UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CHEMOENZYMATIC TRANSFORMATIONS OF INDOLE AND AZAINDOLE CONTAINING NATURAL PRODUCT ANALOGS USING PROMISCUOUS PRENYLTRANSFERASES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ERIC D. GARDNER

Norman, Oklahoma

2021

CHEMOENZYMATIC TRANSFORMATIONS OF INDOLE AND AZAINDOLE CONTAINING NATURAL PRODUCT ANALOGS USING PROMISCUOUS PRENYLTRANSFERASES

A DISSERTATION APPROVED FOR THE

DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY

BY THE COMMITTEE CONSISTING OF

Dr. Shanteri Singh, Chair

Dr. Edgar A. O'Rear

Dr. Daniel T. Glatzhofer

Dr. Wai Tak Yip

Dedication

This dissertation is dedicated to my wife, Jessi Gardner. Her unending love and support began as we simultaneously joined the OU Chemistry graduate program and has continuously played an essential role in my development as a person on the infinite pursuit of knowledge and truth.

The truth will set you free.

Abstract

The overall aim of my work was to expand the practical utility of prenyltransferase-based biocatalysis as a tool in organic synthesis, natural product diversification, and drug discovery efforts. Aromatic prenyltransferases (PTs) serve a key role in the biosynthesis of countless bioactive natural products (NPs), by catalyzing the transfer of a 5 carbon prenyl (dimethylallyl) moiety from an alkyl pyrophosphate "donor" onto an aromatic "acceptor" substrate via electrophilic aromatic substitution. Many PTs have been observed to have remarkably broad substrate scopes, capable of accepting allylic, benzylic, heterobenzylic, and diene containing alkyl donors and transferring them to a range of phenolic and indole-containing substrates. Having evolved to carry out highly regioselective and chemoselective alkylation in complex chemical environments, PTs represent a powerful untapped source of late-stage functionalization catalysts for both natural and synthetic drug diversification efforts.

Chapter 1 introduces and highlights the significance of indole-containing NPs and the diverse spectrum of biological activity that arises from substituted indole scaffolds. A brief overview of known synthetic methods for indole functionalization is covered, focusing on the challenges associated with functionalizing the benzenoid portion of indole. Subsequently, indole-modifying enzymes are introduced, and the structure, function, and mechanism of indole PTs is described as a foundation to this dissertation work.

Chapter 2 demonstrates the utility of PT-based natural product diversification, in which the cytotoxic prenylated tryptophan-containing cyclic dipeptide tryprostatin B (TPS-B) was synthesized and chemoenzymatically alkylated using the promiscuous PT CdpNPT. Using a library of 66 synthetic alkyl-pyrophosphate donors, 24 unique donors were accepted by CdpNPT in the presence of TPS-B to generate novel NP analogs. 11 of these chemoenzymatically produced TPS-B analogs were isolated for structure elucidation via nuclear magnetic resonance (NMR) to reveal high selectivity for indole C6 alkylation. Cytotoxicity assays revealed that the TPS-B analogs produced in this work have a potency similar to their parent NPs. This work demonstrates that PT-based biocatalysts can be used for the late-stage diversification of NPs which provides direct access to NP analogs not accessible through current synthetic methods.

Chapter 3 outlines the synthesis and utilization of a series of azaindole-substituted tryptophan analogs (Aza-Trp) to probe the compatibility of tryptophan-PTs with medicinally relevant indole isosteres. Synthetic tryptophan-mimetic substrates containing additional aromatic N atoms at the 2, 4, 5, 6, and 7 positions were prepared and screened as substrates for the indole C4-PT FgaPT2. These results identified 4-azatryprophan and 5-azatryptophan as previously unreported substrate classes for aromatic PTs. After structural elucidation of the prenylated products, we discovered FgaPT2 catalyzed N-prenylation of the 6-membered ring of these azaindole substrates to form a cationic N-prenylpyridinium products. Not only is this the first report of chemoenzymatic prenylation of azaindole substrates, but this work uncovered a previously undocumented PT-catalyzed reaction.

Chapter 4 reports the synthesis and evaluation of azaindole-substituted tryptophan-proline cyclic dipeptides (Aza-CyWP) as substrates for aromatic PTs, with the goal of chemoenzymatically producing azaindole-containing analogs of this privileged NP core scaffold with enhanced aqueous solubility and altered bioactivity. We discovered that

the indole-C2 PT which has the native function of TPS-B biosynthesis, FtmPT1, was capable of prenylating all 5 azaindole-containing substrates. However, these synthetic substrates were found to alter the regioselectivity of prenylation in an azaindole iso-dependent manner, resulting in a total of 7 fully characterized N1, C2, and C3 prenylated products. Additionally, the highly promiscuous PT utilized in **Chapter 2**, CdpNPT, was found to accept the 7-azaindole containing Aza-CyWP, to produce cyclized C3-reverse prenylated and N1 prenylated products. Using the common medicinal chemistry strategy of isosteric replacement, azaindole containing NP analogs were successfully generated using indole PTs. This work represents the first step toward the application of PT-based biocatalytic functionalization of synthetic heterocyclic scaffolds commonly utilized in structure-activity relationship studies.

Acknowledgements

First and foremost, I would like to thank Dr. Shanteri Singh for being an amazing mentor. Your immense, sustained effort and excellent leadership has allowed me to develop into not only a high-functioning multidisciplinary scientist, but also a better person overall. You are an outstanding and genuine inspiration for the future of multidisciplinary science. Thank you for the perspectives, lessons, and experiences you have given me, as they will surely have a lasting impact on my life and career. I would also like to thank members of the Singh lab, past (Dr. Erin Skull, Dr. Chandrasekhar Bandari, Tejaswi Bavineni, and Dr. Prashant Mandal) and present (Dustin Dimas, Bryce Johnson, and Vikas Kumar) for their intellectually stimulating conversations, strong sense of humor and comradery, and overall support that enabled me to call this community my home. I specifically thank Dustin Dimas for being a fantastic friend, undergraduate mentor, colleague, and for playing a huge role in co-managing the organic synthesis side of the Singh lab.

Secondly, I sincerely thank Dr. Susan Nimmo for her extensive mentorship and support through the Nuclear Magnetic Resonance facility. The experience and training I received working as her assistant enabled me to develop my passion for NMR and research support. I am forever grateful for her support and inspiration, as I wouldn't have been able to develop into the chemist I am today without Susan's help. She served as a pillar of light, able to illuminate the darkest days for countless organic chemistry students. She made a lasting positive impact on myself and many others that has helped fine-tune the trajectory of my scientific career.

Next, I would like to thank my friends I have made since joining the Chemistry department at OU. They have been a godsend for my mental health and wellbeing, and this amazing community that we formed provided some of the most enjoyable, thought provoking, and supportive relationships I have ever known. I thank Quentin Avila, Dr. Steven Schlitzer, Theo Rusmore, Rob Fogel, Dustin Dimas, and many others for the emotional and technical support I needed to survive the Ph. D. program.

Lastly, I would like to thank my family for supporting my interests and wellbeing throughout my whole life and encouraging me to do the work needed to reach my full potential. I am extremely thankful for my parents, David and Lisa Gardner, for believing in me even when I didn't believe in myself. Without their extensive love and support, I would have never made it into graduate school, let alone finished. These academic pursuits led me to Oklahoma where I met Jessi Gardner, who has helped me succeed in more ways than I could possibly fit in this document.

I am deeply thankful to everyone who helped me find the right path, and I aspire to help others as much as you all have helped me.

Table of Contents

Table of Figures:

Chapter 1. Introduction

1.1 The Importance of Indole

The indole moiety is ubiquitous in nature. Indole is the constituent heterocycle in tryptophan (Trp), and is therefore found in essentially all living systems. The human body utilizes several indole-based neurotransmitters such as serotonin (5-hydroxytryptamine, 5-HT) for regulating mood and blood flow, and melatonin for regulating the circadian rhythm. The biosynthetic availability of tryptophan has given rise to a broad spectrum of potent naturally occurring indole alkaloids (Figure 1.1) such as the anti-tumor agent vinblastine, the hallucinogen lysergic acid, the anti-hypertensive/anti-psychotic reserpine, and Trp containing peptide antibiotics such as daptomycin. The privileged indole scaffold is a key structural feature of bioactive compounds with diverse activity, such as antibiotic, anti-fungal, anti-tumor, anti-inflammatory, anti-viral, anti-psychotics, hallucinogens, and more.[1] The indole moiety is a privileged scaffold due its ability to bind to multiple receptors resulting in a wide spectrum of bioactivity. Heterocycles are key features of countless approved drugs and bioactive natural products. As of 2010, 80% of drugs sold in the US contain a heterocycle.[1] One of the most frequently encountered and thoroughly studied heterocycles is indole. Chemists have been exploring indole synthesis methods for more than a century, and through this it became clear that the medicinal relevance of the indole scaffold is extremely high.[2]

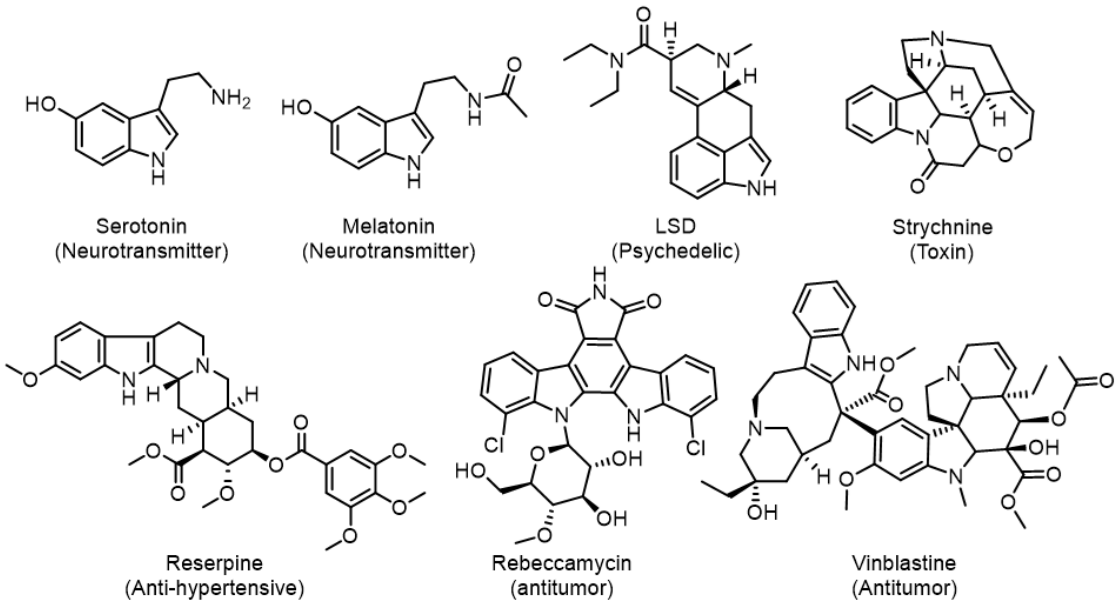


Figure 1.1: Representative examples of biologically active indole containing natural product scaffolds.

1.2 Indole Substitution Effects

Several studies have demonstrated that the derivatization of indole moiety within compounds can alter their receptor selectivity, metabolic stability, potency, etc., which has enabled the development of many approved pharmaceuticals.[1, 3] For example, drugs acting on various serotonin receptors have found use in treating nausea, anxiety, psychosis, and severe headaches. The hallucinogen dimethyltryptamine (DMT), the active component in the traditional medicine ayahuasca, is a nonselective serotonin agonist. A whole class of approved anti-migraine drugs known as triptans has been developed by modifying the indole C5 of DMT, resulting in selective serotonin agonists which lack the perceptual and cognitive effects of the parent drug.

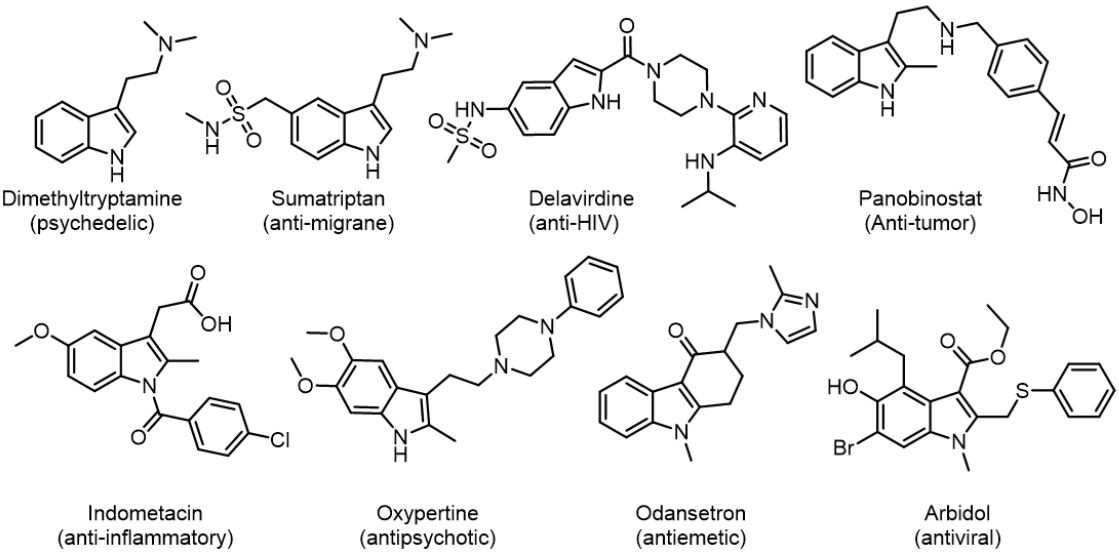


Figure 1.2: Select indole containing drugs showcasing the diverse bioactivity accessible using the privileged indole scaffold.

A majority of bioactive indoles feature substituents at the C3 and/or C2 positions. For example, the anticancer drug panobinostat, the antipsychotic oxypertine, and the antiemetic odansetron all feature small hydrophobic substituents at C2 (Figure 1.2). These three drugs also have C3 substituents featuring a linker containing polar groups, followed by an aromatic ring. The anti-HIV drug delavirdine also shares this trend, however it is present on C2. Despite the similarities in the core scaffold, these compounds have dramatically different bioactivities.

1.3 Methods of Indole Functionalization

1.3.1 Synthetic Functionalization Methods

The synthesis of indole containing pharmaceuticals typically begins with indole substituent pre-installed on an aromatic ring prior to synthesis of the indole ring. Installation of indole substituents prior to indole synthesis is an effective strategy to explore substituent effects on simple indole containing drugs. This strategy becomes less effective

as the structural complexity increases, due to the number of steps required to arrive at the desired product.

Synthetic strategies of modifying existing indole rings has been thoroughly studied, however some significant limitations exist. Countless methods of indole N1, C2, and C3 functionalization have been developed because of the high electron density of the pyrrole ring, that increases its overall reactivity. Alkylation of N1 can be achieved by simply treating with a base followed by an alkyl halide. C3 is the most electron rich site, readily attacking electrophiles such as the iminium intermediate of the Mannich reaction to form gramine.[4] The synthetic methods to functionalize the benzene portion of indole are currently extremely limited.[5] Methods of functionalizing C4 and C7 frequently rely on coordination of the catalyst to specific directing groups installed on either C3 (Figure 1.4) or N1 (Figure 1.3), respectively.[5-7]

1.3.2 Regioselective Functionalization of Indole using N1 Directing Groups

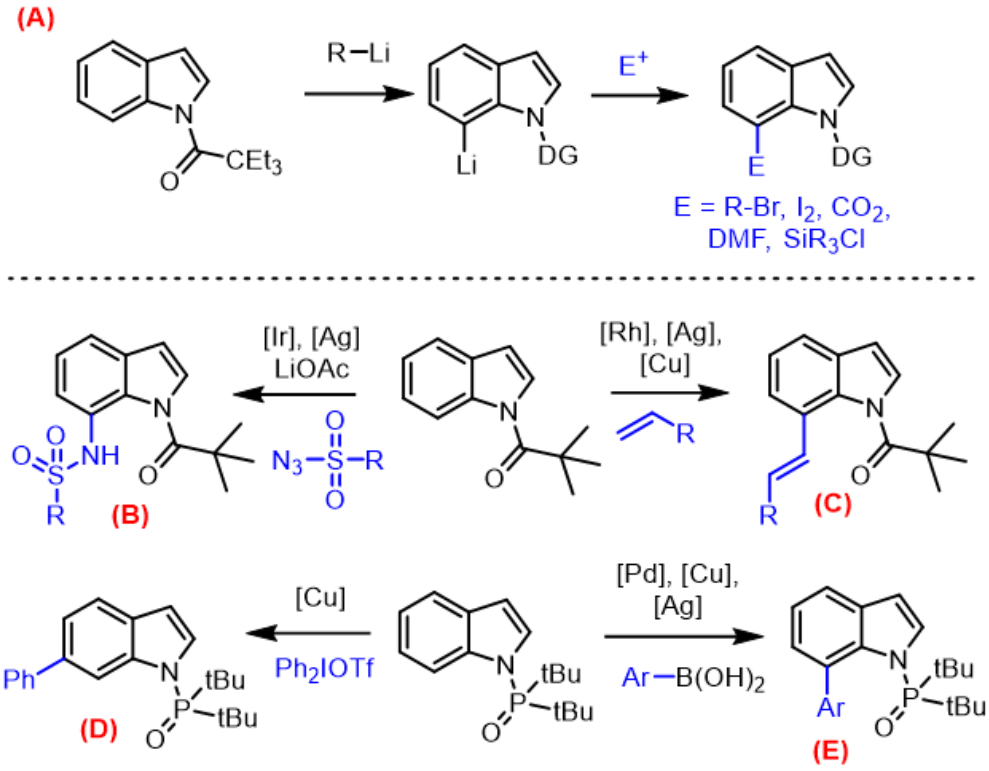


Figure 1.3: Representative indole functionalization methods exploiting directing groups present on N1.
A) Method for direct C7 lithiation and subsequent electrophilic substitution. **B)** Iridium catalyzed C7 selective sulfonamidation of N-pivaloyl indoles. **C)** Selective C7 allylation of N-pivaloyl indoles catalyzed by a combination of [Rh], [Ag], and [Cu]. **D)** A rare example of C6 selective indole functionalization using hypervalent iodine aromatic donor in the presence of an N-ditertbutylphosphine directing group. **E)** Regioselective C4 arylation of N-ditertbutylphosphinyl indoles using aryl boronic acids in the presence of [Pd], [Cu], and [Ag] based catalysts.

Most directing groups (DG) installed on N1 are both sterically demanding and contain an oxygen capable of coordinating to catalyst or reagent. One of the first indole C7 alkylation methods discovered was a direct C7 lithiation which enables capture of electrophiles like CO_2, DMF, iodine, alkyl bromides and trialkylsilyl chlorides (Figure 1.3-A).[8] This approach was later expanded to incorporate Suzuki-Miyaura cross coupling.[9] The frequently encountered pivaloyl DG has been used for indole C7 functionalization. A

C7 selective sulfonamidation of N-pivaloyl indoles has been reported using a combination of $Cp^*Ir(OAc)_2$ and $AgNTf_2$ (Figure 1.3-B).[7] Similarly, this substrate class can undergo C7 selective alkene coupling in the presence of $[RhCp^*Cl_2]_2$, $AgNTf_2$, and $Cu(OAc)•H_2O$ (Figure 1.3-C).[10] Another analogous DG capable of similar reactivity is the ditertbutylphosphine group. Selective indole C7 arylation using arylboronic acids has been demonstrated using $Pd(OAc)_2$, $Cu(OTf)_2$, CuO, and Ag_2O in the presence of a 2-chloropyridine ligand (Figure 1.3-E).[11] Remarkably, the same DG has been capable of indole C6 arylation in the presence of CuO and diaryliodonium triflate (Figure 1.3-D).[12]

1.3.3 Regioselective Functionalization of Indole using C3 Directing Groups

Functionalization of indole C4 has been carried out most commonly with acyl directing groups on the C3 position. These methods typically require protecting groups on the N1 position of indole. The installation of the pivaloyl DG on C3 can be executed with $AlEt_2Cl$ and PivCl, and removal of the DG can be accomplished using TsOH in glycol. C4 arylation of these indoles can be achieved using $Pd(PPh_3)_2Cl_2$, Ag_2O, DBU, and an aryl iodide (Figure 1.4-C).[13] Remarkably C5 selective arylation was achieved by the same group using a combination of CuTc and diphenyliodonium triflate (Figure 1.4-E).[13] Lanke et al. discovered in 2016 that directing group electronics effects have a dramatic influence on catalyst regioselectivity. N-methylindole featuring a trifluoroacetyl group on C3 produced exclusively C4 alkenylation in the presence of Rh (III), $AgSbF_6$, and $Cu(OAc)_2•H2O$ (Figure 1.4-F).[14] Comparatively, switching the DG to an acyl group and replacing Rh (III) with Ru (II) yields exclusively C2 alkenylated indoles (Figure 1.4-D).[14]

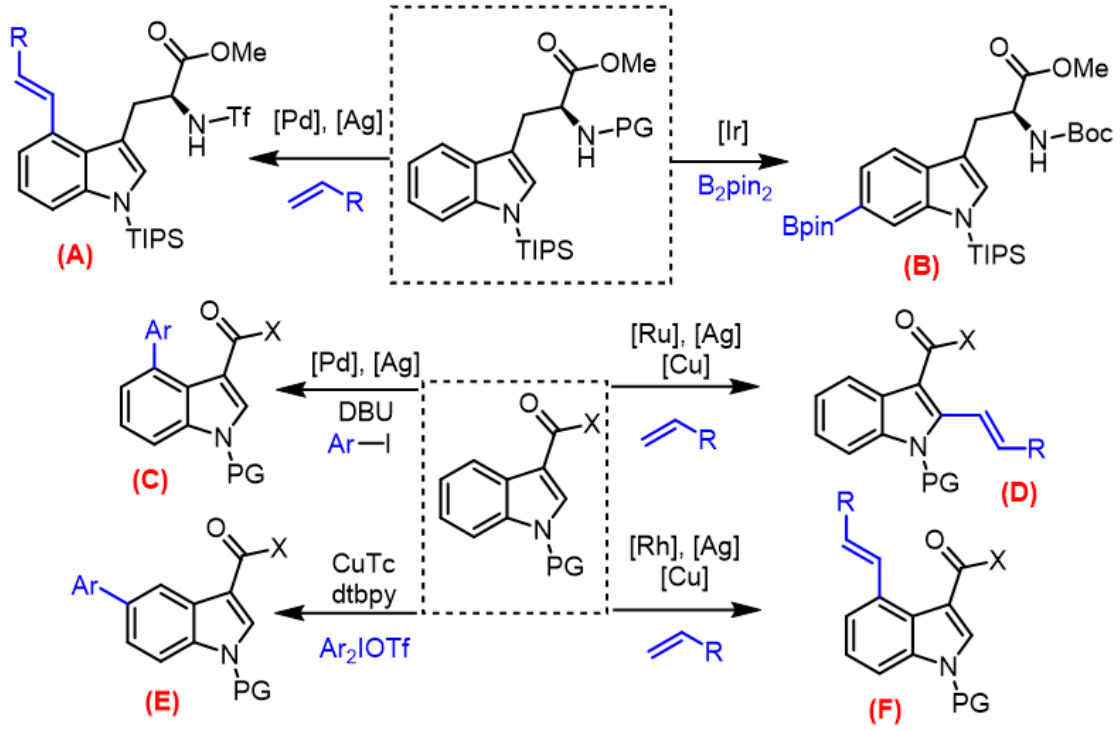


Figure 1.4: Representative organometallic indole functionalization methods
Reaction **(B)** features unique ligand controlled regioselectivity while **(A), (C), (D)** and **(F)** rely on directing groups installed on C3.

Few methods exist for selective C5 and C6 indole functionalization compared to the other positions. This selectivity problem can partially be explained by the similar reactivity of these positions, and their distance from N1 and C3 directing groups. The few existing methods suffer from extremely limited functional group tolerance on both the indole itself and the functionalities being installed.[13, 15] Baran *et at.* developed a remarkable example of iridium-catalyzed ligand-controlled selective C-6 tryptophan borylation (Figure 1.4-B), however even this groundbreaking directing group-free method requires several protecting groups and is unable to tolerate tryptophan containing cyclic dipeptides.[16]

Current late stage modification techniques used in pharmaceutical development frequently utilize photo-redox or radical based reactions,[17] however this approach is unable to reliably functionalize complex natural products.

1.4 Challenges in Natural Product Diversification

Natural products (NPs) are a long-standing source of medicine. A majority of currently approved drugs are natural products, derived from natural products, or inspired from natural product pharmacophores.[18] Compared to their fully synthetic counterparts, NPs generally have higher structural complexity, higher functional group density, and more chiral centers. This high complexity is conducive to high target specificity, and therefore less prone to produce unwanted side-effects. Unfortunately for synthetic chemists, this increased complexity usually comes with increased or altered reactivity. Every additional functional group is another potential point of interaction during a transformation, increasing the probability of generating unwanted side products. Chemically modifying extracted natural products to generate analogs is known as semi-synthesis. For example, the successful semisynthetic derivatization of morphine has led to the development of many approved analgesics such as oxycodone, hydrocodone, diacetylmorphine (heroin), and more.[19, 20] These derivatives were generated by transforming pre-existing synthetic handles found on natural products extracted from opium, thus limiting analog synthesis to specific portions of the molecule. There is a great need to enhance the scope of late-stage natural product modification to expand our knowledge of these pharmacophores and enable the development of new bioactive compounds.

1.5 Chemoenzymatic Indole Functionalization

1.5.1 Synthetic Utility of Tryptophan Synthase

In the context of indole late-stage modification, chemoenzymatic transformations have shown great potential. Enzymes can recognize specific moieties within complex chemical environments and perform reactions with high regioselectivity. Biocatalysts are attractive options in pharmaceutical development due to their lack of toxic and/or expensive metals, decreased production of side products, and overall environmental friendliness.[21, 22] The biosynthetic enzymes which produce and modify indole containing natural products can be utilized to perform challenging transformations. Tryptophan synthase for example has been used to produce unnatural amino acids with excellent enantioselectivity from substituted indoles, azaindoles, and other sulfur and selenium containing heterocycles.[23] (Figure 1.5) Protein engineering methods have successfully expanded the substrate scope of tryptophan synthase to produce β-alkyl Trp analogs.[24] Tryptophan synthase has also been engineered to react with non-indole nucleophiles such as nitro alkanes in order to produce unusual γ-nitro amino acids.[25] Similarly, mutant Trp synthase has been developed which are capable of forming new quaternary carbon centers from 3-methyloxindole.[26]

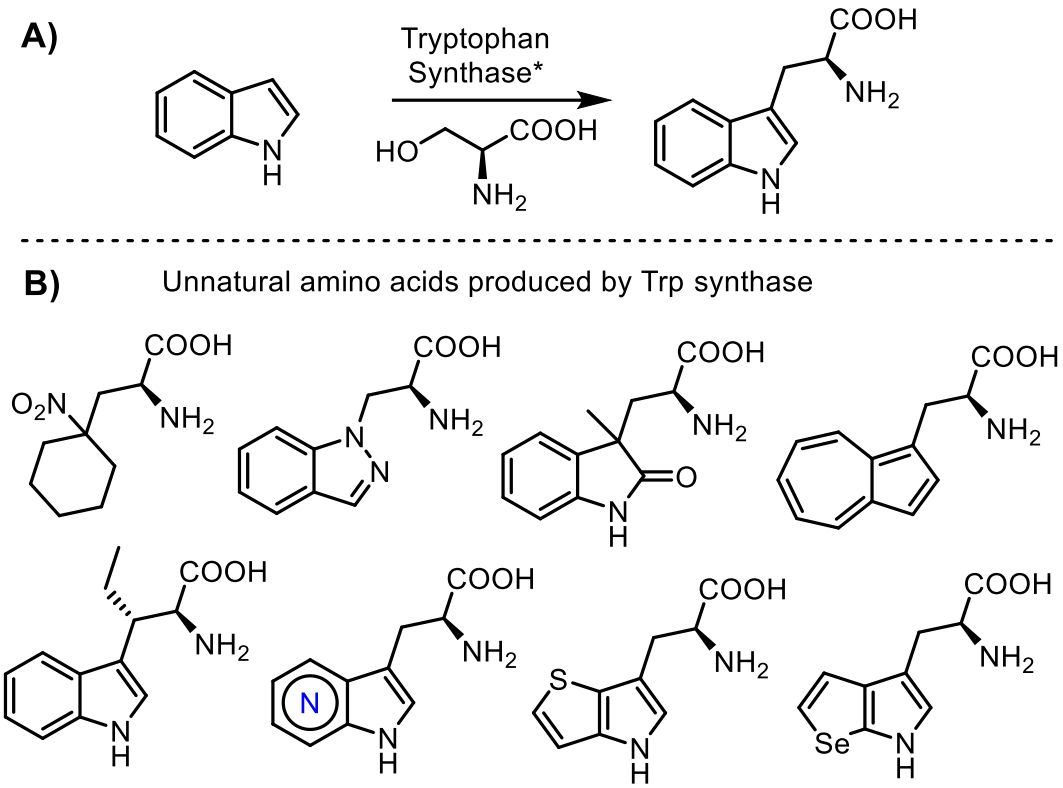


Figure 1.5: Reactions catalyzed by tryptophan synthase TrpB
A) Native biosynthetic role of tryptophan synthase. **B)** Representitive unnatural amino acids produced by native or mutant tryptophan synthases from their corresponding aromatic ring scaffold or nitroalkane. (bottom)

1.5.2 Biocatalytic Late-Stage Functionalization of Indole Substrates

Tryptophan halogenases such as PyrH, ThdH, and RebH are proficient in regioselective tryptophan halogenation at the C5, C6, and C7 positions, respectively.[27, 28] Directed evolution has been successful in producing several RebH variants that selectively halogenate C6 and C5 of tryptamine, while native RebH is selective for C7.[29] Other engineered RebH variants have demonstrated a much larger substrate tolerance, accepting larger indole scaffolds such as yohimbine (Figure 1.6-C).[30] Similarly, an Fe(II)/2-oxogluterate dependant oxygenase WeIO5 has been successfully used in the late stage aliphatic halogenation of the fischerindole scaffold (Figure 1.6-D).[11] A biocatalytic Trp nitrating system has recently been developed to access synthetically challenging

enantiopure nitrotryptophans.[31] This nitration system relies on the engineered P450 TB14 in combination with BsNOS which generates NO from L-Arg (Figure 1.6-A). In the context of drug diversification, indole alkylating enzymes with relaxed substrate specificities are desired so they can be utilized for a wide spectrum of complex indole substrates. Protein engineering methods have proved useful in altering the regioselectivity, substrate specificity, and stability of these enzymes.[24, 31, 32]

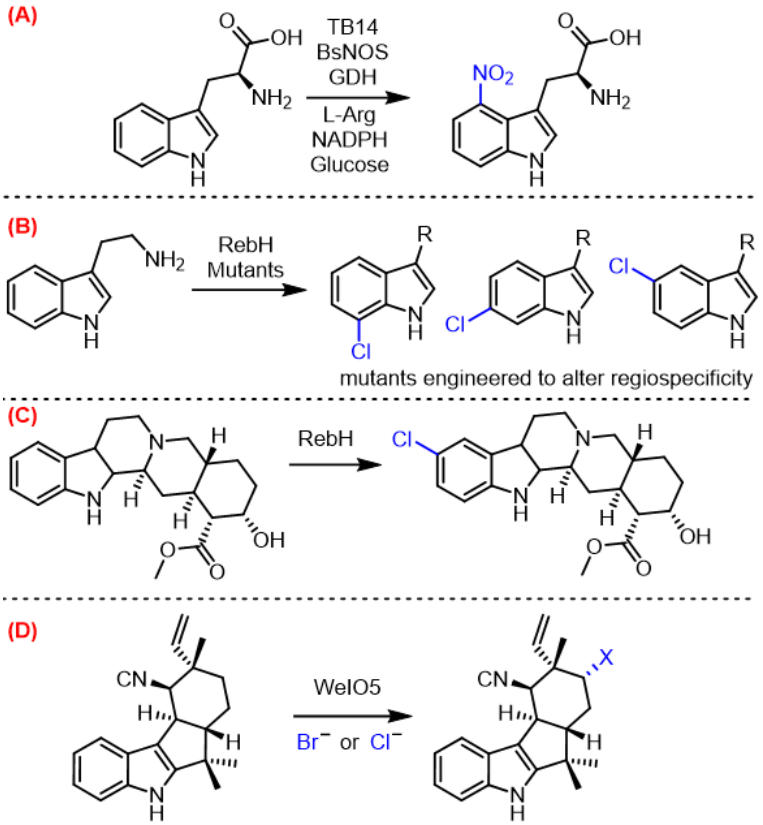


Figure 1.6: Representative examples of biocatalytic late stage modification of indole containing natural products.
A) Biocatalytic production of enantiopure 4-nitrotryptophan using P450 TB12 in combination with BsNOS for in situ generation of NO from arginine **B)** Mutants of the promiscuous indole halogenase RebH has been engineered to produce various regioisomers of chlorotryptamine **C)** Selective indole C5 halogenation of a complex natural product catalyzed by RebH **D)** Regio- and enantioselective halogenation of an unactivated aliphatic moiety within an indole containing terpenoid scaffold by the Fe(II)/2-oxogluterate dependent oxygenase WelO5

1.6 Aromatic Prenyltransferases

1.6.1 Overview of Aromatic Prenyltransferases

Biocatalytic indole functionalization by Trp-synthase, halogenases, and oxygenases have proven useful, however their capacity for late stage functionalization of NPs is limited due to their high substrate specificity. A class of alkylating enzymes known as prenyl transferases show great promise in the future of biocatalysts due to their remarkably broad substrate scope. Prenyltransferases (PTs) are enzymes that catalyze the transfer of a prenyl group to an acceptor molecule. Prenylation is ubiquitous in biological systems. These prenylated molecules include biosynthetic intermediates, proteins, and bioactive natural products. Many soluble prenyltransferases have been identified in bacteria and fungi which prenylate aromatic substrates. These soluble aromatic prenyltransferases are the main focus of this study. Their native function is primarily in the biosynthesis of secondary metabolites (Figure 1.7). The addition of prenyl groups can be a useful late-stage modification of secondary metabolites to alter their lipophilicity and biological activity. Prenyl groups can also serve as building blocks to be incorporated into the carbon skeleton of natural products by subsequent cyclization or oxidation.[33]

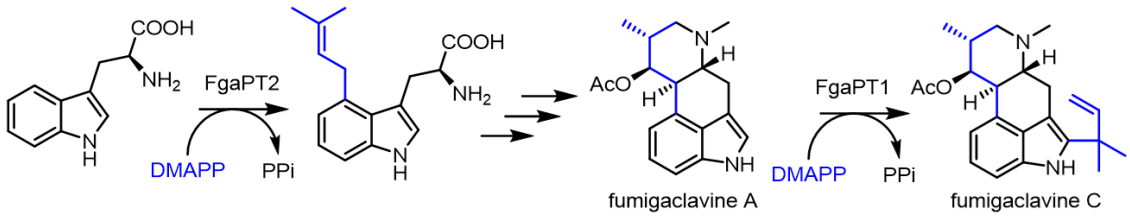


Figure 1.7: The role of prenyltransferases in the biosynthesis of fumigaclavine C
The first prenyl group installed by FgaPT2 is used to build the carbon skeleton downstream, while the late stage prenylation by FgaPT1 catalyzes reverse C2 prenylation. DMAPP derived carbons are shown in blue.

1.6.2 Structure of Aromatic Prenyltransferases

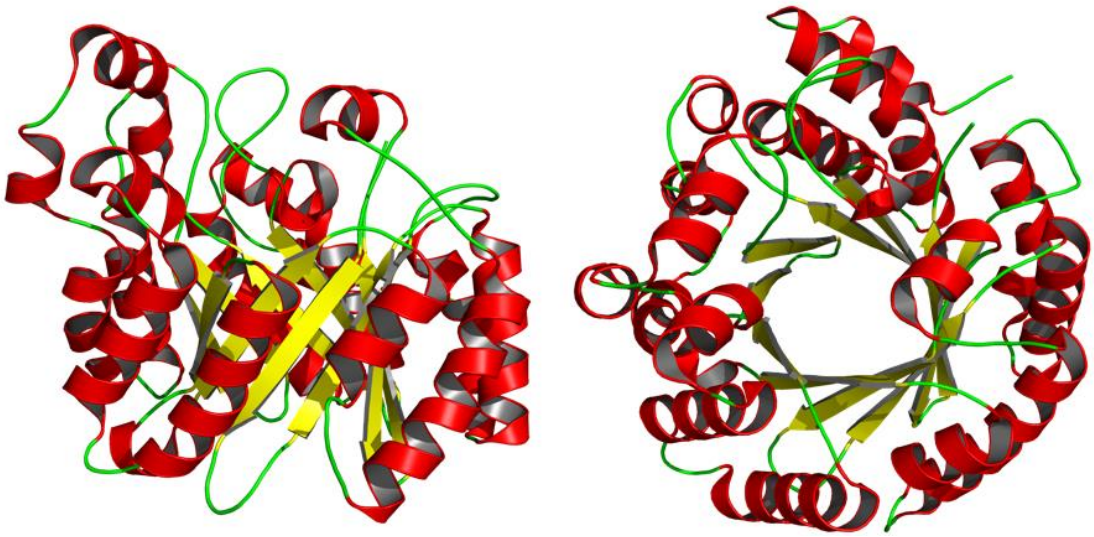


Figure 1.8: Side and top view of the αββα prenyltransferases motif.
The structure of CdpNPT is shown with α-helices and β-sheets colored red and yellow, respectively. The reaction chamber is within the anti-parallel β-barrel. (PDB: 4E0U)[34]

This class of enzymes all share a common ααββ structural motif known as the PT-barrel, also known as the ABBA PT-barrel (Figure 1.8). One end of this barrel is the pyrophosphate "donor" binding site, and the other end is the aromatic "acceptor" binding pocket. Within this class there is a large family of dimethylallyl tryptophan synthases (DMATS) which prenylate the indole ring of tryptophan and tryptophan containing compounds. This dissertation will be referring to them as indole PTs. Other "non-indole" soluble aromatic PTs such as NphB, EpzP and CloQ utilize aromatic compounds such as phenols, naphthalenes, phenazines, and flavonoids.[35] Many of these soluble aromatic PTs exhibit considerable substrate promiscuity which is critical in the development of synthetically useful biocatalysts. These enzymes are tools capable of regiospecifically functionalizing complex compounds in ways not currently achievable via synthetic means.

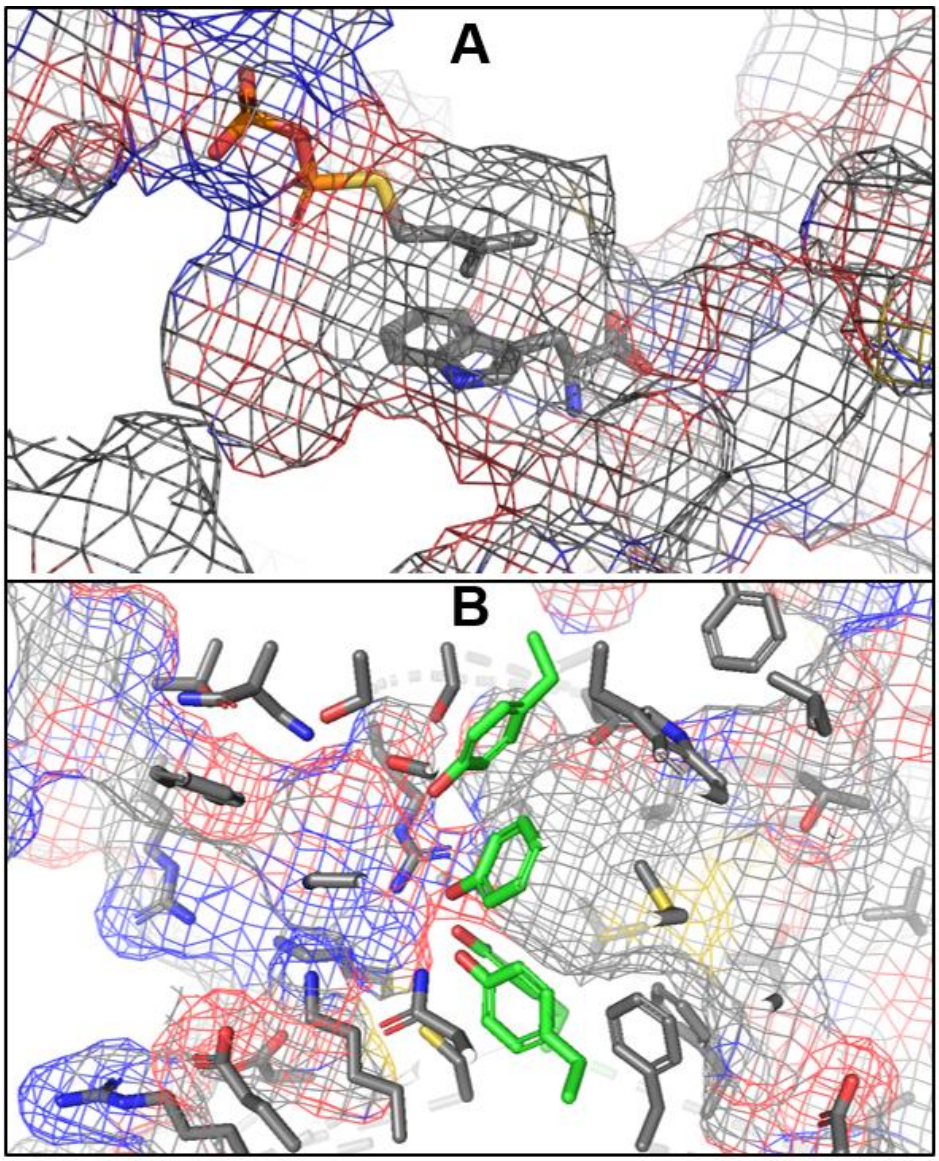


Figure 1.9: Internal structure and mechanism of prenyltransferases.
A) Visualization of the π-stacking interaction of the prenyl group directly above the Trp acceptor in FgaPT2. **B)** Internal structure of the CdpNPT binding pocket with the polar pyrophosphate binding site (left side), the highly conserved 4 tyrosine solvent shield (Green), and the large hydrophobic acceptor binding site (right side).[36]

1.6.3 Mechanism of Aromatic Prenyltransferases

PTs function by transferring a prenyl group derived from prenyl pyrophosphate, referred to as the "donor", to an aromatic "acceptor" molecule. The smallest prenyl donor, the 5 carbon dimethylallyl pyrophosphate (DMAPP), can be enzymatically elongated by

condensation with isopentenyl pyrophosphate (IPP) to form 10, 15, and 20 carbon isoprene groups known as geranyl, farnesyl, and geranylgeranyl, respectively.

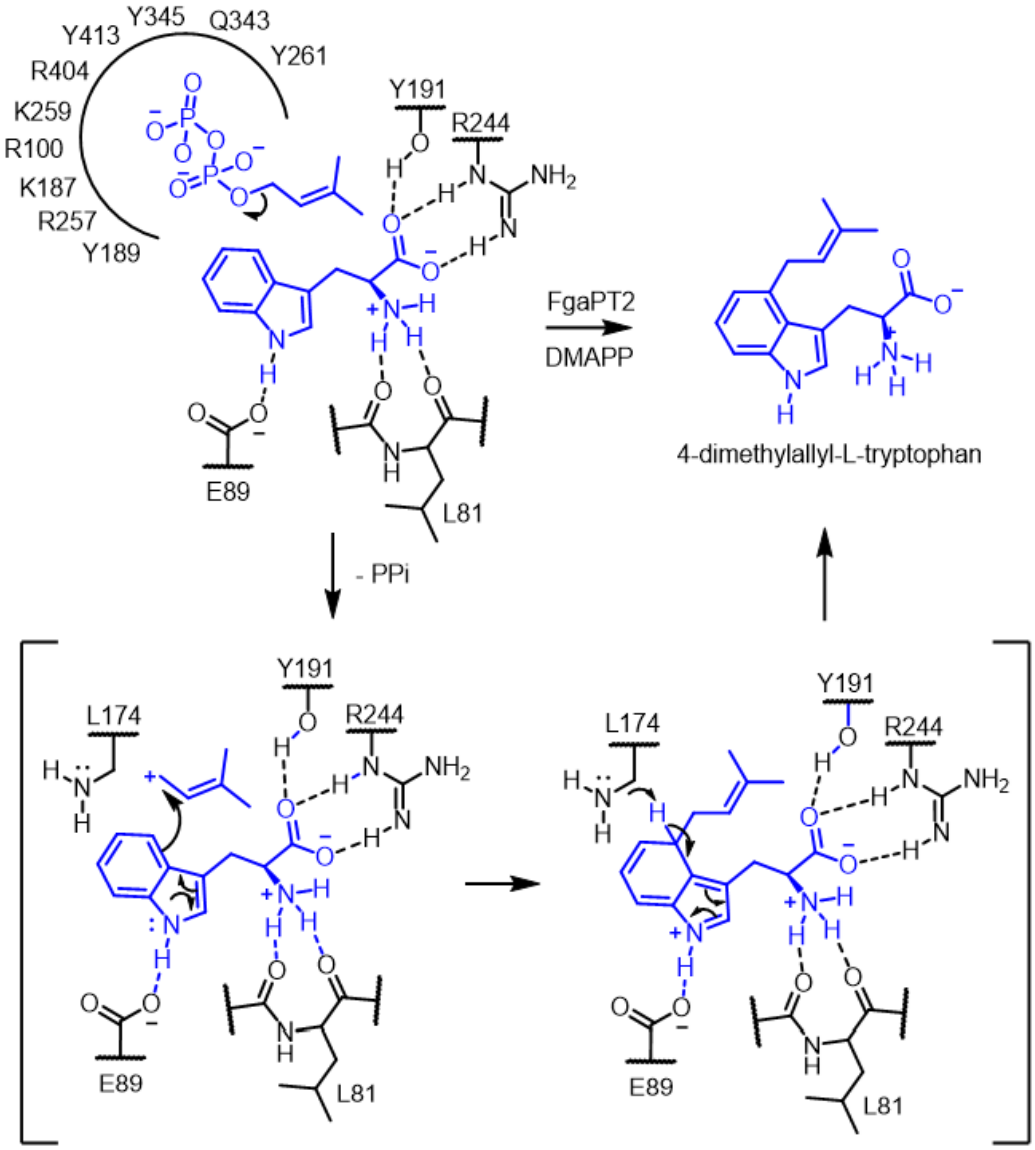


Figure 1.10: Proposed mechanism of FgaPT2 catalyzed 4-prenylation of Trp.

The pyrophosphate moiety functions as a leaving group, which upon dissociation generates a carbocation. The native donor, dimethylallyl pyrophosphate (DMAPP), can stabilize this carbocation by delocalizing the positive charge from the initial allylic carbocation between a highly stabilized tertiary allylic carbocation (Figure 1.10). This

carbocation is generated in a solvent-shielded reaction chamber, sandwiched between stabilizing aromatic amino acid side chains and the aromatic acceptor (Figure 1.9). The aromatic acceptor then captures the carbocation through an electrophilic aromatic substitution reaction, followed by deprotonation to restore aromaticity. Once alkylation is complete, the product is released from the reaction chamber.

1.6.4 Substrate Scope of Aromatic Prenyltransferases

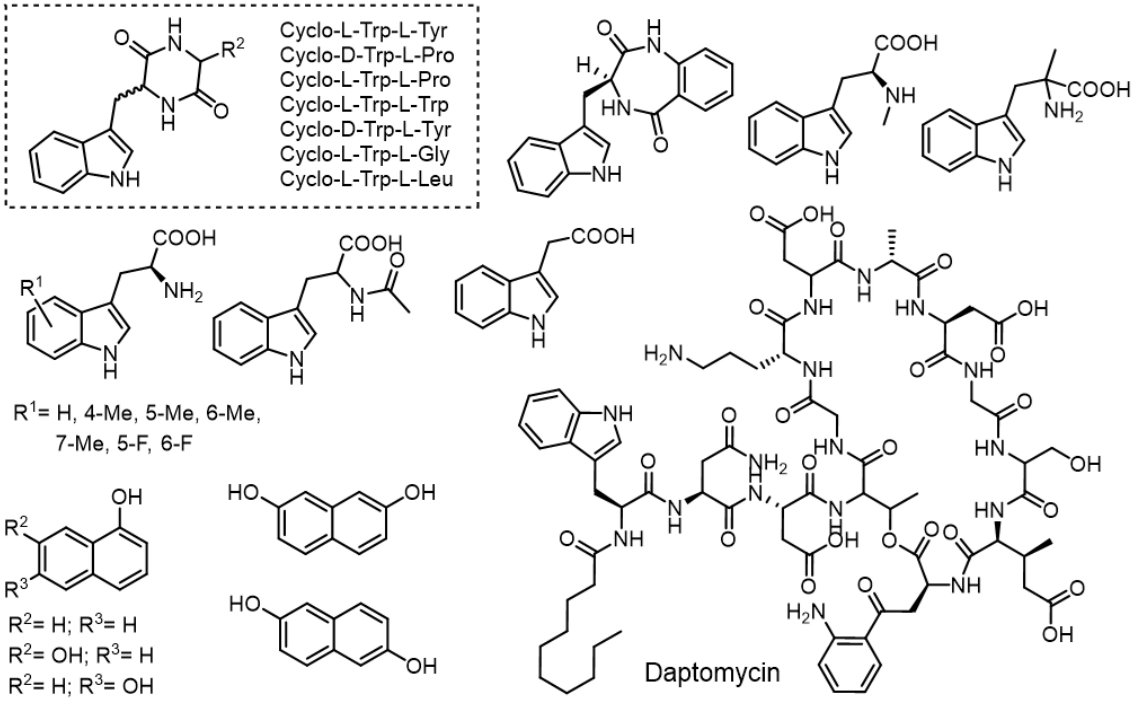


Figure 1.11 Previously reported acceptor substrates for CdpNPT

Enzymes generally are typically known for their remarkable substrate specificity; however, many indole PTs have demonstrated remarkable promiscuity in both acceptor and donor substrates. The remarkably broad accepter tolerance has been explored by several other research groups. CdpNPT for example likely has tryptophan containing cyclic dipeptides as its native acceptor, however it has been shown to accept a wide variety of other Trp containing cyclic dipeptides, including several containing unnatural D amino

acids (Figure 1.11).[37, 38] Several hydroxynapthlenes and methyltryptophan substrates have also been accepted by CdpNPT (Figure 1.11). Most remarkably, CdpNPT was shown to accept the cyclic lipodepsipeptide antibiotic daptomycin as a substrate, producing several new active daptomycin analogs.[39] Daptomycin contains 13 amino acids, while CdpNPT's native substrates are cyclic dipeptides. This example of remarkable acceptor promiscuity demonstrates the potential for developing PTs into broadly applicable indole functionalization catalysts.

The donor substrate scope of indole PTs has been explored to a lesser degree due to the labor-intensive process of alkyl pyrophosphate synthesis and purification. The Singh lab has produced a library of >80 unnatural pyrophosphate donors to probe the donor substrate scope of aromatic PTs (Figure 1.12), as reported previously.[36, 40, 41] Remarkably, a wide variety of allylic, benzylic, and heterobenzylic pyrophosphates have been accepted by PTs such as FgaPT2, NphB, and CdpNPT. The carbocation stabilization provided to the allylic and benzylic carbocations currently appears to be necessary for the reaction to take place.

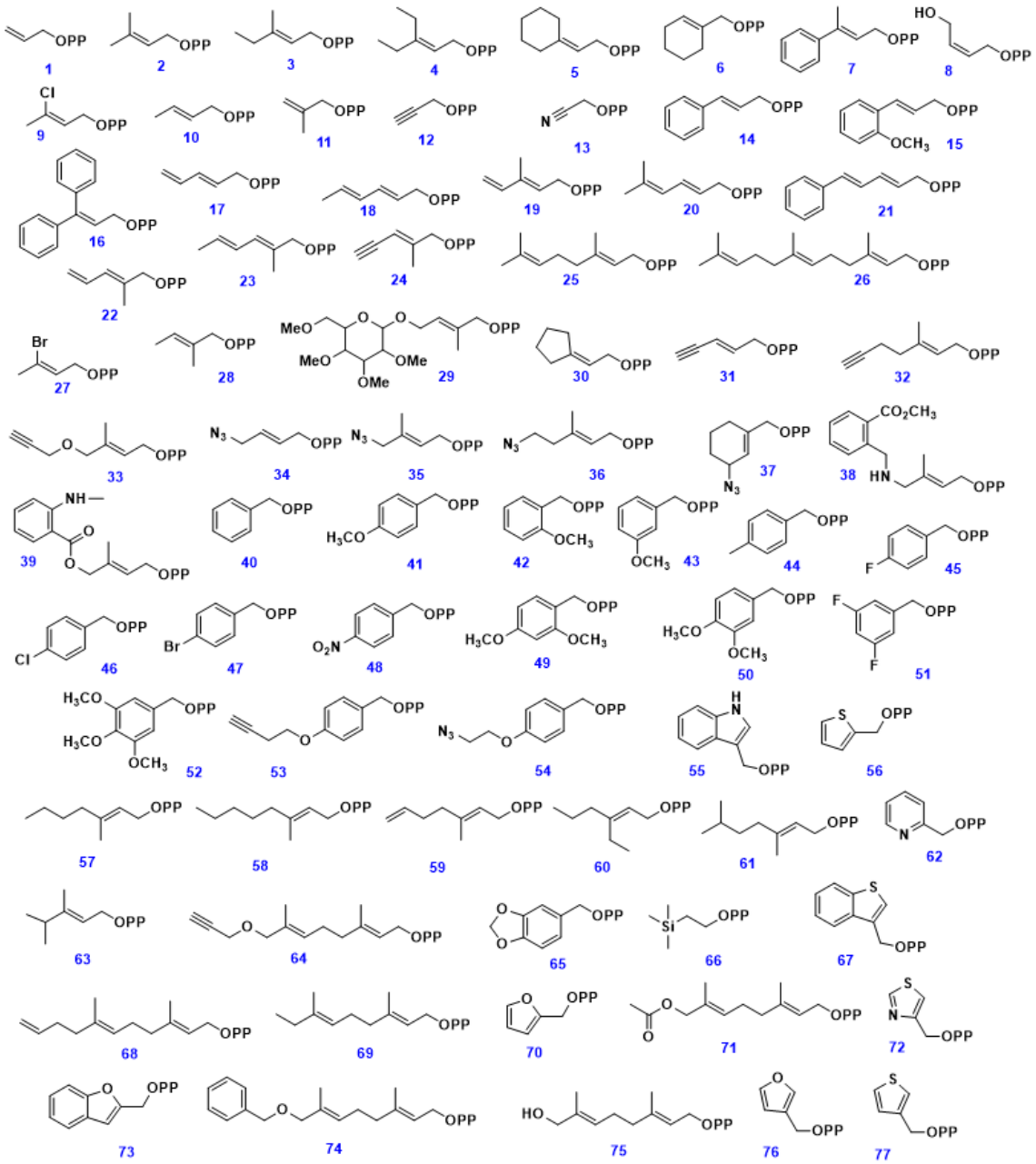


Figure 1.12: Synthetic pyrophosphate donor library used in this study
Synthetic methods for Alkyl-PP library generation were reported previously.[36, 40, 41]

1.7 Research Goals and Approach

The goal of this study is to harness the power of indole PTs to directly modify bioactive indole containing compounds and test the limits of their substrate scope. This research aims to advance the field of PT-based biocatalytic late-stage drug diversification by utilizing a library of unnatural donors to generate libraries of previously inaccessible

drug analogs. Multidisciplinary research such as this requires a combination of synthetic organic chemistry, biochemical methods, HPLC, and extensive NMR based structural analysis. A key aspect of this work was to expand the library of unnatural donors available to us to enable deeper exploration of the donor substrate scope of every PT studied in the Singh lab. As an application of this drug diversification method, my thesis focuses on the studies of chemoenzymatically diversified anti-cancer NP's such as tryprostatin. Additionally, this work highlights the synthesis of azaindole substituted tryptophans (Aza-Trp) and their tryprostatin-inspired diketopiperazines in a quest to understand the effect of indole isosteric replacement on the activity of tryprostatin, and to explore their potential as substrates for indole PT's. The next goal of this work was to expand the acceptor substrate scope beyond indoles. Azaindoles were determined to be likely PT substrate candidates because they are similar to indoles in size and can still participate in the important NH binding interactions within the binding pocket. Azaindoles also are present in many pharmaceuticals, potentially enabling enzymatic derivatization of these drugs.

Despite the massive number of known bioactive indole compounds, synthetic late-stage modification methods remain highly limited. Late-stage functionalization methods enable direct derivatization of complex products, elegantly producing drug analogs that would otherwise only be possible through total synthesis. Late-stage indole modification methods are highly desirable to enable direct derivatization of existing active compounds. This work has demonstrated the ability of indole PT's to perform late-stage functionalization of unnatural, pharmaceutically relevant azaindole substrates, substantially improving the potential industrial utility of indole PT's for drug manufacturing and diversification. Through protein engineering efforts in combination

with thorough substrate scope determination, aromatic PT's could be tailored to perform a wide variety of otherwise impossible chemical transformations with exceptional regiospecificity and chemoselectivity.

1.8 References for Chapter 1

1. Sravanthi, T. V.; Manju, S. L., Indoles — A promising scaffold for drug development. *European Journal of Pharmaceutical Sciences* **2016,** *91*, 1-10.

2. Taber, D. F.; Tirunahari, P. K., Indole synthesis: a review and proposed classification. *Tetrahedron* **2011,** *67* (38), 7195-7210.

3. Chadha, N.; Silakari, O., Indoles as therapeutics of interest in medicinal chemistry: Bird's eye view. *European Journal of Medicinal Chemistry* **2017,** *134*, 159-184.

4. Brehm, W. J.; Lindwall, H. G., THE PREPARATION OF MANNICH BASES RELATED TO GRAMINE. *The Journal of Organic Chemistry* **1950,** *15* (3), 685-687.

5. Leitch, J. A.; Bhonoah, Y.; Frost, C. G., Beyond C2 and C3: Transition-Metal-Catalyzed C–H Functionalization of Indole. *Acs Catal* **2017,** *7* (9), 5618-5627.

6. Robbins, D. W.; Boebel, T. A.; Hartwig, J. F., Iridium-Catalyzed, Silyl-Directed Borylation of Nitrogen-Containing Heterocycles. *Journal of the American Chemical Society* **2010,** *132* (12), 4068-4069.

7. Song, Z.; Antonchick, A. P., Iridium(iii)-catalyzed regioselective C7-sulfonamidation of indoles. **2016,** *14* (21), 4804-4808.

8. Fukuda, T.; Maeda, R.; Iwao, M., Directed C-7 lithiation of 1-(2,2-diethylbutanoyl)indoles. *Tetrahedron* **1999,** *55* (30), 9151-9162.

9. Hartung, C. G.; Fecher, A.; Chapell, B.; Snieckus, V., Directed ortho-Metalation Approach to C-7-Substituted Indoles. Suzuki−Miyaura Cross Coupling and the Synthesis of Pyrrolophenanthridone Alkaloids. *Organic letters* **2003,** *5* (11), 1899-1902.

10. Xu, L.; Zhang, C.; He, Y.; Tan, L.; Ma, D., Rhodium-Catalyzed Regioselective C7-Functionalization of N -Pivaloylindoles. *Angewandte Chemie International Edition* **2016,** *55* (1), 321-325.

11. Yang, Y.; Qiu, X.; Zhao, Y.; Mu, Y.; Shi, Z., Palladium-Catalyzed C–H Arylation of Indoles at the C7 Position. *Journal of the American Chemical Society* **2016,** *138* (2), 495-498.

12. Yang, Y.; Shi, Z., Regioselective direct arylation of indoles on the benzenoid moiety. *Chem Commun (Camb)* **2018,** *54* (14), 1676-1685.

13. Yang, Y.; Gao, P.; Zhao, Y.; Shi, Z., Regiocontrolled Direct C−H Arylation of Indoles at the C4 and C5 Positions. *Angewandte Chemie International Edition* **2017,** *56* (14), 3966-3971.

14. Lanke, V.; Bettadapur, K. R.; Prabhu, K. R., Electronic Nature of Ketone Directing Group as a Key To Control C-2 vs C-4 Alkenylation of Indoles. *Organic letters* **2016,** *18* (21), 5496-5499.

15. Demopoulos, V. J.; Nicolaou, I., Electrophilic Substitution of Indole on the Benzene Moiety: A Synthesis of 5-Acyl- and 5-Aroylindoles. *Synthesis* **1998,** *1998* (10), 1519-1522.

16. Feng, Y.; Holte, D.; Zoller, J.; Umemiya, S.; Simke, L. R.; Baran, P. S., Total Synthesis of Verruculogen and Fumitremorgin A Enabled by Ligand-Controlled C–H Borylation. *Journal of the American Chemical Society* **2015,** *137* (32), 10160-10163.

17. Moir, M.; Danon, J. J.; Reekie, T. A.; Kassiou, M., An overview of late-stage functionalization in today's drug discovery. *Expert Opinion on Drug Discovery* **2019,** *14* (11), 1137-1149.

18. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs from 1981 to 2014. *J Nat Prod* **2016,** *79* (3), 629-61.

19. Mazák, K.; Hosztafi, S.; Kraszni, M.; Noszál, B., Physico-chemical profiling of semisynthetic opioids. *Journal of Pharmaceutical and Biomedical Analysis* **2017,** *135*, 97-105.

20. Murphy, B.; Šnajdr, I.; Machara, A.; Endoma-Arias, M. A. A.; Stamatatos, T. C.; Cox, D. P.; Hudlický, T., Conversion of Thebaine to Oripavine and Other Useful Intermediates for the Semisynthesis of Opiate-Derived Agents: Synthesis of Hydromorphone. *Advanced Synthesis & Catalysis* **2014,** *356* (11-12), 2679-2687.

21. Truppo, M. D., Biocatalysis in the Pharmaceutical Industry: The Need for Speed. *ACS Med Chem Lett* **2017,** *8* (5), 476-480.

22. Adrio, J. L.; Demain, A. L., Microbial enzymes: tools for biotechnological processes. *Biomolecules* **2014,** *4* (1), 117-139.

23. Phillips, R. S., Synthetic applications of tryptophan synthase. *Tetrahedron: Asymmetry* **2004,** *15* (18), 2787-2792.

24. Boville, C. E.; Scheele, R. A.; Koch, P.; Brinkmann-Chen, S.; Buller, A. R.; Arnold, F. H., Engineered Biosynthesis of β-Alkyl Tryptophan Analogues. *Angewandte Chemie International Edition* **2018,** *57* (45), 14764-14768.

25. Romney, D. K.; Sarai, N. S.; Arnold, F. H., Nitroalkanes as Versatile Nucleophiles for Enzymatic Synthesis of Noncanonical Amino Acids. *Acs Catal* **2019,** *9* (9), 8726-8730.

26. Dick, M.; Sarai, N. S.; Martynowycz, M. W.; Gonen, T.; Arnold, F. H., Tailoring Tryptophan Synthase TrpB for Selective Quaternary Carbon Bond Formation. *Journal of the American Chemical Society* **2019,** *141* (50), 19817-19822.

27. Moritzer, A.-C.; Minges, H.; Prior, T.; Frese, M.; Sewald, N.; Niemann, H. H., Structure-based switch of regioselectivity in the flavin-dependent tryptophan 6-halogenase Thal. *Journal of Biological Chemistry* **2019,** *294* (7), 2529-2542.

28. Gruß, H.; Belu, C.; Bernhard, L. M.; Merschel, A.; Sewald, N., Fluorogenic Diversification of Unprotected Bromotryptophan by Aqueous Mizoroki–Heck Cross-Coupling. *Chemistry – A European Journal* **2019,** *25* (23), 5880-5883.

29. Andorfer, M. C.; Park, H. J.; Vergara-Coll, J.; Lewis, J. C., Directed evolution of RebH for catalyst-controlled halogenation of indole C–H bonds. *Chem Sci* **2016,** *7* (6), 3720-3729.

30. Payne, J. T.; Poor, C. B.; Lewis, J. C., Directed Evolution of RebH for Site-Selective Halogenation of Large Biologically Active Molecules. *Angewandte Chemie International Edition* **2015,** *54* (14), 4226-4230.

31. Zuo, R.; Ding, Y., Direct Aromatic Nitration System for Synthesis of Nitrotryptophans in Escherichia coli. *ACS Synthetic Biology* **2019,** *8* (4), 857-865.

32. Romney, D. K.; Murciano-Calles, J.; Wehrmüller, J. E.; Arnold, F. H., Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogues. *Journal of the American Chemical Society* **2017,** *139* (31), 10769-10776.

33. Tanner, M. E., Mechanistic studies on the indole prenyltransferases. *Natural Product Reports* **2015,** *32* (1), 88-101.

34. Schuller, J. M.; Zocher, G.; Liebhold, M.; Xie, X.; Stahl, M.; Li, S.-M.; Stehle, T., Structure and Catalytic Mechanism of a Cyclic Dipeptide Prenyltransferase with Broad Substrate Promiscuity. *Journal of Molecular Biology* **2012,** *422* (1), 87-99.

35. Winkelblech, J.; Fan, A.; Li, S.-M., Prenyltransferases as key enzymes in primary and secondary metabolism. *Applied Microbiology and Biotechnology* **2015,** *99* (18), 7379-7397.

36. Bandari, C.; Scull, E. M.; Bavineni, T.; Nimmo, S. L.; Gardner, E. D.; Bensen, R. C.; Burgett, A. W.; Singh, S., FgaPT2, a biocatalytic tool for alkyl-diversification of indole natural products. *MedChemComm* **2019**.

37. Zou, H.; Zheng, X.; Li, S.-M., Substrate Promiscuity of the Cyclic Dipeptide Prenyltransferases from Aspergillus fumigatus. *Journal of Natural Products* **2009,** *72* (1), 44-52.

38. Yin, W. B.; Ruan, H. L.; Westrich, L.; Grundmann, A.; Li, S. M., CdpNPT, an N-prenyltransferase from Aspergillus fumigatus: overproduction, purification and biochemical characterisation. *Chembiochem* **2007,** *8* (10), 1154-61.

39. Elshahawi, S. I.; Cao, H.; Shaaban, K. A.; Ponomareva, L. V.; Subramanian, T.; Farman, M. L.; Spielmann, H. P.; Phillips Jr, G. N.; Thorson, J. S.; Singh, S., Structure and specificity of a permissive bacterial C-prenyltransferase. *Nat Chem Biol* **2017,** *13* (4), 366-368.

40. Bandari, C.; Scull, E. M.; Masterson, J. M.; Tran, R. H. Q.; Foster, S. B.; Nicholas, K. M.; Singh, S., Determination of Alkyl-Donor Promiscuity of Tyrosine-O-Prenyltransferase SirD from Leptosphaeria maculans. *ChemBioChem* **2017,** *18* (23), 2323-2327.

41. Scull, E. M.; Bandari, C.; Johnson, B. P.; Gardner, E. D.; Tonelli, M.; You, J.; Cichewicz, R. H.; Singh, S., Chemoenzymatic synthesis of daptomycin analogs active against daptomycin-resistant strains. *Applied Microbiology and Biotechnology* **2020**.

Chapter 2. Indole C6 Functionalization of Tryprostatin B Using Promiscuous Prenyltransferase CdpNPT

Abstract

Tryprostatin A and B are prenylated, tryptophan-containing, diketopiperazine natural products, displaying cytotoxic activity through different mechanisms of action. The presence of the 6-methoxy substituent on the indole moiety of tryprostatin A (TPS-A) was shown to be essential for dual inhibition of topoisomerase II and tubulin polymerization. However, the inability to perform late-stage modification of the indole ring has limited the structure-activity relationship studies of this class of natural products. Herein, we describe an efficient chemoenzymatic approach for the late-stage modification of tryprostatin B (TPS-B) using an indole prenyltransferase, CdpNPT from *Aspergillus fumigatus* expressed in recombinant *Escherichia coli*, that generates novel analogs functionalized with allylic, benzylic, heterobenzylic, and diene moieties. Notably, this biocatalytic functionalization revealed high selectivity for the indole C6 position, demonstrating that prenyltransferase-based late-stage diversification enables direct access to previously inaccessible natural product analogs.

Allocation of Contribution

The following chapter is reproduced in part from "Indole C6 Functionalization of Tryprostatin B Using Prenyltransferase CdpNPT." **Gardner, E. D.**; Dimas, D. A.; Finneran, M. C.; Brown, S. M.; Burgett, A. W.; Singh, S., Catalysts 2020, 10 (11), 1247, (DOI: 10.3390/catal10111247).

I produced the results presented in this chapter with the following exceptions. Mr. Matthew C. Finneran from the Burgett research group at University of Oklahoma conducted the cytotoxicity assays described in **Section 2.2.9** and compiled the data to produce **Figure 2.8**. Mr. Dustin Dimas contributed to the total synthesis and purification of Tryprostatin B in addition to the production and purification of CdpNPT utilized in chemoenzymatic scale-up reactions. Ms. Sara Brown contributed to this work by purifying a portion of CdpNPT used for scale-up reactions, and purifying **TPS-61** via RP-HPLC.

2.1 Introduction

Indole diketopiperazines (DKPs) represent a large class of biologically active natural products (NPs), isolated predominately from fungi, [1, 2] with members exhibiting a wide variety of biological activities such as anticancer (tryprostatin, malbrancheamide),[3] herbicidal (thaxtomin A),[4] serotonergic (barettin),[5] antimicrobial (sclerotiamide),[6] antioxidant (neoechinulin A),[7] and immunomodulatory (cristatin A)[8] activities (Figure 2.1). These DKPs are biosynthesized through the cyclization of tryptophan with other amino acids (proline, leucine, histidine, phenylalanine, etc.), followed by modification by enzymes such as aromatic prenyltransferases (PTs), cytochrome P450s, and halogenases that build the final NP scaffold. Among the DKPs, tryprostatins (TPS-A and TPS-B, Figure 1) are Trp-Pro cyclic dipeptides with an indole C2-prenyl group, originally isolated from *Aspergillus fumigatus* BM 939.[9] The distinguishing feature between the two compounds being the presence (TPS-A) or absence (TPS-B) of a methoxy substitution at the indole C6-position. Both compounds are known to display anticancer activities as a result of the indole C2 prenyl group, which has been linked to their inhibition of topoisomerase II.[10] However, TPS-A has also been shown to inhibit tubulin polymerization and breast cancer resistance protein (BCRP, ABCG2),[11] and effective inhibition of BCRP by TPS-A has been shown to restore the efficacy of clinically used chemotherapeutics when tested on BCRP+ breast cancer cell lines.[12] The proposed mechanism of action (MOA), where TPS-A induces tubulin inhibition without directly binding to tubulin, was found to be distinct from the typical tubulin inhibitors such as paclitaxel or vinblastine.[9] Thus, the 6-methoxy substituent of TPS-A appears to enable binding to a new range of biological targets, highlighting the importance of the C6 indole functionalization of the TPS scaffold.

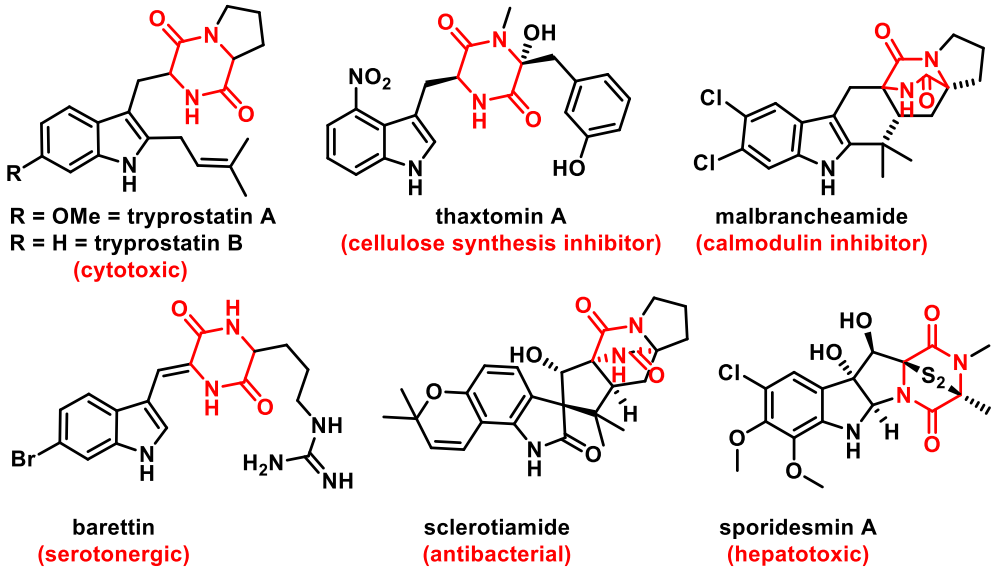


Figure 2.1: Representative bioactive indole diketopiperazines with diverse activities.

However, the developed synthetic methods to modify indole DKPs thus far have been largely focused on indole functionalization at the C2, C3 and N1 positions.[11] Among the limited methods to functionalize the benzene portion of indole, functionalization at C4 and C7 typically rely on coordination of the catalyst to specific directing groups installed on either C3 or N1, respectively.[13-15] This method however, suffers from low functional group tolerance and limited substrate scope.[14-22] The few described methods for C5 and C6 indole functionalization provide limited functional group tolerance.[19, 23] Recently, Baran et al. developed a remarkable example of iridium-catalyzed ligand-controlled selective C6 tryptophan borylation, however, even this groundbreaking directing group-free method requires protecting groups and is unable to tolerate tryptophan containing cyclic dipeptides.[24] Our aim is to address this urgent need to develop methodology capable of diversifying the unreactive positions of the indole ring of these complex NP scaffolds to access a massive untapped chemical space.

Within this context, aromatic PTs show great promise as biocatalysts both in a general sense and more specifically as tools for indole modification. They demonstrate broad substrate specificities for both alkyl-donors and acceptors, transferring a variety of allylic, benzylic, diene, and heterocyclic groups from their corresponding pyrophosphates on to both native and non-native acceptors.[25-31] Specific PTs have also been demonstrated to possess practical utility beyond generalized promiscuity. For example, the naphterpin PT NphB has been shown to N-alkylate the antibacterial sulfabenzamide,[26] and the indole PT CdpNPT has been demonstrated to modify the indole N1, C2, C3, C5- and C6-positions of the macrocyclic peptide antibiotic daptomycin, resulting in alkylated analogs with potency against daptomycin-resistant bacterial strains.[28, 32] Thus, the increased activity of the novel daptomycin analogs has provided strong evidence of the PTs' utility in drug design, while the alkylation of the indole C5- and C6-positions by CdpNPT demonstrates synthetic utility in modifying chemically inert benzenoid positions on indole rings.

Based on these findings and the demonstrated importance of indole C6 functionalization on the TPS scaffold, we hypothesized that CdpNPT could be used to generate benzenoid-functionalized TPS-analogs from TPS-B. Herein, we report the chemoenzymatic synthesis of TPS-B analogs using a library of synthetic alkyl donors, which revealed CdpNPT's ability to transfer several allylic, dienyl, benzylic and heterocylic groups predominantly onto C6 of TPS-B with high regioselectivity. Furthermore, the new TPS analogs were assayed for cytotoxic activity against the human leukemia cancer cell line K562. Modification of the indole C6 position with alkyl, dienyl, and benzylic groups did not substantially change the cytotoxicity activity of tryprostatins, which provides additional SAR information for indole DKPs. This study demonstrates the

utility of aromatic PTs as a tool for functionalization of inert positions of indole DKPs, providing access to previously inaccessible natural product analogs.

2.2 Materials and Methods

2.2.1 General Materials

Unless otherwise stated, all chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA), Acros (New Jersey, USA), Alfa-Aesar (Ward Hill, MA, USA), Ambeed (Arlington Heights, IL, USA), or AK Scientific (Union City, CA, USA) and were of reagent grade or better. The PD-10 column and Ni-NTA columns used for protein purification were purchased from GE Healthcare (Piscataway, NJ).

2.2.2 General Methods

All synthetic reactions were conducted in oven-dried glassware under a nitrogen atmosphere with anhydrous solvents, unless otherwise noted. Flash column chromatography was performed with silica gel (SiliCycle Inc, P60, particle size 40-63 μm). HPLC was conducted using an Agilent 1220 system equipped with a diode array detector.

2.2.3 In vitro CdpNPT Assay

The recombinant CdpNPT was overproduced in *Escherichia Coli* BL21(DE3) cells transformed with the codon optimized synthetic CdpNPT gene in a pET28a vector. The resulting N-terminal His_6-fusion protein was purified to homogeneity via Ni-NTA affinity chromatography as previously reported.[32] Analytical scale *in vitro* CdpNPT reactions were conducted on a 20 μL scale containing 10 μM CdpNPT, 1 mM TPS-B, 1.5 mM alkyl-PP donor, 25 mM Tris pH 7.8, and 5 mM $CaCl_2$ and were incubated at 35°C for 18 hours. Reactions were quenched by the addition of 40 μL cold methanol followed by centrifugation (9000 x g for 30 min) to remove precipitated protein. Product formation was

then analyzed via RP-HPLC using Method A. Negative controls lacking enzyme or donor were included in the screening, and the percent conversion for each reaction was calculated using the peak integration at 280 nm. Specifically, the integrated total product areas were divided by the sum of the areas of integrated products and remaining substrate.

2.2.4 Enzymatic Scale-up Reactions

All scale up reactions were performed on a 10 μM or 15 μM scale in a 50 mL falcon tube containing 1 mM TPS-B, 1.5 mM alkyl-PP donor, 25 mM Tris, and 5 mM $CaCl_2$ at pH 7.8. TPS-B was added as a 10 mM solution in DMSO to give a final DMSO content of 10%. CdpNPT was added for an initial concentration of 10 μM. Reaction mixtures were incubated at 35°C and progress was monitored via RP-HPLC using HPLC Method A. When low product formation was observed, additional enzyme was added, and incubation was continued for another 8-10 hours. Once completed, reaction products were extracted 4 times with 15 mL volumes of ethyl acetate, which were then combined and concentrated in vacuo. Individual products were purified from the concentrates using semi-preparative RP-HPLC with HPLC Method B.

2.2.5 Determination of Structure

High-resolution mass spectrometry (HRMS) confirmation of products was conducted on an Agilent 6545-QTOF W/ 1290 HPLC mass spectrometer at the University of Oklahoma Department of Chemistry and Biochemistry. NMR spectra were obtained on a Varian VNMRS 500 MHz or a Varian INOVA 600 MHz instrument at the NMR facility of the Department of Chemistry and Biochemistry of the University of Oklahoma using 99.8 % $CDCl_3$ (Cambridge Isotope Laboratories, MA, USA). 1H and 13C chemical shifts were referenced to internal solvent resonances. Peak multiplicities were designated by s

(singlet), d (doublet), t (triplet), q (quartet), m (multiplet), and br (broad), and chemical shifts were reported in parts per million (ppm). All NMR spectra were recorded at ambient temperature and processed using MestReNova software.

2.2.6 HPLC Method A

Analytical scale reactions were analyzed via RP-HPLC employing a Gemini-NX C-18 (5 µm, 4.6 mm x 250 mm) column (Phenomenex, Torrance, CA, USA) [gradient of 25% B to 100% B over 25 min, 100% B for 3 min, 100% B to 25% B over 0.1 min, 25% B for 7 min (A = ddH_2O with 0.1% TFA; B = acetonitrile); flow rate = 1 mL min-1; A_{280}].

2.2.7 HPLC Method B

Semi-preparative RP-HPLC was conducted using a Gemini-NX C18 (5 µm, 10 mm x 250 mm) column (Phenomenex, Torrance, CA, USA) to purify scaled up TPS analogs as well as TPS-A. [gradient of 25% B to 40% B over 3 min, 40% B to 100% B over 22 min, 100% B for 3 min, 100% B to 25% B in 0.1 min, 25% B for 7 min (A = ddH_2O with 0.1% formic acid; B = acetonitrile); flow rate = 2 mL min-1; A_{280}].

2.2.8 Synthesis of Tryprostatin A and B

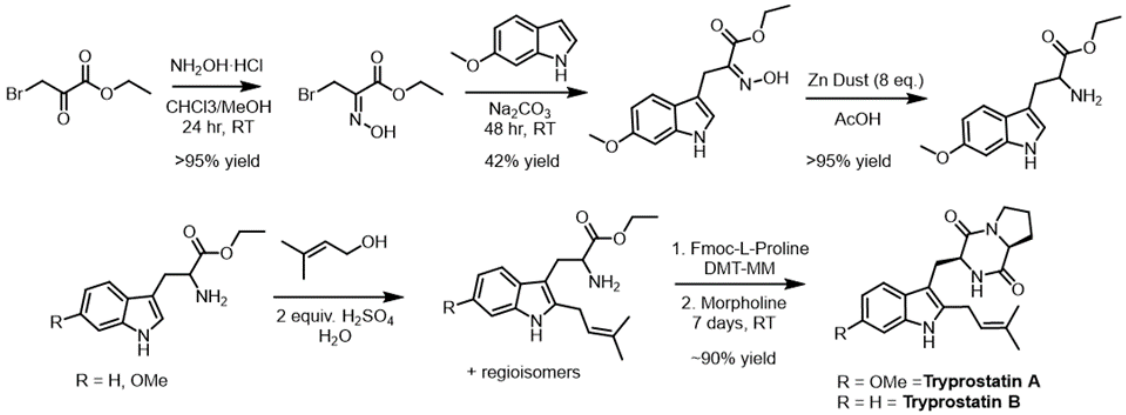


Figure 2.2: Synthetic Schemes for the synthesis of 6-methoxytryptophan, TPS-A and TPS-B.

Racemic 6-methoxy-tryptophan was prepared by the known reaction of ethyl bromo pyruvate oxime with 6-methoxy-indole followed by reduction via zinc dust in acetic acid.[33] TPS-A and TPS-B were both synthesized using a previously reported biomimetic prenylation reaction of their corresponding tryptophan ethyl esters, followed by DMT-MM mediated proline coupling and cyclization (Figure 2.2).[34]

2.2.9 Cell Titer-Blue Viability Assay

A cell titer-blue viability assay was performed in triplicate to evaluate TPS analog cytotoxicity. K-562 cells were cultured in RPMI Media (Thermo 22400105) containing 10% Hyclone (Fischer Sci SH3006603) and 1% penicillin-streptomycin (Thermo 15140122). Cells were incubated at 37 °C in 5.0% CO_2. Cells were seeded out to 2,000 cells per well in 75 µL of media within a 96-well plate. Taxol and Tryprostatin compounds were serial diluted in media to a 4X concentration, and 25 µL of each compound-containing solution were subsequently added to the 75 µL of cells to make a 1X concentration. Cells were incubated with compounds for 48 h at 37 °C in 5.0% CO_2, followed by a 1.5 h incubation with Cell Titer-Blue. Fluorescence was measured using the Promega GloMax microplate reader at an excitation wavelength of 544 nm and an emission wavelength of 590 nm. Cell growth was normalized to untreated cells, and GI_{50} values were calculated using the four-parameter-dose-response curve within GraphPad Prism software.

2.3 Results and Discussion

2.3.1 Design of alkyl pyrophosphate donors

The alkyl-pyrophosphate (alkyl-PP) donor library developed for the present study is shown in Figure 2.3, and each compound was designed to represent steric and electronic configuration within the established requirements of PT donor substrates. To undergo

successful electrophilic aromatic substitution in a PT reaction, the pyrophosphate (OPP) of the alkyl-PP donor must dissociate to generate a stabilized carbocation.[35] Benzylic and allylic carbocations are stabilized through resonance and inductive effects to promote OPP dissociation, with electron donating substituents further increasing the reactivity.[36-38] As such, both allylic (1-44) and benzylic (45-60) analogs were well-represented in the final library. Within the allylic alkyl-PPs, structural and electronic differences were established through variations in carbon chain length, branching patterns, and the inclusion of heteroatoms such as nitrogen, oxygen, chlorine, and bromine. The benzylic analogs were distinguished through differences in the location, extent, and electron donation/withdrawal potential of various substituents on the benzene ring. Additionally, a subset of the alkyl-PPs were heterocyclic in nature (61-66), maintaining the required allylic configuration while varying in the identity of the heteroatom, the size of the ring system, and the location of allylic attachment. The final library of 66 synthetic alkyl-PP donors (Figure 2.3) [25-28] was then used to determine the substrate scope of CdpNPT catalyzed TPS-B functionalization of TPS-B.

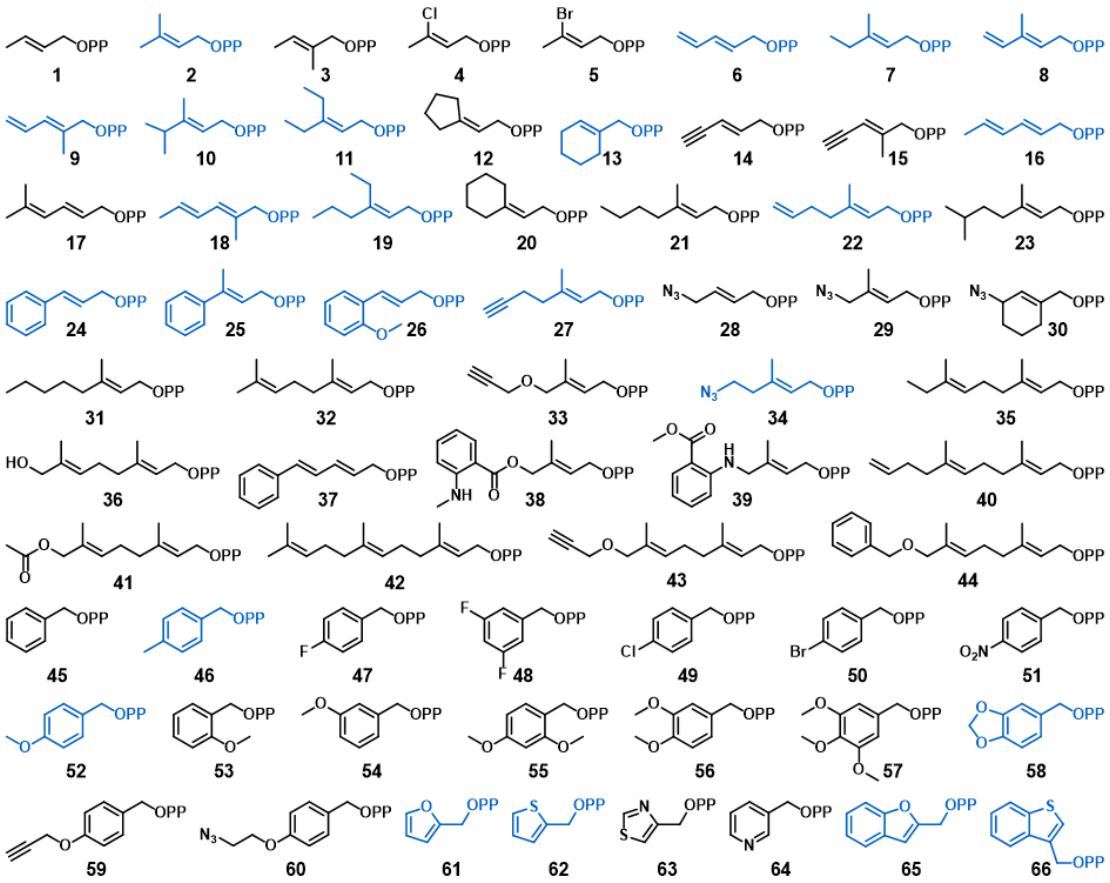


Figure 2.3: Alkyl-PP donor library used for CdpNPT catalyzed TPS-B Modification
Alkyl-PP donors that were transferred to TPS-B are highlighted in blue. Synthetic methods for Alkyl-PP library preperation were reported previously.[25, 27, 28]

2.3.2 Donor Library Screening

Standard uniform assay conditions (1.5 mM alkyl-PP analog, 1 mM TPS-B, 10 µM CdpNPT, 25 mM Tris, 5 mM $CaCl_2$, pH 7.8, 18 h at 35 °C) were adopted to facilitate an initial assessment of CdpNPT's donor scope with TPS-B as the acceptor. Once reactions were complete, the protein was precipitated with cold methanol and removed by centrifugation. The clarified reactions were then analyzed using a reverse phase high-pressure liquid chromatography (RP-HPLC) endpoint assay, and the production of alkyl-TPS derivatives was determined by the difference in retention between TPS-B and alkyl-

TPS. All positive reactions underwent subsequent confirmation by high-resolution electrospray ionization mass spectra (ESI-HRMS) (Appendix 4: HRMS Data).

The initial assessment of CdpNPT catalyzed reactions revealed 24 out of the 66 screened alkyl-PP donors produced new products in the presence of CdpNPT and TPS-B (Figure 2.4 and Table 2.2). Ten of these positive reactions showed >50% conversion (7, 24, 58, 65, 2, 10, 11, 52, 46, 25), 8 displayed moderate conversion (20-50%; 62, 16, 8, 9, 13, 26, 66, 6) and an additional 6 reactions produced low conversion (<20%; 61, 19, 18, 34, 27, 22).

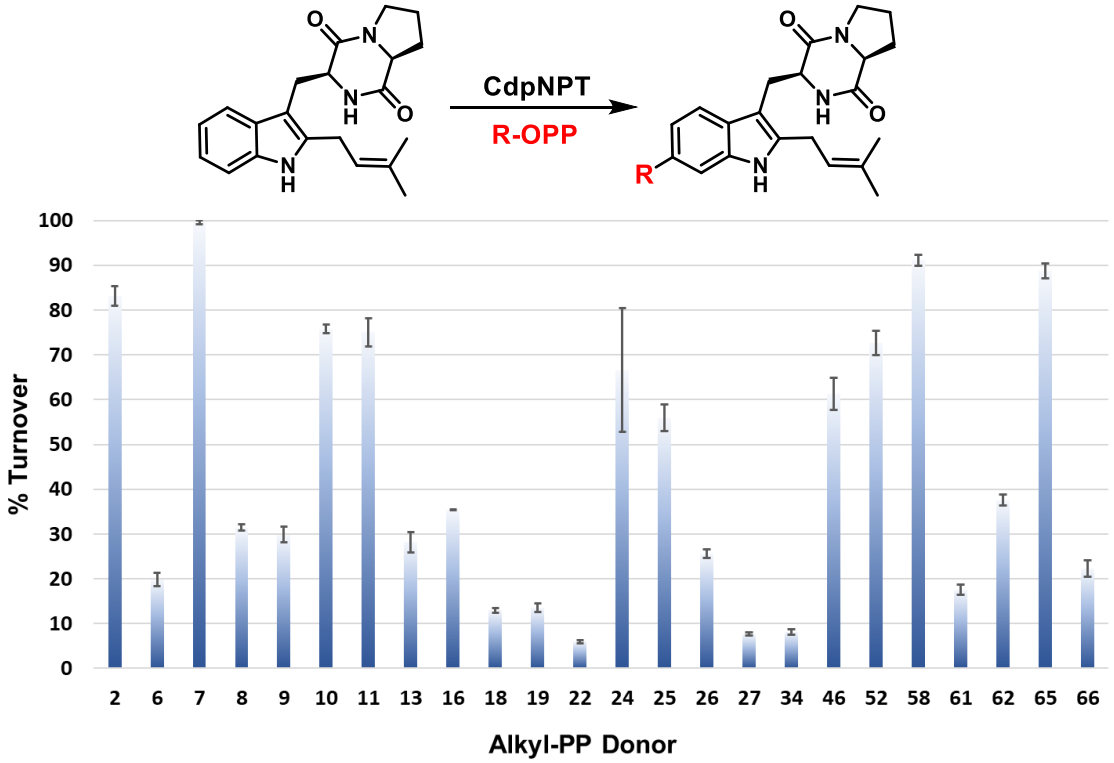


Figure 2.4: Alkyl donor profile of analytical-scale CdpNPT-catalyzed reactions with TPS-B.
The average percent conversion of all positive reactions is shown with the associated standard deviations (n=2) as determined by RP-HPLC at 280 nm. Each reaction was carried out in a 20-µL volume and contained 1.5 mM alkyl-PP, 1 mM TPS-B, and 10 µM purified CdpNPT in a reaction buffer (25 mM Tris pH 7.8, and 5 mM CaCl2) incubated at 35 °C for 18 h. No product formation was observed in the absence of CdpNPT or alkyl-PP.

2.3.3 Allylic Donors

Among the allylic alkyl-PP analogs, DMAPP (**2**) represents the natural donor for CdpNPT-catalyzed reactions, in which it forms a highly stabilized carbocation intermediate through the delocalization of its positive charge between primary and tertiary carbons (Figure 2.5). Thus, alkyl-PP analogs with similar size and planarity to 2 typically gave comparable conversion (**7, 10, 11**). However, longer-chain analogs **19, 22, 27**, and **34** were taken less efficiently, while the 10-carbon geranyl pyrophosphate (**32**) and the 15-carbon farnesyl pyrophosphate (**42**) generated no product. Therefore, the yields of product formation of the CdpNPT catalyzed reactions with TPS-B in general decreased with the increase in the chain length of the alkyl-PP analog. Based on the observed trends, the maximum donor length accepted by CdpNPT was around seven atoms from the pyrophosphate in the linear direction. Further, the conjugated aromatic donors, **24** and **25** were well accepted by CdpNPT, likely due to the formation of a delocalized carbocation and the overall planarity of the substrate. This high turnover is notable considering **24** and **25** are seven carbons long, which is the maximum tolerated size observed in the allylic donors. Thus, these results support the hypothesis that, while the capacity of a donor to delocalize the carbocation intermediate is critical for their efficient utility, steric constraints within the active site play a major role in their acceptability by CdpNPT.

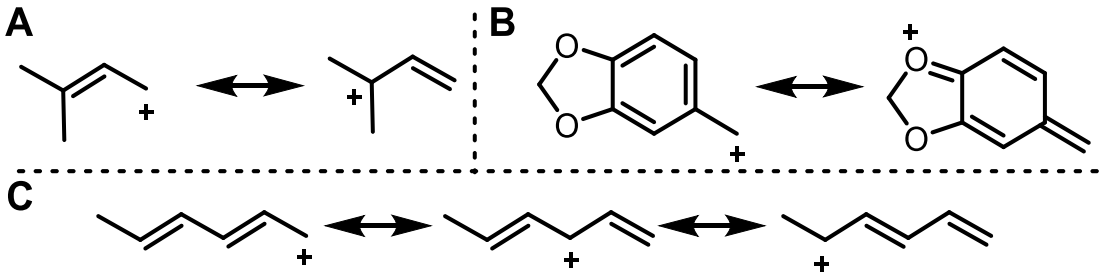


Figure 2.5: Depiction of the resonance stabilization of representitative donor moieties.

2.3.4 Benzylic and Heterocyclic Donors

The carbocation stabilization on benzylic pyrophosphates takes place as a combination of resonance and inductive effects. While CdpNPT did not accept the *ortho*- and *meta*-substituted benzyl pyrophosphates (**48, 53, 54**), it was very efficient in transferring benzylic donors featuring electron donating groups at the *para* position, including the methyl (**46**) and methoxy (**52**) groups. Notably, the unsubstituted benzyl pyrophosphate **45** did not generate a detectable product, suggesting, carbocation stabilizing substituents are necessary for CdpNPT to transfer benzylic moieties. Even more interesting was piperonyl pyrophosphate, (**58**) which produced the highest product yield among the aromatic (benzylic or heterocyclic) donors. This is even more surprising given that the di-methoxy analog **56** showed no discernable turnover. This phenomenon is likely due to steric differences between the two moieties. In both cases, the oxygen atom *para* to the benzyl tail provides similar resonance stabilization to the resulting carbocations (Figure 4B). The drastic difference in activity therefore likely results from the conformational constraints of the benzodioxole ring reducing the entropic barrier for binding when compared to the more flexible and bulkier methoxy groups of **56**. These observations again highlight the potential steric limitations of the binding pocket. Several heterocyclic analogs were also successfully transferred by CdpNPT such as furan (**61**), benzofuran (**65**), thiophene (**62**), and benzothiophene (**66**). The heteroatoms present in these donors can donate their lone pair of electrons to delocalize and thereby stabilize the positive charge more effectively than a simple benzyl carbocation.

2.3.5 Other Functionalized Donors

The synthetic utility of PT catalyzed late stage functionalization would be greatly enhanced if these enzymes efficiently transferred donors containing polar functional groups or other synthetic handles such as azides or alkynes. Azide or alkyne containing compounds are particularly desirable because they enable click chemistry-based derivatization. Of the alkyne- and azide-substituted donors, only **27** and **34** were accepted by CdpNPT, both of which showed poor conversions (<10%, Table 2.2). This was likely the result of their length, as well as the unnatural linear rigidity imparted by the alkyne or azide group. The intolerance of these rigid linear groups is further noted in the shorter chain alkyne and azide donors, **14**,**15**, **28** and **29**, which generated no detectable product under the standard reaction conditions. The low conversion exclusively shown in intermediate length alkyne/azide examples is likely due to the increased conformational flexibility provided by the additional methylene. The azide donor **34** is the only example of a donor exceeding seven atoms in the linear chain that generated products, and also the only accepted donor containing a polar functional group. These exceptions could possibly be a result of hydrogen bonding interactions occurring at the solvent-exposed edge of the donor binding pocket.

The diene containing donors **6**, **8**, **9**, and **16** were accepted by CdpNPT, though a lower degree of regioselectivity was observed. This decrease in regioselectivity could be a result of the multiple carbocation resonance forms and increased conformational flexibility compared to the other donor classes. (Figure 2.5) Overall, this screening study revealed a high degree of donor promiscuity for CdpNPT. HPLC analysis revealed the production of

a single product was observed in 17 (**10, 11, 13, 16, 19, 22, 24, 25, 26, 27, 34, 46, 52, 58, 65, 66**) of the 25 accepted substrates.

2.3.6 Scale-up and Characterization of CdpNPT Reaction Products

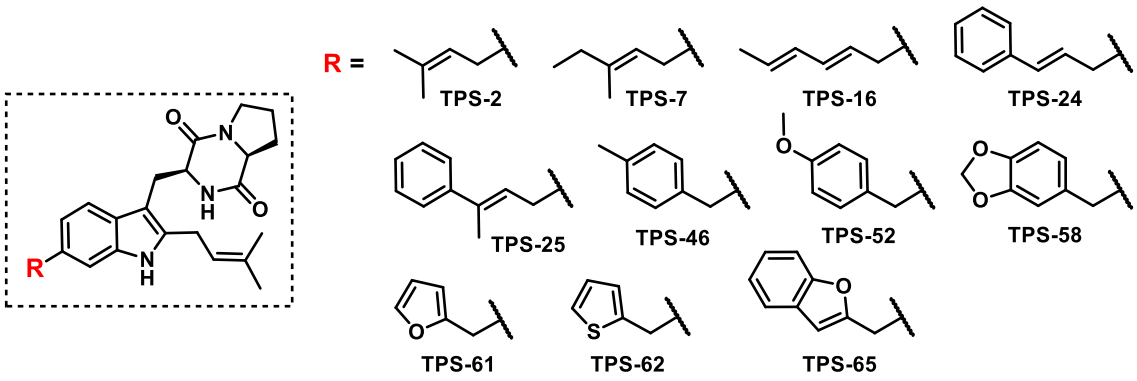


Figure 2.6: NMR-characterized structures of alkyl-TPS-B analogs
Where the number following TPS represent the corresponding alkyl-PP used in the reaction. Alkyl-PP analogs **2, 7, 61,** and **62** produced 10-20% C-5 alkylated product. The regioisomers for **TPS-2**, **TPS-7**, and **TPS-61** were separated by semi-preparative RP-HPLC, however **TPS-62** remained inseparable.

