre-coated with 0.25 mm 230–400 mesh silica gel impregnated with a fluorescent indicator (254 nm). TLC plates were visualized by exposure to short wave ultraviolet light (254 nm) and an aqueous solution of ceric ammonium molybdate (CAM) followed by heating on a hot plate (~250 ºC). Organic solutions were concentrated at 29–30 °C on rotary evaporators capable of achieving a minimum pressure of ~2 torr.

Materials. Commercial reagents and solvents were used as received with the following exceptions: dichloromethane, acetonitrile, tetrahydrofuran, methanol, pyridine, toluene, and triethylamine were purchased from J.T. Baker (Cycletainer™) and were purified by the method of Grubbs et al. under positive argon pressure.[3] Nitroethane was distilled over calcium hydride and stored over 4 Å molecular sieves. Silver hexafluoroantimonate was purchased from Strem Chemicals, Inc. All other solvents and chemicals were purchased from Sigma–Aldrich.

Instrumentation. Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded with Varian inverse probe INOVA-500 and Varian INOVA-500 spectrometers, or a Bruker AVANCE III 400 spectrometer. Chemical shifts are reported in parts per million on the δ scale, and are referenced from the residual protium in the NMR solvent ($CHCl_3$: δ 7.26, CD_3SOCD_2H: δ 2.50, CD_2HOD: δ 3.31).[4] Data are reported as follows: chemical shift [multiplicity (s = singlet, d = doublet, t = triplet, m = multiplet), coupling constant(s) in Hertz, integration, assignment]. Carbon-13 nuclear magnetic resonance (^{13}C NMR) spectra were recorded with a Varian INOVA-500 spectrometer or a Bruker AVANCE III

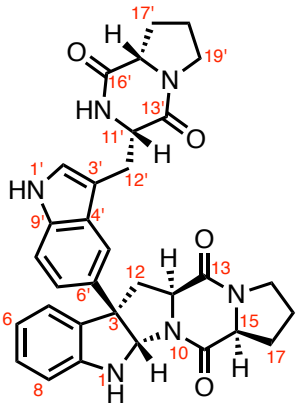
[2] Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923.

[3] Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J. *Organometallics* **1996**, *15*, 1518.

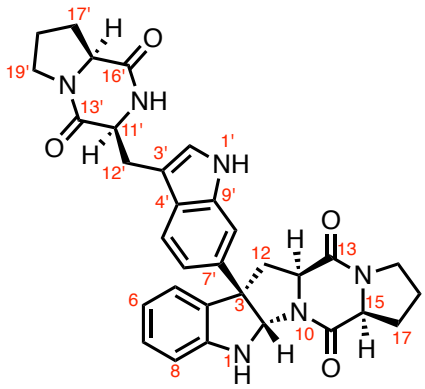
[4] G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg. *Organometallics* **2010**, *29*, 2176.

400 spectrometer and are reported in parts per million on the δ scale, and are referenced from the carbon resonances of the solvent ($CDCl_3$: δ 77.16, DMSO-d_6: δ 39.52, CD_3OD: δ 49.00). Data are reported as follows: chemical shift (assignment). Data are reported as follows: chemical shift (assignment). Infrared data (IR) were obtained with a Perkin-Elmer 2000 FTIR, and are reported as follows: frequency of absorption (cm^{-1}), intensity of absorption (s = strong, m = medium, w = weak, br = broad). Optical rotations were measured on a Jasco-1010 polarimeter. UV-Vis spectrophotometric data were acquired on a Varian Cary 50 Bio UV-Vis spectrophotometer. Preprative HPLC was performed on a Waters system with the 1525 Binary HPLC Pump, 2489 UV/Vis Detector, 3100 Mass Detector, System Fluidics Organizer, and 2767 Sample Manager components. We thank Liam P. Kelly for obtaining the mass spectrometric data at the Department of Chemistry's Instrumentation Facility, Massachusetts Institute of Technology. High resolution mass spectra (HRMS) were recorded on a Bruker Daltonics APEXIV 4.7 Tesla FT-ICR-MS using either an electrospray (ESI) or direct analysis in real time (DART) ionization source.

Positional Numbering System. In assigning the ^{1}H and ^{13}C NMR data of all intermediates en route to our total syntheses of (–)-naseseazine C and (–)-*iso*-naseseazine C, we wished to employ a uniform numbering scheme. This numbering system is also consistent with that employed in the isolation papers of (–)-naseseazine C[5] and *iso*-naseseazine B.[6]



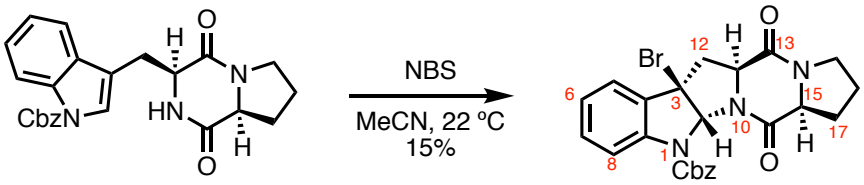
(–)-naseseazine C (**7**)



(–)-*iso*-naseseazine C (**8**)

[5] Buedenbender, L.; Grkovic, T.; Duffy, S.; Kurtbotke, D. I.; Avery, V. M.; Carroll, A. R. *Tet. Lett*, **2016**, *57*, 5893.

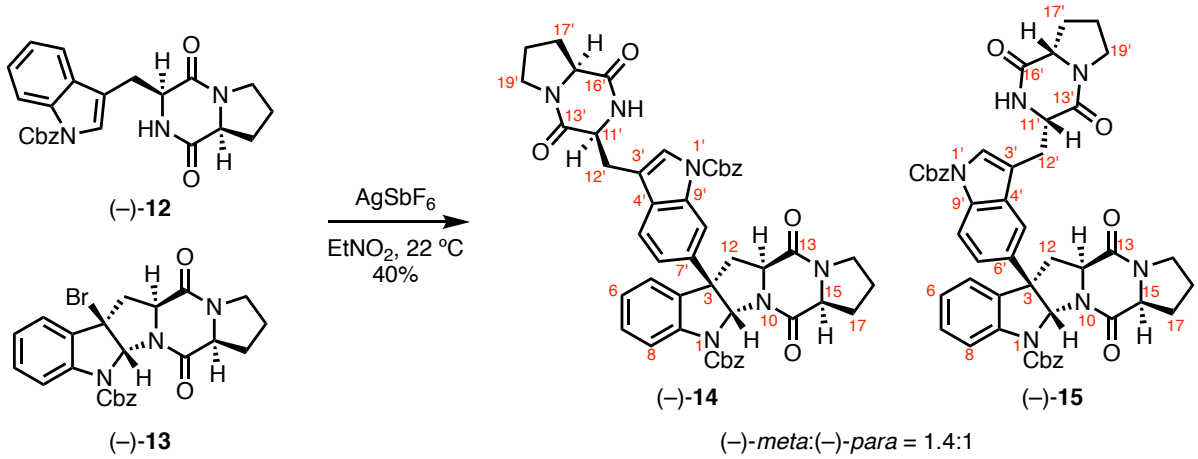
[6] Xiong, Z.-Q.; Liu, Q.-X.; Pan, Z.-L.; Zhao, N.; Feng, Z.-X.; Wang, Y. *Arch. Microbiol*, **2015**, *197*, 299.



(–)-**12** (–)-**13**

Diketopiperazine bromide (–)-13: A sample of *N*-bromosuccinimide (641 mg, 3.60 mmol, 3.00 equiv) was added in one portion to a stirred suspension of diketopiperazine **14** (500 mg, 1.20 mmol, 1 equiv) in acetonitrile (60 mL). After 25 min, a solution of saturated aqueous sodium thiosulfate solution (30 mL) was added to the reaction mixture. After 10 min, the reaction mixture was diluted with ethyl acetate (30 mL), and washed with saturated aqueous sodium thiosulfate solution (30 mL). The aqueous layer was extracted with ethyl acetate (3 × 20 mL) and the combined organic layers were dried over anhydrous sodium sulfate, were filtered, and were concentrated under reduced pressure. The crude reaction mixture was purified by flash column chromatography on silica gel (eluent: 30→35% ethyl acetate in hexanes), including a second chromatographic purification (eluent: 30→35% ethyl acetate in hexanes) to afford *exo*-diketopiperazine bromide **(–)-15** (90.5 mg, 14.8%) as an off-white foam along with an equal amount (~15%) of the corresponding *endo*-diastereomer (*exo*-DKP:*endo*-DKP, 1:1). Structural assignments were made using additional information from gCOSY, HSQC, HMBC and nOe experiments. ^{1}H NMR (MHz, $CDCl_3$, 20 ºC): δ 7.69 (app-br-s, 1H, C_8**H**), 7.48 (app-d, *J* = 7.3 Hz, 2H, Ph_{Cbz}-*o*-**H**), 7.43 (d, *J* = 7.7 Hz, 1H, C_5**H**), 7.41–7.30 (m, 4H, C_7**H**, Ph_{Cbz}-*m*-**H**, Ph_{Cbz}-*p*-**H**), 7.16 (app-t, *J* = 7.5 Hz, 1H, C_6**H**), 6.57 (s, 1H, C_2**H**), 5.43 (d, *J* = 12.2 Hz, 1H, $Ph_{Cbz}C$**H**$_a$), 5.31 (d, *J* = 12.3 Hz, 1H, $Ph_{Cbz}C$**H**$_b$), 4.04 (app-t, *J* = 8.2 Hz, 1H, C_{15}**H**), 3.97 (dd, *J* = 5.9, 11.4 Hz, 1H, C_{11}**H**), 3.62–3.45 (m, 2H, C_{19}**H**), 3.35 (dd, *J* = 5.9, 13.2 Hz, 1H, C_{12}**H**$_a$), 3.01 (dd, *J* = 11.0, 13.2 Hz, 1H, C_{12}**H**$_b$), 2.35–2.23 (m, 1H, C_{17}**H**$_a$), 2.25–2.14 (m, 1H, C_{17}**H**$_b$), 2.10–1.97 (m, 1H, C_{18}**H**$_a$), 1.95–1.80 (m, 1H, C_{18}**H**$_b$). ^{13}C NMR (100 MHz, $CDCl_3$, 20 ºC): δ 165.8 (**C**$_{16}$), 164.2 (**C**$_{13}$), 153.0 (N**C**=O_{Cbz}), 141.4 (**C**$_9$), 135.8 (Ph_{Cbz}-*i*-**C**), 132.0 (**C**$_4$), 131.2 (Ph_{Cbz}), 128.7 (**C**$_7$), 128.5 (Ph_{Cbz}-*o*-**C**), 128.4 (Ph_{Cbz}), 125.1 (**C**$_6$), 123.9 (**C**$_5$), 117.7 (**C**$_8$), 83.3 (**C**$_2$), 68.6, 60.4 (2C, **C**$_{11}$, **C**$_{15}$), 59.0 (**C**$_3$), 45.6 (**C**$_{19}$), 43.1 (**C**$_{12}$), 27.8 (**C**$_{17}$), 23.1 (**C**$_{18}$). FTIR (thin

film) cm^{-1}: 2957 (w), 2887 (w), 1716 (s), 1674 (s), 1329 (m). HRMS (ESI) (*m*/*z*):calc'd for $C_{24}H_{23}BrN_3O_4$ $[M+H]^+$: 496.0866, found: 496.0858. $[\alpha]_D^{24}$: –168 (*c* = 0.46, CH_2Cl_2). TLC (70% ethyl acetate in hexanes), R*f*: 0.22 (UV, CAM).



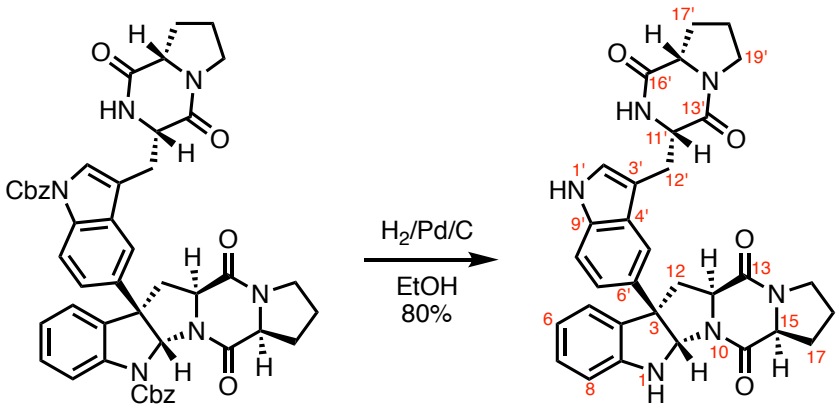
***meta*-(−)-*N*^{in},*N*$^{in'}$-Dicarboxybenzyl *iso*-Naseseazine C (14) and *para*-(−)-*N*^{in},*N*$^{in'}$-Dicarboxybenzyl Naseseazine C (15):** Diketopiperazine bromide **(–)-13** (100 mg, 0.196 mmol, 1 equiv) and proline diketopiperazine **12** (245 mg, 0.588 mmol, 3.00 equiv) were azeotropically dried from anhydrous benzene (3 × 0.5 mL) and the residue was suspended in freshly distilled nitroethane (9.8 mL). Silver hexafluoroantimonate (202 mg, 0.588 mmol, 3.00 equiv) was then added as a solid in a single portion. After 4.5 h, the reaction mixture was diluted with dichloromethane (15 mL) and washed with a mixture of saturated aqueous sodium bicarbonate and saturated aqueous sodium thiosulfate (10:1 v/v, 15 mL). The resulting aqueous layer was extracted with dichloromethane (3 × 10 mL) and the combined organic layers were dried over anhydrous sodium sulfate, were filtered, and were concentrated under reduced pressure. The crude reaction mixture was purified by flash column chromatography on silica gel (eluent: 10→50% acetone in dichloromethane) to afford a regioisomeric mixture of (–)-**15** and (–)-**14** (65.6 mg, 40.1%, (−)-**17**:(−)-**16**, 1.4:1) as a white foam. Regioisomers (−)-**15** and (−)-**14** were separated for the purpose of full and independent characterization by preparative HPLC [Luna preparative HPLC column, C18, 5 μm, 21.2 × 250 mm; 7.0 mL/min; gradient, 5%→100% acetonitrile in water with 0.1% formic acid, 60 min; t_R((−)-**15**) = 48.3 min, t_R((−)-**14**) = 50.4 min]. Structural assignments were made using additional information from gCOSY, HSQC, HMBC and nOe experiments.

***para*-(−)-*N*^{in},*N*$^{in'}$-Dicarboxybenzyl Naseseazine C (15):** ^{1}H NMR (500 MHz, $CDCl_3$, 20 °C): δ 7.99 (br-s, 1H, $C_{8'}$**H**), 7.71 (br-s, 1H, C_8**H**), 7.53 (s, 1H, $C_{2'}$**H**), 7.48–7.43 (m, 3H,

$C_{5'}$**H**, Ph_{Cbz}-**H**), 7.43–7.35 (m, 5H, Ph_{Cbz}-**H**), 7.34–7.26 (m, 4H, C_7**H**) Ph_{Cbz}-**H**), 7.22 (d, *J* = 7.4 Hz, 1H, C_5**H**), 7.10 (app-t, *J* = 7.5 Hz, 1H, C_6**H**), 7.04 (d, *J* = 8.8 Hz, 1H, $C_{7'}$**H**), 6.61 (s, 1H, C_2**H**), 5.69 (s, 1H, $N_{10'}$**H**), 5.42 (app-s, 2H, $Ph_{Cbz'}C$**H**$_2$), 5.39 (d, *J* = 12.2 Hz, 1H, $Ph_{Cbz'}C$**H**$_a$), 5.28 (d, *J* = 12.3 Hz, 1H, $Ph_{Cbz'}C$**H**$_b$), 4.31 (dd, *J* = 3.5, 10.7 Hz, 1H, $C_{11'}$**H**), 4.15 (dd, *J* = 5.8, 11.3 Hz, 1H, C_{11}**H**), 4.10 (app-q, *J* = 7.8 Hz, 2H, C_{15}**H**, $C_{15'}$**H**), 3.69–3.47 (m, 5H, $C_{12'}$**H**$_a$, C_{19}**H**, $C_{19'}$**H**), 3.16 (dd, *J* = 5.8, 12.9 Hz, 1H, C_{12}**H**$_a$), 2.91–2.82 (m, 2H, C_{12}**H**$_b$, $C_{12'}$**H**$_b$), 2.36–2.27 (m, 2H, $C_{17'}$**H**), 2.27–2.11 (m, 1H, C_{17}**H**$_a$), 2.11–1.95 (m, 3H, C_{17}**H**$_b$, C_{18}**H**), 1.96–1.84 (m, 2H, $C_{18'}$**H**). ^{13}C NMR (125 MHz, $CDCl_3$, 20 ºC): δ 169.5 (**C**$_{16'}$), 166.0 (**C**$_{16}$), 165.4 (**C**$_{13}$), 164.9 (**C**$_{13'}$), 153.1 (**C**$_{Cbz}$=O), 150.4 (**C**$_{Cbz'}$=O), 141.5 (**C**$_9$), 136.6 (**C**$_{6'}$), 136.1 (Ph_{Cbz}-*i*-**C**), 135.1 (**C**$_4$), 134.8 (2C, **C**$_{9'}$, $Ph_{Cbz'}$-*i*-**C**), 130.0 (**C**$_{4'}$), 129.3 (**C**$_7$), 129.0 (2C, Ph_{Cbz}-**C**), 128.8 (Ph_{Cbz}-**C**), 128.6 (Ph_{Cbz}-**C**), 128.3 (2C, Ph_{Cbz}-**C**), 125.0 (**C**$_{2'}$), 124.8 (**C**$_6$), 124.0 (**C**$_5$), 123.6 (**C**$_{7'}$), 117.6 (**C**$_8$), 116.3 (**C**$_{8'}$), 116.0 (**C**$_{3'}$), 115.4 (**C**$_{5'}$), 81.9 (**C**$_2$), 69.2 ($Ph_{Cbz'}$-**C**H$_2$), 68.3 (Ph_{Cbz}-**C**H$_2$), 60.7 (**C**$_{11}$ or **C**$_{15'}$), 60.6 (**C**$_{11}$ or **C**$_{15'}$), 59.3 (**C**$_{15}$), 58.2 (**C**$_3$), 53.9 (**C**$_{11'}$), 45.6 (**C**$_{19}$ or **C**$_{19'}$), 45.5 (**C**$_{19}$ or **C**$_{19'}$), 39.6 (**C**$_{12}$), 28.4 (**C**$_{17'}$), 27.9 (**C**$_{17}$), 26.5 (**C**$_{12'}$), 23.1 (**C**$_{18'}$), 22.8 (**C**$_{18}$). FTIR (thin film) cm^{-1}: 3472 (br-m), 2956 (w), 1716 (s), 1668 (s), 1479 (m), 1401 (m). HRMS (ESI) (*m*/*z*): calc'd for $C_{48}H_{45}N_6O_8$ $[M+H]^+$: 833.3293, found: 833.3286. $[\alpha]_D^{24}$: –128 (*c* = 0.24, CH_2Cl_2). TLC (70% acetone in hexanes), R*f*: 0.24 (UV, CAM).

***meta*-(−)-*N*^{in},*N*$^{in'}$-Dicarboxybenzyl *iso*-Naseseazine C (15):** ^{1}H NMR (MHz, $CDCl_3$, 20 ºC): δ 8.19 (br-s, 1H, $C_{8'}$**H**), 7.73 (br-s, 1H, C_8**H**), 7.50 (s, 1H, $C_{2'}$**H**), 7.48–7.38 (m, 6H, Ph_{Cbz}-**H**), 7.38–7.34 (m, 2H, $C_{5'}$**H**, Ph_{Cbz}-**H**), 7.34–7.26 (m, 4H, C_7**H**, Ph_{Cbz}-**H**), 7.19 (d, *J* = 7.5 Hz, 1H, C_5**H**), 7.08 (d, *J* = 7.5 Hz, 1H, $C_{6'}$**H**), 7.07–7.02 (m, 1H, C_6**H**), 6.64 (s, 1H, C_2**H**), 5.57 (s, 1H, $N_{10'}$**H**), 5.41 (d, *J* = 12.3 Hz, 1H, $Ph_{Cbz}C$**H**$_a$), 5.39 (app-s, 2H, $Ph_{Cbz'}C$**H**$_2$), 5.28 (d, *J* = 12.3 Hz, 1H, $Ph_{Cbz}C$**H**$_b$), 4.29 (dd, *J* = 2.9, 10.2 Hz, 1H, $C_{11'}$**H**), 4.18–4.14 (m, 1H, C_{11}**H**), 4.14–4.08 (m, 1H, $C_{15'}$**H**), 4.05 (app-t, *J* = 8.0 Hz, 1H, C_{15}**H**), 3.67–3.50 (m, 5H, $C_{12'}$**H**$_a$, C_{19}**H**, $C_{19'}$**H**), 3.15 (dd, *J* = 5.8, 12.9 Hz, 1H, C_{12}**H**$_a$), 2.92–2.81 (m, 2H, C_{12}**H**$_b$, $C_{12'}$**H**$_b$), 2.37–2.28 (m, 2H), 2.27–2.15 (m, 1H, $C_{17'}$**H**), 2.11–1.98 (m, 3H, C_{17}**H**$_b$, C_{18}**H**), 1.98–1.85 (m, 2H, $C_{18'}$**H**). ^{13}C NMR (MHz, $CDCl_3$, 20 ºC): δ 169.4 (**C**$_{16'}$), 166.0 (**C**$_{16}$), 165.4 (**C**$_{13}$), 165.0 (**C**$_{13'}$), 153.1 (**C**$_{Cbz}$=O), 150.3 (**C**$_{Cbz'}$=O), 141.5 (**C**$_9$), 138.8 (**C**$_{7'}$), 136.2 (2C, **C**$_{4'}$, Ph_{Cbz}-*i*-**C**), 134.9 (3C, **C**$_4$, **C**$_{9'}$, $Ph_{Cbz'}$-*i*-**C**), 129.3 (**C**$_7$), 129.0 (2C, Ph_{Cbz}-**C**), 128.8 (Ph_{Cbz}-**C**), 128.6 (Ph_{Cbz}-**C**), 128.3 (Ph_{Cbz}-**C**), 128.2 (Ph_{Cbz}-**C**), 124.8

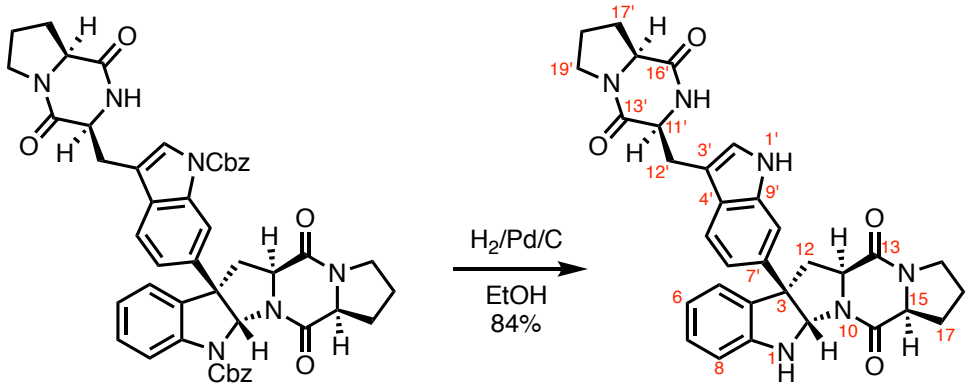
(**C**$_{2'}$), 124.7 (**C**$_{7'}$), 124.0 (**C**$_5$), 121.1 (**C**$_6$), 119.5 (**C**$_{5'}$), 117.6 (**C**$_8$), 115.7 (**C**$_{3'}$), 113.0 (**C**$_{8'}$), 81.8 (**C**$_2$), 69.1 (Ph$_{Cbz'}$-**C**H$_2$), 68.3 (Ph$_{Cbz}$-**C**H$_2$), 60.7 (**C**$_{11}$ or **C**$_{15'}$), 60.6 (**C**$_{11}$ or **C**$_{15'}$), 59.3 (**C**$_{15}$), 58.6 (**C**$_3$), 54.0 (**C**$_{11'}$), 45.6 (**C**$_{19}$ or **C**$_{19'}$), 45.5 (**C**$_{19}$ or **C**$_{19'}$), 39.4 (**C**$_{12}$), 28.4 (**C**$_{17'}$), 28.0 (**C**$_{17}$), 26.5 (**C**$_{12'}$), 23.1 (**C**$_{18'}$), 22.7 (**C**$_{18}$). FTIR (thin film) cm^{-1}: 3472 (br-m), 2955 (w), 1713 (s), 1667 (s), 1480 (m), 1398 (m). HRMS (ESI) (*m*/*z*): calc'd for $C_{48}H_{45}N_6O_8$ $[M+H]^+$: 833.3293, found: 833.3278. $[\alpha]_D^{24}$: –166 (*c* = 0.26, $CHCl_3$). TLC (70% acetone in hexanes), R*f*: 0.28 (UV, CAM).



(–)-**14** (–)-naseseazine C (**4**)

(−)-Naseseazine C (4): Palladium on activated charcoal (3.3 mg, 3.1 μmol, 20 mol%) was added to a solution of para(−)-N^{in},$N^{in'}$-dicarboxybenzyl naseseazine C **14** (12.9 mg, 30.4 mmol, 1 equiv) in degassed (N_2 stream, 5 min) ethanol (1.6 mL). A stream of hydrogen gas was passed through the heterogeneous mixture for 2 min by discharge of a balloon equipped with a needle extending into the reaction mixture. After stirring the reaction mixture at 22 °C for 21 h under an atmosphere of hydrogen gas, the solution was degassed (N_2 stream, 5 min) and filtered through Celite. The filter cake was washed with methanol (3 × 2 mL) and the combined filtrates were concentrated under reduced pressure. The crude reaction mixture was purified by flash column chromatography on silica gel (eluent: 70→ 90% acetone in dichloromethane) to afford (–)-**4** (7.0 mg, 80%). Structural assignments were made using additional information from gCOSY, HSQC, HMBC and nOe experiments. ^{1}H NMR (MHz, CD_3OD, 20 ºC): δ 7.64 (d, *J* = 1.8 Hz, 1H, $C_{5'}$**H**), 7.24 (d, *J* = 8.6 Hz, 1H, $C_{8'}$**H**), 7.19 (d, *J* = 7.4 Hz, 1H, C_5**H**), 7.13 (dd, *J* = 1.9, 8.6 Hz, 1H, $C_{7'}$**H**), 7.08 (s, 1H, $C_{2'}$**H**), 7.06 (td, *J* = 1.2, 7.6 Hz, 1H, C_7**H**), 6.75 (td, *J* = 1.0, 7.5 Hz, 1H, C_6**H**), 6.71 (d, *J* = 8.1 Hz, 1H, C_8**H**), 5.70 (s, 1H, C_2**H**), 4.39 (td, *J* = 1.8, 4.9 Hz, 1H, $C_{11'}$**H**), 4.31–4.23 (m, 2H, C_{11}**H**, C_{15}**H**), 3.97 (ddd, *J* = 1.9, 6.4, 10.9 Hz, 1H, $C_{15'}$**H**), 3.54–3.47 (m, 2H, C_{19}**H**), 3.45–3.39 (m, 1H, $C_{19'}$**H**$_a$), 3.28 (app-d, *J* = 5.2 Hz, 2H, $C_{12'}$**H**), 3.26–3.21 (m, 1H, $C_{19'}$**H**$_b$), 3.12 (dd, *J* = 6.1, 12.7 Hz, 1H, C_{12}**H**$_a$), 2.95–2.88 (m, 1H, C_{12}**H**$_b$), 2.33–2.25 (m, 1H, C_{17}**H**$_a$), 2.10–2.00 (m, 2H, C_{17}**H**$_b$, C_{18}**H**$_a$), 1.98–1.87 (m, 2H, $C_{17'}$**H**$_a$, C_{18}**H**$_b$), 1.71–1.60 (m, 1H, $C_{18'}$**H**$_a$), 1.48–1.40 (m, 1H, $C_{18'}$**H**$_b$), 0.98–0.88 (m, 1H, $C_{17'}$**H**$_b$). ^{1}H NMR (MHz, DMSO-d_6, 20 ºC): δ 10.84 (s, 1H,

$N_{1'}$**H**), 7.74 (s, 1H, $N_{10'}$**H**), 7.63 (s, 1H, $C_{5'}$**H**), 7.22, (d, *J* = 8.6 Hz, 1H, $C_{8'}$**H**), 7.17 (s, 1H, $C_{8'}$**H**), 7.16 (d, *J* = 8.7 Hz, 1H, C_5**H**), 7.03 (d, *J* = 8.6 Hz, 1H, $C_{7'}$**H**), 6.98 (app-t, *J* = 7.7 Hz, 1H, C_7**H**), 6.65 (d, *J* = 8.6 Hz, C_8**H**), 6.64 (s, 1H, N_1**H**), 6.61 (app-t, *J* = 7.4 Hz, 1H, C_6**H**), 5.63 (s, 1H, C_2**H**), 4.35–4.27 (m, 2H, $C_{11'}$**H**, C_{15}**H**), 4.16 (dd, *J* = 6.1, 11.4 Hz, 1H, C_{11}**H**), 4.09 (app-t, *J* = 8.4 Hz, 1H, $C_{15'}$**H**), 3.45–3.33 (m, 3H, C_{19}**H**, $C_{19'}$$\mathbf{H}_a$), 3.33–3.23 (m, 1H, $C_{19'}$$\mathbf{H}_b$), 3.22 (dd, *J* = 4.8, 14.9 Hz, 1H, $C_{12'}$$\mathbf{H}_a$), 3.10–3.02 (m, 2H, $C_{12}$$\mathbf{H}_a$, $C_{12'}$$\mathbf{H}_b$), 2.68 (app-t, *J* = 12.0 Hz, 1H, $C_{12}$$\mathbf{H}_b$), 2.19–2.08 (m, 1H, $C_{17}$$\mathbf{H}_a$), 2.04–1.97 (m, 1H, $C_{17'}$$\mathbf{H}_a$), 1.97–1.86 (m, 2H, $C_{17}$$\mathbf{H}_b$, $C_{18}$$\mathbf{H}_a$), 1.85–1.76 (m, 1H, $C_{18}$$\mathbf{H}_b$), 1.75–1.68 (m, 1H, $C_{18'}$$\mathbf{H}_a$), 1.68–1.59 (m, 1H, $C_{18'}$$\mathbf{H}_b$), 1.53–1.42 (m, 1H, $C_{17'}$$\mathbf{H}_b$). ^{13}C NMR (MHz, CD_3OD, 20 ºC): δ 170.7 (**C**$_{16'}$), 168.5 (**C**$_{16}$), 168.0 (**C**$_{13}$), 167.4 (**C**$_{13'}$), 150.6 (**C**$_9$), 136.9 (**C**$_{9'}$), 135.3 (**C**$_{6'}$), 134.2 (**C**$_4$), 129.6 (**C**$_7$), 128.5 (**C**$_{4'}$), 126.2 (**C**$_{2'}$), 125.2 (**C**$_5$), 121.1 (**C**$_{7'}$), 120.5 (**C**$_6$), 116.3 (**C**$_{5'}$), 112.6 (**C**$_{8'}$), 111.1 (**C**$_8$), 109.9 (**C**$_{3'}$), 83.6 (**C**$_2$), 61.9 (**C**$_{15}$), 61.7 (**C**$_{11}$), 60.8 (**C**$_3$), 60.0 (**C**$_{15'}$), 57.2 (**C**$_{11'}$), 46.3 (**C**$_{19}$), 46.0 (**C**$_{19'}$), 41.1 (**C**$_{12}$), 29.1 (2C, **C**$_{12'}$, **C**$_{17'}$), 28.9 (**C**$_{17}$), 23.8 (**C**$_{18}$), 22.6 (**C**$_{18'}$). ^{13}C NMR (MHz, DMSO-*d*$_6$, 20 ºC)[1]: δ 169.2 (**C**$_{16'}$), 166.0 (**C**$_{16}$), 165.6 (**C**$_{13'}$), 165.3 (**C**$_{13}$), 149.3 (**C**$_9$), 134.8 (**C**$_{9'}$), 134.1 (**C**$_{6'}$), 133.0 (**C**$_4$), 128.1 (**C**$_7$), 127.1 (**C**$_{4'}$), 124.8 (**C**$_{2'}$), 123.8 (**C**$_5$), 119.2 (**C**$_{7'}$), 118.1 (**C**$_6$), 114.6 (**C**$_{5'}$), 111.4 (**C**$_{8'}$), 109.6 (**C**$_{3'}$), 109.5 (**C**$_8$), 81.1 (**C**$_2$), 60.1 (**C**$_{11}$), 59.7 (**C**$_{15}$), 58.7 (**C**$_3$), 58.5 (**C**$_{15'}$), 55.1 (**C**$_{11'}$), 44.7 (2C, **C**$_{19}$, **C**$_{19'}$), 39.9 (**C**$_{12}$), 27.7 (**C**$_{17'}$), 27.4 (**C**$_{17}$), 25.4 (**C**$_{12'}$), 22.6 (**C**$_{18}$), 22.0 (**C**$_{18'}$). FTIR (thin film) cm^{-1}: 3310 (br-m), 2926 (m), 1661 (s), 1435 (s), 1301 (w), 1150 (w). HRMS (ESI) (*m*/*z*): calc'd for $C_{32}H_{33}N_6O_4$ $[M+H]^+$:565.2558, found: 565.2551. $[\alpha]_D^{24}$: –180 (*c* = 0.16, CH_3OH).[2] TLC (80% acetone in dichloromethane), R*f*: 0.11 (UV, CAM).



(–)-**15** (–)-*iso*-naseseazine C (**16**)

(−)-*iso*-Naseseazine C (16): Palladium on activated charcoal (6.5 mg, 6.1 μmol, 20.0 mol%) was added to a solution of meta(−)-N^{in},$N^{in'}$-dicarboxybenzyl *iso*-naseseazine C (–)-**15** (25.2 mg, 30.3 μmol, 1 equiv) in degassed (N_2 stream, 5 min) ethanol (3 mL). A stream of hydrogen gas was passed through the heterogeneous mixture for 2 min by discharge of a balloon equipped with a needle extending into the reaction mixture. After stirring the reaction mixture at 22 °C for 64 h under an atmosphere of hydrogen gas, the solution was degassed (N_2 stream, 5 min) and filtered over Celite. The filter cake was washed with methanol (3 × 2 mL) and the combined filtrates were concentrated under reduced pressure. The crude reaction mixture was purified by flash column chromatography on silica gel (eluent: 70→ 90% acetone in dichloromethane) to afford (–)-**8** (14.3 mg, 83.6%). Structural assignments were made using additional information from gCOSY, HSQC, HMBC and nOe experiments. ^{1}H NMR (MHz, CD_3OD, 20 ºC): δ 7.49 (d, *J* = 8.4 Hz, 1H, $C_{5'}$**H**), 7.31 (d, *J* = 1.6 Hz, 1H, $C_{8'}$**H**), 7.11 –7.01 (m, 2H, C_5**H**, $C_{6'}$**H**), 7.06 (s, 1H, $C_{2'}$**H**), 7.05 (td, *J* = 1.1, 7.1 Hz, 1H, C_7**H**), 6.72 (td, *J* = 1.0, 7.5 Hz, 1H, C_6**H**), 6.70 (d, *J* = 7.7 Hz, 1H, C_8**H**), 5.64 (s, 1H, C_2**H**), 4.35 (td, *J* = 1.8, 5.0 Hz, 1H, $C_{11'}$**H**), 4.21–4.15 (m, 2H, C_{11}**H**, C_{15}**H**), 3.92 (ddd, *J* = 1.9, 6.5, 11.0 Hz, 1H, $C_{15'}$**H**), 3.54–3.44 (m, 2H, C_{19}**H**), 3.44–3.37 (m, 1H, $C_{19'}$**H**$_a$), 3.26 (app-d, *J* = 4.5 Hz, 2H, $C_{12'}$**H**), 3.24–3.19 (m, 1H, $C_{19'}$**H**$_b$), 3.05 (dd, *J* = 6.0, 12.6 Hz, 1H, C_{12}**H**$_a$), 2.85–2.78 (m, 1H, C_{12}**H**$_b$), 2.29–2.21 (m, 1H, C_{17}**H**$_a$), 2.08–1.97 (m, 2H, C_{17}**H**$_b$, C_{18}**H**$_a$), 1.95–1.85 (m, 2H, $C_{17'}$**H**$_a$, C_{18}**H**$_b$), 1.67–1.56 (m, 1H, $C_{18'}$**H**$_a$), 1.46–1.30 (m, 1H, $C_{18'}$**H**$_b$), 0.96–0.86 (m, 1H, $C_{17'}$**H**$_b$). ^{1}H NMR (MHz, DMSO-d_6, 20 ºC): δ 10.79 (d, *J* = 2.5 Hz, 1H, $N_{1'}$**H**), 7.68 (s, 1H,

$N_{10'}$**H**), 7.50 (d, *J* = 8.4 Hz, 1H, $C_{5'}$**H**), 7.23 (d, *J* = 1.6 Hz, 1H, $C_{8'}$**H**), 7.17 (d, *J* = 7.4 Hz, 1H, C_5**H**), 7.15 (d, *J* = 2.4 Hz, 1H, $C_{2'}$**H**), 7.08 (dd, *J* = 1.8, 8.3 Hz, 1H, $C_{6'}$**H**), 6.99 (app-t, *J* = 7.8 Hz, 1H, C_7**H**), 6.65 (d, *J* = 4.7 Hz, 1H, C_8**H**), 6.65 (s, 1H, N_1**H**), 6.63 (d, *J* = 7.5 Hz, 1H, C_6**H**), 5.56 (s, 1H, C_2**H**), 4.31 (ddd, *J* = 1.8, 6.9, 9.1 Hz, 1H, C_{15}**H**), 4.26 (app-t, *J* = 5.3 Hz, 1H, $C_{11'}$**H**), 4.17 (ddd, *J* = 3.2, 6.6, 10.1 Hz, 1H, C_{11}**H**), 4.05 (dd, *J* = 6.9, 9.7 Hz, 1H, $C_{15'}$**H**), 3.45–3.30 (m, 3H, $C_{19'}$**H**$_a$, C_{19}**H**), 3.25 (td, *J* = 4.3, 7.7, 8.3 Hz, 1H, $C_{19'}$**H**$_b$), 3.20 (app-dd, *J* = 4.9, 15.0 Hz, 1H, $C_{12'}$**H**$_a$), 3.03 (app-ddd, *J* = 6.1, 13.8, 21.2 Hz, 2H, $C_{12'}$**H**$_b$, C_{12}**H**$_a$), 2.70–2.61 (m, 1H, C_{12}**H**$_b$), 2.19–2.07 (m, 1H, $C_{17'}$**H**$_a$), 2.01–1.86 (m, 3H, C_{17}**H**, C_{18}**H**$_a$), 1.86–1.73 (m, 1H, C_{18}**H**$_b$), 1.73–1.63 (m, 1H, $C_{18'}$**H**$_a$), 1.63–1.56 (m, 1H, $C_{18'}$**H**$_b$), 1.46–1.35 (m, 1H, $C_{17'}$**H**$_b$). ^{13}C NMR (MHz, CD_3OD, 20 ºC): δ 170.8 (**C**$_{16'}$), 168.5 (**C**$_{16}$), 168.0 (**C**$_{13}$), 167.4 (**C**$_{13'}$), 150.6 (**C**$_9$), 138.2 (**C**$_{7'}$), 137.9 (**C**$_{9'}$), 133.8 (**C**$_4$), 129.7 (**C**$_7$), 127.6 (**C**$_{4'}$), 126.1 (**C**$_{2'}$), 125.1 (**C**$_5$), 120.4 (**C**$_{5'}$ or **C**$_6$), 120.3 (**C**$_{5'}$ or **C**$_6$), 118.0 (**C**$_{6'}$), 111.1 (**C**$_8$), 109.5 (**C**$_{3'}$), 109.3 (**C**$_{8'}$), 83.5 (**C**$_2$), 61.8 (**C**$_{11}$), 61.6 (**C**$_{15}$), 60.7 (**C**$_3$), 60.0 (**C**$_{15'}$), 57.1 (**C**$_{11'}$), 46.3 (**C**$_{19}$), 46.0 (**C**$_{19'}$), 40.9 (**C**$_{12}$), 29.1 (**C**$_{12'}$ or **C**$_{17'}$), 29.0 (**C**$_{12'}$ or **C**$_{17'}$), 28.8 (**C**$_{17}$), 23.8 (**C**$_{18}$), 22.5 (**C**$_{18'}$). ^{13}C NMR (MHz, DMSO-*d*$_6$, 20 ºC): δ 169.1 (**C**$_{16'}$), 166.1 (**C**$_{16}$), 165.5 (**C**$_{13'}$), 165.3 (**C**$_{13}$), 149.4 (**C**$_9$), 136.8 (**C**$_{7'}$), 135.9 (**C**$_{9'}$), 132.4 (**C**$_4$), 128.2 (**C**$_7$), 126.1 (**C**$_{4'}$), 124.8 (**C**$_{2'}$), 124.0 (**C**$_5$), 119.0 (**C**$_{5'}$), 118.1 (**C**$_6$), 116.4 (**C**$_{6'}$), 109.5 (**C**$_8$), 109.2 (**C**$_{3'}$), 108.0 (**C**$_{8'}$), 81.5 (**C**$_2$), 60.2 (**C**$_{11}$), 59.6 (**C**$_{15}$), 58.8 (**C**$_3$), 58.4 (**C**$_{15'}$), 55.2 (**C**$_{11'}$), 44.8 (**C**$_{19}$ or **C**$_{19'}$), 44.6 (**C**$_{19}$ or **C**$_{19'}$), 39.9 (**C**$_{12}$), 27.6 (**C**$_{17'}$), 27.4 (**C**$_{17}$), 25.7 (**C**$_{12'}$), 22.6 (**C**$_{18}$), 21.9 (**C**$_{18'}$). FTIR (thin film) cm^{-1}: 3310 (br-m), 2955 (w), 1653 (s), 1457 (m), 1340 (w), 1056 (w). HRMS (ESI) (*m*/*z*): calc'd for $C_{32}H_{33}N_6O_4$ $[M+H]^+$: 565.2558, found: 565.2554. $[\alpha]_D^{24}$: –312 (*c* = 0.42, CH_3OH).[3] TLC (80% acetone in dichloromethane), R*f*: 0.11 (UV, CAM).

Supplementary Table 2-4. Comparison of our ^{1}H NMR data for synthetic (–)-naseseazine C (4) and (–)-*iso*-naseseazine C (16) as well as biocatalytic (–)-naseseazine C (4) with literature data (CD_3OD):[7]

Assignment	**Wang Report**[11] "New *iso*-naseseazine B", ^{1}H NMR, 600 MHz, CD_3OD	**This Work (Synthetic)** (–)-naseseazine C (**4**) ^{1}H NMR, 500 MHz, CD_3OD, 20 ºC	**This Work (Synthetic)** (–)-*iso*-naseseazine C (**16**) ^{1}H NMR, 500 MHz, CD_3OD, 20 ºC	**This Work (Biocatalytic)** (–)-naseseazine C (**4**) ^{1}H NMR, 600 MHz, CD_3OD
N1	–	–	–	–
C2	5.71 (s)	5.70 (s)	5.64 (s)	5.76 (s)
C3	–	–	–	–
C4	–	–	–	–
C5	7.19 (d, *J* = 7.5 Hz)	7.19 (d, *J* = 7.4 Hz)	7.11–7.01 (m)	7.15 (d, *J* = 7.7 Hz)
C6	6.76 (dt, *J* = 0.8, 7.5 Hz)	6.75 (td, *J* = 1.0, 7.5 Hz)	6.72 (td, *J* = 1.0, 7.5 Hz)	6.71 (t, *J* = 7.5 Hz)
C7	7.07 (dt, *J* = 1.1, 7.5 Hz)	7.06 (td, *J* = 1.2, 7.6 Hz)	7.05 (td, *J* = 1.1, 7.1 Hz)	7.03 (td, *J* = 1.3, 7.7 Hz)
C8	6.72 (d, *J* = 7.8 Hz)	6.71 (d, *J* = 8.1 Hz)	6.70 (d, *J* = 7.7 Hz)	6.68 (d, *J* = 7.9 Hz)
C9	–	–	–	–
N10	–	–	–	–
C11	4.32 (app-t, *J* = 11.4 Hz)	4.31–4.23 (m)	4.21–4.15 (m)	4.27 (td, *J* = 2.0, 8.4 Hz)
C12	2.94 (app-t, *J* = 11.4 Hz) 3.12 (dd, *J* = 6.2, 12.6 Hz)	2.95–2.88 (m) 3.12 (dd, *J* = 6.1, 12.7 Hz)	2.85–2.78 (m) 3.05 (dd, *J* =6.0, 12.6 Hz)	2.89 (dd, *J* = 11.2, 12.7 Hz) 3.09 (dd, *J* = 6.0, 12.7 Hz)
C13	–	–	–	–
N14	–	–		–
C15	4.28 (app-t, *J*	4.31–4.23 (m)	4.21–4.15 (m)	4.24–4.21 (m)

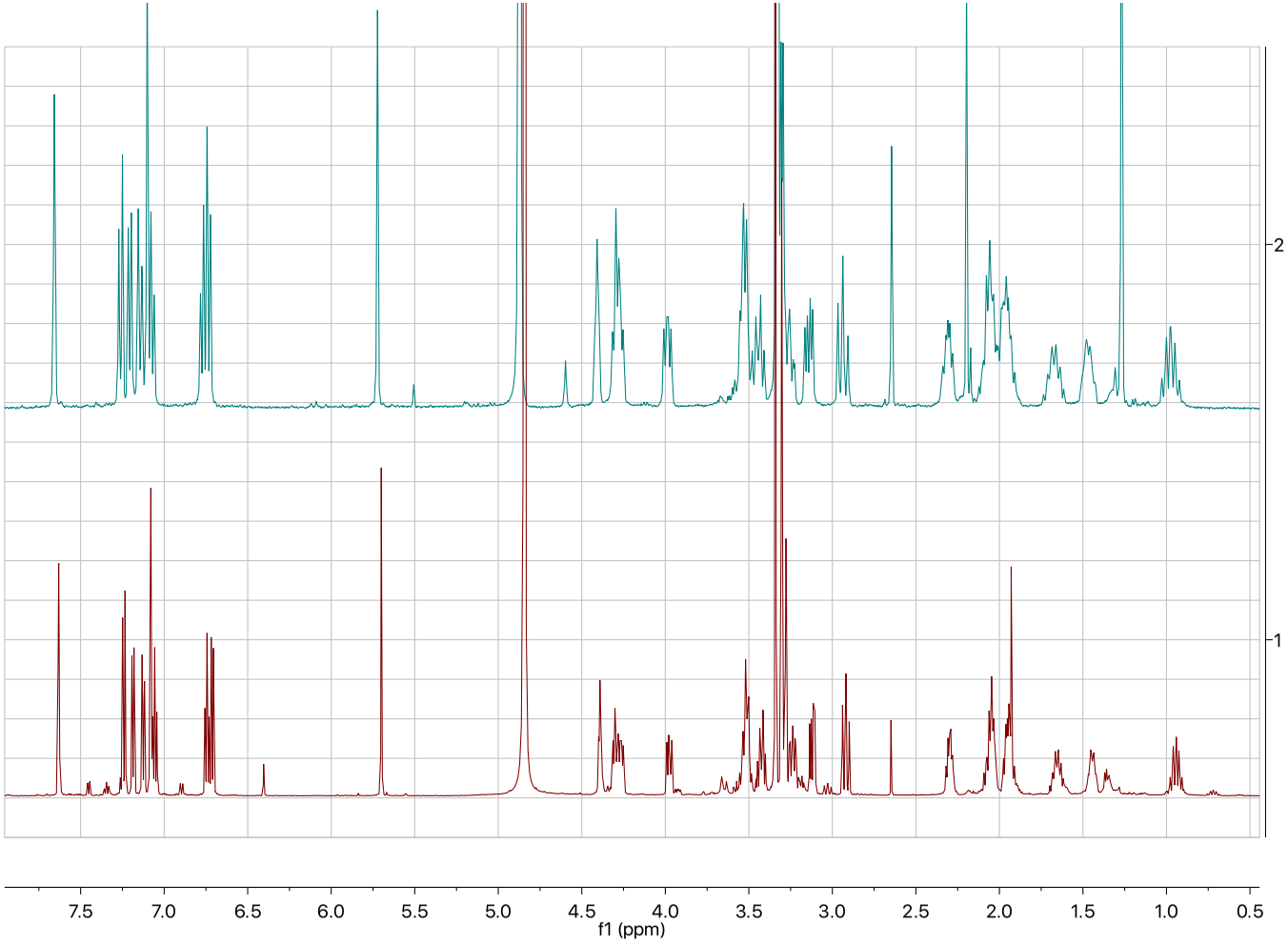
[7] The comparison of the ^{1}H and ^{13}C NMR chemical shifts of the reported "new (–)-*iso*-naseseazine B" with both synthetic regioisomers (–)-**7** and (–)-**8** in CD_3OD shows that the data for "new (–)-*iso*-naseseazine B" is most consistent with synthetic (–)-naseseazine C (**7**).

