Towards Practical Biocatalysis:
Characterization and Engineering of Synthetically Useful Enzymes

by

Allwin Daniel McDonald

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Chemistry)

at the
UNIVERSITY OF WISCONSIN-MADISON
2022

Date of final oral examination: 08/17/2022

The dissertation is approved by the following members of the Final Oral Committee:

Andrew R. Buller, Assistant Professor, Chemistry
Brian F. Pfleger, Professor, Chemical and Biological Engineering
Tim S. Bugni, Professor, Pharmacy
Tina Wang, Assistant Professor, Chemistry

Dedication

A reaction a day
Keeps the thesis at bay
So make sure to make time for two.

Through the good and the bad
Grad school made me so glad,
Truly, **it was all thanks to you**.

I dedicate this thesis to my great-grandmother, Lucille, who was a fount of both joy and wisdom, without whom I would not be who I am today.

Acknowledgements

Obtaining a Ph.D. takes a great deal of time, motivation, and commitment. I have heard it said that a doctoral degree is worth the effort to achieve it. I agree. Luckily, much of the energy needed to complete my degree was supplied by my mentors, colleagues, friends, and family. I will take some moments here to acknowledge those who have aided me in my Ph.D. journey.

First, I want to acknowledge my advisor, Professor Andrew Buller. Andrew's excitement and passion for science is contagious, and my five years learning with and alongside Andrew have only increased my own interest in pursuing science. I have grown, both as a scientist and as a person, thanks to Andrew's mentorship throughout my doctoral studies. Andrew's high standards have pushed me to excel and to pursue 'excellence' instead of stopping at 'good.' The training I received as a member of the Buller Lab will carry with me always.

As an inaugural member of the Buller Lab, I also started my Ph.D. alongside wonderful colleagues. Prasanth, Lydia, and Jon are a group of stupendous scientists and human beings, and I am privileged to have spent five years learning and conducting science with them. I cannot imagine what sailing the scientific seas would have been if not for their guidance and support. I also want to acknowledge the equally wonderful members of the Buller Buddies who joined since the lab was founded. First, I want to acknowledge Peyton, my stalwart collaborator and comrade on our long-term project together. It has been wonderful to watch you grow as a scientist and just excel at all you put your mind to. For risk of spending pages acknowledging all the members of the lab, I want to mention just a few more, such as my 'archnemesis' Meghan and colleague Tony as folks whom I will dearly missing having both life and scientific conversations with. To my mentees, Aadhishre and Sam, I feel truly blessed to have been able to be a part of your own scientific journeys and I cannot wait to watch your success down the road.

I enjoyed the privilege of having had many friends who supported me through my time here in Madison. Rachel has been my friend and running buddy since my first days in Madison, and I want to acknowledge this friendship. A shout-out here also deserves to go to my running crew of Rachel, Susan, and Josh! You guys have helped put lots of pep in my step throughout the years. I will also highlight my climbing partner Sidd, as well as fellow climbing comrades Nick, Jenny, Nate, and Christine and thank them for their continuous motivational support. I would like to thank my friend and roommate Tahoe for helping me survive lockdown and providing many laughs and lively conversations.

Lastly, I want to dedicate time to thank my family. My brother, Seth, has always been there for me, and I look forward to continuing to travel and explore the world with you. My sister, Lauren, has grown up so much during my tenure in graduate school, and I wish you the best as you pursue a dental career! My parents have been so supportive of my Ph.D. pursuits, and I thank you for your love and support. Additionally, to my grandparents and aunts and uncles, I want to say, ‘thank you!’ A caring family has made my life that much better.

Table of Contents

Abstract

Enzymes are attractive catalysts for synthetic applications owing to their high chemo-, regio-, and stereoselectivity, mild reaction conditions, and biodegradability. However, the use of enzymes for biocatalysis is hampered by this same specificity, limiting overall utility. This work has focused on exploring and expanding biocatalyst substrate scopes. Using a broadly active tryptophan decarboxylase enzyme as a model system, I have developed methodology to facilitate protein engineering. Here, I discuss the development of this protein engineering method, the application of this method to engineer enzymes for new activities, and the usage of the engineered enzymes to access medicinally relevant tryptamines and 1,2-amino alcohols.

Chapter 1

Introduction

Chapter 1: Introduction

1. 1. Evolutionary pressures dictate enzymatic function

Evolutionary pressures over millions of years have fine-tuned the functions of enzymes, resulting in enzymes with specific activities suited for their cellular roles. Understanding and defining these roles has long been a chief goal of the field of enzymology. However, deciphering the connection between amino acid sequence and enzymatic activity is challenging (Fig 1).[1] One way enzymologists study the role of a specific residue in an enzyme is by swapping in other amino acids in its place.[2] Most mutations are deleterious to the original enzyme function,[3] and even small amino acid perturbations can greatly impact the ability of enzymes to effect a given transformation. Should a mutation significantly diminish activity, this is a signal that the given amino acid position is vital for enzymatic function or for 3D structure, while a minor effect implicates the site as being less critical. Conversely, some amino acid changes will have little to no impact on activity. In this way, mutation to the existing sequence can impart information about the structure-function relationship of the enzyme of interest. Accurately predicting the role of individual amino acids on enzyme activity is non-trivial, and the causal relationship between a

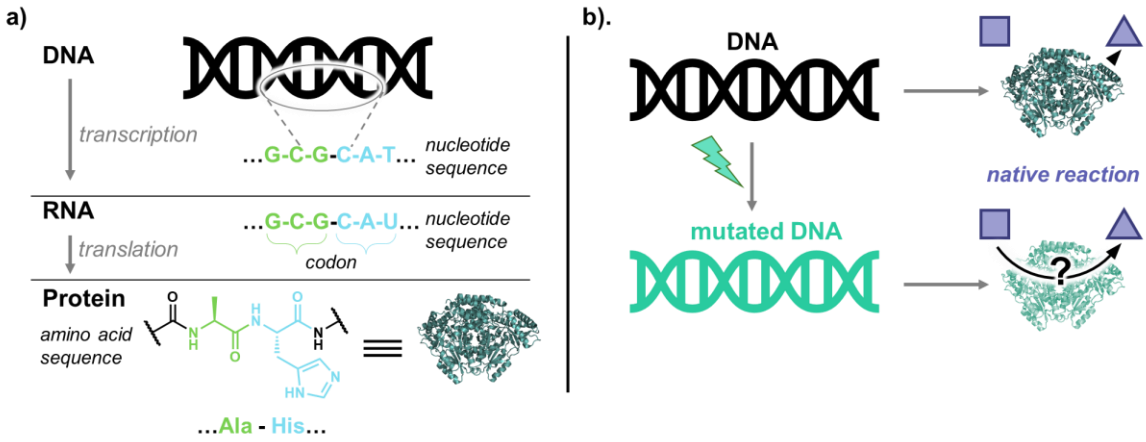


Figure 1. The central dogma of biology. a) Proteins have their amino acid sequences genetically encoded in deoxyribonucleic acids (DNA) as nucleotides. This DNA sequence is transcribed into a ribonucleic acid (RNA) sequence, with groups of three nucleotides (a codon) acting as the code for subsequent translation into proteins as amino acid sequences. **b)** The genetic sequence of DNA can be mutated, resulting in a different nucleotide sequence. The encoded protein from this mutated DNA sequence will have uncharacterized activity on its native reaction.

given mutation and the resulting change in enzymatic activity is often obscure (Fig 1B).[1] As further mutations are made, new functions can emerge that are distinct from the parent activity and occupy a different location in protein sequence space.

One reason behind the difficulty in deciphering structure-function relationships is the vastness of amino acid sequence space. As an example, the median length of proteins in yeast is ~400, which corresponds to $1x10^{52}$ possible sequence combinations of the 20 standard amino acids.[4] Nature evolves enzyme function not by randomly sampling this sequence space but often by stepwise or pairwise mutation to a DNA sequence of an existing enzyme. By changing the DNA code of the parent, the final amino acid sequence of the enzyme will be subtly altered – or mutated – from the original (Fig 1A). This astronomical number of potential enzymes is such that we could never test every individual sequence. Indeed, even after millions of years, only a tiny fraction of this vast sequence space has been explored by Nature, leaving room for new functions to be evolved continually.

1. 2. Chemists can exploit and engineer enzyme function

Enzymes are responsible for the chemistry of life. The transformations that enzymes catalyze create bonds, break bonds, and switch bonds between atoms, resulting in the diversity of molecules found in Nature. Due to the inherent chirality of the amino acids that make up enzymes and the unique three-dimensional architecture of an enzyme active site, the reactions they catalyze are highly selective. High chemo-, regio-, and stereoselectivity are all hallmarks of enzymatic catalysis,[5,6] and are characteristics that are vital to the fulfillment of their biological roles. Although many of these characteristics were evolved for a cellular context, chemists have long appreciated the power of enzymatic catalysis for non-native chemical transformations.

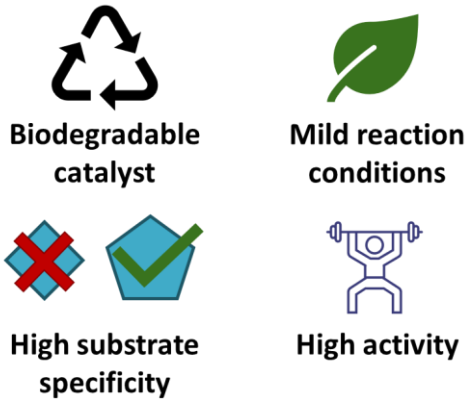


Figure 2. Advantages of biocatalysts.

The field of biocatalysis is devoted to the application of enzymes to make molecules in non-native contexts. This interdisciplinary pursuit leverages the high activity and specificity of enzymes for synthesis (Fig 2).[7,8] Environmental and safety factors, such as the biodegradability of enzymes and the ability to use water as a solvent, also promote biocatalysts as green alternatives to more traditional synthetic chemistry.[5,9] The use of enzymes as biocatalysts has a long history, going back to the 1850s, where Louis Pasteur exploited the stereospecific biological degradation of tartaric acid to obtain enantiopure (-)-tartaric acid.[10] Enzymes, such as hydroxynitrile lyases and lipases, have a long history of applications for stereoselective

transformations[11,12] However, the same specificity that draws chemists to enzymes is often accompanied by a limitation in the scope of substrates that can engage an enzyme. Often, an enzyme will have high activity with only a handful of closely related substrates. If activity on different substrates is desired, an alternative is to identify existing enzymes in Nature with similar activity to investigate if a different enzyme is a more suitable biocatalyst.[13,14]

In the post-genomic era, chemists can now mutate an enzyme's sequence to improve the desired property. This strategy of intentionally altering enzyme sequence to improve an enzymatic characteristic is called protein engineering (Fig 3). Protein engineering allows the enzymatic function to be altered, expanded, and improved for both native and non-native reactions, increasing the utility of biocatalysts. Additionally, properties such as protein expression and stability can be improved, further increasing their utility.[15,16] The process of protein engineering starts with a parent gene sequence that encodes the protein of interest. Genetic mutations are then introduced, resulting in a library or collection of mutants. These DNA fragments are transformed into a host system (in many cases, *E. coli,* a commonly used bacterium to produce proteins). These cells then express the target protein, which is next screened for the desired property. This constitutes one round of protein engineering. If further engineering is warranted, the most improved protein variant can be used as the new parent sequence for subsequent rounds of engineering. This iterative process is called directed evolution and mimics natural evolution, allowing for larger deviations in activity over the course of engineering.[16,17] Seminal examples of directed evolution have improved enzyme

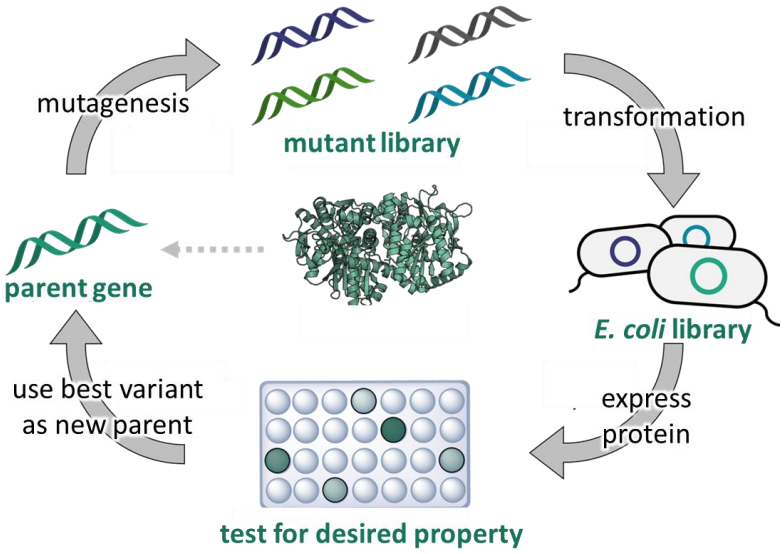


Figure 3. Protein engineering cycle. Biocatalysts can be improved through protein engineering.

stability,[18] improved enzyme activity,[19] and resulted in wholly new functions.[20,21] Through this mode of protein engineering, the available repertoire of potential biocatalytic transformations is greatly expanded.

Many commercially relevant biocatalysts have been heavily engineered to increase their utility. Engineered enzymes such as glucose isomerase or lactase, which break down sugars; proteases, which break down proteins; and lipases, which break down fats, are staples of the food and detergent industries (Fig 4).[22] The stereoselectivity of biocatalysts is often leveraged to interconvert between functional groups. These functional group interconversions (FGI's) have found applications in small molecule synthesis, such as for medicinal compounds, due to their high regio- and stereoselectivity. The biosyntheses of sitagliptin, molnupiravir, and islatravir (among others) relied upon engineered enzymes and highlight the power of biocatalytic transformations for these 'simple' chemical steps (Fig 5).[23–25]

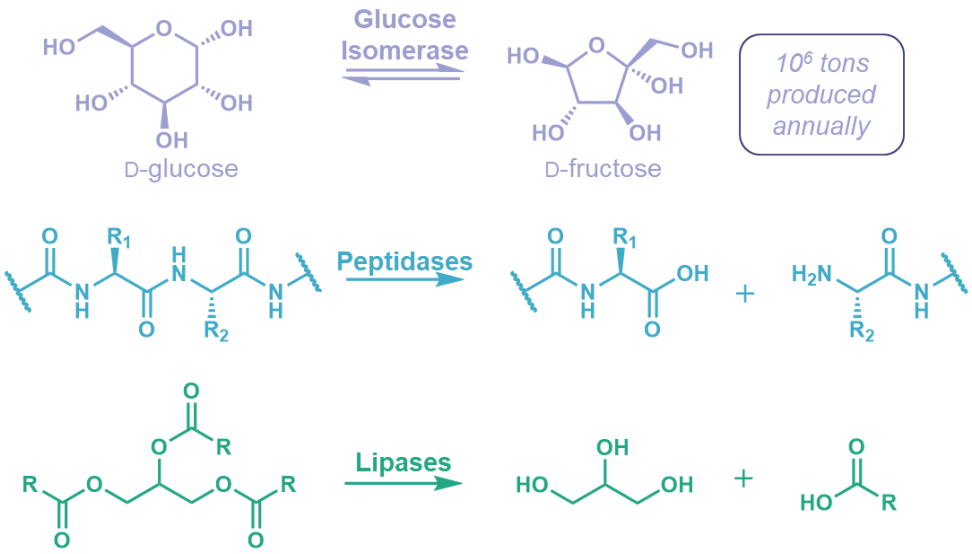


Figure 4. Common commercially used biocatalysts. Specific (glucose isomerase) or generic (peptidase/lipase) reactions are depicted for these common commercial biocatalysts.

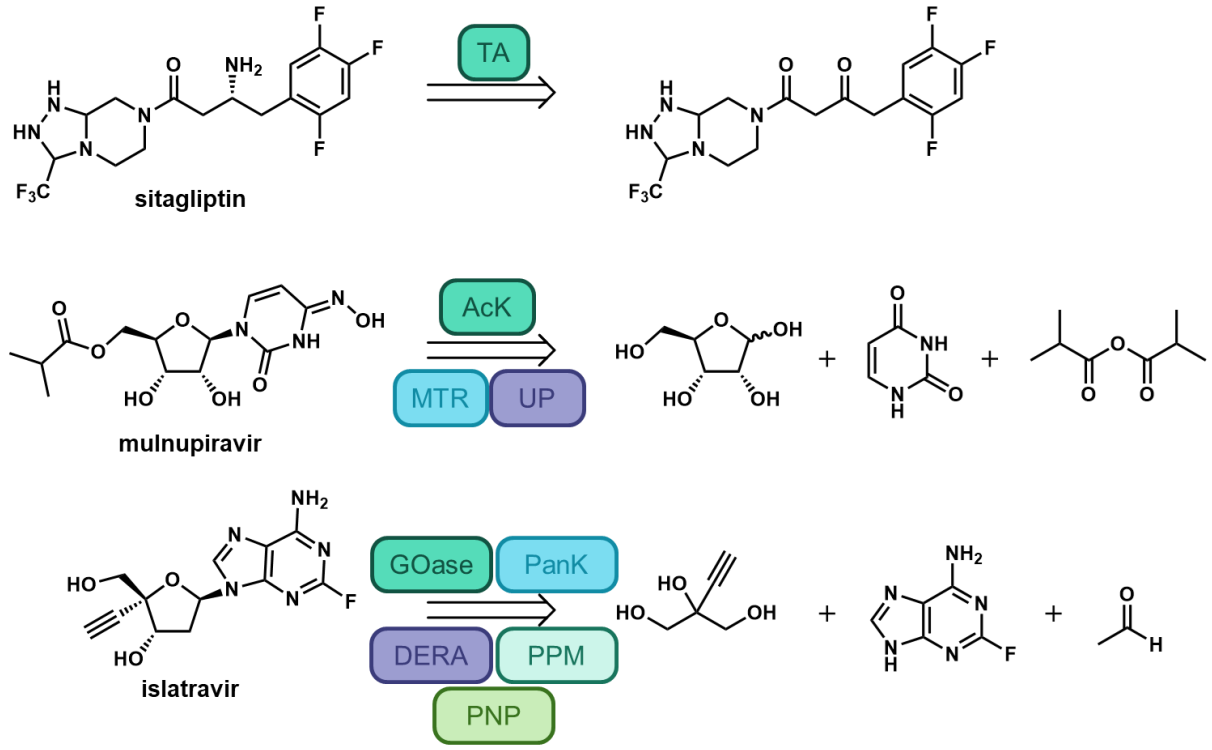


Figure 5. Notable commercial biocatalytic syntheses. Simplified retrosynthetic analysis of three pharmaceutical compounds. Engineered biocatalysts used for these transformations are depicted in colored boxes. TA = transaminase; AcK = acetyl kinase; MTR = 5-S-methylthioribose kinase; UP = uridine phosphorylase; GOase = galactose oxidase; PanK = pantothenate kinase; DERA = deoxyribose 5-phosphate aldolase; PPM = phosphopentomutase; PNP = purine nucleoside phosphorylase.

1. 3. Amino acid synthases perform desirable transformations

Carbon-to-carbon (C-C) bond formation is an example of an often more challenging chemical step than FGIs. This difficulty arises from the specific conditions to generate reactive yet water-stable carbon nucleophiles and electrophiles. Such reactions require finely tuned enzymatic active sites that shield reactive intermediates from unwanted reactions with water.[26] By leveraging the hydrophobic protein interior as a microreactor, enzymes can catalyze reactions that are otherwise challenging in aqueous solution. Therefore, instead of being used to tailor the functional groups of an assembled carbon framework, C-C bond-forming enzymes have the potential to build complex molecules from simple building blocks. Many C-C bond

forming enzymes, such as aldolases and hydroxynitrile lyases, have been explored for their synthetic utility.[11,27] Amino acid synthases are a class of enzymes that form C-C or C-X bonds and are natively responsible for the biosynthesis of biological amino acids. The naturally occurring variety of both standard and non-standard amino acids that these enzymes produce underscores their synthetic potential to produce compounds relevant to the medicinal field.

Amino acid synthases have found widespread use as biocatalysts. Many such synthases use the colored pyridoxal-5'-phosphate (PLP) cofactor and have biological roles in central metabolism (Fig 6). Cofactors are small molecules that will preferentially bind to enzymes and aid in catalysis. Often, these binding interactions are so specific that such cofactors are simply considered part of the enzyme scaffold. PLP is a widespread cofactor in biology and is crucial for amino acid biosynthesis in primary metabolism.[28] PLP-dependent enzymes perform numerous types of transformations, such as transamination, decarboxylation, and oxidation.[29] However, their role as synthases involved in C-C bond formation is among their most powerful transformations.

Here I will discuss some of the more prominent synthases used for non-standard amino acid syntheses, although numerous promiscuous amino acid synthases have been reported in recent years.[30–34] Tryptophan synthase (TrpS) is one of the most notable synthases in this class. TrpS is a heterodimeric enzyme comprised of two copies of α- and β-subunits. The α-subunit of the TrpS enzyme natively converts indole-3-glycerol phosphate into indole, and then the β-subunit couples the amino acid serine (Ser) with indole to produce tryptophan (Trp).[35] The native enzyme can also react directly with some indole analogs to produce non-standard Trps.[36] This normally heterodimeric enzyme was engineered for standalone activity with just the β-subunit, known as TrpB.[37] TrpB was evolved for further improved activity with diverse indole analogs to generate a wide array of Trp analogs.[19] TrpB also has been shown to produce non-Trp amino acids,[38] highlighting this enzyme as a versatile and broadly active biocatalyst. This

enzyme's ability to generate complex and valuable amino acids further underscores its synthetic potential.

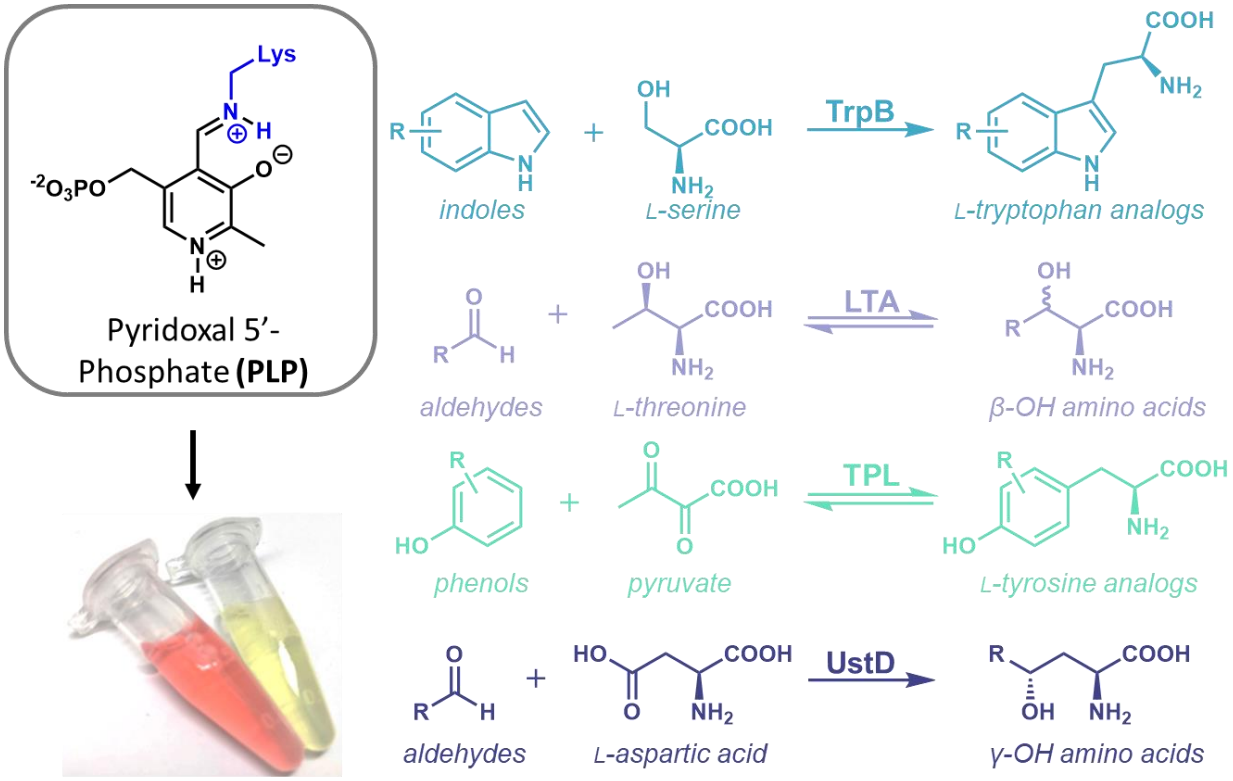


Figure 6. Promiscuous PLP-dependent amino acid synthases. The colored solutions in the bottom-left are depictions of colors of enzyme-bound PLP.

L-threonine aldolases (TA's) have also been used to generate a wide array of non-standard amino acids. While TA's natively break-down L-threonine (Thr) into acetaldehyde and glycine, common biocatalytic routes drive this reaction to run in the "reverse" direction, using different aldehyde substrates and excess glycine to synthesize a wide array of β-OH amino acids.[39] These synthases are promiscuous, accepting a diverse set of aldehydes and even non-glycine amino acids, but often suffer from lower diastereoselectivity at the β-OH position.[40] Since β-OH amino acids are building blocks for several medicinally valuable compounds, protein engineering has been exploited to improve the selectivity of the TA reaction, but the

reversible and thermodynamical nature of this reaction is a general limitation. Nevertheless, LTA's have remained an enduring part of the biocatalyst toolbox due to their high substrate tolerance and ability to install the bioactive β-OH moiety.

Other synthases have additionally seen usage as promiscuous amino acid synthases. Tyrosine phenol lyase (TPL) is an enzyme that natively breaks down tyrosine into phenol and pyruvate. Similar to LTA, the reaction of TPL can be run in the reverse direction, synthesizing various tyrosine analogs from substituted phenols.[41] Additionally, a recently discovered enzyme, UstD, was found to catalyze a promiscuous decarboxylative aldol reaction to form a diverse set of γ-hydroxy amino acids.[31] This enzyme was found to have broad activity on many types of aldehydes, and the prerequisite decarboxylation of aspartate prevented reversibility of this reaction, aiding the formation of the desired γ-hydroxy amino acids.

1. 4. Amino acid decarboxylases produce valuable compounds

Manipulation of the amino acid scaffold can lead to medicinally relevant classes of bioactive compounds, and a common mode of amino acid transformation is decarboxylation. Decarboxylation results in bioactive monoamines that serve as neurotransmitters and hormones, notably serotonin and epinephrine (Fig 7). PLP-dependent amino acid decarboxylases (EC 4.1.1; DC's) catalyze these decarboxylations and are a well-studied class of enzymes that have established biological roles in human health and disease.[42–44] This class of enzyme is found across Nature, in biosynthetic pathways in bacteria, plants, and animals.[45] Often, these enzymes act in catabolism or secondary metabolism pathways, possessing high substrate specificity and activity. Recently, numerous studies investigating bacterial DC's have found previously uncharacterized enzymes with broadly promiscuous substrate scopes, underscoring the biocatalytic potential of this class of catalyst, especially when coupled with promiscuous amino acid synthases.

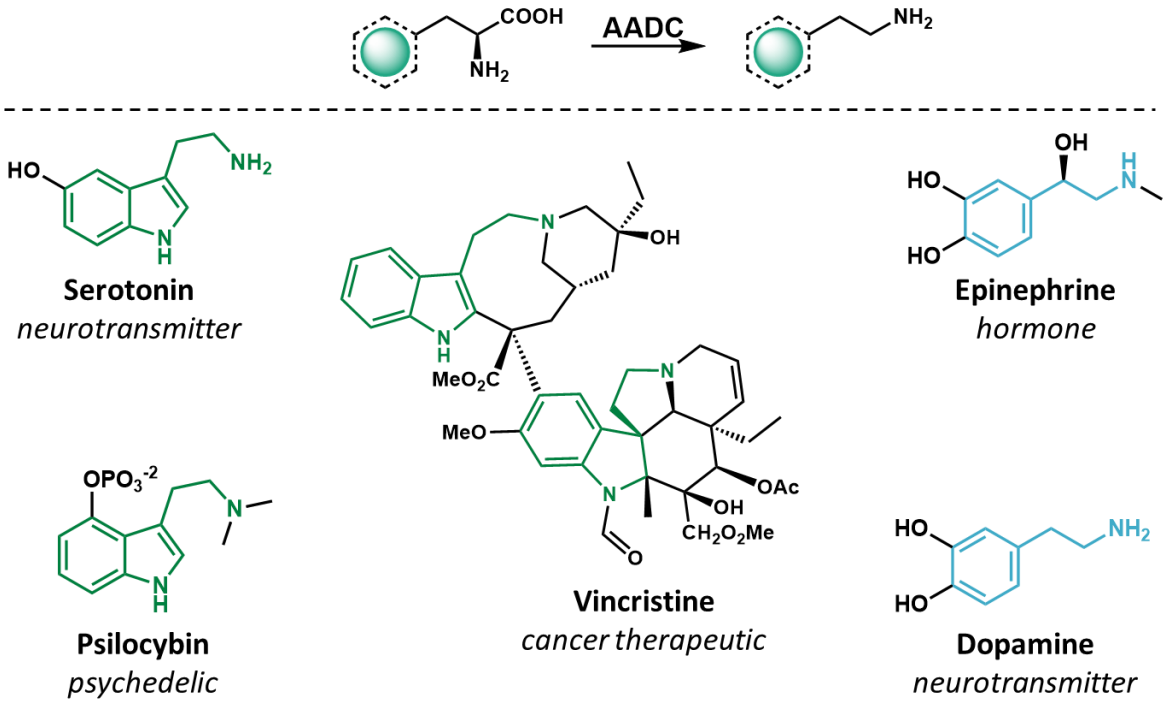


Figure 7. Bioactive naturally occurring monoamines produced via decarboxylation of amino acids. Tryptamine cores are highlighted in green and phenylethylamines are highlighted in blue.

The general mechanism of PLP-dependent DC's has been extensively studied (Fig 8).[46] Initially, the PLP cofactor is bound to a conserved lysine (Lys) residue in the DC active site as the internal aldimine. The amino acid substrate then enters the active site and transamination with the Lys occurs, resulting in the external aldimine. The bound amino acid is decarboxylated, and the resulting negative charge is delocalized into the pyridinium to form the quinonoid intermediate. The quinonoid is subsequently reprotonated with facial selectivity to form the product-bound external aldimine intermediate.[47] At this point, the internal Lys can engage in transimination with the imine, releasing the product. Of note, the amino acid sidechain is not conjugated with the quinonoid intermediate, and so does not participate in the electronic stabilization of the carbanion intermediate. Instead, DC's engage in diverse and poorly understood modes of substrate binding to select one or a small number of amino acid side

chains to bind into the enzyme active site. This underlying selectivity determines the class of amino acid that can be decarboxylated across different classes of DC's.

Aromatic amino acid decarboxylases (AADC's) are a class of DC's that convert the aromatic amino acids (tryptophan, tyrosine, and phenylalanine) into the corresponding primary amines. These classes of decarboxylases are among the most studied, owing to their impact on human health and the regulation of neurotransmitters such as serotonin and melatonin.[42] Due to their central role in acting as 'gatekeepers' into secondary metabolism, many studied AADC's have evolved towards stringent substrate specificity (Fig 9). Notable work from the O'Connor and Weng labs has investigated the low substrate tolerance of plant AADC's.[48,49] This low substrate tolerance has often been cited as a key limitation of AADC biocatalytic utility.

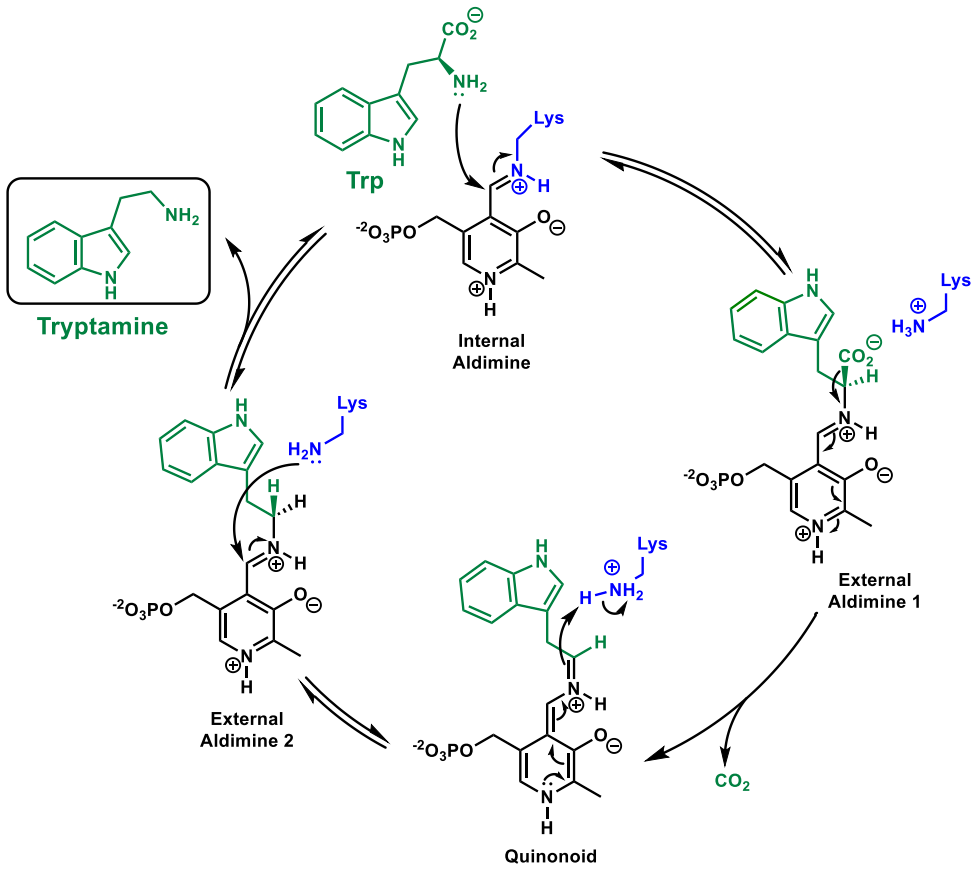


Figure 8. Mechanism of PLP-dependent amino acid decarboxylation.

1. 5. Amino acid decarboxylases have underexplored biocatalytic potential

Although relatively rare, there have been multiple reports of using AADC's for the biosynthesis of small molecules.[46,50] Tryptamines are an often-targeted class of monoamine compounds, owing to their bioactivity and structural complexity and are produced via decarboxylation of tryptophan (Trp). Tryptamines are also the dedicated intermediate to the wider class of indole alkaloid natural products, and thus many biosynthetic pathways making use of tryptophan decarboxylases (TDC's) have been described (Fig 9). The O'Connor group has extensively studied the role of *Catharanthus roseus* TDC (*Cro*TDC) in the production of indole alkaloids *in vivo*.[51] Via enzymatic halogenation of the native Trp substrate, the substrate tolerance of *Cro*TDC in the native indole alkaloid pathway has been investigated.[52] The activity of *Cro*TDC was found to be significantly lower with halogenated substrate analogs, limiting the flux of the pathway to produce halogenated indole alkaloid scaffolds. Significantly, this flux barrier was overcome not by the engineering of the tryptophan decarboxylase, but via the engineering of the tryptophan halogenase to halogenate tryptamine preferentially over Trp.[53] By limiting the *in situ* production of non-native Trp analogs, the overall flux of the pathway increased.

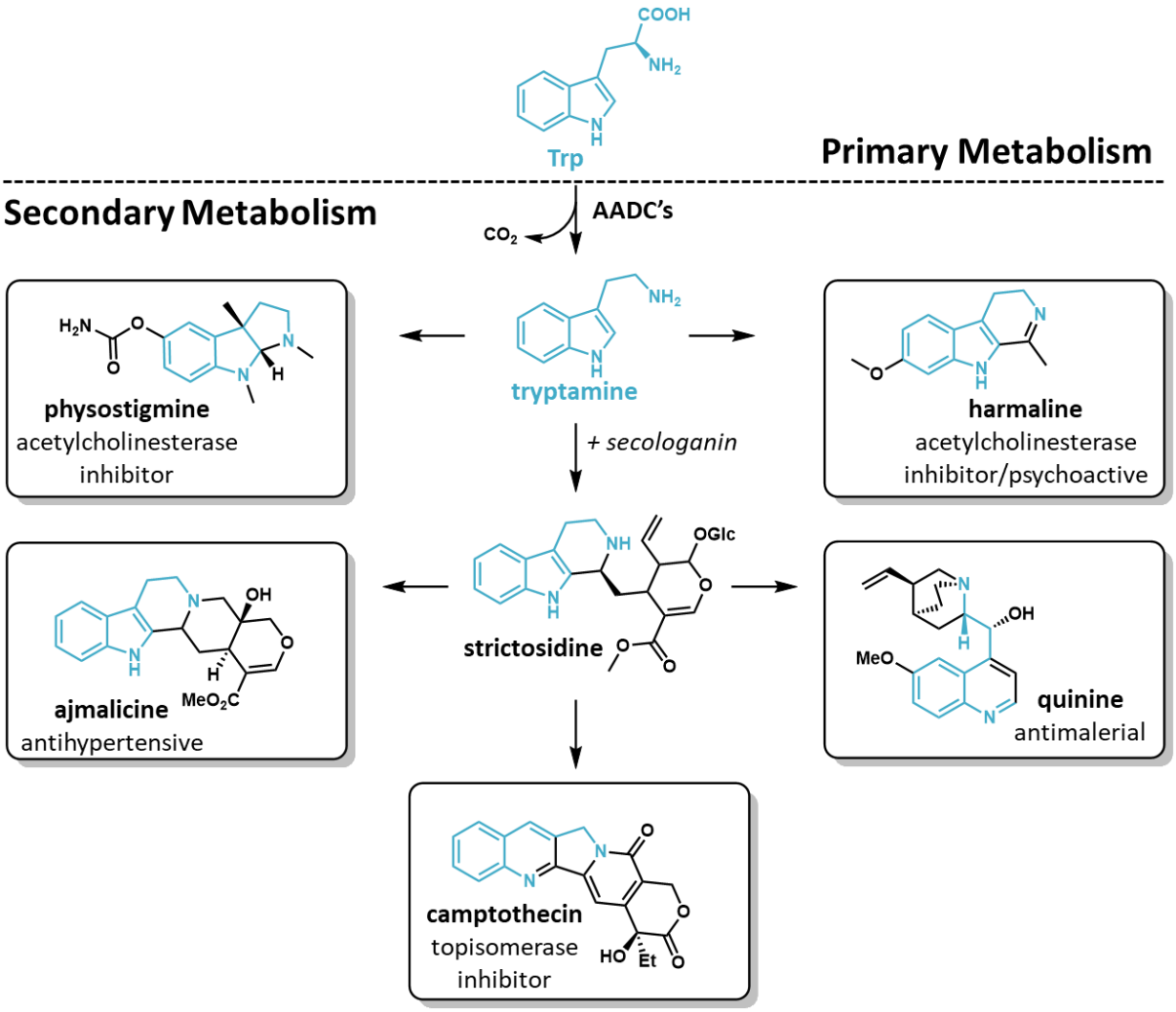


Figure 9. Bioactive indole alkaloids derived from tryptamine. *Figure inspired by Ehrenworth et al. 2017.*[45]

*Cro*TDC was effectively used in the *de novo* synthesis of strictosidine in yeast at 0.5 mg/L,[54] but such decarboxylases have been generally avoided in the production of indole alkaloid analogs. A key limitation of using TDC's for *in vivo* biosynthesis is due to the inherent cytotoxicity of TDC's in decomposing native Trp, which results in tryptamine accumulation. This excess tryptamine has been repeatedly observed to accumulate at much higher levels than targeted tryptamine analogs, primarily due to enzyme preference for native Trp.[55,56] Often, direct feeding of exogenously-produced tryptamine analogs is preferred to avoid native Trp

decarboxylation, circumventing the need for a TDC and increasing overall product titers.[57,58] In fact, a report by the Tang lab used a similar strategy of feeding in exogenous tryptamine for the production of 50 mg/L strictosidine in yeast, efficiently bypassing the decarboxylation step in their biosynthetic pathway.[59]

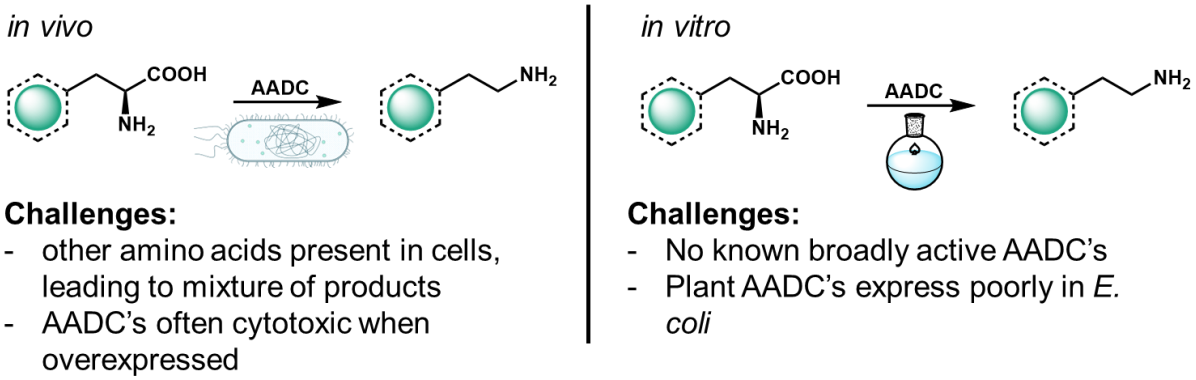


Figure 10. Comparison of *in vivo* vs. *in vitro* AADC biocatalysis.

Plant-derived decarboxylases have seldom been used for *in vitro* biocatalytic reactions due to their stringent specificities. However, new classes of aromatic amino acid decarboxylases have recently been discovered, primarily belonging to the general class of bacterial/fungal AADC's. These enzymes have less defined biochemical functions but have been observed to possess much wider substrate tolerance than their plant AADC counterparts, increasing their *in vitro* potential. Kalb et al. 2016 described the characterization of a fungal tryptophan decarboxylase (*Cs*TDC), which had high activity on native Trp and 5'-OH Trp but did not describe activity on other Trp analogs.[60] A bacterial decarboxylase from *Bacillus atropheus* (AADC-BA) was investigated and found to also possess promiscuous activity on numerous aromatic amino acids.[61,62] AADC-BA was additionally used in the large-scale production of serotonin from 5-OH Trp. Dragulska 2014 successfully employed a phenylalanine decarboxylase from *Streptococcus faecalis* for the synthesis of various tryptamine isotopologues from tryptophans.[63] While these initial results are promising, more work is needed

to determine the utility of fungal and bacterial decarboxylases for the preparative-scale synthesis of tryptamine analogs.

A notable but rare native tryptamine analog is 4-hydroxytryptamine. 4-hydroxytryptamine is the bioactive motif found in psilocin and psilocybin and has received significant interest owing to these molecules' psychoactive properties. The fungal decarboxylases in these biosynthetic pathways have recently been elucidated, and work by Hoffmeister and Weng has focused on understanding the pathways towards 4-hydroxylated tryptamines *en route* to psilocybin.[64,65] The 4-OH-tryptophan decarboxylase PsiD has been used for the preparative synthesis of 4-hydroxylated tryptamines, although only at the 1-3 mg scale.[66,67]

A powerful approach to using DC's is in an *in vitro* cascade format, utilizing the thermodynamic driving force of decarboxylation to improve flux. Griengl and coworkers successfully exploited this aspect of a tyrosine decarboxylase (YDC) in a combined cascade with an LTA to produce aryl 1,2-amino alcohols.[68,69] The YDC prevented the reverse LTA reaction from occurring and applied additional stereocontrol over the reaction, resulting in good e.e.'s of the final amino alcohols. In addition, this cascade was used for the preparative-scale syntheses of a handful of 1,2-amino alcohols, showcasing the potential of decarboxylase biocatalysis when not limited to *in vivo* settings.

There have been few cases of protein engineering these classes of decarboxylases. Kalb 2016 reported active-site engineering of *Cs*TDC for increased substrate tolerance for tyrosine and DOPA over Trp.[60] Additionally, engineering of plant DC's has been found to alter substrate specificities between native aromatic amino acids.[49] Notably, these engineering efforts have been focused on expanding activity on natively present amino acids and have not been targeted towards decarboxylating diverse non-standard amino acids. A more focused study into the engineerability of this class of decarboxylases would better lay out the strengths and weaknesses for preparative scale biosynthesis.

1. 6. Preface to remaining chapters

Aromatic amino acid decarboxylases represent an underexplored class of biocatalyst. Broadly active decarboxylases are as-of-yet unknown and further study into these enzymes is warranted. The remaining chapters of this thesis describe our work investigating and engineering a bacterial tryptophan decarboxylase and using this decarboxylase as a model system to improve protein engineering methodology. Chapter 2 details our characterization of the substrate scope of a recently discovered bacterial tryptophan decarboxylase from *Ruminococcus gnavus* (*Rgn*TDC). Chapter 3 covers our systematic exploration of using substrate multiplexed screening (SUMS) to engineer *Rgn*TDC. In Chapter 4, we apply SUMS to *Rgn*TDC to gain activity on β-OH amino acids for cascade biocatalysis. Finally, Chapter 5 provides an overview on work investigating the potential of SUMS to facilitate substrate scope expansion during engineering and additional future directions. This thesis work thus highlights the biosynthetic potential of amino acid decarboxylases and their ability to advance protein engineering methodology.

Below are the main projects I have contributed to for this thesis work, which have resulted in the following publications:

1. **McDonald, A.D.**; Perkins, L.J.; Buller, A.R. “Facile *in vitro* biocatalytic production of diverse tryptamines.” *ChemBioChem.* 2019. 20, 1939–1944.
2. Thompson, C.M.; **McDonald, A.D.**; Yang, H.; Cavagnero, S.; Buller, A.R. “Modular control of L-tryptophan isotopic substitution via an efficient biosynthetic cascade.” *Org. Biomol. Chem*. 2020. 18, 4189–4192.
3. **McDonald, A.D.***; Higgins, P.M.*; Buller, A.R. “Substrate multiplexed protein engineering facilitates promiscuous biocatalytic synthesis.” (*Accepted at Nat. Comm.).*

1. 7. References

1. Wrenbeck, E. E., Azouz, L. R. & Whitehead, T. A. Single-mutation fitness landscapes for an enzyme on multiple substrates reveal specificity is globally encoded. *Nat. Commun.* **8**, 1–10 (2017).

2. Morrison, K. L. & Weiss, G. A. Combinatorial alanine-scanning. *Curr. Opin. Chem. Biol.* **5**, 302–307 (2001).

3. Liberles, D. A. *et al.* The interface of protein structure, protein biophysics, and molecular evolution. *Protein Sci.* **21**, 769–785 (2012).

4. De Godoy, L. M. F. *et al.* Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455**, 1251–1254 (2008).

5. Sheldon, R. A. & Woodley, J. M. Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* **118**, 801–838 (2018).

6. Hammer, S. C., Knight, A. M. & Arnold, F. H. Design and evolution of enzymes for non-natural chemistry. *Curr. Opin. Green Sustain. Chem.* **7**, 23–30 (2017).

7. Wu, S., Snajdrova, R., Moore, J. C., Baldenius, K. & Bornscheuer, U. T. Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angew. Chemie - Int. Ed.* **60**, 88–119 (2021).

8. Poppe, L. & Vértessy, B. G. The Fourth Wave of Biocatalysis Emerges- The 13 th International Symposium on Biocatalysis and Biotransformations. *ChemBioChem* **19**, 284–287 (2018).

9. Goodwin, N. C., Morrison, J. P., Fuerst, D. E. & Hadi, T. Biocatalysis in Medicinal Chemistry: Challenges to Access and Drivers for Adoption. *ACS Med. Chem. Lett.* **10**, 1363–1366 (2019).

10. Kulshrestha, S., Tyagi, P., Sindhi, V. & Yadavilli, K. S. Invertase and its applications – A brief review. *J. Pharm. Res.* **7**, 792–797 (2013).

11. Dadashipour, M. & Asano, Y. Hydroxynitrile lyases: Insights into biochemistry, discovery, and engineering. *ACS Catal.* **1**, 1121–1149 (2011).

12. Chandra, P., Enespa, Singh, R. & Arora, P. K. *Microbial lipases and their industrial applications: A comprehensive review*. *Microbial Cell Factories* vol. 19 (BioMed Central, 2020).

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

14. Fisher, B. F., Snodgrass, H. M., Jones, K. A., Andorfer, M. C. & Lewis, J. C. Site-Selective C-H Halogenation Using Flavin-Dependent Halogenases Identified via Family-Wide Activity Profiling. *ACS Cent. Sci.* **5**, 1844–1856 (2019).

15. 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. *J. Am. Chem. Soc.* **139**, 10769–10776 (2017).

16. Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **16**, 379–394 (2015).

17. Arnold, F. H. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chemie - Int. Ed.* **57**, 4143–4148 (2018).

18. Chen, K. & Arnold, F. H. Enzyme engineering for nonaqueous solvents: random mutagenesis to enhance activity of subtilisin E in polar organic media. *Bio/Technology* **9**, 1073–1077 (1991).

19. Watkins-Dulaney, E., Straathof, S. & Arnold, F. Tryptophan Synthase: Biocatalyst Extraordinaire. *ChemBioChem* vol. 22 5–16 (2021).

20. Nestl, B. M., Nebel, B. A. & Hauer, B. Recent progress in industrial biocatalysis. *Curr. Opin. Chem. Biol.* **15**, 187–193 (2011).

21. Risso, V. A. *et al.* Enhancing a de novo enzyme activity by computationally-focused ultra-low-throughput screening. *Chem. Sci.* **11**, 6134–6148 (2020).

22. Chapman, J., Ismail, A. E. & Dinu, C. Z. Industrial applications of enzymes: Recent advances, techniques, and outlooks. *Catalysts* **8**, 20–29 (2018).

23. Savile, C. K. *et al.* Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science (80-.).* **329**, 305–310 (2010).

24. McIntosh, J. A. *et al.* Engineered Ribosyl-1-Kinase Enables Concise Synthesis of Molnupiravir, an Antiviral for COVID-19. *ACS Cent. Sci.* **7**, 1980–1985 (2021).

25. Huffman, M. A. *et al.* Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science (80-.).* **366**, 1255–1259 (2019).

26. Isom, D. G., Castañeda, C. A., Cannon, B. R., Velu, P. D. & García-Moreno E, B. Charges in the hydrophobic interior of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 16096–16100 (2010).

27. Windle, C. L., Müller, M., Nelson, A. & Berry, A. Engineering aldolases as biocatalysts. *Curr. Opin. Chem. Biol.* **19**, 25–33 (2014).

28. Percudani, R. & Peracchi, A. A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep.* **4**, 850–854 (2003).

29. Liang, J., Han, Q., Tan, Y., Ding, H. & Li, J. Current advances on structure-function relationships of pyridoxal 5'-phosphate-dependent enzymes. *Front. Mol. Biosci.* **6**, (2019).

30. Rocha, J. F., Pina, A. F., Sousa, S. F. & Cerqueira, N. M. F. S. A. PLP-dependent enzymes as important biocatalysts for the pharmaceutical, chemical and food industries: A structural and mechanistic perspective. *Catal. Sci. Technol.* **9**, 4864–4876 (2019).

31. Ellis, J. M. *et al.* Biocatalytic synthesis of non-standard amino acids by a decarboxylative aldol reaction. *Nat. Catal.* **5**, 136–143 (2022).

32. Smith, J. L., Harrison, I. M., Bingman, C. A. & Buller, A. R. Investigation of β-Substitution Activity of O-Acetylserine Sulfhydrolase from Citrullus vulgaris. *ChemBioChem* **202200157**, 1–8 (2022).

33. Doyon, T. J. *et al.* Scalable and Selective β -Hydroxy- α -Amino Acid Synthesis Catalyzed by Promiscuous l -Threonine Transaldolase ObiH. *ChemBioChem* **22**, 1–10 (2021).

34. Hai, Y., Chen, M., Huang, A. & Tang, Y. Biosynthesis of Mycotoxin Fusaric Acid and Application of a PLP-Dependent Enzyme for Chemoenzymatic Synthesis of Substituted l -Pipecolic Acids. *J. Am. Chem. Soc.* **142**, 19668–19677 (2020).

35. Skye, G. E., Potts, R. & Floss, H. G. Stereochemistry of the Tryptophan Synthetase Reaction. *J. Am. Chem. Soc.* **96**, 1593–1595 (1974).

36. Corr, M. J., Smith, D. R. M. & Goss, R. J. M. One-pot access to l-5,6-dihalotryptophans and l-alknyltryptophans using tryptophan synthase. *Tetrahedron* **72**, 7306–7310 (2016).

37. Buller, A. R. *et al.* Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. *Proc. Natl. Acad. Sci.* **112**, 14599–14604 (2015).

38. Romney, D. K., Sarai, N. S. & Arnold, F. H. Nitroalkanes as Versatile Nucleophiles for Enzymatic Synthesis of Noncanonical Amino Acids. *ACS Catal.* **9**, acscatal.9b02089 (2019).

39. Fesko, K. Threonine aldolases: perspectives in engineering and screening the enzymes with enhanced substrate and stereo specificities. *Appl. Microbiol. Biotechnol.* **100**, 2579–2590 (2016).

40. Blesl, J. *et al.* Application of Threonine Aldolases for the Asymmetric Synthesis of α-Quaternary α-Amino Acids. *ChemCatChem* **10**, 3453–3458 (2018).

41. Martínez-Montero, L., Schrittwieser, J. H. & Kroutil, W. Regioselective Biocatalytic Transformations Employing Transaminases and Tyrosine Phenol Lyases. *Top. Catal.* **62**, 1208–1217 (2019).

42. Shih, D. F. *et al.* Aromatic L-Amino Acid Decarboxylase (AADC) Is Crucial for Brain Development and Motor Functions. *PLoS One* **8**, 1–16 (2013).

43. Williams, B. B. *et al.* Discovery and characterization of gut microbiota decarboxylases that can produce the neurotransmitter tryptamine. *Cell Host Microbe* **16**, 495–503 (2014).

44. Hadjiconstantinou, M. & Neff, N. H. Enhancing aromatic L-amino acid decarboxylase activity: Implications for L-DOPA treatment in Parkinson's disease. *CNS Neurosci. Ther.* **14**, 340–351 (2008).

45. Ehrenworth, A. M. & Peralta-Yahya, P. Accelerating the semisynthesis of alkaloid-based

drugs through metabolic engineering. *Nature Chemical Biology* vol. 13 249–258 (2017).

46. Woo, S., Jong, H. & Shin, S. Aromatic L - amino acid decarboxylases : mechanistic features and microbial applications. *Appl. Microbiol. Biotechnol.* 4445–4458 (2022) doi:10.1007/s00253-022-12028-4.

47. Houck, D. R. & Floss, H. G. PREPARATION OF STEREOSPECIFICALLY α- AND ß-TRITIATED TRYPTAMINE AND THE STEREOCHEMISTRY OF AROMATIC L-AMINO ACID DECARBOXYLASES. *J. Nat. Prod.* **44**, 759–762 (1981).

48. Runguphan, W., Qu, X. & O'Connor, S. E. Integrating carbon-halogen bond formation into medicinal plant metabolism. *Nature* **468**, 461–467 (2010).

49. Torrens-spence, M. P., Chiang, Y., Smith, T., Vicent, M. A. & Wang, Y. Structural basis for divergent and convergent evolution of catalytic machineries in plant aromatic amino acid decarboxylase proteins. *PNAS* **117**, 10806–10817 (2020).

50. Kourist, R., Guterl, J., Miyamoto, K. & Sieber, V. Enzymatic Decarboxylation — An Emerging Reaction for Chemicals Production from Renewable Resources. *ChemCatChem* **6**, 689–701 (2014).

51. Tatsis, E. C. *et al.* A three enzyme system to generate the Strychnos alkaloid scaffold from a central biosynthetic intermediate. *Nat. Commun.* **8**, (2017).

52. Runguphan, W., Qu, X. & O'Connor, S. E. Integrating carbon–halogen bond formation into medicinal plant metabolism. *Nature* **468**, 461–464 (2010).

53. Glenn, W. S., Nims, E. & O'Connor, S. E. Reengineering a tryptophan halogenase to preferentially chlorinate a direct alkaloid precursor. *J. Am. Chem. Soc.* **133**, 19346–19349 (2011).

54. Brown, S., Clastre, M., Courdavault, V. & O'Connor, S. E. De novo production of the plant-derived alkaloid strictosidine in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3205–3210 (2015).

55. Park, S. *et al.* Production of serotonin by dual expression of tryptophan decarboxylase and tryptamine 5-hydroxylase in Escherichia coli. *Appl. Microbiol. Biotechnol.* **89**, 1387–1394 (2011).

56. Mora-Villalobos, J. A. & Zeng, A. P. Synthetic pathways and processes for effective production of 5-hydroxytryptophan and serotonin from glucose in Escherichia coli. *J. Biol. Eng.* **12**, 1–12 (2018).

57. Lee, H. Y., Yerkes, N. & O'Connor, S. E. Aza-Tryptamine Substrates in Monoterpene Indole Alkaloid Biosynthesis. *Chem. Biol.* **16**, 1225–1229 (2009).

58. Runguphan, W. & O'Connor, S. E. Diversification of monoterpene indole alkaloid analogs through cross-coupling. *Org. Lett.* **15**, 2850–2853 (2013).