To assess the regiospecificity of CdpNPT catalyzed alkyl-transfer, 11 representative reactions displaying structural diversity and high product yield in the analytical scale reactions (**2, 7, 16, 24, 25, 46, 52, 58, 61, 62, 65)** were scaled up at 10-15 µmol scale using standardized conditions (Figure 2.6). The resulting alkyl-TPS-B analogs were first extracted using ethyl acetate and then purified via RP-HPLC. Product formation was confirmed via HRMS (Appendix 4: HRMS Data), and the regiospecificity was determined via ^{1}H and 2D NMR experiments acquired in $CDCl_3$. Diagnostic ^{1}H peaks for indole substitution patterns are well known in literature, and these spectra were useful for initial determination of each reaction's regiospecificity. Typically, the H4 is the most downfield indole CH signal at 7.4 ppm. For C6 alkylated products, the H4 peak appears as a doublet centered at this chemical shift, but C5 substituted products feature an H4 singlet shifted upfield due to the electron donating substituent (Appendix 3: NMR Assignments

and Spectra). COSY, HSQC, and HMBC correlations were used to unequivocally verify the position of indole alkylation (Figure 2.10), and our analyses revealed the dominant products to be indole C6 alkylated in every case, with the occasional minor product being the C5 regioisomer (**2**, **7**, **61**, **62**) as 10-20% of the reaction products (Figure 2.7). The donors that produced a mixture of regioisomers were **2**, **7**, the conjugated dienes (**6, 8**, **9**, **18**), and the five-membered heterocycles (**61**, **62**). The common features between the less regioselective analogs is their small size and/or flexibility. The thiophene (**62**) and furan (**61**) pyrophosphates could in theory bind in two different orientations, with the heteroatom facing either side of the binding pocket, due to the small size and planarity of the donor substrate. These two competing binding conformations explain the observed decrease of regiospecificity, generating the C-5 alkylated side product from the less favored conformation. Additionally, the smaller carbocations would have a greater ability to rotate before alkylation, decreasing the overall regiospecificity of their respective reactions. The larger, less flexible alkyl donors (**10, 11, 13**, **19, 24**, **27, 34**) all produced a single product likely due to the conformational constraints within the binding pocket. Based on the NMR confirmed products, all benzylic and cinnamyl donors (**25**, **26**, **46**, **52**, **58**, **65**, **66**) produced exclusively C6 alkylated products. (Appendix 3: NMR Assignments and Spectra)

2.3.7 Cytotoxicity Studies

A structurally diverse set of scaled up alkyl-TPS analogs with allylic, dienely, benzylic, and heterobenzylic substituants (products of alkyl-PPs **2**, **24**, **25**, **46**, **52**, **58**, **61**, **62**, **65**) were tested for cytotoxicity against human leukemia K562 cells. (Table 2.1) Surprisingly, most of the alkylated analogs had similar potency to the parent compound.

This remarkable tolerance of bulky substituents indicates that the diketopiperazine is perhaps the most dominant pharmacophore responsible for the observed cytotoxicity.

Compound	K562 GI_{50} (μM)
TPS-A	97 ± 21
TPS-B	78 ± 31
TPS-2	54 ± 15
TPS-24	50 ± 29
TPS-25	56 ± 19
TPS-46	37 ± 8.8
TPS-52	77 ± 22
TPS-58	44 ± 18
TPS-61	89 ± 29
TPS-62	100 ± 33
TPS-65	53 ± 19
Taxol	.0058 ± .0039

Table 2.1: Anticancer activity of TPS analogs against human leukemia K562 cells. The averaged GI_{50} values with associated standard deviation of compounds from two experimental replicates are shown.

2.4 Conclusions

This work has revealed the promiscuity of CdpNPT towards alkyl-donors, combined with the high regiospecificity of the resulting reactions makes CdpNPT a powerful biocatalyst for functionalizing the C6 position of TPS-B. Few synthetic methods are known for alkylating indole C6, and the existing methods are not suitable for late stage modification of natural products based on the complexity of the molecules. However, using prenyltransferase based late stage modification, we have developed a route to reach these previously inaccessible TPS-B analogs. Upon refining the substrate scope and regiospecificity of promiscuous PT's like CdpNPT, NphB, and FgaPT2, these powerful biocatalysts are destined to become popular catalysts for synthetic late stage modification and pharmaceutical production.

2.5 Appendix 1: HPLC Data

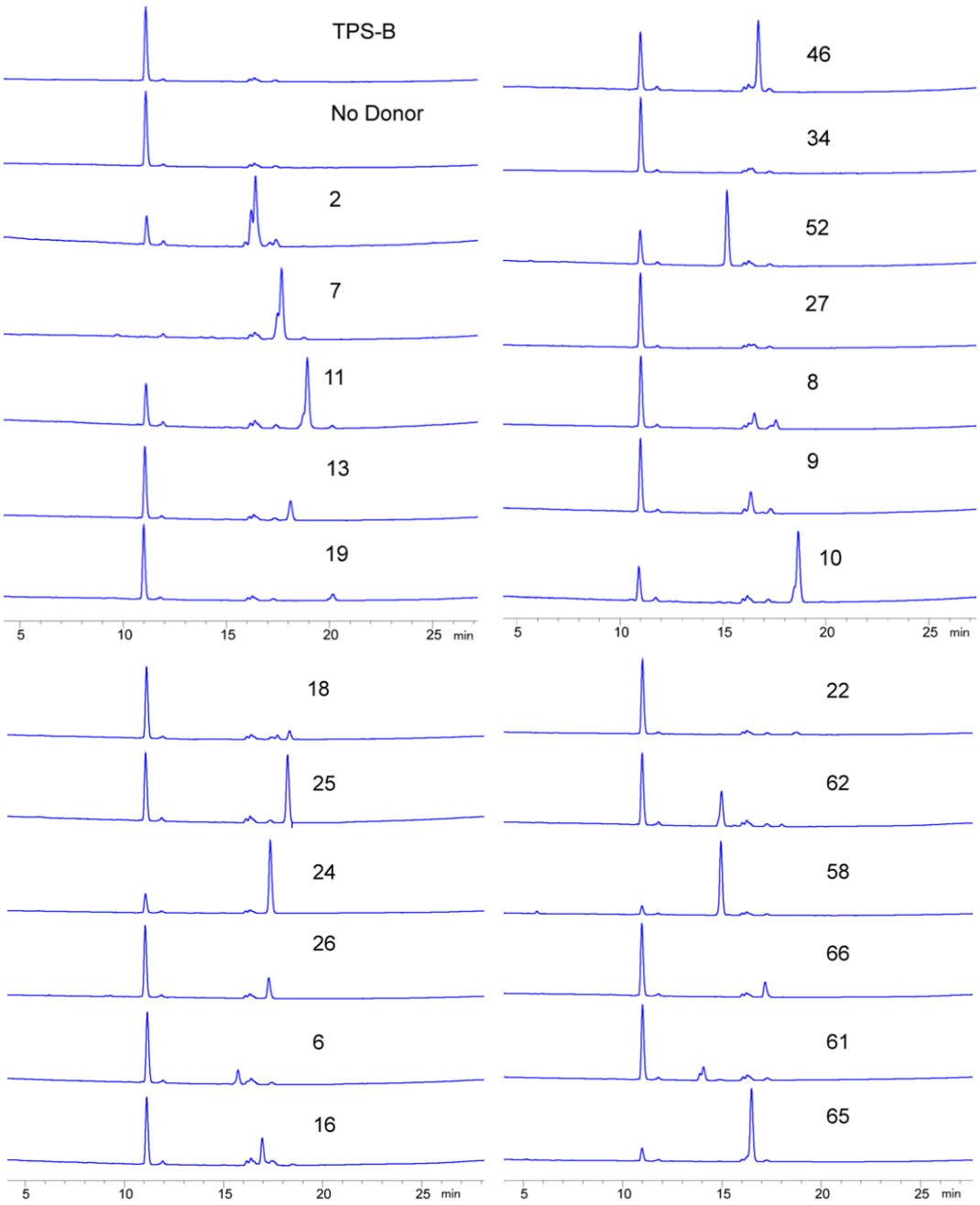


Figure 2.7: HPLC traces of analytical scale reactions containing TPS-B and CdpNPT Numerical labels indicate the alkyl-PP used in that reaction. TPS-B standard and enzymatic control reaction without the donor are indicated. The TPS-B used in these screenings is ~95% pure based on HPLC.

OPP #	Trial 1 % Turnover	Trial 2 % Turnover	Mean % Turnover	Std Dev.
2	80.93	85.40	83.17	2.23
6	18.32	21.34	19.83	1.51
7	99.15	99.95	99.55	0.40
8	30.72	32.12	31.42	0.70
9	28.13	31.64	29.88	1.76
10	74.84	76.86	75.85	1.01
11	71.88	78.25	75.06	3.19
13	25.83	30.50	28.16	2.33
16	35.52	35.28	35.40	0.12
18	12.45	13.51	12.98	0.53
19	12.63	14.44	13.54	0.90
22	5.60	6.34	5.97	0.37
24	80.40	52.77	66.59	13.81
25	52.97	58.86	55.91	2.95
26	24.63	26.63	25.63	1.00
27	7.36	8.00	7.68	0.32
34	7.45	8.79	8.12	0.67
46	57.66	64.90	61.28	3.62
52	69.94	75.40	72.67	2.73
58	89.96	92.27	91.11	1.15
61	16.45	18.73	17.59	1.14
62	36.29	38.85	37.57	1.28
65	87.13	90.40	88.76	1.64
66	20.38	24.08	22.23	1.85

Table 2.2: Calculated percent turnover determined via HPLC in duplicate
Percent turnover was calculated by dividing the sum of all TPS-B and product peak areas by the peak area of remaining TPS-B.

2.6 Appendix 2: Cytotoxicity Assay Data

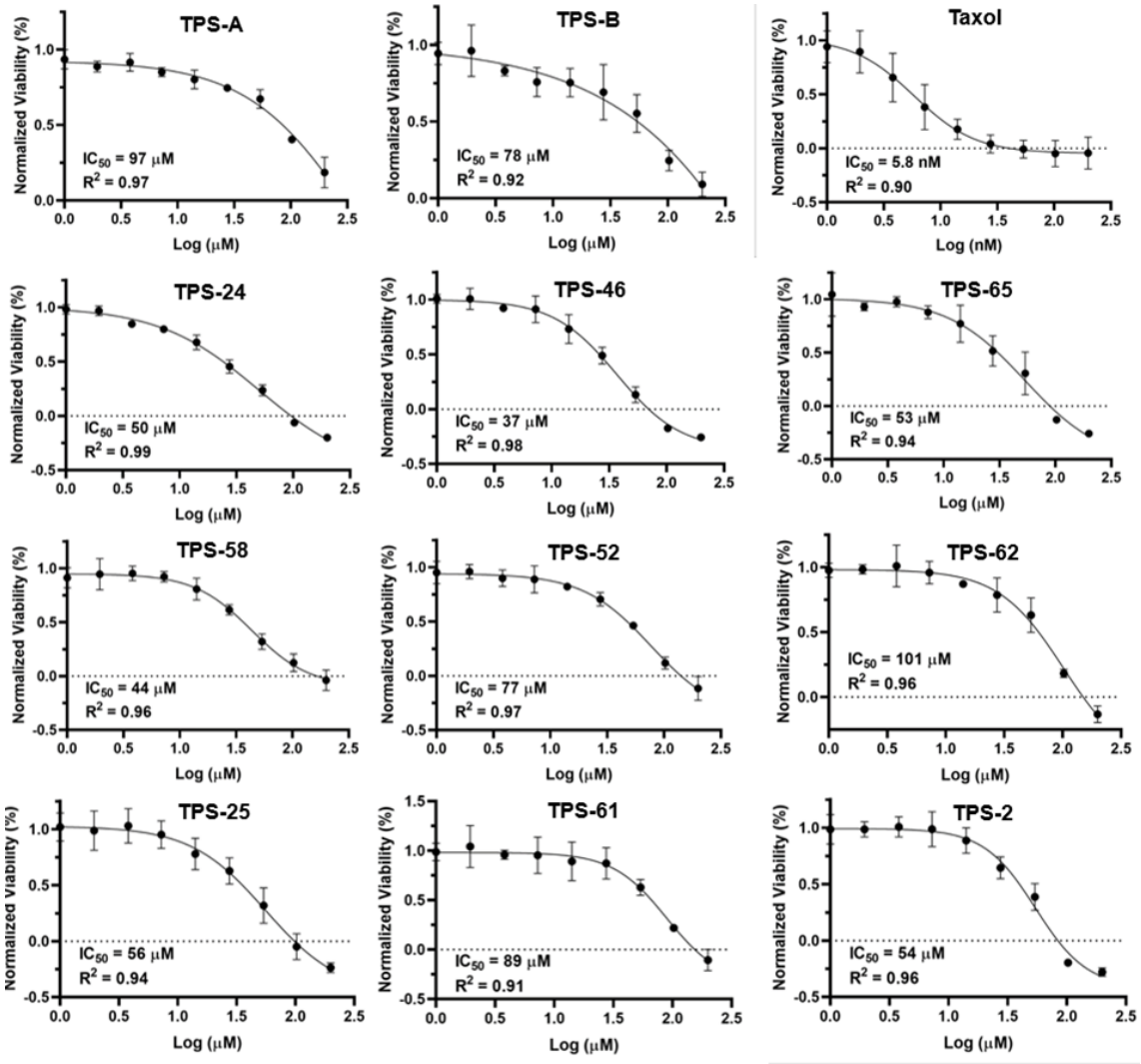
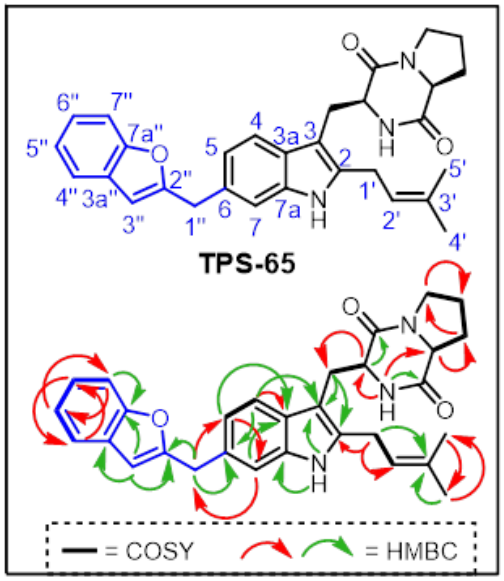


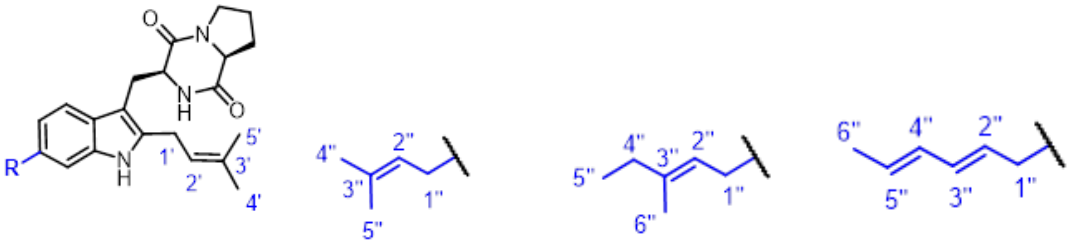
Figure 2.8: Cytotoxicity assay results for TPS analogs run in triplicate against K563 leukemia cells, with TPS-A, TPS-B, and Taxol as positive controls.

2.7 Appendix 3: NMR Assignments and Spectra



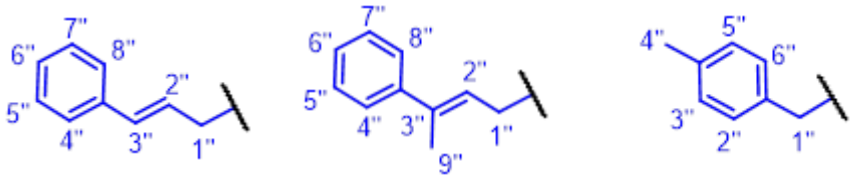
Segment	Position	δC (in ppm), type	δH (in ppm)
L-Trp	NH		5.62 (s)
	α	54.5, CH	4.33 (dd)
	β	25.6, CH_2	2.93 (m), 3.66 (d)
	1		7.98 (s)
	2	136.5, C	
	3	104.5, C	
	3a	126.8, C	
	4	117.9, CH	7.4 (d)
	5	121.3, CH	7.04 (d)
	6	131.0, C	
	7	111.1, CH	7.21 (s)
	7a	136.1, C	
	C=O		
Proline	α	59.3, CH	4.04 (t)
	β	28.4, CH_2	2.02 (m), 2.32 (dt)
	γ	22.7, CH_2	1.89 (m), 2.02 (m)
	δ	45.4, CH_2	3.58 (m), 3.65 (m)
	C=O	165.8, C	
C2-Prenyl	1'	25.1, CH_2	3.44 (qd)
	2'	119.8, CH	5.28 (m)
	3'	135.5, C	
	4'	18.0, CH_3	1.73 (s)
	5'	25.8, CH_3	1.76 (s)
C6-Sub	1''	35.2, CH_2	4.18 (s)
	2''	158.6, C	
	3''	103.3, CH	6.36 (s)
	3a''	128.9, C	
	4''	111.0, CH	7.38 (d)
	5''	123.4, CH	7.18 (td)
	6''	122.6, CH	7.15 (td)
	7''	120.5, CH	7.43 (d)
	7a''	155.0, C	

Figure 2.9: Representative NMR correlations used to assign TPS analogs



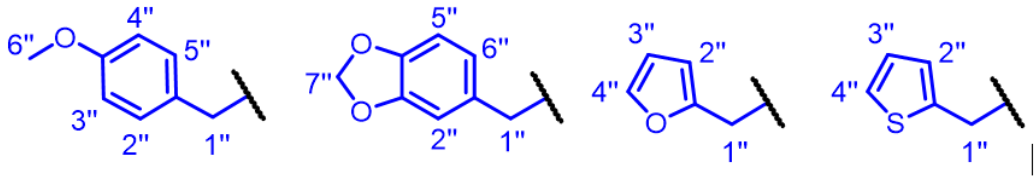
		TPS-2		TPS-7		TPS-16	
Segment	**Position**	**δC , type**	**δH**	**δC , type**	**δH**	**δC , type**	**δH**
L-Trp	NH		5.61 (s)		5.61 (s)		5.61 (s)
	α	54.6, CH	4.36 (m)	54.6, CH	4.34 (dd)	54.6, CH	4.35 (dd)
	β	25.5 CH_2	2.93 (dd), 3.66 (m)	25.6, CH_2	2.92 (dd), 3.64 (m)	25.7, CH_2	2.93 (dd), 3.65 (m)
	1		7.82 (s)		7.81 (s)		
	2						
	3						
	3a						
	4	117.7, CH	7.37 (d)	117.6, CH	7.37 (d)	117.6, CH	7.38 (d)
	5	121, CH	6.94 (d)	121.0, CH	6.94 (d)	121.1, CH	6.94 (d)
	6						
	7	110.1, CH	7.12 (s)	110.1, CH	7.11 (s)	110.5, CH	7.13
	7a						
	C=O						
Proline	α	59.3, CH	4.05 (t)	59.6, CH	4.05 (t)	59.3, CH	4.06 (t)
	β	28.4, CH_2	2.05 (m), 2.35 (m)	28.3, CH_2	2.04 (m), 2.33 (m)	28.4, CH_2	2.05 (m), 2.35 (m)
	γ	22.7, CH_2	1.92 (m), 2.05 (m)	22.6, CH_2	1.91 (m), 2.04 (m)	22.9, CH_2	1.92 (m), 2.05 (m)
	δ	45.5, CH_2	3.59 (m), 3.70 (m)	45.4, CH_2	3.59 (m), 3.67 (m)	45.4, CH_2	3.59 (m), 3.68 (m)
	C=O						
C2-Prenyl	1'	25.2, CH_2	3.46 (d)	25.1, CH_2	3.45 (m)	25.2, CH_2	3.46 (m)
	2'	119.8, CH	5.31 (t)	119.8, CH	5.29 (t)	119.6, CH	5.31 (t)
	3'						
	4'	17.8, CH_3	1.75 (s)	18.0, CH_3	1.74 (s)	18.0, CH_3	1.75 (s)
	5'	25.7, CH_3	1.76 (s)	25.7, CH_3	1.78 (s)	25.8, CH_3	1.79 (s)
C6-Sub	1"	36.5, CH_2	3.42 (d)	34.3, CH_2	3.44 (m)	39.1, CH_2	3.48 (m)
	2"	124.2, CH	5.37 (t)	122.3, CH	5.36	131.0, CH	5.72 (dt)
	3"					131.4, CH	6.06 (m)
	4"	17.8, CH_3	1.75 (s)	32.4, CH_2	2.05 (q)	131.4, CH	6.06 (m)
	5"	25.7, CH_3	1.79 (s)	12.7, CH_3	1..02 (t)	127.7, CH	5.63 (m)
	6"			16.0, CH_3	1.73 (s)	18.0, CH_3	1.73 (d)

Table 2.3: NMR assignments for TPS-2, TPS-7, and TPS-16



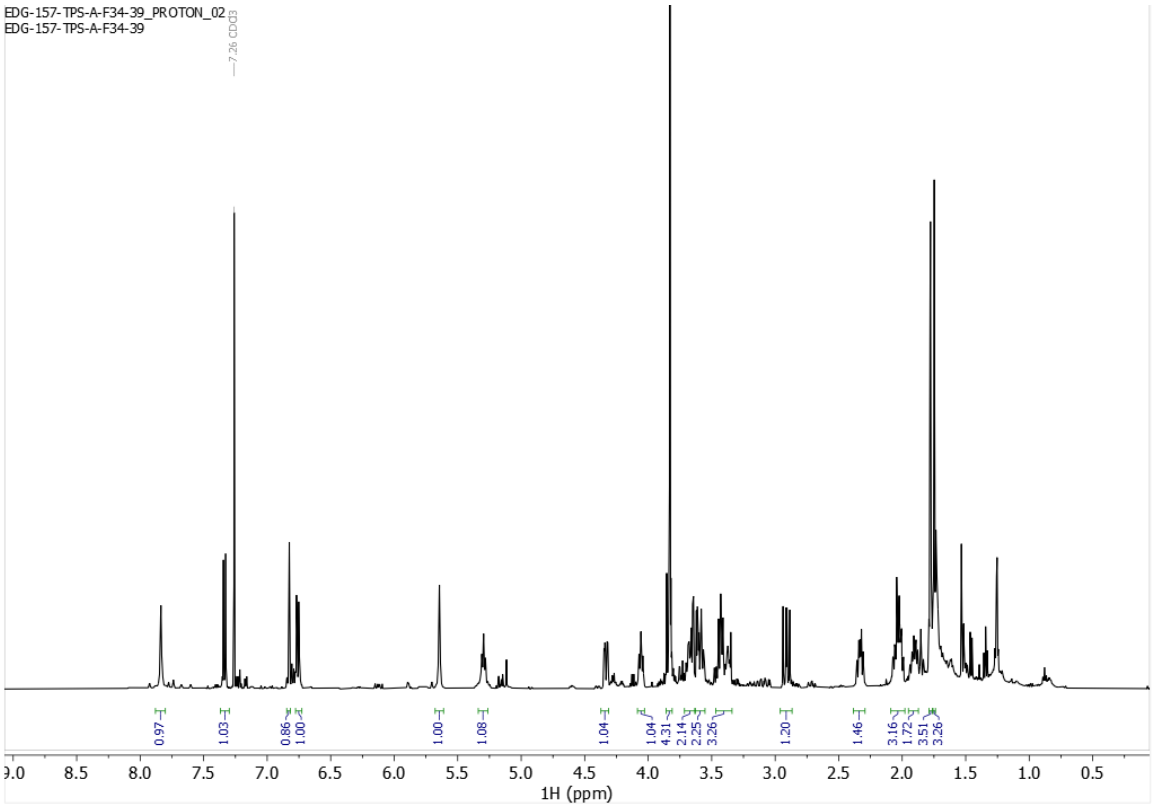
		TPS-24		TPS-25		TPS-46	
Segment	**Position**	**δC , type**	**δH**	**δC , type**	**δH**	**δC , type**	**δH**
L-Trp	NH		5.62 (s)		5.5 (s)		5.6 (s)
	α	54.5, CH	4.35 (dd)	54.7, CH	4.36 (dd)	54.5, CH	4.34 (dd)
	β	25.7, CH_2	2.92 (m), 3.66 (m)	25.7, CH_2	2.93 (dd), 3.64 (m)	25.7, CH_2	2.91 (m), 3.65 (m)
	1		7.95 (s)		7.83 (s)		7.87 (s)
	2	135.9, C				136.0, C	
	3	104.4, C				104.4, C	
	3a	126.5, C				126.3, C	
	4	117.5, CH	7.40 (d)	117.8, CH	7.39 (d)	117.6, CH	7.36 (d)
	5	120.9, CH	6.99 (d)	121.1, CH	6.99 (dd)	121.3, CH	6.95 (d)
	6	134.1, C				135.7, C	
	7	110.5, CH	7.17 (s)	110.3, CH	7.17 (s)	110.9, CH	7.06 (s)
	7a	135.9, C					
	C=O						
Proline	α	59.2, CH	4.04 (t)	59.3, CH	4.05 (t)	59.3, CH	4.04 (t)
	β	28.4, CH_2	2.02 (m), 2.32 (dt)	28.4, CH_2	2.03 (m), 2.33 (dt)	28.3, CH_2	2.02 (m), 2.32 (dt)
	γ	22.8, CH_2	1.89 (m), 2.02 (m)	22.7, CH_2	1.91 (m), 2.03 (m)	22.6, CH_2	1.90 (m), 2.02 (m)
	δ	45.4, CH_2	3.58 (m), 3.66 (m)	45.5, CH_2	3.58 (m), 3.67 (m)	45.4, CH_2	3.58 (m), 3.66 (m)
	C=O	165.9, C				165.9, C	
C2-Prenyl	1'	25.2, CH_2	3.44 (qd)	25.2, CH_2	3.45 (m)	25.0, CH_2	3.43 (qd)
	2'	119.6, CH	5.29 (t)	119.9, CH	5.29 (t)	119.8, CH	5.26 (t)
	3'	135.4, C				135.4, C	
	4'	18.0, CH_3	1.73 (s)	18.0, CH_3	1.74 (s)	17.9, C	1.72 (s)
	5'	25.8, CH_3	1.76 (s)	25.8, CH_3	1.78 (s)	25.7	1.76 (s)
C6-Sub	1"	39.5, CH_2	3.62 (d)	35.2, CH_2	3.65 (d)	41.7, CH_2	4.02 (s)
	2"	130.2, CH	6.44 (m)	127.4, CH	6.02 (t)	128.9, CH	7.08 (s)
	3"	129.2, CH	6.4 (m)			128.9, CH	7.08 (s)
	4"	137.7, C		125.8, CH	7.41 (m)	21.0, CH_3	2.30 (s)
	5"	125.9, CH	7.35 (d)	128.2, CH	7.29 (t)	128.9, CH	7.08 (s)
	6"	128.2, CH	7.27 (t)	126.7, CH	7.21 (t)	128.9, CH	7.08 (s)
	7"	126.8, CH	7.18 (t)	128.2, CH	7.29 (t)		
	8"	128.2, CH	7.27 (t)	125.8, CH	7.41 (m)		
	9"	125.9, CH	7.35 (d)	16.0, CH_3	2.16 (s)		

Table 2.4: NMR assignments for TPS-24, TPS-25, and TPS-46

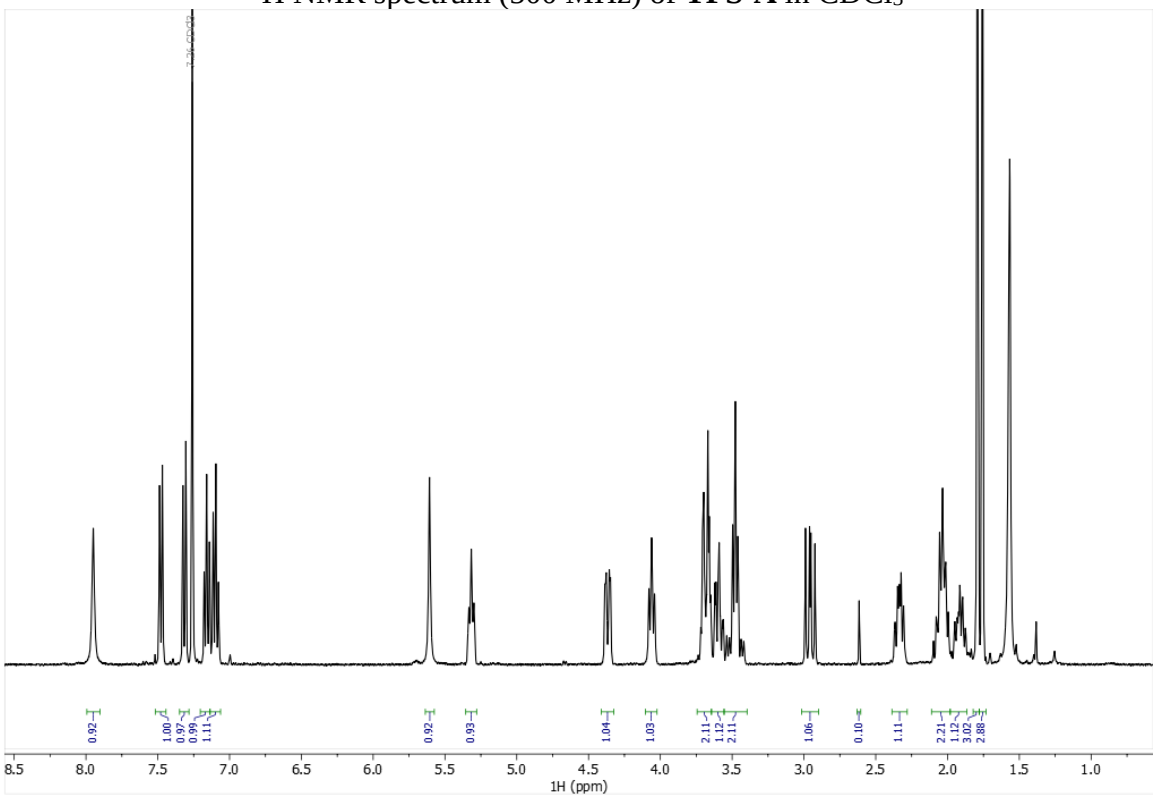


		TPS-52		TPS-58		TPS-61	TPS-62	
Segment	**Position**	**δC , type**	**δH**	**δC , type**	**δH**	**δH**	**δH**	**δC , type**
L-Trp	NH		5.62 (s)			5.60 (s)	5.61 (s)	
	α	54.5, CH	4.33 (dd)	54.5, CH	4.33 (dd)	4.35 (d)	4.36 (dd)	54.6, CH
	β	25.0, CH_2	2.91 (dd), 3.64 (m)	25.7, CH_2	2.91 (m), 3.63 (d)	2.93 (m), 3.66 (m)	2.93 (dd), 3.63 (m)	25.5, CH_2
	1		7.82 (s)		7.91 (s)	7.86 (s)	7.85 (s)	
	2			136.0, C				
	3			104.5, C				
	3a			126.3, C				
	4	117.7, CH	7.37 (d)	117.5, CH	7.37 (d)	7.40 (d)	7.40 (d)	117.7, CH
	5	121.2, CH	6.95 (d)	121.1, CH	6.93 (dd)	7.00 (d)	7.02 (dd)	121.1, CH
	6			135.6, C				
	7	110.6, CH	7.05 (s)	110.7, CH	7.07 (s)	7.16 (s)	7.16 (s)	110.7, CH
	7a			136.1, C				
	C=O							
Proline	α	59.7, CH	4.04 (t)	59.3, CH	4.04 (t)	4.05 (s)	4.05 (t)	59.3, CH
	β	28.5, CH_2	2.03 (m), 2.32 (m)	28.4, CH_2	2.02 (m), 2.32 (dt)	2.04 (m), 2.33 (m)	2.03 (m), 2.31 (m)	28.5, CH_2
	γ	22.6, CH_2	1.90 (m), 2.03 (m)	22.7, CH_2	1.89 (m), 2.02 (m)	1.90 (m), 2.33 (m)	1.89 (m), 2.01 (m)	22.8, CH_2
	δ	45.4, CH_2	3.58 (m), 3.66 (m)	45.4, CH_2	3.58 (m), 3.66 (m)	3.60 (m), 3.66 (m)	3.58 (t), 3.67 (m)	45.4, CH_2
	C=O			165.9, C				
C2-Prenyl	1'	25.2, CH_2	3.43 (m)	25.2, CH_2	3.43 (qd)	3.45 (m)	3.44 (t)	25.0, CH_2
	2'	119.5, CH	5.28 (t)	119.8, CH	5.27 (m)	5.29 (t)	5.30 (t)	119.7, CH
	3'			135.5, C				
	4'	18.1, CH_3	1.73 (s)	18.1, CH_3	1.73 (s)	1.74 (s)	1.74 (s)	17.7, CH_3
	5'	25.7, CH_3	1.76 (s)	25.9, CH_3	1.76 (s)	1.78 (s)	1.78 (s)	25.8, CH_3
C6-Sub	1"	41.1, CH_2	4.00 (s)	41.8, CH	3.96 (s)	4.05 (s)	4.23 (s)	36.3, CH_2
	2"	129.7, CH	7.11 (d)	109.2, CH	6.66 (m)	6.00 (d)	6.81 (dd)	125.1, CH
	3"	113.8, CH	6.81 (d)	145.7, C		6.28 (d)	6.92 (dd)	126.8, CH
	4"	113.8, CH	6.81 (d)	147.5, C		7.32 (s)	7.13 (dd)	123.9, CH
	5"	129.7, CH	7.11 (d)	108.0, CH	6.71 (d)			
	6"	55.2, CH_3	3.77 (s)	121.5, CH	6.66 (m)			
	7"			100.8, CH_2	5.88 (s)			

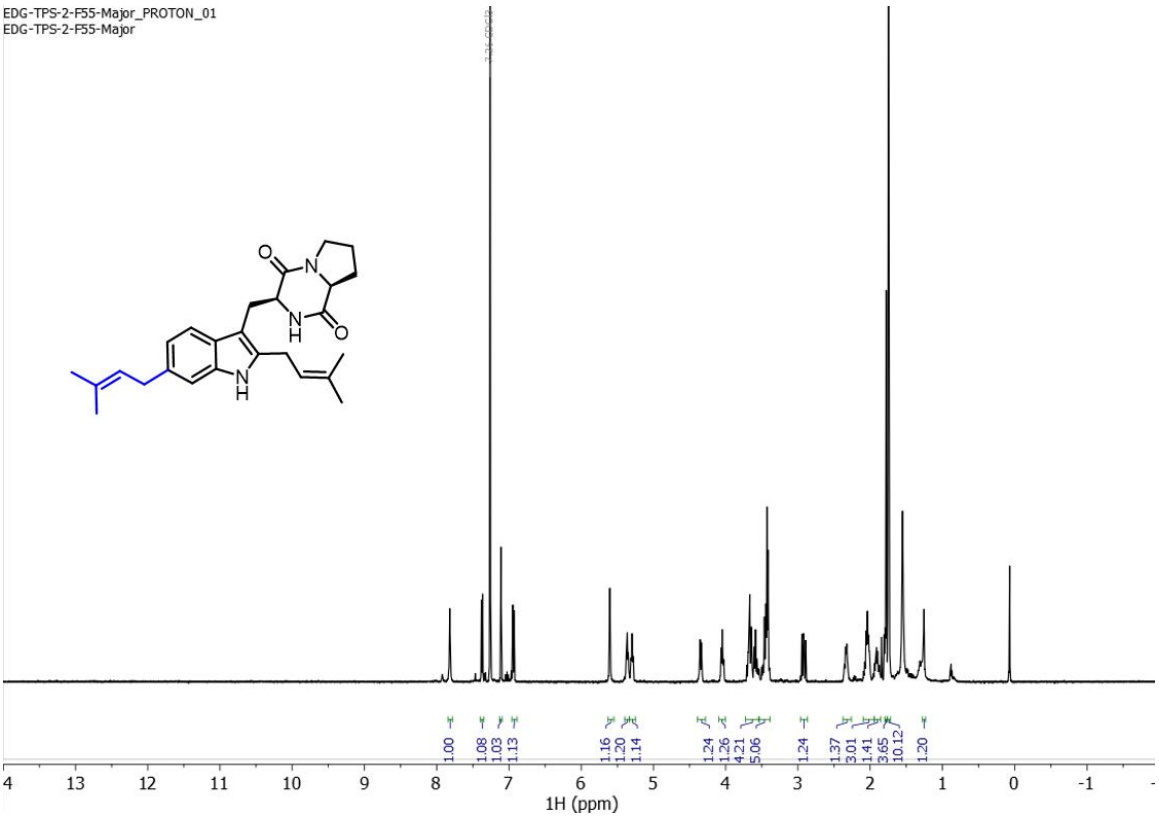
Table 2.5: NMR assignments for TPS-52, TPS-58, TPS-61, and TPS-62



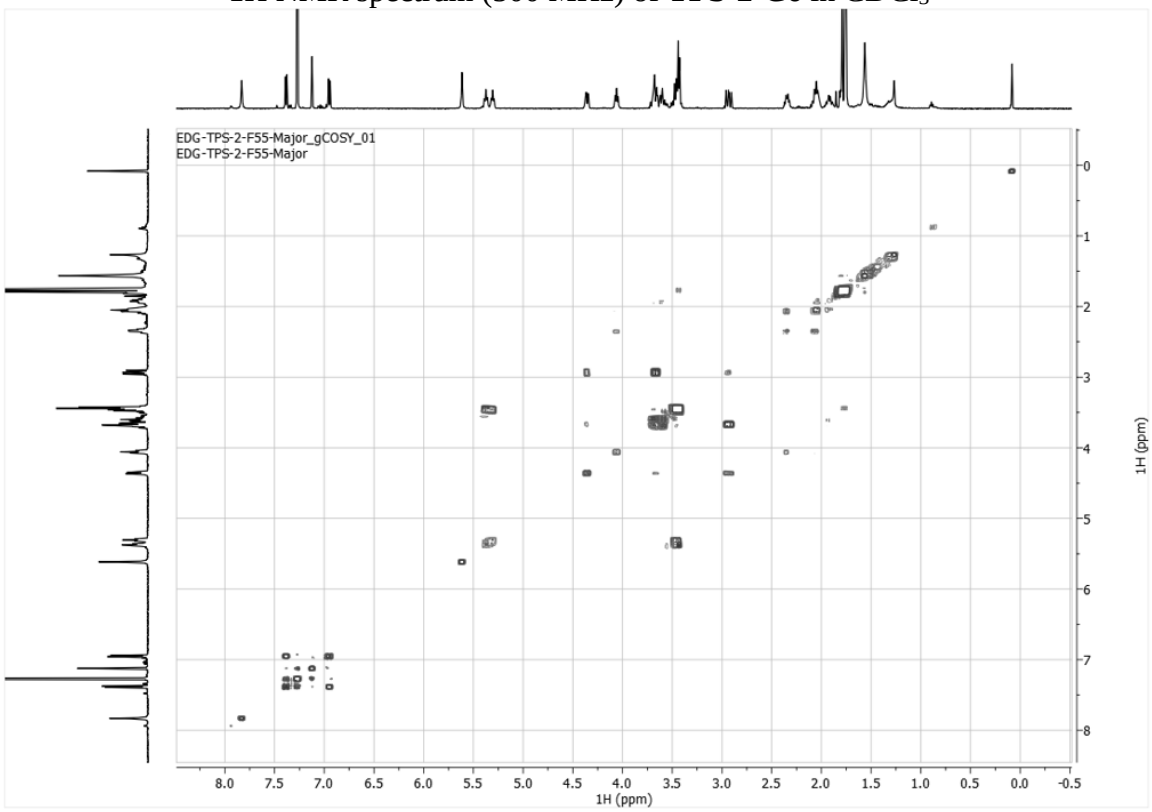
^{1}H NMR spectrum (500 MHz) of **TPS-A** in $CDCl_3$



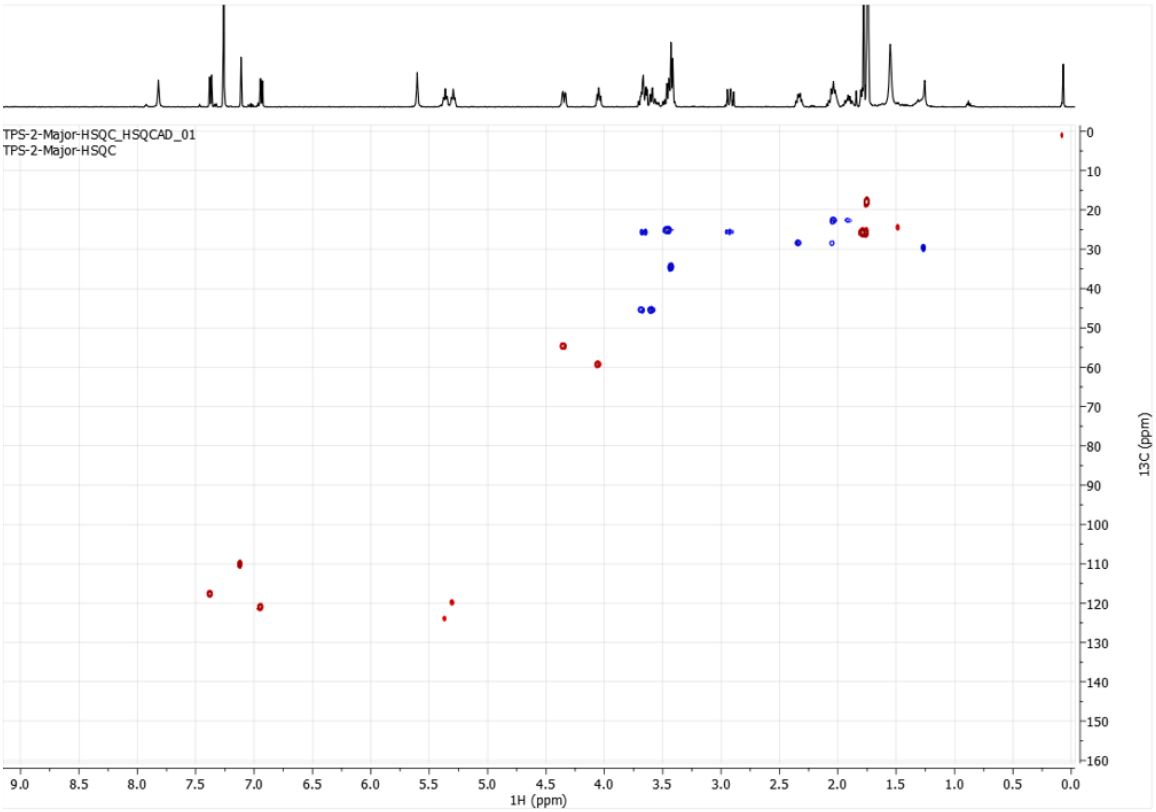
^{1}H NMR spectrum (400 MHz) of **TPS-B** in $CDCl_3$



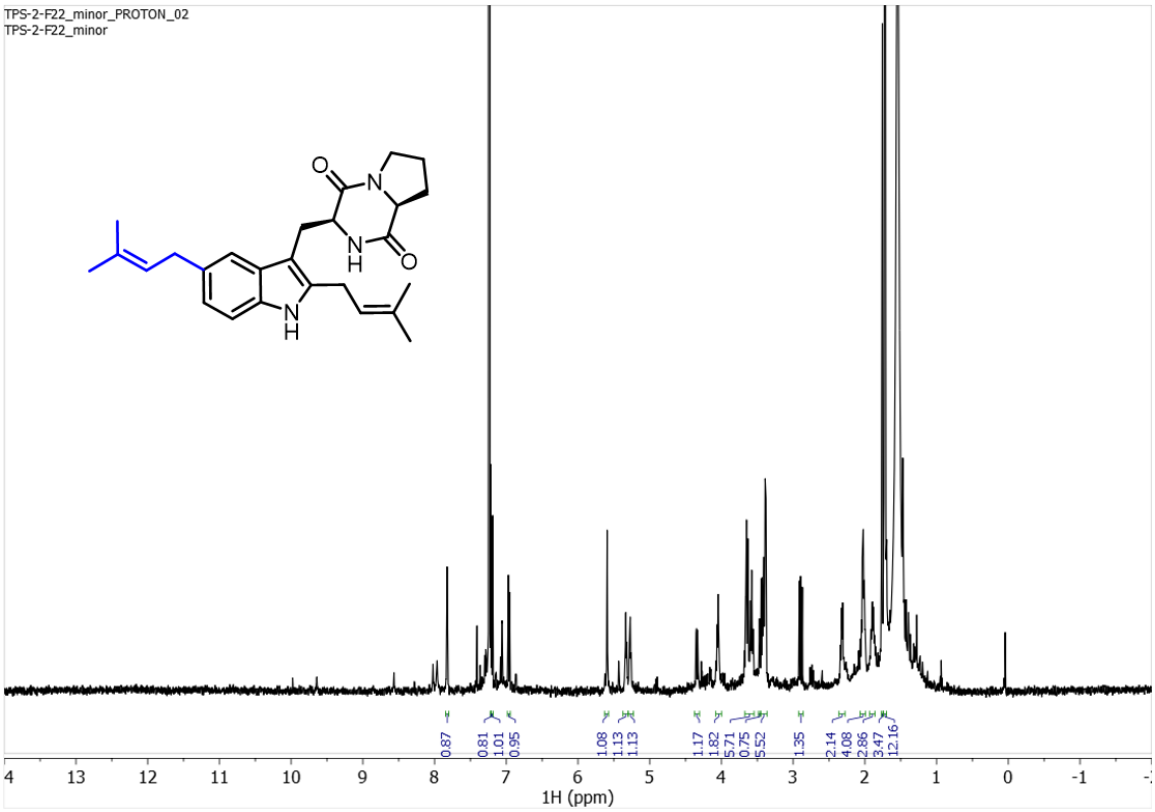
1H NMR spectrum (500 MHz) of **TPS-2-C6** in $CDCl_3$



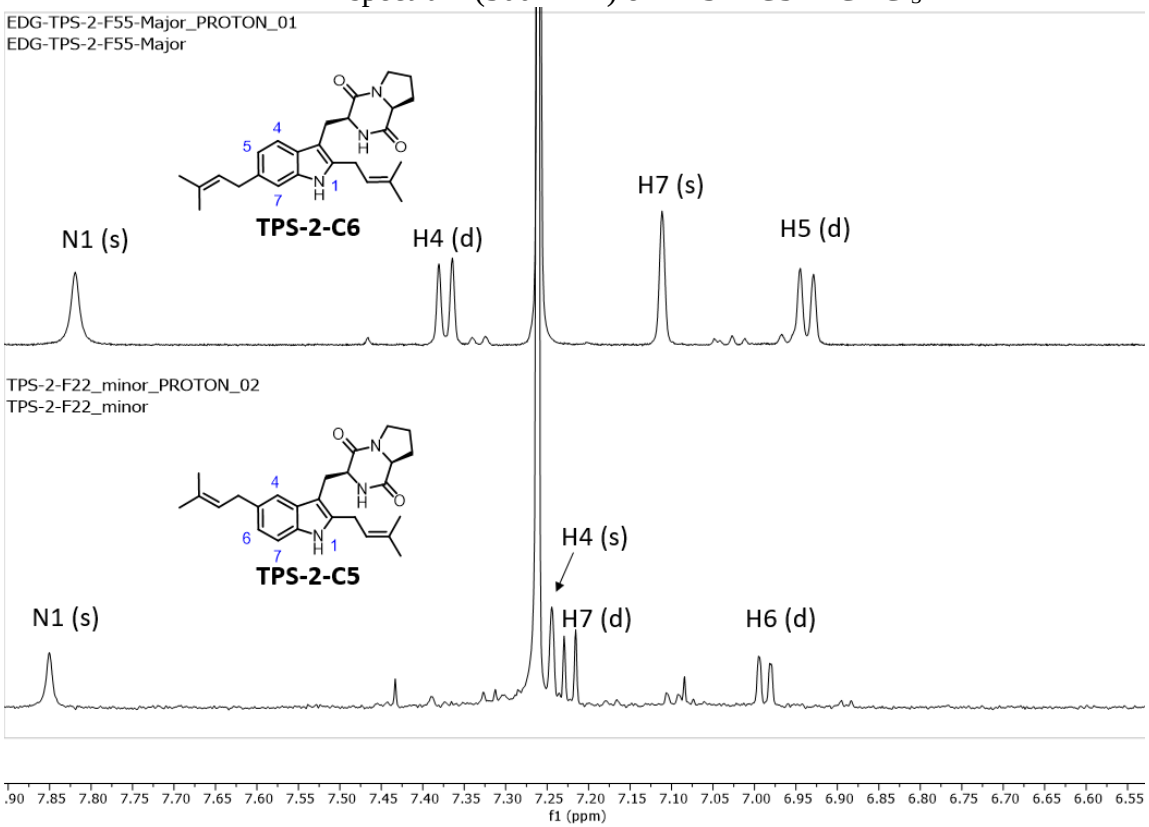
COSY NMR spectrum (500 MHz) of **TPS-2-C6** in $CDCl_3$



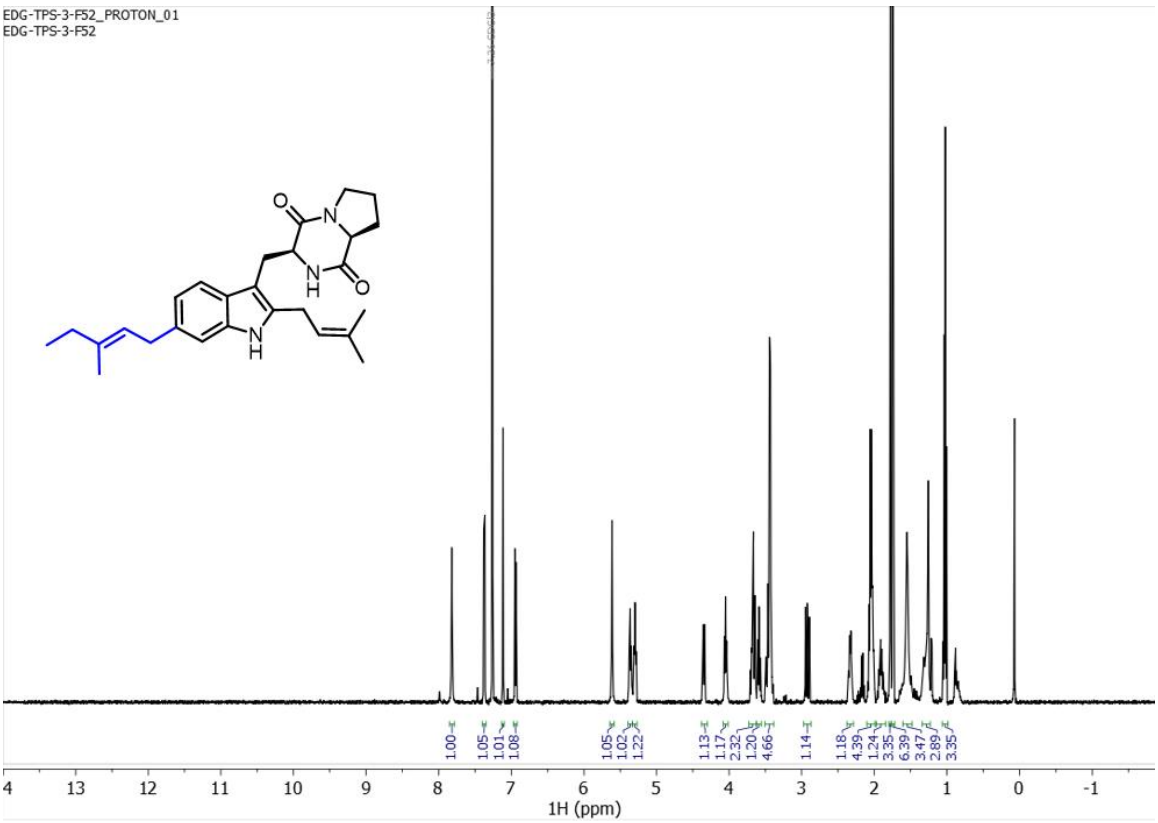
HSQC NMR spectrum (500 MHz) **of TPS-2-C6** in CDCl3



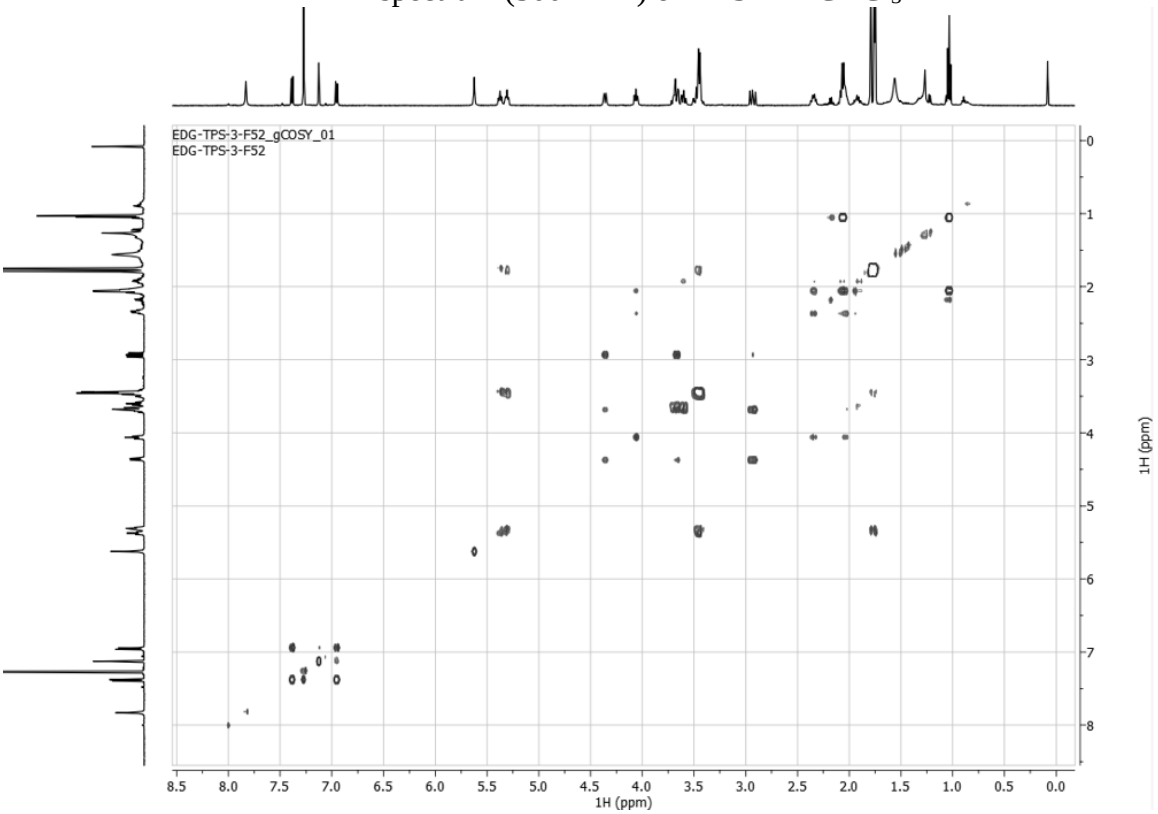
^{1}H NMR spectrum (500 MHz) **of TPS-2-C5** in $CDCl_3$



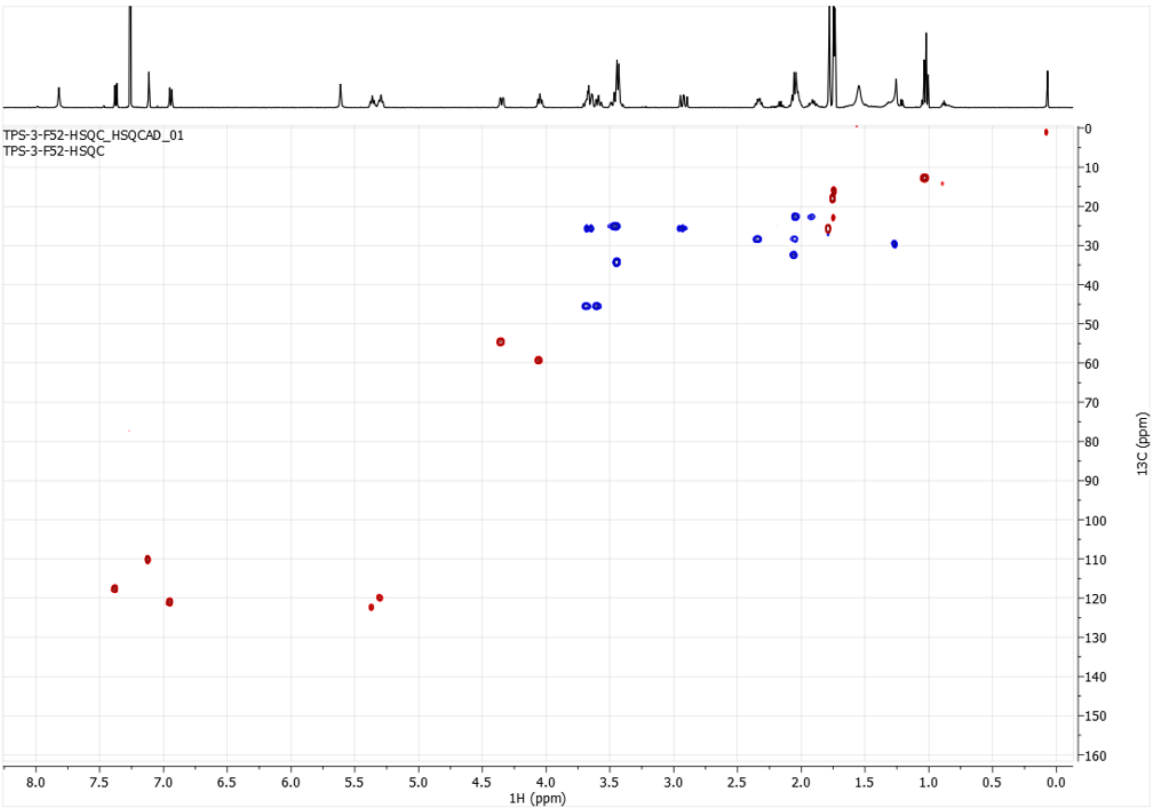
Stacked ^{1}H NMR Spectra (500 MHz) of isolated **TPS-2** regioisomers in $CDCl_3$



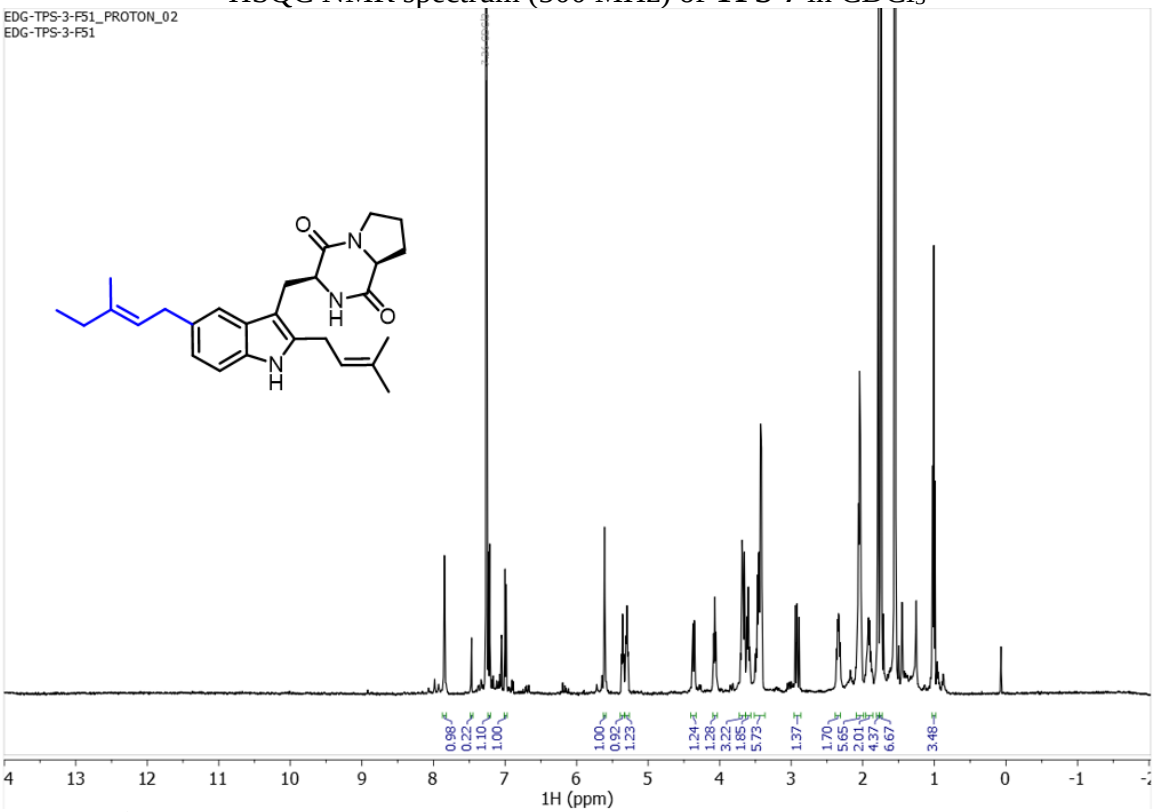
^{1}H NMR spectrum (500 MHz) of **TPS-7** in $CDCl_3$



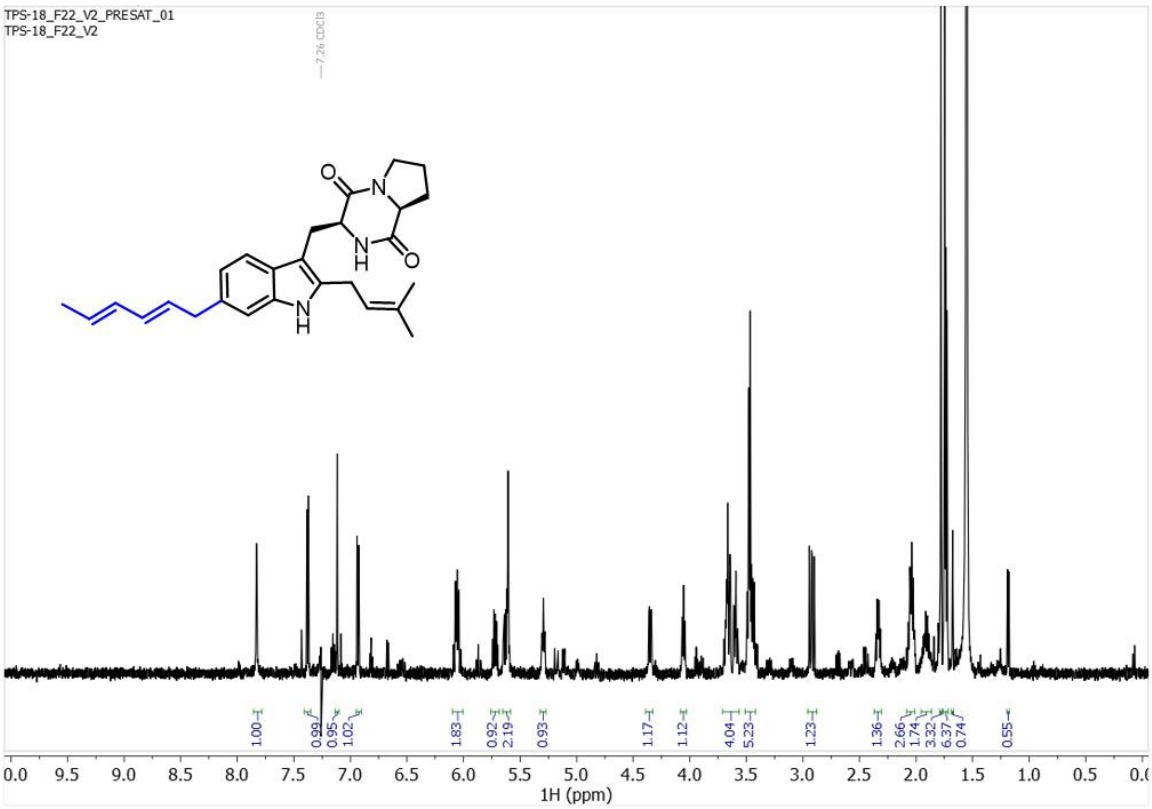
COSY spectrum (500 MHz) of **TPS-7** in $CDCl_3$



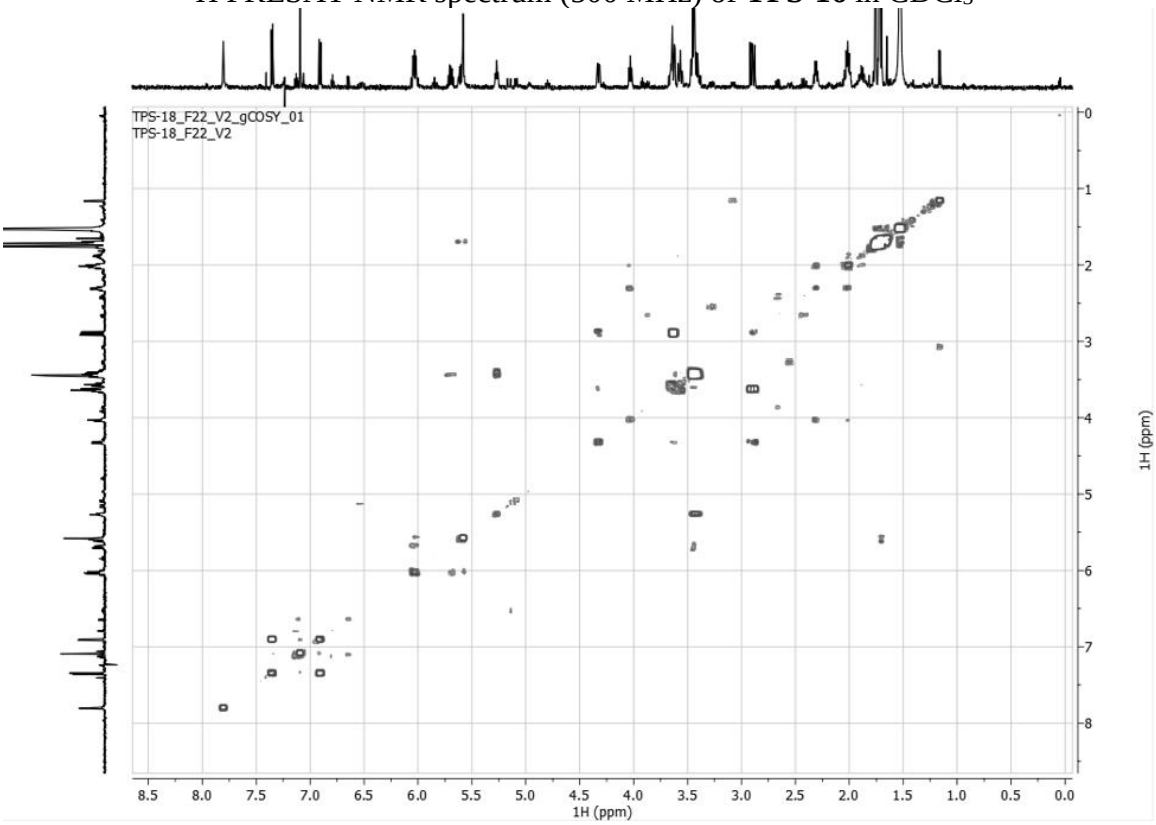
HSQC NMR spectrum (500 MHz) of **TPS-7** in $CDCl_3$



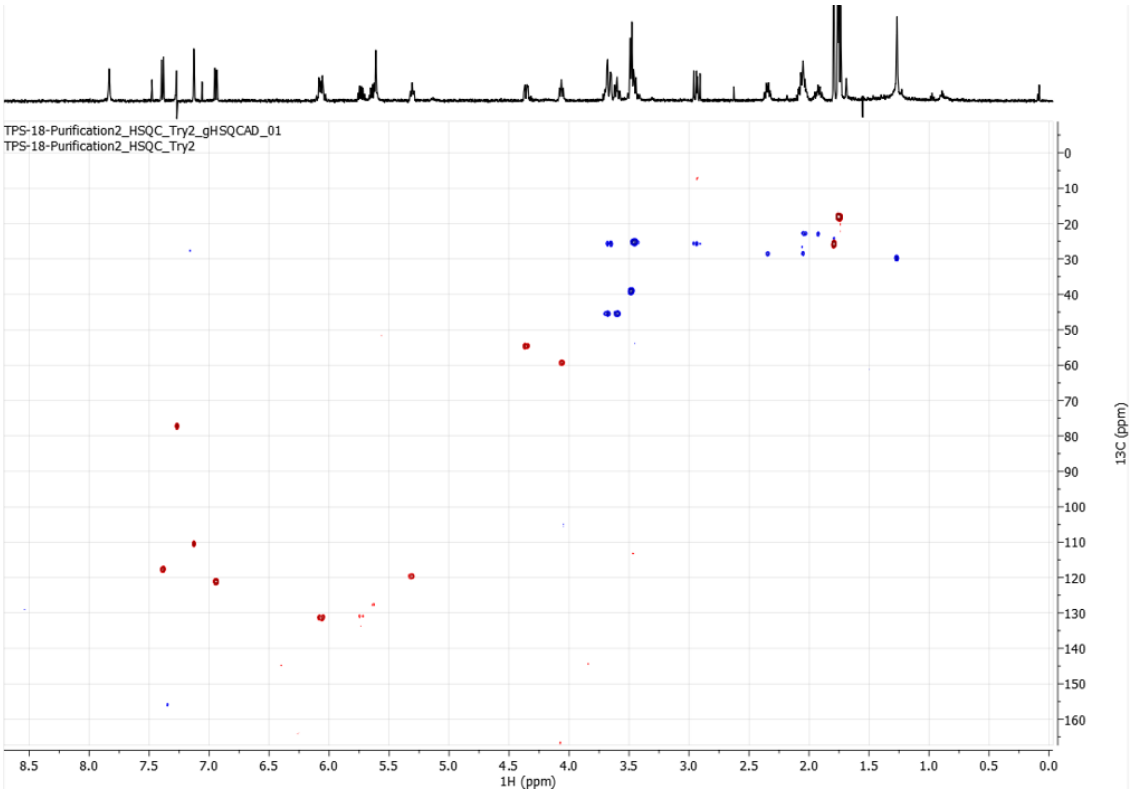
^{1}H NMR spectrum (500 MHz) of **TPS-7-C5** (minor product) in $CDCl_3$.



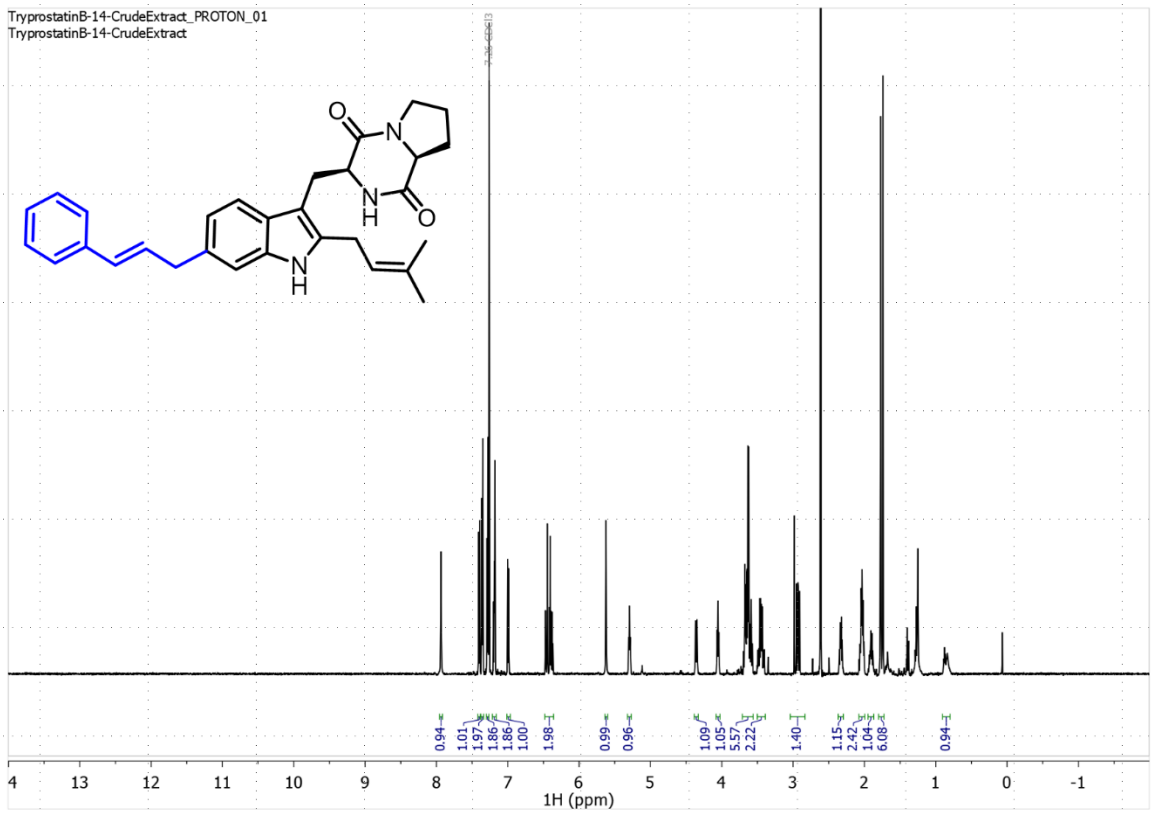
^{1}H PRESAT NMR spectrum (500 MHz) of **TPS-16** in $CDCl_3$



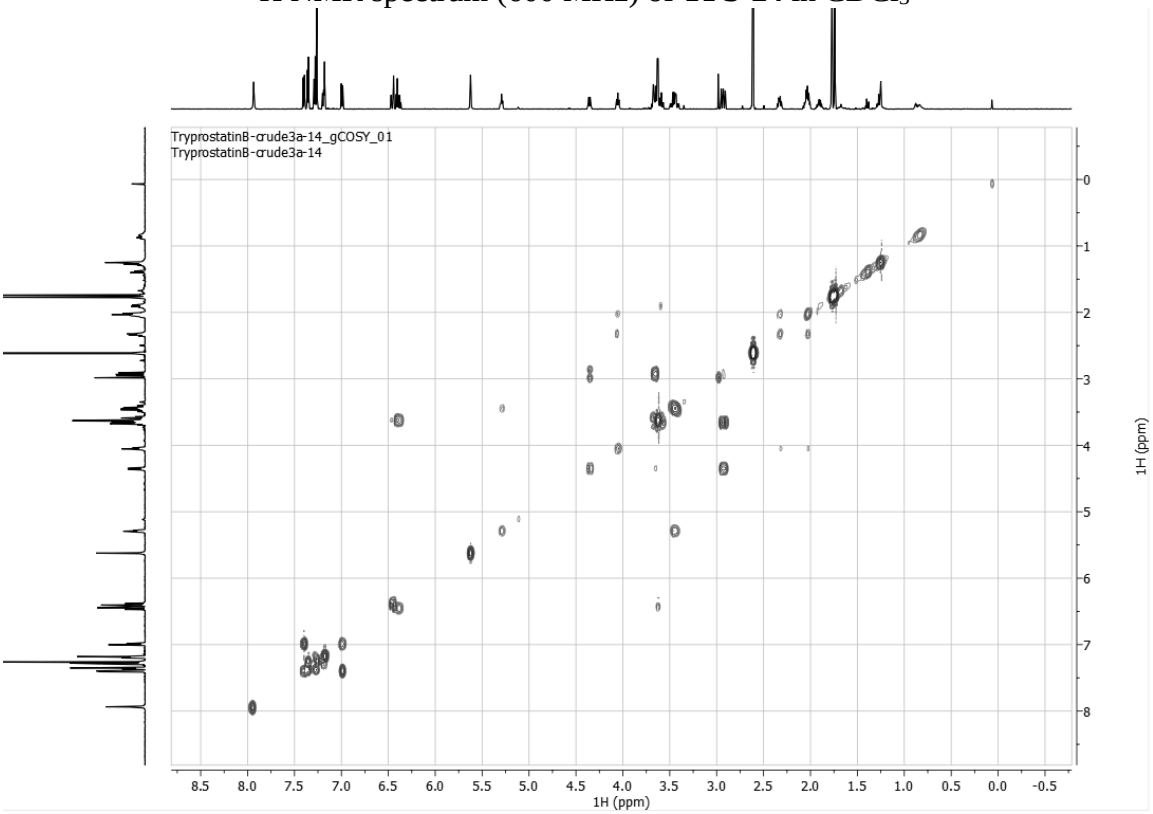
COSY NMR spectrum (500 MHz) of **TPS-16** in $CDCl_3$



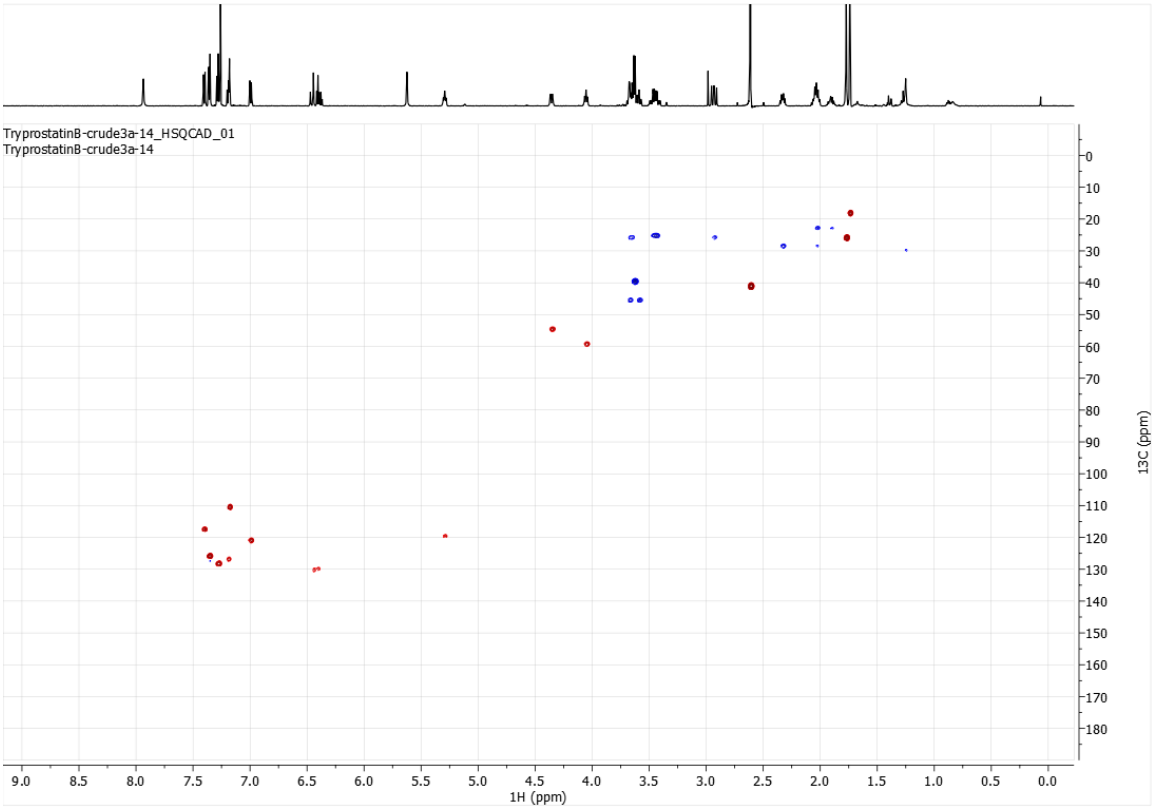
HSQC NMR spectrum (500 MHz) of **TPS-16** in $CDCl_3$



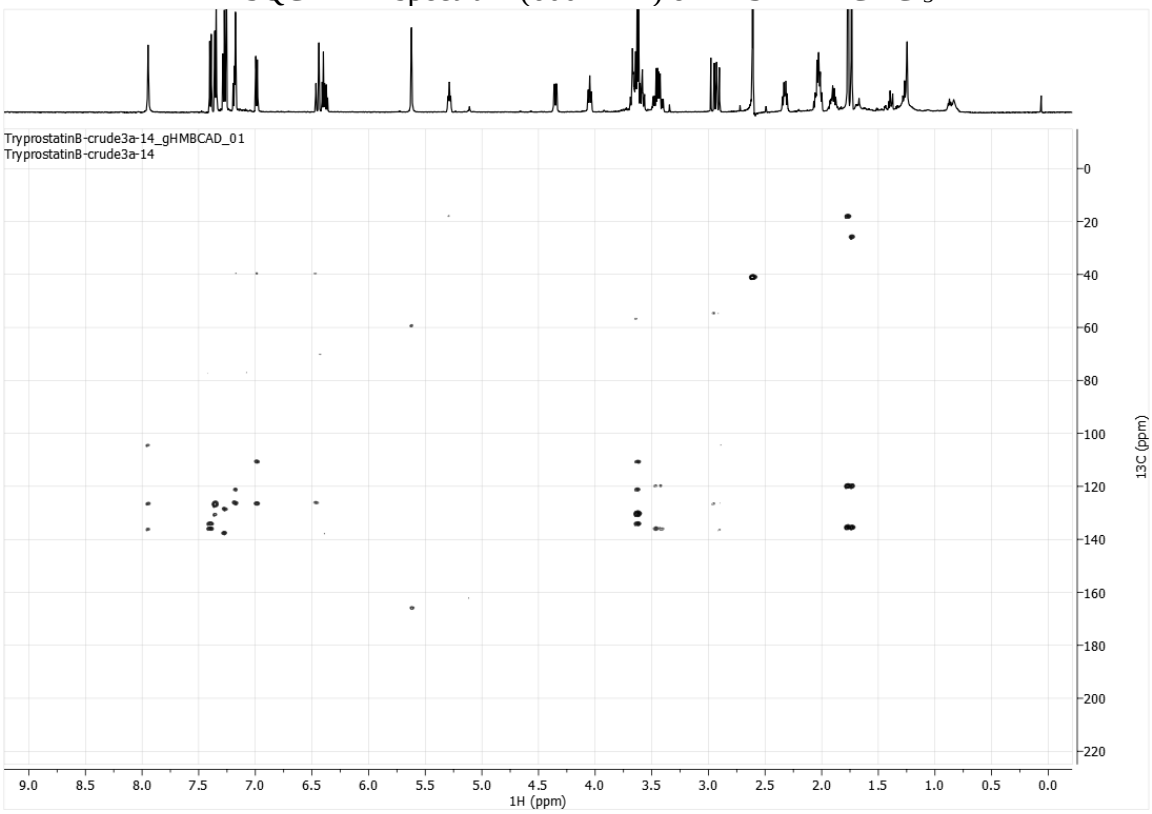
^{1}H NMR spectrum (600 MHz) of **TPS-24** in $CDCl_3$



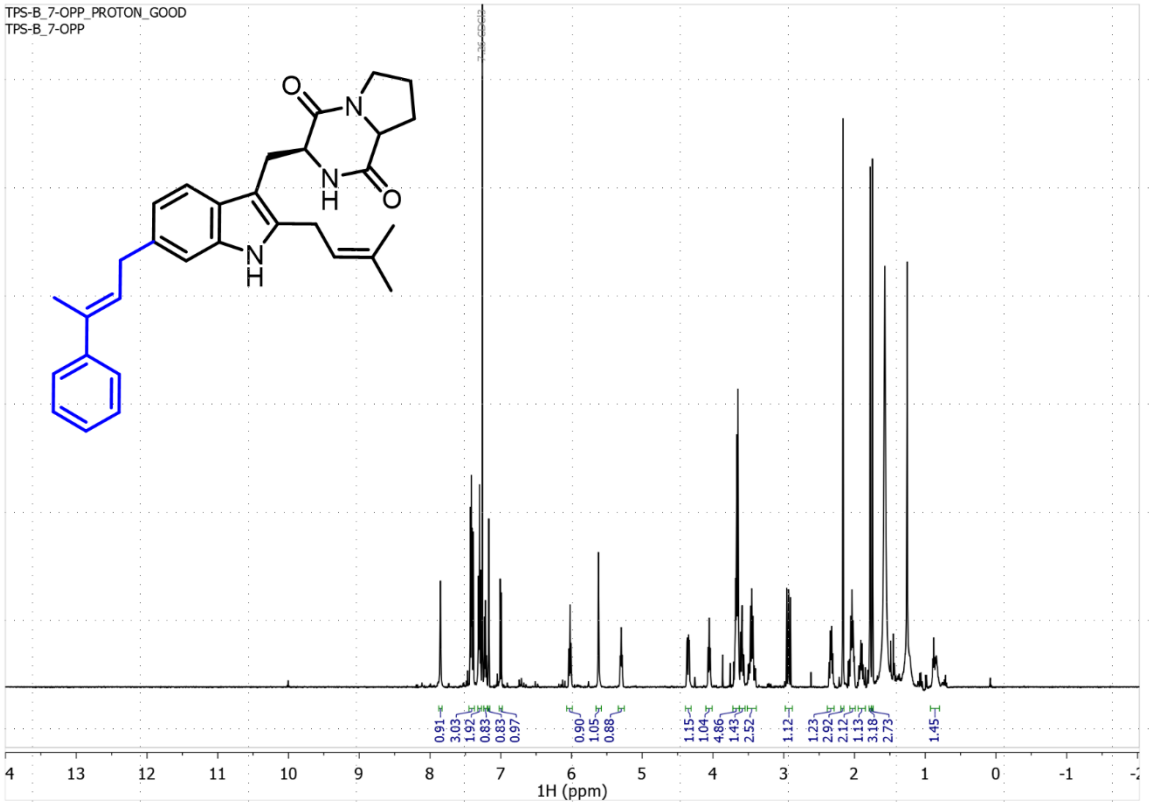
COSY NMR spectrum (600 MHz) of **TPS-24** in $CDCl_3$



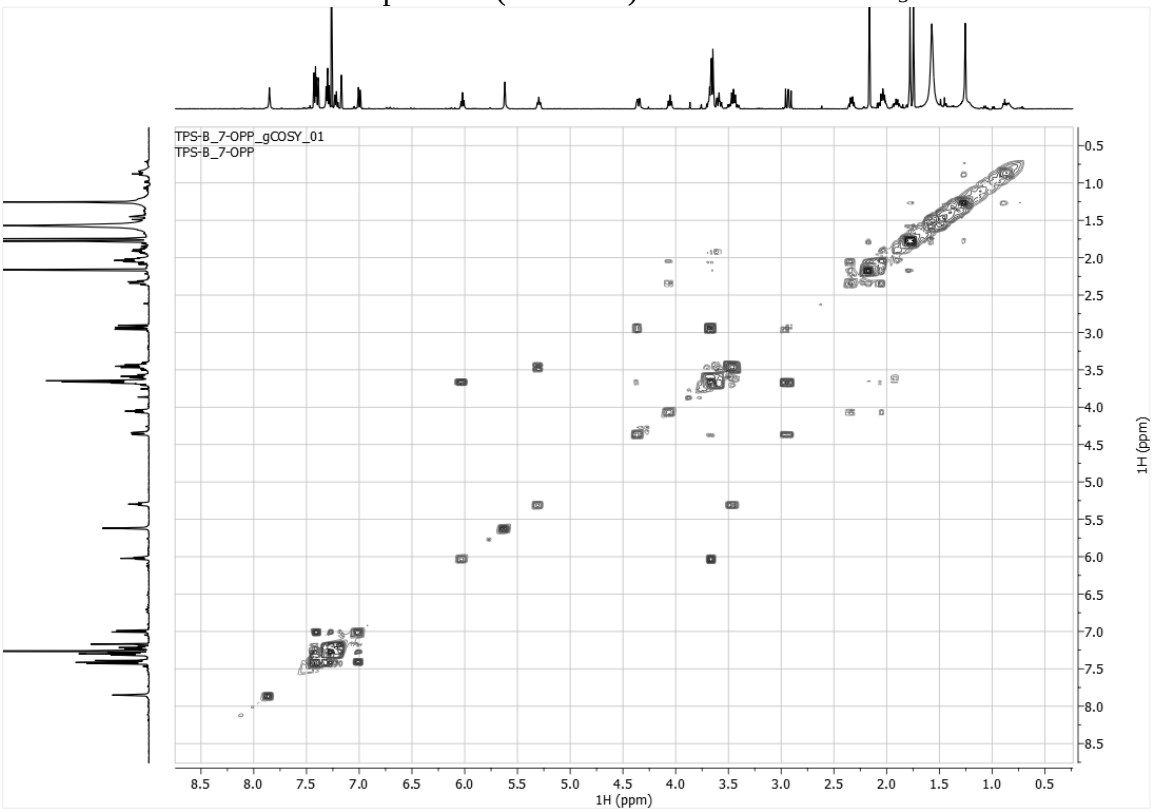
HSQC NMR spectrum (600 MHz) of **TPS-24** in $CDCl_3$



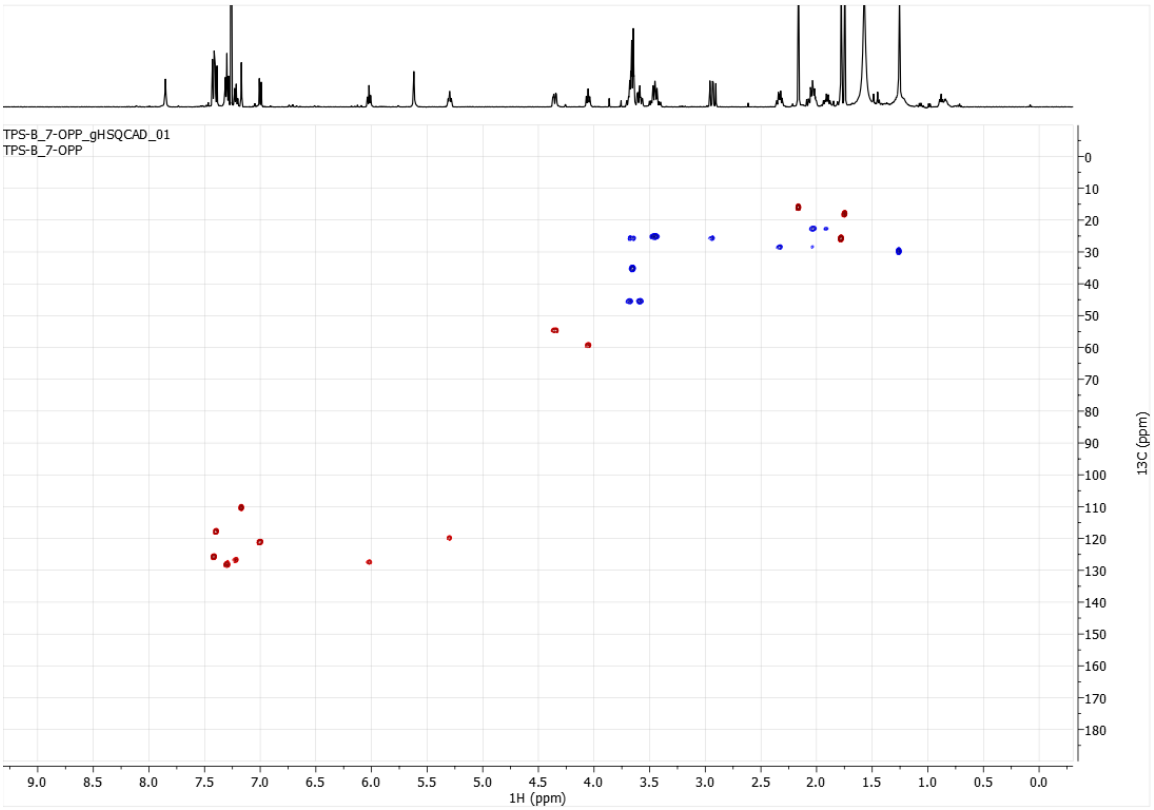
HMBC NMR spectrum (600 MHz) of **TPS-24** in $CDCl_3$



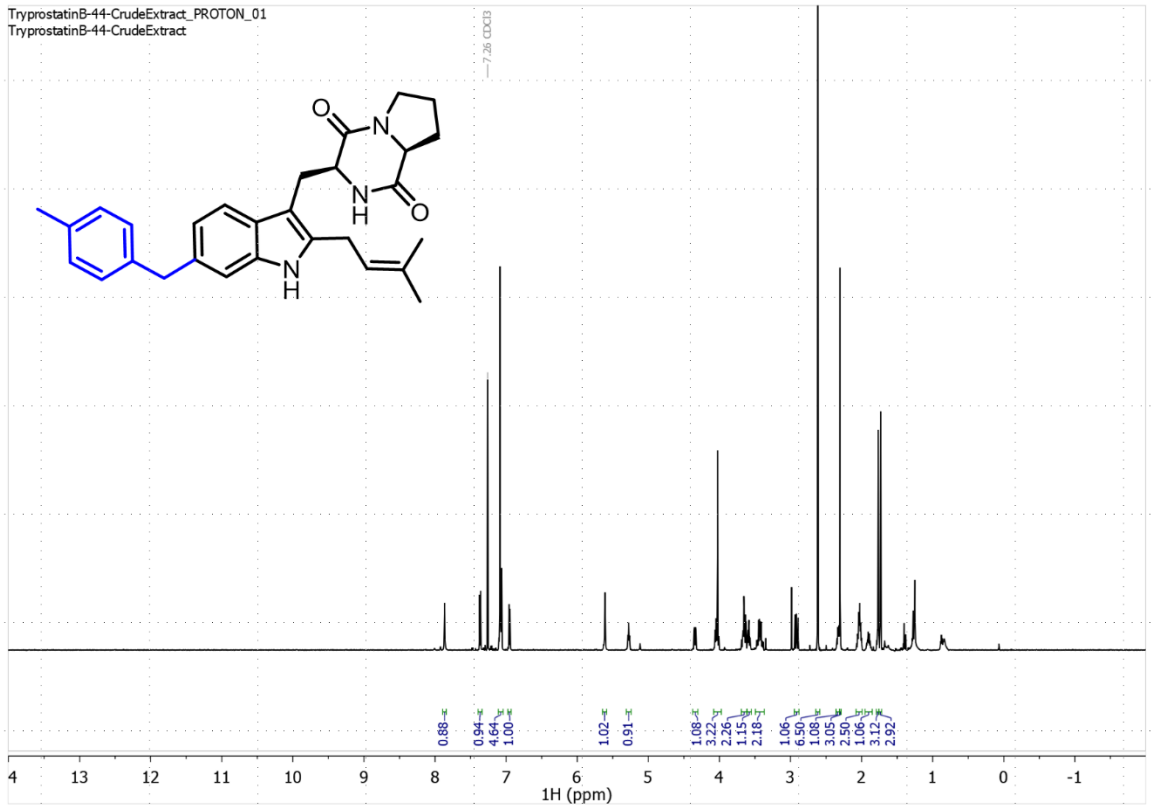
^{1}H NMR spectrum (500 MHz) of **TPS-25** in $CDCl_3$



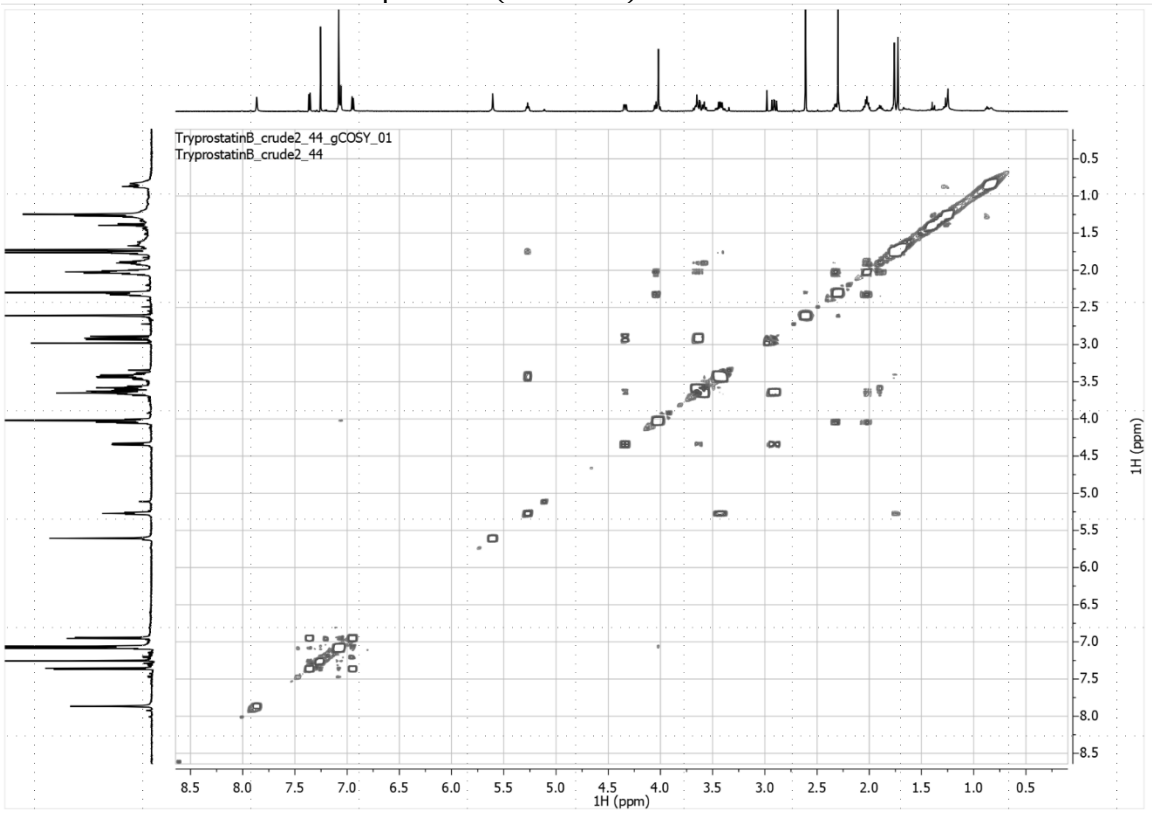
COSY NMR spectrum (500 MHz) of **TPS-25** in $CDCl_3$