	=8.2 Hz)			
C16	–	–	–	–
C17	2.06 (m) 2.31 (m)	2.10–2.00 (m) 2.33–2.25 (m)	2.08–1.97 (m) 2.29–2.21 (m)	2.07–2.04 (m) 2.31–2.22 (m)
C18	1.95 (m) 2.06 (m)	1.98–1.87 (m) 2.10–2.00 (m)	1.95–1.85 (m) 2.08–1.97 (m)	1.98–1.85 (m) 2.05–1.97 (m)
C19	3.52 (m)	3.54–3.47 (m)	3.54–3.44 (m)	3.52–3.45 (m)
N1'	–	–	–	–
C2'	7.09 (s)	7.08 (s)	7.06 (s)	7.05 (s)
C3'	–	–	–	–
C4'	–	–	–	–
C5'	7.24 (d, *J* = 8.6 Hz)	7.64 (d, *J* = 1.8 Hz)[8]	7.49 (d, *J* = 8.4 Hz)	7.60 (d, *J* = 1.8 Hz)
C6'	7.12 (dd, *J* = 1.6, 8.6 Hz)	–[13]	7.11–7.01 (m)	–[13]
C7'	–	7.13 (dd, *J* = 1.9, 8.6 Hz)[13]	–	7.09 (dd, *J* = 1.8, 8.5 Hz)
C8'	7.64 (d, *J* = 1.3 Hz)	7.24 (d, *J* = 8.6 Hz)[13]	7.31 (d, *J* = 1.6 Hz)	7.21 (d, *J* = 8.6 Hz)
C9'	–	–	–	–
N10'	–	–	–	–
C11'	4.40 (br-t, *J* = 4.8 Hz)	4.39 (td, *J* = 1.8, 4.9 Hz)	4.35 (td, *J* = 1.8, 5.0 Hz)	4.36 (td, *J* = 1.7, 5.0 Hz)
C12'	3.33 (m)	3.28 (app-d, *J* = 5.2 Hz)	3.26 (app-d, *J* = 4.5 Hz)	3.25 (d, *J* = 5.1 Hz)
C13'	–	–	–	–
N14'	–	–	–	–
C15'	3.97 (ddd, *J* = 1.5, 6.4, 10.8 Hz)	3.97 (ddd, *J* = 1.9, 6.4, 10.9 Hz)	3.92 (ddd, *J* = 1.9, 6.5, 11.0 Hz)	3.94 (ddd, *J* = 1.8, 6.4, 11.0 Hz)
C16'	–	–	–	–
C17'	0.89 (m) 1.95 (m)	0.98–0.88 (m) 1.98–1.87 (m)	0.96–0.86 (m) 1.95–1.85 (m)	0.99–0.87 (m) 1.98–1.85 (m)
C18'	1.42 (m) 1.65 (m)	1.48–1.40 (m) 1.71–1.60 (m)	1.46–1.30 (m) 1.67–1.56 (m)	1.47–1.38 (m) 1.75–1.55 (m)

[8] Our assignments of these resonances are supported by key HSQC and HMBC correlations.

C19'	3.28 (m) 3.43 (m)	3.26–3.21 (m) 3.45–3.39 (m)	3.24–3.19 (m) 3.44–3.37 (m)	3.24–3.18 (m) 3.43–3.36 (m)

Supplementary Figure 2-3. Stacked ^{1}H NMR spectra of synthetic (top) and biocatalytic (bottom) (–)-naseseazine C (7) (CD_3OD):



Supplementary Table 2-5. Comparison of our ^{13}C NMR data for (–)-naseseazine C (4) and (–)-*iso*-naseseazine C (16) with literature data (CD_3OD):[9]

Assignment	**Wang Report**[11] "New *iso*-naseseazine B" ^{13}C NMR, 150 MHz, CD_3OD	**This Work** (–)-naseseazine C (**4**) ^{13}C NMR, 125 MHz, CD_3OD, 20 ºC	**This Work (Biocatalytic)** (–)-naseseazine C (**4**) ^{13}C NMR, 150 MHz, CD_3OD, 20 ºC	**This Work** (–)-*iso*-naseseazine C (**16**) ^{13}C NMR, 125 MHz CD_3OD, 20 ºC	**Chemical Shift Difference** Δδ = δ (–)-naseseazine C (this work) – δ (Wang report)	**Chemical Shift Difference** Δδ = δ (–)-*iso*-naseseazine C (this work) – δ (Wang report)	**Chemical Shift Difference** Δδ = δ **Biocatalytic** (–)-naseseazine C (this work) – δ (Wang report)
C2	83.6	83.6	83.6	83.5	0	–0.1	0
C3	60.8	60.8	60.8	60.7	0	–0.1	0
C4	134.2	134.2	134.2	133.8	0	–0.4	0
C5	125.2	125.2	125.2	125.1	0	–0.1	0
C6	120.5	120.5	120.5	120.4 or 120.3	0	–0.1 or –0.2	0
C7	129.7	129.6	129.7	129.7	–0.1	0	-0.1
C8	111.2	111.1	111.11	111.1	–0.1	–0.1	-0.1
C9	150.6	150.6	150.6	150.6	0	0	0
C11	62.0	61.9	61.9	61.8	–0.1	–0.2	-0.1
C12	41.1	41.1	41.1	40.9	0	–0.2	0
C13	168.1	168.0	168.1	168.0	–0.1	–0.1	0

[9] The comparison of the ^{1}H and ^{13}C NMR chemical shifts of the reported "new (–)-*iso*-naseseazine B" with both synthetic regioisomers (–)-**7** and (–)-**8** in CD_3OD shows that the data for "new (–)-*iso*-naseseazine B" is most consistent with synthetic (–)-naseseazine C (**7**).

C15	61.7	61.9	61.7	61.6	0.2	–0.1	0
C16	168.6	168.5	168.6	168.5	–0.1	–0.1	0
C17	28.9	28.9	28.9	28.8	0	–0.1	0
C18	23.9	23.8	23.9	23.8	–0.1	–0.1	0
C19	46.4	46.3	46.3	46.3	–0.1	–0.1	-0.1
C2'	126.3	126.2	126.2	126.1	–0.1	–0.2	-0.1
C3'	109.8	109.9	109.9	109.5	0.1	–0.3	0.1
C4'	128.5	128.5	128.5	127.6	0	–0.9	0
C5'	112.6	116.3	116.3	120.4 or 120.3	3.7[10]	7.8 or 7.7	3.7[15]
C6'	121.1	135.3	135.3	118.0	14.2[15]	–3.1	14.2[15]
C7'	135.3	121.1	121.1	138.2	–14.2[15]	2.9	-14.2[15]
C8'	116.3	112.6	112.6	109.3	–3.7[15]	–7.0	-3.7[15]
C9'	136.9	136.9	137	137.9	0	1.0	0.1
C11'	57.3	57.2	57.2	57.1	–0.1	–0.2	-0.1
C12'	29.3	29.1	29.1	29.1 or 29.0	–0.2	–0.2 or –0.3	-0.2
C13'	167.4	167.4	167.4	167.4	0	0	0
C15'	60.1	60.0	60.1	60.0	–0.1	0.1	0
C16'	170.7	170.7	170.7	170.8	0	0.1	0
C17'	29.1	29.1	28.9	29.1 or 29.0	0	0 or –0.1	-0.2
C18'	22.5	22.6	22.6	22.5	0.1	0	0.1
C19'	46.03	46.0	46	46.0	0	0	0

[10] Our assignments of these resonances are supported by key HSQC and HMBC correlations.

Supplementary Table 2-6. Comparison of our ^{13}C NMR data for (–)-naseseazine C (4) and (–)-*iso*-naseseazine C (16) with literature data (CD_3OD):[11]

Wang Report[11] "New *iso*-naseseazine B" ^{13}C NMR, 150 MHz, CD_3OD	**This Work** (–)-naseseazine C (**4**) ^{13}C NMR, 125 MHz, CD_3OD, 20 ºC	**This Work** (–)-*iso*-naseseazine C (**16**) ^{13}C NMR, 125 MHz CD_3OD, 20 ºC	**Chemical Shift Difference** Δδ = δ **synthetic** (–)-naseseazine C (this work) – δ (Wang report)	**Chemical Shift Difference** Δδ = δ **synthetic** (–)-*iso*-naseseazine C (this work) – δ (Wang report)
170.7	170.7	170.8	0	0.1
168.6	168.5	168.5	–0.1	–0.1
168.1	168.0	168.0	–0.1	–0.1
167.4	167.4	167.4	0	0
150.6	150.6	150.6	0	0
136.9	136.9	138.2	0	1.3
135.3	135.3	137.9	0	2.6
134.2	134.2	133.8	0	–0.4
129.7	129.6	129.7	–0.1	0
128.5	128.5	127.6	0	–0.9
126.3	126.2	126.1	–0.1	–0.2
125.2	125.2	125.1	0	–0.1
121.1	121.1	120.4 or 120.3	0	–0.7 or –0.8
120.5	120.5	120.4 or 120.3	0	–0.1 or –0.2
116.3	116.3	118.0	0	1.7

[11] The comparison of the ^{1}H and ^{13}C NMR chemical shifts of the reported "new (–)-*iso*-naseseazine B" with both synthetic regioisomers (–)-**7** and (–)-**8** in CD_3OD shows that the data for "new (–)-*iso*-naseseazine B" is most consistent with synthetic (–)-naseseazine C (**7**).

112.6	112.6	111.1	0	–1.5
111.2	111.1	109.5	–0.1	–1.7
109.8	109.9	109.3	0.1	–0.5
83.6	83.6	83.5	0	–0.1
62.0	61.9	61.8	–0.1	–0.2
61.7	61.9	61.6	0.2	–0.1
60.8	60.8	60.7	0	–0.1
60.1	60.0	60.0	–0.1	–0.1
57.3	57.2	57.1	–0.1	–0.2
46.4	46.3	46.3	–0.1	–0.1
46.03	46.0	46.0	–0.03	–0.03
41.1	41.1	40.9	0	–0.2
29.3	29.1	29.1 or 29.0	–0.2	–0.2 or –0.3
29.1	29.1	29.1 or 29.0	0	0 or –0.1
28.9	28.9	28.8	0	–0.1
23.9	23.8	23.8	–0.1	–0.1
22.5	22.6	22.5	0.1	0

Supplementary Table 2-7. Comparison of our ^{1}H NMR data for (–)-naseseazine C (4) and (–)-*iso*-naseseazine C (16) with literature data (DMSO-d_6):[12]

Assignment	**Carroll Report**[10] naseseazine C ^{1}H NMR, 800 MHz, DMSO-d_6	**This Work** (–)-naseseazine C (**4**) ^{1}H NMR, 500 MHz, DMSO-d_6, 20 ºC	**This Work** (–)-*iso*-naseseazine C (**16**) ^{1}H NMR, 500 MHz, DMSO-d_6, 20 ºC
N1	6.56 (s)	6.64 (s)	6.65 (s)
C2	5.62 (s)	5.63 (s)	5.56 (s)
C3	–	–	–
C4	–	–	–
C5	7.16 (dd, *J* = 1.2, 7.4 Hz)	7.16 (d, *J* = 8.7 Hz)	7.17 (d, *J* = 7.4 Hz)
C6	6.61 (td, *J* = 1.0, 7.4 Hz)	6.61 (app-t, *J* = 6.4 Hz)	6.63 (d, *J* = 4.7 Hz)
C7	6.98 (td, *J* = 1.2, 7.7 Hz)	6.98 (app-t, *J* = 7.7 Hz)	6.99 (app-t, *J* = 7.8 Hz)
C8	6.65 (dd, *J* =1.5, 7.8 Hz)	6.65 (d, *J* = 8.6 Hz)	6.65 (d, *J* = 4.7 Hz)
C9	–	–	–
N10	–	–	–
C11	4.16 (m)	4.16 (dd, *J* = 6.1, 11.4 Hz)	4.17 (ddd, *J* = 3.2, 6.6, 10.1 Hz)
C12	2.69 (m) 3.04 (m)	2.68 (app-t, *J* = 12.0 Hz) 3.10–3.02 (m)	2.70–2.61 (m) 3.03 (app-ddd, *J* = 6.1, 13.8, 21.2 Hz)
C13	–	–	–
N14	–	–	–
C15	4.29 (t)	4.35–4.27 (m)	4.31 (ddd, *J* = 1.8, 6.9, 9.1 Hz)
C16	–	–	–
C17	1.94 (m)	1.97–1.86 (m)	2.01–1.86 (m)

[12] The comparison of the ^{1}H and ^{13}C NMR chemical shifts of the reported naseseazine C with both synthetic regioisomers (–)-**7** and (–)-**8** in DMSO-d_6 shows that the data for naseseazine C is most consistent with synthetic (–)-naseseazine C (**7**).

	2.14 (m)	2.19–2.08 (m)	
C18	1.81 (m) 1.91 (m)	1.85–1.76 (m) 1.97–1.86 (m)	2.01–1.86 (m) 1.86–1.73 (m)
C19	3.35 (m) 3.42 (m)	3.45–3.33 (m)	3.45–3.30 (m)
N1'	10.82 (d, *J* =2.4 Hz)	10.84 (s)	10.79 (d, *J* = 2.5 Hz)
C2'	7.17 (d, *J* = 2.4 Hz)	7.17 (s)	7.15 (d, *J* = 2.4 Hz)
C3'	–	–	–
C4'	–	–	–
C5'	7.62 (d, *J* = 1.8 Hz)	7.63 (s)	7.50 (d, *J* = 8.4 Hz)
C6'	–	–	7.08 (dd, *J* = 1.8, 8.3 Hz)
C7'	7.03 (dd, *J* = 1.8, 8.5 Hz)	7.03 (d, *J* = 8.6 Hz)	–
C8'	7.22 (d, *J* = 8.5 Hz)	7.22 (d, *J* = 8.6 Hz)	7.23 (d, *J* = 1.6 Hz)
C9'	–	–	–
N10'	7.70 (s)	7.74 (s)	7.68 (s)
C11'	4.30 (m)	4.35–4.27 (m)	4.26 (app-t, *J* = 5.3 Hz)
C12'	3.08 (m) 3.21 (m)	3.10–3.02 (m) 3.22 (dd, *J* = 4.8, 14.9 Hz)	3.03 (app-ddd, *J* = 6.1, 13.8, 21.2 Hz) 3.20 (app-dd, *J* = 4.9, 15.0 Hz)
C13'	–	–	–
N14'	–	–	–
C15'	4.08 (m)	4.09 (app-t, *J* = 8.4 Hz)	4.05 (dd, *J* = 6.9, 9.7 Hz)
C16'	–	–	–
C17'	1.44 (m) 2.01 (m)	1.53–1.42 (m) 2.04–1.97 (m)	1.46–1.35 (m) 2.19–2.07 (m)
C18'	1.61 (m) 1.69 (m)	1.68–1.59 (m) 1.75–1.68 (m)	1.63–1.56 (m) 1.73–1.63 (m)
C19'	3.28 (m) 3.40 (m)	3.34–3.23 (m) 3.45–3.33 (m)	3.25 (td, *J* = 4.3, 7.7, 8.3 Hz) 3.45–3.30 (m)

Supplementary Table 2-8. Comparison of our ^{13}C NMR data for (–)-naseseazine C (4) and (–)-*iso*-naseseazine C (16) with literature data (DMSO-d_6):[13]

Assignment	**Carroll Report**[10] naseseazine C ^{13}C NMR, 200 MHz, DMSO-d_6	**This Work** (–)-naseseazine C (**4**) ^{13}C NMR, 125 MHz DMSO-d_6, 20 ºC	**This Work** (–)-*iso*-naseseazine C (**16**) ^{13}C NMR, 125 MHz DMSO-d_6, 20 ºC	**Chemical Shift Difference** Δδ = δ (–)-naseseazine C (this work) – δ (Carroll report)	**Chemical Shift Difference** Δδ = δ (–)-*iso*-naseseazine C (this work) – δ (Carroll report)
C2	81.1	81.1	81.5	0	–0.4
C3	59.7	58.7	58.8	–1.0[14]	0.9
C4	134.1	133.0	132.4	–1.1[16]	1.7
C5	125.8	123.8	124.0	–2.0[16]	1.8
C6	118.1	118.1	118.1	0	0
C7	128.1	128.1	128.2	0	–0.1
C8	111.4	109.5	109.5	–1.9[16]	1.9
C9	149.3	149.3	149.4	0	–0.1
C11	60.1	60.1	60.2	0	–0.1
C12	40.0	39.9	39.9	–0.1	0.1
C13	165.3	165.3	165.3	0	0
C15	58.5	59.7	59.6	1.2[16]	–1.1
C16	169.1	166.0	166.1	–3.1[16]	3.0
C17	27.4	27.4	27.4	0	0

[13] The comparison of the ^{1}H and ^{13}C NMR chemical shifts of the reported naseseazine C with both synthetic regioisomers (–)-**7** and (–)-**8** in DMSO-d_6 shows that the data for naseseazine C is most consistent with synthetic (–)-naseseazine C (**7**).

[14] Our assignment of these resonances is supported by key HSQC and HMBC correlations.

C18	22.5	22.6	22.6	0.1	–0.1
C19	44.7	44.7	44.8 or 44.6	0	0.1 or –0.1
C2'	124.8	124.8	124.8	0	0
C3'	109.6	109.6	109.2	0	0.4
C4'	127.1	127.1	126.1	0	1.0
C5'	114.6	114.6	119.0	0	–4.4
C6'	134.8	134.1	116.4	–0.7[16]	18.4
C7'	119.2	119.2	136.8	0	–17.6
C8'	110.4	111.4	108.0	1.0[16]	2.4
C9'	135.9	134.8	135.9	–1.1[16]	0
C11'	55.1	55.1	55.2	0	–0.1
C12'	25.4	25.4	25.7	0	–0.3
C13'	166.0	165.6	165.5	–0.4[16]	0.5
C15'	58.8	58.5	58.4	–0.3[16]	0.4
C16'	165.6	169.2	169.1	3.6[16]	–3.5
C17'	27.7	27.7	27.6	0	0.1
C18'	21.9	22.0	21.9	0.1	0
C19'	44.7	44.7	44.8 or 44.6	0	0.1 or –0.1

Supplementary Table 2-9. Comparison of our ^{13}C NMR data for (–)-naseseazine C (4) and (–)-*iso*-naseseazine C (16) with literature data (DMSO-d_6):[15]

Carroll Report[10] naseseazine C ^{13}C NMR, 200 MHz, DMSO-d_6	**This Work** (–)-naseseazine C (**7**) ^{13}C NMR, 125 MHz DMSO-d_6, 20 ºC	**This Work** (–)-*iso*-naseseazine C (**8**) ^{13}C NMR, 125 MHz DMSO-d_6, 20 ºC	**Chemical Shift Difference** Δδ = δ (–)-naseseazine C (this work) – δ (Carroll report)	**Chemical Shift Difference** Δδ = δ (–)-*iso*-naseseazine C (this work) – δ (Carroll report)
169.1	169.2	169.1	–0.1	0
166.0	166.0	166.1	0	0.1
165.6	165.6	165.5	0	–0.1
165.3	165.3	165.3	0	0
149.3	149.3	149.4	0	0.1
135.9	–	136.8	N/A[16]	0.9
134.8	134.8	135.9	0	1.1
134.1	134.1	132.4	0	–1.7
–	133.0	–	N/A[17]	–
128.1	128.1	128.2	0	0.1
127.1	127.1	–	0	N/A[18]
125.8	–	126.1	N/A[18]	0.3
124.8	124.8	124.8	0	0

[15] The comparison of the ^{1}H and ^{13}C NMR chemical shifts of the reported naseseazine C with both synthetic regioisomers (–)-**7** and (–)-**8** in DMSO-d_6 shows that the data for naseseazine C is most consistent with synthetic (–)-naseseazine C (**7**).

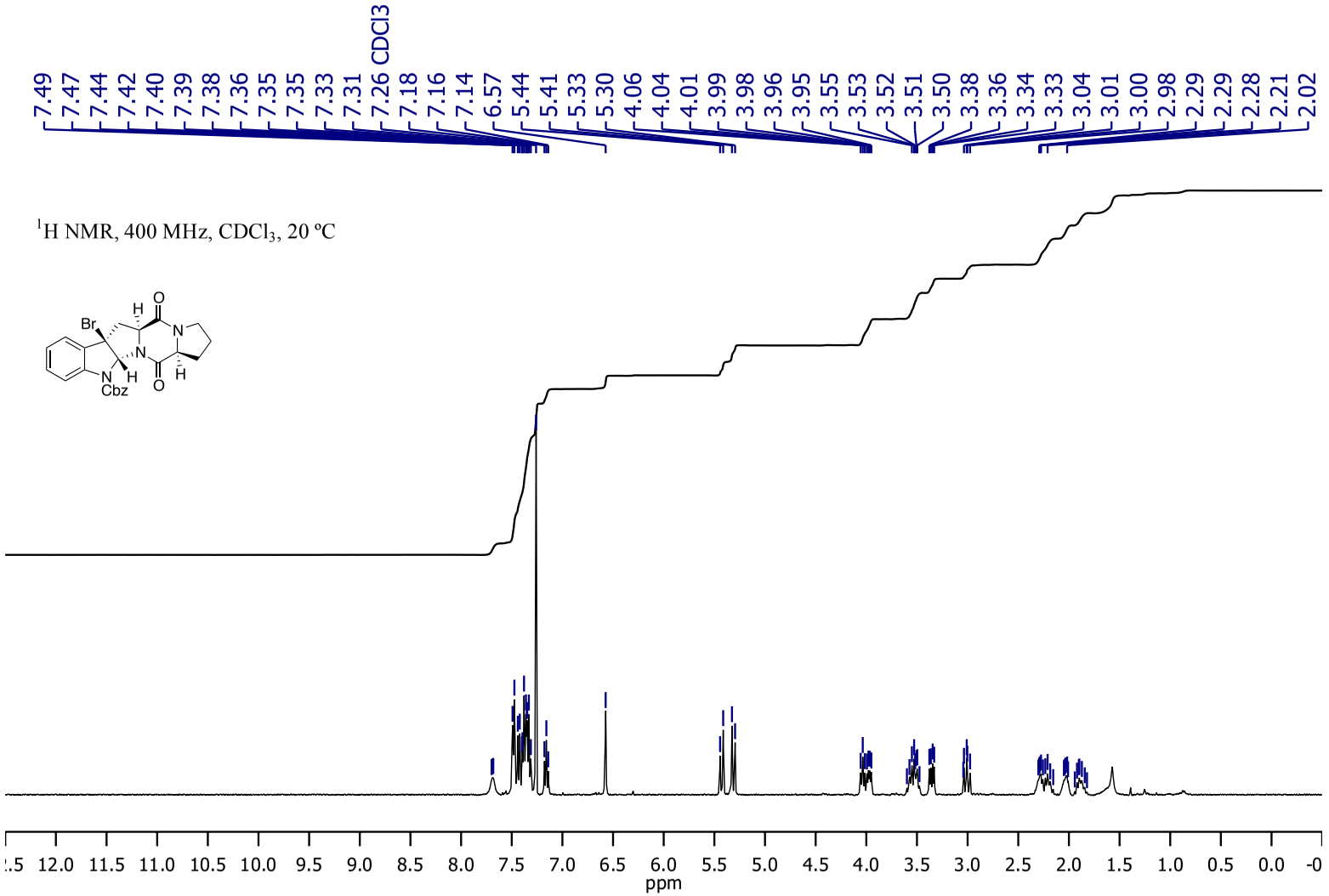
[16] We do not observe signals at 145.9, 125.8, and 110.4 ppm in our data, which has an excellent signal-to-noise ratio.

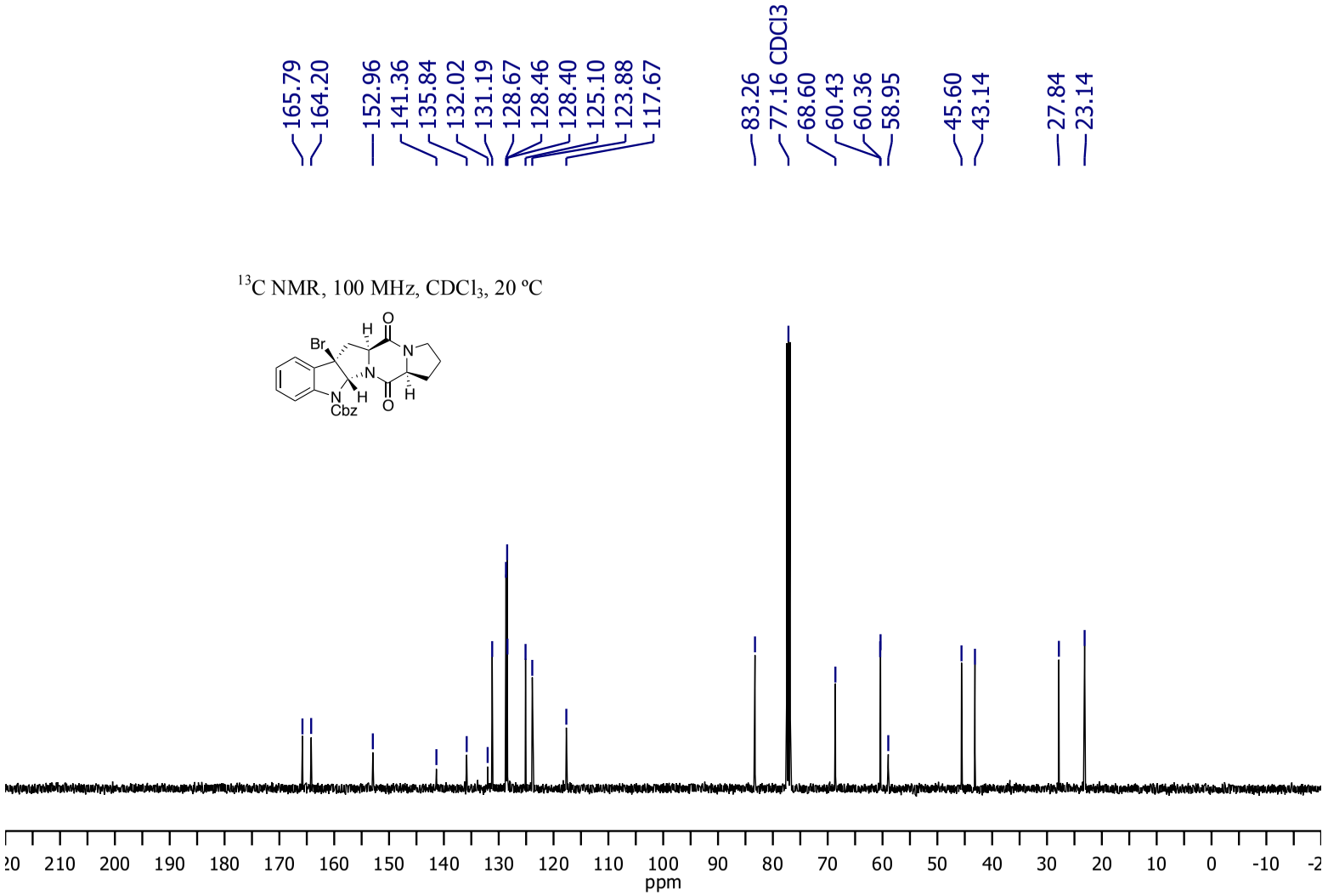
[17] We observe signals at 133.0, 123.8, and 109.5 ppm in our ^{13}C NMR spectrum, which has an excellent signal-to-noise ratio.

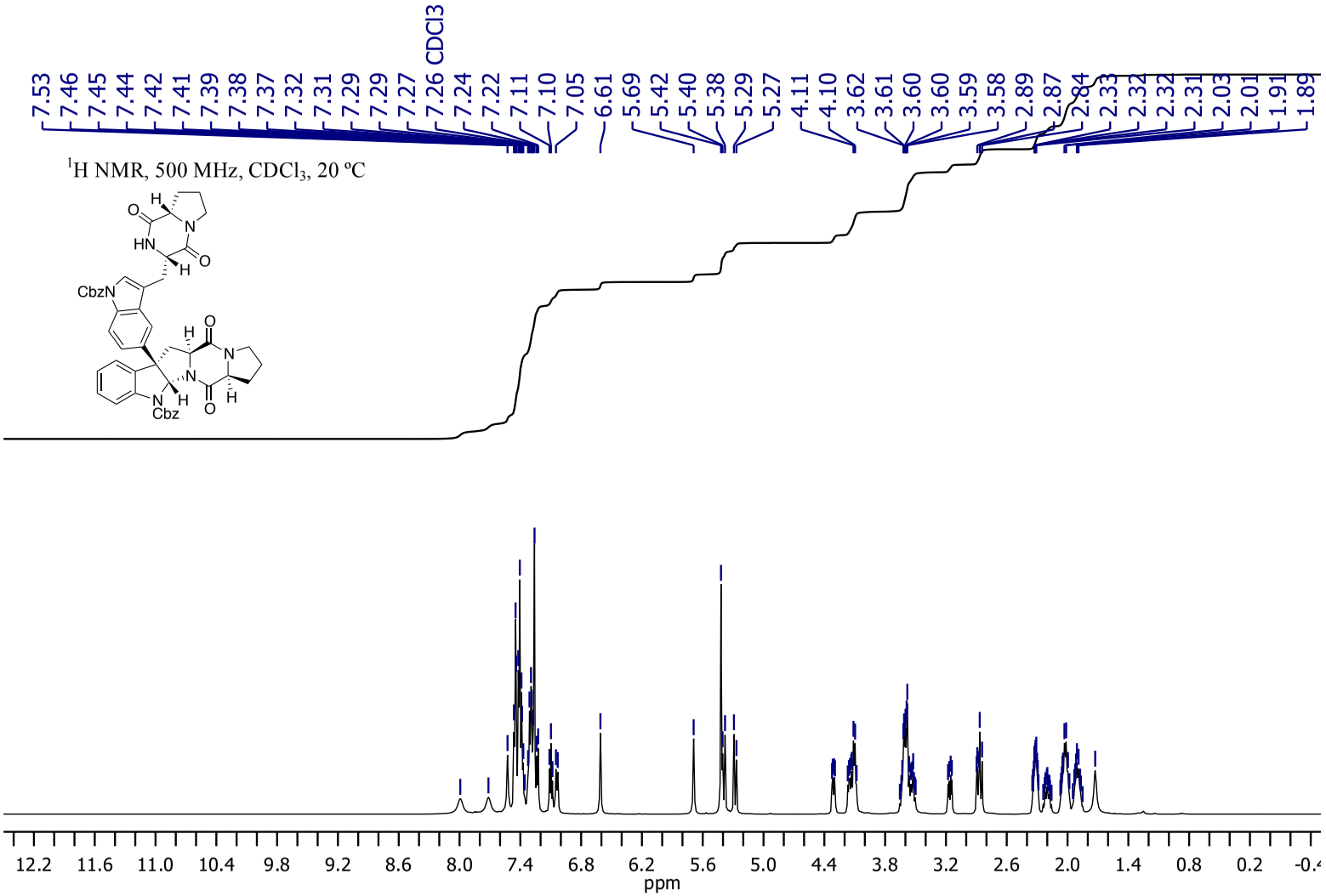
[18] We do not observe signals at 127.1 and 111.4 ppm in our data, which has an excellent signal-to-noise ratio.

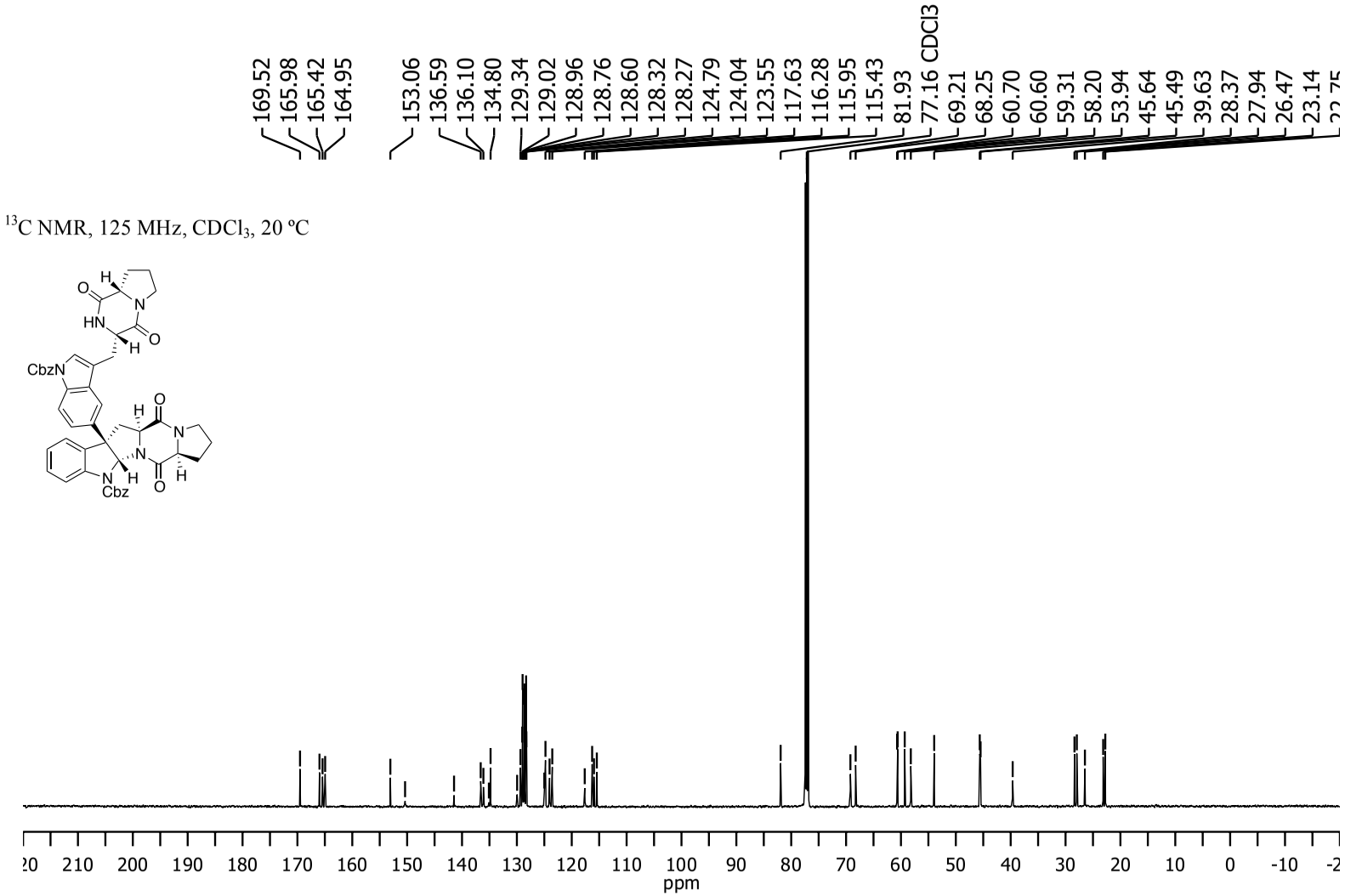
–	123.8	124.0	N/A[17]	N/A[19]
119.2	119.2	119.0	0	–0.2
118.1	118.1	118.1	0	0
114.6	114.6	116.4	0	1.8
111.4	111.4	–	0	N/A[20]
110.4	–	109.5	N/A[18]	–0.9
109.6	109.6	109.2	0	–0.4
–	109.5	108.0	N/A[19]	N/A[21]
81.1	81.1	81.5	0	0.4
60.1	60.1	60.2	0	0.1
59.7	59.7	59.6	0	–0.1
58.8	58.7	58.8	–0.1	0
58.5	58.5	58.4	0	–0.1
55.1	55.1	55.2	0	0.1
44.7	44.7	44.8 or 44.6	0	0.1 or –0.1
44.7	44.7	44.8 or 44.6	0	0.1 or –0.1
40.0	39.9	39.9	–0.1	–0.1
27.7	27.7	27.6	0	–0.1
27.4	27.4	27.4	0	0
25.4	25.4	25.7	0	0.3
22.5	22.6	22.6	0.1	0.1
21.9	22.0	21.9	0.1	0

[19] We observe signals at 124.0 and 108.0 ppm in our ^{13}C NMR spectrum, which has an excellent signal-to-noise ratio.

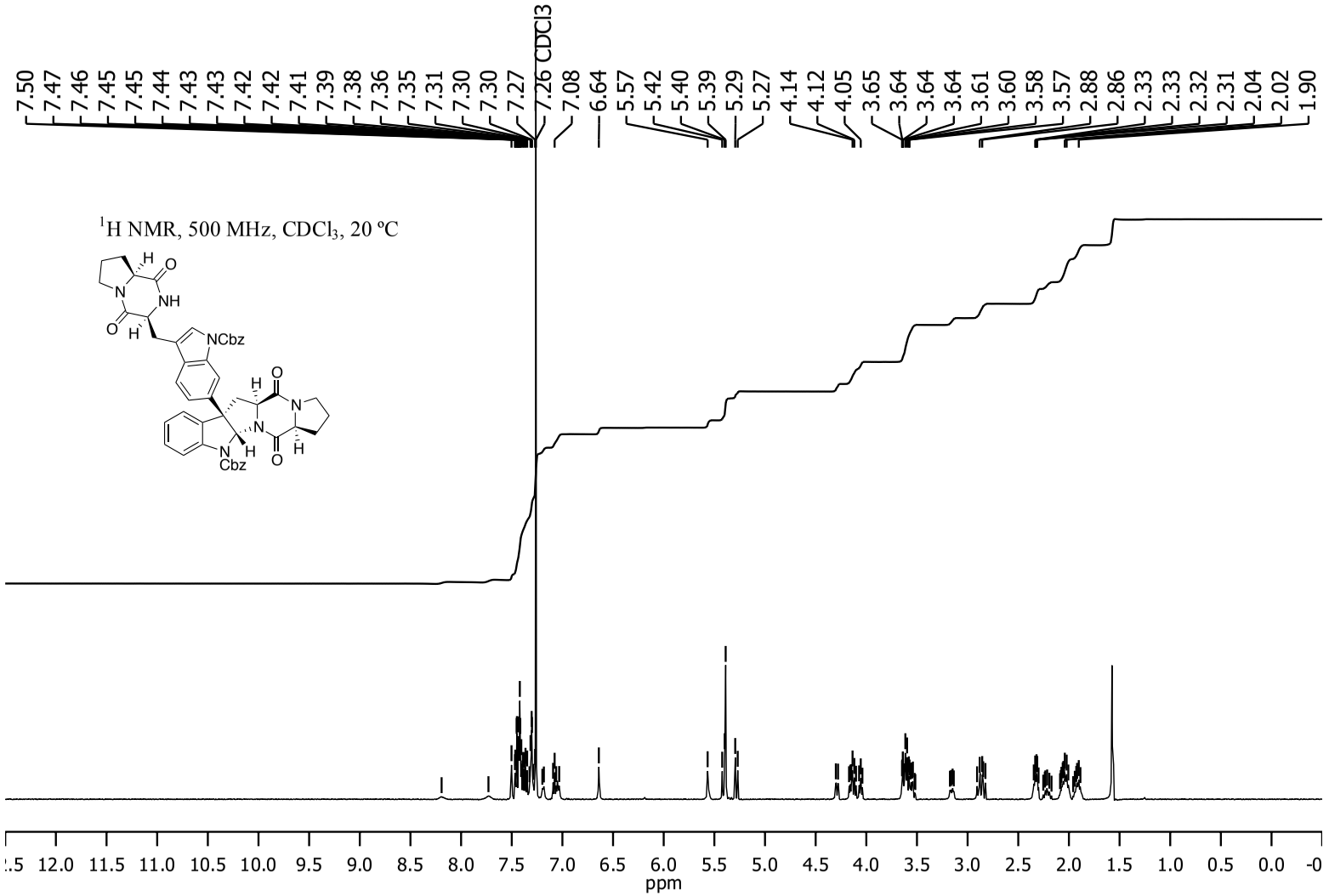

7.49
7.47
7.44
7.42
7.40
7.39
7.38
7.36
7.35
7.35
7.33
7.31
7.26 CDCl3
7.18
7.16
7.14
6.57
5.44
5.41
5.33
5.30
4.06
4.04
4.01
3.99
3.98
3.96
3.95
3.55
3.53
3.52
3.51
3.50
3.38
3.36
3.34
3.33
3.04
3.01
3.00
2.98
2.29
2.29
2.28
2.21
2.02
1H NMR, 400 MHz, CDCl3, 20 ºC
Br
H
O
N
N
H
N
H
O
H
Cbz
ppm

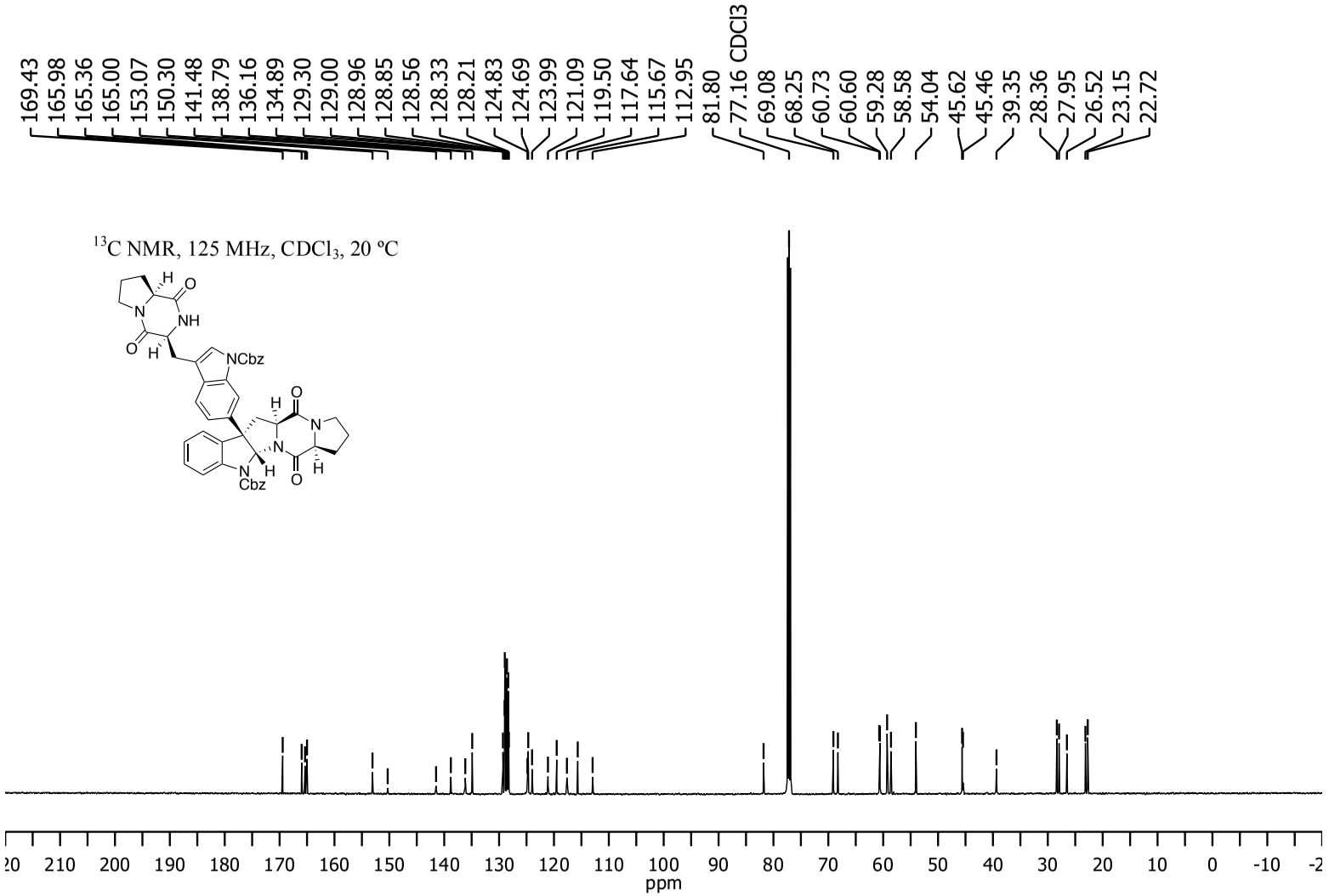

165.79
164.20
152.96
141.36
135.84
132.02
131.19
128.67
128.46
128.40
125.10
123.88
117.67
83.26
77.16 CDCl3
68.60
60.43
60.36
58.95
45.60
43.14
27.84
23.14
13C NMR, 100 MHz, CDCl3, 20 ºC
Br
H
O
N
N
H
N
H
Cbz
O
H
ppm