59. Misa, J., Billingsley, J. M., Niwa, K., Yu, R. K. & Tang, Y. Engineered Production of

Strictosidine and Analogues in Yeast. *ACS Synth. Biol.* **11**, 1639–1649 (2022).

60. Kalb, D., Gressler, J. & Hoffmeister, D. Active-Site Engineering Expands the Substrate Profile of the Basidiomycete L-Tryptophan Decarboxylase CsTDC. *ChemBioChem* **17**, 132–136 (2016).

61. Choi, Y., Han, S., Kim, J., Jang, Y. & Shin, J. Biochemical characterization and synthetic application of aromatic L -amino acid decarboxylase from Bacillus atrophaeus. *Appl. Microbiol. Biotechnol.* **105**, 2775–2785 (2021).

62. Han, S. W., Choi, Y. R. & Shin, J. S. Biocatalytic Decarboxylation of Aromatic l-Amino Acids with In Situ Removal of Both Products for Enhanced Production of Biogenic Amines. *Catal. Letters* **151**, 2996–3003 (2021).

63. Dragulska, S. & Kańska, M. Enzymatic synthesis of tryptamine and its halogen derivatives selectively labeled with hydrogen isotopes. *J. Radioanal. Nucl. Chem.* **299**, 759–763 (2014).

64. Torrens-Spence, M. P., Liu, C. T., Pluskal, T., Chung, Y. K. & Weng, J. K. Monoamine Biosynthesis via a Noncanonical Calcium-Activatable Aromatic Amino Acid Decarboxylase in Psilocybin Mushroom. *ACS Chem. Biol.* **13**, 3343–3353 (2018).

65. Fricke, J., Blei, F. & Hoffmeister, D. Enzymatic Synthesis of Psilocybin. *Angew. Chemie Int. Ed.* **56**, 12352–12355 (2017).

66. Fricke, J. *et al.* Enzymatic Route toward 6-Methylated Baeocystin and Psilocybin. *ChemBioChem* **20**, 2824–2829 (2019).

67. Blei, F., Baldeweg, F., Fricke, J. & Hoffmeister, D. Biocatalytic Production of Psilocybin and Derivatives in Tryptophan Synthase-Enhanced Reactions. *Chem. - A Eur. J.* **24**, 10028–10031 (2018).

68. Steinreiber, J. *et al.* Synthesis of Aromatic 1 ,2-Amino Alcohols Utilizing a Bienzymatic Dynamic Kinetic Asymmetric Transformation. *Adv. Synth. Catal.* **349**, 1379–1386 (2007).

69. Steinreiber, J. *et al.* Overcoming Thermodynamic and Kinetic Limitations of Aldolase-Catalyzed Reactions by Applying Multienzymatic Dynamic Kinetic Asymmetric Transformations. *Angew. Chem. Int. Ed.* 1624–1626 (2007) doi:10.1002/anie.200604142.

Chapter 2

Characterization of a promiscuous tryptophan decarboxylase

Content in this chapter is adapted from the following published work:

McDonald, A.D.; Perkins, L.J.; Buller, A.R. “Facile *in vitro* biocatalytic production of diverse tryptamines.” *ChemBioChem.* 2019. 20, 1939–1944.

Several of the tryptamine preparative syntheses were performed by Lydia J. Perkins, who was additionally a valuable intellectual contributor to this work.

Chapter 2: Characterization of a promiscuous tryptophan decarboxylase

2. 1. Introduction

Tryptamines are a prominent class of natural products that include potent bioactive compounds such as the neurotransmitter serotonin and psilocybin, the precursor to the active component of the hallucinogenic *Psilocybe* mushrooms (Fig 1a).[1,2] A common motif among alkaloid natural products is modification of the indole ring, which can have a profound impact on the molecule's pharmacological properties. Similarly, the triptan class of anti-migraine drugs, such as rizatriptan, are tryptamine analogs and are made through organic synthesis.[3] In Nature, formation of tryptamine also serves as the starting point in the biosynthesis of more sophisticated natural products, such as vincristine and camptothecin.[3] Given the importance of tryptamines, this class of compounds has been a frequent target for practical demonstrations of synthetic methodology, many of which begin with a substituted indole as the starting material. Classic synthetic approaches include reductive alkylation[4,5] or, more recently, C-H activation methodology (Fig 1c).[6,7] While these traditional routes have proven quite successful, it is also desirable to develop enzymatic approaches, as they often better conform to the principles of green chemistry and have the potential to be coupled with cellular metabolism.[8,9] To date, biocatalytic methodology to produce substituted tryptamines has lagged far behind traditional organic synthesis.

Nature's strategy to make modified tryptamines often starts with L-tryptophan (Trp) and a tailoring enzyme that performs a specific modification of the indole ring. The Trp derivatives are subsequently decarboxylated and the resultant tryptamines metabolized further into fully elaborated natural products.[1] This approach can be exploited synthetically to produce a few tryptamine analogs, but the generation of further chemical diversity quickly becomes cumbersome, as a separate enzyme is required for each modification (Fig 1c) [10] For example, the genetic incorporation of Trp halogenases with distinct regiospecificities into the medicinal plant *Catharanthus roseus* afforded access to the halogenated alkaloids.[11] However these efforts were

hindered by the relatively poor efficiency of the *Catharanthus roseus* TDC (*Cro*TDC) with non-natural substrates. Indeed, most enzymes that decarboxylate Trp are polyspecific for other, smaller aromatic amino acids and are sluggish with non-standard amino acids that are larger than Trp.[12] Moreover, there are many desirable functional groups, such as fluoroindoles, that cannot be made through any currently known enzymatic route.

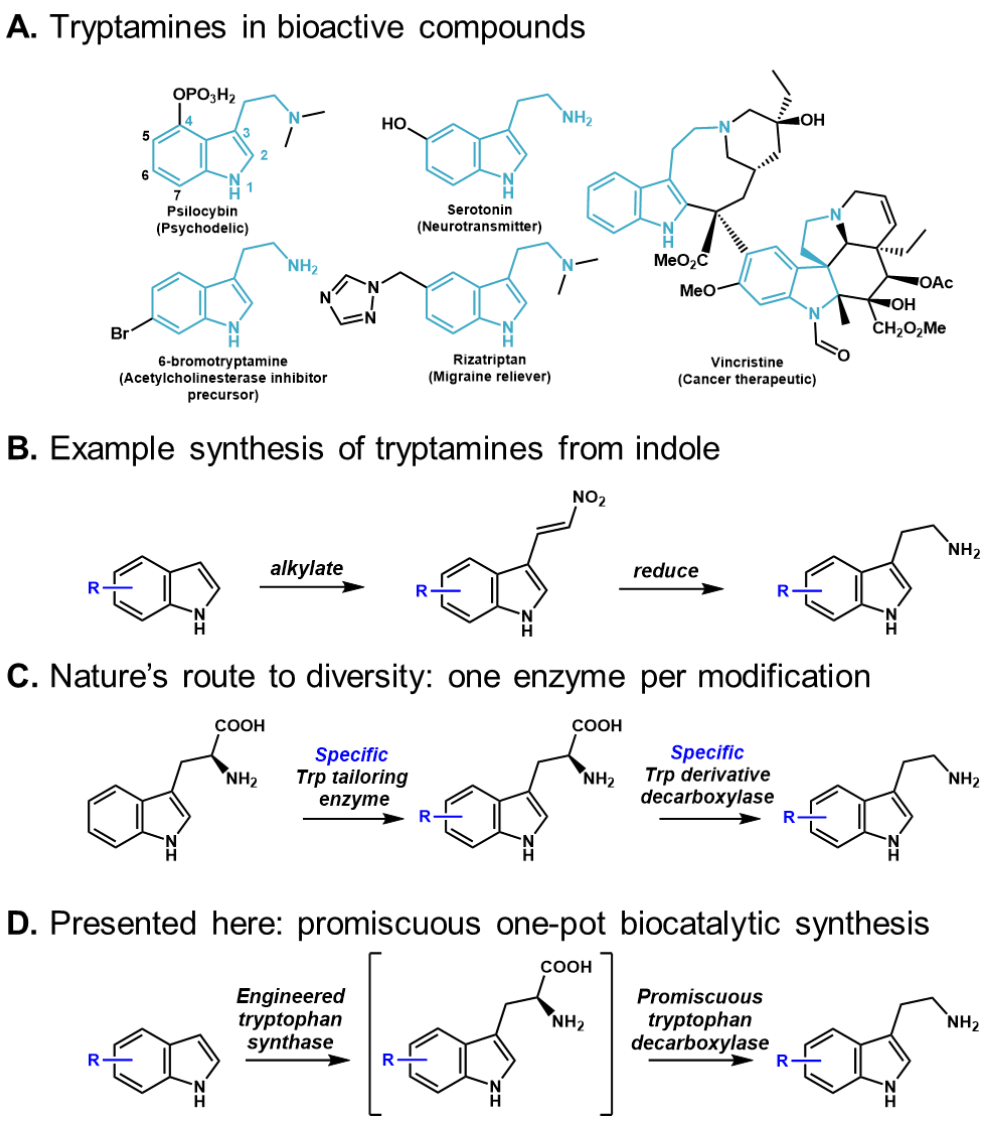


Figure 1. (Bio)synthetic routes to pharmacologically relevant tryptamines.

An alternative biocatalytic approach to produce modified tryptamines is to tap into the outstanding chemical diversity that is already commercially available for indoles. That is, rather

than relying on highly specific tailoring enzymes to introduce chemical diversity, pre-modified indole building blocks can be converted into their corresponding Trp analog through the promiscuous action of a Trp synthase enzyme.[13–15] Natively, Trp synthase uses the pyridoxal 5'-phosphate (PLP) cofactor to condense L-serine (Ser) and indole into Trp. Many substituted indoles and even other amino acids can participate in this versatile reaction, affording access to complex starting materials for downstream modification.[16–19] Recently, these efforts have been made even easier by the evolution of the β-subunit of tryptophan synthase (TrpB) (EC 4.2.1.20) for efficient stand-alone function.[20–24] These engineered TrpBs can operate at high temperatures, which is helpful in solubilizing hydrophobic indoles, and provide access to highly complex Trp analogs in good yields. Theoretically, these diverse tryptophans could be converted to the analogous tryptamines through the action of a promiscuous tryptophan decarboxylase (Fig 1d). Recently, the decarboxylase PsiD from the psilocybin biosynthetic pathway was shown to act on both its native substrate, 4-hydroxytryptophan, as well as two additional regioisomers that were prepared using a TrpB enzyme.[25] These enzymes were combined in a short and efficient cascade, but the scalability and generality of this route for producing tryptamines with alternative functionalities is not yet known.

In this chapter, I cover the comprehensive characterization of the Trp decarboxylase from the gut microbe *Ruminococcus gnavus* (*Rgn*TDC) (EC 4.1.1.28) as the first broadly promiscuous member of its enzyme family. This enzyme was previously shown to be specific for standard Trp over other aromatic amino acids and an X-ray crystal structure with a substrate analog was determined, revealing the indole binding pocket.[26] Structural analysis suggested that this region of the active site could accommodate substitution on the indole ring and therefore function on a range of substrates, hinting at biocatalytic potential.

2. 2. Results and Discussion

2. 2. 1. Kinetic characterization of *Rgn*TDC

To begin investigation of *Rgn*TDC's biocatalytic potential, we overexpressed *Rgn*TDC in *E. coli* BL21 (DE3) cells and observed that cell cultures grew to unusually low cell density in rich media, ~ 4 g per L culture. We suspected *Rgn*TDC was exerting a toxic effect by depleting cells of Trp, and so exogenous indole (0.5 mM) was added at the time of induction. This procedure significantly increased cell densities to ~11 g per L culture, presumably by supplementing the native *E. coli* Trp synthase enzyme with indole. As isolated, *Rgn*TDC copurified with covalently bound PLP and was reliably procured in large quantities with yields of > 100 mg purified enzyme per L culture (Fig 2a).

We measured the progress of the native *Rgn*TDC reaction *in vitro*, which showed an impressive 20,000 turnovers in 4 h (Fig 2B). The total turnover number (TTN) with Trp was improved 4-fold with 10 additional equivalents of the PLP cofactor, up to ~80,000 turnovers in 4 h (Figure 2b). One hypothesis to explain this loss of cofactor is due to a competing oxidative deamination reaction, which has been studied for other PLP-dependent decarboxylases.[27] This degradation process can be mitigated with the addition of extra PLP, which swaps out the damaged cofactor and enables catalysis to continue.

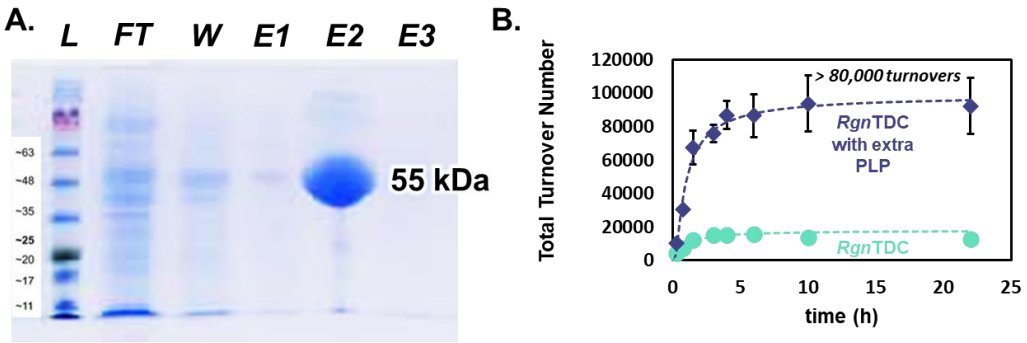


Figure 2. Initial expression and activity testing of *Rgn*TDC. A. Expression of *Rgn*TDC: L = ladder; FT = flow-through; W = wash; E = elutions. **B.** Progress curve of *Rgn*TDC-catalyzed Trp decaboxylation.

In our hands, *Rgn*TDC operated with a catalytic efficiency (k_{cat}/K_M) of 6,900 $M^{-1}sec^{-1}$ with native Trp, ~7-fold lower than reported previously. However, the K_M of 1.6 mM was in good agreement with the original report (Fig 3b).[26] Serendipitously, this catalytic efficiency is also similar to that observed for *Cro*TDC, enabling direct comparison. *Cro*TDC has a much lower K_M with Trp (52 µM) (Table 1). In contrast, the catalytic efficiency of *Rgn*TDC is driven by k_{cat}, which is ~130 times higher than that of *Cro*TDC, accounting for the high TTN of the *R. gnavus* enzyme.

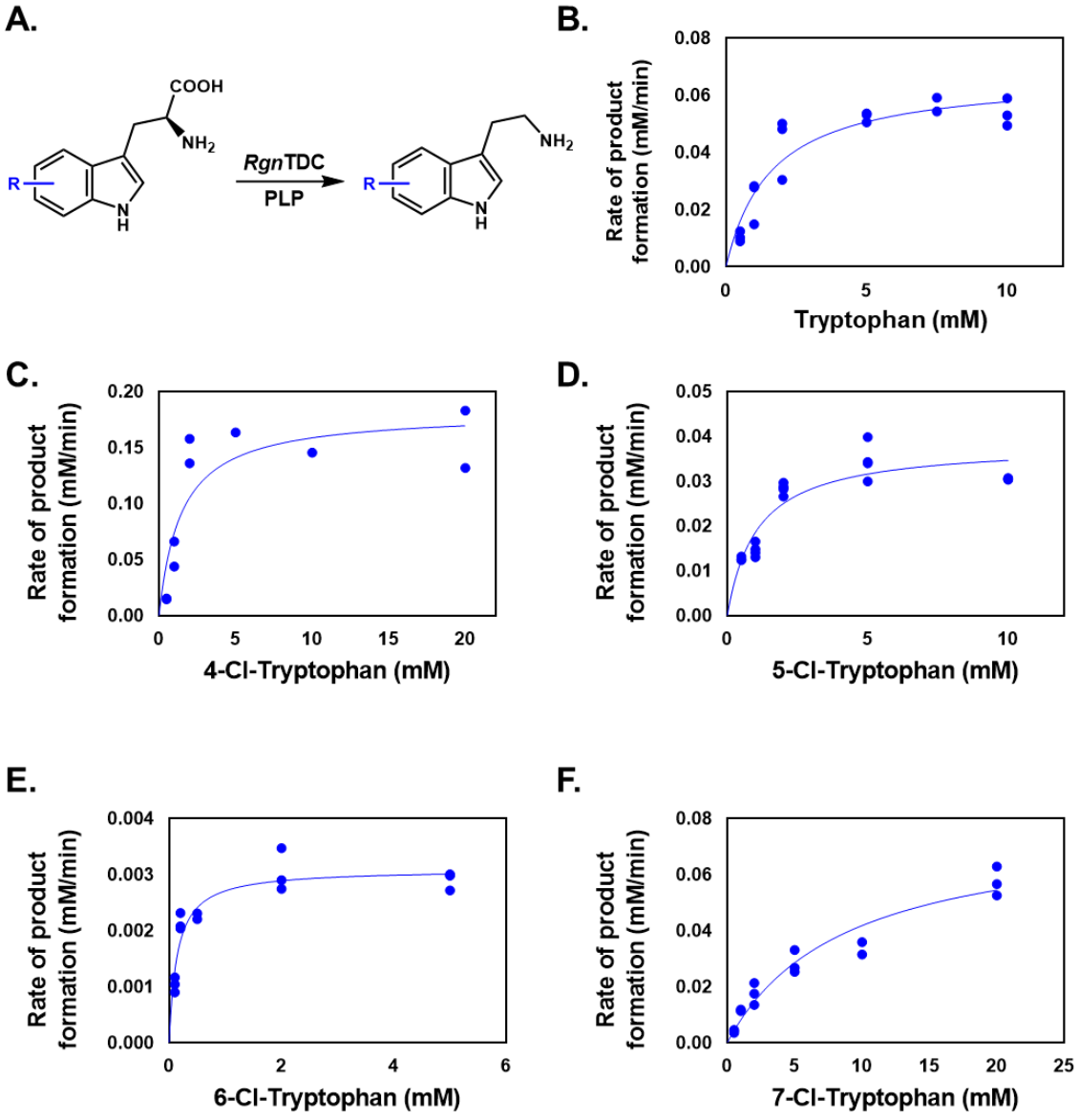


Figure 3. Michaelis-Menten curves of tryptamine analog decarboxylation by *Rgn*TDC.

Previously characterized TDC's have had poor activity on substrates that bear substitutions on the indole ring and are therefore only of practical use with amino acids that are the size of Trp or smaller.[28,29] The kinetic parameters of *Cro*TDC with two chlorotryptophans have been measured, which revealed that the relatively poor efficiency of this reaction is driven by large changes to K_M, which limits *in vivo* production of chlorinated indole alkaloids in *C. roseus* (Table 1).[11,30]

Table 1. Kinetic parameters of *Cro*TDC (from ref. 11) and *Rgn*TDC (this work).

TDC	Substrate	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} sec^{-1})	Fold-decrease in activity relative to Trp
*Cro*TDC	**L-Trp**	52	0.085	1560	-
	5-Cl-Trp	499	0.027	55	28
	7-Cl-Trp	538	0.042	77	20
RgnTDC	**L-Trp**	1.6 ± 0.4	11.2 ± 0.9	6900	-
	4-Cl-Trp	1.5 ± 0.7	2.3 ± 0.3	1600	**4.3**
	5-Cl-Trp	1.2 ± 0.3	0.68 ± 0.05	570	**14**
	6-Cl-Trp	0.14 ± 0.03	0.52 ± 0.02	3600	**1.9**
	7-Cl-Trp	9.0 ± 2.0	10 ± 1	1200	**5.8**

To test the promiscuity of the *Rgn*TDC, we sought to measure its kinetic parameters with a complete set of 4-, 5-, 6-, and 7-chlorotryptophans. The requisite enantiopure chlorotryptophans, however, were not readily available. We therefore prepared each Trp analog from their corresponding indole precursor through a simple biocatalytic reaction using engineered TrpB enzymes from *Pyrococcus furiosus* (*Pf*TrpB's) with high stand-alone activity.[20–24] Surprisingly, the kinetic parameters of *Rgn*TDC with each chlorotryptophan showed only modest

decreases in catalytic efficiency when compared to the native Trp substrate (Table 1, Fig 3c-f). The K_M for the chlorotryptophans ranged from 5-fold higher to 11-fold lower, whereas the k_{cat} values were either similar or lower compared to Trp. The overall effect is a smaller decrease in catalytic efficiency with substituted tryptophans when compared to *Cro*TDC, indicating that *Rgn*TDC possesses relatively high native promiscuous activity. More importantly, for synthetic purposes, we reason that high K_M values of *Rgn*TDC are less indicative of preparative-scale activity than k_{cat}, since the enzyme can be kinetically saturated *in vitro* simply by adding more substrate. Indeed, we measured TTNs for each chlorotryptophan and observed an excellent correlation between the k_{cat} and the TTN for each substrate (Fig 4a). Thse TTNs did not correlate well with k_{cat}/K_M's, however (Fig 4b). In this light, *Rgn*TDC is a particularly attractive catalyst as it maintains a high k_{cat} with each of the chlorotryptophan substrates and performs thousands of turnovers with four substrates that make distinct steric demands of the active site.

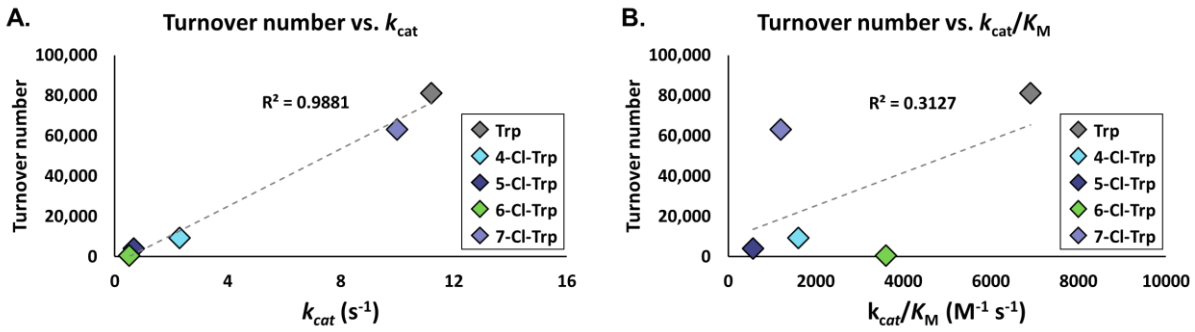


Figure 4. Comparison of *Rgn*TDC kinetic parameters vs. total turnover number (TTN).

2. 2. 2. Characterization of *Rgn*TDC substrate scope

To investigate how well *Rgn*TDC performs with diverse indole ring substitutions, we prepared a panel of substituted tryptophans using the *Pf*TrpB enzymes, relying heavily on the enzyme, *Pf*TrpB^{2B9} (see Methods for details).[21] This engineered TrpB variant is highly promiscuous and thus acts as a 'generalist' catalyst for production of tryptophan analogs.[21] Engineering and detailed study of this engineered catalyst has been reported and the overall engineering efforts for TrpB enzymes were recently reviewed. [21,31,32] The collection of tryptophans encompasses a range of sterically and electronically diverse substituents for each of the 4−7 positions. Many of these Trp analogs have poor aqueous solubility at high (0.1 M) concentrations, so reactions were carried out with 10 mM substrate in aqueous phosphate buffer with 5% methanol as a cosolvent. Generally, *Rgn*TDC had the highest activity with the 7-substituted tryptophans, consistent with the enzyme's high k_{cat} with 7-chlorotryptophan (Table 1, Fig 5a). The 5- and 6-substituted tryptophans were less well accommodated by the enzyme. We analyzed the TTN with each substrate as a function of the group van der Waals radius of the substituent, which revealed the relative steric sensitivity of *Rgn*TDC to substitution at each position (Fig 6). *Rgn*TDC was most sensitive to steric perturbation at the 4-position. For the largest 4-substituent we tested, 4-cyanotryptophan, only trace product formation (< 10 turnovers) was observed. Satisfyingly, when we measured activity on a series of 4-halotryptophans we noticed an order of magnitude increase in activity as the size of the halide substituent decreases, up to ~96,000 TTN with 4-fluorotryptophan (Fig 6a).

In contrast to substitution elsewhere on the ring, a methyl substitution at the 2-position was poorly tolerated and *Rgn*TDC performed just 275 turnovers in 4 h (Fig 5b). When a methyl group is added to the β-position of Trp, no reactivity was observed. However, we assayed *Rgn*TDC with a di-substituted Trp analog and were pleased to observe formation of product

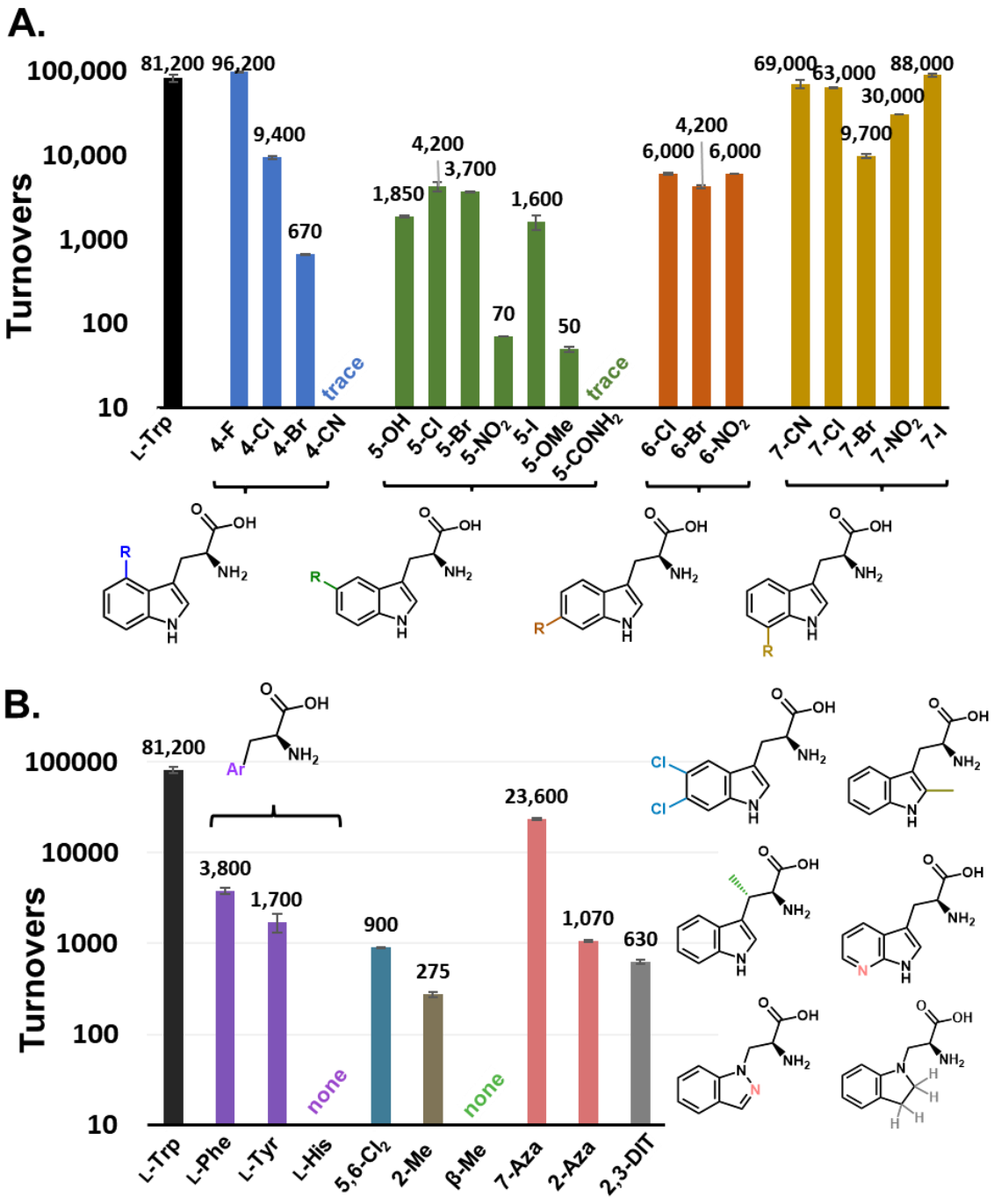


Figure 5. Total turnover number analysis of *Rgn*TDC with various Trp analogs.

corresponding to 900 turnovers with 5,6-dichlorotryptophan as a substrate, which is lower than the activity on either of the mono-substituted analogs.

Native TDC enzymes are often regarded as possessing low activity on amino acids larger than Trp, and the broad promiscuity of *Rgn*TDC was therefore quite surprising. We sought to probe the substrate scope further, testing for activity with substrates that are more structurally distinct from Trp. In line with previous studies, *Rgn*TDC shows activity on L-tyrosine,[26] but with 50-fold lower activity relative to Trp. However, there is no observable decarboxylation of L- histidine. In addition, no activity was observed with D-tryptophan. Amino acids bearing indole isosteres were also prepared with *Pf*TrpB^{2B9} (Fig 5b). Indazole and indoline react with TrpB at N1 to yield 2-azaisotryptophan and 2,3-dihydro-isotryptophan (DIT), respectively, and with which *Rgn*TDC has modest activity. In contrast, 7-aza-tryptophan is a highly reactive substrate, consistent with its steric profile being nearly identical to the native Trp substrate.

We sought to rationalize the structural basis for the reactivity profile of *Rgn*TDC using information from the previously-determined structure with a suicide-inhibitor bound in the active site (PDB ID: 4OBV).[26] Using these data, we modelled the covalent Trp-PLP external aldimine adduct that forms prior to decarboxylation and whose geometry with the carboxylate oriented perpendicular to the PLP ring can be readily inferred through Dunathan's stereoelectronic hypothesis.[33] Addition of a chloro group to the 4-position of Trp clearly leads to a steric clash with the carboxylate (Fig 6a). Given the steep decrease in activity with progressively larger substituents, the complex appears to distort away from an optimally reactive orientation to relieve this steric clash. Bulky side chain residues in the *Rgn*TDC active site similarly appear to limit the size of the groups available at the 5- and 6-positions and cause some structural rearrangement of the active site to facilitate catalysis (Fig 6b-6c). There is a noticeable sparsity

of residues flanking the 7-position of the indole ring, which is consistent with *Rgn*TDC's prodigious ability to produce 7-substituted tryptamines (Fig 6d).

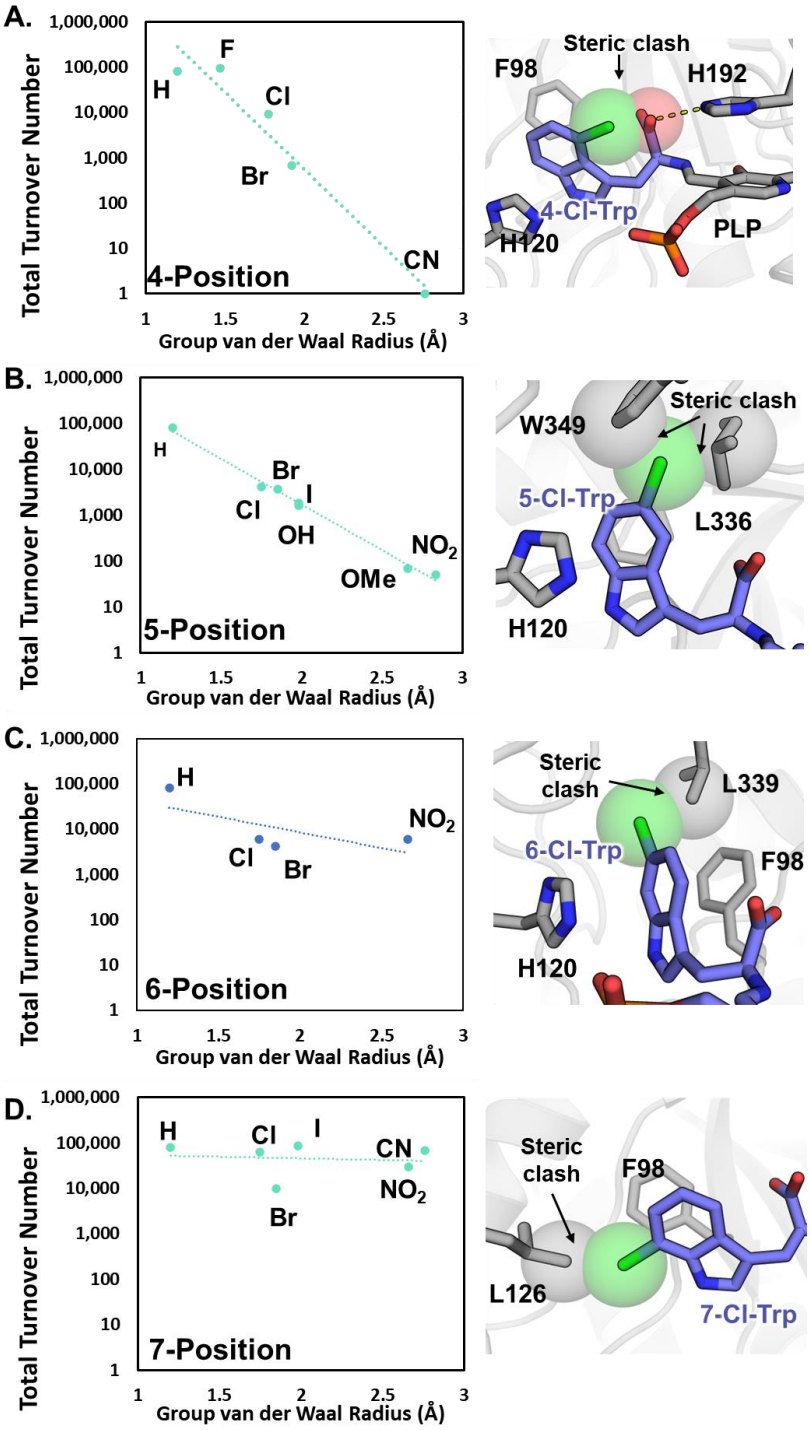


Figure 6. Comparison of van der Waal radii of indole substitutions and observed total turnover numbers for *Rgn*TDC.

2. 2. 3. Investigation of biocatalytic utility of *Rgn*TDC for tryptamine synthesis

This study of *Rgn*TDC required synthesis of substituted Trp substrates that were themselves produced in a biocatalytic reaction. Naturally, we considered whether *Rgn*TDC could be combined with a Trp synthase, *Pf*TrpB^{2B9}, into an efficient two-enzyme sequence to produce tryptamines from the corresponding indole and serine. Although both enzymes have high activity on a range of substrates, they operate at different optimal temperatures (37 vs 75 ^{o}C), which presents an inherent challenge. We first tested the enzymes in a cascade reaction at 37 ^{o}C, which resulted in full conversion of indole to tryptamine (5000 turnovers, Fig 7), demonstrating that the two enzymes can operate efficiently under the same conditions.

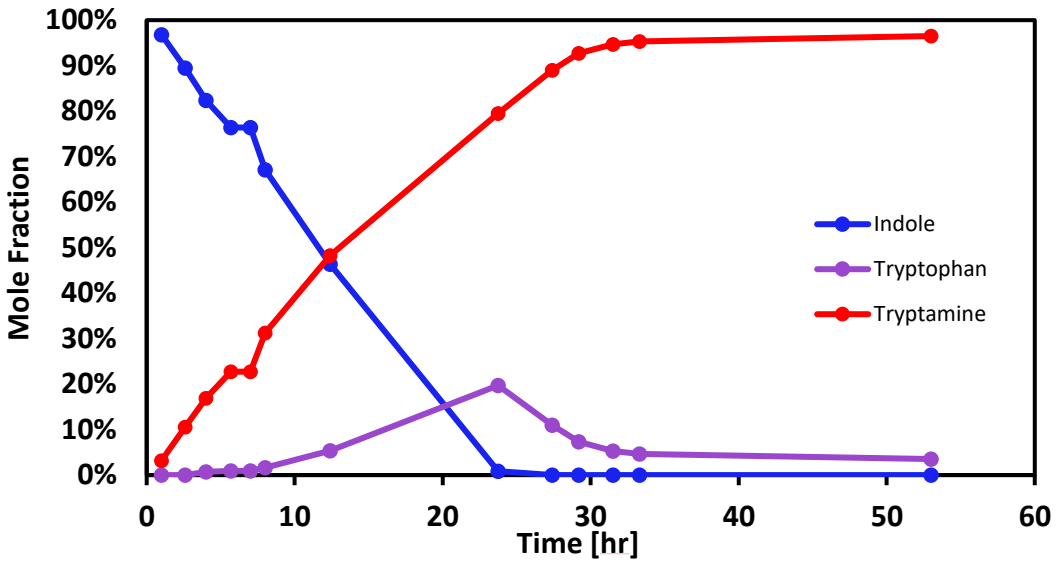


Figure 7. Progress curve of *Pf*TrpB^{2B9} and *Rgn*TDC cascade reaction with indole.

Although TrpB enzymes have recently been engineered for higher activity at lower temperature,[24] the solubility of substituted indoles drops precipitously at these temperatures, hindering preparative scale reactions. Therefore, we adopted a sequential one-pot, two-enzyme reaction setup. First, *Pf*TrpB^{2B9} acts at 75 °C on 0.01 M substituted indole and three equivalents of Ser for 12 h. The reaction vessel was then cooled to 37 °C and *Rgn*TDC was added along

with 10 additional equivalents of PLP. Decarboxylation was allowed to continue for several hours (times depend on the substrate, see Materials and Methods for details) prior to purification by a simple acid/base extraction procedure. We used this two-enzyme sequence to produce six different tryptamines on 1.0 mmol scale with a range of yields (Figure 8). Isolated yield tracked well with the efficiency of the analytical scale reactions. While the one-pot reaction procedure was still sufficient for substrates poorly tolerated by *Rgn*TDC, these reactions suggest that further improvements could be obtained through protein engineering.

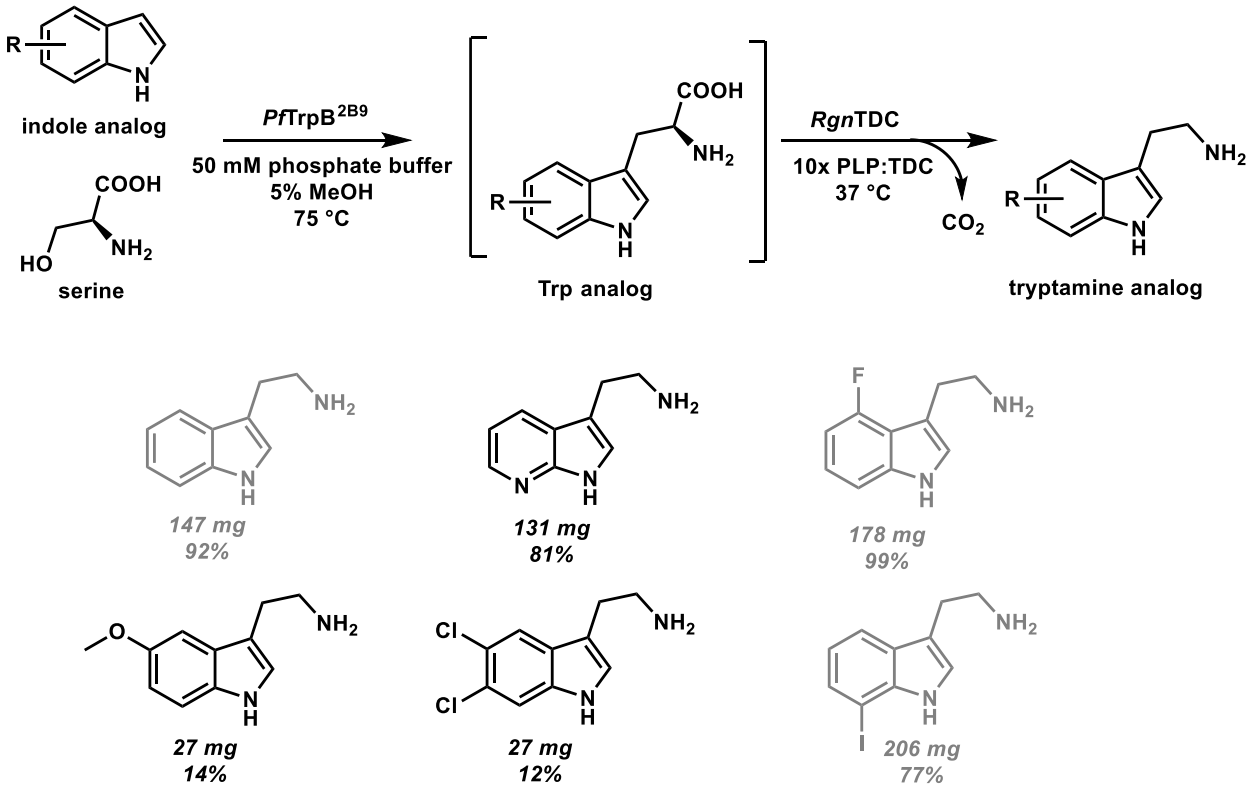


Figure 8. Preparative scale one-pot tryptamine syntheses using TrpB^{2B9} and *Rgn*TDC.
*Tryptamines in gray were synthesized by Lydia Perkins.

2. 3. Conclusion

We have shown that the TDC previously isolated from *R. gnavus* is the first broadly promiscuous enzyme for the *in vitro* synthesis of tryptamines. Through kinetic experiments, we determined that the catalytic efficiency of *Rgn*TDC is driven by a high k_{cat}, which in turn governs reactivity of the catalyst *in vitro*. We observed this enzyme performs up to 96,000 turnovers and

this high activity was leveraged to prepare > 20 different tryptamine analogs. Hence, native *Rgn*TDC breaks the mold of the phylogenetically-distinct, but more well-known aromatic amino acid decarboxylases. This promiscuity suggests the potential of *Rgn*TDC as a useful enzyme for *in vitro* biocatalysis. Both the *Rgn*TDC and the engineered TrpB enzymes can be procured in high quantities (> 100 mg enzyme per L culture) and combine to provide easy access to diverse tryptamines from commercially abundant starting materials. In contrast to previous methodology to produce tryptamines, the procedures detailed here do not require significant training in organic synthesis or specialist equipment that is not already found in most biochemistry labs. Hence, these enzymatic tools will enable other biochemists to synthesize sophisticated analogs of metabolic intermediates and spur new synthetic biology efforts to engineer biosynthesis.

2. 4. Materials and Methods

General Experimental Methods

Chemicals and reagents were purchased from commercial suppliers (Sigma-Aldrich, VWR, Chem-Impex International, Alfa Aesar) and used without further purification unless otherwise noted. *E. coli* cells were electroporated with an Eppendorf E-porator at 2500 V. New Brunswick I26R, 120 V/60 Hz shakers (Eppendorf) were used for cell growth. Cell disruption via sonication performed with a Sonic Dismembrator 550 (Fisher Scientific). Optical density measurements were collected on a UV-2600 Shimadzu spectrophotometer (Shimadzu). UPLC/LC-MS data were collected on an Acquity UHPLC with an Acquity QDA MS detector (Waters) using an ACQUITY UPLC CSH Phenyl-Hexyl column (Waters) or an Intrada Amino Acid column (Imtakt). Preparative flash chromatographic separations were performed on an Isolera One Flash Purification system (Biotage). NMR data were collected on Bruker 400 or 500 MHz spectrometers. High resolution mass data were collected with a Q Extractive Plus Orbitrap (NIH 1S10OD020022-1) instrument with the samples ionized by ESI. Fits of kinetic data were performed with PRISM 8 Graphpad software.

Cloning, expression, and purification of RgnTDC

A codon-optimized copy of the *Ruminococcus gnavus* tryptophan decarboxylase (TDC) gene was purchased as a gBlock from Integrated DNA Technologies. This DNA fragment was cloned into a pET22b vector by the Gibson Assembly method.[34] BL21 *E. coli* cells were subsequently transformed with the resulting cyclized DNA product *via* electroporation. After 30 min of recovery in Luria-Burtani (LB) media at 37 °C, cells were plated onto LB plates with 100 µg/mL ampicillin (AMP) and incubated overnight. Single colonies were used to inoculate 2 x 5 mL TB + 100 µg/mL AMP, which were grown overnight at 37 °C. Expression cultures, typically 1 L of Terrific Broth (TB) + 100 µg/mL ampicillin were inoculated from these starter cultures and shaken (180 rpm) at 37 °C. After 3 hours (OD600 = 1.5), the expression cultures were chilled on ice. After 90 min on ice, expression was induced with 1 mM IPTG, and the cultures were supplemented with 0.5 mM indole. Cultures were expressed overnight at 20 °C with shaking at 180 rpm. Cells were then harvested by centrifugation at 4300 x*g* at 4 °C for 30 min. Cell pellets were used fresh or frozen at -20 °C until lysis.

To purify *Rgn*TDC, cell pellets were thawed on ice and then resuspended in lysis buffer (50 mM potassium phosphate buffer (pH = 8.0), 1 mg/ml Hen Egg White Lysozyme (GoldBio), 0.2 mg/ml DNaseI (GoldBio), 1 mM MgCl2, and 150 µM pyridoxal 5′-phosphate (PLP)). A volume of 4 mL of lysis buffer per gram of wet cell pellet was used. After 30 min of shaking at 37 °C, the lysis suspension was disrupted using sonication (20 min; 0.8 s on, 0.2 s off at a power setting of 5). The resulting lysate was then spun down at 75,600 x*g* to pellet cell debris. Ni/NTA GoldBio beads were added to the supernatant and incubated on ice for 40 min prior to purification by Ni-affinity chromatography. The column was washed with 3 column volumes of 20 mM imidazole, 100 mM NaCl, 50 mM potassium phosphate buffer (pH = 8.0), and eluted with 250 mM imidazole, 100 mM NaCl, 50 mM potassium phosphate buffer (pH = 8.0). Elution of the desired protein product was monitored by the disappearance of its bright yellow color (resulting

from the PLP cofactor) from the column. The protein product was dialyzed to < 1 μM imidazole, dripped into liquid nitrogen to flash freeze, and stored at -80 °C for no more than 1 month before use, as activity was found to decrease following long periods at -80 °C. The concentration of protein was determined by Bradford assay. Generally, this procedure yielded 100 – 200 mg per L culture.

Reaction condition optimization

PLP concentration for the decarboxylation of tryptophan

Various concentrations of pyridoxal-5'-phosphate (PLP) cofactor were screened to maximize *Rgn*TDC activity. The tested reaction conditions were 10 mM L-Trp, 0 – 0.001 mM PLP, 0 – 0.0001 mM *Rgn*TDC (100,000 max TTN), in 100 μl of 50 mM potassium phosphate buffer (pH = 8.0). We observed that decarboxylation activity continues to improve with higher excess of PLP (Fig 9). Due to the high activity observed at 10x molar equivalents excess PLP to *Rgn*TDC, this concentration was used for future reactions.

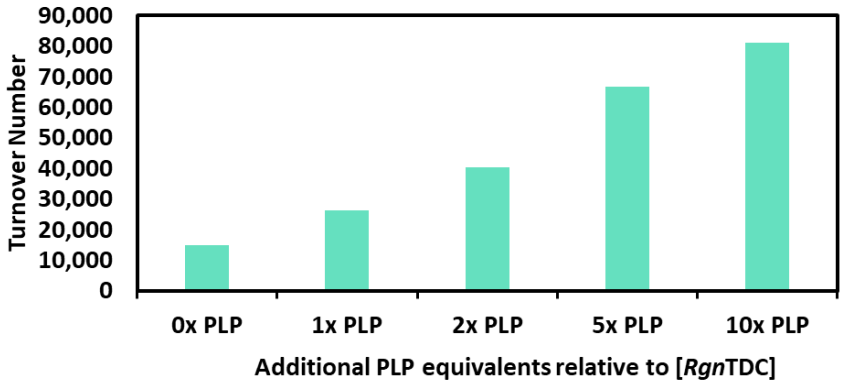


Figure 9. Effect of additional PLP on *Rgn*TDC-catalyzed Trp decarboxylation. '0x PLP' refers to the as-isolated protein.

Decarboxylation cosolvent screening

On an analytical (100 µL) scale, cosolvents were screened to aid in solubilizing the more hydrophobic tryptophan analogs. The reaction conditions were 10 mM L-Trp, 0.7 µM pyridoxal-5'-phosphate (PLP), 0.07 µM *Rgn*TDC (0.0007% mol catalyst, 150,000 max TTN), 0 – 25 µL methanol or acetonitrile in 1.7 mL Eppendorf tubes, diluted to a total reaction volume of 100 µL with potassium phosphate buffer (pH = 8.0). It was found that 5% methanol was the optimal cosolvent concentration for tryptophan decarboxylation (Fig 10a).

Tryptophan synthase cosolvent screening

On an analytical (100 µL) scale, cosolvents were screened to aid in solubilizing hydrophobic indoles. The reaction conditions were 10 mM 5-bromoindole, 30 mM serine, 2 µM *Pf*TrpB^{2B9} (0.02% mol catalyst, 5000 Max TTN), 0–20 µL methanol or acetonitrile in 1.7 mL Eppendorf tubes, diluted to a total reaction volume of 100 µL with potassium phosphate buffer (pH = 8). It was found that 5% methanol was the optimal cosolvent concentration (Fig 10b).

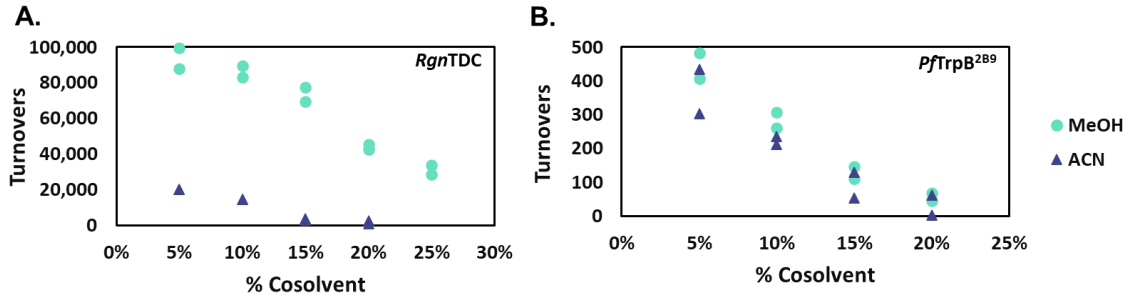


Figure 10. Effect of cosolvent on *Rgn*TDC and *Pf*TrpB^{2B9}. Methanol is the preferred solvent for subsequent biocatalytic transformations.

Biocatalytic synthesis of tryptophan analogs

*Pf*TrpB^{2B9} (and other TrpB enzymes used for tryptophan analog preparations) were purified in identical fashion to TDC, but with sonication replaced by a heat treatment step after lysis at 75 °C for 15 min. These enzymes were also flash frozen and stored at -80 °C after

purification. For preparative reactions, solid indole analog (0.2–1.0 mmol) and serine (0.6–3.0 mmol) were added to a 15 mL glass pressure flask. Solids were dissolved in 0.5 mL of methanol and then diluted with 10 mL of 50 mM potassium phosphate buffer (pH = 8.0). The appropriate volume of *Pf*TrpB^{2B9} solution was thawed on ice, centrifuged at 15,000 x*g* for 5 min to pellet aggregated protein, and then added to bring the total enzyme concentration to 2 μM (0.02% mol catalyst, 5000 Max TTN). The reaction flask was sealed and then heated to 75 °C for 16 hours. The progress of the reaction was monitored by UPLC-MS.

After the reaction had reached full conversion, reaction solutions were evaporated to ~ 1 mL and the tryptophans subsequently purified with reverse phase C18 column chromatography. Tryptophan-containing fractions were pooled, evaporated to remove solvent, resuspended in water, and lyophilized overnight (or until dry). The resultant solid was weighed to determine the yield. Typically, tryptophans were isolated in poor to good yields (20 – 70%).

Michaelis-Menten kinetic experiments

*Rgn*TDC was thawed on ice and then centrifuged at 15,000 x*g* for 5 min to pellet aggregated protein. The supernatant was then diluted 1:100 in 50 mM potassium phosphate buffer (pH = 8.0). Reactions conditions were 0.5–20 mM tryptophan analog with 10 molar equivalents of pyridoxal-5'-phosphate (PLP) cofactor in 1.7 mL Eppendorf tubes. Reactions were then diluted to a total reaction volume of 100 μL with potassium phosphate buffer (pH = 8.0). The enzyme and substrate solutions were equilibrated to 37 °C prior to mixing. Final enzyme concentrations were 0.1–1.25 μM. All reactions were run at least in duplicate at 37 °C. At various time points after addition of *Rgn*TDC, 20 μL of the reaction solution were quenched in 180 μL quenching solution (equal volumes of acetonitrile and 1 M HCl). Quenched reactions were centrifuged at 15,000 x*g* for 10 min prior to LC-MS injection of the supernatant. Enzymatic activity was quantified by integrating the substrate and product UV absorbance peaks at 280 nm.

Analytical scale decarboxylation reactions by *Rgn*TDC

*Rgn*TDC was thawed on ice and then centrifuged at 15,000 x g for 5 min to pellet aggregated protein and the supernatant extracted. The *Rgn*TDC solution was then diluted to workable concentrations using 50 mM potassium phosphate buffer (pH = 8.0). Reactions were carried out with 10 mM substrate concentration, at 10 molar equivalents of PLP to enzyme, in 50 mM potassium phosphate buffer (pH = 8.0). Reactions were carried out in 1.7 mL Eppendorf tubes at 37 °C for 4 hours. After 4 hours, reactions were quenched with 100 µL of quenching solution (equal volume of acetonitrile and 1 M HCl). Quenched reactions were centrifuged at 15,000 x g for 10 min prior to UPLC-MS injection. Reaction conversion was observed by substrate to product ratio via UV-VIS absorbance peak areas at 280 nm.

Preparative-scale decarboxylation of tryptophan analogs by *Rgn*TDC:

*Rgn*TDC was thawed on ice and then centrifuged at 15,000 x*g* for 5 min to pellet aggregated protein. Solid tryptophan analog was weighed out (0.1–0.3 mmol) into a 20 mL scintillation vial and diluted to 10 mL with 50 mM potassium phosphate buffer (pH = 8.0). The resulting mixture was then heated at 50 °C to solubilize tryptophan analogs. Once no more solid tryptophan was visible, the solution was cooled to room temperature. Ten molar equivalents of pyridoxal-5'-phosphate (PLP) relative to *Rgn*TDC was then added followed by addition of *Rgn*TDC (0.02–1% mol catalyst, 5000 to 100 max TTN). Reactions took place at 37 °C for 16 h. Reaction progress was monitored by UPLC-MS.

After reaction completion, the reaction solution was alkalized with 4 M KOH until pH > 12 to denature catalyst and render the tryptamine soluble in organic solvent. The tryptamine product was then extracted 3 times with 15 ml ethyl acetate, using centrifugation as needed to break emulsions. The organic extracts were then combined, dried over $MgSO_4$, filtered, and the organic solvent was evaporated to yield purified tryptamine analog.

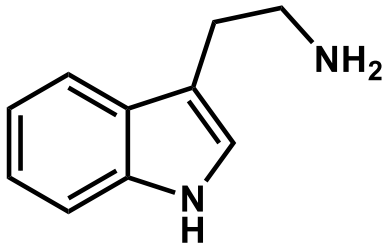
For serotonin and tyramine synthesis, reaction products were not separable via organic extraction, and were instead purified by a preparative C18 column isocratic separation with 1% methanol in water. Product fractions were pooled, evaporated to remove organic solvent, and dried via lyophilization.

Sequential one-pot enzymatic reactions from indole analogs:

One mmol of indole and 3 mmol (315 mg) serine were dissolved in 0–5 mL of methanol, as needed, in a 150 mL glass high pressure vial. The reaction mixture was diluted to 100 mL with 50 mM potassium phosphate buffer (pH = 8.0). The appropriate volume of *Pf*TrpB^{2B9} solution (thawed on ice and then centrifuged at 15,000 x*g* for 5 min to pellet aggregated protein) was added to bring the total enzyme concentration to 0.4 – 4 μM (500–5000 max TTN). The pressure vial was sealed to prevent evaporation and heated to 75 °C in a water bath overnight. The next day the vial was removed from the water bath and allowed to cool to room temperature on the benchtop. Once cooled, 100 μL of 20 mM aqueous PLP solution were added to bring the total PLP concentration to 20 μM. An appropriate volume of *Rgn*TDC (2 – 10 μM; 500–5000 Max TTN) was thawed on ice and then centrifuged at 15,000 x*g* for 5 min to pellet aggregated protein, added to the reaction mixture, and the sealed vial was incubated at 37 °C. The progress of the reaction was monitored by UPLC-MS. Once the reaction had reached an apparent 100% conversion, or after 12 hours, the reaction mixture was acidified with concentrated HCl until pH paper showed that the pH was < 2. The aqueous layer was washed 3 times with 100 mL ethyl acetate, using centrifugation as needed to break emulsions. The aqueous layer was then made basic with 6 M NaOH until the pH > 12. The basic layer was extracted 4 times with 100 mL ethyl acetate. The organic extracts from base were combined, dried over MgSO4, filtered, and the organic solvent was evaporated to yield purified tryptamine analog.

Structural Characterization of Tryptamine Analogs

Tryptamine (1): Following the method for the one-pot sequential reaction above, tryptamine was prepared from indole (117 mg, 1.0 mmol) and serine (105 mg, 3 mmol) in 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 2 μM (0.02% mol catalyst, 5000 max TTN) *Pf*TrpB^{2B9} and 2 μM (0.02% mol catalyst, 5000 max TTN) *Rgn*TDC. Product was obtained as an off-white powder, 147 mg (92% yield).