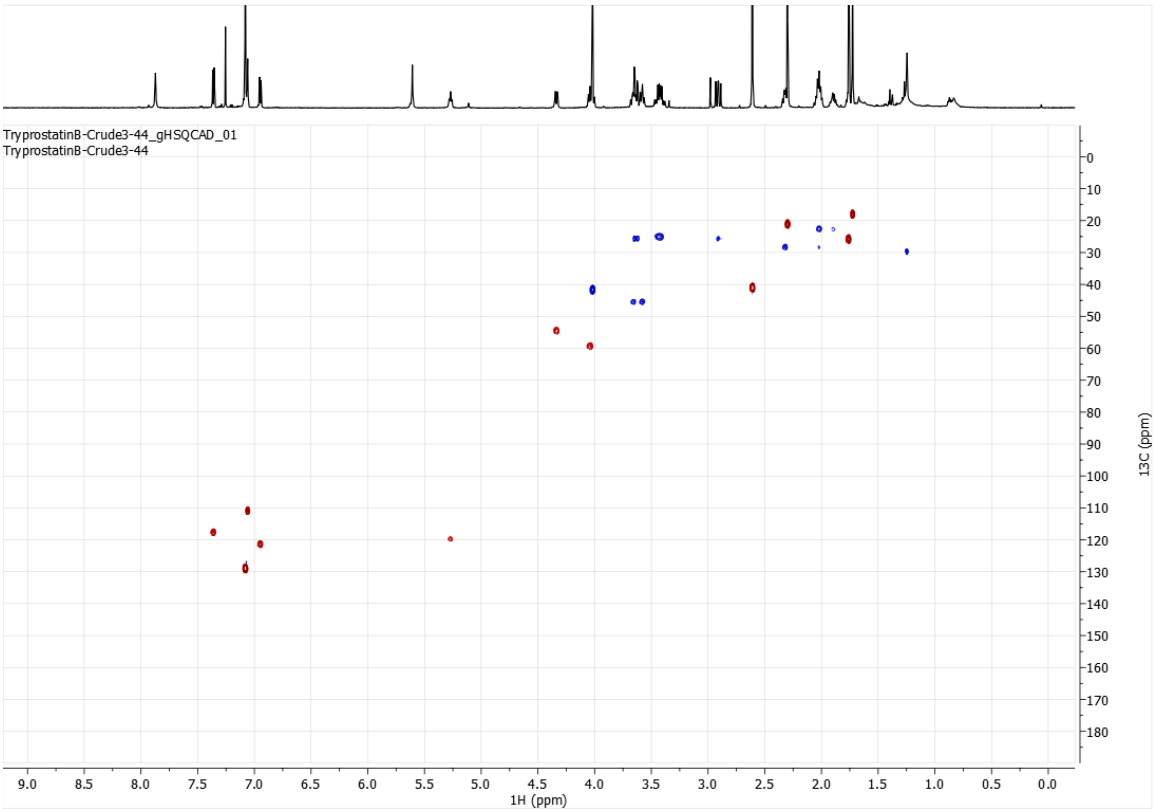
HSQC NMR spectrum (500 MHz) of **TPS-25** in $CDCl_3$



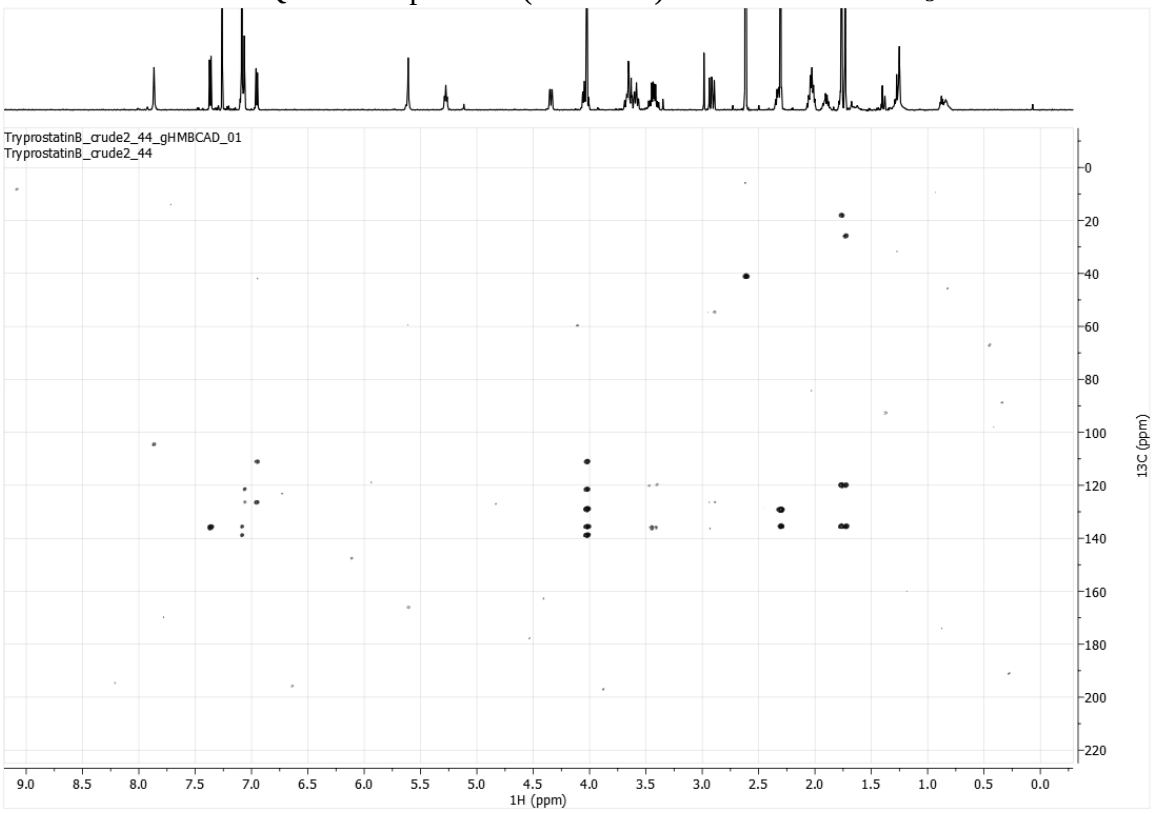
^{1}H NMR spectrum (600 MHz) of **TPS-46** in $CDCl_3$



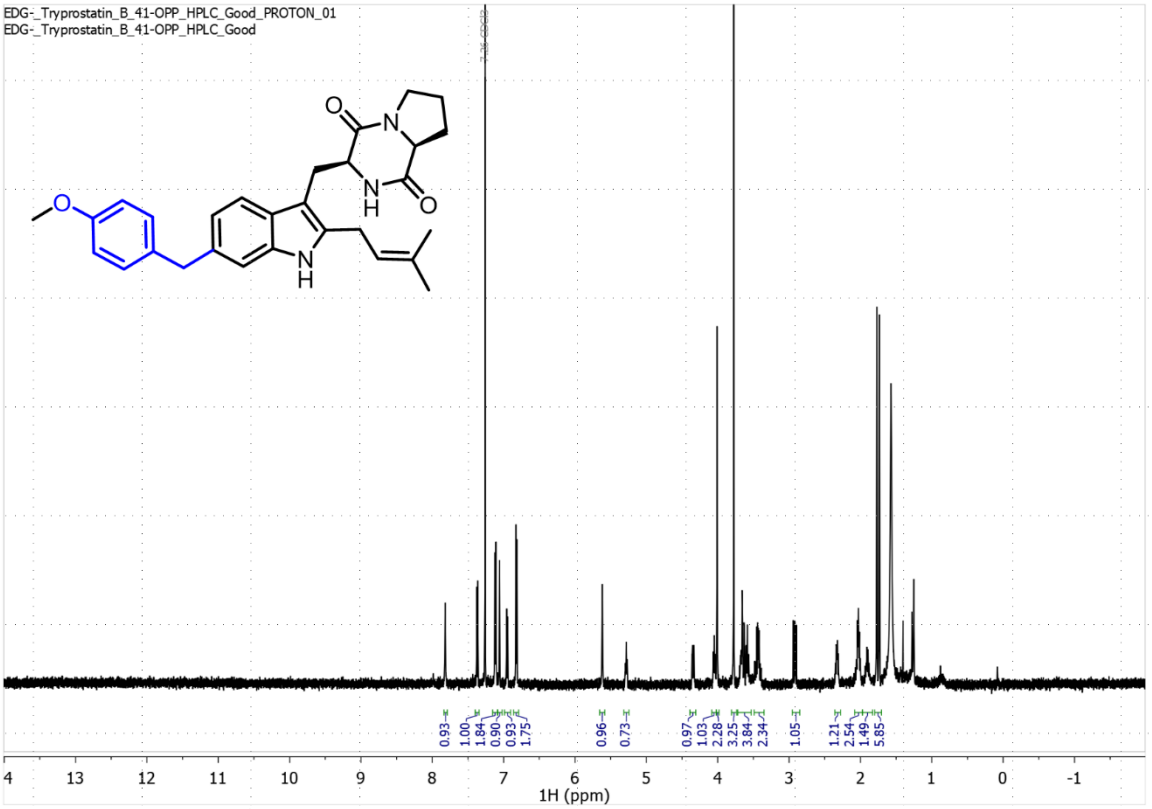
COSY NMR spectrum (600 MHz) of **TPS-46** in $CDCl_3$



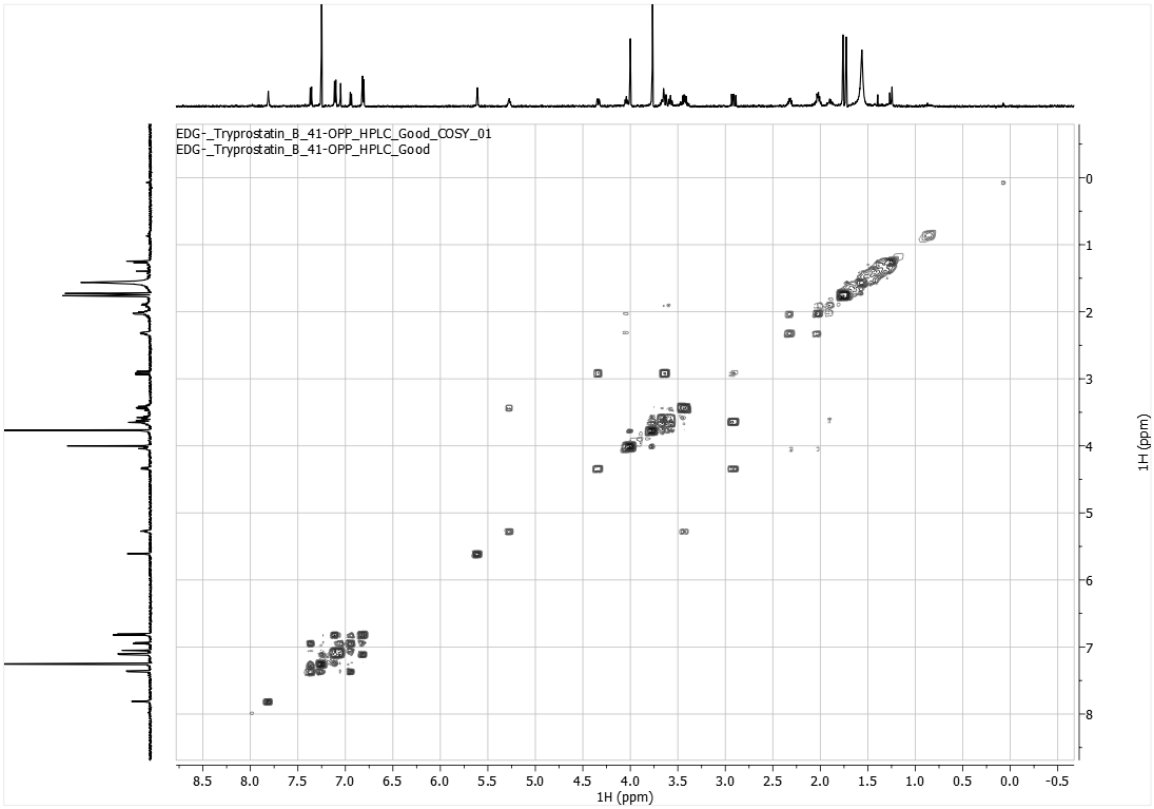
HSQC NMR spectrum (600 MHz) of **TPS-46** in $CDCl_3$



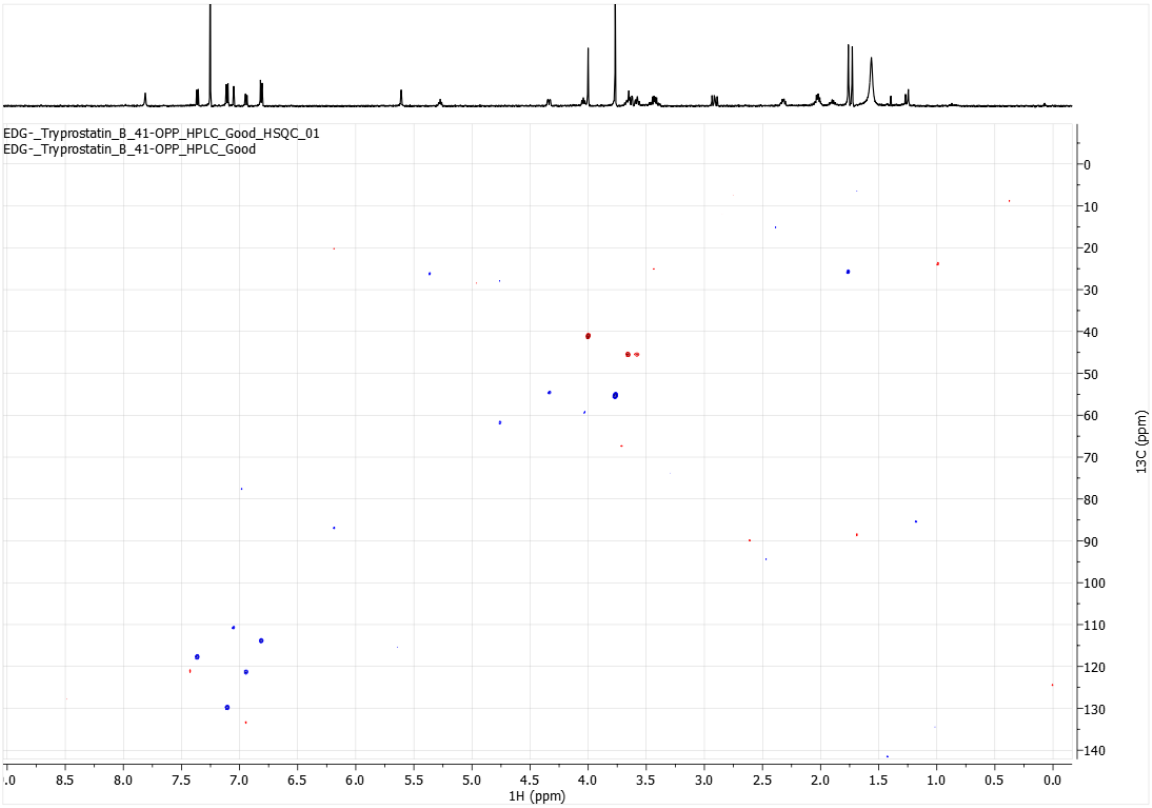
HMBC NMR spectrum (600 MHz) of **TPS-46** in $CDCl_3$



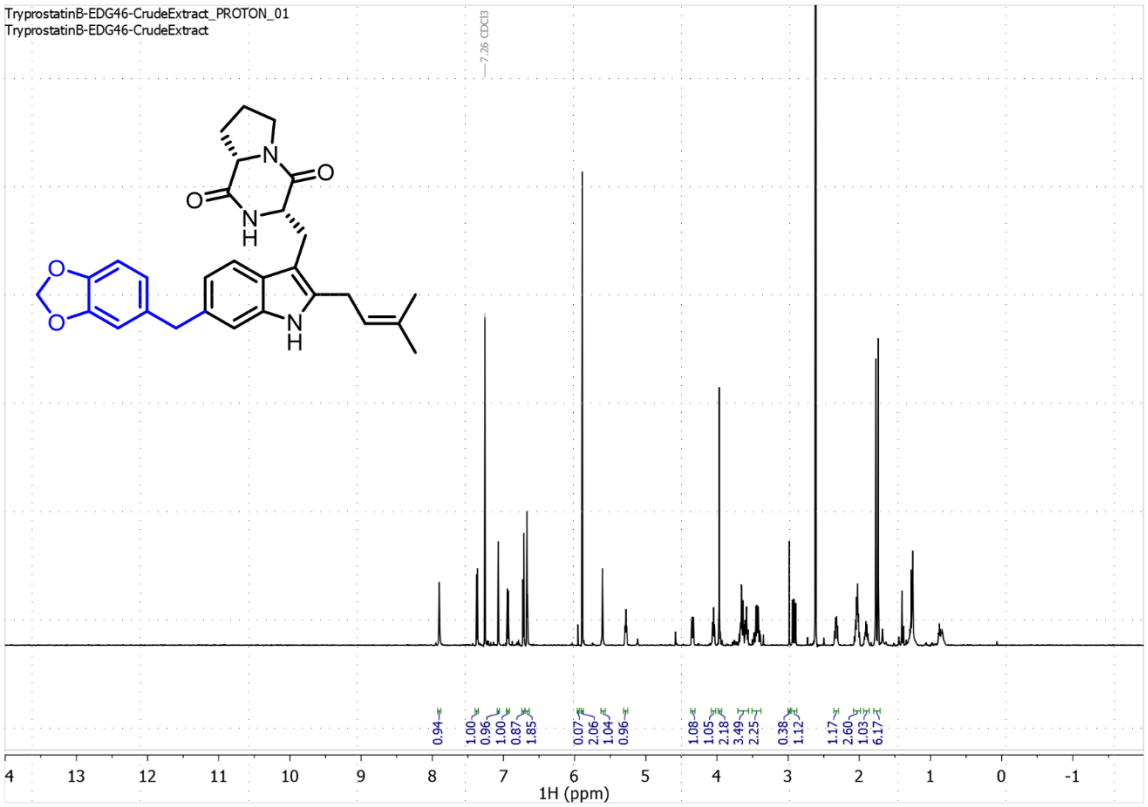
^{1}H NMR spectrum (600 MHz) of **TPS-52** in CDCl3



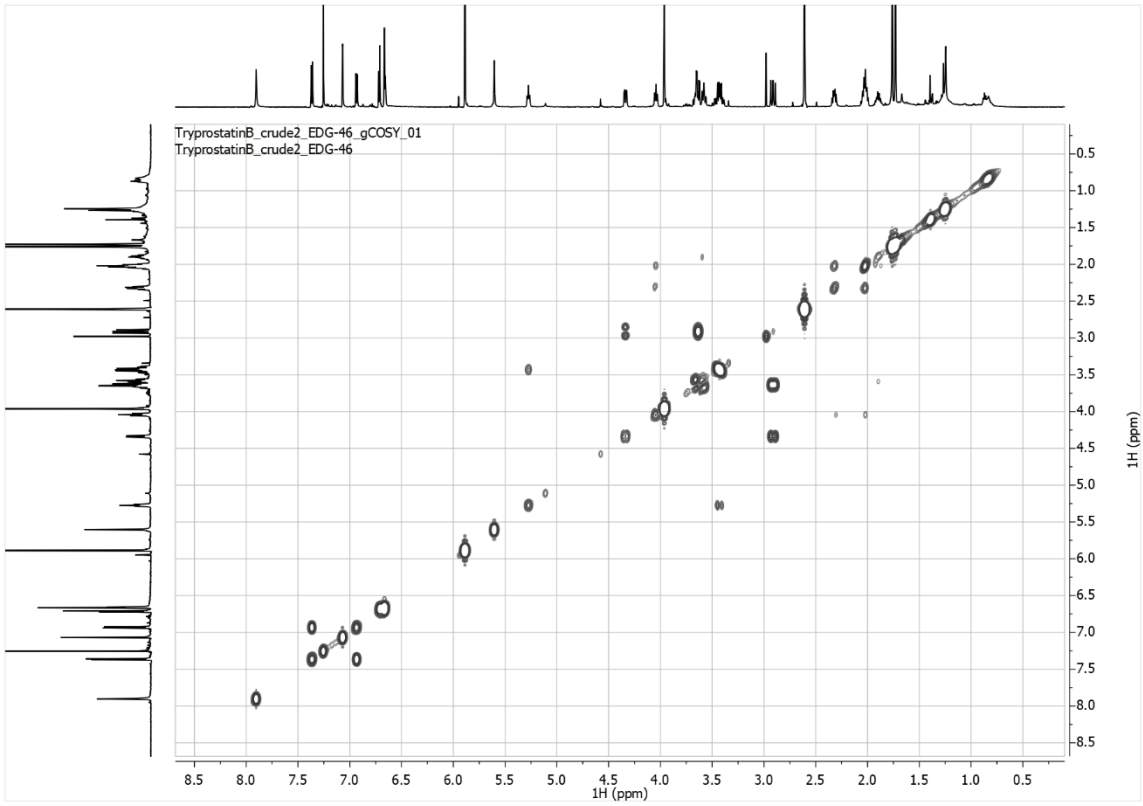
COSY NMR spectrum (600 MHz) of **TPS-52** in $CDCl_3$



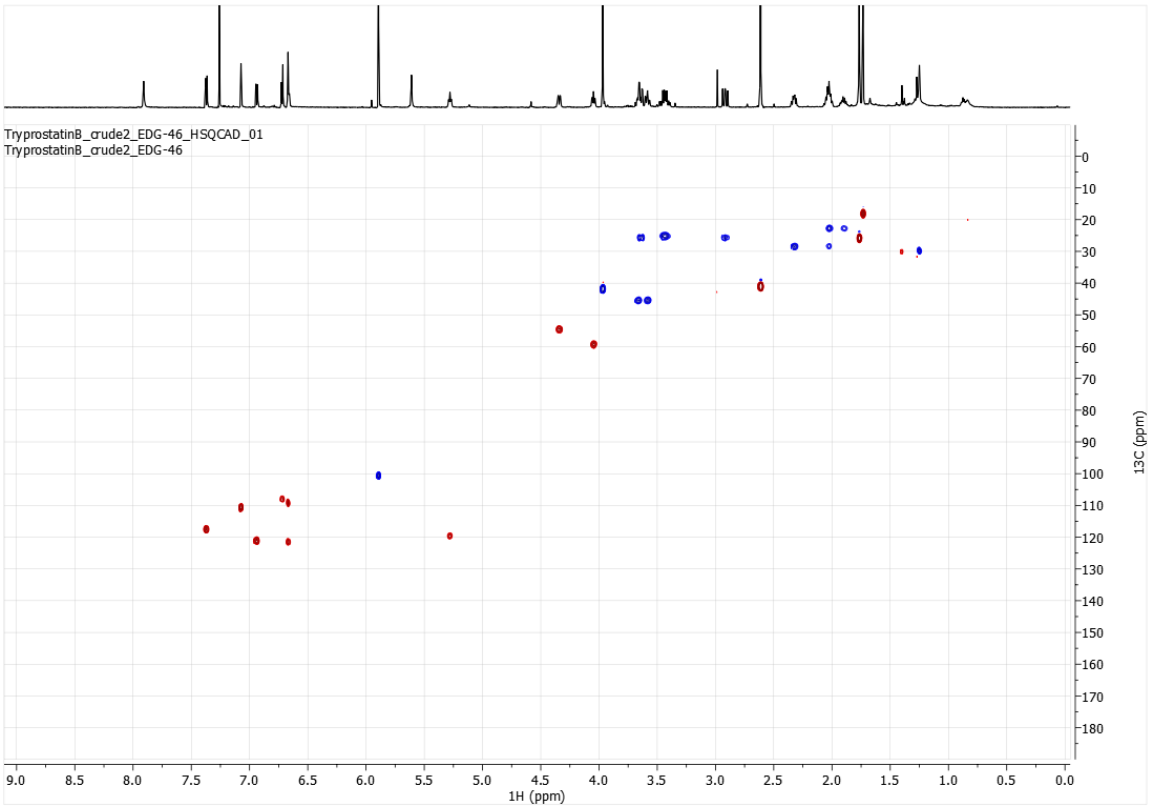
HSQC NMR spectrum (600 MHz) of **TPS-52** in $CDCl_3$



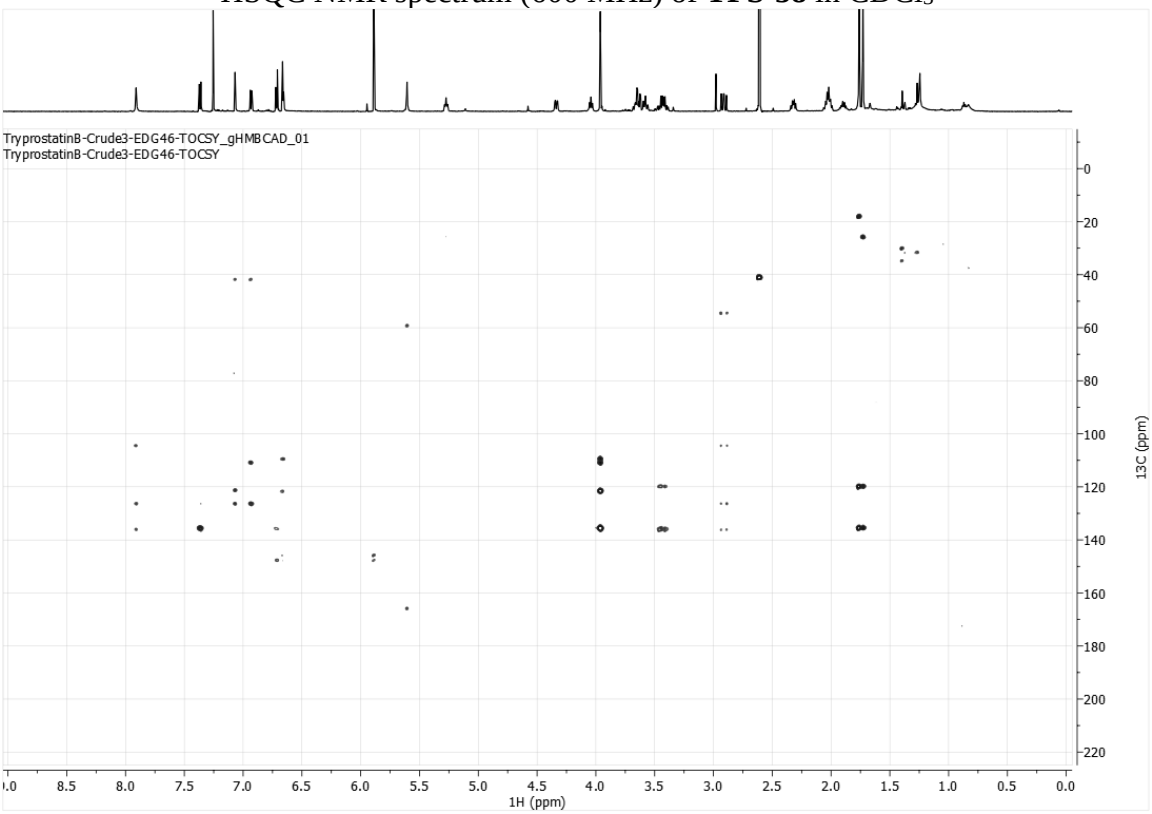
^{1}H NMR Spectrum (600 MHz) of **TPS-58** in $CDCl_3$



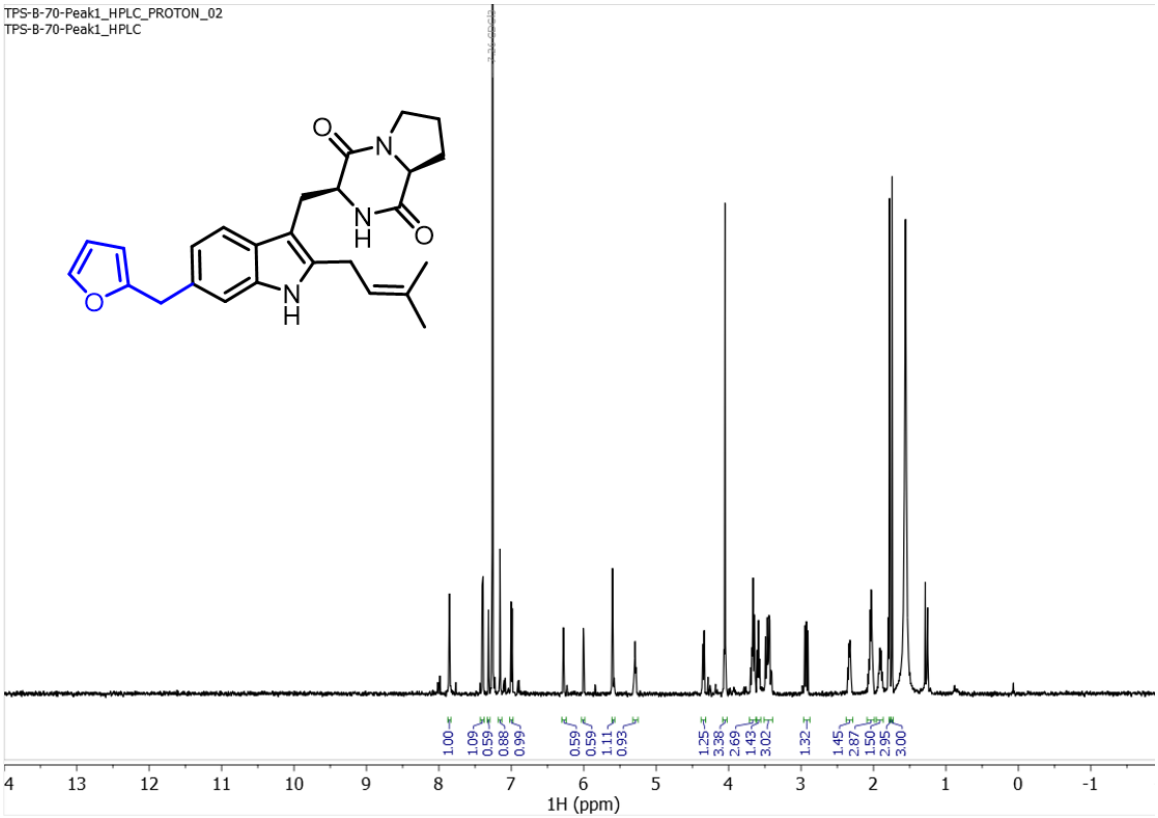
COSY NMR spectrum (600 MHz) of **TPS-58** in $CDCl_3$



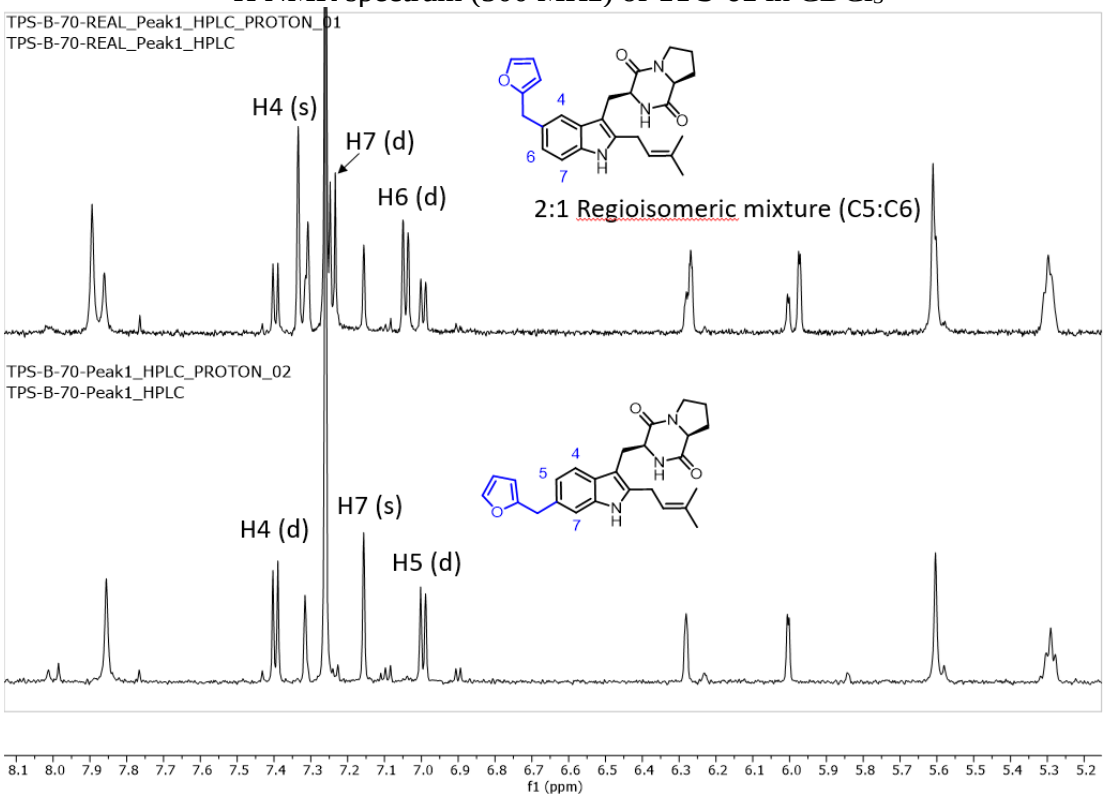
HSQC NMR spectrum (600 MHz) of **TPS-58** in $CDCl_3$



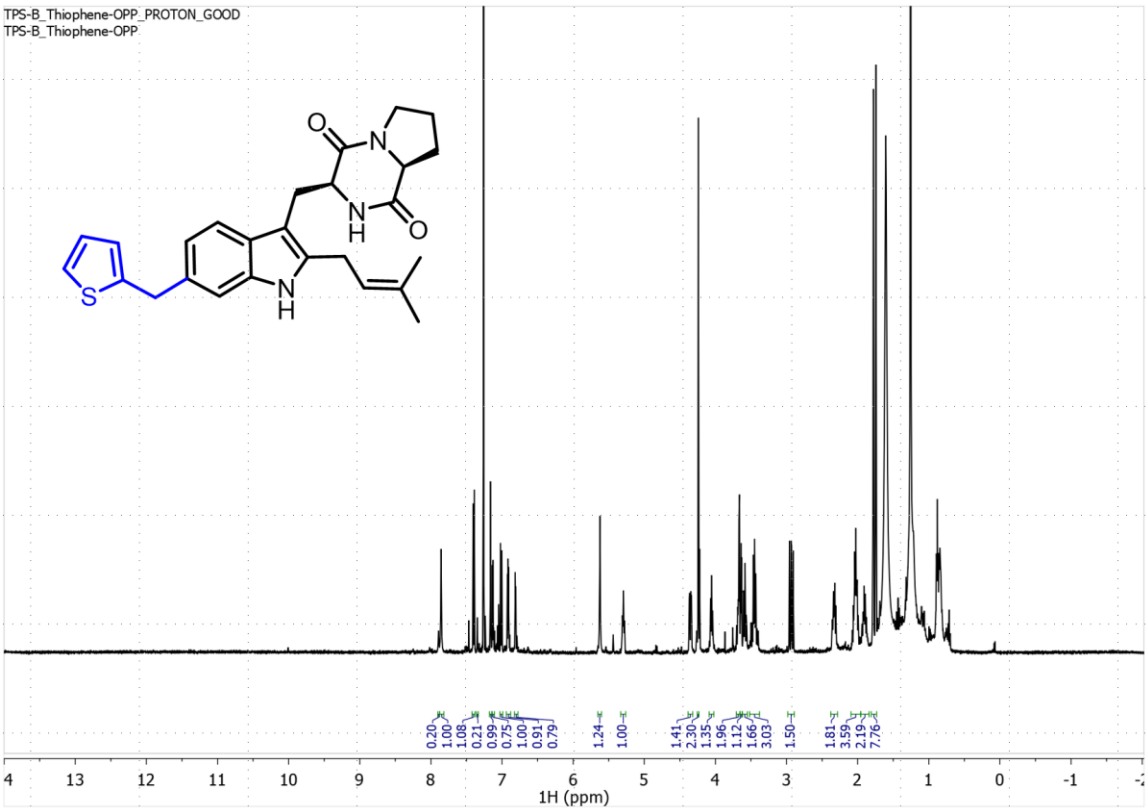
HMBC NMR spectrum (600 MHz) of **TPS-58** in $CDCl_3$



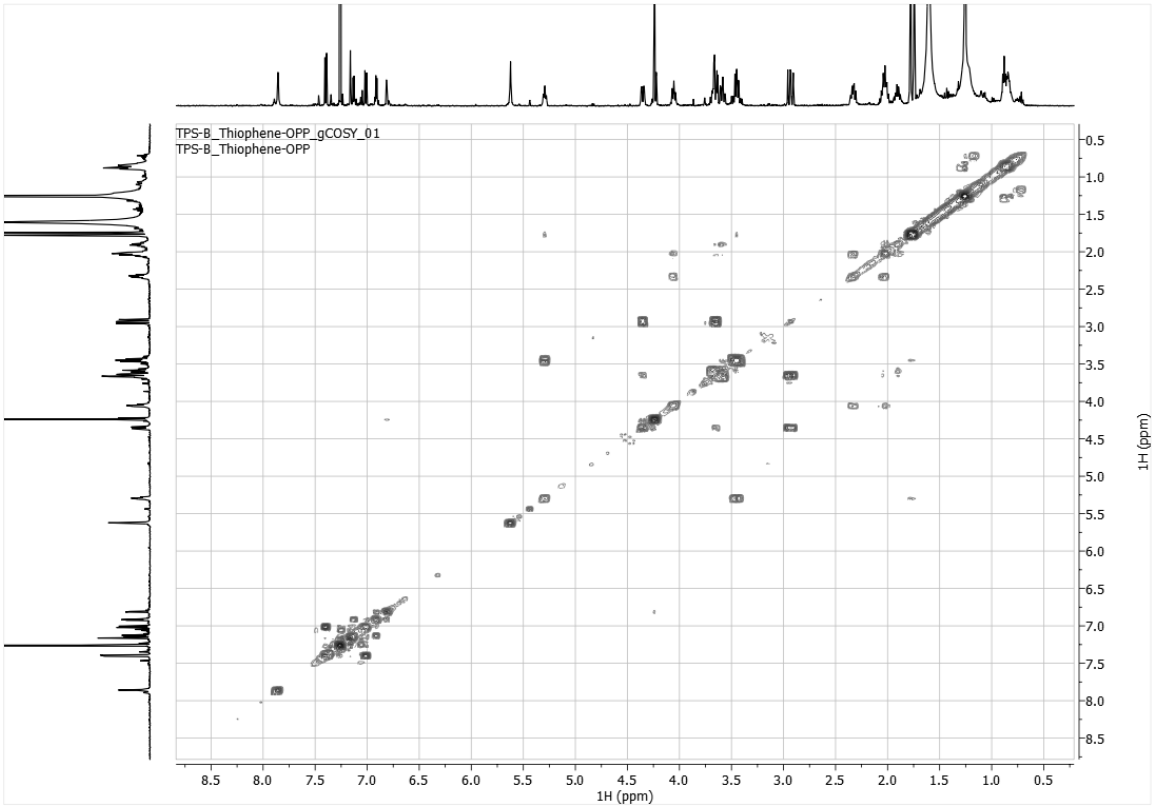
^{1}H NMR spectrum (500 MHz) of **TPS-61** in $CDCl_3$



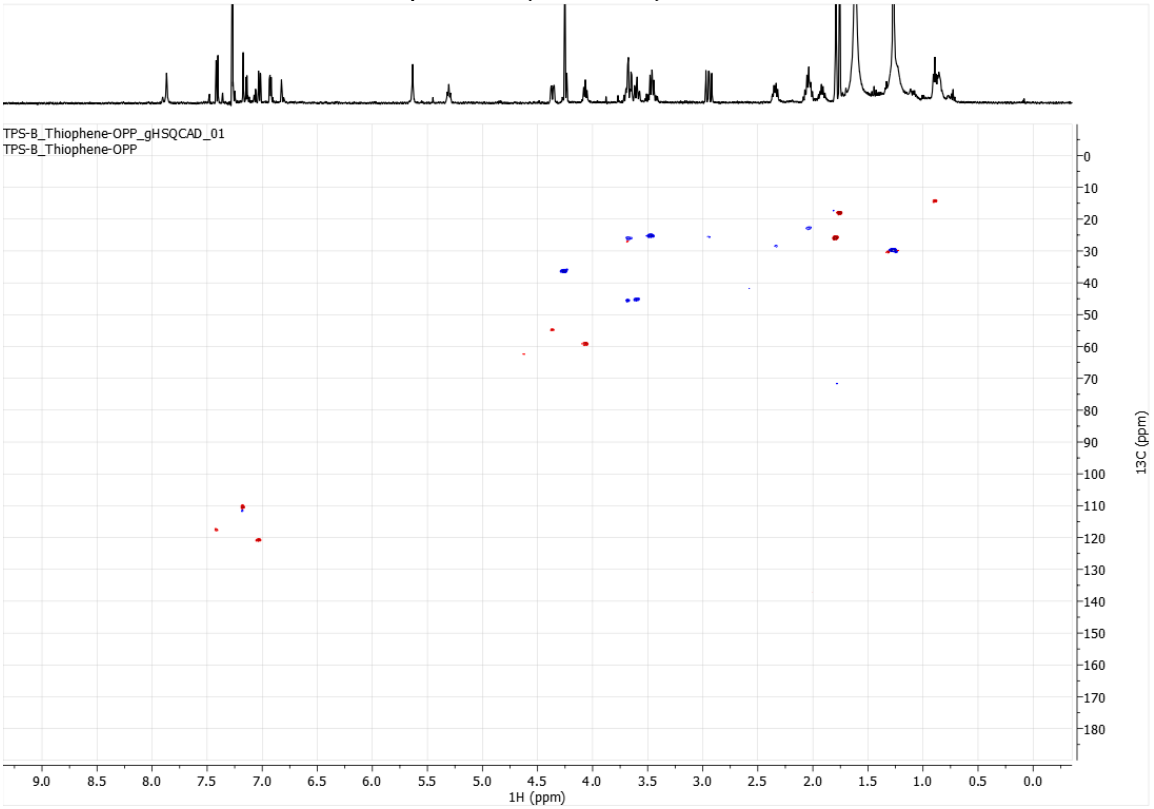
Stacked ^{1}H spectra of **TPS-61** fractions after partial purification.



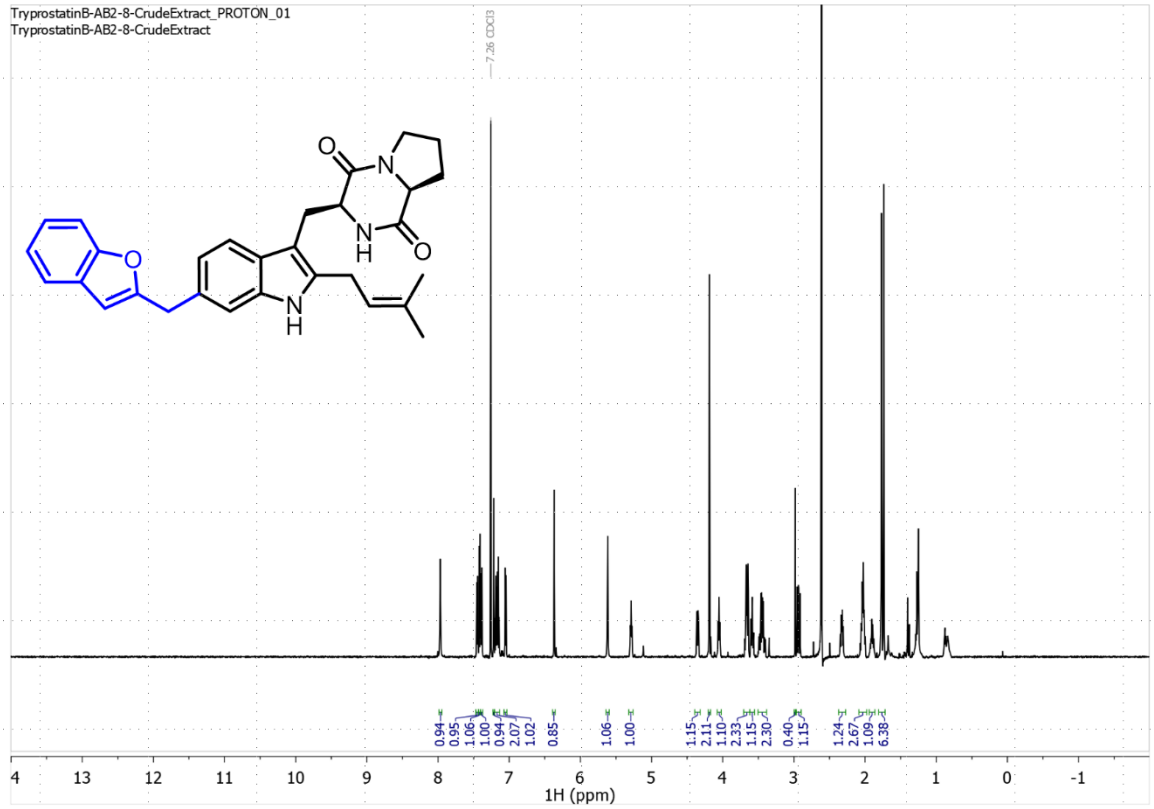
^{1}H NMR spectrum (500 MHz) of **TPS-62** in $CDCl_3$. The minor C5 alkylated product is 20% of the products based on integrations.



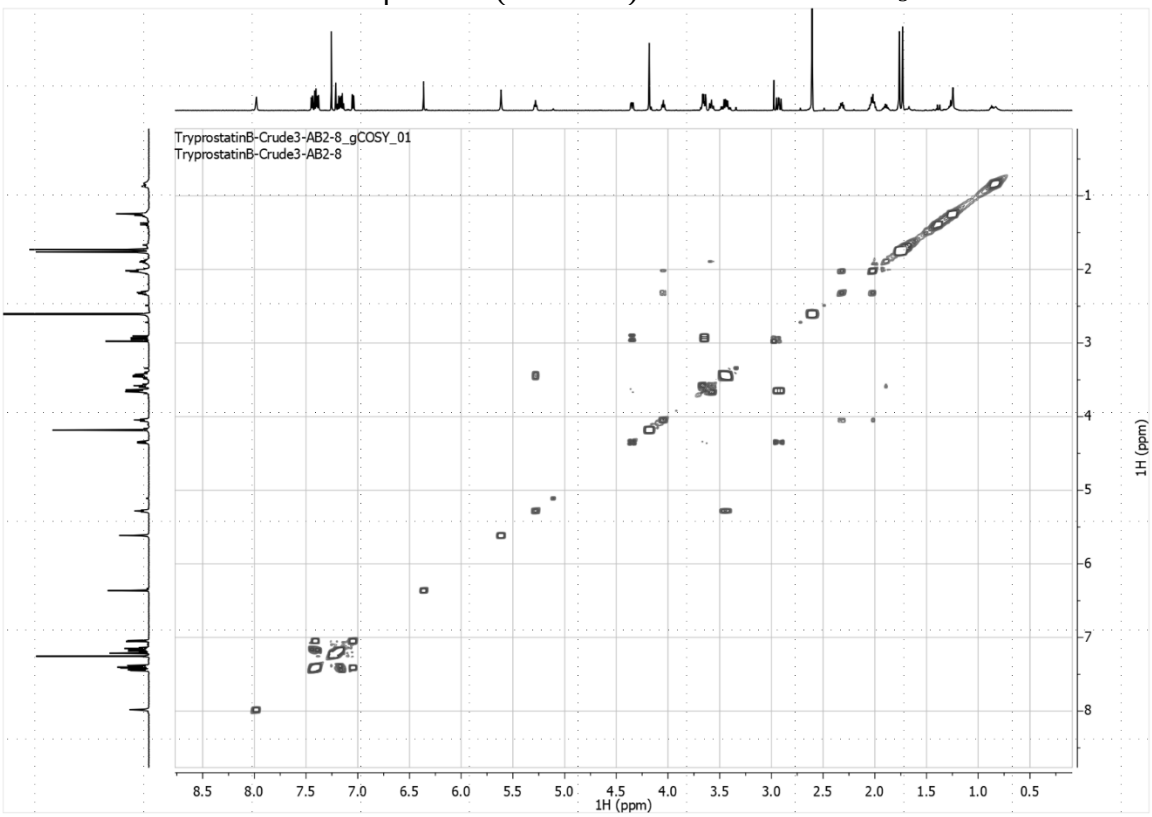
COSY NMR spectrum (500 MHz) of **TPS-62** in $CDCl_3$



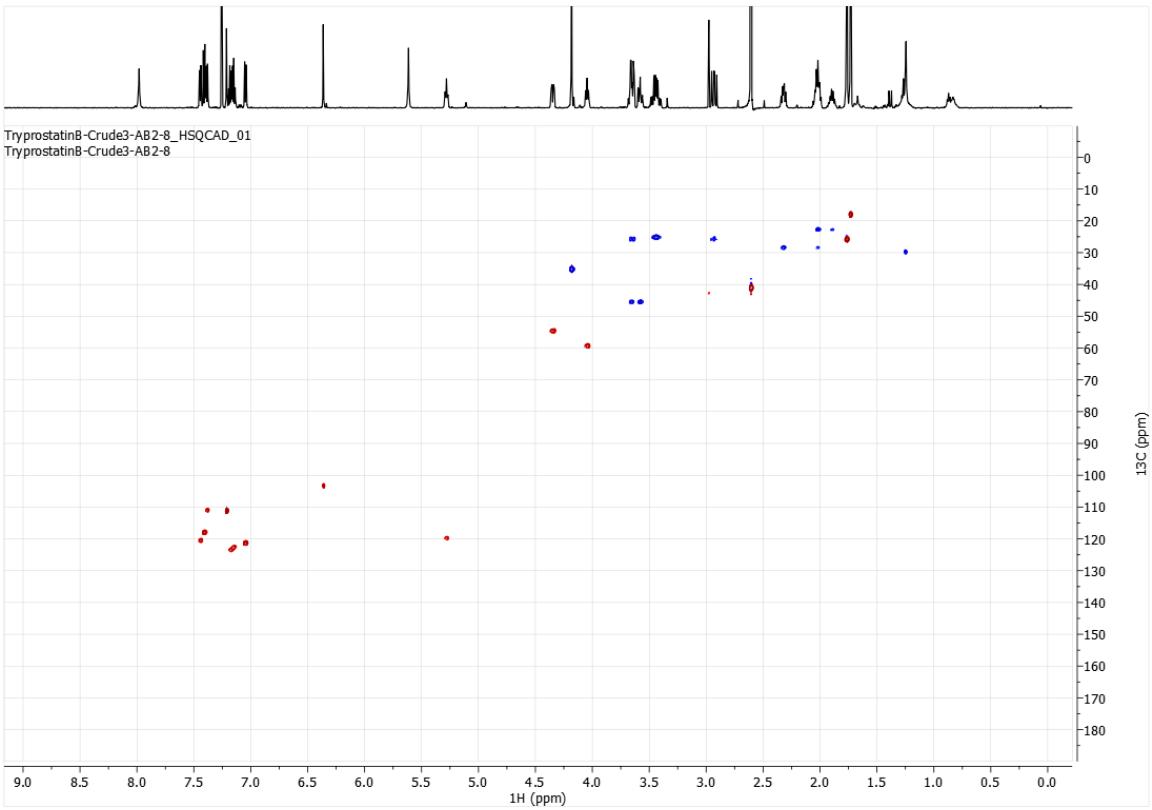
HSQC NMR spectrum (500 MHz) of **TPS-62** in $CDCl_3$



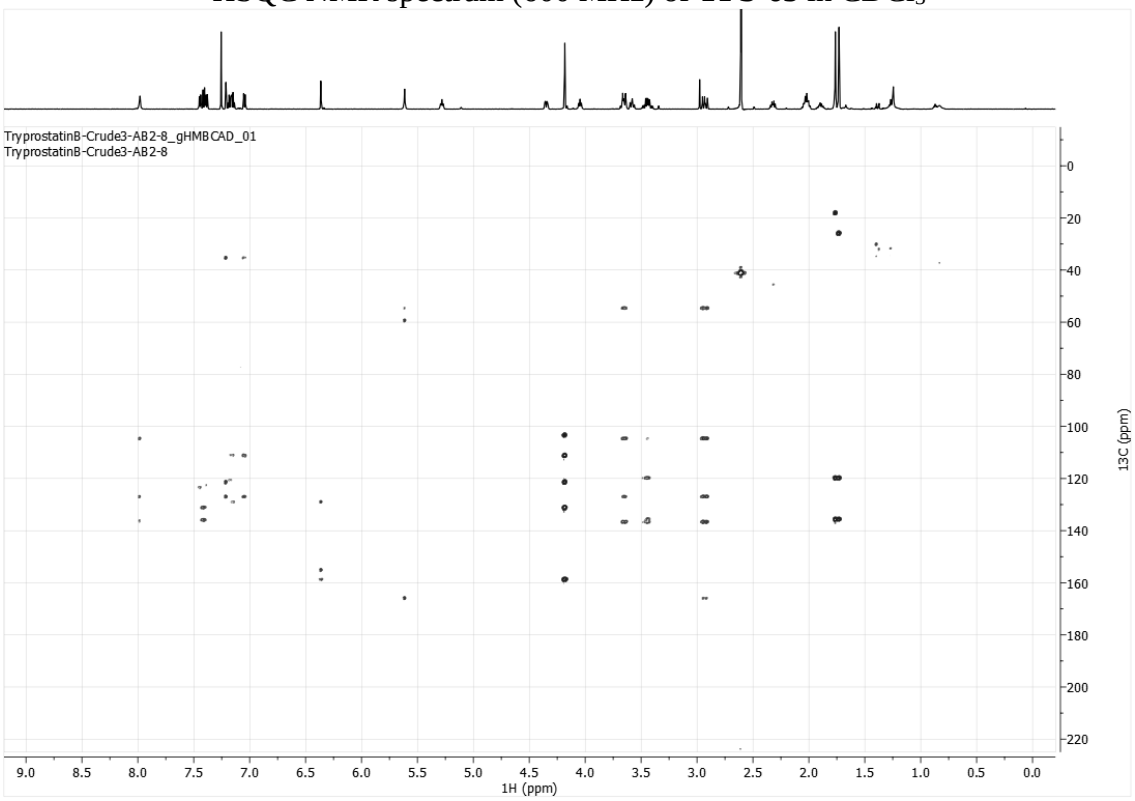
^{1}H NMR spectrum (600 MHz) of **TPS-65** in $CDCl_3$



COSY NMR spectrum (600 MHz) of **TPS-65** in $CDCl_3$



HSQC NMR spectrum (600 MHz) of **TPS-65** in $CDCl_3$

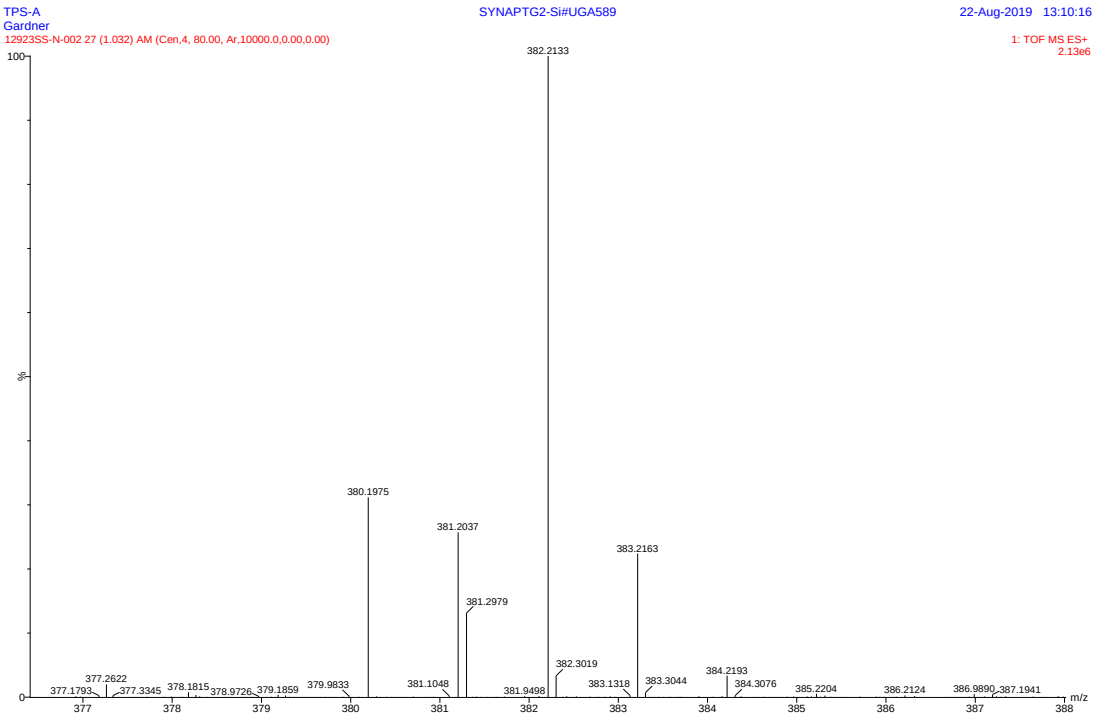


HMBC NMR spectrum (600 MHz) of **TPS-65** in $CDCl_3$

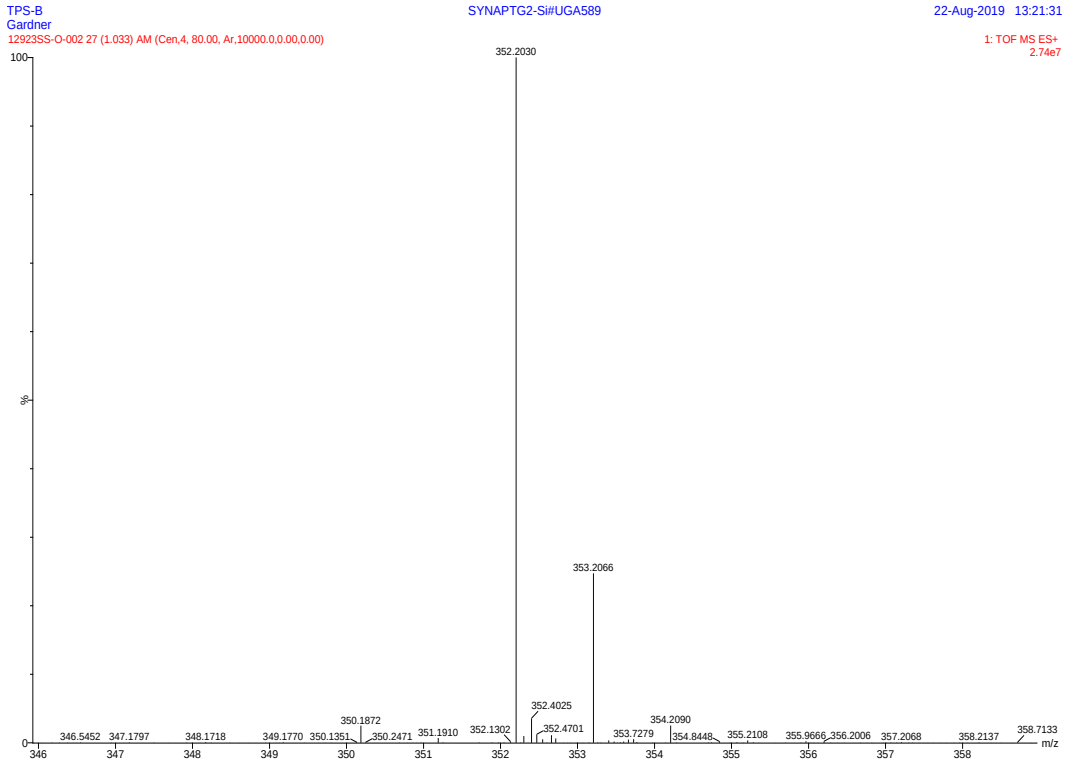
2.8 Appendix 4: HRMS Data

TPSB	Formula	Calculated	Observed
TPS-A	$C_{22}H_{27}N_3O_3$	382.2125	382.2133
TPS-B	$C_{21}H_{25}N_3O_2$	352.20194	352.203
TPS-2	$C_{26}H_{33}N_3O_2$	420.26453	420.2659
TPS-6	$C_{26}H_{31}N_3O_2$	418.241627	418.2498
TPS-7	$C_{27}H_{35}N_3O_2$	434.272927	434.2819
TPS-8	$C_{27}H_{33}N_3O_2$	432.257277	432.265
TPS-9	$C_{27}H_{33}N_3O_2$	432.257277	432.2652
TPS-10	$C_{28}H_{37}N_3O_2$	448.288577	448.2971
TPS-11	$C_{28}H_{37}N_3O_2$	448.288577	448.2964
TPS-13	$C_{28}H_{35}N_3O_2$	446.272927	446.2809
TPS-16	$C_{27}H_{33}N_3O_2$	432.257277	432.2653
TPS-18	$C_{28}H_{35}N_3O_2$	446.272927	446.2810
TPS-19	$C_{29}H_{39}N_3O_2$	462.304226	462.3118
TPS-22	$C_{29}H_{37}N_3O_2$	460.288577	460.2964
TPS-24	$C_{30}H_{33}N_3O_2$	468.26453	468.2651
TPS-25	$C_{31}H_{35}N_3O_2$	482.28018	482.2802
TPS-26	$C_{31}H_{35}N_3O_3$	498.267842	498.2758
TPS-27	$C_{29}H_{35}N_3O_2$	458.272927	458.2814
TPS-34	$C_{27}H_{34}N_6O_2$	475.274324	475.2823
TPS-46	$C_{29}H_{33}N_3O_2$	456.26453	456.2655
TPS-52	$C_{29}H_{33}N_3O_3$	472.25945	472.2602
TPS-58	$C_{29}H_{31}N_3O_4$	486.23871	486.2389
TPS-61-	$C_{26}H_{29}N_3O_3$	432.22815	432.2292
TPS-61-	$C_{26}H_{29}N_3O_3$	432.22815	432.2294
TPS-62	$C_{26}H_{29}N_3O_2S$	448.20531	448.2061
TPS-65	$C_{30}H_{31}N_3O_3$	482.2438	482.2442
TPS-66	$C_{30}H_{31}N_3O_2S$	498.213699	498.2216

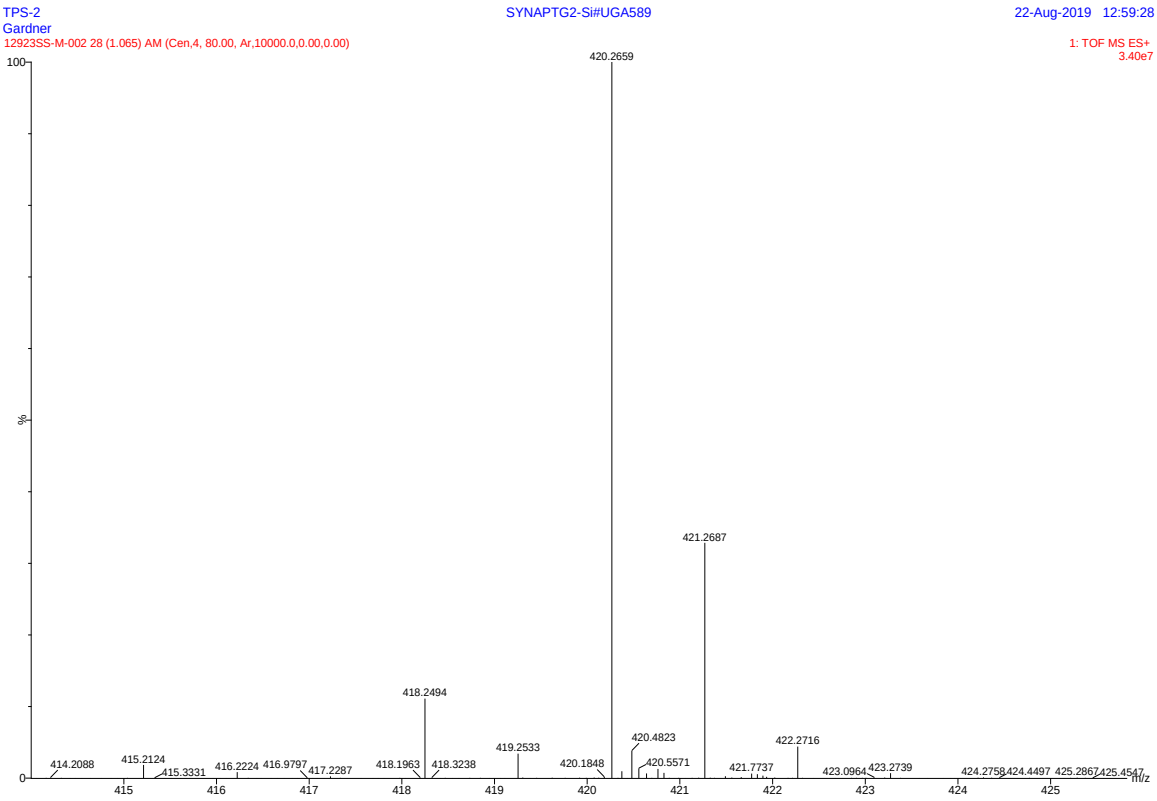
Table 2.6: HRMS data of the TPS analogs produced by CdpNPT and their corresponding alkyl donor