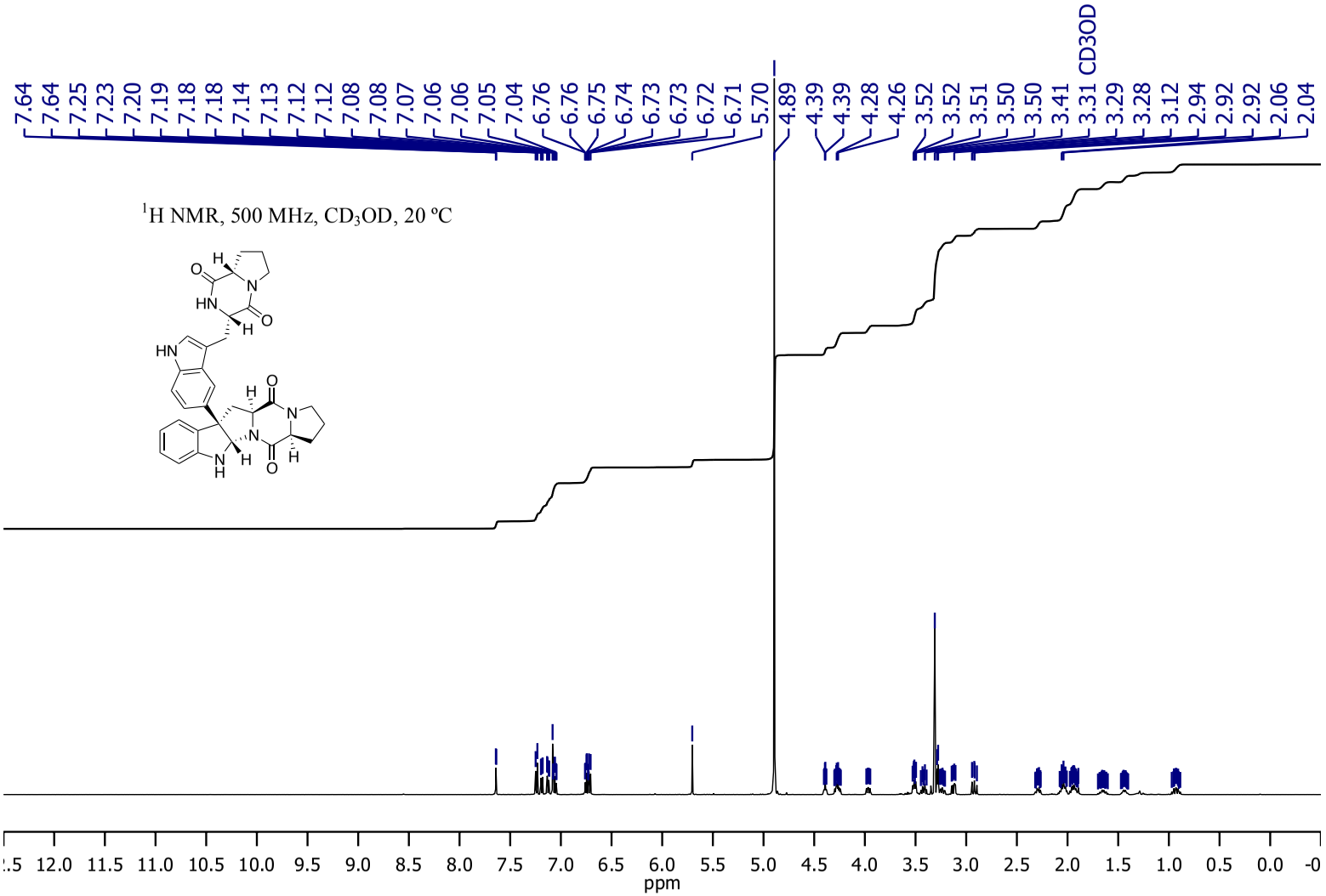

7.53
7.46
7.45
7.44
7.42
7.41
7.39
7.38
7.37
7.32
7.31
7.29
7.29
7.27
7.26 CDCl3
7.24
7.22
7.11
7.10
7.05
6.61
5.69
5.42
5.40
5.38
5.29
5.27
4.11
4.10
3.62
3.61
3.60
3.60
3.59
3.58
2.89
2.87
2.84
2.33
2.32
2.32
2.31
2.03
2.01
1.91
1.89
^{1}H NMR, 500 MHz, $CDCl_3$, 20 ºC
CbzN
Cbz
12.2
11.6
11.0
10.4
9.8
9.2
8.6
8.0
7.4
6.8
6.2
5.6
5.0
4.4
3.8
3.2
2.6
2.0
1.4
0.8
0.2
ppm

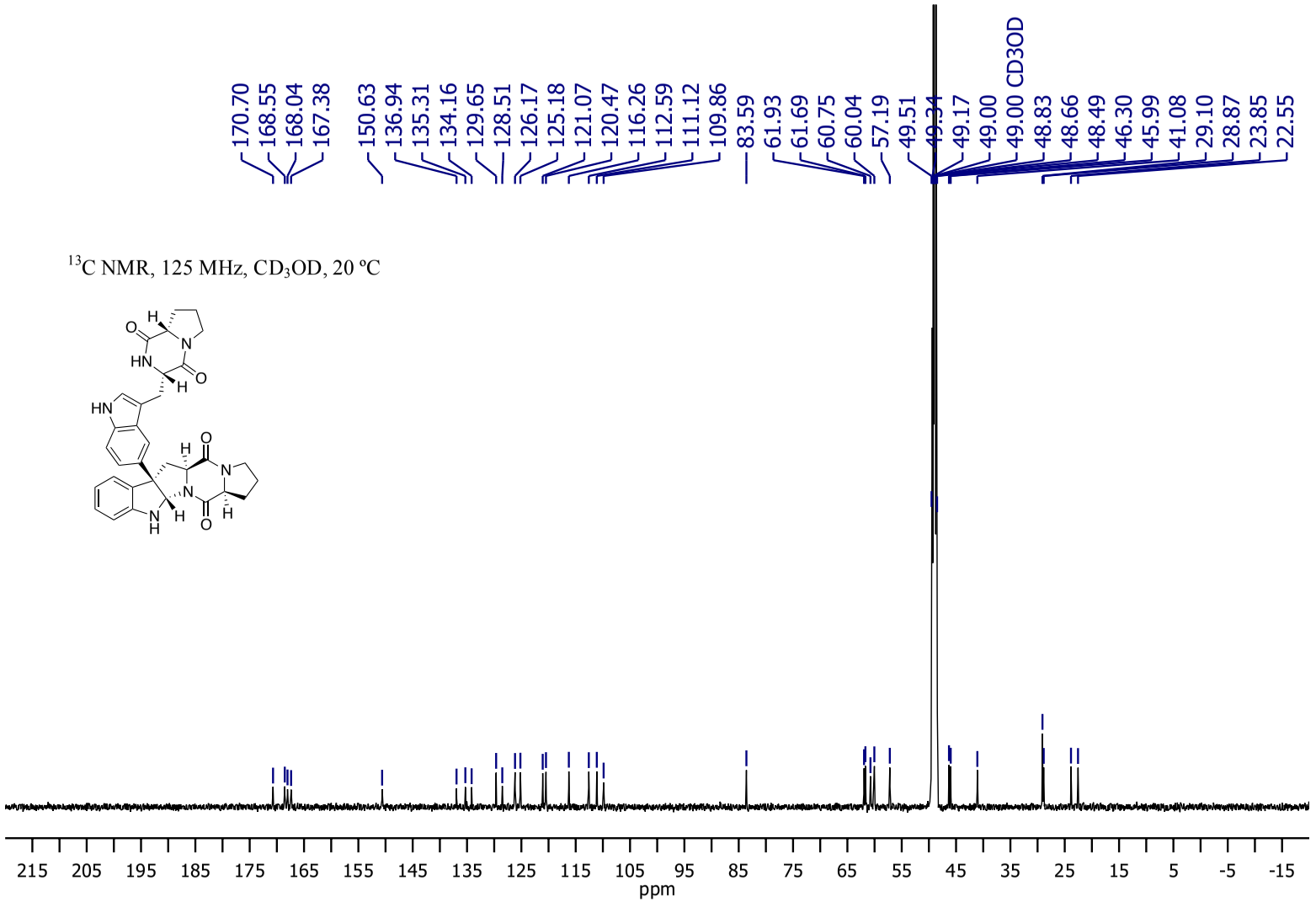


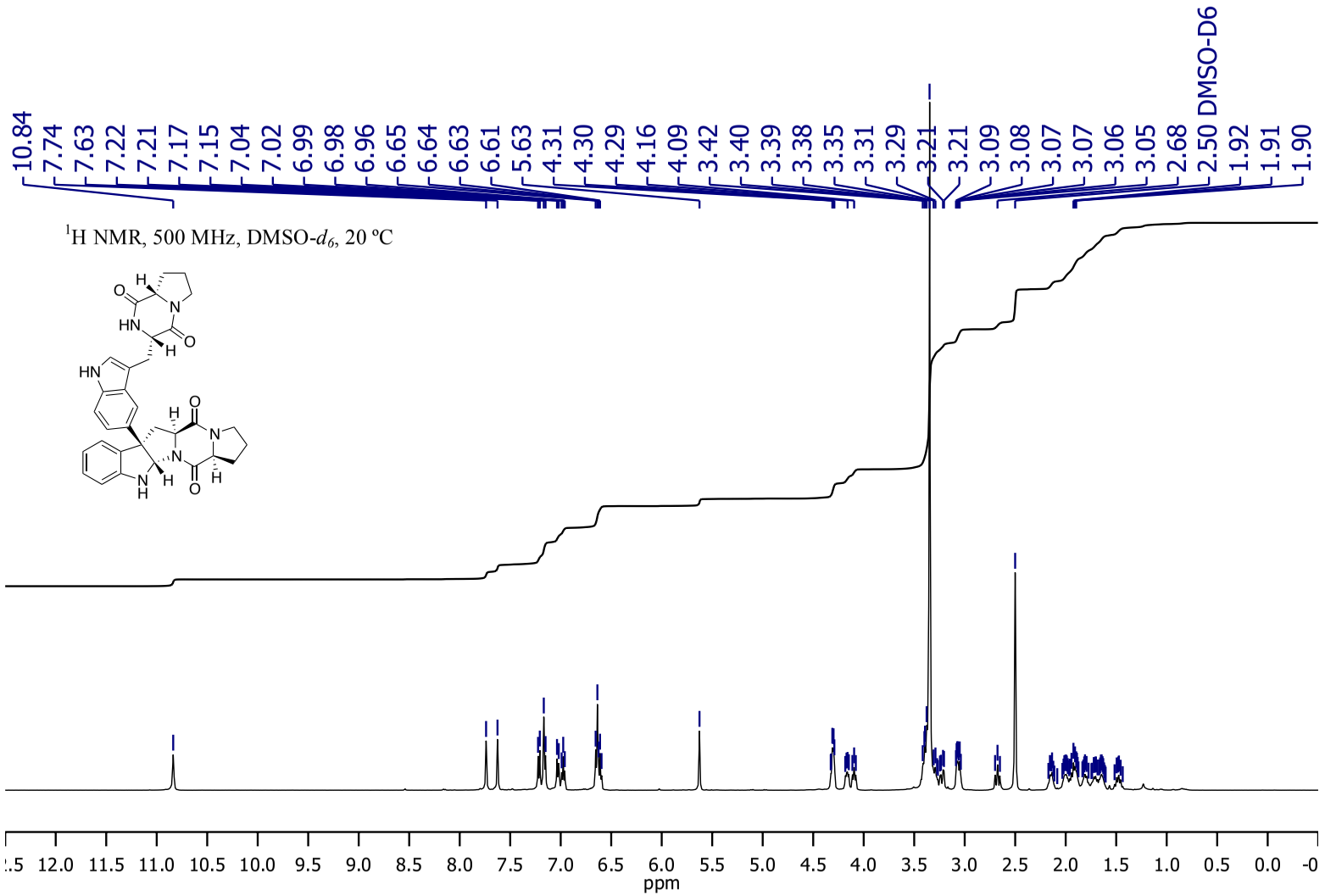
^{13}C NMR, 125 MHz, $CDCl_3$, 20 ºC

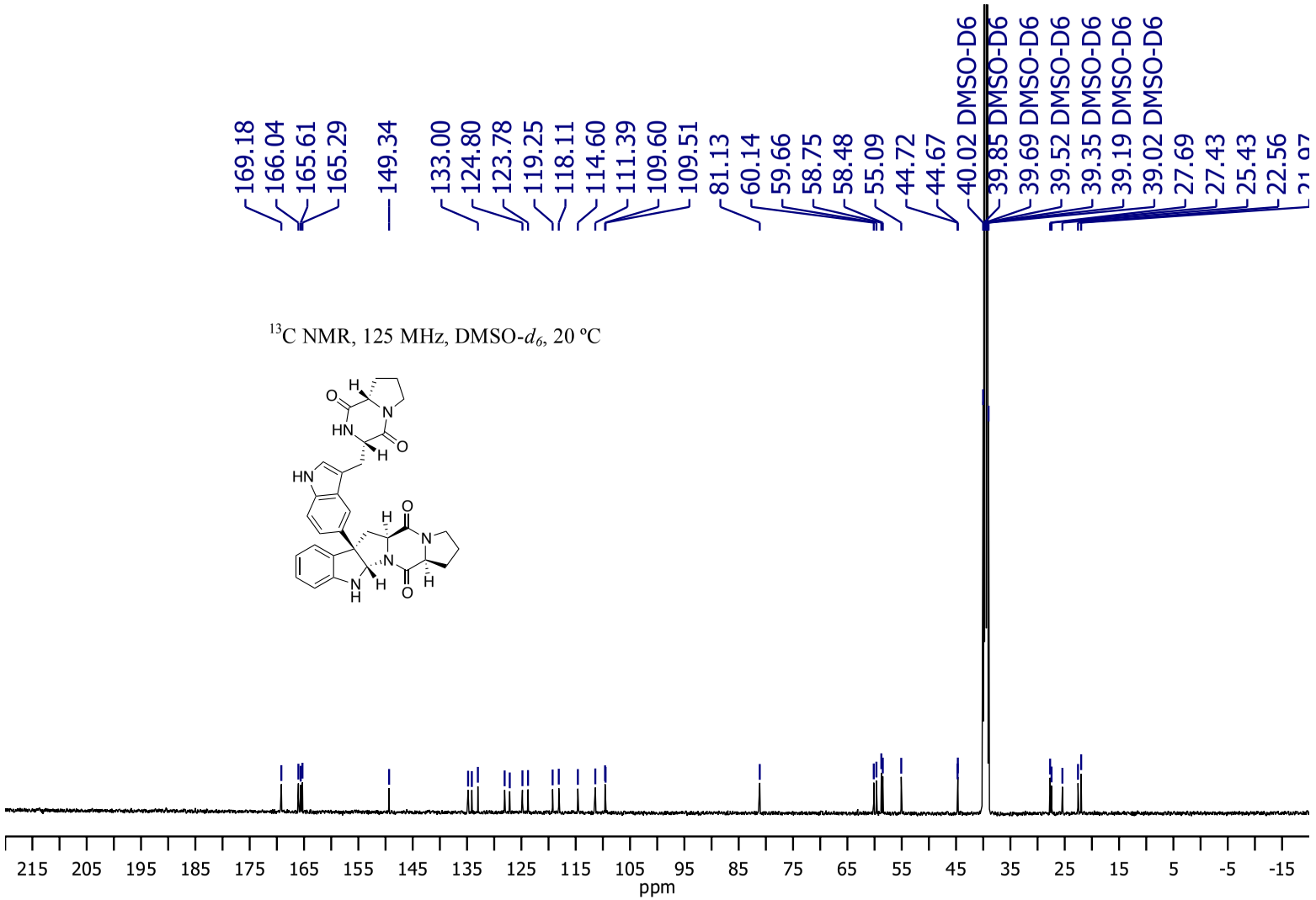

7.50
7.47
7.46
7.45
7.45
7.44
7.43
7.43
7.42
7.42
7.41
7.39
7.38
7.36
7.35
7.31
7.30
7.30
7.27
7.26 CDCl3
7.08
6.64
5.57
5.42
5.40
5.39
5.29
5.27
4.14
4.12
4.05
3.65
3.64
3.64
3.64
3.61
3.60
3.58
3.57
2.88
2.86
2.33
2.33
2.32
2.31
2.04
2.02
1.90
^{1}H NMR, 500 MHz, $CDCl_3$, 20 ºC
NCbz
Cbz
ppm

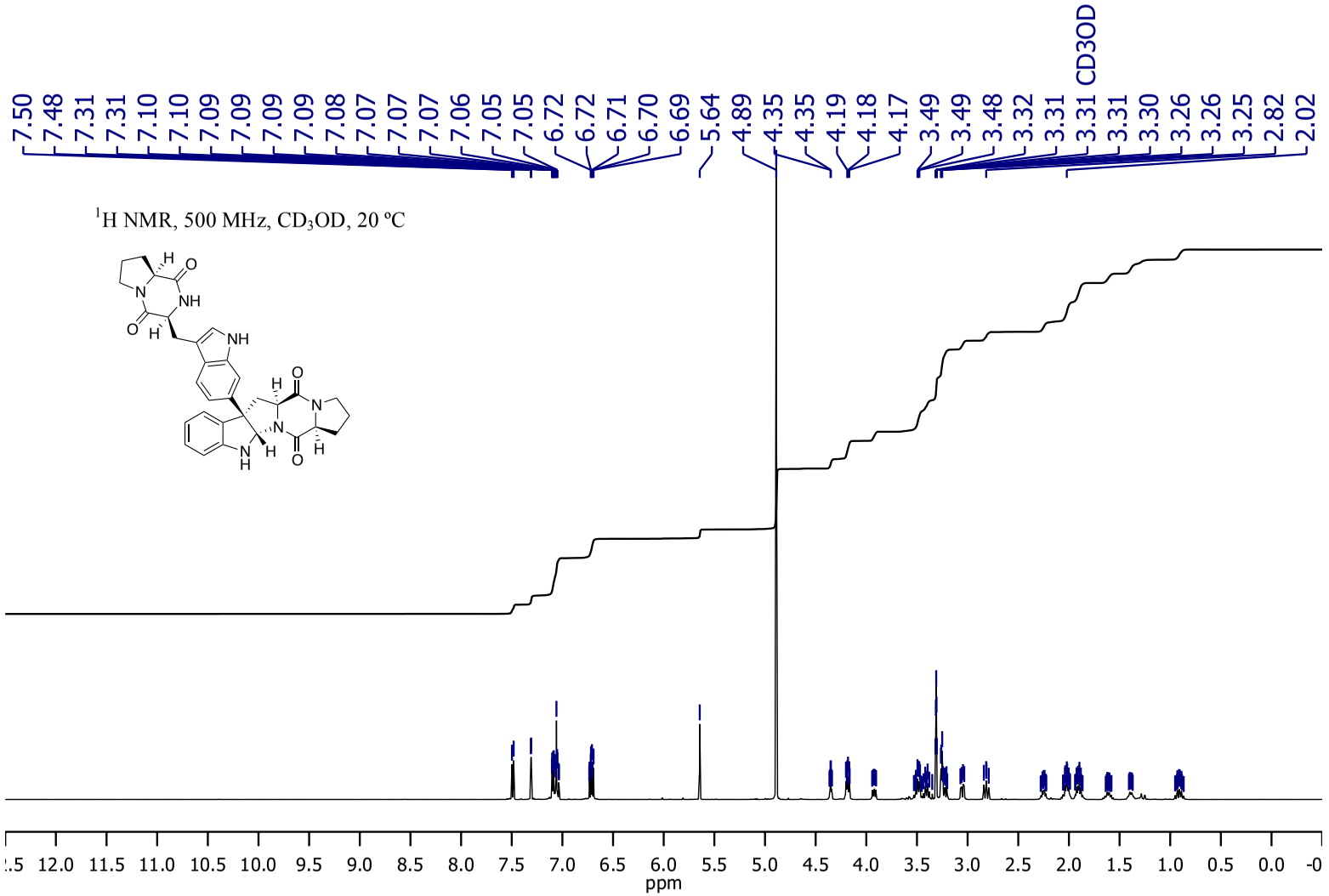

169.43
165.98
165.36
165.00
153.07
150.30
141.48
138.79
136.16
134.89
129.30
129.00
128.96
128.85
128.56
128.33
128.21
124.83
124.69
123.99
121.09
119.50
117.64
115.67
112.95
81.80
77.16 CDCl3
69.08
68.25
60.73
60.60
59.28
58.58
54.04
45.62
45.46
39.35
28.36
27.95
26.52
23.15
22.72
13C NMR, 125 MHz, CDCl3, 20 ºC
NCbz
Cbz
ppm

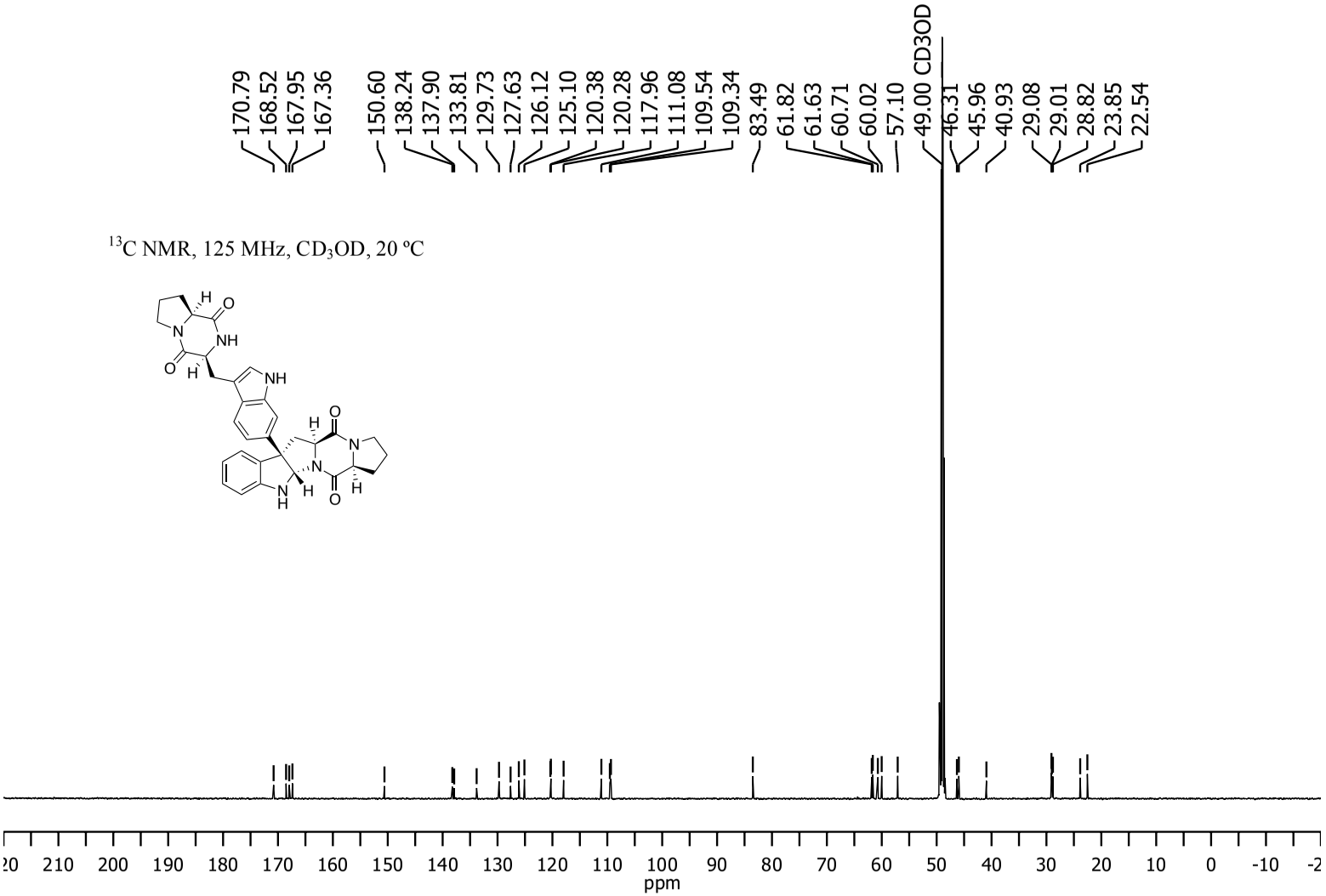

7.64
7.64
7.25
7.23
7.20
7.19
7.18
7.18
7.14
7.13
7.12
7.12
7.08
7.08
7.07
7.06
7.06
7.05
7.04
6.76
6.76
6.75
6.74
6.73
6.73
6.72
6.71
5.70
4.89
4.39
4.39
4.28
4.26
3.52
3.52
3.51
3.50
3.50
3.41
3.31 CD3OD
3.29
3.28
3.12
2.94
2.92
2.92
2.06
2.04
1H NMR, 500 MHz, CD3OD, 20 ºC
ppm

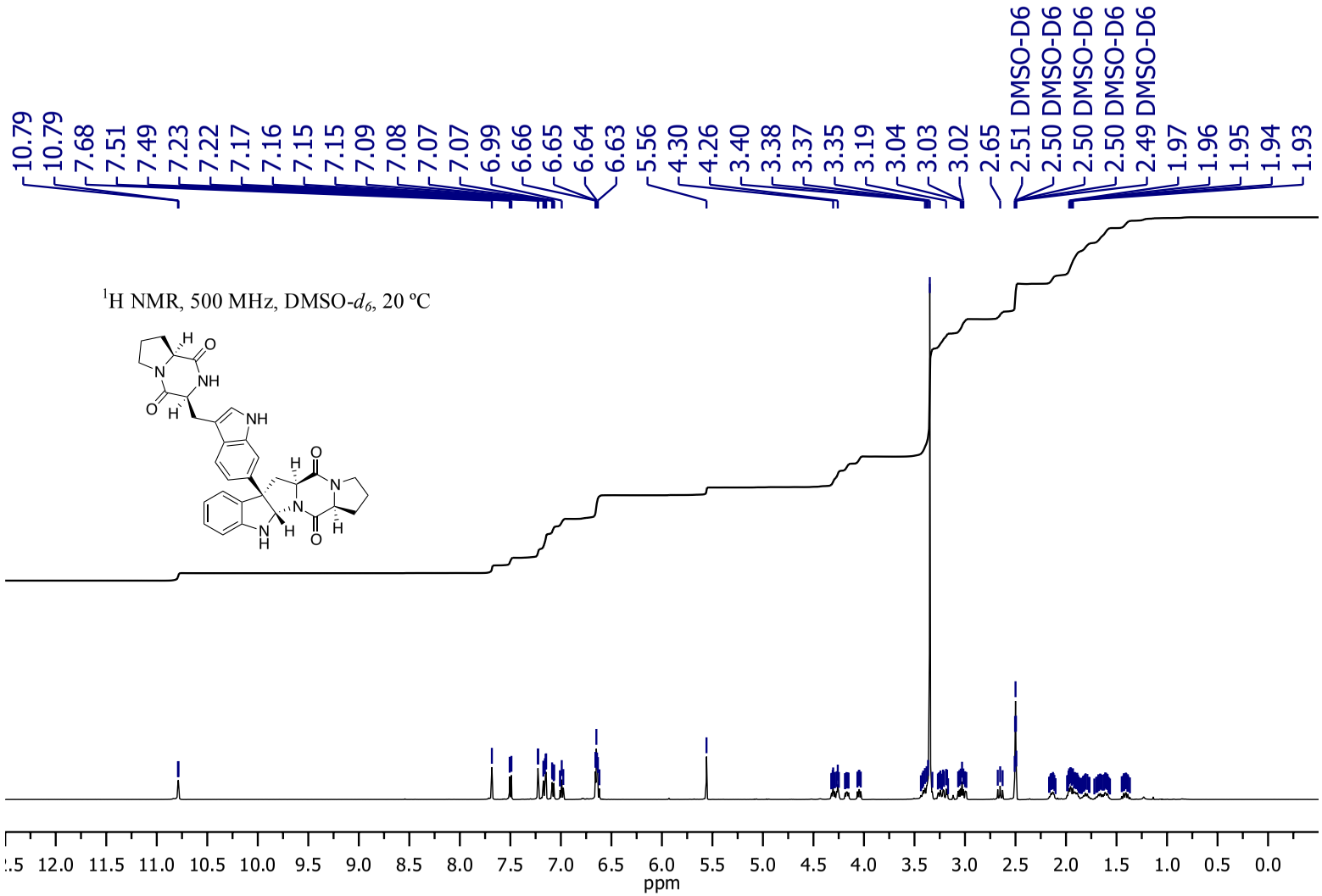

170.70
168.55
168.04
167.38
150.63
136.94
135.31
134.16
129.65
128.51
126.17
125.18
121.07
120.47
116.26
112.59
111.12
109.86
83.59
61.93
61.69
60.75
60.04
57.19
49.51
49.34
49.17
49.00
49.00 CD3OD
48.83
48.66
48.49
46.30
45.99
41.08
29.10
28.87
23.85
22.55
^{13}C NMR, 125 MHz, CD_3OD, 20 ºC
215 205 195 185 175 165 155 145 135 125 115 105 95 85 75 65 55 45 35 25 15 5 -5 -15
ppm

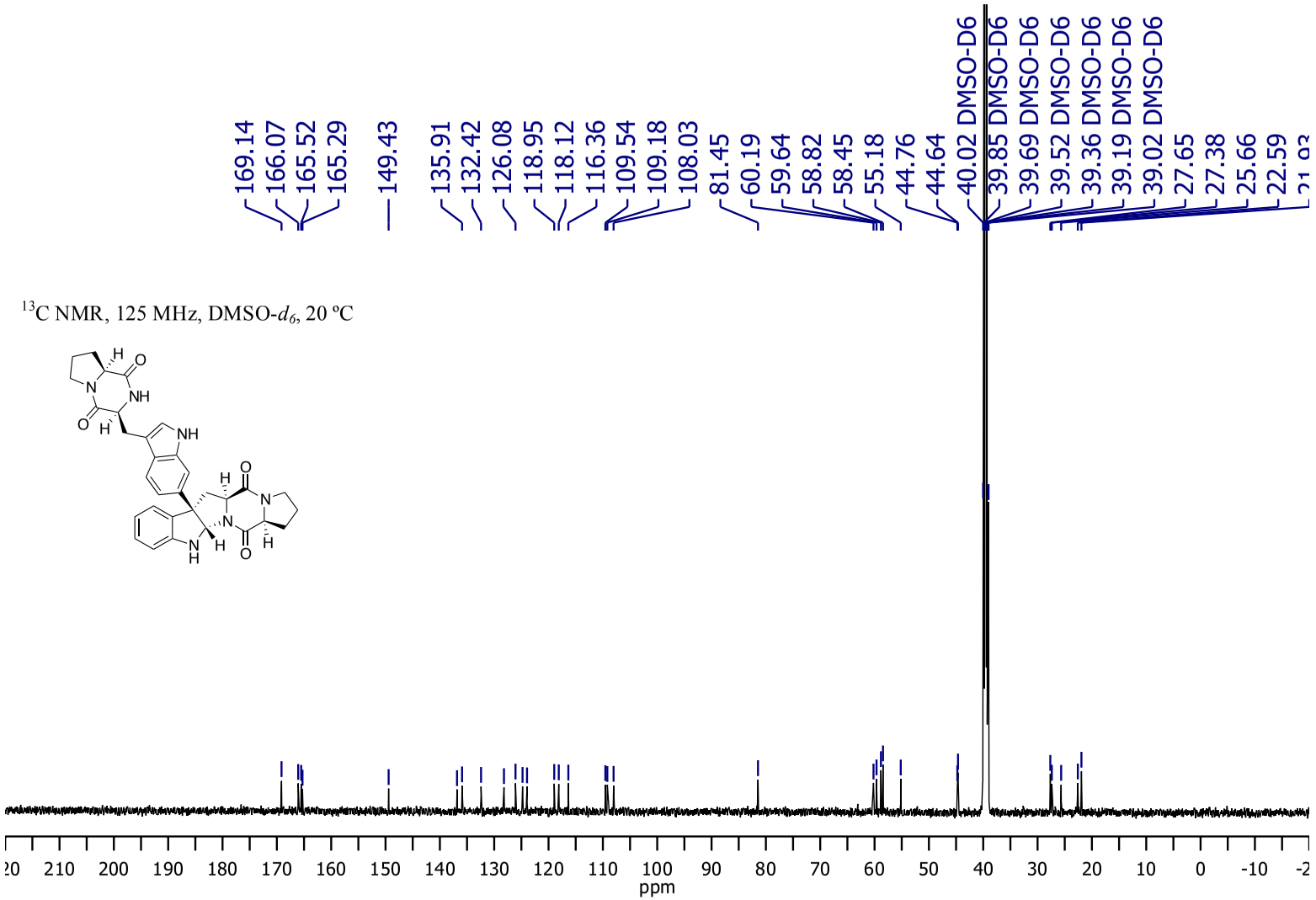

10.84
7.74
7.63
7.22
7.21
7.17
7.15
7.04
7.02
6.99
6.98
6.96
6.65
6.64
6.63
6.61
5.63
4.31
4.30
4.29
4.16
4.09
3.42
3.40
3.39
3.38
3.35
3.31
3.29
3.21
3.21
3.09
3.08
3.07
3.07
3.06
3.05
2.68
2.50 DMSO-D6
1.92
1.91
1.90
^{1}H NMR, 500 MHz, DMSO-d_6, 20 ºC
ppm


169.18
166.04
165.61
165.29
149.34
133.00
124.80
123.78
119.25
118.11
114.60
111.39
109.60
109.51
81.13
60.14
59.66
58.75
58.48
55.09
44.72
44.67
40.02 DMSO-D6
39.85 DMSO-D6
39.69 DMSO-D6
39.52 DMSO-D6
39.35 DMSO-D6
39.19 DMSO-D6
39.02 DMSO-D6
27.69
27.43
25.43
22.56
21.07
^{13}C NMR, 125 MHz, DMSO-d_6, 20 ºC
ppm

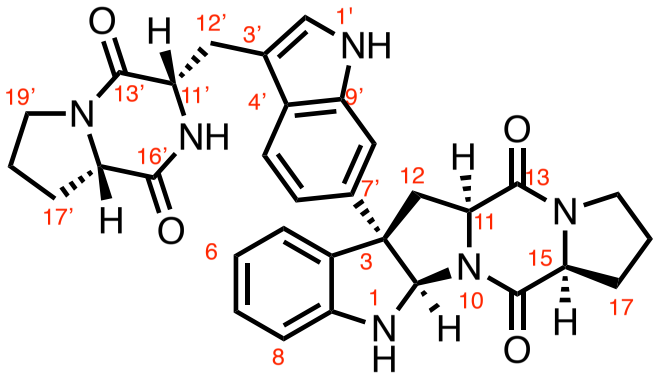

7.50 7.48 7.31 7.31 7.10 7.10 7.09 7.09 7.09 7.09 7.08 7.07 7.07 7.07 7.06 7.05 7.05 6.72 6.72 6.71 6.70 6.69 5.64 4.89 4.35 4.35 4.19 4.18 4.17 3.49 3.49 3.48 3.32 3.31 3.31 CD3OD 3.31 3.30 3.26 3.26 3.25 2.82 2.02
^{1}H NMR, 500 MHz, CD_3OD, 20 ºC
ppm


170.79
168.52
167.95
167.36
150.60
138.24
137.90
133.81
129.73
127.63
126.12
125.10
120.38
120.28
117.96
111.08
109.54
109.34
83.49
61.82
61.63
60.71
60.02
57.10
49.00 CD3OD
46.31
45.96
40.93
29.08
29.01
28.82
23.85
22.54
13C NMR, 125 MHz, CD3OD, 20 ºC
ppm


10.79
10.79
7.68
7.51
7.49
7.23
7.22
7.17
7.16
7.15
7.15
7.09
7.08
7.07
7.07
6.99
6.66
6.65
6.64
6.63
5.56
4.30
4.26
3.40
3.38
3.37
3.35
3.19
3.04
3.03
3.02
2.65
2.51 DMSO-D6
2.50 DMSO-D6
2.50 DMSO-D6
2.50 DMSO-D6
2.49 DMSO-D6
1.97
1.96
1.95
1.94
1.93
^{1}H NMR, 500 MHz, DMSO-d_6, 20 ºC
ppm


169.14
166.07
165.52
165.29
149.43
135.91
132.42
126.08
118.95
118.12
116.36
109.54
109.18
108.03
81.45
60.19
59.64
58.82
58.45
55.18
44.76
44.64
40.02 DMSO-D6
39.85 DMSO-D6
39.69 DMSO-D6
39.52 DMSO-D6
39.36 DMSO-D6
39.19 DMSO-D6
39.02 DMSO-D6
27.65
27.38
25.66
22.59
^{13}C NMR, 125 MHz, DMSO-d_6, 20 ºC
ppm

Positional Numbering System. For the purpose of comparison the ^{1}H and ^{13}C NMR data of the employed numbering system is consistent with that employed in the isolation papers of (+)-naseseazine B[20] and total synthesis of (+)-naseseazine B.



(+)-naseseazine B

[20] Buedenbender, L.; Grkovic, T.; Duffy, S.; Kurtbotke, D. I.; Avery, V. M.; Carroll, A. R. *Tet. Lett*, **2016**, *57*, 5893.

Supplementary Table 2-10. Comparison of our ^{1}H NMR data for (+)-naseseazine B (1) with synthetic literature:[1]

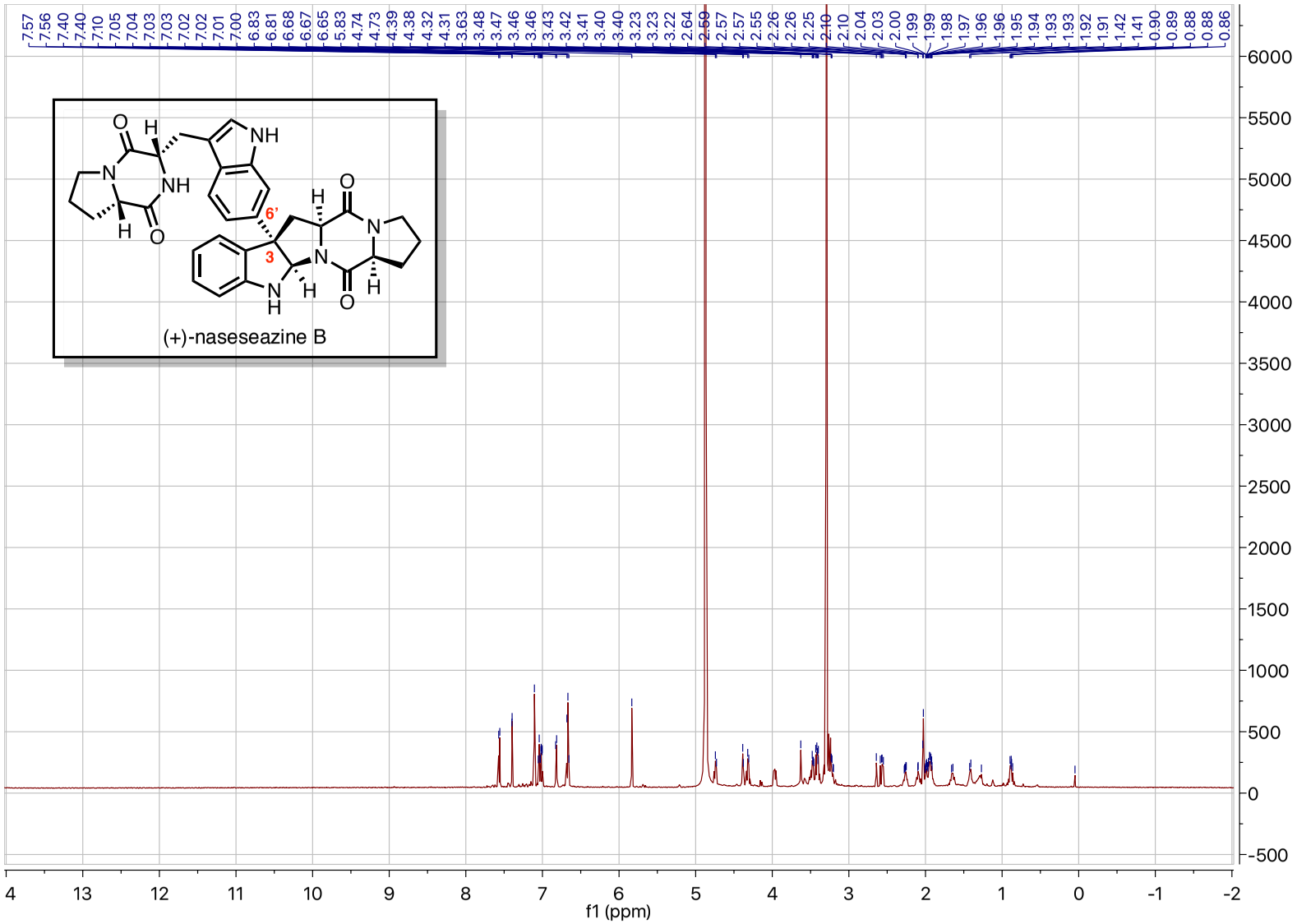
Assignment	Movassaghi Synthetic (+)-naseseazine B (**1**) ^{13}C NMR, 126 MHz DMSO-d_6, 20 ºC	This Work (+)-naseseazine B (**1**) ^{13}C NMR, 151 MHz DMSO-d_6, 20 ºC
N1	-	-
C2	5.84 (s)	5.83 (s)
C3	-	-
C4	-	-
C5	6.83 (d, J = 7.2)	6.82 (d, J = 7.2)
C6	6.68 (app-t, J = 7.3)	6.67 (app-t, J = 7.3)
C7	7.06 (app-dt, J = 1.2, 7.9)	7.04 (app-dt, J = 1.2, 7.9)
C8	6.69 (d, J = 7.7)	6.67 (d, J = 7.7)
C9	-	-
N10	-	-
C11	4.71 (app-t, J = 8.6)	4.74 (app-t, J = 8.6)
C12	3.35−3.21 (m) 2.57 (dd, J = 10.2, 13.7)	3.44 (m) 2.57 (dd, J = 10.2, 13.7)
C13	-	-
N14	-	-
C15	4.29 (app-t, J = 7.7)	4.32 (app-t, J = 7.7)
C16	-	-
C17	2.30−2.22 (m) 2.15−2.04 (m)	2.26 (dtd) 2.10 (ddd)

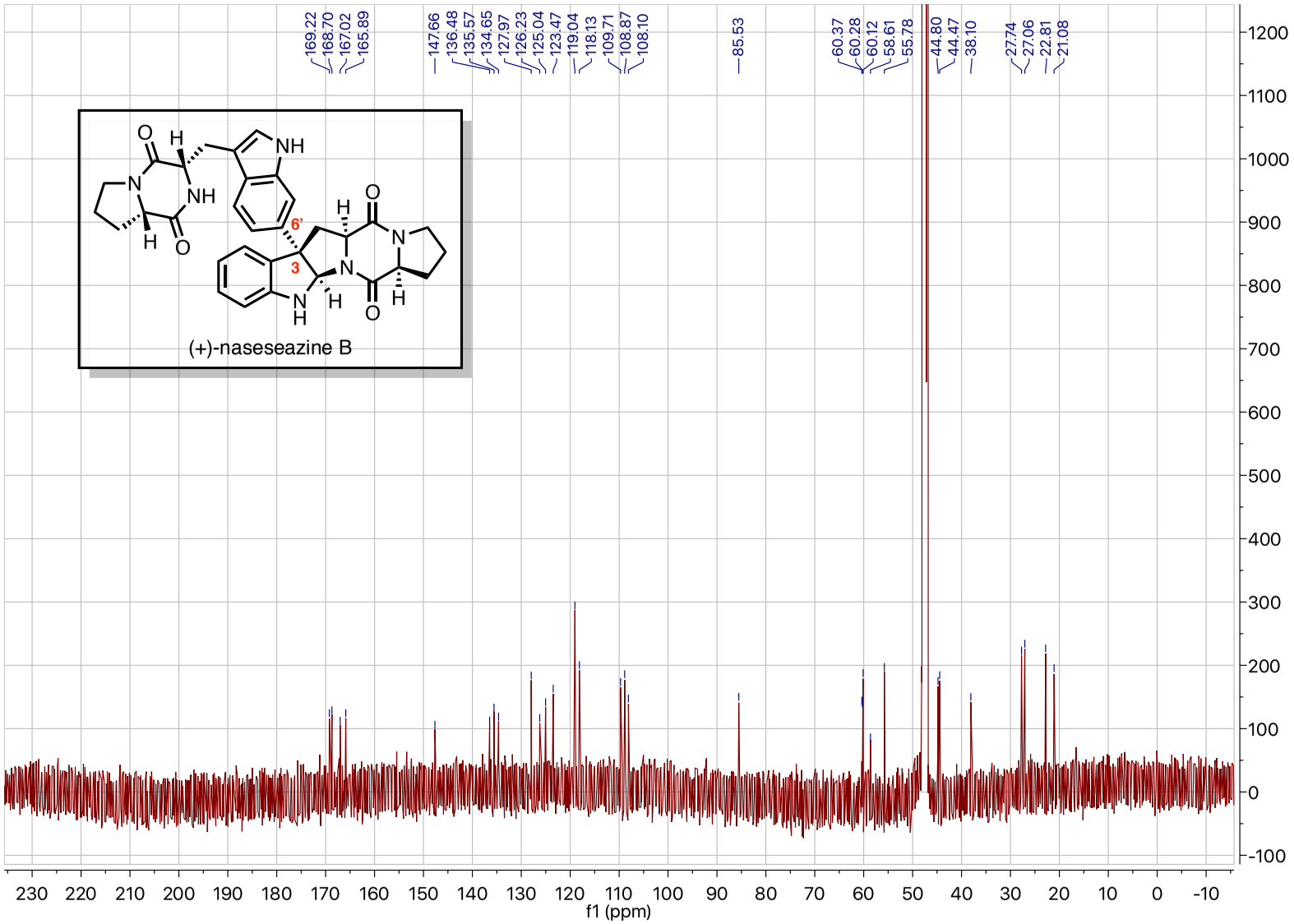
C18	2.04−1.87 (m) 2.04−1.87 (m)	1.96 (m)
C19	3.52−3.37 (m) 3.52−3.37 (m)	3.44 (m)
N1	-	-
C2’	7.12 (s)	7.10 (s)
C3	-	-
C4	-	-
C5	7.57 (d, J = 7.4)	7.56 (d, J = 7.4)
C6	7.01 (dd, J = 1.6, 8.5)	7.01 (dd, J = 1.6, 8.5)
C7	-	-
C8	7.41 (d, J =1.3)	7.40 (d, J =1.3)
C9	-	-
N10	-	-
C11	4.39 (app-t, J = 4.6)	4.32 (app-t, J = 4.6)
C12	3.35−3.21 (m) 3.35−3.21 (m)	3.44 (m)
C13	-	-
N14	-	-
C15	3.98 (ddd, J = 1.4, 6.3, 10.7)	3.97 (ddd, J = 1.4, 6.3, 10.7)
C16	-	-
C17	2.04−1.87 (m) 0.97−0.87 (m)	1.96 (m) .89 (m)
C18	1.73−1.60 (m) 1.49−1.40 (m)	1.65 (m) 1.42 (d)
C19	3.52−3.37 (m) 3.35−3.21 (m)	3.44 (m)

Supplementary Table 2-11. Comparison of our ^{13}C NMR data for (+)-naseseazine B (1) with synthetic literature:[1]

Assignment	**Movassaghi Synthetic** (+)-naseseazine B (**1**) ^{13}C NMR, 125 MHz DMSO-d_6, 20 ºC	**This Work** (+)-naseseazine B (**1**) ^{13}C NMR, 125 MHz DMSO-d_6, 20 ºC	**Chemical Shift Difference** Δδ = δ (+)-naseseazine B(this work) – δ (Movassaghi report)
C18'	22.7	22.5	-0.2
C18	24.4	24.3	-0.1
C17	28.6	28.5	-0.1
C12'	29.3	29.2	-0.1
C17'	29.3	29.2	-0.1
C12	39.7	39.6	-0.2
C19'	46.1	45.9	-0.2
C19	46.4	46.3	-0.1
C11'	57.3	57.2	-0.1
C15'	60.2	60.1	-0.1
C11	61.7	61.6	-0.1
C3	61.9	61.7	-0.2
C15	61.9	61.8	-0.1
C2	87.1	87.0	-0.1
C3'	109.8	109.6	-0.3
C8'	110.5	110.3	-0.2
C8	111.3	111.2	-0.1
C6'	119.8	119.6	-0.2
C6	120.6	120.5	-0.1
C5'	120.6	120.5	-0.1

C5	125.1	124.9	-0.2
C2'	126.6	126.5	-0.1
C4'	127.8	127.7	-0.1
C7	129.6	129.4	-0.2
C4	136.2	136.1	-0.1
C7'	137.2	137.0	-0.2
C9'	138.1	137.9	-0.2
C9	149.3	149.1	-0.2
C13'	167.5	167.3	-0.2
C13	168.6	168.5	-0.1
C16	170.3	170.2	-0.2
C16'	170.9	170.7	-0.2


(+)-naseseazine B
f1 (ppm)


169.22
168.70
167.02
165.89
147.66
136.48
135.57
134.65
127.97
126.23
125.04
123.47
119.04
118.13
109.71
108.87
108.10
85.53
60.37
60.28
60.12
58.61
55.78
44.80
44.47
38.10
27.74
27.06
22.81
21.08
(+)-naseseazine B
f1 (ppm)

References:

1. Zerbino, D. R.; Birney, E., *Gen. Res.* **2008,** *18* (5), 821-829.
2. Notredame, C.; Higgins, D. G.; Heringa, J., *J. Mol. Biol.* **2000,** *302* (1), 205-17.
3. Robert, X.; Gouet, P., *Nucl. Acids Res.* **2014,** *42* (Web Server issue), W320-4.
4. Omura, T.; Sato, R., *J. Biol. Chem.* **1964,** *239*, 2370-8.
5. Khatri, Y.; Hannemann, F.; Ewen, K. M.; Pistorius, D.; Perlova, O.; Kagawa, N.; Brachmann, A. O.; Muller, R.; Bernhardt, R., *Chem. Biol.* **2010,** *17* (12), 1295-1305.
6. Gertz, E. M.; Yu, Y. K.; Agarwala, R.; Schaffer, A. A.; Altschul, S. F., *BMC Biol.* **2006,** *4*, 4
7. Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923.
8. Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J. *Organomet.* **1996**, *15*, 1518.
9. G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg. *Organomet.* **2010**, *29*, 2176.
10. Buedenbender, L.; Grkovic, T.; Duffy, S.; Kurtbotke, D. I.; Avery, V. M.; Carroll, A. R. *Tet. Lett.*, **2016**, 57, 5893.
11. Xiong, Z.-Q.; Liu, Q.-X.; Pan, Z.-L.; Zhao, N.; Feng, Z.-X.; Wang, Y. *Arch. Microbiol.*, **2015**, 197, 299.