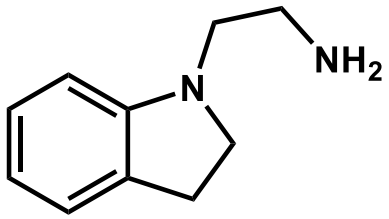


^{1}H NMR (500 MHz, DMSO-d_6) δ 10.79 (s, 1H), 7.51 (d, *J* = 7.8 Hz, 1H), 7.33 (d, *J* = 8.0 Hz, 1H), 7.13 (d, *J* = 2.3 Hz, 1H), 7.06 (ddd, *J* = 8.0, 6.9, 1.2 Hz, 1H), 6.96 (ddd, *J* = 7.9, 6.9, 1.0 Hz, 1H), 2.86 – 2.71 (m, 4H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 136.25, 127.32, 122.58, 120.78, 118.32, 118.08, 112.49, 111.30, 42.65, 29.39.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{13}N_{2;}$ calculated 161.1073, observed 161.1072

2,3-dihydro-isotryptamine (2): Following the method for the tryptophan analog decarboxylation above, 2,3-dihydro-isotryptamine was prepared from 2,3-dihydro-isotryptophan (29.9 mg, 0.145 mmol) in 50 mM phosphate buffer (pH = 8.0) using 30 μM (0.2% mol catalyst 500 max TTN) *Rgn*TDC. The product was isolated as a light tan oil, 19.1 mg (81% yield).

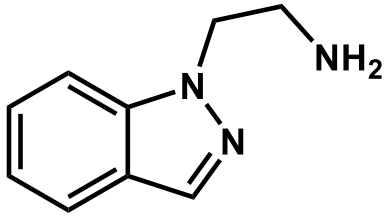


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.04 (dd, *J* = 7.3, 1.3, 1H), 7.01 (td, *J* = 7.7, 1.3, 1H), 6.63 (td, *J* = 7.4, 1.0, 1H), 6.56 (d, *J* = 7.9, 1H), 3.33 (d, *J* = 8.3, 1H), 3.15 (t, *J* = 6.3, 1H), 2.97 – 2.89 (m, 2H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 136.25, 127.32, 122.58, 120.78, 118.32, 118.08, 112.49, 111.30, 42.65, 29.39.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{15}N_{2;}$ calculated 163.1230, observed 163.1229.

2-azaisotryptamine (3): Following the method for the tryptophan analog decarboxylation above, 2-azaisotryptamine was prepared from 2-azaisotryptophan (19.0 mg, 0.092 mmol) in 50 mM phosphate buffer (pH = 8.0) using 30 μM (0.2% mol catalyst, 500 max TTN) *Rgn*TDC. This product was isolated as a white powder, 13.6 mg (91% yield).

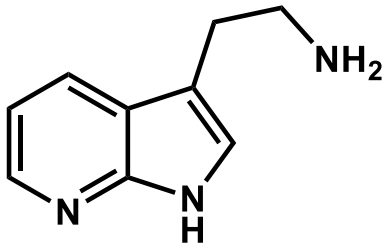


^{1}H NMR (400 MHz, Methanol-d_4) δ = 8.03 (d, *J* = 1.0, 1H), 7.75 (dt, *J* = 8.1, 1.0, 1H), 7.59 (dq, *J* = 8.5, 1.0, 1H), 7.41 (ddd, *J* = 8.4, 7.0, 1.1, 1H), 7.16 (ddd, *J* = 7.9, 6.9, 0.9, 1H), 4.46 (t, *J* = 6.3, 2H), 3.11 (t, *J* = 6.3, 2H).

^{13}C NMR (101 MHz, Methanol-d_4) δ = 141.30, 134.38, 127.83, 125.26, 122.16, 121.94, 110.26, 51.69, 42.38.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_{3;}$ calculated 162.1026, observed 162.1026.

7-azatryptamine (4): Following the method for the one-pot sequential reaction above, 7-azatryptamine was prepared from 7-azaindole (118.1 mg, 1.0 mmol) and serine (105 mg, 3 mmol) in 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 2 μM (5000 max TTN) *Pf*TrpB^{2B9} and 2 μM (0.02% mol catalyst, 5000 max TTN) *Rgn*TDC. The product was isolated as a tan, waxy solid, 130.5 mg (81% yield).

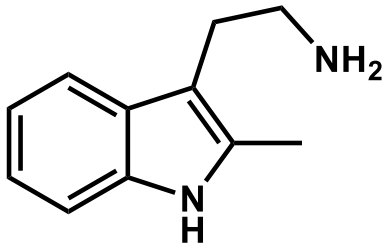


^{1}H NMR (500 MHz, DMSO-d_6) δ = 11.35 (s, 1H), 8.18 (d, *J* = 4.5, 1H), 7.94 (d, *J* = 7.7, 1H), 7.25 (d, *J* = 2.9, 1H), 7.02 (dd, *J* = 7.7, 4.6, 1H), 2.80 (dt, *J* = 31.8, 6.7, 4H).

^{13}C NMR (126 MHz, DMSO) δ 169.48, 149.11, 142.81, 142.74, 126.99, 126.87, 123.63, 119.97, 115.24, 115.17, 111.79, 111.35, 42.73, 29.34, 25.66, 23.14.

MS/ESI m/z for $[M+H]^+$; $C_9H_{12}N_3$; calculated 162.1026 , observed 162.1025.

2-methyltryptamine (5): Following the method for the one-pot sequential reaction above, 2-methyltryptamine was prepared from 2-methylindole (26.24 mg, 0.200 mmol) in 15 mL of 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 27 μM (0.2% mol catalyst, 500 max TTN) *Pf*TrpB^{2B9} and using 13 μM (0.1% mol catalyst,1000 max TTN) *Rgn*TDC. This product was isolated as an off-white oil, 16.1 mg (46% yield).

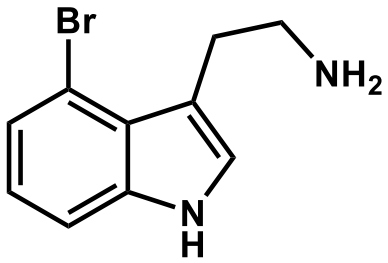


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.43 (dt, *J* = 7.7, 1.0, 1H), 7.23 (dt, *J* = 8.0, 1.0, 1H), 7.00 (ddd, *J* = 8.1, 7.0, 1.3, 1H), 6.95 (ddd, *J* = 8.0, 7.0, 1.2, 1H), 2.97 – 2.84 (m, *J* = 3.9, 4H), 2.38 (s, 3H), 1.89 (d, *J* = 6.9, 1H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 137.20, 133.47, 129.84, 121.46, 119.54, 118.30, 111.37, 108.14, 42.82, 27.22, 11.36.

MS/ESI m/z for $[M+H]^+$; $C_{11}H_{15}N_{2;}$ calculated 175.1230, observed 175.1229.

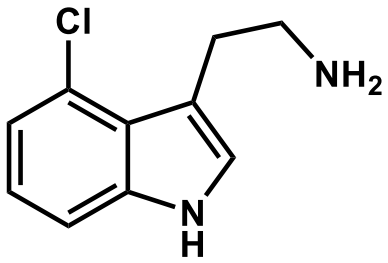
4-bromotryptamine (6): Following the method for the tryptophan analog decarboxylation above, 4-bromotryptamine was prepared from 4-bromotryptophan (22.1 mg, 0.092 mmol) in 50 mM phosphate buffer (pH=8.0) using 20 μM (0.2% mol catalyst 500 max TTN) *Rgn*TDC. This product was isolated as a pale-yellow oil, 1.7 mg (8% yield).



^{1}H NMR (400 MHz, Methanol-d_4) δ = 7.33 (dd, J = 8.2, 0.8, 1H), 7.18 (s, 1H), 7.20 – 7.14 (m, 1H), 6.95 (t, J = 7.9, 1H), 3.21 (t, J = 7.2, 2H), 3.12 – 3.06 (m, 2H). This result matches literature spectrum for this compound.[5]

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2Br$; calculated 239.0178, observed 239.0177.

4-chlorotryptamine (7): Following the method for the tryptophan analog decarboxylation above, 4-chlorotryptamine was prepared from 4-chlorotryptophan (9.8 mg, 0.050 mmol) in 50 mM phosphate buffer (pH=8) using 2.5 μM (0.05% mol catalyst, 2000 max TTN) *Rgn*TDC. This product was isolated as an off-white oil, 8.6 mg (88% yield).

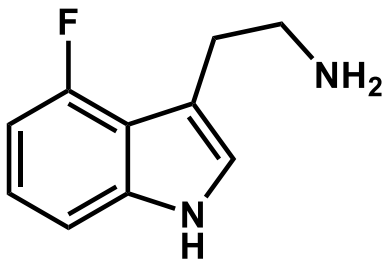


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.28 (dd, *J* = 7.9, 1.0, 1H), 7.14 (s, 1H), 7.02 (t, *J* = 7.8, 1H), 6.97 (dd, *J* = 7.6, 1.0, 1H), 3.15 (t, *J* = 7.1, 2H), 3.02 (t, *J* = 7.1, 2H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 139.95, 126.69, 125.79, 125.06, 122.97, 120.64, 113.03, 111.43, 44.04, 29.70.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2Cl$; calculated 195.0684, observed 195.0683.

4-fluorotryptamine (8): Following the method for the one-pot sequential reaction above, 4-fluorotryptamine was prepared from 4-fluoroindole (135 mg, 1.0 mmol) and serine (105 mg, 3 mmol) in 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 2 μM (0.02% mol catalyst, 5000 max TTN) *Pf*TrpB^{2B9} and 2 μM (0.02% mol catalyst, 5000 max TTN) *Rgn*TDC. The product was isolated as a white powder, 178 mg (quantitative yield).



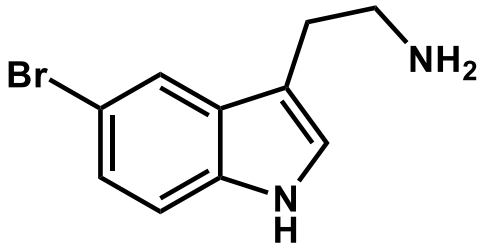
^{1}H NMR (500 MHz, DMSO-d_6) δ = 11.09 (s, 1H), 7.14 (t, 2H), 6.99 (td, *J* = 7.9, 5.2, 1H), 6.68 (dd, *J* = 11.6, 7.7, 1H), 2.81 (s, 4H).

^{13}C NMR (126 MHz, DMSO-d_6) δ = 157.52, 155.58, 139.37, 139.27, 123.32, 121.32, 121.25, 115.54, 115.38, 110.89, 110.86, 107.96, 107.93, 103.22, 103.06, 43.10, 43.09, 30.63.

^{19}F NMR (377 MHz, DMSO-d_6) δ -124.69.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2F$; calculated 179.0979, observed 179.0978.

5-bromotryptamine (9): Following the method for the one-pot sequential reaction above, 4-bromotryptamine was prepared from 5-bromoindole (39.21 mg, 0.2 mmol) and serine (63.1 mg, 0.6 mmol) in 100 mL of 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 4 μM (0.2% mol catalyst, 500 max TTN) *Tma*TrpB^{2F3} and 2 μM (0.1% mol catalyst, 1000 max TTN) *Rgn*TDC. This product was isolated as a pale-yellow oil, 20.8 mg (43% yield).

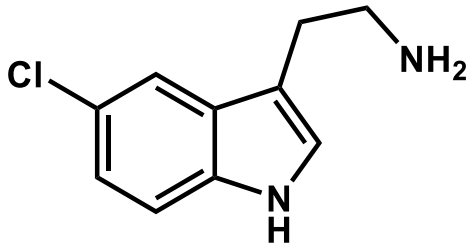


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.70 (d, *J* = 1.9, 1H), 7.27 (d, *J* = 8.5, 1H), 7.18 (dd, *J* = 8.6, 1.9, 1H), 7.13 (s, 1H), 2.98 (t, *J* = 7.1, 2H), 2.91 (t, *J* = 6.7, 2H), .

^{13}C NMR (126 MHz, Methanol-d_4) δ = 136.85, 130.40, 125.35, 125.16, 121.79, 114.00, 112.90, 112.41, 42.50, 27.86.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2Br$; calculated 239.0178, observed 239.0176.

5-chlorotryptamine (10): Following the method for the one-pot sequential reaction above, 5-chlorotryptamine was prepared from 5-chloroindole (45.5 mg, 0.3 mmol) and serine (94.6 mg, 0.9 mmol) in 10 mL of 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 60 μM (0.2% mol catalyst, 500 max TTN) *Pf*TrpB^{2B9} and 30 μM (0.1% mol catalyst, 1000 max TTN) *Rgn*TDC. This product was isolated as a pale tan powder, 30.5 mg (52% yield).

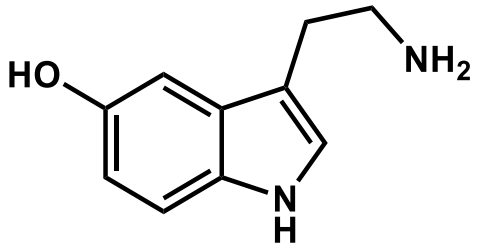


^{1}H NMR (500 MHz, Acetonitrile-d_3) δ = 9.53 (s, 1H), 7.57 (d, *J* = 2.1, 1H), 7.35 (d, *J* = 8.7, 1H), 7.12 (s, 1H), 7.08 (dd, *J* = 8.6, 2.1, 1H), 2.87 (q, *J* = 5.3, 3.6, 4H), 1.64 (s, 2H).

^{13}C NMR (126 MHz, Acetonitrile-d_3) δ = 135.90, 129.74, 125.28, 124.74, 122.15, 118.85, 114.16, 113.56, 43.47, 30.02.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2Cl$; calculated 195.0684, observed 195.0683.

Serotonin (11): Following the method for the tryptophan analog decarboxylation above, serotonin was prepared from 5-hydroxytryptophan (34.8 mg, 0.158 mmol) in 50 mM phosphate buffer (pH = 8.0) using 30 μM (0.2% mol catalyst, 500 max TTN) *Rgn*TDC. The reaction solution was then evaporated to remove excess solvent and purified via a reverse phase C18 column. This product was isolated as a dark grey powder, 20.9 mg (60% yield).

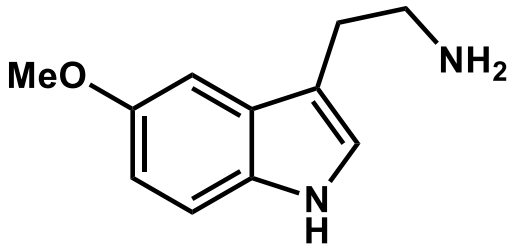


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.18 (d, *J* = 8.6, 1H), 7.06 (s, 1H), 6.92 (d, *J* = 2.3, 1H), 6.69 (dd, *J* = 8.6, 2.4, 1H), 3.07 (t, *J* = 7.1, 2H), 2.95 (t, *J* = 7.1, 2H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 151.35, 133.25, 129.11, 124.76, 112.86, 112.63, 110.80, 103.29, 41.85, 26.73.

NMR spectra match literature precedent.[35]

5-methoxytryptamine (12): Following the method for the one-pot sequential reaction above, 5-methoxytryptamine was prepared from 5-methoxyindole (147.1 mg, 1.0 mmol) and serine (105 mg, 3 mmol) in 50 mM phosphate buffer (pH = 8.0) using 4 μM (0.04% mol catalyst, 2500 max TTN) *Pf*TrpB^{2B9} and 20 μM (0.2% mol catalyst, 500 max TTN) *Rgn*TDC. 5-OMe-tryptophan was unable to be fully converted to the corresponding tryptamine as per analysis by UPLC-MS, however some product formed and was isolated as a tan waxy solid, 27 mg (14% yield).

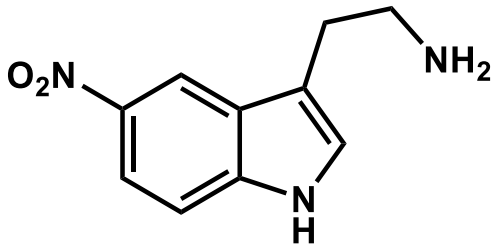


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.25 (d, *J*=8.7, 1H), 7.08 – 7.04 (m, 2H), 6.78 (dd, *J*=8.8, 2.5, 1H), 3.85 (s, 3H), 2.92 (ddd, *J*=20.1, 6.2, 0.9, 3H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 153.52, 132.10, 127.63, 122.90, 111.83, 111.49, 111.17, 99.87, 54.88, 41.67, 28.19.

MS/ESI: m/z for $[M+H]^+$; $C_{11}H_{15}N_2O$; calculated 191.1179, observed 191.1177.

5-nitrotryptamine (13): Following the method for the one-pot sequential reaction above, 5-nitrotryptamine was prepared from 5-nitroindole 32.4 mg, 0.2 mmol) and serine (63.1 mg, 0.6 mmol) in 100 mL of 50 mM phosphate buffer (pH = 8.0) using 4 μM (0.04% mol catalyst, 2500 max TTN) *Pf*TrpB^{5G8} and 10 μM (0.5% mol catalyst, 200 max TTN) *Rgn*TDC. UPLC-MS analysis showed incomplete conversion of indole to tryptophan as well as tryptophan to tryptamine. The product was isolated as a yellow powder, 13.9 mg (34% yield).

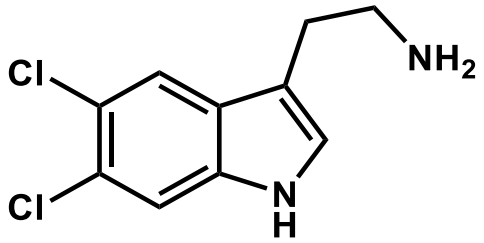


^{1}H NMR (400 MHz, Methanol-d_4) δ = 8.61 (d, *J* = 2.2, 1H), 8.06 (dd, *J* = 9.0, 2.3, 1H), 7.48 (d, *J* = 9.0, 1H), 7.38 (s, 1H), 3.23 – 3.05 (m, 4H).

^{13}C NMR ^{13}C NMR (101 MHz, Methanol-d_4) δ = 142.51, 141.30, 127.87, 127.77, 118.07, 116.48, 114.74, 112.62, 41.94, 26.13.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_3O_2$; calculated 206.0924, observed 206.0924.

5,6-dichlorotryptamine (14): Following the method for the one-pot sequential reaction above, 5,6-dichlorotryptamine was prepared from 5,6-dichloroindole (185 mg, 1.0 mmol) and serine (105 mg, 3 mmol) in 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 4 μM (0.04% mol catalyst, 2500 max TTN) *Pf*TrpB^{2B9} and 20 μM (0.2% mol catalyst, 500 max TTN) *Rgn*TDC. The product was isolated as tan powder, 27 mg (12% yield).

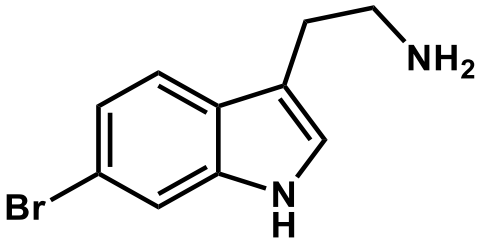


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.72 (s, 1H), 7.52 (s, 1H), 7.22 (s, 1H), 3.02 (dt, *J* = 48.1, 7.3, 4H), 1.90 (s, 2H).

^{13}C NMR (126 MHz, Methanol-d_4) δ 135.66, 127.03, 125.15, 124.72, 122.26, 118.94, 112.54, 110.54, 40.54, 25.05.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{11}N_2Cl_2$; calculated 229.0294, observed 229.0294.

6-bromotryptamine (15): Following the method for the one-pot sequential reaction above, 6-bromotryptamine was prepared from 6-bromoindole (39.21 mg, 0.20 mmol) and serine (63.05 mg, 0.6 mmol) in 15 mL of 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 27 μM (0.2% mol catalyst, 500 max TTN) *Pf*TrpB^{2B9} and 13 μM (0.1% mol catalyst, 1000 max TTN) *Rgn*TDC. This product was isolated a white powder, 26.1 mg (54% yield).

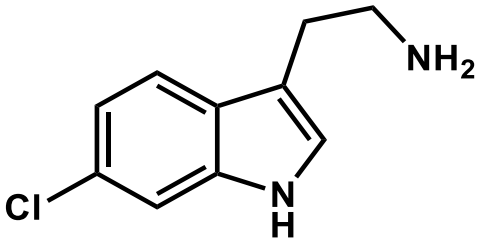


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.50 (d, *J* = 1.7, 1H), 7.44 (d, *J* = 8.4, 1H), 7.10 (dd, *J* = 8.4, 1.7, 1H), 7.08 (s, 1H), 2.95 – 2.83 (m, 4H), 1.90 (s, 0H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 139.03, 127.70, 124.49, 122.71, 120.74, 115.79, 115.11, 113.73, 42.98, 29.15.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2Br$; calculated 239.0178, observed 239.0177.

6-chlorotryptamine (16): Following the method for the one-pot sequential reaction above, 6-chlorotryptamine was prepared from 6-chloroindole (30.3 mg, 0.20 mmol) and serine (63.05 mg, 0.6 mmol) in 15 mL of 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 13 μM (0.1% mol catalyst, 1000 max TTN) *Pf*TrpB^{2B9} and 13 μM (0.1% mol catalyst, 1000 max TTN) *Rgn*TDC. This product was isolated a pale orange powder, 27.8 mg (71% yield).

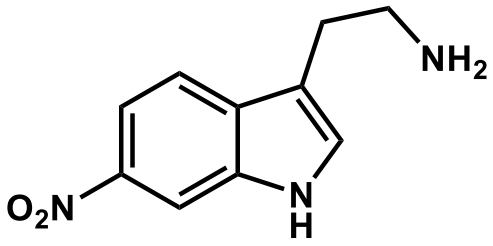


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.51 (d, *J* = 8.4, 1H), 7.36 (d, *J* = 1.8, 1H), 7.16 (s, 1H), 7.01 (dd, *J* = 8.5, 1.9, 1H), 3.10 (t, *J* = 7.2, 2H), 3.01 (t, *J* = 7.3, 2H), 1.90 (s, 2H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 179.05, 137.26, 127.13, 125.73, 123.62, 118.98, 118.78, 110.84, 110.51, 40.42, 24.84.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2Cl$; calculated 195.0684, observed 195.0683.

6-nitrotryptamine (crude) (17): Following the method for the one-pot sequential reaction above, 6-nitrotryptamine was prepared from 6-nitroindole (32.43 mg, 0.20 mmol) and serine (63.05 mg, 0.6 mmol) in 100 mL of 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 4 μM (0.2% mol catalyst, 500 max TTN) *Pf*TrpB^{0A9} and 2 μM (0.05% mol catalyst, 2000 max TTN) *Rgn*TDC. This product was isolated a bright yellow powder, 27.5 mg (67% yield), with a small amount of 6-nitrotryptophan impurity.



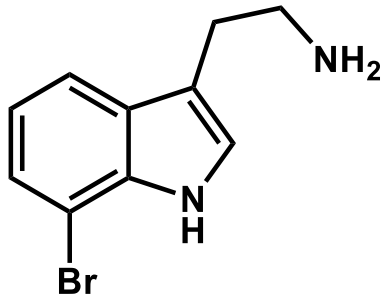
6-nitrotryptamine (purified): It was observed that 6-nitrotryptophan co-purified with 6-nitrotryptamine. This level of purity was acceptable for our purposes, but in order to show that a 100% pure product can be isolated, we purified the crude 6-nitrotryptamine product via a reverse phase C18 column and dried via evaporation and lyophilization. The purified product was isolated as a bright yellow powder, 8.0 mg (19% yield).

^{1}H NMR (500 MHz, Methanol-d_4) δ = 8.33 (d, *J*=2.0, 1H), 7.94 (dd, *J*=8.8, 2.1, 1H), 7.69 (d, *J*=8.8, 1H), 7.48 (s, 1H), 2.96 (d, *J*=2.6, 4H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 144.14, 136.53, 133.42, 130.35, 119.38, 114.91, 114.90, 109.15, 43.02, 28.85.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_3O_2$; calculated 206.0924, observed 206.0924.

7-bromotryptamine (18): Following the method for the one-pot sequential reaction above, 7-bromotryptamine was prepared from 7-bromoindole (39.21 mg, 0.20 mmol) and serine (63.05 mg, 0.6 mmol) in 100 mL of 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 4 μM (0.2% mol catalyst, 500 max TTN) *Pf*TrpB^{2B9} and 1 μM (0.05% mol catalyst, 2000 max TTN) *Rgn*TDC. This product was isolated as an off-white powder, 24.7 mg (51% yield).



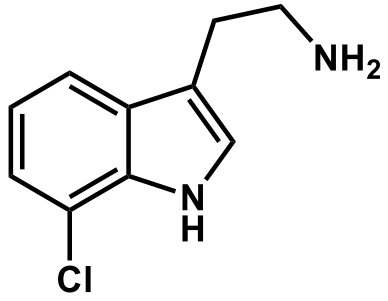
^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.54 (dd, *J* = 7.8, 0.9, 1H), 7.27 (d, *J* = 7.8, 1H), 7.17 (s, 1H), 6.93 (t, *J* = 7.8, 1H), 3.02 – 2.89 (m, 4H), 1.90 (s, 1H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 136.62, 130.24, 124.99, 124.87, 124.84, 120.92, 118.73, 114.33, 105.62, 42.69, 28.71.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2Br$; calculated 239.0178, observed 239.0176.

\

7-chlorotryptamine (19): Following the method for the tryptophan analog decarboxylation above, 7-chlorotryptamine was prepared from 4-chlorotryptophan (26.3 mg, 0.110 mmol) in 50 mM phosphate buffer (pH = 8.0) using 2.2 μM (0.2% mol catalyst, 500 max TTN) *Rgn*TDC. This product was isolated as a white powder, 20.1 mg (93% yield).

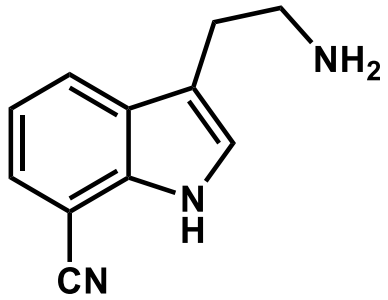


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.49 (dd, *J* = 7.9, 0.9, 1H), 7.16 (s, 1H), 7.11 (d, *J* = 7.5, 1H), 6.98 (t, *J* = 7.8, 1H), 2.98 – 2.88 (m, 4H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 133.73, 129.15, 123.41, 120.46, 119.10, 116.81, 116.36, 113.02, 41.43, 27.55.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2Cl$; calculated 195.0684, observed 195.0684.

7-cyanotryptamine (20): Following the method for the one-pot sequential reaction above, 7-cyanotryptamine was prepared from 7-cyanoindole (28.43 mg, 0.20 mmol) and serine (63.05 mg, 0.6 mmol) in 10 mL of 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 40 μM (0.2% mol catalyst, 500 max TTN) *Pf*TrpB^{2B9} and 20 μM (0.1% mol catalyst, 1000 max TTN) *Rgn*TDC. This product was isolated as a tan powder, 19.8 mg (53% yield).

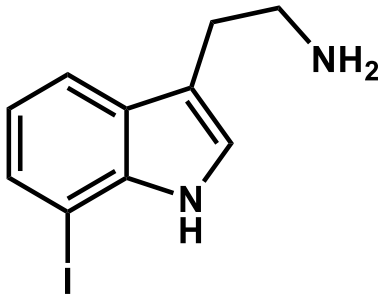


^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.88 (dd, *J* = 8.1, 1.1, 1H), 7.48 (dd, *J* = 7.4, 1.0, 1H), 7.25 (s, 1H), 7.13 (t, *J* = 7.7, 1H), 2.99 – 2.91 (m, 4H), 1.90 (s, 1H).

^{13}C NMR (126 MHz, Methanol-d_4) δ = 137.85, 129.83, 127.38, 125.79, 125.04, 119.73, 118.31, 114.87, 95.29, 42.82, 28.61.

MS/ESI m/z for $[M+H]^+$; $C_{11}H_{12}N_{3;}$ calculated 186.1026, observed 186.1026.

7-iodotryptamine (21): Following the method for the one-pot sequential reaction above, 7-iodotryptamine was prepared from 7-iodoindole (243 mg, 1.0 mmol) and serine (105 mg, 3 mmol) in 50 mM phosphate buffer (pH = 8.0) with 5% MeOH as a cosolvent using 2 μM (0.02% mol catalyst, 5000 max TTN) *Pf*TrpB^{2B9} and 2 μM (0.02% mol catalyst, 5000 max TTN) *Rgn*TDC. The product was isolated as a white powder, 206 mg (77%).

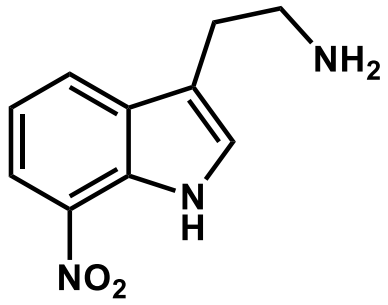


^{1}H NMR (500 MHz, DMSO-d_6) δ 10.76 (s, 1H), 7.56 (dd, J = 7.8, 0.9 Hz, 1H), 7.47 (dd, J = 7.3, 0.9 Hz, 1H), 7.18 (s, 1H), 6.80 (t, J = 7.6 Hz, 1H), 2.78 (dt, J = 37.6, 7.2 Hz, 4H).

^{13}C NMR (126 MHz, DMSO) δ 138.41, 130.22, 128.39, 124.17, 120.60, 119.02, 114.52, 77.42, 42.85, 29.68.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_2I$; calculated 287.0040, observed 287.0035.

7-nitrotryptamine (22): Following the method for the one-pot sequential reaction above, 5-nitrotryptamine was prepared from 5-nitroindole (32.4 mg, 0.2 mmol) and serine (63.1 mg, 0.6 mmol) in 100 mL of 50 mM phosphate buffer (pH = 8.0) using 4 μM (0.05% mol catalyst, 2500 max TTN) *Pf*TrpB^{5G8} and 10 μM (0.5% mol catalyst, 200 max TTN) *Rgn*TDC. UPLC-MS analysis showed incomplete conversion of the indole to tryptophan as well as incomplete conversion of tryptophan to tryptamine. The product was isolated as a yellow powder, 5.5 mg (13% yield).

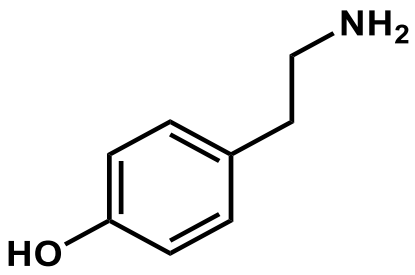


^{1}H NMR (500 MHz, Methanol-d_4) δ = 8.14 (dd, *J* = 8.0, 1.0, 1H), 8.03 (dd, *J* = 7.7, 1.0, 1H), 7.33 (s, 1H), 7.21 (t, *J* = 7.9, 1H), 2.99 (s, 5H).

^{13}C NMR (126 MHz, MeOD) δ = 134.31, 133.21, 130.64, 127.71, 126.76, 119.72, 119.24, 115.29, 42.87, 28.49.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{12}N_3O_2$; calculated 206.0924, observed 206.0924.

Tyramine (23): Following the method for the tryptophan analog decarboxylation above, tyramine was prepared from tyrosine (40 mg, 0.221 mmol) in 50 mM phosphate buffer (pH = 8.0) using 40 μM (0.2% mol catalyst, 500 max TTN) *Rgn*TDC. After quenching the reaction, the reaction solution was evaporated to remove excess solvent and purified via a reverse phase C18 column. This product was isolated as yellow powder, 15.7 mg (51% yield).



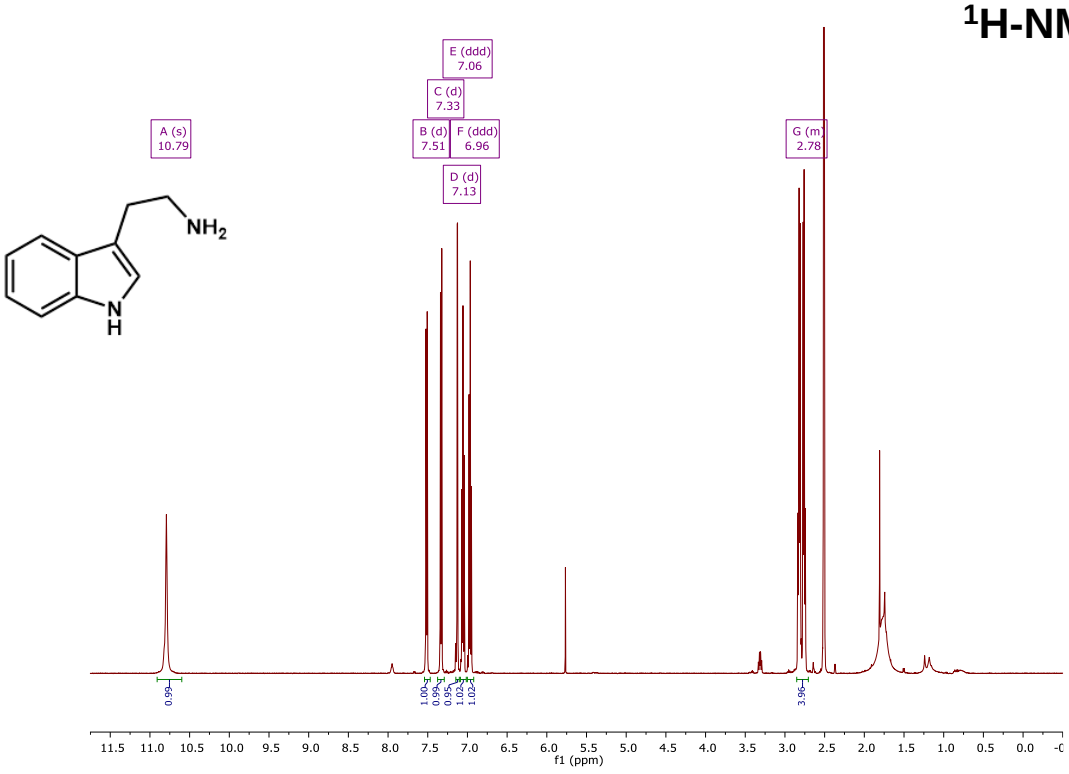
^{1}H NMR (500 MHz, Methanol-d_4) δ = 7.06 – 7.01 (m, 1H), 6.75 – 6.70 (m, 1H), 2.89 (t, *J* = 7.4, 1H), 2.70 (t, *J* = 7.3, 1H).

^{13}C NMR (126 MHz, MeOD) δ = 155.80, 129.32, 129.16, 115.02, 42.29, 36.09.

MS/ESI m/z for $[M+H]^+$; $C_8H_{12}NO$; calculated 138.0913, observed 138.0913

2-(1H-Indol-3-yl)ethanamine (1)

^{1}H-NMR



^{13}C-NMR

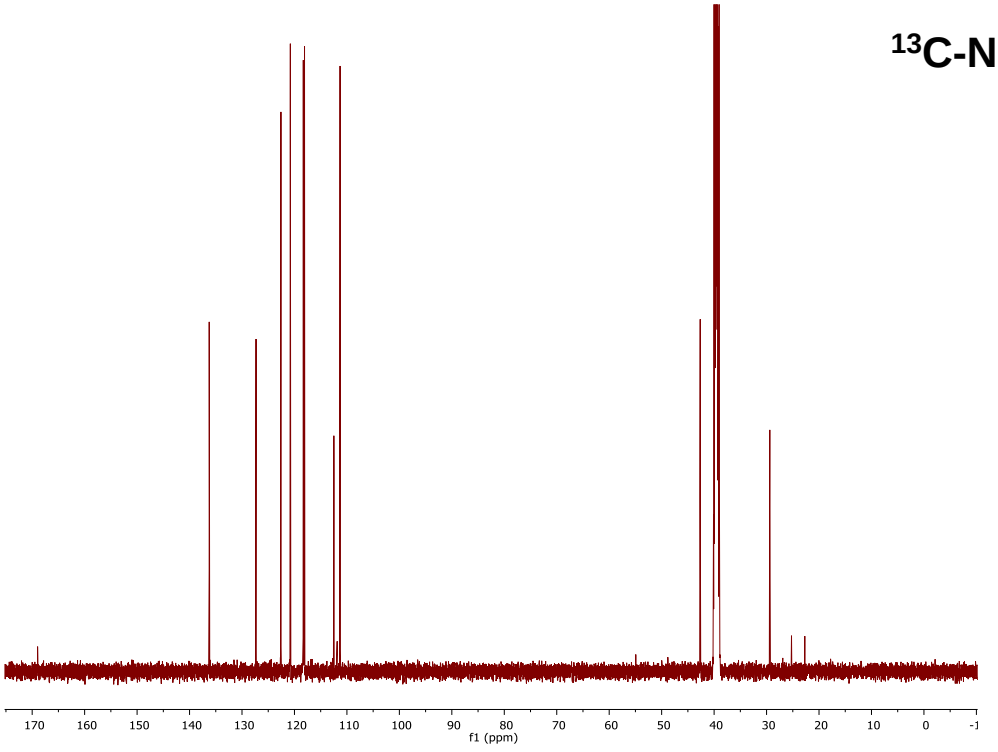
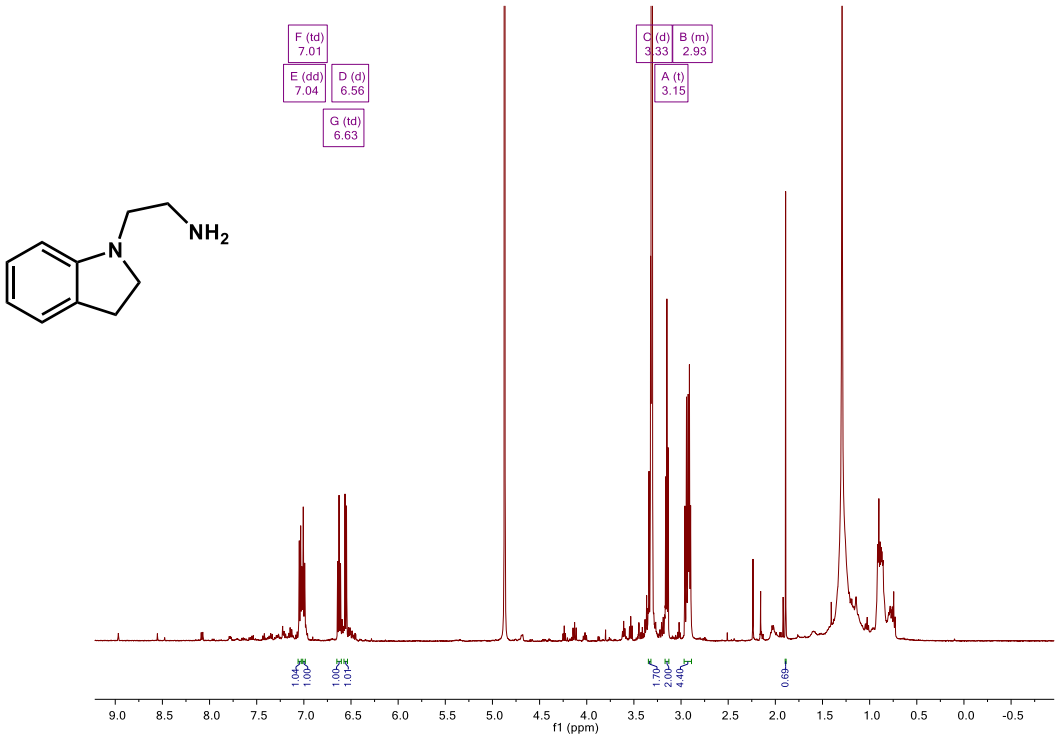


Figure 11. ^{1}H and ^{13}C NMR spectra of 2-(1H-Indol-3-yl)ethanamine (1).

2-(2,3-dihydro-indol-1-yl)-ethylamine (2)

^{1}H-NMR



^{13}C-NMR

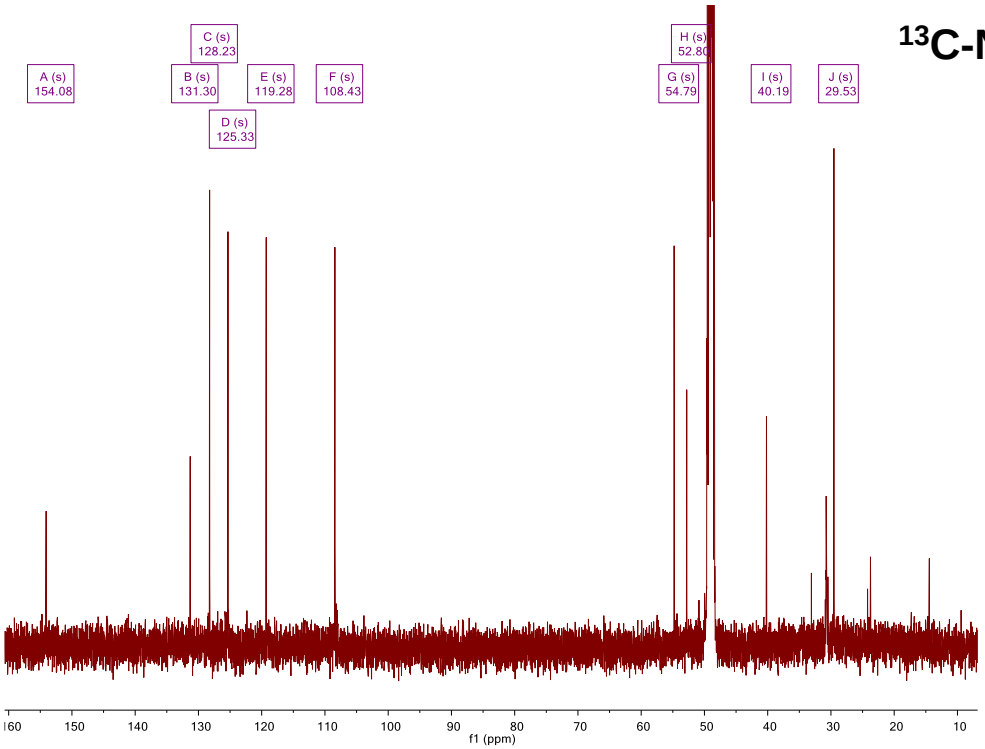
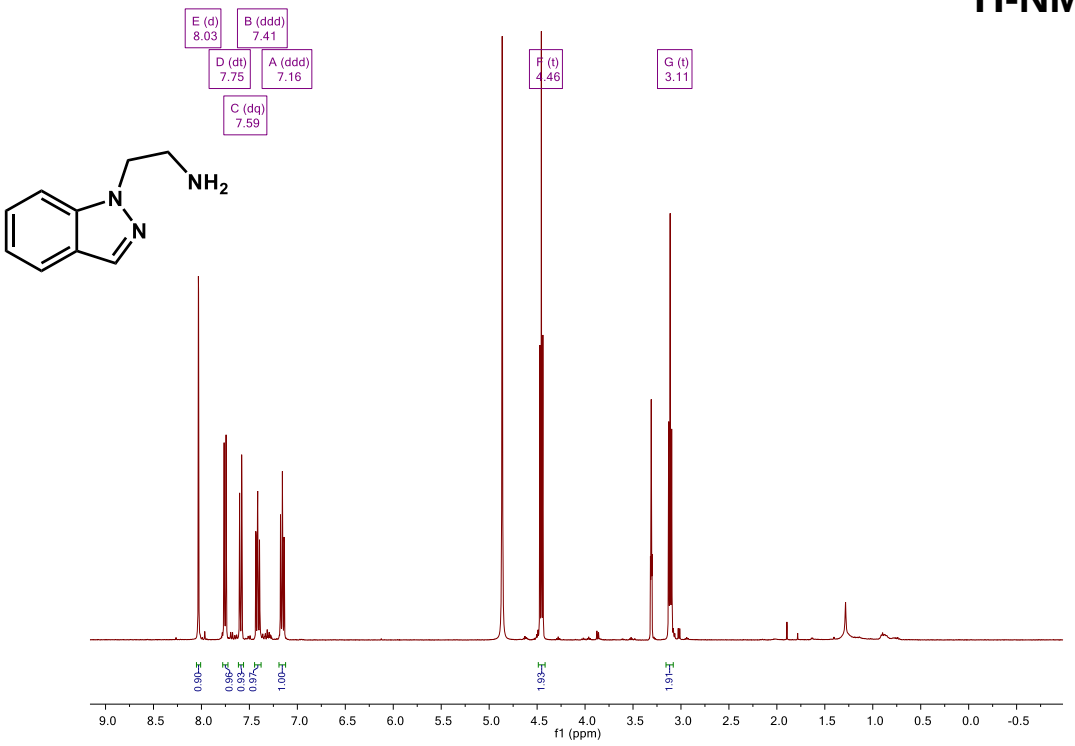


Figure 12. ^{1}H and ^{13}C NMR spectra of 2-(2,3-dihydro-indol-1-yl)-ethylamine (2).

2-(1H-indazol-1-yl)ethylamine (3)

^{1}H-NMR



^{13}C-NMR

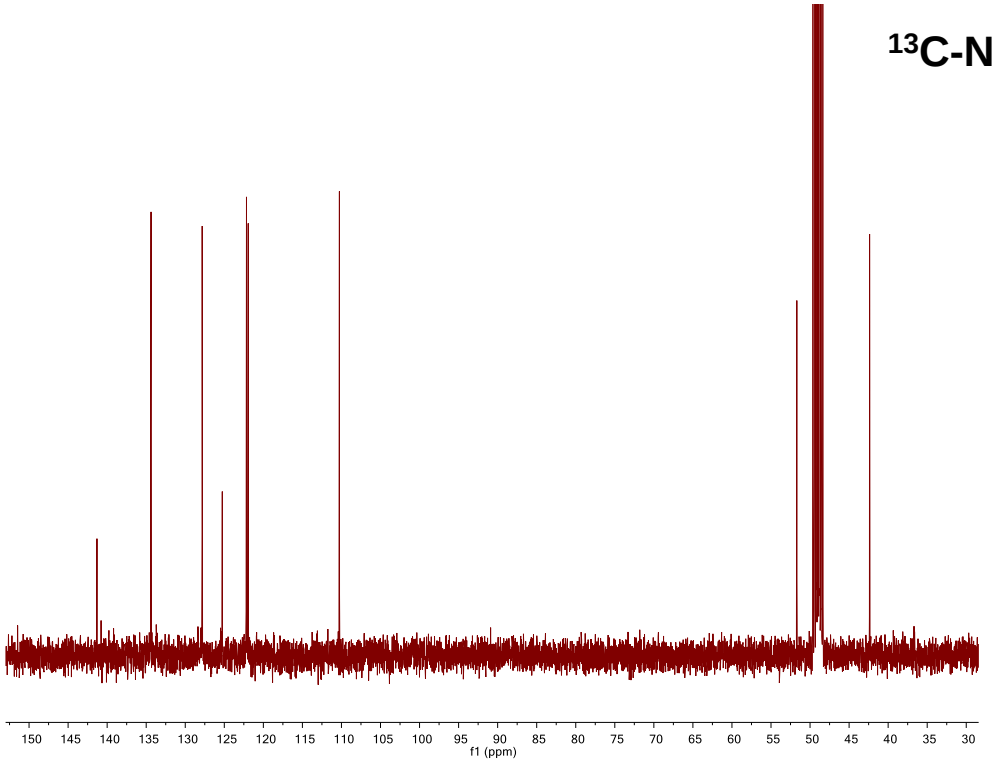
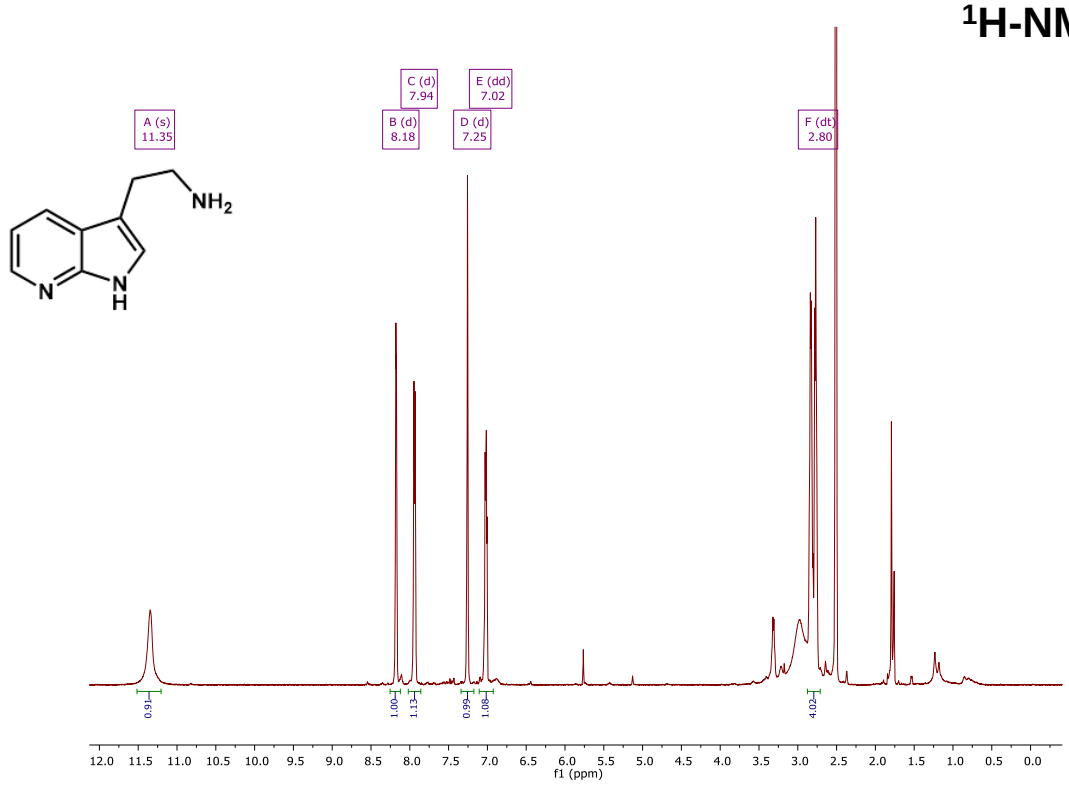


Figure 13. ^{1}H NMR and ^{13}C spectra of 2-(1H-indazol-1-yl)ethylamine (3).

2-(1H-pyrrollo[2,3-b]pyridin-3-yl)ethanamine (4)

^{1}H-NMR



^{13}C-NMR

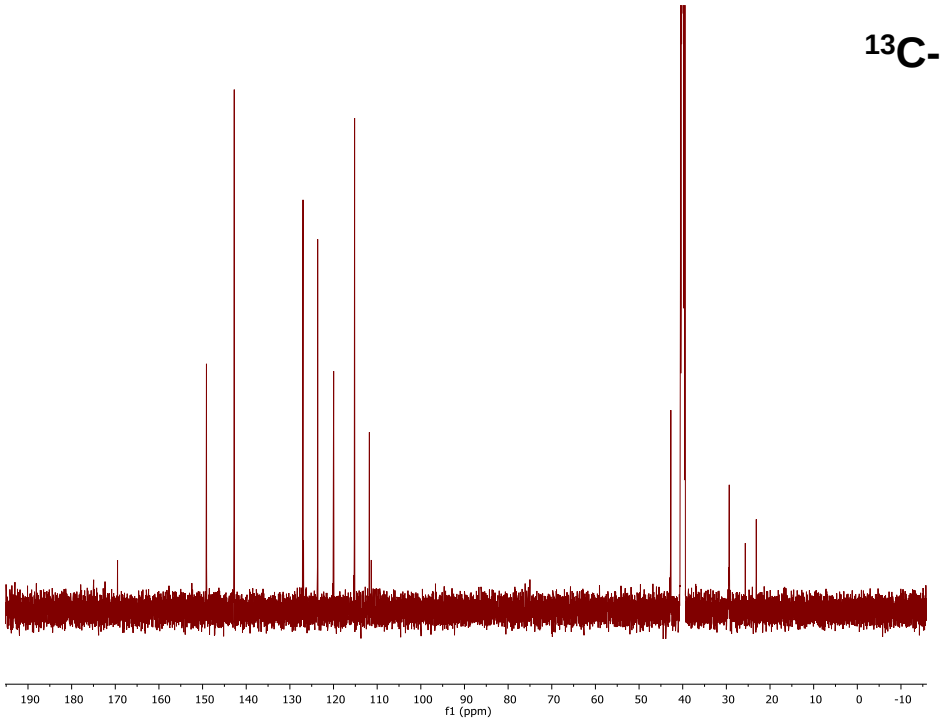
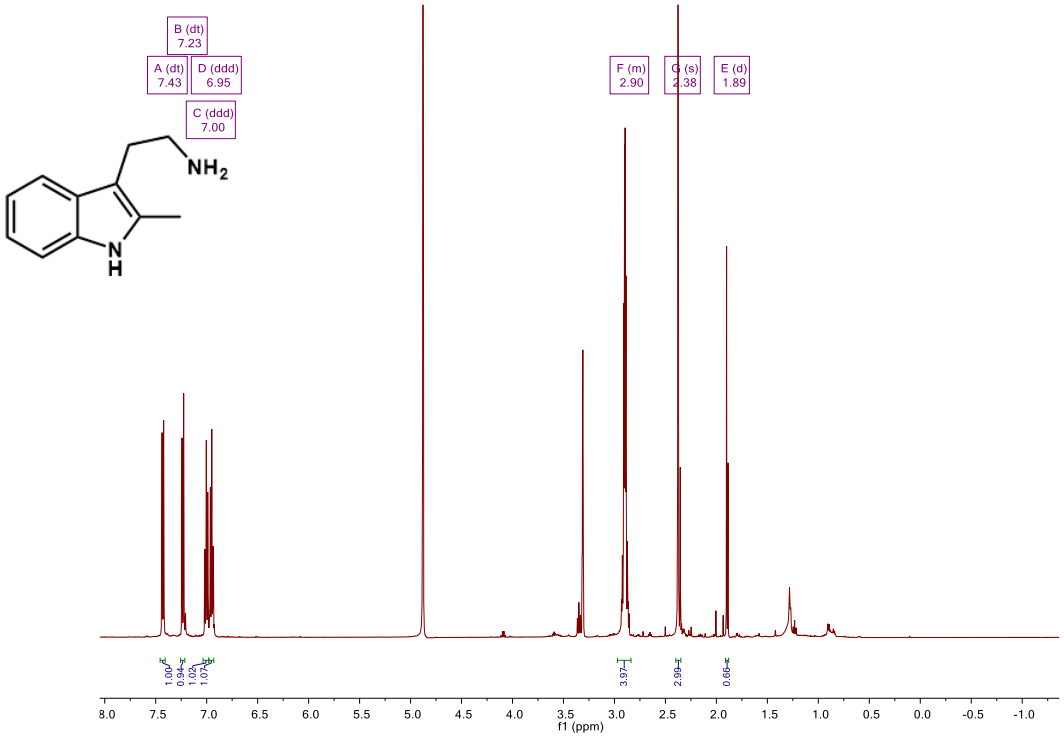


Figure 14. ^{1}H NMR and ^{13}C spectra of 2-(1H-pyrrollo[2,3-b]pyridin-3-yl)ethanamine (4).

2-(2-methyl-1H-indol-3-yl)ethanamine (5)

^{1}H-NMR



^{13}C-NMR

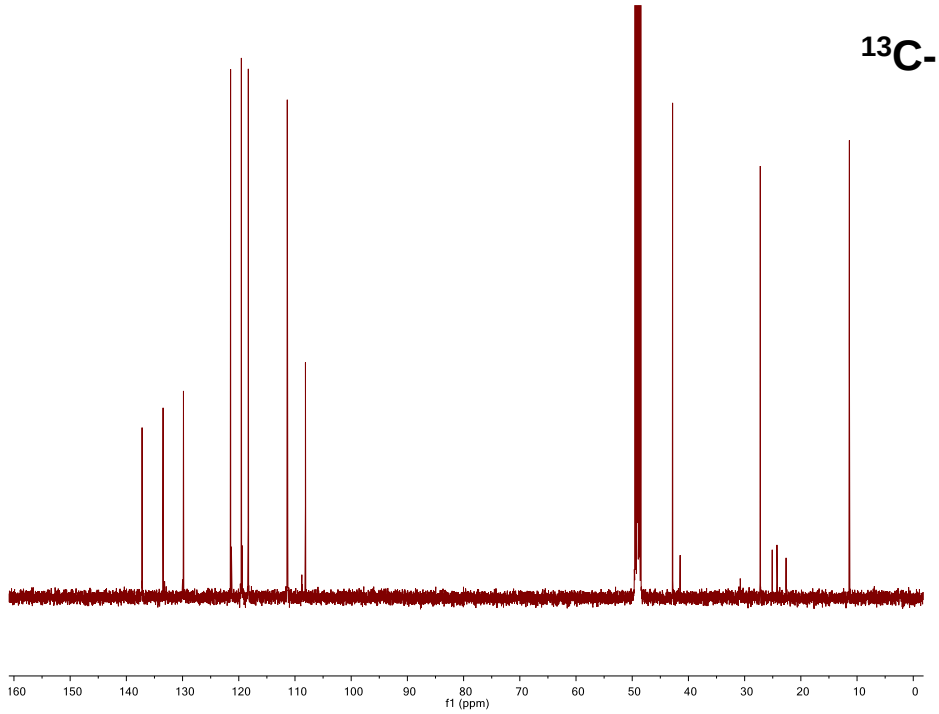


Figure 15. ^{1}H NMR and ^{13}C spectra of 2-(2-methyl-1H-indol-3-yl)ethanamine (5).