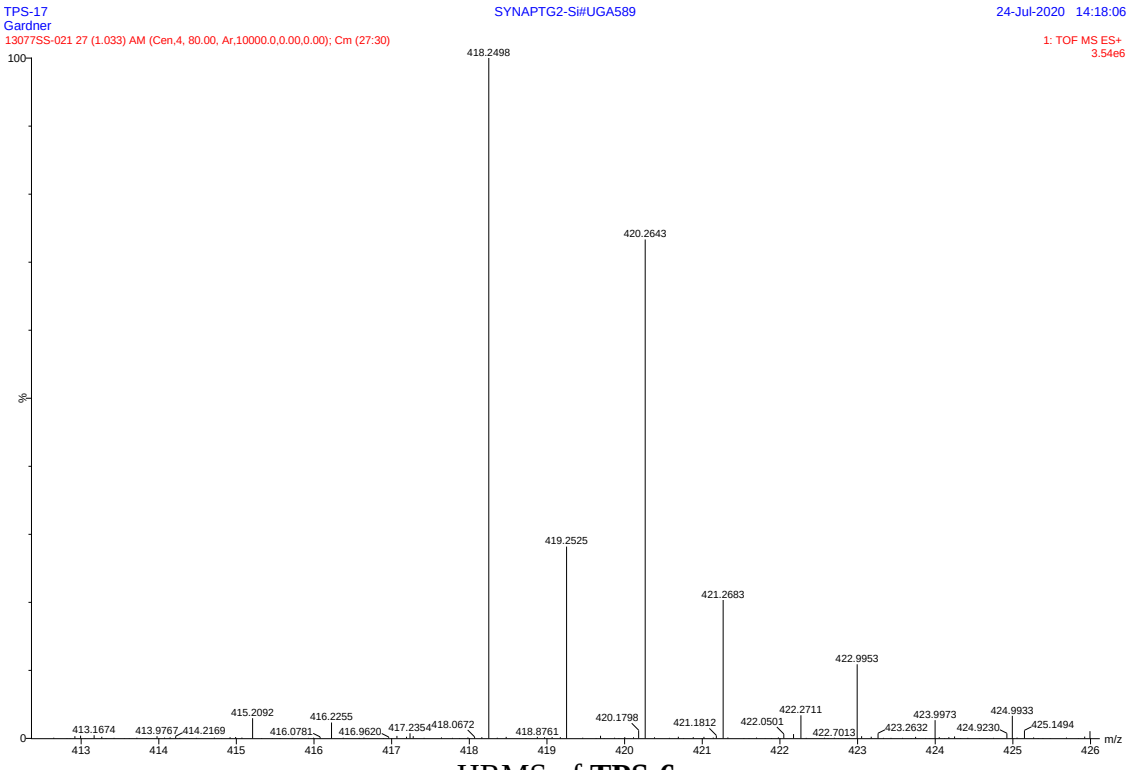
HRMS of **TPS-A**



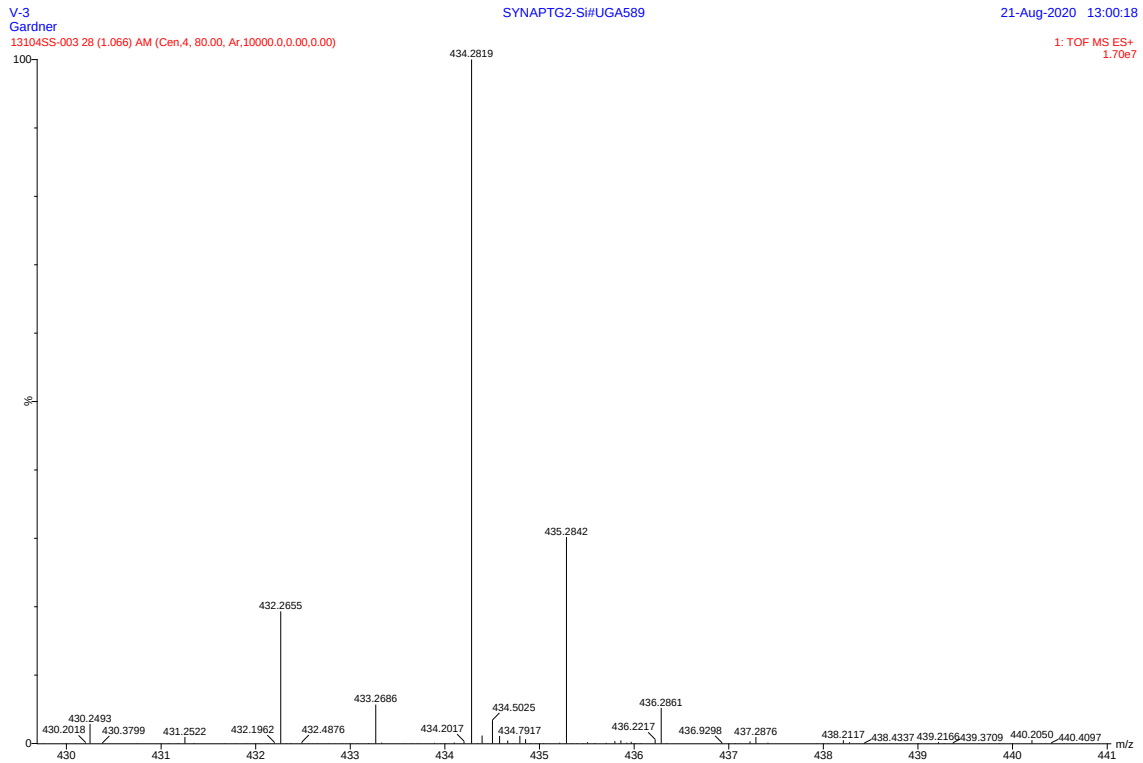
HRMS of **TPS-B**



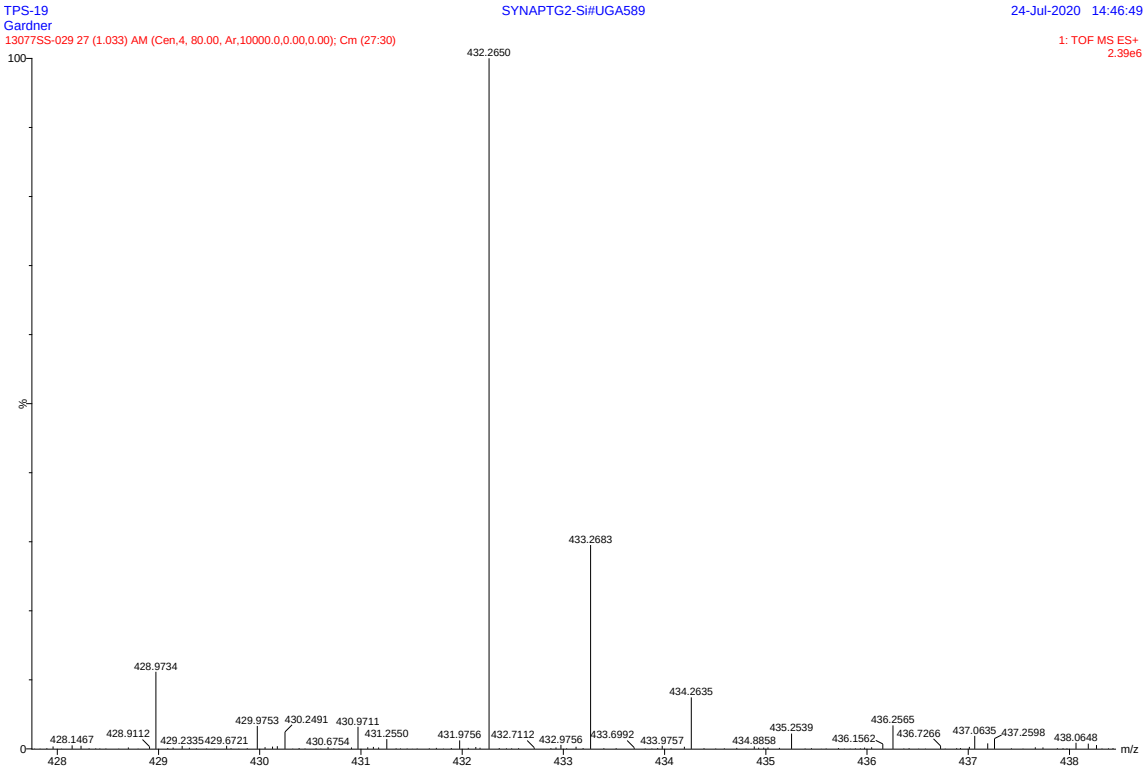
HRMS of **TPS-2**



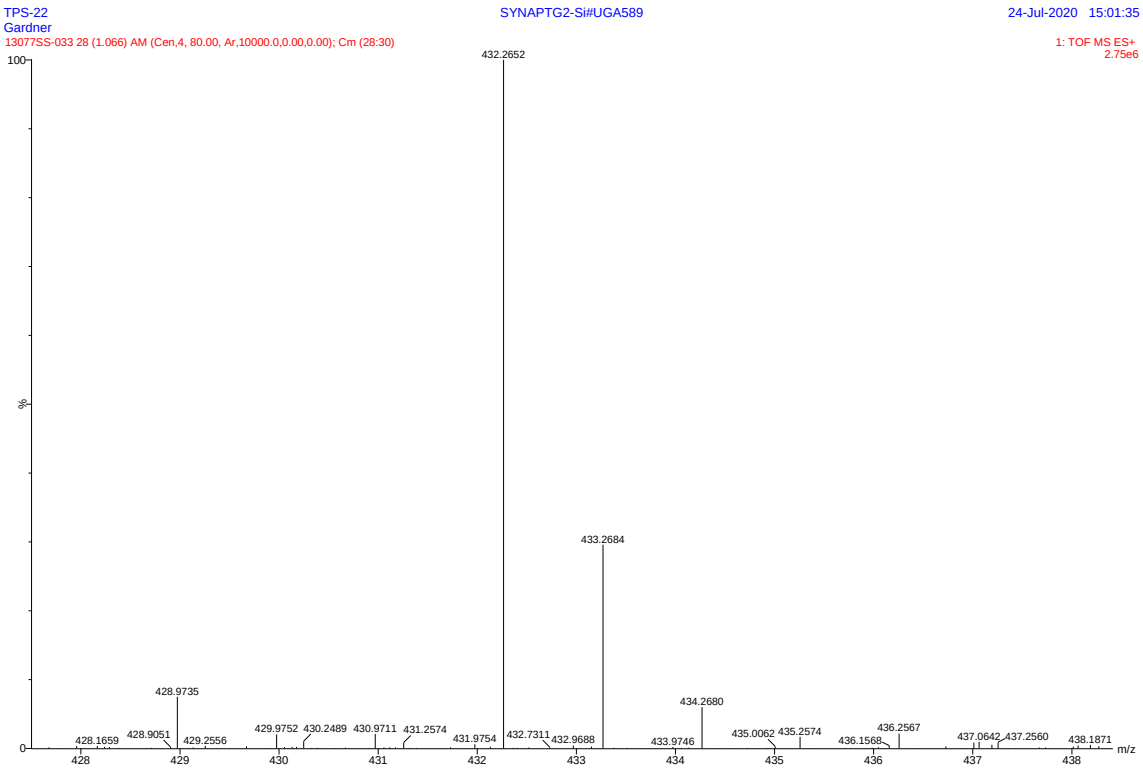
HRMS of **TPS-6**



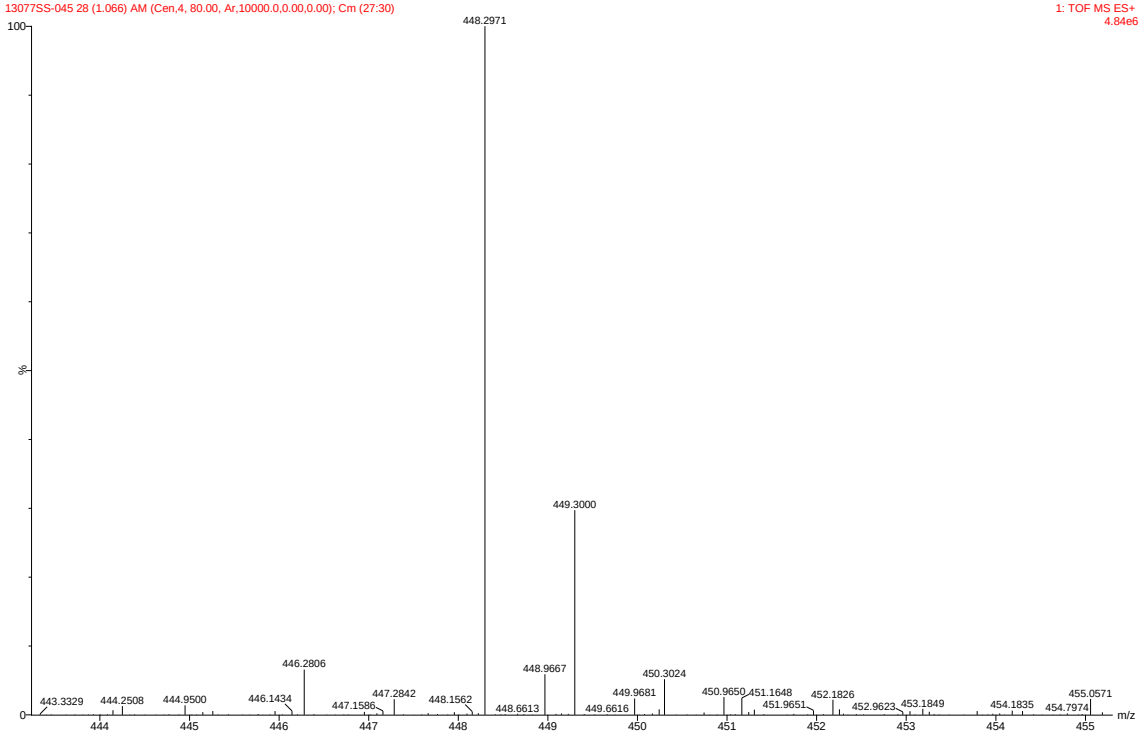
HRMS of **TPS-7**



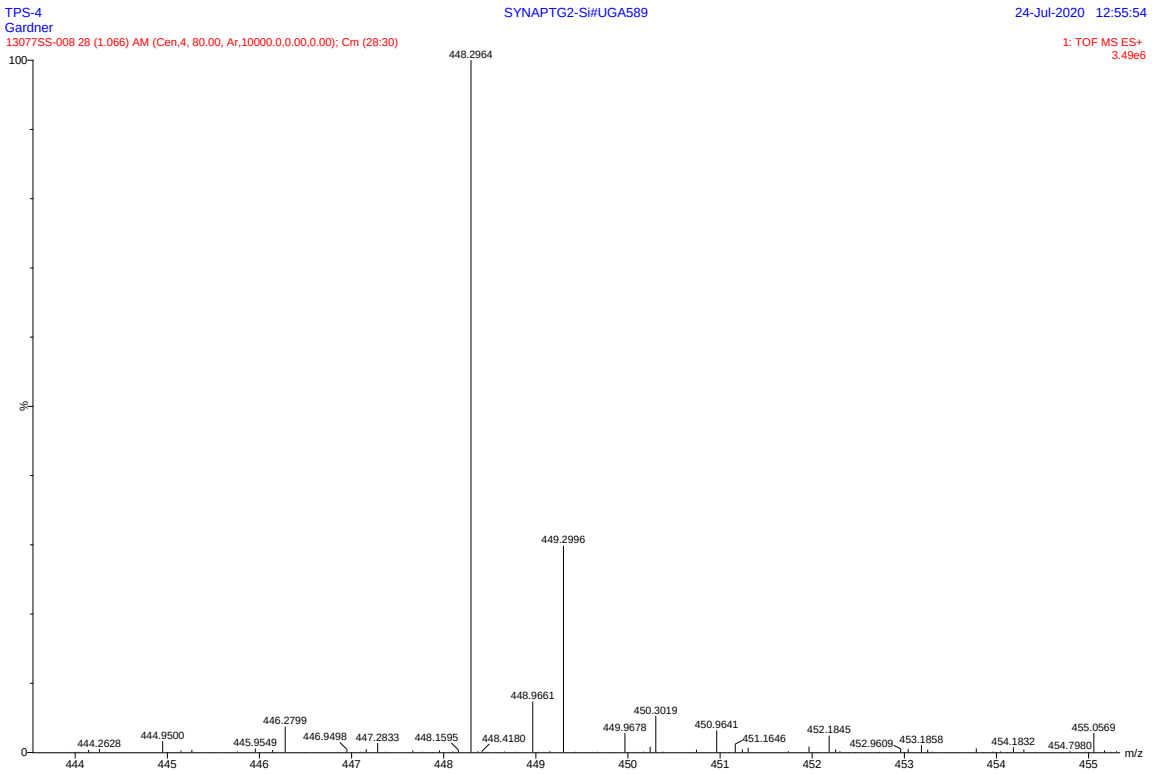
HRMS of **TPS-8**



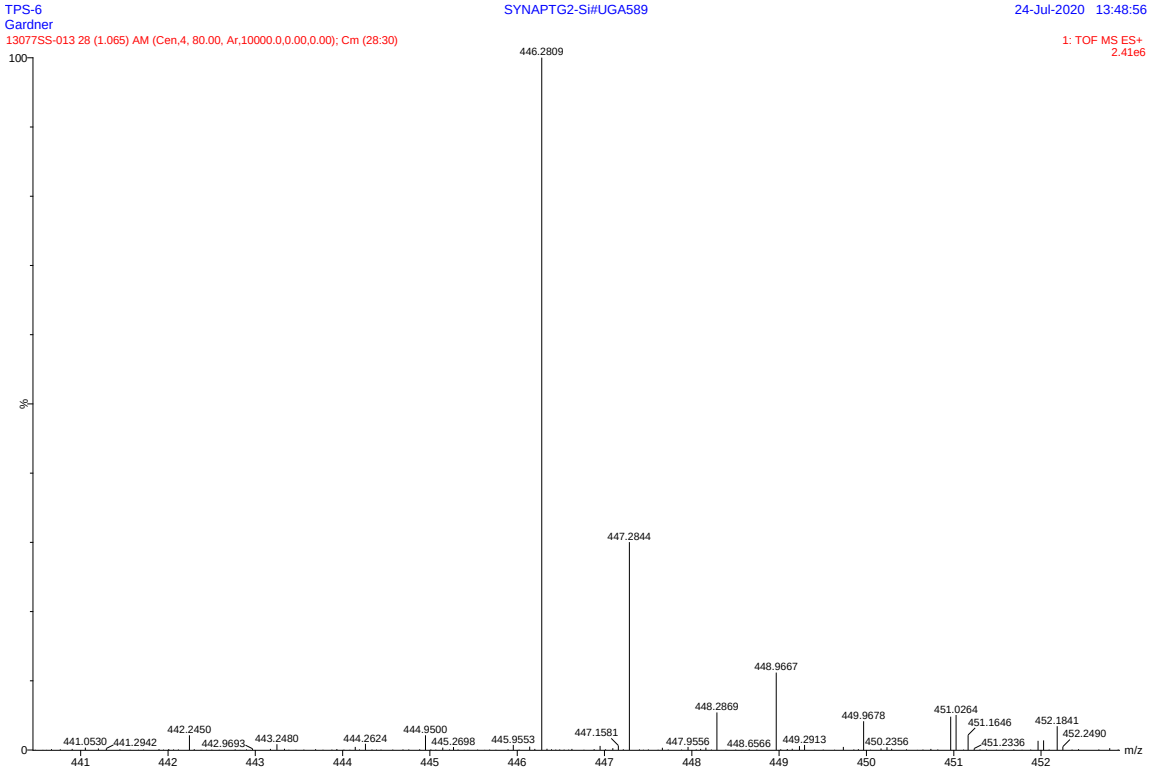
HRMS of **TPS-9**



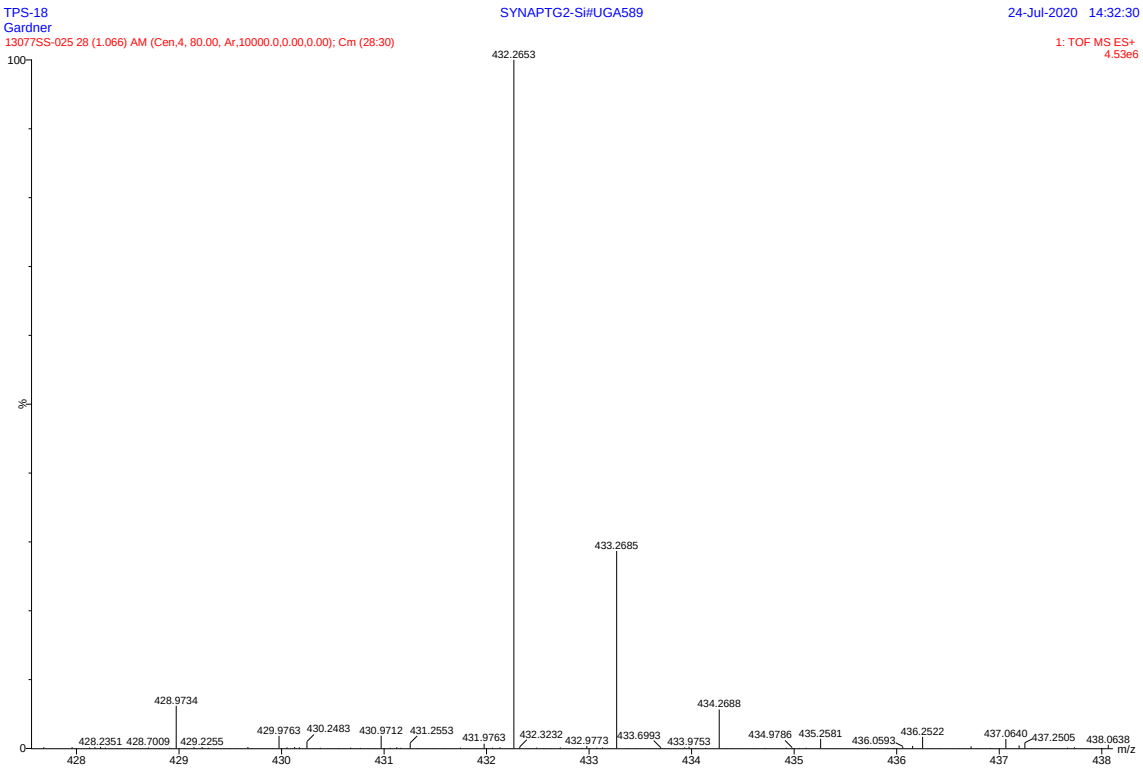
HRMS of **TPS-10**



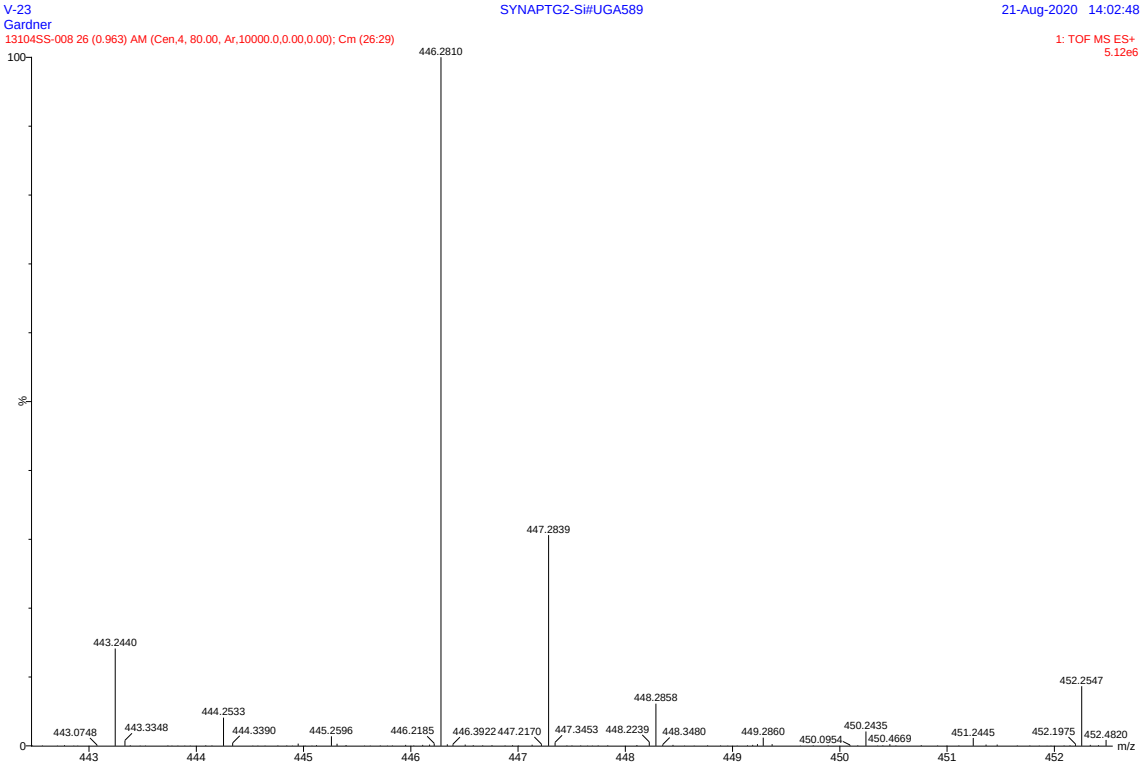
HRMS of **TPS-11**



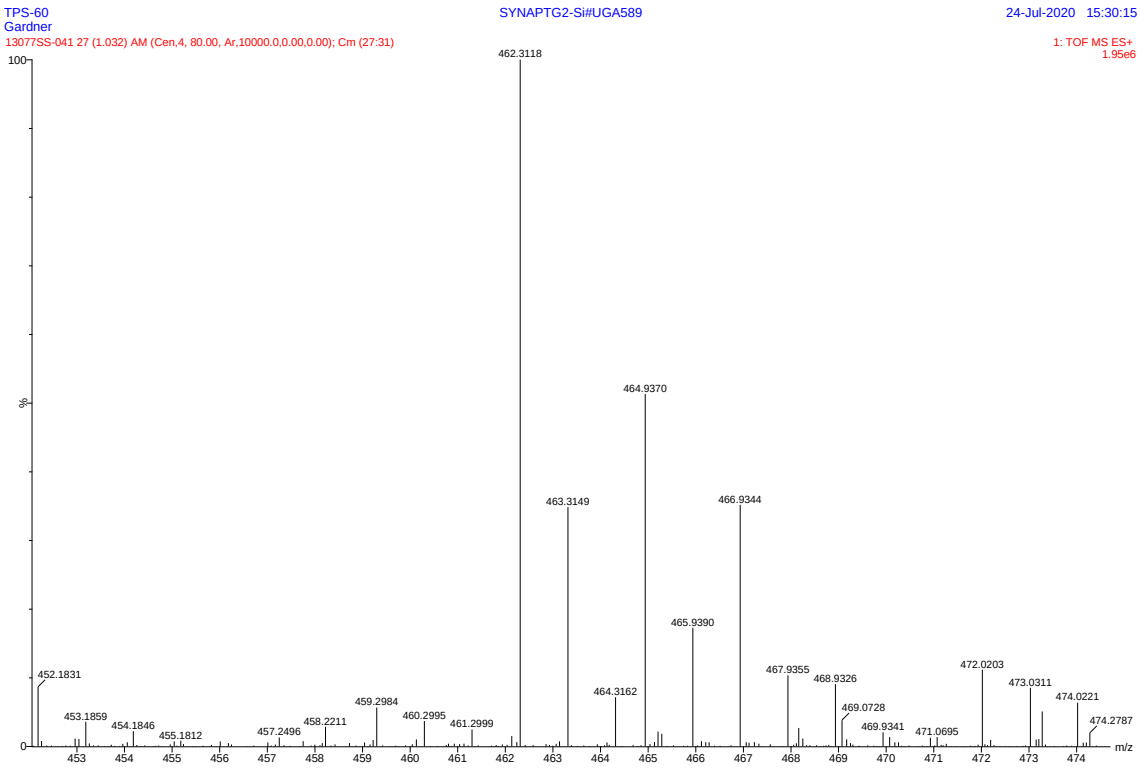
HRMS of **TPS-13**



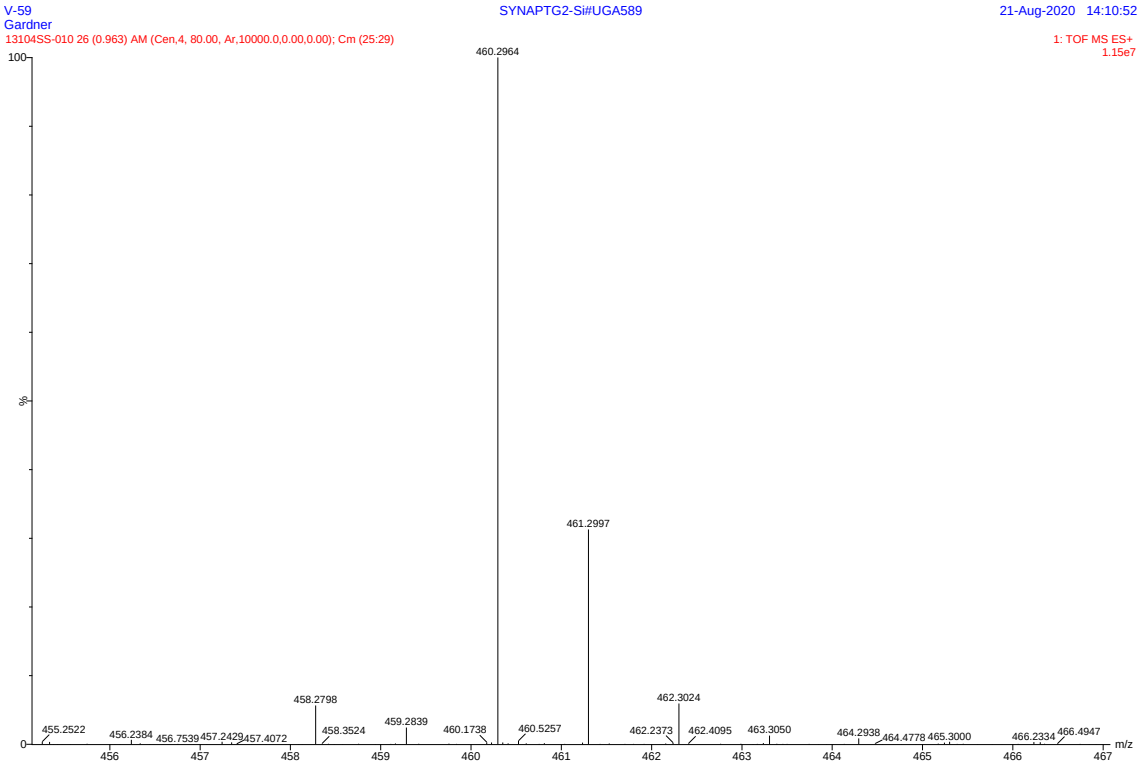
HRMS of **TPS-16**



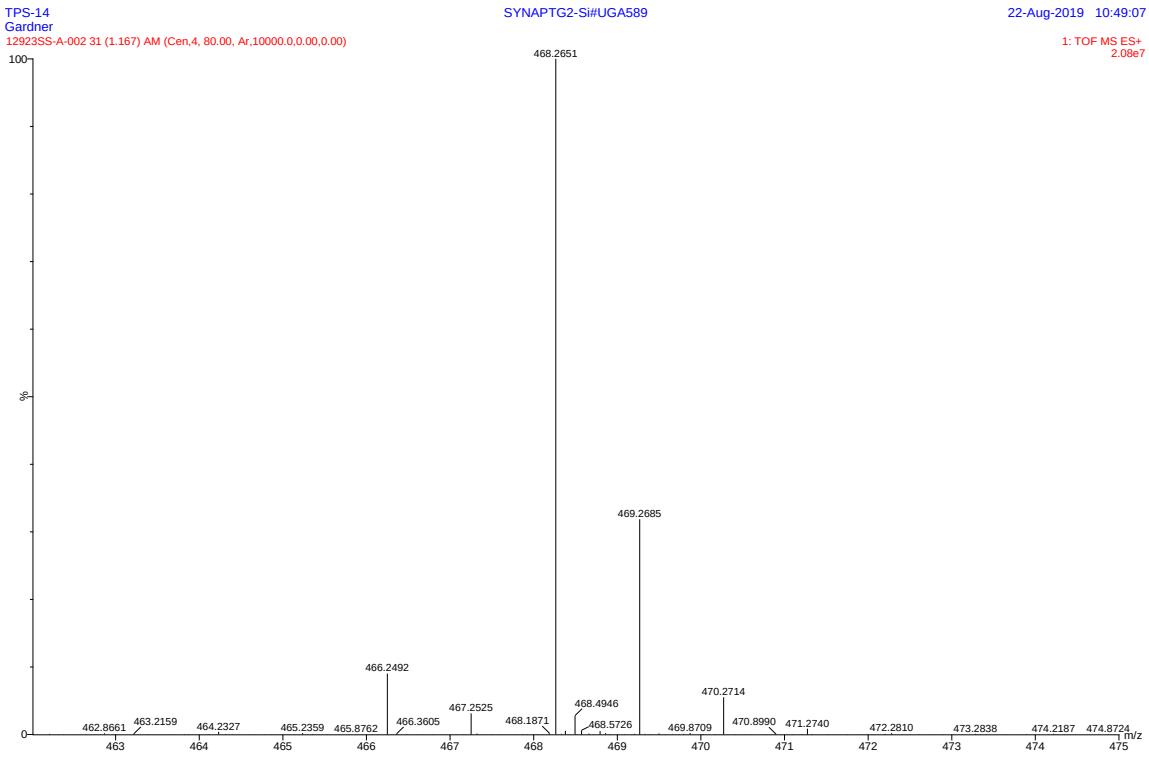
HRMS of **TPS-18**



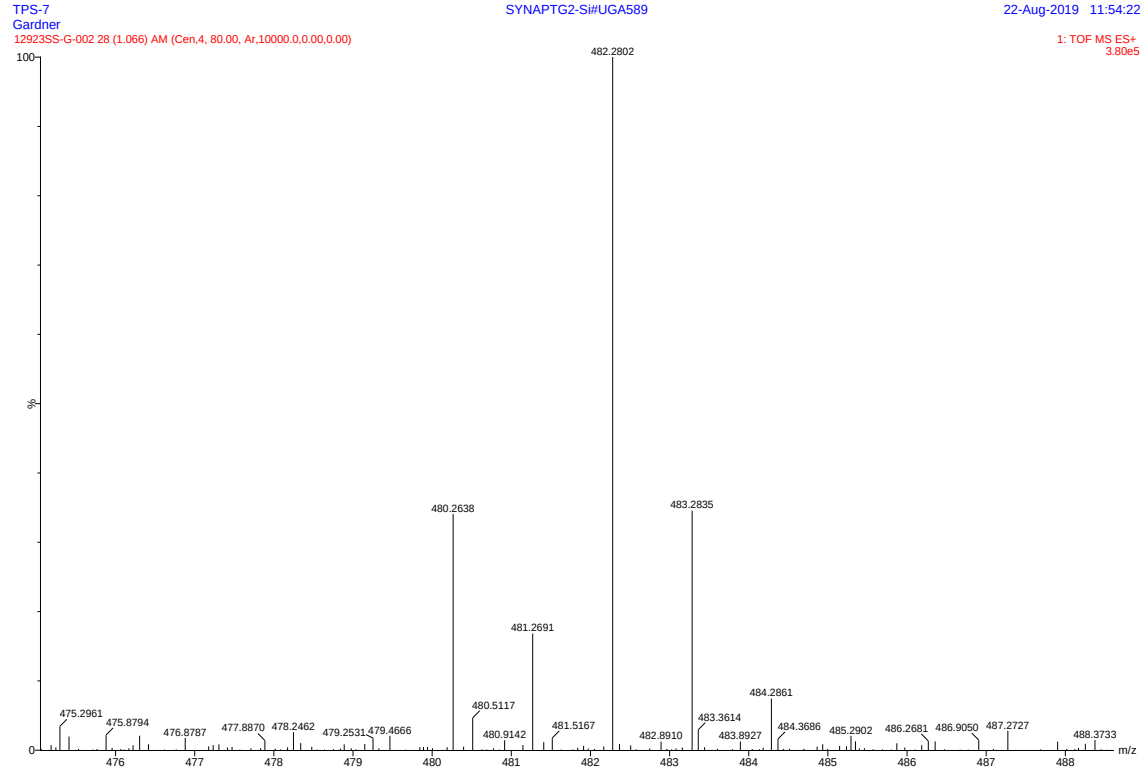
HRMS of **TPS-19**



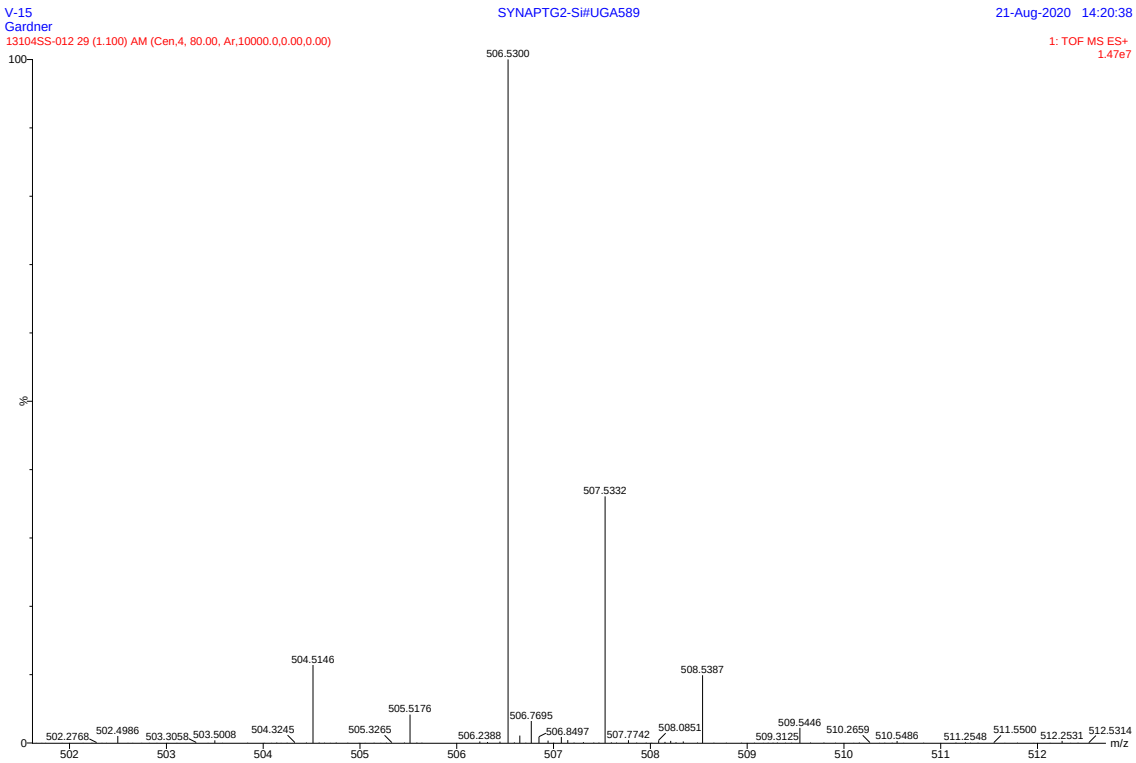
HRMS of **TPS-22**



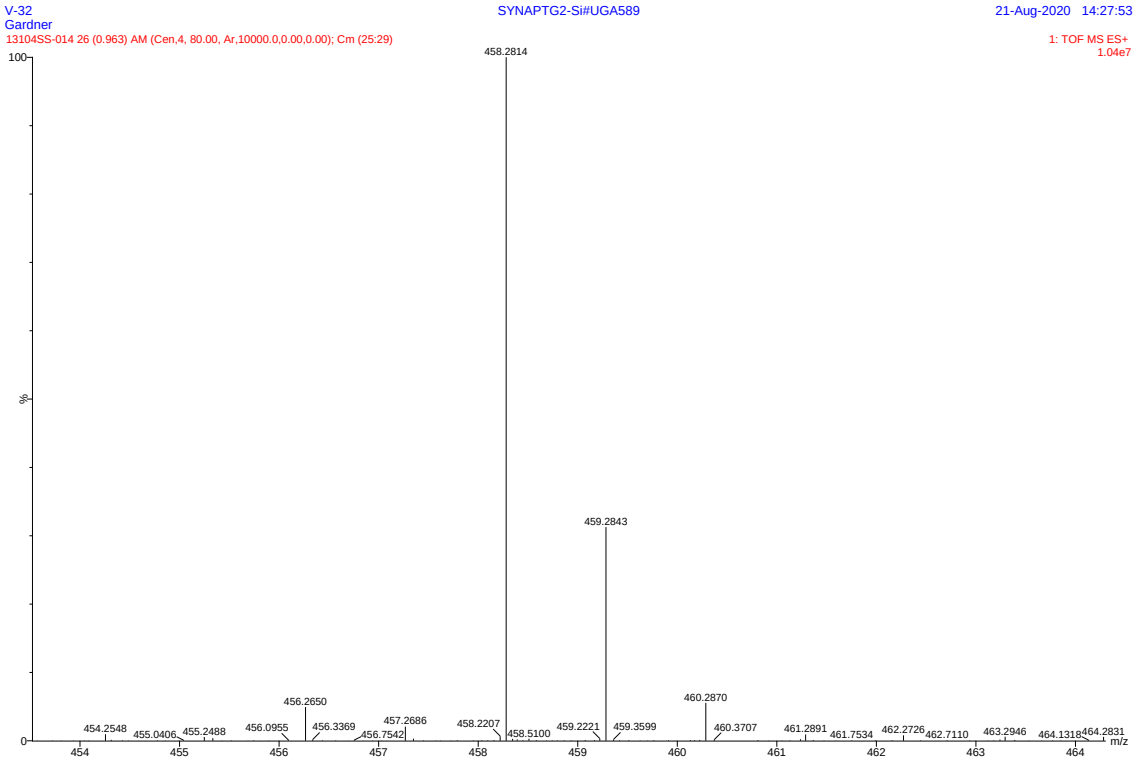
HRMS of **TPS-24**



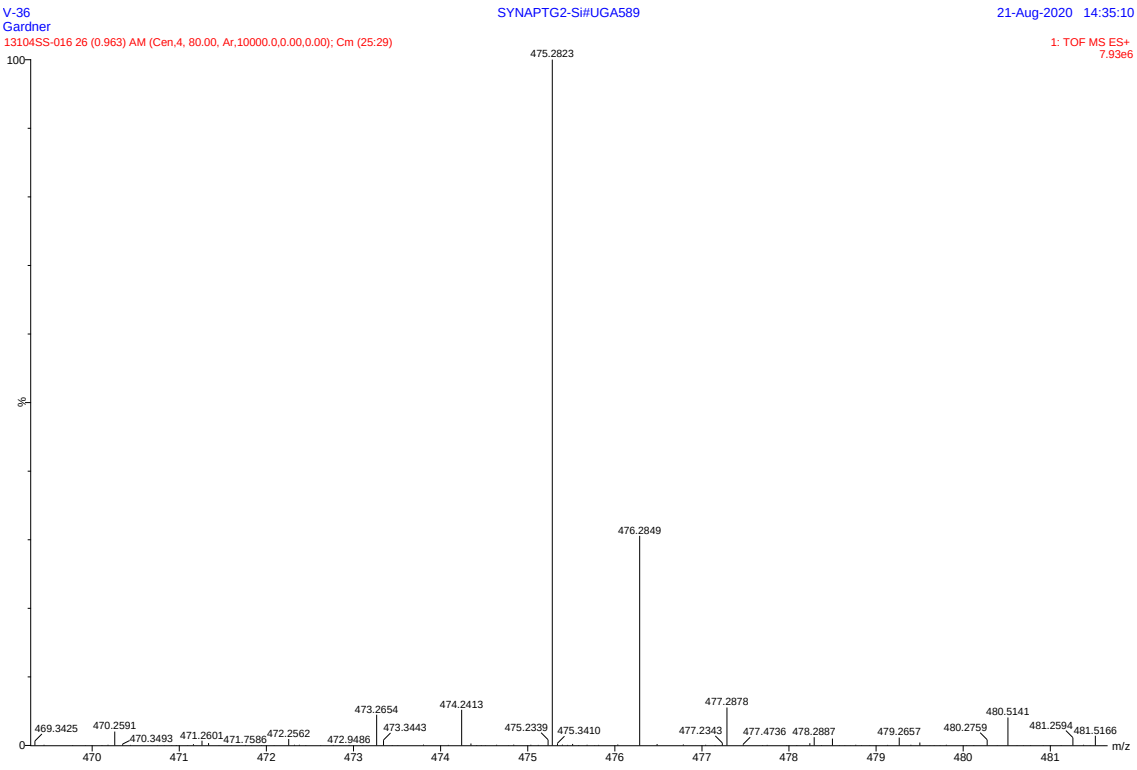
HRMS of **TPS-25**



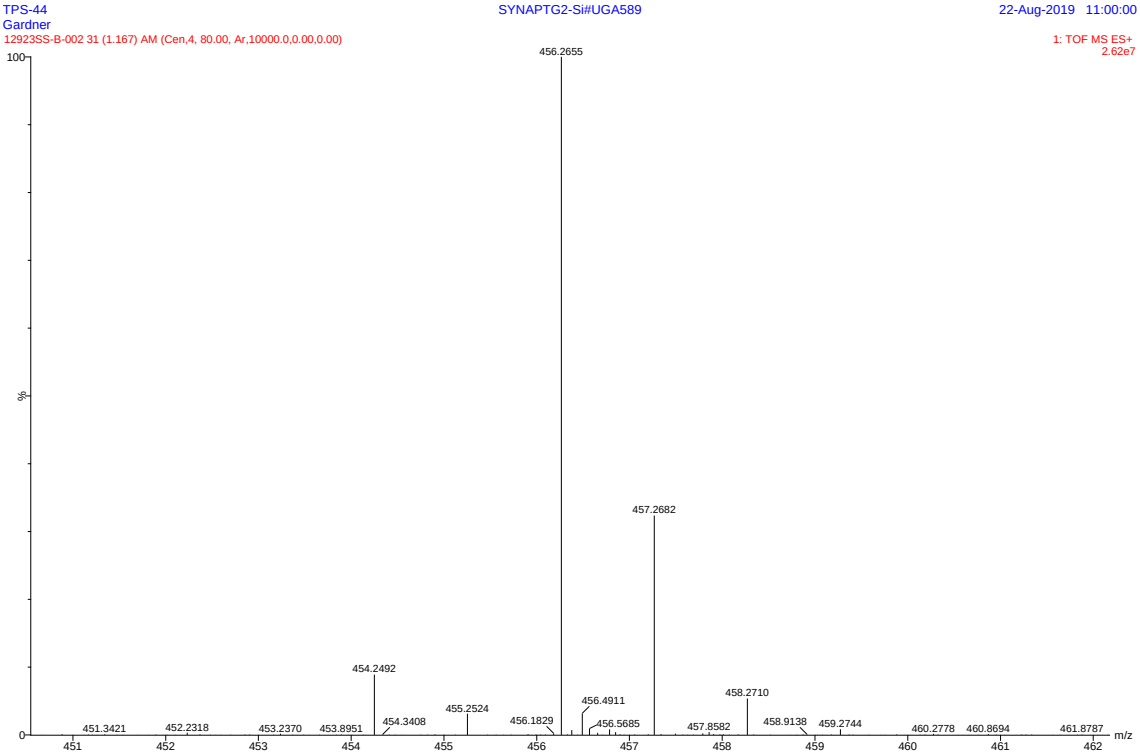
HRMS of **TPS-26**



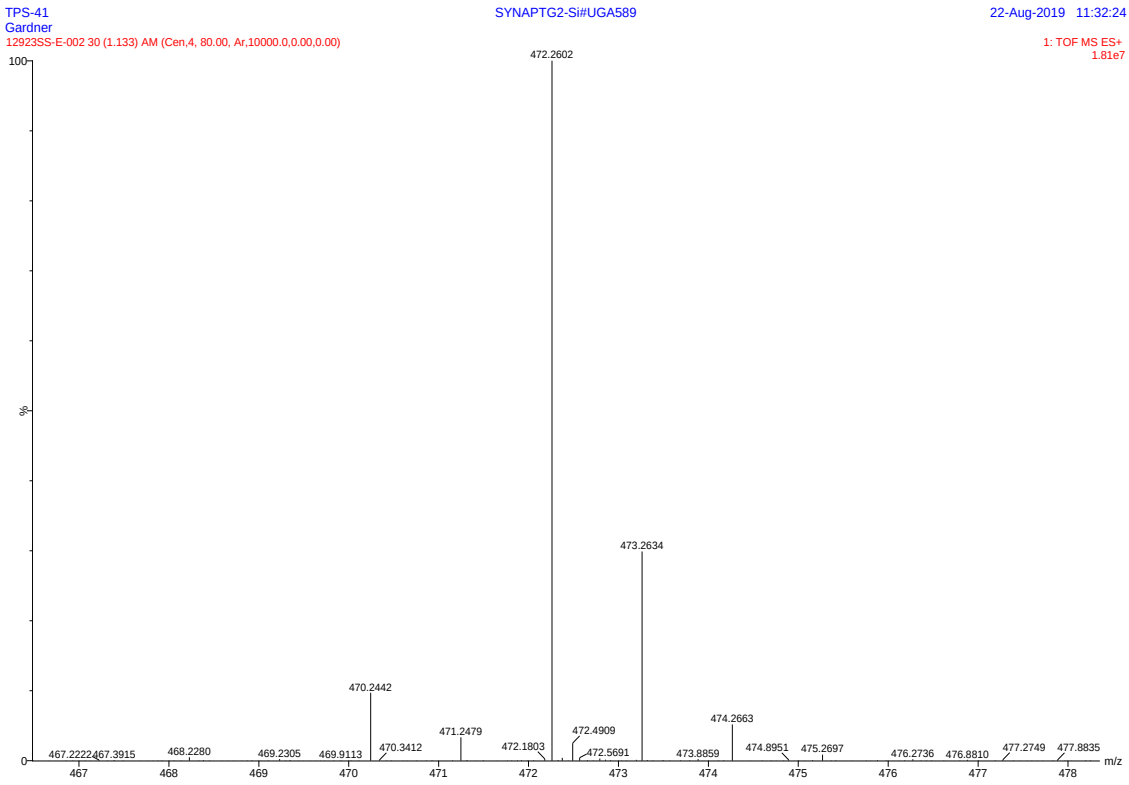
HRMS of **TPS-27**



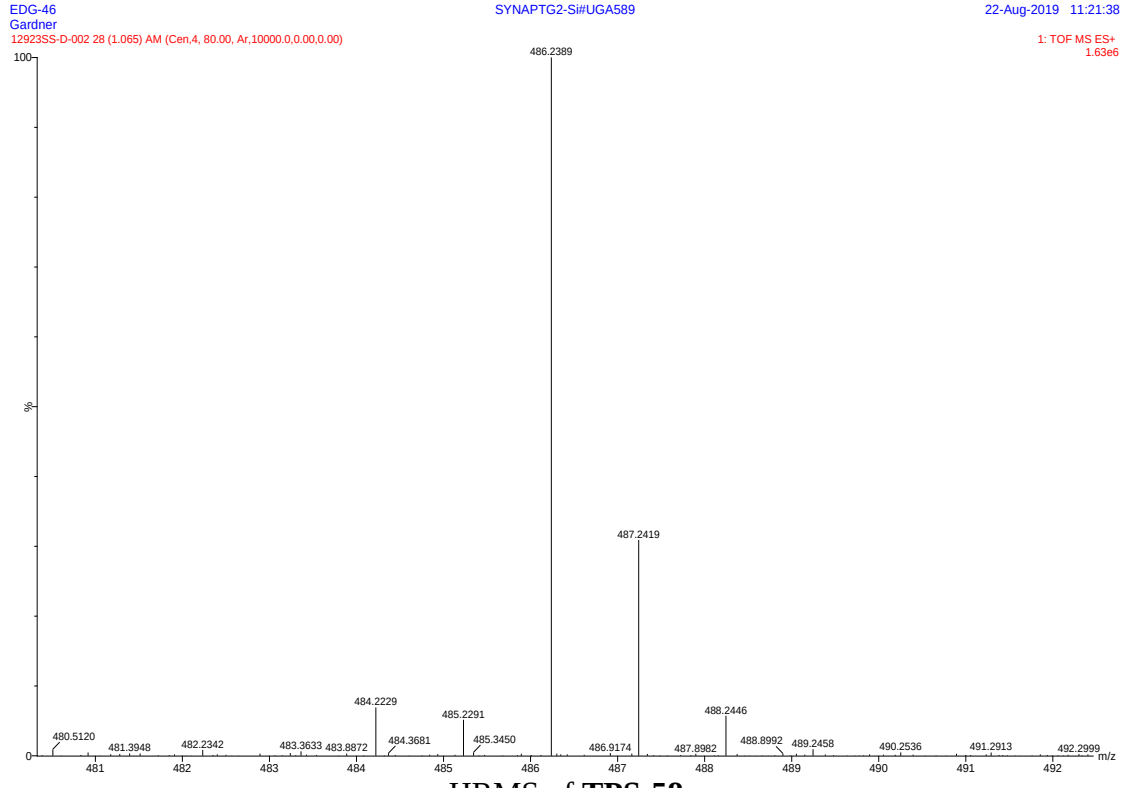
HRMS of **TPS-34**



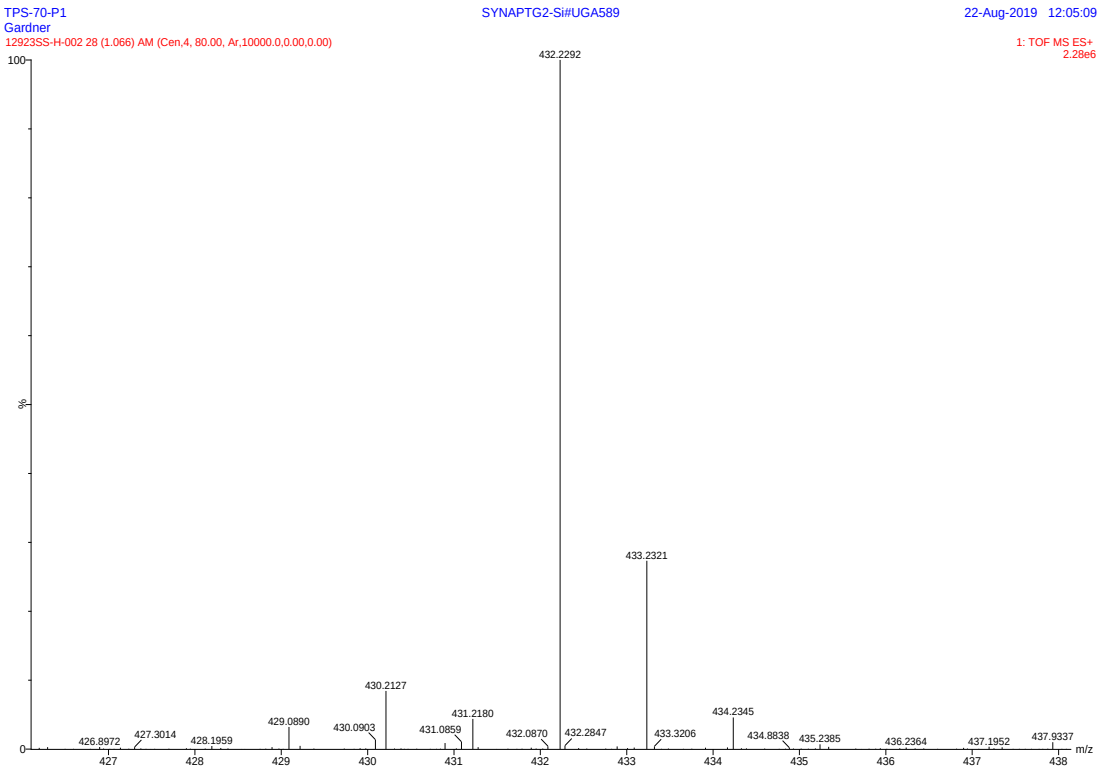
HRMS of **TPS-46**



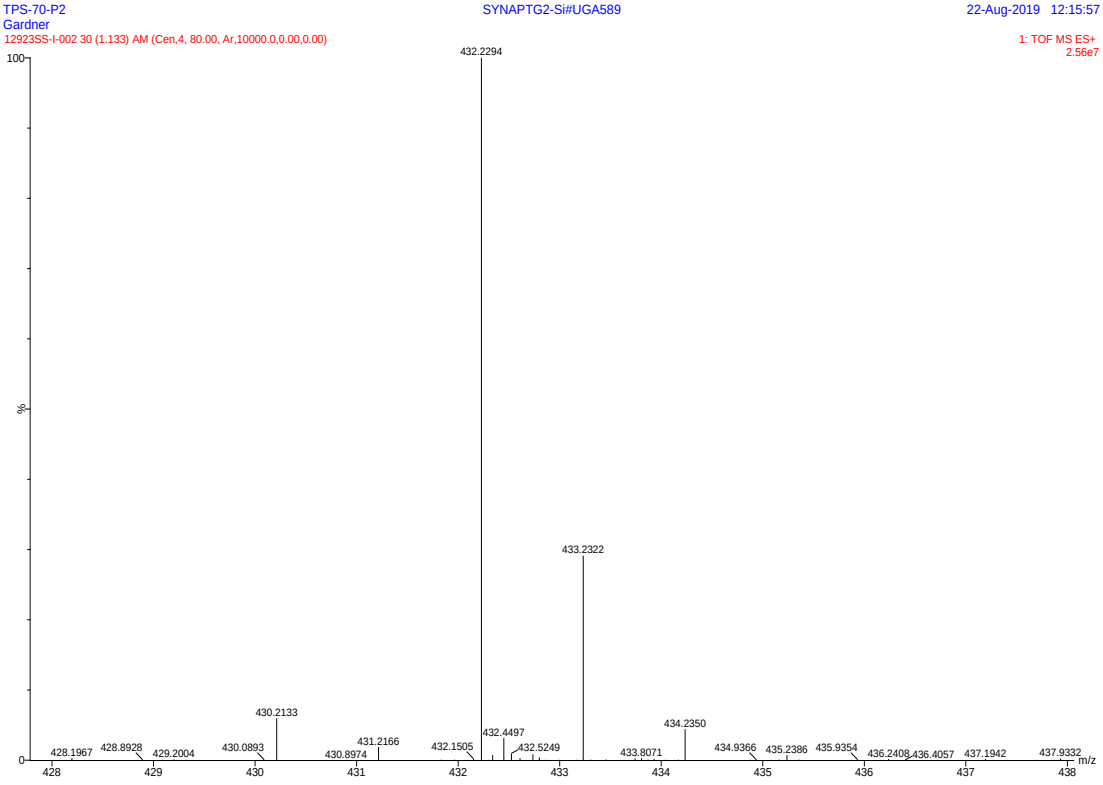
HRMS of **TPS-52**



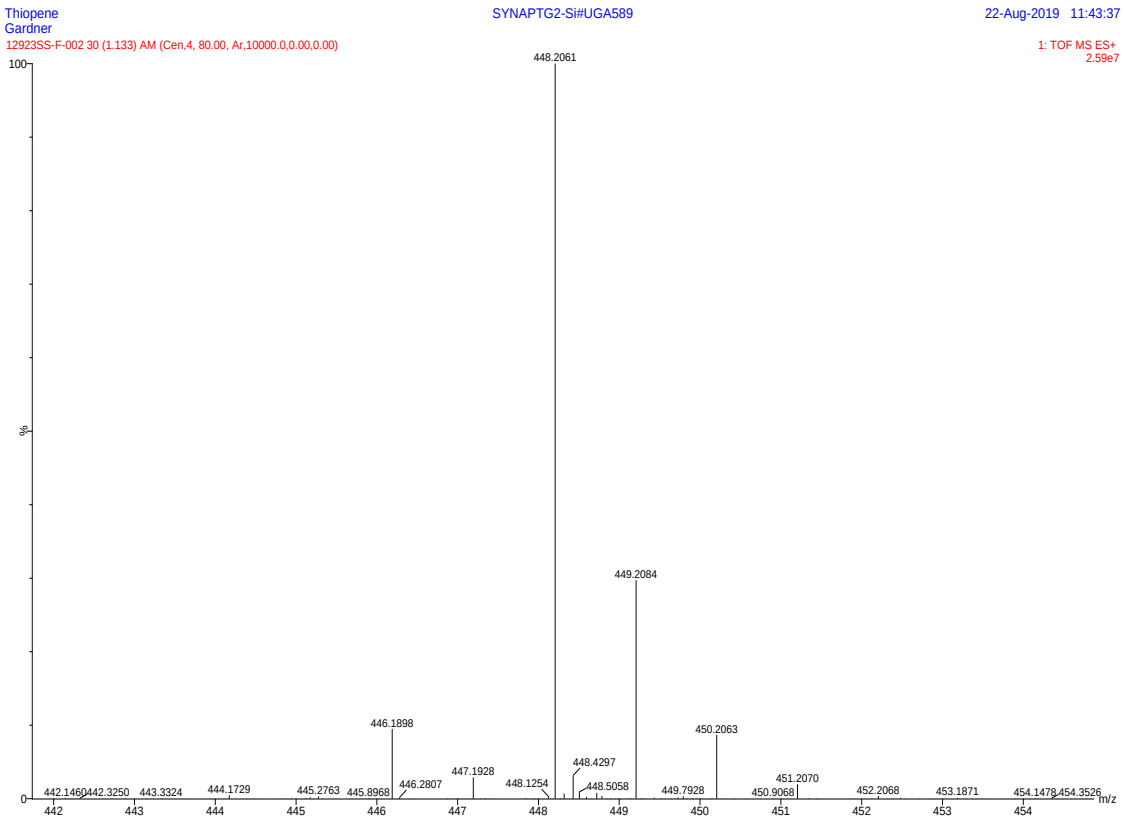
HRMS of **TPS-58**



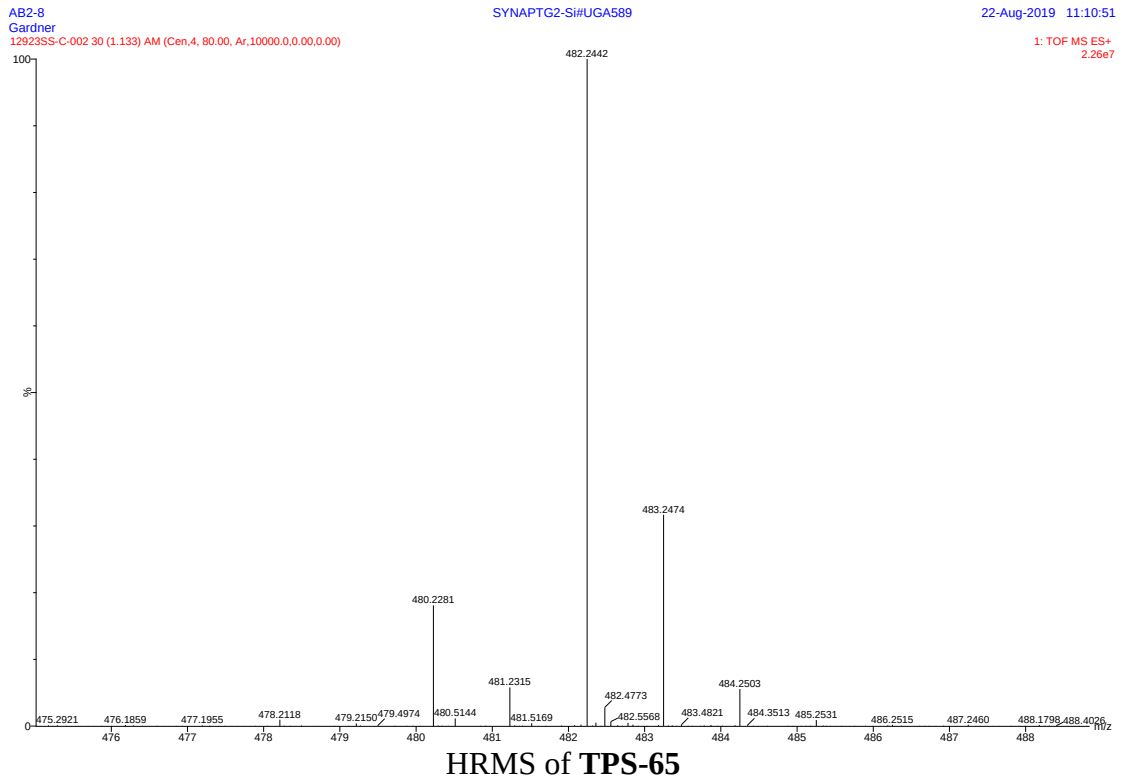
HRMS of **TPS-61-P1**



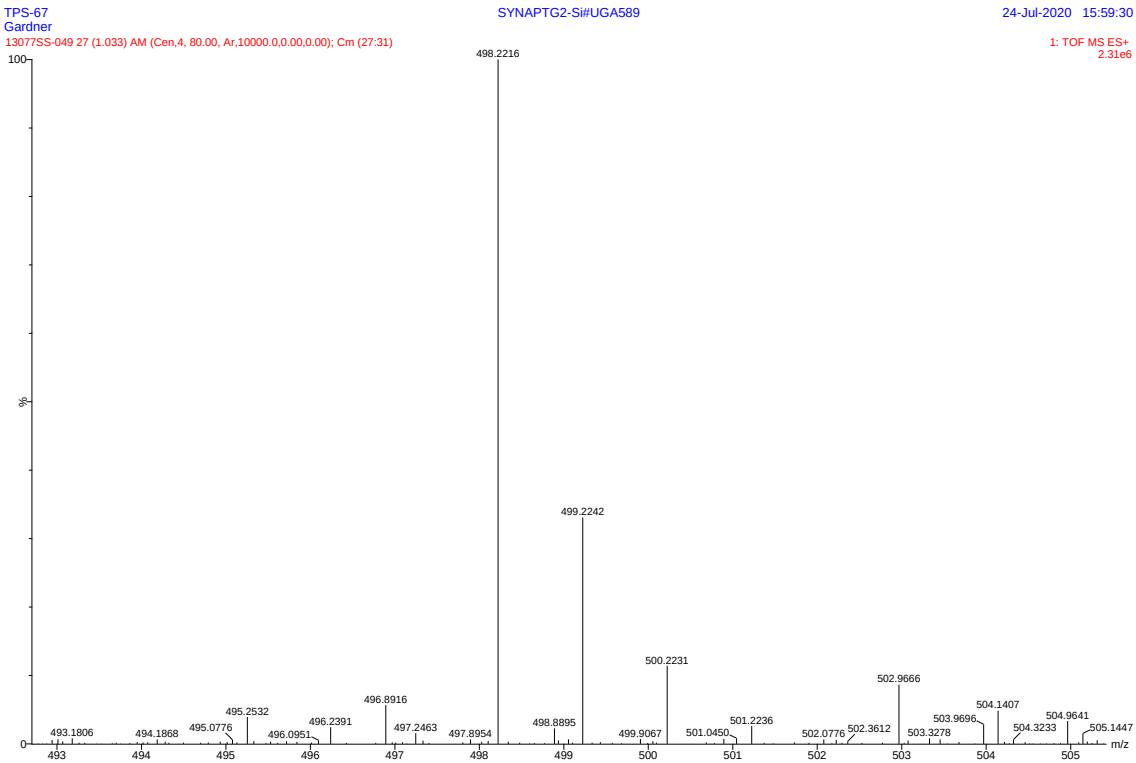
HRMS of **TPS-61-P2**



HRMS of **TPS-62**



HRMS of **TPS-65**



HRMS of **TPS-66**

2.9 References for Chapter 2

1. Evidente, A.; Kornienko, A.; Cimmino, A.; Andolfi, A.; Lefranc, F.; Mathieu, V.; Kiss, R., Fungal metabolites with anticancer activity. *Nat. Prod. Rep.* **2014,** *31* (5), 617-627.

2. Ma, Y. M.; Liang, X. A.; Kong, Y.; Jia, B., Structural Diversity and Biological Activities of Indole Diketopiperazine Alkaloids from Fungi. *J Agric Food Chem* **2016,** *64* (35), 6659-71.

3. Cui, C.-B.; Kakeya, H.; Okada, G.; Onose, R.; Osada, H., Novel Mammalian Cell Cycle Inhibitors, Tryprostatins A, B and Other Diketopiperazines Produced by Aspergillus fumigatus. I. Taxonomy, Fermentation, Isolation and Biological Properties. *The Journal of Antibiotics* **1996,** *49* (6), 527-533.

4. Bischoff, V.; Cookson, S. J.; Wu, S.; Scheible, W. R., Thaxtomin A affects CESA-complex density, expression of cell wall genes, cell wall composition, and causes ectopic lignification in Arabidopsis thaliana seedlings. **2009,** *60* (3), 955-965.

5. SjöGren, M.; Jonsson, P. R.; DahlströM, M.; LundäLv, T.; Burman, R.; GöRansson, U.; Bohlin, L., Two Brominated Cyclic Dipeptides Released by the Coldwater Marine Sponge Geodia barretti Act in Synergy As Chemical Defense. **2011,** *74* (3), 449-454.

6. Lavey, N. P.; Coker, J. A.; Ruben, E. A.; Duerfeldt, A. S., Sclerotiamide: The First Non-Peptide-Based Natural Product Activator of Bacterial Caseinolytic Protease P. *Journal of Natural Products* **2016,** *79* (4), 1193-1197.

7. Kuramochi, K.; Ohnishi, K.; Fujieda, S.; Nakajima, M.; Saitoh, Y.; Watanabe, N.; Takeuchi, T.; Nakazaki, A.; Sugawara, F.; Arai, T.; Kobayashi, S., Synthesis and

Biological Activities of Neoechinulin A Derivatives: New Aspects of Structure–Activity Relationships for Neoechinulin A. *CHEMICAL & PHARMACEUTICAL BULLETIN* **2008,** *56* (12), 1738-1743.

8. Ravikanth, V.; Niranjan Reddy, V. L.; Ramesh, P.; Prabhakar Rao, T.; Diwan, P. V.; Khar, A.; Venkateswarlu, Y., An immunosuppressive tryptophan-derived alkaloid from Lepidagathis cristata. *Phytochemistry* **2001,** *58* (8), 1263-1266.

9. Usui, T.; Kondoh, M.; Cui, C. B.; Mayumi, T.; Osada, H., Tryprostatin A, a specific and novel inhibitor of microtubule assembly. *The Biochemical journal* **1998,** *333 (Pt 3)* (Pt 3), 543-548.

10. Wollinsky, B.; Ludwig, L.; Hamacher, A.; Yu, X.; Kassack, M. U.; Li, S.-M., Prenylation at the indole ring leads to a significant increase of cytotoxicity of tryptophan-containing cyclic dipeptides. *Bioorganic & Medicinal Chemistry Letters* **2012,** *22* (12), 3866-3869.

11. Jain, H. D.; Zhang, C.; Zhou, S.; Zhou, H.; Ma, J.; Liu, X.; Liao, X.; Deveau, A. M.; Dieckhaus, C. M.; Johnson, M. A.; Smith, K. S.; Macdonald, T. L.; Kakeya, H.; Osada, H.; Cook, J. M., Synthesis and structure–activity relationship studies on tryprostatin A, an inhibitor of breast cancer resistance protein. *Bioorganic & Medicinal Chemistry* **2008,** *16* (8), 4626-4651.

12. Woehlecke, H.; Osada, H.; Herrmann, A.; Lage, H., Reversal of breast cancer resistance protein–mediated drug resistance by tryprostatin A. *International Journal of Cancer* **2003,** *107* (5), 721-728.

13. Robbins, D. W.; Boebel, T. A.; Hartwig, J. F., Iridium-Catalyzed, Silyl-Directed Borylation of Nitrogen-Containing Heterocycles. *Journal of the American Chemical Society* **2010,** *132* (12), 4068-4069.

14. Song, Z.; Antonchick, A. P., Iridium(iii)-catalyzed regioselective C7-sulfonamidation of indoles. **2016,** *14* (21), 4804-4808.

15. Leitch, J. A.; Bhonoah, Y.; Frost, C. G., Beyond C2 and C3: Transition-Metal-Catalyzed C–H Functionalization of Indole. *Acs Catal* **2017,** *7* (9), 5618-5627.

16. Fukuda, T.; Maeda, R.; Iwao, M., Directed C-7 lithiation of 1-(2,2-diethylbutanoyl)indoles. *Tetrahedron* **1999,** *55* (30), 9151-9162.

17. Kona, C. N.; Nishii, Y.; Miura, M., Iridium-Catalyzed Direct C4- and C7-Selective Alkynylation of Indoles Using Sulfur-Directing Groups. *Angewandte Chemie International Edition* **2019,** *58* (29), 9856-9860.

18. Yang, Y.; Qiu, X.; Zhao, Y.; Mu, Y.; Shi, Z., Palladium-Catalyzed C–H Arylation of Indoles at the C7 Position. *Journal of the American Chemical Society* **2016,** *138* (2), 495-498.

19. Yang, Y.; Gao, P.; Zhao, Y.; Shi, Z., Regiocontrolled Direct C−H Arylation of Indoles at the C4 and C5 Positions. *Angewandte Chemie International Edition* **2017,** *56* (14), 3966-3971.

20. Xu, L.; Zhang, C.; He, Y.; Tan, L.; Ma, D., Rhodium-Catalyzed Regioselective C7-Functionalization of N -Pivaloylindoles. *Angewandte Chemie International Edition* **2016,** *55* (1), 321-325.

21. Lu, X.; He, S.-J.; Cheng, W.-M.; Shi, J., Transition-metal-catalyzed CH functionalization for late-stage modification of peptides and proteins. *Chinese Chemical Letters* **2018,** *29* (7), 1001-1008.

22. Shah, T. A.; De, P. B.; Pradhan, S.; Punniyamurthy, T., Transition-metal-catalyzed site-selective C7-functionalization of indoles: advancement and future prospects. *Chemical Communications* **2019,** *55* (5), 572-587.

23. Demopoulos, V. J.; Nicolaou, I., Electrophilic Substitution of Indole on the Benzene Moiety: A Synthesis of 5-Acyl- and 5-Aroylindoles. *Synthesis* **1998,** *1998* (10), 1519-1522.

24. Feng, Y.; Holte, D.; Zoller, J.; Umemiya, S.; Simke, L. R.; Baran, P. S., Total Synthesis of Verruculogen and Fumitremorgin A Enabled by Ligand-Controlled C–H Borylation. *Journal of the American Chemical Society* **2015,** *137* (32), 10160-10163.

25. Bandari, C.; Scull, E. M.; Bavineni, T.; Nimmo, S. L.; Gardner, E. D.; Bensen, R. C.; Burgett, A. W.; Singh, S., FgaPT2, a biocatalytic tool for alkyl-diversification of indole natural products. *MedChemComm* **2019**.

26. Johnson, B. P.; Scull, E. M.; Dimas, D. A.; Bavineni, T.; Bandari, C.; Batchev, A. L.; Gardner, E. D.; Nimmo, S. L.; Singh, S., Acceptor substrate determines donor specificity of an aromatic prenyltransferase: expanding the biocatalytic potential of NphB. *Applied Microbiology and Biotechnology* **2020**.

27. Bandari, C.; Scull, E. M.; Masterson, J. M.; Tran, R. H. Q.; Foster, S. B.; Nicholas, K. M.; Singh, S., Determination of Alkyl-Donor Promiscuity of Tyrosine-O-Prenyltransferase SirD from Leptosphaeria maculans. *ChemBioChem* **2017,** *18* (23), 2323-2327.

28. Scull, E. M.; Bandari, C.; Johnson, B. P.; Gardner, E. D.; Tonelli, M.; You, J.; Cichewicz, R. H.; Singh, S., Chemoenzymatic synthesis of daptomycin analogs active against daptomycin-resistant strains. *Applied Microbiology and Biotechnology* **2020**.

29. Chen, R.; Gao, B.; Liu, X.; Ruan, F.; Zhang, Y.; Lou, J.; Feng, K.; Wunsch, C.; Li, S.-M.; Dai, J.; Sun, F., Molecular insights into the enzyme promiscuity of an aromatic prenyltransferase. *Nat Chem Biol* **2017,** *13* (2), 226-234.

30. Tarcz, S.; Xie, X.; Li, S.-M., Substrate and catalytic promiscuity of secondary metabolite enzymes: O-prenylation of hydroxyxanthones with different prenyl donors by a bisindolyl benzoquinone C- and N-prenyltransferase. *RSC Adv.* **2014,** *4* (35), 17986-17992.

31. Zou, H.; Zheng, X.; Li, S.-M., Substrate Promiscuity of the Cyclic Dipeptide Prenyltransferases from Aspergillus fumigatus. *Journal of Natural Products* **2009,** *72* (1), 44-52.

32. Elshahawi, S. I.; Cao, H.; Shaaban, K. A.; Ponomareva, L. V.; Subramanian, T.; Farman, M. L.; Spielmann, H. P.; Phillips Jr, G. N.; Thorson, J. S.; Singh, S., Structure and specificity of a permissive bacterial C-prenyltransferase. *Nat Chem Biol* **2017,** *13* (4), 366-368.

33. Kumar, K.; Wang, P.; Sanchez, R.; Swartz, E. A.; Stewart, A. F.; DeVita, R. J., Development of Kinase-Selective, Harmine-Based DYRK1A Inhibitors that Induce Pancreatic Human β-Cell Proliferation. *Journal of Medicinal Chemistry* **2018,** *61* (17), 7687-7699.

34. Tanaka, S.; Shiomi, S.; Ishikawa, H., Bioinspired Indole Prenylation Reactions in Water. *Journal of Natural Products* **2017,** *80* (8), 2371-2378.

35. Tanner, M. E., Mechanistic studies on the indole prenyltransferases. *Natural Product Reports* **2015,** *32* (1), 88-101.

36. Liebhold, M.; Li, S. M., Regiospecific benzylation of tryptophan and derivatives catalyzed by a fungal dimethylallyl transferase. *Organic letters* **2013,** *15* (22), 5834-7.

37. Liebhold, M.; Xie, X.; Li, S.-M., Breaking Cyclic Dipeptide Prenyltransferase Regioselectivity by Unnatural Alkyl Donors. *Organic letters* **2013,** *15* (12), 3062-3065.

38. Liebhold, M.; Xie, X.; Li, S. M., Expansion of enzymatic Friedel-Crafts alkylation on indoles: acceptance of unnatural beta-unsaturated allyl diphospates by dimethylallyl-tryptophan synthases. *Organic letters* **2012,** *14* (18), 4882-5.

Chapter 3. Synthesis and Evaluation of Azaindole Containing Tryptophan analogs as Novel Indole PT Substrates

Abstract

Azaindoles represent quintessential isosteres of indole in the context of pharmaceutical lead optimization. Introducing additional aromatic nitrogen atoms to an indole moiety increases hydrophilicity and frequently has a profound impact on target selectivity and potency due to new active site interactions, while maintaining a similar steric profile. Considering the high prevalence of azacycles in pharmaceuticals, and the vast number of indole-containing bioactive natural products, few methods exist to rapidly access and derivatize azaindole-substituted natural product analogs. In this work, promiscuous aromatic prenyltransferases (PTs) which are key late-stage tailoring enzymes for countless indole-containing NPs, will be explored as a highly innovative method of functionalizing azaindole-containing substrate-mimetics. A series of aza-tryptophans (AzaTrp) featuring N-substitution of every aromatic CH position of the indole ring were synthesized as substrate-mimetics for the Trp alkylating PTs FgaPT2, CdpNPT, and FtmPT1. We have demonstrated for the first time that these prenyltransferases are able to prenylate azaindole containing substrate analogs, with altered regioselectivity based on the position of N-substitution. Remarkably, FgaPT2 was found to produce cationic N-prenyl pyridinium products in the case of 4-azatryptophan (**4A-W**) and 5-azatryptophan (**5A-W**), which not only represents a new substrate class of indole PTs, but a previously unobserved prenylation mode. The discovery that nitrogenous indole bioisosteres can be accepted by aromatic prenyltransferases provides access to previously inaccessible bioactive natural product analogs.

Allocation of Contribution

I produced the results presented in this chapter with the following exceptions. Mr. Brady Dehnert performed the multigram synthesis and purification of **7A-W** intermediate **7e** from 7-azaindole. Mr. Sam Huang synthesized, purified, and characterized **5A-W** intermediates **5b**, **5d**, and **5e**. The contents and results described in this chapter have not yet been published.

3.1 Introduction to Azaindoles

3.1.1 Indole and isosteric replacement

The indole moiety is a feature of countless natural products due to its ubiquitous presence in all living cells as the heteroaromatic moiety of tryptophan. The biosynthetic availability of tryptophan has made indole an extremely common heterocyclic scaffold within bioactive secondary metabolites. As natural products continue to inspire the development of modern medicine, chemists develop strategies to modify these pharmacophores to generate novel derivatives with improved physical and biological properties. Within the context of medicinal chemistry, bioisosteres are chemical moieties with similar physical and chemical properties which can be substituted with one another to produce a broadly similar biological activity. Maintaining relatively small deviations in size, shape, and electronic profile of the substituent are the parameters most critical maintaining biological activity.

3.1.2 Azaindoles and their Significance

While a large number of indole isosteres are known, azaindoles have become a particularly attractive scaffold to many branches of chemical and biochemical research. Not only do azaindoles represent and excellent isosteric replacement for medicinal chemists[1-3], they possess intrinsically high fluorescence and can also serve as versatile organometallic ligands[4-8]. Azaindoles have remarkable luminescent properties which have been implemented for organic light emitting diodes[8-10], catalysis[6, 11], and various biochemical/protein interaction probing methods[12-17].

Compared to the other azaindole isomers, 7-azaindole (7-AI) has been the most extensively studied.[18] 7-azaindole is known to undergo photoinduced proton-mediated

excited state tautomerization and has a redshifted absorbance and emission profiles compared to indole. This altered spectral profiles enables selective excitation and detection of 7-AI moieties with minimal interference from naturally occurring amino acids within biological systems. The quantum yield of 7-AI is highly dependent on the polarity of its chemical environment, with a quantum yield of 0.01 in a pH 7 aqueous solution, compared to a quantum yield of 0.25 in acetonitrile.[13] This variability enables the use of 7-AI as a fluorescent probe for microenvironments, such as protein conformations and protein-protein interactions. The nonspecific incorporation 7-aza-Trp (**7A-W**) into proteins in the place of Trp has been achieved by treating an initial cell culture mass with **7A-W** (or 7-aza-indole) at the time of protein induction, which circumvents the toxicity of the unnatural amino acid.[17] Remarkably, this relatively simple strategy can produce "spectrally enhanced" proteins with >85% incorporation of **7A-W**. The highly variable fluorescence can be used to explore protein-ligand and protein-protein interactions, since displacement of a **7A-W** solvent shell with a hydrophobic interaction/environment greatly increases fluorescence.[12, 13, 19] The azatryptophans represent an optimal series of fluorescent tryptophan isosteres that impart novel chemical and photophysical properties without major conformational disruption.

3.1.3 The Importance of Azaindoles in Pharmaceutical Lead Optimization

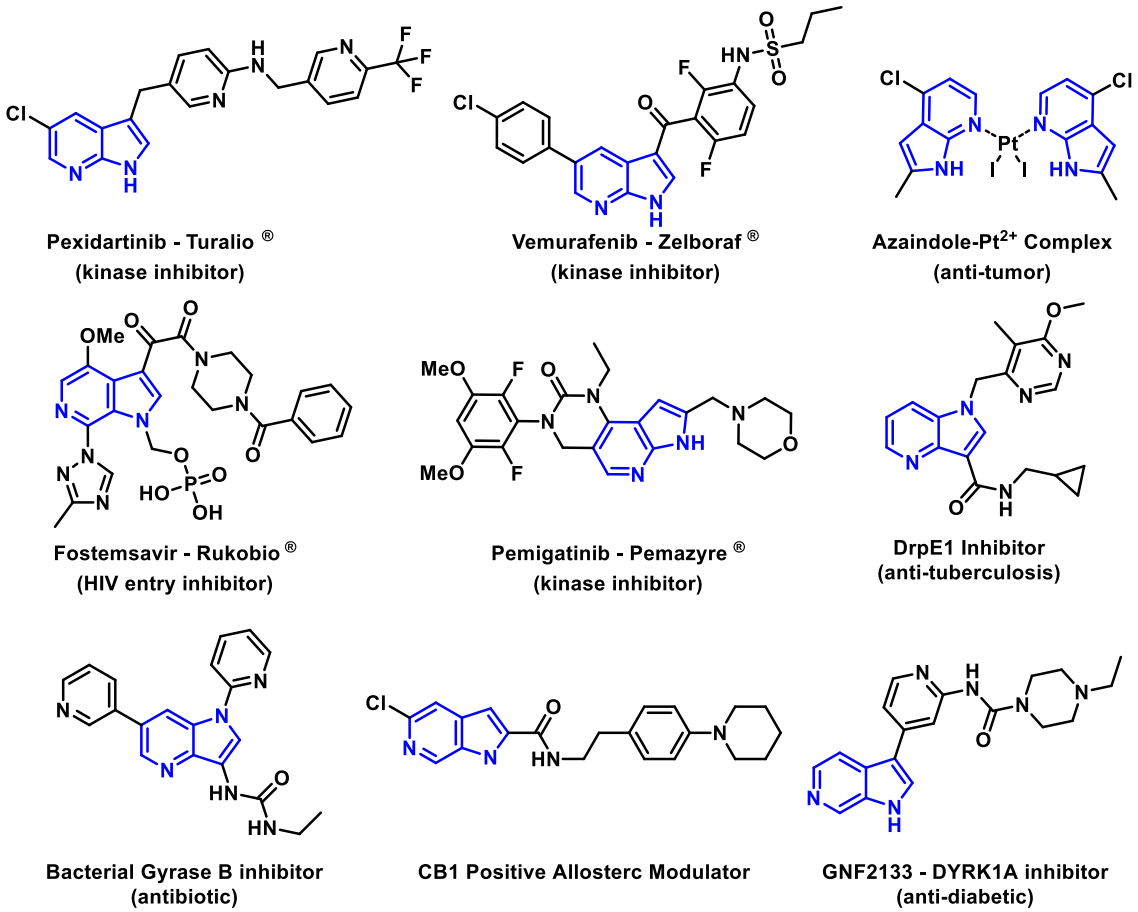


Figure 3.1: Representative azaindole containing compounds with diverse bioactivity

In the context of pharmaceutical lead optimization strategies, the isosteric replacement of aromatic CH moieties with N atoms enables compounds to engage in new interactions without substantially changing the size and shape of the scaffold. The similar steric demands indole and azaindole is advantageous, as any dramatic changes in binding affinity would likely arise from previously unavailable binding interactions compared to the indole parent compound. The incorporation of azacyclic isosteres such as azaindoles can enable new hydrogen bonding interactions with the drug target, which can dramatically enhance binding affinity and target selectivity. This lead optimization strategy has been increasingly used to improve a variety of important pharmacological characteristics such

as potency, selectivity, solubility, and metabolic stability.[3] There have been many developments reported in the area of isosteric replacement using azaindoles in medicinal chemistry resulting in useful activities including kinase inhibitors,[1, 20] anti-tumor agents,[4, 21, 22] anti-HIV,[23] anti-tuberculosis,[24, 25] antibiotics,[26] anti-diabetic,[27] and cannabinoid receptor modulators (Figure 3.1).[28] A representative study where 7-azaindole was incorporated into the cytotoxic natural product rebeccamycin revealed that 7-azarebeccamycin analogs display a higher degree of cytotoxic selectivity between cancer cell lines, and likely possess altered mechanisms of action.[29] However, isosteric replacement within the core scaffolds of NPs is rarely accessible through late stage synthetic modification, thus requiring labor intensive total synthesis of the natural product analogs.

3.1.4 Chemoenzymatic Transformations of Azaindole-Containing Substrates

Thus far, only a couple of enzymes have been demonstrated to utilize azaindoles as substrates. These include: Tryptophan synthase from *S. typhimurium* and other organisms has been shown to produce azatryptophans in the presence of 4-aza, 5-aza, 6-aza, and 7-azaindoles,[30-32] the viral halogenase VirX1, which can iodinate a variety of heterocyclic scaffolds including 6-azaindole, quinolines, isoquinolines, and benzopyrazines,[33] and the strictosidine synthase, that can accept azatryptamines to generate azaindole containing analogs of the tetrahydro-β-carboline isositsrikine.[34] Recent studies have demonstrated aromatic prenyltransferases (PTs) to be highly promiscuous towards both their alkyl donors and aromatic acceptors[35-44]. Given the biosynthetic pathways of many biologically relevant indole-containing NPs involve PTs as one of the enzymes, we wanted to evaluate the ability of PTs to accept azaindoles as their substrates.

3.1.5 Industrial Biocatalytic Applications and Advantages

Recently, biocatalytic approaches for drug synthesis and modification have gained substantial interest.[45, 46] In the pharmaceutical industry, thousands of tons of beta lactam antibiotics are produced using penicillin G amidase.[47] Nitrile hydratases are used in the chemical industry to produce the industrial monomer acrylamide from acrylonitrile on a several hundred thousand tons per year scale.[47] While the traditional organo-metallic synthetic methods have substantial limitations on their functional group tolerance, regioselectivity, and chemoselectivity, especially in the context of complex molecules, biocatalytic methods are attractive options for drug synthesis, that reduces the reliance on expensive and/or toxic metal catalysts or waste products, in addition, enzymes evolved to carry out enantio- and regioselective reactions on highly complex molecules.[45, 46, 48-50]

3.1.6 Aims and Significance

Thus, exploring the applicability of biocatalysts for the late-stage modifications of azaindole molecules for structure activity relationship (SAR) studies has great potential in developing novel drug candidates. Additionally, we hypothesize that substrates could be engineered to unlock new alkylation modes for PTs by the incorporation of unnatural nucleophiles into the acceptor substrate. By introducing a nucleophilic pyridine-like N atom near the native alkylation site of tryptophan-mimetic substrates, we hope to capture the dimethylallyl carbocation intermediate in novel manner. Pushing biosynthetic enzymes to perform unnatural reactions has improved several industrial scale manufacturing methods. To actualize the full potential of PT's as biocatalytic tools, their substrate scope and reactivity must be creatively explored in studies such as this.

3.2 Materials and Methods

3.2.1 General Materials

Unless otherwise stated, all chemicals and reagents were purchased from Acros (New Jersey, USA), Sigma-Aldrich (St. Louis, MO, USA), Ambeed (Arlington Heights, IL, USA), or AK Scientific (Union City, CA, USA) and were reagent grade or better. The Ni-NTA columns and PD-10 columns used for protein purification were purchased from GE Healthcare (Piscataway, NJ).

3.2.2 General Methods

All reported reactions were conducted in oven-dried glassware under an anhydrous nitrogen atmosphere with anhydrous solvents, unless otherwise noted. Column chromatography purification was performed using silica gel (SiliCycle Inc, P60, particle size 40-63 μm). HPLC analysis and purification was conducted using an Agilent 1220 system equipped with a diode array detector.

3.2.3 Structural Elucidation via NMR and HRMS

High-resolution mass spectrometry (HRMS) confirmation of products and substrates was conducted on an Agilent 6545-QTOF W/ 1290 HPLC mass spectrometer at the University of Oklahoma Department of Chemistry and Biochemistry. NMR Spectra were collected on a Varian VNMRS 500 MHz, a Varian INOVA 600 MHz, a Varian Mercury VX-300, or a Varian VNMRS 400 MHz NMR spectrometer at the NMR facility of the Department of Chemistry and Biochemistry of the University of Oklahoma. All NMR spectra were obtained in 400-600 μL volumes of D_3COD (99.8%), DMSO-d_6 (99.9%), $CDCl_3$ (99.8%), or D_2O (99.9%) (Cambridge Isotope Laboratories, MA, USA), and processed using MestReNova software. Structural elucidation and NMR peak

assignment of chemoenzymatically prenylated products was performed using ^{1}H-COSY, ^{1}H-^{13}C HSQC, and ^{1}H-^{13}C HMBC experiments.

3.2.4 Synthesis of 2-Azatryptophan (2A-W)

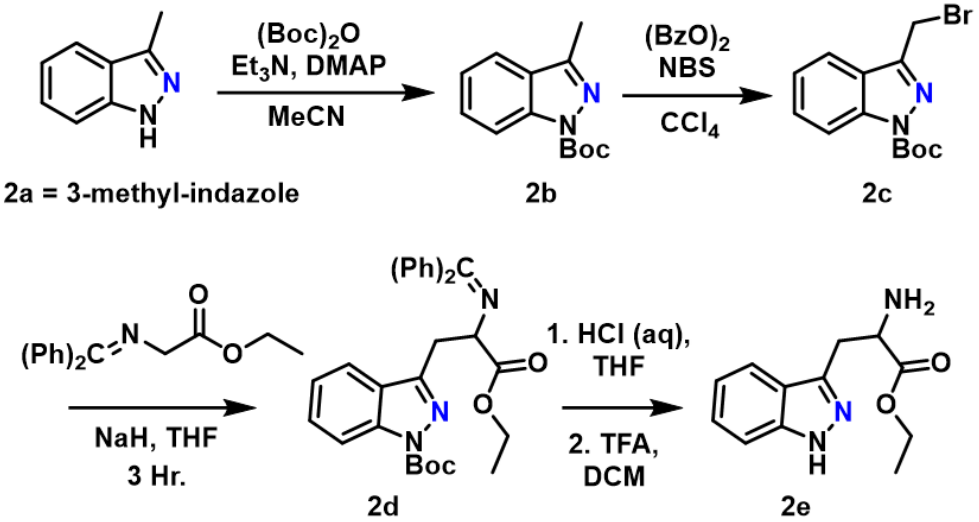


Figure 3.2: Synthetic scheme for the synthesis of 2-Aza-Tryptophan

3-methyl-indazole Boc protection (2b): A cooled solution of 3-methyl-indazole (600 mg, 4.54 mmol) in 20 mL anhydrous acetonitrile was prepared and treated with DMAP (1.11 g, 9.08 mmol), triethylamine (1.27 mL, 9.08 mmol), and $(Boc)_2O$ (1.98 g, 9.08 mmol). The reaction mixture was stirred for 1 hour at 0°C before allowing to warm to room temperature. After 2 hours, TLC indicated consumption of starting material, and the solvent was removed in vacuo. The crude residue was dissolved in EtOAc and washed with H_2O. The aqueous layer was extracted twice with ethyl acetate, and the combined organic layers were dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo. The crude product was purified by flash column chromatography on silica gel using EtOAc:hexane. Pure product was recovered as a pale orange oil (1.03 g, 97%). This literature reported procedure was used and the product NMR spectra matched the reported NMR data.[51]

Synthesis of 3-(bromomethyl)-1-(*tert*-butoxycarbonyl)indazole (2c): A solution was prepared containing Boc protected 3-methyl-indazole (716 mg, 3.08 mmol) and N-bromosuccinamide (658 mg, 3.70 mmol) in 16 mL anhydrous CCl_4. The reaction mixture was then treated with benzoyl peroxide (74 mg, 0.31 mmol) and heated at reflux for 4 hours. The mixture was cooled to room temperature, filtered through celite, and concentrated in vacuo. The resulting residue was purified via silica gel chromatography using a gradient from 100% hexane to 10% EtOAc in hexane. Unreacted starting material was successfully recovered during purification. This literature reported procedure was used and the product NMR spectra matched the reported NMR data.[51]

Synthesis of tert-butyl 3-(2-((diphenylmethylene)amino)-3-ethoxy-3-oxopropyl)-1H-indazole-1-carboxylate (2d): A solution containing NaH (23.1 mg, 0.964 mmol, 60% suspension in mineral oil) in 6 mL anhydrous THF under dry nitrogen was treated with ethyl N-(diphenylmethylene)glycinate (257.7 mg, 0.964 mmol) and stirred at 0°C for 30 minutes. A solution of **2c** in 2 mL THF was then added, and the resulting mixture was stirred for 3 hours at 0°C. The reaction was quenched with saturated aqueous NH_4Cl and extracted three times with ethyl acetate. The combined organic layers were dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo. The resulting residue was purified using flash column chromatography on silica gel eluting with a hexane/EtOAc gradient. After removal of the solvent in vacuo, 200 mg (62.5%) of pure **2d** was obtained. ^{1}H NMR (600 MHz, Chloroform-*d*) δ 8.05 (d, *J* = 8.4 Hz, 1H), 7.61 (d, *J* = 8.0 Hz, 1H), 7.53 – 7.49 (m, 2H), 7.45 – 7.41 (m, 2H), 7.33 (dt, *J* = 13.8, 7.4 Hz, 2H), 7.28 – 7.23 (m, 2H), 7.14 (t, *J* = 7.5 Hz, 1H), 6.73 (d, *J* = 7.4 Hz, 2H), 4.63 (dd, *J* = 8.9, 4.6 Hz, 1H), 4.26 – 4.12 (m,

2H), 3.65 (dd, *J* = 14.3, 4.6 Hz, 1H), 3.61 (dd, *J* = 14.3, 8.9 Hz, 1H), 1.68 (s, 9H), 1.25 (t, *J* = 7.1 Hz, 3H).

Synthesis of ethyl 2-amino-3-(1H-indazol-3-yl)propanoate (2e): The protected 2-aza-tryptophan ester **2d** (200 mg, 0.4 mmol) was dissolved in THF to make a 0.1 M solution, which was then treated with 5 M HCl (10 eq), and stirred at room temperature overnight. The reaction mixture was diluted with ethyl acetate, neutralized with saturated $NaHCO_3$, and extracted 3 times with ethyl acetate to yield 63 mg (47%) of the Boc-protected **2e** after purification via silica column chromatography. Quantitative deprotection of the N-Boc functionality was achieved by dissolving 20 mg the purified material in 1 mL anhydrous DCM and subsequent dropwise addition of 2 mL triflouroacetic acid at 0°C. The reaction was then allowed to warm to room temperature and stirred overnight. The solvent was removed via nitrogen flow, and the resulting residue was dissolved in aqueous K_2CO_3 and extracted three times with ethyl acetate. Combined organic layers were dried over Na_2SO_4, filtered, and concentrated in vacuo to yield pure 2-aza-tryptophan ethyl ester **(2e)**. ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.66 (d, *J* = 8.1 Hz, 1H), 7.41 (d, *J* = 8.4 Hz, 1H), 7.37 – 7.29 (m, 1H), 7.16 – 7.09 (m, 1H), 4.12 (qd, *J* = 7.1, 5.1 Hz, 2H), 4.05 (s, 1H), 3.47 (dd, *J* = 14.9, 4.6 Hz, 1H), 3.34 (dd, *J* = 14.9, 7.6 Hz, 1H), 1.17 (t, *J* = 7.1 Hz, 3H).

3.2.5 Synthesis Azatryptophans via Aza-Gramine Intermediates

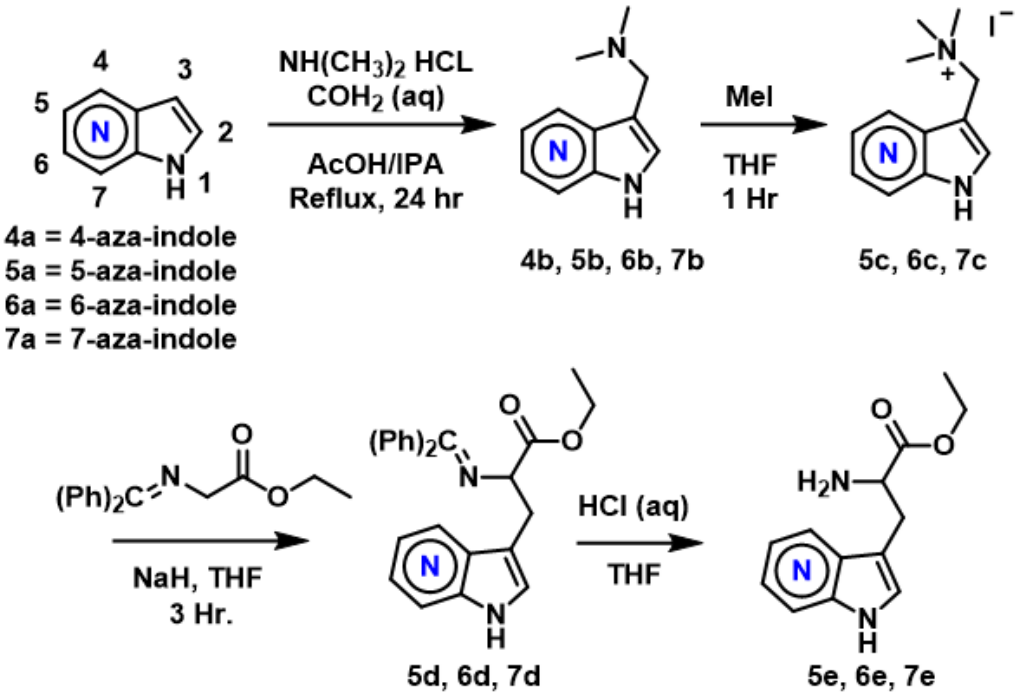


Figure 3.3: Synthetic scheme for 5A-W, 6A-W, and 7A-W employing Aza-gramine intermediates (4b, 5b, 6b, 7b)

General Procedure for aza-gramine synthesis (4b, 5b, 6b, 7b): A solution of dimethylamine (40 wt % in H_2O, 0.7 mL, 5.5 mmol), acetic acid (0.28 mL, 4.9 mmol), and isopropanol (1 mL) was prepared and cooled to 0°C in an ice bath. The solution was then treated with 37% aqueous formaldehyde (0.363 mL, 4.88 mmol) and stirred at 0°C for 30 min. A solution containing azaindole (**4a, 5a, 6a, or 7a**) (500mg, 4.24 mmol) in isopropanol (2 mL) was then added to the reaction mixture, with the resulting solution stirred at 0°C for 30 min. The solution was then heated to 90°C overnight. The reaction mixture was allowed to warm to room temperature, diluted with 10 mL water, and basified with 5 M NaOH. The aqueous solution was extracted with $CHCl_3$:IPA (3:1) (3 x 10 mL). The combined organic layers were dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo to yield a yellow oil which slowly crystalized upon standing at room temperature.

If contaminates were detected via NMR, purification via column chromatography was carried out using $CHCl_3$:MeOH:NH_4OH (90:10:1 followed by 40:10:1). Isolated yields of aza-gramines **(4b, 5b, 6b, 7b)** were >80%. Later experiments found that repeated extraction of the crude reaction mixture using large volumes (100 mL) of diethyl ether (in place of the $CHCl_3$:IPA mixture previously mentioned) yielded pure crystalline products which could be taken forward without purification. The relatively large volumes of diethyl ether required for effective extraction can become disadvantageous at multigram scales.

N,N-dimethyl-1-(1H-pyrrolo[3,2-b]pyridin-3-yl)methanamine (4b): Isolated as a white crystalline solid using the procedure described above. ^{1}H NMR (300 MHz, Methanol-d_4) δ 8.29 (dd, J = 4.8, 1.5 Hz, 1H), 7.79 (dd, J = 8.2, 1.5 Hz, 1H), 7.54 (s, 1H), 7.14 (dd, J = 8.2, 4.8 Hz, 1H), 3.79 (s, 2H), 2.27 (s, 6H).

N,N-dimethyl-1-(1H-pyrrolo[3,2-c]pyridin-3-yl)methanamine (5b): Isolated as a viscous clear oil using the procedure described above. ^{1}H NMR (400 MHz, Methanol-d_4) δ 8.91 (s, 0H), 8.16 (d, J = 5.8 Hz, 1H), 7.42 (d, J = 5.8 Hz, 1H), 7.37 (s, 1H), 3.76 (s, 2H), 2.32 (d, J = 1.1 Hz, 6H).

N,N-dimethyl-1-(1H-pyrrolo[2,3-c]pyridin-3-yl)methanamine (6b): Isolated as a white crystalline solid using the procedure described above. ^{1}H NMR (300 MHz, Methanol-d_4) δ 10.07 – 9.96 (m, 1H), 9.40 (d, J = 5.5 Hz, 1H), 8.92 – 8.83 (m, 1H), 8.73 (s, 1H), 3.47 (s, 6H).

N,N-dimethyl-1-(1H-pyrrolo[2,3-b]pyridin-3-yl)methanamine (7b): Isolated as a white crystalline solid using the procedure described above. ^{1}H NMR (500 MHz, Methanol-d_4) δ 8.19 (dd, J = 4.8, 1.5 Hz, 1H), 8.10 (dd, J = 7.9, 1.5 Hz, 1H), 7.37 (s, 1H), 7.12 (dd, J = 7.9, 4.8 Hz, 1H), 3.68 (s, 2H), 2.28 (s, 6H).

Procedure for aza-gramine methylation (5c, 6c, 7c): A 0.15 M solution of 5-aza-gramine (420mg, 2.4 mmol) was prepared in anhydrous THF under a nitrogen atmosphere and cooled to 0°C. The reaction mixture was treated with methyl iodide (1.5 eq, 510 mg) and stirred for 5 minutes at 0°C before warming to room temperature. After stirring at room temperature for 1 hour, the mixture was diluted to a triple volume using pentane. The resulting suspension was filtered, washed with pentane, and taken forward without purification. Isolated yield of 5-aza-gramine methiodide was 549 mg (72.2%). Using this protocol, isolated yield for 6-aza-gramine methiodide using was 57.2 %, and the yield for 7-aza-gramine methiodide was nearly quantitative. This procedure was based on previously published methods.[52]

Synthesis of ethyl 2-((diphenylmethylene)amino)-3-(1H-pyrrolo[3,2-c]pyridin-3-yl)propanoate (5d): Ethyl N-(diphenylmethylene)glycinate (168 mg, 0.63 mmol) was dissolved in 1.5 mL dry DMF and cooled to 0°C. NaH (15 mg, 0.63 mmol, 60% suspension in mineral oil) was then added and stirred for 15 min. Half of the 5-aza-gramine methiodide **(5c)** was then added as a suspension in 1.5 mL DMF. After stirring at 0°C for an hour, the remaining **5c** was added as a powder in one portion. The reaction mixture was stirred at 0°C for 6 hours and then left at 4°C overnight without stirring in a refrigerator. The reaction was quenched with saturated aqueous sodium bicarbonate and extracted 3 times with EtOAc. Organic layers were dried, filtered and removed in vacuo. Purification via silica gel column was done with an initial gradient from 25% EtOAc in hexanes, to 100% EtOAc, followed by $CHCl_3$:MeOH:NH_4OH (90:10:1) to elute the product. Azaindole containing products fluoresce brightly under a 365 nm TLC visualization lamp, allowing for rapid identification of azaindole containing fractions. The fluorescent product was collected to

yield 75 mg (29.9%) **5d** as a brown viscous liquid. ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.53 (s, 1H), 8.23 (d, *J* = 5.8 Hz, 1H), 7.61 – 7.56 (m, 3H), 7.33 (dd, *J* = 21.7, 7.3 Hz, 4H), 7.22 (q, *J* = 7.5, 6.4 Hz, 3H), 7.04 (s, 1H), 6.68 (d, *J* = 7.2 Hz, 2H), 4.40 (dd, *J* = 8.3, 5.1 Hz, 1H), 4.24 – 4.10 (m, 2H), 3.47 (dd, *J* = 14.3, 5.1 Hz, 1H), 3.28 (dd, *J* = 14.3, 8.3 Hz, 1H), 1.22 (t, *J* = 7.1 Hz, 3H).

Synthesis of ethyl 2-((diphenylmethylene)amino)-3-(1H-pyrrolo[2,3-c]pyridin-3-yl)propanoate (6d): The benzophenone protected **6A-W** intermediate **6d** was prepared using the above synthetic method for **5d.** The product **6d** was isolated as a viscous brown-red oil with yields ranging from 16-25%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.81 (s, 1H), 8.06 (d, *J* = 5.6 Hz, 1H), 7.60 (dd, *J* = 8.3, 1.3 Hz, 2H), 7.42 – 7.37 (m, 1H), 7.35 – 7.29 (m, 3H), 7.27 – 7.25 (m, 2H), 7.22 (t, *J* = 7.8 Hz, 2H), 6.67 (d, *J* = 7.4 Hz, 2H), 4.40 (dd, *J* = 8.4, 4.8 Hz, 1H), 4.20 (p, *J* = 7.1 Hz, 2H), 3.45 (dd, *J* = 14.4, 4.7 Hz, 1H), 3.31 (dd, *J* = 14.4, 8.5 Hz, 1H), 1.26 (t, *J* = 7.1 Hz, 3H).

Synthesis of ethyl 2-((diphenylmethylene)amino)-3-(1H-pyrrolo[2,3-b]pyridin-3-yl)propanoate (7d): A solution containing NaH (490 mg, 20.43 mmol, 60% suspension in mineral oil) in 100 mL anhydrous THF under dry nitrogen was treated with ethyl N-(diphenylmethylene)glycinate (5.46 g, 20.43 mmol) and stirred at 0°C for 30 minutes. A suspension of 7-aza-gramine methiodide (5.4 g, 17.03 mmol) in 50 mL anhydrous THF was then added and stirred for 6 hours at 0°C, and left overnight at 4°C. The reaction was quenched with a saturated solution of sodium bicarbonate and extracted 3 times with ethyl acetate. Combined organic layers were dried over anhydrous Na_2SO_4, filtered, and concentrated in vacuo. Purification was achieved via silica gel chromatography with a gradient from 25% EtOAc in hexane to 100% ethyl acetate to yield

3.66 g (54%) clear brown viscous liquid. ^{1}H NMR spectrum of the obtained product matches literature reported spectra.[52]

Synthesis of ethyl 2-amino-3-(1H-pyrrolo[2,3-c]pyridin-3-yl)propanoate (6e): The protected 6-aza-tryptophan ester **6d** (36 mg, 0.09 mmol) was dissolved in THF to make a 0.1 M solution, which was then treated with 2 M HCl (3.5 eq), and stirred at room temperature. The mixture was quenched with saturated sodium bicarbonate and extracted twice with diethyl ether to remove the benzophenone. The aqueous layer was further basified with excess K_2CO_3, extracted 3 times with EtOAc, and 3 times with $CHCl_3$/Isopropanol (3:1). The combined organic layers were dried over Na_2SO_4, filtered, and concentrated. The crude product was purified via silica gel column chromatography starting with EtOAc, followed by a gradient of $CHCl_3$:MeOH:NH_4OH (90:10:1 → 40:10:1) to elute the product. Fluorescent fractions with high relative polarity were combined to provide 17 mg pure 6-aza-tryprophan ethyl ester **6e** (80% yield) as a yellow oil. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.78 (s, 1H), 8.23 (d, *J* = 5.6 Hz, 1H), 7.55 (d, *J* = 5.5 Hz, 1H), 7.30 (s, 1H), 4.14 (qd, *J* = 7.1, 4.8 Hz, 2H), 3.80 (dd, *J* = 7.3, 5.2 Hz, 1H), 3.24 (dd, *J* = 14.5, 5.2 Hz, 1H), 3.09 (dd, *J* = 14.5, 7.3 Hz, 1H), 1.22 (t, *J* = 7.1 Hz, 3H).

Synthesis of ethyl 2-amino-3-(1H-pyrrolo[3,2-c]pyridin-3-yl)propanoate (5e): The title compound was obtained using the conditions listed above (Synthesis of **6e**) with comparable yields. The crude reaction mixture containing **5d** was also subject to these hydrolytic conditions, enabling simplified purification of the highly polar **5e.** ^{1}H NMR (600 MHz, Methanol-d_4) δ 8.76 (s, 1H), 8.10 (d, *J* = 5.9 Hz, 1H), 7.36 (d, *J* = 5.8 Hz, 1H), 7.22 (s, 1H), 4.04 (t, *J* = 7.1 Hz, 2H), 3.76 (t, *J* = 6.3 Hz, 1H), 3.18 (d, *J* = 6.3 Hz, 2H), 1.10 (t, *J* = 7.2 Hz, 3H).

Synthesis of ethyl 2-amino-3-(1H-pyrrolo[2,3-b]pyridin-3-yl)propanoate (7e):

The benzophenone-protected 7-aza-tryptophan ester (**7d**) (3.66 g, 9.21 mmol) was hydrolyzed using a previously reported procedure.[52] The desired product was obtained as a transparent brown oil (1.54 g, 71%). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 9.66 (s, 1H), 8.31 (d, *J* = 4.9 Hz, 1H), 7.95 (d, *J* = 7.8 Hz, 1H), 7.21 (s, 1H), 7.09 (dd, *J* = 7.9, 4.7 Hz, 1H), 4.14 (qd, *J* = 7.2, 2.2 Hz, 2H), 3.79 (dd, *J* = 7.3, 5.2 Hz, 1H), 3.23 (dd, *J* = 14.4, 5.2 Hz, 1H), 3.07 (dd, *J* = 14.4, 7.3 Hz, 1H), 1.22 (t, *J* = 7.2 Hz, 3H).

3.2.6 Synthesis of 4-Azatryptophan (4A-W)

Synthesis of ethyl 2-amino-3-(1H-pyrrolo[3,2-b]pyridin-3-yl)propanoate (4e):

To a flask containing 7 mL xylenes was added 4-aza-gramine **4b** (700 mg, 4 mmol), diethyl acetamidomalonate (868.9 mg, 4 mmol), and NaOH (50 mg, 1.25 mmol). The solution was then refluxed under nitrogen for 5 hours. The hot mixture was filtered and cooled. The precipitate that formed upon cooling was filtered, washed with hexane, and dried in vacuo to give a pure intermediate product. The residue was dissolved in 15.6 mL EtOH and treated with 1.56 mL water containing KOH (281mg, 5 mmol). After stirring at room temperature for 3 hours, the reaction was acidified with HCl and concentrated in vacuo. Water (7 mL) was added to the reaction mixture, the pH was adjusted to 1 with conc. HCl, and the solvent was removed using a stream of nitrogen. The solid residue was heated at 130°C under nitrogen for 1 hour to decarboxylate the intermediate. The residue was then dissolved in 20% aqueous HCl and heated at reflux for 22 hours. After cooling, the reaction mixture was basified with K_2CO_3 and washed with $CHCl_3$:IPA (3:1). The mother liquor was then saturated with K_2CO_3 and extracted with isopropanol three times. The isopropanol was removed in vacuo to yield a tan/white solid 4-aza-tryptophan.

Additional product was obtained by acidification of the mother liquor with conc. HCl, and subsequent removal of the solvent with low heat and nitrogen flow to produce a salt cake. The salt cake was extracted 3 times with 20 mL portions of boiling ethanol. The ethanol was evaporated to give a residue that was recrystallized by dissolution in 5 mL H_2O, basification with 4 drops 30% aqueous ammonia, and then diluted with 5 mL acetone. The mixture was left overnight at 4°C and the solid precipitate was recovered as pure 4-aza-tryptophan. The basified 4-aza-tryptophan has low solubility in water and DMSO, and moderate solubility in methanol, ethanol, and isopropanol. The free **4A-W** was converted into the ethyl ester **4e** by overnight reflux in a solution of thionyl chloride in ethanol. This esterification gave quantitative yield of 4-aza-trp ethyl ester **4e**.

3.2.7 Saponification of Aza-Trp Ethyl Esters

Synthesis of Aza-Trp from Aza-Trp Ethyl Esters: Saponification of the Aza-Trp esters was achieved by first dissolving 20 mg of the Aza-Trp ethyl ester (**2e, 5e, 6e, 7e**) in 4 mL ethanol. To the resulting solution was added 2 mL 1 M NaOH, followed by overnight stirring at 40°C in a sealed vial. The solution was then concentrated to ~2 mL via nitrogen flow, and 2 mL saturated $NaHCO_3$ was then added. Due to the extreme polarity of the products, traditional extraction solvents were ineffective. The aqueous solution was saturated with K_2CO_3 and extracted 3x with Isopropanol. The isopropanol layers were combined, dried over anhydrous K_2CO_3, filtered, and concentrated. The resulting residues were purified by semi-preparative RP-HPLC using Gemini-NX, C-18 (5 μm, 10 × 250 mm) column (Phenomenex, Torrance, California, USA) on an Agilent 1220 system equipped with a DAD detector using [gradient of 1% B for 3 min,1% B to 25% B for 15 min, 25% B to 100% B for 5 min, 100% B for 4 min, 100% B to 1% B in 0.5 min, 1% B for 6 min

(A = ddH_2O with 0.1% TFA or formic Acid; B = acetonitrile); flow rate = 2.5 ml min^{-1}; A_{230}]. 1D and 2D NMR characterization of the resulting products was performed in D_2O with a 5 mM product concentration.

2-amino-3-(1H-indazol-3-yl)propanoic acid (2A-W): ^{1}H NMR (500 MHz, Deuterium Oxide) δ 7.87 (d, *J* = 8.2 Hz, 1H), 7.61 (d, *J* = 8.6 Hz, 1H), 7.50 (t, *J* = 7.4 Hz, 1H), 7.25 (t, *J* = 7.5 Hz, 1H), 3.90 (dd, *J* = 7.5, 5.2 Hz, 1H), 3.49 (dd, *J* = 14.9, 5.0 Hz, 1H), 3.36 (dd, *J* = 14.9, 8.0 Hz, 1H). HRMS-ESI: Calc for $C_{10}H_{12}N_3O_2$ $[M+H]^+$: 206.09295; Found: 206.0920.

2-amino-3-(1H-pyrrolo[3,2-b]pyridin-3-yl)propanoic acid (4A-W): ^{1}H NMR (500 MHz, Deuterium Oxide) δ 8.47 (dd, *J* = 8.3, 1.1 Hz, 1H), 8.41 (dd, *J* = 5.9, 1.1 Hz, 1H), 7.95 (s, 1H), 7.59 (dd, *J* = 8.3, 5.9 Hz, 1H), 4.20 (dd, *J* = 7.2, 5.8 Hz, 1H), 3.54 – 3.40 (m, 2H). HRMS-ESI: Calc for $C_{10}H_{12}N_3O_2$ $[M+H]^+$: 206.09295; Found: 206.0927.

2-amino-3-(1H-pyrrolo[3,2-c]pyridin-3-yl)propanoic acid (5A-W): ^{1}H NMR (500 MHz, Deuterium Oxide) δ 9.05 (s, 1H), 8.46 (s, 1H), 8.28 (d, *J* = 6.7 Hz, 1H), 7.88 (d, *J* = 6.7 Hz, 1H), 7.70 (s, 1H), 4.08 (t, *J* = 5.9 Hz, 1H), 3.57 – 3.46 (m, 2H). HRMS-ESI: Calc for $C_{10}H_{12}N_3O_2$ $[M+H]^+$: 206.09295; Found: 206.0927.

2-amino-3-(1H-pyrrolo[2,3-c]pyridin-3-yl)propanoic acid (6A-W): ^{1}H NMR (500 MHz, Deuterium Oxide) δ 8.95 (s, 1H), 8.46 (s, 1H), 8.18 (d, *J* = 6.5 Hz, 1H), 8.09 (d, *J* = 6.5 Hz, 1H), 8.03 (s, 1H), 4.08 (t, *J* = 6.1 Hz, 1H), 3.50 (dd, *J* = 16.8, 6.1 Hz, 2H). HRMS-ESI: Calc for $C_{10}H_{12}N_3O_2$ $[M+H]^+$: 206.09295; Found: 206.0927.

2-amino-3-(1H-pyrrolo[2,3-b]pyridin-3-yl)propanoic acid (7A-W): ^{1}H NMR (500 MHz, Deuterium Oxide) δ 8.73 (dd, *J* = 8.0, 1.2 Hz, 1H), 8.40 (dd, *J* = 6.0, 1.2 Hz, 1H),

7.66 (s, 1H), 7.60 (dd, *J* = 8.0, 6.0 Hz, 1H), 4.23 (t, *J* = 6.1 Hz, 1H), 3.52 (dd, *J* = 14.3, 6.2 Hz, 2H). HRMS-ESI: Calc for $C_{10}H_{12}N_3O_2$ $[M+H]^+$: 206.09295; Found: 206.0927

3.2.8 Overexpression and Purification of FgaPT2

The recombinant FgaPT2 was overproduced in *Escherichia coli* BL21(DE3) cells transformed with the codon-optimized synthetic *FgaPT2* gene in a pET28a vector. The resulting FgaPT2 with N-terminal His_6-fusion protein was purified to homogeneity *via* nickel-nitrilotriacetic acid (affinity chromatography as previously reported.[38, 40] For protein purification, 4 x 1 L cultures of LB medium supplemented with kanamycin (50 µg/mL) were inoculated with 0.1% (v/v) of an overnight *E. Coli* BL21(DE3) seed culture and subsequently incubated at 37°C with shaking (220 rpm). Cultures were induced with isopropyl-β-D-thiogalactopyranoside (IPTG, 0.5 mM final concentration) once the optical density at 600 nm (OD_{600}) reached ~0.6-0.8. Following IPTG addition, cultures were allowed to grow for an additional 16 hours at 22°C. Cells were harvested via centrifugation and resuspended in ~20 mL lysis buffer (10 mM imidazole, 50 mM sodium monobasic phosphate, 300 mM NaCl, pH 8.0) per 1 L culture and stored at -80°C until used. During the 24 hours preceding cell lysis and purification, the cell pellet was allowed to thaw and were promptly refrozen at -80°C two times. Cells were allowed to thaw a third time and were subsequently lysed by sonication (SONICATOR 100 W 10 x 10 s pulses, 20 seconds between pulses). Insoluble particulate was removed by centrifugation at 15,000 x *g* for 1 hour. The supernatant was collected, passed through a 0.22 µm syringe filter, and the desired N-His_6-FgaPT2 fusion protein was purified via HiTrap Ni-NTA

3.2.9 Analytical scale FgaPT2 assay

Analytical scale FgaPT2 reactions were conducted at a 50 µL scale containing 21 µM FgaPT2, 1 mM Aza-Trp substrate, 1.5 mM DMAPP, 50 mM Tris (pH 7.8), and 5 mM $MgCl_2$. The reaction mixtures were incubated at 35°C for 18 hours and then quenched with the addition of 100 µL cold methanol. The precipitated protein was removed via centrifugation (9000 x g for 30 min). Analytical scale reactions were analyzed on an Agilent 1220 HPLC system equipped with a DAD, using a Gemini-NX C-18 (5 µm, 4.6 mm x 250 mm) column (Phenomenex, Torrance, CA, USA) [1% B for 3 min, gradient of 1% B to 5% B over 7 min, gradient of 5% B to 25% B over 10 min, gradient of 25% B to 100% B over 10 min, 100% B for 4 min, 100% B to 1% B over 1 min, 1% B for 5 min (A = ddH_2O with 0.1% TFA; B = acetonitrile); flow rate = 1 mL min^{-1}; A_{230} + A_{254} + A_{280} + A_{316} + A_{340}]. For each reaction, percent turnover was based upon the integration of each compound peak and calculated by dividing the integrated area of the remaining starting material by the integrated area of the negative control. Calibration curves were used to verify the integrity of this calculation, accounting for the unique absorbance profiles of different azaindole isomers, and absorbance profile variability between their corresponding prenylated regioisomers.

3.2.10 Enzymatic Scale Up Reactions for AzaTrp Prenylation

Scale up reactions of AzaTrp prenylation were conducted in a volume of 20 mL consisting of 1 mM **4A-W** or **5A-W**, 1.25 mM DMAPP, 50 mM Tris (pH 7.8) and 5 mM $MgCl_2$. Reactions were initiated by the addition of 4 mg purified FgaPT2 (~500 µL) and incubated at 35°C for 16-48 h. Reaction progress was monitored on analytical HPLC as described above by taking 50 uL aliquots at regular time intervals. An additional 4 mg

enzyme was added after 24 hours for reactions containing **4A-W** due to poor conversion (<25%), and incubated for an additional 24 hours at 35°C. Subsequently, the reaction mixture was concentrated via lyophilization to <5mL, saturated with K_2CO_3, then extracted 3 times with 10 mL portions of isopropanol. The combined organic layers were dried over K_2CO_3, filtered via syringe filter, and evaporated via nitrogen flow. The residue was redissolved in 1 mL methanol and purified by semi-preparative RP-HPLC using Gemini-NX, C-18 (5 μm, 10 × 250 mm) column (Phenomenex, Torrance, California, USA) on an Agilent 1220 system equipped with a DAD detector using the following method. [gradient of 1% B for 3 min,1% B to 25% B for 15 min, 25% B to 100% B for 5 min, 100% B for 4 min, 100% B to 1% B in 0.5 min, 1% B for 6 min (A = ddH_2O with 0.1% TFA or formic Acid; B = acetonitrile); flow rate = 2.5 ml min^{-1}; A_{230}]

3.3 Results and Discussion

3.3.1 Synthesis of 2-Azatryptophan

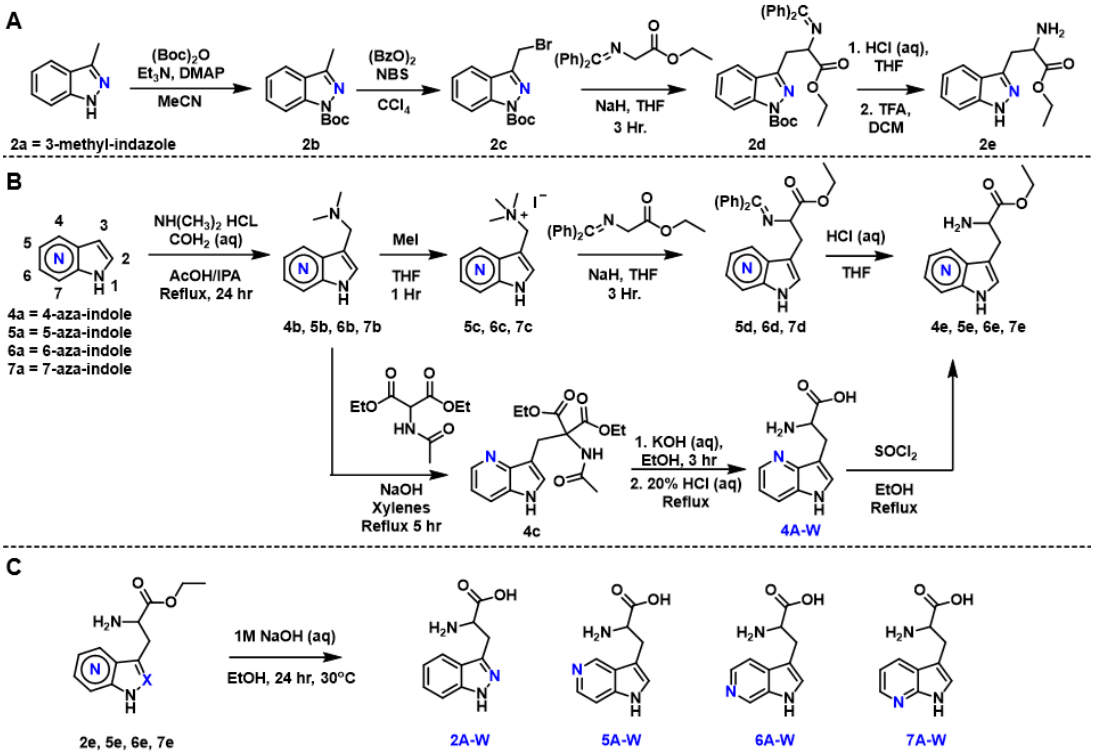


Figure 3.4: Combined synthetic schemes for Aza-Trp isomers

The azaindole substituted substrate analogs were prepared from commercially available azaindole building blocks. (Figure 3.4) The 2-azatryptophan was prepared from N-Boc-protected 3-methyl-indazole following a radical bromination of the methyl group using known methods.[51] The amino acid moiety was introduced via substitution of the benzophenone imine of glycine ethyl ester, followed by subsequent imine hydrolysis using aqueous HCl. (Figure 3.4) Compared to indole, which has the highest electron density at C3, the adjacent nitrogen atoms within the indazole scaffold withdrawal electron density from C3, effectively preventing typical indole reactions such as gramine formation (**4b**, **5b**, **6b**, **7b**). The synthetic strategy used in this work was inspired by a reported method[52]

of transforming unsubstituted 7-azaindole **(7a)** to the corresponding **7A-W** using well known 'classic' indole transformations.

3.3.2 Synthetic Transformations of Aza-Gramines to Aza-Tryptophans

For synthetic methodology development, one goal of this work was to determine if this reported **7A-W** synthesis route could be used to generate other Aza-Trp isomers. The synthesis of 7-aza-tryptophan (**7A-W**) was accomplished using a previously reported method, involving synthesis of the 7-aza-gramine (**7b**) from 7-azaindole, followed by methylation of the gramine to form a leaving group, and subsequent displacement of the trimethylammonium moiety with a protected amino acid functionality (N-(diphenylmethylene)glycine ethyl ester). Since the azaindoles containing nitrogen atoms on the benzene ring have significant electronic similarities to indole, the aza-gramine synthesis of 4-, 5-, and 6-aza substituted tryptophan analogs (**4A-W**, **5A-W, 6A-W**) was pursued using similar procedure as that of **7A-W** using commercially available azaindole building blocks.

The azaindoles containing nitrogen atoms on the benzene ring have significant electronic similarities to indole, and thus were capable of aza-gramine formation in the presence of dimethylamine and formaldehyde in high yields. (Figure 3.4-B) The aza-gramine intermediates were methylated to form an insoluble quaternary ammonium salt (**5c**, **6c**, **7c**). The salts were then subjected to the protected glycine ethyl ester under basic conditions to yield protected Aza-Trp intermediates **5d**, **6d**, and **7d**. Upon imine hydrolysis, the benzophenone byproduct was readily extracted from the reaction mixture using diethyl ether, and the highly fluorescent azatryptophan esters (**5e**, **6e**, **7e**) which were extracted

using 3:1 $CHCl_3$:IPA and subsequently purified via silica gel chromatography using $CHCl_3$:MeOH:NH_4OH as the eluent.

The 4-azagramine however was found to be robustly resistant to this alkylation method. The methylation of 4-azagramine produced an insoluble brick red sticky solid, while the other quaternary salts generated fine powders. An alternative synthetic route was required to synthesize the **4A-W**.[53] The 4-azagramine **4b** was refluxed in xylenes in the presence of acetamidomalonate and catalytic NaOH to generate intermediate **4c**, which was then saponified, and subsequently decarboxylated/deacylated via extended reflux in 20% aqueous HCl. The pure 4-azatryptophan (**4A-W**) was recovered via recrystallization, and then esterified to form **4e**. Notably, this process of converting 4-azaindole into the pure 4-azatryptophan does not require column chromatography, since the 4-azagramine and 4-azatryptophan were readily recrystallized. While the **4A-W** was recovered by recrystallization, the other Aza-Trp isomers were obtained by saponification of the Aza-Trp ethyl ester intermediates (**2e, 5e, 6e, 7e**) by overnight stirring in 2:1 Ethanol:1 M NaOH (aq) at 30°C and subsequent RP-HPLC purification (Figure 3.4). This work provides a new method of synthesizing **5A-W** as well as **6A-W**.

3.3.3 Analytical Scale Screening of Aza-Trp Substrates with PT's

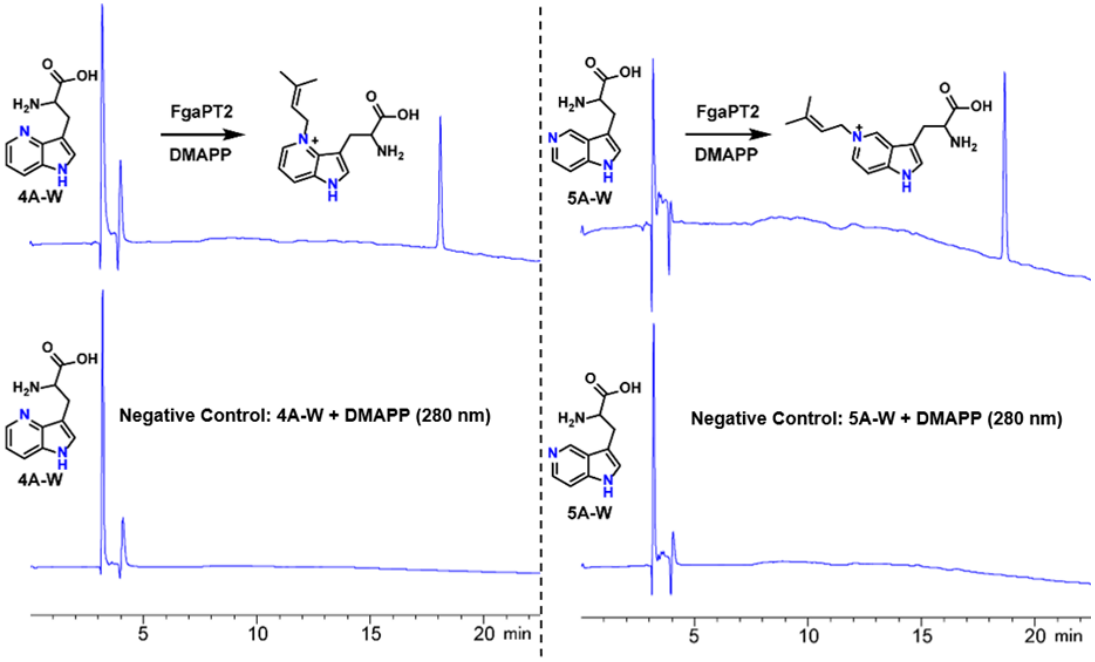


Figure 3.5: Analytical scale RP-HPLC traces of PT catalyzed prenylation of AzaTrp substrates

In order to access the ability of Aza-Trp (**2A-W, 4A-W, 5A-W, 6A-W, 7A-W**) to act as substrates for prenyltransferases (PTs), three PTs were selected. The selected PTs, all of which were originally isolated from *Aspergillus fumigatus*[54-56] include FgaPT2, a Trp C4-PT, CdpNPT, a cyclic-Trp-Pro dipeptide C3-reverse PT, and FtmPT1, a brevianamide C2-PT. All Aza-Trp analogs were evaluated with all three PTs under standard reaction conditions in the presence of DMAPP as the prenyl group donor. Appropriate controls were included to distinguish between enzymatic and non-enzymatic transformations. Formation of products were monitored using differences in retention time between the substrate and products using RP-HPLC at 280 nm. Among the aza-Trp analogs, only **4A-W** and **5A-W** served as substrates for FgaPT2, and produced 23.6% and 71% prenylated products respectively (Figure 3.5). All other PT's screened with Aza-Trp as their acceptor substrates did not produce observable products under these reaction conditions.

3.3.4 Structural Determination of Chemoenzymatically Prenylated Aza-Trp

Structural determination of the reaction products was carried out using NMR spectroscopy including 1D ^{1}H, COSY, HSQC, and HMBC methods. All isolated Aza-Trp reaction products were NMR characterized in D_2O. The Aza-Trp analogs and their corresponding products have remarkably high polarity and were only able to be isolated as ionized species due to their two distinct basic N atoms, and the unprotected carboxylic acid moiety. This limited solubility in d_6-DMSO, thus requiring the use of D_2O as the NMR solvent.