Chapter 3
Evaluation of Substrate Flexibility of Bacterial DKP Dimerases NascB and NznB, and Discovery of NascB Homologues with Novel Chemoselectivity

3.1 Abstract

In this chapter we explore substrate flexibility of diketopiperazine dimerizing cytochromes P450, NascB and NznB described in Chapter 2. A small library of diketopiperazines was synthesized to profile the initial reactivity of NascB and NznB against non-native substrates. NascB and NznB were unable to homodimerize diketopiperazines with sterically bulky amino acid side chains (Leu, Phe, Tyr, Trp), and a bioinformatics campaign in an effort to discover homologues with new or broader selectivity was undertaken. This campaign resulted in the discovery of two cytochromes with new selectivity, NascB-F5053 (from *Streptomyces* sp. NRRL-F5055) whose native reaction generated a mixture of (–)-naseseazine C (**1**) and (–)-aspergilazine A (**2**), and NascB-S1868 (from *Streptomyces* sp. NRRL-S1868) which formed (–)-aspergilazine A (**2**) exclusively. All four cytochromes were screened for their dimerization potential against a library of > 30 diketopiperazines, and a biotransformation approach was developed to scale reactions which form unique dimeric products. This work was enabled through a joint effort from myself, Dr. Yogan Khatri who assisted in design of the gene constructs, as well as Robert Hohlman who assisted in the synthesis of substrates.

3.2 Introduction

The biological activity of diketopiperazine dimers has inspired the development of efficient methods towards the total synthesis of this class of compounds. While these methods have enabled the synthesis of many of the compounds, there are no general

methods that allow for the synthesis of a variety of stereochemistries or connectivities from the same starting materials. Because of this, it is particularly challenging to generate libraries of these compounds, identify key pharmacophores and identify structure activity relationships within the molecular architecture. NascB and NznB are both site- and stereodivergent cytochromes, and from the same diketopiperazine monore generate two stereochemically, regiochemically, and topologically different dimeric structures.

Based on the flexibility displayed in the *in vitro* simulation of biosynthetic crosstalk seen in Chapter 1, we expected that both NascB and NznB may be able to catalyze the dimerization of a variety of DKP substrates and could be used to biocatalytically generate a library of different homo- and heterodimeric structures. The orthogonality of certain functional groups to biological reactions such as alkynes also act as functional handles for late stage diversification of dimers, and we sought to evaluate the ability for NascB and NznB to incorporate these functional groups into dimeric scaffolds for secondary diversification of these scaffolds.

3.3 Research and Discussion

To screen the reactivity profile of NascB and NznB, a small library of diketopiperazines were synthesized in three steps with no intermittent purification (Experimental). This set of substrates was screened against both the NascB and NznB cytochromes P450 for their ability to homodimerize under conditions optimized to fully convert native substrate **4** with the lowest loading of NascB biocatalyst. However in nearly all cases these reactions could be pushed to complete conversion with increased catalyst loadings.

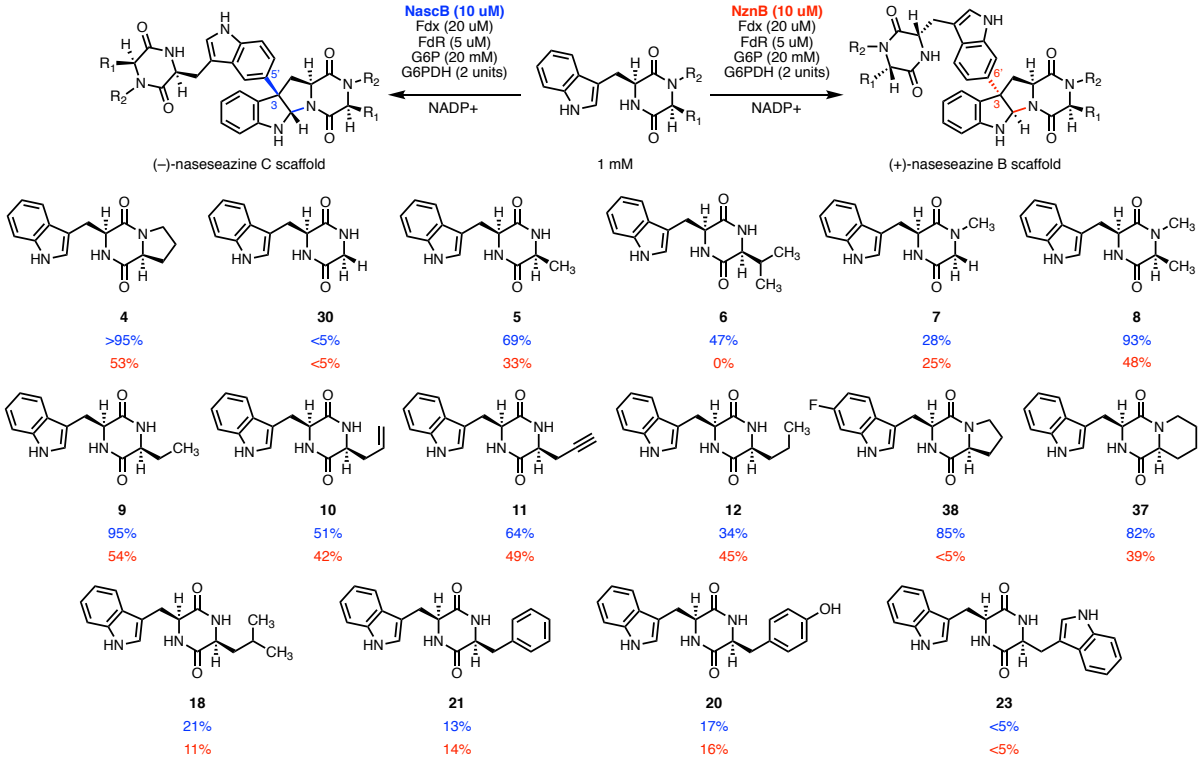


Figure 3-1. Percent conversions of homodimerizations catalyzed by NascB and NznB.

NascB consistently showed higher conversions at equivalent concentrations to NznB, and we hypothesize that this is due to the stability of this catalyst, as NznB is more sensitive to denaturation during purification and can be seen precipitating under reaction conditions. Notably, the alkynyl diketopiperazine **11** was accepted as a substrate and was readily converted to a dimeric product and we are currently evaluating further functionalization of these dimeric products in copper-catalyzed [3+2] azide-alkyne cycloaddition "click" reactions. [1] C6-fluorinated substrate **38** was converted by NascB and completely stopped reactivity for NznB, as C6 is the site for dimerization. Substrates with bulkier amino acid side chains, **18**, **20**, **21** and **23**, consistently showed little to no conversion to dimeric products. Based on the cyclic nature of native substrate **4**, and minimal steric bulk of native substrate **5**, we hypothesized that the enzyme active site was not evolved to recognize amino acids with these bulky side chains.

3.3.1 Bioinformatic Discovery and *in vitro* analysis of NascB homologues

Dimerization of large substrates (**18, 20, 21,** and **23**) failed, and in hopes of discovering a cytochrome with a broader substrate scope or new selectivity we mined

the NCBI database for homologues of cytochromes NascB and NznB. To identify NascB/NznB homologues, we used their upstream CDPSs, NascA and NznB, well as all known CDPS-driven gene clusters and using BLASTp[2] we identified the top 100 homologues for each CDPS, and then eliminated duplicates by hand. This curated CDPS library of 489 members was then used to assemble a sequence similarity network (SSN) [3] and group CDPS-driven clusters with their respective reactivity/metabolites and predictively identify NascB and NznB homologues.

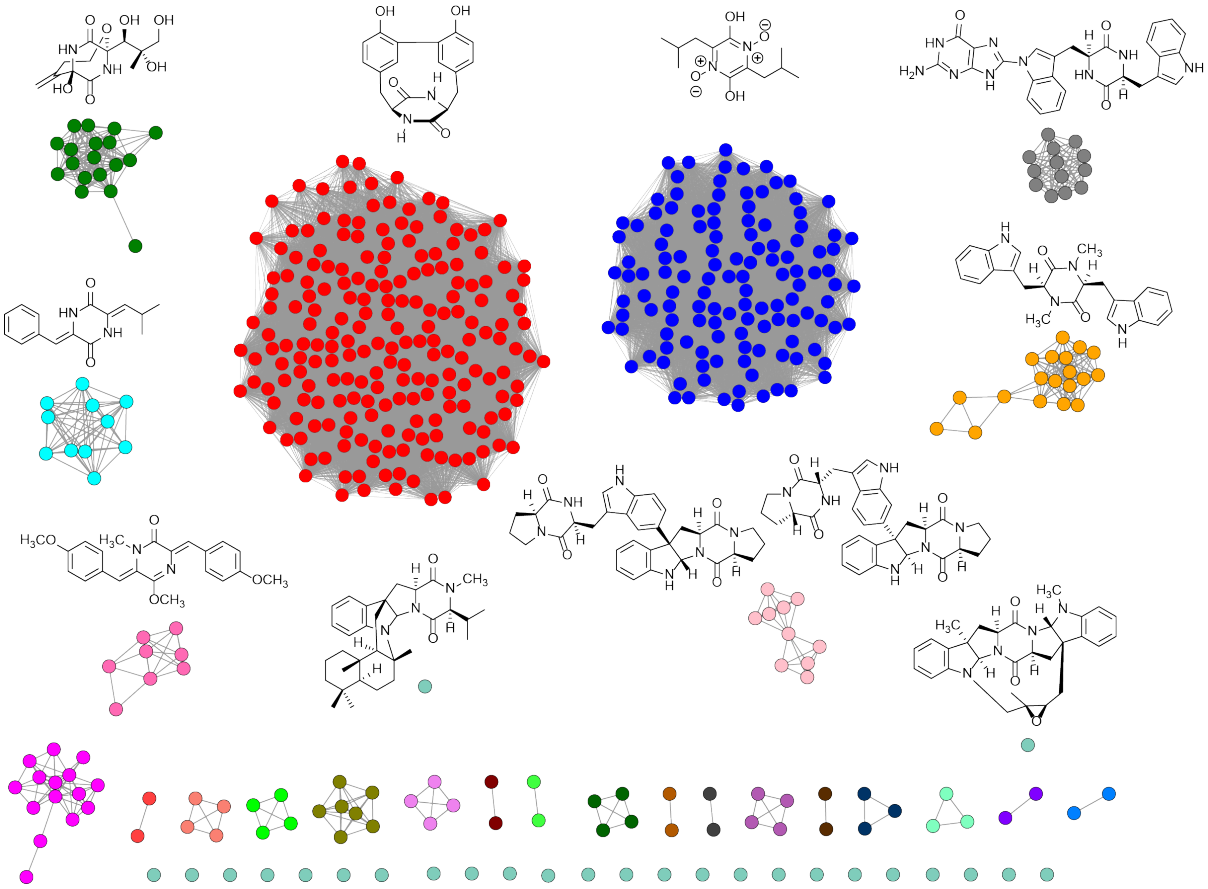


Figure 3-2. Curated Sequence Similarity Network (SSN) of 498 CDPS with products from characterized biosynthetic gene clusters above each cluster.

Analysis of the SSN identified two *Streptomyces* strains deposited in the NRRL (NRRL-F5053 and NRRL-S1868) both of which possessed NascB (hereon refered to as NascB-CMB-MQ030) homologues 87% and 84% identity respectively. These strains were obtained from the NRRL, genomic DNA harvested and the NascB homologues

NascB-F5053 and NascB-S1868 were cloned, expressed and their substrate scope characterized. NznB homologues from these strains were also identified, however these NznB homologues had extremely high sequence homology (99% identity) and as such we hypothesized they had similar reactivities and selectivities and these cytochromes were not explored further.

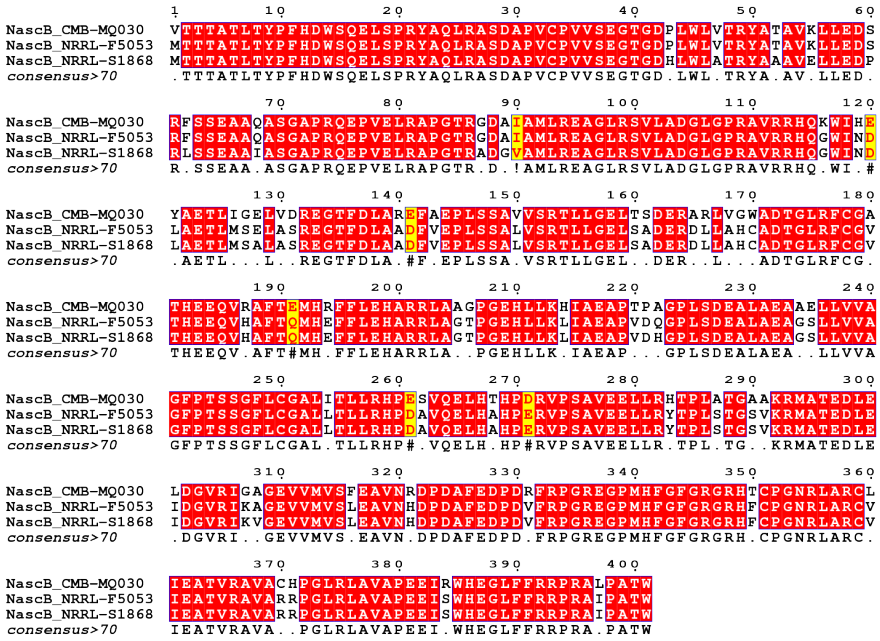


Figure 3-3. Sequence alignment of NascB homologues from NRRL-F5053 and NRRL-S1868.

3.3.3 Expression and Characterization of NascB Homolgoues

When subjected to reaction conditions optimized for NascB-CMB-MQ030 and **4**, NascB-F5053 generated (–)-naseseazine C (**1**) as the major product produced along with a less polar minor product whose mass also corresponded to a dimer. When this reaction was performed with NascB-S1868 this new product, **2**, was produced exclusively. To identify this product, this reaction was scaled up (10mg) and major product purified by HPLC. NMR characterization of this product revealed there was no distinctive aminal proton corresponding to a cyclized dimer. When NMR solvent was exchanged to DMSO-d_6 the spectra corresponded favorably to reported spectra from both isolation and total synthesis reports for uncyclized, unsymmetrical homodimer (–)-aspergilazine A (**2**).[4-7]

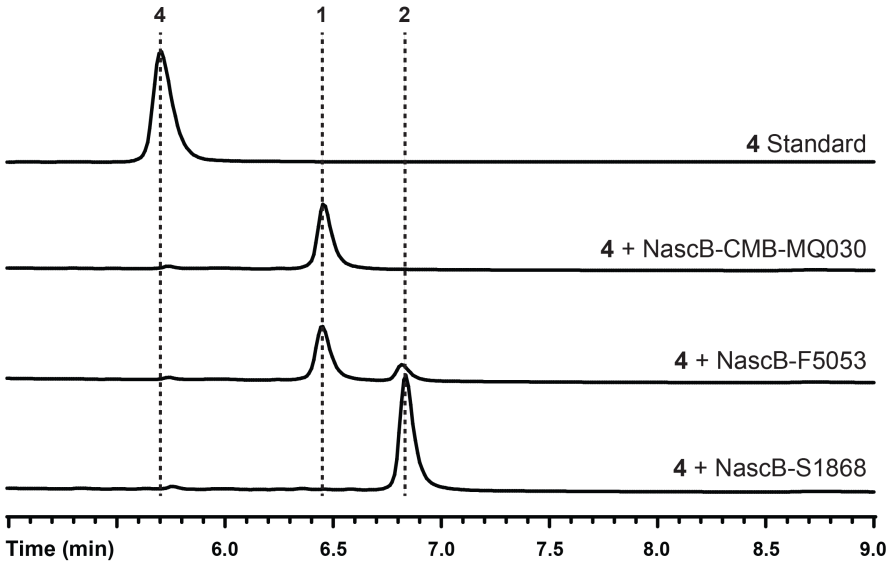
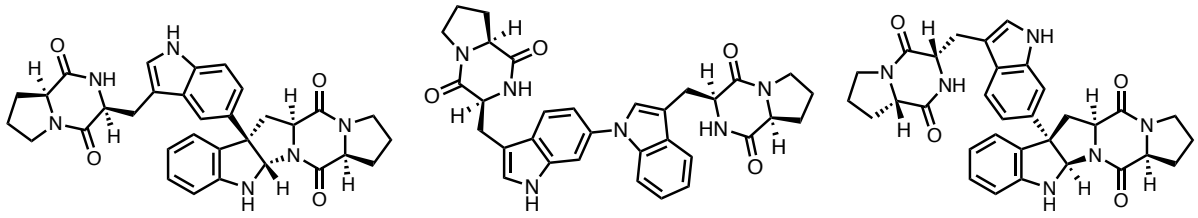


Figure 3-4. HPLC traces showing conversion of **4** by NascB_CMB-MQ030, NascB-F5053, and NascB-S1868.

While initially isolated from a fungal source,[7] (–)-aspergilazine A (**2**) is among a rare class of natural products to be isolated from both a fungal and bacterial[5] source. The bacterial isolation report of **2** from *Streptomyces* sp. SMA-1 also reports the co-isolation of (+)-naseseazine B (**3**) and (–)-naseseazine C (**1**) (incorrectly assigned as "isonaseseazine B"). Strain NRRL-F5053 possesses an NznB homologue, and as seen in Figure 3-4, NascB-F5053 generates both (–)-naseseazine C (**1**) as well as (–)-aspergilazine A (**3**), and possesses the biosynthetic potential to produce all three of these metabolites suggesting that NRRL-F5053 is likely a close relative to this *Streptomyces* sp. SMA-1.



(–)-naseseazine C (**1**) (–)-aspergilazine A (**2**) (+)-naseseazine B (**3**)

Figure 3-5. Diketopiperazine dimers isolated from Streptomyces sp. SMA-1.

Given the change in chemoselectivity between intermolecular Csp^3–Csp^2 bond formation catalyzed by NascB-CMB-MQ030, and intermolecular Csp^2–Nsp^2 bond formation seen in reactions with NascB-F5053 and NascB-S1868, we hypothesized that NascB-F5053 and NascB_S1868 may possess different selectivities in terms of their substrate scope. To examine this, the initial set of substrates used to profile NascB-CMB-MQ030 and NznB were used to screen the reactivity of NascB-F5053 and NascB_S1868.

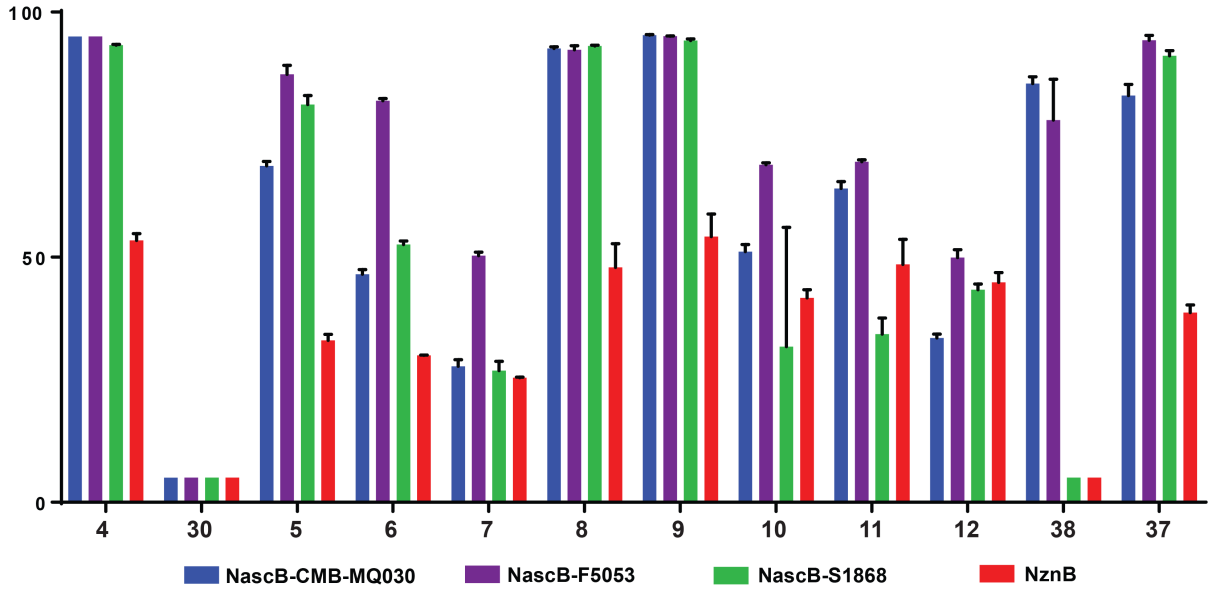


Figure 3-6. Grouped bar graph representation of percent conversions from reaction of homodimerization substrates with NascB-CMB-MQ030, NznB, NascB-F5053 and NascB-S1868. Each entry is the result of 3 reactions with standard deviation shown.

In this homodimerization assay, NascB-F5053 and NascB-S1868 consistently catalyze higher conversions of all substrates compared to NascB-CMB-MQ030 and NznB. Examination of the HPLC traces of these reactions reveal that not only do NascB-F5053 and NascB-S1868 have increased activity, but these cytochromes also generate an increased number of products in many cases. Homodimerization of **9** by NascB-F5053 and NascB-S1868 leads to formation of an additional product, and each cytochrome generates differential product distributions relative to each other (Figure 7).

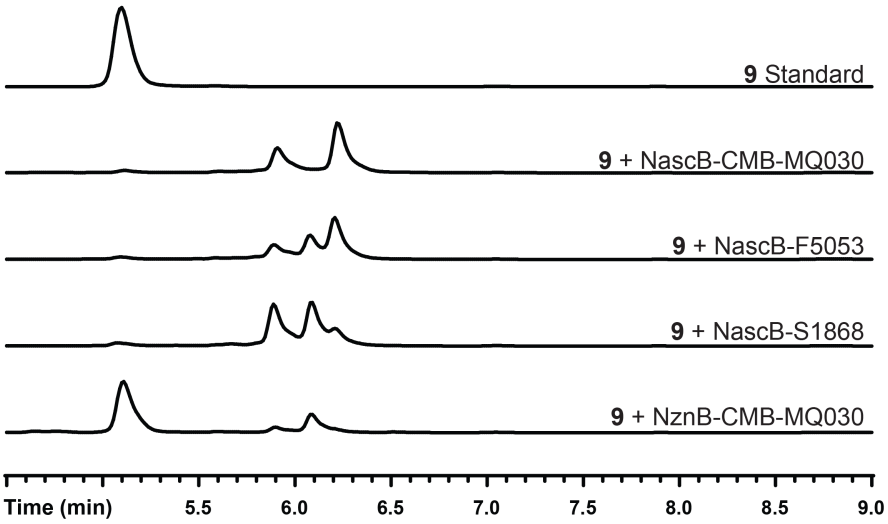


Figure 3-7. HPLC traces showing differential product profiles for reaction of **9** with NascB-CMB-MQ030, NascB-F5053, NascB-S1868, and NznB.

To further profile this differential reactivity, a larger subset of DKP substrates were synthesized and subjected to a heterodimerization reaction conditions with **4** as the common monomer to all reactions (Figure 8).

3.3.4 Heterodimerization Substrate Scope of NascB-CMB-MQ030, NascB-F5053, NascB-S1868, and NznB.

The majority of known diketopiperazine dimers are composed of a dimerization of two units of the same diketopiperazine.[8] There are several examples of isolated DKP dimers, which are the result of a heterodimerization of two different diketopiperazine monomers,[9-11] including the initial isolation report of the (+)-naseseazines, as (+)-naseseazine A is a non-symmetrical heterodimer of **4** and **5**. To explore the potential of NascB-CMB-MQ030, NascB-F5053, NascB-S1868, and NznB to heterodimerize two diketopiperazines, a library of over 30 members were constructed and were subjected to heterodimerization reactions with native substrate **4** (Figure 3-8).

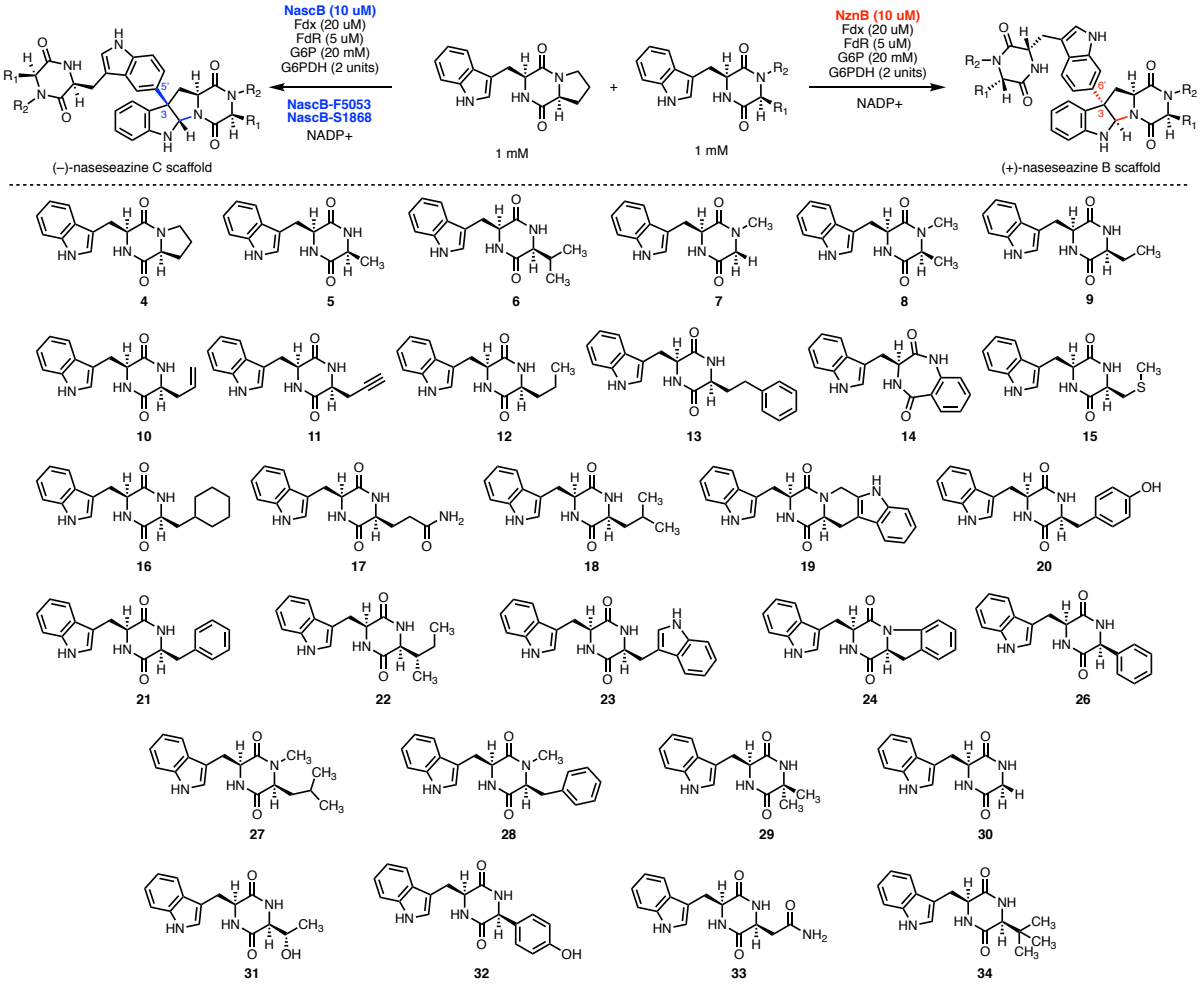


Figure 3-8. Full panel of diketopiperazines synthesized and assayed by cytochromes.

Many of these reactions resulted in the formation of complex mixtures, as many of the DKP substrates have very similar polarity, in many cases starting materials were not fully consumed. Each substrate is capable of generating at least one homodimer, and multiple isomeric heterodimers can be formed, all of which have relatively similar polarity. Regardless of HPLC column selection, solvent, pH, and method development, in the majority of reactions substrates and/or products co-eluted and did not allow for direct quantification of products formed.

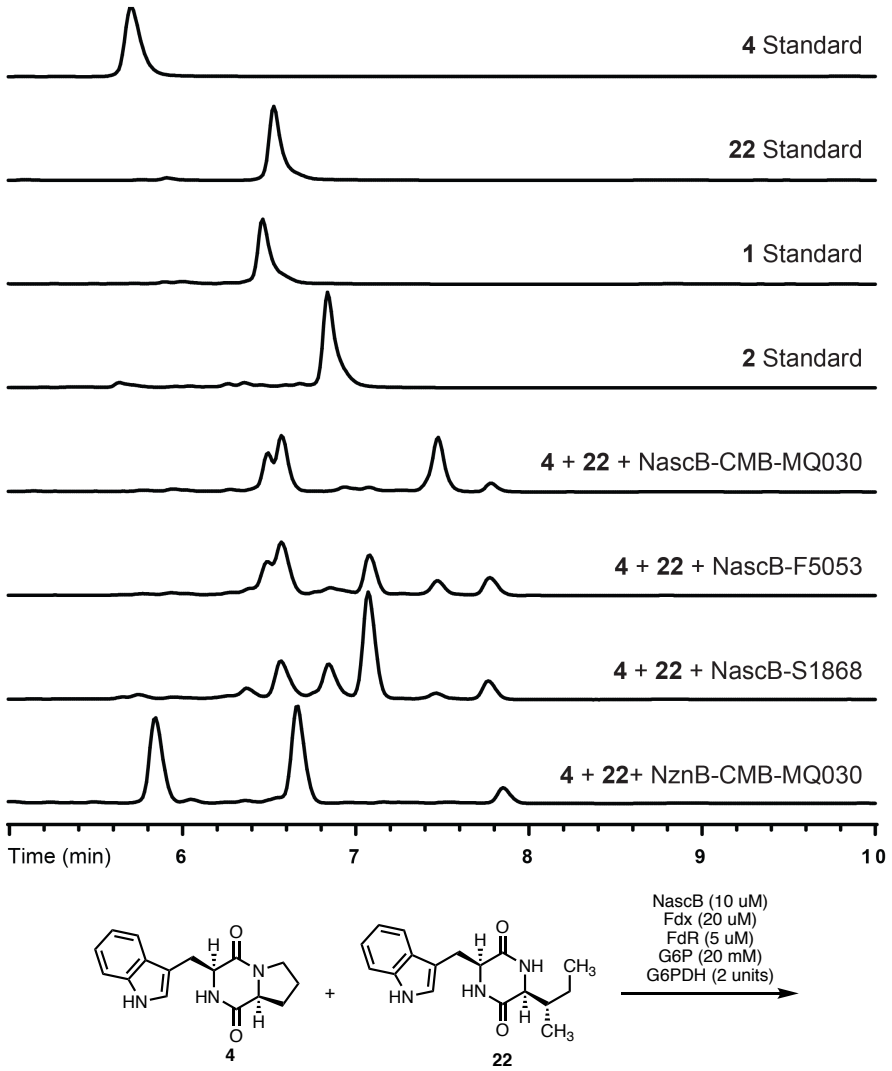


Figure 3-9. Representative HPLC traces from heterodimerization of **4** and **22**, percent conversion cannot be calculated for reactions such as these due to co-eluting of substrates and products.

To obtain a general understanding of the substrate selectivity for each enzyme, we opted to pursue a binary approach to detecting the formation of heterodimers using mass spectrometry summarized in Figure 10.

Cytochrome	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
NascB-CMB-MQ030																	
NascB-F5053																	
NascB-S1868																	
NznB																	

Cytochrome	20	21	22	23	24	25	26	27	28	30	31	32	33
NascB-CMB-MQ030													
NascB-F5053													
NascB-S1868													
NznB													

Figure 3-10. Binary tables for heterodimerization reactions. Green cells indicate at least one dimer detectable by MS, as well as correlating to a new peak in the UV chromatogram, red cells indicate no dimer detected. Orange squares are represent sulfur containing dimers which underwent further oxidation from thioether to sulfoxide.

As seen in the homodimerization assay, NascB-F5053 and NascB-S1868 exhibited increased flexibility substrate flexibility. These cytochromes were more capable of generating heterodimers from diketopiperazines with sterically bulky side chains (**16**-**20**, **23**, **28-30**, and **33**) while NascB-CMB-MQ030 was unable to convert these substrates. While conversions were very low for NznB, there were clear dimers formed from larger substrates as well. Also of note, **24** inhibited any dimerization from occurring (including homodimerization of **4**), while **19** did not inhibit dimerization, but also did not participate in homo- or heterodimerizations.

The NascB-F5053 catalyzed heterodimerization reaction of **4** and **23**, results in a single heterodimeric product, **37** **(**Figure 3-11**)**, generated exclusively and homodimerization of **4** is inhibited and little to no (–)-naseseazine C (**1**) is observed.

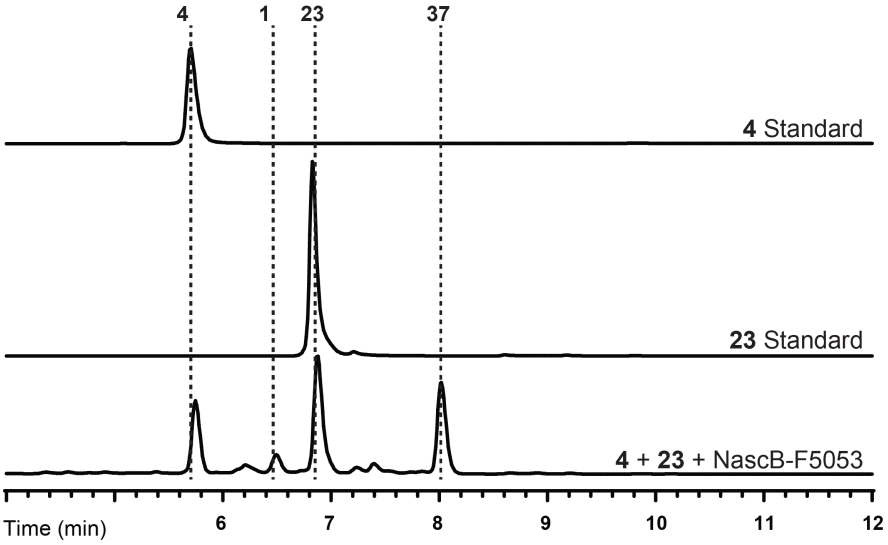


Figure 3-11. HPLC trace of NascB-F5053 catalyzed heterodimerization of **4** and **23**.

Despite their activity in heterodimerization reactions, substrates such as **23** are much bulkier, and possess greater degrees of conformational flexibility due to their acyclic nature. We hypothesize homodimerization of these substrates is not observed due to the inability for the cytochrome P450 active site to accommodate two units of these larger substrates in conformations conducive to catalysis. However, as these larger monomers are able to selectively form heterodimers with native monomer, **4,** there must be at least one portion of the binding pocket with the ability to accommodate these larger DKP monomers. To gain further insight into this pocket selectivity, we strove to scale the heterodimerization of **4** and **23** and characterize the structure of the product. However, attempts to scale this reaction *in vitro* suffered from low, and inconsistent conversions. To be able to perform this reaction on a scale that would enable resolving the structure of **37** we developed a biotransformation approach that would be more amenable to preparatory scale.

3.3.5 Development of a biotransformation platform for scalable synthesis of diketopiperazine dimers

Homo- and heterodimerization substrate scope experiments demonstrated biocatalytic potential for NascB-CMB-MQ030 and its homologues to generate a diversity of topologically diverse products from a variety of diketopiperazines substrates (Figures 3-1 and 3-10). Unfortunately many of these reactions generate multiple products, and/or suffer from modest to low conversions under optimized *in vitro* reaction conditions. We strove to develop an operationally simple method to scale up these reactions in order to access these products for spectroscopic characterization and as a platform to synthesize libraries for biological evaluation. Scalability of *in vitro* reactions is troublesome for cytochrome P450 enzymes in particular, as associated electron transfer proteins must also be expressed and purified, and in many cases are used, most optimally, in stoichiometric excess relative to the cytochrome. Biotransformation has been shown to be an effective strategy for biocatalytic oxidation reactions,[12-13] and circumvents protein purification through co-expression of redox partners (or utilization of redox parters native to heterologous host), and feeding substrates directly to a living culture. We initially attempted to generate a P450-RhFRED fusion[14] for

biotransformation, however expression of this chimera resulted exclusively in insoluble protein. A two-plasmid system was developed, wherein the plasmid that encodes for expression of our P450 of interest is co-transformed with a second a dicistronic plasmid which encodes genes for both the *S. oleracea* ferredoxin (Fdx) and ferredoxin reductase (FdR). Both plasmids are co-transformed into *E. coli* heterologous host C41DE3, and protein expression is induced with IPTG. Post induction, cells are washed and resuspended in minimal buffer along substrate to be dimerized (Figure 12).

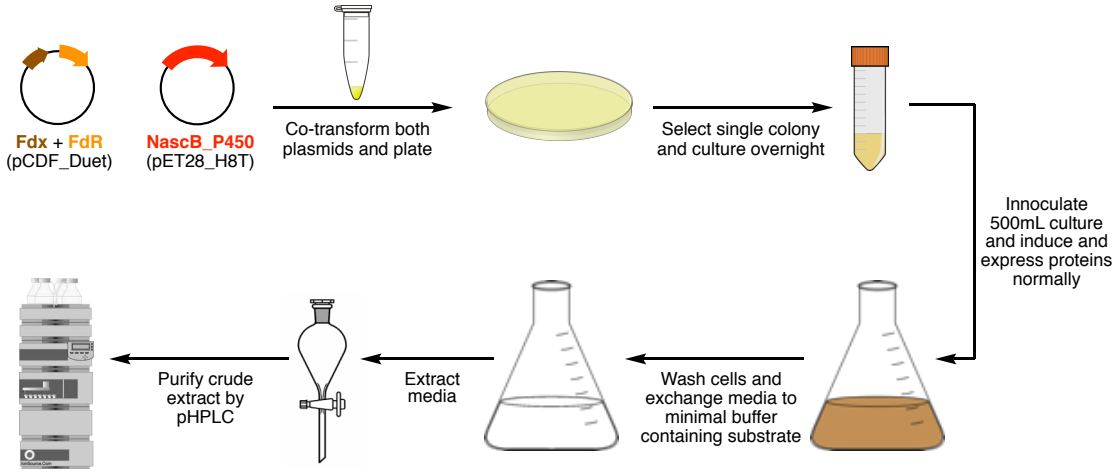


Figure 3-12. General biotransformation protocol for DKP dimerization.

Biotransformations using native substrate **4** went to full conversion of substrate to product within 5 hours reaction time, and per liter of culture we were able to generate 10-15 mg of dimeric DKP product using NascB-F5053. The polar diketopiperazine monomers and dimers are intransigent to liquid-liquid extraction and result in very low recovery from aqueous media directly. Instead, solid phase extraction using XAD-16 resin was employed, resulting in high extraction efficiency and full recovery of even the most polar of DKP products from unwanted media components such as glycerol.

As noted from heterodimerization reactions, **4** and **23** generated a new single heterodimeric product, albeit in low yield and low conversion making it particularly unsuitable for *in vitro* scale up. This reaction was chosen as a test case for biotransformation. Under unoptimized conditions, we were able to generate 11 mg of purified product from a 1L biotransformation reaction and were able to elucidate the

structure of this product as **37**. Given the structure of this new product, we propose that the pocket that accommodates the southern diketopiperazine has a broader flexibility and is able to accommodate larger diketopiperazines in heterodimerization reactions, while the pocket that binds the northern monomer may be more selective for its substrate.

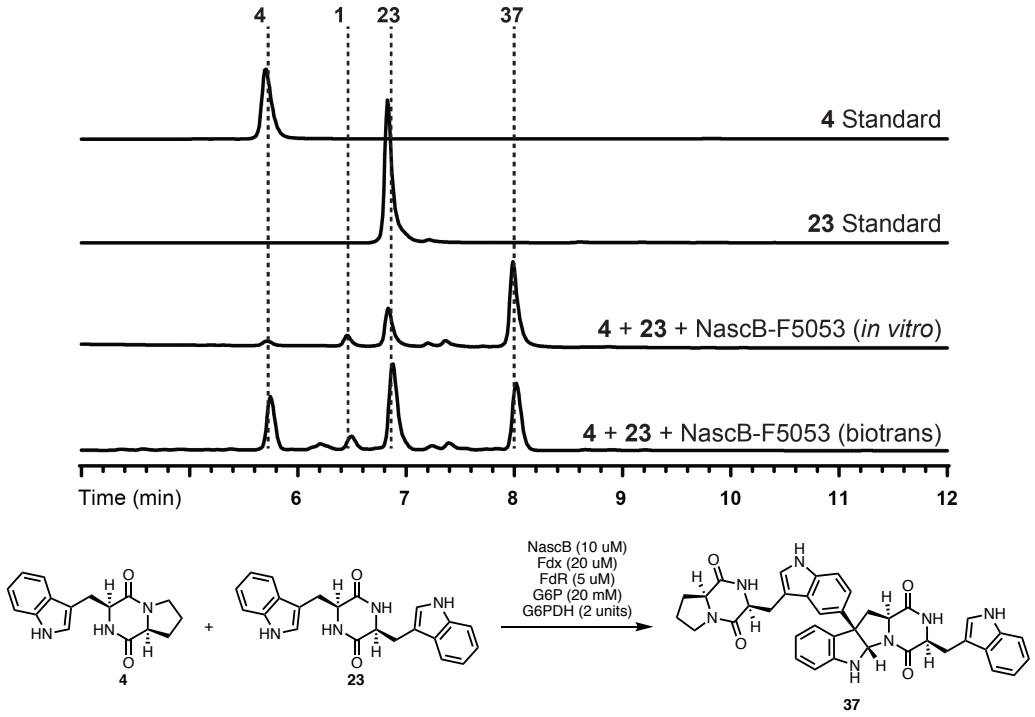


Figure 3-13. HPLC trace of biotransformation of **4** and **23** and structure of product **37**.

3.4 Conclusion

We synthesized a panel of diketopiperazines and profiled the substrate scope of previously described bacterial cytochromes, NascB and NznB, for their enzymatic homodimerization activity. In an effort to discover cytochromes with improved or divergent reactivity, we probed the NCBI database for putative homologues and discovered NascB homologues NascB-F5053 and NascB-S1868. When **4** was subjected to reaction conditions with these cytochromes a new metabolite, was produced, and when scaled up and characterized the structure of this product was determined to be (–)-aspergilazine A (**2**). A larger series of DKPs were synthesized and all four cytochromes were screened for their activity in both homodimerization and heterodimerization reactions.

From these heterodimerization reactions we were able to make inferences regarding the binding pockets of NascB-F5053, as only a single product was produced in the heterodimerization reaction of **4** and **23**. As the southern monomer (acceptor) is that composed of two bulky tryptohan units (derived from DKP **23**), we hypothesize that the binding site for this DKP is more flexibile, while the binding pocket for the northern monomer (donor) must be more selective for its substrate.

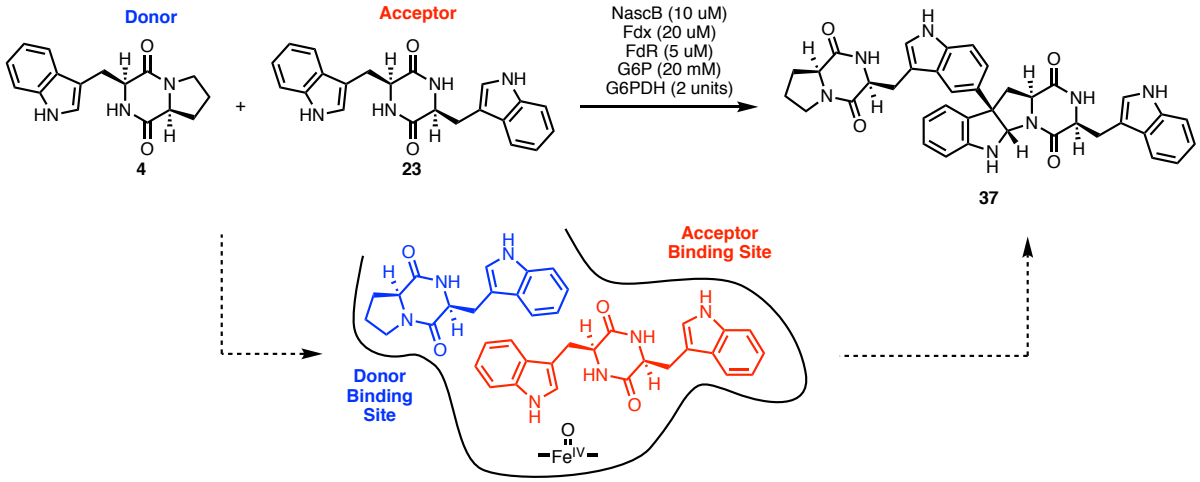


Figure 3-14. Hypothetical representation of binding site where acceptor binding site can accommodate larger substrates while donor site is more selective for smaller substrates.

To efficiently access quantities of material suitable for straightforward spectroscopic analysis, a biotransformation approach was developed and implemented on the low yielding heterodimerization of **4** and **23**, and allowed for 11mg of heterodimer **37** to be isolated from a single biotransformation.