2-(4-bromo-1H-indol-3-yl)ethanamine (6)

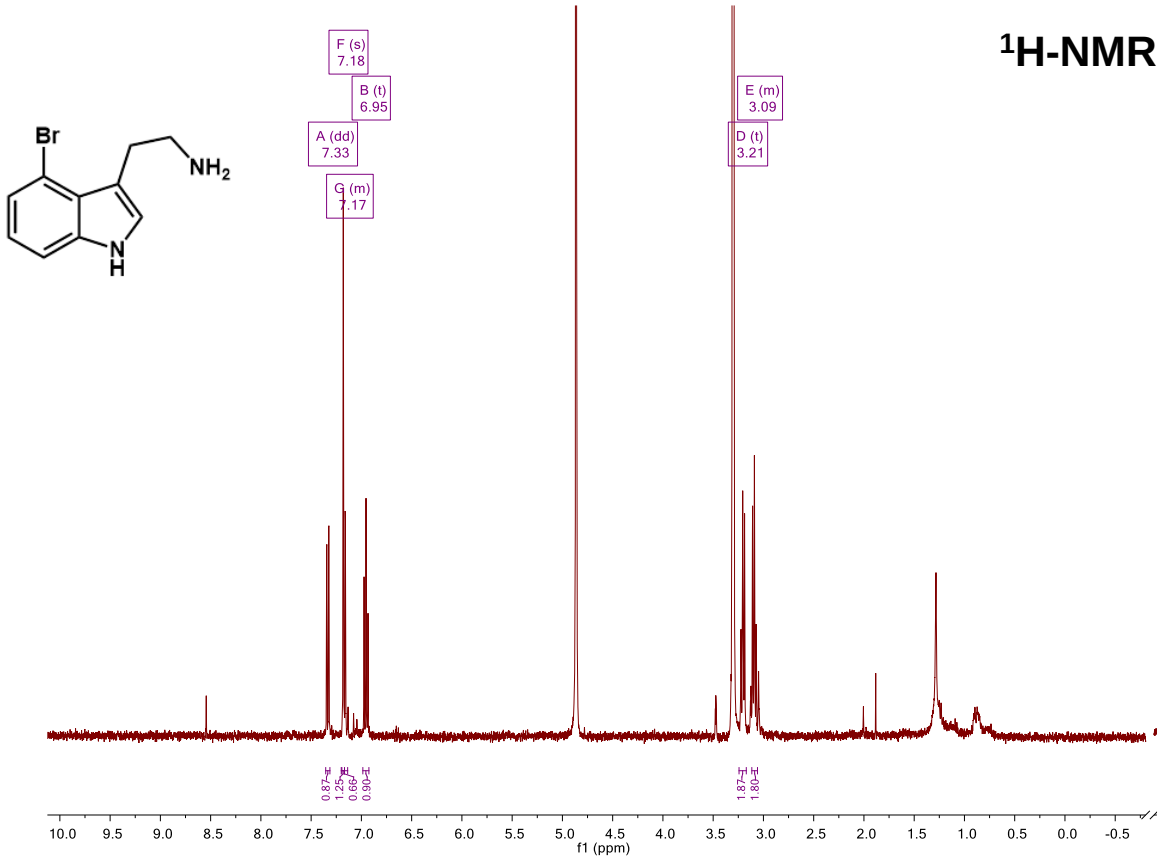


Figure 16. ^{1}H NMR spectrum of 2-(4-bromo-1H-indol-3-yl)ethanamine (6).

2-(4-chloro-1H-indol-3-yl)ethanamine (7)

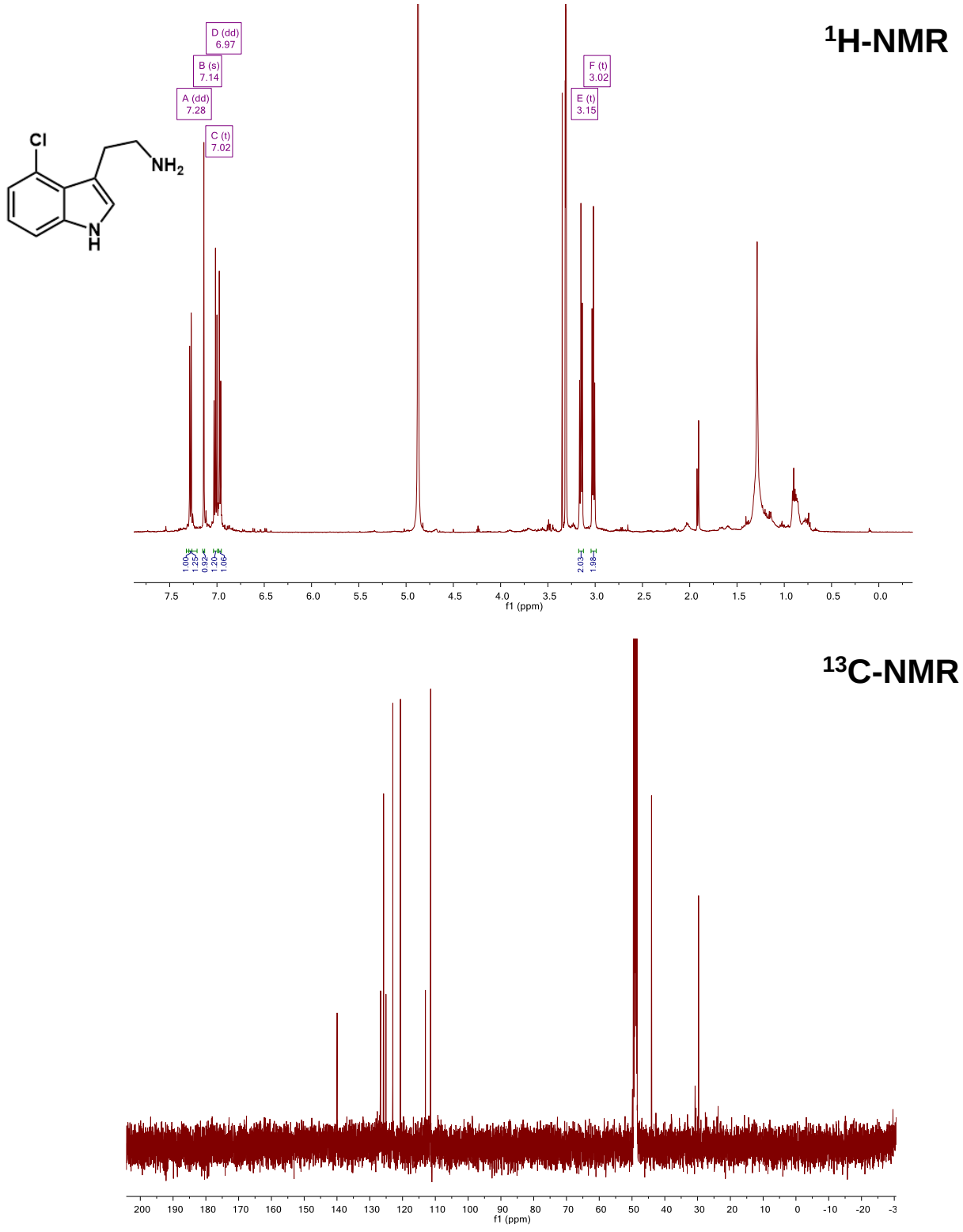
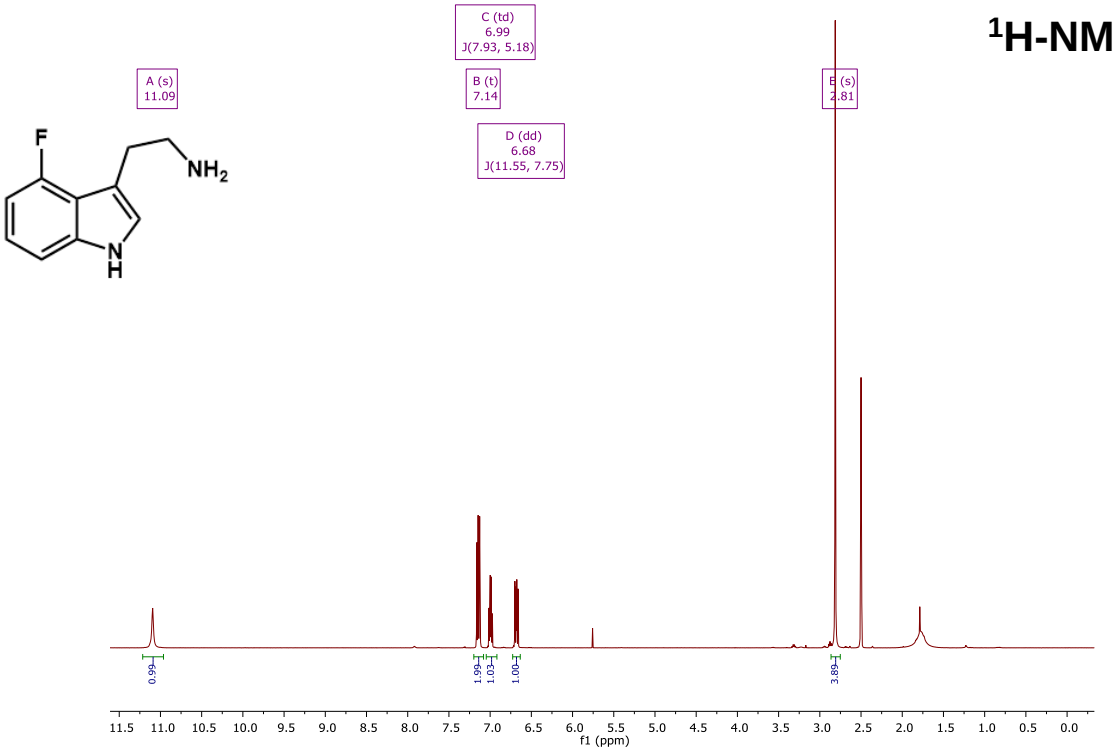


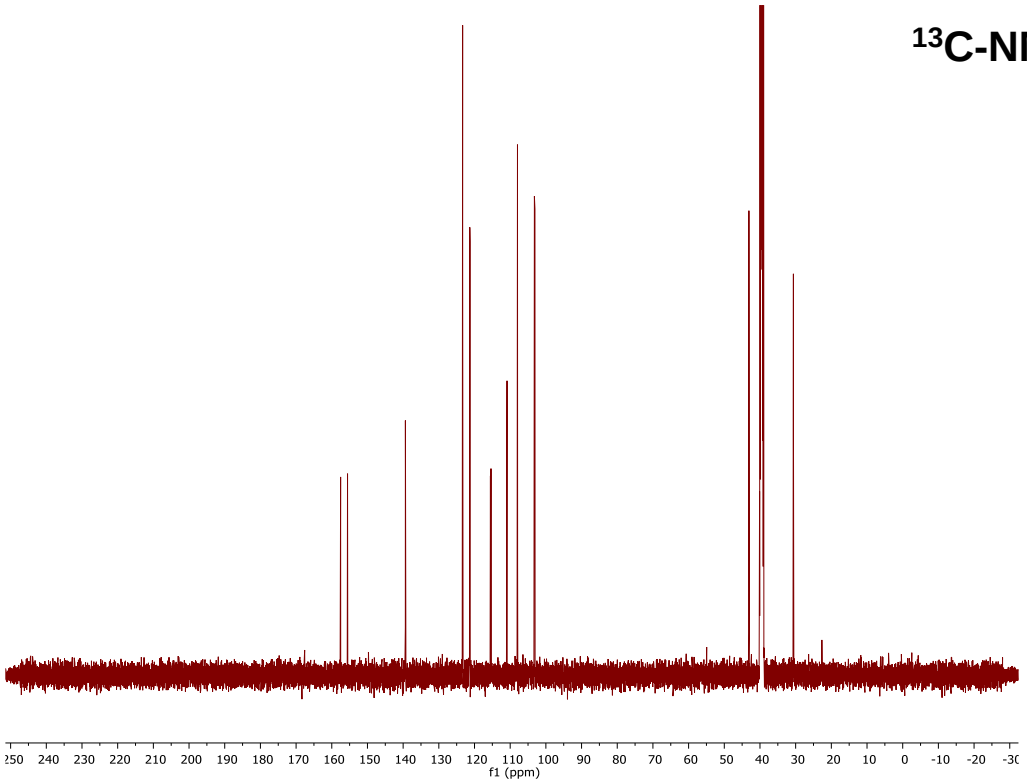
Figure 17. ^{1}H NMR and ^{13}C spectra of 2-(4-chloro-1H-indol-3-yl)ethanamine (7).

2-(4-fluoro-1H-indol-3-yl)ethanamine (8)

^{1}H-NMR



^{13}C-NMR



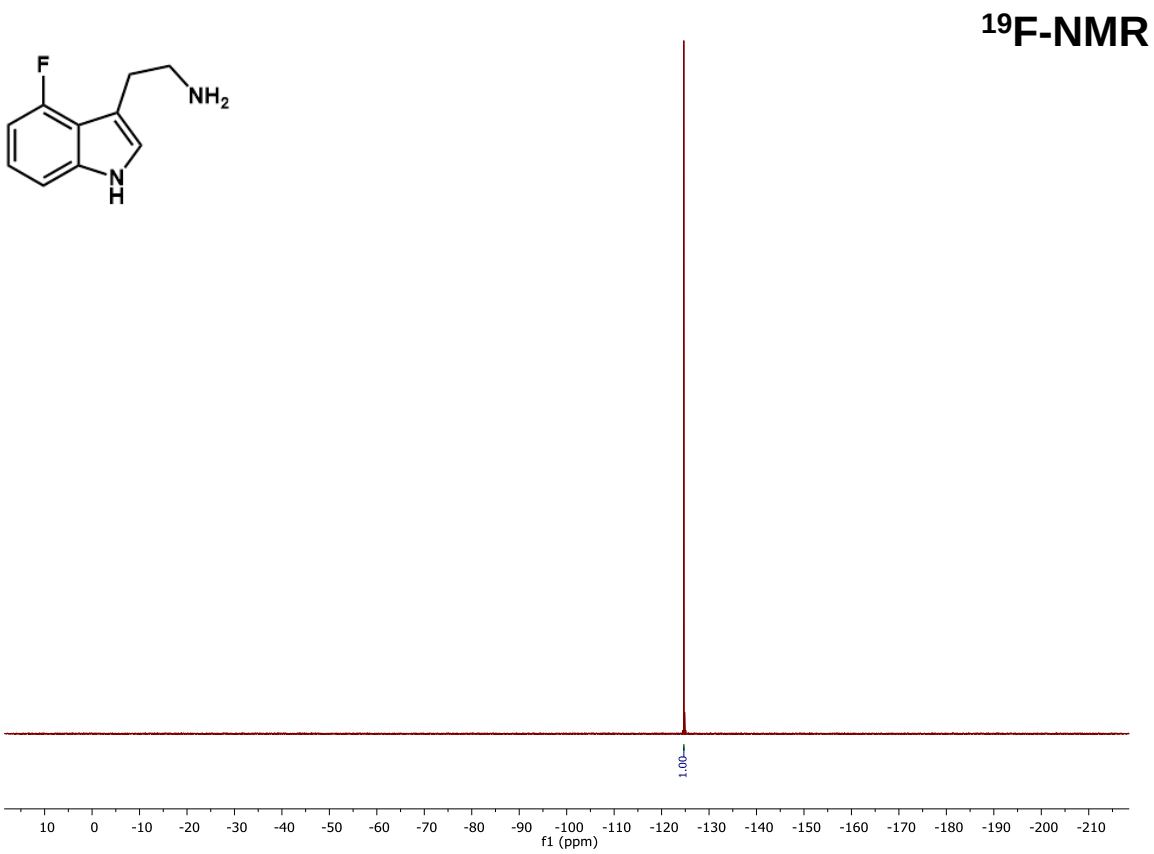


Figure 18. ^{1}H NMR, ^{13}C, and ^{19}F spectra of 2-(4-fluoro-1H-indol-3-yl)ethanamine (8).

2-(5-bromo-1H-indol-3-yl)ethanamine (9)

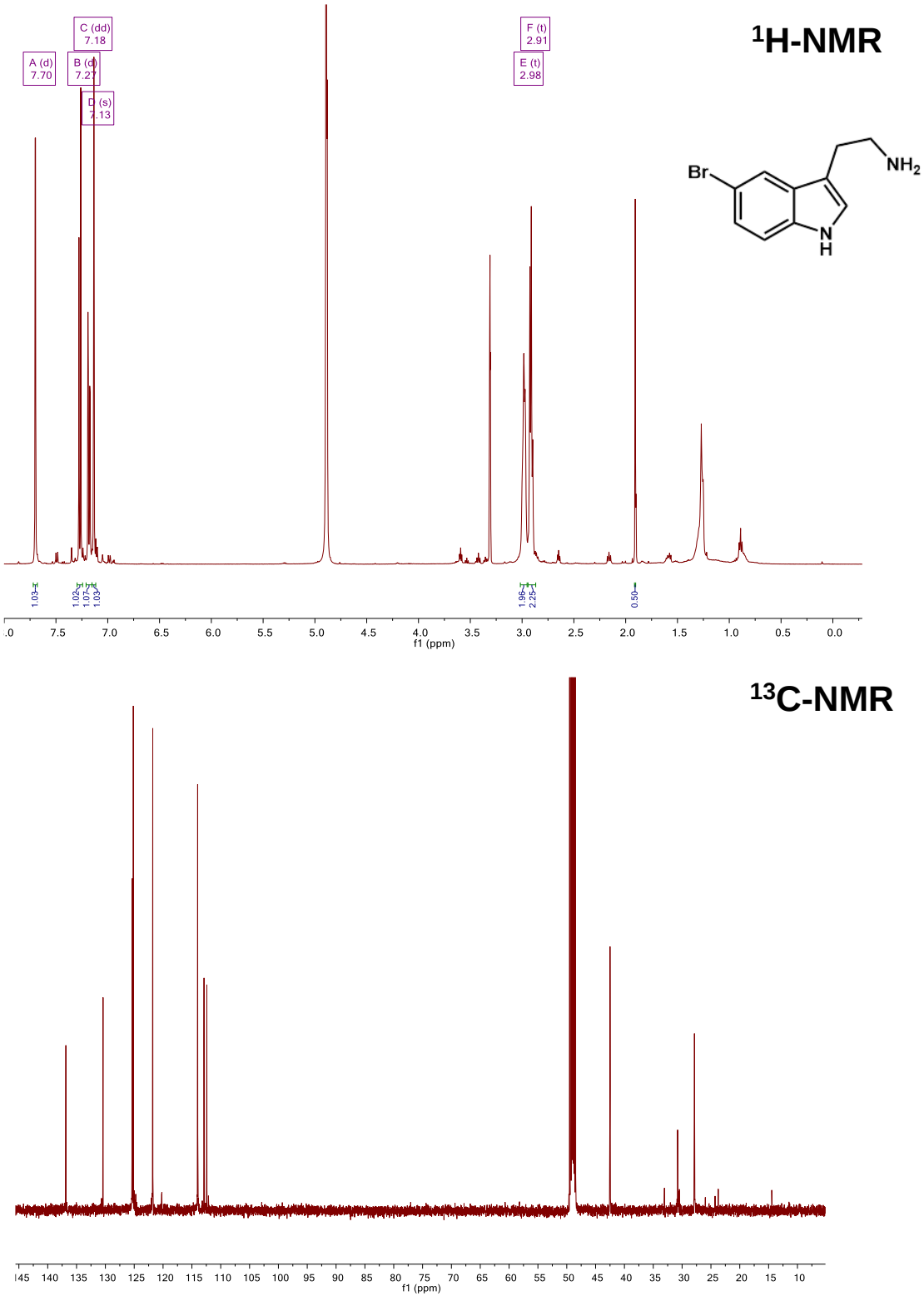


Figure 19. ^{1}H NMR and ^{13}C spectra of 2-(5-bromo-1H-indol-3-yl)ethanamine (9).

2-(5-chloro-1H-indol-3-yl)ethanamine (10)

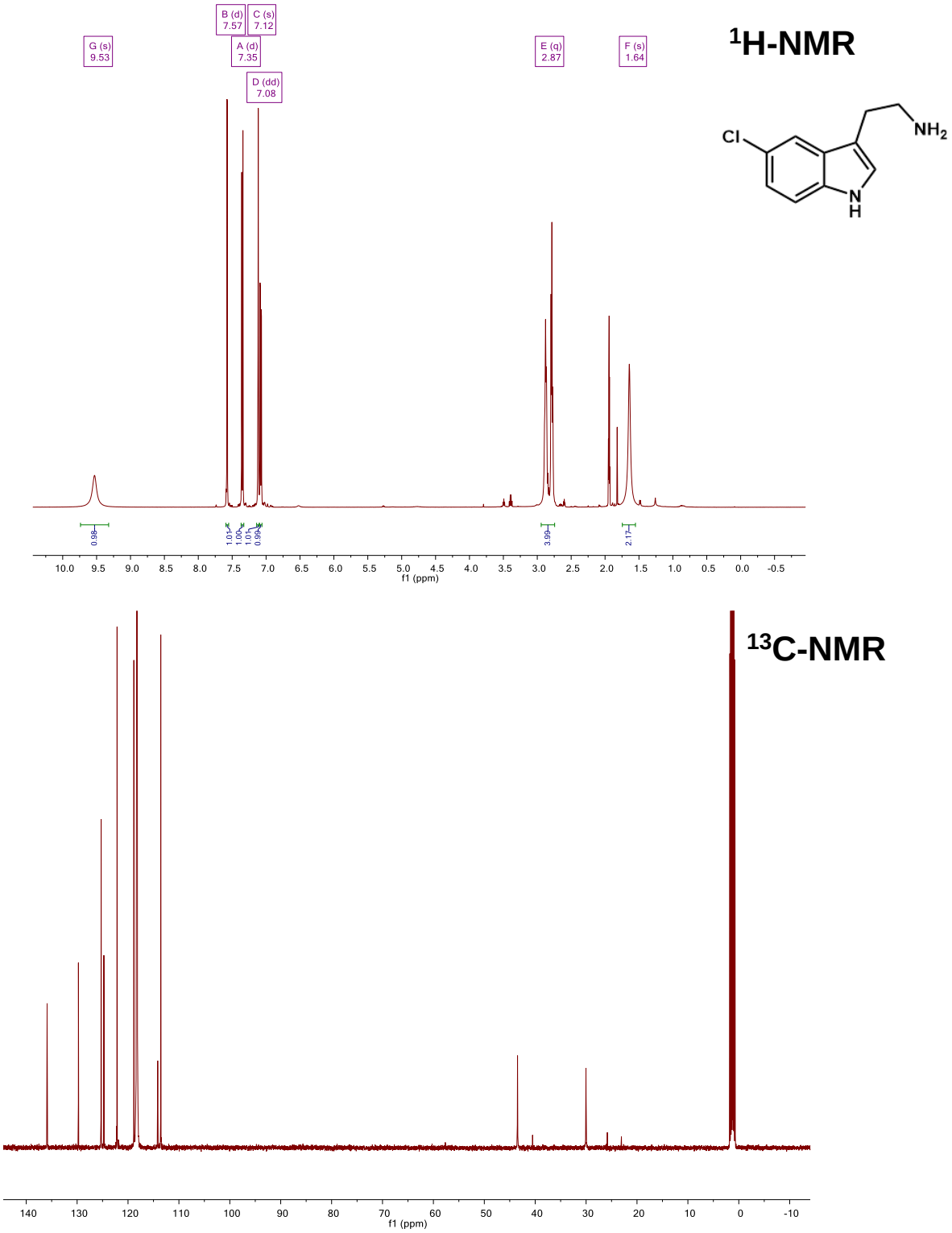


Figure 20. ^{1}H NMR and ^{13}C spectra of 2-(5-chloro-1H-indol-3-yl)ethanamine (10).

3-(2-aminoethyl)-1H-indol-5-ol (11)

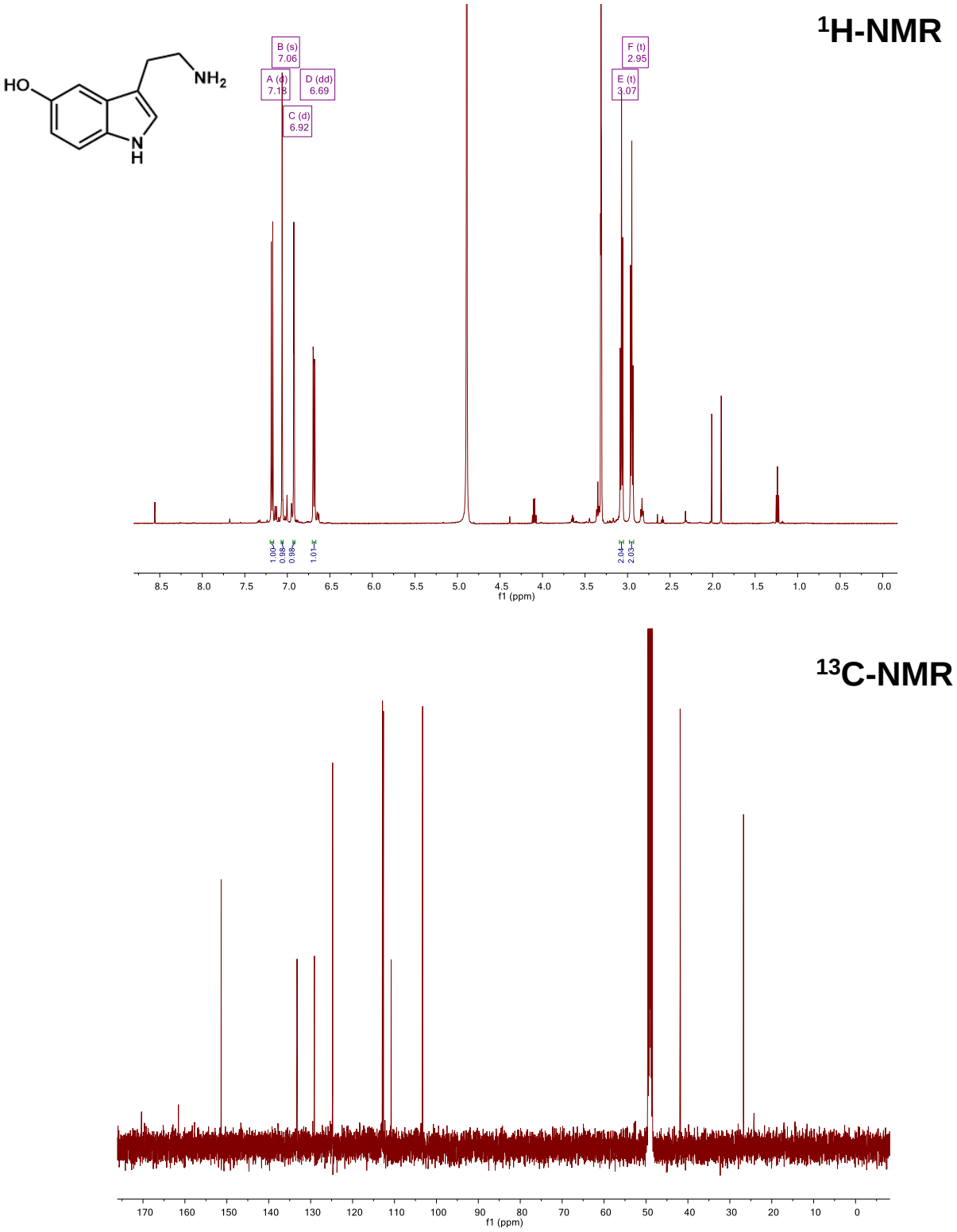


Figure 21. ^{1}H NMR and ^{13}C spectra of 3-(2-aminoethyl)-1H-indol-5-ol (11).

2-(5-methoxy-1H-indol-3-yl)ethanamine (12)

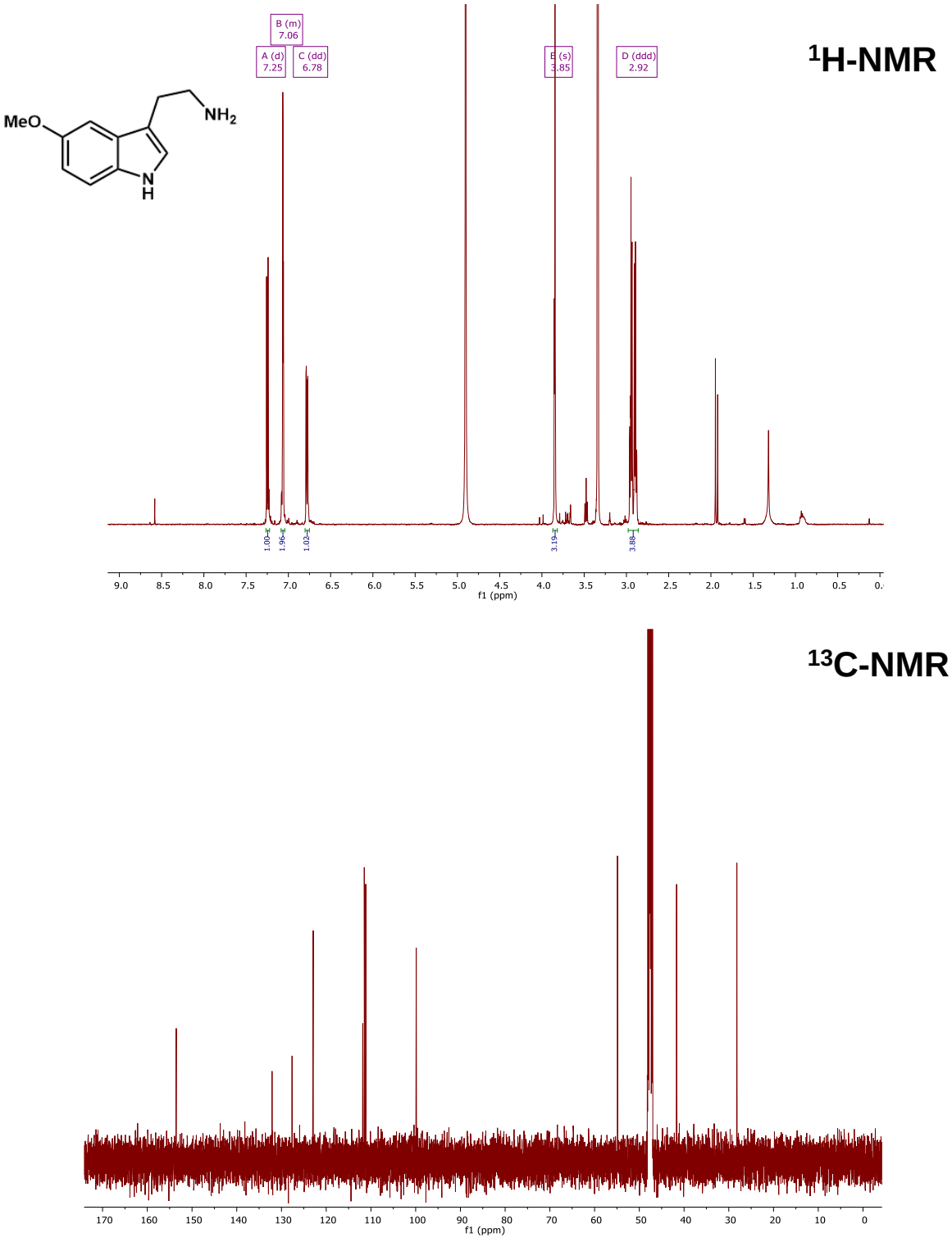
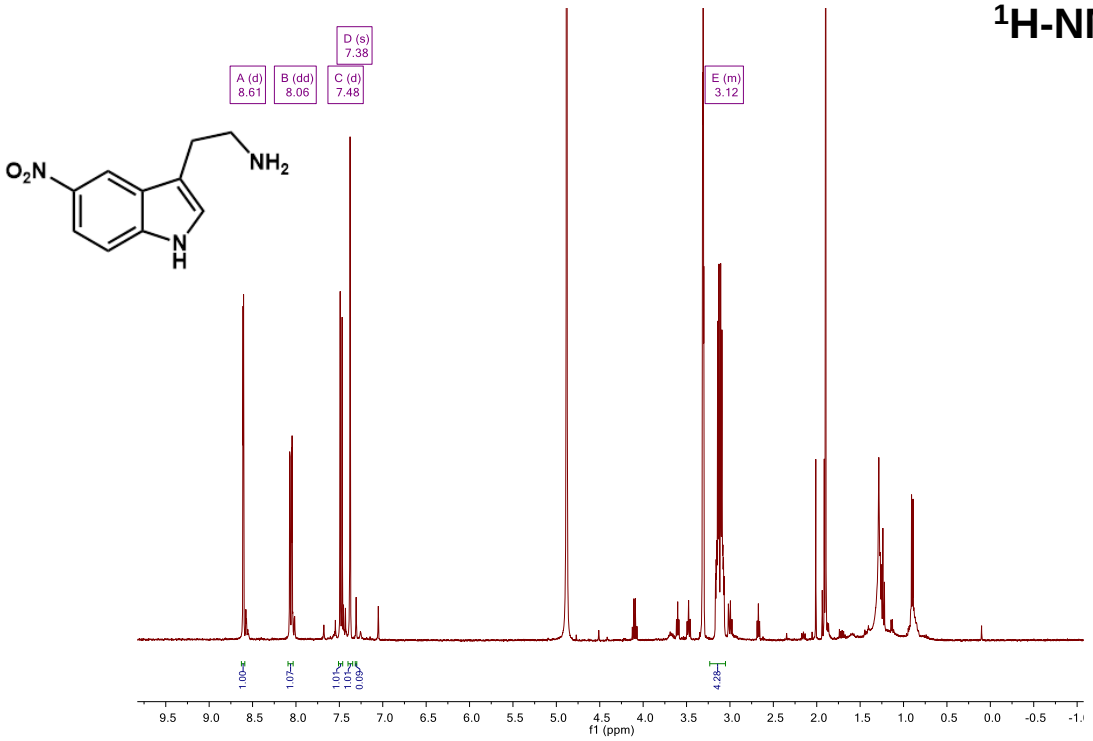


Figure 22. ^{1}H NMR and ^{13}C spectra of 2-(5-methoxy-1H-indol-3-yl)ethanamine (12).

2-(5-nitro-1H-indol-3-yl)ethanamine (13)

^{1}H-NMR



^{13}C-NMR

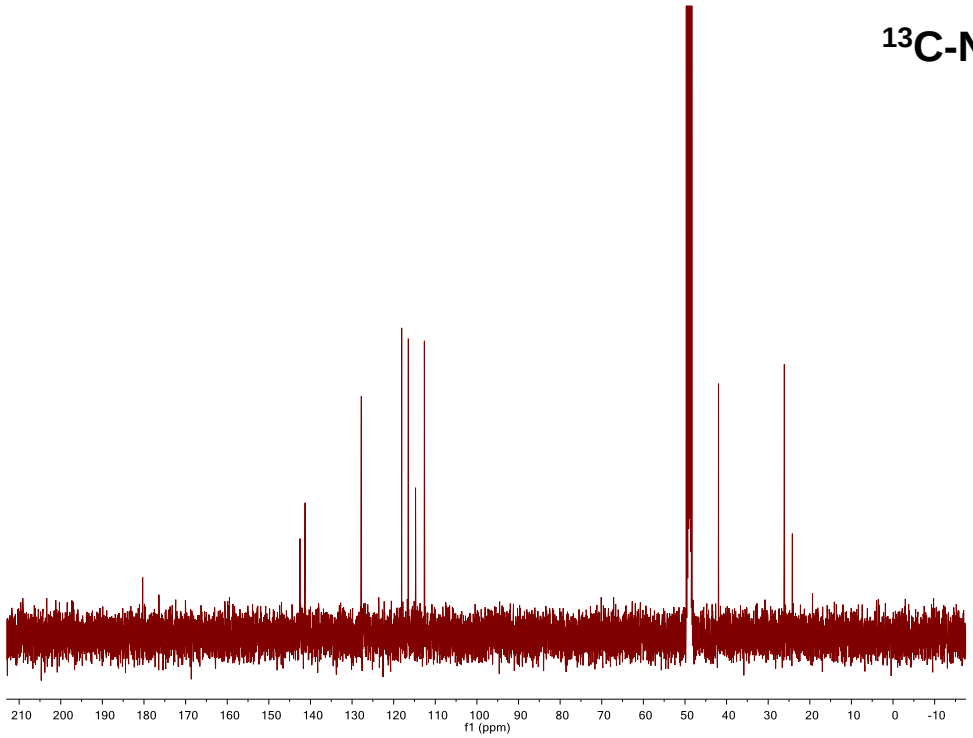
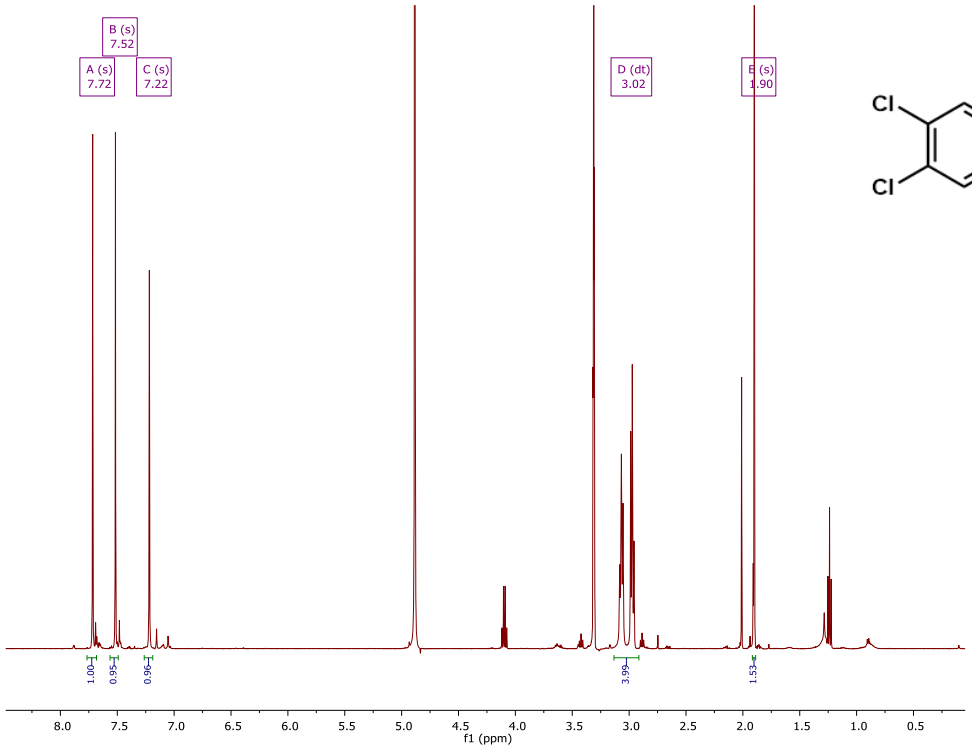


Figure 23. ^{1}H NMR and ^{13}C spectra of 2-(5-nitro-1H-indol-3-yl)ethanamine (13).

2-(5,6-dichloro-1H-indol-3-yl)ethanamine (14)

^{1}H-NMR



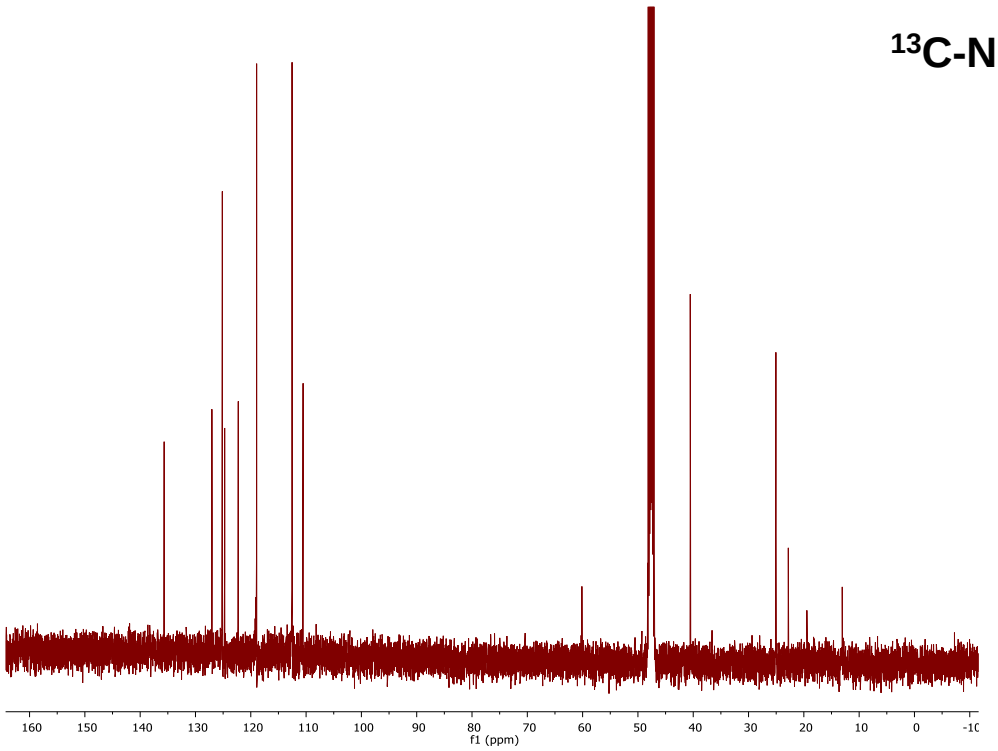


Figure 24. ^{1}H NMR and ^{13}C spectra of 2-(5,6-dichloro-1H-indol-3-yl)ethanamine (14).

2-(6-bromo-1H-indol-3-yl)ethanamine (15)

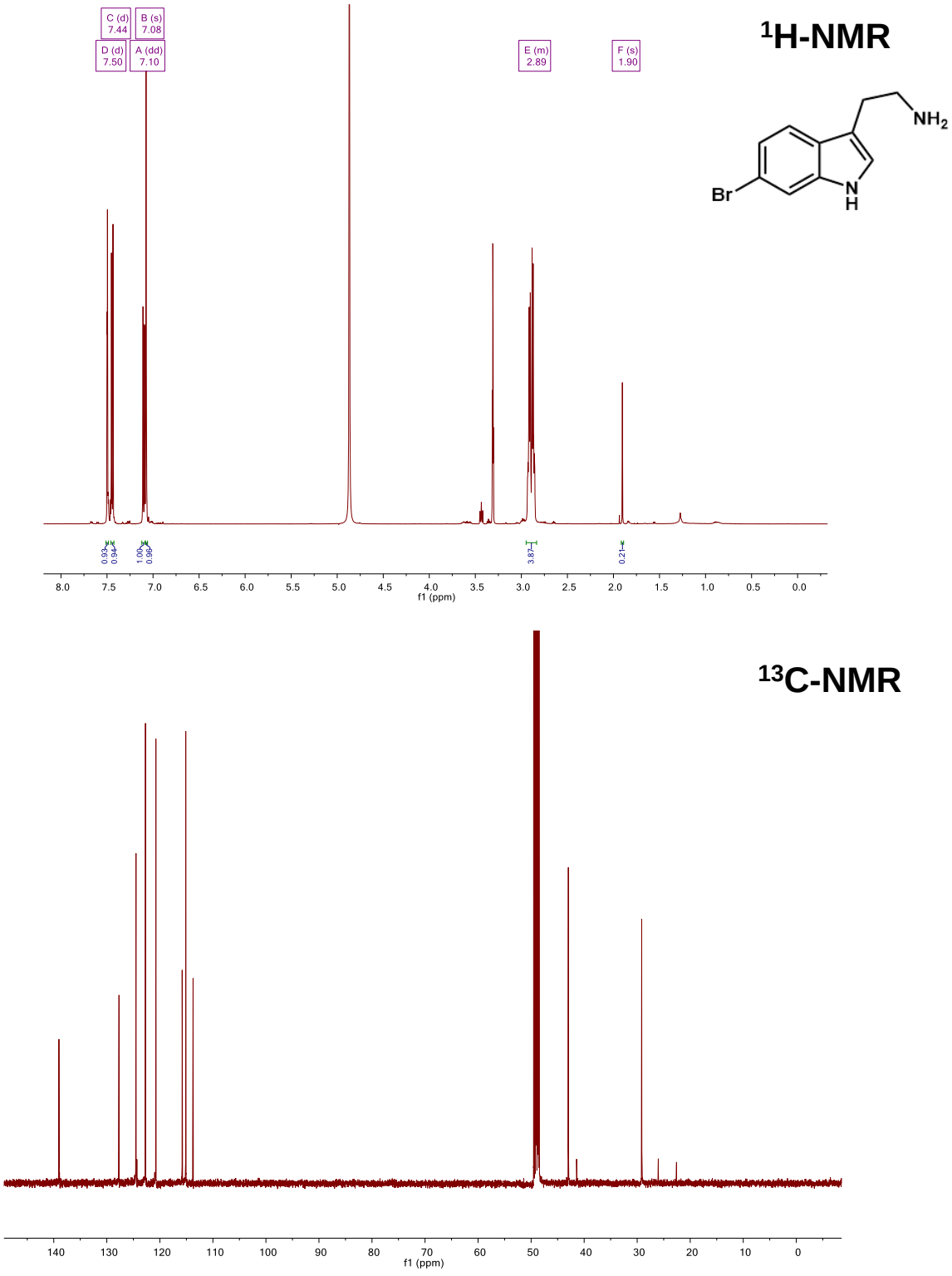


Figure 25. ^{1}H NMR and ^{13}C spectra of 2-(6-bromo-1H-indol-3-yl)ethanamine (15). 2-(6-chloro-1H-indol-3-yl)ethanamine (16)

^{1}H-NMR

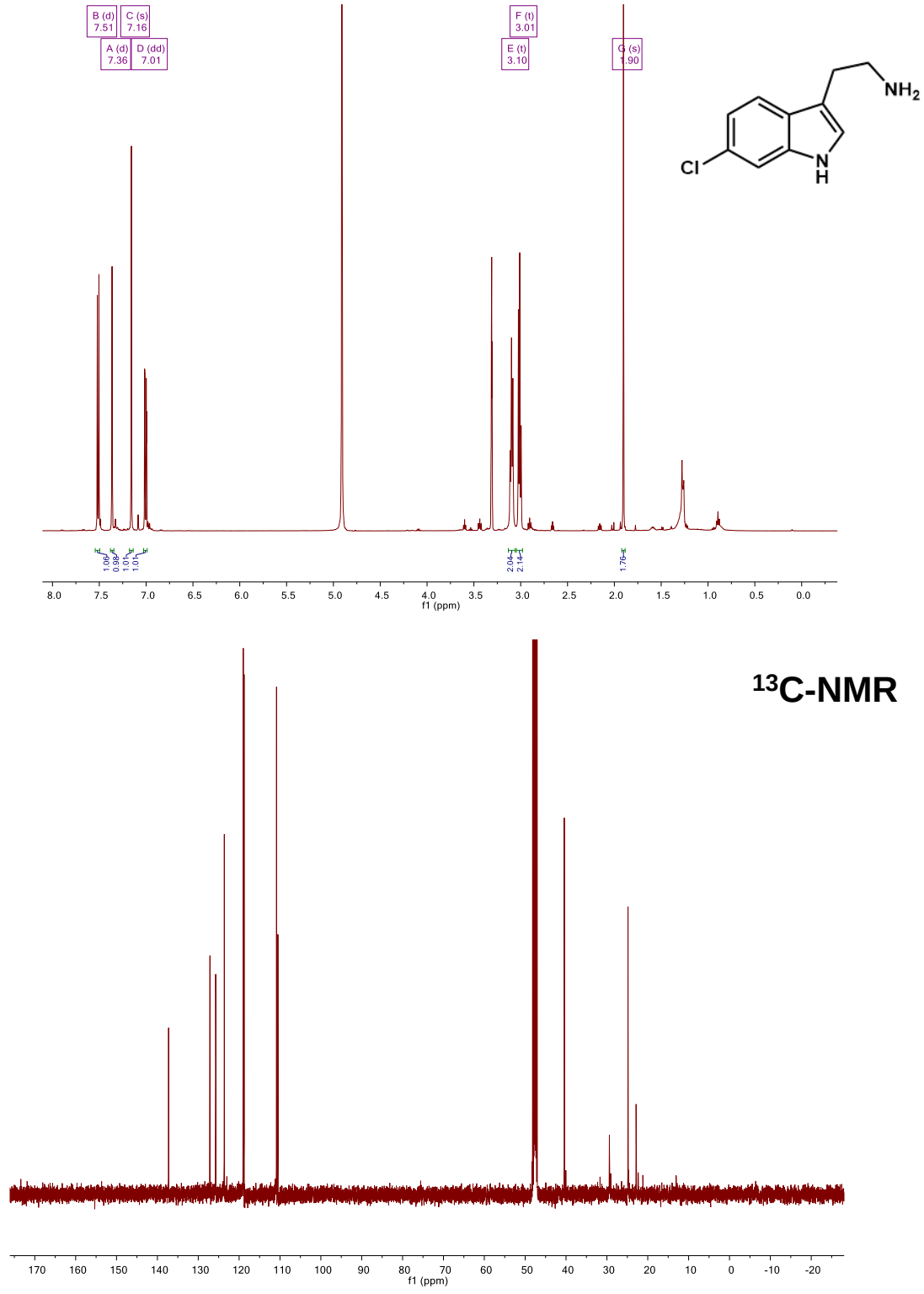


Figure 26. ^{1}H NMR and ^{13}C spectra of 2-(6-chloro-1H-indol-3-yl)ethanamine (16).

2-(6-nitro-1H-indol-3-yl)ethanamine (crude) (17)

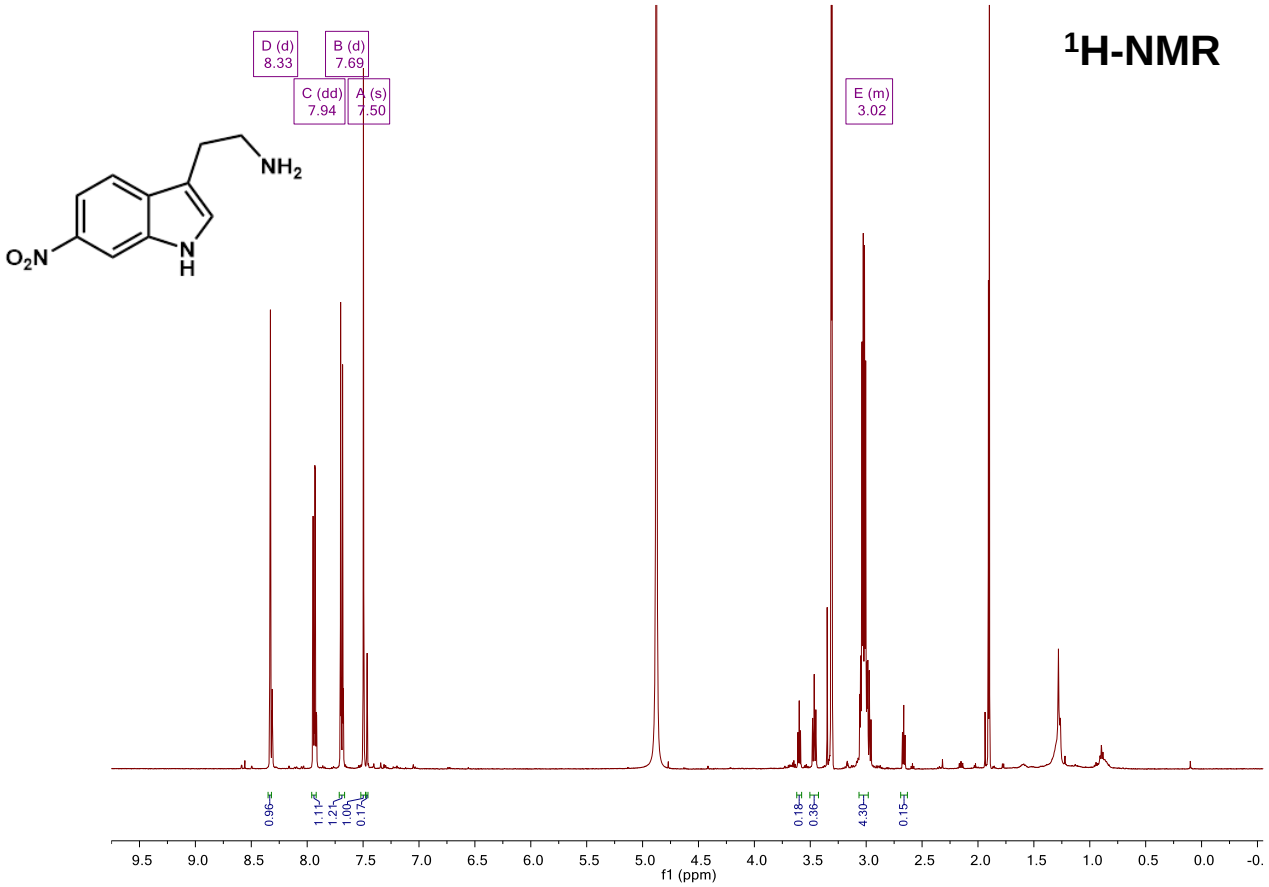


Figure 27. ^{1}H NMR spectrum of 2-(6-nitro-1H-indol-3-yl)ethanamine (crude) (17).

2-(6-nitro-1H-indol-3-yl)ethanamine (purified) (17)

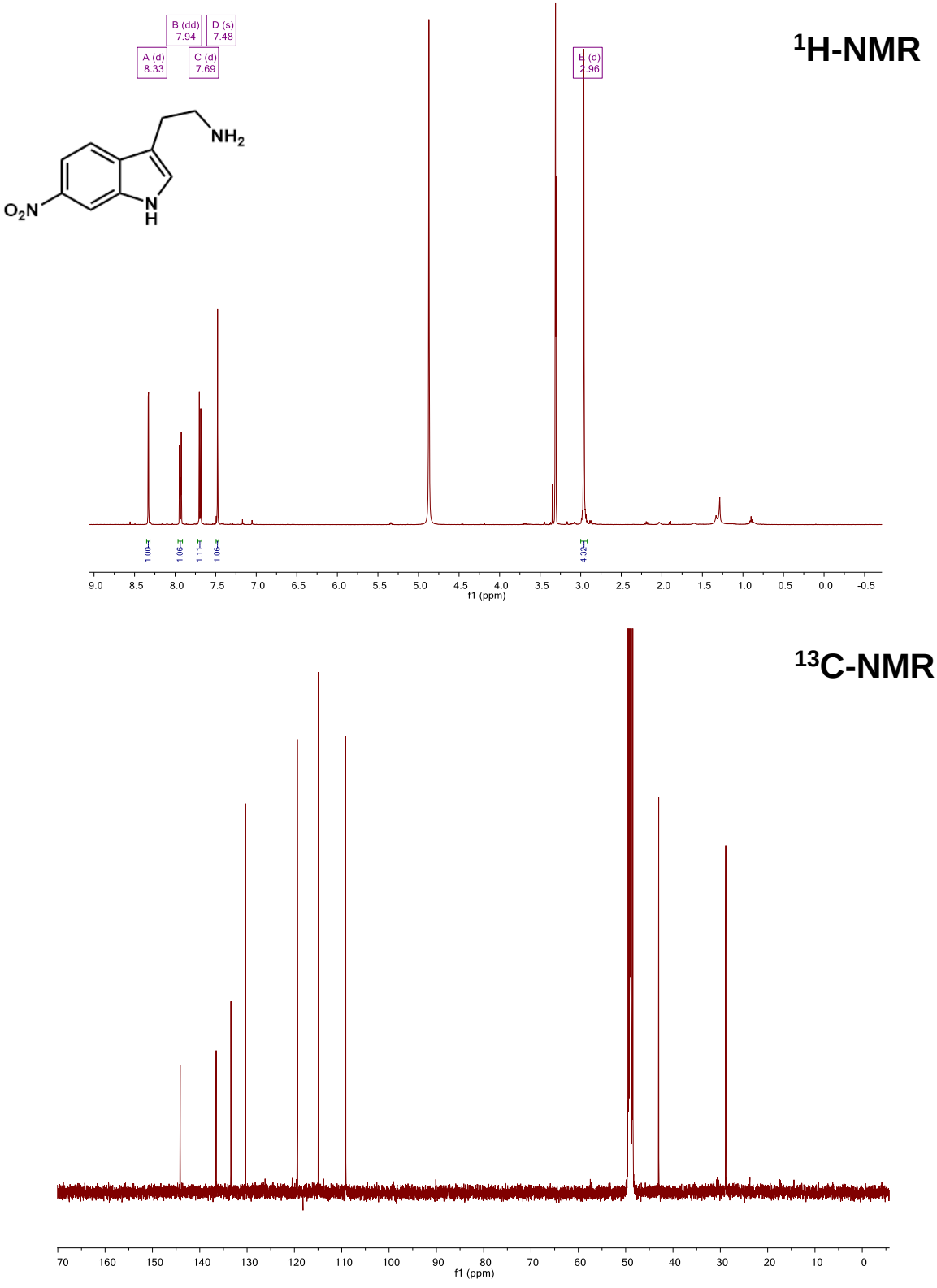


Figure 28. ^{1}H and ^{13}C NMR spectra of 2-(6-nitro-1H-indol-3-yl)ethanamine (purified) (17).