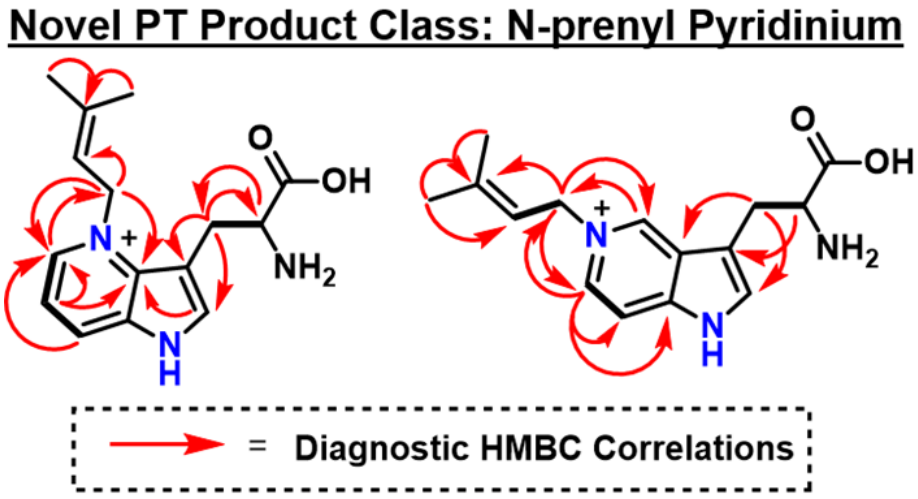


Figure 3.6: NMR characterized products for FtmPT2 catalyzed prenylation of 4A-W and 5A-W

FgaPT2 is known to efficiently prenylate the indole C4 position of tryptophan. In the case of two synthetic Aza-Trp isomers, **4A-W** and **5A-W**, the basic pyridine N atom is oriented near the C1' position of DMAPP, bearing a nucleophilic lone pair perfectly positioned to capture the dimethylallyl carbocation intermediate. The two Aza-Trp isomers that showed product formation via RP-HPLC on an analytical scale (**4A-W** and **5A-W**)

were scaled up (as described in Section 3.2.10), extracted with isopropanol, and purified via semi-prep RP-HPLC.

Observed HMBC correlations are shown as red arrows in Figure 3.6. Substitution of aromatic carbon positions was quickly ruled out by ^{1}H and HSQC experiments featuring the same number of aromatic CH signals as the Aza-Trp substrates. Extensive HMBC correlations between the prenyl C1' CH_2 and pyridine ring carbons were observed, while no correlations were found between the azaindole-C2 position and the prenyl group. These observations unequivocally demonstrate that FgaPT2 catalyzes pyridine N-prenylation in the presence of DMAPP and either **4A-W** or **5A-W.** To the best of our knowledge, these products are the only known examples of aromatic prenyltransferases alkylating an aromatic nitrogen atom, forming a N-prenylpyridinium product. Additionally, this study demonstrates for the first time that azaindoles are substrates of indole PT's.

3.3.5 Comparison of Azaindole isomer-dependent Properties

The position of the nitrogen isosteric replacement has been clearly demonstrated to alter reactivity within the binding site. It is known that the most electron rich position of indazole is C5, which is fundamentally different from pyrrole fused heterocycles. The lack of a pyrrole ring in the indazole substrates prevents effective comparison to the other azaindoles regarding the reactivity of the 5 membered ring. Compared to indole, all the azaindole isomers have significantly reduced electron density. Aza-substitution on the 4, 5, 6 and 7 positions of indole ring function as nucleophilic bases, imparting significantly altered reactivity compared to indole in the context of electrophilic aromatic substitution. The basic pyridine nitrogen within **4a**, **5a**, **6a**, and **7a** have a pKa of 6.9, 8.3, 7.9, and 4.6 respecatively.[34, 57] The decreased ability of azaindoles to participate in electrophilic

aromatic substitution is highly relevant to PT based catalysis, as nearly all indole PTs rely on this alkylation mechanism.

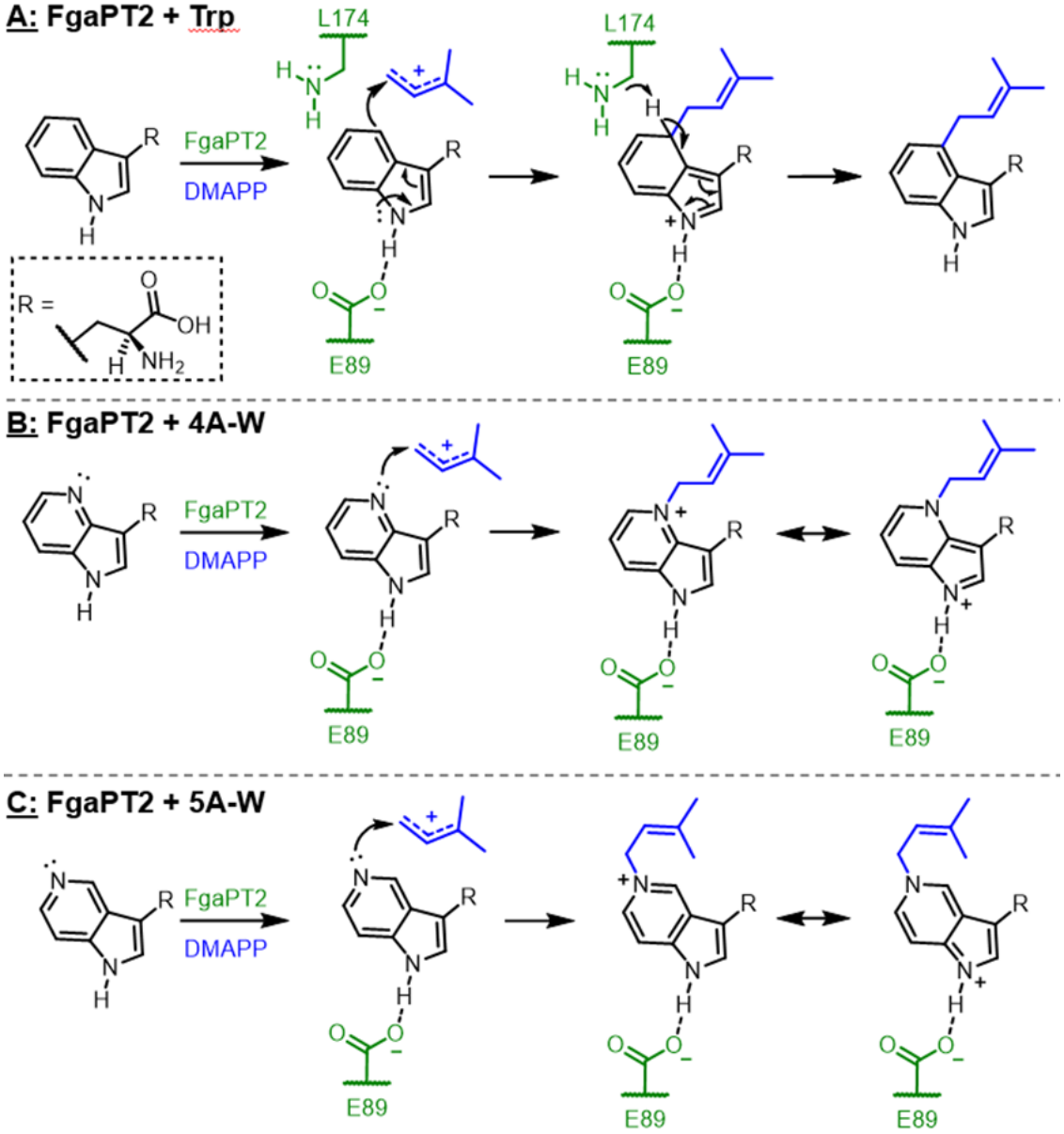


Figure 3.7: Proposed mechanism for FgaPT2 catalyzed prenylation of its native substrate (A), 4A-W (B), and 5A-W (C)

The acceptor binding site for PTs is hydrophobic in nature and optimized for carbocation stabilization and capture. The observation that FgaPT2, which typically alkylates C4 of Trp, was able to N-alkylate **4A-W** and **5A-W** could be explained through

a predominantly steric-based argument. However, it was also observed during scale up that **5A-W** reached maximum conversion (~70%) within 4 hours, while **4A-W** was only capable of reaching ~40% conversion after 48 hours, with an additional dose of enzyme after 24 hours. The mechanism of FgaPT2 catalyzed Trp prenylation relies on E89 for orienting the substrate and stabilizing the intermediate, and the basic L174 for deprotonating the C4 position to restore aromaticity (Figure 3.7-A). These accepted Aza-Trp substrates undergo N4 (Figure 3.7-B) or N5 (Figure 3.7-C) prenylation through an alternative mechanism, direct nucleophilic attack of the aromatic N atom lone pair onto the dimethylallyl carbocation. This directly forms the observed products without generating an arenium intermediate required for the native C4 prenylation of Trp. The native cation-stabilizing effect of E89 remain highly applicable to these unnatural substrate transformations.

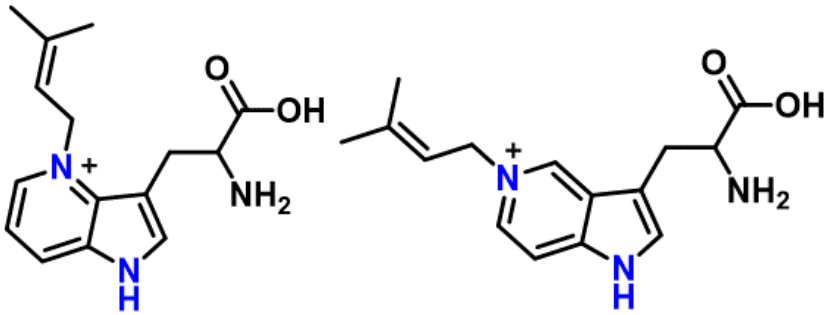
3.3.6 Delocalization and Stabilization by Azaindoles

Analogous to the native prenylation mechanism, a positive charge is formed on the acceptor moiety and subsequently delocalized to the N-1 position to generate a stabilizing salt bridge with E89. Several factors are hypothesized to contribute to the higher percent conversion of **5A-W** compared to **4A-W**. Compared to 4-azaindole, 5-azaindole has a more nucleophilic N atom which upon N-alkylation generates a more thermodynamically stable cationic species due to enhanced resonance delocalization of the positive charge to the "para" oriented N1 (Figure 11-C), analogous to the common catalytic base 4-dimethylaminopyridine. This enhanced product stabilization is likely the primary thermodynamic driving force behind the higher % conversion observed for FgaPT2 catalyzed N-prenylation of **5A-W** compared to **4A-W.**

3.4 Conclusions

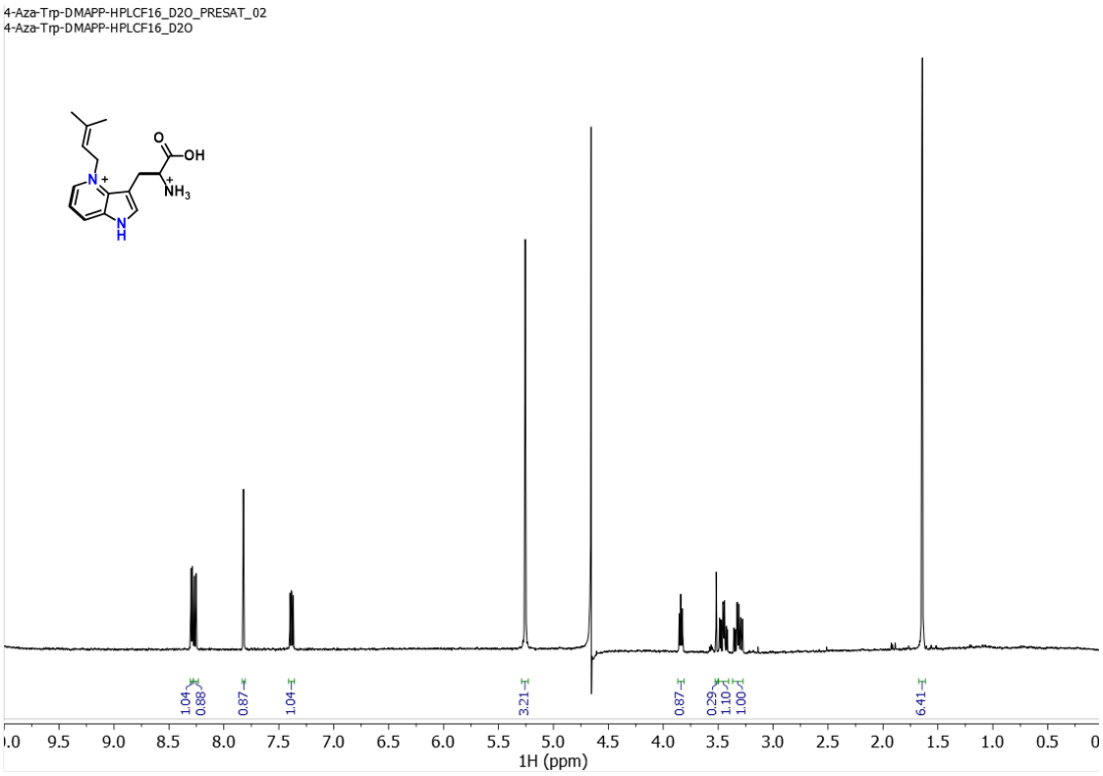
This study has demonstrated that the aromatic prenyltransferases FgaPT2 can tolerate isosteric replacement of the acceptor indole moiety with a variety of azaindole isomers. Notably, with **4A-W** and **5A-W**, N-prenylation of a pyridine ring was observed as a PT catalyzed reaction product for the first time. Selective N-alkylation of a pyridine moiety in the presence of an unprotected primary amine is extremely challenging to achieve using modern synthetic methods, making an excellent example for the advantages of PT-based biocatalytic transformations. These results have also demonstrated the utility of PT application through precursor directed biosynthesis methods for natural product analog synthesis and diversification. The results of this study will help guide future substrate and protein engineering efforts to improve the synthetic utility of PTs in drug discovery.

3.5 Appendix 1: NMR Assignments and Spectra

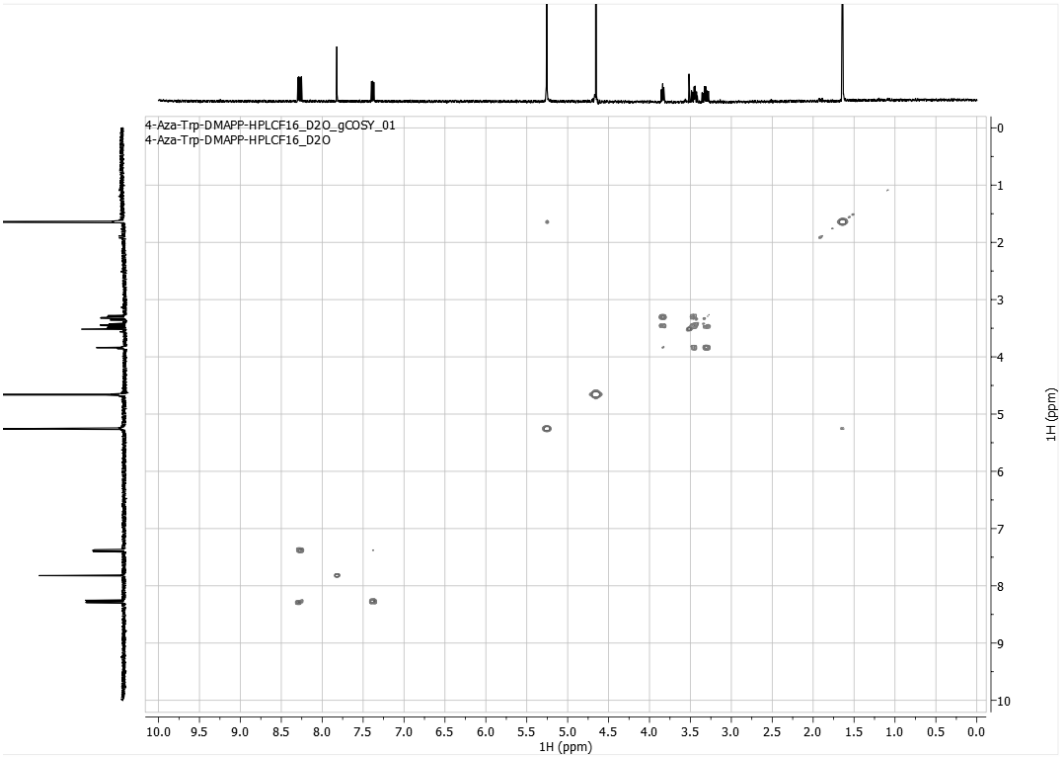


		4-Aza-Trp-DMA (D_2O)		5-Aza-Trp-DMA (D_2O)	
Segment	**Position**	***δ*C , type**	***δ*H**	***δ*C , type**	***δ*H**
L-Trp	NH				
	α	55.5, CH	3.84 (t)	54.5, CH	4.00 (t)
	β	27.1, CH_2	3.31 (td) ,3.45 (td)	25.6, CH_2	3.41 (m)
	1				
	2	136.1, CH	7.82 (s)	131.5, CH	7.63 (d)
	3	104.2, C		111.0, C	
	3a	133.9, C		124.3, C	
	4			136.8, CH	8.98 (s)
	5	137.4, CH	8.29 (d)		
	6	117.2, CH	7.38 (dd)	133.6, CH	8.17 (d)
	7	127.5, CH	8.26 (d)	109.7, CH	7.78 (d)
	7a	134.7, C		141.6, C	
	C=O	172.7, C		173.6, C	
N-Prenyl	1'	54.5, CH_2	5.26 (s)	57.6, CH_2	5.10 (d)
	2'	115.2, CH	5.26 (s)	116.2, CH	5.49 (t)
	3'	143.3, C		143.3, C	
	4'	24.4, CH_3	1.65 (s)	24.8, CH_3	1.79 (s)
	5'	17.5, CH_3	1.65 (s)	17.3, CH_3	1.81 (s)

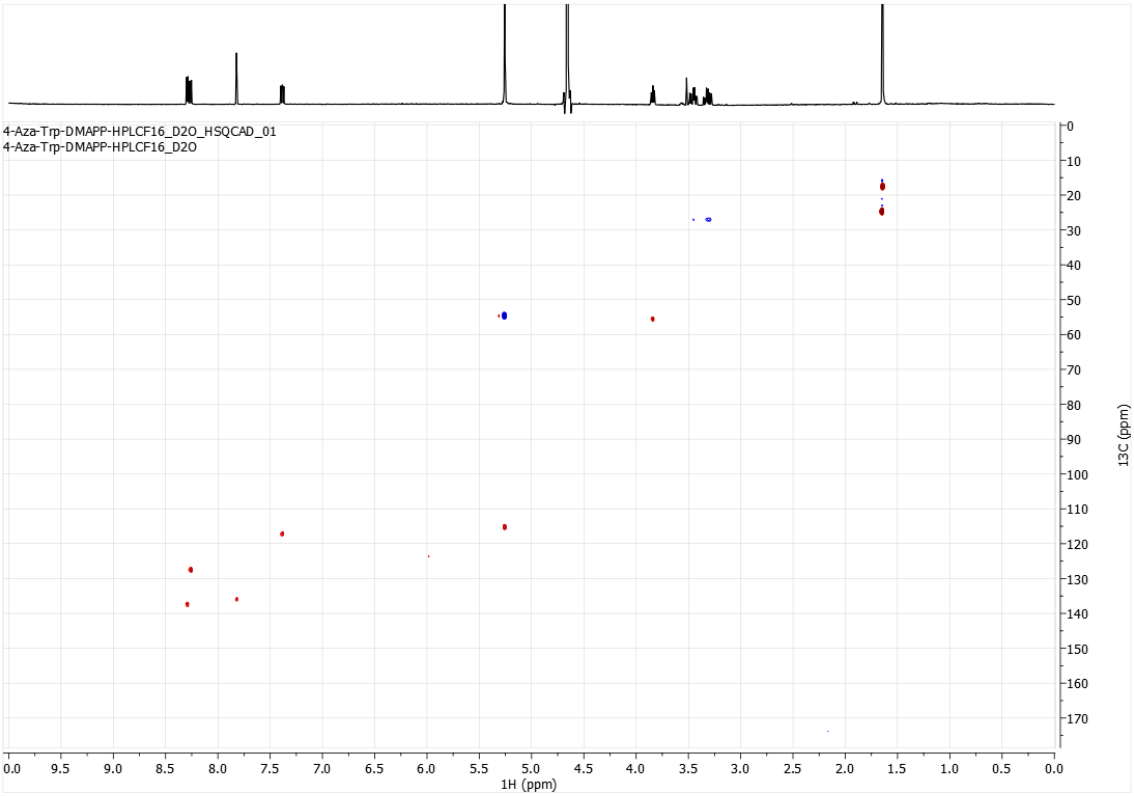
Table 3.1: NMR assignments for 4-Aza-Trp-DMA and 5-Aza-Trp-DMA



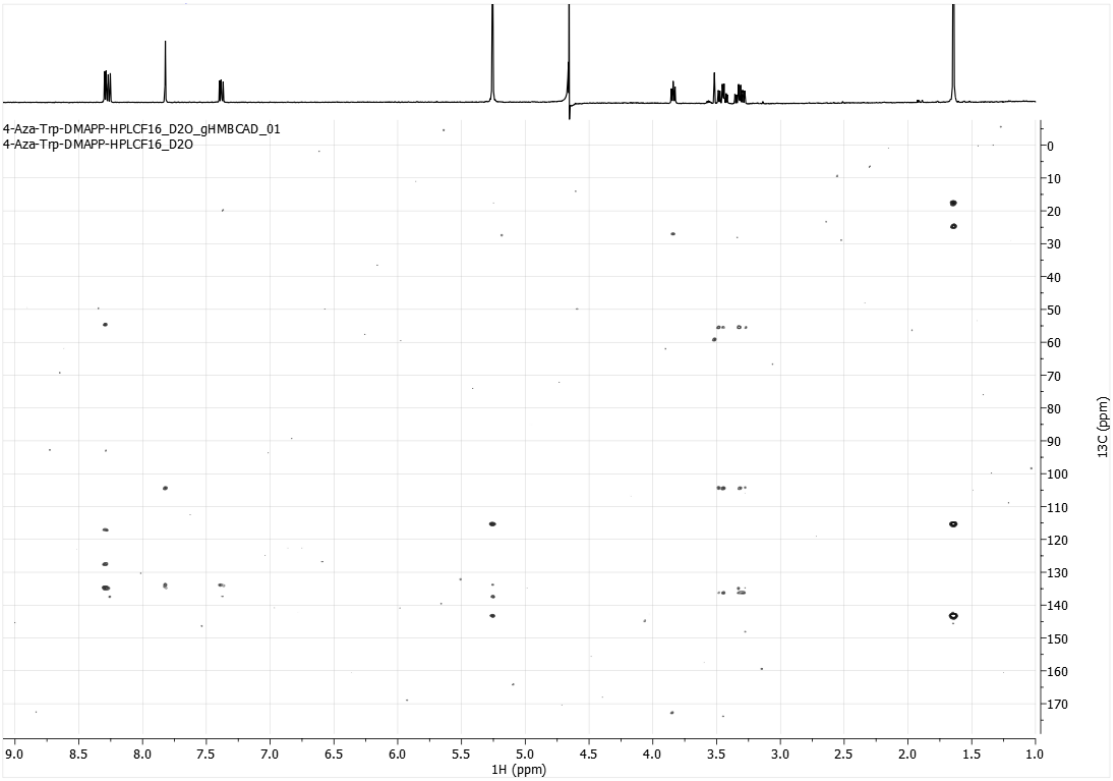
^{1}H NMR spectrum (500 MHz) of **4-Aza-Trp-DMA** in D_2O



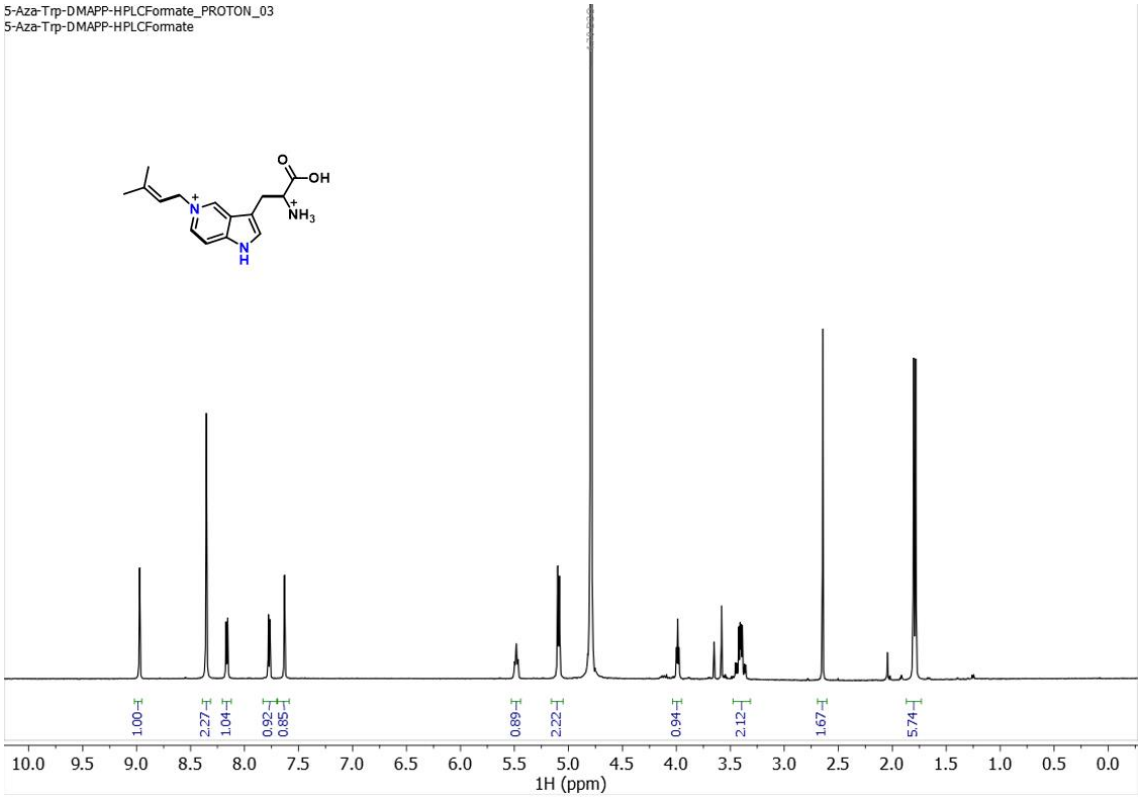
^{1}H-COSY spectrum (500 MHz) of **4-Aza-Trp-DMA** in D_2O



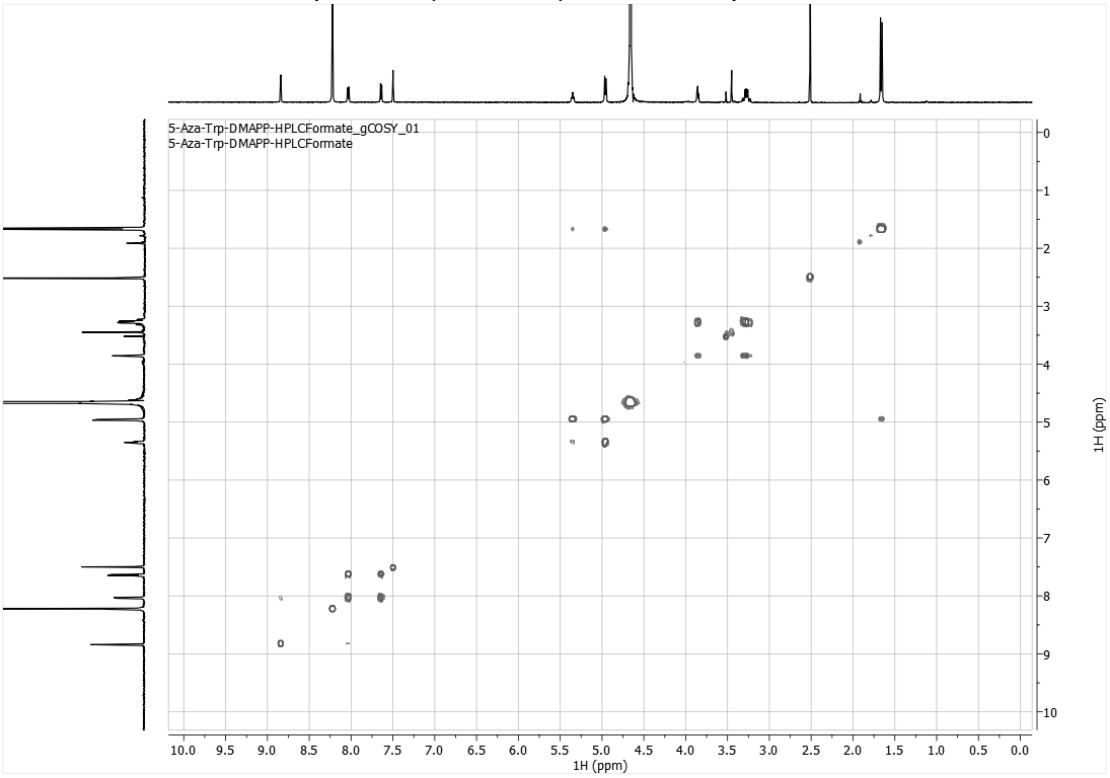
HSQC spectrum (500 MHz) of **4-Aza-Trp-DMA** in D_2O



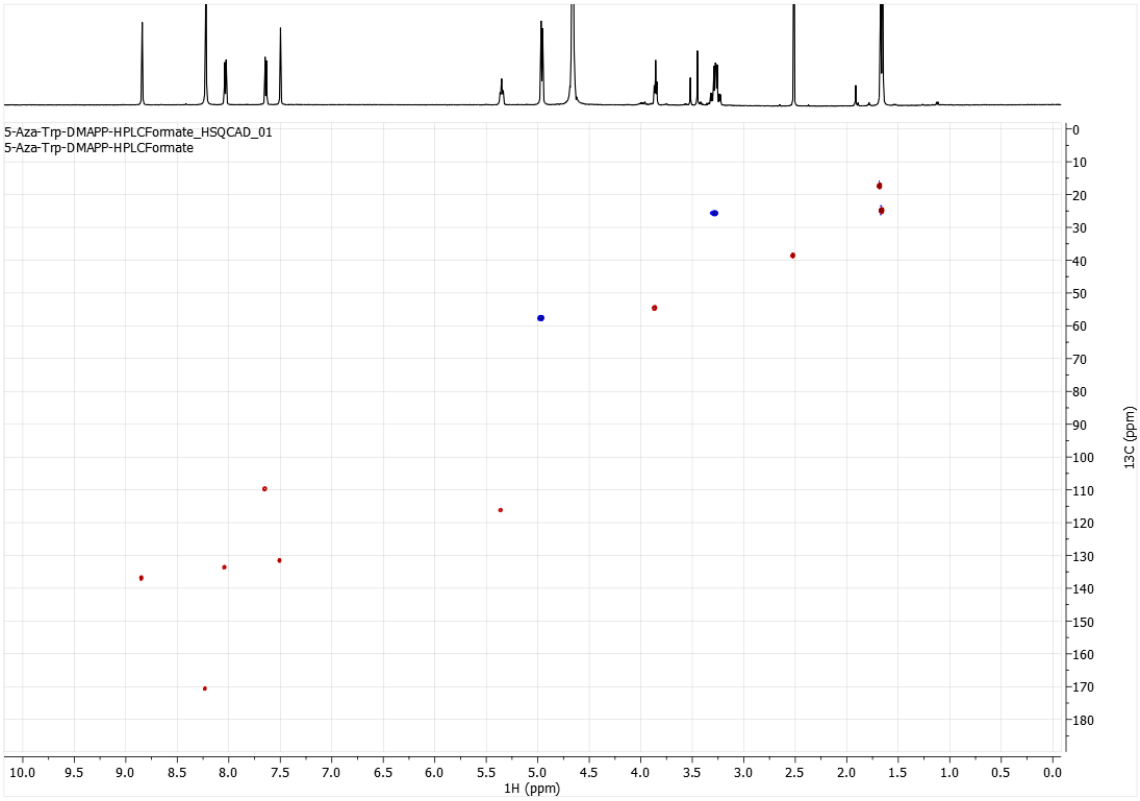
HMBC spectrum (500 MHz) of **4-Aza-Trp-DMA** in D_2O



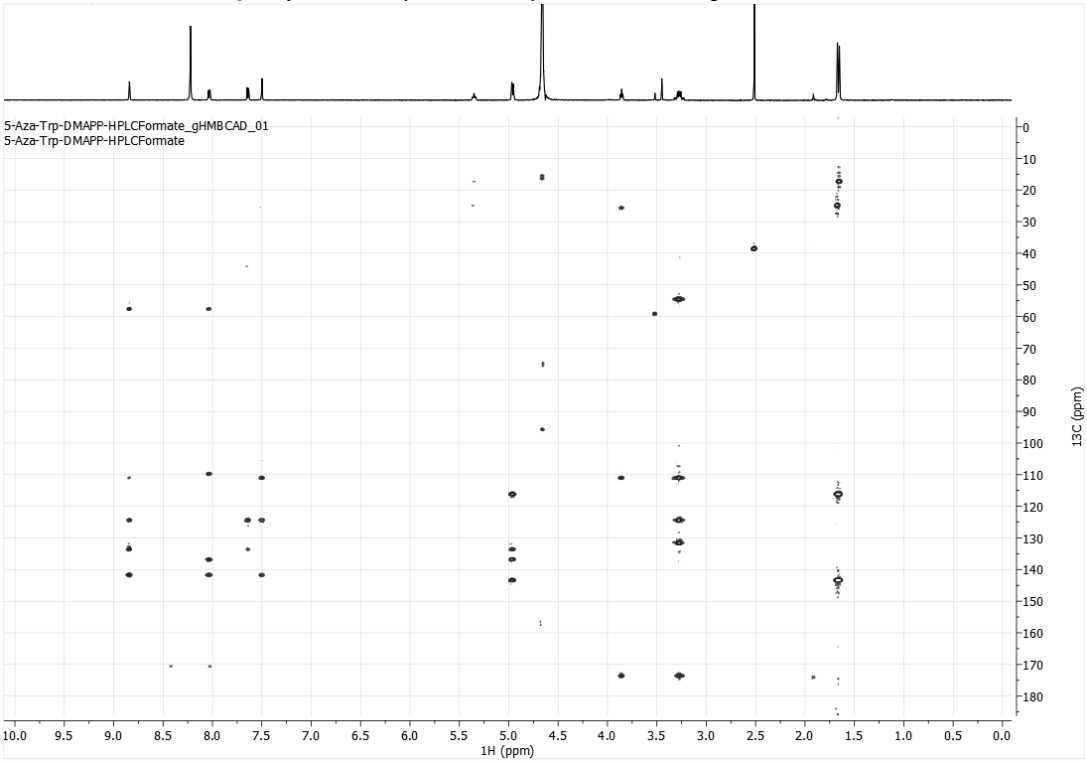
1H NMR spectrum (500 MHz) of 5-Aza-Trp-DMA in D2O



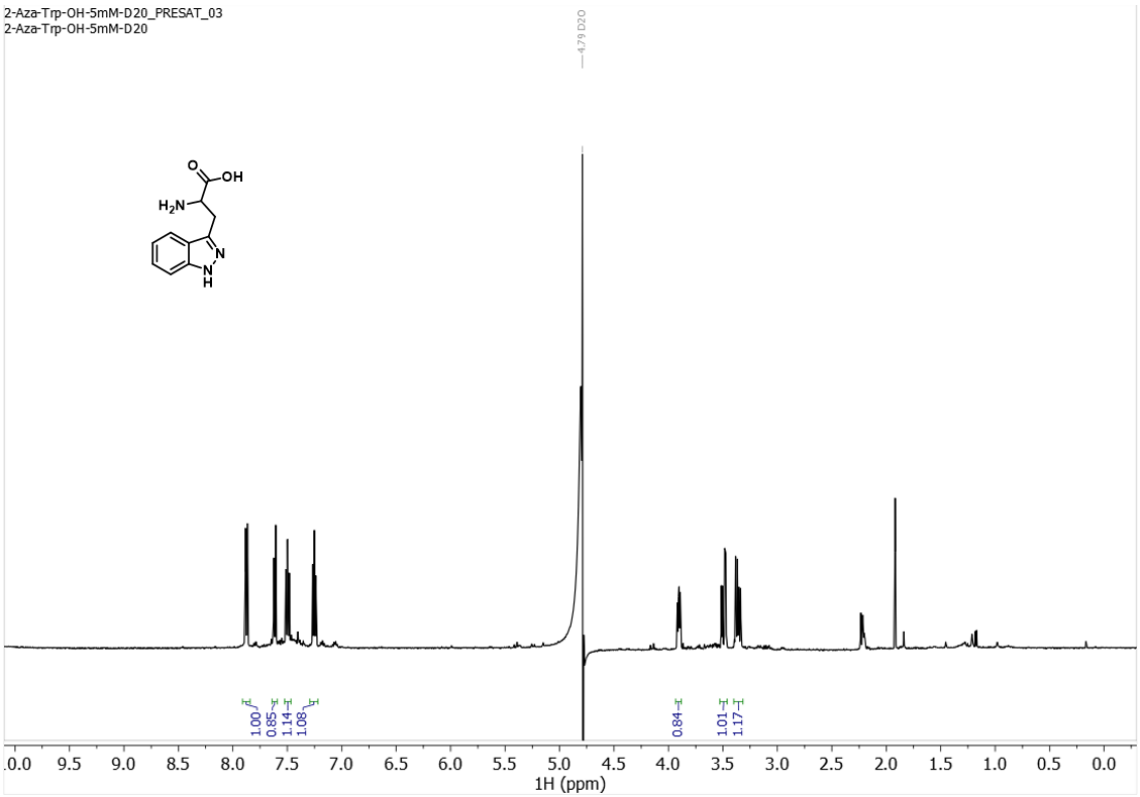
^{1}H-COSY spectrum (500 MHz) of **5-Aza-Trp-DMA** in D_2O



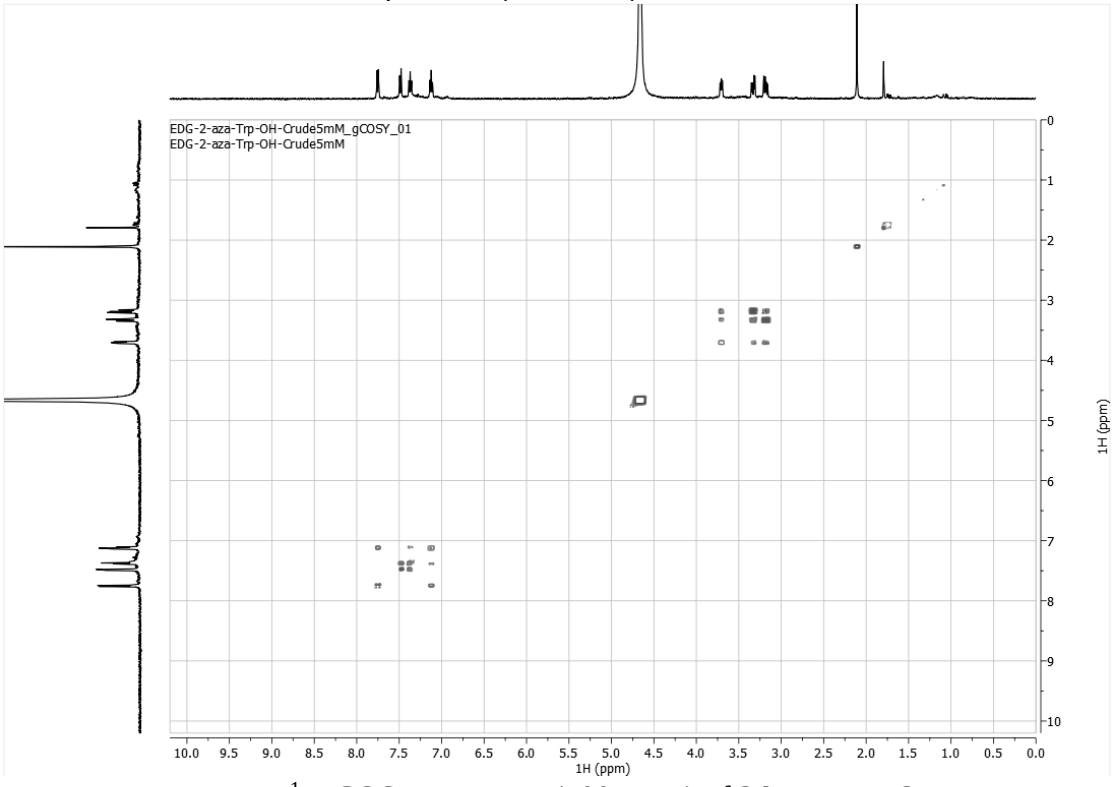
HSQC spectrum (500 MHz) of **5-Aza-Trp-DMA** in D_2O



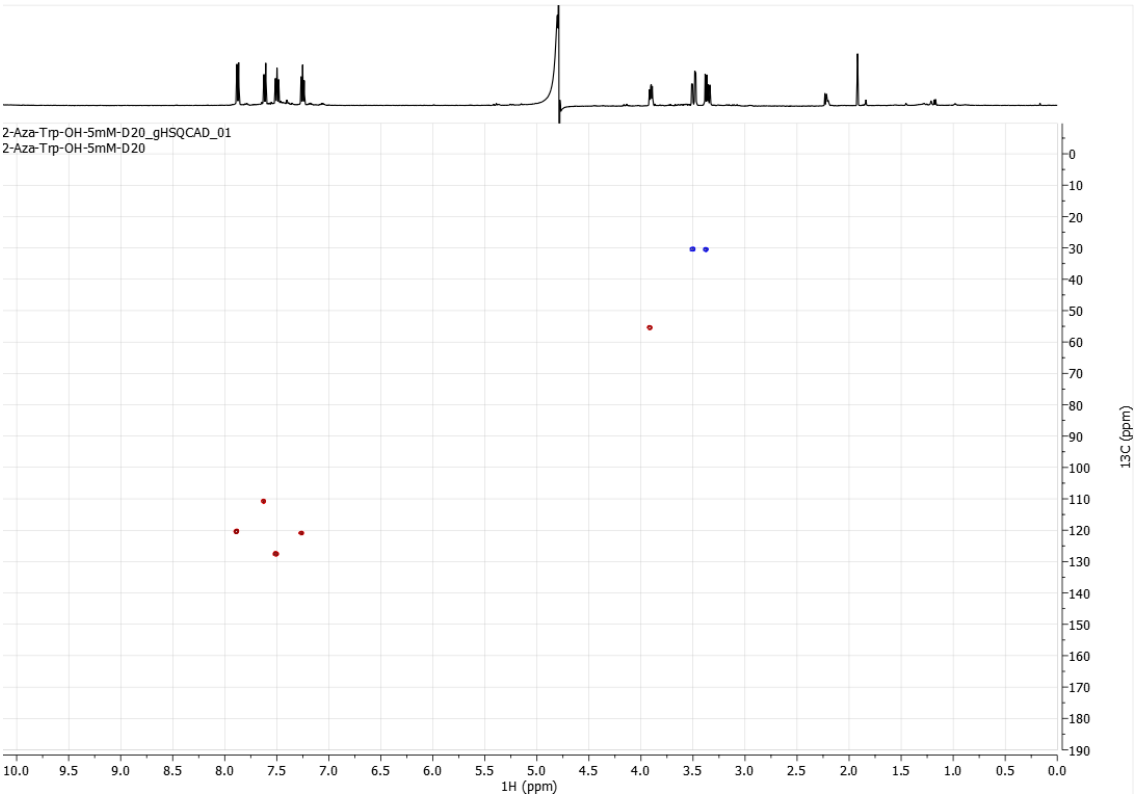
HMBC spectrum (500 MHz) of **5-Aza-Trp-DMA** in D_2O



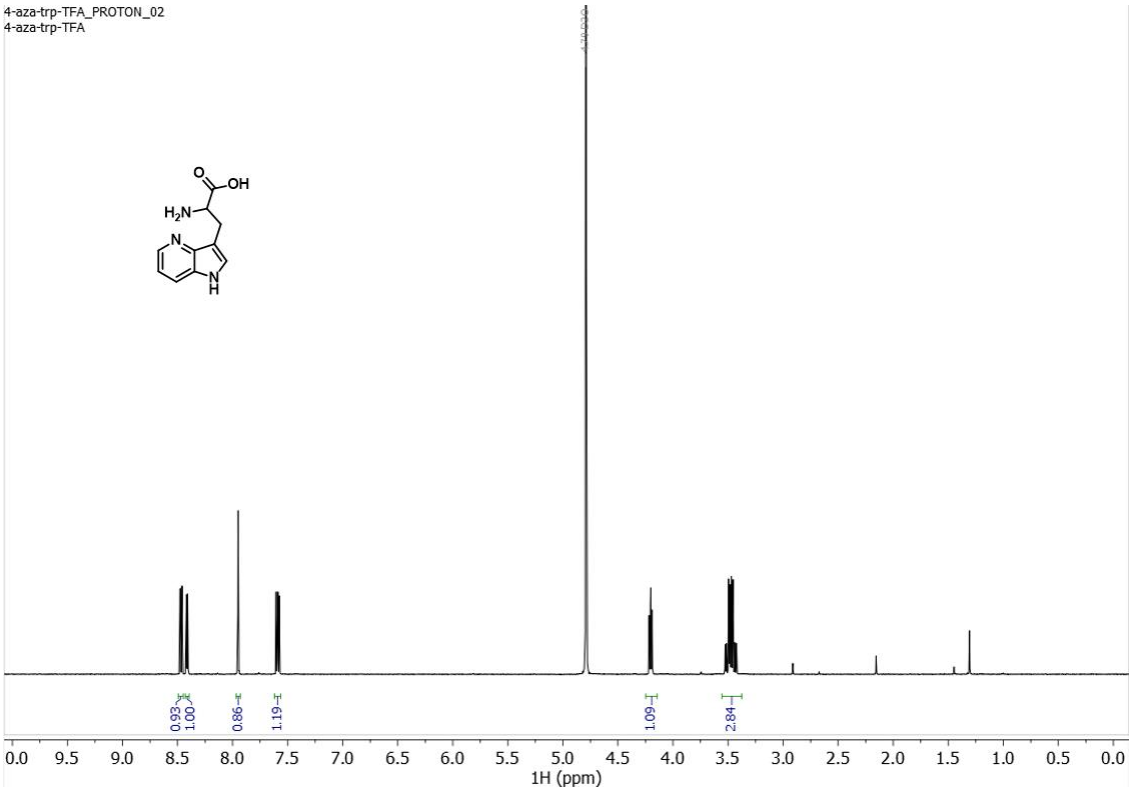
^{1}H NMR spectrum (500 MHz) of **2A-W** in D_2O



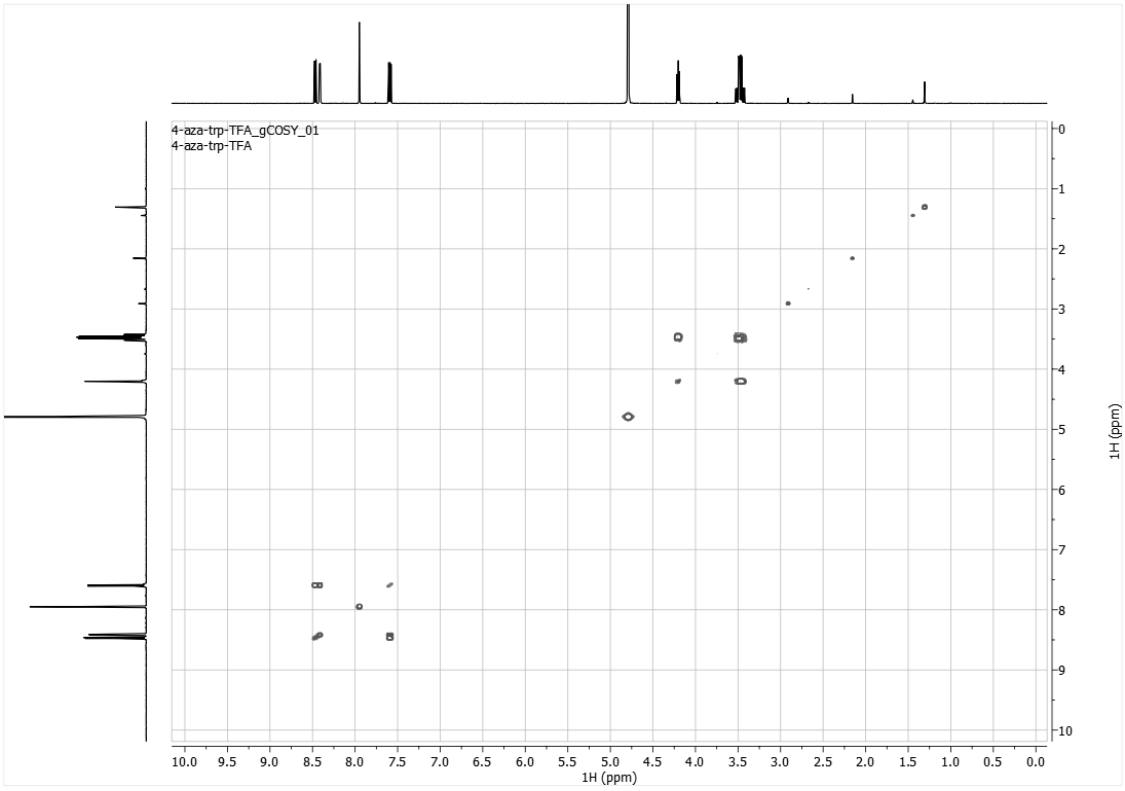
^{1}H-COSY spectrum (500 MHz) of **2A-W** in D_2O



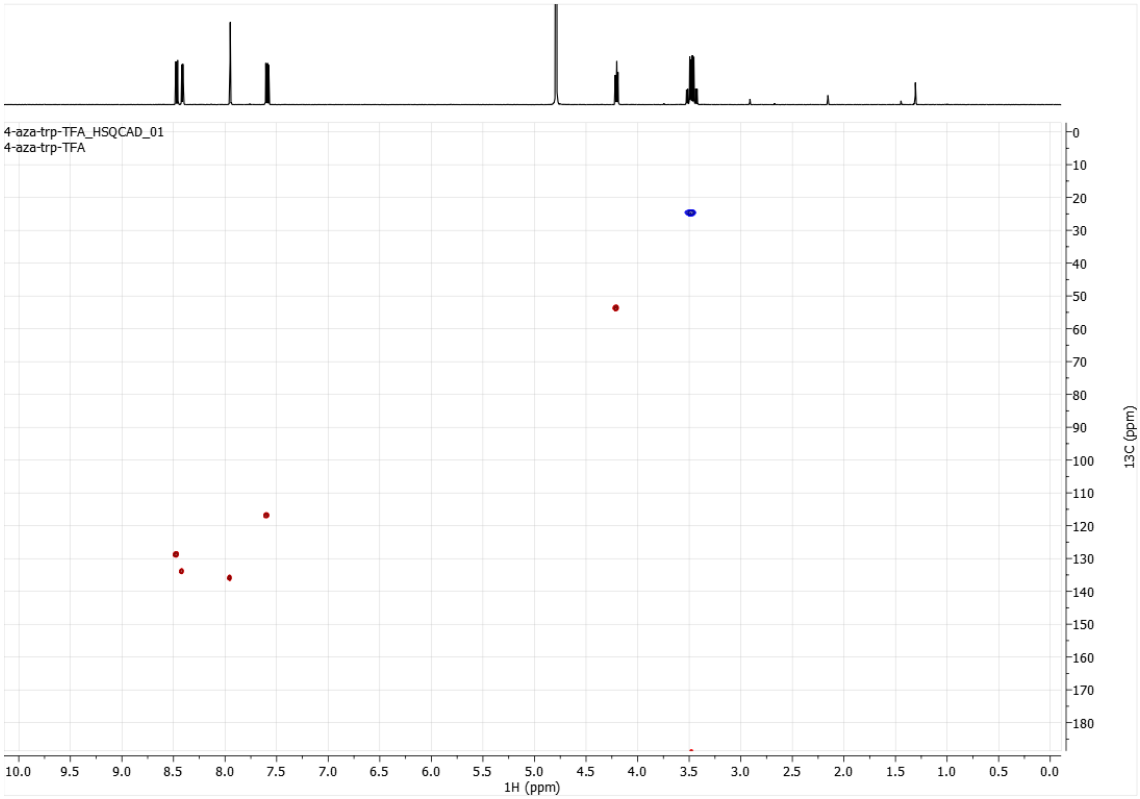
HSQC spectrum (500 MHz) of **2A-W** in D_2O



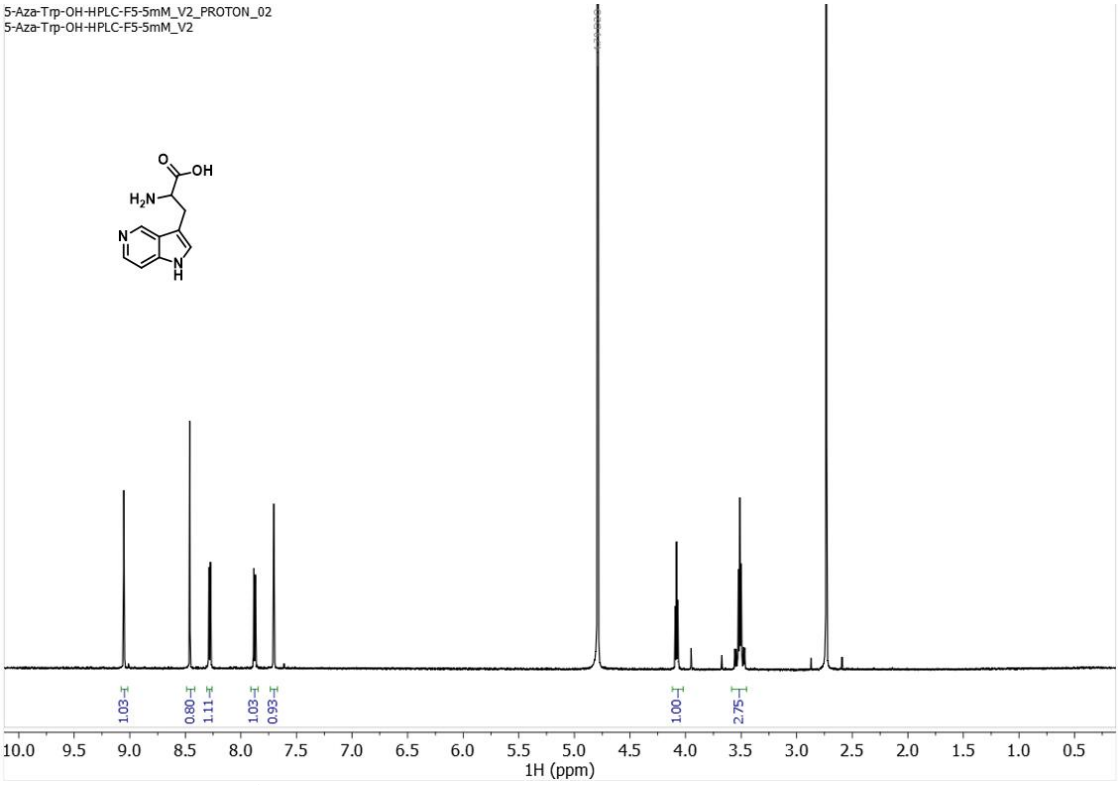
^{1}H NMR spectrum (500 MHz) of **4A-W** in D_2O



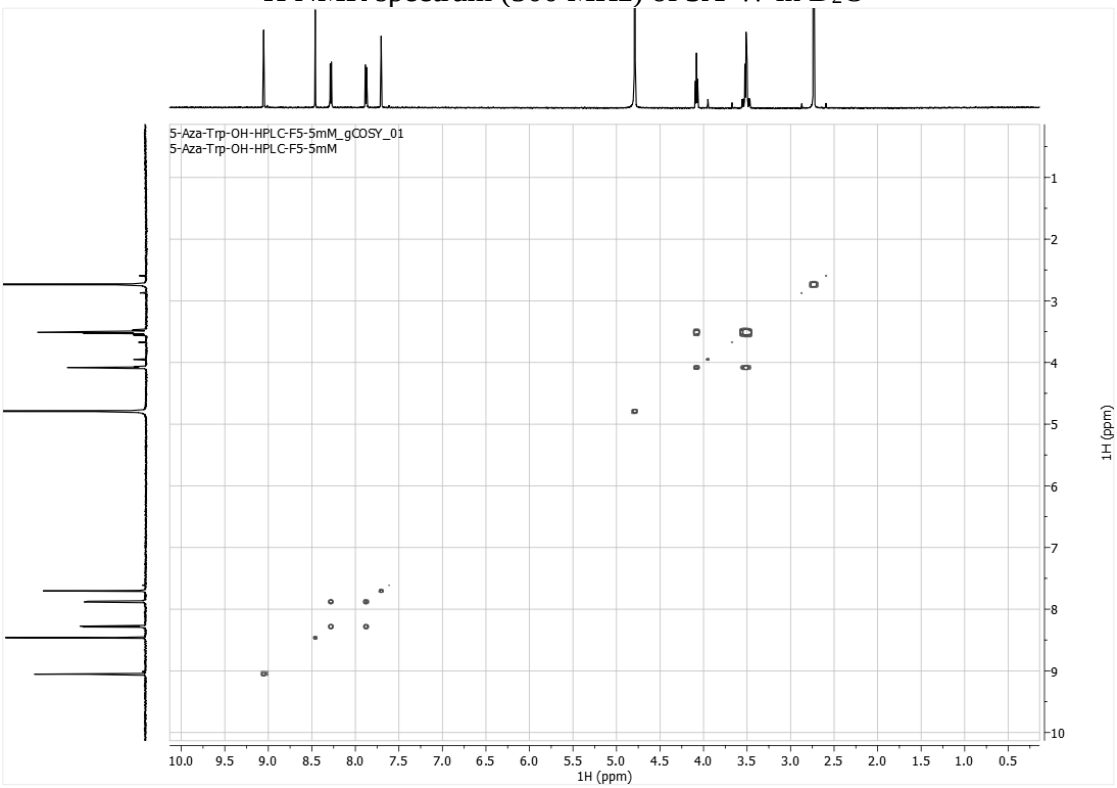
^{1}H-COSY spectrum (500 MHz) of **4A-W** in D_2O



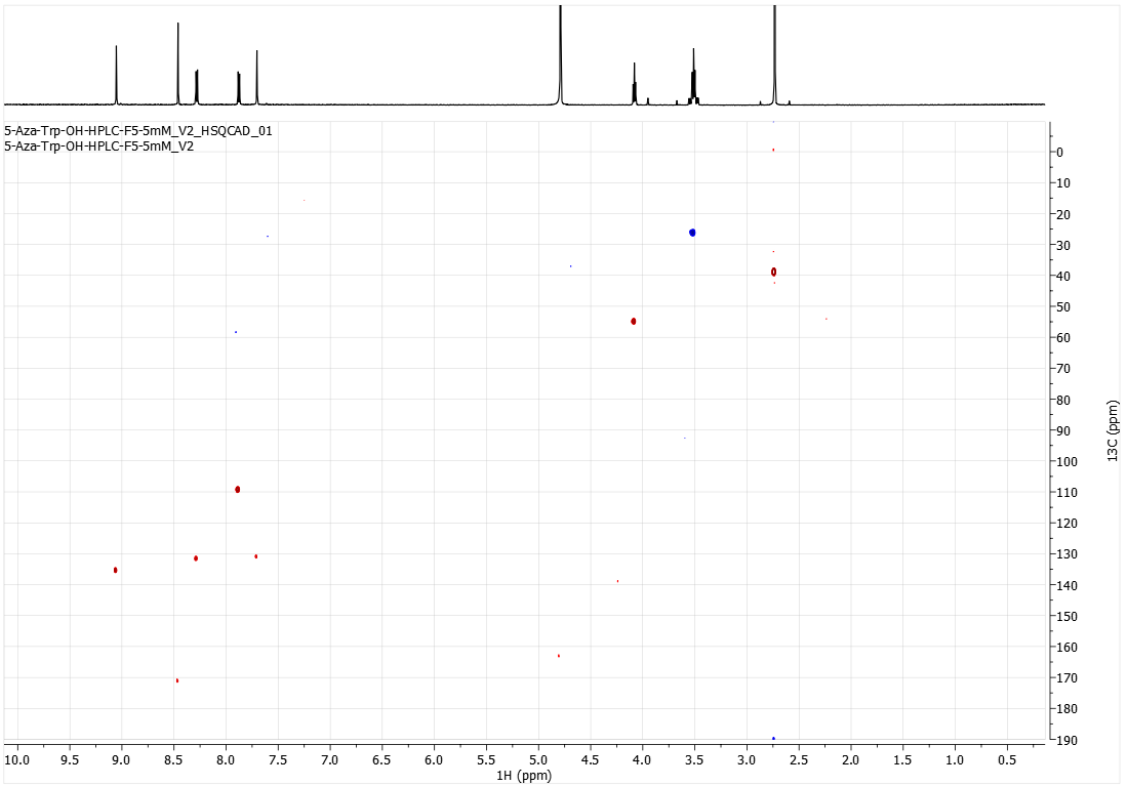
HSQC spectrum (500 MHz) of **4A-W** in D_2O



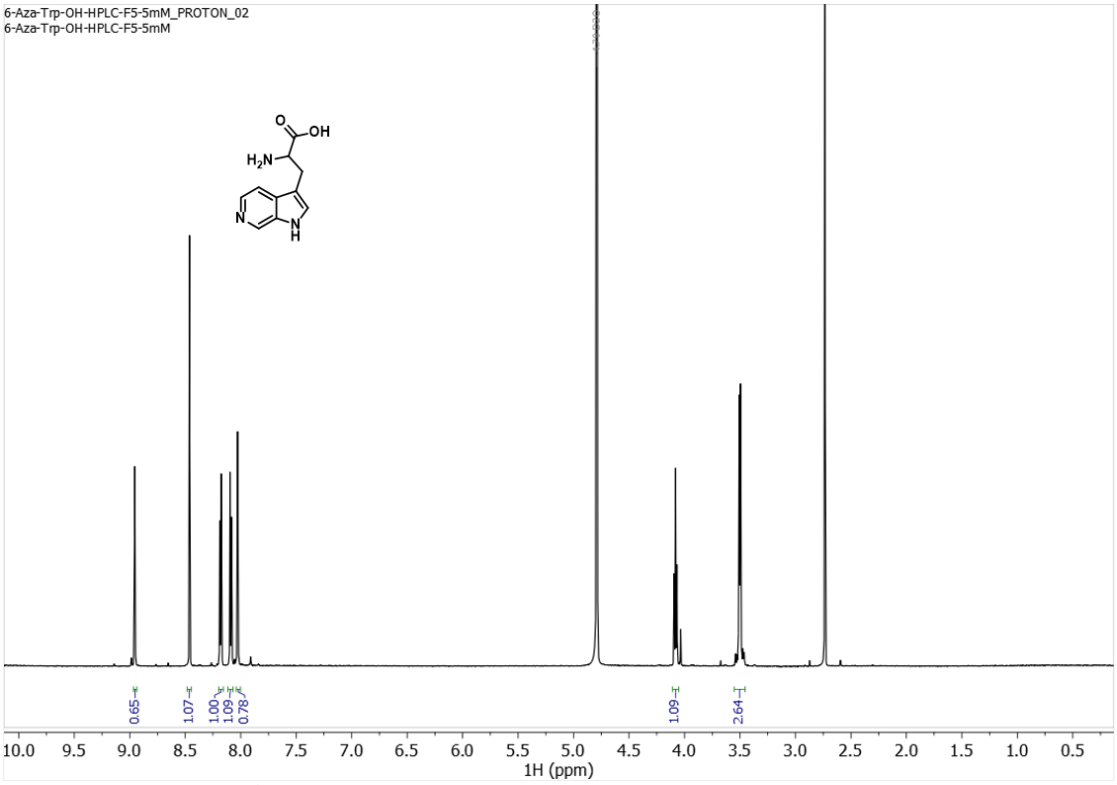
^{1}H NMR spectrum (500 MHz) of **5A-W** in D_2O



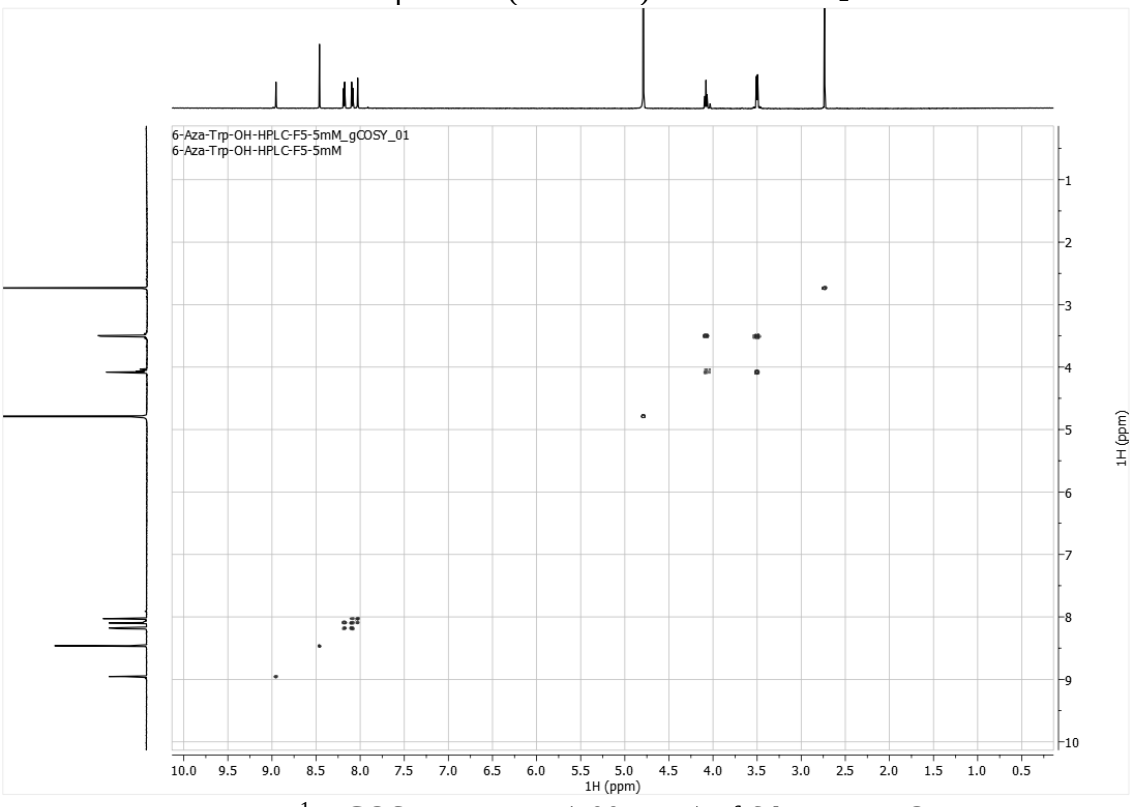
^{1}H-COSY spectrum (500 MHz) of **5A-W** in D_2O



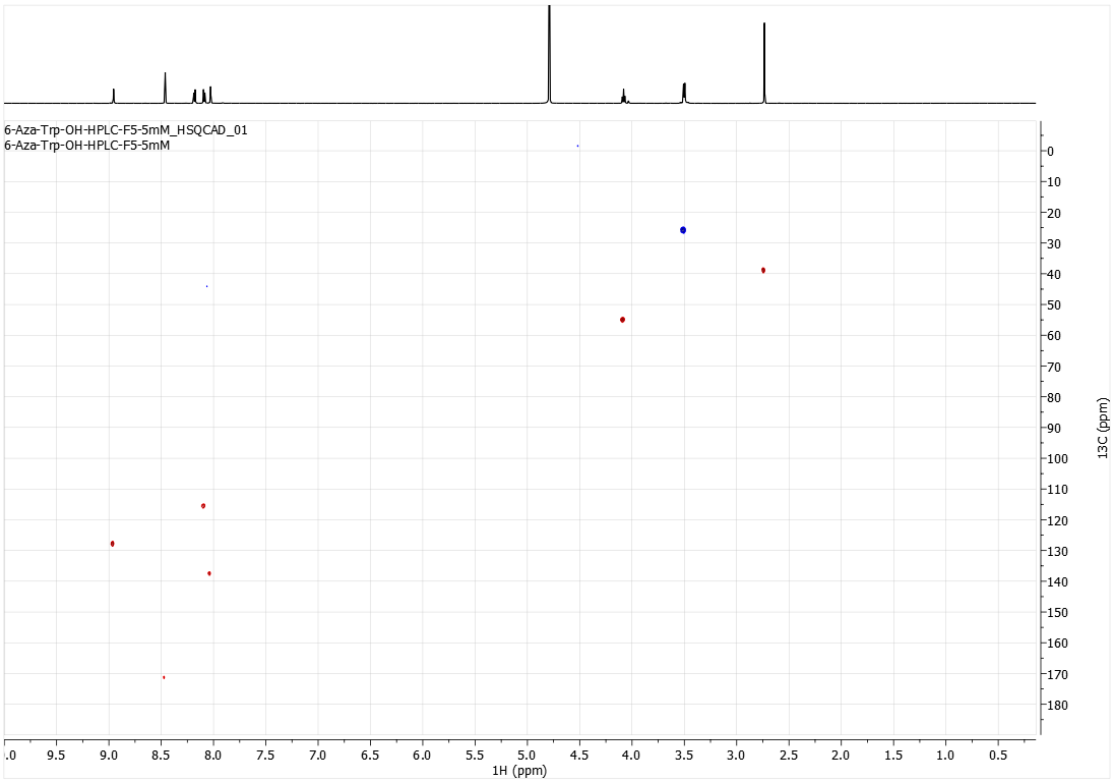
HSQC spectrum (500 MHz) of **5A-W** in D_2O



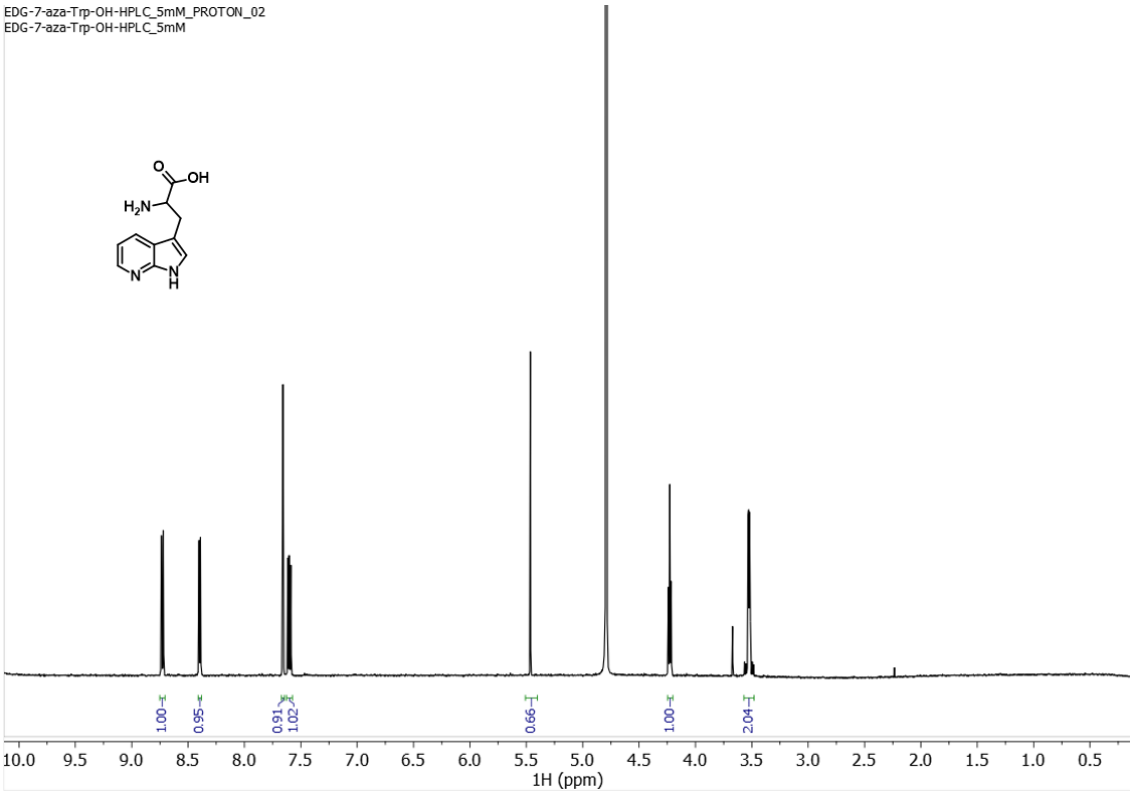
^{1}H NMR spectrum (500 MHz) of **6A-W** in D_2O



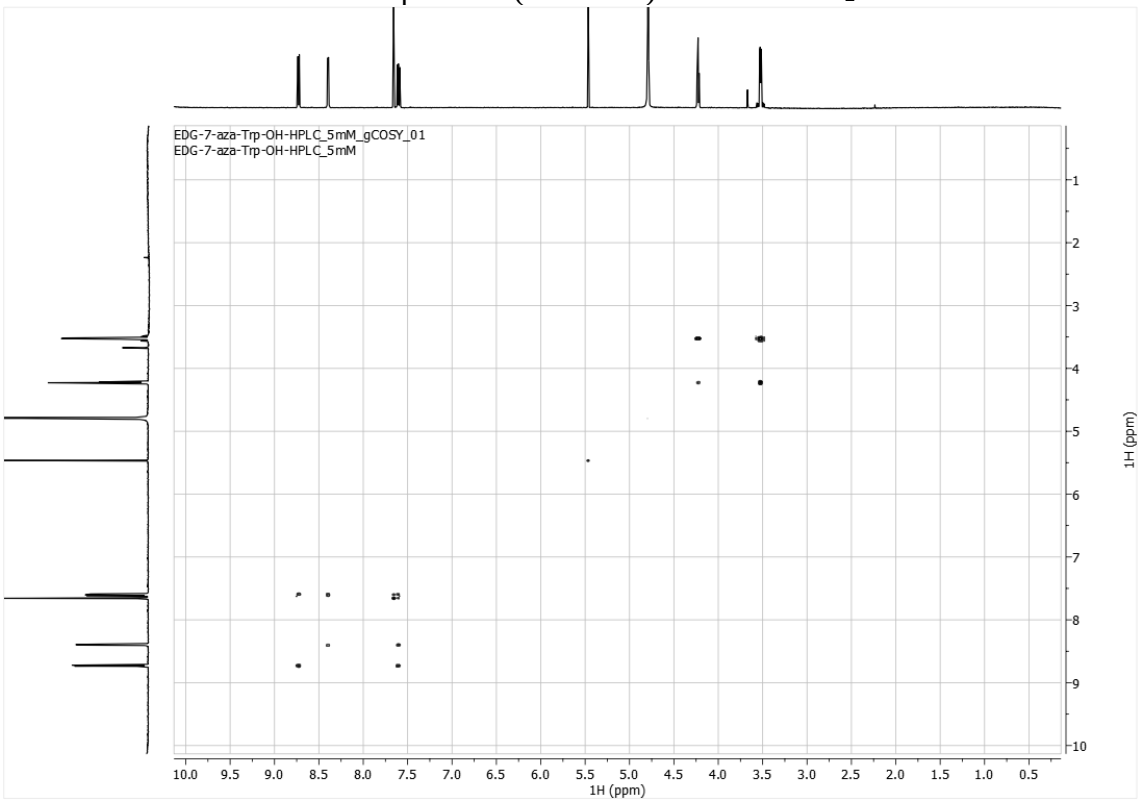
^{1}H-COSY spectrum (500 MHz) of **6A-W** in D_2O



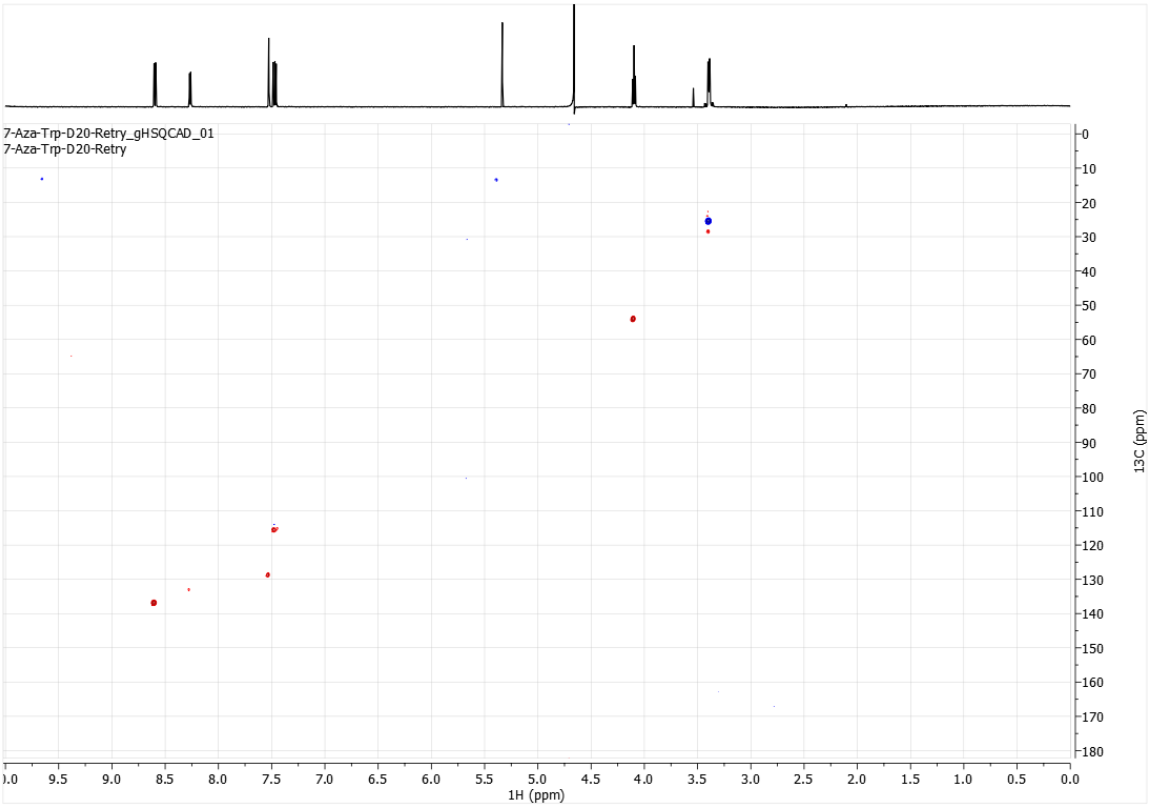
HSQC spectrum (500 MHz) of **6A-W** in D_2O



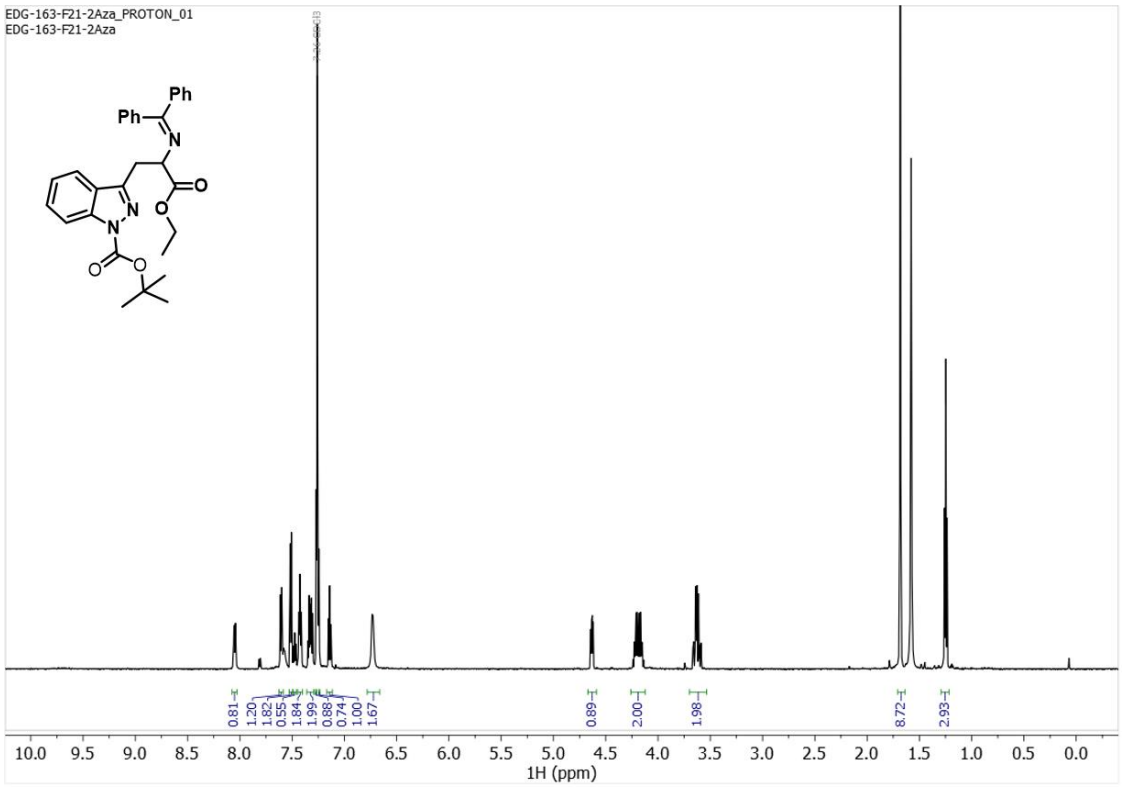
^{1}H NMR spectrum (500 MHz) of **7A-W** in D_2O



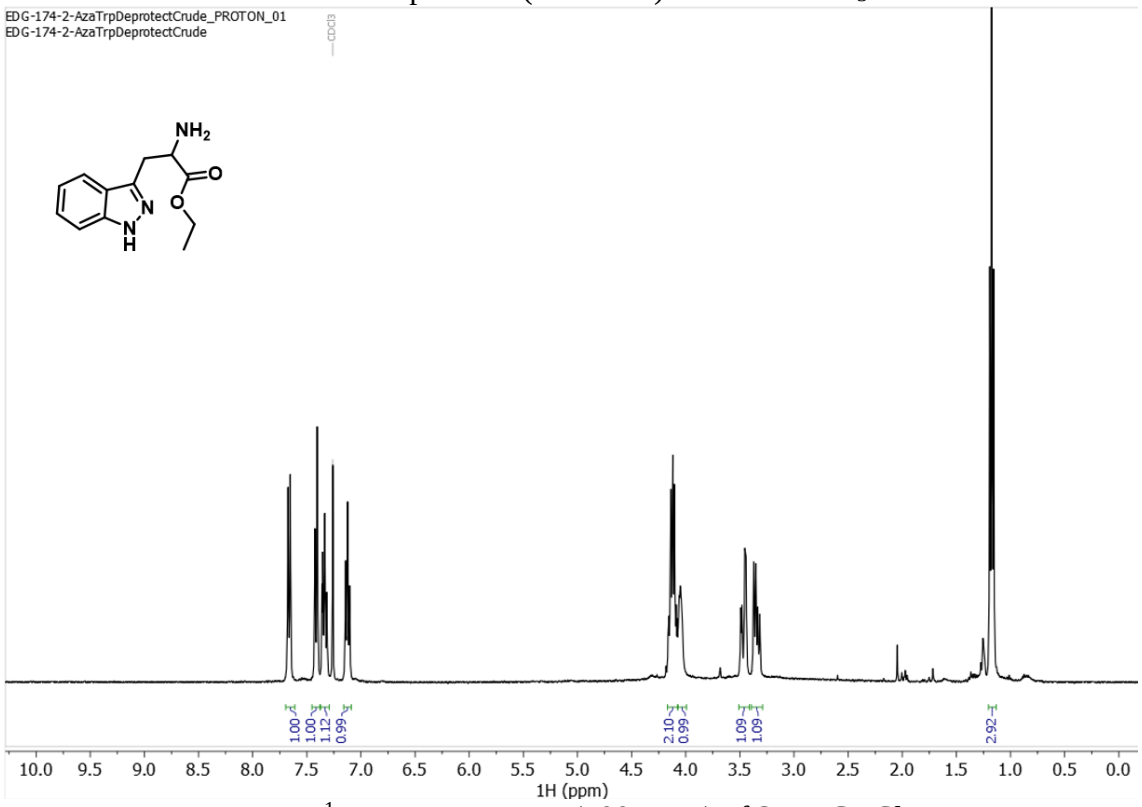
^{1}H-COSY spectrum (500 MHz) of **7A-W** in D_2O



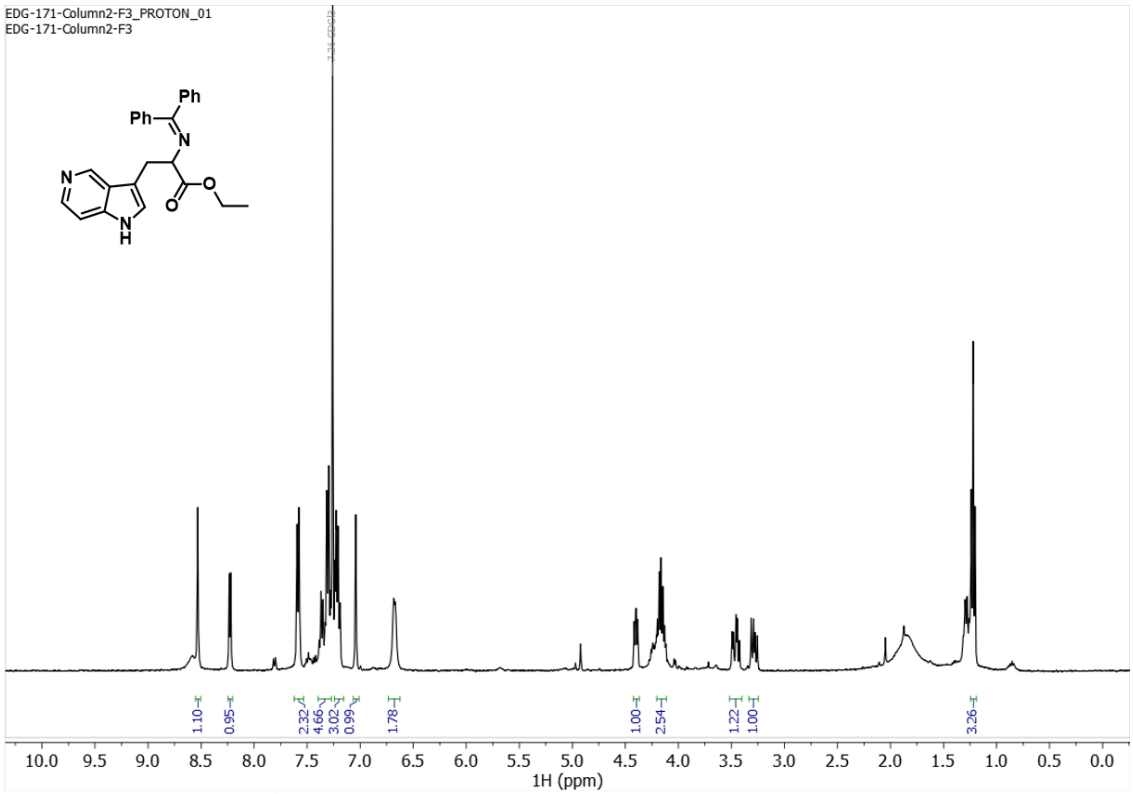
HSQC spectrum (500 MHz) of **7A-W** in D_2O



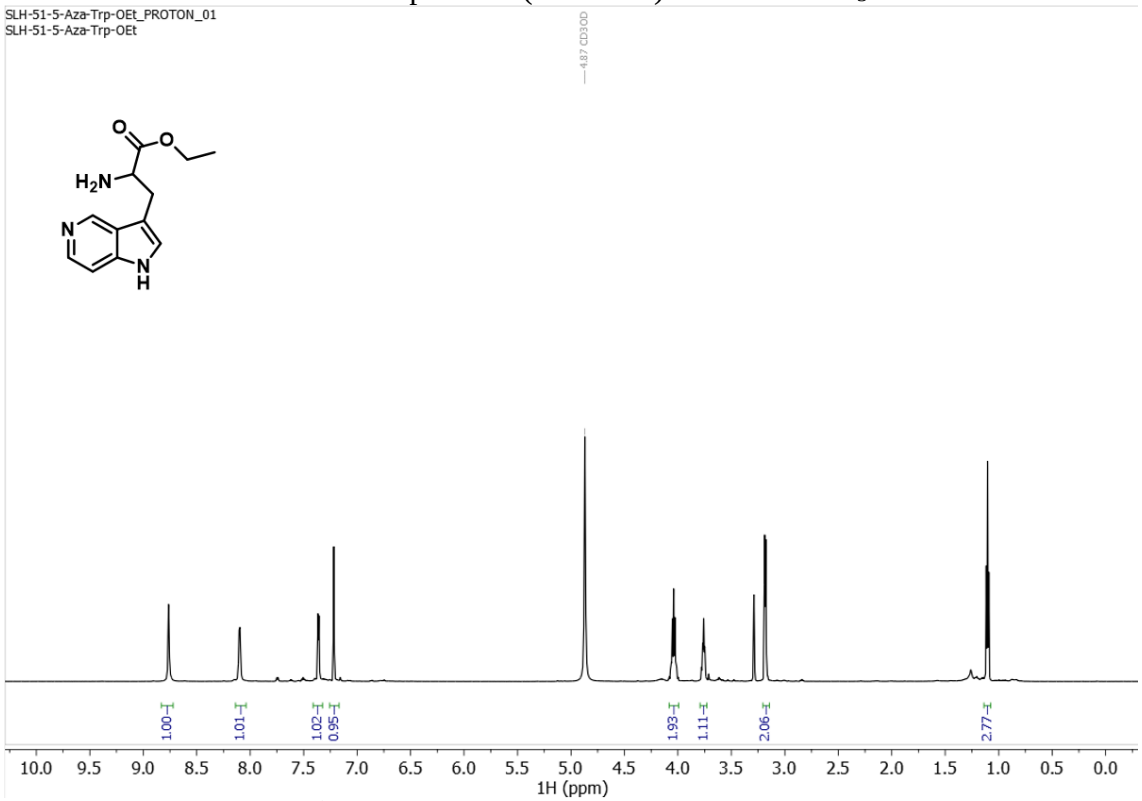
^{1}H NMR spectrum (400 MHz) of **2d** in $CDCl_3$



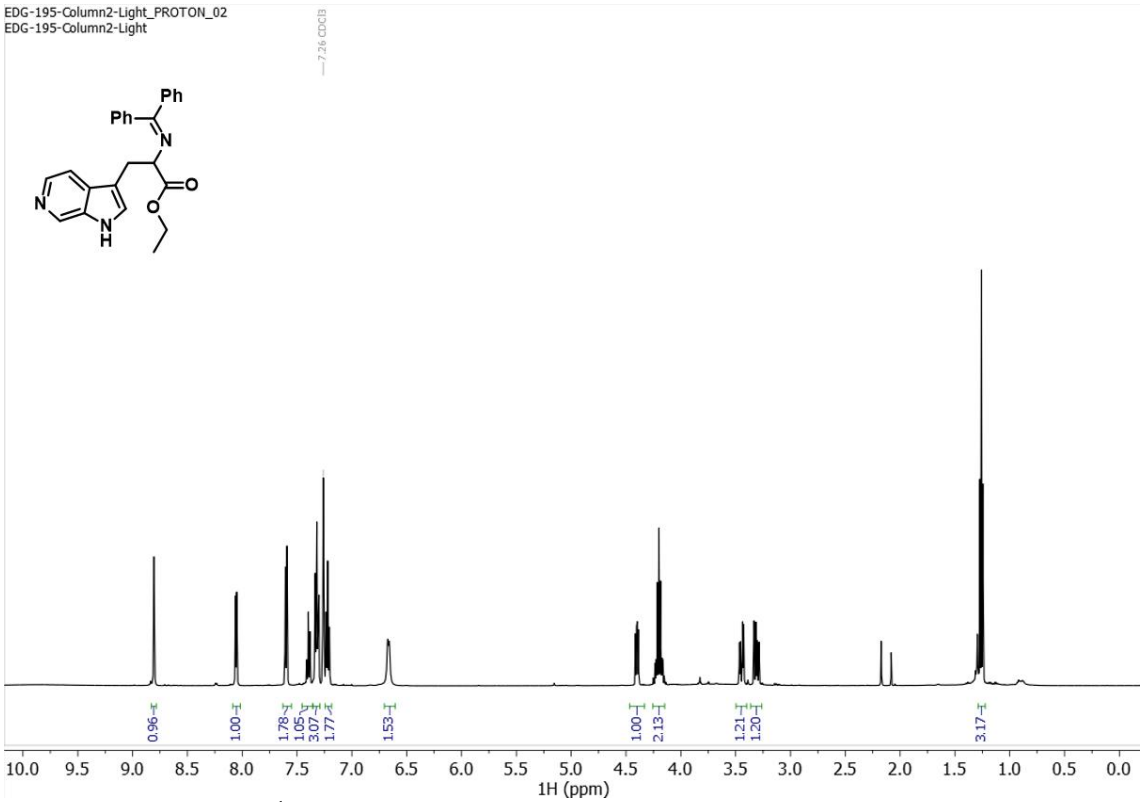
^{1}H NMR spectrum (400 MHz) of **2e** in $CDCl_3$



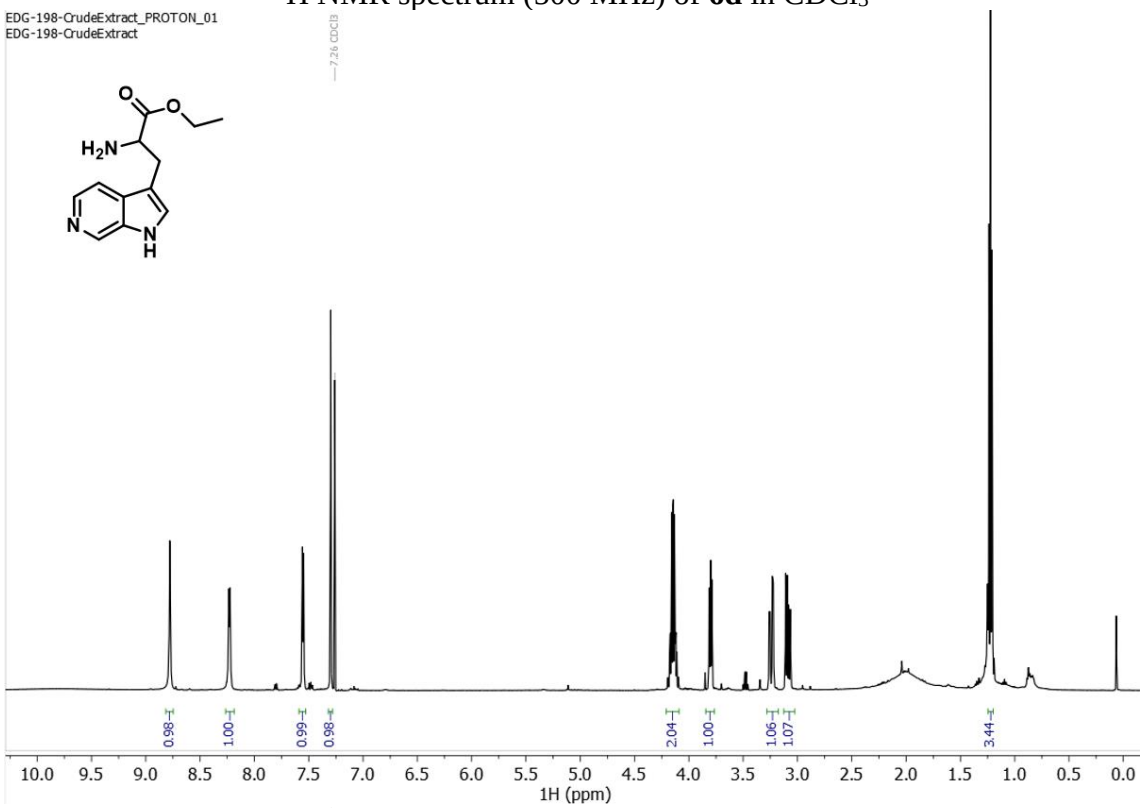
^{1}H NMR spectrum (400 MHz) of **5d** in $CDCl_3$



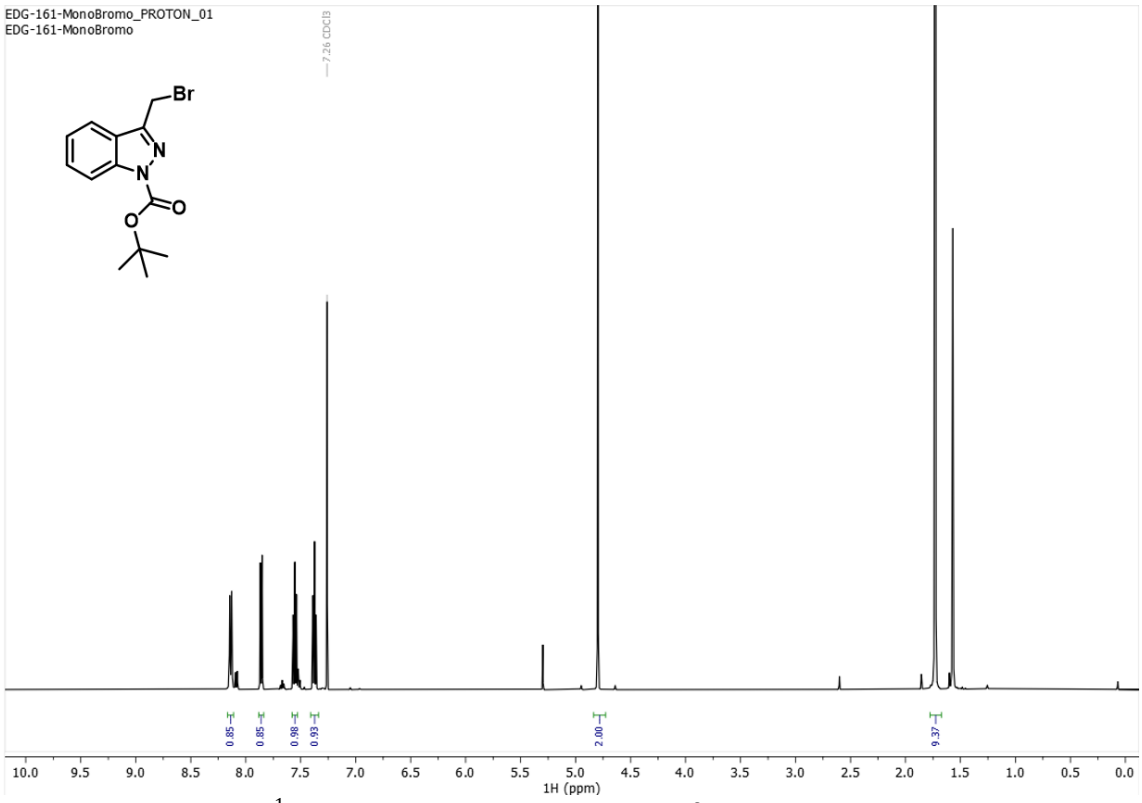
^{1}H NMR spectrum (600 MHz) of **5e** in CD_3OD