3.5 References

1. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., *Angew Chem Int Ed Engl* **2001,** *40* (11), 2004-2021.
2. Jacob, A.; Lancaster, J.; Buhler, J.; Harris, B.; Chamberlain, R. D., *ACM Trans Reconfigurable Technol Syst* **2008,** *1* (2), 9.
3. Song, N.; Joseph, J. M.; Davis, G. B.; Durand, D., *PLoS Comput Biol* **2008,** *4* (4), e1000063.
4. Chuang, K. V.; Kieffer, M. E.; Reisman, S. E., *Org. Lett.* **2016,** *18* (18), 4750-3.
5. Xiong, Z. Q.; Liu, Q. X.; Pan, Z. L.; Zhao, N.; Feng, Z. X.; Wang, Y., *Arch. Microbiol.* **2015,** *197* (2), 299-309.
6. Boyd, E. M.; Sperry, J., *Org. Lett.* **2014,** *16* (19), 5056-9.
7. Cai, S.; Kong, X.; Wang, W.; Zhou, H.; Zhu, T.; Li, D.; Gu, Q., *Tet. Lett.* **2012,** *53* (21), 2615-2617.
8. Borthwick, A. D., *Chem. Rev.* **2012,** *112* (7), 3641-716.
9. Ding, G.; Jiang, L.; Guo, L.; Chen, X.; Zhang, H.; Che, Y., *J. Nat. Prod.* **2008,** *71* (11), 1861-5.
10. Li, X. B.; Li, Y. L.; Zhou, J. C.; Yuan, H. Q.; Wang, X. N.; Lou, H. X., *J Asian Nat. Prod. Res.* **2015,** *17* (2), 182-7.
11. Varoglu, M.; Corbett, T. H.; Valeriote, F. A.; Crews, P., *J Org Chem* **1997,** *62* (21), 7078-7079.
12. Baker Dockrey, S. A.; Lukowski, A. L.; Becker, M. R.; Narayan, A. R. H., *Nat Chem* **2018,** *10* (2), 119-125.
13. Payne, J. T.; Poor, C. B.; Lewis, J. C., *Angew Chem Int Ed Engl* **2015,** *54* (14), 4226-30.
14. Narayan, A. R.; Jimenez-Oses, G.; Liu, P.; Negretti, S.; Zhao, W.; Gilbert, M. M.; Ramabhadran, R. O.; Yang, Y. F.; Furan, L. R.; Li, Z.; Podust, L. M.; Montgomery, J.; Houk, K. N.; Sherman, D. H., *Nat Chem* **2015,** *7* (8), 653-60.

3.5 Experimental

Supplementary Table 3-1. Primers used in cloning NascB-F5053 and NascB-S1868

Primer #	Gene	Primer	5'→3'
1	NascB-F5053 NascB-S1868	Forward	ctgagaatctctacttccaaggcgctagc GTGACCACCACCACCGCCACGCTGAC CT
2	NascB-F5053 NascB-S1868	Reverse	cctttcgggctttgttagcagccggatc TCACCAGGTGGCGGGGATCGCCCGC

General procedures for overexpression, purification and spectral characterization of cytochrome P450s

The expression of pET28b-NznB+09 was performed by transforming the plasmid into the competent cell *E. coli* strain C41 (DE3) and selected on Luria-Bertani (LB) medium plates containing100 µg/mL kanamycin. A single colony was grown overnight in LB broth containing the same concentration of the antibiotics. The main culture was prepared by inoculating 1% of each overnight culture into a 2 L baffled flask containing 500 ml of Terrific Broth (TB) containing 100 µg/mL kanamycin. The cultures were incubated at 37°C and 250 rpm for 3-4 h unless the absorbance A_{590nm} = 1. The expression of NznB was induced by the addition of 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 0.5 mM delta-5-aminolevulinic acid (w/v) as well as 1mM thiamine. The cultures were incubated further for 30 h, 18°C at 160 rpm. Cells were harvested by centrifugation at 5000×*g* at 4°C for 10 min and the cell pellet was stored at –80°C until purification of the protein.

The cell pellet was resuspended in 5% culture volume of lysis buffer (50 mM kPi, pH 8.0 containing 200 mM $(NH_4)_2SO_4$, 20% glycerol, 0.5 mM EDTA, 10 mM beta-mercaptoethanol and 1mM phenylmethane sulfonyl fluoride (PMSF) and disrupted by sonication. The cell lysate was centrifugated (35,000 rpm for 35 min) and filtered (0.2 µM Millipore filter). The soluble His_8-tagged NznB was purified by affinity

chromatography using Ni-NTA (Qiagen) column and the collected fractions were analyzed by SDS-PAGE. The suitable red fractions were pooled and dialyzed at 4°C three times using 50 mM Tris, pH 7.4 containing 50 mM kPi, pH 8.0 containing 200 mM $(NH_4)_2SO_4$, 20% glycerol, 0.5 mM EDTA, and 1 mM dithiothreitol (DTT) against a total of 6 L buffer.

UV–Vis spectra for the purified cytochomes were recorded at room temperature in buffer (50 mM Tris-HCl buffer, pH 7.4 containing 5% glycerol) was used for the spectral measurements of the oxidized and reduced form. NasB was reduced by the addition of a small amount of sodium dithionite. The concentration of the P450s was estimated by CO-difference spectra, assuming ε(459–490) = 91 mM−1 cm−1 according to the method of Omura and Sato.

Assembly of Sequence Similarity Network

The top 100 BlastP hits from the following characterized CDPS were compiled and duplicates were removed.

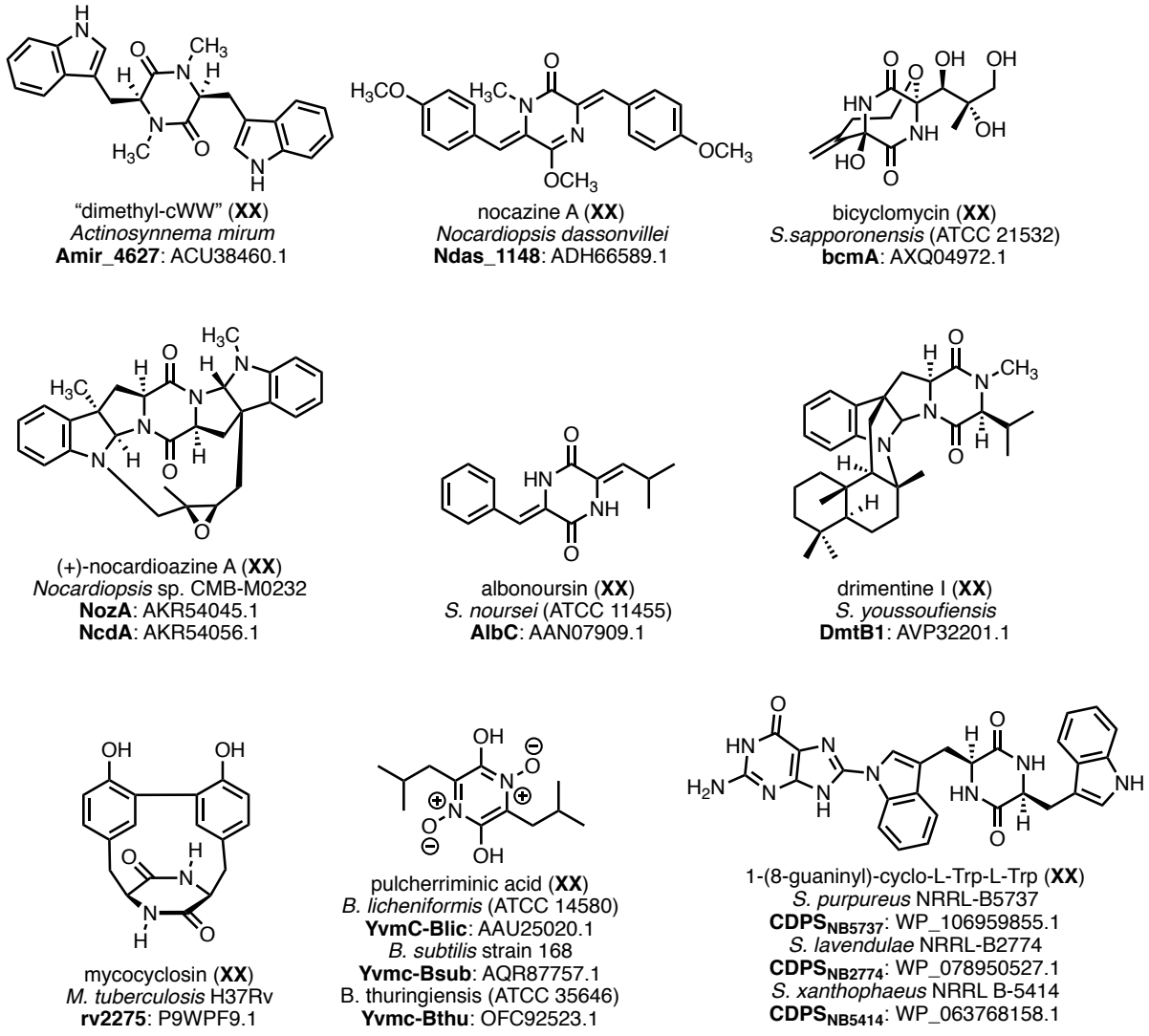


Figure S3-1. Natural products from characterized gene clusters and CDPS (in bold) used to construct CDPS library and sequence similarity network.

Optimized reaction conditions for analytical scale P450 reactions

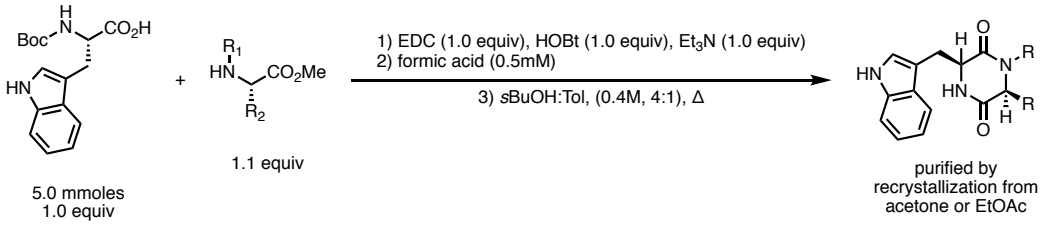
The conversion of DKPs by P450s were carried out with the heterologous redox electron partners Fdx and FdR from *S. olerecea*. The *in vitro* reaction mixture included 50 mM Tris-HCl, pH 7.4, 10 μM P450, 20 μM Fdx and 3 μM FdR at the end volume of

250 µL. Without exceeding 1% of the total reaction volume, 1mM of DKP dissolved in DMSO was added (in the case of heterodimerization reaction, both substrates were added to a final concentration of 1mM). The reaction was initiated by adding NADPH (1 mM). Two of the control reactions including all the contents except for NADPH or P450 were used as controls. The reaction was incubated at 30°C for 2 h agitating at 600 rpm in a thermoshaker (Multi-thermoshaker, Benchmark). Reactions were quenched by the addition of 3 volumes of methanol, and centrifuged at 17,000rpm and directly used for HPLC analysis. Reactions were resolved using a linear gradient of 15-75% acetonitrile: water (0.1% formic acid) over 16 min (2 mL/min flow rate, 40ºC) on a Phenomenex Luna 5µ C-18(2) 100A, 250 × 4.60mm 5 micron column. Percent conversions were calculated using a 5 point standard curve.

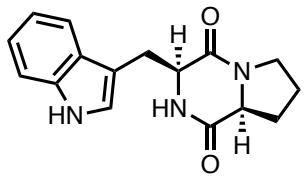
General Procedure for Biotransformation Reactions

A single colony from a fresh transformation of P450 and sFdx/sFdR plamsid was grown overnight in LB broth. The main culture was prepared by inoculating 1% of each overnight culture into a 2 L baffled flask containing 500 ml of Terrific Broth (TB) containing appropriate antibiotics. The cultures were incubated at 37°C and 250 rpm for 3-4 h unless the absorbance A_{590nm} = 1. The expression of P450 and electron transfer proteins was induced by the addition of 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 0.5 mM delta-5-aminolevulinic acid (w/v) as well as 1mM thiamine. The cultures were incubated further for 30 h, 18°C at 160 rpm. Cells were harvested by centrifugation at 5000×*g* at 4°C for 10. Pellets were resuspended in a minimal buffer at 4ºC (50mM kPi (pH = 8.0), and 20% glycerol, 500mL buffer/500mL media) and this was repeated 3 times. Pellets were resuspended in a final 500mL of buffer also containing 100mM EDTA, and substrate(s) were added to a final concentration of 1mM as a solution/mixture in 40% β-hydroxycyclodextrin. Reaction was monitored by HPLC until no further conversion, and then cells were removed by centrifugation. Media was treated with 20g/L of washed XAD-16 resin and incubated overnight with shaking. Resin was then removed via filtration and crude product was extracted off of resin by soaking resin in 500mL methanol. Methanol wash was then concentrated and residue was filtered through celite and directly utilized for purification by prep-HPLC.

General procedure for synthesis of diketopiperazine substrates (XX-XX)



4: Synthesized via general procedure, 34% yield over 3 steps.

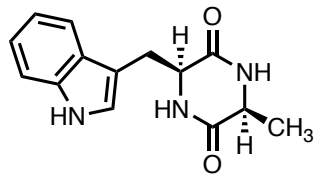


4

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.83 (s, 1H), 7.73 (s, 1H), 7.54 (d, *J* = 7.9 Hz, 1H), 7.29 (d, *J* = 8.1 Hz, 1H), 7.15 (d, *J* = 2.2 Hz, 1H), 7.02 (t, *J* = 7.5 Hz, 1H), 6.93 (t, *J* = 7.4 Hz, 1H), 1.94 (dtd, *J* = 9.8, 6.9, 2.7 Hz, 1H), 1.66 (ddd, *J* = 19.1, 9.1, 5.6 Hz, 1H), 1.59 (ddp, *J* = 12.2, 8.4, 4.6, 4.1 Hz, 1H), 1.35 (dtd, *J* = 12.1, 10.3, 7.8 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 169.43, 165.92, 136.37, 127.77, 124.82, 121.30, 119.08, 118.65, 111.65, 109.70, 58.85, 55.66, 45.03, 28.11, 26.23, 22.31.

5: Synthesized via general procedure, 23% yield over 3 steps.



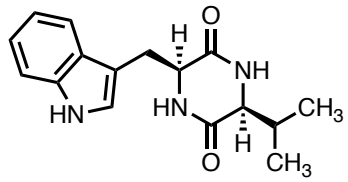
5

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.87 (s, 1H), 8.00 (s, 1H), 7.89 (s, 1H), 7.54 (d, *J* = 7.9 Hz, 1H), 7.28 (d, *J* = 8.0 Hz, 1H), 7.04 – 6.97 (m, 2H), 6.91 (t, *J* = 7.4 Hz, 1H), 4.08 (d, *J* = 2.7 Hz, 1H), 3.56 (dd, *J* = 7.0, 2.0 Hz, 1H), 3.32 (s, 1H), 3.21 (dd, *J* = 14.4, 4.1

Hz, 1H), 2.98 (dd, *J* = 14.4, 4.6 Hz, 1H), 2.47 (t, *J* = 1.9 Hz, 1H), 0.38 (d, *J* = 7.0 Hz, 3H).

^{13}C NMR (151 MHz, dmso) δ 168.14, 167.16, 136.20, 128.22, 124.98, 121.22, 119.40, 118.80, 111.51, 108.91, 55.82, 50.20, 29.27, 19.98.

6: Synthesized via general procedure, 52% yield over 3 steps.

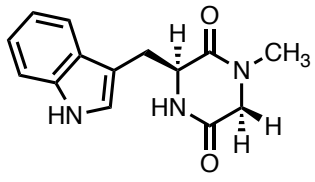


6

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.84 (s, 1H), 7.97 (s, 1H), 7.85 (d, *J* = 2.3 Hz, 1H), 7.57 (d, *J* = 7.9 Hz, 1H), 7.25 (d, *J* = 8.0 Hz, 1H), 7.04 (d, *J* = 2.3 Hz, 1H), 6.98 (t, *J* = 7.5 Hz, 1H), 6.90 (t, *J* = 7.4 Hz, 1H), 4.10 (d, *J* = 2.3 Hz, 1H), 3.51 – 3.34 (m, 1H), 3.18 (dd, *J* = 14.4, 5.0 Hz, 1H), 3.04 (dd, *J* = 14.5, 4.6 Hz, 1H), 2.47 (d, *J* = 1.9 Hz, 1H), 1.61 (dtd, *J* = 14.0, 7.1, 3.8 Hz, 1H), 0.57 (d, *J* = 7.1 Hz, 4H), 0.13 (d, *J* = 6.8 Hz, 3H).

^{13}C NMR (151 MHz, dmso) δ 167.78, 166.70, 136.44, 128.35, 124.91, 121.07, 119.32, 118.59, 111.43, 109.23, 59.71, 55.60, 31.52, 29.17, 18.76, 16.52.

7: Synthesized via general procedure, 21% yield over 3 steps.

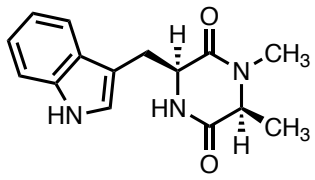


7

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.93 (s, 1H), 8.19 – 8.13 (m, 1H), 7.39 (d, *J* = 7.9 Hz, 1H), 7.31 (d, *J* = 8.1 Hz, 1H), 7.04 – 7.00 (m, 1H), 6.98 (d, *J* = 2.3 Hz, 1H), 6.96 – 6.90 (m, 1H), 4.05 (q, *J* = 3.9 Hz, 1H), 3.32 (d, *J* = 4.7 Hz, 1H), 3.30 – 3.20 (m, 2H), 2.94 (dd, *J* = 14.4, 4.6 Hz, 1H), 2.51 – 2.45 (m, 1H), 2.37 (d, *J* = 17.0 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 166.73, 165.53, 136.42, 127.59, 125.46, 121.42, 119.11, 118.76, 111.68, 108.29, 55.87, 50.84, 33.22, 30.55.

8: Synthesized via general procedure, 23% yield over 3 steps.

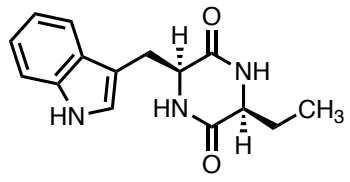


8

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.87 (s, 1H), 8.18 (s, 1H), 7.45 (d, *J* = 8.0 Hz, 1H), 7.27 (d, *J* = 8.0 Hz, 1H), 6.99 (d, *J* = 8.9 Hz, 2H), 4.10 (d, *J* = 3.8 Hz, 1H), 3.54 (q, *J* = 7.1 Hz, 1H), 3.23 (dd, *J* = 14.4, 3.8 Hz, 1H), 2.95 (dd, *J* = 14.5, 4.5 Hz, 1H), 2.47 (d, *J* = 2.9 Hz, 1H), 0.18 (d, *J* = 6.8 Hz, 3H).

^{13}C NMR (151 MHz, dmso) δ 167.51, 165.65, 136.26, 128.11, 125.08, 121.33, 119.33, 118.76, 111.58, 108.73, 56.73, 56.06, 31.77, 30.56, 17.19.

9: Synthesized via general procedure, 43% yield over 3 steps.

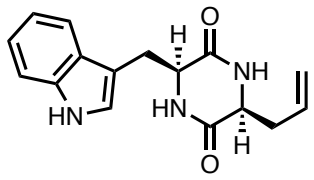


9

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.92 – 10.76 (m, 1H), 8.01 (s, 1H), 7.92 (s, 1H), 7.56 (d, *J* = 7.9 Hz, 1H), 7.26 (d, *J* = 8.0 Hz, 1H), 7.02 (d, *J* = 2.2 Hz, 1H), 6.99 (t, *J* = 7.5 Hz, 1H), 6.90 (t, *J* = 7.4 Hz, 1H), 4.09 (s, 1H), 3.51 (s, 1H), 3.21 (dd, *J* = 14.5, 4.4 Hz, 1H), 3.00 (dd, *J* = 14.5, 4.6 Hz, 1H), 1.17 – 1.02 (m, 1H), 0.81 (dp, *J* = 14.1, 7.2 Hz, 1H), 0.13 (t, *J* = 7.4 Hz, 3H).

^{13}C NMR (151 MHz, dmso) δ 167.57, 167.21, 136.34, 128.30, 124.95, 121.11, 119.40, 118.64, 111.43, 109.05, 55.74, 55.33, 29.22, 26.43, 8.46.

10: Synthesized via general procedure, 32% yield over 3 steps.

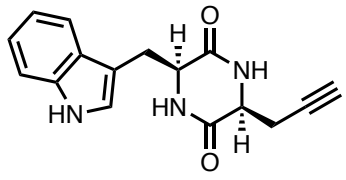


10

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.89 (d, *J* = 2.4 Hz, 1H), 8.08 (s, 1H), 7.75 (d, *J* = 2.5 Hz, 1H), 7.54 (d, *J* = 7.9 Hz, 1H), 7.29 (d, *J* = 8.1 Hz, 1H), 7.00 (dd, *J* = 7.3, 5.0 Hz, 2H), 6.91 (d, *J* = 7.5 Hz, 1H), 5.00 (ddt, *J* = 17.1, 10.2, 7.1 Hz, 1H), 4.59 (dd, *J* = 10.1, 2.3 Hz, 1H), 4.37 (dd, *J* = 17.0, 2.2 Hz, 1H), 4.09 (d, *J* = 2.8 Hz, 1H), 3.68 – 3.51 (m, 1H), 3.23 (dd, *J* = 14.4, 4.1 Hz, 1H), 2.97 (dd, *J* = 14.4, 4.7 Hz, 1H), 1.94 – 1.68 (m, 1H), 1.17 (dt, *J* = 14.4, 7.5 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 167.23, 166.73, 136.29, 133.66, 128.28, 125.12, 121.19, 119.47, 118.76, 118.05, 111.48, 108.89, 55.78, 54.10, 38.55, 29.31.

11: Synthesized via general procedure, 28% yield over 3 steps.

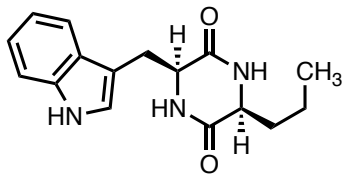


11

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.86 (d, *J* = 2.5 Hz, 1H), 8.00 (dd, *J* = 29.8, 2.2 Hz, 2H), 7.52 (d, *J* = 7.9 Hz, 1H), 7.30 (d, *J* = 8.1 Hz, 1H), 7.11 (d, *J* = 2.3 Hz, 1H), 7.02 (t, *J* = 7.5 Hz, 1H), 6.93 (t, *J* = 7.4 Hz, 1H), 4.10 (s, 1H), 3.83 (s, 1H), 3.24 – 3.05 (m, 2H), 2.68 (s, 1H), 2.14 (ddd, *J* = 16.8, 4.8, 2.7 Hz, 1H), 1.92 (ddd, *J* = 16.8, 5.5, 2.6 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 167.51, 165.97, 136.35, 128.07, 124.69, 121.28, 119.11, 118.80, 111.59, 109.29, 80.70, 73.87, 55.62, 53.54, 30.20, 23.75.

12: Synthesized via general procedure, 29% yield over 3 steps.

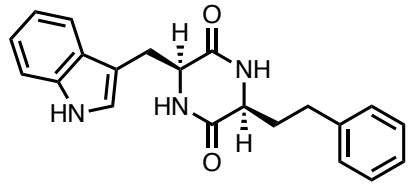


12

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.86 (s, 1H), 8.02 (s, 1H), 7.90 (d, *J* = 2.3 Hz, 1H), 7.54 (d, *J* = 7.9 Hz, 1H), 7.27 (d, *J* = 8.1 Hz, 1H), 7.02 – 6.97 (m, 2H), 6.89 (t, *J* = 7.4 Hz, 1H), 4.08 (q, *J* = 3.6 Hz, 1H), 3.45 (q, *J* = 4.5, 3.1 Hz, 1H), 3.23 (dd, *J* = 14.4, 4.0 Hz, 1H), 2.97 (dd, *J* = 14.4, 4.7 Hz, 1H), 2.47 (t, *J* = 2.0 Hz, 1H), 1.00 – 0.76 (m, 1H), 0.65 – 0.29 (m, 5H).

^{13}C NMR (151 MHz, dmso) δ 167.41, 167.36, 136.36, 128.31, 125.04, 121.10, 119.44, 118.69, 111.41, 108.99, 55.84, 54.36, 36.07, 29.28, 17.15, 14.00.

13: Synthesized via general procedure, 41% yield over 3 steps.

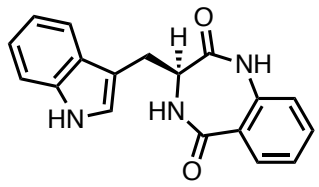


13

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.86 (d, *J* = 2.3 Hz, 1H), 8.14 (d, *J* = 2.1 Hz, 1H), 8.06 (d, *J* = 2.2 Hz, 1H), 7.60 (d, *J* = 7.9 Hz, 1H), 7.22 – 7.13 (m, 3H), 7.07 (t, *J* = 7.4 Hz, 1H), 7.03 (d, *J* = 2.2 Hz, 1H), 6.98 (t, *J* = 7.5 Hz, 1H), 6.92 (t, *J* = 7.4 Hz, 1H), 6.75 (d, *J* = 7.5 Hz, 2H), 4.13 (d, *J* = 4.5 Hz, 1H), 3.57 (d, *J* = 5.9 Hz, 1H), 3.40 – 3.22 (m, 2H), 2.99 (dd, *J* = 14.4, 4.6 Hz, 1H), 1.78 (ddq, *J* = 25.4, 13.8, 6.8, 5.2 Hz, 2H), 1.21 (tt, *J* = 11.0, 5.2 Hz, 1H), 0.88 (tt, *J* = 12.3, 6.0 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 167.46, 167.05, 141.84, 136.26, 128.46, 128.44, 128.32, 125.92, 125.16, 121.07, 119.48, 118.73, 111.61, 108.87, 55.81, 54.14, 35.47, 29.93, 29.22.

14: Synthesized via general procedure, 26% yield over 3 steps.

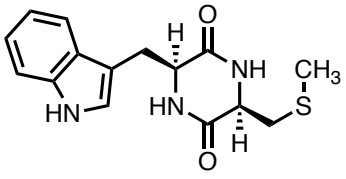


14

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.83 (d, *J* = 2.4 Hz, 1H), 10.41 (s, 1H), 8.40 (d, *J* = 5.9 Hz, 1H), 7.62 (d, *J* = 7.8 Hz, 1H), 7.46 (t, *J* = 7.1 Hz, 2H), 7.29 (d, *J* = 8.0 Hz, 1H), 7.21 (d, *J* = 2.3 Hz, 1H), 7.14 (t, *J* = 7.5 Hz, 1H), 7.07 (d, *J* = 8.1 Hz, 1H), 7.03 – 6.96 (m, 1H), 3.85 (dt, *J* = 9.0, 5.6 Hz, 1H), 3.20 (dd, *J* = 14.9, 5.2 Hz, 1H), 2.97 (dd, *J* = 14.9, 9.3 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 171.96, 168.01, 137.21, 136.46, 132.65, 130.80, 127.36, 126.65, 124.74, 124.31, 121.38, 121.34, 118.75, 118.67, 111.81, 110.05, 23.95.

15: Synthesized via general procedure, 41% yield over 3 steps.



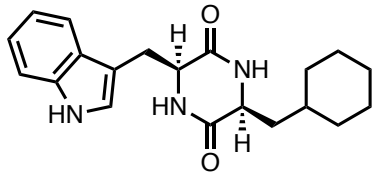
15

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.88 (d, *J* = 2.4 Hz, 1H), 8.14 (d, *J* = 2.1 Hz, 1H), 7.96 (d, *J* = 2.1 Hz, 1H), 7.56 (d, *J* = 7.9 Hz, 1H), 7.26 (d, *J* = 8.1 Hz, 1H), 7.02 – 6.96 (m, 2H), 6.90 (t, *J* = 7.5 Hz, 1H), 4.11 (d, *J* = 4.3 Hz, 1H), 3.64 (d, *J* = 5.8 Hz, 1H), 3.25 (dd, *J* = 14.4, 3.8 Hz, 1H), 2.96 (dd, *J* = 14.4, 4.6 Hz, 1H), 1.70 (d, *J* = 1.1 Hz, 3H), 1.57

(ddq, *J* = 28.7, 18.2, 6.0, 5.1 Hz, 2H), 1.18 (ddt, *J* = 15.5, 10.7, 5.5 Hz, 1H), 1.01 – 0.82 (m, 1H).

^{13}C NMR (151 MHz, dmso) δ 167.37, 166.77, 136.24, 128.23, 125.08, 121.09, 119.45, 118.72, 111.66, 108.82, 55.78, 53.42, 33.13, 29.20, 28.04, 14.67.

16: Synthesized via general procedure, 45% yield over 3 steps.

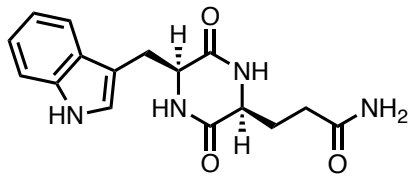


16

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.88 (d, *J* = 2.3 Hz, 1H), 8.01 (d, *J* = 2.5 Hz, 1H), 7.94 (d, *J* = 2.7 Hz, 1H), 7.92 (s, 0H), 7.52 (d, *J* = 7.9 Hz, 1H), 7.26 (d, *J* = 8.1 Hz, 1H), 7.04 – 6.96 (m, 2H), 6.89 (t, *J* = 7.5 Hz, 1H), 4.05 (d, *J* = 3.4 Hz, 1H), 3.40 (dt, *J* = 9.5, 3.6 Hz, 1H), 3.22 (dd, *J* = 14.4, 4.0 Hz, 1H), 2.95 (dd, *J* = 14.4, 4.7 Hz, 1H), 1.59 – 1.30 (m, 4H), 1.10 – 0.87 (m, 5H), 0.67 (ddd, *J* = 13.7, 9.2, 4.5 Hz, 1H), 0.39 (q, *J* = 11.0, 10.4 Hz, 2H), -0.11 (ddd, *J* = 14.0, 9.6, 4.8 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 168.03, 167.55, 162.72, 136.33, 128.19, 125.02, 121.10, 119.42, 118.73, 111.55, 108.93, 55.96, 52.00, 42.55, 36.21, 33.28, 32.32, 31.68, 31.20, 31.13, 29.58, 26.42, 26.09, 25.73.

17: Synthesized via general procedure, 21% yield over 3 steps.



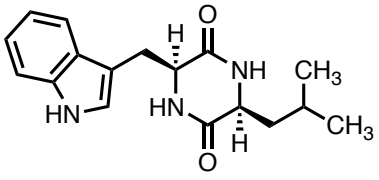
17

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.86 (s, 1H), 8.36 (d, *J* = 7.5 Hz, 1H), 7.73 (s, 1H), 7.46 (d, *J* = 7.8 Hz, 1H), 7.31 (d, *J* = 8.1 Hz, 1H), 7.13 (d, *J* = 2.2 Hz, 1H), 7.08 – 7.01 (m, 1H), 6.96 (t, *J* = 7.4 Hz, 1H), 4.50 (td, *J* = 8.0, 5.8 Hz, 1H), 4.00 (dd, *J* = 8.9, 3.9 Hz,

1H), 3.57 (d, *J* = 1.2 Hz, 3H), 3.15 (dd, *J* = 14.7, 5.6 Hz, 1H), 3.05 (dd, *J* = 14.6, 8.5 Hz, 1H), 2.31 – 2.14 (m, 1H), 2.03 (qdd, *J* = 16.2, 9.8, 6.6 Hz, 2H), 1.76 (ddt, *J* = 13.5, 9.3, 4.6 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 177.84, 173.15, 172.62, 136.52, 127.41, 124.24, 121.42, 118.88, 118.37, 111.88, 109.70, 55.61, 53.46, 52.35, 29.45, 27.33, 25.59.

18: Synthesized via general procedure, 53% yield over 3 steps.

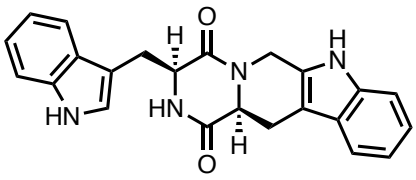


18

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.89 (d, *J* = 2.4 Hz, 1H), 8.03 (d, *J* = 2.6 Hz, 1H), 7.93 (d, *J* = 2.8 Hz, 1H), 7.52 (d, *J* = 7.9 Hz, 1H), 7.27 (d, *J* = 8.1 Hz, 1H), 7.04 – 6.96 (m, 2H), 6.89 (t, *J* = 7.4 Hz, 1H), 4.06 (d, *J* = 3.4 Hz, 1H), 3.36 (dt, *J* = 8.1, 3.6 Hz, 1H), 3.23 (dd, *J* = 14.4, 4.0 Hz, 1H), 2.95 (dd, *J* = 14.4, 4.7 Hz, 1H), 1.22 – 1.08 (m, 1H), 0.60 (ddd, *J* = 13.6, 9.2, 4.6 Hz, 1H), 0.49 (d, *J* = 6.5 Hz, 3H), 0.39 (d, *J* = 6.6 Hz, 3H), -0.06 (ddd, *J* = 14.0, 9.4, 5.1 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 167.88, 167.51, 136.33, 128.21, 125.06, 121.18, 119.42, 118.76, 111.52, 108.92, 55.96, 52.77, 44.11, 29.54, 23.28, 23.12, 21.69.

19: Synthesized via general procedure, 42% yield over 3 steps.



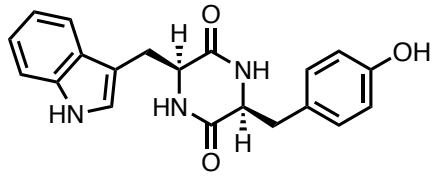
19

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.84 – 10.69 (m, 1H), 8.40 (d, *J* = 2.4 Hz, 0H), 7.51 – 7.46 (m, 0H), 7.23 (d, *J* = 8.0 Hz, 0H), 7.11 (d, *J* = 7.7 Hz, 1H), 6.98 (t, *J* = 7.5 Hz, 0H),

6.92 – 6.87 (m, 1H), 6.81 – 6.71 (m, 1H), 5.24 (d, *J* = 16.5 Hz, 1H), 4.02 (d, *J* = 16.5 Hz, 1H), 3.92 (dd, *J* = 11.8, 4.1 Hz, 1H), 3.25 (dd, *J* = 14.4, 4.0 Hz, 1H), 3.05 (dd, *J* = 14.4, 4.5 Hz, 1H), 2.57 (dd, *J* = 14.6, 4.1 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 166.13, 165.10, 136.25, 136.15, 129.59, 128.15, 126.67, 124.49, 121.15, 121.12, 119.01, 118.88, 118.64, 117.79, 111.36, 111.33, 108.46, 105.90, 56.27, 55.83, 30.64, 26.10.

20: Synthesized via general procedure, 46% yield over 3 steps.

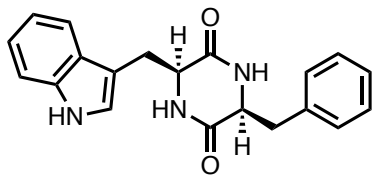


20

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.88 (d, *J* = 2.4 Hz, 1H), 9.14 (s, 1H), 7.81 (d, *J* = 2.6 Hz, 1H), 7.62 (d, *J* = 2.7 Hz, 1H), 7.45 (d, *J* = 7.9 Hz, 1H), 7.29 (d, *J* = 8.1 Hz, 1H), 7.12 – 7.02 (m, 1H), 7.01 – 6.93 (m, 2H), 6.56 (d, *J* = 8.0 Hz, 2H), 6.49 (d, *J* = 8.0 Hz, 2H), 4.29 (d, *J* = 4.5 Hz, 1H), 3.98 – 3.82 (m, 1H), 3.79 – 3.66 (m, 1H), 3.46 (p, *J* = 5.9 Hz, 1H), 2.76 (dd, *J* = 14.5, 4.3 Hz, 1H), 2.45 – 2.25 (m, 2H), 1.76 (dd, *J* = 13.6, 7.0 Hz, 1H), 1.42 – 1.21 (m, 2H), 0.99 (dd, *J* = 6.2, 1.1 Hz, 3H), 0.80 (td, *J* = 7.4, 1.1 Hz, 3H).

^{13}C NMR (151 MHz, dmso) δ 167.16, 166.66, 156.39, 136.48, 131.13, 127.89, 126.82, 124.79, 121.29, 119.15, 118.83, 115.30, 111.74, 109.31, 67.57, 56.30, 55.65, 32.14, 30.39, 23.53, 10.54.

21: Synthesized via general procedure, 41% yield over 3 steps.

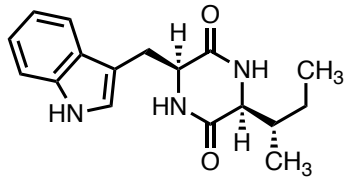


21

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.87 (d, J = 2.5 Hz, 1H), 7.90 (d, J = 2.5 Hz, 1H), 7.70 (d, J = 2.7 Hz, 1H), 7.46 (d, J = 7.9 Hz, 1H), 7.29 (d, J = 8.1 Hz, 1H), 7.22 – 7.10 (m, 3H), 7.04 (t, J = 7.5 Hz, 1H), 6.96 (t, J = 7.5 Hz, 1H), 6.93 (d, J = 2.2 Hz, 1H), 6.69 – 6.65 (m, 2H), 3.94 (s, 1H), 3.82 (t, J = 4.1 Hz, 1H), 2.77 (dd, J = 14.5, 4.4 Hz, 1H), 2.46 (ddd, J = 39.5, 13.9, 5.3 Hz, 2H), 1.82 (dd, J = 13.5, 7.1 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 167.22, 166.59, 136.94, 136.45, 130.13, 128.46, 127.94, 126.79, 124.84, 121.31, 119.19, 118.84, 111.76, 109.21, 56.04, 55.68, 30.12.

22: Synthesized via general procedure, 55% yield over 3 steps.

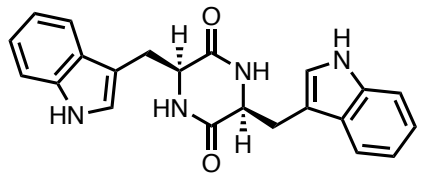


22

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.82 (d, J = 2.5 Hz, 1H), 7.99 (d, J = 2.1 Hz, 1H), 7.82 (d, J = 2.1 Hz, 1H), 7.57 (d, J = 7.9 Hz, 1H), 7.26 (s, 0H), 7.02 (d, J = 2.2 Hz, 1H), 6.98 (dd, J = 8.1, 6.8 Hz, 1H), 6.89 (t, J = 7.4 Hz, 1H), 3.50 – 3.47 (m, 1H), 3.21 (dd, J = 14.5, 4.5 Hz, 1H), 3.01 (dd, J = 14.4, 4.6 Hz, 1H), 1.27 (ddt, J = 10.4, 7.3, 3.6 Hz, 1H), 0.63 – 0.21 (m, 8H).

^{13}C NMR (151 MHz, dmso) δ 167.59, 166.65, 136.40, 128.42, 125.02, 121.03, 119.41, 118.60, 111.37, 109.19, 59.30, 55.60, 38.30, 28.87, 23.27, 14.87, 12.06.

23: Synthesized via general procedure, 52% yield over 3 steps.

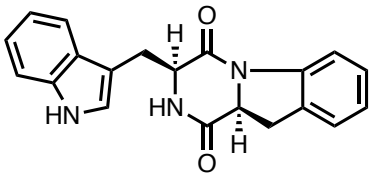


23

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.82 (d, *J* = 2.4 Hz, 1H), 7.69 (d, *J* = 2.6 Hz, 1H), 7.32 (d, *J* = 7.9 Hz, 1H), 7.26 (d, *J* = 8.1 Hz, 1H), 7.01 (t, *J* = 7.5 Hz, 1H), 6.92 (t, *J* = 7.4 Hz, 1H), 6.56 (d, *J* = 2.3 Hz, 1H), 3.88 – 3.79 (m, 1H), 3.32 (s, 1H), 2.67 (dd, *J* = 14.3, 4.2 Hz, 1H), 2.14 (dd, *J* = 14.3, 6.7 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 167.14, 136.46, 127.76, 124.83, 121.21, 118.96, 118.77, 111.68, 109.17, 55.69, 30.42.

24: Synthesized via general procedure, 41% yield over 3 steps.

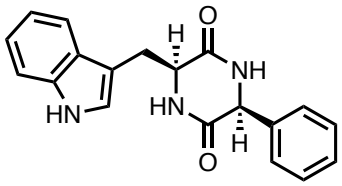


24

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.89 (dd, *J* = 8.0, 2.3 Hz, 1H), 9.05 (s, 1H), 8.70 (dd, *J* = 123.2, 7.5 Hz, 1H), 7.44 (t, *J* = 8.2 Hz, 2H), 7.32 (t, *J* = 7.6 Hz, 1H), 7.21 (t, *J* = 8.3 Hz, 1H), 7.19 – 7.12 (m, 2H), 7.04 (dd, *J* = 9.3, 5.8 Hz, 1H), 6.99 (t, *J* = 7.5 Hz, 1H), 6.95 (t, *J* = 7.4 Hz, 1H), 4.95 (dd, *J* = 11.0, 4.0 Hz, 1H), 4.51 – 4.41 (m, 1H), 3.44 (dd, *J* = 16.8, 11.0 Hz, 1H), 3.21 – 3.01 (m, 3H), 2.93 (dd, *J* = 16.8, 4.0 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 172.56, 172.49, 170.59, 161.10, 158.87, 141.92, 136.45, 130.31, 127.90, 127.48, 126.05, 125.26, 124.25, 124.12, 124.05, 121.38, 118.84, 118.43, 118.35, 111.83, 110.13, 109.46, 58.02, 53.79, 53.58, 52.42, 52.30, 33.04, 27.29.

26: Synthesized via general procedure, 22% yield over 3 steps.

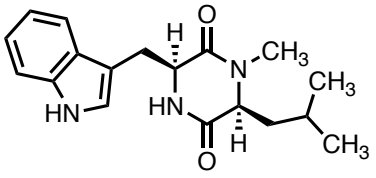


26

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.89 (d, J = 2.4 Hz, 1H), 8.27 (dd, J = 22.8, 2.3 Hz, 2H), 7.54 (d, J = 7.9 Hz, 1H), 7.37 (d, J = 8.1 Hz, 1H), 7.11 – 7.06 (m, 1H), 7.06 – 6.98 (m, 2H), 6.94 (t, J = 7.5 Hz, 2H), 6.89 (t, J = 7.4 Hz, 1H), 6.35 (d, J = 7.6 Hz, 2H), 4.73 (d, J = 2.1 Hz, 1H), 4.22 (dd, J = 4.5, 2.4 Hz, 1H), 3.28 (dd, J = 14.3, 4.3 Hz, 1H), 3.01 (dd, J = 14.3, 4.6 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 166.70, 165.51, 139.44, 136.42, 128.34, 128.06, 127.63, 127.57, 125.19, 121.22, 119.48, 118.83, 111.52, 108.92, 58.86, 55.92, 29.13.

27: Synthesized via general procedure, 27% yield over 3 steps.

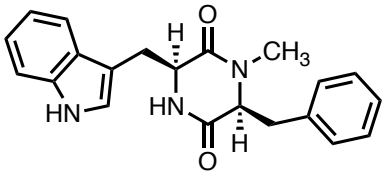


27

^{1}H NMR (599 MHz, dmso) δ 10.90, 10.89, 8.13, 8.13, 7.46, 7.45, 7.27, 7.26, 7.01, 7.00, 7.00, 6.99, 6.99, 6.92, 6.91, 6.91, 6.89, 4.11, 4.10, 4.10, 4.09, 3.43, 3.42, 3.41, 3.41, 3.24, 3.23, 3.21, 3.21, 2.97, 2.96, 2.95, 2.94, 2.59.

^{13}C NMR (151 MHz, dmso) δ 167.19, 166.15, 136.46, 128.24, 124.94, 121.25, 119.27, 118.72, 111.65, 108.99, 59.36, 56.32, 41.57, 32.49, 30.30, 24.80, 23.06, 21.74.

28: Synthesized via general procedure, 25% yield over 3 steps.



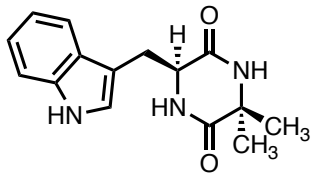
28

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.89 (d, J = 2.4 Hz, 1H), 7.97 (d, J = 3.0 Hz, 1H), 7.40 (d, J = 7.9 Hz, 1H), 7.30 (d, J = 8.0 Hz, 1H), 7.20 (t, J = 7.4 Hz, 2H), 7.15 (t, J = 6.9 Hz, 1H), 7.08 – 7.02 (m, 1H), 6.96 (t, J = 7.4 Hz, 1H), 6.93 (d, J = 2.2 Hz, 1H), 6.78 – 6.74 (m, 2H), 3.99 (t, J = 5.4 Hz, 1H), 3.93 (dt, J = 7.0, 3.6 Hz, 1H), 2.66 (dd, J = 14.4, 4.3

Hz, 1H), 2.61 (s, 3H), 2.51 (dd, *J* = 14.0, 5.0 Hz, 1H), 2.10 (t, *J* = 6.4 Hz, 1H), 2.07 (t, *J* = 6.0 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 166.13, 166.04, 137.44, 136.57, 129.96, 128.72, 127.77, 127.07, 124.77, 121.35, 118.96, 118.83, 111.84, 109.39, 63.17, 55.88, 38.15, 33.23, 30.98.

29: Synthesized via general procedure, 24% yield over 3 steps.

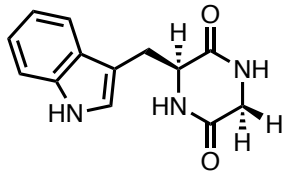


29

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.86 (d, *J* = 2.5 Hz, 1H), 7.94 (d, *J* = 12.4 Hz, 2H), 7.56 (d, *J* = 7.9 Hz, 1H), 7.27 (d, *J* = 8.0 Hz, 1H), 7.01 (d, *J* = 2.3 Hz, 1H), 6.99 (t, *J* = 7.5 Hz, 1H), 6.90 (t, *J* = 7.4 Hz, 1H), 4.12 (td, *J* = 4.2, 1.8 Hz, 1H), 3.24 (dd, *J* = 14.4, 3.9 Hz, 1H), 2.98 (dd, *J* = 14.4, 4.6 Hz, 1H), 1.11 (s, 3H), 0.39 (s, 3H).

^{13}C NMR (151 MHz, dmso) δ 170.61, 166.65, 136.17, 128.24, 125.00, 121.18, 119.48, 118.76, 111.48, 108.90, 56.12, 55.12, 29.27, 28.73, 27.98.

30: Synthesized via general procedure, 30% yield over 3 steps.

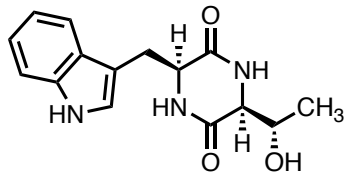


30

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.90 (d, *J* = 2.4 Hz, 1H), 8.08 (d, *J* = 2.7 Hz, 1H), 7.75 (d, *J* = 2.8 Hz, 1H), 7.52 (d, *J* = 7.9 Hz, 1H), 7.30 (d, *J* = 8.1 Hz, 1H), 7.04 – 7.00 (m, 2H), 6.92 (t, *J* = 7.4 Hz, 1H), 3.99 (q, *J* = 4.1 Hz, 1H), 3.28 (dd, *J* = 17.3, 3.0 Hz, 1H), 3.21 (dd, *J* = 14.5, 4.7 Hz, 1H), 2.98 (dd, *J* = 14.5, 4.6 Hz, 1H), 2.75 (d, *J* = 17.2 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 168.38, 166.12, 136.36, 127.91, 125.03, 121.33, 119.13, 118.87, 111.61, 108.78, 55.88, 44.28, 29.61.

31: Synthesized via general procedure, 34% yield over 3 steps.