2-(7-bromo-1H-indol-3-yl)ethanamine (18)

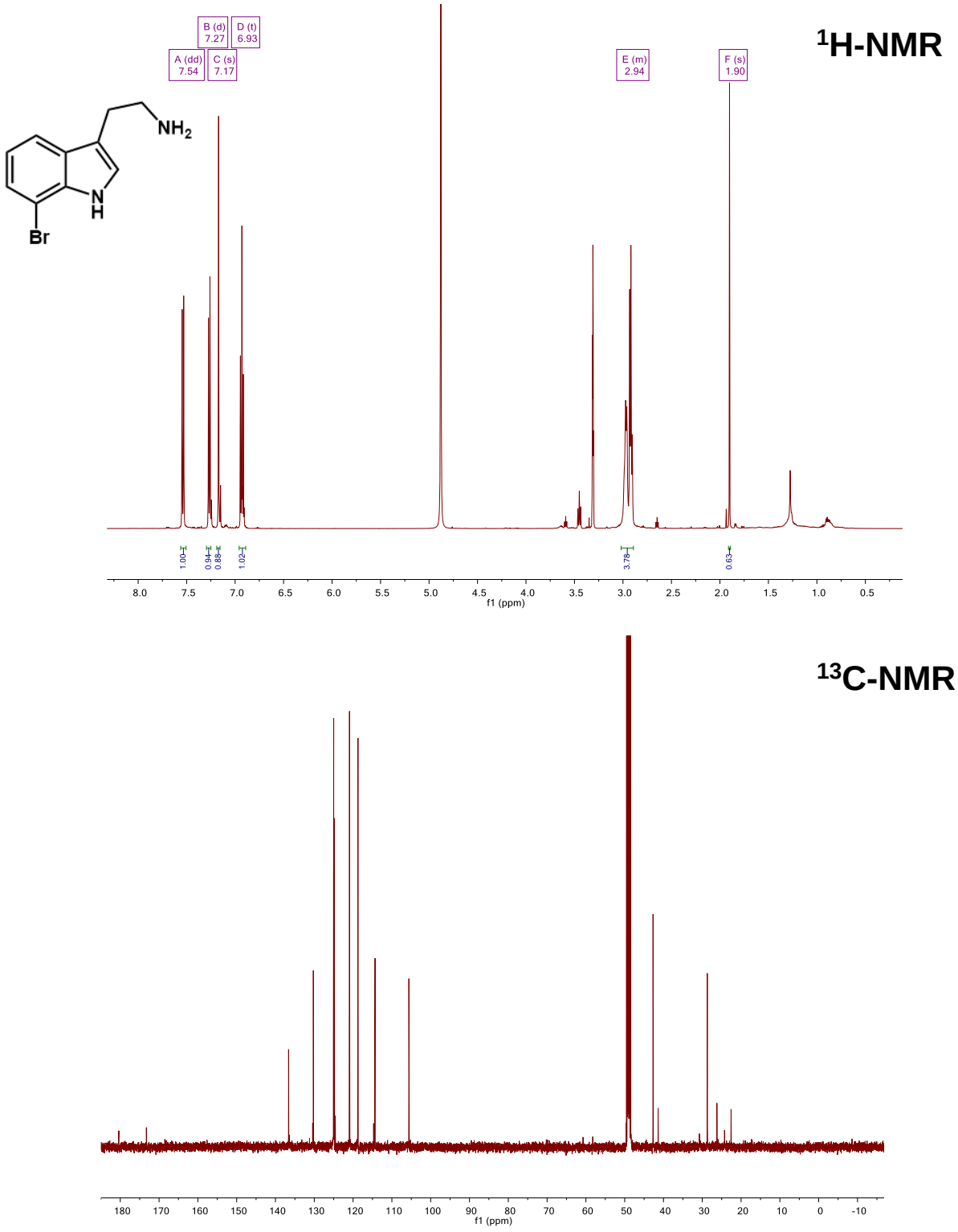


Figure 29. ^{1}H and ^{13}C NMR spectra of 2-(7-bromo-1H-indol-3-yl)ethanamine (18).

2-(7-chloro-1H-indol-3-yl)ethanamine (19)

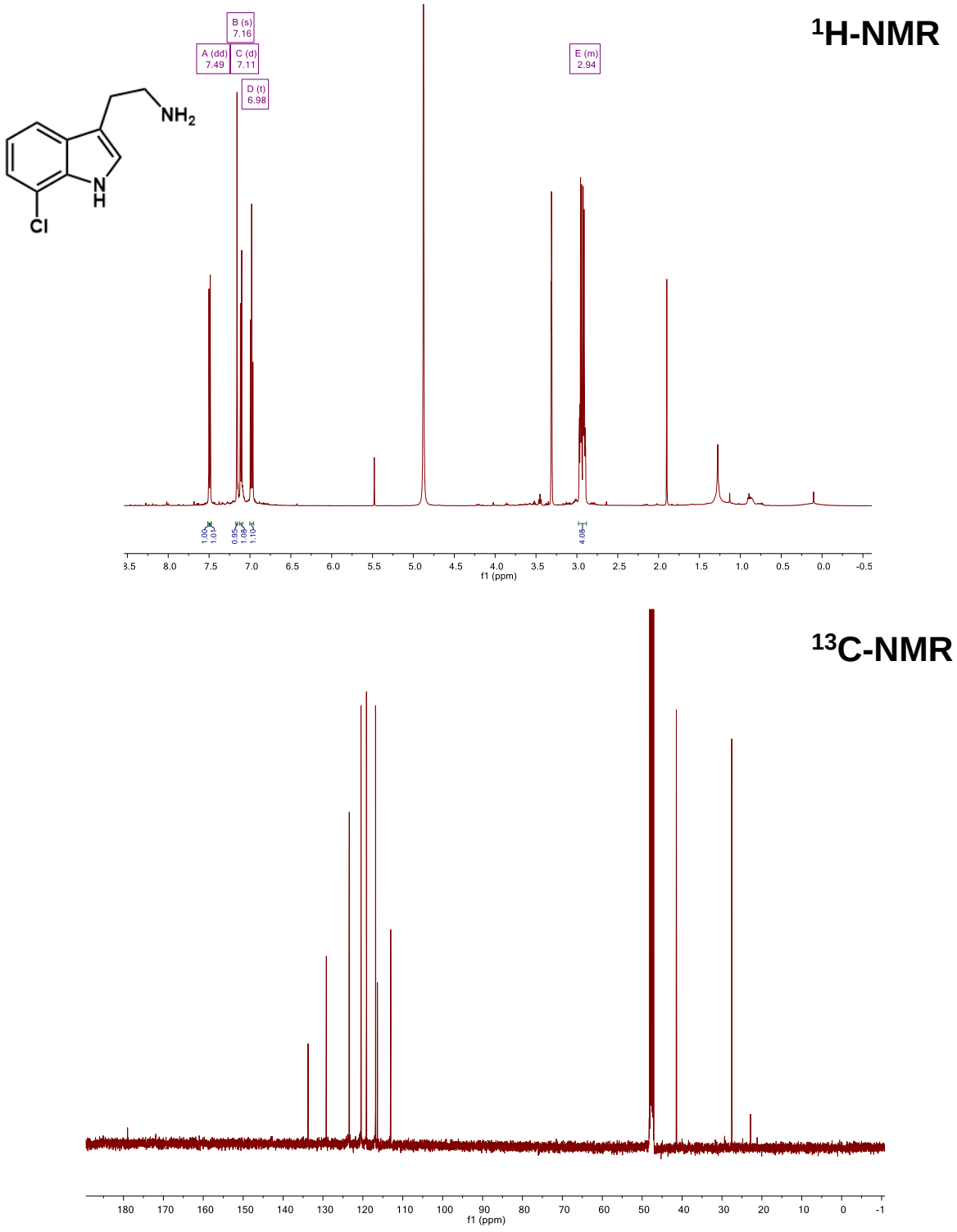


Figure 30. ^{1}H and ^{13}C NMR spectra of 2-(7-chloro-1H-indol-3-yl)ethanamine (19).

2-(7-cyano-1H-indol-3-yl)ethanamine (20)

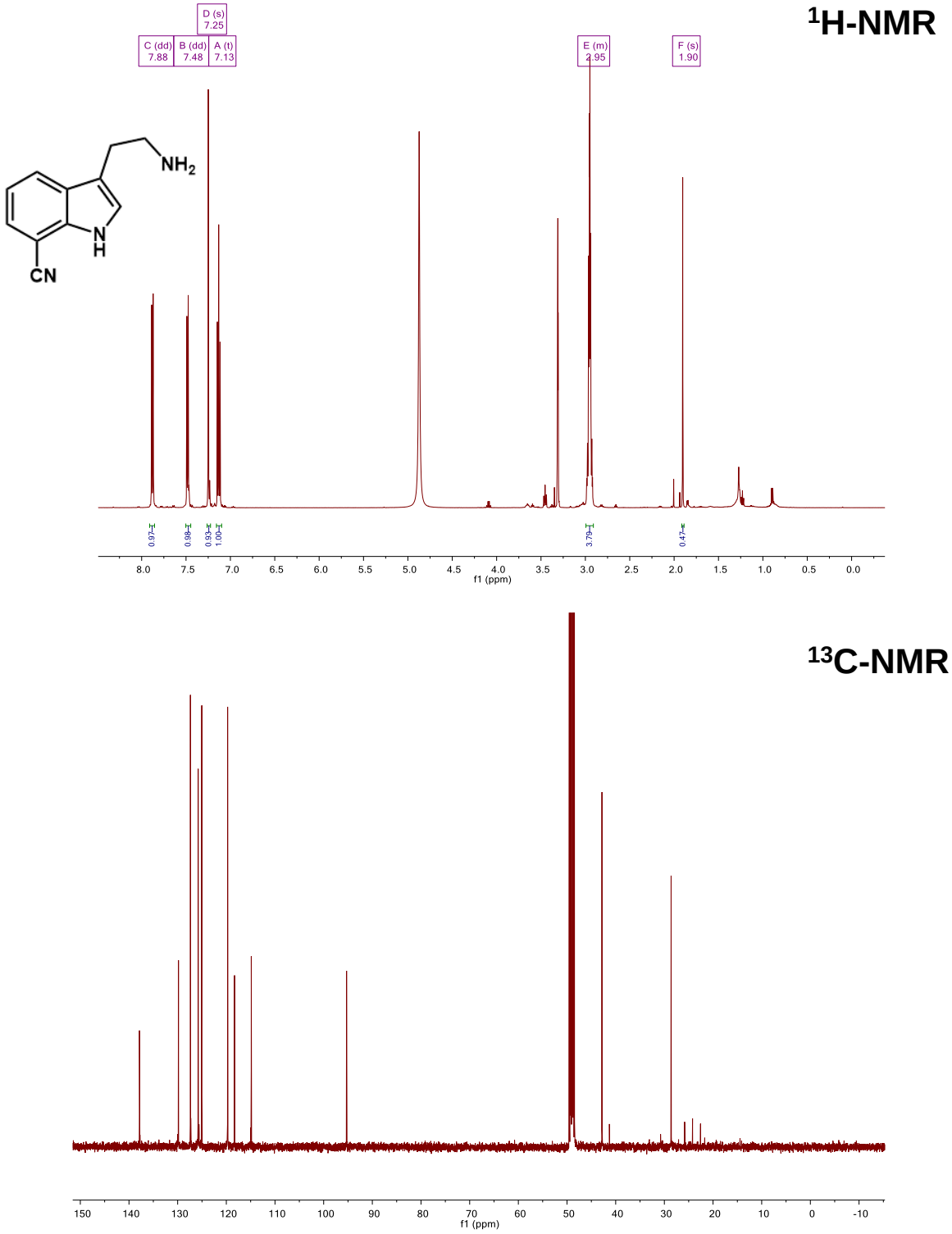
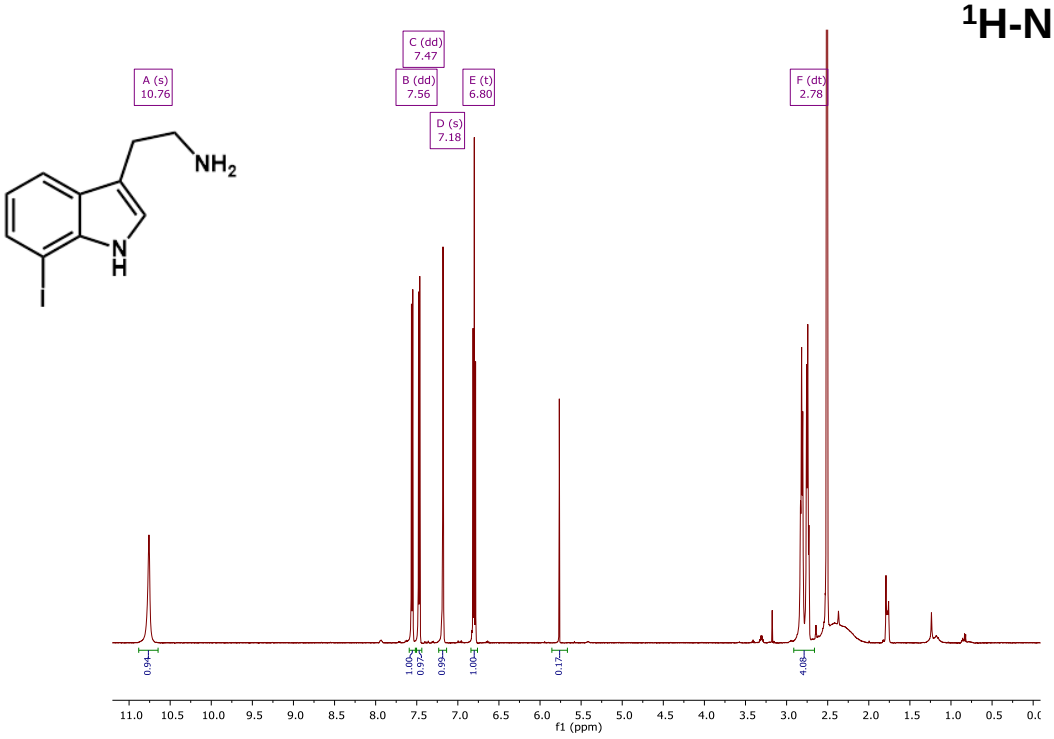


Figure 31. ^{1}H and ^{13}C NMR spectra of 2-(7-cyano-1H-indol-3-yl)ethanamine (20).

2-(7-iodo-1H-indol-3-yl)ethanamine (21)

^{1}H-NMR



^{13}C-NMR

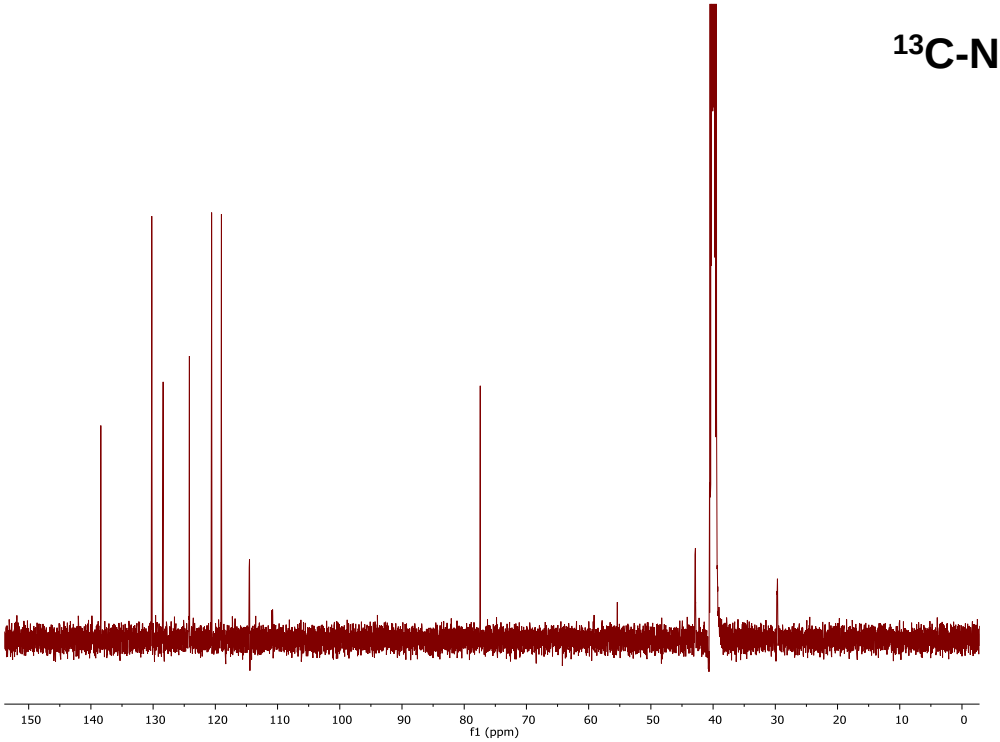


Figure 32. ^{1}H and ^{13}C NMR spectra of 2-(7-iodo-1H-indol-3-yl)ethanamine (21).

2-(7-nitro-1H-indol-3-yl)ethanamine (22)

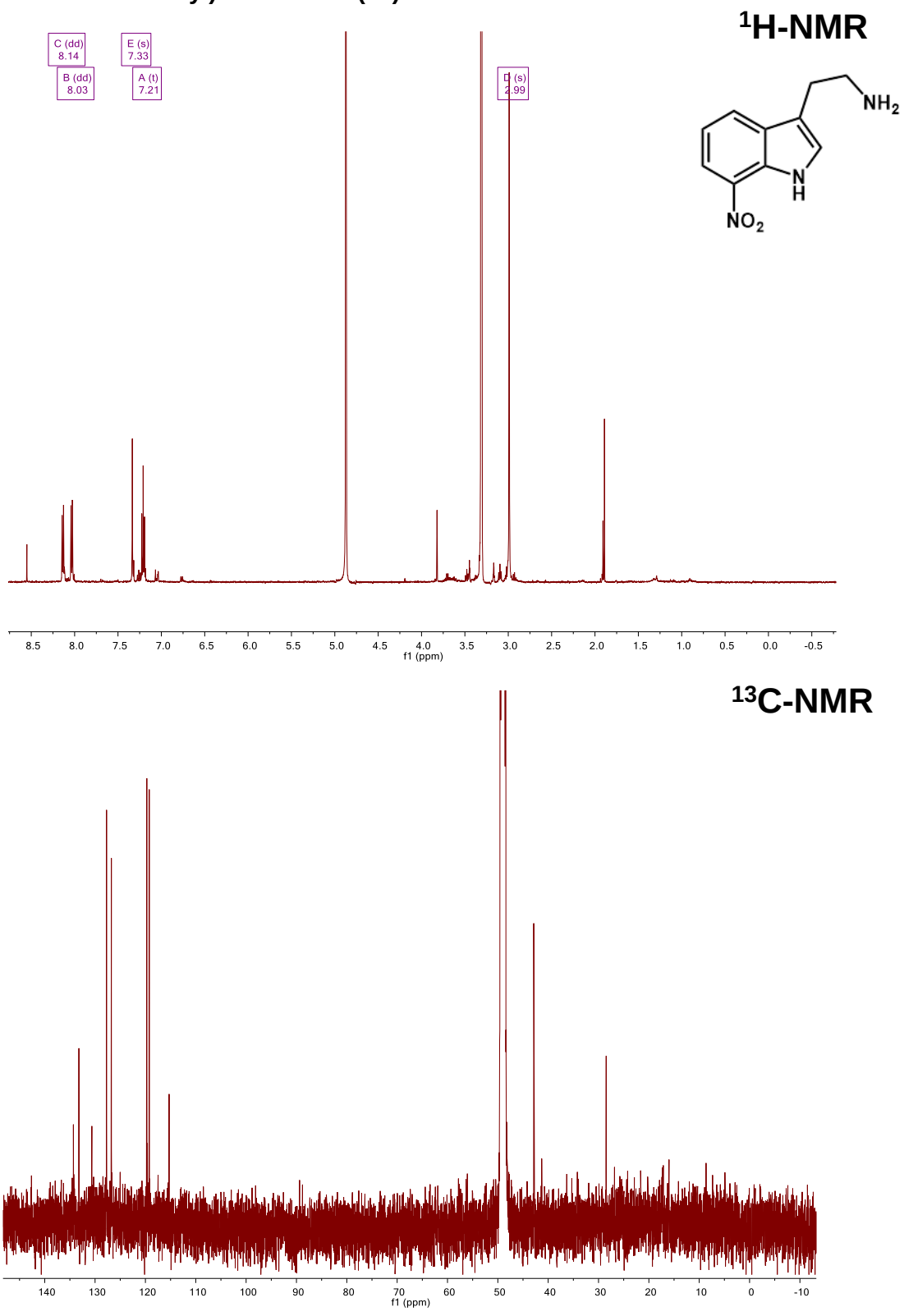
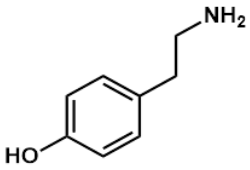


Figure 33. ^{1}H and ^{13}C NMR spectra 2-(7-nitro-1H-indol-3-yl)ethanamine (22).

4-(2-aminoethyl)phenol (23)



^{1}H-NMR

^{13}C-NMR

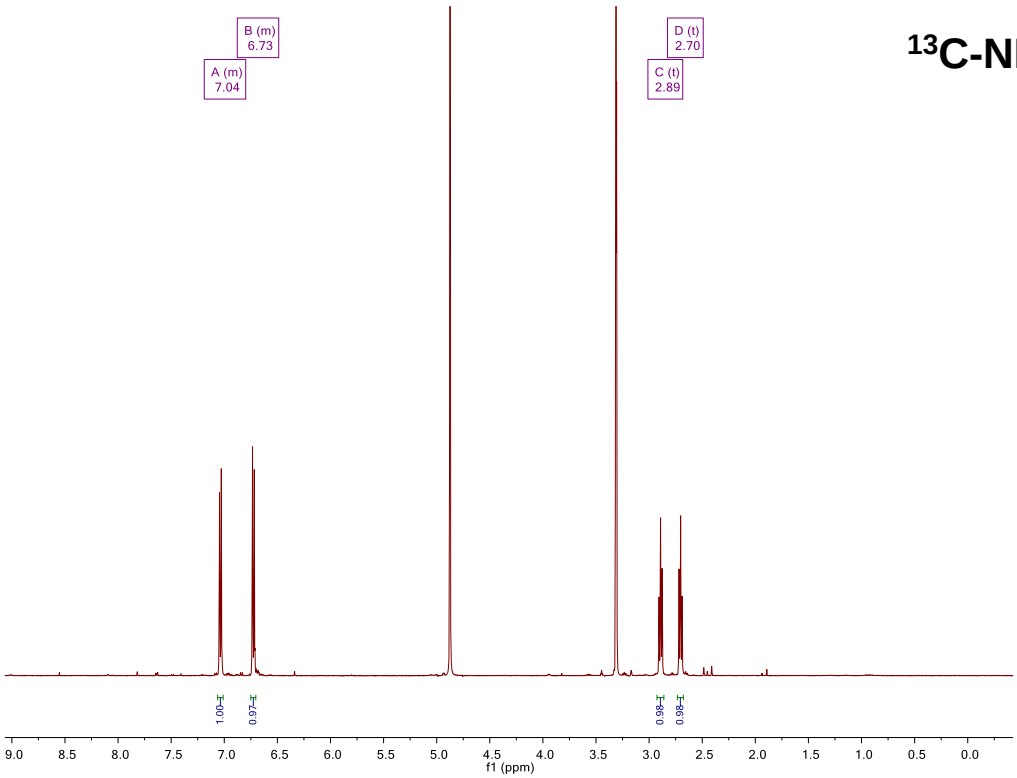


Figure 34. ^{1}H and ^{13}C NMR spectra 4-(2-aminoethyl)phenol (23).

2. 5. References

1. Fricke, J., Blei, F. & Hoffmeister, D. Enzymatic Synthesis of Psilocybin. *Angew. Chemie Int. Ed.* **56**, 12352–12355 (2017).
2. Chan, R. H. A. & Mendelson, J. E. Hallucinogens. in *The Effects of Drug Abuse on the Human Nervous System* vol. 101 533–552 (2013).
3. Ehrenworth, A. M. & Peralta-Yahya, P. Accelerating the semisynthesis of alkaloid-based drugs through metabolic engineering. *Nature Chemical Biology* vol. 13 249–258 (2017).
4. Righi, M. *et al.* Synthesis of tryptamine derivatives via a direct, one-pot reductive alkylation of indoles. *J. Org. Chem.* **77**, 6351–6357 (2012).
5. Olsen, E. K. *et al.* Marine AChE inhibitors isolated from Geodia barretti: Natural compounds and their synthetic analogs. *Org. Biomol. Chem.* **14**, 1629–1640 (2016).
6. Bartolucci, S., Mari, M., Bedini, A., Piersanti, G. & Spadoni, G. Iridium-Catalyzed Direct Synthesis of Tryptamine Derivatives from Indoles: Exploiting N - Protected β - Amino Alcohols as Alkylating Agents. *J. Org. Chem.* **80**, 3217–3222 (2015).
7. Muratore, M. E. *et al.* Enantioselective Brønsted acid-catalyzed N-acyliminium cyclization cascades. *J. Am. Chem. Soc.* **131**, 10796–10797 (2009).
8. Sheldon, R. A. & Woodley, J. M. Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* **118**, 801–838 (2018).
9. Chang, M. C. Y. & Keasling, J. D. Production of isoprenoid pharmaceuticals by engineered microbes. *Nat. Chem. Biol.* 1–8 (2006) doi:10.1038/nchembio836.
10. Latham, J., Brandenburger, E., Shepherd, S. A., Menon, B. R. K. & Micklefield, J. Development of Halogenase Enzymes for Use in Synthesis. *Chem. Rev.* **118**, 232–269 (2018).
11. Runguphan, W., Qu, X. & O'Connor, S. E. Integrating carbon–halogen bond formation into medicinal plant metabolism. *Nature* **468**, 461–464 (2010).
12. Kalb, D., Gressler, J. & Hoffmeister, D. Active-Site Engineering Expands the Substrate Profile of the Basidiomycete L-Tryptophan Decarboxylase CsTDC. *ChemBioChem* **17**, 132–136 (2016).
13. Smith, D. R. M. *et al.* The First One-Pot Synthesis of L-7-Iodotryptophan from 7-Iodoindole and Serine, and an Improved Synthesis of Other L-7-Halotryptophans. *Org. Lett.* **16**, 2622–2625 (2014).
14. Corr, M. J., Smith, D. R. M. & Goss, R. J. M. One-pot access to l-5,6-dihalotryptophans and l-alknyltryptophans using tryptophan synthase. *Tetrahedron* **72**, 7306–7310 (2016).
15. Phillips, R. S. Synthetic applications of tryptophan synthase. *Tetrahedron Asymmetry* vol. 15 2787–2792 (2004).

16. Francis, D., Winn, M., Latham, J., Greaney, M. F. & Micklefield, J. An Engineered Tryptophan Synthase Opens New Enzymatic Pathways to β-Methyltryptophan and Derivatives. *ChemBioChem* **18**, 382–386 (2017).

17. Buller, A. R., Van Roye, P., Murciano-Calles, J. & Arnold, F. H. Tryptophan Synthase Uses an Atypical Mechanism To Achieve Substrate Specificity. *Biochemistry* **55**, 7043–7046 (2016).

18. Roy, A. D., Grüschow, S., Cairns, N. & Goss, R. J. M. Gene expression enabling synthetic diversification of natural products: Chemogenetic generation of pacidamycin analogs. *J. Am. Chem. Soc.* **132**, 12243–12245 (2010).

19. Corr, M. J. *et al.* Sonogashira diversification of unprotected halotryptophans, halotryptophan containing tripeptides; and generation of a new to nature bromo-natural product and its diversification in water. *Chem. Sci.* **8**, 2039–2046 (2017).

20. Buller, A. R. *et al.* Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. *Proc. Natl. Acad. Sci.* **112**, 14599–14604 (2015).

21. Herger, M. *et al.* Synthesis of β-Branched Tryptophan Analogues Using an Engineered Subunit of Tryptophan Synthase. *J. Am. Chem. Soc.* **138**, 8388–8391 (2016).

22. Murciano-Calles, J., Romney, D. K., Brinkmann-Chen, S., Buller, A. R. & Arnold, F. H. A Panel of TrpB Biocatalysts Derived from Tryptophan Synthase through the Transfer of Mutations that Mimic Allosteric Activation. *Angew. Chemie - Int. Ed.* **55**, 11577–11581 (2016).

23. 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. *J. Am. Chem. Soc.* **139**, 10769–10776 (2017).

24. Boville, C. E., Romney, D. K., Almhjell, P. J., Sieben, M. & Arnold, F. H. Improved Synthesis of 4-Cyanotryptophan and Other Tryptophan Analogues in Aqueous Solvent Using Variants of TrpB from Thermotoga maritima. *J. Org. Chem.* **83**, 7447–7452 (2018).

25. Blei, F., Baldeweg, F., Fricke, J. & Hoffmeister, D. Biocatalytic Production of Psilocybin and Derivatives in Tryptophan Synthase-Enhanced Reactions. *Chem. - A Eur. J.* **24**, 10028–10031 (2018).

26. Williams, B. B. *et al.* Discovery and characterization of gut microbiota decarboxylases that can produce the neurotransmitter tryptamine. *Cell Host Microbe* **16**, 495–503 (2014).

27. Woo, S., Jong, H. & Shin, S. Aromatic L - amino acid decarboxylases : mechanistic features and microbial applications. *Appl. Microbiol. Biotechnol.* 4445–4458 (2022) doi:10.1007/s00253-022-12028-4.

28. Bertoldi, M., Moore, P. S., Maras, B., Dominici, P. & Voltattorni, C. B. Mechanism-based Inactivation of Dopa Decarboxylase by Serotonin *. *J. Biol. Chem.* **271**, 23954–23959 (1996).

29. Nakazawa, H., Kumagai, H. & Yamada, H. Aromatic L-Amino Acid Decarboxylase from Micrococcus percitreus Purification , Crystallization and Properties zation and properties of the aromatic L-amino described that S-adenosyl-L-methionine. *Agric. Bio. Chem.* **45**, 2543–2552 (1981).

30. Glenn, W. S., Nims, E. & O'Connor, S. E. Reengineering a tryptophan halogenase to preferentially chlorinate a direct alkaloid precursor. *J. Am. Chem. Soc.* **133**, 19346–19349 (2011).

31. Buller, A. R. *et al.* Directed evolution mimics allosteric activation by stepwise tuning of the conformational ensemble. *J. Am. Chem. Soc.* **140**, 7256–7266 (2018).

32. Almhjell, P. J., Boville, C. E. & Arnold, F. H. Engineering enzymes for noncanonical amino acid synthesis. *Chem. Soc. Rev.* **47**, 8980–8997 (2018).

33. Dunathan, H. C. Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **55**, 712–6 (1966).

34. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

35. Bartolucci, S., Mari, M., Gregorio, G. Di & Piersanti, G. Observations concerning the synthesis of tryptamine homologues and branched tryptamine derivatives via the borrowing hydrogen process : synthesis of psilocin , bufotenin , and serotonin. *Tetrahedron* **72**, 2233–2238 (2016).

Chapter 3

Development of substrate multiplexed screening technology and its application to engineering of a tryptophan decarboxylase

Content in this chapter is adapted from the following published work:

McDonald, A.D.*; Higgins, P.M.*; Buller, A.R. “Substrate multiplexed protein engineering facilitates promiscuous biocatalytic synthesis.” (*Accepted at Nat. Comm.).*

As a stalwart collaborator in both thinking about substrate multiplexed screening and experimental design, I want to acknowledge Peyton M. Higgins as an invaluable contributor to this work.

Chapter 3: Development of substrate multiplexed screening technology and its application to engineering of a tryptophan decarboxylase

3. 1. Introduction

Biocatalysts are prized for their ability to perform well-defined transformations. Unfortunately, the use of enzymes in chemical synthesis is often hampered by their small or poorly understood substrate scopes.[1] Using traditional protein engineering approaches, activity can readily be increased on a model compound.[2–4] Most engineering advances have centered on smart library design[5–8] and screening speed,[9–12] and engineering campaigns using diverse approaches have, at times, led to promiscuous catalysts.[2,12–14] However, the substrate scopes of intermediates along evolutionary lineages are typically unknown. Consequently, when protein engineering does yield a catalyst with a limited scope, evolution is tediously repeated to generate activity with additional substrates.[8,15–17] Screening for activity on a single substrate necessarily overlooks mutations that are activating for substrates not included in the screen and can inadvertently lead to enzymes with high activity but narrow substrate scopes.[16–18] Methods that directly assess catalyst promiscuity would overcome this recurring pitfall and enable the development and application of biocatalysts for organic synthesis, both as single enzymes and in multi-enzyme cascade settings.

An alternative to single-substrate screening is to obtain information on catalyst promiscuity by screening with multiple substrates, either iteratively or in competition. Previously, these approaches have gone by various names including fingerprinting, substrate cocktail, multi-substrate, or multiplexed assays.[18–22] To avoid confusion as to whether substrates were assayed in separate parallel reactions or in competition, we use the term substrate multiplexed screening (SUMS) to specifically refer to methods where substrates are in direct competition (Fig 1a). A classic application of SUMS has been for characterization of native enzyme specificity, defined as the extent to which an enzyme distinguishes between substrates.[21,23,24] In pioneering work, the Reymond group showed how careful assay design to maintain initial

velocity conditions can enable the high-throughput characterization of lipase and esterase substrate specificities without measuring kinetic parameters for each individual substrate.[19] Similarly, the Kries group has advanced the use of a multiplexed assay the substrate specificity of the adenylation domains of non-ribosomal peptide synthetases.[25] In each of these cases, enzymologists take exquisite care to maintain initial velocity conditions, else the connection between product abundance and the underlying kinetic parameters is lost.

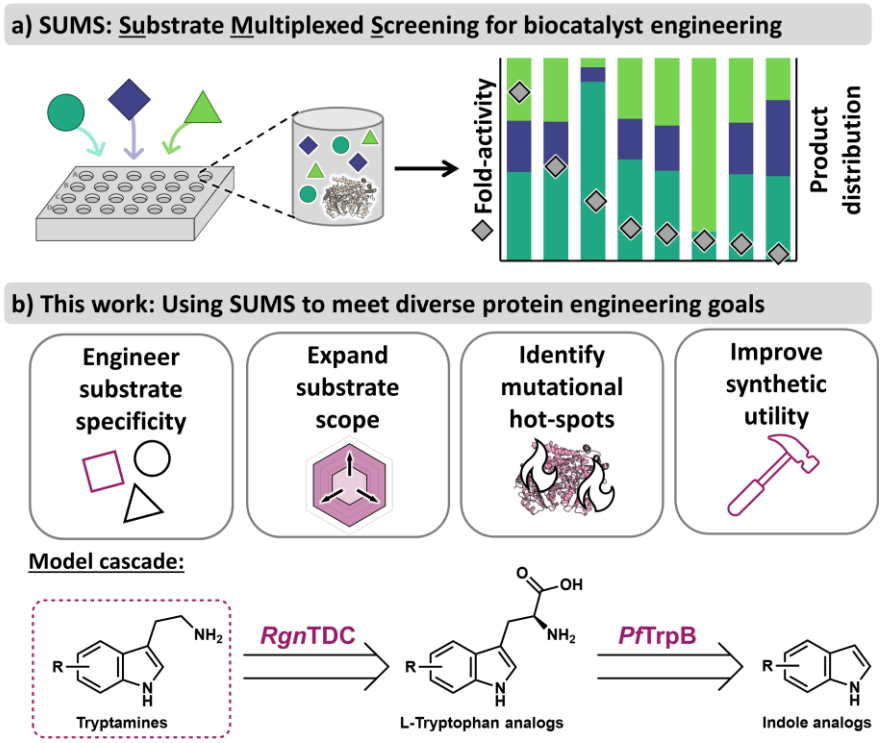


Figure 1. Substrate multiplexed screening is a powerful protein engineering methodology.

Within the broader bioengineering community, there are some scenarios where SUMS has long been the default mode of operation. Specifically, all screens and selections that take place under *in vivo* conditions, or with simulated mixtures of cytosolic metabolites, necessarily incorporate substrate competition. By monitoring formation of multiple products, some groups have directly read out on catalyst specificity in order to avoid variants with undesirable off-target

activity.[21] For example, the Williams group engineered the acyltransferase domain of a modular polyketide synthase to alter the substrate specificity away from the native substrate.[26] Alternatively, Weeks and Wells intentionally engineered more promiscuous subtiligase variants by screening with a peptide mixture that simulates the proteome of *E. coli* and detecting product formation using mass spectrometry (MS).[27] Critically, in these and other cases,[28–32] each enzyme's envisioned use is in an environment where substrate competition will occur.

Direct measurement of substrate specificity is also a recurring, albeit a less prominent, feature of protein engineering for organic synthesis. Although not commonly described as such, all screens with racemic substrates are, as defined here, substrate multiplexed screens. Resolution of enantiomeric products requires chiral chromatography or subsequent product functionalization, which are often too cumbersome for high throughput (HT)-screening. A standard approach, then, is to first screen variants for total activity on a racemic mixture and enantiospecificity is only measured for highly active variants.[33] To circumvent this limitation, researchers have made clever use of pseudo-enantiomers to directly determine the enantioselectivity of variants in higher throughput formats.[34,35] These screening efforts focused on stereospecificity are conceptually related to, but distinct from, the more general case where researchers might screen enzymes on non-stereoisomeric substrates with the intention of monitoring changes in scope.

To date, examples of SUMS methods to expand the substrate scope of an enzyme remain comparatively rare. Jakoblinnert *et al.* screened carbonyl reductase variants using mixtures of three to four substrates and identified a single mutation that improved activity on four previously poor substrates. In this case, NADH depletion was measured and changes to substrate specificity were established in subsequent single substrate assays.[36] Junker *et al.* successfully applied a two-substrate SUMS method to simultaneously evaluate aldolase variants for changes in both activity and specificity with aldehyde and ketone-derived

nucleophiles.[37] Recently, Knorrscheidt *et al*. demonstrated how a SUMS method using MISER-GCMS and a cocktail of three substrates could successfully identify mutations that altered the activity, specificity, and regioselectivity of an unspecific peroxygenase.[38] Given the ubiquity of SUMS in metabolic engineering and chemical biology, the success of these examples for biocatalysis, and the increasing interest in engineering enzymes for organic synthesis, it is striking that SUMS remains a specialized and uncommon approach in this field.

The transition from using a single substrate to screening with multiple substrates introduces significant and poorly understood complexities that we hypothesize have hindered wider adoption of SUMS. The kinetics of substrate competition impact assay results, as both substrates and products may act as inhibitors of the enzyme being engineered.[39] Because competing substrates exit the initial velocity regime at different rates, quantitative specificity information may be lost. Relatedly, it is not immediately apparent how well measurements in a multiplexed setting will correlate with synthetic utility for single substrate reactions. Although data from SUMS are intrinsically richer and might be leveraged in unique ways to guide engineering, data analysis and presentation become more challenging as the number of potential products increases.[37,38]

Here, we present a systematic exploration of the advantages and disadvantages of using SUMS to engineer enzymes for organic synthesis. We chose to investigate the L-tryptophan (Trp) decarboxylase from *Ruminococcus gnavus* (*Rgn*TDC) as a model system for this work. *Rgn*TDC is an enzyme that natively catalyzes the decarboxylation of Trp to form tryptamine (Fig 1b).[40–42] *Rgn*TDC is an exceptional decarboxylase with many Trp analogs but we previously demonstrated that this enzyme struggles with the highly bioactive 4- and 5-substituted substrates. We directly compare the effectiveness of SUMS and single substrate screening for site-saturation mutagenesis and show that SUMS reveals counter-intuitive trends in substrate promiscuity. We then utilize engineered *Rgn*TDC variants along with a tryptophan

synthase to synthesize substituted tryptamines from L-serine (Ser) and indole analogs. Together, these efforts show how SUMS can be used to immediately assess the substrate scope of variant libraries, form new hypotheses about enzyme function, and lead to synthetically useful enzymes for the efficient construction of bioactive molecules.

3. 2. Results and Discussion

3. 2. 1. Consideration of kinetics underlying substrate competition

Before we began screening libraries, we investigated the many variables of SUMS, such as substrate choice, relative substrate concentrations, and assay duration, that can impact the observed product profiles. To connect the SUMS output to the underlying kinetics on single substrates, we used the *Rgn*TDC reaction as a model system (Fig 2). For a unimolecular reaction under initial velocity conditions with equimolar substrates in competition with one another, the product abundances will be exactly proportional to the catalytic efficiencies (k_{cat}/K_M) of the individual reactions in isolation (Fig 2a).[39,43,44] As has been described, this relationship holds true even when the individual substrate concentrations exceed their K_M's.[39,43] We measured traditional Michaelis-Menten parameters for *Rgn*TDC with a variety of substituted Trp analogs (Table 1). When *Rgn*TDC was subjected to a mixture of three Trp analogs (Trp, 4-Br-Trp, and 6-Cl-Trp), the resulting product distribution reflected the relative catalytic efficiencies of each substrate (Fig 2b). As has long been appreciated in enzymology, such multiplexed activity measurements are a true measure of specificity and provide rich kinetic information about enzyme function.[19] However, these relationships are restricted to initial velocity conditions and are an incomplete measure of synthetic utility.

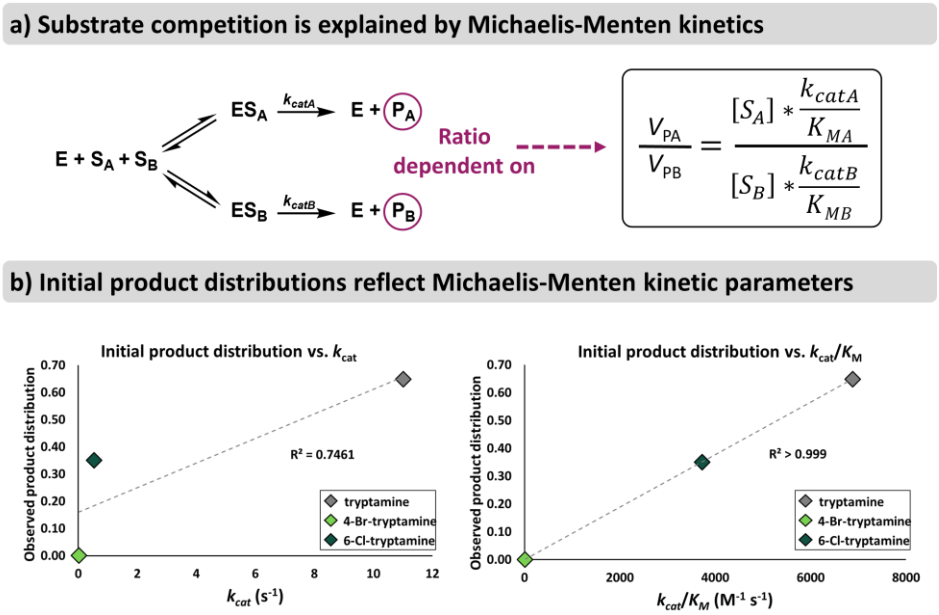


Figure 2. *Rgn*TDC serves as a model system to investigate underlying kinetic phenomena of substrate competition.

To capture enzyme stability effects and achieve high conversions, effective screening conditions for biocatalysis applications often utilize longer reaction times beyond the initial velocity regime. When reactions are run to higher conversion, the product profile becomes uncoupled from the Michaelis-Menten kinetics and is, instead, a heuristic readout of reactivity (Fig 3a). Additionally, both substrates and products can inhibit enzyme activity. Often, but not always, more reactive substrates act as strong competitive inhibitors of activity on poor substrates.[43,44] By decreasing the relative concentrations of reactive substrates, conversion with less active substrates can be observed (Fig 3b). Regardless, we posit that by screening on a mixture with both highly and poorly reactive substrates, we can identify catalysts that retain the ability to operate at high turnover numbers as well as identify desirable increases in activity with multiple sluggish substrates. This information is useful for organic synthesis, where enzymes would ideally react with a wide range of substrates and drive reactions to high yield. By altering the substrate composition and reaction time, screening conditions can quickly be tuned to match diverse engineering goals (Fig 2b).

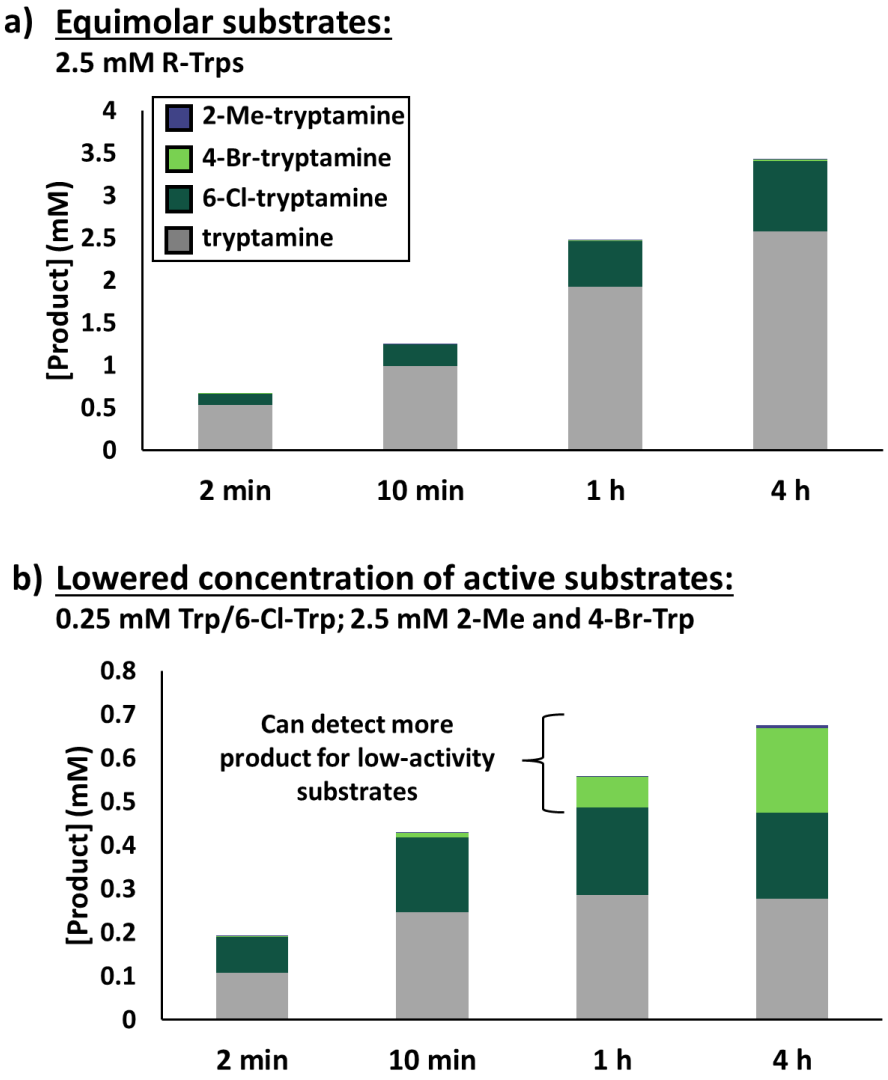


Figure 3. Relative product distributions change over time. a) Product distribution from a reaction with *Rgn*TDC and four equimolar Trp analogs over time. Reaction conditions: 2.5 mM each Trp, 0.25 μM *Rgn*TDC, 37 °C. **b)** Product distribution with Trp and 6-Cl-Trp at 1:10 concentration.

3. 2. 2. SUMS can directly identify mutations that impact substrate promiscuity

We set out to apply a multiplexed method to screen *Rgn*TDC libraries; our engineering goal was to identify mutations that either increased activity on a single substrate or on multiple Trp analogs. We began by engineering for higher *Rgn*TDC activity on 5-substituted Trp analogs, and structure-based modelling suggested the active site residue W349 forms preclusive steric interactions with these substrates (Fig 4c). We screened a site-saturation mutagenesis (SSM)

library, which exchanges the native residue for each other proteinogenic amino acid, at W349 with a mixture of five Trp substrates. For most of the substrates, we found that many mutations increased activity, and that increases in activity varied among the different substrates (Fig 3a). The structurally conservative mutations W349Y and W349F increased activity most with 5-OMe-Trp relative to other substrates, whereas the W349S mutation had the highest activity increase with 5-OEt-Trp and produced the most total product. From this screen, W349K was identified as the most generally improved variant because it produced only slightly less 5-OEt-tryptamine than W349S and formed the most product with all other substrates.

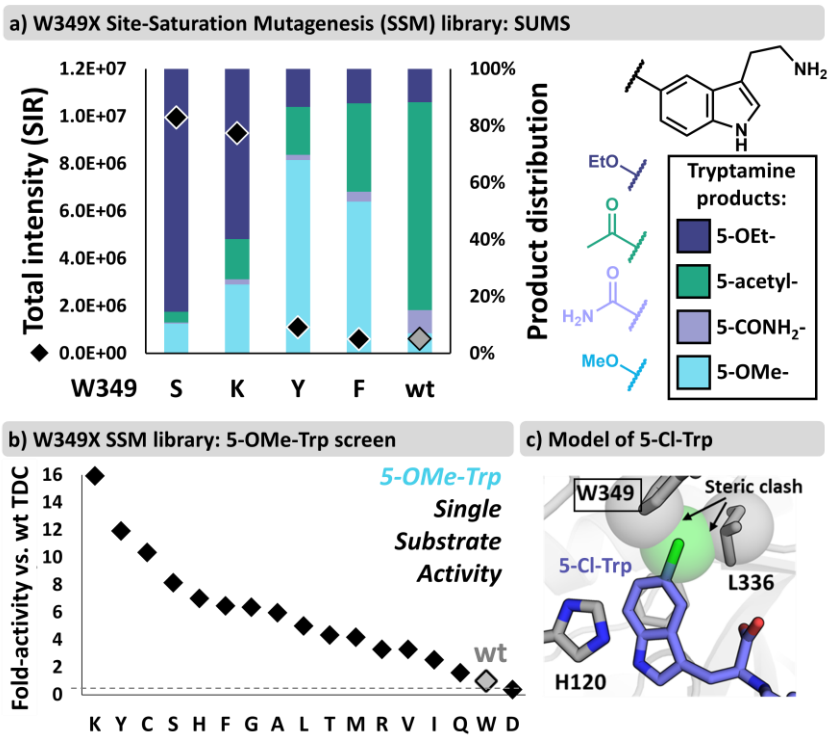


Figure 3. *Rgn*TDC site-saturation mutagenesis library screening results at W349X with either substrate competition or with a single substrate.

To contrast the promiscuity information from SUMS with traditional approaches, we performed a single-substrate screen with 5-OMe-Trp on the same W349 library (Fig 3b). As before, we found that most mutations increased activity with 5-OMe-Trp. However, there was a

poor correlation between activity on 5-OMe-Trp and general activation on 5-substituted Trp analogs. Although W349K was the most activating mutation in both screens, mutations such as W349Y appeared to be highly reactive with 5-OMe-Trp but only poorly tolerated other Trp analogs. These results illustrate how SUMS can immediately identify shifts in both substrate promiscuity and activity with no greater screening effort than would be required for a comparable single substrate screening approach.

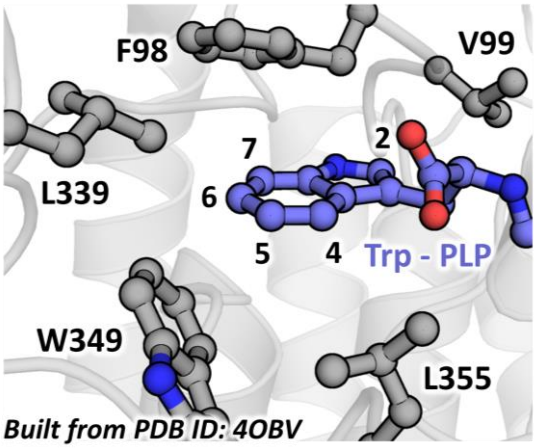


Figure 4. Active-site model of *Rgn*TDC.

While detailed structural analysis revealed W349 as a conspicuous site for improved activity on 5-substituted Trp analogs, such specific hypotheses are not readily formed with all enzymes. We reasoned that SUMS could also be deployed in a setting where there is no specific hypothesis as to which residues govern activity with specific substrates. To simulate this more common scenario, we screened a mixture of Trp analogs that were each substituted at a different position against a set of nine active site SSM libraries (Fig 4, Fig S6-S14). From these screens, we found that mutation at two positions, L126 and H120, had only modest impacts on activity and promiscuity. Mutation at L336 and T356 resulted in many catalytically feeble enzymes, and the variants that retained activity had promiscuity profiles that were similar to

wild-type. For the other sites, mutation caused large changes to apparent promiscuity while retaining significant catalytic activity. For example, we observed > 50-fold activity increases with several enzyme-substrate pairs, such as L355M with 4-Br-Trp and F98V with 2-Me-Trp (Fig 5a). Screening with this more diverse substrate mixture also revealed that W349K maintains high activity with non-5-substituted-Trp substrates like 6-Cl-Trp. Other mutations, such as V99A and L339V, were less strongly activating for 2-Me-Trp and 4-Br-Trp but retained broad activity for substituted Trp analogs.

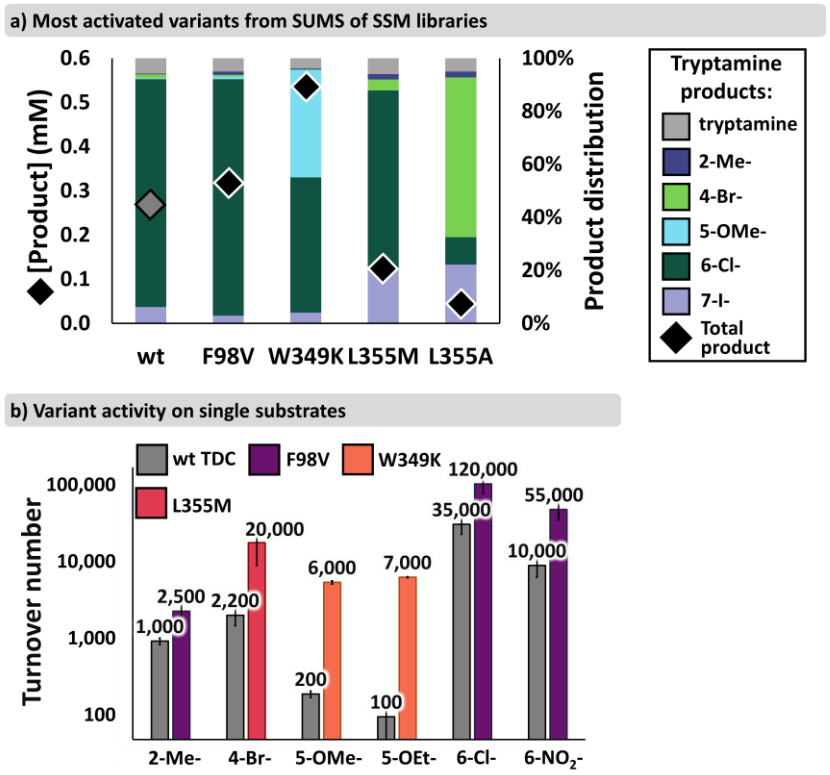


Figure 5. Screening results from *Rgn*TDC active-site site-saturation mutagenesis (SSM) libraries and subsequent validation for activity on single substrates.

3. 2. 3. Variants identified from SUMS have improved single substrate activity

An essential step during protein engineering is to validate the activity of any hits detected during library screening. Since our ultimate goal was to use engineered *Rgn*TDC variants to synthesize tryptamine analogs under single substrate conditions, we validated library hits using

single substrate reactions and purified enzymes. Importantly, although there are many confounding factors that make activity in competition distinct from activity on pure substrates, we found that turnover numbers from these single substrate reactions trended well with multiplexed screening results, with the engineered variants showing large increases in single substrate activity (Fig 5b). Hence, in the same effort necessary to improve activity with one substrate, SUMS enabled the parallel engineering of *Rgn*TDC variants for improved activity with multiple challenging substrates.

To understand the kinetic determinants of substrate promiscuity shifts for *Rgn*TDC variants, we measured Michaelis-Menten parameters for a subset of substrates (Table 1). All activated variants showed higher k_{cat} values with their more reactive substrates when compared to wild-type. The one exception being F98V with 2-Me-Trp. The Michaelis-Menten kinetic reactions were run with only 10-fold excess PLP relative to enzyme, and we found F98V to display low activity with [PLP] < 1 μM (Figure 17). Turnover reactions run with higher [PLP] indicated this variant is activated over wt *Rgn*TDC in terms of TTNs, but not in terms of k_{cat}. However, there was also significant variation in K_M values for activated *Rgn*TDC variants, and such effects were difficult to rationalize from simple structural analysis. The W349K mutation, for example, accelerates decarboxylation of 5-OMe-Trp exclusively by increasing k_{cat}, with minimal impact to K_M values (Table 1). Molecular modeling indicates 4-substituted Trp analogs would form deleterious steric clashes with L355, and rational approaches to engineering would prescribe mutation to smaller sidechains. One such mutation, L355A, improved activity on 4-Br-Trp, but the L355M mutation was even more activating and had a decreased K_M for both Trp and 4-Br-Trp compared to wild-type *Rgn*TDC. We highlight these unexpected findings as an advantage of interrogating active site libraries with SUMS, as such mutations could have been missed entirely by screening with the wrong pairings of substrate and mutational site.