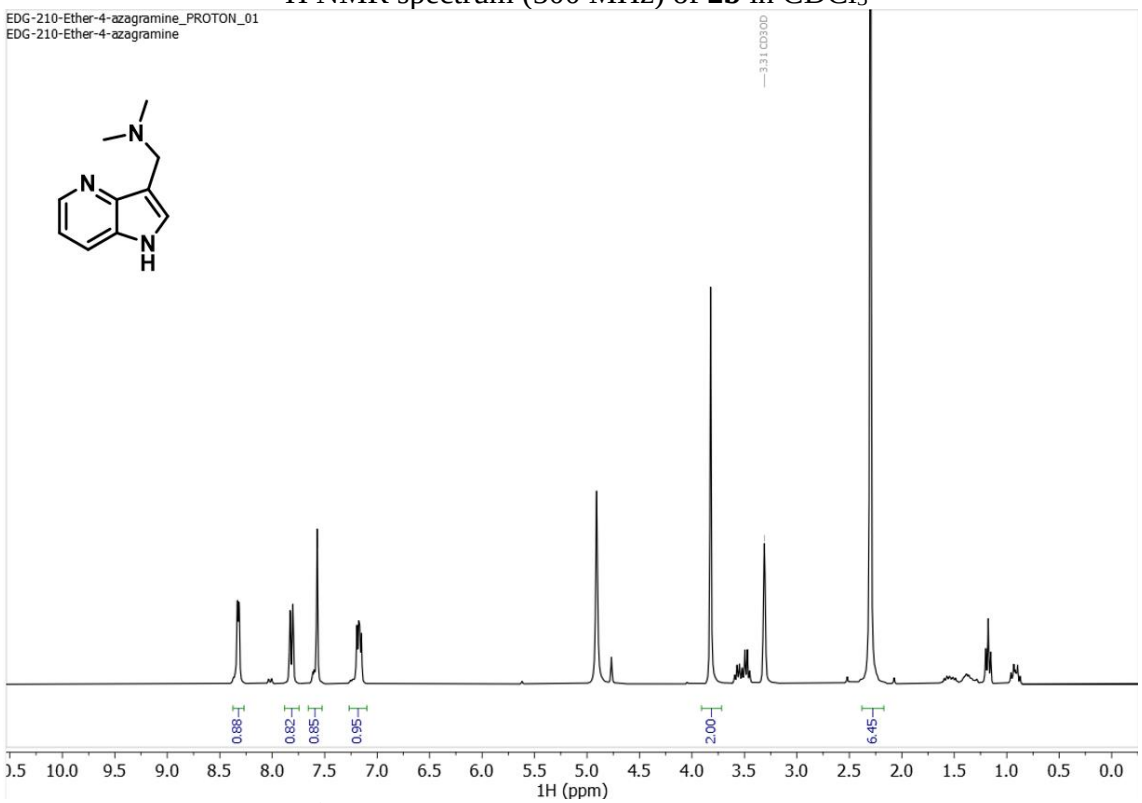
^{1}H NMR spectrum (500 MHz) of **6d** in $CDCl_3$



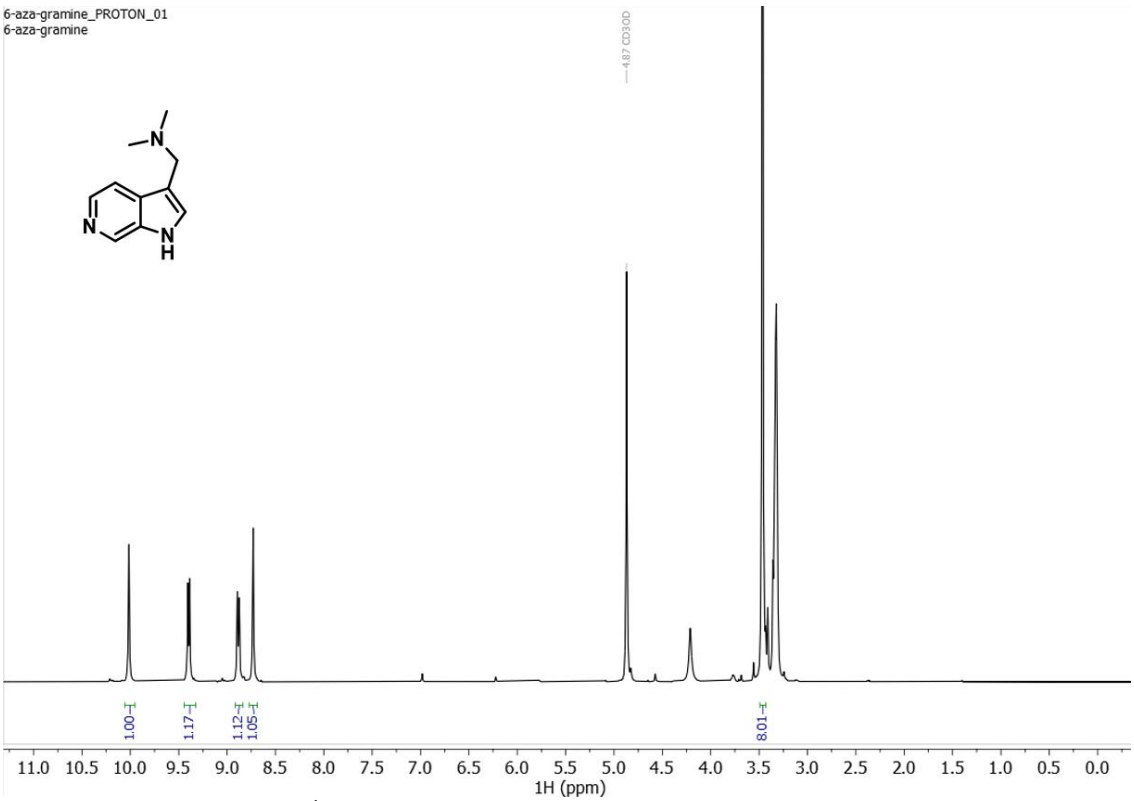
^{1}H NMR spectrum (500 MHz) of **6e** in $CDCl_3$



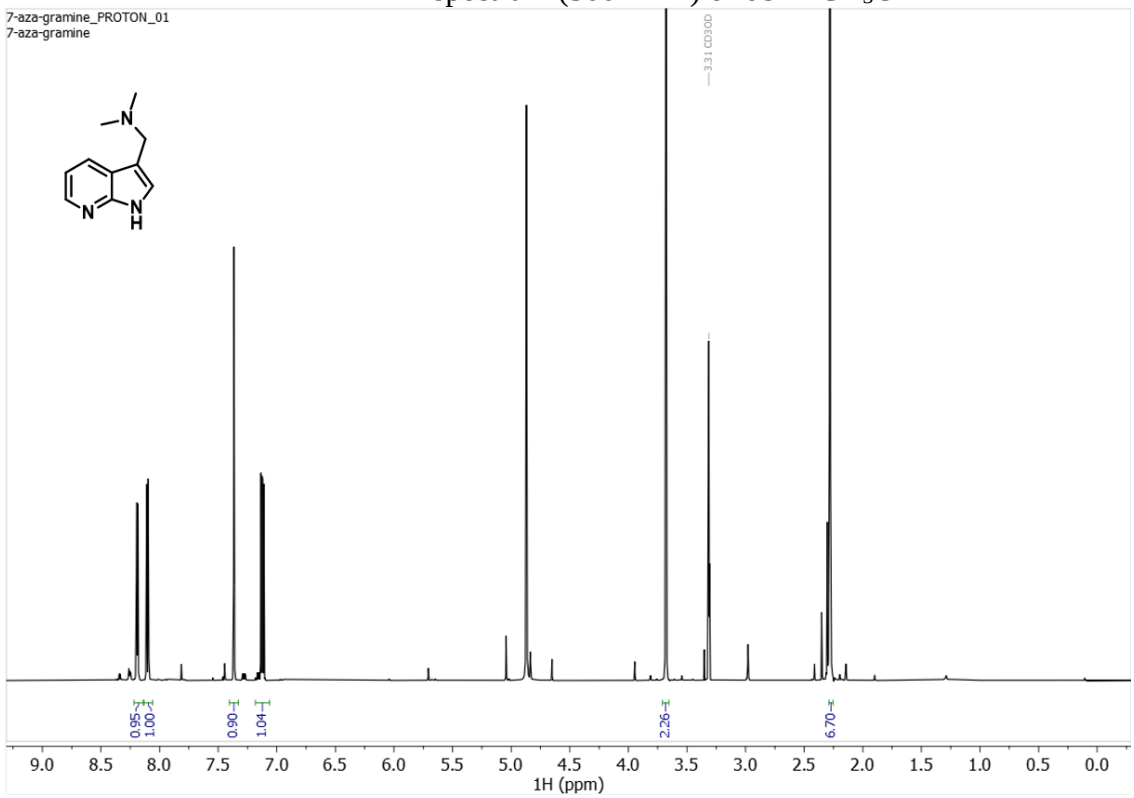
^{1}H NMR spectrum (500 MHz) of **2b** in $CDCl_3$



^{1}H NMR spectrum (300 MHz) of **4b** in CD_3OD

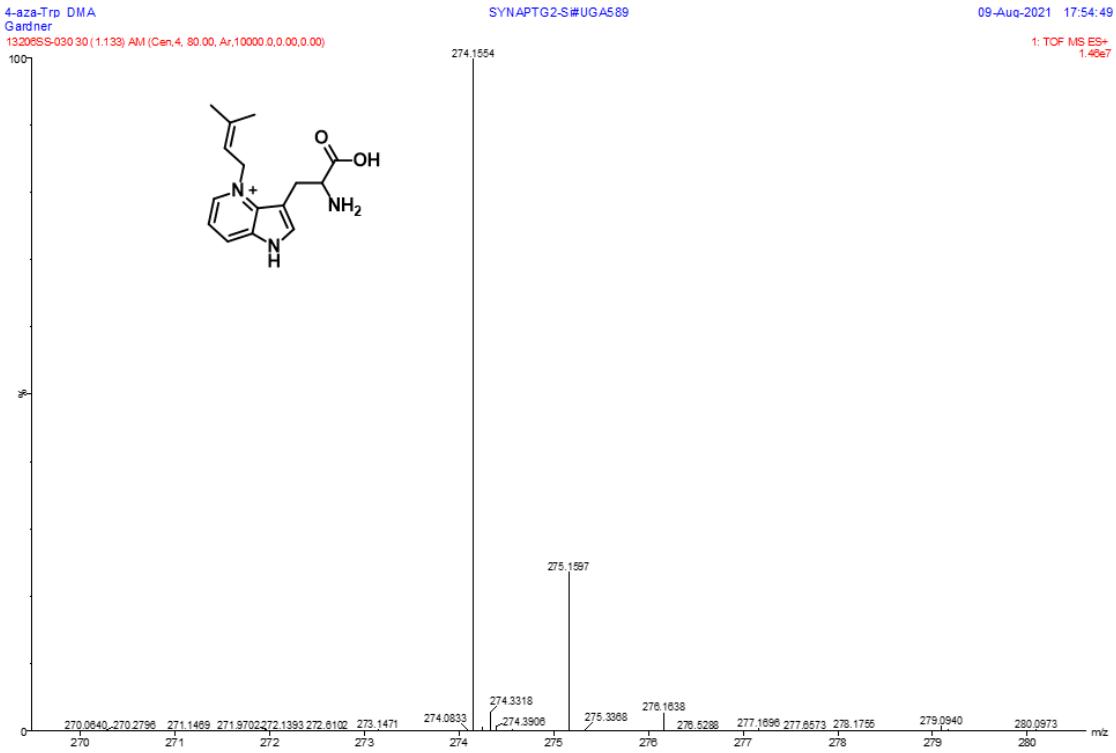


^{1}H NMR spectrum (300 MHz) of **6b** in CD_3OD

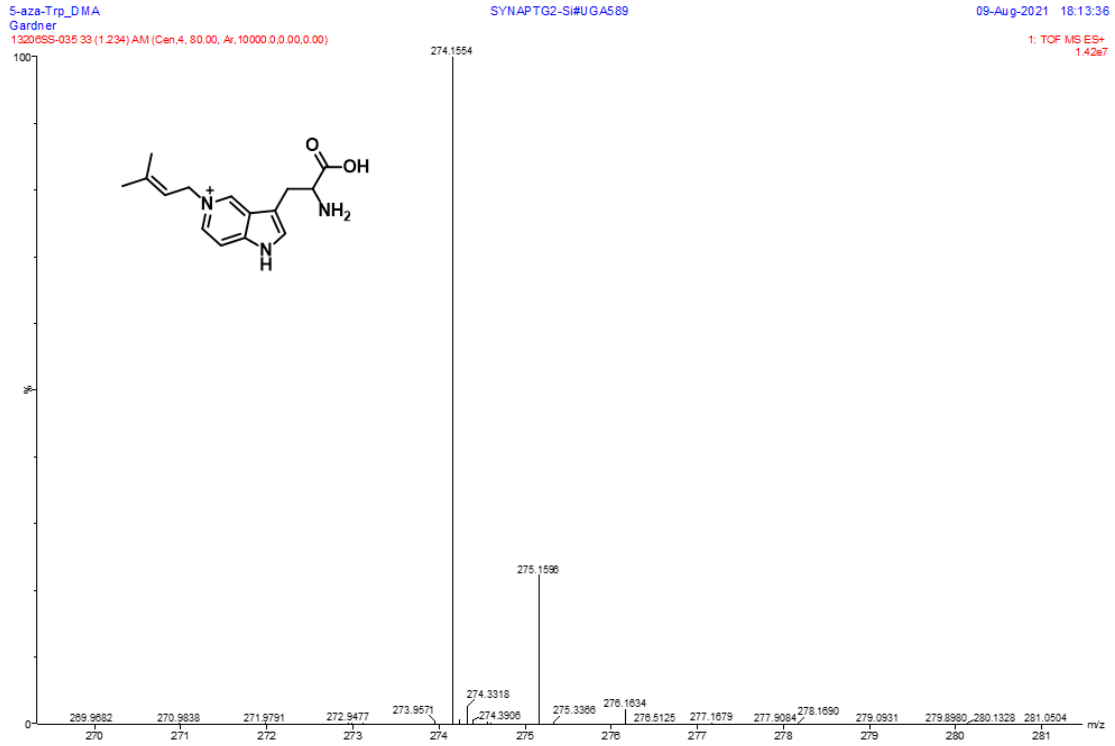


^{1}H NMR spectrum (500 MHz) of 7**b** in CD_3OD

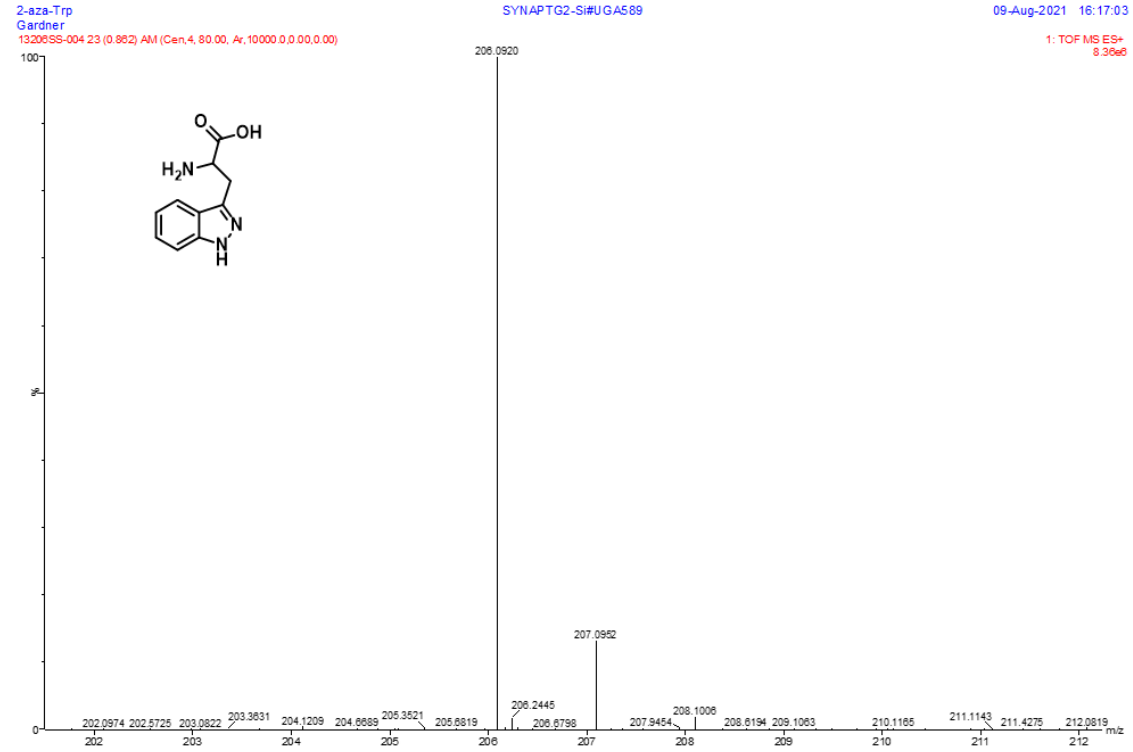
3.6 Appendix 2: HRMS Data



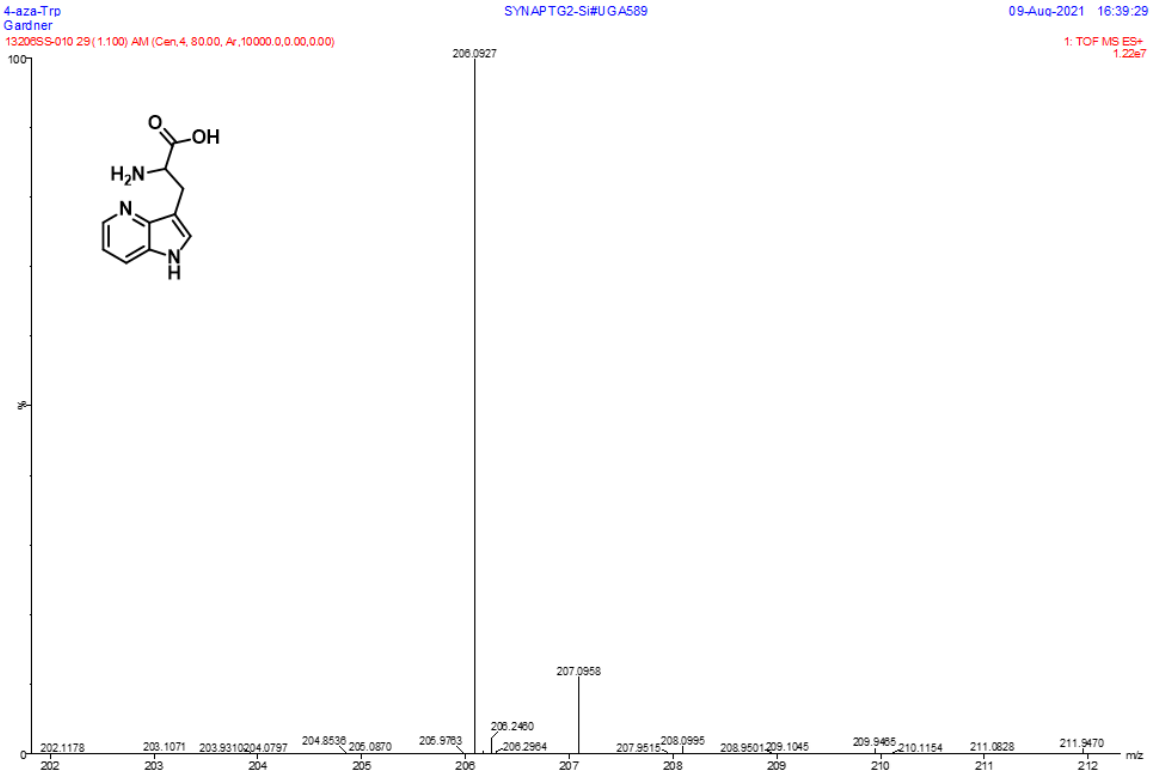
HRMS data for **4A-W-DMA**



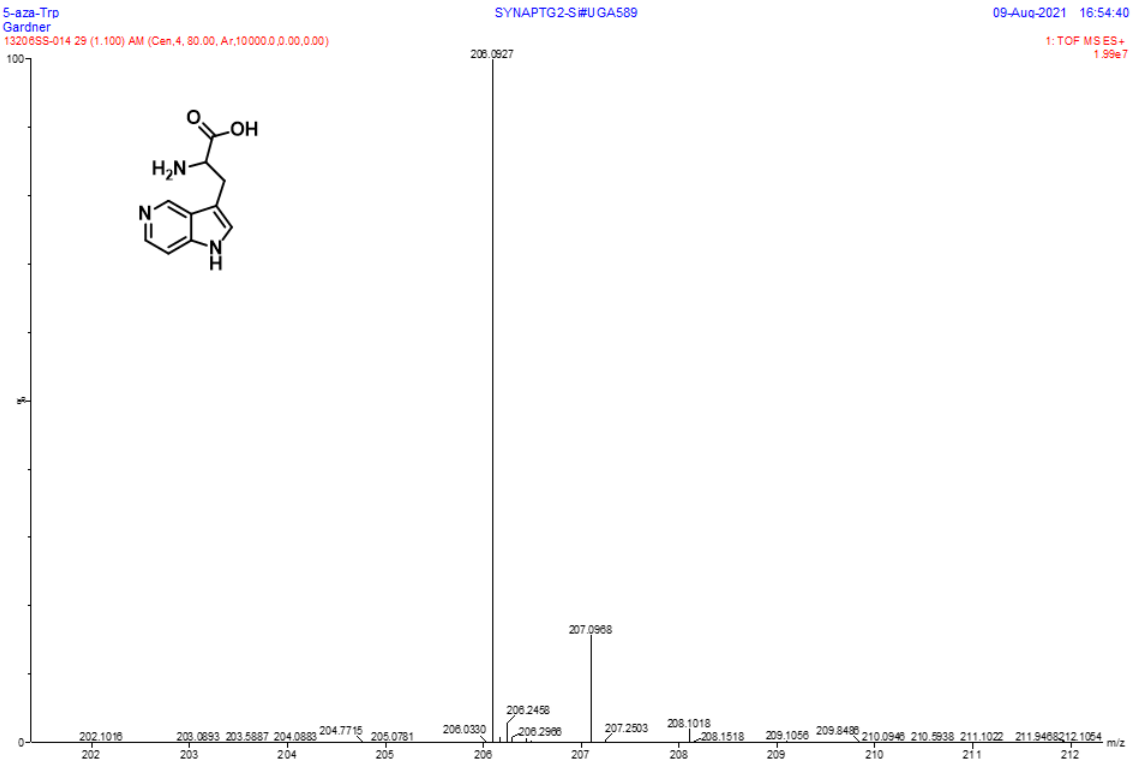
HRMS data for **5A-W-DMA**



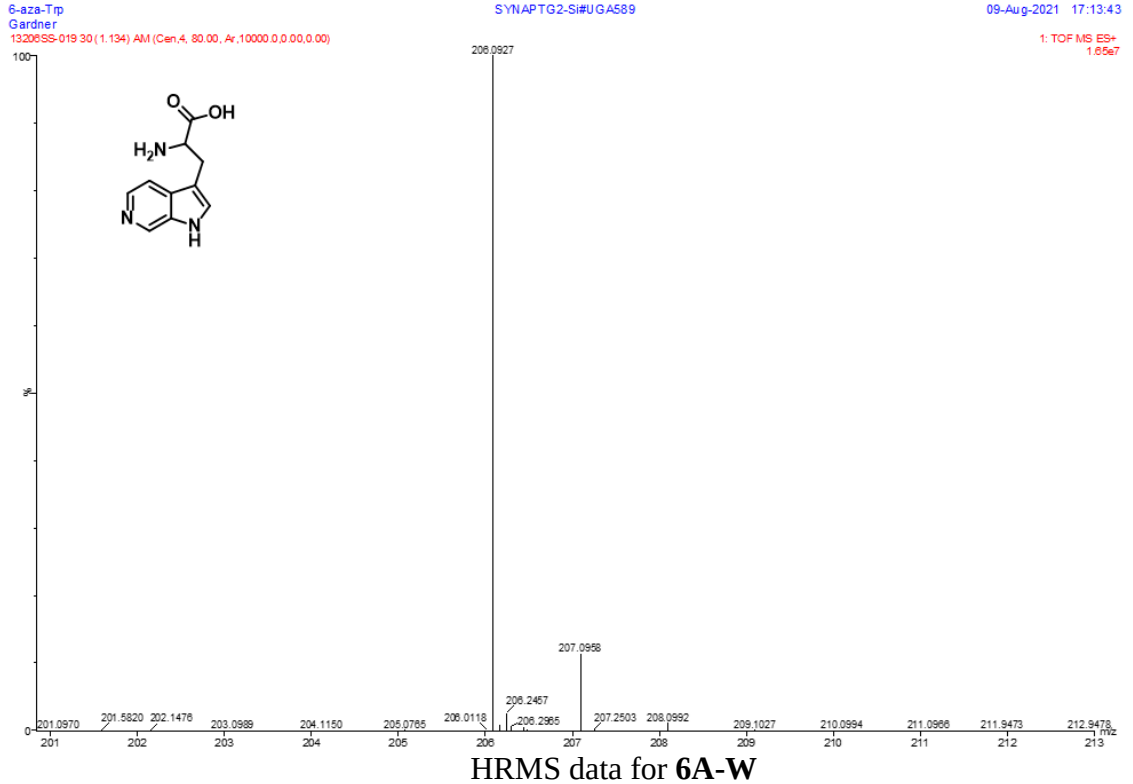
HRMS data for **2A-W**



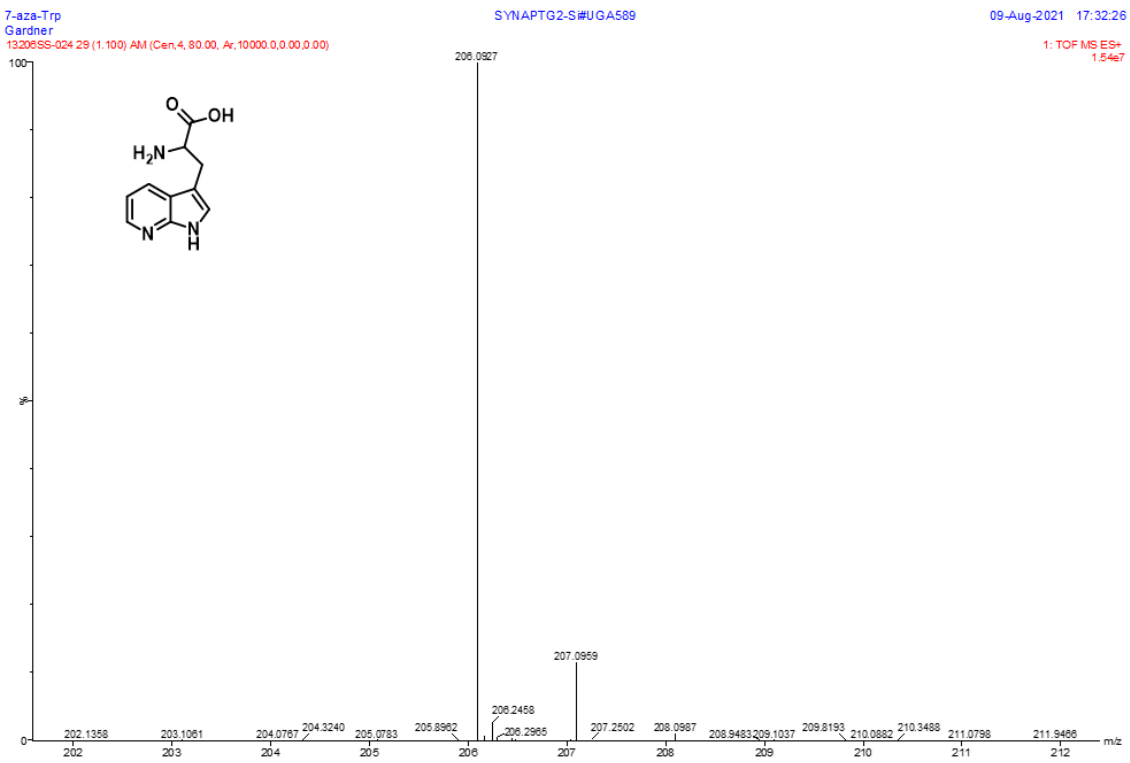
HRMS data for **4A-W**



HRMS data for **5A-W**



HRMS data for **6A-W**



HRMS data for **7A-W**

3.7 References for Chapter 3

1. Mérour, J.-Y.; Buron, F.; Plé, K.; Bonnet, P.; Routier, S., The Azaindole Framework in the Design of Kinase Inhibitors. *Molecules* **2014,** *19* (12), 19935-19979.

2. Motati, D. R.; Amaradhi, R.; Ganesh, T., Azaindole therapeutic agents. *Bioorganic & Medicinal Chemistry* **2020,** *28* (24), 115830.

3. Pennington, L. D.; Moustakas, D. T., The Necessary Nitrogen Atom: A Versatile High-Impact Design Element for Multiparameter Optimization. *Journal of Medicinal Chemistry* **2017,** *60* (9), 3552-3579.

4. Štarha, P.; Vančo, J.; Trávníček, Z.; Hošek, J.; Klusáková, J.; Dvořák, Z., Platinum(II) Iodido Complexes of 7-Azaindoles with Significant Antiproliferative Effects: An Old Story Revisited with Unexpected Outcomes. *PLOS ONE* **2016,** *11* (12), e0165062.

5. Zhao, S.-B.; Wang, S., Luminescence and reactivity of 7-azaindole derivatives and complexes. *Chemical Society Reviews* **2010,** *39* (8), 3142.

6. Pati, B. V.; Sagara, P. S.; Ghosh, A.; Mohanty, S. R.; Ravikumar, P. C., Ruthenium-Catalyzed Cross Dehydrogenative Annulation of N-(7-Azaindole)benzamides with Maleimides: One-Step Access to Highly Functionalized Pyrroloisoquinoline. *The Journal of Organic Chemistry* **2021,** *86* (9), 6551-6565.

7. Tsoureas, N.; Nunn, J.; Bevis, T.; Haddow, M. F.; Hamilton, A.; Owen, G. R., Strong agostic-type interactions in ruthenium benzylidene complexes containing 7-azaindole based scorpionate ligands. *Dalton Trans.* **2011,** *40* (4), 951-958.

8. Zhao, S.-B.; Mcormick, T.; Wang, S., Ambient-Temperature Metal-to-Ligand Charge-Transfer Phosphorescence Facilitated by Triarylboron: Bnpa and Its Metal Complexes. *Inorganic Chemistry* **2007,** *46* (26), 10965-10967.

9. Martin, C.; Borreguero, C.; Kennes, K.; Van Der Auweraer, M.; Hofkens, J.; De Miguel, G.; García-Frutos, E. M., Simple Donor–Acceptor Luminogen Based on an Azaindole Derivative as Solid-State Emitter for Organic Light-Emitting Devices. *ACS Energy Letters* **2017,** *2* (12), 2653-2658.

10. Lee, J.; Liu, Q.-D.; Bai, D.-R.; Kang, Y.; Tao, Y.; Wang, S., 2,3,4,5-Tetrafunctionalized Siloles: Syntheses, Structures, Luminescence, and Electroluminescence. *Organometallics* **2004,** *23* (26), 6205-6213.

11. Lapcinska, S.; Dimitrijevs, P.; Lapcinskis, L.; Arsenyan, P., Visible Light-Mediated Functionalization of Selenocystine-Containing Peptides. *Advanced Synthesis & Catalysis* **2021,** *363* (13), 3318-3328.

12. De Filippis, V.; De Boni, S.; De Dea, E.; Dalzoppo, D.; Grandi, C.; Fontana, A., Incorporation of the fluorescent amino acid 7-azatryptophan into the core domain 1-47 of hirudin as a probe of hirudin folding and thrombin recognition. *Protein Science* **2004,** *13* (6), 1489-1502.

13. Shen, J.-Y.; Chao, W.-C.; Liu, C.; Pan, H.-A.; Yang, H.-C.; Chen, C.-L.; Lan, Y.-K.; Lin, L.-J.; Wang, J.-S.; Lu, J.-F.; Chun-Wei Chou, S.; Tang, K.-C.; Chou, P.-T., Probing water micro-solvation in proteins by water catalysed proton-transfer tautomerism. *Nature Communications* **2013,** *4* (1).

14. Lepthien, S.; Hoesl, M. G.; Merkel, L.; Budisa, N., Azatryptophans endow proteins with intrinsic blue fluorescence. *Proceedings of the National Academy of Sciences* **2008,** *105* (42), 16095-16100.

15. Kwon, O.-H.; Zewail, A. H., Double proton transfer dynamics of model DNA base pairs in the condensed phase. *Proceedings of the National Academy of Sciences* **2007,** *104* (21), 8703-8708.

16. Kaur, K.; Chaudhary, S.; Singh, S.; Mehta, S. K., An azaindole–hydrazine imine moiety as sensitive dual cation chemosensor depending on surface plasmon resonance and emission properties. *Sensors and Actuators B: Chemical* **2016,** *222*, 397-406.

17. Alexander Ross, J. B.; Szabo, A. G.; Hogue, C. W. V., [8] Enhancement of protein spectra with tryptophan analogs: Fluorescence spectroscopy of protein-protein and protein-nucleic acid interactions. In *Methods in Enzymology*, Elsevier: 1997; pp 151-190.

18. Popowycz, F.; Routier, S.; Joseph, B.; Mérour, J.-Y., Synthesis and reactivity of 7-azaindole (1H-pyrrolo[2,3-b]pyridine). *Tetrahedron* **2007,** *63* (5), 1031-1064.

19. Broos, J.; Ter Veld, F.; Robillard, G. T., Membrane Protein−Ligand Interactions inEscherichia coliVesicles and Living Cells Monitored via a Biosynthetically Incorporated Tryptophan Analogue†. *Biochemistry* **1999,** *38* (31), 9798-9803.

20. Irie, T.; Sawa, M., 7-Azaindole: A Versatile Scaffold for Developing Kinase Inhibitors. *Chemical and Pharmaceutical Bulletin* **2018,** *66* (1), 29-36.

21. Han, Y.; Dong, W.; Guo, Q.; Li, X.; Huang, L., The importance of indole and azaindole scaffold in the development of antitumor agents. *European Journal of Medicinal Chemistry* **2020,** *203*, 112506.

22. Prudent, R.; Vassal-Stermann, É.; Nguyen, C.-H.; Mollaret, M.; Viallet, J.; Desroches-Castan, A.; Martinez, A.; Barette, C.; Pillet, C.; Valdameri, G.; Soleilhac, E.; Di Pietro, A.; Feige, J.-J.; Billaud, M.; Florent, J.-C.; Lafanechère, L., Azaindole derivatives are inhibitors of microtubule dynamics, with anti-cancer and anti-angiogenic activities. *British Journal of Pharmacology* **2013,** *168* (3), 673-685.

23. Chen, K.; Risatti, C.; Bultman, M.; Soumeillant, M.; Simpson, J.; Zheng, B.; Fanfair, D.; Mahoney, M.; Mudryk, B.; Fox, R. J.; Hsaio, Y.; Murugesan, S.; Conlon, D. A.; Buono, F. G.; Eastgate, M. D., Synthesis of the 6-Azaindole Containing HIV-1 Attachment Inhibitor Pro-Drug, BMS-663068. *The Journal of Organic Chemistry* **2014,** *79* (18), 8757-8767.

24. Chatterji, M.; Shandil, R.; Manjunatha, M. R.; Solapure, S.; Ramachandran, V.; Kumar, N.; Saralaya, R.; Panduga, V.; Reddy, J.; Kr, P.; Sharma, S.; Sadler, C.; Cooper, C. B.; Mdluli, K.; Iyer, P. S.; Narayanan, S.; Shirude, P. S., 1,4-Azaindole, a Potential Drug Candidate for Treatment of Tuberculosis. *Antimicrobial Agents and Chemotherapy* **2014,** *58* (9), 5325-5331.

25. Shirude, P. S.; Shandil, R.; Sadler, C.; Naik, M.; Hosagrahara, V.; Hameed, S.; Shinde, V.; Bathula, C.; Humnabadkar, V.; Kumar, N.; Reddy, J.; Panduga, V.; Sharma, S.; Ambady, A.; Hegde, N.; Whiteaker, J.; Mclaughlin, R. E.; Gardner, H.; Madhavapeddi, P.; Ramachandran, V.; Kaur, P.; Narayan, A.; Guptha, S.; Awasthy, D.; Narayan, C.; Mahadevaswamy, J.; Vishwas, K.; Ahuja, V.; Srivastava, A.; Prabhakar, K.; Bharath, S.; Kale, R.; Ramaiah, M.; Choudhury, N. R.; Sambandamurthy, V. K.; Solapure, S.; Iyer, P. S.; Narayanan, S.; Chatterji, M., Azaindoles: Noncovalent DprE1

Inhibitors from Scaffold Morphing Efforts, Kill Mycobacterium tuberculosis and Are Efficaciousin Vivo. *Journal of Medicinal Chemistry* **2013,** *56* (23), 9701-9708.

26. Zhang, J.; Yang, Q.; Cross, J. B.; Romero, J. A. C.; Poutsiaka, K. M.; Epie, F.; Bevan, D.; Wang, B.; Zhang, Y.; Chavan, A.; Zhang, X.; Moy, T.; Daniel, A.; Nguyen, K.; Chamberlain, B.; Carter, N.; Shotwell, J.; Silverman, J.; Metcalf, C. A.; Ryan, D.; Lippa, B.; Dolle, R. E., Discovery of Azaindole Ureas as a Novel Class of Bacterial Gyrase B Inhibitors. *Journal of Medicinal Chemistry* **2015,** *58* (21), 8503-8512.

27. Liu, Y. A.; Jin, Q.; Zou, Y.; Ding, Q.; Yan, S.; Wang, Z.; Hao, X.; Nguyen, B.; Zhang, X.; Pan, J.; Mo, T.; Jacobsen, K.; Lam, T.; Wu, T. Y.-H.; Petrassi, H. M.; Bursulaya, B.; Didonato, M.; Gordon, W. P.; Liu, B.; Baaten, J.; Hill, R.; Nguyen-Tran, V.; Qiu, M.; Zhang, Y.-Q.; Kamireddy, A.; Espinola, S.; Deaton, L.; Ha, S.; Harb, G.; Jia, Y.; Li, J.; Shen, W.; Schumacher, A. M.; Colman, K.; Glynne, R.; Pan, S.; Mcnamara, P.; Laffitte, B.; Meeusen, S.; Molteni, V.; Loren, J., Selective DYRK1A Inhibitor for the Treatment of Type 1 Diabetes: Discovery of 6-Azaindole Derivative GNF2133. *Journal of Medicinal Chemistry* **2020,** *63* (6), 2958-2973.

28. Immadi, S. S.; Dopart, R.; Wu, Z.; Fu, B.; Kendall, D. A.; Lu, D., Exploring 6-Azaindole and 7-Azaindole Rings for Developing Cannabinoid Receptor 1 Allosteric Modulators. *Cannabis and Cannabinoid Research* **2018,** *3* (1), 252-258.

29. Marminon, C.; Pierré, A.; Pfeiffer, B.; Pérez, V.; Léonce, S.; Joubert, A.; Bailly, C.; Renard, P.; Hickman, J.; Prudhomme, M., Syntheses and Antiproliferative Activities of 7-Azarebeccamycin Analogues Bearing One 7-Azaindole Moiety. *Journal of Medicinal Chemistry* **2003,** *46* (4), 609-622.

30. Phillips, R. S., Synthetic applications of tryptophan synthase. *Tetrahedron: Asymmetry* **2004,** *15* (18), 2787-2792.

31. Romney, D. K.; Murciano-Calles, J.; Wehrmüller, J. E.; Arnold, F. H., Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogues. *Journal of the American Chemical Society* **2017,** *139* (31), 10769-10776.

32. Sloan, M. J.; Phillips, R. S., Enzymatic synthesis of aza-l-tryptophans: The preparation of 5- and 6-Aza-l-tryptophan. *Bioorganic & Medicinal Chemistry Letters* **1992,** *2* (9), 1053-1056.

33. Gkotsi, D. S.; Ludewig, H.; Sharma, S. V.; Connolly, J. A.; Dhaliwal, J.; Wang, Y.; Unsworth, W. P.; Taylor, R. J. K.; Mclachlan, M. M. W.; Shanahan, S.; Naismith, J. H.; Goss, R. J. M., A marine viral halogenase that iodinates diverse substrates. *Nature Chemistry* **2019,** *11* (12), 1091-1097.

34. Lee, H.-Y.; Yerkes, N.; O'Connor, S. E., Aza-Tryptamine Substrates in Monoterpene Indole Alkaloid Biosynthesis. *Chemistry & Biology* **2009,** *16* (12), 1225-1229.

35. Johnson, B. P.; Scull, E. M.; Dimas, D. A.; Bavineni, T.; Bandari, C.; Batchev, A. L.; Gardner, E. D.; Nimmo, S. L.; Singh, S., Acceptor substrate determines donor specificity of an aromatic prenyltransferase: expanding the biocatalytic potential of NphB. *Applied Microbiology and Biotechnology* **2020**.

36. Scull, E. M.; Bandari, C.; Johnson, B. P.; Gardner, E. D.; Tonelli, M.; You, J.; Cichewicz, R. H.; Singh, S., Chemoenzymatic synthesis of daptomycin analogs active against daptomycin-resistant strains. *Applied Microbiology and Biotechnology* **2020**.

37. Bandari, C.; Scull, E. M.; Masterson, J. M.; Tran, R. H. Q.; Foster, S. B.; Nicholas, K. M.; Singh, S., Determination of Alkyl-Donor Promiscuity of Tyrosine-O-Prenyltransferase SirD from Leptosphaeria maculans. *ChemBioChem* **2017,** *18* (23), 2323-2327.

38. Bandari, C.; Scull, E. M.; Bavineni, T.; Nimmo, S. L.; Gardner, E. D.; Bensen, R. C.; Burgett, A. W.; Singh, S., FgaPT2, a biocatalytic tool for alkyl-diversification of indole natural products. *MedChemComm* **2019**.

39. Gardner, E. D.; Dimas, D. A.; Finneran, M. C.; Brown, S. M.; Burgett, A. W.; Singh, S., Indole C6 Functionalization of Tryprostatin B Using Prenyltransferase CdpNPT. *Catalysts* **2020,** *10* (11), 1247.

40. Elshahawi, S. I.; Cao, H.; Shaaban, K. A.; Ponomareva, L. V.; Subramanian, T.; Farman, M. L.; Spielmann, H. P.; Phillips Jr, G. N.; Thorson, J. S.; Singh, S., Structure and specificity of a permissive bacterial C-prenyltransferase. *Nat Chem Biol* **2017,** *13* (4), 366-368.

41. Pan, L. L.; Song, L. F.; Miao, Y.; Yang, Y.; Merz, K. M., Jr., Mechanism of Formation of the Nonstandard Product in the Prenyltransferase Reaction of the G115T Mutant of FtmPT1: A Case of Reaction Dynamics Calling the Shots? *Biochemistry* **2017,** *56* (24), 2995-3007.

42. Liebhold, M.; Xian, M.; Li, S. M., Breaking Cyclic Dipeptide Prenyltransferase Regioselectivity by Unnatural Alkyl Doners. *Organic letters* **2013,** *15* (12), 3062-3065.

43. Liebhold, M.; Xie, X.; Li, S. M., Expansion of enzymatic Friedel-Crafts alkylation on indoles: acceptance of unnatural beta-unsaturated allyl diphospates by dimethylallyl-tryptophan synthases. *Organic letters* **2012,** *14* (18), 4882-5.

44. Winkelblech, J.; Liebhold, M.; Gunera, J.; Xie, X.; Kolb, P.; Li, S. M., Tryptophan C5-, C6- and C7-Prenylating Enzymes Displaying a Preference for C-6 of the Indole Ring in the Presence of Unnatural Dimethylallyl Diphosphate Analogues. *Advanced Synthesis & Catalysis* **2015,** *357* (5), 975-986.

45. Truppo, M. D., Biocatalysis in the Pharmaceutical Industry: The Need for Speed. *ACS Med Chem Lett* **2017,** *8* (5), 476-480.

46. Clouthier, C. M.; Pelletier, J. N., Expanding the organic toolbox: a guide to integrating biocatalysis in synthesis. *Chemical Society Reviews* **2012,** *41* (4), 1585-1605.

47. Sheldon, R. A.; Brady, D.; Bode, M. L., The Hitchhiker's guide to biocatalysis: recent advances in the use of enzymes in organic synthesis. *Chemical Science* **2020,** *11* (10), 2587-2605.

48. Tibrewal, N.; Tang, Y., Biocatalysts for natural product biosynthesis. *Annu Rev Chem Biomol Eng* **2014,** *5*, 347-66.

49. Baker Dockrey, S. A.; Lukowski, A. L.; Becker, M. R.; Narayan, A. R. H., Biocatalytic site- and enantioselective oxidative dearomatization of phenols. *Nat Chem* **2018,** *10* (2), 119-125.

50. Aleku, G. A.; Nowicka, B.; Turner, N. J., Biocatalytic Potential of Enzymes Involved in the Biosynthesis of Isoprenoid Quinones. *ChemCatChem* **2018,** *10* (1), 124-135.

51. Crestey, F.; Collot, V.; Stiebing, S.; Rault, S., A new and efficient synthesis of 2-azatryptophans. *Tetrahedron* **2006,** *62* (33), 7772-7775.

52. Füller, J. J.; Röpke, R.; Krausze, J.; Rennhack, K. E.; Daniel, N. P.; Blankenfeldt, W.; Schulz, S.; Jahn, D.; Moser, J., Biosynthesis of Violacein, Structure and Function of

l-Tryptophan Oxidase VioA from Chromobacterium violaceum. *Journal of Biological Chemistry* **2016,** *291* (38), 20068-20084.

53. Azimov, V. A.; Uritskaya, M. Y.; Yakhontov, L. N., Synthesis of racemic α-amino acids of the series of 4- and 7-azaindoles. *Pharmaceutical Chemistry Journal* **1968,** *2* (11), 605-607.

54. Yin, W. B.; Ruan, H. L.; Westrich, L.; Grundmann, A.; Li, S. M., CdpNPT, an N-prenyltransferase from Aspergillus fumigatus: overproduction, purification and biochemical characterisation. *Chembiochem* **2007,** *8* (10), 1154-61.

55. Unsold, I. A.; Li, S. M., Overproduction, purification and characterization of FgaPT2, a dimethylallyltryptophan synthase from Aspergillus fumigatus. *Microbiology* **2005,** *151* (Pt 5), 1499-505.

56. Grundmann, A.; Li, S. M., Overproduction, purification and characterization of FtmPT1, a brevianamide F prenyltransferase from Aspergillus fumigatus. *Microbiology* **2005,** *151* (Pt 7), 2199-207.

57. Adler, T. K.; Albert, A., The Biological and Physical Properties of the Azaindoles. *Journal of Medicinal Chemistry* **1963,** *6* (5), 480-483.

Chapter 4: Chemoenzymatic Diversification of Azaindole Substituted Diketopiperazine Natural Product Analogs

Abstract

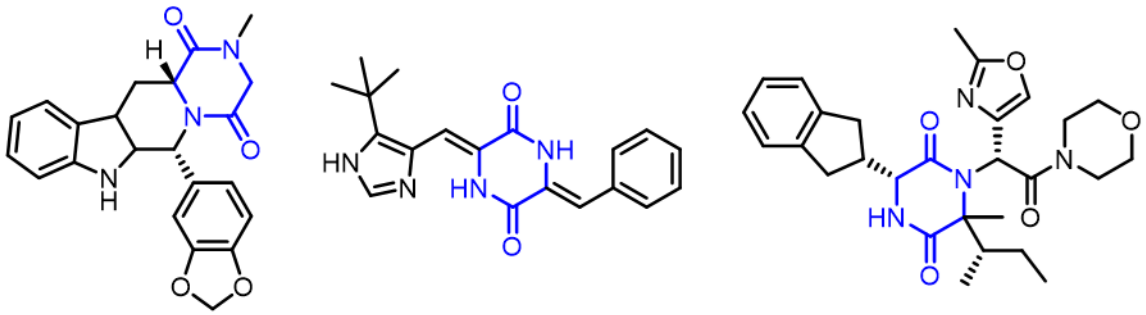
Prenylated indole containing diketopiperazines (DKP) are a large family of natural products with a wide spectrum of bioactivity. The therapeutic utility and structure activity relationship of these natural products remains largely unexplored due to limitations in synthetic methodology. In this study, unnatural azaindole containing DKP's were synthesized and chemoenzymatically alkylated using the indole C2 prenyltransferase FtmPT1. While the substrate scope for aromatic prenyltransferases is typically limited to indole or phenolic substrates, we observed chemoenzymatic prenylation of five novel azaindole substrate analogs. The additional aromatic nitrogen atoms altered the regioselectivity of the enzyme, producing novel N1 prenylated and cyclized C3 prenylated azaindole-containing natural product analogs. The discovery that nitrogenous indole bioisosteres can be accepted by aromatic prenyltransferases provides access to previously inaccessible bioactive natural product analogs.

Allocation of Contribution

I conceptualized this project and produced the results presented in this chapter with the following exceptions. Mr. Dustin Dimas synthesized and purified **7A-CyWP** on a gram scale which was subsequently used for enzymatic scale-up reactions. The contents and results described in this chapter have not yet been published.

4.1 Introduction

A majority of FDA approved drugs are natural products (NPs) or derived thereof,[1] of which 59% contain a nitrogenous heterocycle.[2] Prominent among the heterocyclic NPs are tryptophan-proline cyclic dipeptides (CyWP), a sub-set of compounds within the indolediketopiperazines (IDP) class, which have shown to display a diverse range of biological activities.[3] The 2,5-diketopiperazine (DKP) core is readily accessible through the condensation and cyclization of amino acids, which is the first committed step in the biosynthesis of countless secondary metabolites. Cyclization of small peptide-based compounds dramatically decreases the metabolic liability of their linear counterparts. FDA approved DKP drugs include the PDE5 inhibitor tadalafil (Cialis®), and several more have reached phase III clinical trials such as the tubulin polymerization inhibitor plinabulin, and the oxytocin receptor antagonist retosiban.



Tadalafil (Cialis®)
PDE5 Inhibitor

Plinabulin (Phase III)
Anti-cancer

Retosiban (Phase III)
Oxytocin Antagonist

Figure 4.1: Representitave 2,5-diketopiperazines which have received FDA approval or reached phase III clinical trials.

4.1.1 Bioactive Indole Diketopiperazine (IDP) Natural Products

Thousands of naturally occurring IDPs have been characterized, many of which have biosynthetic origins similar to tryprostatin, with biological activities spanning a diverse spectrum of targets. Verruculogin, for example, is a cytotoxic and neurotoxic

compound structurally related to the tryprostatin family featured in chapter 2 (Figure 4.2). Verruculogin inhibits mouse tsFT210 cells in the M phase, exhibiting an MIC value of 12.2 uM.[4] The tryprostatins, spirotryprostatins, and verruculogen have all been isolated from *Aspergillus fumigatus*. Verruculogen has also been isolated from several *Penicillium* species.[3]

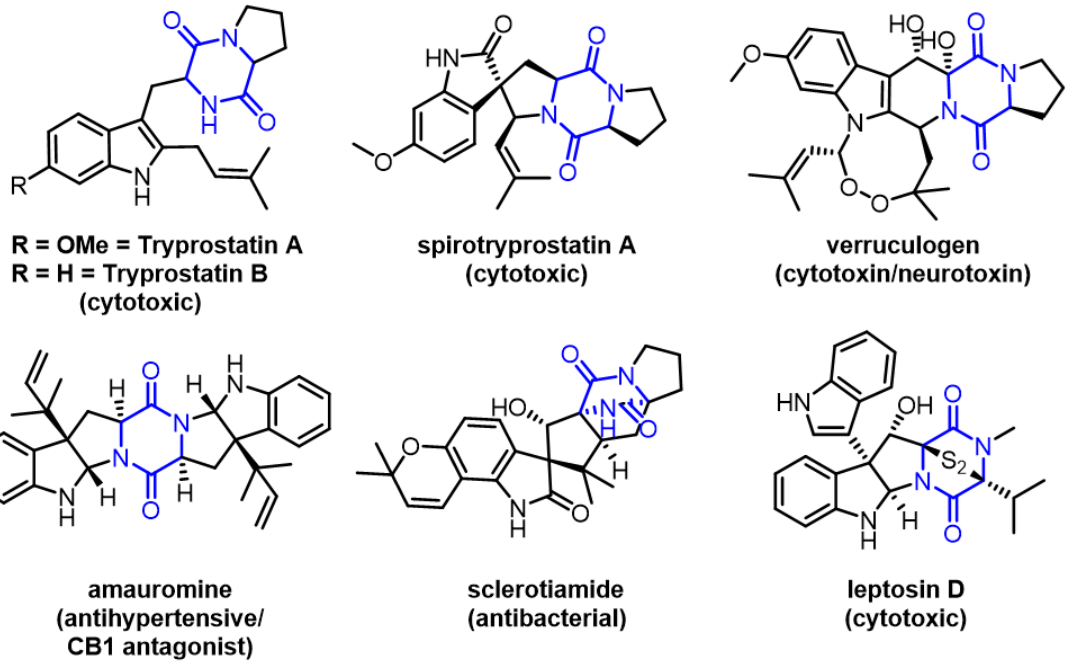
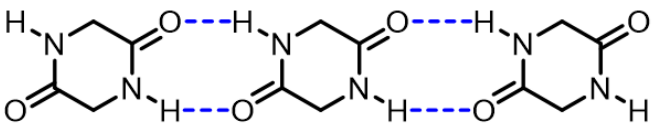


Figure 4.2:Representitave bioactive IDP natural products

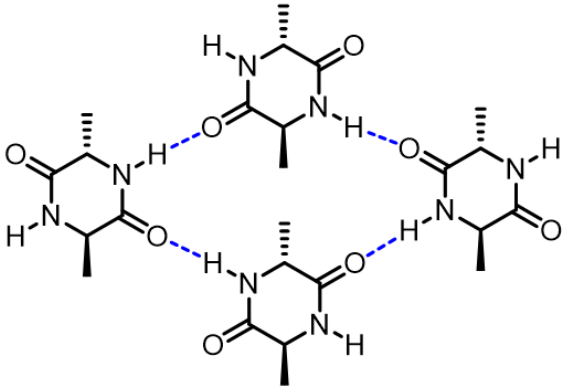
Sclerotiamide is a NP isolated from *A. sclerotiorum* which was initially noted to have an insecticidal effect against the corn earworm, however recent studies have revealed that it acting as the first known non-peptide-based activator of caseinolytic protease P (ClpP). Activation of ClpP induces nonspecific degradation of cellular proteins and peptides, making sclerotiamide an antibiotic with a unique mechanism of action.[5] Leptosin D is a disulfide containing IDP isolated from *Leptosphaeria* which displays anticancer activity against P388 cells with an ED50 of 0.086 ug/mL.[6] Amauromine is an IDP with unique bioactivity isolated from the fungus *Auxarthron reticulatum*. It was originally

identified as having strong vasodilating effects, and later found to be a potent and selective CB1 antagonist.[7] This privileged scaffold is an attractive source of potential new bioactive compounds.

4.1.2 Drawbacks of the IDP Scaffold



Tape H-Bonding



Layer H-Bonding

Figure 4.3 Typical H-bonding modes of 2,5-DKP in the solid state

Despite the remarkably broad spectrum of bioactivity observed in IDP NPs, the relatively high structural complexity and low natural abundance limits in vivo studies.[8] Additionally, this class of compounds frequently suffers from low aqueous solubility which is a major problem in the early stages bioactivity studies.[9] The unusually poor aqueous solubility of IDPs is presumably due to the combination of π-π stacking and intermolecular hydrogen bonding networks which both promote tight crystal packing (Figure 4.3).[10, 11] Therefore, optimization of IDP scaffolds with enhanced solubility has potential to enhance their pharmacological and pharmaceutical properties. Replacement of the core

hydrophobic indole moiety with more hydrophilic bioisosteres is a well-established strategy to further expand and optimize this diverse class of bioactive molecules.

4.1.3 Isosteric Replacement in IDP Scaffolds

The similar steric demands indole and azaindole is advantageous, as any dramatic changes in binding affinity would likely arise from previously unavailable binding interactions compared to the indole parent compound. The incorporation of azacyclic isosteres such as azaindoles can enable new hydrogen bonding interactions with the drug target, which can dramatically enhance binding affinity and target selectivity. However, most biologically active IDPs feature a highly functionalized indole core, therefore requiring labor intensive total synthesis to produce azaindole-substituted analogs. With the current limitations of late-stage functionalization methods, there is a need for innovative methods of natural product diversification.

Poly-functionalized azaindole scaffolds are most commonly accessed by late-stage generation of the azaindole moiety from functionalized precursors, a strategy which is effective for the total synthesis of specific target compounds, but inevitably imposes restrictions on functional group tolerance and synthetic practicality.[12-16] Most synthetic methods for functionalizing preexisting azaindole moieties is dependent on directed metalation groups to provide regiospecificity.[17-19] Alternative strategies of azaindole functionalization typically require halogenated or otherwise activated azaindoles capable of participating in organometallic coupling reactions.[20-23] Considering the prevalence of azacycles in pharmaceuticals, there is a high demand innovative methodology for the chemoselective and regioselective functionalization of azacyclic scaffolds.

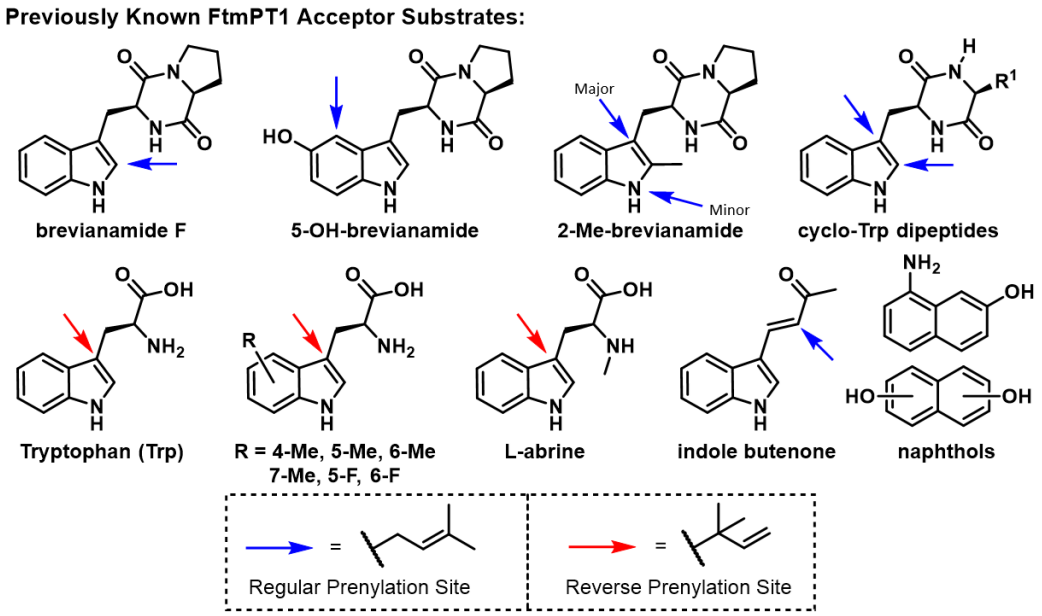


Figure 4.4: Literature reported prenyl acceptor substrates for FtmPT1

Several previous studies have explored the acceptor substrate promiscuity of indole PTs such as FgaPT2, CdpNPT, and FtmPT1, the scope of which includes Trp-containing cyclic dipeptides,[24-26] 5-hydroxy-brevianamide,[26] 2-methyl-brevianamide,[26] Trp,[27] Methyl-Trp,[27] Fluoro-Trp,[27] various hydroxynaphthlenes,[28] and an indole butanone.[29] (Figure 4.4) Initial studies of FtmPT1 catalyzed prenylation of Trp and methyl/fluoro-Trp analogs reported N-1 reverse prenylated Trp products,[27] however later mechanistic studies have given firm evidence that the products are instead a cyclized C-3 reverse prenylated pyrroloindoline.[26] Another well studied promiscuous PT known to prenylate indole containing CyWP substrates is CdpNPT. The native transformation catalyzed by CdpNPT is C3 reverse prenylation of the indole moiety, which produces cyclized pyrroloindoline products via nucleophilic attack of the C2 position after prenylation-induced dearomatization. In addition to many of the known acceptor substrates for FtmPT1, CdpNPT has demonstrated a remarkable ability to functionalize more complex Trp containing substrates such as the FDA approved lipodepsipeptide antibiotic daptomycin,

as well as the natural product Tryprostatin B.[30-32] CdpNPT not only features an expanded acceptor scope, it has also demonstrated compatibility with a broad spectrum of unnatural allylic, benzylic, heterobenzylic, and diene containing pyrophosphate donors. Notably, several daptomycin analogs chemoenzymatically synthesized using CdpNPT and a library of synthetic pyrophosphate doners were highly active against daptomycin resistant strains of *Staphylococcus aureus* (MRSA).[31]

Broadly, these PT substrate scope investigations have revealed that in many cases, modifications to the native substrate alter the prenylation regiospecificity. Tryprostatin B for example, a C2 prenylated CyWP, was well accepted by CdpNPT, but the alkylation site was almost exclusively on the indole C6 position.[30] While a massive number of non-native indole containing substrates have been successfully alkylated by indole PTs, alteration of the core aromatic scaffold remains underexplored.

4.1.4 Aims and Significance

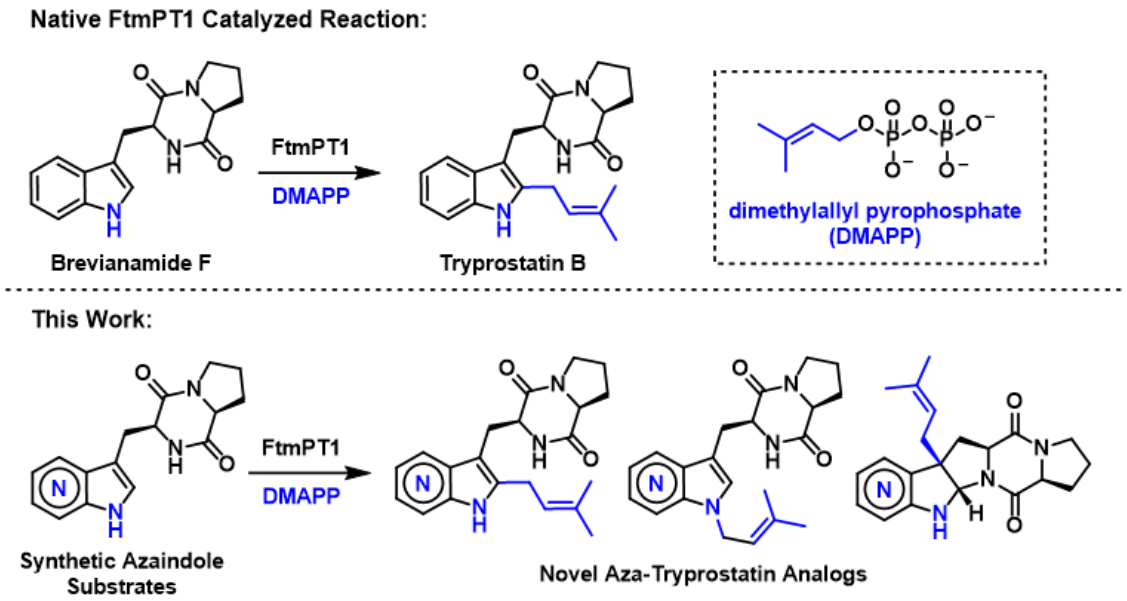


Figure 4.5: Native biosynthetic role of FtmPT1 (top) compared to the reactions observed in this study (bottom)

Considering the pharmaceutical relevance of azacycles, there is high demand to expand the catalytic utility of indole PTs into unnatural azacyclic scaffolds to provide environmentally friendly access to previously inaccessible chemical space. The goal of this study was to evaluate the ability of indole PTs to accept azaindole-substituted IDP analogs to rapidly generate azaindole-containing natural product analogs (Figure 4.5). This late-stage diversification strategy serves as a novel approach for generating new NP-inspired scaffolds, by exploiting the low substrate specificity of the PT's within the biosynthetic pathway of the parent compound. Through synthesizing a series of AzaTrp substituted CyWP substrates, the compatibility of indole PTs with azaindole-containing substrates will be explored to enhance the catalytic utility of PTs as late-stage functionalization catalysts. This substrate-directed biocatalytic diversification strategy enables facile access to previously inaccessible NP analogs, and therefore new structure-activity relationship studies for pharmaceutical development.

4.2 Materials and Methods

4.2.1 General Materials

Unless otherwise stated, all chemicals and reagents were purchased from Acros (New Jersey, USA), Sigma-Aldrich (St. Louis, MO, USA), Ambeed (Arlington Heights, IL, USA), or AK Scientific (Union City, CA, USA) and were reagent grade or better. The Ni-NTA columns and PD-10 columns used for protein purification were purchased from GE Healthcare (Piscataway, NJ).

4.2.2 General Methods

All reported reactions were conducted in oven-dried glassware under an anhydrous nitrogen atmosphere with anhydrous solvents, unless otherwise noted. Column

chromatography purification was performed using silica gel (SiliCycle Inc, P60, particle size 40-63 µm). HPLC analysis and purification was conducted using an Agilent 1220 system equipped with a diode array detector (DAD). High-resolution mass spectrometry (HRMS) confirmation of products and substrates was conducted on an Agilent 6545-QTOF W/ 1290 HPLC mass spectrometer at the University of Oklahoma Department of Chemistry and Biochemistry. NMR Spectra were collected on a Varian VNMRS 500 MHz, a Varian INOVA 600 MHz, a Varian Mercury VX-300, or a Varian VNMRS 400 MHz NMR spectrometer at the NMR facility of the Department of Chemistry and Biochemistry of the University of Oklahoma. All NMR spectra were obtained in 400-600 µL volumes of Methanol-d_4 (99.8%), DMSO-d_6 (99.9%), $CDCl_3$ (99.8%), or D_2O (99.9%) (Cambridge Isotope Laboratories, MA, USA), and processed using MestReNova (Version 12.0.3 and 14.2.1, Mestrelab Research, S.L, Compostela, Spain) software. Structural elucidation and NMR peak assignment of chemoenzymatically prenylated products was performed using 1H-COSY, 1H-13C HSQC, and 1H-13C HMBC experiments.

4.2.3 HPLC Methods

HPLC Method 1: Used for the purification of **2f** and **2A-N1** [gradient of 10% B to 100% B over 25 min, 100% B for 4 min, 100% B to 10% B in 0.1 min, 10% B for 6 min (A = ddH_2O with 0.1% formic acid; B = acetonitrile); flow rate = 2 mL min^{-1}; A_{280}].

HPLC Method 2: Used for the purification of **4f** and **5-C3** [gradient of 10% B to 100% B over 30 min, 100% B for 7 min, 100% B to 10% B in 0.1 min, 10% B for 8 min (A = ddH_2O with 0.1% formic acid; B = acetonitrile); flow rate = 1.5 mL min^{-1}; A_{280}].

HPLC Method 3: Used for the purification of **4-C3, 4-N1** and **5f** [gradient of 5% B to 70% B over 18 min, 70% B to 100% B over 2 min, 100% B for 5 min, 100% B to

25% B in 0.1 min, 5% B for 5 min (A = ddH_2O with 0.1% ammonia; B = acetonitrile); flow rate = 2 mL min^{-1}; A_{280}, A_{340}].

HPLC Method 4: Used for the purification of **6f** [gradient of 20% B to 100% B over 30 min, 100% B for 5 min, 100% B to 20% B in 0.2 min, 20% B for 5 min (A = ddH_2O with 0.1% ammonia; B = acetonitrile); flow rate = 2 mL min^{-1}; A_{280}].

HPLC Method 5: Used for the purification of **7f, 7-N1**, **7-C2,** and **7-C3** [gradient of 10% B to 60% B over 20 min, 60% B to 100% B in 5 min, 100% B for 5 min, 100% B to 10% B in 0.1 min, 10% B for 5 min (A = ddH_2O with 0.1% ammonia; B = acetonitrile); flow rate = 2 mL min^{-1}; A_{280}].

HPLC Method 6: Used for the purification of **6-C3** [gradient of 10% B to 100% B over 30 min, 100% B for 7 min, 100% B to 10% B in 0.1 min, 10% B for 8 min (A = ddH_2O with 0.1% formic acid; B = acetonitrile); flow rate = 1.5 mL min^{-1}; A_{280}].

HPLC Method 7: Used for the analytical scale chemoenzymatic reactions containing AzaCyWP substrates. [gradient of 1% B to 100% B over 25 min, 100% B for 3 min, 100% B to 1% B over 0.1 min, 1% B for 7 min (A = ddH_2O with 0.1% TFA; B = acetonitrile); flow rate = 1 mL min^{-1}; A_{230} + A_{254} + A_{280} + A_{316} + A_{340}].

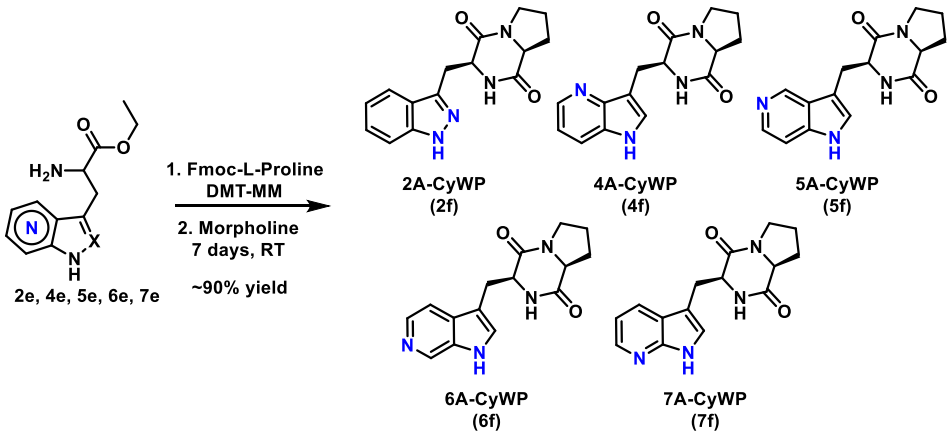


Figure 4.6: Synthesis of Aza-CyWP Substrates

4.2.4 Synthesis of Aza-Trp-Pro cyclic dipeptides (2f, 4f, 5f, 6f, 7f)

Fmoc-L-proline (25.5 mg, 0.135 mmol) was added to 3 mL EtOH containing 6-Aza-Trp ethyl ester **6e** (30 mg, 0.128 mmol) and the mixture was cooled to 0°C in an open flask. DMT-MM (39.2 mg, 0.142 mmol) was then added in one portion to the reaction mixture. The reaction mixture was stirred for an additional 2 hours at 0°C before removal of the solvent in vacuo. Morpholine (4 mL) was then added to the residue and stirred at room temperature for 7 days. (Figure 3) Solvent was then removed by nitrogen flow, and the crude residue was redissolved in aqueous K_2CO_3. The mother liquor was then extracted three times with $CHCl_3$:Isopropanol (3:1), and the organic layers were combined, dried over anhydrous K_2CO_3, filtered, and concentrated in vacuo. The crude product was purified via silica gel column chromatography starting with EtOAc, followed by a gradient of $CHCl_3$:MeOH:NH_4OH (90:10:1 → 40:10:1) to elute the product. Desired (L,L) diastereomers are known to be less polar than their corresponding (D,L) diastereomers. Reverse phase HPLC purification was then carried out using the specified HPLC methods.

(8aS)-3-((1H-indazol-3-yl)methyl)hexahydropyrrolo[1,2-a]pyrazine-1,4-dione (2f): Semi-preparative RP-HPLC was conducted using a Gemini-NX C18 (5 µm, 10 mm x 250 mm) column (Phenomenex, Torrance, CA, USA) and **HPLC Method 1** (described above) to purify **2f** as a single diastereomer. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.75 (d, *J* = 8.1 Hz, 1H), 7.48 (d, *J* = 8.4 Hz, 1H), 7.40 (t, *J* = 7.6 Hz, 1H), 7.27 (s, 1H), 7.17 (t, *J* = 7.5 Hz, 1H), 4.46 (dt, *J* = 10.5, 2.5 Hz, 1H), 4.11 (t, *J* = 7.7 Hz, 1H), 3.93 (dd, *J* = 15.7, 3.3 Hz, 1H), 3.74 – 3.55 (m, 4H), 3.25 (dd, *J* = 15.7, 10.4 Hz, 1H), 2.42 – 2.28 (m, 1H), 2.13 –

1.99 (m, 4H), 1.98 – 1.84 (m, 1H), 1.26 (t, *J* = 5.0 Hz, 2H). HRMS-ESI: Calc for $C_{15}H_{16}N_4O_2$ $[M+H]^+$: 285.135151; Found: 285.1350.

(8aS)-3-((1H-pyrrolo[3,2-b]pyridin-3-yl)methyl)hexahydropyrrolo[1,2-a]pyrazine-1,4-dione (4f): Semi-preparative RP-HPLC was conducted using a Gemini-NX C18 (5 µm, 10 mm x 250 mm) column (Phenomenex, Torrance, CA, USA) and **HPLC Method 2** (described above) to purify **4f** as a single diastereomer. ^{1}H NMR (500 MHz, d_6-DMSO) δ 11.23 (s, 1H), 8.88 (s, 1H), 8.32 (dd, J = 4.6, 1.4 Hz, 1H), 7.79 (dd, J = 8.2, 1.4 Hz, 1H), 7.58 (s, 1H), 7.14 (dd, J = 8.2, 4.6 Hz, 1H), 4.27 (dt, J = 7.4, 2.0 Hz, 1H), 4.16 (t, J = 7.7 Hz, 1H), 3.51 (dd, J = 15.1, 3.3 Hz, 1H), 3.45 – 3.36 (m, 2H), 2.95 (dd, J = 14.9, 8.8 Hz, 1H), 2.10 (dt, J = 11.1, 6.5 Hz, 1H), 1.92 – 1.72 (m, 3H). HRMS-ESI: Calc for $C_{15}H_{16}N_4O_2$ $[M+H]^+$: 285.135151; Found: 285.1349.

(8aS)-3-((1H-pyrrolo[3,2-c]pyridin-3-yl)methyl)hexahydropyrrolo[1,2-a]pyrazine-1,4-dione (5f): Semi-preparative RP-HPLC was conducted using a Gemini-NX C18 (5 µm, 10 mm x 250 mm) column (Phenomenex, Torrance, CA, USA) and **HPLC Method 3** (described above) to purify **5f** as a mixture of diastereomers**.** ^{1}H NMR (600 MHz, Methanol-d_4) δ 9.17 (s, 1H), 9.00 (s, 1H), 8.44 (s, 3H), 8.28 (d, *J* = 6.6 Hz, 2H), 7.82 (d, *J* = 6.7 Hz, 1H), 7.78 (d, *J* = 6.5 Hz, 1H), 7.64 (s, 1H), 7.57 (s, 1H), 4.57 (s, 1H), 4.24 (t, *J* = 5.5 Hz, 1H), 4.19 – 4.13 (m, 1H), 3.61 – 3.52 (m, 2H), 3.47 (dd, *J* = 14.8, 5.9 Hz, 1H), 3.42 – 3.32 (m, 3H), 3.29 (d, *J* = 4.8 Hz, 1H), 3.22 (t, *J* = 10.9 Hz, 1H), 3.12 (dd, *J* = 11.1, 6.5 Hz, 1H), 2.16 (dt, *J* = 13.0, 6.8 Hz, 1H), 2.11 (dt, *J* = 12.8, 6.6 Hz, 1H), 1.91 (q, *J* = 8.7 Hz, 1H), 1.81 (p, *J* = 9.3 Hz, 1H), 1.77 – 1.68 (m, 2H), 1.61 (p, *J* = 10.0 Hz, 1H),

1.43 – 1.33 (m, 1H). HRMS-ESI: Calc for $C_{15}H_{16}N_4O_2$ $[M+H]^+$: 285.135151; Found: 285.1355.

(8aS)-3-((1H-pyrrolo[2,3-c]pyridin-3-yl)methyl)hexahydropyrrolo[1,2-a]pyrazine-1,4-dione (6f): Semi-preparative RP-HPLC was conducted using a Gemini-NX C18 (5 µm, 10 mm x 250 mm) column (Phenomenex, Torrance, CA, USA) and **HPLC Method 4** (described above) to purify **6f** as a single diastereomer. ^{1}H NMR (600 MHz, d_6-DMSO) δ 11.36 (s, 1H), 8.67 (s, 1H), 8.05 (d, *J* = 5.5 Hz, 1H), 7.96 (s, 1H), 7.57 (d, *J* = 5.5 Hz, 1H), 7.39 (d, *J* = 2.3 Hz, 1H), 4.33 (t, *J* = 5.1 Hz, 1H), 4.05 (t, *J* = 8.0 Hz, 1H), 3.77 (dd, *J* = 8.4, 5.7 Hz, 1H), 3.24 (ddd, *J* = 12.0, 9.1, 3.8 Hz, 1H), 3.21 – 3.11 (m, 2H), 2.96 (dt, *J* = 11.3, 6.0 Hz, 1H), 2.64 – 2.57 (m, 1H), 2.01 – 1.89 (m, 2H), 1.74 – 1.50 (m, 3H), 1.30 (dt, *J* = 20.0, 10.5 Hz, 1H). HRMS-ESI: Calc for $C_{15}H_{16}N_4O_2$ $[M+H]^+$: 285.135151; Found: 285.1348.

(8aS)-3-((1H-pyrrolo[2,3-b]pyridin-3-yl)methyl)hexahydropyrrolo[1,2-a]pyrazine-1,4-dione (7f): Semi-preparative RP-HPLC was conducted using a Gemini-NX C18 (5 µm, 10 mm x 250 mm) column (Phenomenex, Torrance, CA, USA) and **HPLC Method 5** (described above) to purify **7f products**. The **7f** obtained after chromatographic separation was determined to be a 2:1 mixture of diastereomers (L,L/D,L) as shown by NMR analysis. ^{1}H NMR (500 MHz, d_6-DMSO) (L,L-Diastereomer Peaks)**:** δ 11.33 (s, 1H), 8.17 – 8.10 (m, 1H), 7.96 (dd, *J* = 7.9, 1.6 Hz, 1H), 7.88 (s, 1H), 7.23 (d, *J* = 2.4 Hz, 1H), 6.98 (td, *J* = 7.8, 4.6 Hz, 1H), 4.30 (t, *J* = 5.0 Hz, 1H), 4.01 (dd, *J* = 9.8, 7.1 Hz, 1H), 3.38 – 3.25 (m, 1H), 3.20 (ddd, *J* = 13.1, 9.3, 4.2 Hz, 1H), 3.16 – 3.05 (m, 2H), 1.93 (dtd, *J* = 12.2, 6.8, 2.7 Hz, 1H), 1.74 – 1.46 (m, 2H), 1.27 (ddt, *J* = 12.2, 10.4, 5.2 Hz, 1H). HRMS-ESI: Calc for $C_{15}H_{16}N_4O_2$ $[M+H]^+$: 285.135151; Found: 285.1341.

4.2.5 Analytical scale prenyltransferase reaction screening

Analytical scale FtmPT1 reactions were conducted in a total volume of 20 µL containing 45 µM FtmPT1, 1 mM substrate **(2f, 4f, 5f, 6f, or 7f)**, 1.5 mM DMAPP, 25 mM Tris (pH 7.8), and 5 mM $CaCl_2$. Analytical scale CdpNPT reactions were conducted in a total volume of 20-50 µL containing 200 µM CdpNPT, 1 mM substrate **(2f, 4f, 5f, 6f, or 7f)**, 1.5 mM DMAPP, 50 mM Tris (pH 7.8), and 5 mM $MgCl_2$. Negative control reactions without enzyme were conducted in parallel to verify that products did not emerge via non-biocatalytic mechanisms. The reaction mixtures were incubated at 35°C for 18 hours and then quenched with the addition of 40 µL cold methanol. The precipitated protein was removed via centrifugation (9000 x g for 30 min). Analytical scale reactions were analyzed on an Agilent 1220 HPLC system equipped with a DAD, using reverse-phase Gemini-NX C18 (5 µm, 4.6 mm x 250 mm) column (Phenomenex, Torrance, CA, USA) using **HPLC Method 7** for Aza-CyWP substrates. For each reaction, percent turnover was calculated by dividing the area under the unreacted substrate peak by that of total amount of substrate used for the reaction, as calculated from the negative control and/or using calibration curves in cases where the absorbance profiles of products differed from their substrates. All positive reactions were subsequently confirmed using high-resolution mass spectrometry (HRMS) as shown in Appendix 2 (4.6).

4.2.6 Enzymatic Scale Up Reactions for Aza-CyWP Prenylation

Scale up reactions of Aza-CyWP prenylation were conducted in a volume of 15-20 mL consisting of 1 mM Aza-CyWP**,** 1.5 mM DMAPP, 25 mM Tris (pH 7.8) and 5 mM $CaCl_2$. Reactions were initiated by the addition of 10 mg purified FtmPT1 or 20 mg of CdpNPT (~350 µL) at 35°C for 16-18 h. Reaction progress was monitored on analytical

HPLC as described above by taking 50 µL aliquots at regular time intervals. An additional 10 mg enzyme was added if poor conversion (<25%) was observed, followed by another 24 h incubation at 35°C. Subsequently, the reactions were basified with K_2CO_3 and extracted 3 times with 15 mL volumes of EtOAc. Brief centrifugation was used to accelerate organic layer separation. The combined organic fractions were dried over Na_2SO_4 and concentrated under reduced pressure. Putative products were subsequently isolated by semi-preparative RP-HPLC using Gemini-NX, C-18 (5 µm, 10 × 250 mm) column (Phenomenex, Torrance, California, USA) on an Agilent 1220 system equipped with a DAD detector using Semi-preparative **HPLC Method 1** for **2f** scale up reactions, **HPLC Method 3** for **4f** scale up reactions (**4-N1** and **4-C3** isolation**)**, **HPLC Method 2** for **5f** scale up reactions (**5-C3** isolation), **HPLC Method 6** for **6f** scale up reactions (**6-C3** isolation), and **HPLC Method 5** for **7f** scale up reactions (**7-N1, 7-C3,** and **7-C2** isolation).

4.3 Results and Discussions

To evaluate the ability of Aza-CyWP analogs (**2f, 4f, 5f, 6f, 7f**) to act as substrates for prenyltransferases (PTs), three PTs were selected. The selected PTs, all from *Aspergillus fumigatus*[25, 33, 34] include FgaPT2, a Trp C4-PT, CdpNPT, a cyclic-Trp-Pro dipeptide C3-reverse PT, and FtmPT1, a cyclic-Trp-Pro C2-PT from the biosynthetic pathways of fumigaclavine C, aszonalenin, and fumitremorgin B respectively.[25, 34-36] All aza-CyWP analogs were evaluated with all three PTs under standard reaction conditions in the presence of DMAPP as the prenyl group donor. Appropriate controls were included to distinguish between enzymatic and non-enzymatic transformations. Formation of products were monitored using differences in retention time between the substrate and products using RP-HPLC. This initial screen revealed FgaPT2, did not accept Aza-CyWP analogs

as substrates, which is consistent with enzymes' reported low prenylation efficiency with CyWP substrate.[35]

4.3.1 Analytical Scale Aza-CyWP Substrate Screening with CdpNPT

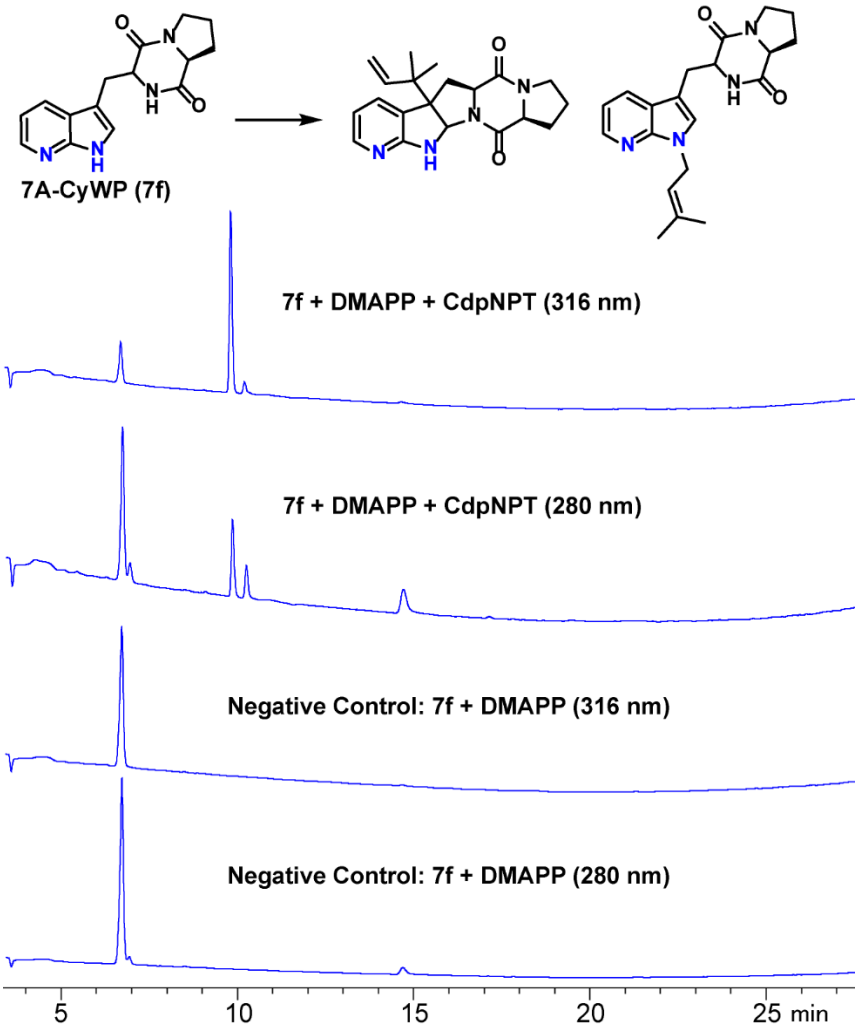


Figure 4.7 RP-HPLC traces of analytical scale reactions containing 7f in the presence of CdpNPT and DMAPP at 316 nm and 280 nm
Negative controls which lack enzyme displayed below as reference.

Surprisingly, despite the known promiscuity of CdpNPT, accepted only **7f** as substrate resulting in two products with a cumulative product formation of 28.7%. The **7f** was used as a mixture of diastereomers, which decreased conversion based on the known diastereomeric selectivity of CdpNPT.[27] The major product displayed altered absorbance

profile compared to the substrate, indicative of dearomatization of the pyrrole ring (Figure 4.7), which is consistent with the native reaction of CdpNPT.[26, 31] The unexpected selectivity for **7f** over the other Aza-CyWP isomers likely arises from the variations in C3 electron density, with **7f** being the most like the native substrate, thus allowing **7f** to proceed through analogous transformations.

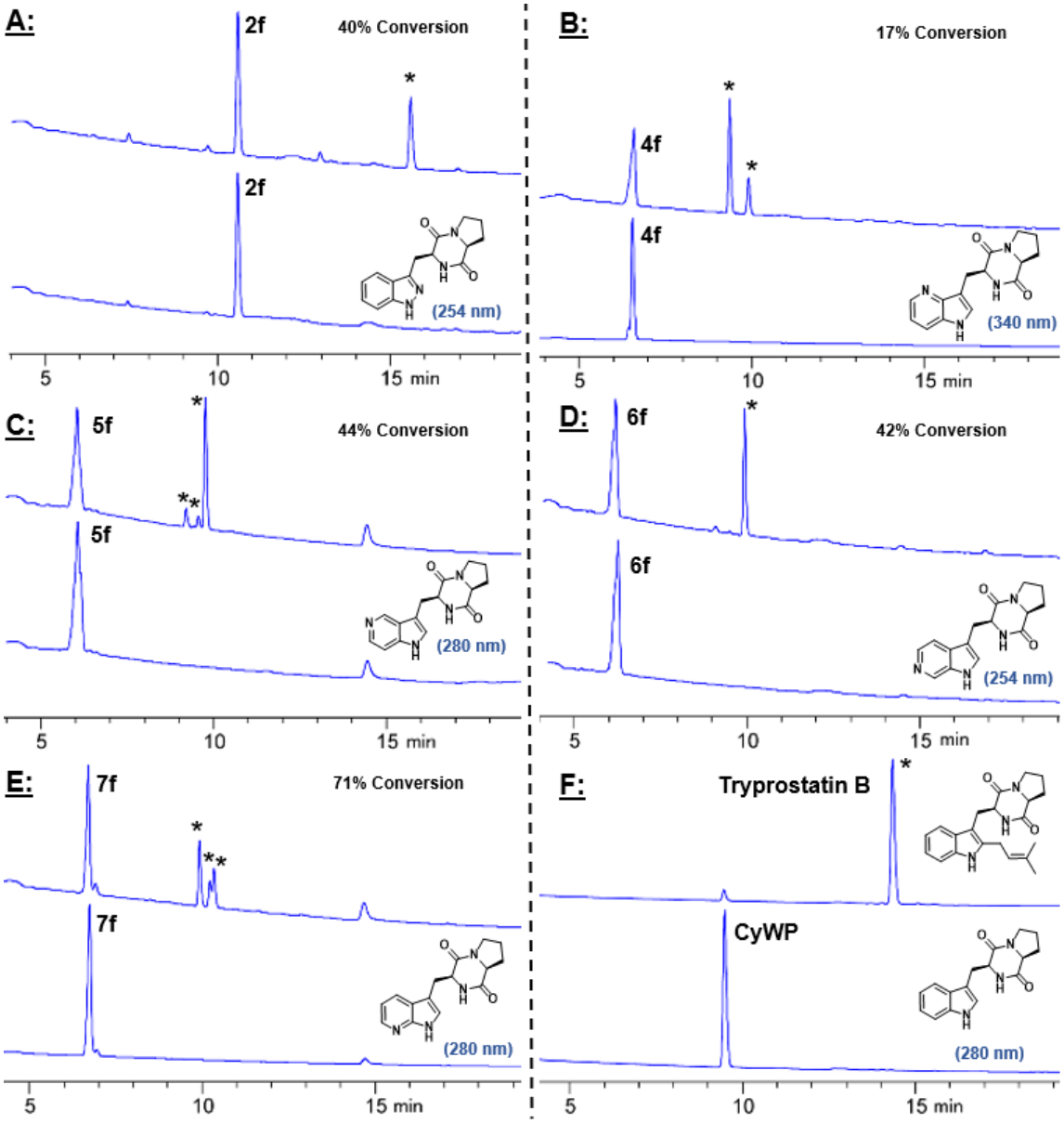


Figure 4.8: RP-HPLC traces of analytical scale reactions containing Aza-CyWP analogs in the presence of FtmPT1 and DMAPP (Panel A-F).

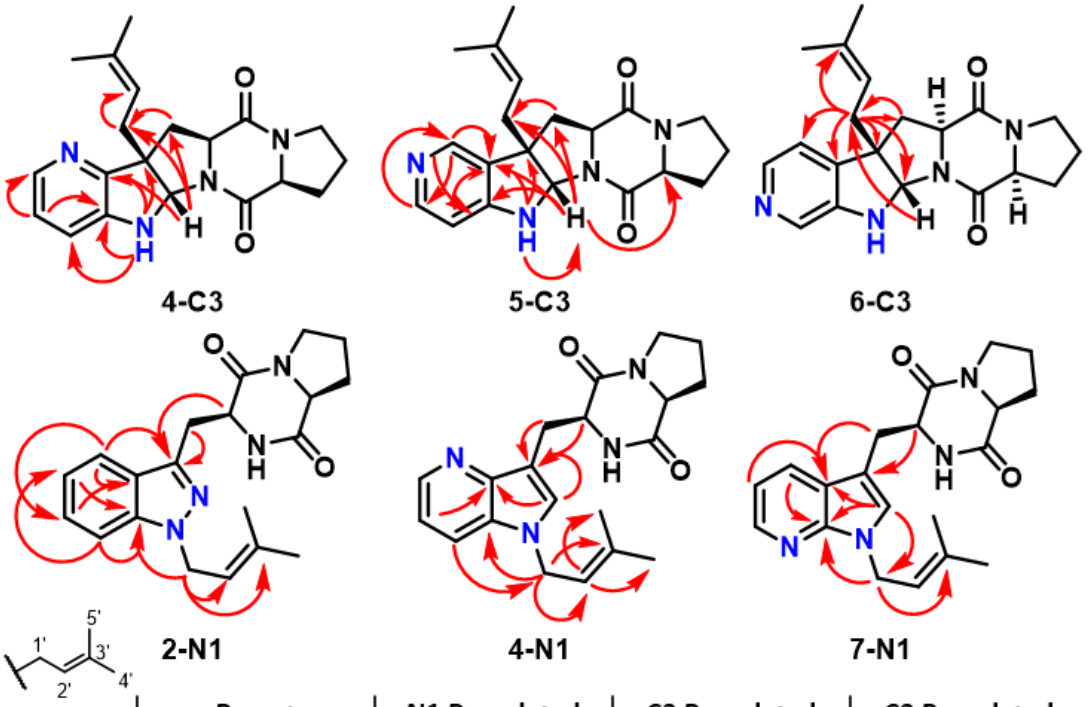
Interestingly, the C2-PT FtmPT1 accepted all Aza-CyWP as substrates with DMAPP as prenyl donor under standard reaction conditions. Consistent with the substrate choice of FtmPT1, which prenylates Trp moieties only within the context of cyclic-dipeptides,[24, 37, 38] it did not accept any aza-Trp analogs. The analysis of analytical scale reactions revealed the formation of a single major product for **2f**, and **6f**, and 2-3 products for **4f**, **5f**, and **7f**. (Figure 4.8) The calculated percent conversion for **2f, 4f, 5f, 6f,** and **7f** under these conditions was found to be 40%, 17%, 44%, 42% and 71%, respectively. Percent conversion for **5f** and **7f,** which exist as a mixture of diastereomers, was calculated to represent the percent conversion of the accepted (L,L) diastereomer to enable more accurate comparison between the Aza-CyWP substrates, since the (D,L) diastereomers are not accepted by FtmPT1. It was noted that the major prenylated products (**4-C3**, **5-C3**, and **6-C3** described below) in reactions containing **4f**, **5f**, and **6f** had substantially altered absorbance maxima compared to the starting material, likely indicating the formation of C3 alkylated dearomatized products.[24, 26]

4.3.2 Enzymatic Scale-up Reactions and Purification

To determine the regiospecificity of PT-catalyzed prenyl-transfer, the 6 reactions which showed turnover on analytical scale were scaled up. Scale up reactions were performed using 5-7 mg of Aza-CyWP's under standard reaction conditions (see 4.2.6 Materials and Methods). The prenylated Aza-CyWP products were extracted in ethyl acetate (EtOAc), and subsequently, the partially pure products were purified using semipreparative reverse-phase high-performance liquid chromatography (RP-HPLC) (see 4.2.6 Materials and Methods). In cases where the reactions yielded multiple products

(FtmPT1 + **4f**/**5f**/**7f**), 0.1% formic acid aqueous phase was replaced with a basic 0.1% ammonia aqueous phase to enhance product separation.

4.3.3 Structural Elucidation of FtmPT1 Catalyzed Aza-CyWP Prenylated Products



	Parent		N1 Prenylated		C3 Prenylated		C2 Prenylated	
Peak	δH	δC	δH	δC	δH	δC	δH	δC
N1	11.2				6.8-7.5		11.2	
C2	~7.6	128	~8	130-140	5.2	77		136*
C3		108		108		55		105*
C1'			4.8-5.4	41-54	2.4	35	3.4	25

Figure 4.9: **Product structures isolated from FtmPT1 catalyzed reactions, showing HMBC correlations as red arrows.**

All five Aza-CyWP analogs acted as substrates for FtmPT1 reaction. Upon scale up and characterization of enzymatically prenylated products, a single major product was fully characterized from **2f, 5f,** and **6f** containing reactions**.** Scale up reactions containing **4f** and **7f** produced a separable mixture of two regioisomers in quantities sufficient for NMR characterization. The **2f** substrate produced a single N1 prenylated product, as supported by lack of N1 proton signals, the chemical shift values of prenyl CH_2 (C1'/H1'

= 46.1/4.94 ppm), and HMBC correlations between the prenyl H1' and the indazole C7a position (Figure 4.9). N-prenylation has been consistently observed to shift the prenyl group's H1' chemical shift down field (>4.7 ppm), further supporting this structural assignment. The chemical shifts of the prenyl H1' show a consistent trend, with N-prenylated H1' at 4.8-5.4 ppm, C2-prenylated H1' at ~3.4, and C3-prenylated H1' at ~2.35 ppm (Figure 4.9). N1 prenylated products were isolated from reactions containing **2f** and **7f** as the major product, and from reactions containing **4f** as a minor product. Surprisingly, the expected C2 prenylated product was only isolated from the reaction containing **7f.** The enzyme substrates **5f** and **7f** were used as a mixture of diastereomers, producing diastereomerically pure prenylated products in the presence of FtmPT1. This observation supports the hypothesis that the (L)-(L) diastereomer of **5f** and **7f** were exclusively accepted by FtmPT1.

4.3.4 Cyclized C3 Prenylated Azaindole Products

An unexpected C3 alkylated cyclized product was observed as the major product for **4f**, **5f**, and **6f**. The diagnostic NMR signal for the C3 alkylated products is the dramatic upfield shift of the C2 ^{13}C resonance from ~125-145 ppm, to ~73-78 ppm, and their associated H2 signal shifting from >7 ppm to ~5.3 ppm. These cyclized products also have dramatically altered C3 chemical shifts (55 ppm). Normal prenylation of C3 is clear due to the presence of the C2' methine signal, and the absence of terminal alkene signals indicative of reverse prenylation. This prenylation induced cyclization dearomatizes the pyrrole ring, altering the absorbance spectrum observed via HPLC. Native FtmPT1 catalyzed C3 normal prenylation/cyclization has only been previously reported in the case of 2-methyl-CyWP.[26] Overall, every Aza-CyWP substrate tested was successfully

prenylated by FtmPT1 to produce a total of 7 fully characterized products that were alkylated at C3, N1, and C2. (Figure 4.10) Determination of the absolute stereochemistry of the product chiral centers was achieved using diagnostic NOE interactions. The (l,l) configuration of the diketopiperazine ring was confirmed by 2D-ROESY experiments displaying clear interactions between the α protons of the Aza-Trp and l-Proline. In the case of cyclized products, the *syn-cis* configuration of the new chiral centers at C2 and C3 was confirmed by observed interactions between H2 and H2' and the absence of interactions between H2 and the Aza-Trp α proton.