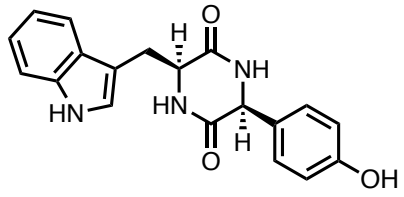


31

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.84 (d, *J* = 2.4 Hz, 1H), 8.02 (d, *J* = 3.0 Hz, 1H), 7.77 (d, *J* = 3.0 Hz, 1H), 7.51 (d, *J* = 7.8 Hz, 1H), 7.30 (d, *J* = 8.1 Hz, 1H), 7.12 (d, *J* = 2.3 Hz, 1H), 7.10 – 6.98 (m, 1H), 6.93 (t, *J* = 7.4 Hz, 1H), 4.97 (d, *J* = 5.9 Hz, 1H), 4.29 (dd, *J* = 4.6, 0.8 Hz, 1H), 3.89 (td, *J* = 7.6, 6.4, 3.2 Hz, 1H), 3.74 (pd, *J* = 6.4, 2.9 Hz, 1H), 3.49 (t, *J* = 3.1 Hz, 1H), 3.48 – 3.44 (m, 1H), 3.24 – 3.16 (m, 2H), 1.29 (dddd, *J* = 16.2, 13.3, 10.5, 6.4 Hz, 2H), 0.99 (d, *J* = 6.1 Hz, 2H), 0.92 (d, *J* = 6.5 Hz, 3H), 0.80 (t, *J* = 7.4 Hz, 2H).

^{13}C NMR (151 MHz, dmso) δ 168.32, 166.89, 136.58, 128.06, 124.42, 121.20, 118.96, 118.69, 111.65, 110.13, 67.61, 67.57, 60.82, 56.20, 32.14, 31.62, 23.53, 20.11, 10.54.

32: Synthesized via general procedure, 22% yield over 3 steps.

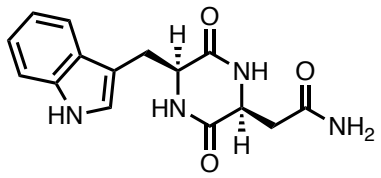


32

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.91 (d, *J* = 2.5 Hz, 1H), 8.20 (d, *J* = 2.4 Hz, 1H), 8.16 (d, *J* = 2.3 Hz, 1H), 7.54 (d, *J* = 7.9 Hz, 1H), 7.35 (d, *J* = 8.1 Hz, 1H), 7.02 (t, *J* = 7.5 Hz, 1H), 7.00 (d, *J* = 2.2 Hz, 1H), 6.89 (t, *J* = 7.5 Hz, 1H), 6.37 – 6.31 (m, 2H), 6.17 – 6.11 (m, 2H), 4.59 (t, *J* = 1.8 Hz, 1H), 4.17 (dd, *J* = 4.4, 2.3 Hz, 1H), 3.25 (dd, *J* = 14.3, 4.5 Hz, 1H), 2.99 (dd, *J* = 14.3, 4.6 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 166.60, 165.99, 156.92, 136.39, 129.84, 128.72, 128.34, 125.16, 121.19, 119.47, 118.80, 114.84, 111.48, 108.97, 58.40, 55.92, 29.12.

33: Synthesized via general procedure, 17% yield over 3 steps.

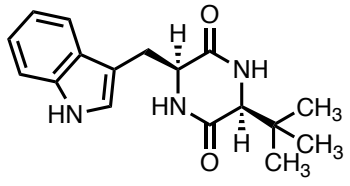


33

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.87 (d, *J* = 2.6 Hz, 1H), 7.92 (d, *J* = 2.1 Hz, 1H), 7.64 (d, *J* = 2.1 Hz, 1H), 7.54 (d, *J* = 7.9 Hz, 1H), 7.31 (d, *J* = 8.1 Hz, 1H), 7.09 (d, *J* = 2.3 Hz, 1H), 7.02 (t, *J* = 7.5 Hz, 1H), 6.93 (t, *J* = 7.4 Hz, 1H), 6.88 (d, *J* = 8.2 Hz, 2H), 4.14 – 4.07 (m, 1H), 3.97 (ddd, *J* = 9.9, 4.5, 1.8 Hz, 1H), 3.18 (dd, *J* = 14.6, 5.5 Hz, 1H), 3.07 (dd, *J* = 14.6, 4.5 Hz, 1H), 2.14 (dd, *J* = 15.6, 4.4 Hz, 1H), 1.45 (dd, *J* = 15.6, 8.3 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 171.84, 167.65, 167.34, 136.31, 128.08, 124.84, 121.31, 119.24, 118.83, 111.68, 109.28, 55.62, 51.71, 38.63, 29.10.

34: Synthesized via general procedure, 44% yield over 3 steps.

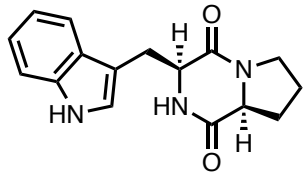


34

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.89 (d, *J* = 2.5 Hz, 1H), 8.02 (d, *J* = 2.8 Hz, 1H), 7.89 (s, 1H), 7.52 (d, *J* = 7.9 Hz, 1H), 7.29 (d, *J* = 8.0 Hz, 1H), 7.08 – 6.97 (m, 2H), 6.92 (t, *J* = 7.4 Hz, 1H), 4.01 (q, *J* = 4.2 Hz, 1H), 3.22 (dd, *J* = 14.4, 4.7 Hz, 1H), 2.98 (dd, *J* = 14.4, 4.6 Hz, 1H), 2.89 (q, *J* = 6.9 Hz, 1H), 1.01 (d, *J* = 6.9 Hz, 3H).

^{13}C NMR (151 MHz, dmso) δ 168.90, 168.56, 136.33, 127.93, 124.96, 121.32, 119.15, 118.87, 111.60, 108.83, 56.36, 49.19, 29.58, 18.78.

37: Synthesized via general procedure, 25% yield over 3 steps.

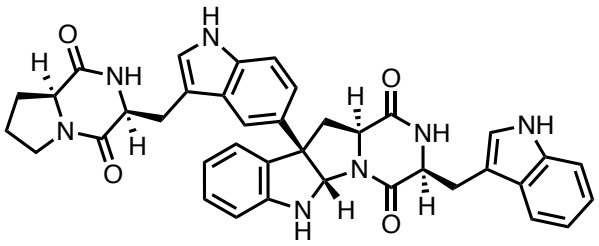


37

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.90 (s, 1H), 8.17 (s, 1H), 7.48 (d, *J* = 7.9 Hz, 1H), 7.27 (d, *J* = 8.1 Hz, 1H), 6.99 (t, *J* = 7.6 Hz, 1H), 6.95 (s, 1H), 6.90 (t, *J* = 7.5 Hz, 1H), 4.18 (d, *J* = 12.8 Hz, 1H), 4.11 (s, 1H), 3.45 – 3.25 (m, 4H), 2.89 (d, *J* = 4.6 Hz, 0H), 2.14 (t, *J* = 12.8 Hz, 1H), 1.34 (d, *J* = 12.7 Hz, 1H), 1.17 – 0.96 (m, 3H), 0.08 (dd, *J* = 18.1, 8.8 Hz, 1H), -0.82 (tt, *J* = 12.8, 6.0 Hz, 1H).

^{13}C NMR (151 MHz, dmso) δ 166.51, 163.69, 136.16, 128.23, 125.38, 121.30, 119.53, 118.75, 111.46, 108.36, 58.18, 55.46, 41.52, 31.14, 30.39, 29.31, 23.78, 23.74.

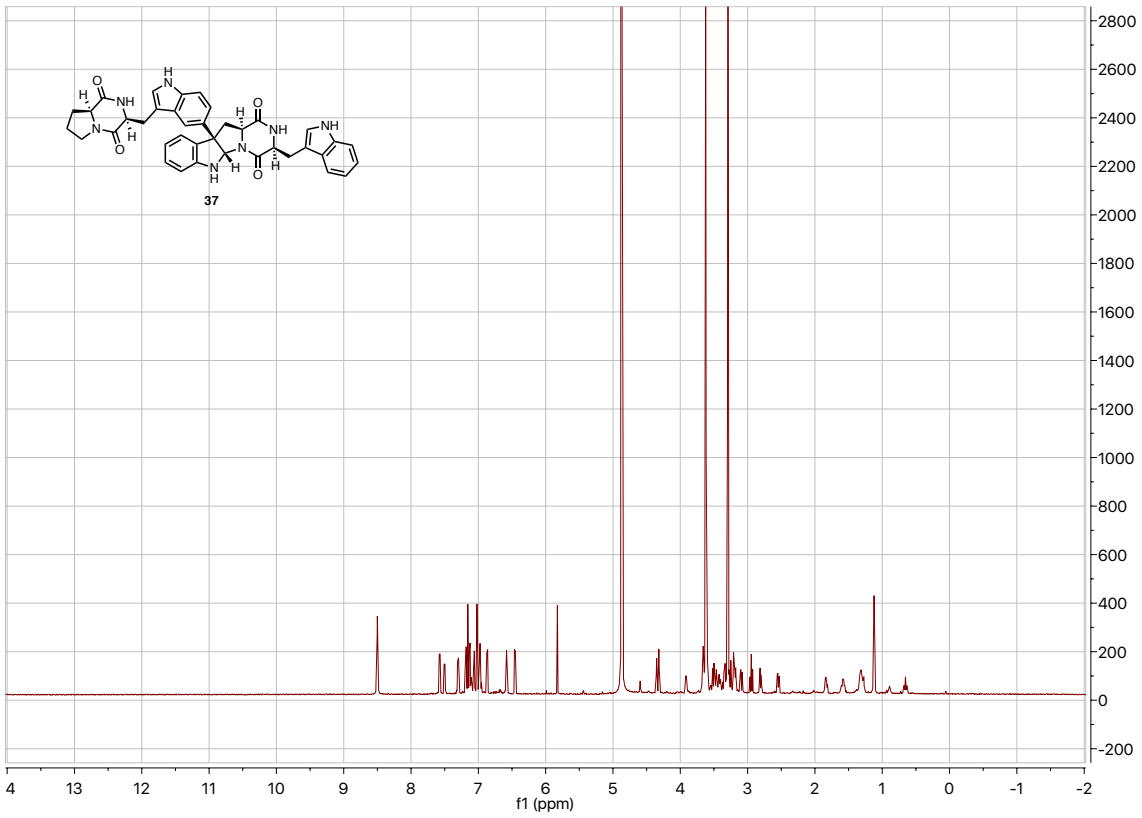
NMR Data for Heterodimer 37



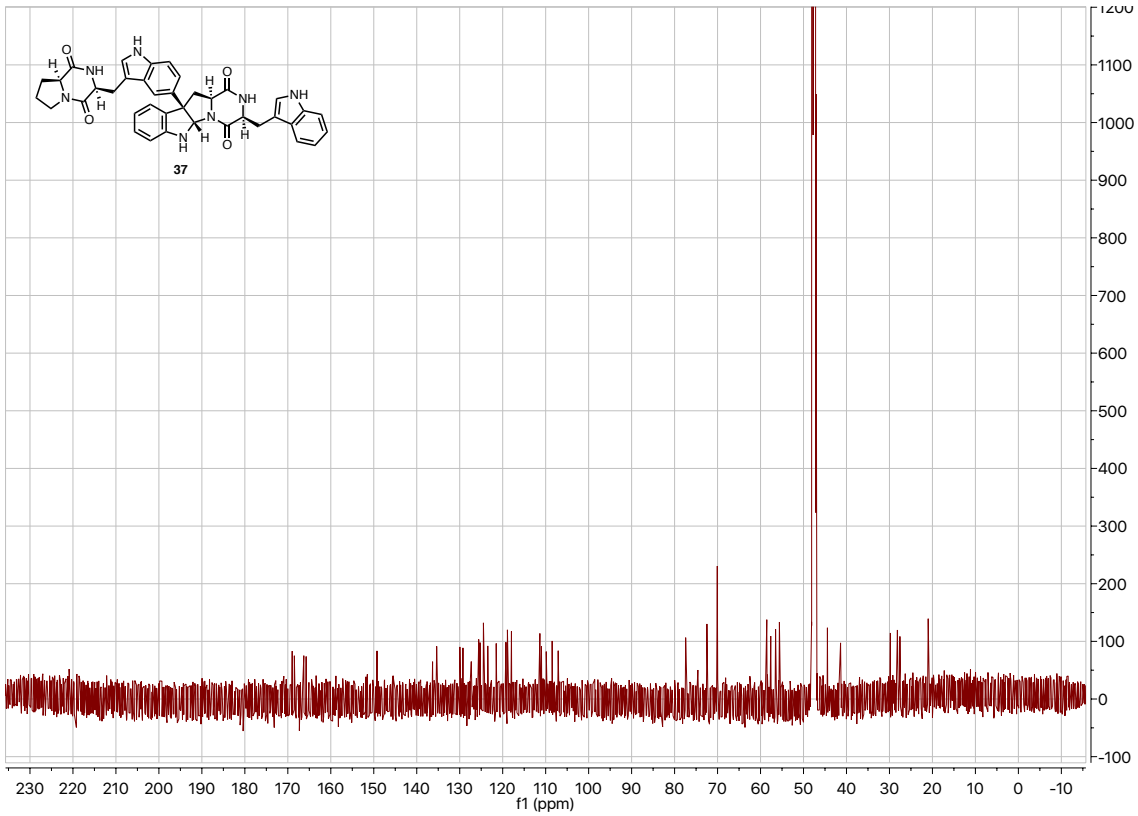
37

^{1}H NMR (599 MHz, Methanol-d_4) δ 8.50 (s, 1H), 7.57 (d, *J* = 7.6 Hz, 1H), 7.51 – 7.49 (m, 1H), 7.30 (d, *J* = 8.1 Hz, 1H), 7.19 (s, 0H), 7.16 (s, 1H), 7.12 (dt, *J* = 14.7, 7.1 Hz, 1H), 7.06 (t, *J* = 7.7 Hz, 1H), 7.02 (s, 0H), 7.01 – 6.95 (m, 1H), 6.87 (d, *J* = 7.5 Hz, 1H), 5.83 (s, 0H), 4.43 – 4.27 (m, 1H), 3.91 (dd, *J* = 11.1, 6.3 Hz, 1H), 3.66 (d, *J* = 5.2 Hz, 1H), 3.54 – 3.40 (m, 1H), 3.39 – 3.30 (m, 0H), 3.25 – 3.12 (m, 1H), 3.09 (dd, *J* = 14.6, 4.2 Hz, 1H), 2.95 (t, *J* = 12.0 Hz, 1H), 2.81 (dd, *J* = 12.0, 5.5 Hz, 1H), 2.54 (dd, *J* = 12.0, 5.5 Hz, 1H), 1.83 (dt, *J* = 12.9, 6.8 Hz, 1H), 1.59 (p, *J* = 10.5, 10.0 Hz, 1H), 1.43 – 1.26 (m, 2H), 1.12 (t, *J* = 4.8 Hz, 2H), 0.76 – 0.58 (m, 1H).

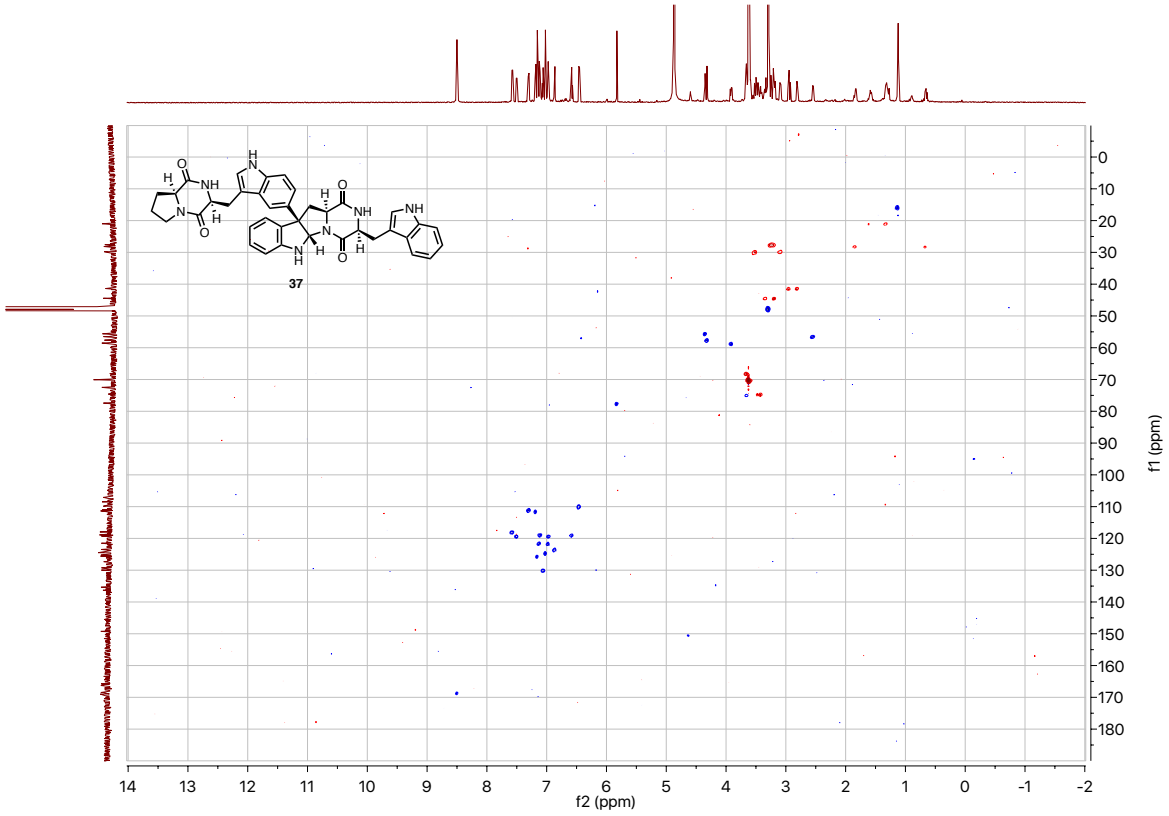
^{13}C NMR (151 MHz, cd_3od) δ 168.98, 166.28, 165.74, 149.25, 135.36, 129.96, 129.25, 125.58, 125.25, 124.48, 123.47, 121.65, 121.51, 119.25, 119.07, 118.90, 118.83, 117.96, 111.34, 110.98, 109.86, 108.49, 107.09, 77.42, 72.48, 70.07, 58.55, 57.62, 56.50, 55.61, 44.44, 41.38, 29.79, 28.15, 27.57, 20.99.



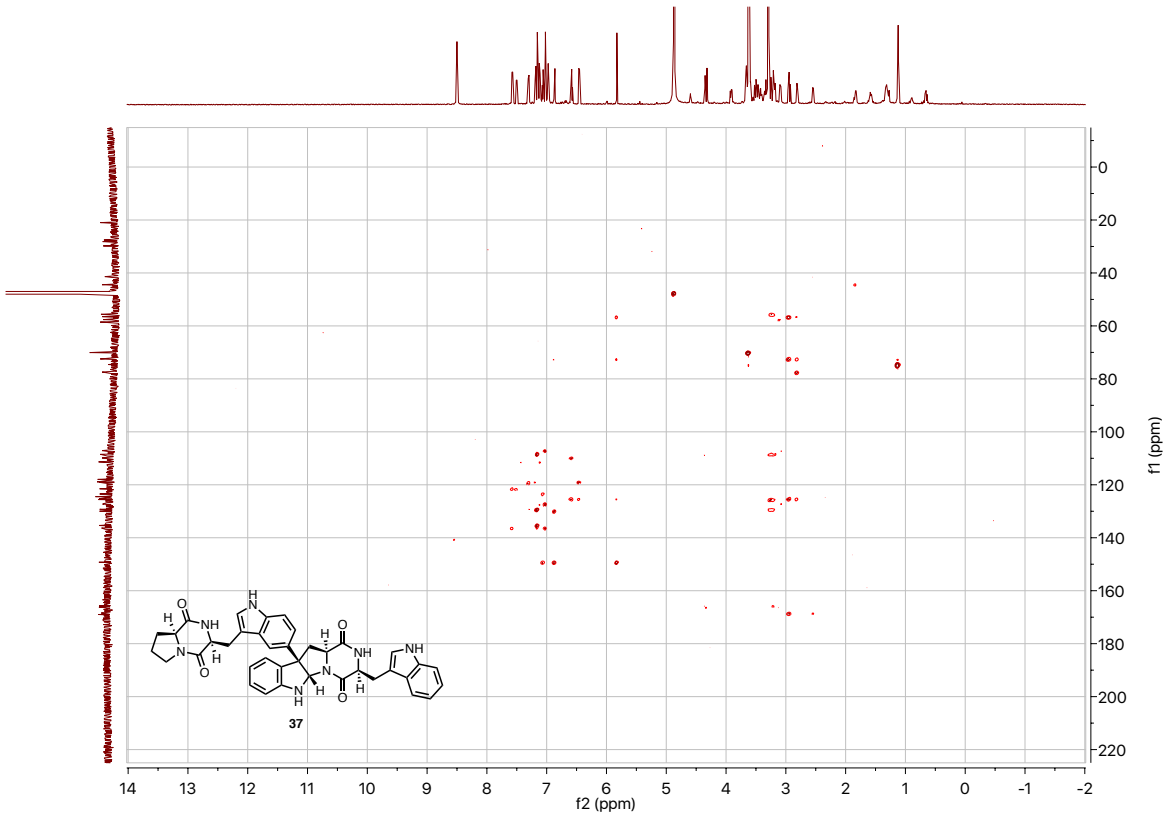
Proton NMR (600MHz) for **37**



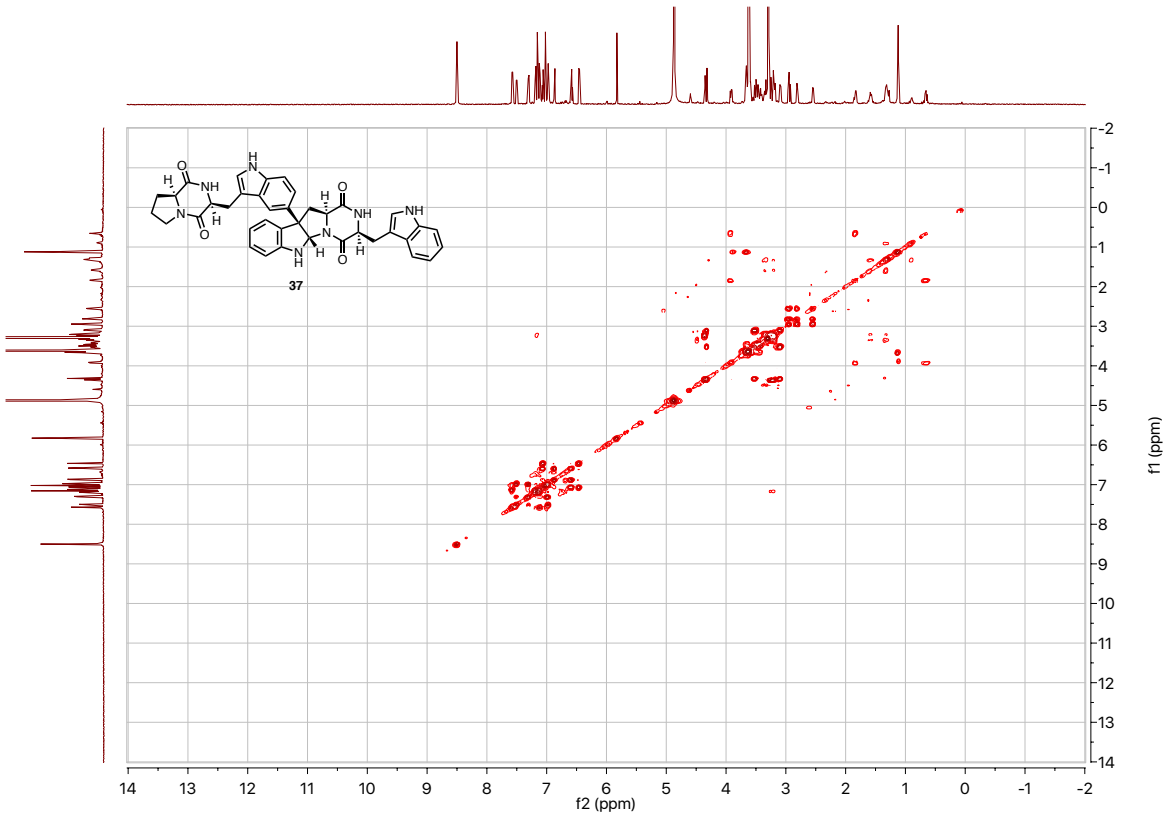
Carbon NMR (151MHz) for **37**



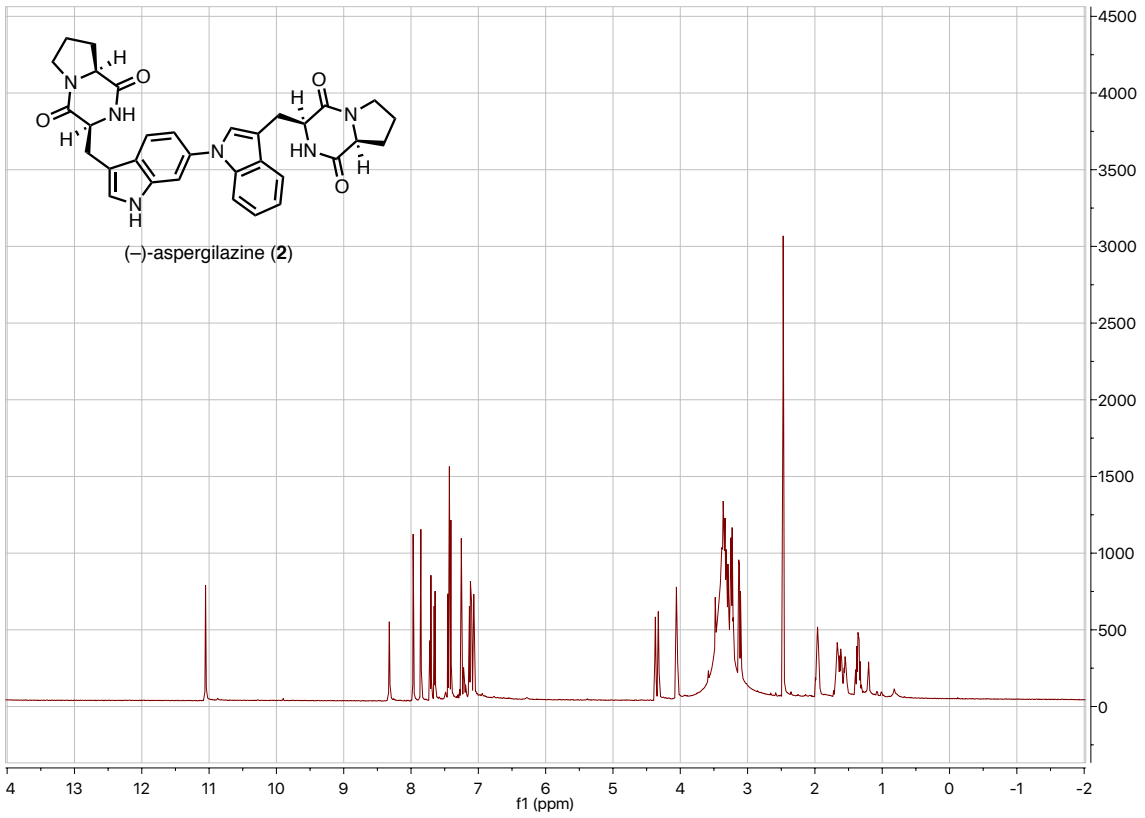
gHSQCAD for **37**



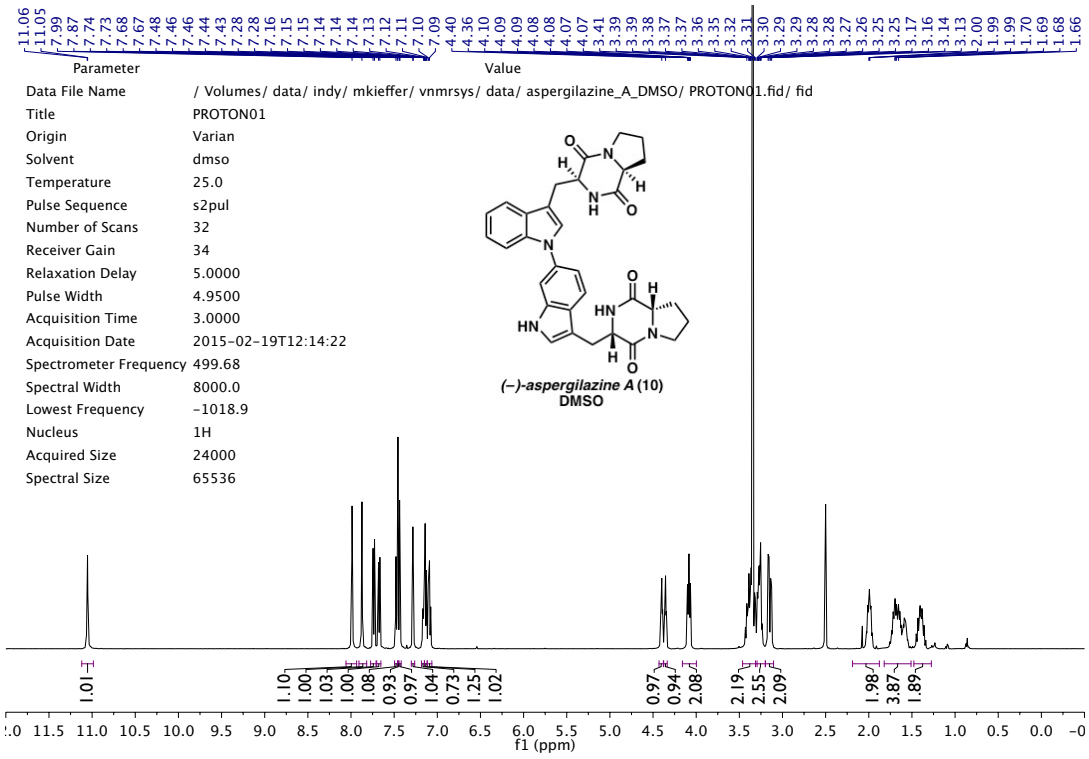
gHMBCAD for **37**



gCOSY for **73**



^{1}H-NMR Spectra for (–)-aspergilazine (2) in DMSO



Literature Spectra for (–)-Aspergilazine (2) from Reisman *et. al.*[1]

References

1. Chuang, K. V.; Kieffer, M. E.; Reisman, S. E., *Org. Lett.* **2016,** *18* (18), 4750-3.

Chapter 4

Preliminary Studies in the Structure and Mechanism of Bacterial Diketopiperazine Dimer Forming Cytochrome P450, NascB-F5053

4.1 Abstract

Chapters 2 and 3 described the initial discovery and characterization of the (+)- and (–)-naseseazines gene clusters, the pected discovery of the (–)-aspergilazine biosynthetic gene cluster, subsequent characterization of the encoded cytochromes P450. However, as stated at the outset of this dissertation, our original goal was not only the discovery these clusters, but to decipher the underlying molecular mechanisms behind the selectivity displayed by the enzymes that catalyze oxidative C–C and C–N bond formation. To that end, we report the structure of NascB-F5053, the cytochrome P450 which forms both the (–)-naseseazine and (–)-aspergilazine scaffold. To that end we obtained a 1.5 Å structure of NascB-F5053 in complex with two units of the native substrate **3** bound in the active site. Preliminary mutational analysis of the active site identified key residues requisite for binding, catalysis and selectivity. Further analysis of possible reaction mechanisms was also interrogated using synthetic probes and computational chemistry. All experimental structural work in this Chapter was performed by Dr. Sean Newmister, and computational work by Dr. Jacob Sanders (UCLA).

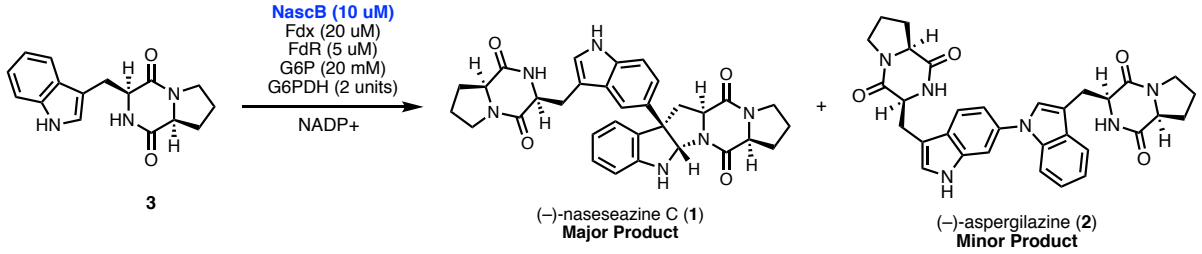


Figure 4-1. Native reaction of NascB-F5053.

Work described in Chapter 4 is ongoing at present.

4.2 Introduction

Dimeric diketopiperazine natural products are a large family of natural products with over 100 members isolated to date. Notably, the biological activity of these natural products is tethered not necessarily to their amino acid composition, but rather the regiochemistry and stereochemistry of dimerization. (+)-WIN 64821 (**4**) and (–)-ditryptophenaline (**5**) are of identical amino acid composition, however the configuration of the 4 contiguous stereocenters in **4** (C2,C2',C3, and C3') are all *R*, while the corresponding carbons in **5** are all *S*[1]. Despite their near identical composition, **4** is a potent substance P antagonist[1-6], while **5** shows little biological activity in the identical assay.

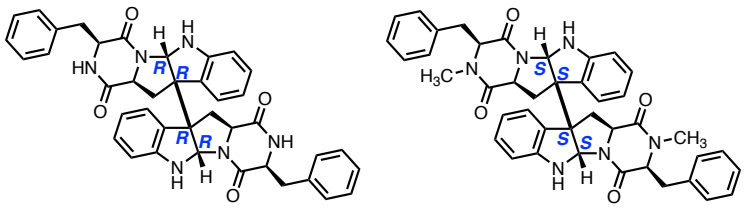


Figure 4-2. Structures of WIN 64821 (**4**) and (–)-ditryptophenaline (**5**).

While great efforts have been made in development of synthetic methods to access these complex molecules, the functionality of key intermediates in these syntheses installed to direct bond formation limit the potential product to a single regio- and stereochemical outcome. Our own studies as well as literature precedent[3, 7] has shown that the biosynthetic machinery that catalyzes dimerization is innately flexible, and capable of accommodating a variety of monomeric building blocks into dimers. In a directed biosynthesis study of the (+)-WIN 64821 producing *Aspergillus* sp. SC319, feeding a variety of amino acid starting materials resulted in the production of 36 analogues of **4** with substitution on the indoline ring (from functionalized tryptophans), substitution on the phenyl ring (functionalized phenylalanine), and complete substitution of phenylalanine for aliphatic amino acids. While this strategy enabled the rapid and facile assembly of complex congeners for structure activity relationships (SAR), this technique again only allows for a single regio- and stereochemical outcome, and is subject to the challenges associated with using a wild-type host for biotransformation.

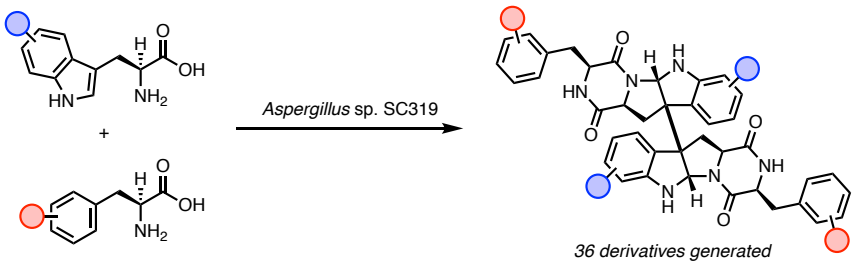


Figure 4-3. Directed biosynthesis of (+)-WIN 64821 analogues.

To access alternative configurations and connectivities, we hope to achieve a fundamental understanding of the molecular underpinning behind the exquisite selectivity displayed by these dimerization enzymes. As reported in Chapters 2 and 3, we have recently discovered cytochromes P450 from bacterial biosynthetic pathways involved in the biosynthesis of the (+)-naseseazine, (–)-naseseazine, and (–)-aspergilazine diketopiperazine dimers. These enzymes offer us a unique opportunity to study their dimerization selectivity, as they each display their own unique stereochemical and site-selectivity. To this end, we sought to crystallize these enzymes in complex with their substrate and/or product and interrogate the residues surrounding the active site to uncover how they may affect the reaction selectivity, as well as uncover the possible mechanism for this oxidative dimerization.

4.3 Results and Discussion

In selecting candidate enzymes for crystallization, we chose to exclude NznB as it was typically obtained in low yields from protein purification (3-4 mg/L), and was the least stable as indicated by its low turnover number and precipitation during purification and *in vitro* reactions. Of the remaining cytochromes, NascB-F5053 was selected as it consistently had the highest turnover numbers in biocatalytic reactions, and was capable of producing both (–)-naseseazine C (**1**) and (–)-aspergilazine (**2**), and crystallography offered a route to understand this enzyme's ability to catalyze two different bond forming cascades. Ultimately, we were able to obtain a 1.5 Å crystal structure (Figure 4-4, left) of NascB-F5053 in complex with two units of substrate **3** through molecular replacement. NascB-F5053 possesses the overall cytochrome P450

fold, and the cytochrome P450 to which it bears the most structural similarity (2.3 Å rmsd, DALI21)[8] is another carbon–carbon bond forming cytochrome, CYP121 from *Mycobacterium tuberculosis* (PDB 4G1X). Unlike CYP121, and other structurally characterized C–C bond forming cytochromes P450,[9-12] both substrates are buried side-by-side, deep in the active site directly above the plane of the heme cofactor.

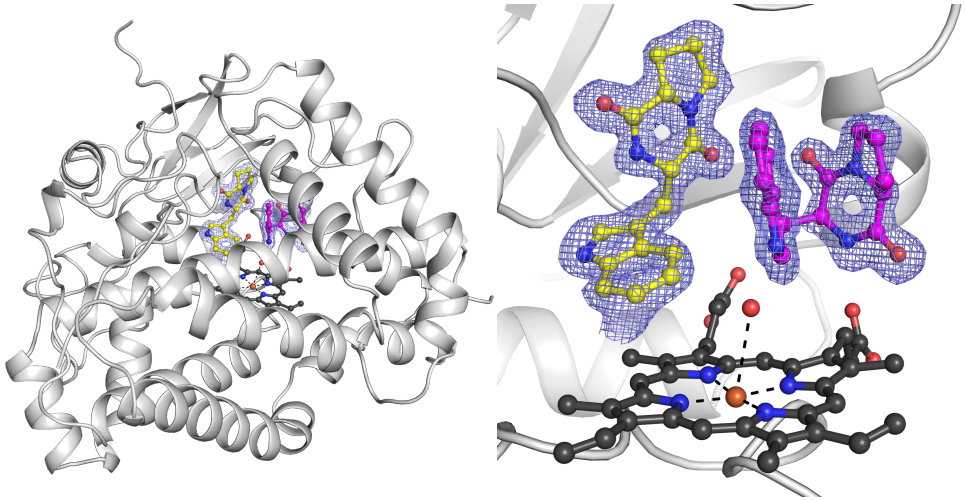


Figure 4-4. Left: Full view of crystal structure NascB-F5053$_{wt.}$
Right: Close up of active site showing unambiguous electron density of substrates shown in blue mesh (F_0-F_s) with substrates modeled in.

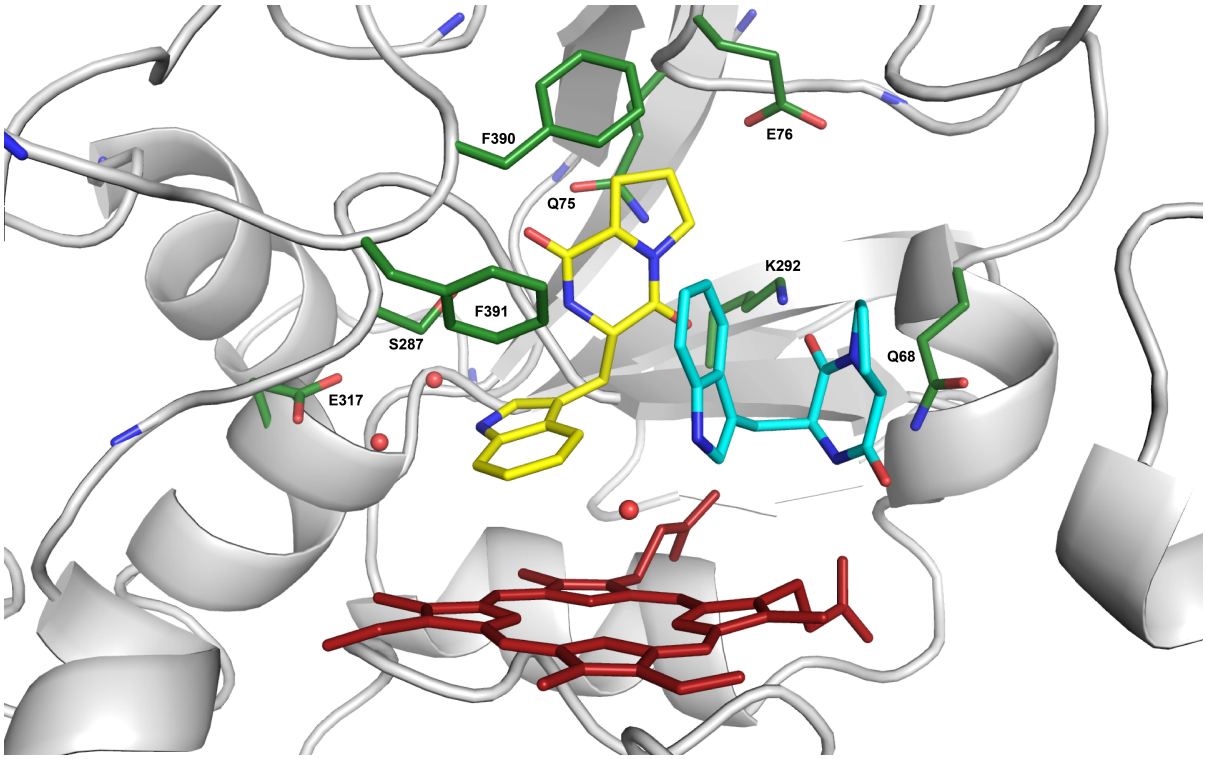


Figure 4-5. Active site of NascB-F5053wt showing two units of **3** bound in the active site and residues within 5 Å of either substrate.

4.3.1 Donor and Acceptor Substrate Binding Pockets

Given the high selectivity displayed in the assembly of non-symmetrical diketopiperazine dimers by NascB-F5053, in the previous chapter we hypothesized that within the active site there must be two dissimilar substrate binding sites in order to position the northern monomer, (referred to as the donor, colored blue in Figure 6) and southern monomer (hereon referred to as the acceptor, colored red in Figure 6), for selective dimerization.

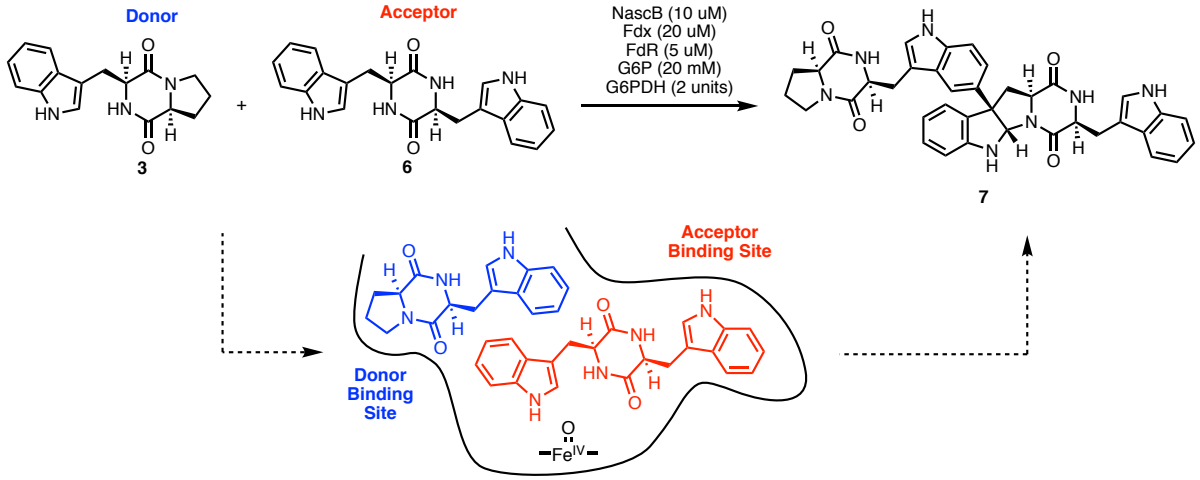


Figure 4-6. Donor and acceptor nomenclature.

4.3.2 Donor binding site

In Chapter 3 we described the heterodimerization reaction of **3** and **6**, wherein a single heterodimer, **7** was formed (Figure 4-6) and from this we gleaned that the acceptor pocket possesses a more relaxed substrate specificity, as it is able to accommodate the bulkier tryptophan side chain, while the donor pocket is more selective for its substrate and unable to accommodate larger DKP monomers. Examination of the donor-binding site (residues within 5 Å of the donor monomer) support this hypothesis, as the donor is completely enveloped by active site residues, with Phe390, Phe391,Gln75 and Glu76 completely surrounding the proline side chain leaving no space for larger amino acid side chains. Ser287 and Lys292 are hydrogen bonding with the carbonyl oxygens of the diketopiperazine, with Lys292 also hydrogen bonding to a carbonyl from the acceptor DKP (Figure 7).

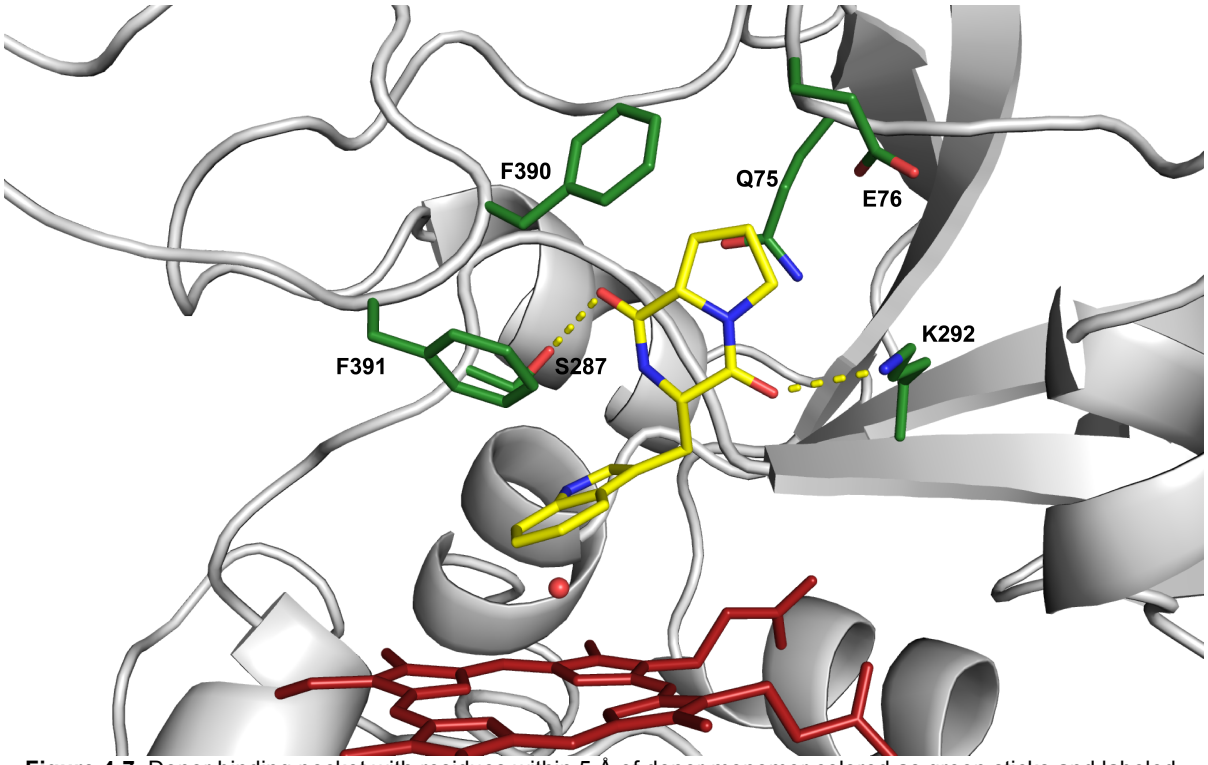


Figure 4-7. Donor binding pocket with residues within 5 Å of donor monomer colored as green sticks and labeled.