Table 1. *Rgn*TDC variants kinetic parameters

Substrate	Variant	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})
Trp	**wt TDC**	11 ± 1	1.6 ± 0.4	6900
	F98V	8.7 ± 1.4	2.8 ± 0.3	3100
	W349K	1.0 ± 0.1	1.4 ± 0.2	700
	L355A	0.14 ± 0.01	0.53 ± 0.45	260
	L355M	3.3 ± 0.3	0.54 ± 0.36	6200
2-Me-Trp	**wt TDC**	0.055 ± 0.008	2.0 ± 0.4	27
	F98V	0.042 ± 0.009	13 ± 0.3	3.3
4-Br-Trp	**wt TDC**	0.37 ± 0.08	3.4 ± 0.3	110
	L355A	0.82 ± 0.1	1.4 ± 0.2	590
	L355M	1.1 ± 0.1	0.7 ± 0.2	1600
5-OMe-Trp	**wt TDC**	0.017 ± 0.002	0.99 ± 0.48	17
	W349K	0.25 ± 0.03	1.1 ± 0.3	240
6-Cl-Trp	**wt TDC**	0.52 ± 0.02	0.14 ± 0.03	3700
	F98V	0.91 ± 0.06	0.60 ± 0.2	1500

Because activity in competition is not identical to activity on isolated substrates, the information gleaned from SUMS is more than the sum of its parts. The application of SUMS here does not rely on accurate measurements of absolute substrate specificity, as screening does not take place under initial velocity conditions. Instead, we use SUMS to identify changes in the ratios of the products. Although some substrates were fully converted to product under our reaction conditions (i.e., Trp to tryptamine), we identified variants with retained or improved activity on these compounds in single substrate reactions.

3. 2. 4. Engineered *Rgn*TDC variants are improved biocatalysts

Last, we sought to demonstrate the practical utility of the enzymes produced via SUMS. Many enzymes are more synthetically useful when employed in tandem reactions or cascades, which can overcome thermodynamic limitations and obviate the need for purification of intermediates.[45] While the use of multiple enzymes in concert can magnify the benefits afforded by biocatalysis, catalysts must have complementary substrate scopes to synthesize a diverse set of products.[46–48] To this end, we demonstrate efficient cascade catalysis through the mmol-scale syntheses of tryptamine analogs, including 5-OMe-tryptamine and 5-OEt-tryptamine, known serotonin receptor agonists,[49] and 2-Me-tryptamine and 4-Br-tryptamine, which were particularly challenging products for cascade reactions using the parent enzymes. For example, 5-OMe-tryptamine was isolated in double the yield and at larger scale than our previously reported synthesis, while using only a tenth as much catalyst. Each product was made in a telescoped biocatalytic cascade with an engineered tryptophan synthase β-subunit (*Pf*2B9^{H275E}) and an engineered *Rgn*TDC variant and isolated with improved yields compared to reactions with the parent enzymes (Fig 6). Although no *Rgn*TDC variant was identified with improved activity for all Trp analogs, the direct assessment of substrate scope provided by SUMS allowed us to rapidly select an improved catalyst for each tryptamine product. These reactions proceeded smoothly and afforded access to a variety of desirable tryptamines on preparative scale.

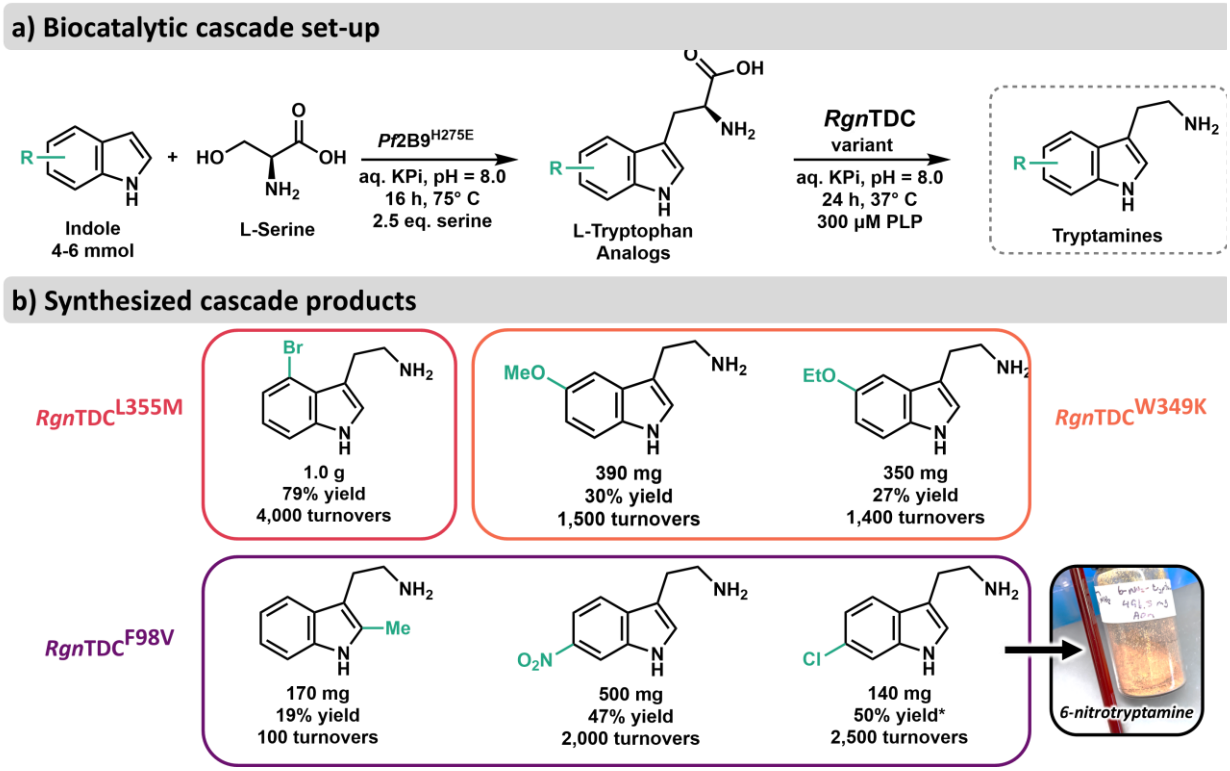


Figure 6. Biocatalytic cascade with engineered *Rgn*TDC variants for improved tryptamine synthesis.

Using the additional information measurable through SUMS, researchers can make more informed choices about which variant(s) to carry forward during directed evolution to maximize biocatalytic utility. A variant that is highly active and selective for a single substrate may be well-suited for subsequent rounds of engineering when a specific reaction is the desired goal. In contrast, mutations that are broadly activating are often more desirable for general synthetic applications. Because selection of activating mutations has historically been made without regard to the potential impact on the scope of an enzyme, the consequences of mutational choice with respect to promiscuity are unknown. It is generally held that there is an activity-specificity tradeoff, but this hypothesis has seldom been tested by intentionally engineering for promiscuity.[50] More importantly, because catalytic perfection is not needed for an enzyme to be useful, it remains to be determined what practical limits there are to evolution for both broad

scope and high catalytic efficiency. We anticipate that SUMS methods will facilitate answering these fundamental questions. More immediately, SUMS remains an easy-to-implement strategy for biocatalyst engineering where enzymes with a broad and/or well-defined substrate scope are desired for efficient chemical synthesis.

3. 3. Conclusions

We provide here an exploration of how to use SUMS to engineer enzymes with improved activity on multiple compounds simultaneously. By directly assessing enzyme activity on substrates in competition, SUMS provides uniquely rich promiscuity information that has hitherto been underutilized. Importantly, just as knowledge of enzyme mechanism is not a pre-requisite for the effective application of directed evolution, *a priori* kinetic knowledge is not required for the design and interpretation of effective multiplexed screens. These results, and the generality of the assay design principles described here, suggest the potential for SUMS to be applied to virtually any class of enzyme. Hence, the ease of implementing SUMS should minimize barriers to adoption of this budding approach.

3. 4. Materials and Methods

3. 4. 1. General Experimental Methods

Chemicals and reagents were purchased from commercial suppliers (Sigma-Aldrich, VWR, Chem-Impex International, Alfa Aesar) and used without further purification. *E. coli* (BL21 (DE3)) cells were used for protein expression. An Eppendorf E-porator (2500 V) was used for transformations. New Brunswick I26R, 120 V/60 Hz shakers (Eppendorf) were used for cell growth. Cell disruption via sonication was performed with a Sonic Dismembrator 550 (Fisher Scientific) sonicator. Optical density measurements were collected on a UV-2600 Shimadzu spectrophotometer (Shimadzu). UPLC-MS data were collected on an Acquity UHPLC with an Acquity QDA MS detector (Waters). Column separations were performed on an Isolera One Flash Purification system (Biotage). NMR data were collected on a Bruker 500 MHz

spectrometer. High resolution mass data were collected with a Q Extractive Plus Orbitrap (NIH 1S10OD020022-1) instrument with the samples ionized by ESI. Kinetic data were fit with PRISM 8 Graphpad software.

Cloning and expression of *Rgn*TDC site-saturation libraries

A codon-optimized copy of the *Ruminococcus gnavus* tryptophan decarboxylase (*Rgn*TDC) gene with a C-terminal His-tag was previously purchased as a gBlock from Integrated DNA Technologies. This DNA fragment was inserted into a pET22b vector via Gibson assembly.[51] BL21 *E. coli* cells were subsequently transformed with the resulting cyclized DNA product *via* electroporation. After 30 min of recovery in Luria-Burtani (LB) media at 37 °C, cells were plated onto LB plates with 100 µg/mL ampicillin (AMP) and incubated overnight. Single colonies were used to inoculate 5 mL TB + 100 µg/mL AMP and grown overnight at 37 °C. pET22b-*Rgn*TDC plasmid was purified via the Zymo Research Plasmid Prep kit and eluted in 20 µL sterile Milli-Q H_2O. Plasmids were stored at -20 °C.

Based on a previously reported model of an *Rgn*TDC-Trp complex (PDB ID: 4OBV),[42] active site residues in the *Rgn*TDC active site were chosen for mutagenesis. The sites were: Phe98, Val99, His120, Leu126, Leu336, Leu339, Trp349, Leu355, and Thr356 (Supplementary Fig. 5). Primers were purchased from Integrated DNA Technologies. For each site of mutation, three primers encoding the degenerate codons NDT, VHG, and TGG at the codon of interest were mixed in a 12:9:1 ratio, respectively.[52]

Each gene library was amplified first as two separate fragments and then combined via polymerase chain assembly (PCA)[53] to form full-length *Rgn*TDC gene mutagenized at the site of interest. The corresponding genes were then inserted into a pET22b vector as described above and then transformed into BL21(DE3) *E. coli* cells and plated on LB + 100 µg/mL AMP agar plates.

Individual colonies were picked and used to inoculate wells containing 600 µl TB + 100 µg/mL AMP in a 96-well plate. One column of the plate was used as an internal control with five wild-type colonies, one sterile control, and two non-*Rgn*TDC pET22b negative controls. Plates were grown at 37 °C for 16 h at 200 RPMs. 20 µl of this starter culture plate was used to inoculate 630 µl TB + 100 µg/mL AMP. These expression plates were grown at 37 °C for 3 h until OD_{600} ~ 1.0. Plates were then placed on ice for 1 h. Expression was induced by addition of 50 µl TB + 100 µg/mL AMP + 14 mM IPTG + 7 mM indole. Plates were then grown at 23 °C for 16 h at 200 RPM and subsequently spun down at 4000 x*g* for 15 min and the supernatant discarded. Plates containing cell pellets were stored at -20 °C.

Screening of *Rgn*TDC site-saturation libraries

Cell pellets were thawed and then resuspended in lysis buffer: 50 mM potassium phosphate buffer (pH = 8.0), 1 mg/mL Hen Egg White Lysozyme (GoldBio), 0.2 mg/mL DNaseI (GoldBio), 1 mM $MgCl_2$, and 300 µM pyridoxal 5′-phosphate (PLP). A volume of 600 µL lysis buffer per well was used. After 45 min of shaking at 37 °C, the resulting lysate was then spun down at 4000 x*g* to pellet cell debris. Then, 180 µL of the resulting supernatant was added to 20 µL of a substrate mixture in a separate reaction plate. Final substrate concentrations are as follows: **W349X 5-substituted-Trp screen**: 2 mM each of 5-methoxytryptophan, 5-ethoxytryptophan, 2-methyl-5-methoxytryptophan, 5-carboxamidotryptophan, and 5-acetyltryptophan; **W349X single-substrate screen**: 2 mM 5-methoxytryptophan; **active-site site-saturation mutagenesis screens**: 2 mM each of 2-methyltryptophan, 4-bromotryptophan, 5-methoxytryptophan, and 6-chlorotryptophan; 0.2 mM 7-iodotryptophan and 0.2 mM tryptophan. Reactions were incubated at 37 °C for 4 h, quenched via addition of 150 µL 1:1 acetonitrile:1 M HCl, and centrifuged at 4000 x*g* for 10 min. 200 µL of the quenched reaction mixture supernatant was filtered into a 96-well plate for UPLC-MS analysis. Data were collected on an Acquity UHPLC with an Acquity QDA MS detector (Waters) using an Intrada Amino Acid

column (Imtakt). Tryptamine product m/z ion counts were used to quantify product formation from the tryptophan reaction mixture from corresponding standard curves (Fig 8-16).

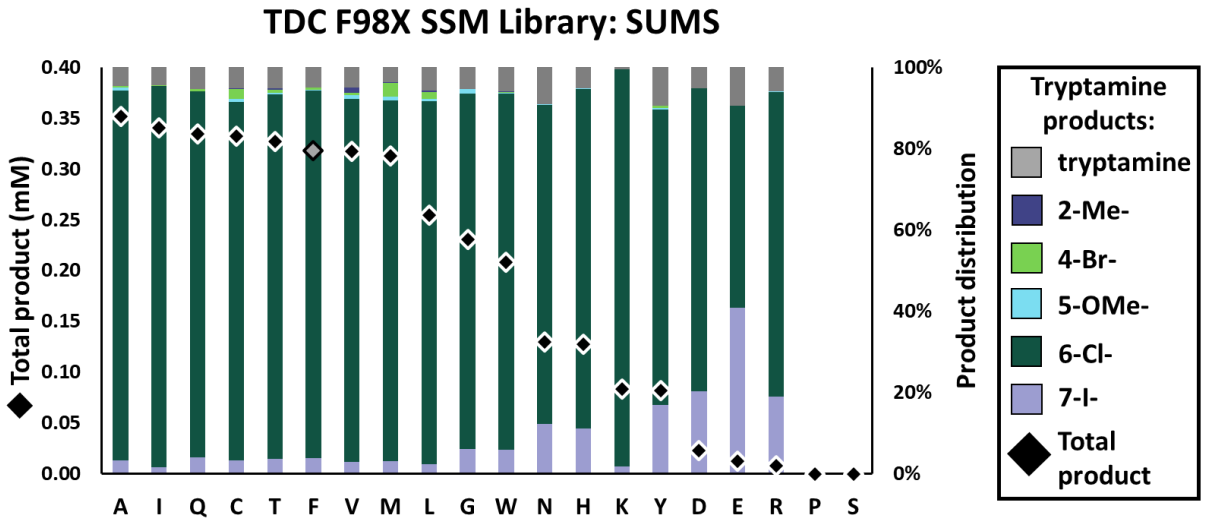


Figure 7. Retention of function (ROF) curve from SUMS of the *Rgn*TDC F98 site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention standard curves. The wild-type sequence is denoted by a grey diamond. Relative product amounts and mM total product were averaged from all wells with the given sequence. (No proline or serine mutations were sequenced from this library). The sampling rate of each mutation is detailed in Table S4.

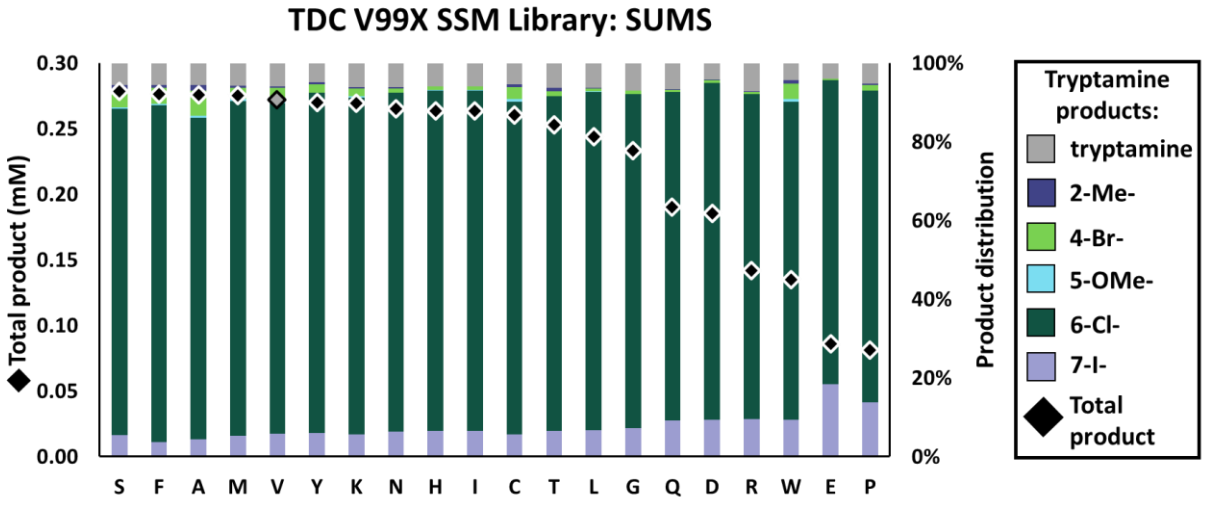


Figure 8. Retention of function (ROF) curve from SUMS of the *Rgn*TDC V99 site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion

retention standard curves. The wild-type sequence is denoted by a grey diamond. Relative product amounts and mM total product were averaged from all wells with the given sequence. The sampling rate of each mutation is detailed in Table S4.

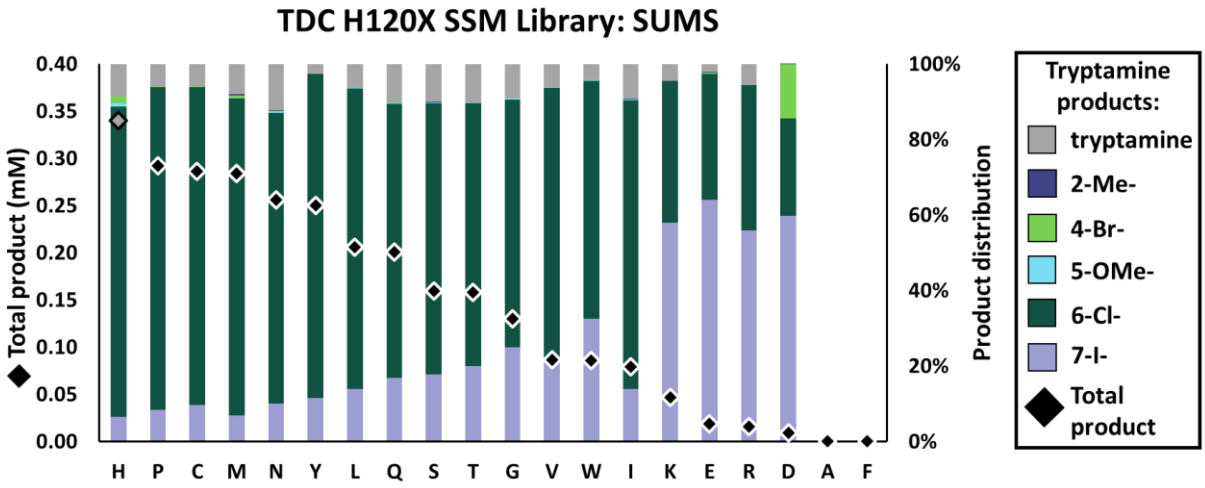


Figure 9. Retention of function (ROF) curve from SUMS of the *Rgn*TDC H120 site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention standard curves. The wild-type sequence is denoted by a grey diamond. Relative product amounts and mM total product were averaged from all wells with the given sequence. (No alanine or phenylalanine mutations were sequenced from the library). The sampling rate of each mutation is detailed in Table S4.

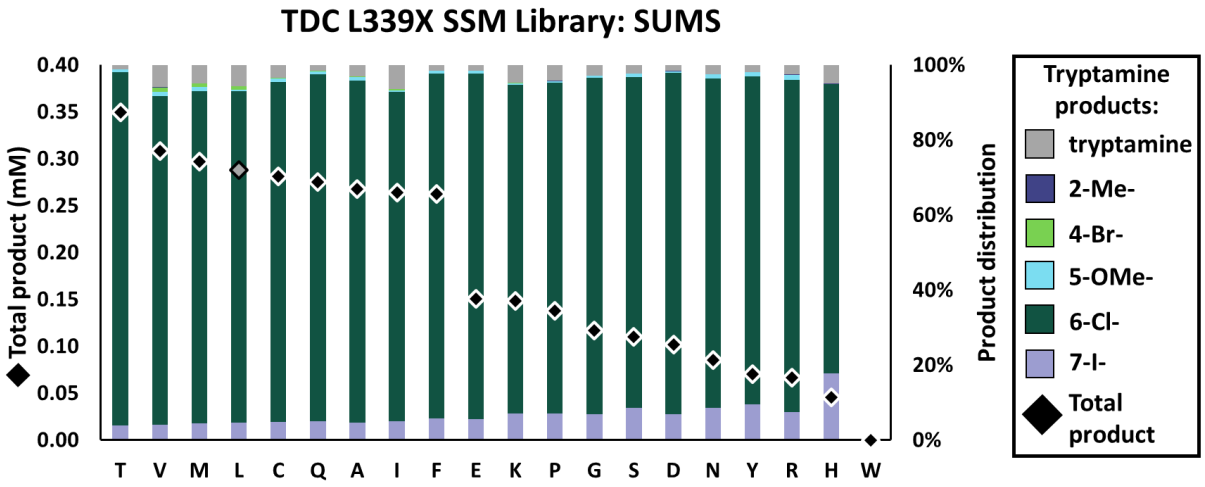


Figure 10. Retention of function (ROF) curve from SUMS of the *Rgn*TDC L339 site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention standard curves. The wild-type sequence is denoted by a grey diamond. Relative product amounts and mM total product were averaged from all wells with the given sequence.

(No tryptophan mutations were sequenced from the library). The sampling rate of each mutation is detailed in Table S4

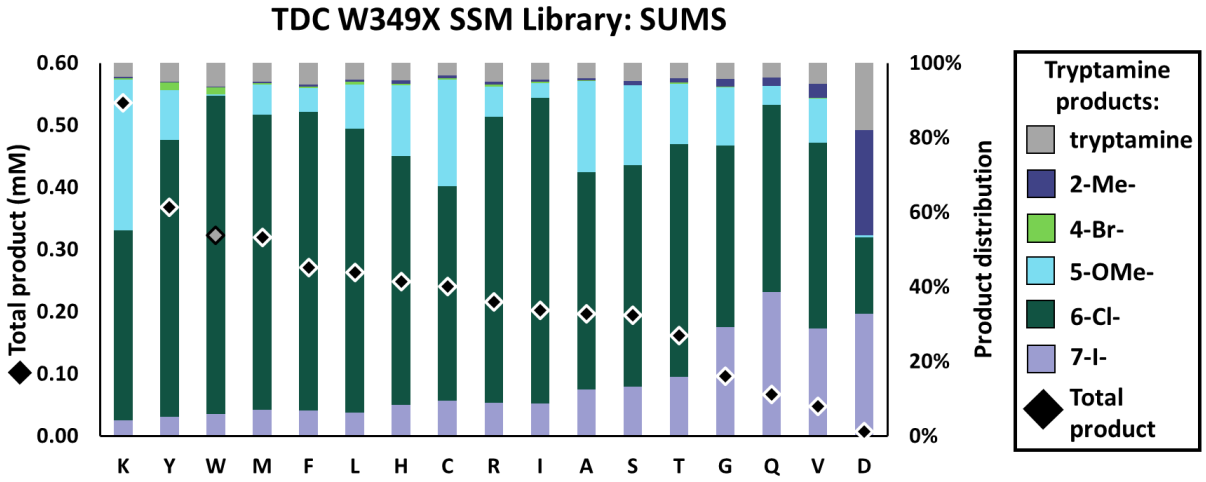


Figure S11. Retention of function (ROF) curve from SUMS of the *Rgn*TDC W349 site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention standard curves. The wild-type sequence is denoted by a grey diamond. Relative product amounts and mM total product were averaged from all wells with the given sequence. The sampling rate of each mutation is detailed in Table S4.

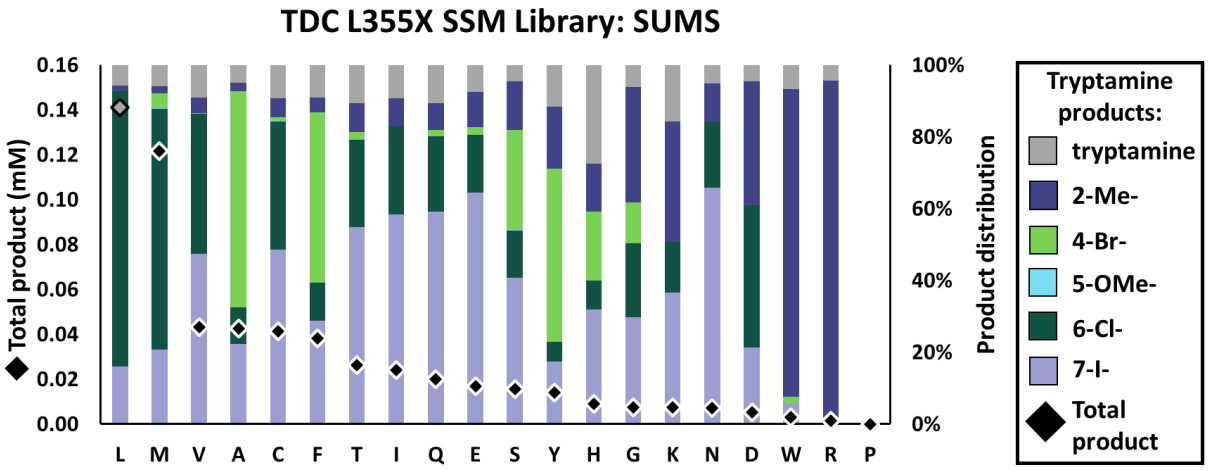


Figure 12. Retention of function (ROF) curve from SUMS of the *Rgn*TDC L355 site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention standard curves. The wild-type sequence is denoted by a grey diamond. Relative product amounts and mM total product were averaged from all wells with the given sequence. (No proline mutations were sequenced from the library). The sampling rate of each mutation is detailed in Table S4.

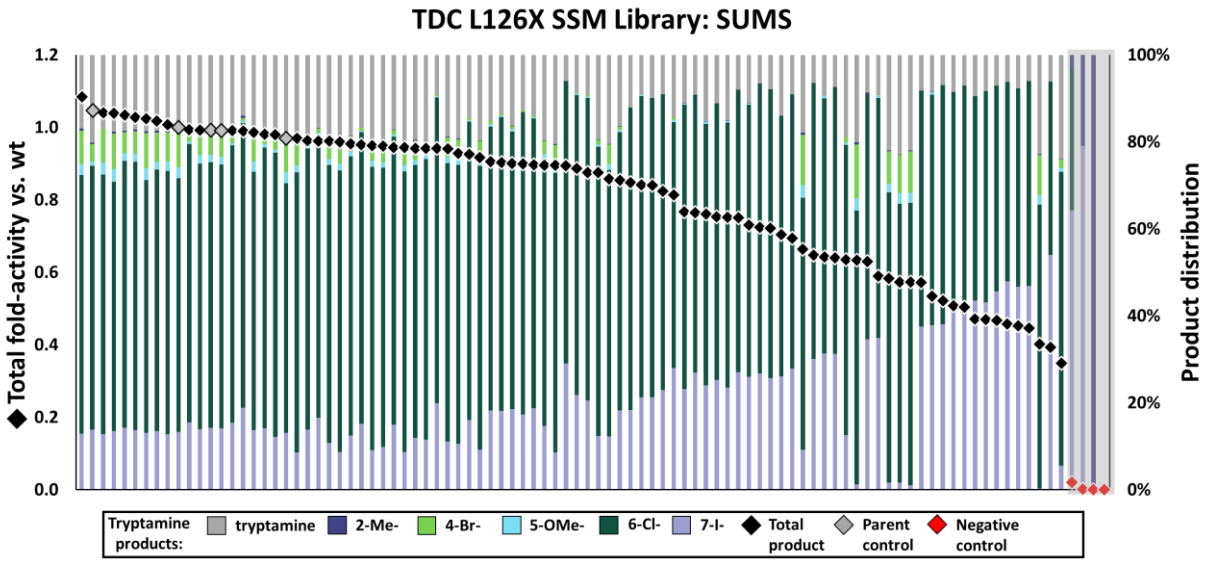


Figure 13. Retention of function (ROF) curve from SUMS of the *Rgn*TDC L126 site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention standard curves. Wild-type sequences are denoted by grey diamonds, while negative controls are shown as red diamonds. Greyed-out section indicates wells where the quality of signal-to-noise was too low and the relative product distributions are dominated by noise and therefore no longer indicative of enzyme promiscuity.

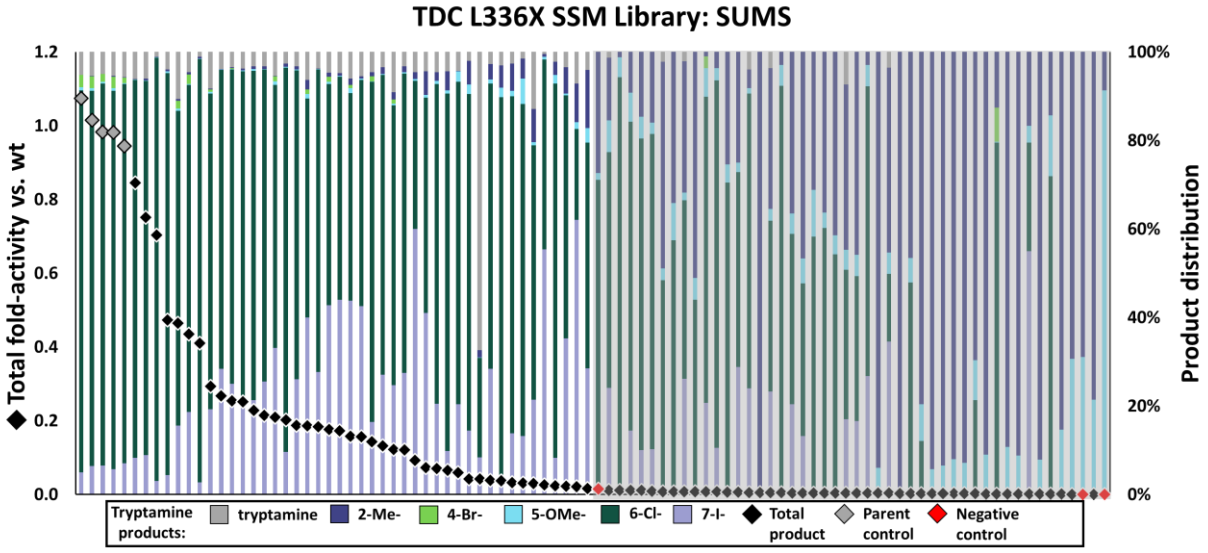


Figure 14. Retention of function (ROF) curve from SUMS of the *Rgn*TDC L336 site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention standard curves. Wild-type sequences are denoted by grey diamonds, while negative controls are shown as red diamonds. Greyed-out section indicates wells where the quality of signal-to-noise was too low and the relative product distributions are dominated by noise and therefore no longer indicative of enzyme promiscuity.

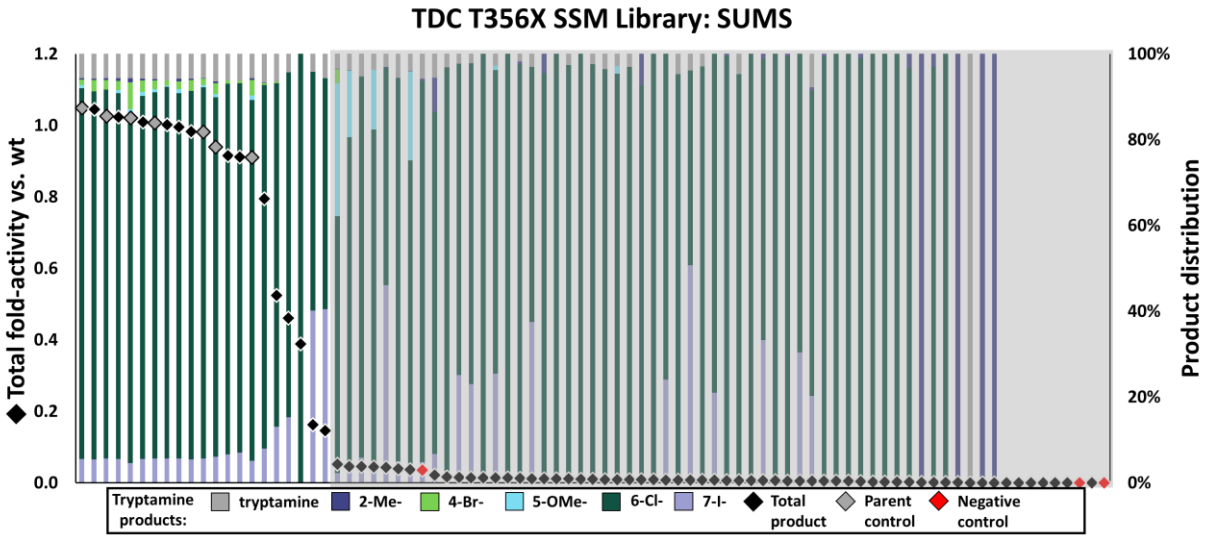


Figure 15. Retention of function (ROF) curve from SUMS of the *Rgn*TDC T356 site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention standard curves. Wild-type sequences are denoted by grey diamonds, while negative controls are shown as red diamonds. Greyed-out section indicates wells where the quality of signal-to-noise was too low and the relative product distributions are dominated by noise and therefore no longer indicative of enzyme promiscuity.

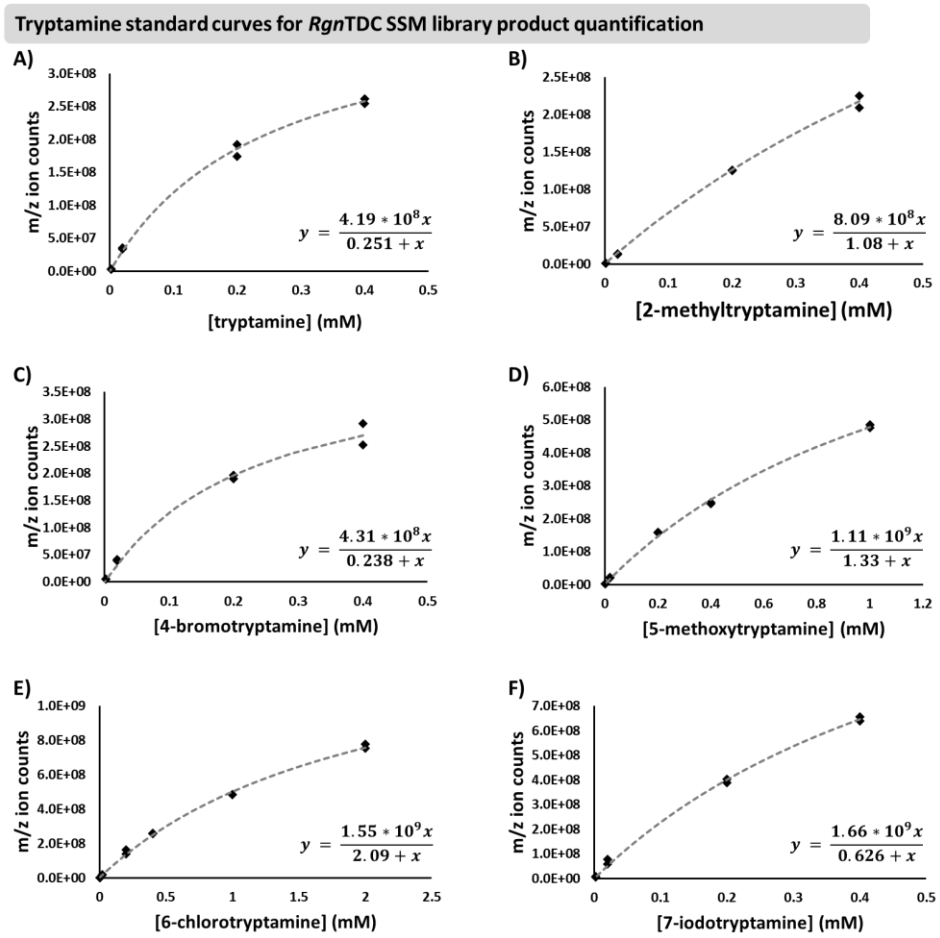


Figure 16. Tryptamine standard curves for analysis of product abundance in *Rgn*TDC SSM libraries. Due to non-linear effects caused by electrospray ionization efficiency, data were fit to hyperbolic equations.

Expression of *Rgn*TDC and *Pf*TrpB variants

Cells from wells containing putatively activated *Rgn*TDC or *Pf*TrpB variants were inoculated into 5 mL TB + AMP and grown for 16 h at 37 °C. Each variant strain culture was used to inoculate 500 mL TB + AMP, which was subsequently grown at 37 °C for 3 h, or until OD_{600} ~ 1.5. Cultures were chilled on ice for 30 min, and then induced with 1 mM IPTG. Cells expressing *Rgn*TDC were additionally supplemented with 0.5 mM indole, as this was found to boost cell growth. Cultures expressed for 16 h at 23 °C. The next morning, cells were pelleted via centrifugation at 4000 x*g* for 15 min. Supernatants were bleached and discarded and pellets were weighed and stored at -20 °C until lysis.

To purify enzyme variants, cell pellets were thawed at room temperature and then resuspended in lysis buffer: 50 mM potassium phosphate buffer (pH = 8.0), 1 mg/mL Hen Egg White Lysozyme (GoldBio), 0.2 mg/mL DNaseI (GoldBio), 1 mM $MgCl_2$, and 300 μM pyridoxal 5′-phosphate (PLP). A volume of 4 mL of lysis buffer per gram of wet cell pellet was used. After 30 min of shaking at 37 °C, *Rgn*TDC lysis suspensions were disrupted using sonication (5 min; 0.8 s on, 0.2 s off at a power setting of 5). *Pf*TrpB lysis suspensions were heat treated at 75 °C for 15 min as an alternative to sonication. The resulting lysate was spun down at 75,000 x*g* to pellet cell debris. Ni/NTA beads (GoldBio) were added to the supernatant and incubated on ice for 30 to 60 min prior to purification by Ni-affinity chromatography. The column was washed with 3 column volumes of wash buffer A (20 mM imidazole, 50 mM potassium phosphate buffer (pH = 8.0)) and with 2 column volumes of wash buffer B (40 mM imidazole, 50 mM potassium phosphate buffer (pH = 8.0); *Pf*TrpB variants only), and the proteins were eluted with an elution buffer (250 mM imidazole, 50 mM potassium phosphate buffer, pH = 8.0). Elution of the desired protein product was monitored by the disappearance of its bright yellow color (resulting from the PLP cofactor) from the column. The protein product was dialyzed to < 1 μM imidazole, dripped into liquid nitrogen to flash freeze, and stored at -80 °C. The concentration of protein was determined by Bradford assay upon thawing. Generally, this procedure yielded 50 – 200 mg per L culture for *Rgn*TDC variants and 200 – 500 mg per L culture for *Pf*TrpB variants.

Turnover analysis of *Rgn*TDC variants

*Rgn*TDC variants were thawed on ice from storage at -80 °C and then centrifuged at 15,000 x*g* for 5 min to pellet aggregated protein. The supernatant was then diluted between 1:1 and 1:100 in 50 mM potassium phosphate buffer (pH = 8.0) (depending on enzyme concentration). Reactions were run in plastic 96-well plates with 10 mM Trp analog, either 33 nM or 1 μM TDC variant, and 1000 equivalents PLP relative to TDC with a total volume of 100 μL in 50 mM potassium phosphate buffer (pH = 8.0). Reactions were incubated for 16 h at 37 °C. Reactions

were quenched with 200 µL acetonitrile and further diluted with 100 µL H_2O. Quenched reactions were centrifuged at 15,000 x*g* for 10 min prior to analysis by UPLC-MS. Enzymatic activity was quantified by integrating the substrate and product UV absorbance peaks at 280 nm. Reactions were run in triplicate, with replicate reactions run on different days with freshly thawed enzyme (Table 2). We note that the TTN for reactions of 7-I-Trp were significantly higher than originally observed in Chapter 2. We attribute this difference to the concentration of PLP for these reactions. [PLP] for the reported TTN's in Chapter 2 was ~1 µM (being 10x relative to [*Rgn*TDC[), while [PLP] for the reactions presented here contained 1000x PLP relative to [*Rgn*TDC]. This addition of extra cofactor increased overall conversion.

Table 2. *Rgn*TDC variant total turnover number with different substrates

	wt TDC	F98V	W349K	L355A	L355M
Trp	90,000 ± 21,000	210,000 ± 30,000	21,000 ± 3,000	3,000 ± 2,000	82,000 ± 19,000
7-I-Trp	250,000 ± 60,000	56,000 ± 11,000	19,000 ± 4,000	8,000 ± 900	160,000 ± 50,000
6-Cl-Trp	35,000 ± 9,000	120,000 ± 30,000	13,000 ± 4,000	1,000 ± 200	41,000 ± 14,000
6-NO2-Trp	10,000 ± 3,000	55,000 ± 15,000	18,000 ± 6,000	300 ± 100	14,000 ± 6,000
5-OMe-Trp	220 ± 20	160 ± 20	6,000 ± 300	60 ± 10	240 ± 20
5-OEt-Trp	90 ± 70	90 ± 10	7,000 ± 100	30 ± 5	90 ± 90
4-Br-Trp	2,200 ± 600	400 ± 300	100 ± 70	10,000 ± 2,000	20,000 ± 10,000
2-Me-Trp	1,000 ± 100	2,500 ± 700	7 ± 0	1,000 ± 100	1,000 ± 240

[PLP] vs. activity for *Rgn*TDC variants

*Rgn*TDC variants were thawed and prepared as described above. 10 mM 6-chlorotryptophan was added to a solution with varying molar equivalents of PLP relative to TDC (500 nM, 5 μM, 50 μM, or 500 μM). TDC was added such that the final concentration was 50 nM and the solution was diluted up to 100 μl with 50 mM potassium phosphate buffer (pH = 8.0). Reactions were incubated for 16 h at 37 °C. Reactions were quenched with 200 μL acetonitrile and diluted with 100 μL H_2O. Quenched reactions were centrifuged at 15,000 x*g* for 10 min prior to UPLC-MS injection of the supernatant. Enzymatic activity was quantified by integrating the

substrate and product UV absorbance peaks at 280 nm. Reactions were run in duplicate, with replicate reactions run on separate days with freshly thawed enzyme (Fig 17)

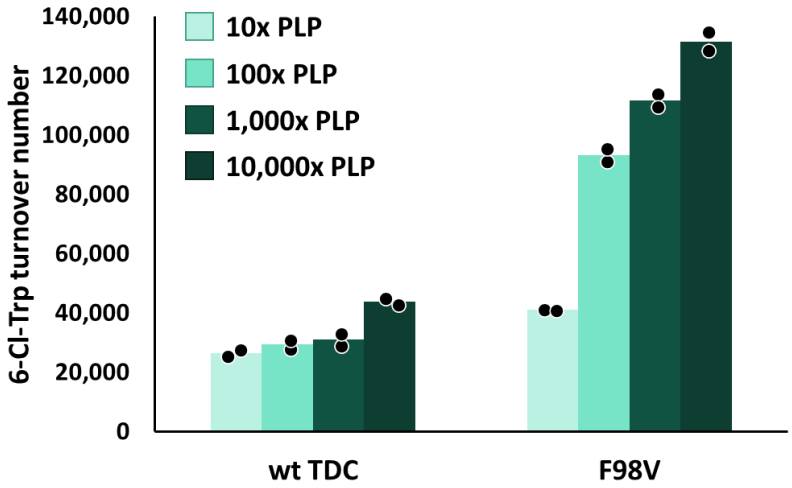


Figure 17. Comparison of effect of additional pyridoxal phosphate (PLP) on conversion of 6-chlorotryptophan (6-Cl-Trp) to 6-chlorotryptamine. Amount of PLP shown is relative to the *Rgn*TDC concentration of 50 nM. Reactions were conducted in duplicate on different days. F98V *Rgn*TDC showed a higher dependence on PLP than wild-type (wt) *Rgn*TDC, potentially indicating weaker cofactor binding or increased turnover-dependent cofactor decomposition.

*Rgn*TDC substrate multiplexed timecourse reaction

*Rgn*TDC variants were thawed and prepared as described above. Reaction conditions included 0.25 µM *Rgn*TDC, 2.5 µM PLP, and one of two sets of substrate mixes: 2.5 mM tryptophan, 2-methyltryptophan, 4-bromotryptophan, and 6-chlorotryptophan; or 0.25 mM tryptophan and 6-chlorotryptophan, 2.5 mM 2-methyltryptophan and 4-bromo-tryptophan. Reactions were filled to 100 µl with 50 mM potassium phosphate pH = 8.0. Reactions were conducted in triplicate and incubated for 4 h at 37 °C. Reactions timepoints were taken by

quenching 10 µl of the reaction solution in 190 µl of acetonitrile, followed by dilution with 200 µl H_2O. Quenched reactions were centrifuged at 15,000 x*g* for 10 min prior to UPLC-MS injection of the supernatant. Product abundance was quantified via single-ion retention and fit to product standard curves (Fig 3).

Michaelis-Menten analysis of *Rgn*TDC variants

*Rgn*TDC variants were thawed on ice from storage at -80 °C and then centrifuged at 15,000 x*g* for 5 min to pellet aggregated protein. The supernatant was then diluted between 1:1 and 1:100 in 50 mM potassium phosphate buffer (pH = 8.0), depending on enzyme concentration. Reaction conditions used 0.1 – 20 mM tryptophan analog with 10 equivalents of PLP cofactor relative to TDC in 1.7 mL Eppendorf tubes with a total volume of 100 µL in 50 mM potassium phosphate buffer (pH = 8.0). The enzyme and substrate solutions were equilibrated to 37 °C prior to mixing. Final enzyme concentrations were 0.1 – 5 µM. At various time points after addition of *Rgn*TDC, 20 µL aliquots of the reaction solution were quenched in 180 µL acetonitrile and diluted with 200 µL H_2O. Quenched reactions were centrifuged at 15,000 x*g* for 10 min prior to UPLC-MS injection of the supernatant. Enzymatic activity was quantified by integrating the substrate and product UV absorbance peaks at 280 nm. All replicate reactions were run on different days with fresh enzyme. Initial velocity slopes were fit using the revised Michaelis-Menten equation as described by Johnson 2019.[54] Michaelis-Menten curves are provided in Figures 18-22.

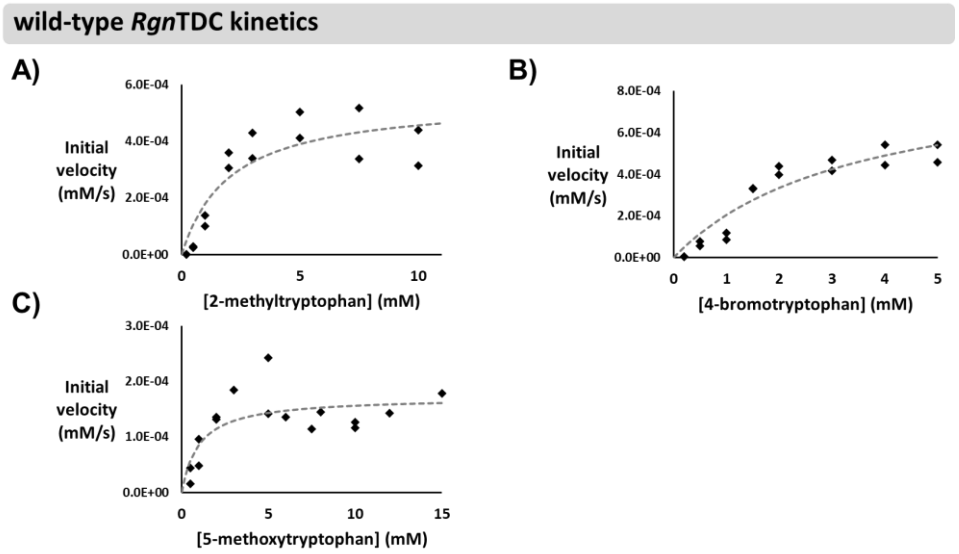


Figure 18. Michaelis-Menten curves for wild-type *Rgn*TDC with 2-methyltryptophan, 4-bromotryptophan, and 5-methoxytryptophan. Relative enzyme concentration used are as follows: **A)** 10 μM *Rgn*TDC, **B)** 2.5 μM *Rgn*TDC, **C)** 10 μM *Rgn*TDC.

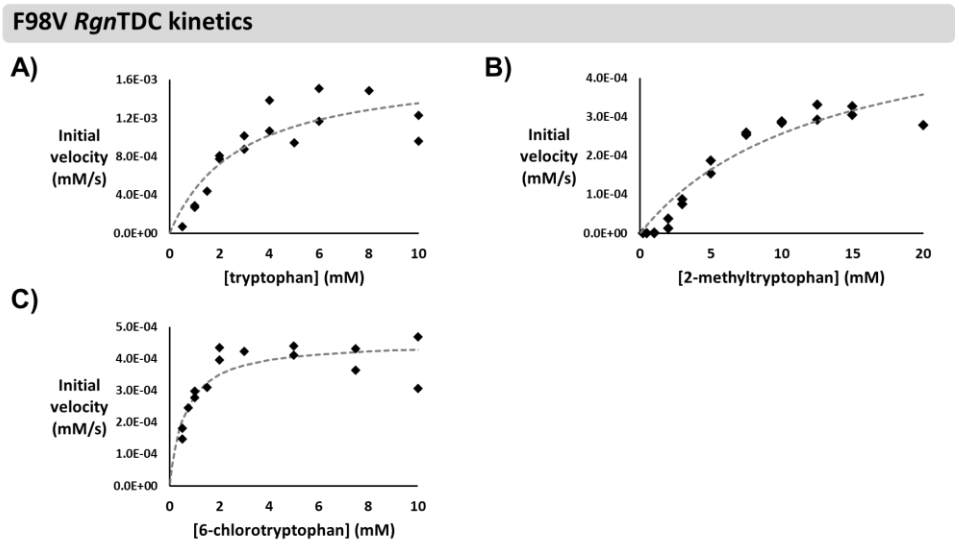


Figure 19. Michaelis-Menten curves for F98V *Rgn*TDC with tryptophan, 2-methyltryptophan, and 6-chlorotryptophan. Relative enzyme concentration used are as follows: **A)** 0.2 μM *Rgn*TDC, **B)** 14 μM *Rgn*TDC, **C)** 0.5 μM *Rgn*TDC.

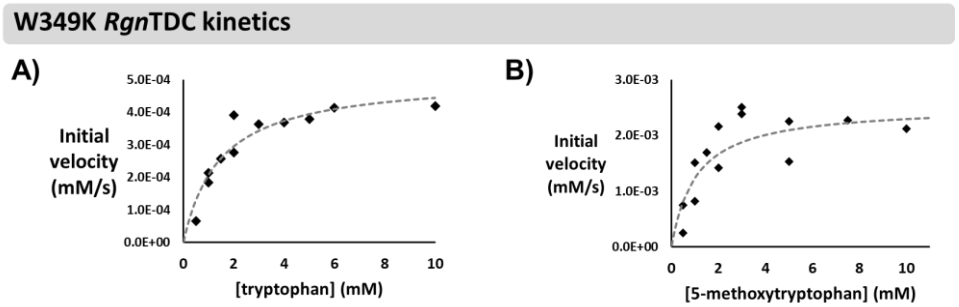


Figure 20. Michaelis-Menten curves for W349K *Rgn*TDC with tryptophan and 5-methoxytryptophan. Relative enzyme concentration used are as follows: **A)** 0.5 μM *Rgn*TDC, **B)** 10 μM *Rgn*TDC.

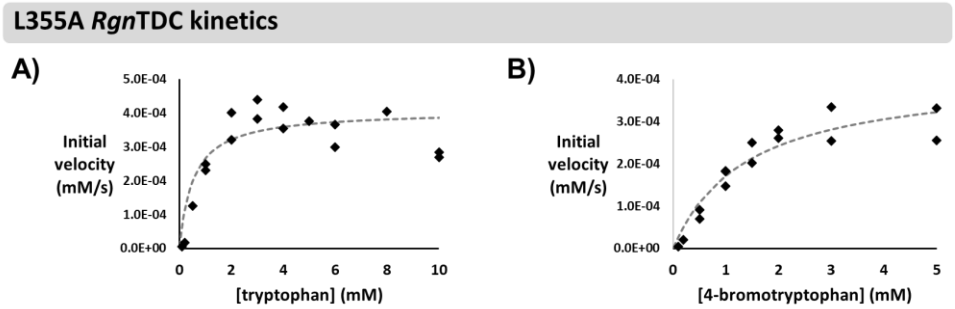


Figure 21. Michaelis-Menten curves for L355A *Rgn*TDC with tryptophan and 4-bromotryptophan. Relative enzyme concentration used are as follows: **A)** 3 μM *Rgn*TDC, **B)** 0.5 μM *Rgn*TDC**.**

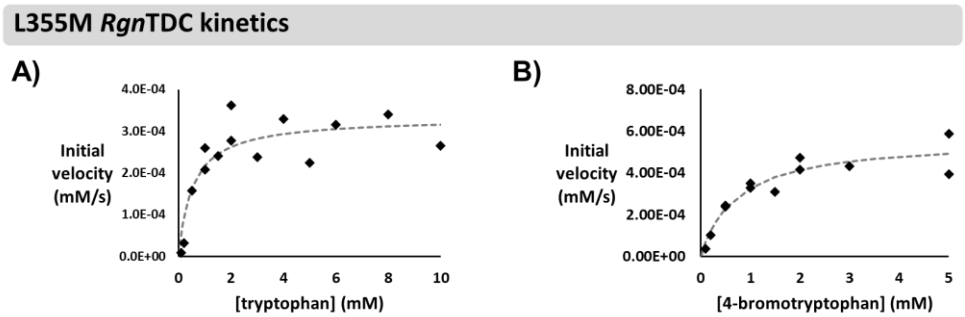


Figure 22. Michaelis-Menten curves for L355M *Rgn*TDC with tryptophan and 4-bromotryptophan. Relative enzyme concentration used are as follows: **A)** 0.1 μM *Rgn*TDC, **B)** 0.5 μM *Rgn*TDC.

Tryptophan analog biosynthesis and purification

For each tryptophan analog, the corresponding indole derivative (~1 mmol) was added to a 100 mL pressure flask and dissolved in MeOH (5 mL). L-Serine (3 mmol) was added, and the resulting solution was diluted to just under 100 mL with 50 mM potassium phosphate buffer (pH = 8.0). PLP was added such that the final concentration was 300 μM. Then, 2B9 was added at 0.1% mol catalyst relative to the indole analog. The solution was incubated at 75 **°C for 16 h. Following UPLC-MS analysis of conversion, the solution was heat-treated at 90 °C for 30 min. Solutions were evaporated down to 10 mL, filtered, and run over C18 via flash chromatography with H_2O/MeOH. Tryptophan products typically eluted between 10-35% MeOH. Product-containing fractions were combined and evaporated down to ~2-3 mL, where the solutions were transferred to pre-tared flasks, evaporated to dryness, resuspended in H_2O, flash-frozen, and lyophilized for 24 – 48 h. Resulting powders were weighed and submitted for ^{1}H-NMR analysis, with final reported yields taking hydration states into account (*all tryptophan analog products were isolated as hydrates*).**

2-methyltryptophan

2-methyltryptophan was isolated as an off-white powder (81.8 mg) in 94% yield. ^{1}H NMR (500 MHz, MeOD) δ 7.64 – 7.58 (m, 1H), 7.31 – 7.25 (m, 1H), 7.06 (td, *J* = 7.1, 1.4 Hz, 1H), 7.02 (td, *J* = 7.5, 1.3 Hz, 1H), 3.86 (dd, *J* = 9.9, 4.2 Hz, 1H), 3.52 – 3.46 (m, 1H), 3.08 (dd, *J* = 15.2, 10.0 Hz, 1H), 2.44 (s, 3H). Spectrum matched previously reported spectra.[55]

2-methyl-5-methoxytryptophan

2-methyl-5-methoxytryptophan was isolated as a tan powder (164 mg) in 66% yield. ^{1}H NMR (500 MHz, MeOD) δ 7.18 (d, *J* = 2.4 Hz, 1H), 7.15 (d, *J* = 8.6 Hz, 1H), 6.69 (dd, *J* = 8.7, 2.4 Hz, 1H), 3.85 (s, 3H), 3.76 (dd, *J* = 9.8, 4.2 Hz, 1H), 3.38 (dd, *J* = 15.0, 4.2 Hz, 1H), 2.97 (dd, *J* = 14.9, 9.8 Hz, 1H), 2.41 (s, 3H). ^{13}C NMR (126 MHz, MeOD) δ 174.08, 153.91, 134.36, 131.06, 128.63, 110.68, 110.27, 104.24, 99.64, 55.86, 54.99, 26.93, 10.18.

MS/ESI m/z for $[M+H]^+$; C13H16N2O3; calculated 247.10882, observed 247.1087.