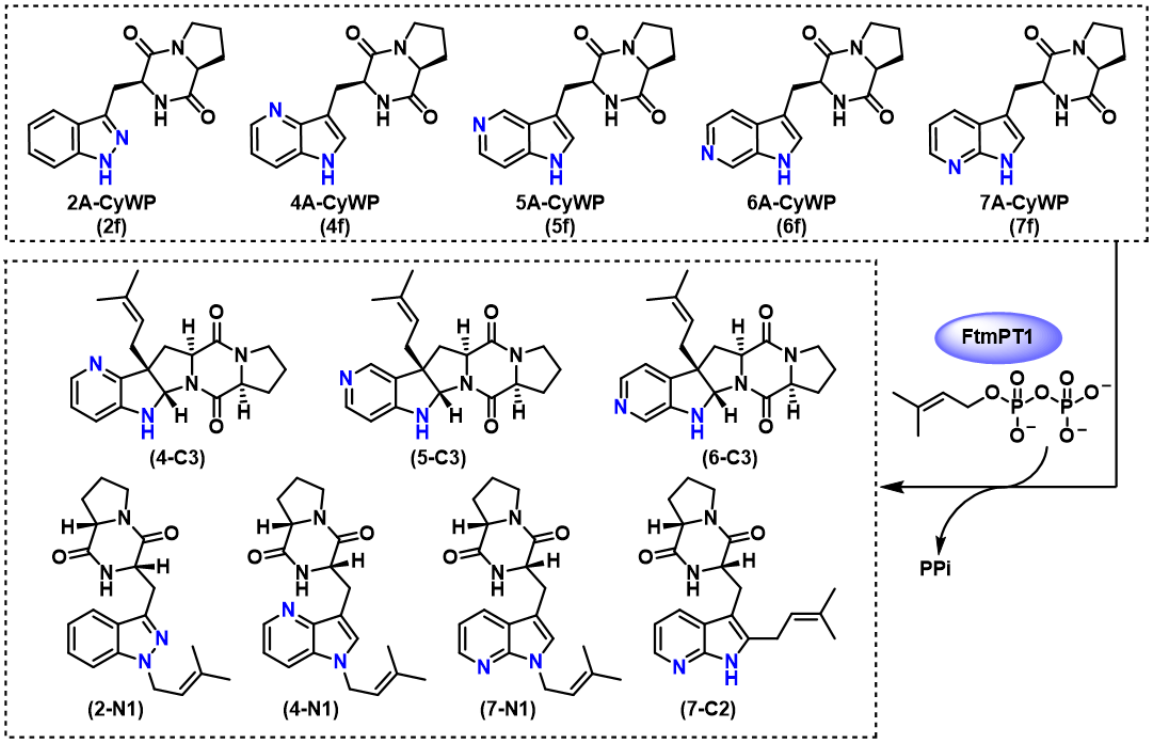


Figure 4.10: Azaindole dipeptide substrates accepted by FtmPT1 and their isolated reaction products.

4.3.5 Structural Elucidation of CdpNPT Catalyzed Aza-CyWP Prenylated Products

Among the prenylated Aza-CyWP products, two prenylated **7A-CyWP** products were isolated from the CdpNPT catalyzed reaction. The major product showed dramatic shifts in the product absorbance profile, which upon NMR analysis revealed it to be pyrrole-ring dearomatized, cyclized C3-reverse prenylated product was indeed formed (Figure 4.11). Having already characterized other cyclized Aza-CyWP products as reference points, the major difference observed in the CdpNPT product was the terminal alkene signal observed via HSQC (C1'/H1' = 114.4/5.04-5.11 ppm). The azaindole C-2 chemical shifts for **7-R-C3** observed via HSQC (C1'/H1' = 73.1/5.42 ppm) were analogous to that observed in the cyclized products (**4-C3**, **5-C3**, **6-C3**) isolated from FtmPT1 scale up reactions. The minor N1-prenylated product **7-N1** was also observed, which was also found as a product of FtmPT1 catalyzed prenylation of **7f**. Assignment of C2 and C3 chiral centers was done using 1D-ROESY experiments, which indicated close spatial proximity between H2 and the prenyl methyl groups.

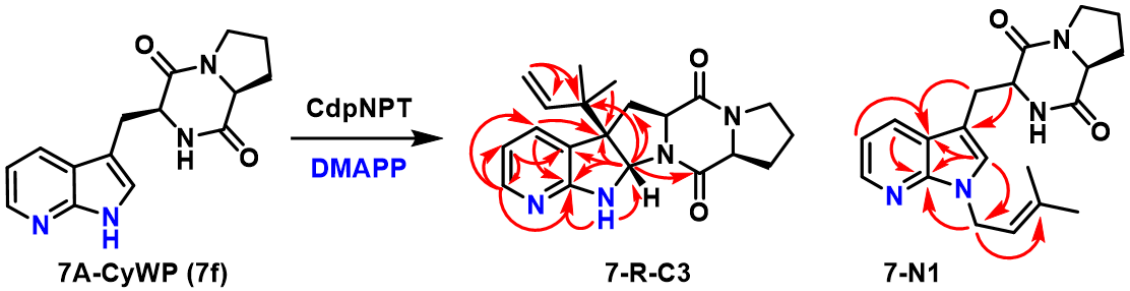


Figure 4.11 NMR characterization of the chemoenzymatically produced products of 7A-CyWP in the presence of CdpNPT and DMAPP

4.3.6 Mechanistic Discussion of Aza-CyWP Prenylation Regioselectivity

The products isolated from FtmPT1 catalyzed prenylation are an excellent data set for mechanistic deduction, since every Aza-CyWP isomer was accepted but produced

different product distributions. Surprisingly, the only azaindole isomer that undergoes the native 2-prenylation reaction by FtmPT1 was **7f**, as a minor product. Within the context of FtmPT1, the **5f** and **6f** analogs have the highest basicity and are the only isomers that produced C3 prenylated products with high regioselectivity. The C3 alkylated product was present as a minor product in the case of **4f**, and not detected for the 7-aza substrate. Compared to the native CyWP substrate, **7f** is the most electronically similar, allowing it to produce the native-type C2 prenylated product. The less electronically similar azaindole isomers (**4f**, **5f**, **6f**) favor the generation of the C3 prenylated cyclized product, with increasing C3 selectivity as the azaindole pKa increases (**5a**: 8.3 ≈ **6a**: 7.9 > **4a**: 6.9 >> **7a**: 4.6). Another parameter to consider is the electron density at the C2 position, which based on ^{13}C chemical shifts is known that C2 electron density follows the trend of **7f** > Indole > **5f** > **4f** > **6f**.[39] While this electron density trend suggests that **7f** should more readily form C2 prenylated products, the relative nucleophilicity of the N1 position must be taken into consideration as a competing prenyl group acceptor in close proximity to C2, resulting in the major product formed from **7f** by FtmPT1 being N1-prenylated. This rational derived from C2 electron density assumes that the dimethylallyl carbocation is being directly attacked by C2.

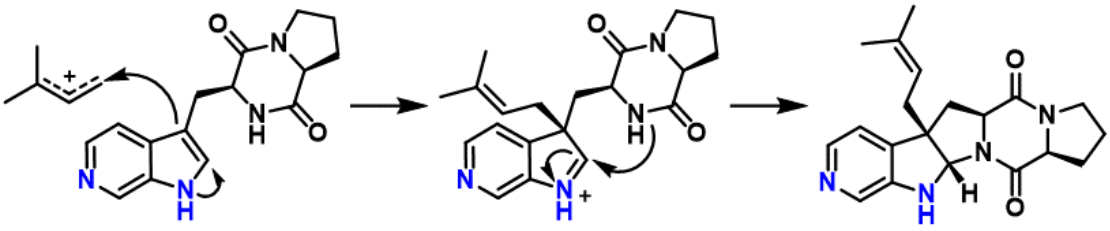


Figure 4.12 Proposed mechanism for the formation of C3 prenylated cyclic products.

4.3.7 Comparison to Previous Mechanistic Studies of FtmPT1

Previously published mechanistic studies suggest that the native C2 prenylation reaction is potentially derived from an initial C3 reverse prenylation and subsequent cope rearrangement to C2.[26, 40] A similar rearrangement could also transfer the C3 reverse-prenyl group to the N1 position, which is the proposed mechanism by which the native CdpNPT reaction generates a minor N1 prenylated product and a C3 reverse prenylated major product. The rearrangement mechanism depends on the relative speed at which the dearomatized C3 prenylated intermediate can cyclize on the transiently formed nucleophilic iminium, versus rearrangement of the prenyl group to restore aromaticity. Crystallographic studies of FtmPT1 have indicated that the C3` position of DMAPP was held near the indole C2 and C3 positions, with the C1` of DMAPP positioned over the 6-membered ring of the indole substrate.[41] While these proximity-based arguments suggest that reverse prenylation would occur, the steric environment of the reaction chamber of FtmPT1 contains a residue sidechain Y435 that would clash with the bulky quaternary carbon produced by reverse prenylation.[41] Therefore, consistent with the native reaction of FtmPT1 and the observed normal prenylated Aza-CyWP products, these reverse prenylation/rearrangement-based mechanisms are unlikely.

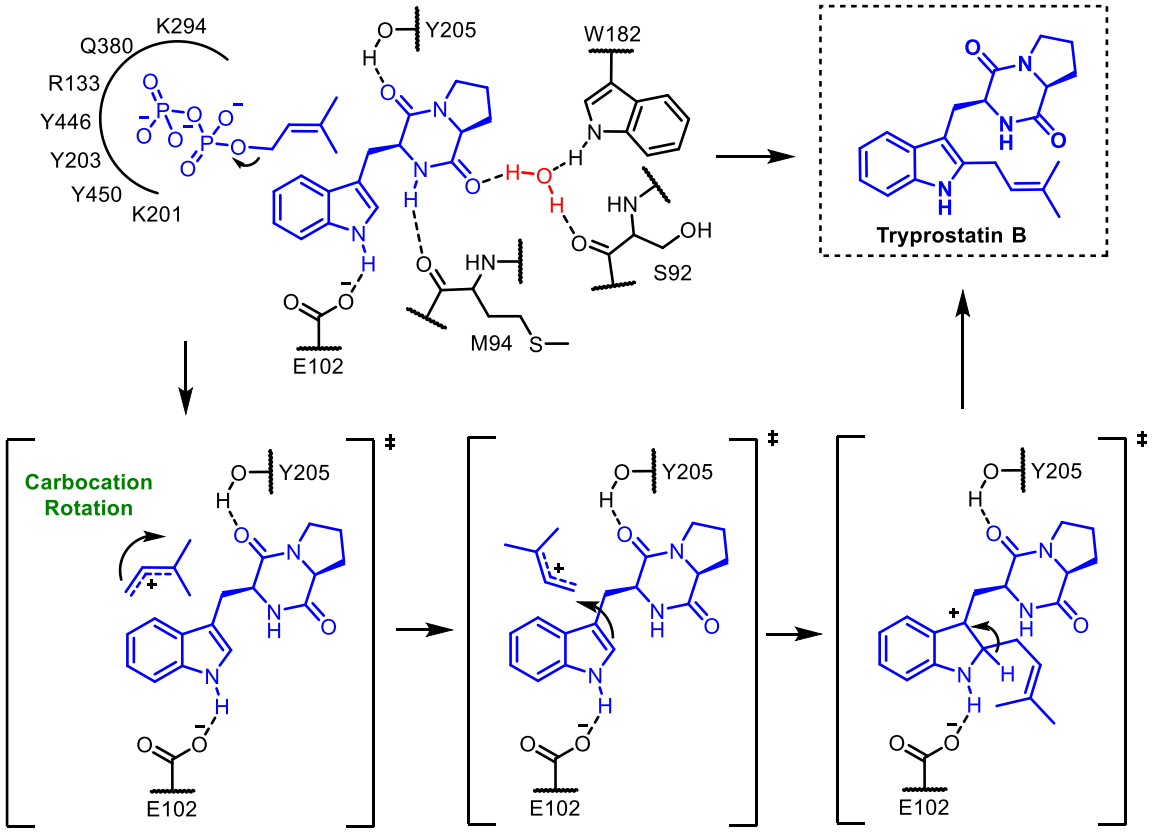


Figure 4.13: Proposed mechanism for FtmPT1 catalyzed C2 prenylation

The planar dimethylallyl carbocation is formed in a solvent-shielded reaction chamber, and further stabilized by π-stacking interactions on either side. Based on the known kinetic parameters of FtmPT1, rotation of the carbocation can easily occur within the time-scale of this reaction (Figure 4.13). Kinetic isotope experiments have revealed that two partially rate-limiting steps exist, the phosphate ester hydrolysis, and the C-C bond formation. The timescale of carbocation rotation is orders of magnitude faster than these partially rate limiting steps, making a strong argument for the carbocation-rotation based mechanism. Once the carbocation is generated, it rotates above the aromatic acceptor until C1` undergoes electrophilic attack of the azaindole ring on C2, C3 or N1.

The steric restraints of the binding pocket have a strong influence on reaction regiospecificity, where the additional nitrogen atoms can participate in new interactions,

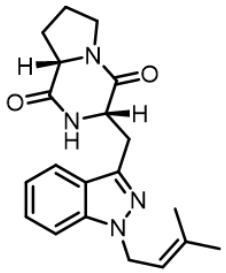
thus altering the binding conformation relative to the native substrate. Additionally, the rate at which cyclization occurs after C3 prenylation likely differs between the azaindole isomers, partially due to the variable nucleophilicity of the C2 iminium intermediate (Figure 4.12, center), but also the enzyme-induced conformation of this reaction intermediate. This combination of electronic and steric effects, keeping in mind the free rotation of the dimethylallyl carbocation, is likely responsible for the observed tendency to prenylate the nucleophilic N1 or C3 positions rather than C2. The results of this study serve as unique mechanistic probe to help further elucidate and control the catalytic action of aromatic PTs.

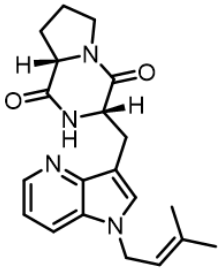
4.4 Conclusions and Future Directions

This study represents a major step for PT-based biocatalysts, having successfully demonstrated that both native CdpNPT as well as FtmPT1 are capable of prenylating azaindole containing substrate-mimetics, revealing substrate dependent regiospecificity. This study demonstrates the possibilities of substrate directed biocatalytic late-stage diversification of bioactive NPs using promiscuous indole PTs. The electronic differences between the azaindole isomers effected the prenylation regiospecificity, providing a total of 4 characterized C3 prenylated cyclized products, 3 characterized N1 prenylated products, and a single C2 prenylated product. Given that the native enzyme FtmPT1 was able to prenylate all 5 azaindole scaffolds screened, it is possible that through protein engineering methods that efficient and regioselective alkylating enzymes could be developed based on the remarkable catalytic scaffolding provided by aromatic PTs. CdpNPT for example is known to accept many unnatural acceptor and donor substrates. In the context of pharmaceutical SAR studies, a diverse library of drug analogs could be

produced via late-stage diversification using a library of donor pyrophosphates, and appropriate PTs. The outstanding regioselectivity and chemoselectivity of PTs, combined with their broad substrate scope, makes then excellent candidates for developing facile methods of late-stage aromatic ring functionalization.

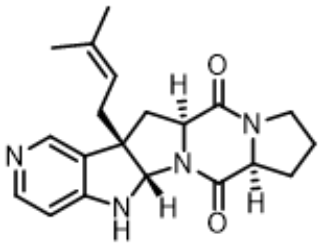
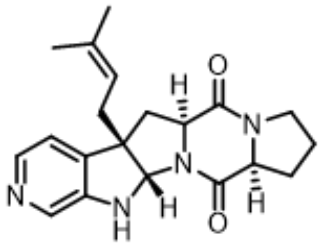
4.5 Appendix 1: NMR Assignments and Spectra





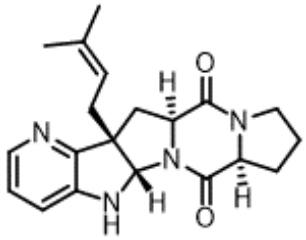
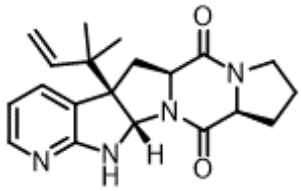
		2-N1 (d_6-DMSO)		4-N1 (d_6-DMSO)	
Segment	**Position**	**δC , type**	**δH**	**δC , type**	**δH**
L-Trp	NH		7.85 (s)		8.19 (s)
	α	53.6, CH	4.54 (t)	55.4, CH	4.51 (t)
	β	27.1, CH_2	3.17 (dd), 3.52 (dd)	25.5, CH_2	3.22 (dd), 3.47 (dd)
	1				
	2			142.7, CH	8.10 (s)
	3	140.5, C		104.4, C	
	3a	122.6, C		137.7, C	
	4	120.4, CH	7.77 (d)		
	5	119.6, CH	7.10 (t)	127.4, CH	8.39 (s)
	6	126.1, CH	7.36 (t)	113.8, CH	7.37 (t)
	7	109.4, CH	7.53 (d)	135.7, CH	8.37 (s)
	7a	139.6, C		138.7, C	
	C=O	165.1, C		165.6, C	
Proline	α	58.3, CH	4.17 (s)	58.5, CH	4.16 (m)
	β	27.6, CH_2	1.65 (m), 2.07 (ddt)	27.4, CH_2	1.76 (m), 2.10 (m)
	γ	21.9, CH_2	1.75 (m)	22.1, CH_2	1.76 (m)
	δ	44.6, CH_2	3.29 (m), 3.40 (m)	44.8, CH_2	3.31 (m)
	C=O				
Prenyl	1'	46.1, CH_2	4.94 (d)	54.1, CH_2	5.42 (m), 5.56 (dd)
	2'	120.0, CH	5.30 (t)	118.5, CH	5.40 (m)
	3'	135.2, C		135.4, C	
	4'	25.2, CH_3	1.67 (s)	25.3, CH_3	1.75 (s)
	5'	17.8, CH_3	1.81 (s)	18.3, CH_3	1.82 (s)

Table 4.1: NMR assignments for 2-N1 and 4-N1

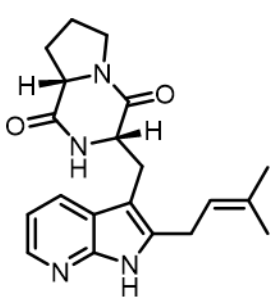
		5-C3 (d_6-DMSO)		6-C3 (d_6-DMSO)	
Segment	**Position**	**δC , type**	**δH**	**δC , type**	**δH**
L-Trp	NH				
	α	59.7, CH	4.13 (dd)	59.4, CH	4.08 (dd)
	β	36.4, CH2	2.18 (t), 2.54 (dd)	37.6, CH_2	2.18 (t), 2.53 (m)
	1		7.50 (s)		6.78 (s)
	2	78.3, CH	5.23 (s)	77.9, CH	5.22 (s)
	3	54.5, C		55.2, C	
	3a	127.7, C		139.9, C	
	4	144.2, CH	8.13 (s)	118.5, CH	7.20 (d)
	5			139.2, CH	7.88 (d)
	6	149.8, CH	7.99 (d)		
	7	104.0, CH	6.43 (d)	130.2, CH	7.85 (s)
	7a	156.2, C		146.2, C	
	C=O	166.1, C		165.2, C	
Proline	α	60.3, CH	4.26 (t)	59.3, CH	4.25 (t)
	β	27.9, CH_2	1.90 (m), 2.12 (m)	27.2, CH_2	1.89 (m), 2.11 (m)
	γ	22.82, CH_2	1.79 (m), 1.89 (m)	22.5, CH_2	1.79 (m), 1.86 (m)
	δ	45.3, CH_2	3.34 (m)	44.6, CH_2	3.33 (m)
	C=O			168.78, C	
Prenyl	1'	36.4, CH_2	2.37 (d)	34.9, CH_2	2.36 (d)
	2'	119.4, CH	5.08 (t)	118.7, CH	5.05 (t)
	3'	134.8, C		134.2, C	
	4'	26.0, CH_3	1.63 (s)	25.5, CH_3	1.62 (s)
	5'	18.3, CH_3	1.46 (s)	17.6, CH_3	1.46 (s)

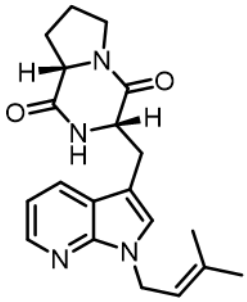
Table 4.2: NMR assignments for 5-C3 and 6-C3

		4-C3 (d_6-DMSO)		7-R-C3 (d_6-DMSO)	
Segment	**Position**	**δC , type**	**δH**	**δC , type**	**δH**
L-Trp	NH				
	α	59.1, CH	4.05 (dd)	59.2, CH	4.07 (dd)
	β	36.8, CH_2	2.16 (m), 2.53 (m)	35.6, CH_2	2.26 (dd), 2.38 (dd)
	1		6.81 (s)		7.27 (s)
	2	76.2, CH	5.26 (s)	73.1, CH	5.42 (s)
	3	55.5, C		59.5, C	
	3a	153.4, C		122.7, C	
	4			132.2, CH	7.49 (dd)
	5	138.3, CH	7.78 (d)	113.0, CH	6.58 (dd)
	6	122.6, CH	6.93 (dd)	147.2, CH	7.82 (dd)
	7	113.9, CH	6.79 (d)		
	7a	144.0, C		162.8, C	
	C=O	165.7, C		165.1, C	
Proline	α	59.5, CH	4.24 (t)	59.7, CH	4.25 (t)
	β	27.2, CH_2	1.90 (m), 2.12 (m)	27.3, CH_2	1.89 (m), 2.13 (dt)
	γ	22.5, CH_2	1.80 (m), 1.89 (m)	22.3, CH_2	1.79 (m), 1.89 (m)
	δ	44.6, CH_2	3.34 (m)	44.6, CH_2	3.32-3.40 (m)
	C=O			165.7, C	
3-Prenyl	1'	34.5, CH_2	2.38 (t)	114.4, CH_2	5.04 (dd), 5.11 (dd)
	2'	118.9, CH	5.04 (t)	143.3, CH	5.96 (dd)
	3'	134.4, C		40.8, C	
	4'	25.7, CH_3	1.59 (s)	21.8, CH_3	1.01 (s)
	5'	17.6, CH_3	1.50 (s)	22.3, CH_3	0.89 (s)

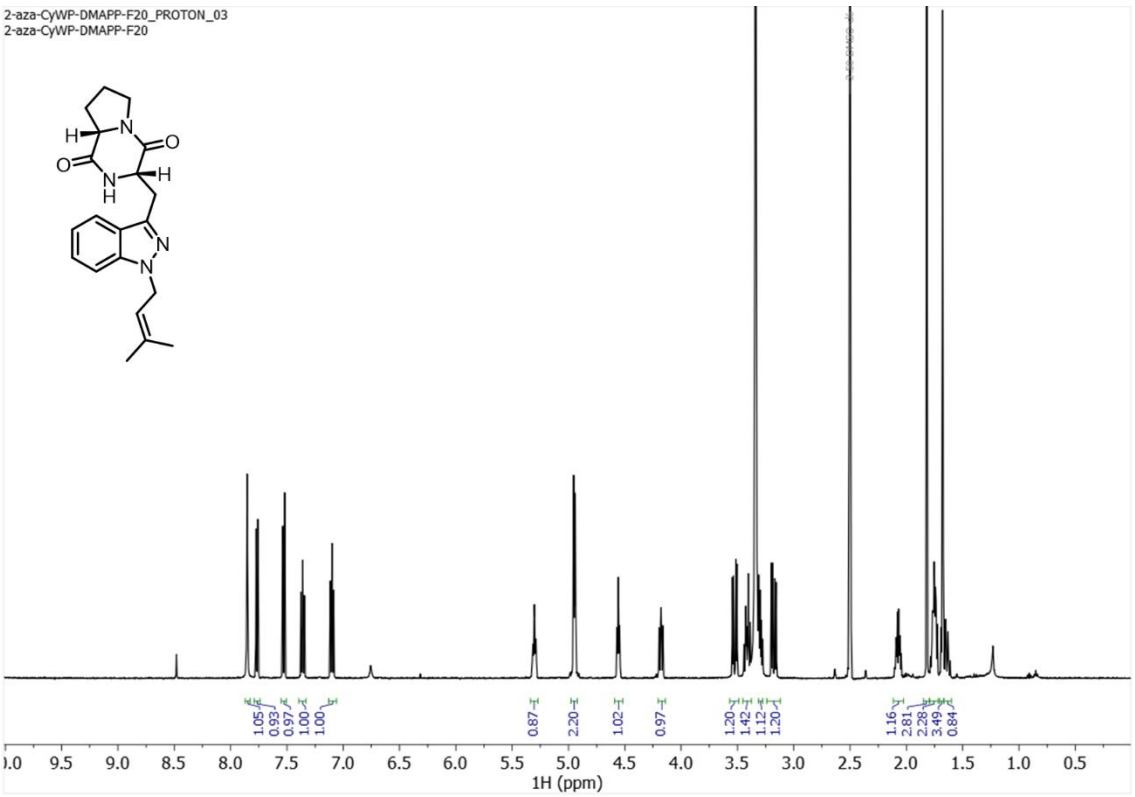
Table 4.3: NMR assignments for 4-C3 and 7-R-C3



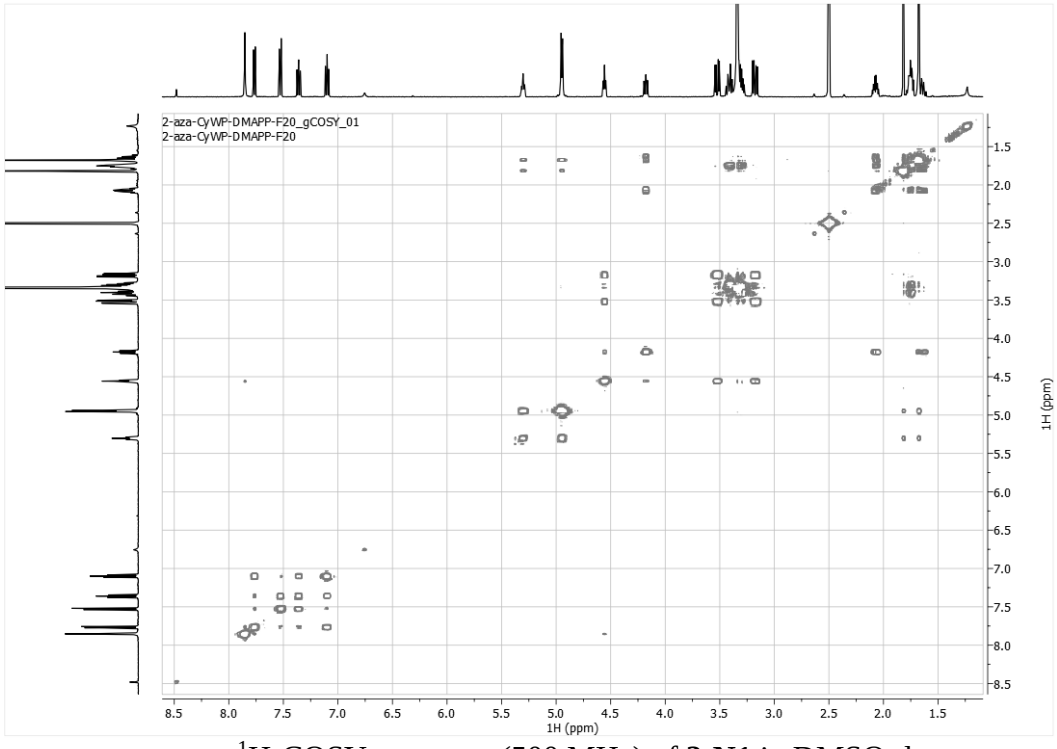


		7-C2 (d_6-DMSO)		7-N1 (d_6-DMSO)	
Segment	**Position**	**δC , type**	**δH**	**δC , type**	**δH**
L-Trp	NH		7.57 (s)		7.98 (s)
	α	55.9, CH	4.24 (t)	55.8, CH	4.31 (t)
	β	26.7, CH_2	3.08 (m)	26.5, CH_2	3.13 (t)
	1		11.29 (s)		
	2			127.6, CH	7.26 (s)
	3			108.0, C	
	3a			120.5, C	
	4	126.2, CH	7.84 (d)	127.8, CH	8.01 (dd)
	5	114.6, CH	6.93 (dd)	115.4, CH	7.04 (dd)
	6	141.2, CH	8.05 (dd)	142.6, CH	8.20 (dd)
	7				
	7a			147.0, CH	
	C=O			165.6, C	
Proline	α	58.8, CH	3.94 (dd)	58.8, CH	4.02 (dd)
	β	28.1, CH_2	0.97 (q), 1.85 (m)	28.4, CH_2	1.24 (m), 1.95 (m)
	γ	21.9, CH_2	1.32 (m), 1.57 (m)	22.1, CH_2	1.59 (m)
	δ	44.3, CH_2	3.10 (m), 3.32 (m)	45.0, CH_2	3.23 (m), 3.32 (m)
	C=O				
3-Prenyl	1'	24.8, CH_2	3.43 (m)	41.5, CH_2	4.78 (dd)
	2'	121.2, CH	5.30 (t)	120.9, CH	5.33 (t)
	3'			136.2, C	
	4'	25.9, CH_3	1.66 (s)	25.8, CH_3	1.69 (s)
	5'	18.3, CH_3	1.70 (s)	18.1, CH_3	1.78 (s)

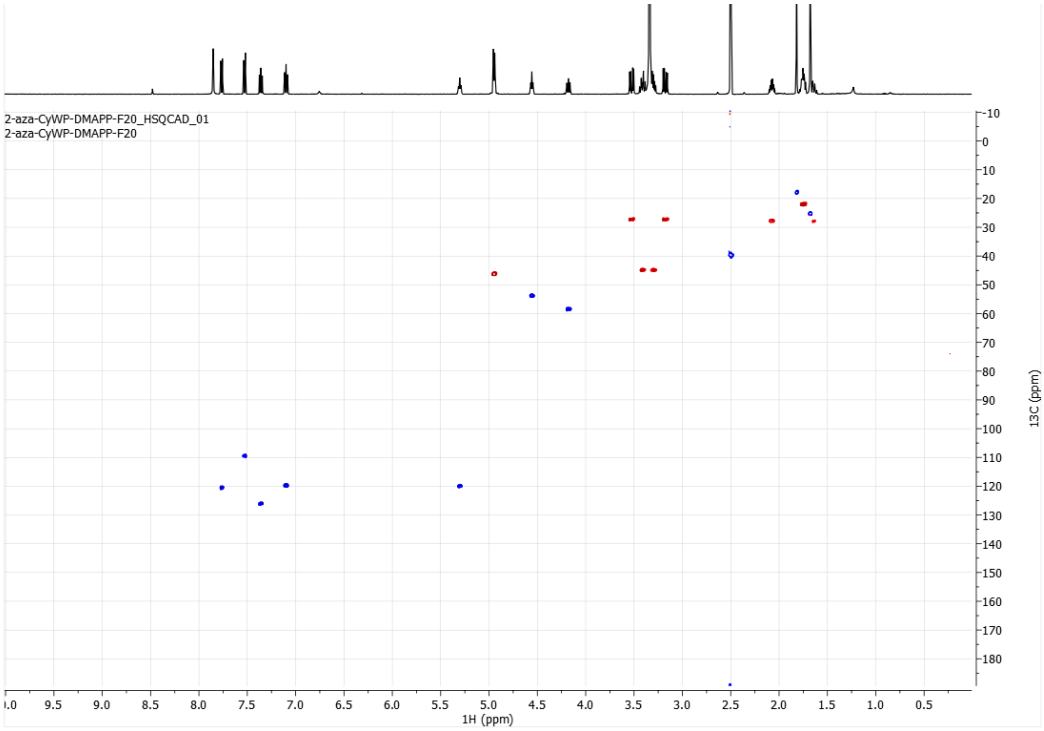
Table 4.4: NMR assignments for 7-C2 and 7-N1



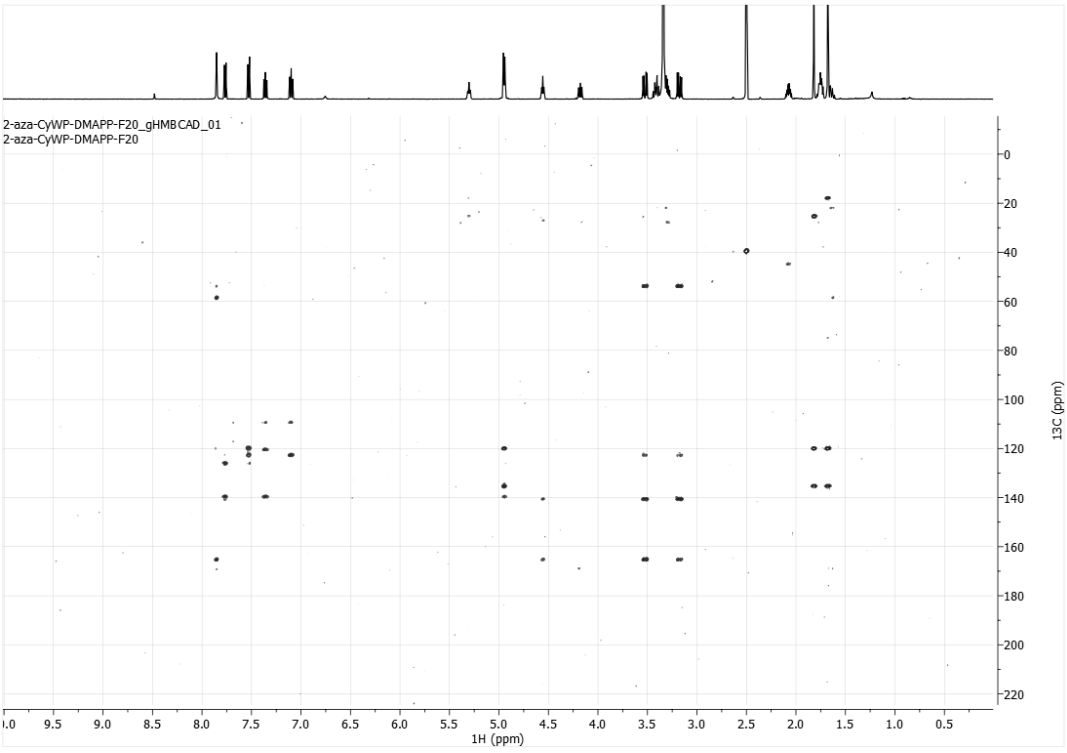
^{1}H NMR spectrum (500 MHz) of **2-N1** in DMSO-d_6



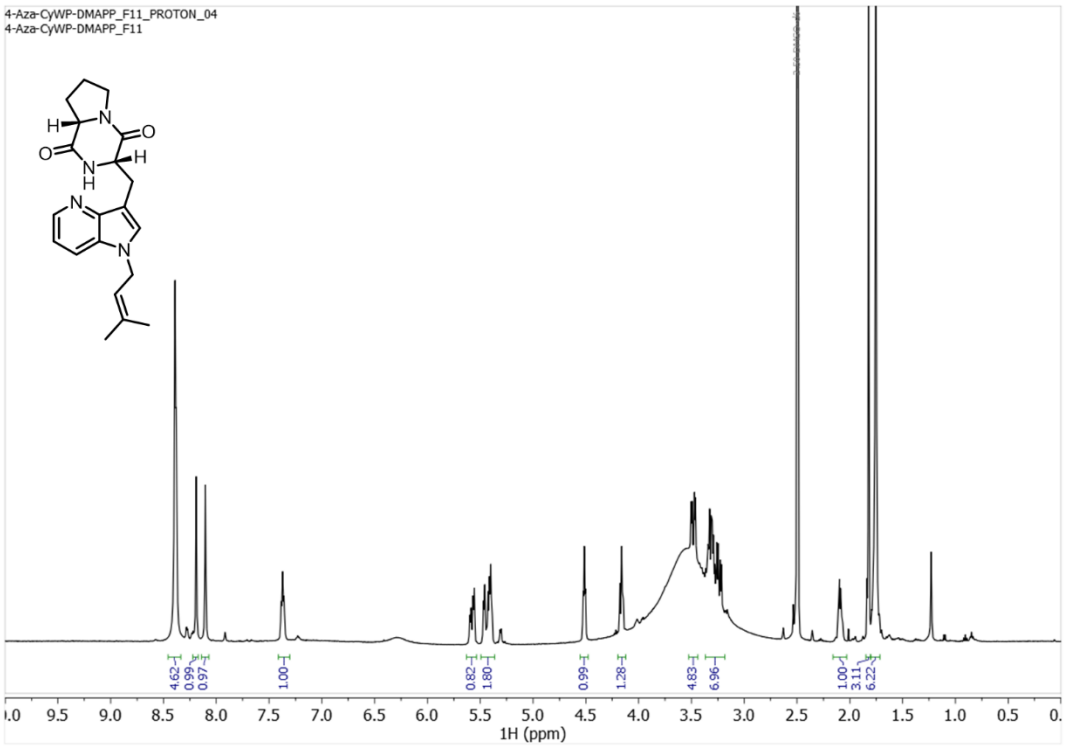
^{1}H-COSY spectrum (500 MHz) of **2-N1** in DMSO-d_6



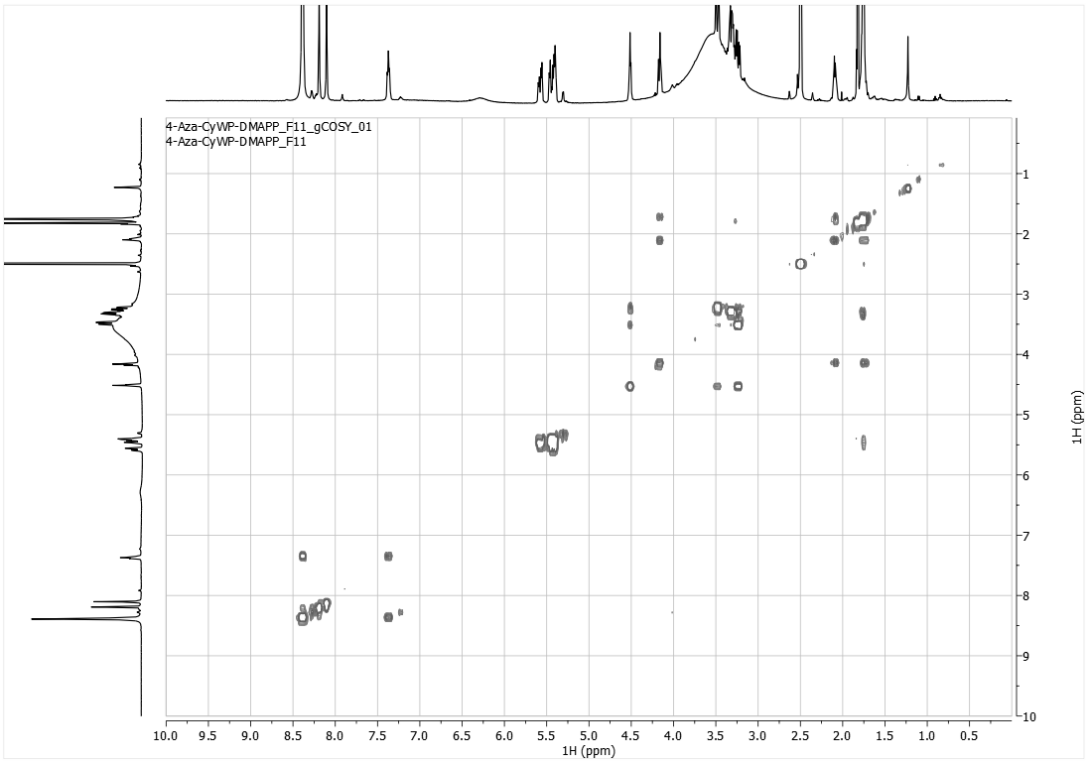
HSQC spectrum (500 MHz) of **2-N1** in DMSO-d_6



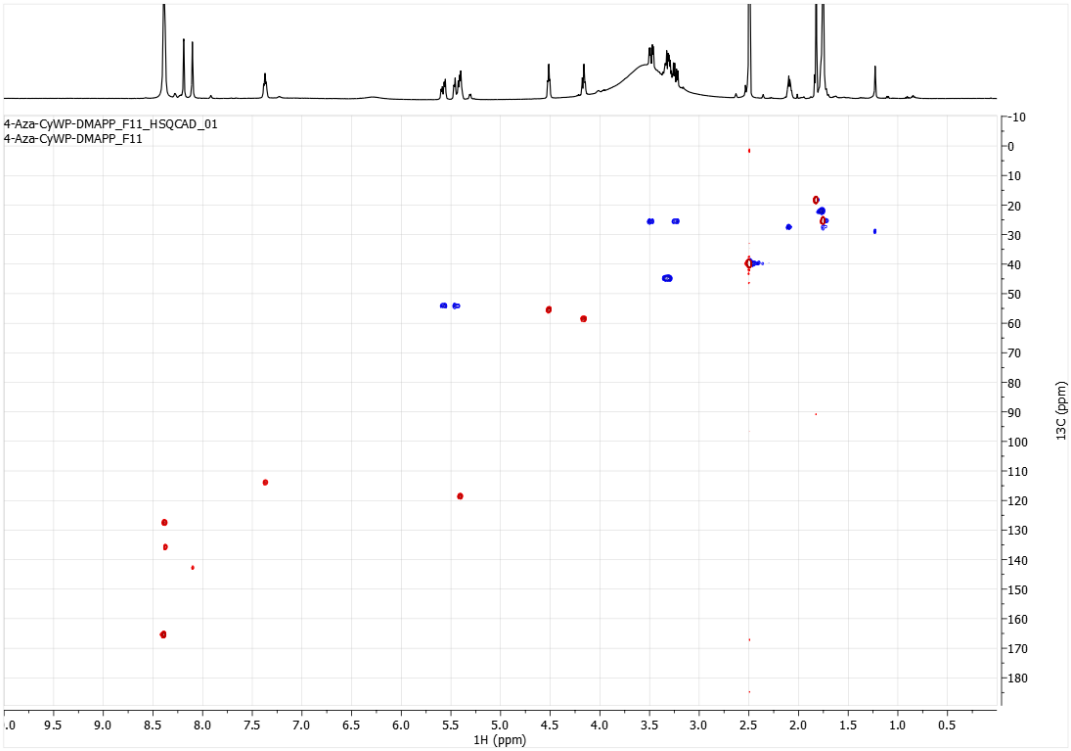
HMBC spectrum (500 MHz) of **2-N1** in DMSO-d_6



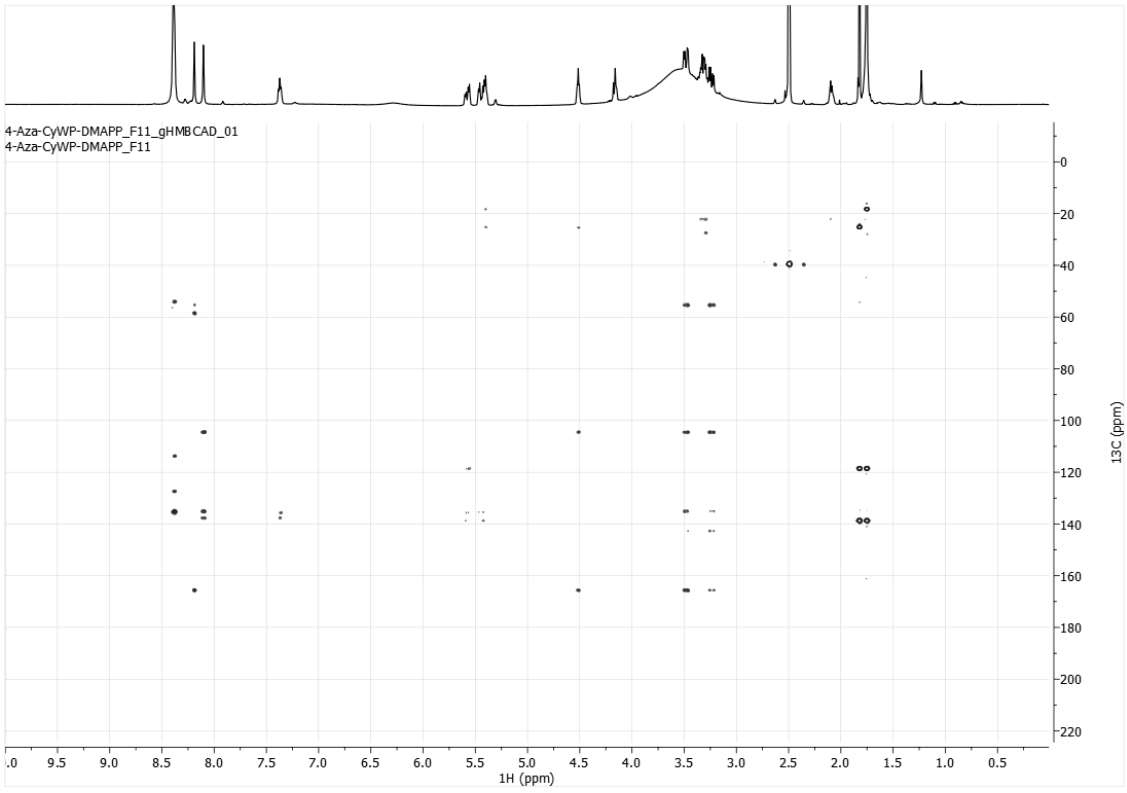
^{1}H NMR spectrum (500 MHz) of **4-N1** in DMSO-d_6



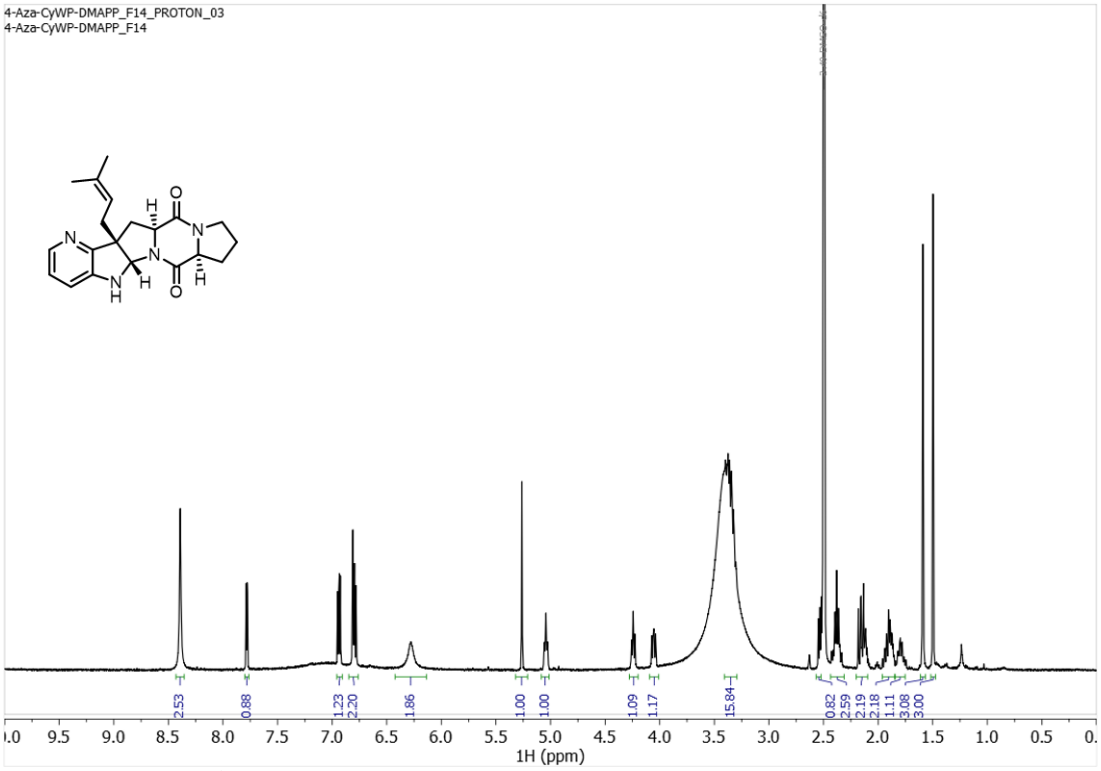
^{1}H-COSY spectrum (500 MHz) of **4-N1** in DMSO-d_6



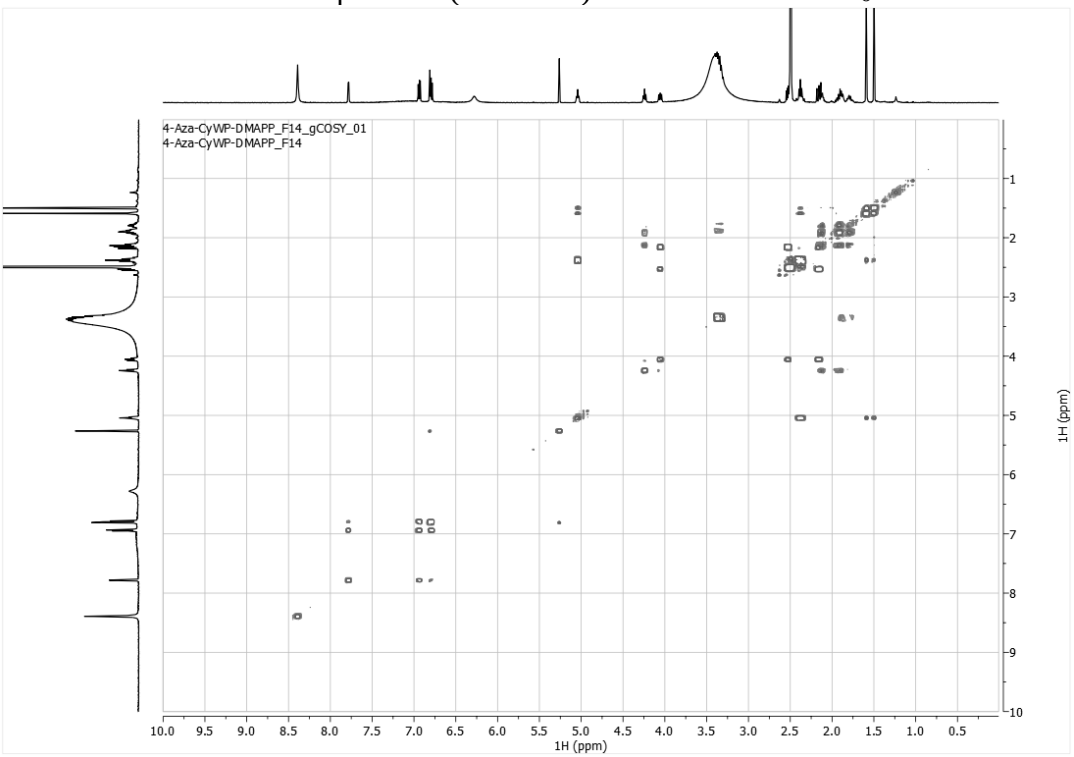
HSQC spectrum (500 MHz) of **4-N1** in DMSO-d_6



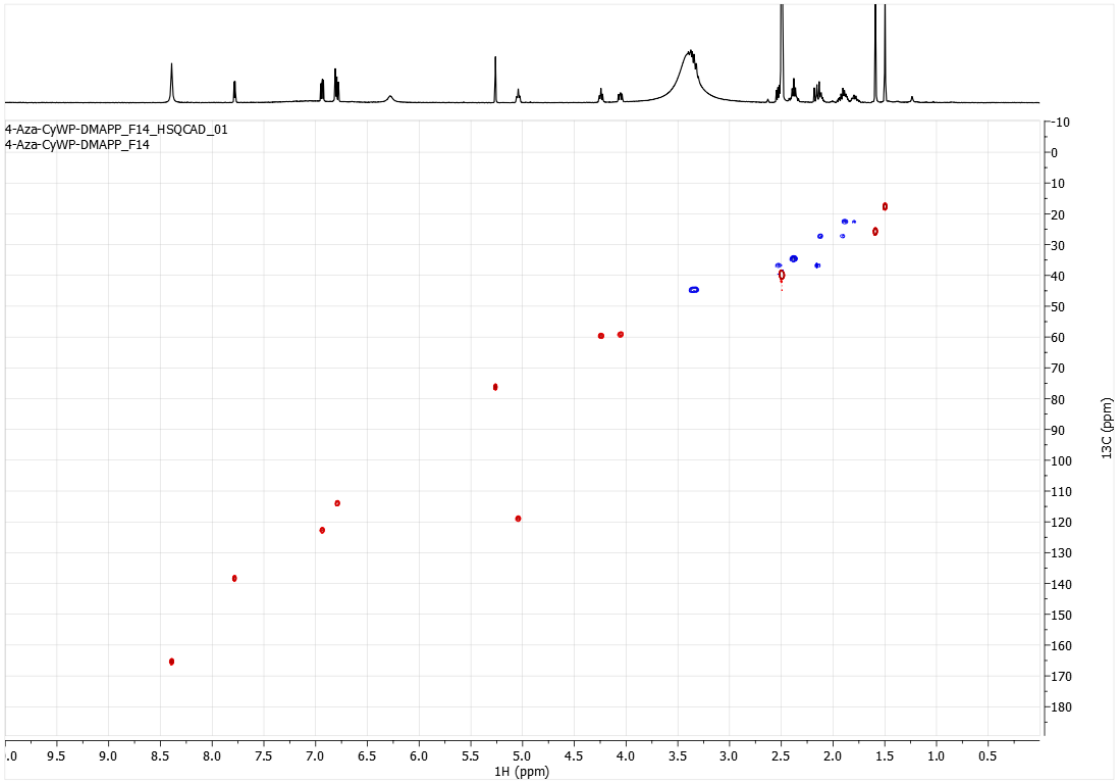
HMBC spectrum (500 MHz) of **4-N1** in DMSO-d_6



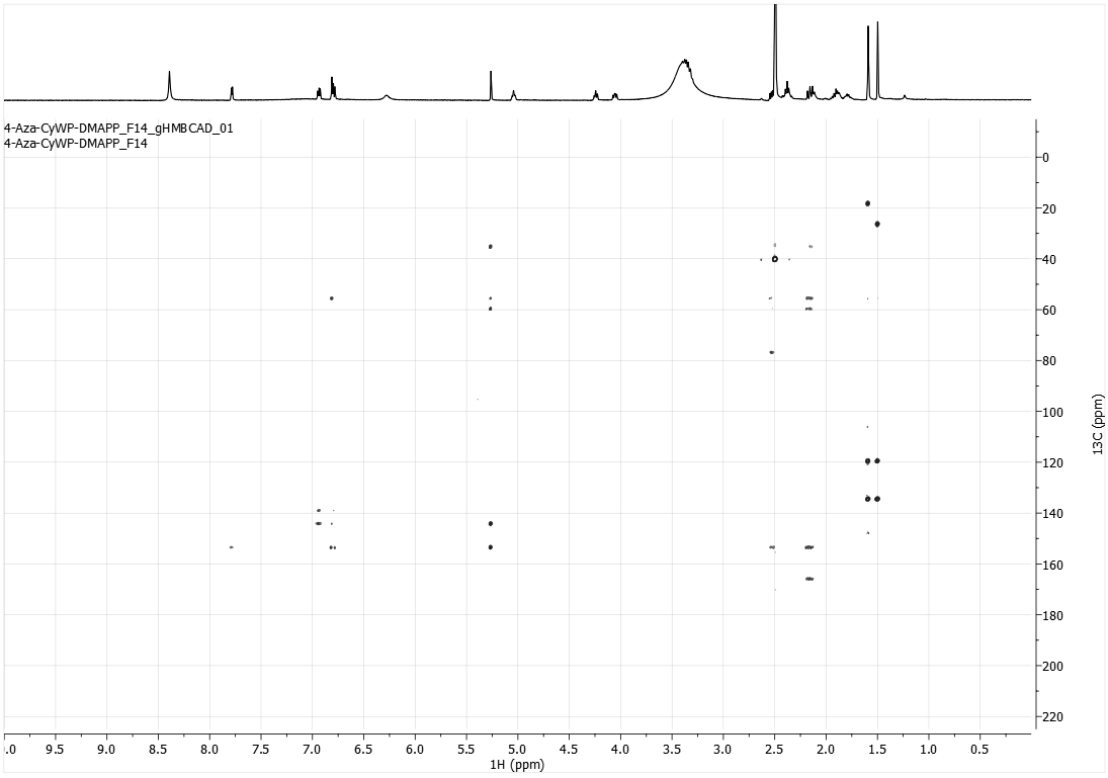
^{1}H NMR spectrum (500 MHz) of **4-C3** in DMSO-d_6



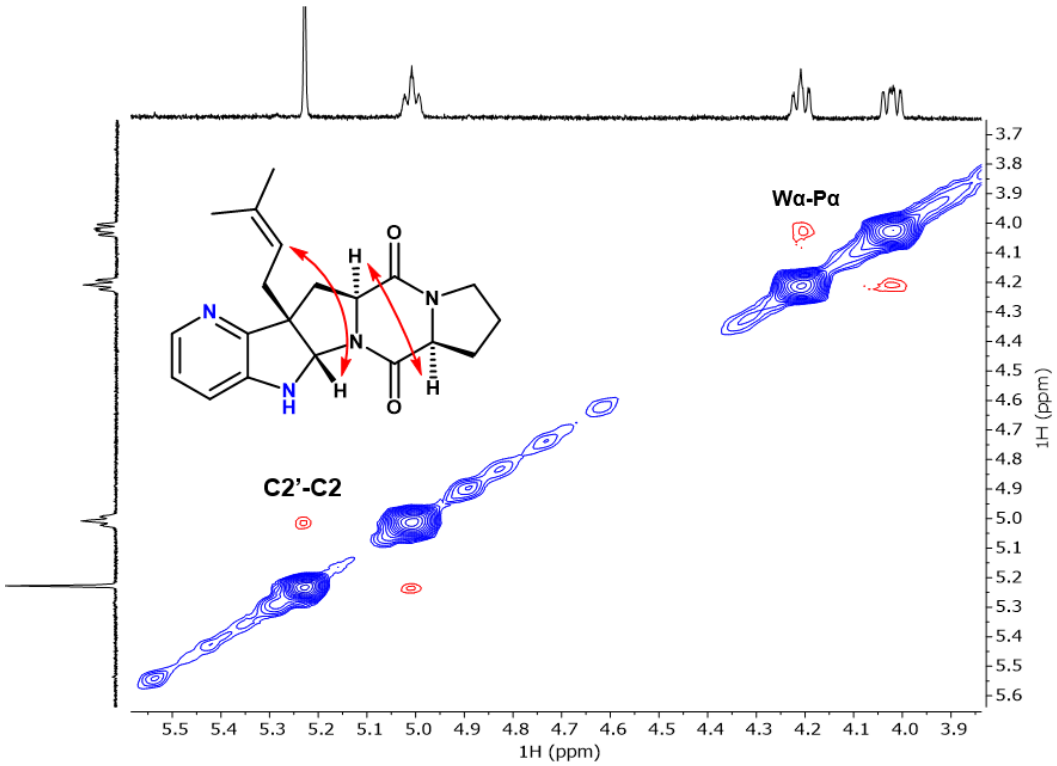
^{1}H-COSY spectrum (500 MHz) of **4-C3** in DMSO-d_6



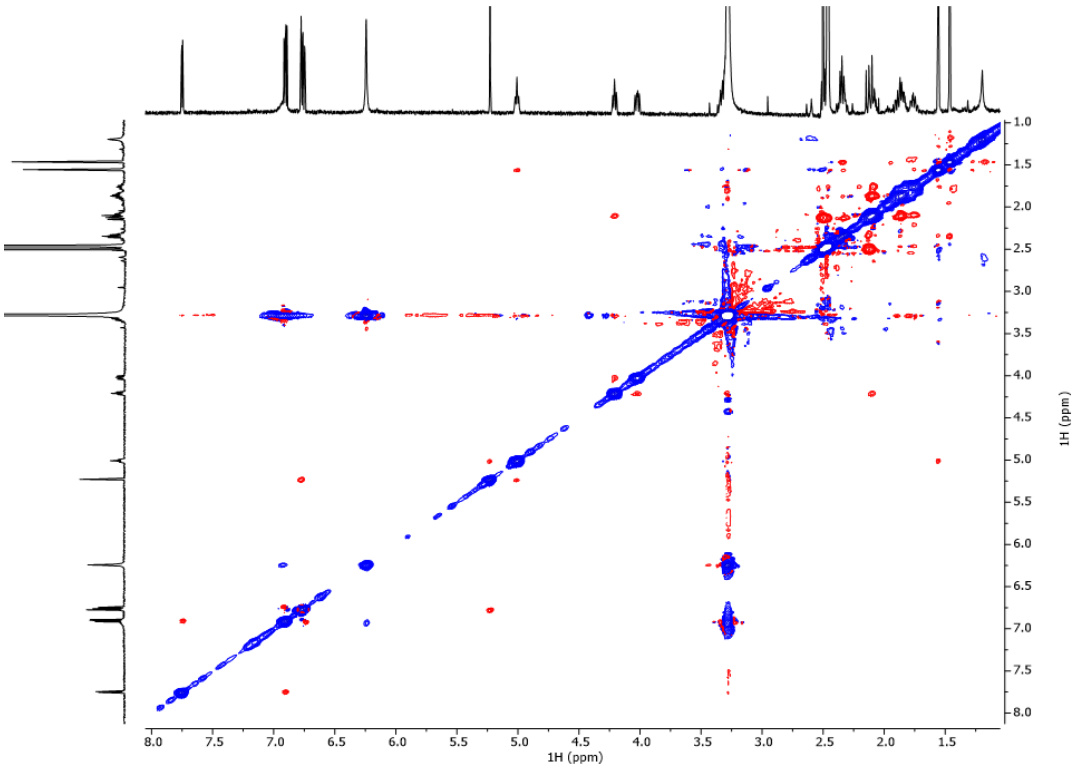
HSQC spectrum (500 MHz) of **4-C3** in DMSO-d_6



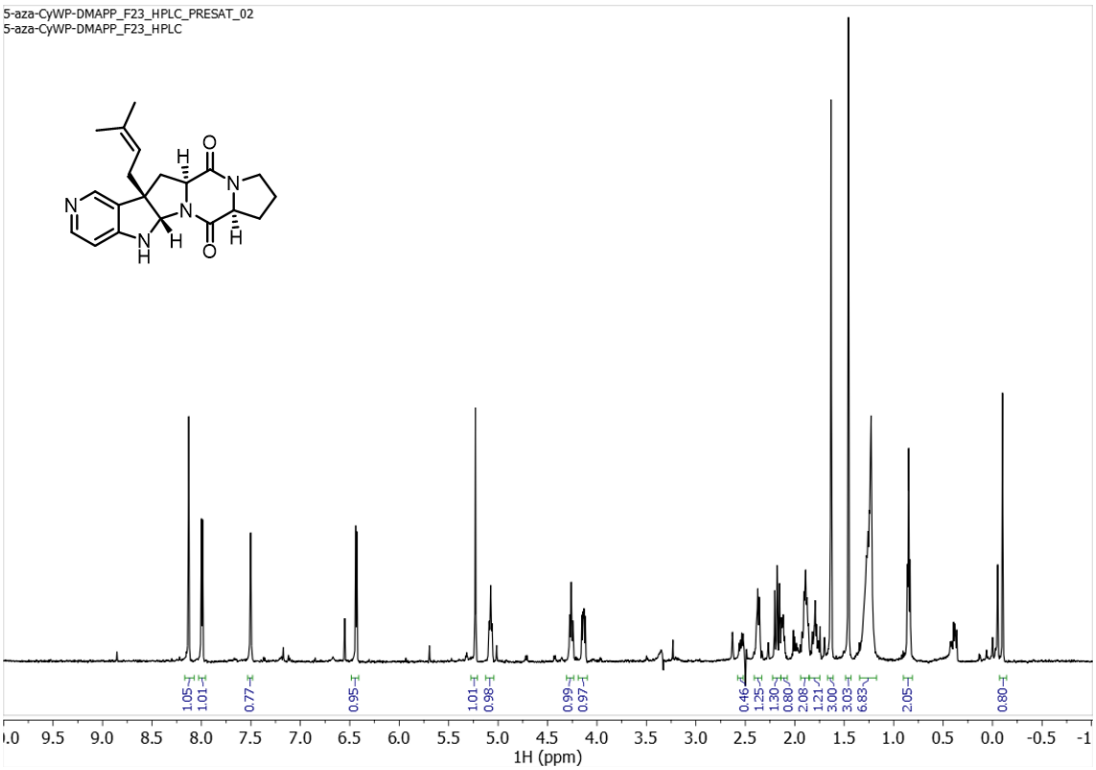
HMBC spectrum (500 MHz) of **4-C3** in DMSO-d_6



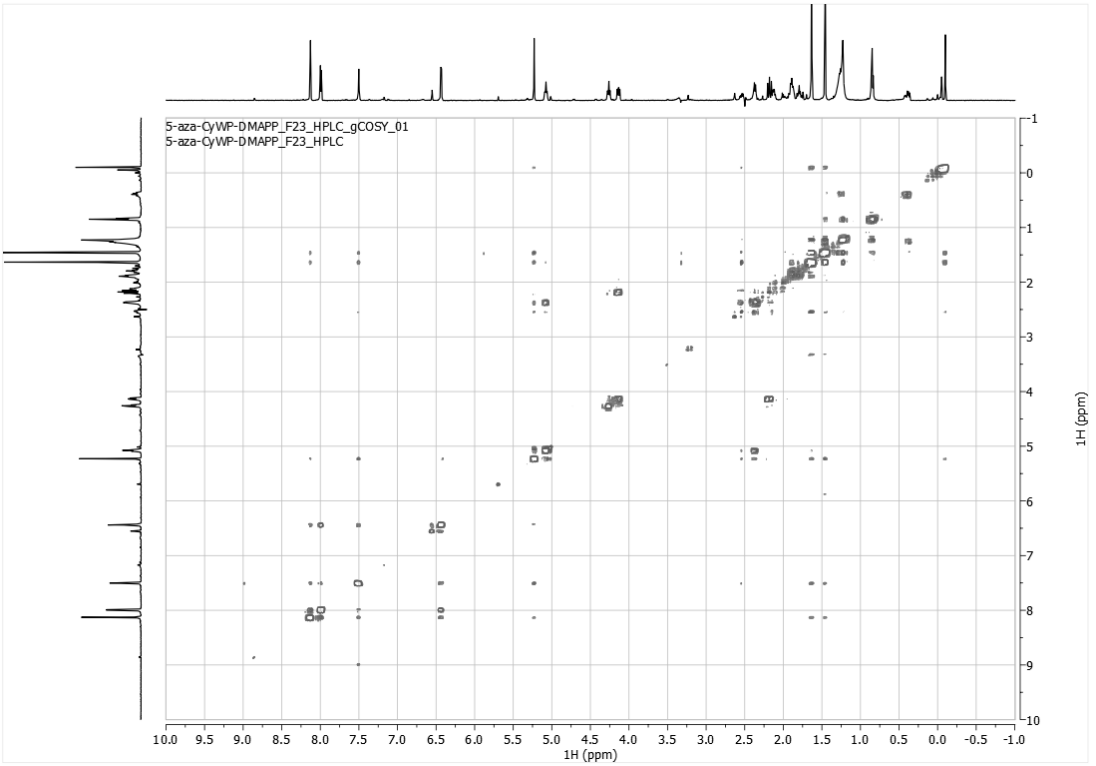
Zoomed in 2D-ROESY spectrum (500 MHz) of **4-C3,** with diagnostic peaks



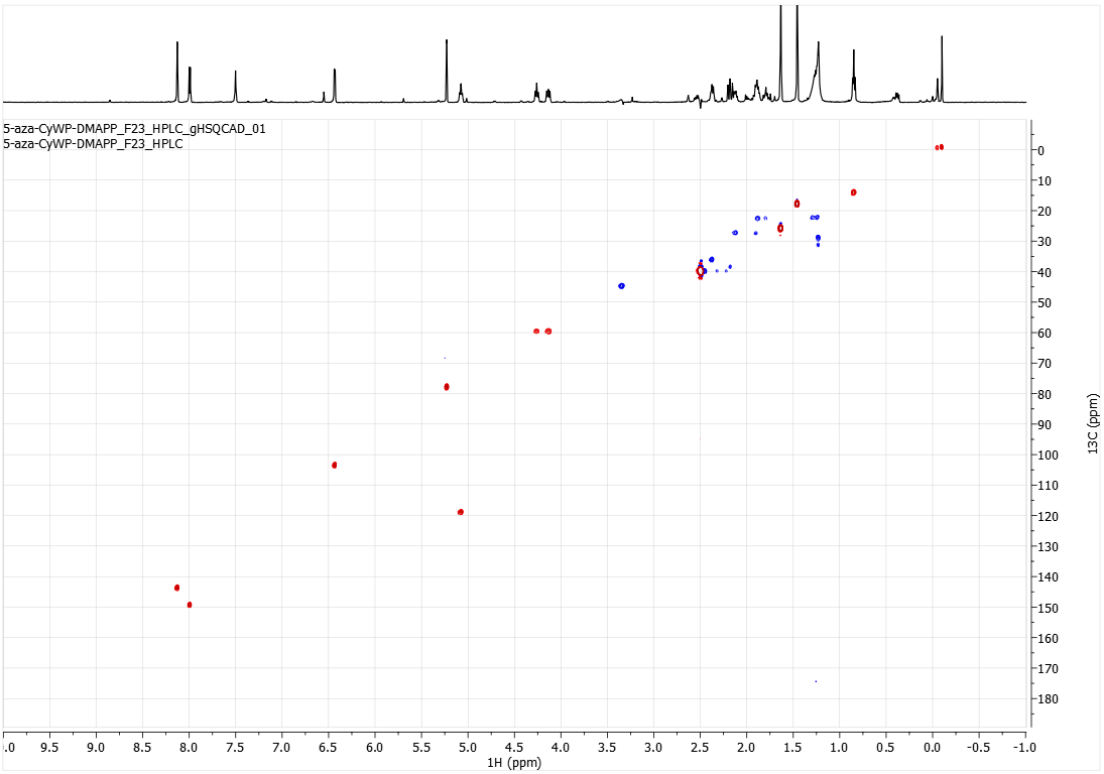
Full 2D-ROESY spectrum (500 MHz) of **4-C3**



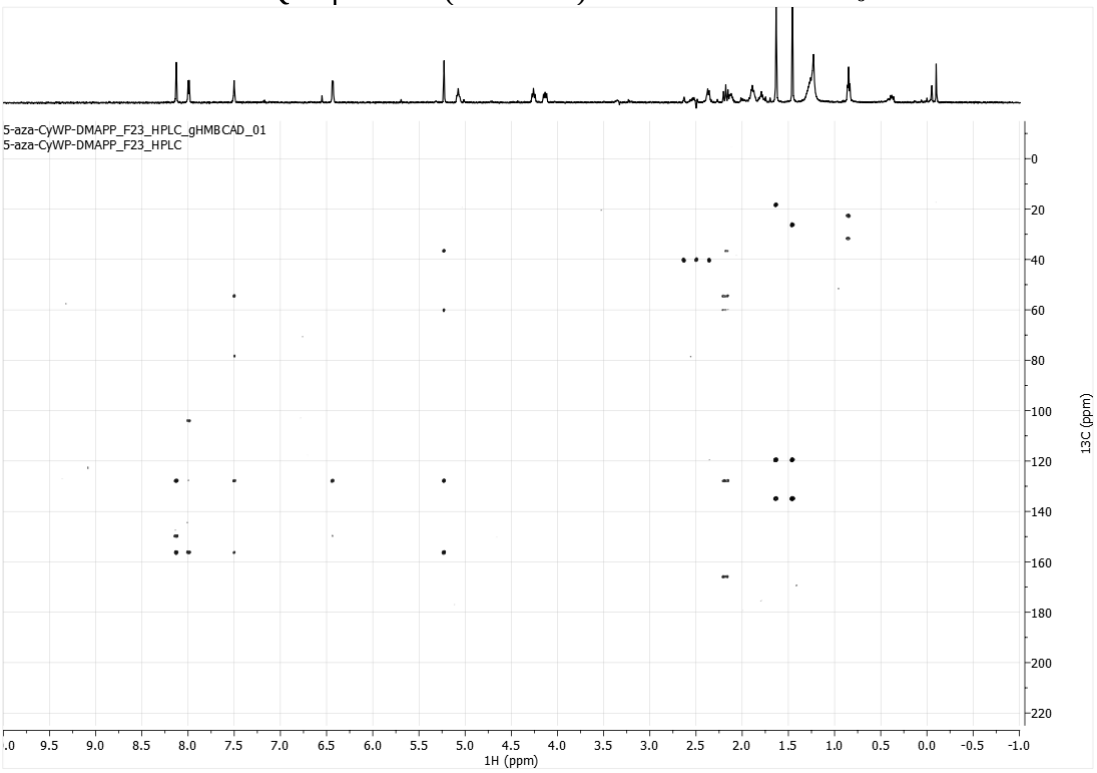
^{1}H NMR spectrum (500 MHz) of **5-C3** in DMSO-d$_6$



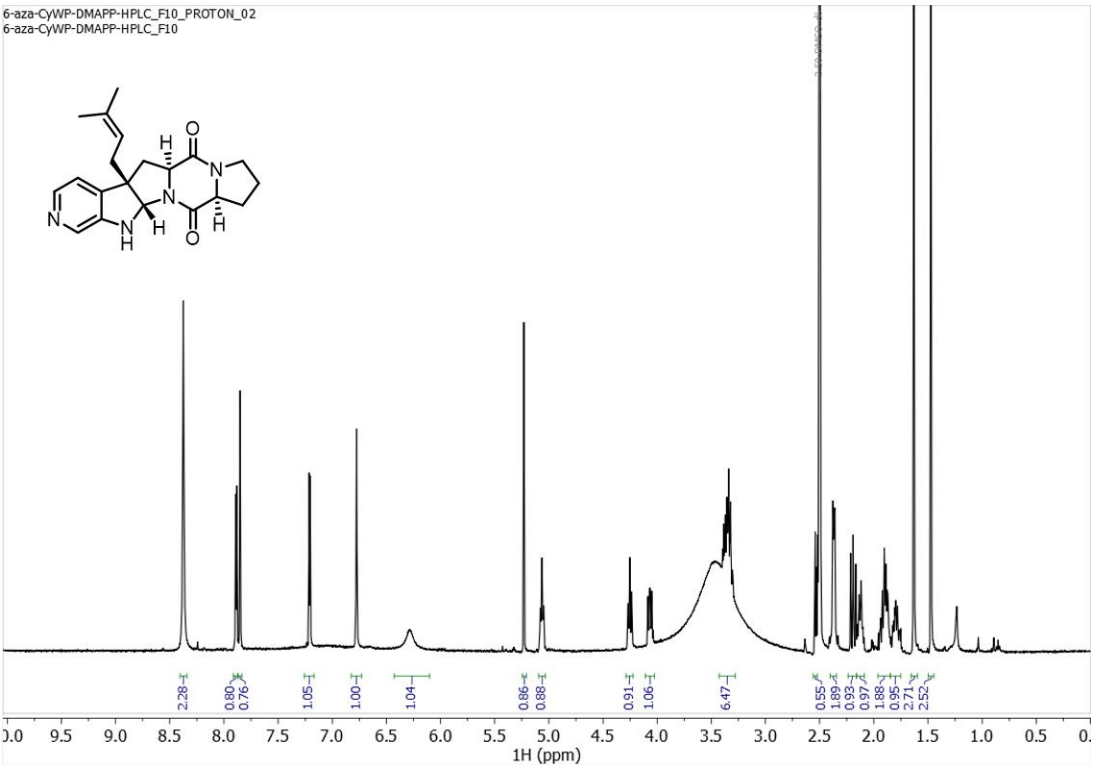
^{1}H-COSY spectrum (500 MHz) of **5-C3** in DMSO-d$_6$



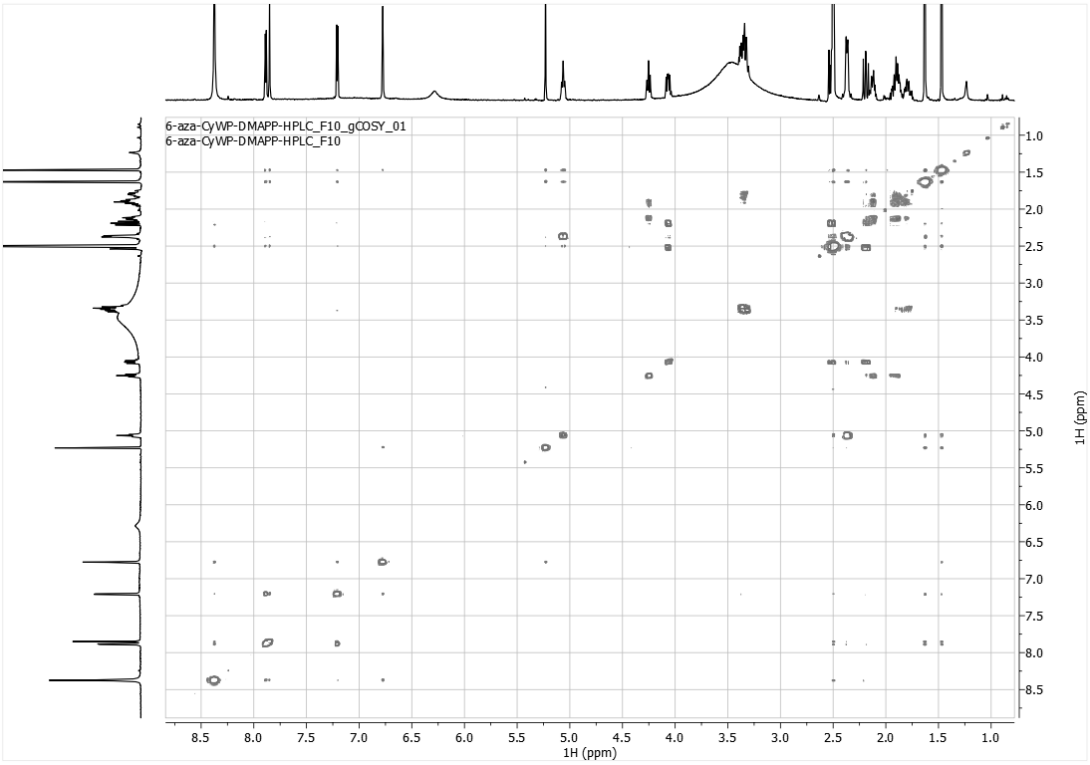
HSQC spectrum (500 MHz) of **5-C3** in DMSO-d_6



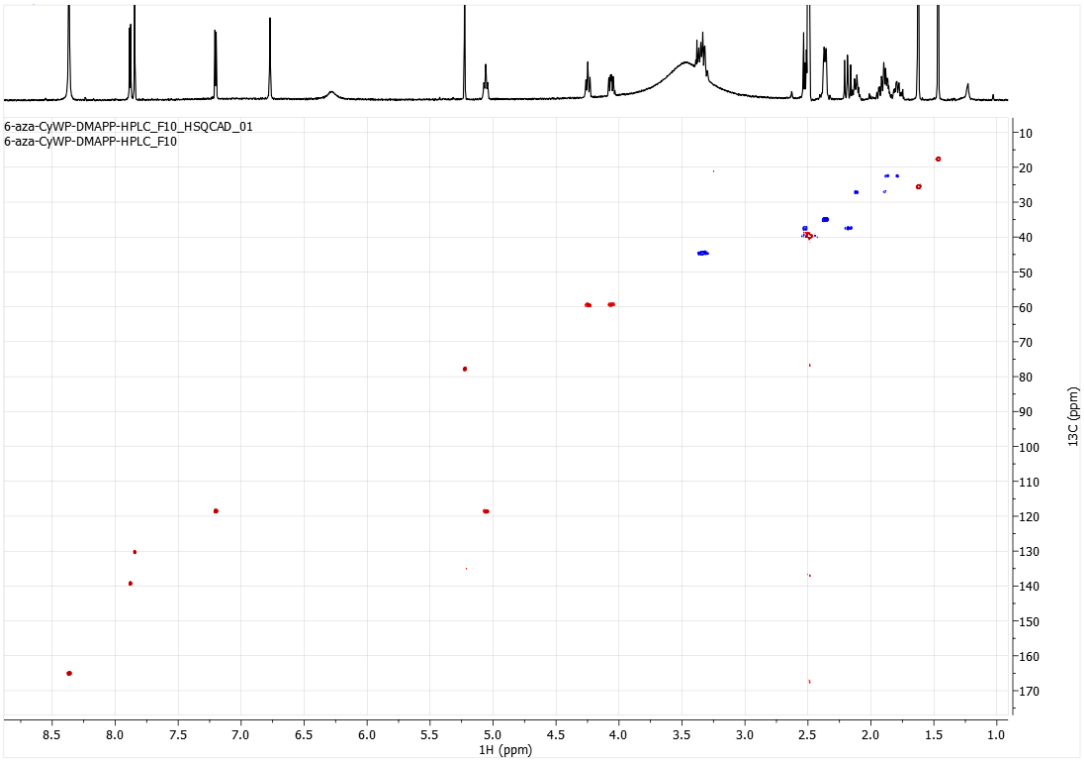
HMBC spectrum (500 MHz) of **5-C3** in DMSO-d_6



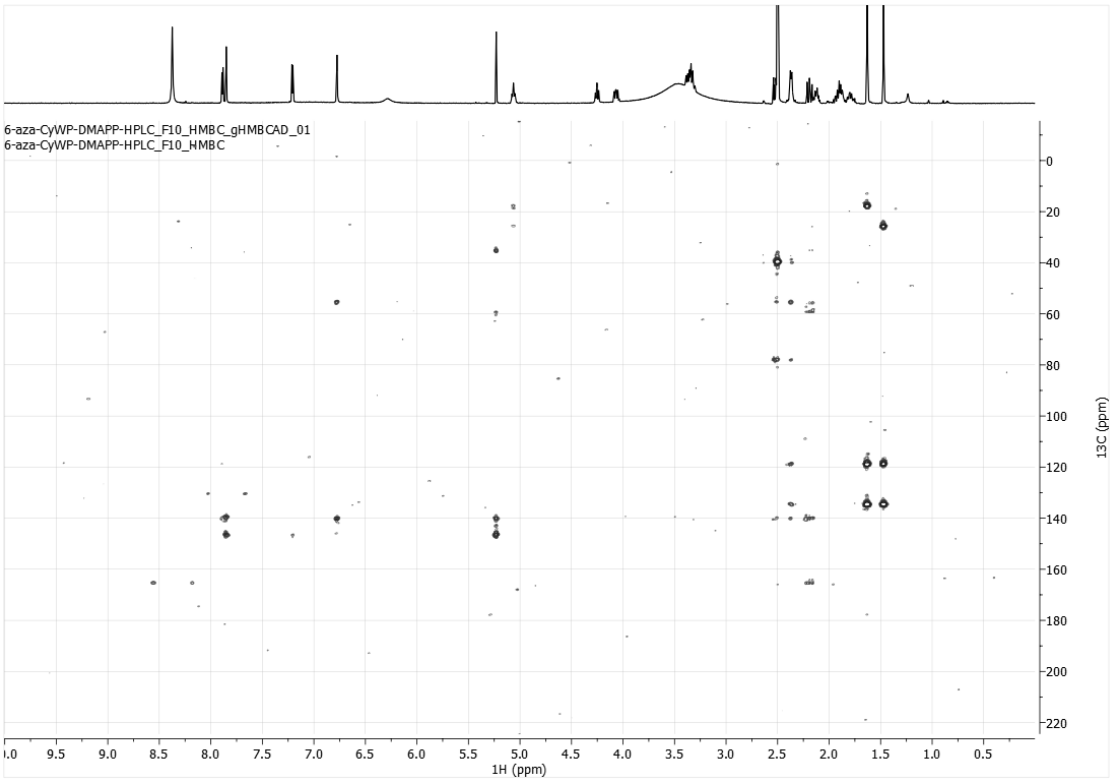
^{1}H NMR spectrum (500 MHz) of **6-C3** in DMSO-d_6



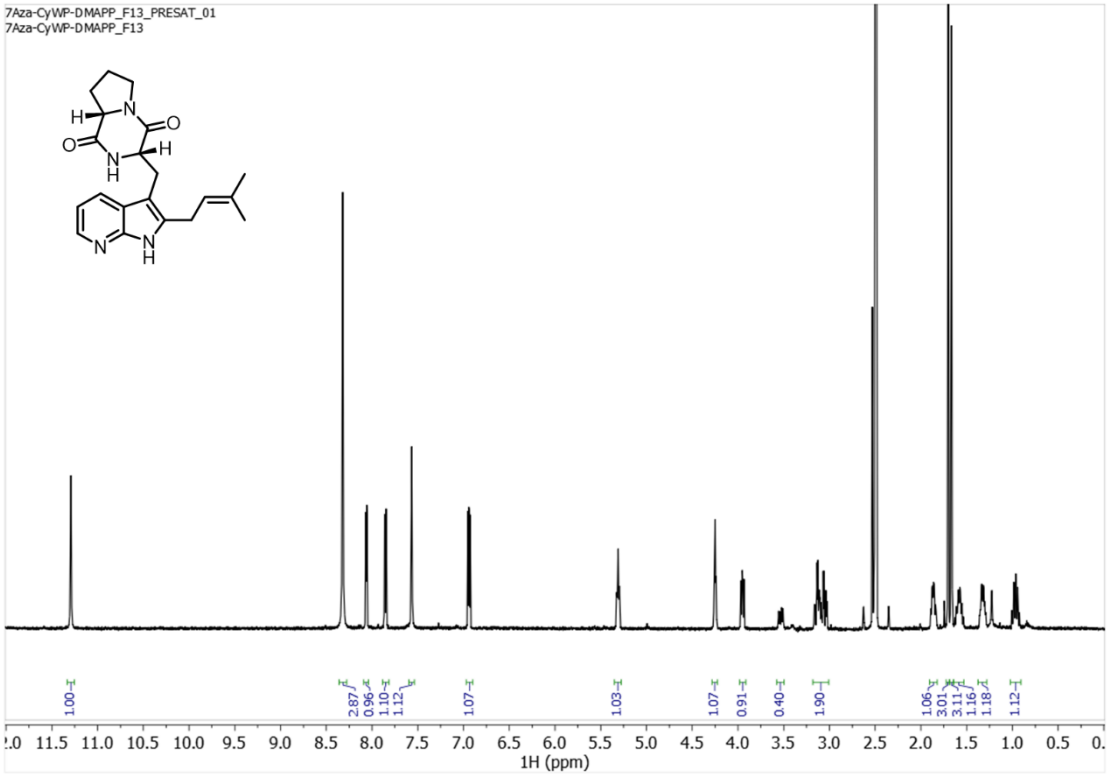
^{1}H-COSY spectrum (500 MHz) of **6-C3** in DMSO-d_6



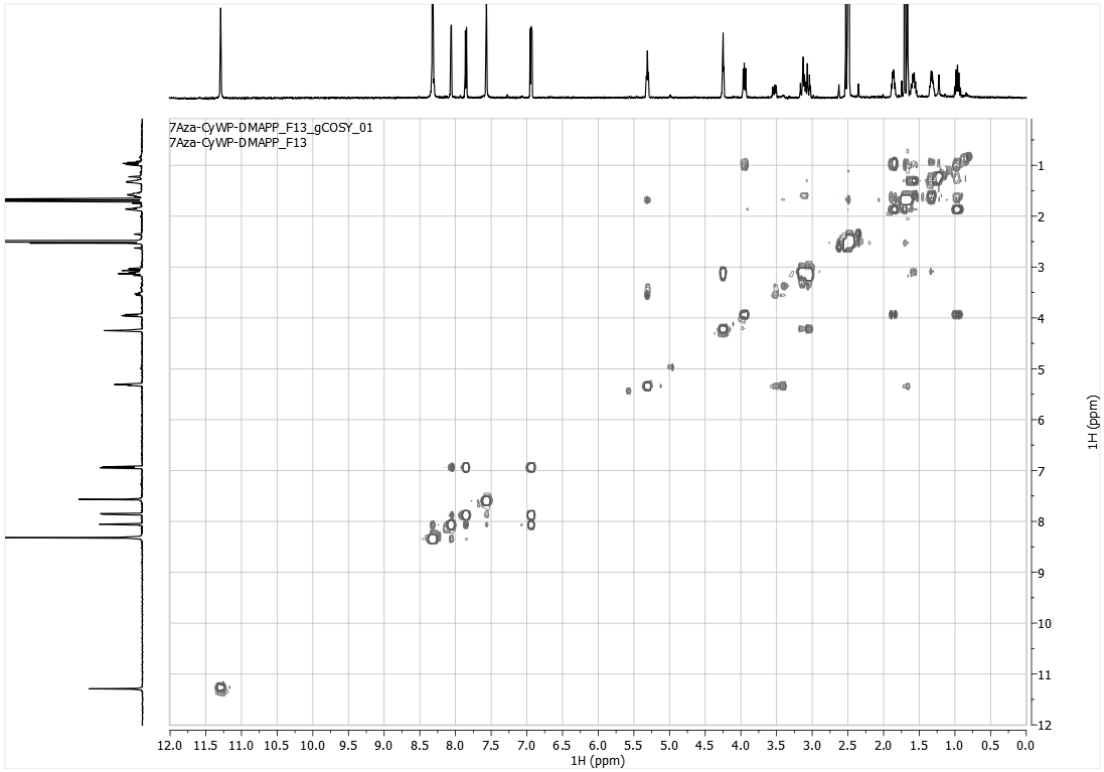
HSQC spectrum (500 MHz) of **6-C3** in DMSO-d_6



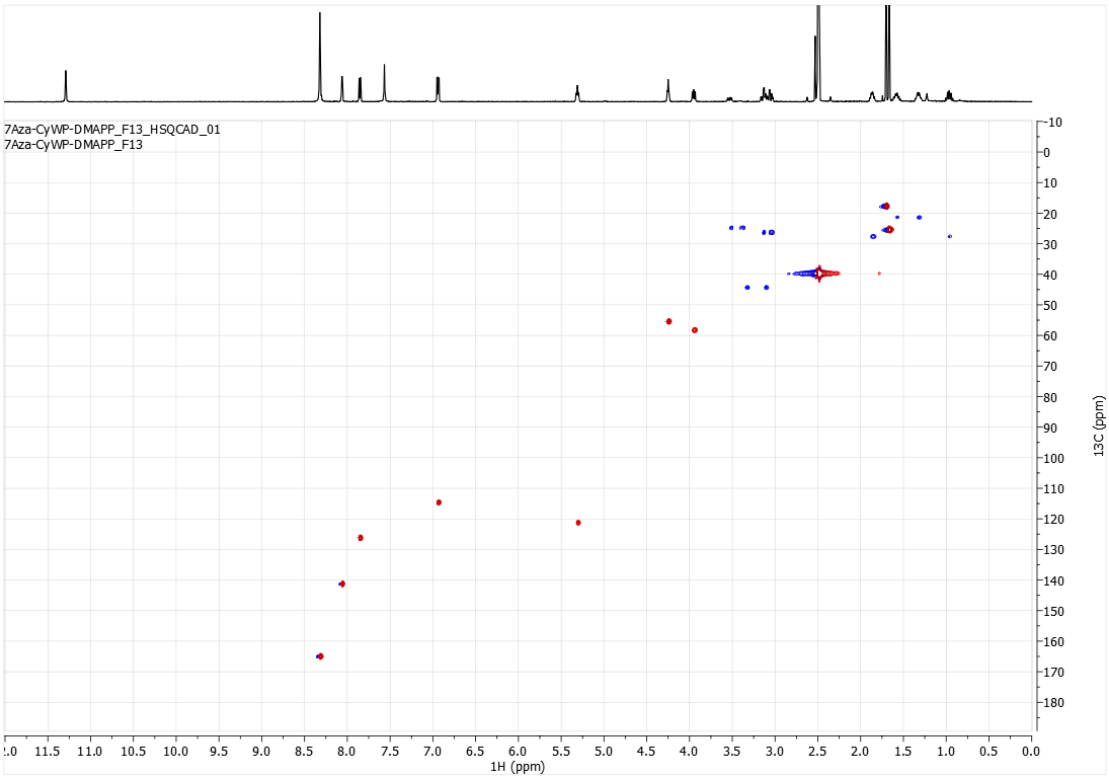
HMBC spectrum (500 MHz) of **6-C3** in DMSO-d_6



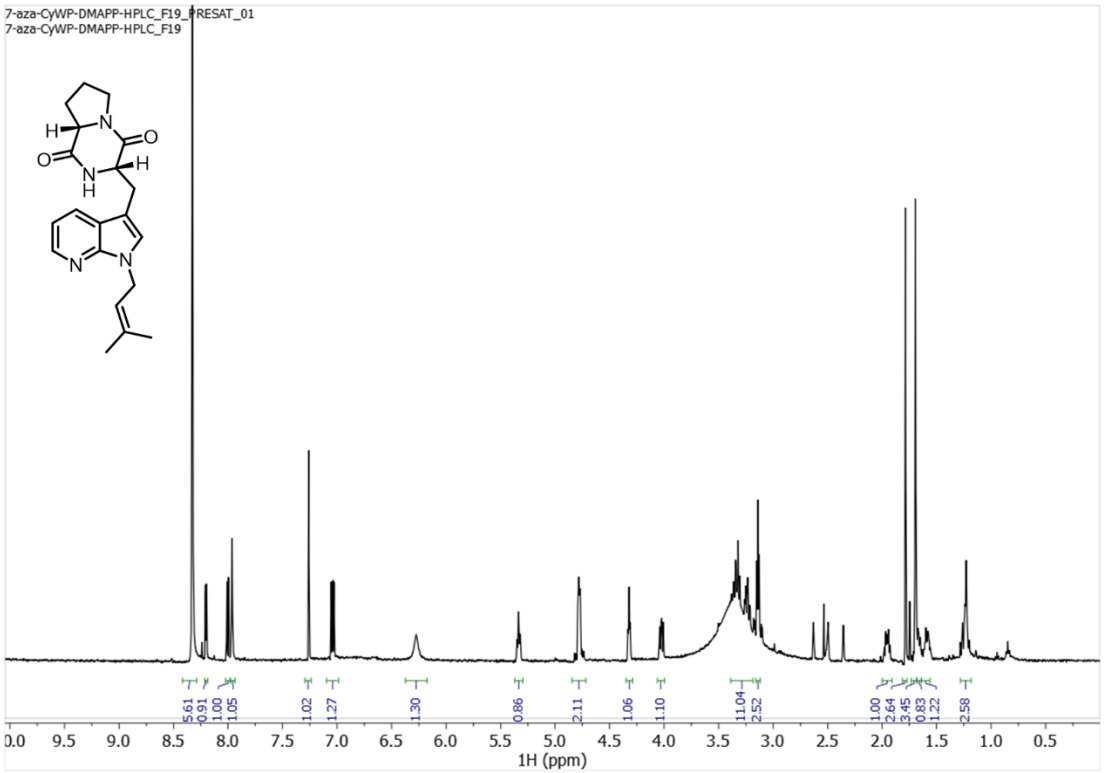
^{1}H NMR spectrum (500 MHz) of **7-C2** in DMSO-d$_6$



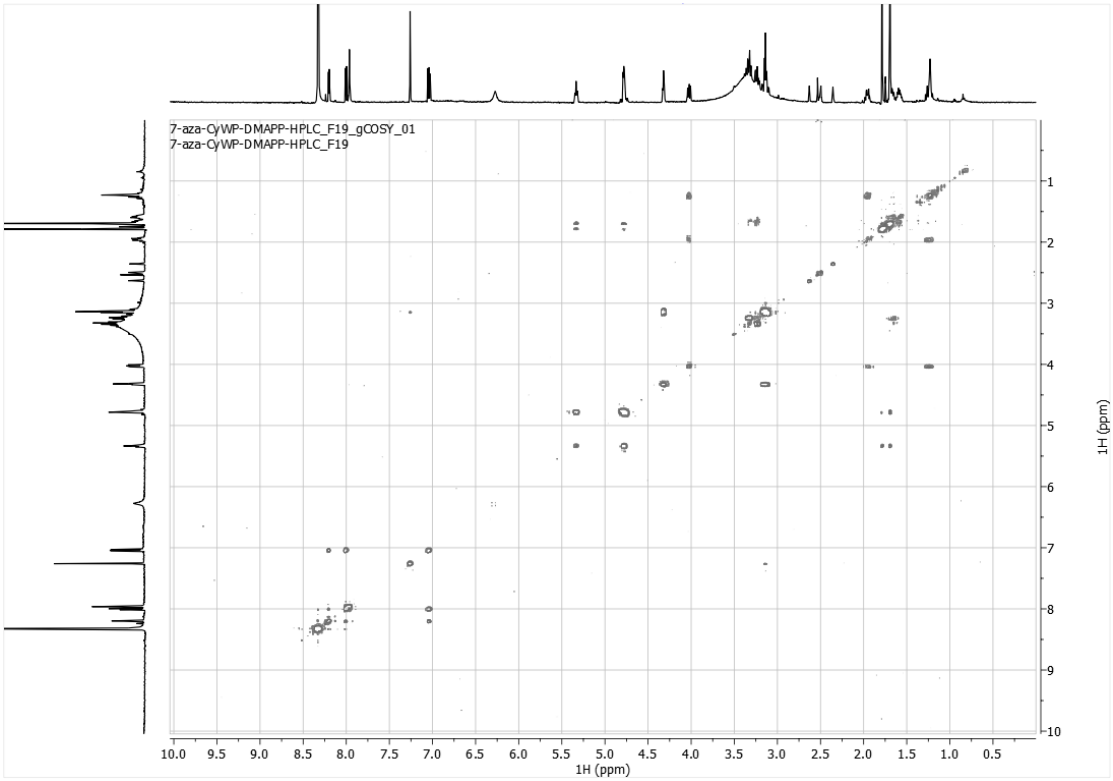
^{1}H-COSY spectrum (500 MHz) of **7-C2** in DMSO-d$_6$



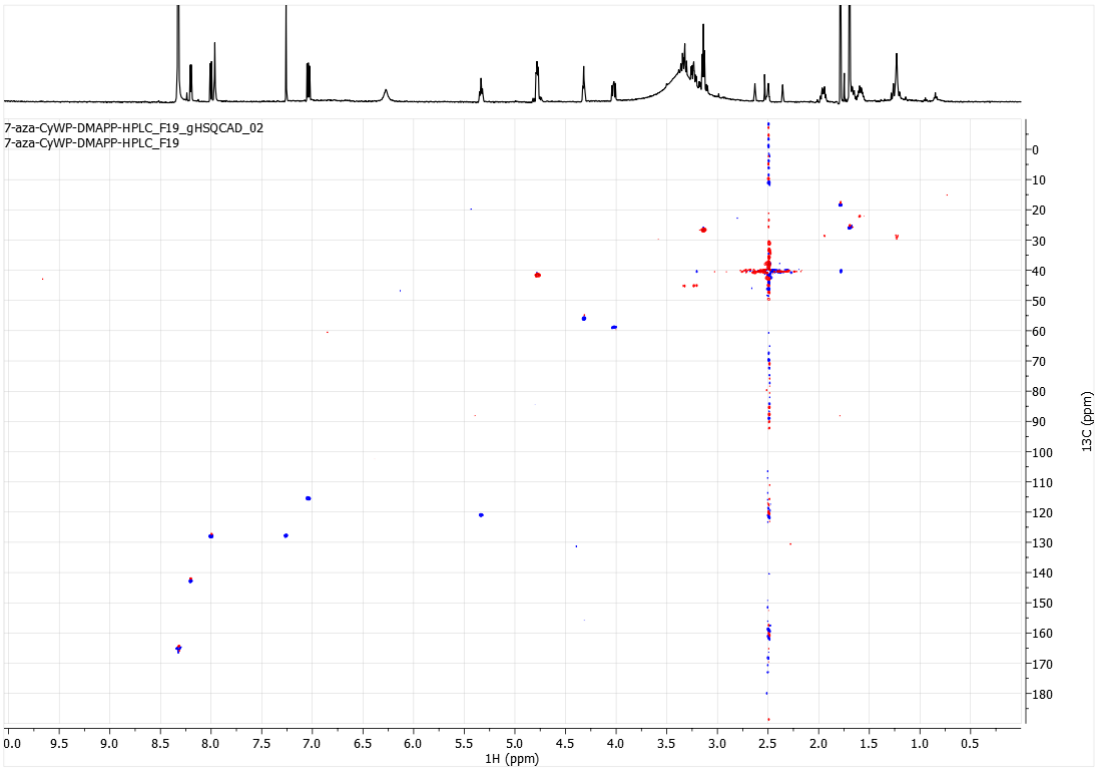
HSQC spectrum (500 MHz) of **7-C2** in DMSO-d_6