4.3.3 Acceptor binding site

Within 5 Å of the acceptor DKP there are only 3 amino acid residues (Figure 4-8). Lys292 is hydrogen bonding to carbonyl oxygens on both the donor and acceptor DKP, acting as a bridging residue between both binding sites. Gln68 lies directly behind the diketopiperazine core, while Phe 391 making edge-to-edge contact with the tryptophan arm of acceptor diketopiperazine. The lack of residues making van der Waals interactions with the acceptor residue may explain its substrate flexibility, as there are no residues around the proline side chain of the acceptor leaving space for larger substrates to occupy the active site.

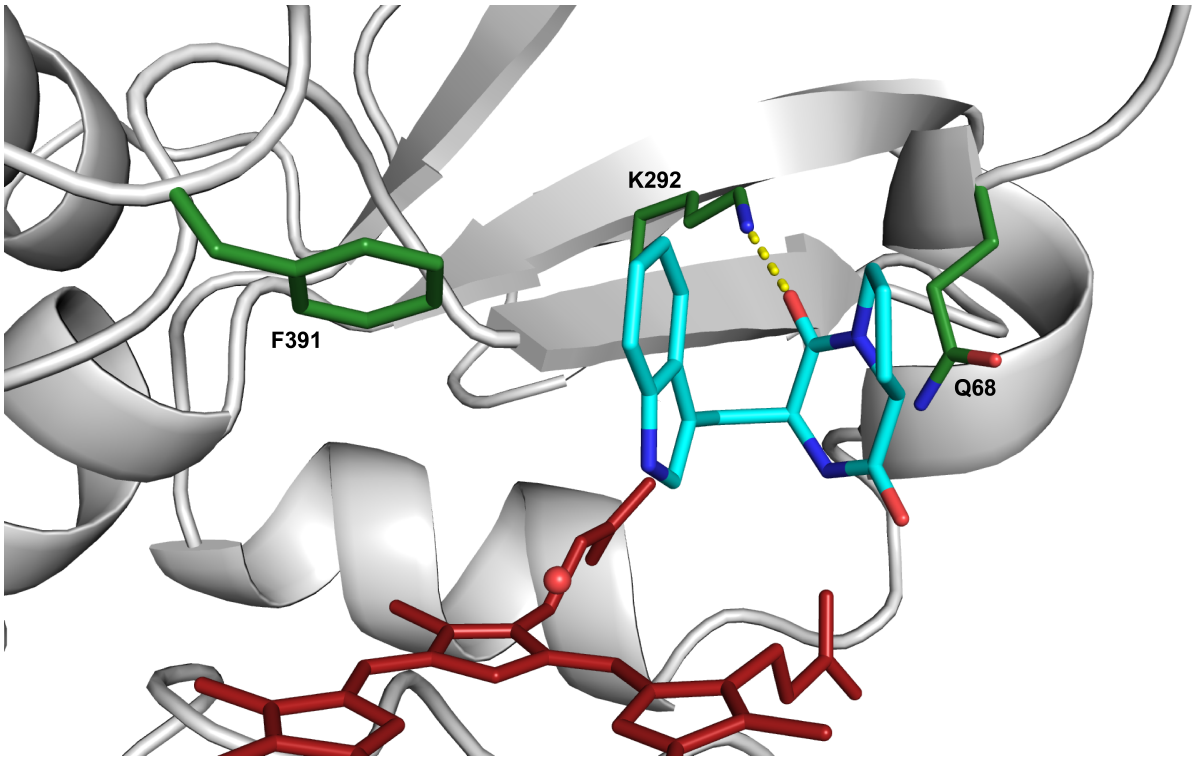


Figure 4-8. Acceptor binding pocket with residues within 5Å of acceptor monomer.

4.3.4. Structure Informed Mechanism

The substrates in the active site adopt a conformation such that the N–H of the indole the acceptor DKP (Figure 4-9, teal), and the C5–H of the donor DKP (Figure 9, yellow) both directed toward the axial water ligand. Based on this complex and literature precedent for bacterial C–C bond forming P450s,[9-12] we propose the following mechanism for dimerization (Figure 4-10). We hypothesize the initial substrate radical is generated from Michaelis complex **A** (Figure 10) via proton-coupled electron transfer (PCET), either through direct interaction with the heme or assisted by hydrogen bonding to yield intermediate **B** (Figure 10). A series of proton transfer steps yields cyclization of the pyrroloindoline to intermediate **C** (Figure 10). Carbon-carbon bond formation and dearomatization of donor yields intermediate **D** (Figure 10), which rearomatizes via PCET to product complex **E** (Figure 10) after product dissociation, the catalytic cycle can recur.

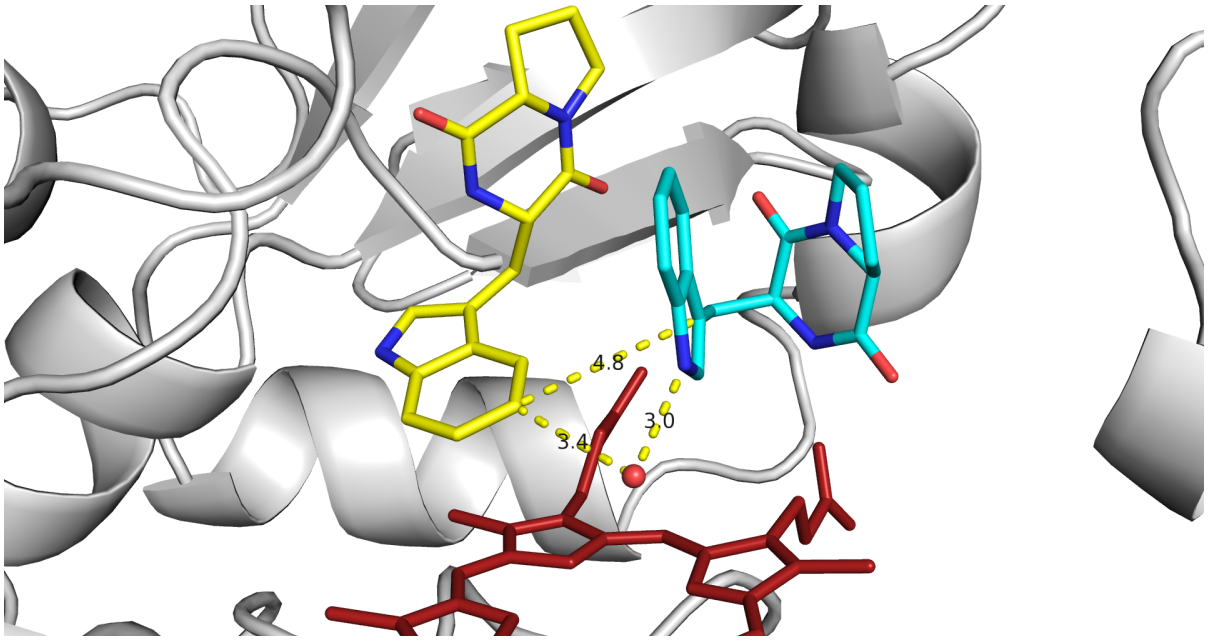


Figure 4-9. Active site showing bond to be formed and bonds (hypothetically) broken during dimerization reaction.

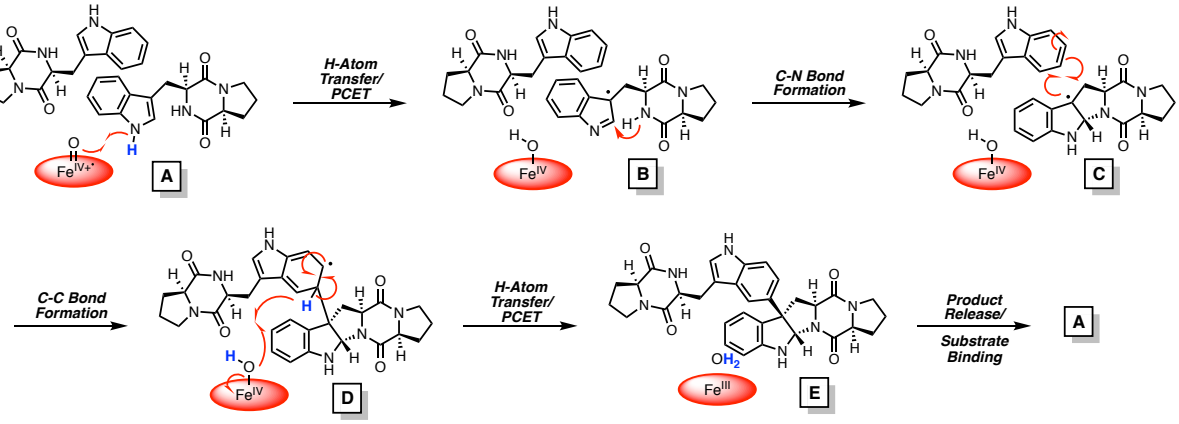


Figure 4-10. Proposed reaction mechanism based on structure.

4.3.5 Molecular Dynamics Simulations

Despite the regiochemistry for (–)-naseseazine being represented in the conformation of both substrates in the active site, and both reactive hydrogen atoms being oriented for catalysis, if the conformation of the substrates found in the crystal structure were to remain static throughout the course of the reaction, the stereochemical configuration of the product would not reflect that found experimentally in the product (–)-naseseazine C (**1**).

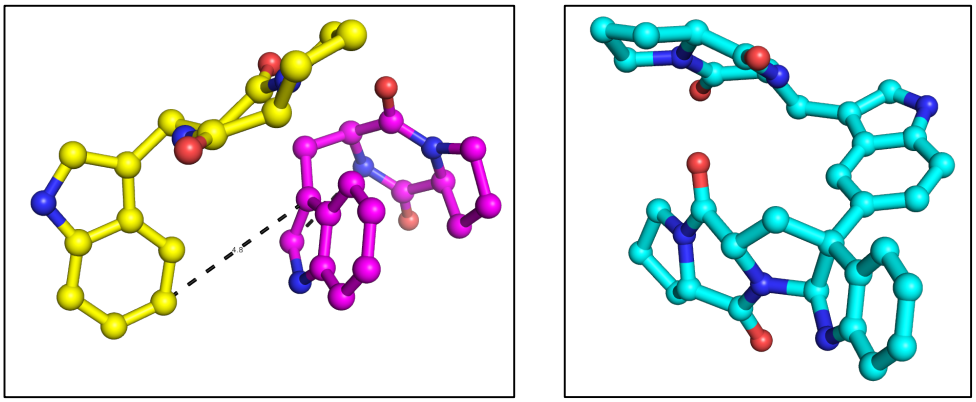


Figure 4-11. Left: conformation of substrates in active site, Right: 3D model of (–)-naseseazine C (**1**).

To evaluate how the protein, and specifically active site might undergo remodeling during the course of catalysis we worked with our collaborators in the Houk lab to model how the substrates may move during the course of a Molecular Dynamics (MD) simulation. The heme cofactor was modeled as both the resting state and and iron-oxo (compound 1) oxidations states, however under neither condition did conformation of the acceptor change dramatically enough to allow for formation of the desired stereochemistry. However, when the iron center was modeled as the compound 1 iron-oxo species, within the first 100 nanoseconds of the simulation, the indole N–H moved away from the iron-oxo and the N–H of the diketopiperazine moved within 2-4 Å of the iron-oxo and remained there for the majority of the simulation (Figure 4-12).

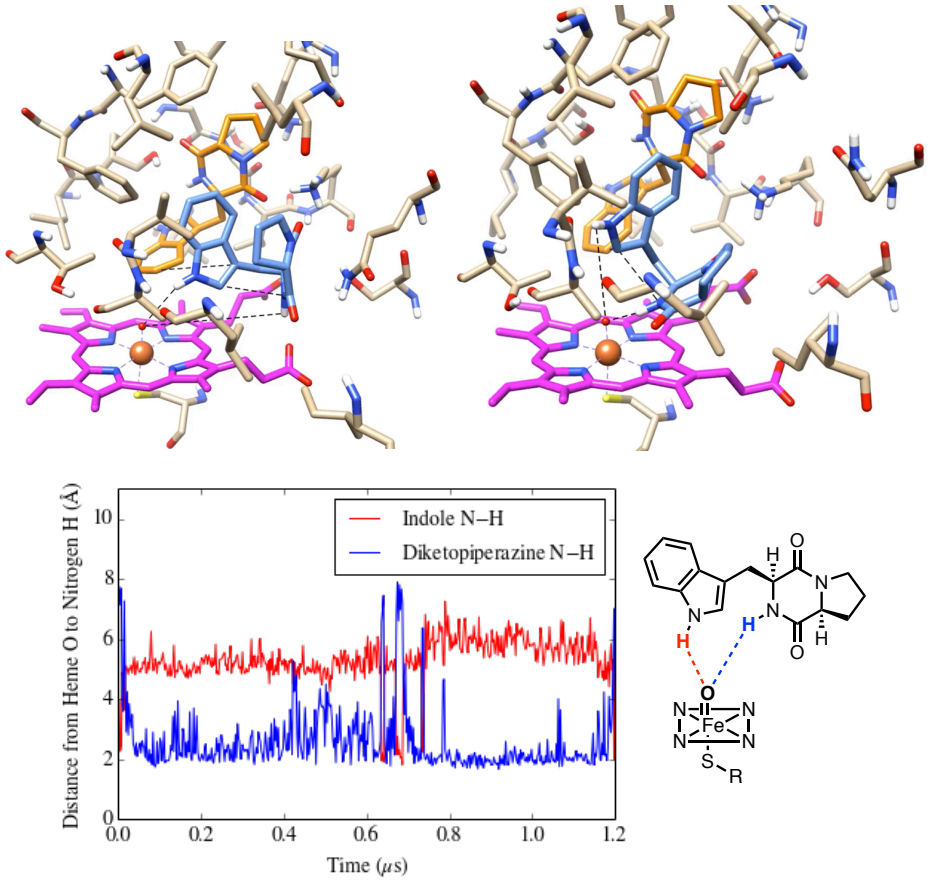


Figure 4-12. Top left: Snapshot of MD simulation at initial conformation, Top right, Snapshot of MD simulation at 300 microseconds. Bottom: Graphical representation of bond distances throughout MD simulation.

Based on this simulation we postulated that there may be a second possible mechanism where the cyclization-dimerization cascade is initiated by H–atom abstraction from the amide N–H from the diketopiperazine core rather than the N–H of the tryptophan indole (Figure 4-13).

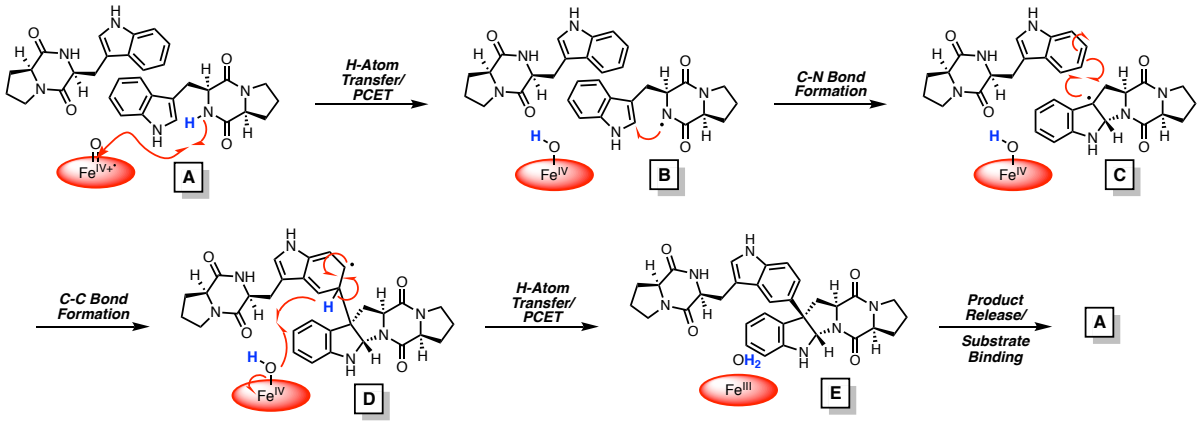


Figure 4-13. Alternative reaction mechanism based on MD simulations.

4.3.6 First Generation Mechanistic Probes

To evaluate the possibility for DKP N-H abstraction a probe molecule was synthesized wherein the indole N–H was "protected" as a simple methyl group. When subjected to standard reaction conditions, no dimer was observed however a single product was formed exclusively whose mass represented a hydroxylated product (Figure 4-14).

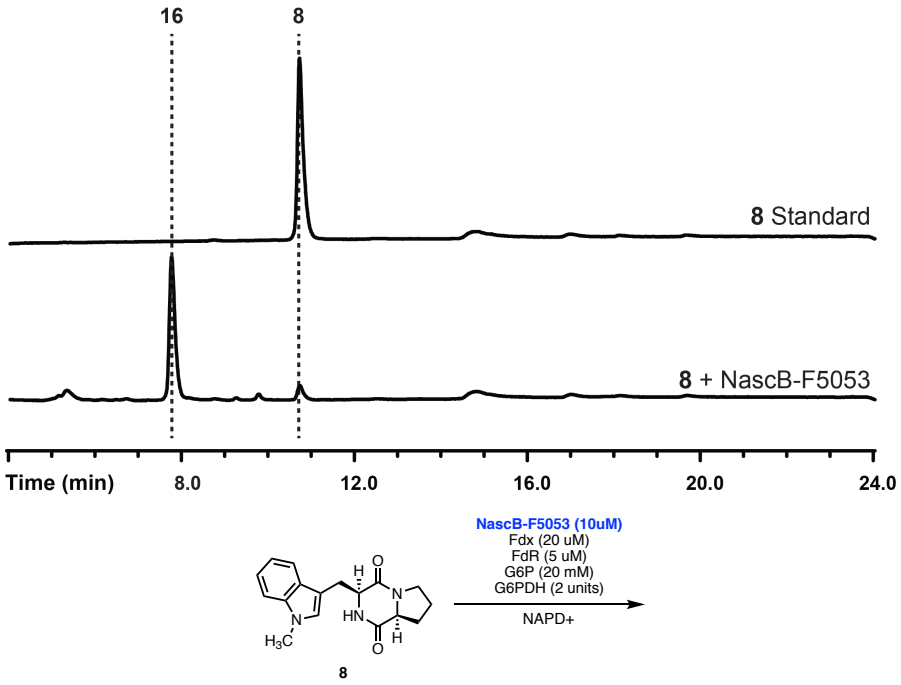


Figure 4-14. HPLC trace from reaction of first generation mechanistic probe, **8**, with NascB-F5053.

We hypothesized this product may be generated via two different paths, either hydroxylation via radical rebound of a radical intermediate, or alternatively via hydration of an intermediate carbocation generated from single electron oxidation of key radical intermediate (Figure 4-15).

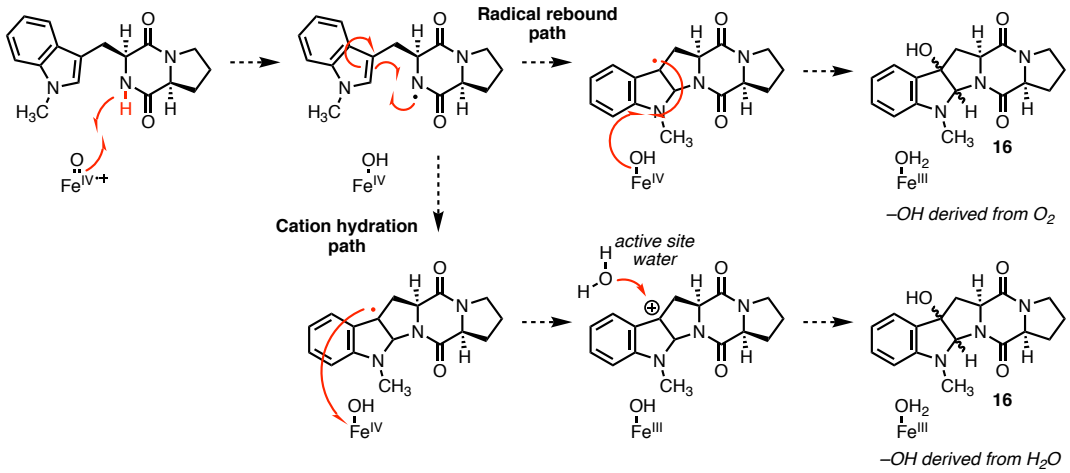


Figure 4-15. Possible mechanisms leading to hydroxylated shunt product **16**.

Close examination of the active site revealed a residue outside of 5 Å, Glu317, which forms an ordered hydrogen-bonding network with the indole N–H of the donor diketopiperazine. Using PROPKA,[13] E317 was shown to be 100% buried with a calculated pKa = 6.61, and under reaction conditions (pH=8.0) may exist as partially deprotonated.

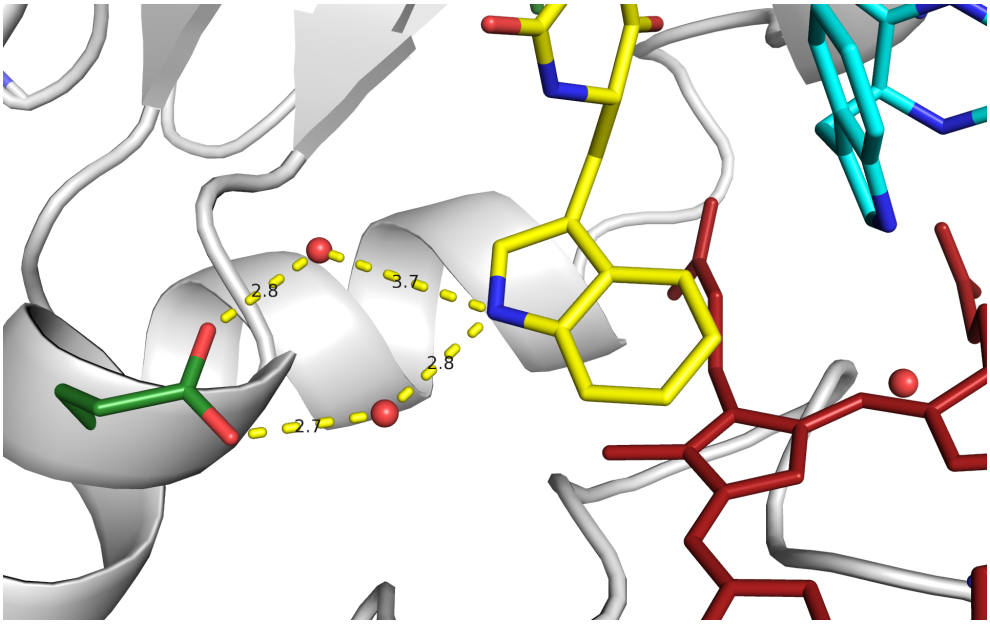


Figure 4-16. Crystal structure highlighting E317 interaction with donor monomer **3**.

Given this interaction, we hypothesized that E317 may be requisite for catalysis. If charged, E317A and the ordered waters may be partially deprotonating the indole N–H of the donor DKP, increasing the electron density of this monomer, activating the indole as a nucleophile in a Friedel-Crafts-type mechanism (Figure 4-17). To determine the role E317 may play in the reaction mechanism the E317A variant was generated and it's reactivity profiled with native substrate **3** and mechanistic probe **8**.

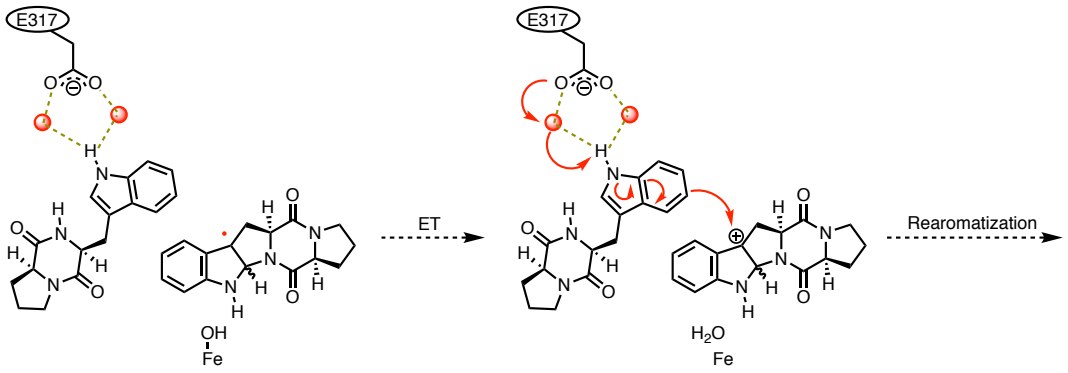


Figure 4-17. E317 may be partially deprotonating indole N–H from donor monomer, increasing electron density of indole and activating it for nucleophilic attack.

The E317A variant is still catalytically competent for both the native homodimerization reaction as well as generating the hydroxylated shunt from mechanistic probe **8** (Figure 18). While E317 may not play a direct role in catalysis, we sought to design a set of two, second-generation probes for heterodimerization. Probing the reaction mechanism for an unsymmetrical dimerization using two units of the same substrate (homodimerization) does not allow steric or electronic differentiation of the electrophile and/or the nucleophile as the same molecule plays both roles. In a selective heterodimerization, where one monomer always assumes the role of nucleophile and the other electrophile, the option to differentiate these substrates emerges and enables studying the reaction mechanism through traditional techniques for radical substitution or electrophilic aromatic substitution such as a Hammett plot.[14]

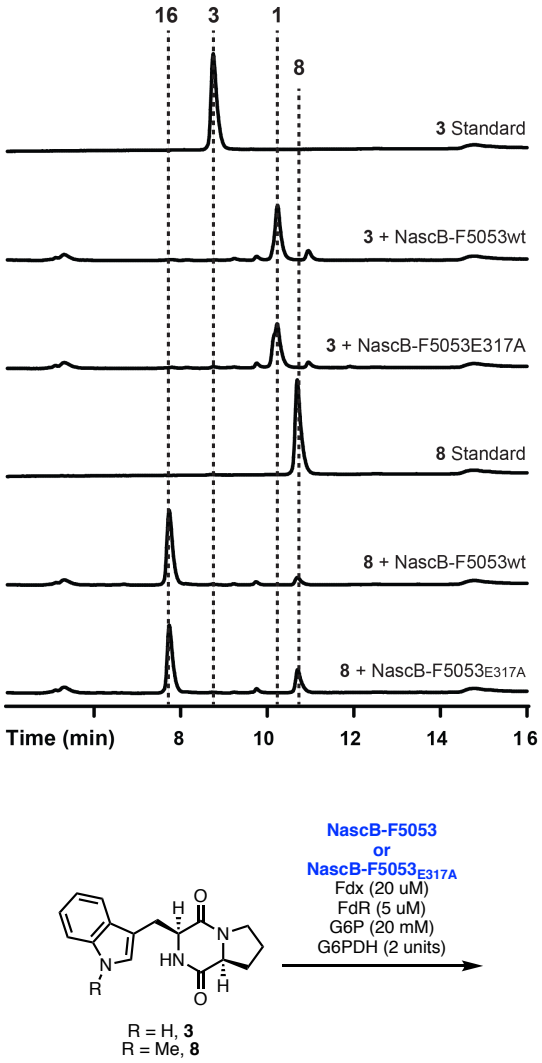


Figure 4-18. HPLC trace showing reaction with wt v.s. E317A with native substrate **3** and probe molecule **8**.

4.3.7 Second Generation Mechanistic Probes

In the reaction of **3** with **6** with NascB-F5053, a single heterodimeric product is formed and the homodimerization of **3** is completely suppressed.

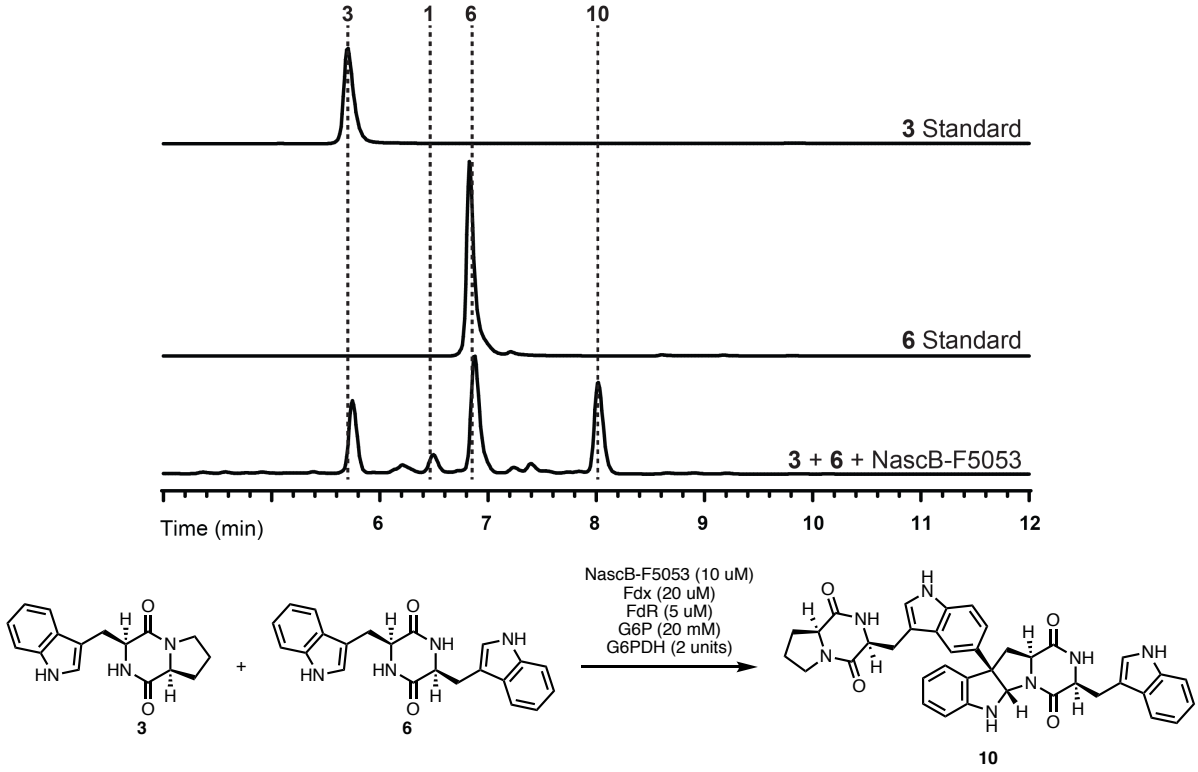


Figure 4-19. HPLC trace of heterodimerization of **3** and **6** exclusively forming a single product, **10**.

We hypothesized that we could use derivatives of these two monomers to evaluate the reaction mechanism. Based on the structure, we hypothesized that Q75 and E76 would preclude **6** from binding in the donor pocket, and would therefore occupy the acceptor pocket. To validate this hypothesis and ensure that each substrate can only occupy a single binding pocket without interchange, we crystallized NascB-F5053 with equivalent quantities of **3** and **6** and found that **3** seems to exclusively occupy the donor pocket and **6** the acceptor pocket (Figure 4-19).

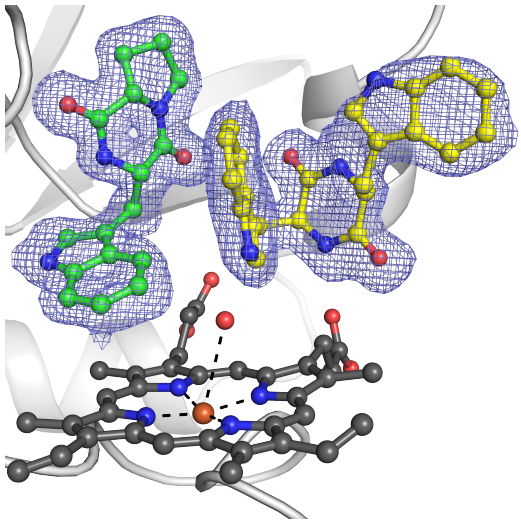


Figure 4-20. Active site of NascB-F5053 crystallized in complex with **3** (green) and **6** (yellow) with electron density (F_0-F_s) shown in blue mesh.

With this new crystal structure validating our binding hypothesis, we constructed second generation probe **9**, which based on the NascB heterocomplex of **3** and **6**, should exclusively occupy the acceptor site and should not be influenced by other donor residues such as E317. When subjected to unoptimized reaction conditions with an equimolar quantity of **3** to occupy the donor site, no dimer was observed and again a hydroxylated shunt product was observed albeit in low conversion. This result represents a proof of concept in the ability to separately interrogate the donor and acceptor pockets with second generation probes, and further functionalized probes are currently being developed concomitant with mutagenic approaches to deciphering the mechanism of NascB-F5053.

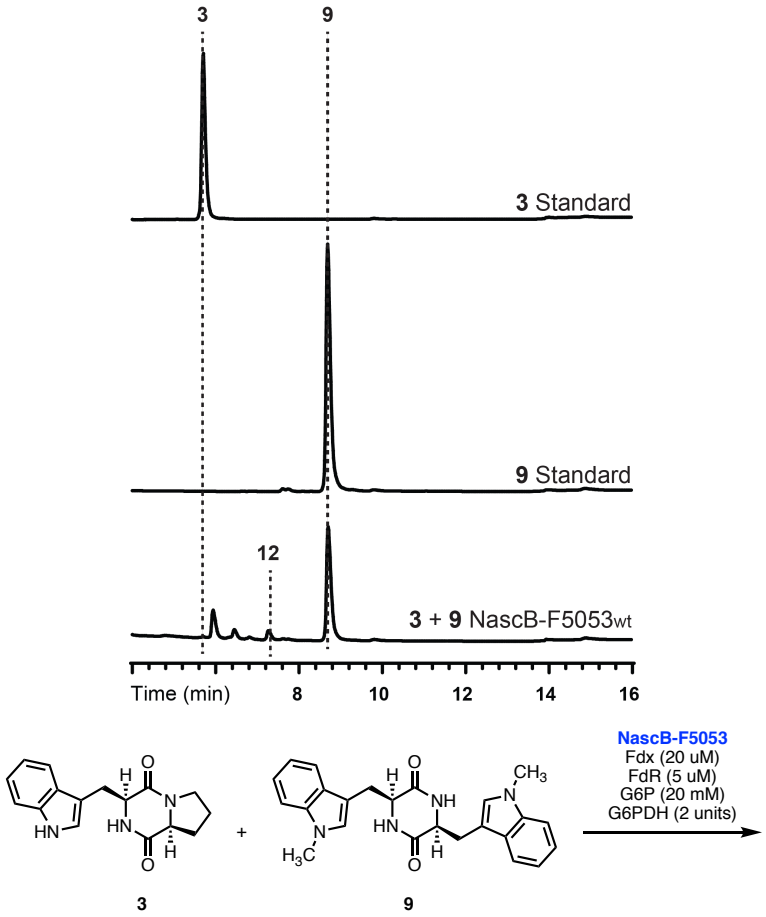


Figure 4-21. HPLC trace for reaction of second-generation probe (**9**) and **3** generates hydroxylated product.

4.3.8. Mutagenesis of Active Site Residues

To gain further insight into residues that may play a role in binding, catalytic activity, and substrate and product selectivity, alanine scanning mutagenesis of NascB-F5053 active site residues within 5Å of either substrate was performed (Figure 4-5). Thr244 was also selected for mutagenesis as this is the conserved proton donor residue required for proton transfer and generation of compound I.[15] NascB-F5053 variants Q68A, Q75A, and K292A could not be isolated, as when expressed under standard conditions overexpressed cytochrome was insoluble.

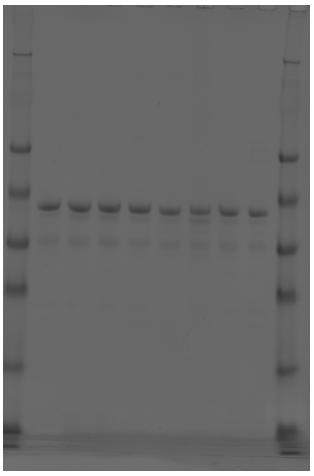


Figure 4-22. SDS-PAGE gel of purified fractions from NascB-F5053 variants, L to R: SeeBlue2 Ladder, NascB-F5053$_{E76A}$, NascB-F5053$_{Q75AE76A}$, NascB-F5053$_{T244A}$, NascB-F5053$_{S287A}$, NascB-F5053$_{E317A}$, NascB-F5053$_{F390A}$, NascB-F5053$_{F391A}$, NascB-F5053$_{F90AF391A,}$ SeeBlue2 Ladder (total protein concentrations all normalized 10 ug/well).

All soluble variants were subjected to standard reaction conditions with native substrate **3** (Figure 4-22). All variants harboring single residue mutations were still capable of producing both (–)-naseseazine C (**1**), as well as (–)-aspergilazine (**2**) (not easily visible in HPLC, but detectable in by MS) albeit in lower yields (Figure 22). However, the variant containing both F390A and F391A mutations, completely lost any dimerization capacity, and binding of 3 to the active site was no longer detectable by UV-Vis spectroscopy. Notably, the T244A variant did not lose its enzymatic activity, a rare trait in cytochromes P450, and may suggest that the substrate may donate a proton in the generation of compound I, as seen in EryF,[15] or NascB-F5053 may have a redundant residue capable of donating a proton.

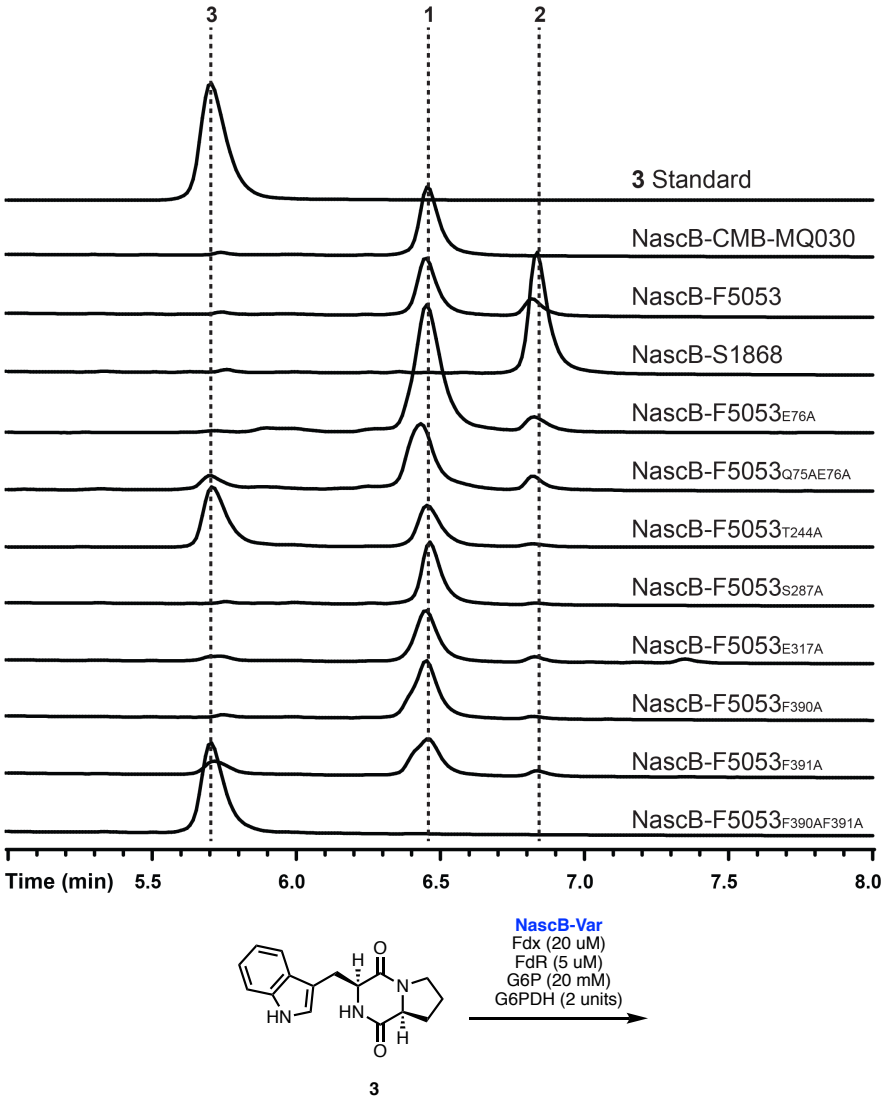


Figure 4-23. HPLC traces of native substrate **3** with NascB-F5053wt and soluble active site variants.

As previously noted NascB-F5053 is able to catalyze heterodimerizations of large substrates such as **6** with native substrate **3**, while NascB-CMB-MQ030 cannot perform this reaction, or does so in very low conversions (Figure 4-23). Sequence alignment of these two enzymes reveals that between the two enzymes only a single active site residue differs, where Ser287 in NascB-F5053 exists as Ala287 in NascB-CMB-MQ030.

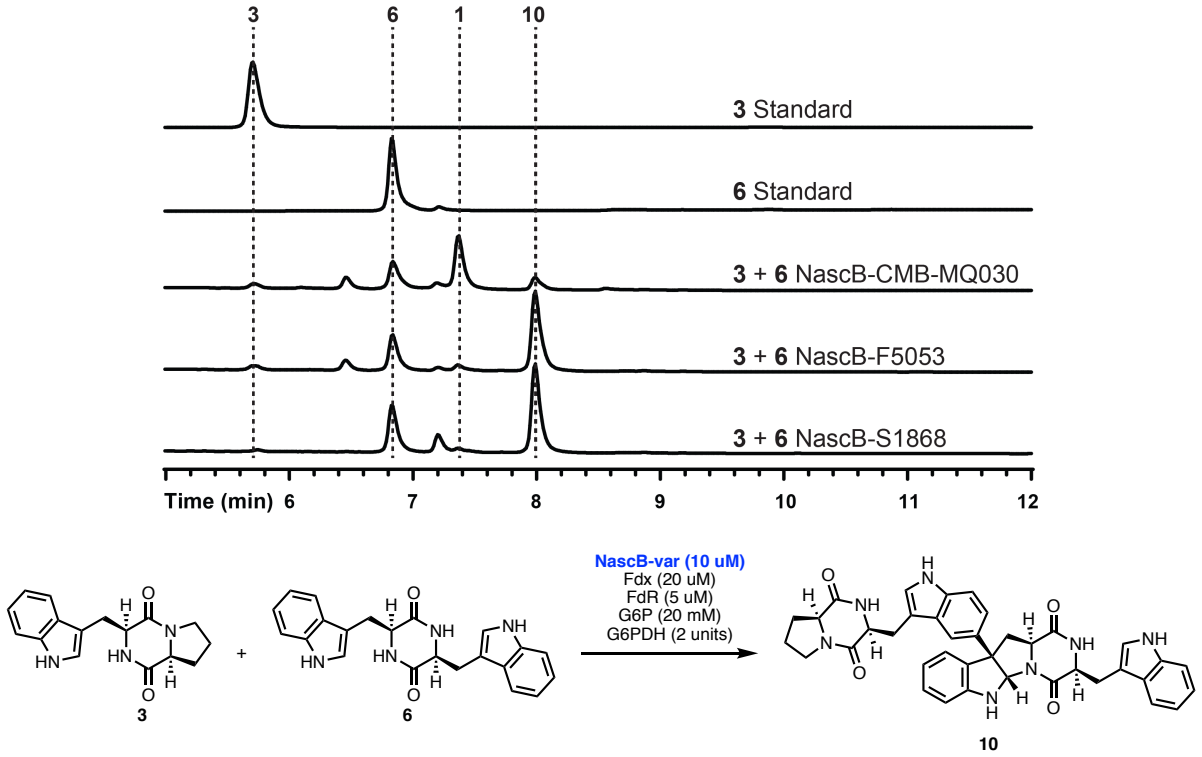


Figure 4-24. HPLC traces of reaction of **3** and **6** with Nasc-CMB-MQ030, NascB-F5053, and NascB-S1868.

To determine if Ser287 or another residue contribute to the expanded reactivity of NascB-F5053, our panel of variants were assayed for their ability to catalyze the heterodimerization of **3** and **6** (Figure 4-25). In contrast to the native reaction, all variants with the exception of NascB-F5053$_{Q75AE76A}$ saw a dramatic decrease in conversion of heterodimer **10**, and NascB-F5053$_{F390AF391A}$ completely lost its enzymatic activity.

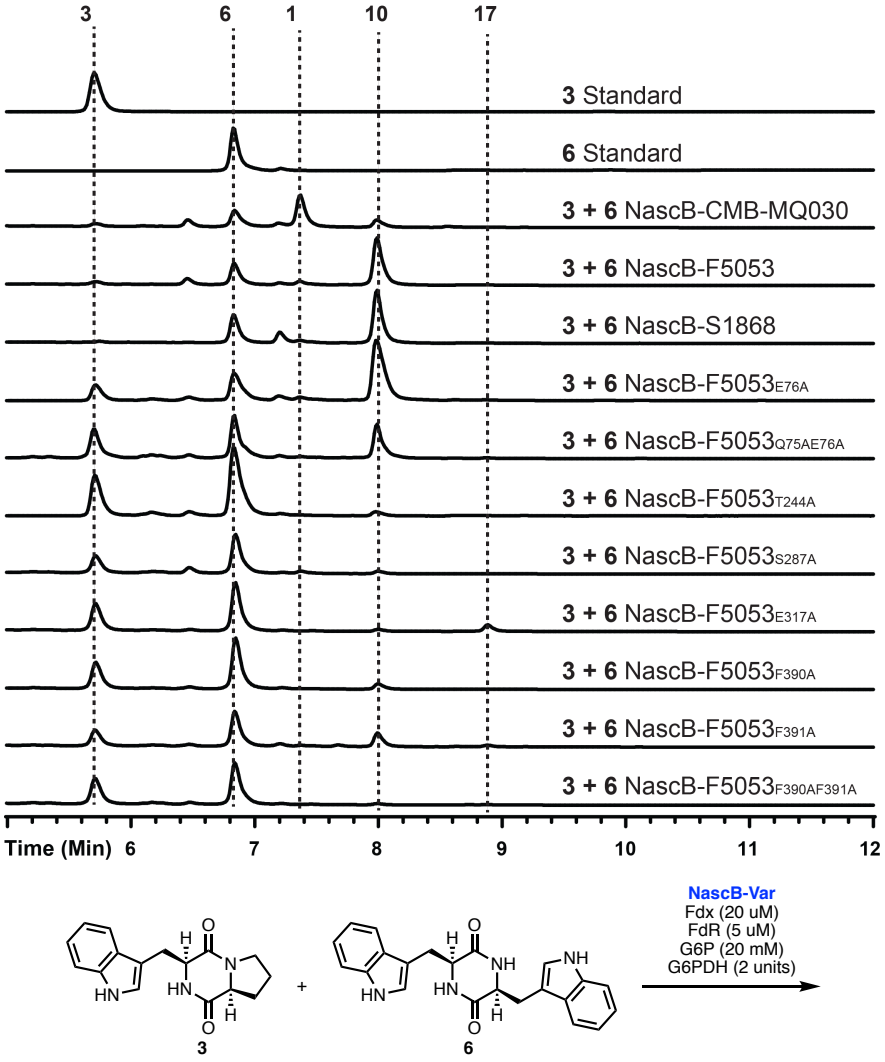


Figure 4-25. HPLC traces of reaction of **3** and **6** with NascB-F5053wt and soluble active site variants.