4-bromotryptophan

4-bromotryptophan was isolated as a white powder (66.9 mg) in 59% yield. ^{1}H NMR (500 MHz, MeOD) δ 7.38 (d, 1H), 7.28 (s, 1H), 7.23 (d, *J* = 7.4 Hz, 1H), 7.00 (t, *J* = 7.9 Hz, 1H), 4.07 (dd, *J* = 10.3, 4.6 Hz, 1H), 3.98 (dd, *J* = 15.1, 4.6 Hz, 1H), 3.13 (dd, *J* = 15.1, 10.3 Hz, 1H). Spectrum matched previously reported spectra.[56]

5-acetyltryptophan

5-acetyltryptophan was isolated as a white powder (162 mg) in 42% yield. ^{1}H NMR (500 MHz, MeOD) δ 8.54 (d, *J* = 1.6 Hz, 1H), 7.85 (dd, *J* = 8.6, 1.7 Hz, 1H), 7.44 (d, *J* = 8.6 Hz, 1H), 7.32 (s, 1H), 3.87 (dd, *J* = 8.7, 4.3 Hz, 1H), 3.54 (dd, *J* = 15.1, 4.3 Hz, 1H), 3.24 (dd, *J* = 15.2, 8.7 Hz, 1H), 2.71 (s, 3H). ^{13}C NMR (126 MHz, D_2O) δ 204.10, 174.67, 139.52, 128.34, 126.68, 126.36, 121.99, 121.84, 111.81, 109.80, 55.11, 26.35, 26.09.

MS/ESI m/z for $[M+H]^+$; C13H14N2O3; calculated 245.0932, observed 245.0930.

5-carboxamidotryptophan

5-carboxamidotryptophan was isolated as a white powder (163 mg) in 52% yield. ^{1}H NMR (500 MHz, MeOD) δ 8.32 (dd, 1H), 7.74 (dd, *J* = 8.6, 1.7 Hz, 1H), 7.44 (dd, *J* = 8.6, 0.7 Hz, 1H), 7.31 (s, 1H), 3.88 (dd, *J* = 9.1, 4.0 Hz, 1H), 3.55 (ddd, *J* = 15.1, 4.1, 1.0 Hz, 1H), 3.19 (dd, *J* = 15.3, 9.1 Hz, 1H). Spectrum matched previously reported spectra.[56]

5-ethoxytryptophan

5-ethoxytryptophan was isolated as a white powder (82.7 mg) in 59% yield. ^{1}H NMR (500 MHz, MeOD) δ 7.26 (d, *J* = 5.3 Hz, 1H), 7.25 (s, 1H), 7.17 (s, 1H), 6.79 (dd, *J* = 8.9, 2.2 Hz, 1H), 4.15 – 4.06 (m, 2H), 3.84 (dd, *J* = 9.3, 4.1 Hz, 1H), 3.14 – 3.05 (m, 1H), 1.42 (t, *J* = 6.9 Hz, 4H). ^{13}C

NMR (126 MHz, D_2O) δ 151.88, 131.88, 127.20, 125.80, 112.78, 112.57, 108.23, 102.41, 65.51, 55.32, 27.44, 14.11.

MS/ESI m/z for $[M+H]^+$; C13H16N2O3; calculated 247.10882, observed 247.1087.

5-methoxytryptophan

5-methoxytryptophan was isolated as a brown powder (556 mg) in 48% yield. ^{1}H NMR (500 MHz, MeOD) δ 7.27 (d, *J* = 3.3 Hz, 1H), 7.25 (d, *J* = 2.8 Hz, 1H), 7.17 (s, 1H), 6.79 (dd, *J* = 8.8, 2.4 Hz, 1H), 3.87 (s, 3H), 3.85 (ddd, *J* = 9.5, 3.9, 1.9 Hz, 1H), 3.52 – 3.46 (m, 1H), 3.11 (dd, *J* = 15.2, 9.6 Hz, 1H). Spectrum matched previously reported spectra.[57]

6-chlorotryptophan

6-chlorotryptophan was isolated as a white powder (63.8 mg) in 67% yield. ^{1}H NMR (500 MHz, MeOD) δ 7.68 (d, *J* = 8.5 Hz, 1H), 7.38 (d, *J* = 1.9 Hz, 1H), 7.23 (s, 1H), 7.04 (dd, 1H), 3.83 (ddd, *J* = 9.4, 5.6, 4.2 Hz, 1H), 3.47 (dd, *J* = 15.2, 4.2 Hz, 1H), 3.15 (dd, 1H). Spectrum matched previously reported spectra.[56]

6-nitrotryptophan

6-nitrotryptophan was isolated as a yellow powder (245 mg) in 98% yield, with some pyruvate impurity. ^{1}H NMR (500 MHz, MeOD) δ 8.36 (d, *J* = 2.0 Hz, 1H), 7.98 (dd, *J* = 8.8, 2.1 Hz, 1H), 7.86 (d, *J* = 8.8 Hz, 1H), 7.57 (s, 1H), 3.79 (dd, *J* = 8.2, 4.6 Hz, 1H), 3.47 – 3.41 (m, 1H), 3.26 – 3.19 (m, 1H). Spectrum matched previously reported spectra.[56]

7-iodotryptophan

7-iodotryptophan was isolated as a white powder (94.3 mg) in 79% yield, with some pyruvate impurity. ^{1}H NMR (500 MHz, MeOD) δ 7.75 (dd, *J* = 7.9, 0.9 Hz, 1H), 7.54 (d, *J* = 7.4 Hz, 1H), 7.30 (s, 1H), 6.88 (t, *J* = 7.7 Hz, 1H), 3.86 (dd, *J* = 9.1, 4.2 Hz, 1H), 3.50 (dd, 1H), 3.16 (dd, 1H). Spectrum matched previously reported spectra.[56]

Cascade synthesis and isolation of tryptamines

4–6 mmol (1.4 mmol for 6-chloroindole) of the corresponding indole analog was added to a 1 L Erlenmeyer flask and dissolved in 20 mL MeOH. 12 mmol serine was added, and the resulting solution was diluted up to just under 500 mL with 50 mM potassium phosphate buffer pH = 8.0. PLP was added such that the final concentration was 300 μM. Then, *Pf*2B9^{H275E} was added at 0.05% mol catalyst relative to the indole analog (0.2 mg/mL). The solution was incubated at 75 **°C for 16 h. Following UPLC-MS analysis of conversion, the solution was cooled to 37 °C, upon which *Rgn*TDC was added** at 0.02 – 0.2% mol catalyst relative to the indole (0.09 – 0.9 mg/mL**). The solutions were incubated at 37 °C for 24 h. Solutions were then evaporated down to 50 – 100 mL. To break emulsions, the solutions were acidified with 6 M HCl until pH < 1, 100 mL ethyl acetate (EtOAc) was added, and the resulting mixtures were centrifuged at 4000 x g for 10 min. These solutions were added to a separatory funnel, the aqueous layer was drained, and the organic layer removed. This was repeated twice more, with 2 mL 6 M HCl added in between extractions. Then, the aqueous layer was alkalized with 6 M NaOH until pH > 12.* Tryptamine products were then extracted 3x with 150 mL EtOAc, with 2 mL 6 M NaOH added in between extractions to the aqueous layer. Organic layers were pooled, dried with sodium sulfate, filtered, and evaporated down to 5-10 mL. Solutions were transferred to 20 mL scintillation vials, evaporated to near dryness (tryptamines were observed as liquids at 50 °C), and dried under vacuum overnight. Dried samples were weighed and submitted for ^{1}H and ^{13}C NMR analysis.**

2-methyltryptamine

F98V *Rgn*TDC was used as 0.2% mol catalyst for the reaction with 2-methylindole. 2-methyltryptamine was isolated as a reddish oil (166 mg) in 19% yield, with some EtOAc impurity. ^{1}H NMR (500 MHz, DMSO) δ 7.40 (dd, *J* = 7.6, 1.1 Hz, 1H), 7.22 (dt, *J* = 8.0, 1.0 Hz,

1H), 7.00 – 6.86 (m, 2H), 2.75 – 2.67 (m, 4H), 2.33 (s, 3H), 2.00 (s, 1H). Spectrum matched previously reported spectra.

4-bromotryptamine

L355M *Rgn*TDC was used as 0.02% mol catalyst for the reaction with 4-bromoindole. 4-bromotryptamine was isolated as a light pink solid (995 mg) in 79% yield. ^{1}H NMR (500 MHz, DMSO) δ 11.16 (s, 1H), 7.36 (dd, *J* = 8.1, 0.9 Hz, 1H), 7.23 (s, 1H), 7.15 (dd, *J* = 7.5, 0.9 Hz, 1H), 6.95 (t, *J* = 7.8 Hz, 1H), 2.97 (dd, 2H), 2.84 (dd, *J* = 7.9, 6.5 Hz, 2H), 1.42 (s, 2H). Spectrum matched previously reported spectra.

5-ethoxytryptamine

W349K *Rgn*TDC was used as 0.02% mol catalyst for the reaction with 5-ethoxyindole. 5-ethoxytryptamine was isolated as a tan solid (353 mg) in 27% yield. ^{1}H NMR (500 MHz, DMSO) δ 10.59 (s, 1H), 7.21 (d, *J* = 8.7 Hz, 1H), 7.07 (d, *J* = 2.3 Hz, 1H), 6.98 (d, *J* = 2.4 Hz, 1H), 6.70 (dd, *J* = 8.7, 2.4 Hz, 1H), 4.01 (q, *J* = 6.9 Hz, 2H), 2.80 (t, *J* = 6.9 Hz, 2H), 2.71 (t, *J* = 7.1 Hz, 2H), 1.42 (s, 2H), 1.34 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (126 MHz, DMSO) δ 152.49, 131.92, 128.15, 123.71, 112.81, 112.34, 111.86, 101.75, 63.84, 43.16, 30.08, 15.42.

MS/ESI m/z for $[M+H]^{+}$; C12H16N2O; calculated 205.1335, observed 205.1333.

5-methoxytryptamine

W349K *Rgn*TDC was used as 0.02% mol catalyst for the reaction with 5-methoxyindole. 5-methoxytryptamine was isolated as a dark brown solid (387 mg) in 30% yield, which co-purified with 70 mg 5-methoxytryptophan. ^{1}H NMR (500 MHz, DMSO) δ 10.60 (s, 1H), 7.22 (d, *J* = 8.7 Hz, 1H), 7.08 (d, *J* = 2.4 Hz, 1H), 6.99 (d, *J* = 2.4 Hz, 1H), 6.71 (dd, *J* = 8.7, 2.5 Hz, 1H), 3.76 (s, 3H), 2.81 (t, *J* = 7.2 Hz, 2H), 2.72 (t, *J* = 7.1 Hz, 2H), 1.68 (s, 2H). Spectrum matched previously reported spectra.

6-chlorotryptamine

F98V *Rgn*TDC was used as 0.02% mol catalyst for the reaction with 6-chloroindole. 6-chlorotryptamine was isolated as an off-white solid (144 mg) in 50% yield. ^{1}H NMR (500 MHz, DMSO) δ 10.97 (s, 1H), 7.52 (d, *J* = 8.4 Hz, 1H), 7.37 (d, *J* = 1.9 Hz, 1H), 7.19 (d, *J* = 2.2 Hz, 1H), 6.98 (dd, *J* = 8.4, 1.9 Hz, 1H), 2.84 (t, *J* = 14.5 Hz, 2H), 2.77 (t, *J* = 7.1 Hz, 2H), 1.85 (s, 3H). Spectrum matched previously reported spectra.

6-nitrotryptamine

F98V *Rgn*TDC was used as 0.02% mol catalyst for the reaction with 6-nitroindole. **Purification was carried out as described above until the solution was basified. The bright yellow solution was turned bright red upon addition of NaOH, and a yellow precipitate was observed. The precipitate was found to be insoluble in both diethyl ether and EtOAc and was filtered from both the aqueous and organic layers. Upon washing with cold water and diethyl ether, a yellow solid was obtained. The remaining aqueous solution was extracted 3x with diethyl ether, resulting in a pale-yellow organic layer. The organic extraction was evaporated dry, and the resulting yellow solid was combined with the filtered solid via resuspension in water. The resuspension was flash frozen and lyophilization was carried out to remove water.** 6-nitrotryptamine was isolated as an initially brown powder (which after 2 weeks turned yellow) (495.5 mg) in 47% yield. Interestingly, this product produces a deep red color when dissolved in solution, which seems to be indicative of the amine group being deprotonated. ^{1}H NMR (500 MHz, ACN) δ 8.32 (d, *J* = 2.2 Hz, 1H), 7.81 (dd, *J* = 8.8, 2.1 Hz, 1H), 7.57 (d, *J* = 8.8 Hz, 1H), 7.52 (s, 1H), 2.90 – 2.82 (m, 4H). Spectrum matched previously reported spectra.

NMR spectra

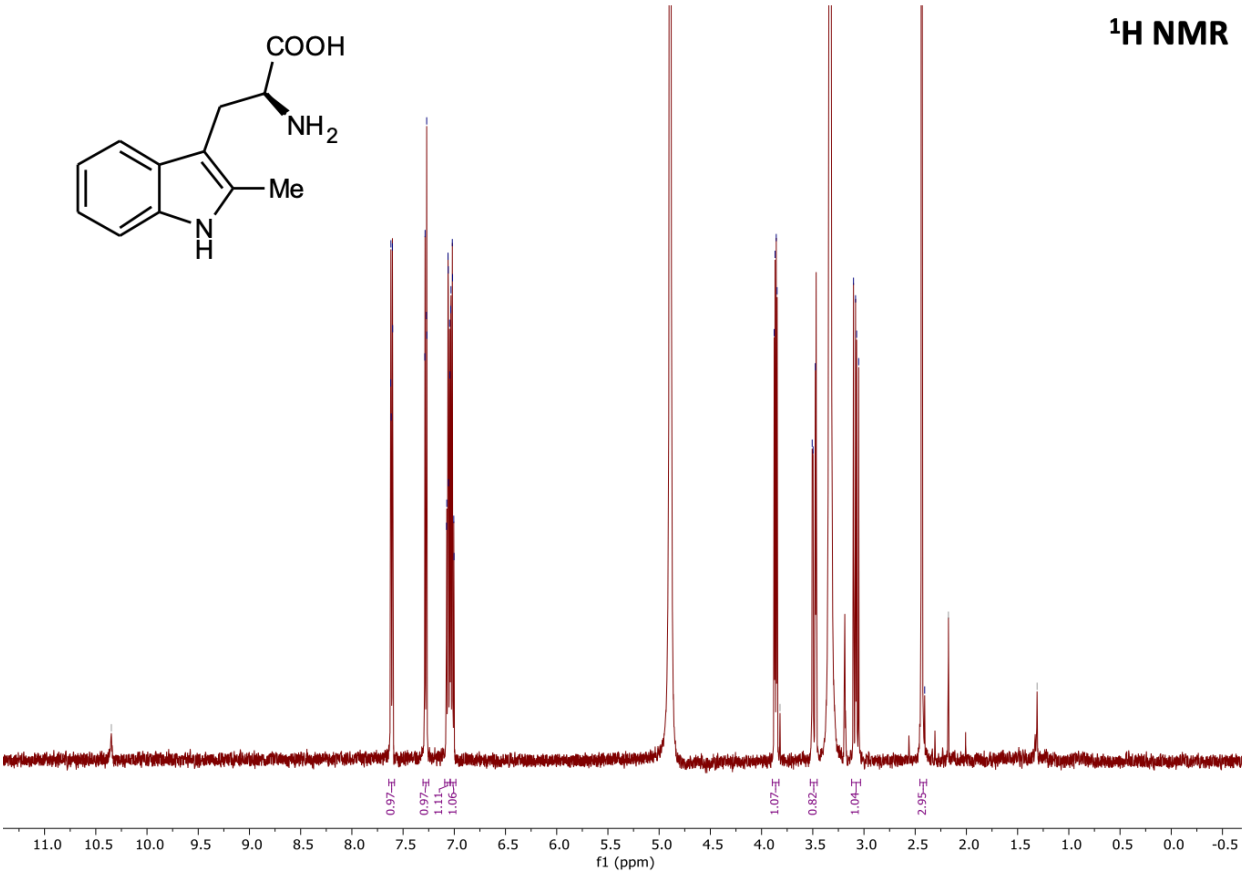


Figure 23. ^{1}H NMR spectrum of (S)-2-amino-3-(2-methyl-1H-indol-3-yl)propanoic acid.

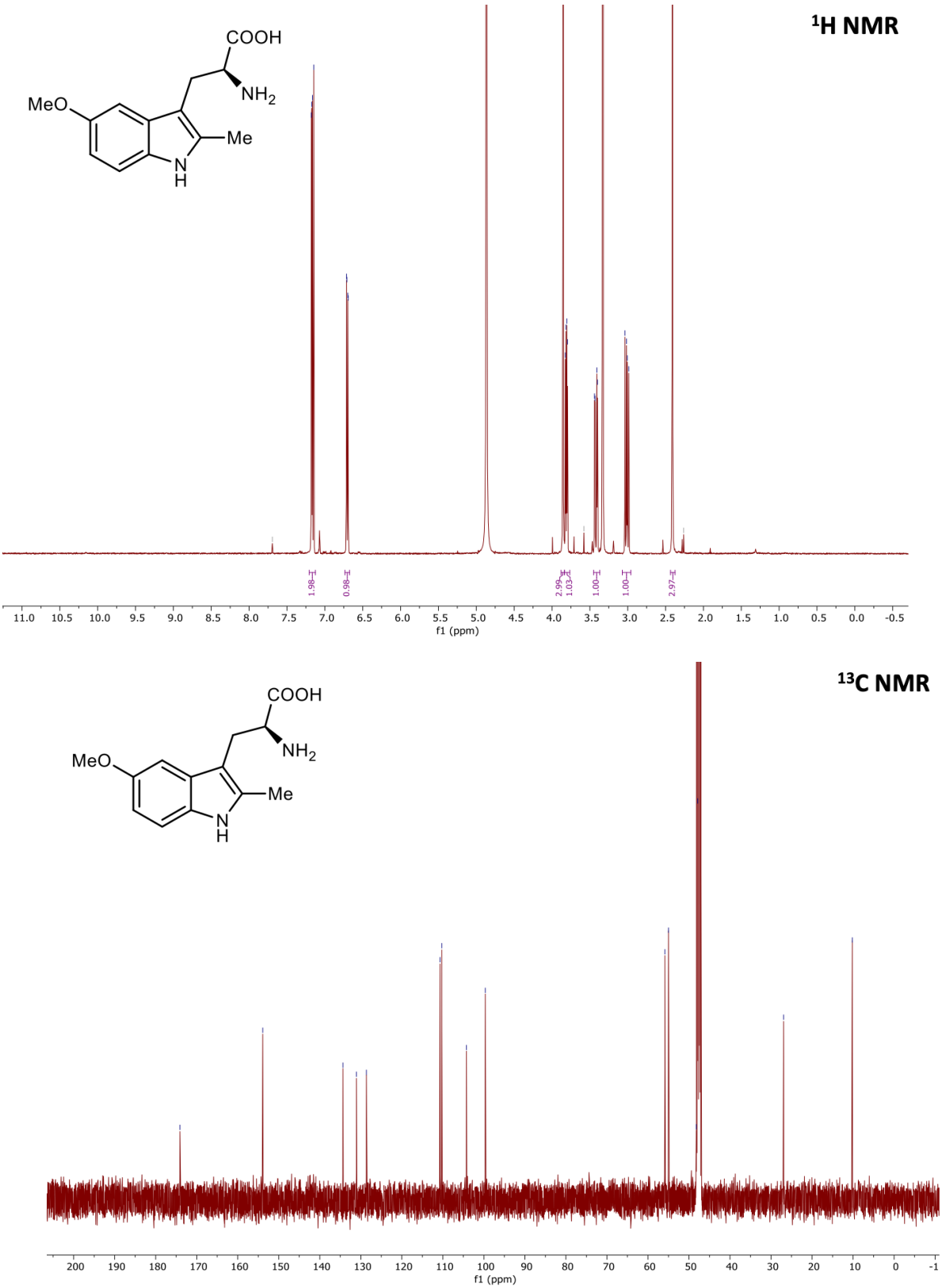


Figure 24. ^{1}H and ^{13}C NMR spectra of (S)-2-amino-3-(5-methoxy-2-methyl-1H-indol-3-yl)propanoic acid.

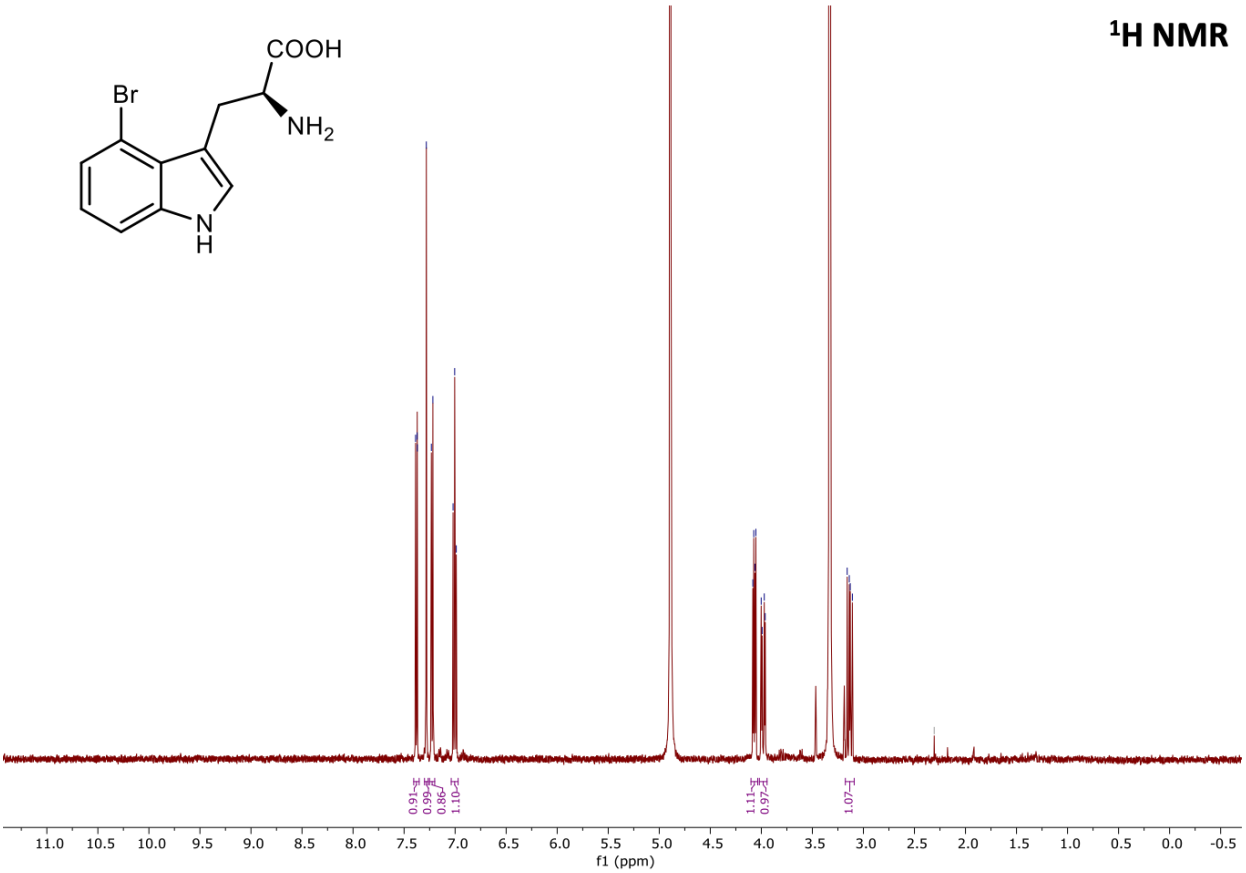
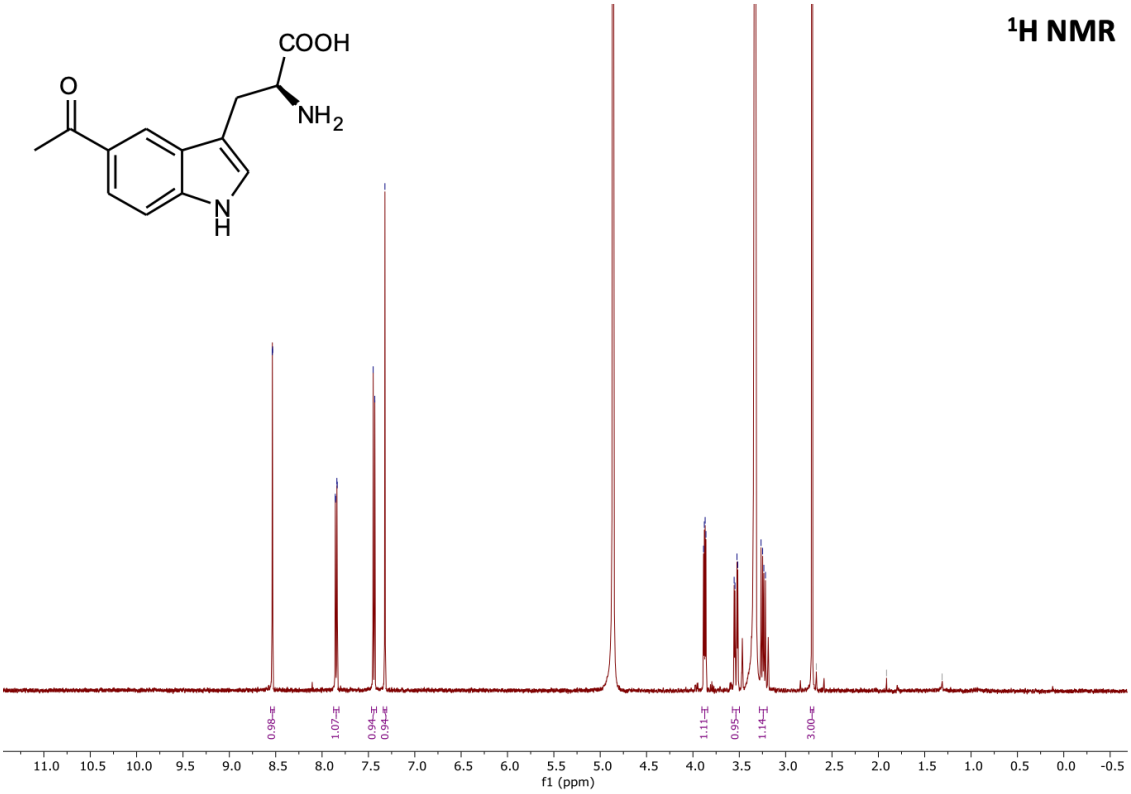


Figure 25. ^{1}H NMR spectrum of (S)-2-amino-3-(4-bromo-1H-indol-3-yl)propanoic acid.



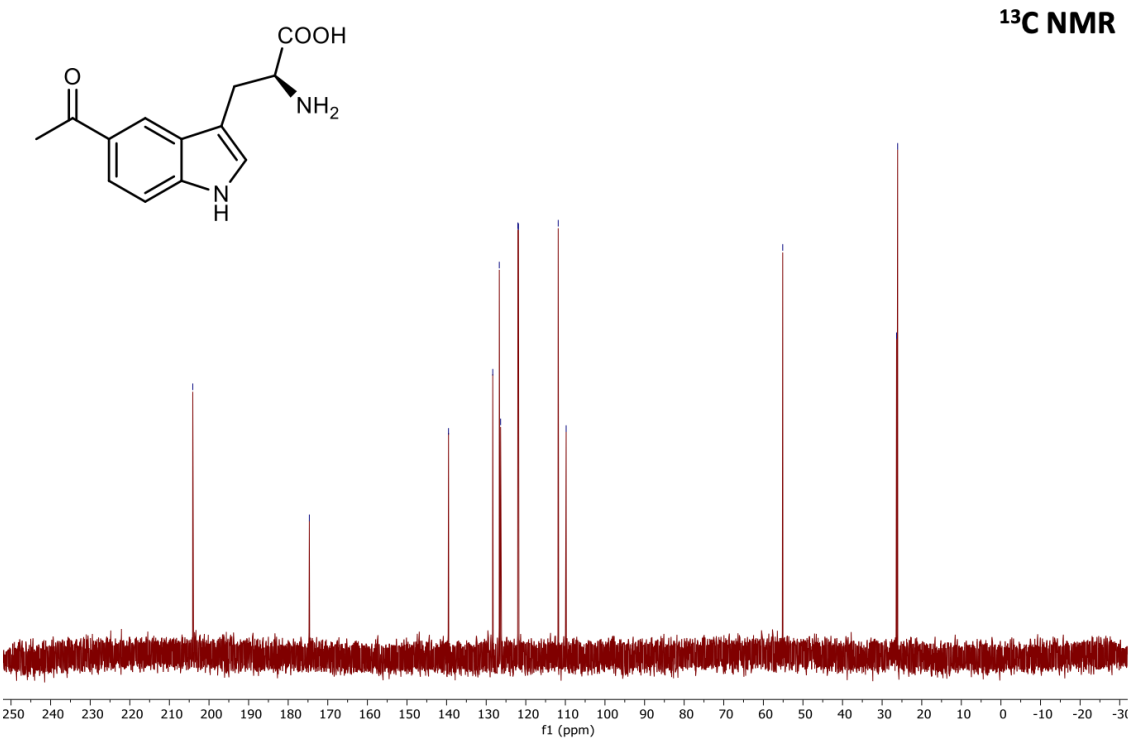


Figure 26. ^{1}H and ^{13}C NMR spectra of (S)-3-(5-acetyl-1H-indol-3-yl)-2-aminopropanoic acid.

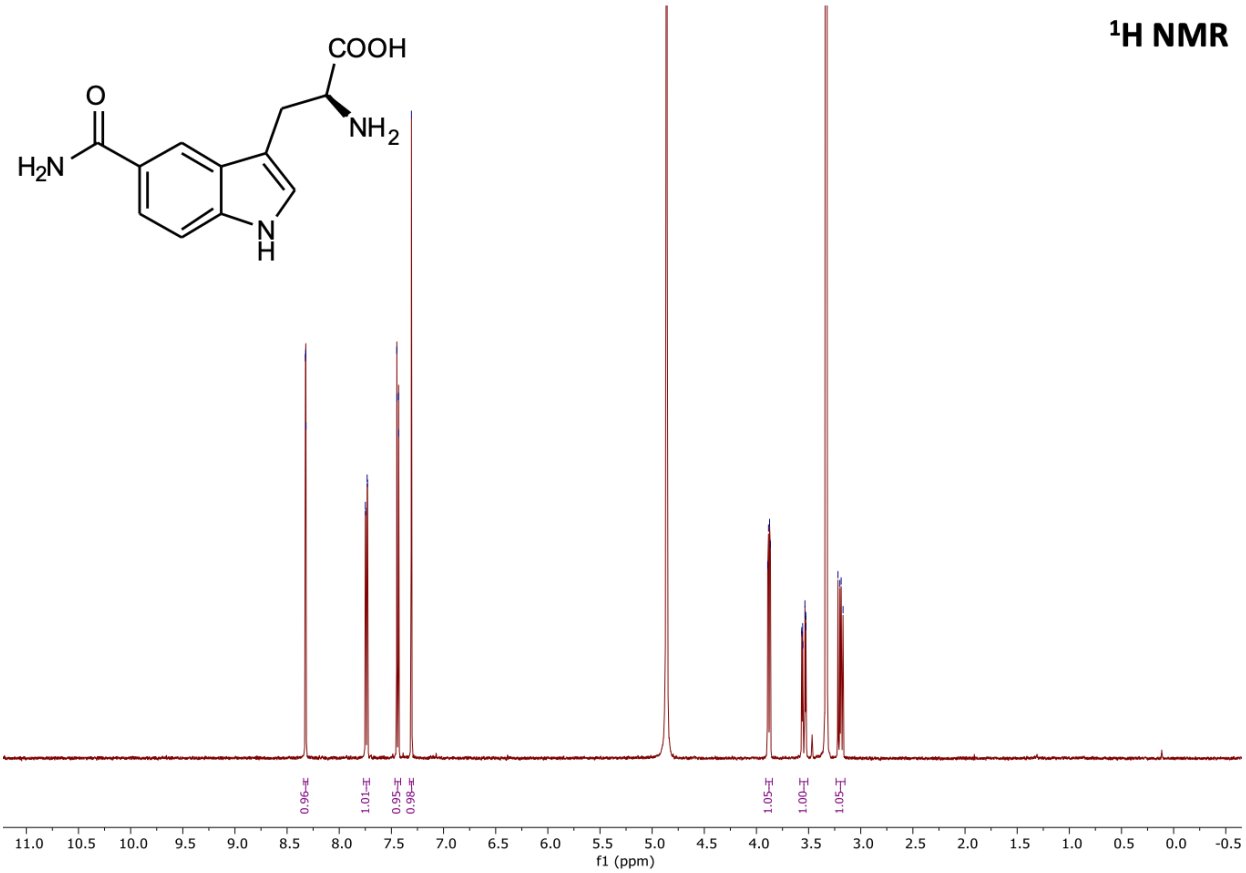
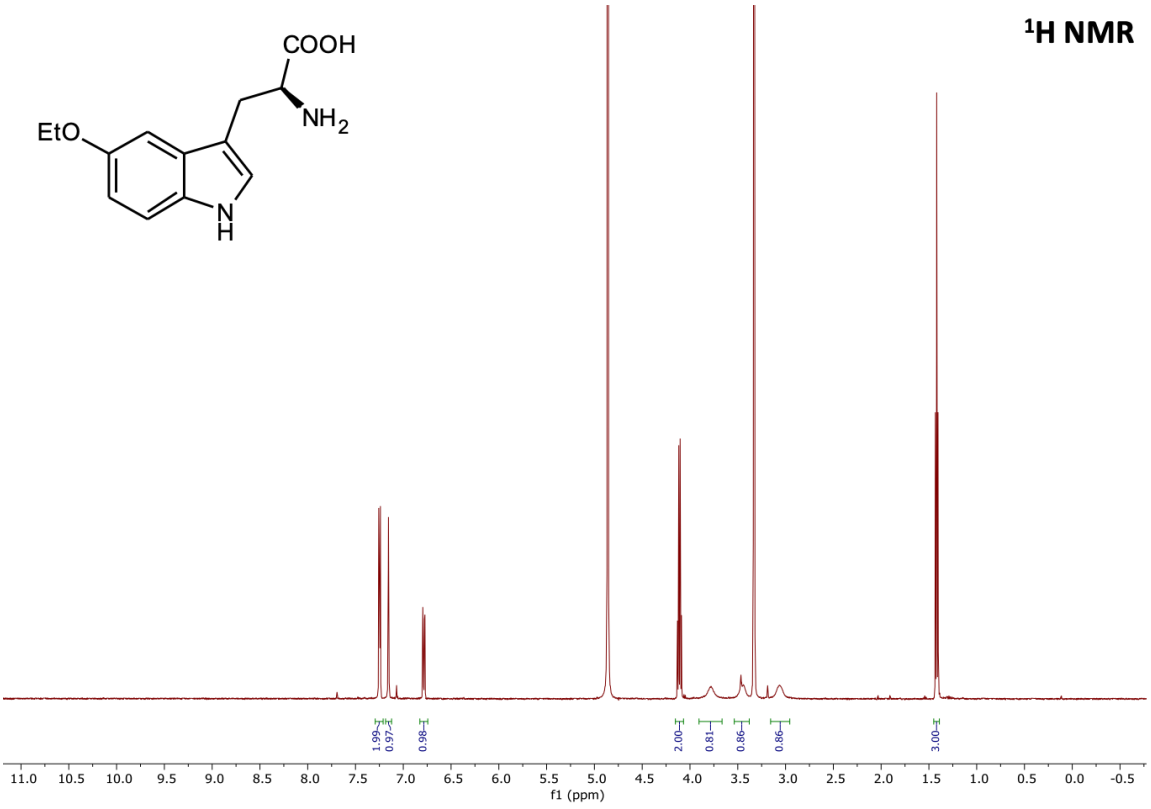


Figure 27. ^{1}H NMR spectrum of (S)-2-amino-3-(5-carbamoyl-1H-indol-3-yl)propanoic acid.


COOH
NH2
EtO
N
H
1H NMR
1.99
0.97
0.98
2.00
0.81
0.86
0.86
3.00
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

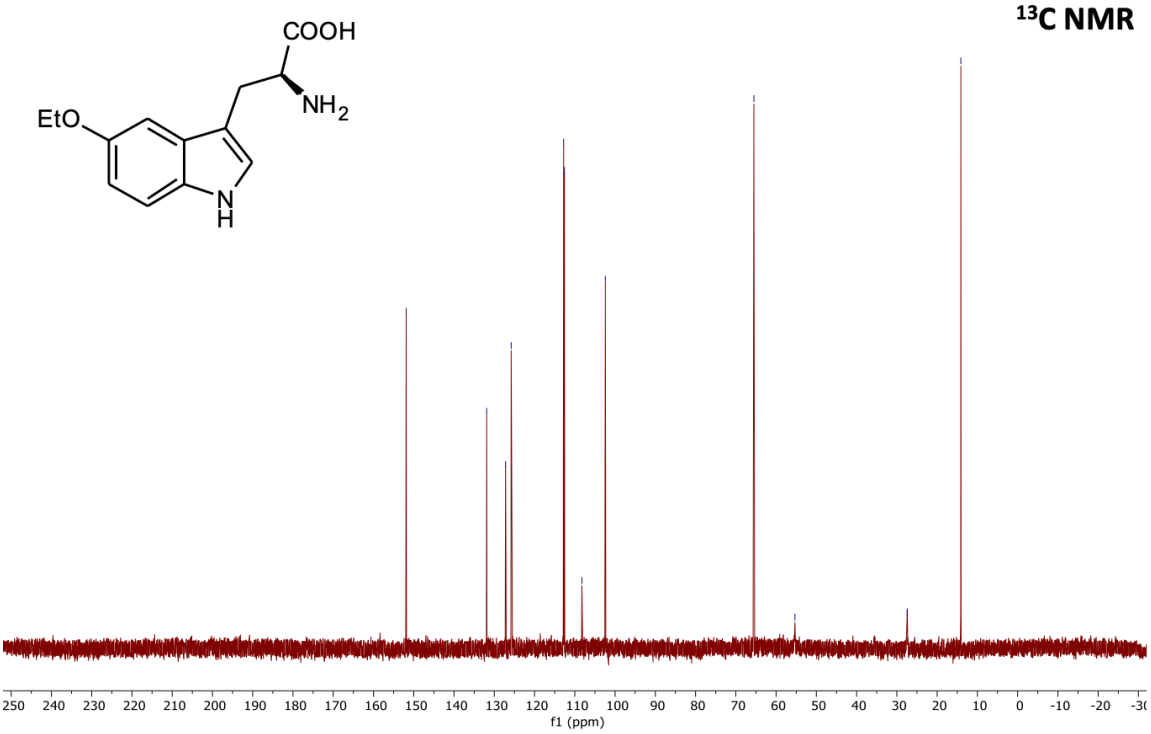

COOH
NH2
EtO
N
H
13C NMR
250 240 230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20
f1 (ppm)

Figure 28. ^{1}H and ^{13}C NMR spectra of (S)-2-amino-3-(5-ethoxy-1H-indol-3-yl)propanoic acid.

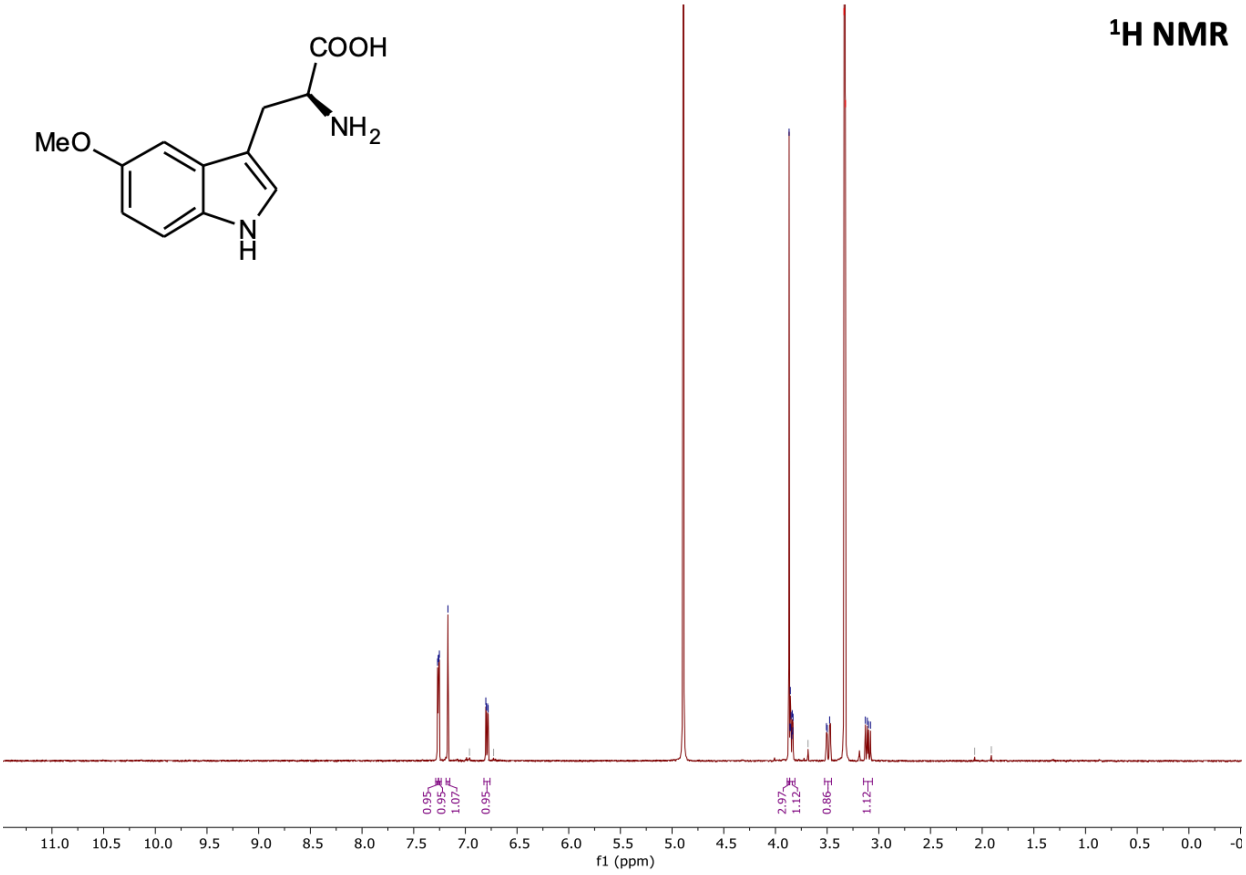


Figure 29. ^{1}H NMR spectrum of (S)-2-amino-3-(5-methoxy-1H-indol-3-yl)propanoic acid.

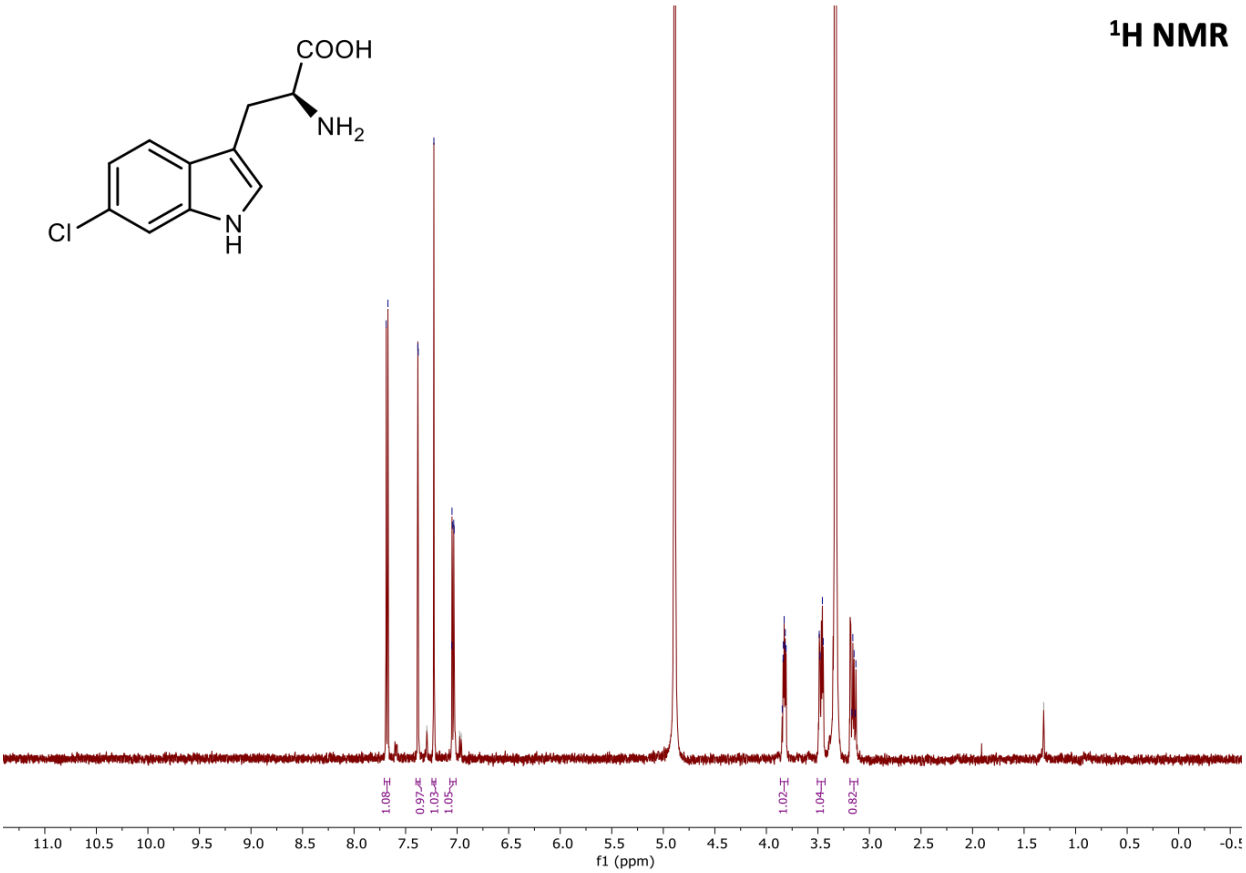


Figure 30. ^{1}H NMR spectrum of (S)-2-amino-3-(6-chloro-1H-indol-3-yl)propanoic acid.

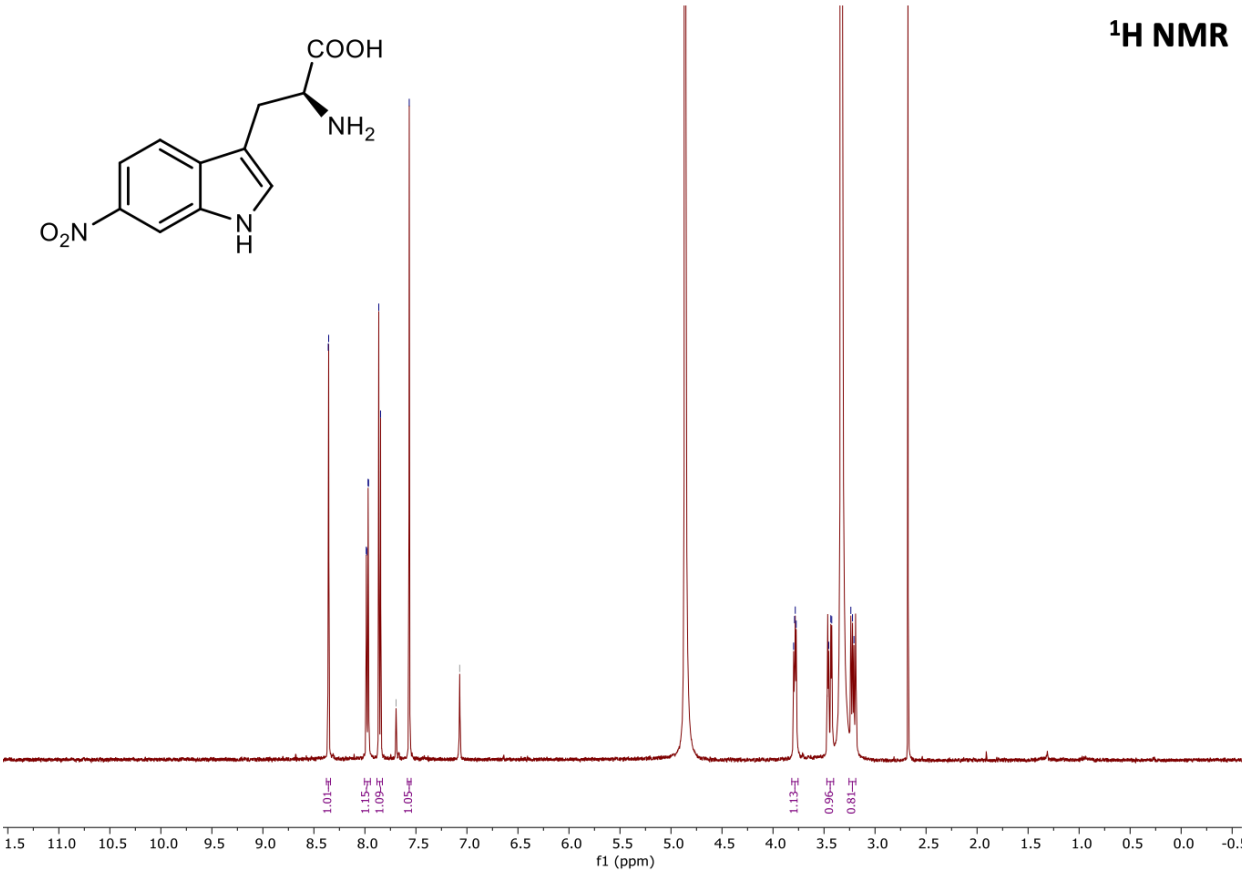


Figure 31. ^{1}H NMR spectrum of (S)-2-amino-3-(6-nitro-1H-indol-3-yl)propanoic acid.

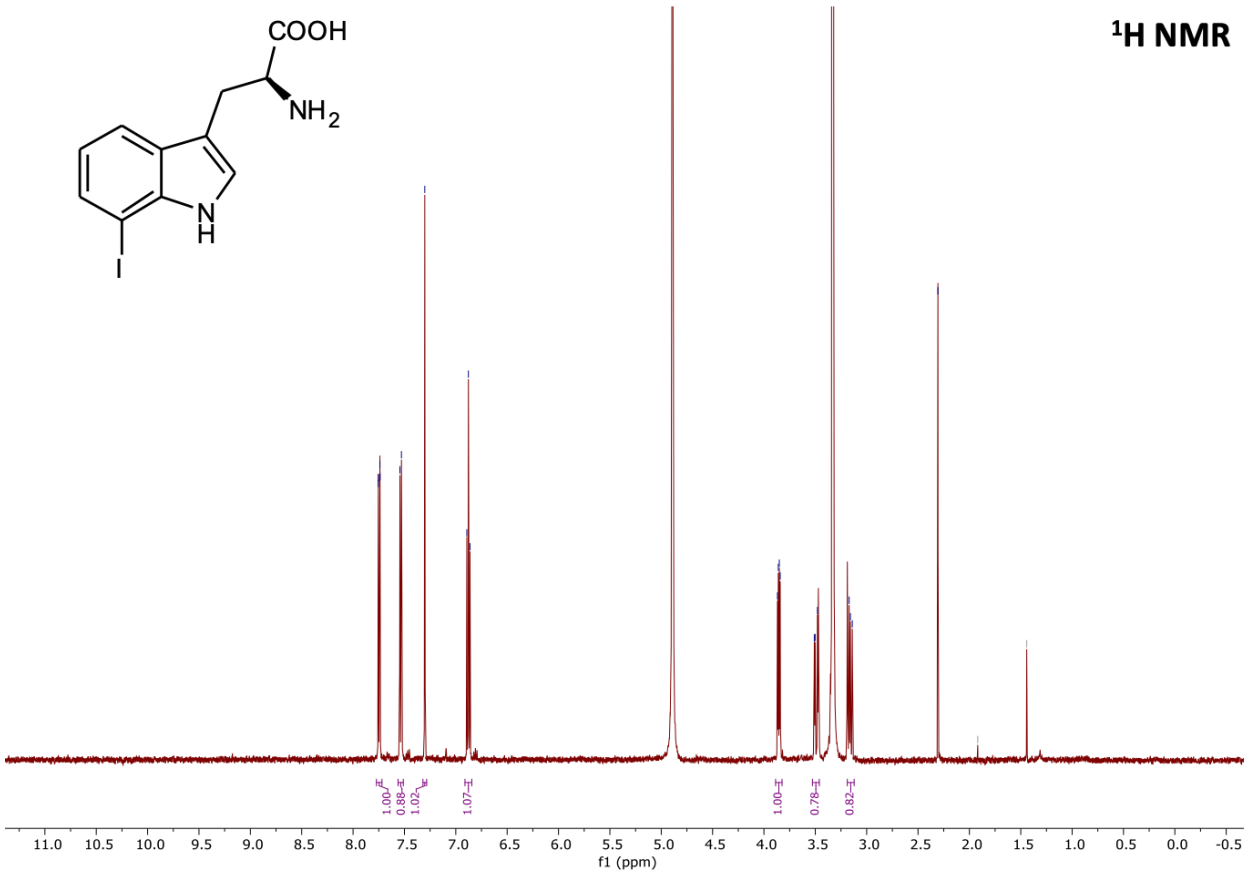


Figure 32. ^{1}H NMR spectrum of (S)-2-amino-3-(7-iodo-1H-indol-3-yl)propanoic acid.

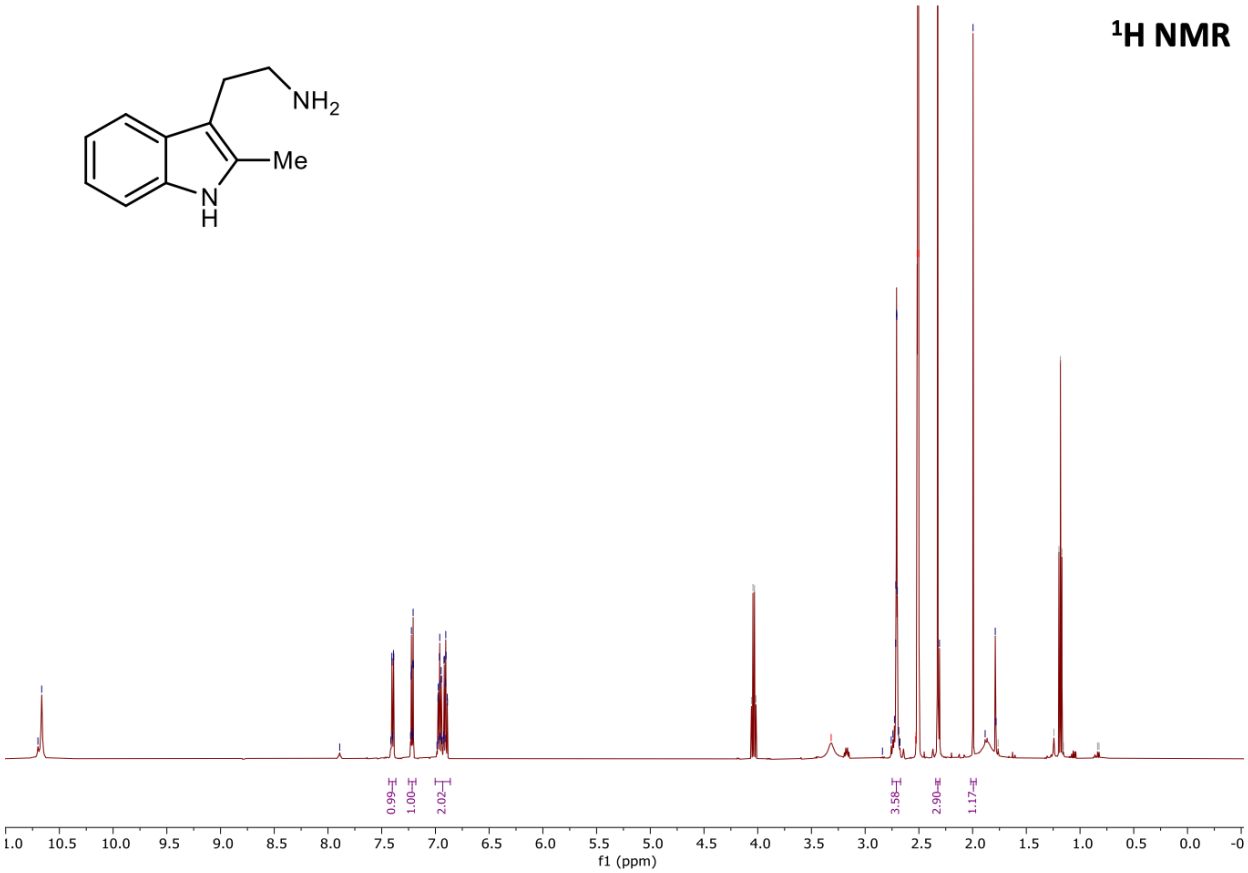


Figure 33. ^{1}H NMR spectrum of 2-(2-methyl-1H-indol-3-yl)ethan-1-amine.