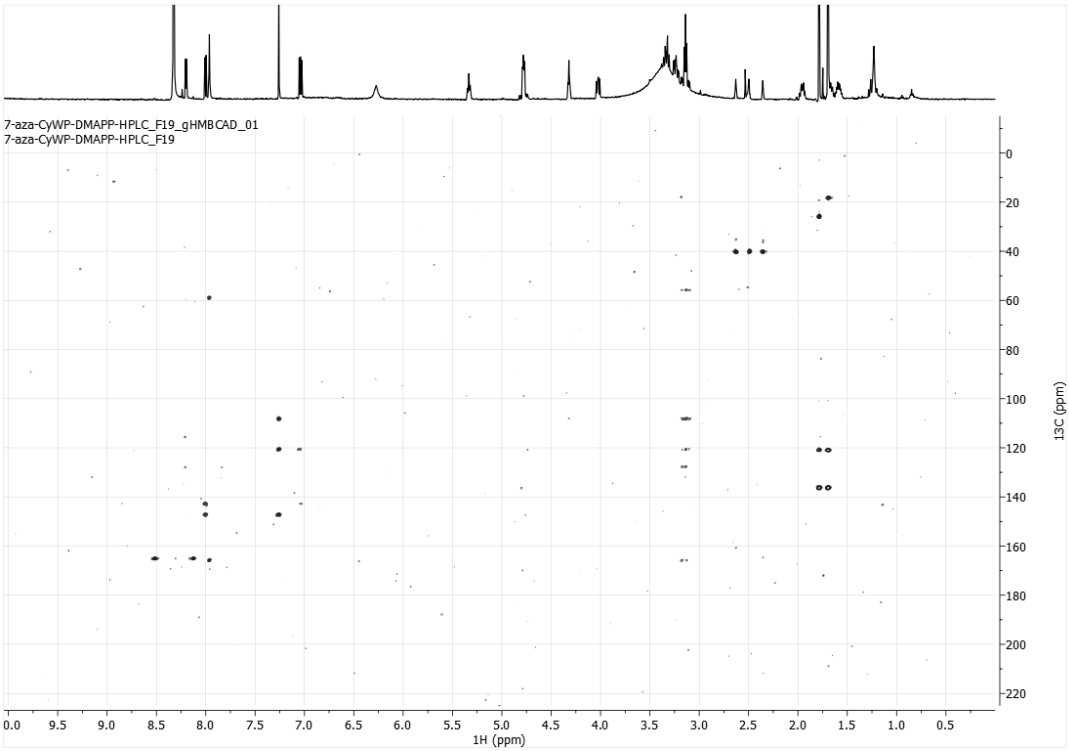
^{1}H NMR spectrum (500 MHz) of **7-N1** in DMSO-d$_6$



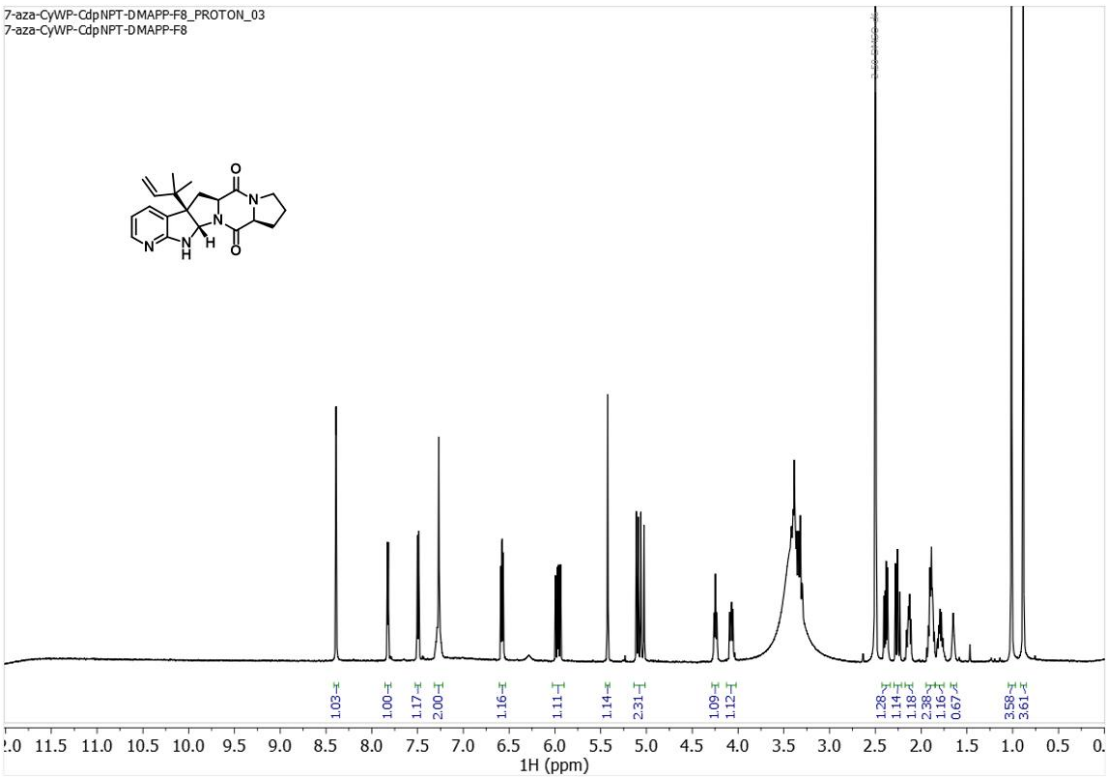
^{1}H-COSY spectrum (500 MHz) of **7-N1** in DMSO-d$_6$



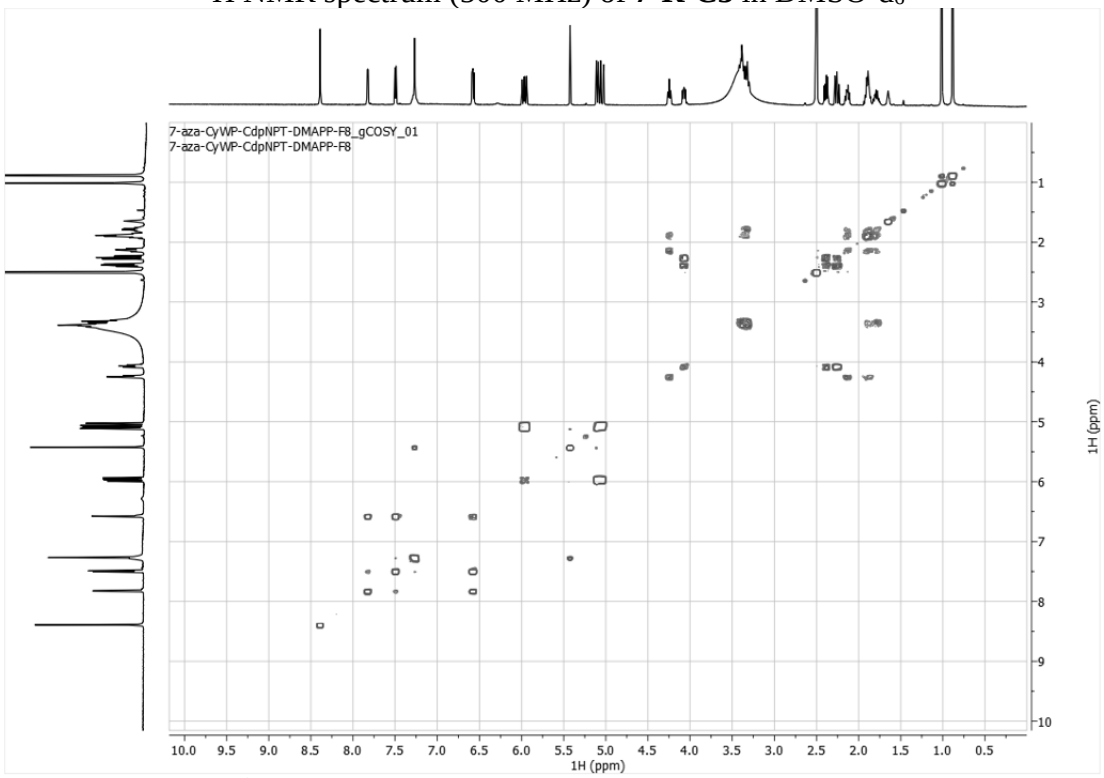
HSQC spectrum (500 MHz) of **7-N1** in DMSO-d_6



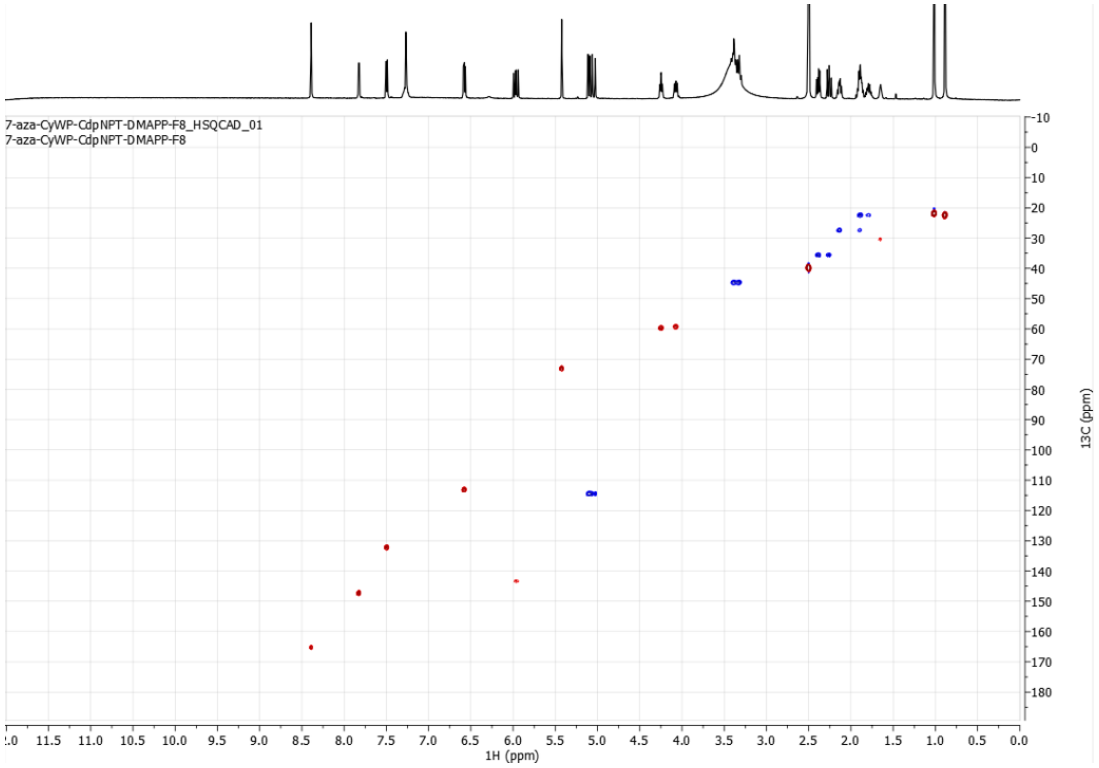
HMBC spectrum (500 MHz) of **7-N1** in DMSO-d_6



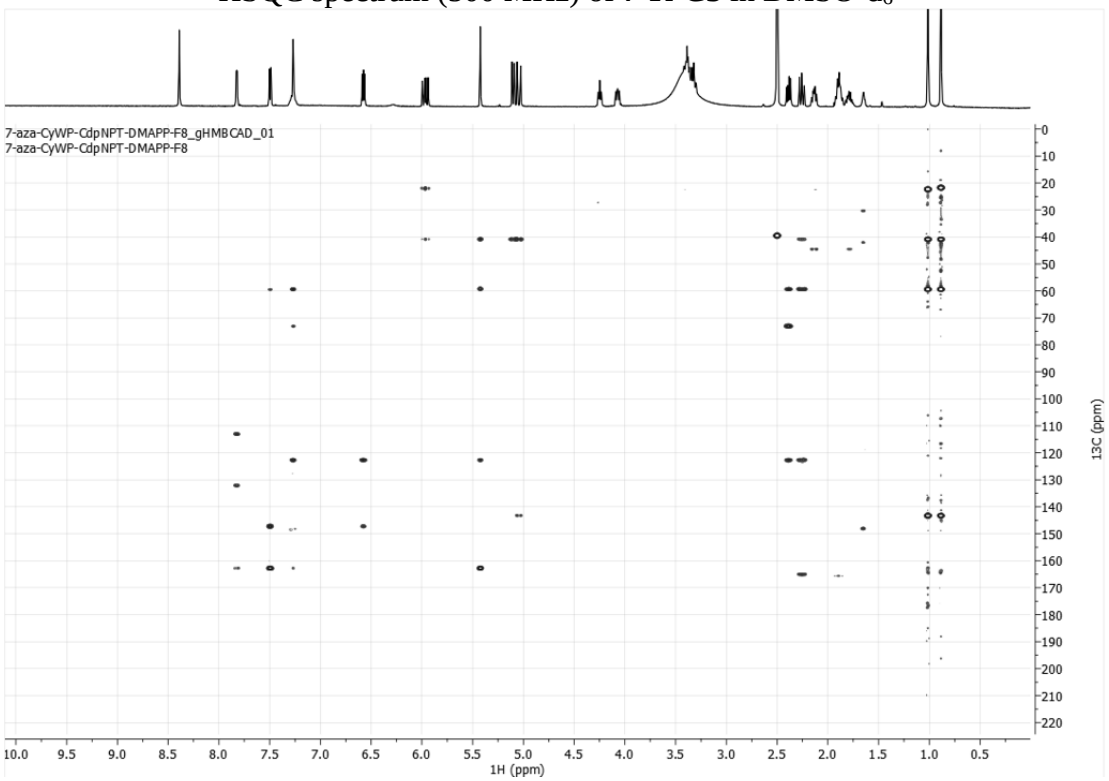
^{1}H NMR spectrum (500 MHz) of **7-R-C3** in DMSO-d_6



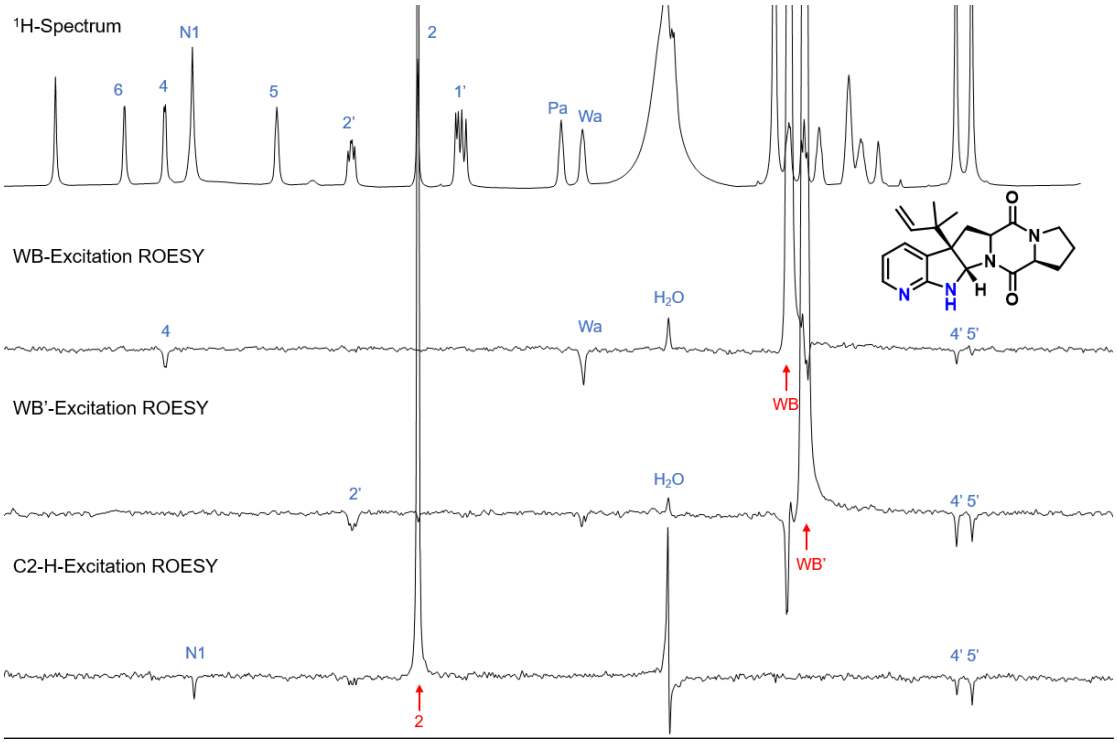
^{1}H-COSY spectrum (500 MHz) of **7-R-C3** in DMSO-d_6



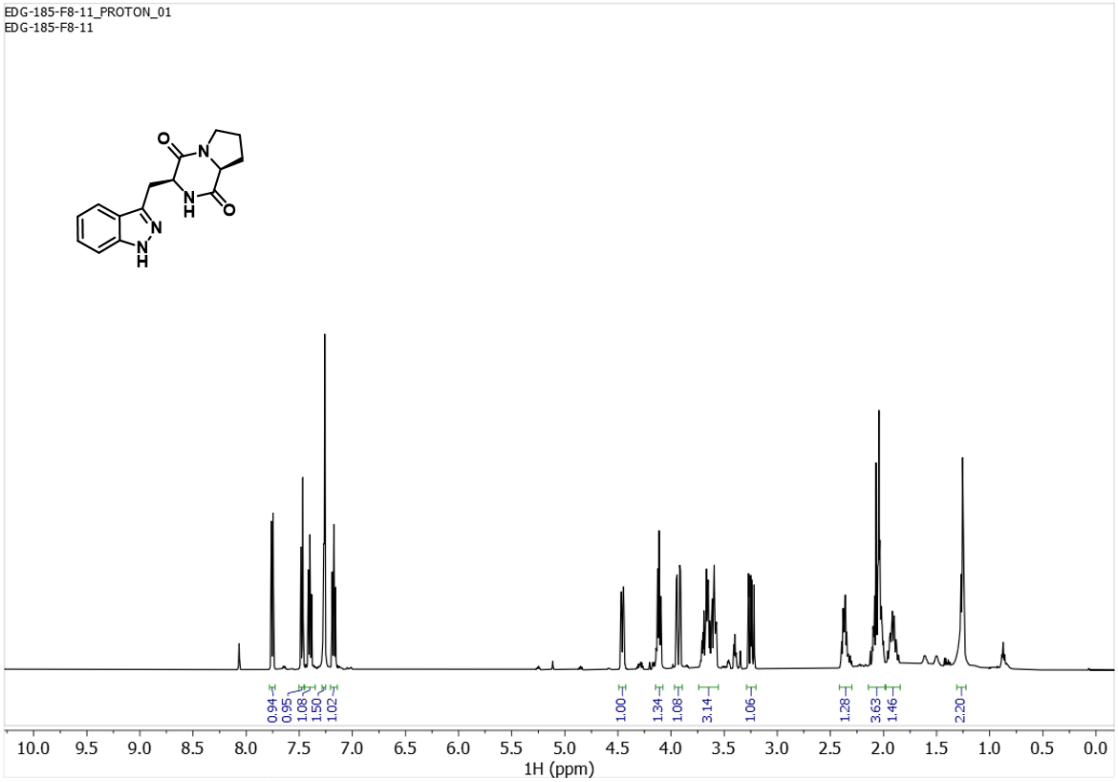
HSQC spectrum (500 MHz) of **7-R-C3** in DMSO-d_6



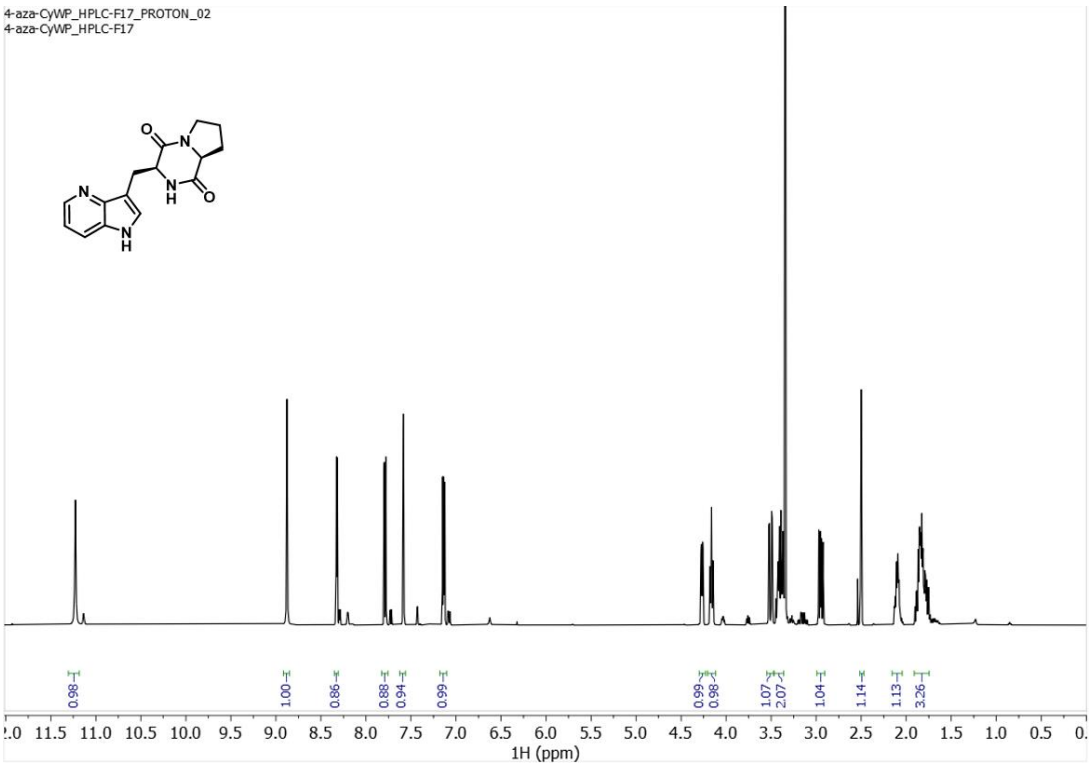
HMBC spectrum (500 MHz) of **7-R-C3** in DMSO-d_6



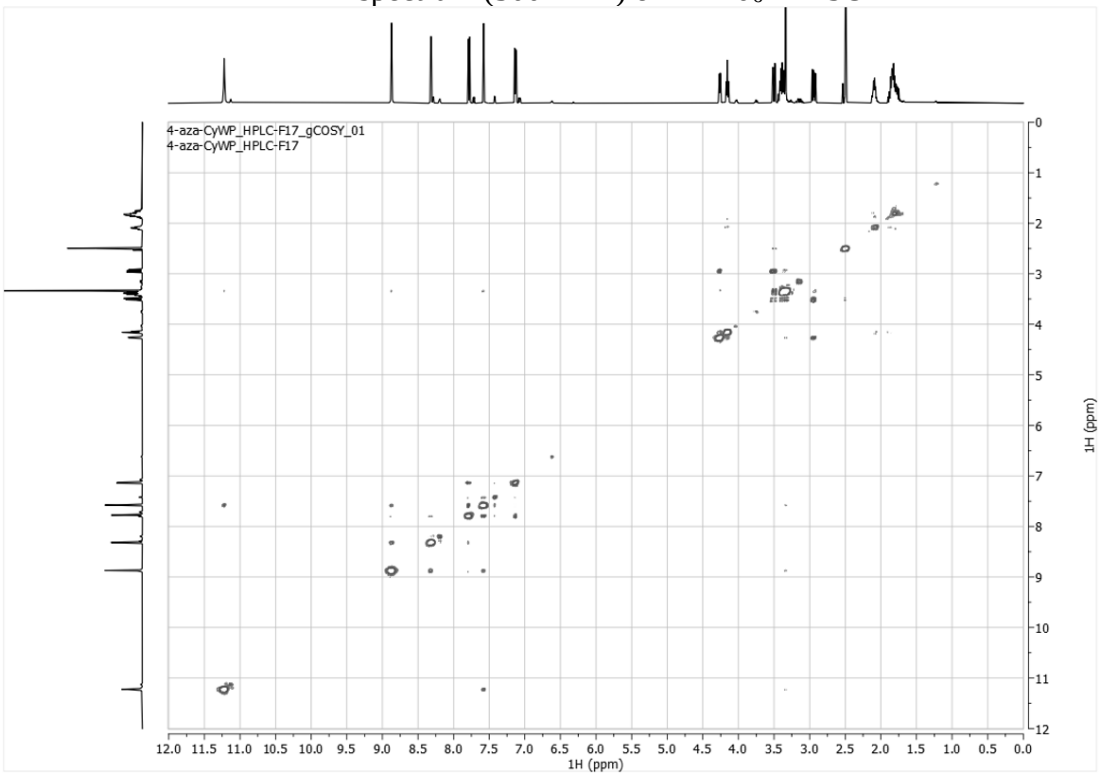
Stacked 1D-ROESY experiments for **7-R-C3**, with a ^{1}H spectrum on top for reference



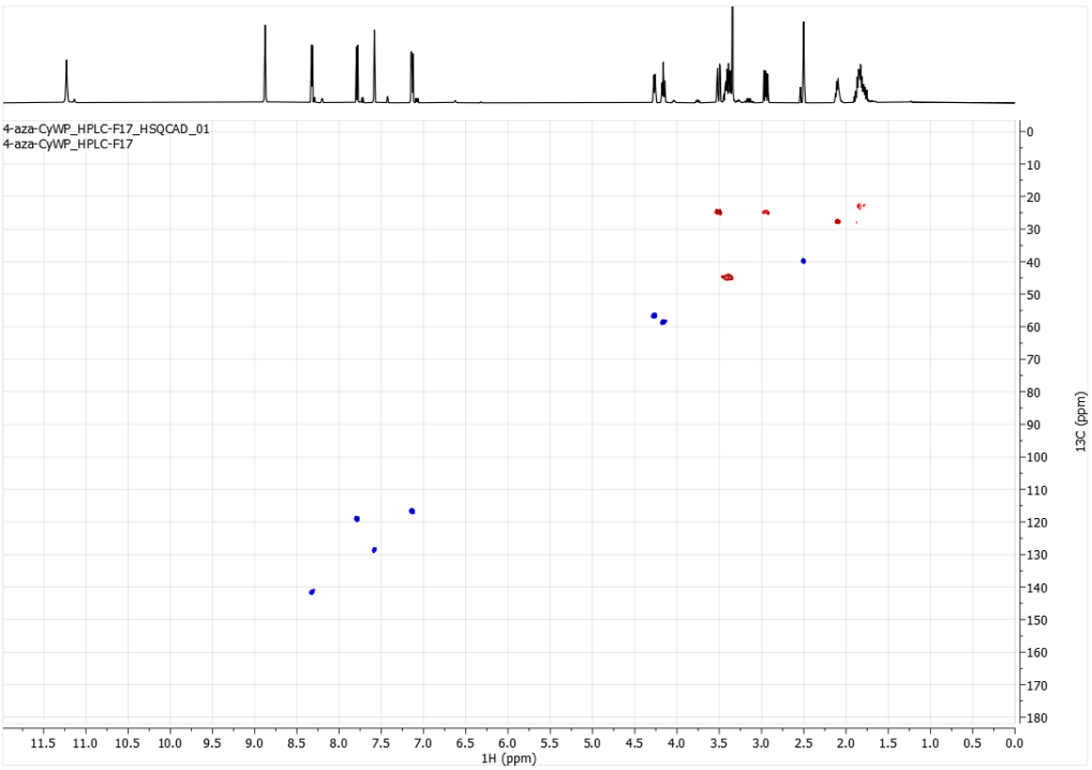
^{1}H NMR spectrum (500 MHz) of **2f** in $CDCl_3$



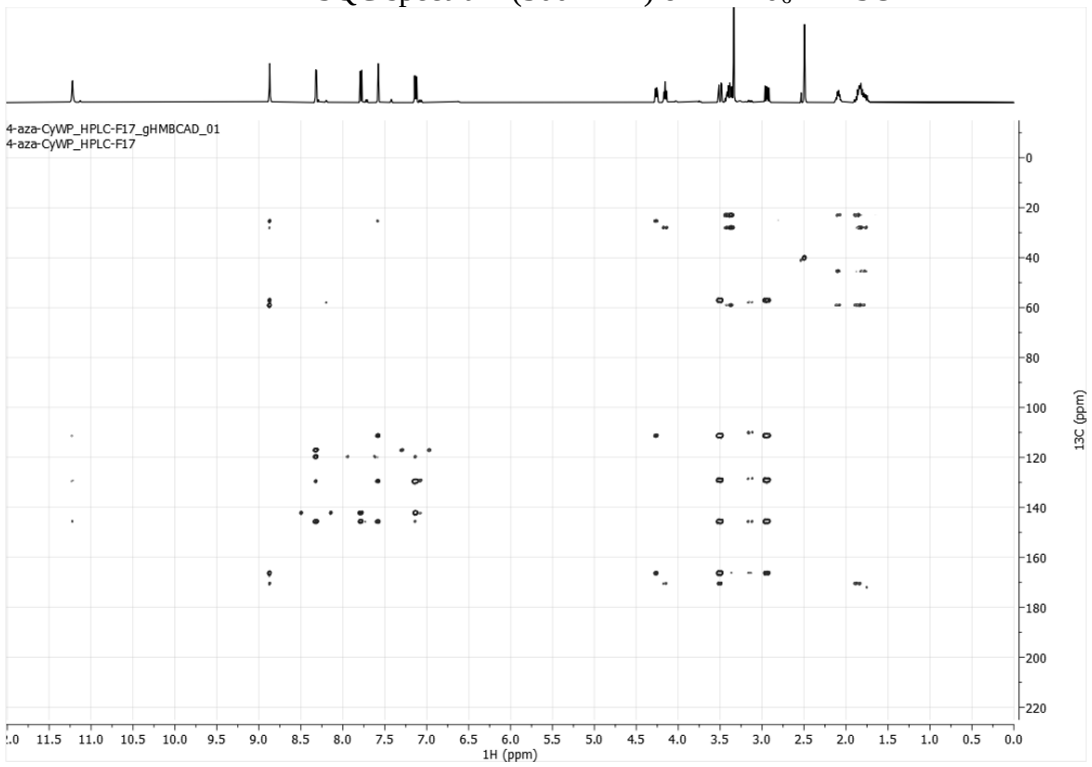
^{1}H NMR spectrum (500 MHz) of **4f** in d_6-DMSO



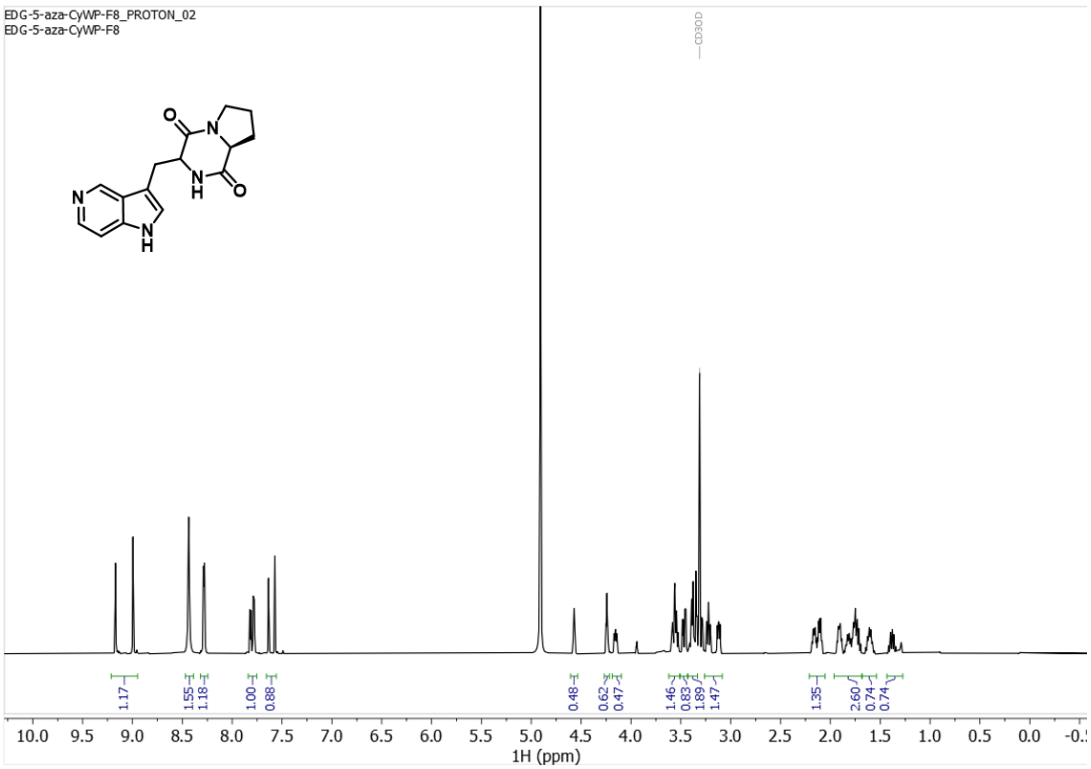
^{1}H-COSY spectrum (500 MHz) of **4f** in d_6-DMSO



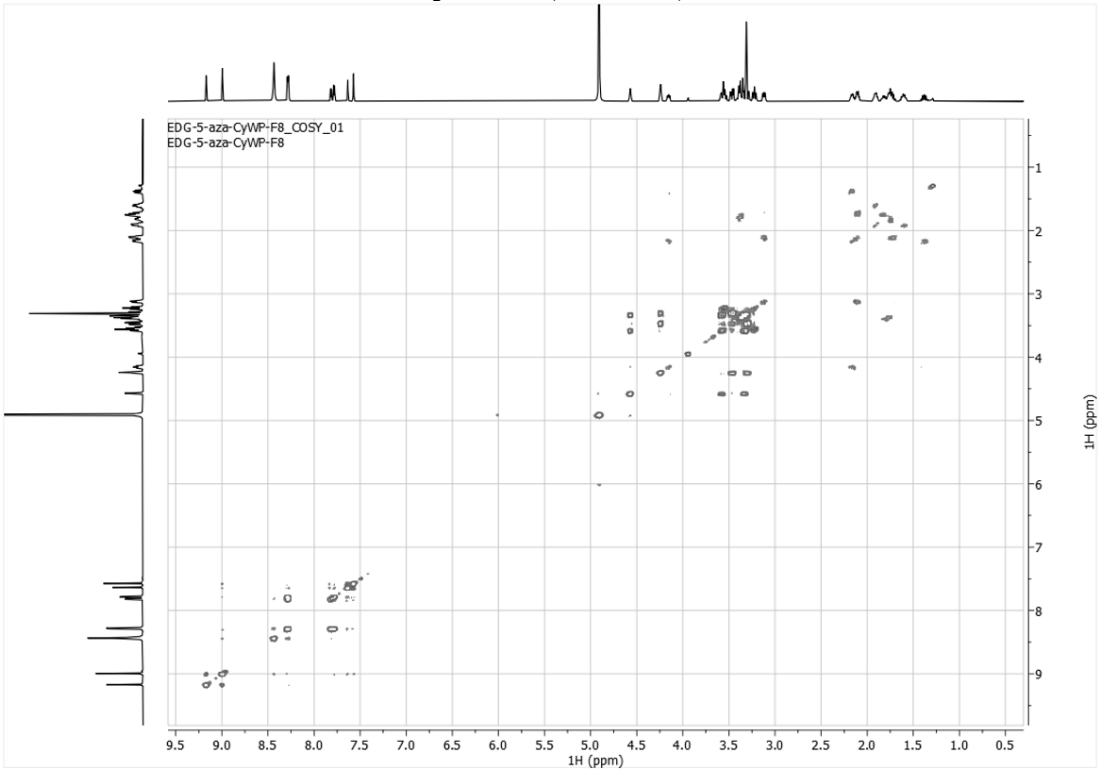
HSQC spectrum (500 MHz) of **4f** in d_6-DMSO



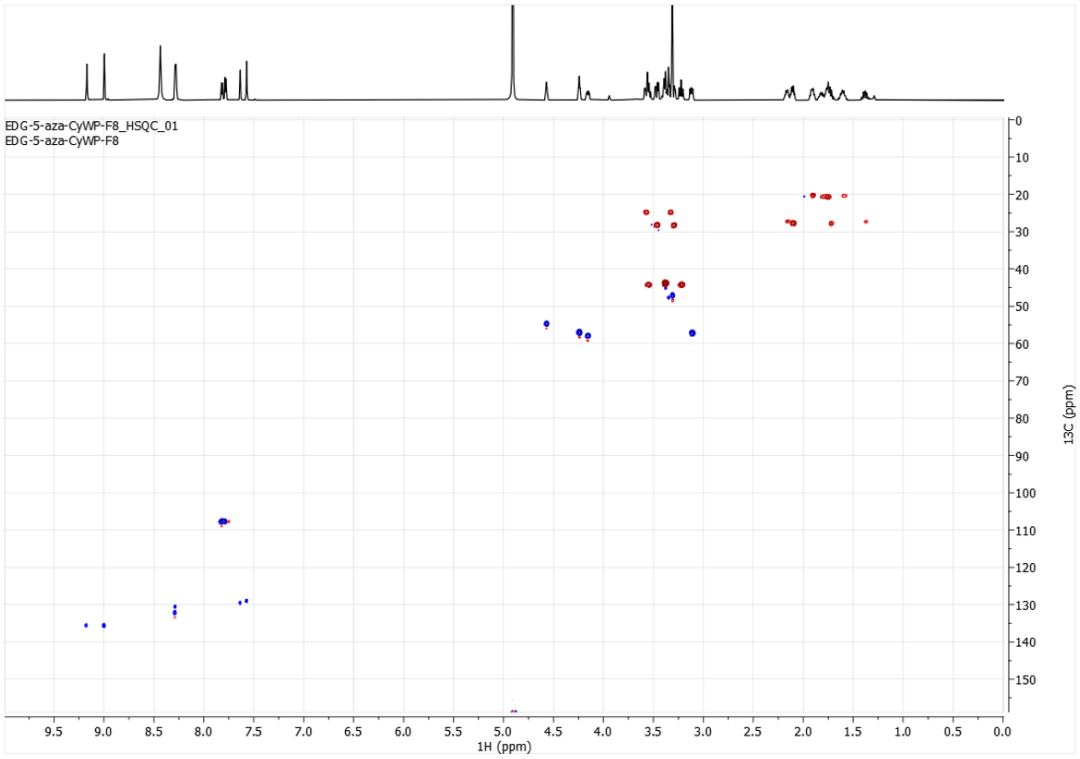
HMBC spectrum (500 MHz) of **4f** in d_6-DMSO



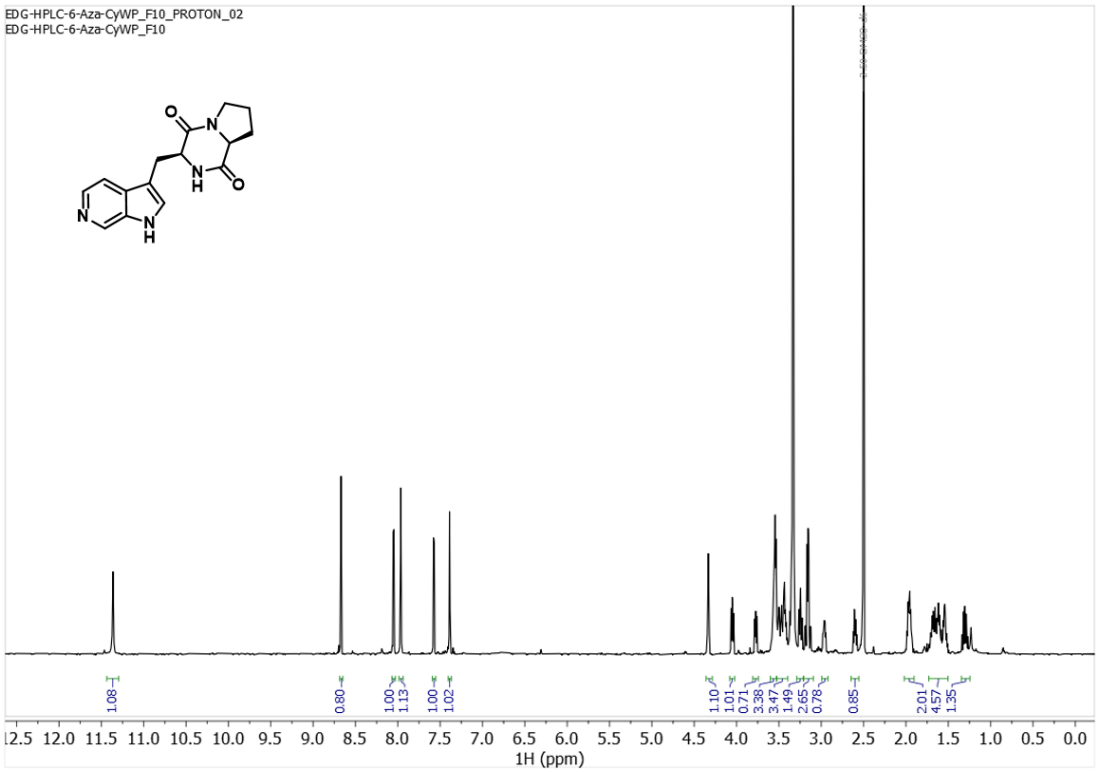
^{1}H NMR spectrum (500 MHz) of **5f** in CD_3OD



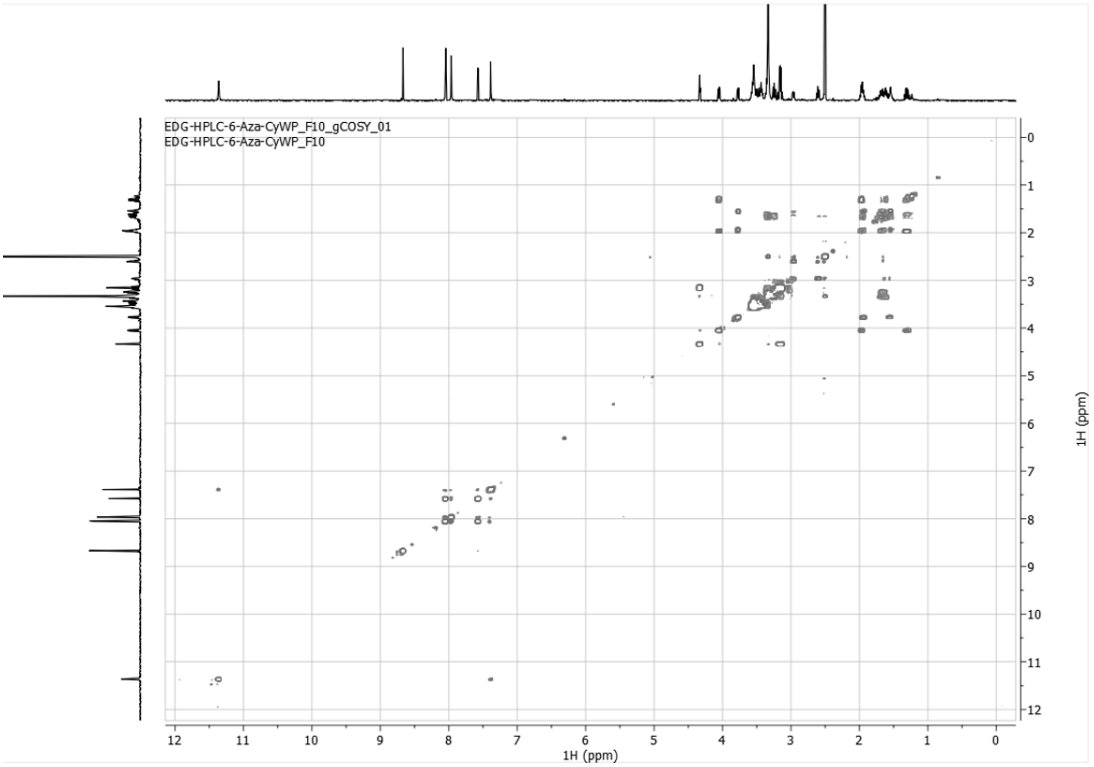
^{1}H-COSY spectrum (500 MHz) of **5f** in CD_3OD



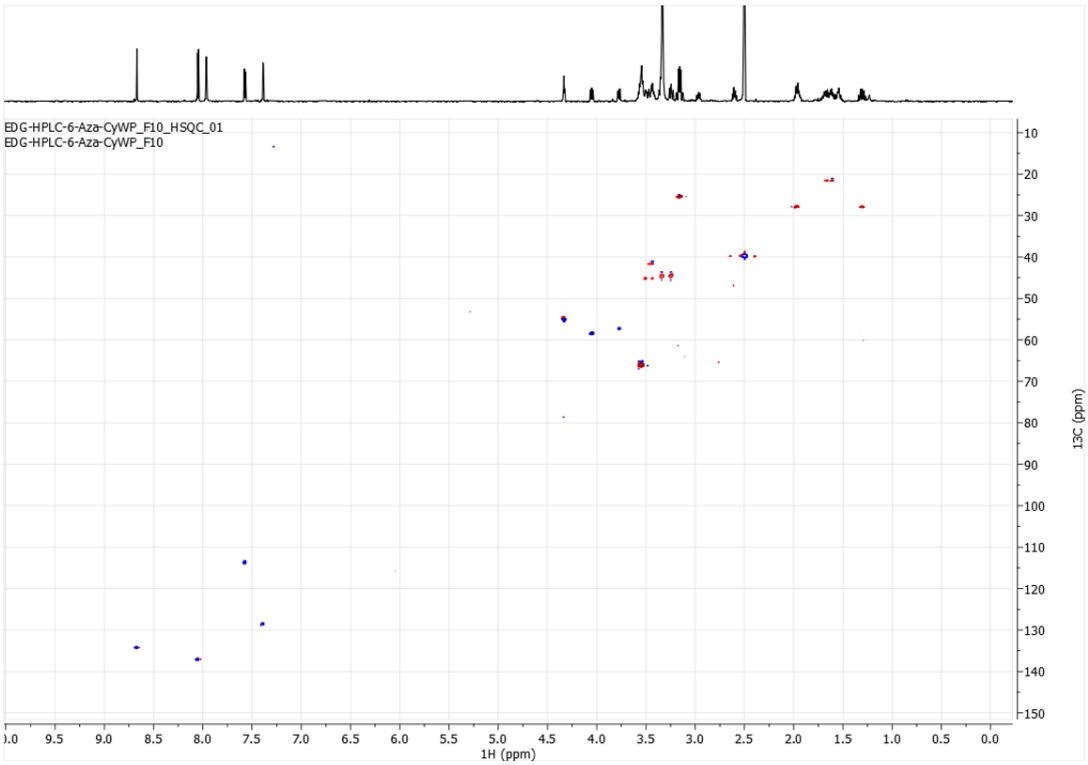
HSQC spectrum (500 MHz) of **5f** in CD_3OD



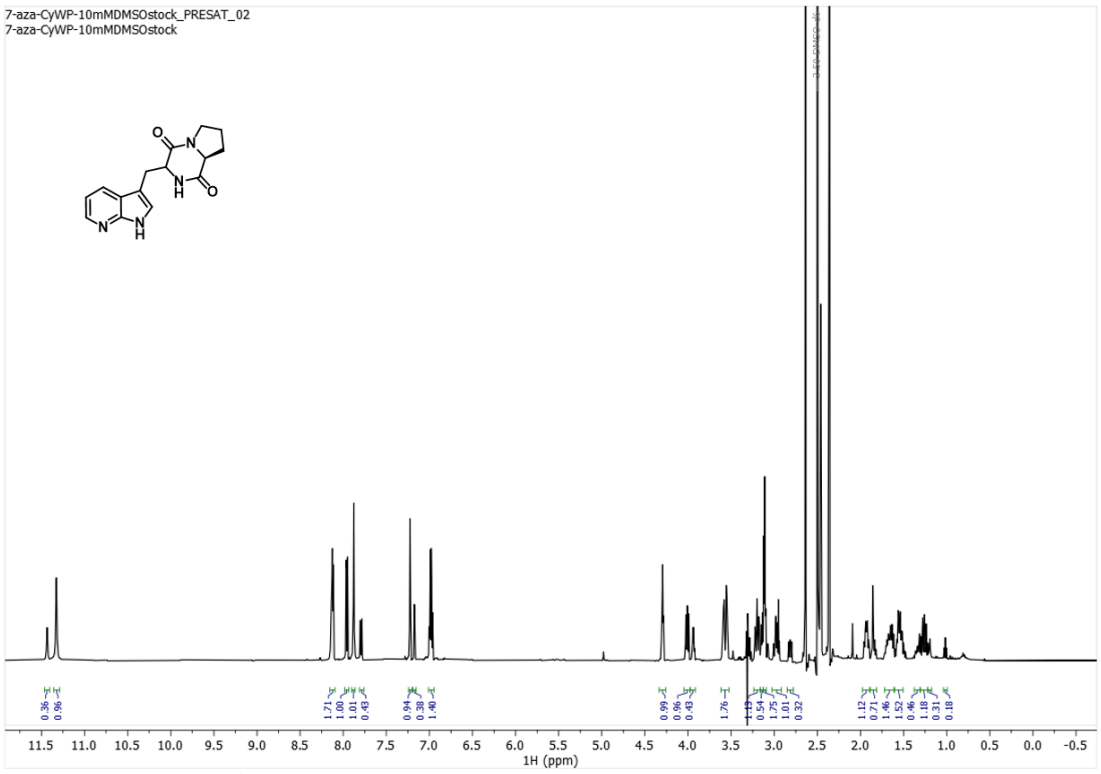
^{1}H NMR spectrum (500 MHz) of **6f** in d_6-DMSO



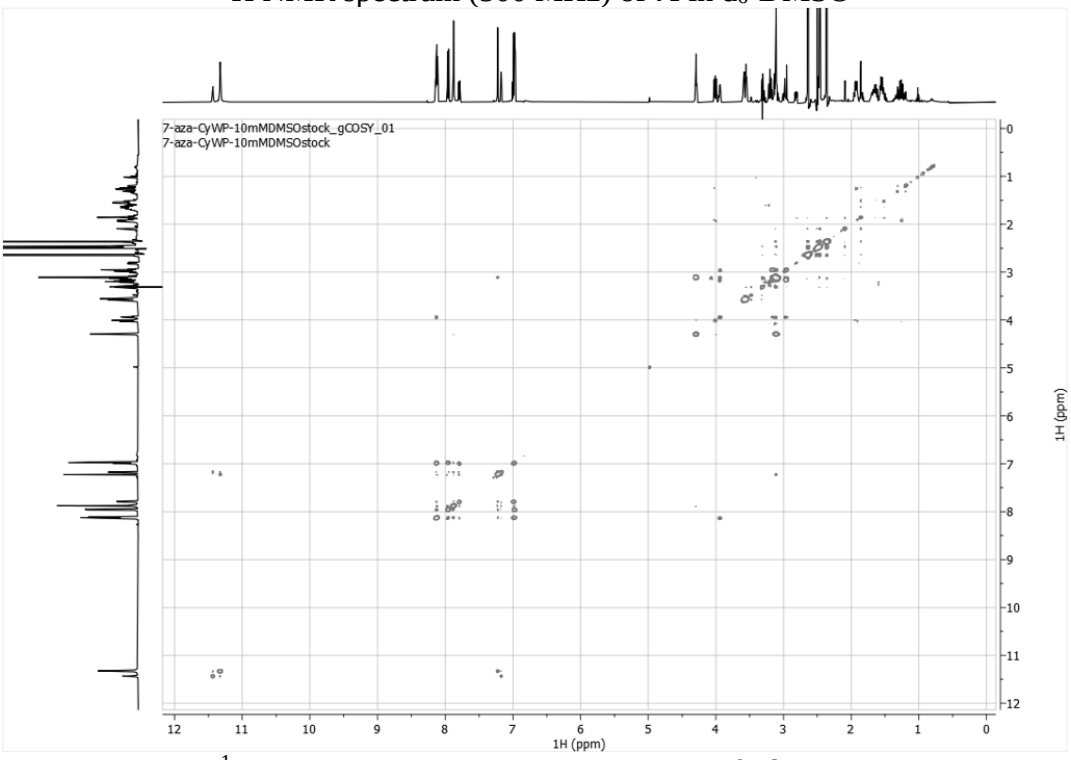
^{1}H-COSY spectrum (500 MHz) of **6f** in d_6-DMSO



HSQC spectrum (500 MHz) of **6f** in d_6-DMSO

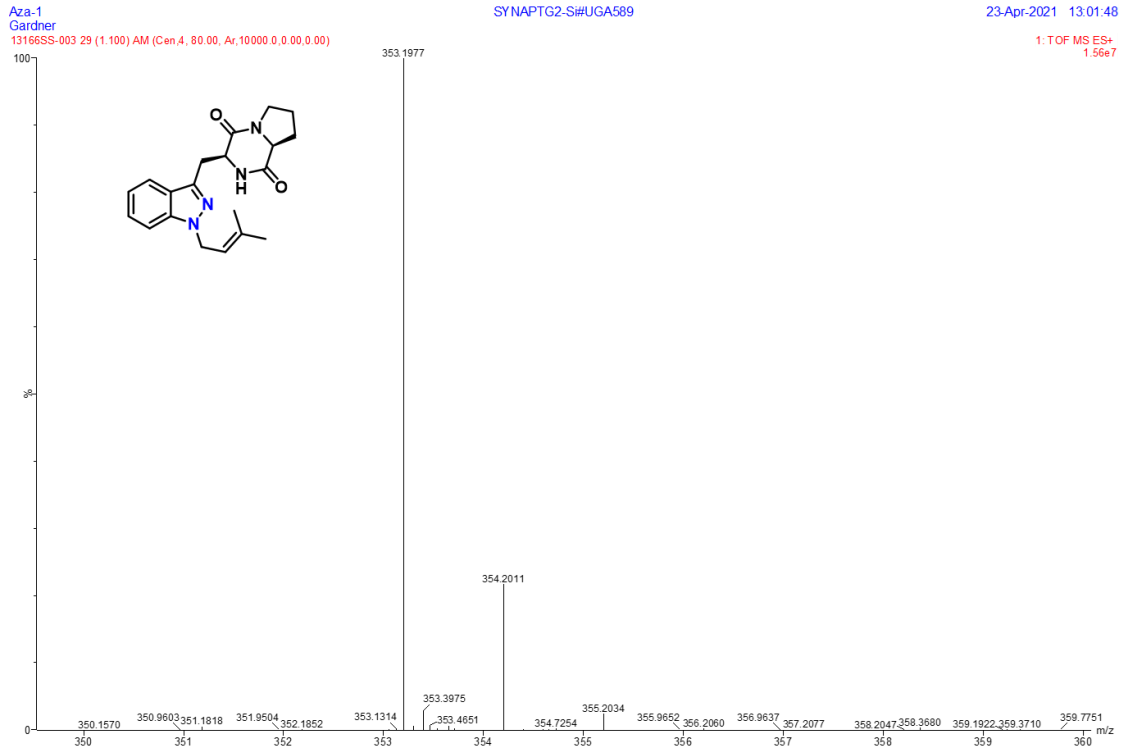


^{1}H NMR spectrum (500 MHz) of **7f** in d_6-DMSO

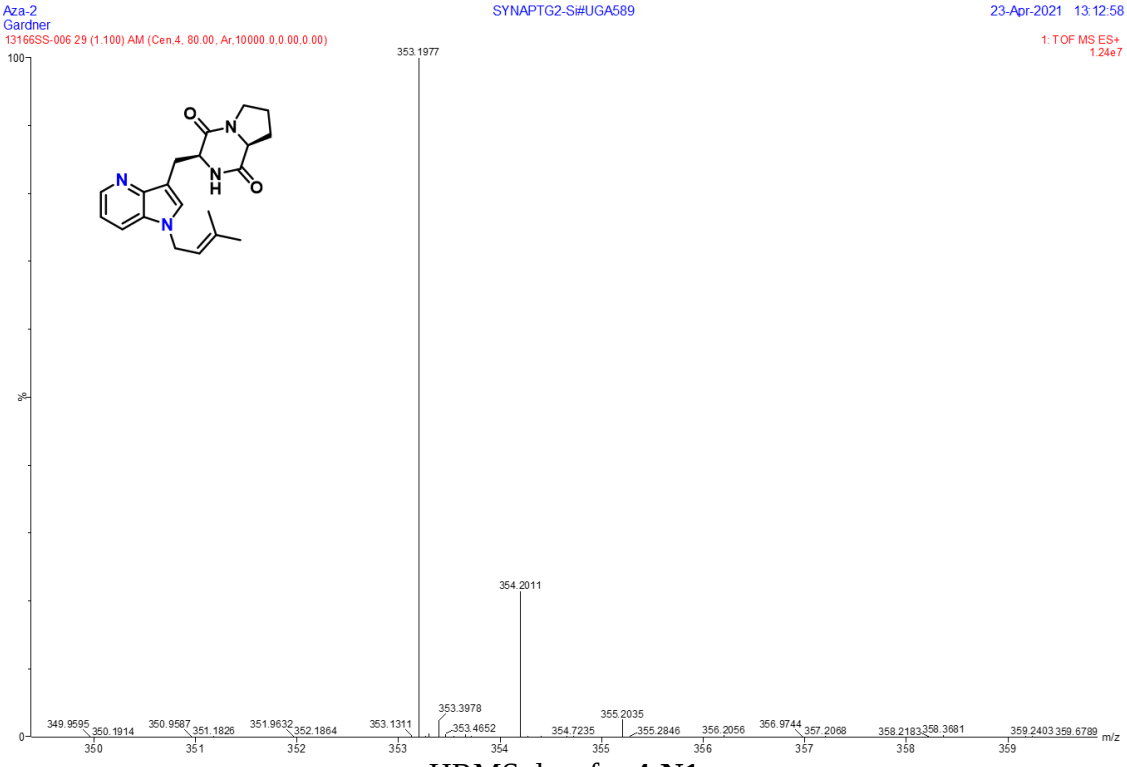


^{1}H-COSY NMR spectrum (500 MHz) of 7**f** in d_6-DMSO

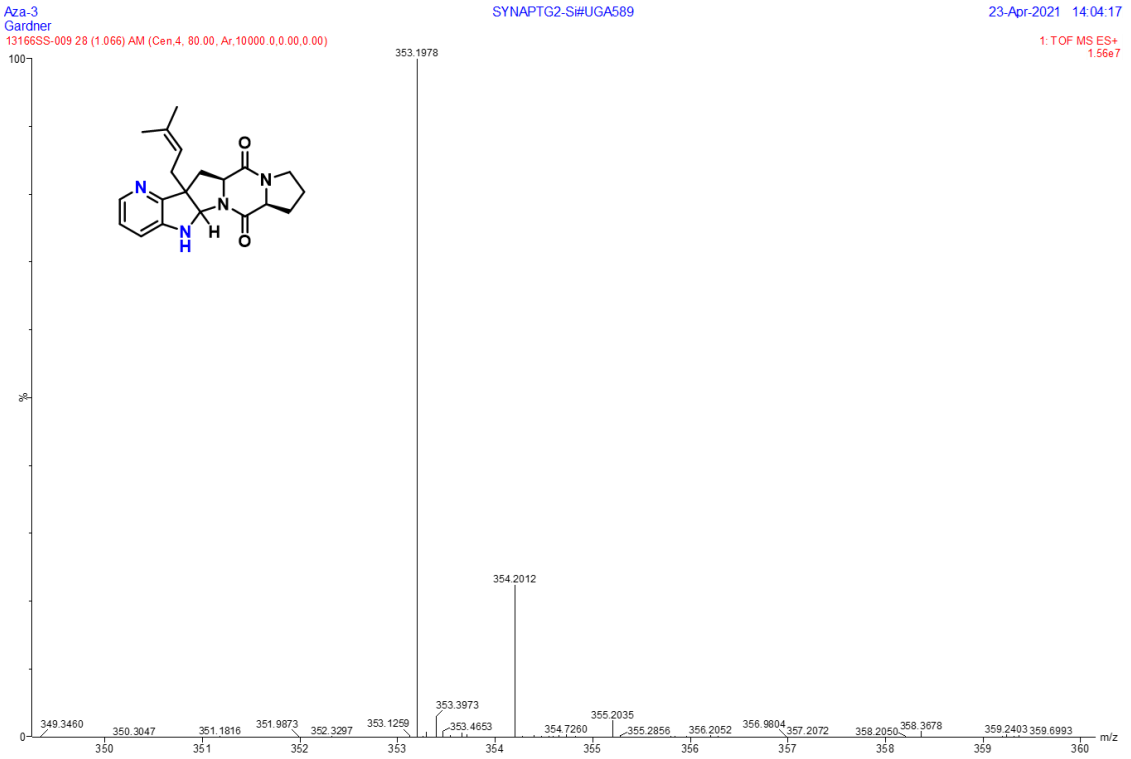
4.6 Appendix 2: HRMS Data



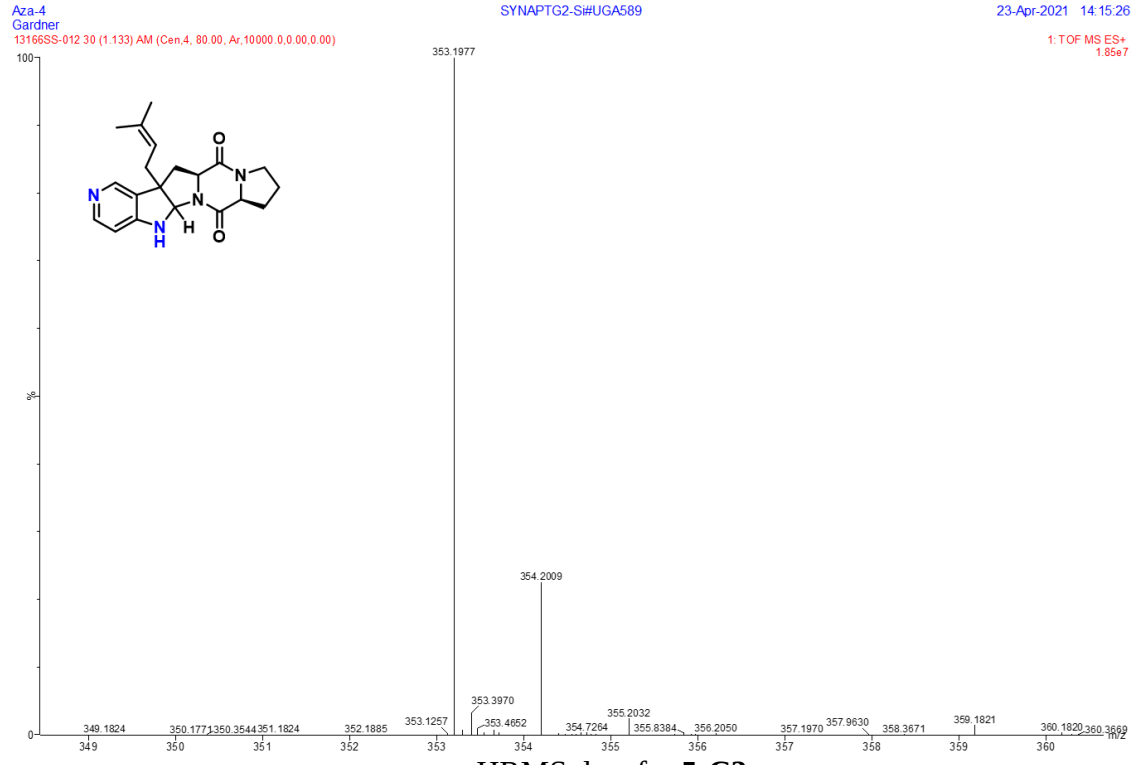
HRMS data for **2-N1**



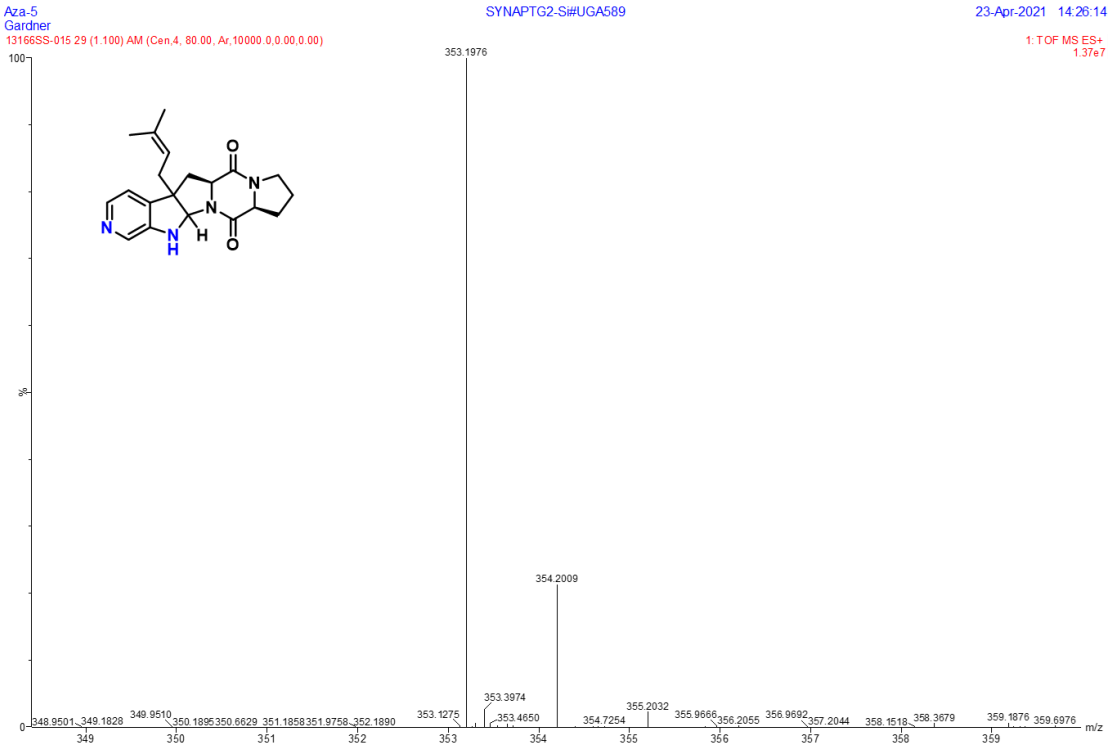
HRMS data for **4-N1**



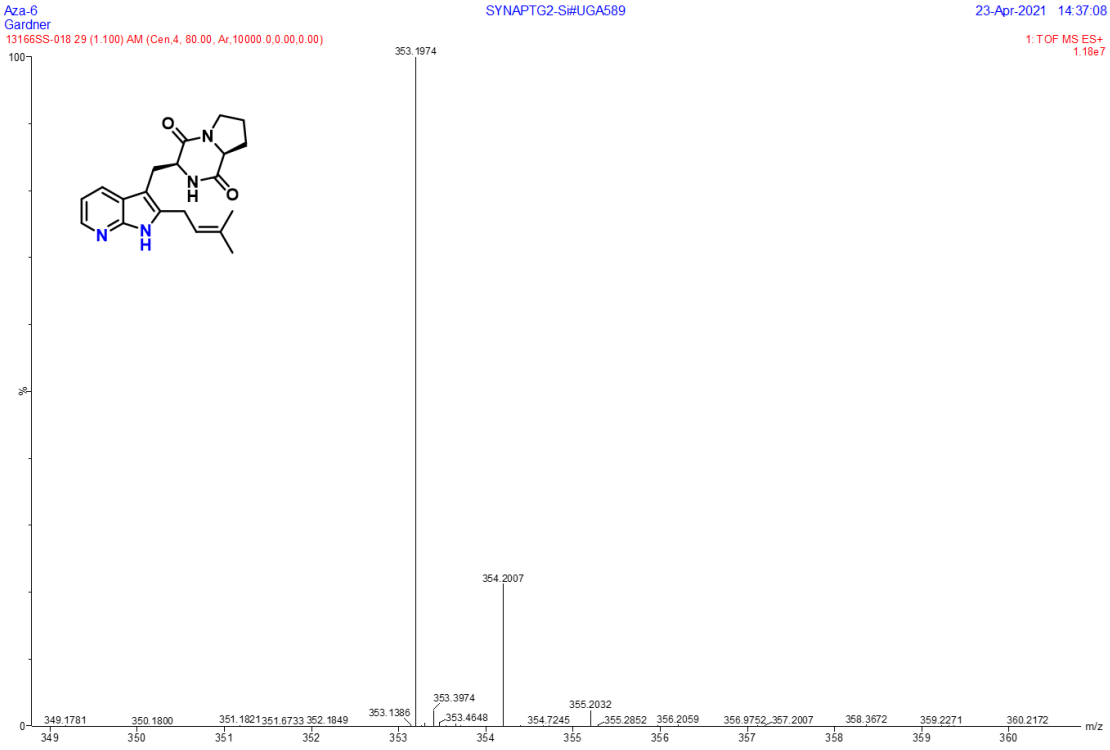
HRMS data for **4-C3**



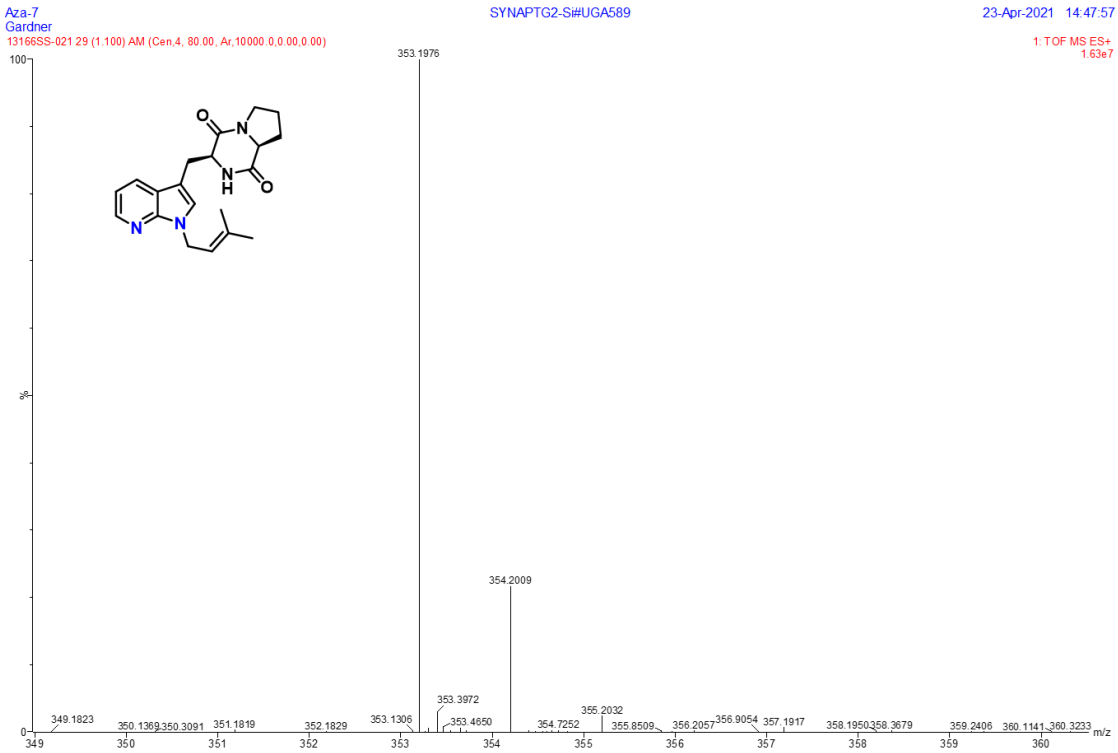
HRMS data for **5-C3**



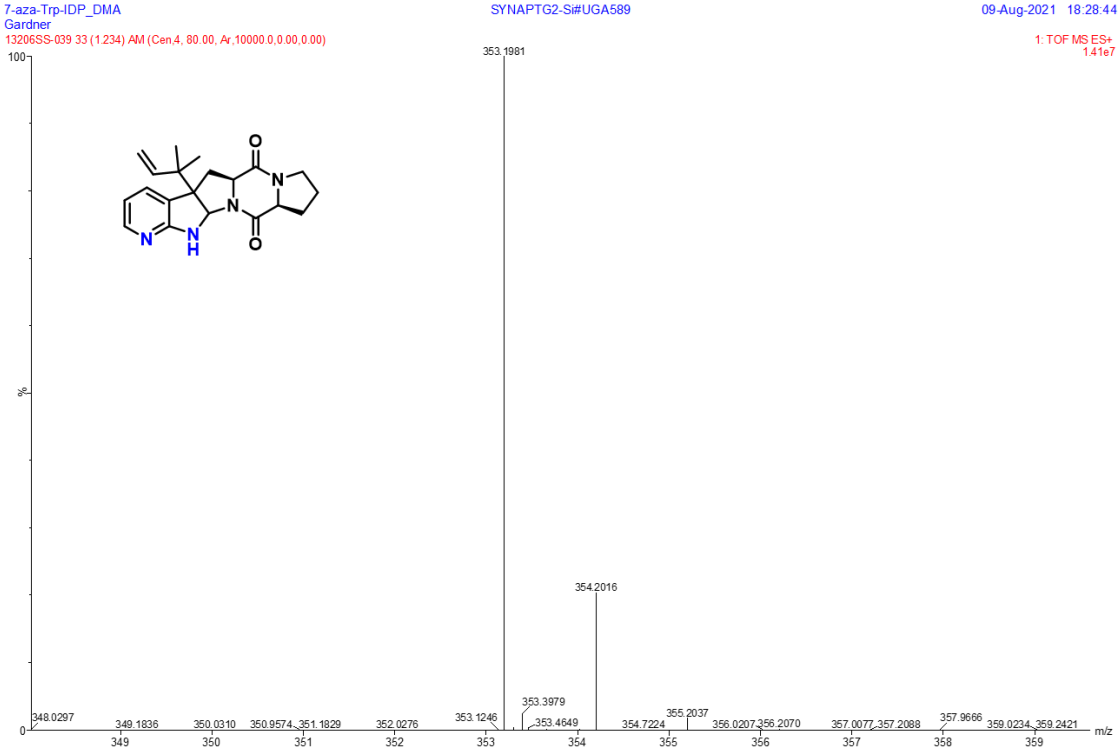
HRMS data for 6-C3



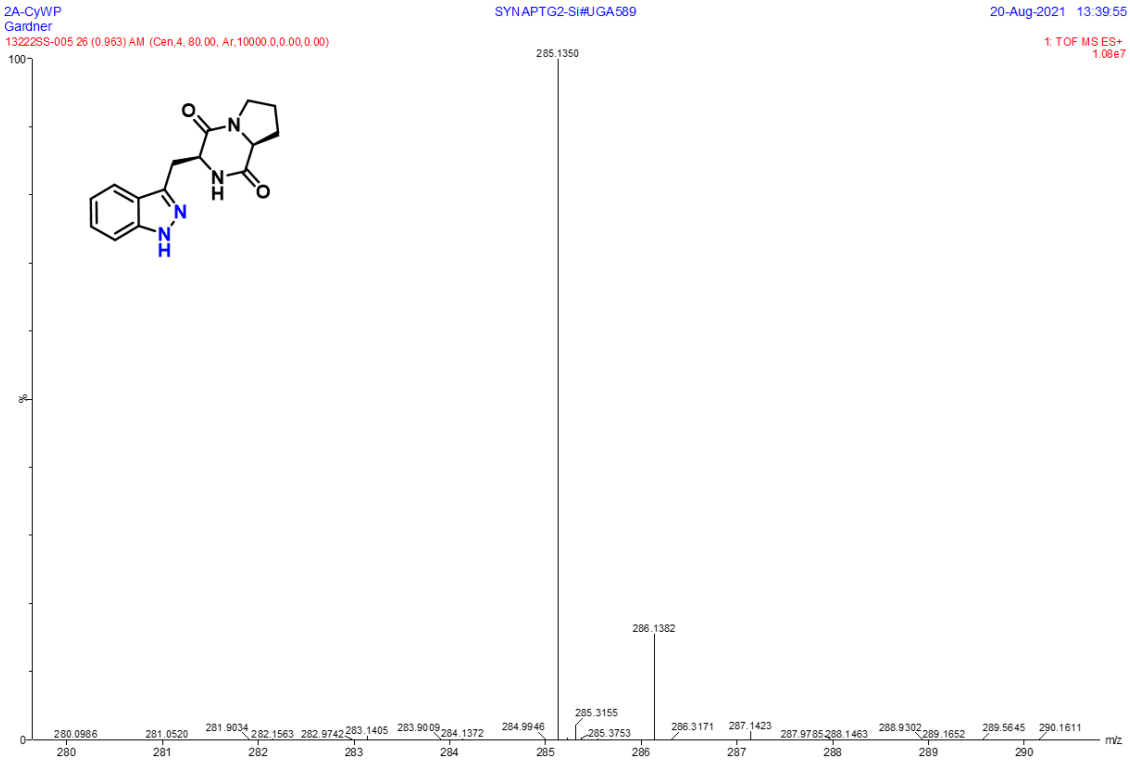
HRMS data for **7-C2**



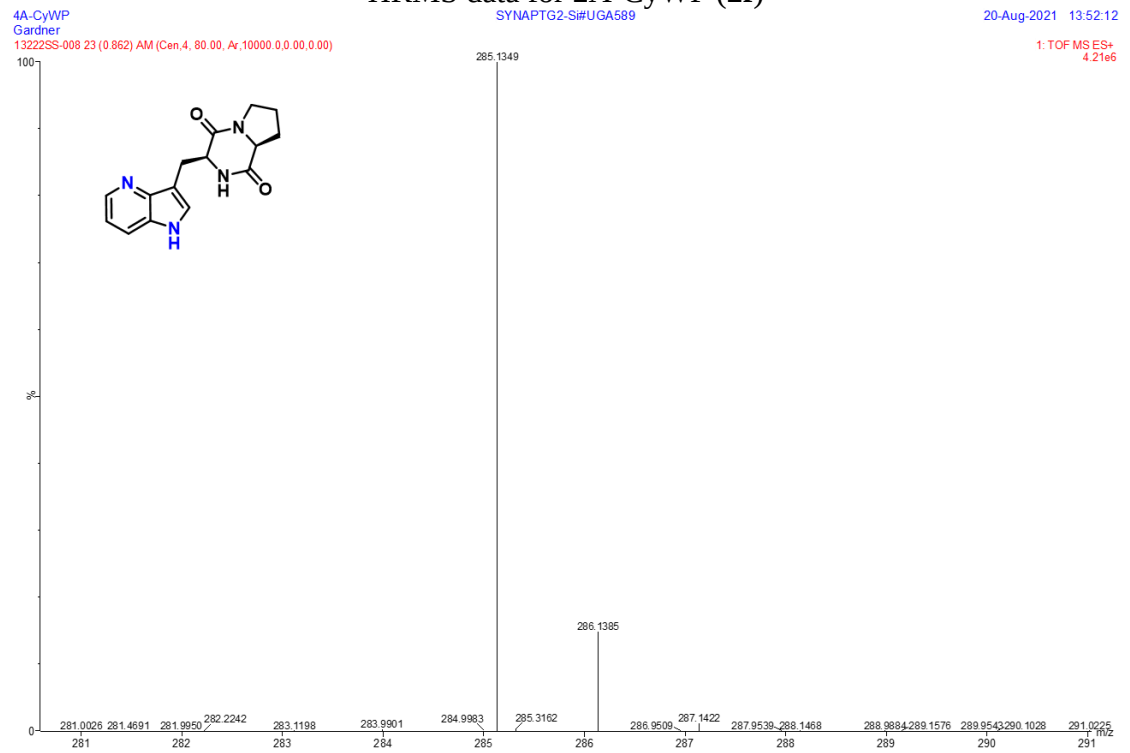
HRMS data for **7-N1**



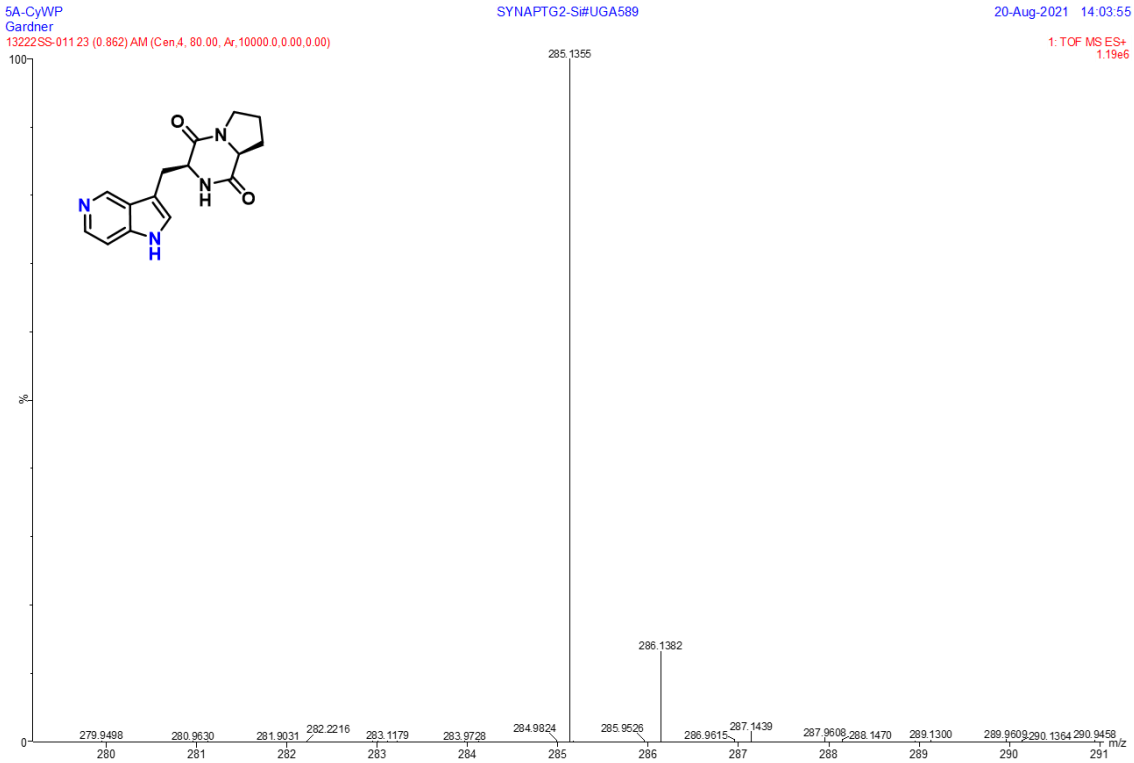
HRMS data for **7-R-C3**



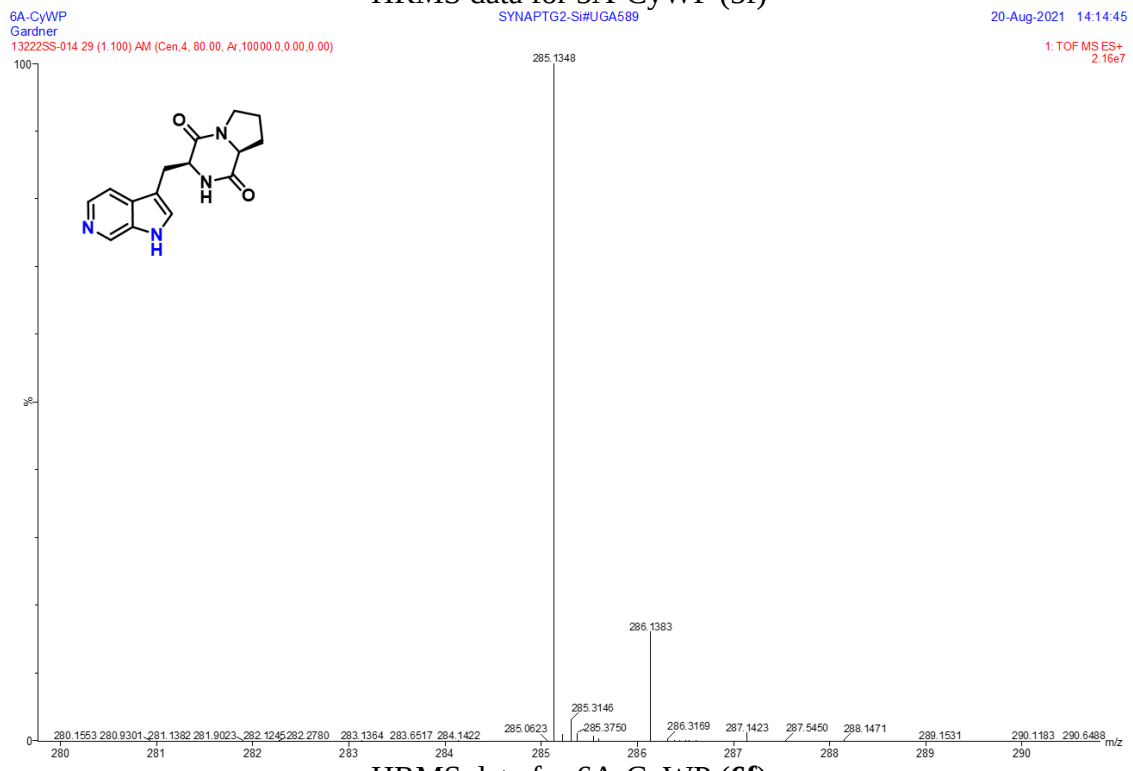
HRMS data for 2A-CyWP (**2f**)



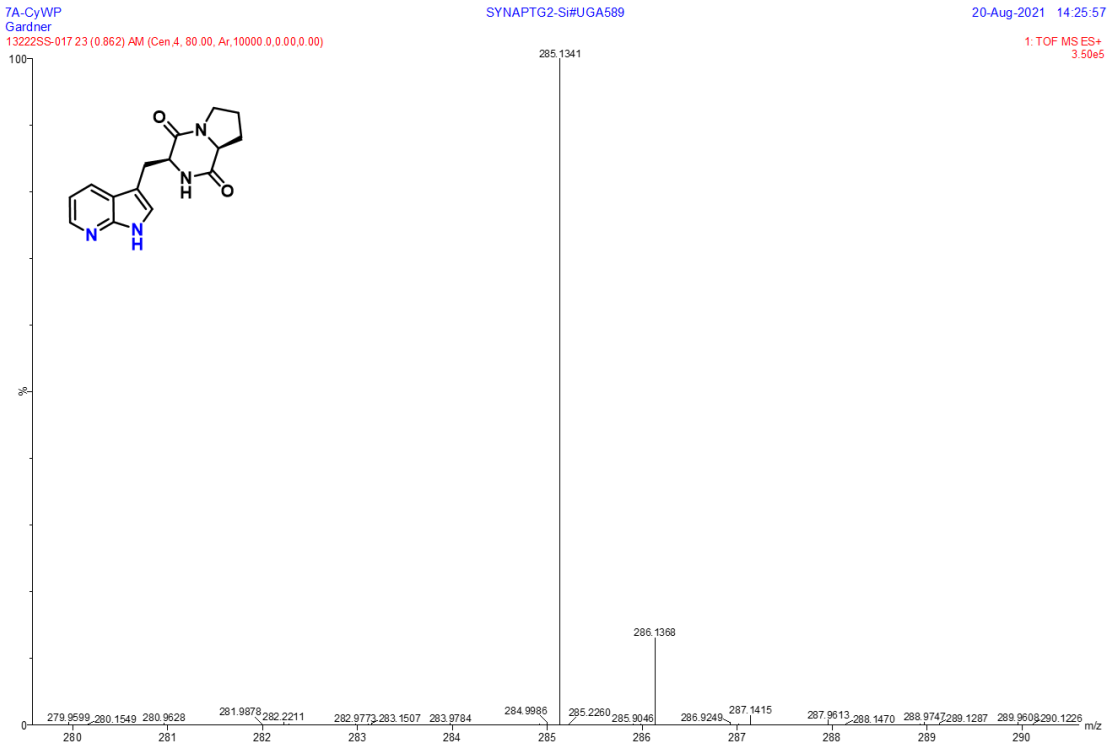
HRMS data for 4A-CyWP (**4f**)



HRMS data for 5A-CyWP (5f)



HRMS data for 6A-CyWP (**6f**)



HRMS data for 7A-CyWP (**7f**)

4.7 References for Chapter 4

1. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs from 1981 to 2014. *J Nat Prod* **2016,** *79* (3), 629-61.

2. Vitaku, E.; Smith, D. T.; Njardarson, J. T., Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen Heterocycles among U.S. FDA Approved Pharmaceuticals. *Journal of Medicinal Chemistry* **2014,** *57* (24), 10257-10274.

3. Ma, Y. M.; Liang, X. A.; Kong, Y.; Jia, B., Structural Diversity and Biological Activities of Indole Diketopiperazine Alkaloids from Fungi. *J Agric Food Chem* **2016,** *64* (35), 6659-71.

4. Cui, C.-B.; Kakeya, H.; Okada, G.; Onose, R.; Osada, H., Novel Mammalian Cell Cycle Inhibitors, Tryprostatins A, B and Other Diketopiperazines Produced by Aspergillus fumigatus. I. Taxonomy, Fermentation, Isolation and Biological Properties. *The Journal of Antibiotics* **1996,** *49* (6), 527-533.

5. Lavey, N. P.; Coker, J. A.; Ruben, E. A.; Duerfeldt, A. S., Sclerotiamide: The First Non-Peptide-Based Natural Product Activator of Bacterial Caseinolytic Protease P. *Journal of Natural Products* **2016,** *79* (4), 1193-1197.

6. Takahashi, C.; Numata, A.; Ito, Y.; Matsumura, E.; Araki, H.; Iwaki, H.; Kushida, K., Leptosins, antitumour metabolites of a fungus isolated from a marine alga. **1994,** (13), 1859.

7. Elsebai, M. F.; Rempel, V.; Schnakenburg, G.; Kehraus, S.; Müller, C. E.; König, G. M., Identification of a Potent and Selective Cannabinoid CB1 Receptor Antagonist from Auxarthron reticulatum. *ACS Med Chem Lett* **2011,** *2* (11), 866-869.

8. Cao, J.; Wang, B.-G., Chemical diversity and biological function of indolediketopiperazines from marine-derived fungi. *Marine Life Science & Technology* **2020,** *2* (1), 31-40.

9. Liao, S.; Qin, X.; Li, D.; Tu, Z.; Li, J.; Zhou, X.; Wang, J.; Yang, B.; Lin, X.; Liu, J.; Yang, X.; Liu, Y., Design and synthesis of novel soluble 2,5-diketopiperazine derivatives as potential anticancer agents. *European Journal of Medicinal Chemistry* **2014,** *83*, 236-244.

10. Ando, S.; Grote, A. L.; Koide, K., Diastereoselective Synthesis of Diketopiperazine Bis-α,β-epoxides. *The Journal of Organic Chemistry* **2011,** *76* (4), 1155-1158.

11. Macdonald, J. C.; Whitesides, G. M., Solid-State Structures of Hydrogen-Bonded Tapes Based on Cyclic Secondary Diamides. *Chemical Reviews* **1994,** *94* (8), 2383-2420.

12. Santos, A.; Mortinho, A.; Marques, M., Metal-Catalyzed Cross-Coupling Reactions on Azaindole Synthesis and Functionalization. *Molecules* **2018,** *23* (10), 2673.

13. Song, J. J.; Reeves, J. T.; Gallou, F.; Tan, Z.; Yee, N. K.; Senanayake, C. H., Organometallic methods for the synthesis and functionalization of azaindoles. *Chemical Society Reviews* **2007,** *36* (7), 1120.

14. Kannaboina, P.; Mondal, K.; Laha, J. K.; Das, P., Recent advances in the global ring functionalization of 7-azaindoles. *Chemical Communications* **2020,** *56* (79), 11749-11762.

15. Subota, A. I.; Volochnyuk, D. M.; Gorlova, A. O.; Grygorenko, O. O., Scalable synthesis and properties of 7-methyl- 4-azaindole. *Heterocyclic Communications* **2017,** *23* (6).

16. Motati, D. R.; Amaradhi, R.; Ganesh, T., Recent developments in the synthesis of azaindoles from pyridine and pyrrole building blocks. *Organic Chemistry Frontiers* **2021,** *8* (3), 466-513.

17. Schneider, C.; David, E.; Toutov, A. A.; Snieckus, V., In Situ Anionic Shielding for Regioselective Metalation: Directed peri and Iterative Metalation Routes to Polyfunctionalized 7-Azaindoles. *Angewandte Chemie International Edition* **2012,** *51* (11), 2722-2726.

18. Dalziel, M. E.; Patel, J. J.; Kaye, M. K.; Cosman, J. L.; Kitching, M. O.; Snieckus, V., Regioselective Functionalization of 7-Azaindole by Controlled Annular Isomerism: The Directed Metalation-Group Dance. *Angewandte Chemie International Edition* **2019,** *58* (22), 7313-7317.

19. L'Heureux, A.; Thibault, C.; Ruel, R., Synthesis of functionalized 7-azaindoles via directed ortho-metalations. *Tetrahedron Letters* **2004,** *45* (11), 2317-2319.

20. Chi, S. M.; Choi, J.-K.; Yum, E. K.; Chi, D. Y., Palladium-catalyzed functionalization of 5- and 7-azaindoles. *Tetrahedron Letters* **2000,** *41* (6), 919-922.

21. Liu, C.-F.; Zhang, G.-T.; Sun, J.-S.; Dong, L., Access to π-conjugated azaindole derivatives via rhodium(iii)-catalyzed cascade reaction of azaindoles and diazo compounds. *Organic & biomolecular chemistry* **2017,** *15* (14), 2902-2905.

22. Baykov, S. V.; Boyarskiy, V. P., Metal-Free Functionalization of Azine N-Oxides with Electrophilic Reagents. *Chem Heterocycl Com+* **2020,** *56* (7), 814-823.

23. Pati, B. V.; Sagara, P. S.; Ghosh, A.; Mohanty, S. R.; Ravikumar, P. C., Ruthenium-Catalyzed Cross Dehydrogenative Annulation of N-(7-Azaindole)benzamides

with Maleimides: One-Step Access to Highly Functionalized Pyrroloisoquinoline. *The Journal of Organic Chemistry* **2021,** *86* (9), 6551-6565.

24. Wollinsky, B.; Ludwig, L.; Xie, X.; Li, S.-M., Breaking the regioselectivity of indole prenyltransferases: identification of regular C3-prenylated hexahydropyrrolo[2,3-b]indoles as side products of the regular C2-prenyltransferase FtmPT1. *Organic & biomolecular chemistry* **2012,** *10* (46), 9262.

25. Grundmann, A.; Li, S. M., Overproduction, purification and characterization of FtmPT1, a brevianamide F prenyltransferase from Aspergillus fumigatus. *Microbiology* **2005,** *151* (Pt 7), 2199-207.

26. Mahmoodi, N.; Tanner, M. E., Potential rearrangements in the reaction catalyzed by the indole prenyltransferase FtmPT1. *Chembiochem* **2013,** *14* (15), 2029-37.

27. Zou, H.; Zheng, X.; Li, S.-M., Substrate Promiscuity of the Cyclic Dipeptide Prenyltransferases from Aspergillus fumigatus. *Journal of Natural Products* **2009,** *72* (1), 44-52.

28. Zhao, W.; Fan, A.; Tarcz, S.; Zhou, K.; Yin, W.-B.; Liu, X.-Q.; Li, S.-M., Mutation on Gly115 and Tyr205 of the cyclic dipeptide C2-prenyltransferase FtmPT1 increases its catalytic activity toward hydroxynaphthalenes. *Applied Microbiology and Biotechnology* **2017,** *101* (5), 1989-1998.

29. Chen, J.; Morita, H.; Wakimoto, T.; Mori, T.; Noguchi, H.; Abe, I., Prenylation of a Nonaromatic Carbon of Indolylbutenone by a Fungal Indole Prenyltransferase. *Organic letters* **2012,** *14* (12), 3080-3083.

30. Gardner, E. D.; Dimas, D. A.; Finneran, M. C.; Brown, S. M.; Burgett, A. W.; Singh, S., Indole C6 Functionalization of Tryprostatin B Using Prenyltransferase CdpNPT. *Catalysts* **2020,** *10* (11), 1247.

31. Scull, E. M.; Bandari, C.; Johnson, B. P.; Gardner, E. D.; Tonelli, M.; You, J.; Cichewicz, R. H.; Singh, S., Chemoenzymatic synthesis of daptomycin analogs active against daptomycin-resistant strains. *Applied Microbiology and Biotechnology* **2020**.

32. Elshahawi, S. I.; Cao, H.; Shaaban, K. A.; Ponomareva, L. V.; Subramanian, T.; Farman, M. L.; Spielmann, H. P.; Phillips Jr, G. N.; Thorson, J. S.; Singh, S., Structure and specificity of a permissive bacterial C-prenyltransferase. *Nat Chem Biol* **2017,** *13* (4), 366-368.

33. Yin, W. B.; Ruan, H. L.; Westrich, L.; Grundmann, A.; Li, S. M., CdpNPT, an N-prenyltransferase from Aspergillus fumigatus: overproduction, purification and biochemical characterisation. *Chembiochem* **2007,** *8* (10), 1154-61.

34. Unsold, I. A.; Li, S. M., Overproduction, purification and characterization of FgaPT2, a dimethylallyltryptophan synthase from Aspergillus fumigatus. *Microbiology* **2005,** *151* (Pt 5), 1499-505.

35. Steffan, N.; Unsöld, I. A.; Li, S.-M., Chemoenzymatic Synthesis of Prenylated Indole Derivatives by Using a 4-Dimethylallyltryptophan Synthase from Aspergillus fumigatus. *ChemBioChem* **2007,** *8* (11), 1298-1307.

36. Yin, W. B.; Cheng, J.; Li, S. M., Stereospecific synthesis of aszonalenins by using two recombinant prenyltransferases. *Organic & biomolecular chemistry* **2009,** *7* (10), 2202-7.

37. Wollinsky, B.; Ludwig, L.; Hamacher, A.; Yu, X.; Kassack, M. U.; Li, S.-M., Prenylation at the indole ring leads to a significant increase of cytotoxicity of tryptophan-containing cyclic dipeptides. *Bioorganic & Medicinal Chemistry Letters* **2012,** *22* (12), 3866-3869.

38. Yu, X.; Xie, X.; Li, S.-M., Substrate promiscuity of secondary metabolite enzymes: prenylation of hydroxynaphthalenes by fungal indole prenyltransferases. *Applied Microbiology & Biotechnology* **2011,** *92* (4), 737-748.

39. Lee, H.-Y.; Yerkes, N.; O'Connor, S. E., Aza-Tryptamine Substrates in Monoterpene Indole Alkaloid Biosynthesis. *Chemistry & Biology* **2009,** *16* (12), 1225-1229.

40. Luk, L. Y. P.; Qian, Q.; Tanner, M. E., A Cope Rearrangement in the Reaction Catalyzed by Dimethylallyltryptophan Synthase? *Journal of the American Chemical Society* **2011,** *133* (32), 12342-12345.

41. Jost, M.; Zocher, G.; Tarcz, S.; Matuschek, M.; Xie, X.; Li, S. M.; Stehle, T., Structure-function analysis of an enzymatic prenyl transfer reaction identifies a reaction chamber with modifiable specificity. *J Am Chem Soc* **2010,** *132* (50), 17849-58.