All reactions were futher analyzed by mass spectrometry, and the NascB-F5053$_{E317A}$ variant generated a second product whose mass indicated a heterodimeric product. This reaction was performed with a higher loading of biocatalyst and the generation of this new product was validated by mass spectrometry. NascB-F5053$_{E317A}$ can be expressed and purified at relatively high titer (5-8mg/L) and as such the reaction was scaled up *in vitro* (10 mg of **3** and 10 mg of **6)** and this product was purified by prep-HPLC to yield 3 mg, and the structure of this compound is being validated through rigourous NMR analysis.

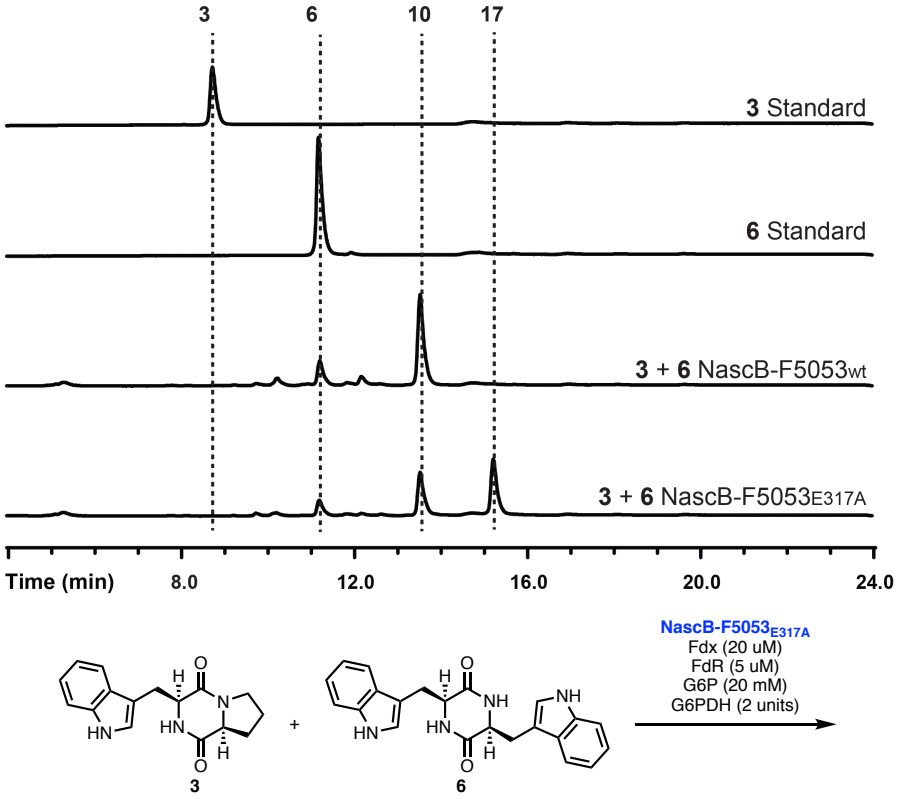


Figure 4-26. HPLC trace of reactions of **3** and **6** with NascB-F5053wt and NascB-F5053$_{E317A}$.

4.4 Discussion

The structure of NascB-F5053 was solved as a complex with two units of its native substrate **3**. Based on this structure we proposed a general mechanism for the dimerization of **3**, which we probed through synthesis of small molecule probes, computational simulations, and mutagenesis. Mutagenesis of active site residues within 5 Å of either substrate did not alter the reactivity profile with regard to the stereochemical or regiochemical outcome of the reaction, however a residue hypothesized to be requisite for catalysis Glu317 was exchanged for Ala, a new product was formed in the heterodimerization reaction of **3** and **6**. Work is currently ongoing in deciphering residues that influence the selectivity of diketopiperazine dimerization.

4.5 References

1. Barrow, C. J.; Sedlock, D. M., *J. Nat. Prod.* **1994,** *57* (9), 1239-44.
2. Barrow, C. J.; Cai, P.; Snyder, J. K.; Sedlock, D. M.; Sun, H. H.; Cooper, R., *Journal of Organic Chemistry* **1993,** *58* (22), 6016-6021.
3. Popp, J. L.; Musza, L. L.; Barrow, C. J.; Rudewicz, P. J.; Houck, D. R., *J Antibiot* **1994,** *47* (8), C5-C5.
4. Sedlock, D. M.; Barrow, C. J.; Brownell, J. E.; Hong, A.; Gillum, A. M.; Houck, D. R., *J Antibiot* **1994,** *47* (4), 391-398.
5. Popp, J. L.; Musza, L. L.; Barrow, C. J.; Rudewicz, P. J.; Houck, D. R., *J Antibiot* **1994,** *47* (4), 411-419.
6. Oleynek, J. J.; Sedlock, D. M.; Barrow, C. J.; Appell, K. C.; Casiano, F.; Haycock, D.; Ward, S. J.; Kaplita, P.; Gillum, A. M., *J Antibiot* **1994,** *47* (4), 399-410.
7. Saruwatari, T.; Yagishita, F.; Mino, T.; Noguchi, H.; Hotta, K.; Watanabe, K., *Chembiochem* **2014,** *15* (5), 656-9.
8. Holm, L.; Rosenstrom, P., *Nucleic Acids Research* **2010,** *38*, W545-W549.
9. Wang, Y.; Chen, H.; Makino, M.; Shiro, Y.; Nagano, S.; Asamizu, S.; Onaka, H.; Shaik, S., *J. Am. Chem. Soc.* **2009,** *131* (19), 6748-62.
10. Zhao, B.; Lamb, D. C.; Lei, L.; Kelly, S. L.; Yuan, H.; Hachey, D. L.; Waterman, M. R., *Biochemistry* **2007,** *46* (30), 8725-33.
11. Zhao, B.; Bellamine, A.; Lei, L.; Waterman, M. R., *Arch Biochem Biophys* **2012,** *518* (2), 127-32.
12. Dumas, V. G.; Defelipe, L. A.; Petruk, A. A.; Turjanski, A. G.; Marti, M. A., *Proteins* **2014,** *82* (6), 1004-21.
13. Rostkowski, M.; Olsson, M. H.; Sondergaard, C. R.; Jensen, J. H., *BMC Struct Biol* **2011,** *11*, 6.
14. Luk, L. Y.; Bunn, S.; Liscombe, D. K.; Facchini, P. J.; Tanner, M. E., *Biochemistry* **2007,** *46* (35), 10153-61.
15. Xiang, H.; Tschirret-Guth, R. A.; Ortiz De Montellano, P. R., *J Biol Chem* **2000,** *275* (46), 35999-6006.

4.5 Experimental

Supplementary Table 4-1. Mutagenic primers used to generate NascB-F5053 variants

Primer #	Gene	Primer 5'→3'
1	NascB-F5053$_{Q68A}$	GGACTCGCGGTTCAGCTCGGAGGCCGCAGCCGCGTC CGGCGCGCCCCGCCAGGAAC
2	NascB-F5053$_{Q75A}$	CACAGGCGTCCGGCGCGCCCCGCGCCGAACCGGTCGAGCT GCGGGCGCCCGGCACCCGG
3	NascB-F5053$_{E76A}$	ACAGGCGTCCGGCGCGCCCCGCCAGGCCCCGGTCGA GCTGCGGGCGCCCGGCACCCGG
4	NascB-F5053$_{Q75AE76A}$	CACAGGCGTCCGGCGCGCCCCGCGCCGCCCCGGTCG AGCTGCGGGCGCCCGGCACCCG
5	NascB-F5053$_{T244A}$	CTCGCTCCTCGTCGTCGCCGGCTTCCCGGCCTCGTCG GGATTCCTGTGCGGCGCGCTGCT
6	NascB-F5053$_{S287A}$	CCGTGGAGGAACTCCTGCGGTACACGCCCCTCGCCAC CGGCTCGGTCAAGCGGATGGCCA
7	NascB-F5053$_{K292A}$	CTGCGGTACACGCCCCTCTCGACCGGCTCGGTCGCC CGGATGGCCACCGAGGACCTGGAG
8	NascB-F5053$_{E317A}$	AAGGCCGGCGAGGTGGTGATGGTCTCCCTCGCCGCC GTCAACCACGACCCGGACGCCTTC
9	NascB-F5053$_{F390A}$	CCCGAGGAGATCAGCTGGCACGAAGGGCTCGCCTTC CGCCGCCCGCGGGCGATCCCCGCC
10	NascB-F5053$_{F391A}$	CCCGAGGAGATCAGCTGGCACGAAGGGCTCTTCGCC CGCCGCCCGCGGGCGATCCCCGCC
11	NascB-F5053$_{F390AF391A}$	CCCGAGGAGATCAGCTGGCACGAAGGGCTCGCCGCC CGCCGCCCGCGGGCGATCCCCGCC

General procedures for overexpression, purification and spectral characterization of cytochrome P450s

The expression of pET28b-NznB+09 was performed by transforming the plasmid into the competent cell *E. coli* strain C41 (DE3) and selected on Luria-Bertani (LB) medium plates containing100 μg/mL kanamycin. A single colony was grown overnight in

LB broth containing the same concentration of the antibiotics. The main culture was prepared by inoculating 1% of each overnight culture into a 2 L baffled flask containing 500 ml of Terrific Broth (TB) containing 100 µg/mL kanamycin. The cultures were incubated at 37°C and 250 rpm for 3-4 h unless the absorbance A_{590nm} = 1. The expression of NznB was induced by the addition of 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 0.5 mM delta-5-aminolevulinic acid (w/v) as well as 1mM thiamine. The cultures were incubated further for 30 h, 18°C at 160 rpm. Cells were harvested by centrifugation at 5000×*g* at 4°C for 10 min and the cell pellet was stored at –80°C until purification of the protein.

The cell pellet was resuspended in 5% culture volume of lysis buffer (50 mM kPi, pH 8.0 containing 200 mM $(NH_4)_2SO_4$, 20% glycerol, 0.5 mM EDTA, 10 mM beta-mercaptoethanol and 1mM phenylmethane sulfonyl fluoride (PMSF) and disrupted by sonication. The cell lysate was centrifugated (35,000 rpm for 35 min) and filtered (0.2 µM Millipore filter). The soluble His_8-tagged NznB was purified by affinity chromatography using Ni-NTA (Qiagen) column and the collected fractions were analyzed by SDS-PAGE. The suitable red fractions were pooled and dialyzed at 4°C three times using 50 mM Tris, pH 7.4 containing 50 mM kPi, pH 8.0 containing 200 mM $(NH_4)_2SO_4$, 20% glycerol, 0.5 mM EDTA, and 1 mM dithiothreitol (DTT) against a total of 6 L buffer.

UV–Vis spectra for the purified cytochomes were recorded at room temperature in buffer (50 mM Tris-HCl buffer, pH 7.4 containing 5% glycerol) was used for the spectral measurements of the oxidized and reduced form. NasB was reduced by the addition of a small amount of sodium dithionite. The concentration of the P450s was estimated by CO-difference spectra, assuming ε(459–490) = 91 mM−1 cm−1 according to the method of Omura and Sato.

Process of crystallization

Crystallization of NascB-F5053.

Single, diffraction quality crystals of the *NascB-F5053*complexes were grown by sitting drop vapor diffusion at 20 °C by mixing 2 µL of 8 mg/mL protein containing 1 mM DKP substrate and 0.2% DMSO with 2 µL of a well solution containing 23% PEG 3350, 100 mM DL-malic acid, 2.5% ethylene glycol. Equimolar concentrations of substrate (0.5 mM) were used for the *NascB-F5053*·**4**·**6** complex. Sitting droplets were nucleated after 4 h from an earlier spontaneous crystallization using a cat whisker. Single crystals grew after 48 hours. 8 µL of a cryoprotecting solution containing 10 mM Tris pH 7.5, 2 % glycerol, 23% PEG 3350, 100 mM DL-malic acid, 18% ethylene glycol, 1 mM DKP substrate, 0.2% DMSO was added directly to the sitting drops and the crystals were harvested using nylon loops and vitrified by rapid plunging into liquid nitrogen. *NascB-F5053* crystallized in space group *P* 1with unit cell dimensions of *a* = 56.1 Å, b = 56.3 Å, *c* = 58.8 Å, α = 91.1°, β = 92.9°, γ = 100.1° and two chains in the asymmetric unit.

Data collection and processing.

X-ray data were collected at 100 K on beamline 23ID-B at the General Medical Sciences and Cancer Institutes Structural Biology Facility at the Advanced Photon Source in Argonne, IL, USA. Diffraction data were integrated and scaled using XDS.[1] Data collection statistics are given in Supplementary Table 4-2.

Molecular replacement, model building and refinement.

The structure of NascB-F5053 was solved using Phenix MR-Rosetta.[2] This resulted in an initial model that could be extended by alternating cycles of manual building in *Coot*[3] and least-squares refinement with Phenix.[4] Final models were validated using MolProbity.[5]

Supplementary Table 4-2. Data Collection Statistics

	cWP	**cWP_cWW**	**cWV**
Wavelength (Å)	1.03	1.03	1.03
Resolution range (Å)	40.83 - 1.49	41.22 - 1.48	34.37 - 1.349

	(1.54 - 1.49)	(1.53 - 1.48)	(1.397 - 1.349)
Space group	P 1	P 1	P 2_1
Unit cell	56.12, 56.28, 58.78 (Å) 91.13, 92.93, 100.10 (°)	56.07, 56.13, 58.48 (Å) 91.36, 93.04, 101.10 (°)	55.80, 56.28, 56.09 (Å) 90, 101.74, 90 (°)
Total reflections	382537 (31077)	324387 (4806)	496898 (42769)
Unique reflections	109037 (10046)	94423 (2902)	74183 (7153)
Multiplicity	3.5 (3.1)	3.4 (1.7)	6.7 (6.0)
Completeness (%)	94.48 (87.34)	80.96 (24.95)	99.05 (95.79)
Mean I/sigma(I)	16.78 (2.39)	20.64 (1.81)	16.69 (1.51)
Wilson B-factor ($Å^2$)	19.96	22.24	18.17
R-merge	0.039 (0.361)	0.030 (0.287)	0.049 (0.901)
R-meas	0.046 (0.439)	0.036 (0.398)	0.053 (0.988)
CC1/2	0.998 (0.85)	0.999 (0.851)	0.999 (0.665)
Reflections used in refinement	109025 (10043)	94401 (2902)	74175 (7151)
R-work	0.168 (0.287)	0.192 (0.333)	0.197 (0.307)
R-free	0.211 (0.340)	0.234 (0.410)	0.229 (0.386)
Number of atoms (total):	6989	6995	3550
macromolecules	6108	6079	3067
ligands	190	196	85
solvent	691	720	398
Protein residues	798	793	396
RMS(bonds)	0.005	0.007	0.007
RMS(angles)	0.91	1.23	1.26
Ramachandran	97.21	97.56	97.68

favored (%)			
Ramachandran allowed (%)	2.79	2.44	2.32
Ramachandran outliers (%)	0.00	0.00	0.00
Average B-factor ($Å^2$)	27.51	28.55	25.32
macromolecules	26.79	28.07	24.59
ligands	19.50	22.09	17.77
solvent	36.01	34.38	32.57

MD Simulations

Molecular Dynamics simulations were performed using the GPU code (*pmemd*)[47] of the AMBER 16 package[48]. Parameters for intermediates and substrates were generated within the *antechamber* module using the general AMBER force field (*gaff*)[49], with partial charges set to fit the electrostatic potential generated at the HF/6-31G(d) level by the RESP model[50]. The charges were calculated according to the Merz–Singh– Kollman scheme[51-52] using the Gaussian 09 package[36]. Each protein was immersed in a pre-equilibrated truncated cuboid box with a 10 Å buffer of TIP3P[53] water molecules using the *leap* module, resulting in the addition of around 15,000 solvent molecules. The systems were neutralized by addition of explicit counter ions (Na+ and Cl−). All subsequent calculations were done using the widely tested Stony Brook modification of the Amber14 force field (*ff14sb*)[54]. A two-stage geometry optimization approach was performed. The first stage minimizes the positions of solvent molecules and ions imposing positional restraints on the solute by a harmonic potential with a force constant of 500 $kcal \cdot mol^{-1} \cdot Å^{-2}$ and the second stage minimizes all the atoms in the simulation cell. The systems were gently heated using six 50 ps steps, incrementing the temperature by 50 K for each step (0–300 K) under constant-volume and periodic-boundary conditions. Water molecules were treated with the SHAKE algorithm such that the angle between the hydrogen atoms was kept fixed. Long- range electrostatic effects

were modelled using the particle-mesh-Ewald method[55]. An 8 Å cutoff was applied to Lennard–Jones and electrostatic interactions. Harmonic restraints of 10 kcal·mol–1 were applied to the solute and the Langevin 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. Each system was then equilibrated without restrains for 2 ns with a 2 fs time step at a constant pressure of 1 atm and temperature of 300 K. After the systems were equilibrated in the NPT ensemble, subsequent MD simulations were performed for an additional 500 ns under an NVT ensemble and periodic-boundary conditions.

Reaction conditions for analytical scale P450 reactions

The conversion of DKPs by P450s were carried out with the heterologous redox electron partners Fdx and FdR from *S. olerecea*. The *in vitro* reaction mixture included 50 mM Tris-HCl, pH 7.4, 10 µM P450, 20 µM Fdx and 3 µM FdR at the end volume of 250 µL. Without exceeding 1% of the total reaction volume, 1mM of DKP dissolved in DMSO was added (in the case of heterodimerization reaction, both substrates were added to a final concentration of 1mM). The reaction was initiated by adding NADPH (1 mM). Two of the control reactions including all the contents except for NADPH or P450 were used as controls. The reaction was incubated at 30°C for 2 h agitating at 600 rpm in a thermoshaker (Multi-thermoshaker, Benchmark). Reactions were quenched by the addition of 3 volumes of methanol, and centrifuged at 17,000rpm and directly used for HPLC analysis. Reactions were resolved using a linear gradient of 15-75% acetonitrile: water (0.1% formic acid) over 16 min (2 mL/min flow rate, 40ºC) on a Phenomenex Luna 5µ C-18(2) 100A, 250 × 4.60mm 5 micron column.

HPLC-MS Data from first generation reaction probes

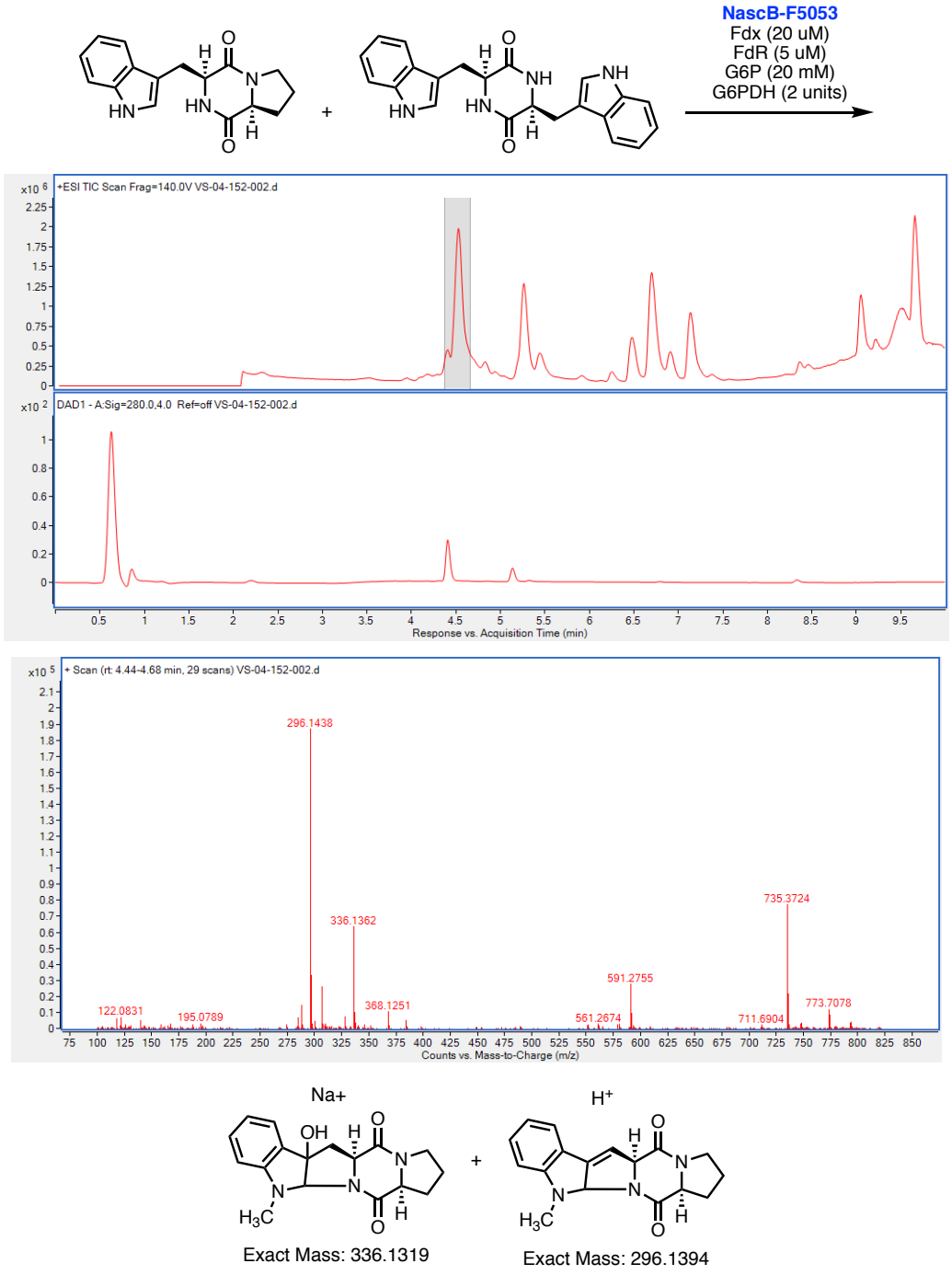
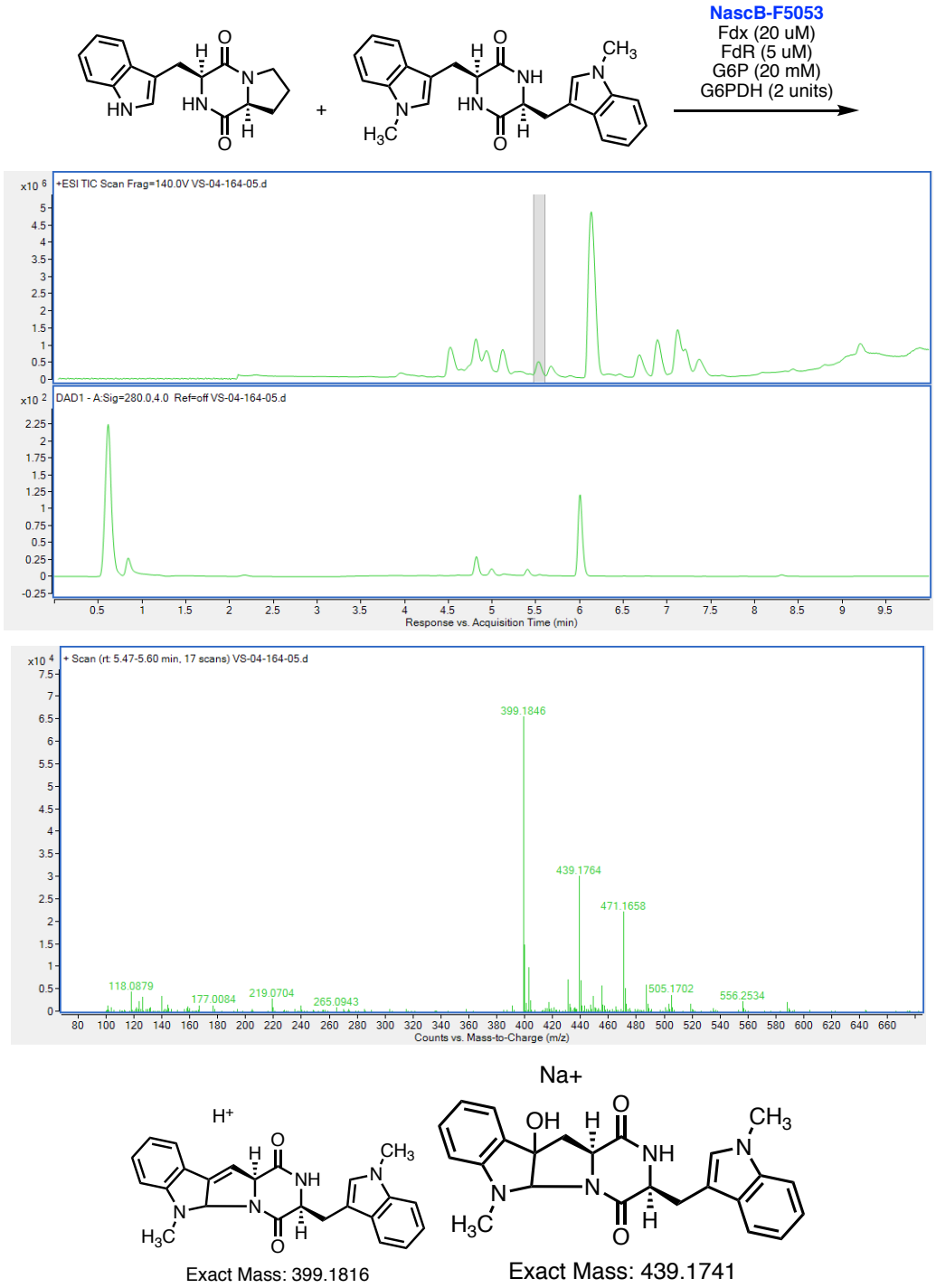


Figure S4-1. Mass Spectra from first generation probe reaction. Hydroxylated mass as well as dehydrated mass are observed.

HPLC-MS Data from second generation reaction probes



Supplemental Figure 4-2. Mass Spectra from second generation probe reaction. Hydroxylated mass as well as dehydrated mass are observed.

References

1. Kabsch, W., *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 133-44.
2. Terwilliger, T. C.; Dimaio, F.; Read, R. J.; Baker, D.; Bunkoczi, G.; Adams, P. D.; Grosse-Kunstleve, R. W.; Afonine, P. V.; Echols, N., *J Struct Funct Genomics* **2012,** *13* (2), 81-90.
3. Emsley, P.; Cowtan, K., *Acta Crystallogr D Biol Crystallogr* **2004,** *60* (Pt 12 Pt 1), 2126-32.
4. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 2), 213-21.
5. Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., *Acta Crystallogr D Biol Crystallogr* **2010,** *66* (Pt 1), 12-21.

Chapter 5
Summary and Future Directions

5.1 Summary

This thesis was focused mainly on the identification ad characterization of the wild-type activity cytochromes P450 involved in the bacterial biosynthesis of non-symmetrical diketopiperazine dimers as discussed in Chapters 2-4. In this chapter, I will briefly discuss the future direction for the diketopiperazine dimerizing cytochromes P450.

5.2 Future Directions

5.2.1 Full substrate scope analysis

Diketopiperazines are extremely synthetically accessible, and due to this accessibility make for great building blocks. In Chapter 3 we were able to rapidly synthesize a handful of monomers and screen them against our 4 cytochromes. However, we quickly ran into a problem analyzing these reactions, as the polarity of these starting materials and products are heavily overlapped and do not allow for quantification of product formed, and in many cases even identification of the number of products formed.

Our goal is to continue to identify more and more diketopiperazine dimerizing cytochromes P450 as well as generate variants with new selectivity that we can probe. However, as we continuously synthesize diketopiperazine monomers for dimerization assays, the number of reactions to be performed and analyzed rapidly grows well beyond the capabilities we have at present. For example, if we were to screen heterodimerization reactions with 50 diketopiperazine monomers, and run reactions where every diketopiperazine is dimerized with every other diketopiperazine, the

number of reactions to be run can be calculated using the binomial theorem (Equation 1) where n = the total number of diketopiperazines and k = the number of monomers in each reaction, which is 2.

$$Number\ of\ reactions\ to\ be\ run = \frac{n!}{(n-k)!\,k!} \qquad \text{(Equation 1)}$$

Solving this equation we see the number of reactions to be run to screen 50 diketopiperazine monomers is 1225 for a single cytochrome. This number however does not represent the number of possible products formed in these reactions. As we saw in many of the homo- and heterodimerization reactions multiple products are formed due to the non-symmetrical nature of the dimerization. Again for a given library of 50 diketopiperazines, if we assume that there are 4 products formed in each heterodimerization reaction of two DKP monomers, A and B (A-A, B-B, A-B, and B-A), the number of possible products becomes the square of the number of substrates and we have a possible of >2500 unique products. To resolve this issue, our goal is to develop a small volume and high-throughput approach to screening reactivity of cytochromes as well as development of analytical methods for separations of products. To enable this approach we are in preliminary discussions with High-Throughput Chemistry facilities at pharmaceutical companies to develop a parallel approach to sceening and product identification, and hopefully biological evaluation of the generated library of products as well.

5.2.2 Identification of enzymatic mechanism for dimerization selectivity

Research from Chapter 4 demonstrated that residues within 5 Å of either substrate in the active site are not the key gate-keepers for selectivity in the non-symmetrical dimerization of diketopiperazine substrates, although some variants were unable to be characterized due to instability which is currently still under optimization. Given the lack of dramatic changes in reactivity in initial alanine scanning mutagenesis, as well as the conformation of the substrates in the active site not representing the stereochemistry of the product, we hypothesize the active site of NascB-F5053 must undergo dramatic remodeling in order to allow for cyclization and dimerization

consistent with the stereochemistry of the product. This remodeling event may involve residues outside of this first interacting sphere (>5 Å). To uncover these residues, we are currently working towards attaining crystal structures of the other cytochromes from this study in complex with their substrates as well as products in hopes of capturing an alternative conformation, which may give clues to these gatekeeping residues. We are also working with our collaborators in the Houk lab, and performing MD simulations to try and identify these residues.

The NznB cytochrome generates isomeric dimers relative to NascB, and ideally we would be able to acquire a crystal structure of this enzyme to identify residues influencing its selectivity. However, the NznB cytochrome is relatively unstable, and upon cleavage of of the N-terminal 8X-His tag with TEV protease, the protein precipitates. Homologues of NznB from the strains reported in this study (NRRL-F5053 and NRRL-S1868) are currently under investigation, however their >99% identity may result in similar problems with solubility. Curiously, in examining our sequence similarity network, we identified another candidate cytochrome with extremely high sequence identity to NznB (Figure 5-1). This gene has been codon optimized and is currently being synthesized for it's evaluation as a putative DKP dimerase.

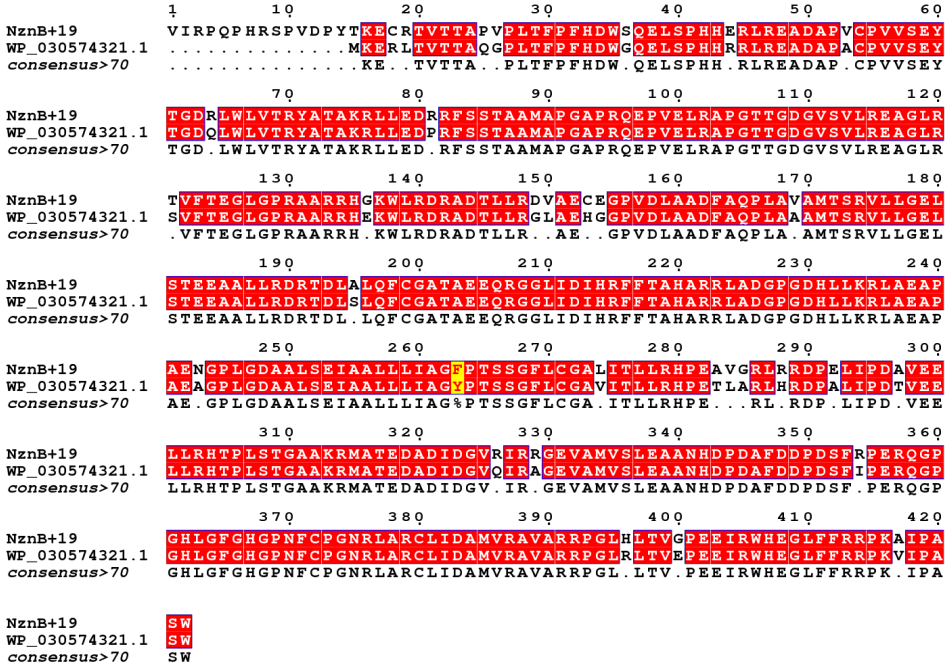


Figure 5-1. Sequence alignment of NznB with putative NznB homologue (WP_030574321.1) from *Streptomyces aureocirculatus*.

5.2.3 Recombination and Engineering of NascB

As referenced above, active site residues appear to play little to no role in selectivity of dimerization. This outcome is not completely unsurprising, as when a sequence alignment is performed, assuming active site residues are the same across homologues, there are minimal differences in active site residues between NascB-CMB-MQ030, NascB-5053, NascB-S1868.

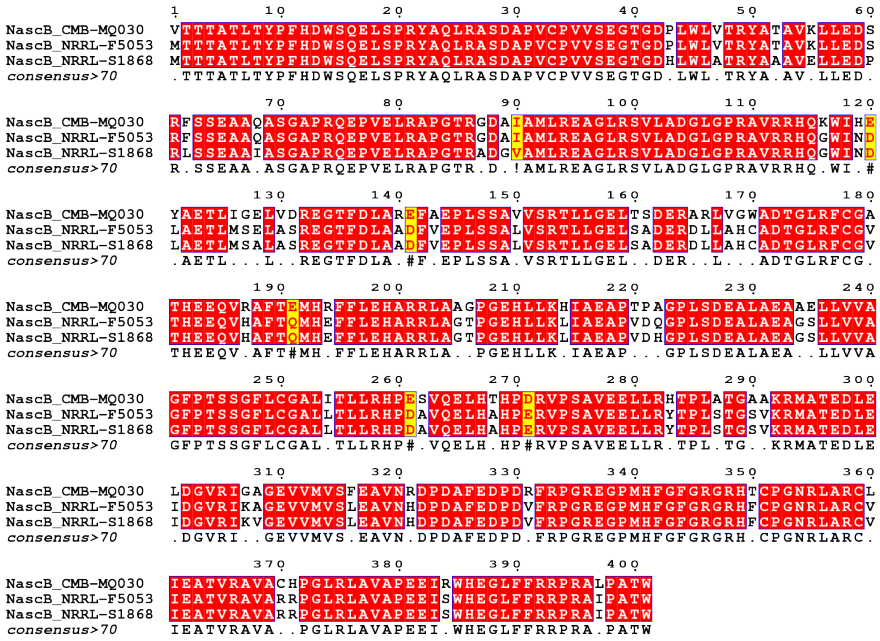


Figure 5-2. Sequence alignment of NascB homologues from NRRL-F5053 and NRRL-S1868.

We hypothesize that regions outside of the active site may be controlling selectivity of dimerization, and to identify these regions, homology models of NascB-CMB-MQ030 and NascB-S1868 were generated with SWISS-MODEL and regions of where one homologue had a unique mutation were highlighted (Figure 5-3). To identify how NascB-CMB-MQ030 can selectively generate (–)-naseseazine C, and how NascB-S1868 can exclusively generate (–)-aspergilazine, while NascB-F5053 generates a mixture of both, we propose to graft these regions of low-no conservation of NascB-S1868 and NascB-CMB-MQ030 onto NascB-F5053 and assay the products of these chimaeras. Along with using the homology model to identify key regions for chimaeragenesis, we also propose to use SCHEMA[1] and other computational tools for identifying key regions for recombination.

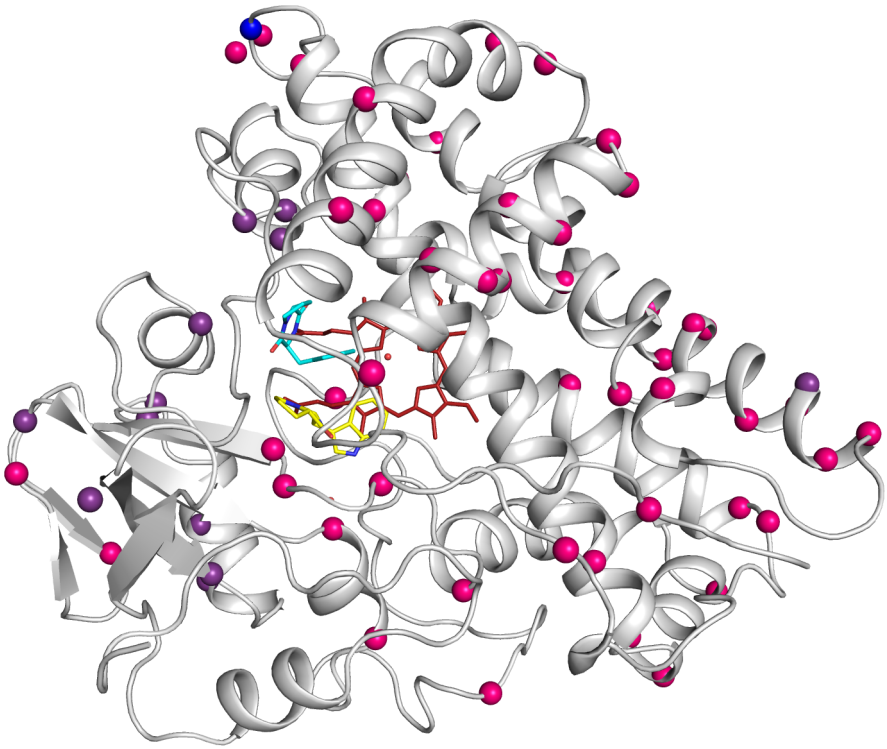


Figure 5-3. Unique variations from each NascB variant are highlighted as colored spheres on the NascB-F5053 structure. Blue = NascB-F5053, Purple = NascB-S1868, Pink = NascB-CMB-MQ030.

5.3.4 Directed Evolution of NascB-F5053

The original goal of this thesis was to not only discover the dimerizing cytochromes, but to engineer them into synthetically useful designer biocatalysts to access any connectivity and stereochemistry of dimer on demand in a catalyst controlled fashion such that a single monomer could be used to construct a variety of dimers. However, as previously mentioned, active site residues are not playing an obvious role in dimerization selectivity, and as such selecting residues for mutagenesis cannot easily be performed rationally. It is well precedented in directed evolution campaigns to engineer new selectivity into enzymes, that residues well outside of the active site are those that have undergone mutagenesis in variants with new selectivity. Given that our enzyme active site does not appear to be influencing selectivity anyway, we have begun investigation into development of a high-throughput platform for directed evolution of NascB-F5053.

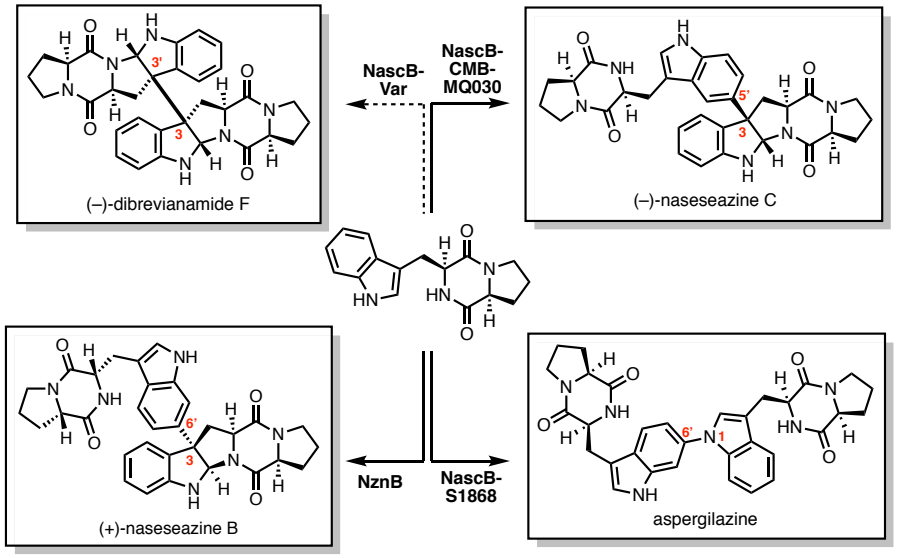


Figure 5-4. Proposed biocatalyst controlled dimerization of a single monomer to a variety of known and natural product dimers.

Indole fuctionalizing enzymes have been engineered with new selectivity via directed evolution, with a key example being C6-indole halogenase ,RebH. Lewis *et. al.* utilized directed evolution to engineer RebH with new selectivity by utilizing a deuterium incorporation approach allowing for rapid screening via mass spectroscopy alleviating the need for separation of similar compounds.[2-6]

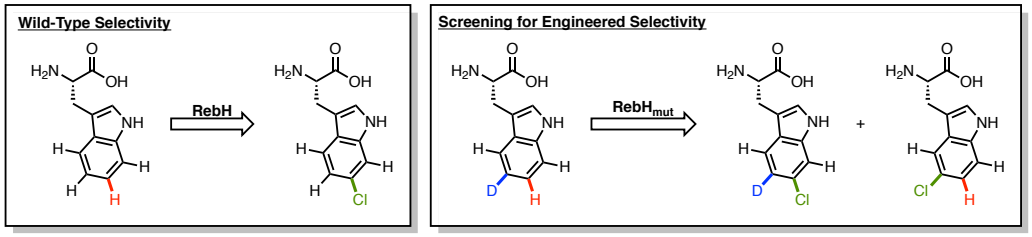


Figure 5-5. Lewis *et. al.* deuterium labeling strategy for directed evolution of RebH.

We considered using this approach, as we could incorporate deuterium into our diketopiperazines using similar methodology, however in examining some of the target products we saw that the protons that would be lost in the course of the reaction would be those exchangeable with solvent and therefore would not be detected/could not be screened for using this approach. Looking back at our initial data from Chapter 2, we saw that fluorination at the position of dimerization completely stopped reactivity, while fluorination *ortho*- to the position of dimerization had little to no effect on percent

conversion. As such we decided to adopt a fluorine incorporation strategy, wherein the position where the carbon-carbon bond would be natively formed would be blocked by a fluorine, and any variant that can form any other bond would generate a dimer (Figure 5-6).

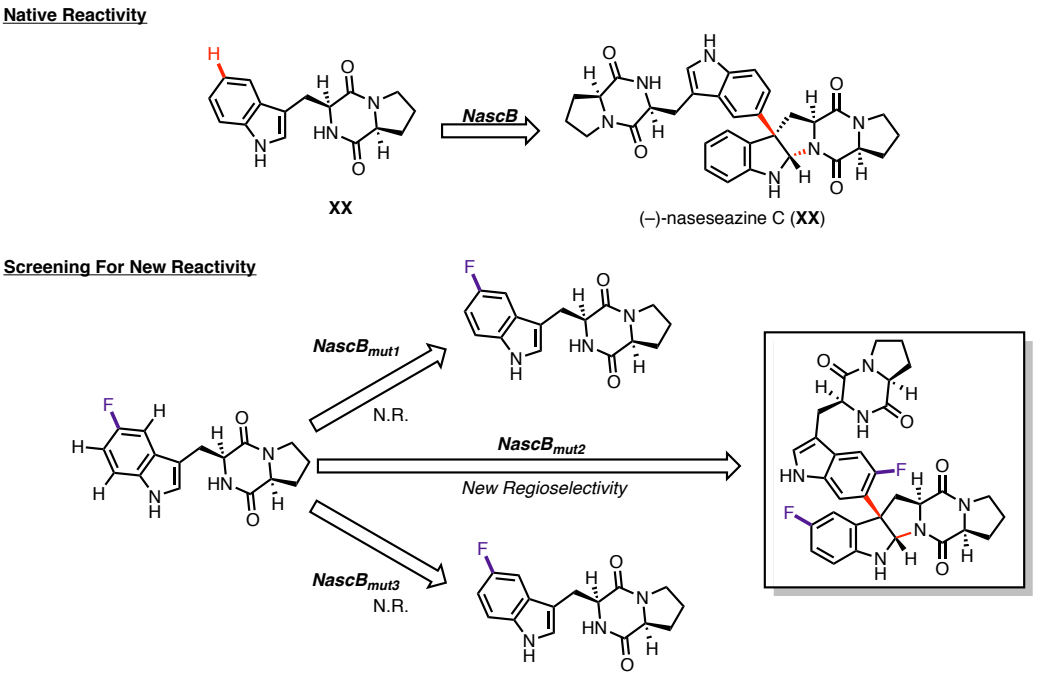


Figure 5-6. Strategy for fluorine incorporated directed evolution strategy for NascB-F5053.

This strategy would allow not only detection of variants with new selectivity, but also through polyfluorination of the arene ring we would be able to detect formation of dimers undetectable through deuterium incorporation such as C3-C3' linked symmetrical dimers (such as dibrevianamide F), as well as C3-N1' linked dimers (Figure 5-7). This also enables a more facile analysis, as rather than discerning a difference in a single AMU as you would for deuterium labeling, the readout for this assay would be binary; dimer or monomer.

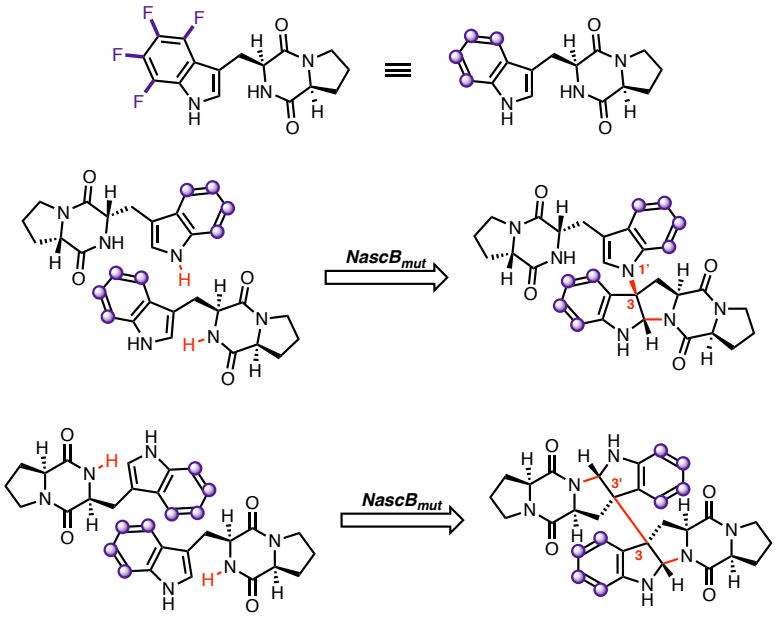


Figure 5-7. Polyfluorination enables detection of dimers whose protons lost in the reaction would be exchangeable with solvent.

References:

1. Endelman, J. B.; Silberg, J. J.; Wang, Z. G.; Arnold, F. H., *Protein Eng Des Sel* **2004,** *17* (7), 589-94.
2. Payne, J. T.; Andorfer, M. C.; Lewis, J. C., *Angew Chem Int Ed Engl* **2013,** *52* (20), 5271-4.
3. Poor, C. B.; Andorfer, M. C.; Lewis, J. C., *Chembiochem* **2014,** *15* (9), 1286-9.
4. Payne, J. T.; Poor, C. B.; Lewis, J. C., *Angew Chem Int Ed Engl* **2015,** *54* (14), 4226-30.
5. Andorfer, M. C.; Park, H. J.; Vergara-Coll, J.; Lewis, J. C., *Chem Sci* **2016,** *7* (6), 3720-3729.
6. Belsare, K. D.; Andorfer, M. C.; Cardenas, F. S.; Chael, J. R.; Park, H. J.; Lewis, J. C., *ACS Synth Biol* **2017,** *6* (3), 416-420.