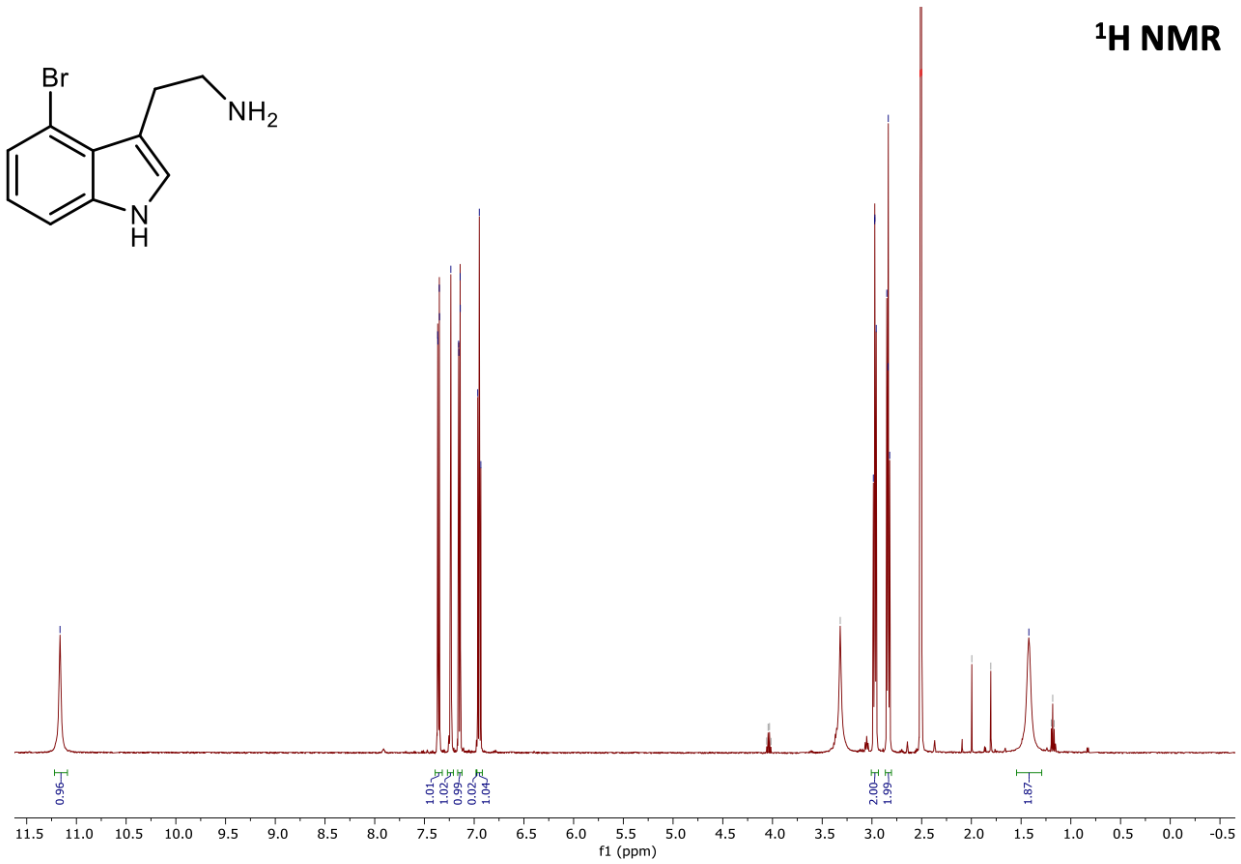


Figure 34. ^{1}H NMR spectrum of 2-(4-bromo-1H-indol-3-yl)ethanamine.

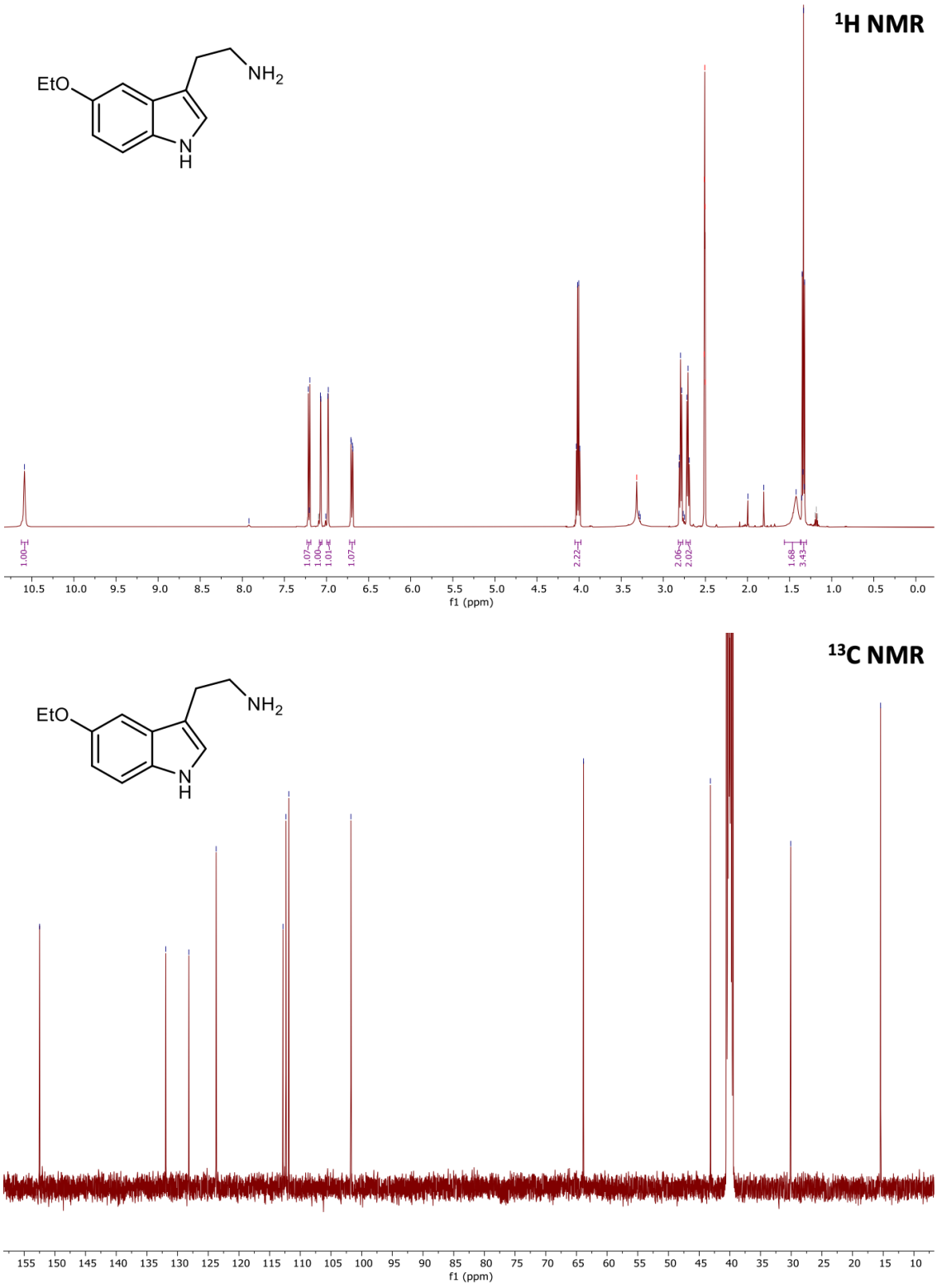


Figure 35. ^{1}H and ^{13}C NMR spectra of 2-(5-ethoxy-1H-indol-3-yl)ethanamine.

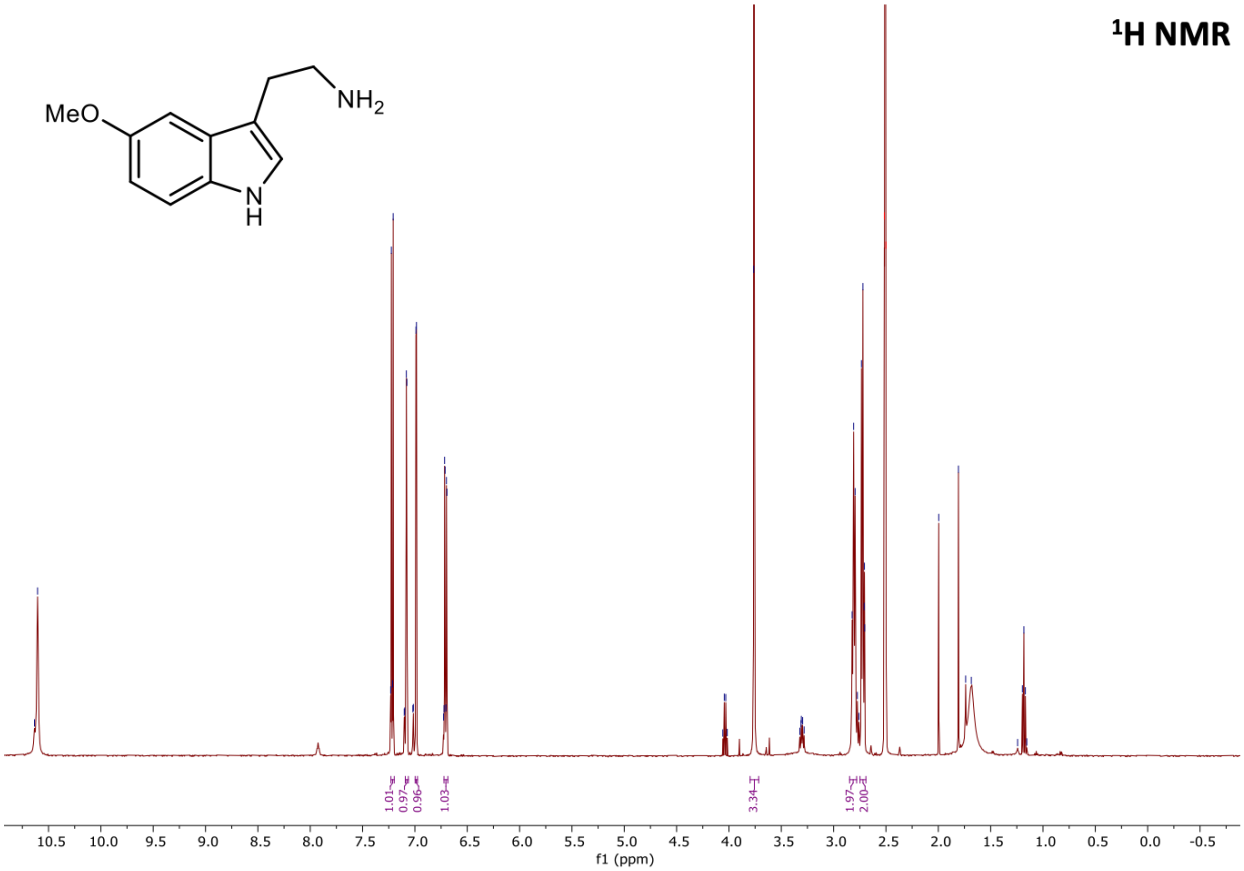


Figure 36. ^{1}H NMR spectrum of 2-(5-methoxy-1H-indol-3-yl)ethanamine.

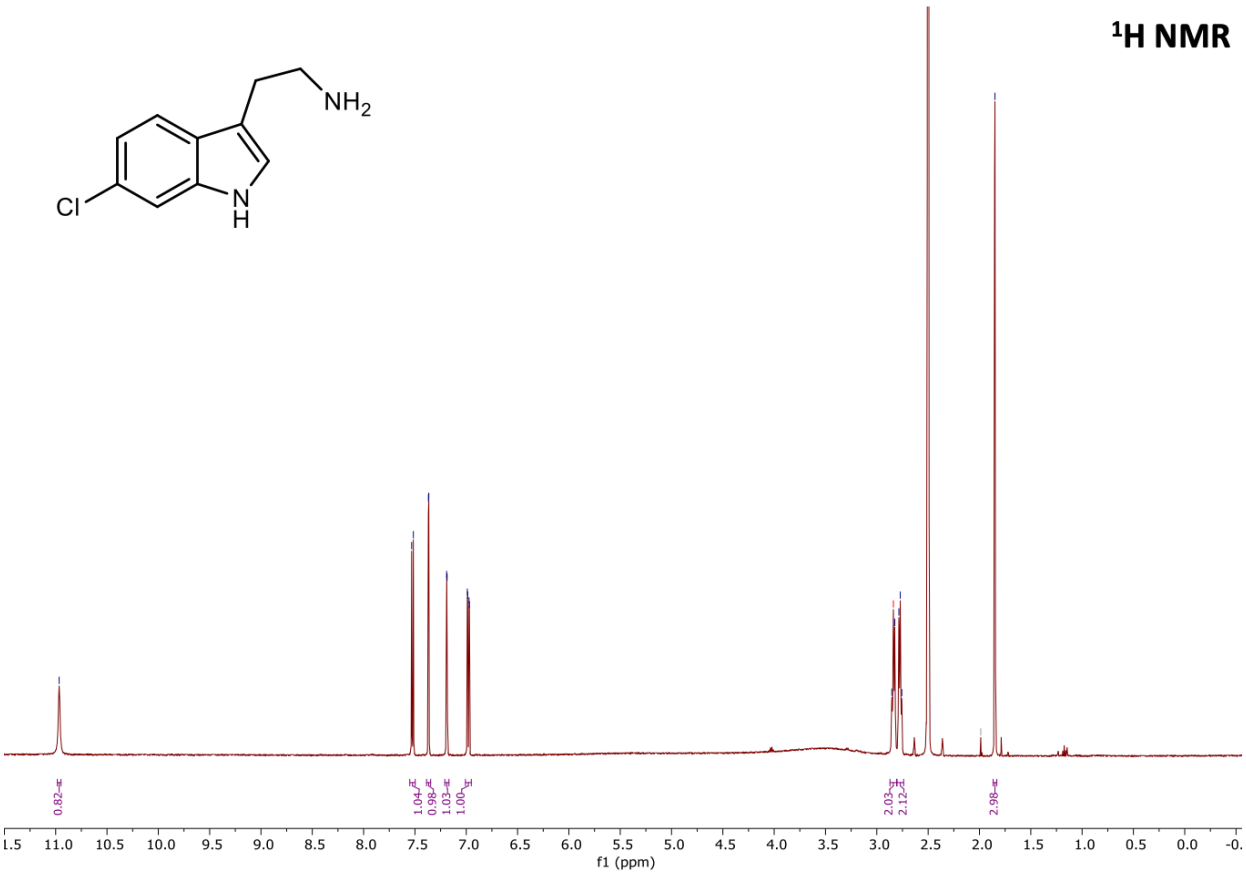


Figure 37. ^{1}H NMR spectrum of 2-(6-chloro-1H-indol-3-yl)ethan-1-amine.

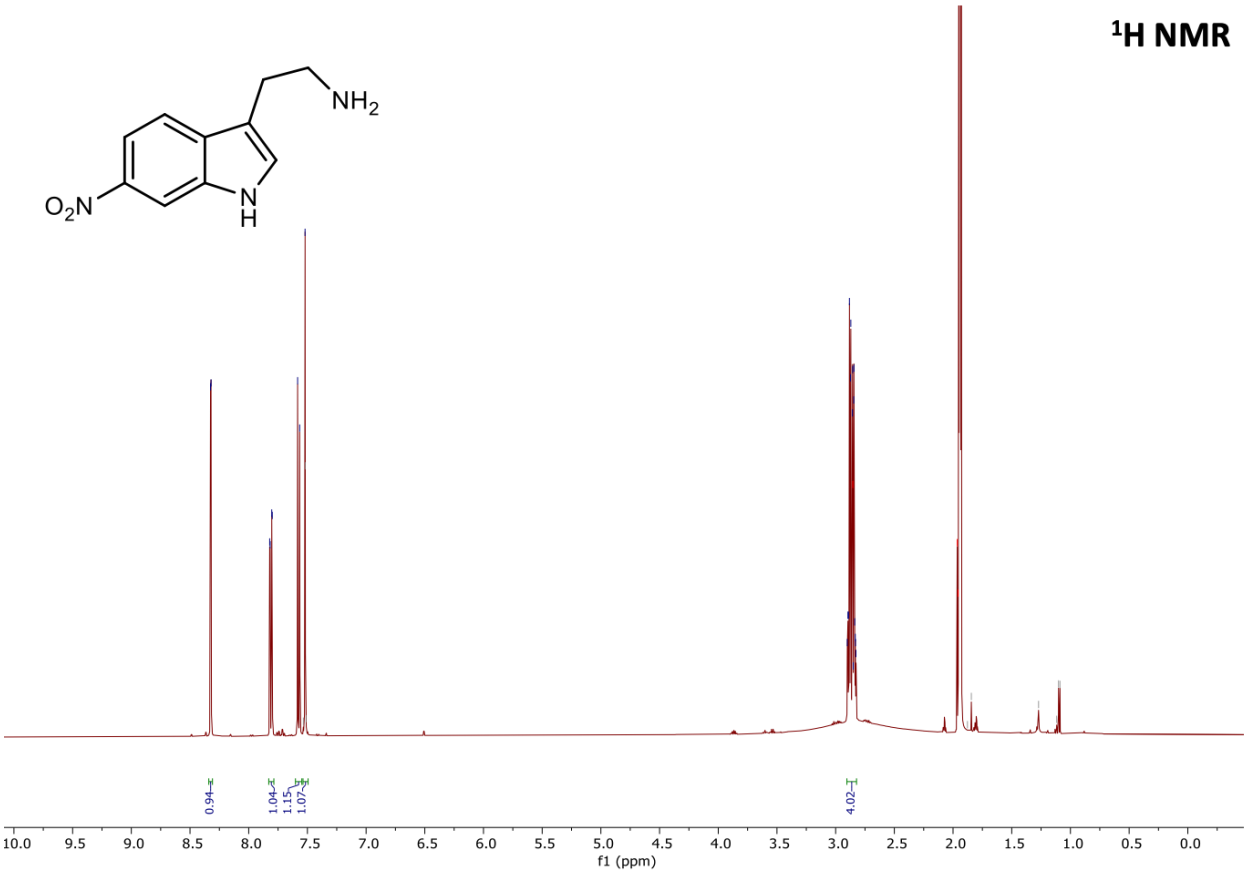


Figure 38. ^{1}H NMR spectrum of 2-(6-nitro-1H-indol-3-yl)ethan-1-amine.

3. 5. References

1. Goodwin, N. C., Morrison, J. P., Fuerst, D. E. & Hadi, T. Biocatalysis in Medicinal Chemistry: Challenges to Access and Drivers for Adoption. *ACS Med. Chem. Lett.* **10**, 1363–1366 (2019).

2. Savile, C. K. *et al.* Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science (80-.).* **329**, 305–310 (2010).

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

4. Huffman, M. A. *et al.* Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science (80-.).* **366**, 1255–1259 (2019).

5. Sandström, A. G., Wikmark, Y., Engström, K., Nyhlén, J. & Bäckvall, J. E. Combinatorial reshaping of the Candida antarctica lipase A substrate pocket for enantioselectivity using an extremely condensed library. *PNAS* **109**, 78–83 (2012).

6. Fox, R. J. *et al.* Improving catalytic function by ProSAR-driven enzyme evolution. *Nat. Biotechnol.* **25**, 338–344 (2007).

7. Amin, N. *et al.* Construction of stabilized proteins by combinatorial consensus mutagenesis. *Protein Eng. Des. Sel.* **17**, 787–793 (2004).

8. Reetz, M. T., Bocola, M., Carballeira, J. D., Zha, D. & Vogel, A. Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test. *Angew. Chemie - Int. Ed.* **44**, 4192–4196 (2005).

9. Diefenbach, X. W. *et al.* Enabling Biocatalysis by High-Throughput Protein Engineering Using Droplet Microfluidics Coupled to Mass Spectrometry. *ACS Omega* **3**, 1498–1508 (2018).

10. McLaren, D. G. *et al.* High-Throughput Mass Spectrometry for Hit Identification: Current Landscape and Future Perspectives. *SLAS Discov.* **26**, 168–191 (2021).

11. Ye, L., Yang, C. & Yu, H. From molecular engineering to process engineering: development of high-throughput screening methods in enzyme directed evolution. *Appl. Microbiol. Biotechnol. 2017 1022* **102**, 559–567 (2017).

12. Obexer, R. *et al.* Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. *Nat. Chem.* **9**, 50–56 (2017).

13. Herger, M. *et al.* Synthesis of β-Branched Tryptophan Analogues Using an Engineered Subunit of Tryptophan Synthase. *J. Am. Chem. Soc.* **138**, 8388–8391 (2016).

14. Tang, Q. *et al.* Directed Evolution of a Halide Methyltransferase Enables Biocatalytic Synthesis of Diverse SAM Analogs. *Angew. Chemie - Int. Ed.* **60**, 1524–1527 (2021).

15. 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.* **7**, 3720–3729 (2016).

16. 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. *J. Am. Chem. Soc.* **139**, 10769–10776 (2017).

17. Romney, D. K., Sarai, N. S. & Arnold, F. H. Nitroalkanes as Versatile Nucleophiles for Enzymatic Synthesis of Noncanonical Amino Acids. *ACS Catal.* **9**, acscatal.9b02089 (2019).

18. Andorfer, M. C. *et al.* Understanding Flavin-Dependent Halogenase Reactivity via Substrate Activity Profiling. *ACS Catal.* **7**, 1897–1904 (2017).

19. Goddard, J., Reymond, J. & Uni, V. Enzyme Activity Fingerprinting with Substrate Cocktails. *JACS* 11116–11117 (2004).

20. Kim, H. *et al.* A multi-substrate screening approach for the identification of a broadly applicable Diels – Alder catalyst. *Nat. Commun.* 1–6 (2019) doi:10.1038/s41467-019-08374-z.

21. Stanisic, A., Husken, A. & Kries, H. HAMA: a multiplexed LC-MS/MS assay for specificity profiling of adenylate-forming enzymes. *Chem. Sci.* **10**, 10395–10399 (2019).

22. Zhang, K., Damaty, S. El & Fasan, R. P450 Fingerprinting Method for Rapid Discovery of Terpene Hydroxylating P450 Catalysts with Diversified Regioselectivity. *JACS* 3242–3245 (2011).

23. Kuo, Y. M., Henry, R. A. & Andrews, A. J. Measuring specificity in multi-substrate/product systems as a tool to investigate selectivity in vivo. *Biochim. Biophys. Acta - Proteins Proteomics* **1864**, 70–76 (2016).

24. Joiner, C. M., Levine, Z. G., Aonbangkhen, C., Woo, C. M. & Walker, S. Aspartate Residues Far from the Active Site Drive O-GlcNAc Transferase Substrate Selection. *J. Am. Chem. Soc.* **141**, 12974–12978 (2019).

25. Stanisic, A. *et al.* HAMA: A multiplexed LC-MS/MS assay for specificity profiling of adenylate-forming enzymes. *Chem. Sci.* **10**, 10395–10399 (2019).

26. Koryakina, I. *et al.* Poly specific trans -Acyltransferase machinery revealed via engineered acyl-coa synthetases. *ACS Chem. Biol.* **8**, 200–208 (2013).

27. Weeks, A. M. & Wells, J. A. Engineering peptide ligase specificity by proteomic identification of ligation sites. *Nat. Chem. Biol.* **14**, 50–57 (2018).

28. Kalkreuter, E. *et al.* Computationally-guided exchange of substrate selectivity motifs in a modular polyketide synthase acyltransferase. *Nat. Commun.* 1–12 (2021) doi:10.1038/s41467-021-22497-2.

29. Xu, L. *et al.* Mapping enzyme catalysis with metabolic biosensing. *Nat. Commun.* **12**, 1–7 (2021).

30. Si, T. *et al.* Profiling of Microbial Colonies for High-Throughput Engineering of Multistep

Enzymatic Reactions via Optically Guided Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *J. Am. Chem. Soc.* **139**, 12466–12473 (2017).

31. Xue, P. *et al.* A mass spectrometry-based high-throughput screening method for engineering fatty acid synthases with improved production of medium-chain fatty acids. *Biotechnol. Bioeng.* **117**, 2131–2138 (2020).

32. Choe, K., Xue, P., Zhao, H. & Sweedler, J. V. MacroMS: Image-Guided Analysis of Random Objects by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **32**, 1180–1188 (2021).

33. Engstrom, K., Nyhlen, J., Sandstrom, A. G. & Bäckvall, J.-E. Directed Evolution of an Enantioselective Lipase with Broad Substrate Scope for Hydrolysis of r -Substituted Esters. *J. Am. Chem. Soc.* **132**, 7038–7042 (2010).

34. Reetz, M. T., Becker, M. H., Klein, H.-W. & Stöckigt, D. A Method for High-Throughput Screening of Enantioselective Catalysts. *Angew. Chemie* **38**, 1758–1760 (1999).

35. Ma, F. *et al.* Efficient molecular evolution to generate enantioselective enzymes using a dual-channel microfluidic droplet screening platform. *Nat. Commun.* **9**, 1–8 (2018).

36. Jakoblinnert, A. *et al.* Reengineered carbonyl reductase for reducing methyl-substituted cyclohexanones. *Protein Eng. Des. Sel.* **26**, 291–298 (2013).

37. Junker, S., Roldan, R., Joosten, H. J., Clapés, P. & Fessner, W. D. Complete Switch of Reaction Specificity of an Aldolase by Directed Evolution In Vitro: Synthesis of Generic Aliphatic Aldol Products. *Angew. Chemie - Int. Ed.* **57**, 10153–10157 (2018).

38. Knorrscheidt, A. *et al.* Simultaneous screening of multiple substrates with an unspecific peroxygenase enabled modified alkane and alkene oxyfunctionalisations. *Catal. Sci. Technol.* 6058–6064 (2021) doi:10.1039/d0cy02457k.

39. Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*. (W. H. Freeman and Company, 1999).

40. Fricke, J., Blei, F. & Hoffmeister, D. Enzymatic Synthesis of Psilocybin. *Angew. Chemie - Int. Ed.* **56**, 12352–12355 (2017).

41. Runguphan, W., Qu, X. & O'Connor, S. E. Integrating carbon–halogen bond formation into medicinal plant metabolism. *Nature* **468**, 461–464 (2010).

42. Williams, B. B. *et al.* Discovery and characterization of gut microbiota decarboxylases that can produce the neurotransmitter tryptamine. *Cell Host Microbe* **16**, 495–503 (2014).

43. Cornish-Bowden, A. Enzyme Specificity : Its Meaning in the General Case. *J. theor. Biol.* **108**, 451–457 (1984).

44. Chou, T. C. & Talalay, P. A simple generalized equation for the analysis of multiple inhibitions of Michaelis-Menten kinetic systems. *J. Biol. Chem.* **252**, 6438–6442 (1977).

45. Schrittwieser, J. H., Velikogne, S., Hall, M. & Kroutil, W. Artificial Biocatalytic Linear

Cascades for Preparation of Organic Molecules. *Chemical Reviews* vol. 118 270–348 (2018).

46. Sattler, J. H. *et al.* Redox Self-Sufficient Biocatalyst Network for the Amination of Primary Alcohols. *Angew. Chemie* **51**, 9156–9159 (2012).

47. Staudt, S. *et al.* Direct Oxidation of Cycloalkanes to Cycloalkanones with Oxygen in Water. *Angew. Chemie* **52**, 2359–2363 (2013).

48. Schrittwieser, J. H., Velikogne, S., Hall, M. & Kroutil, W. Artificial Biocatalytic Linear Cascades for Preparation of Organic Molecules. *Chem. Rev.* **118**, 270–348 (2018).

49. Glennon, R. A. Higher-End Serotonin Receptors: 5-HT5, 5-HT6, and 5-HT7. *J. Med. Chem.* **46**, 2795–2812 (2003).

50. Khersonsky, O. & Tawfik, D. S. Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annual Review of Biochemistry* vol. 79 471–505 (2010).

51. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

52. Kille, S. *et al.* Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2**, 83–92 (2013).

53. TerMaat, J. R., Pienaar, E., Whitney, S. E., Mamedov, T. G. & Subramanian, A. Gene synthesis by integrated polymerase chain assembly and PCR amplification using a high-speed thermocycler. *J. Microbiol. Methods* **79**, 295–300 (2009).

54. Johnson, K. A. New standards for collecting and fitting steady state kinetic data. *Beilstein J. Org. Chem.* **15**, 16–29 (2019).

55. Buller, A. R. *et al.* Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. *Proc. Natl. Acad. Sci.* **112**, 14599–14604 (2015).

56. 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. *J. Am. Chem. Soc.* **139**, 10769–10776 (2017).

57. Murciano-Calles, J., Romney, D. K., Brinkmann-Chen, S., Buller, A. R. & Arnold, F. H. A Panel of TrpB Biocatalysts Derived from Tryptophan Synthase through the Transfer of Mutations that Mimic Allosteric Activation. *Angew. Chemie - Int. Ed.* **55**, 11577–11581 (2016).

Chapter 4

Engineering complementary substrate scopes using substrate multiplexed screening

Aadhishre Kasat and Samantha Bruffy were both longstanding collaborators on this work, and their contributions, both experimentally and scientifically, have been vital for this project.

Chapter 4: Engineering complementary substrate scopes using substrate multiplexed screening

4. 1. Introduction

One-pot biocatalytic cascades are sought-after for practical syntheses since they obviate the need for intermediate isolation and reduce waste. Such multi-catalyst systems can also improve yield by coupling thermodynamically unfavorable reactions to favorable ones and limiting kinetic competition from unfavorable reaction pathways.[1] Due to the high chemoselectivity of enzymes, side reactivity is often low, leading to effective and high-yielding enzymatic cascades. For these reasons, biocatalytic cascades employing enzymes such as lipases, dehydrogenases, or transaminases have been used successfully for a wide variety of syntheses,[1,2] including diverse process scale reactions.[3–5] Notably, many of these enzymatic cascades are devoid of biocatalytic C-C bond forming steps.[4,6–10] Although C-C bond forming reactions are especially valuable targets for biocatalytic cascades, enzymatic cascades to form C-C bonds often suffer from poor substrate scopes.[11–13] With the notable exception of cofactor regeneration, biocatalytic cascades are seldom deployed in ‘generalist’ catalytic modality, where a single cascade is capable of making diverse products. This challenge is seldom met because each enzyme must have sufficient activity on a complementary set of substrates for efficient cascade flux (Fig 1a top). The time and significant resources required to engineer enzymes to operate on new substrates means that, often, bespoke enzymes are generated for each target cascade. This limitation in scope, in turn, hinders the wider adoption of enzymes in organic synthesis. Therefore, development of methods to better control the scope of biocatalysts, particularly those that enable complexity-generating steps to assemble and tailor the carbon backbone, would have a significant impact.

Substrate multiplexed screening (SUMS) is a methodology that may be suited to tackling the challenge of poor substrate scope complementarity between enzymes in a cascade. As previously described in Chapter 3, this method involves screening for activity on multiple

substrates in competition to directly read out on biocatalyst promiscuity (Fig 1a bottom). In this implementation, SUMS relies on selecting diverse substrates with unique product m/z ratios, enabling simultaneous quantitation. Thus, information on both activity on individual substrates and specificity can be derived from a single assay, reducing uncertainties from the guess-and-

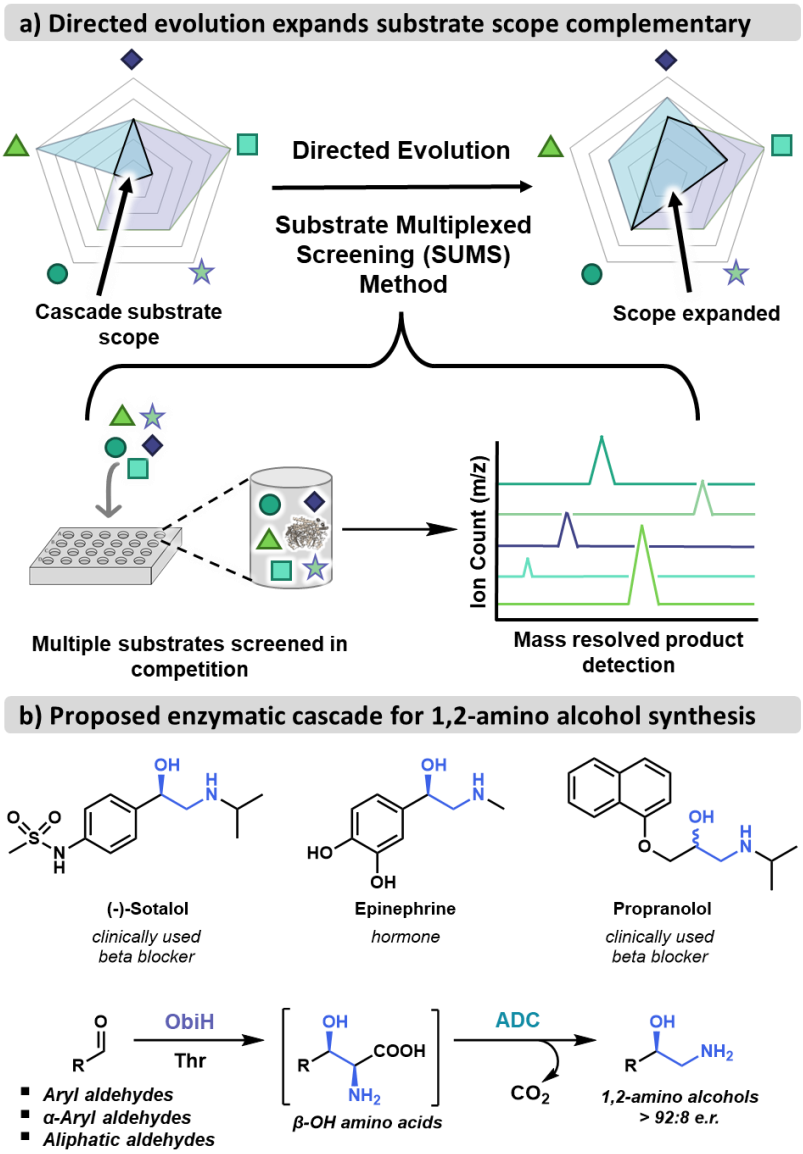


Figure 1. Substrate multiplexed screening (SUMS) has advantages for cascade engineering. a) Screening with multiple substrates can more efficiently expand substrate scopes. **b)** 1,2-amino alcohols are a valuable bioactive motif that can be accessed via the proposed ObiH-ADC (amino acid decarboxylase) cascade.

check workflow of engineering promiscuous enzymes. Therefore, SUMS is positioned as a useful methodology for the development of efficient enzymatic cascades with non-native substrates.

Due to the wide chemical diversity of clinically used 1,2-amino alcohols, most notably in β-adrenergic receptor agonists (β-blockers, Fig 1b), there is continued motivation to expand routes to this class of compounds.[14] Further, their relative simplicity serves as a fertile testing ground for promiscuous biocatalytic synthesis.[15–18] Sehl et al. described a cascade that produced norephedrine stereoisomers in high yields from benzaldehyde using a lyase and transaminase.[13,19] Other C-C bond forming cascades producing 1,2-amino alcohols, such as using aminocyclases or imine reductases, have been successful, but again were limited in substrate diversity.[20,21] Griengl and coworkers developed an enzymatic cascade to access enantio-enriched 1,2-amino alcohols by coupling a promiscuous threonine aldolase with a tyrosine decarboxylase.[22,23] However, the utility of the cascade was undermined due to the limited activity of the decarboxylase, allowing access only to aryl 1,2-amino alcohols.

We sought to develop a robust and highly generalizable C-C bond forming biocatalytic cascade to form 1,2-amino alcohols from achiral aldehydes using an L-threonine transaldolase (LTTA) and an amino acid decarboxylase (ADC). Recent work by our lab and others has characterized the activity and specificity of a promiscuous and highly stereoselective LTTA, ObiH.[24–26] ObiH (also known as ObaG) catalyzes the formation of β-OH amino acids from diverse aryl, α-aryl, and aliphatic aldehydes (Fig 2a). However, no known amino acid decarboxylase has been characterized with a scope that encompasses the wide array of β-OH amino acids produced via ObiH. Tyrosine decarboxylase was previously reported to have activity on some aromatic β-OH amino acids, but decarboxylation activity on β-OH amino acids has not been well explored for other classes of ADC's.[22,23] Therefore, SUMS may serve to intentionally select or engineer an ADC for this non-native activity. To demonstrate this potential

of SUMS, here we implement SUMS to improve the substrate scope overlap of a decarboxylase to form 1,2-amino alcohols in a two-enzyme cascade. scope overlap of an ADC in a 1,2-amino alcohol-generating two-enzyme cascade.

The results here show how SUMS can be deployed in a multi-generational evolutionary campaign to efficiently monitor changes in substrate scope during evolution. Over the course of engineering, we report gain-of-function on aliphatic and α-aryl β-OH amino acids for *Rgn*TDC as well as improvements in activity on aryl β-OH amino acids. We then showcase the utility of the engineered cascade for the preparative syntheses of diverse 1,2-amino alcohols with good yields and high e.r.'s. The strategies that are successful here may serve as a prelude to the development of cascades for the construction of more complex molecular scaffolds.

4. 2. Results and Discussion

4. 2. 1. Identification of a suitable ADC

We first investigated two ADC's as potential candidates *for the ObiH-ADC cascade:* the tryptophan decarboxylase from *Ruminococcus gnavus* (*Rgn*TDC, Fig 2b) and the valine decarboxylase from *Kitasatospora setae (VlmD, Fig 2c). We reasoned that assaying both a large aromatic ADC and a short-chain aliphatic ADC would test distinct limits for active-site steric tolerance for the β-OH moiety. We tested each enzyme for activity with a small set of β-OH amino acids and found that both decarboxylases could tolerate the β-OH moiety. RgnTDC had high* activity on several β-OH phenylalanine (Phe) analogs but had only trace activity on aliphatic β-OH amino acids (Fig 3a). VlmD showed the opposite activity trend, possessing high activity on smaller aliphatic β-OH amino acids, with little to no activity on larger substrates (Fig 3b). Notably, VlmD had high activity on Thr (**1b**), which may limit its utility when used in combination with ObiH in a one-pot cascade.

To investigate the potential of these decarboxylases for cascade synthesis of 1,2-amino alcohols*, we assayed these enzymes along with ObiH to form* (*R*)-2-amino-1-phenylethan-1-ol

(**2b**) from benzaldehyde and Thr (Fig 2d). The ObiH-VlmD cascade only produced trace amounts of the desired product, and instead the decarboxylation product from Thr was the major product, highlighting the specificity challenge of using these enzymes in concert (Fig 2d). Conversely, the ObiH-*Rgn*TDC cascade produced the desired product with minimal Thr decarboxylation. We therefore chose *Rgn*TDC as the ADC of our cascade. The engineering goal for *Rgn*TDC was clear. This enzyme does not natively possess any activity on the aliphatic β-OH amino acid substrates that can be made by ObiH, and we therefore set out to expand the scope of *Rgn*TDC so that it might compliment the scope of ObiH for cascade synthesis.

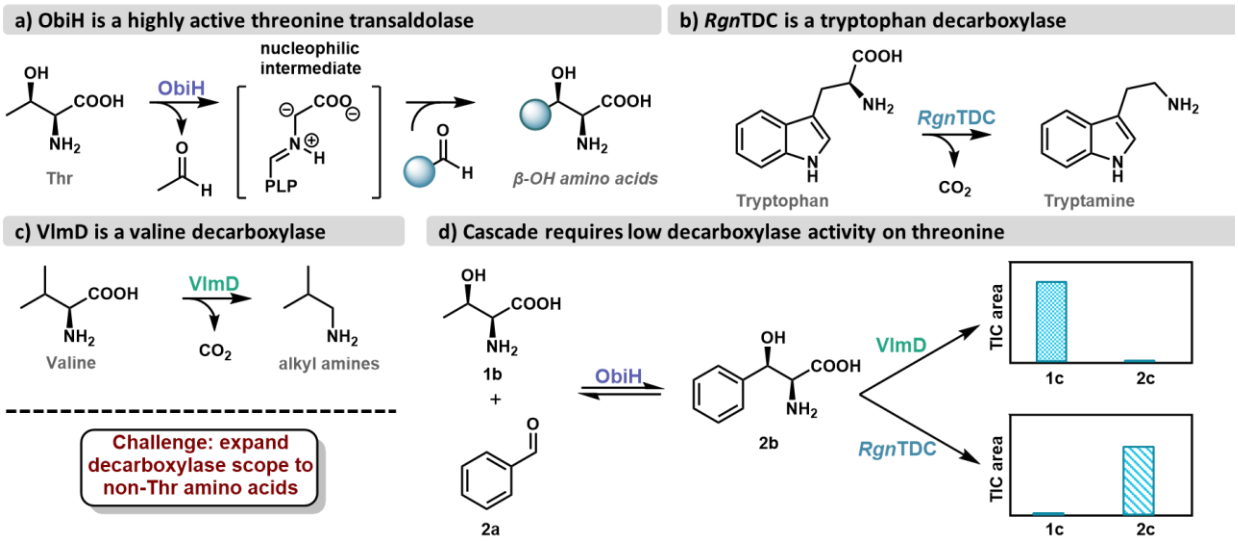


Figure 2. Enzymes under investigation for proposed ObiH – ADC cascade. a) General reaction of ObiH with aldehyde substrates. **b)** Native reaction of *Ruminococcus gnavus* tryptophan decarboxylase (*Rgn*TDC) **c)** Native reaction of *Kitasatospora setae valine decarboxylase (VlmD).* ***d)*** *Product outcomes from either ObiH – VlmD (top) or ObiH – RgnTDC (bottom) cascades from Thr (**1b**) and benzaldehyde (**2a**) to produce the corresponding decarboxylated products **1c** and **2c**.* 10 mM benzaldehyde, 50 mM Thr, 400 µM PLP, 50 mM KPi pH = 8.0, and 10 µM VlmD/*Rgn*TDC (final volume = 100 µl). Reactions were allowed to proceed for 16 h at 37 °C. Full data shown in Fig 12.

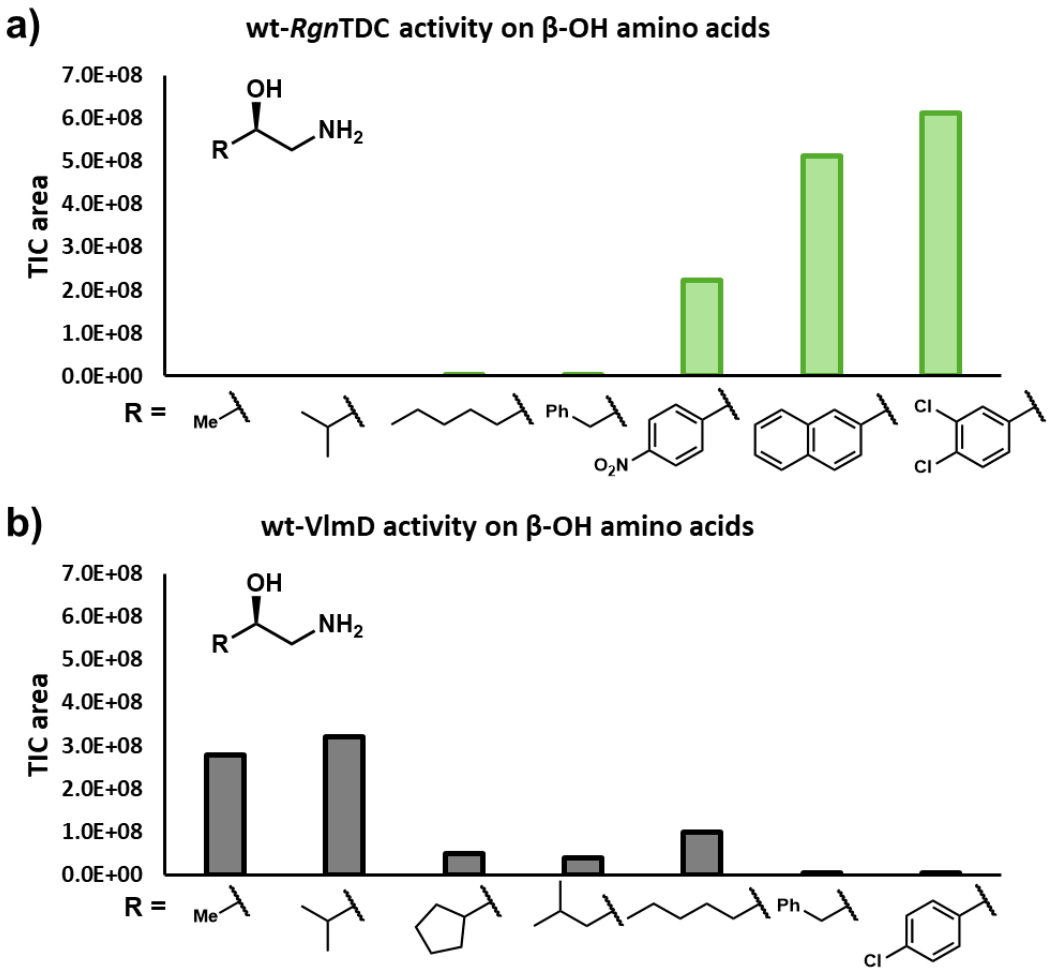


Figure 3. Activity of *Rgn*TDC and VlmD with various β-OH amino acids. 10 mM amino acid, 500 μM PLP, 50 mM KPi pH = 8.0, and 37.5 μM *Rgn*TDC/VlmD (final volume = 100 μl). Reactions were allowed to proceed for 16 h at 37 °C.

4. 2. 2. Engineering of *Rgn*TDC for activity on β-OH amino acids

Both substrate scope and sequence space must be considered for SUMS. We next investigated substrates to include for SUMS. As activity on the aromatic β-OH amino acids was already relatively high, the aromatic β-OH amino acids would be preferentially decarboxylated in the screen, inhibiting our ability to observe gains of activity for less active substrates. We therefore opted not to include this class of compounds in the substrate pool for screening. If engineered variants lost activity on aromatic β-OH amino acids, we would still be able to use the wt-*Rgn*TDC. We therefore assayed for activity improvements on a substrate space of benzylic and

short- and long-chain aliphatic β-OH substrates in our engineering screens. We assayed for gain-of-function on: L-Thr **(1b)**, β-OH homoPhe (**3b**) β-OH leucine (**4b**), and (2*S*,3*R*)-2-amino-3-hydroxyoctanoic acid (**5b**) (Fig 4b). Since the target cascade requires Thr, this amino acid was included in our *Rgn*TDC screens to monitor whether any variant gains activity on this small aliphatic β-OH amino acid. The first site in the protein that we considered for site-saturation mutagenesis (SSM) was His120. This residue is one of two that π-stacks with the native Trp substrate in *Rgn*TDC and bioinformatic analysis showed that this residue is not conserved among decarboxylases.

One variant from the H120X library, H120N, was found to significantly boost activity on both **3b** and **5b** (Fig 13) but possessed no activity on **1b** or **4b** when screened in competition. To better understand how the H120N mutation impacts activity, we assayed purified *Rgn*TDC^{H120N} against a complementary set of amino acids that lack the β-OH moiety: Trp, Phe, L-leucine (Leu), and L-homophenylalanine (homoPhe). We found significantly increased activity compared to wt for both Leu and homoPhe, commensurate with boosts for the hydroxylated homoPhe (Fig 5a). Additionally, we observed only a slight decrease in activity on the aromatic substrates Trp and Phe (Fig 5b). These data suggest that the H120N mutation does not boost activity by specifically engaging the β-OH moiety.

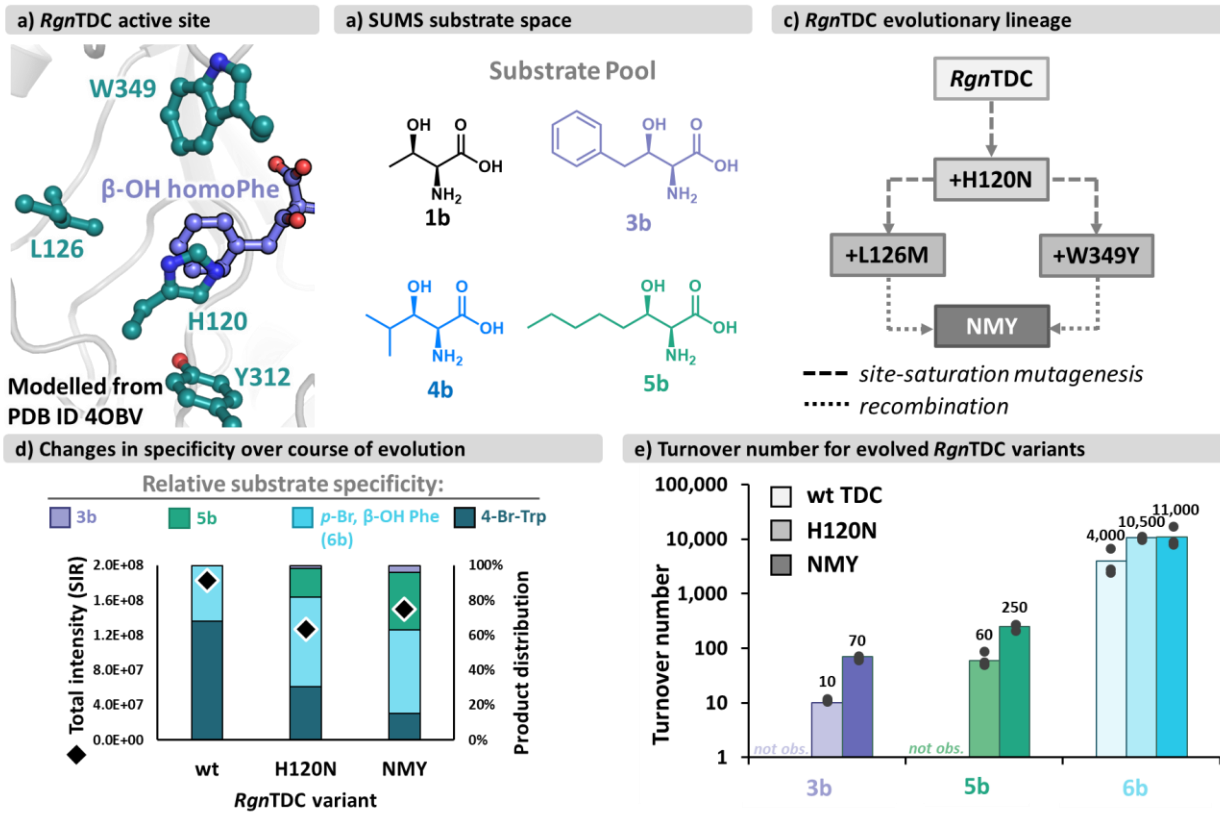


Figure 4. Engineering of *Rgn*TDC. a) Substrates used for screening during engineering. **b)** Depiction of *Rgn*TDC active site with **3b** modelled in (from 4OBV). **c)** Evolutionary trajectory of *Rgn*TDC engineering. 'NMY' = H120N, L126M, and W349Y *Rgn*TDC. **d)** Specificity for various amino acids over the course of *Rgn*TDC engineering from substrate multiplexed reactions. 2 mM **3b**, 2 mM **5b**, 0.2 mM **6b**, and 2 mM **4-bromotryptophan** (4-Br-Trp), 400 µM PLP, 1 µM *Rgn*TDC variant, and 50 mM KPi pH = 7.5 (final volume = 100 µl). Reactions were conducted in duplicate for 16 h at 37 °C. **e)** Turnover numbers for β-OH amino acid for *Rgn*TDC variants. No product was observed for wt-*Rgn*TDC for **3b** or **5b**. Reaction conditions are detailed in Materials and Methods.

The H120N mutation gained activity on **3b**, but the activity was still poor (10 turnovers), prompting additional rounds of evolution using *Rgn*TDC^{H120N} as the parent. We continued the strategy of targeting the actives site with three additional SSM libraries (L126, Y312, W349, Fig 4a, Fig 14-16). Most variants observed in the screens were deactivating, but the W349Y, W349F, and L126M mutations were further activating for **3b** and **5b** (Fig 4c, Fig 14-15). As with wt-*Rgn*TDC, no activity was detected on either **1b** or **4b**. Recombining these active-site

mutations, we found the triple mutant variant H120N/L126M/W349Y (hereafter called *Rgn*TDC^{NMY}) further improved activity on substrates **3b** and **5b**.

We next isolated *Rgn*TDC^{NMY} and investigated its activity as a purified enzyme. We compared the specificity of *Rgn*TDC^{NMY} against wt-*Rgn*TDC and *Rgn*TDC^{H120N} to investigate the specificity of these variants with different amino acids. We included 4-Br-Trp in these measurements to represent the native tryptophan substrate class (Fig 4d). We observed that *Rgn*TDC^{H120N} and *Rgn*TDC^{NMY} both shifted specificity from 4-Br-Trp towards β-OH amino acids, with *Rgn*TDC^{NMY} possessing the largest shifts. In addition, *Rgn*TDC^{NMY} was greatly activating on **3b** and **5b** in single substrate reactions (Fig 4e). Notably, the H120N mutation alone also improved activity on β-OH, *p*-Br-Phe (**6b**), despite this substrate not being present in the screen, but the additional L126M and W349Y mutations did not further increase activity with this substrate.

We again tested whether the boosts in decarboxylase activity were specific to the β-hydroxy group, or were instead more reflective of a general mode of activation. We tested the *Rgn*TDC^{NMY} variant against the same des-hydroxy amino acids that were assayed. We observed that this variant had just a 1.9-fold increase in activity for homoPhe, compared to the 7-fold activity increase for β-OH homoPhe (Fig 4e, Fig 5). These data are evidence of evolution producing a better generalist, as the final variant continues to improve with the lowest-activity substrates and maintains or even slightly improves activity with the previously preferred substrates.

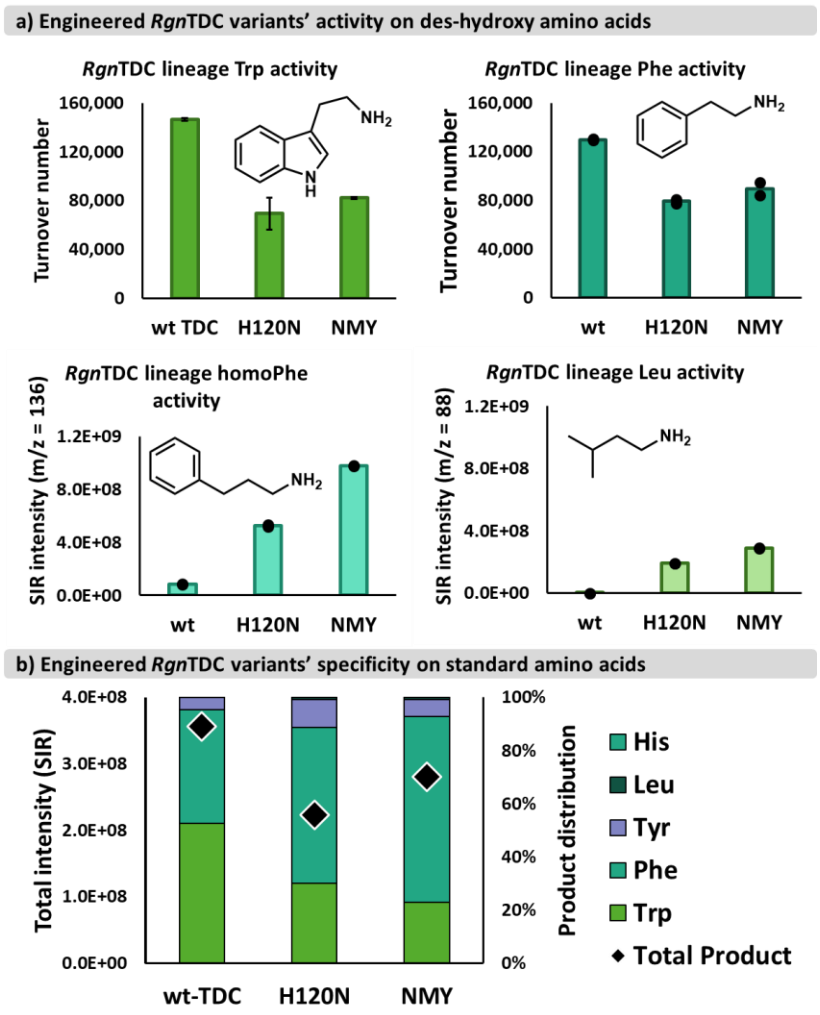


Figure 5. Activity of *Rgn*TDC variants with des-β-OH amino acids. a) Single-substrate activity comparisons. Reaction conditions: 10 mM amino acid, 400 μM PLP, 0.05 μM *Rgn*TDC (for **Trp/Phe**) or 10 μM *Rgn*TDC (for **homoPhe/Leu**), and 50 mM KPi pH = 8.0 (final volume = 100 μl). Reactions were allowed to proceed for 16 h at 37 °C. **b)** Analysis of *Rgn*TDC variant specificities for standard amino acids. Reaction conditions: 2 mM each substrate (**His, Leu, Tyr, Phe, Trp**), 400 μM PLP, 1 μM *Rgn*TDC variant, and 50 mM KPi pH = 8.0 (final volume = 100 μl). Reactions were conducted in duplicate for 10 min at 37 °C. Average product single ion retention (SIR) areas were integrated to quantitate relative amounts of product formed.

Aromatic amino acid decarboxylases (AADCs) have not previously been engineered for activity on aliphatic or β-OH amino acids. We successfully engineered *Rgn*TDC^{NMY} for higher specificity and activity towards β-OH amino acids relative to wt-*Rgn*TDC (Fig 4d). The H120N mutation was an unexpected find, as His at this site is not conserved among aromatic amino

acid decarboxylases, but is present in VlmD, a valine decarboxylase, and other aliphatic amino acid decarboxylases (such as methionine decarboxylase) (Fig 6). The exact role of Asn for this substrate specificity is currently unclear. We also observed that this residue plays a role in the relative specificity between Trp and Phe (Fig 5b), although to a lesser extent. Additionally, we note that both *Rgn*TDC and VlmD natively tolerate the β-OH functionality, albeit with reduced turnovers compared to the des-hydroxy amino acid analogs. Previously, a tyrosine decarboxylase was also reported to possess good activity with β-OH amino acids.[22,23] These data suggest that the β-OH moiety might be well-tolerated throughout different decarboxylase classes. This work underscores that many facets of substrate discrimination in amino acid decarboxylases are as-of-yet unknown, prompting further investigation into these biocatalytically relevant enzymes.

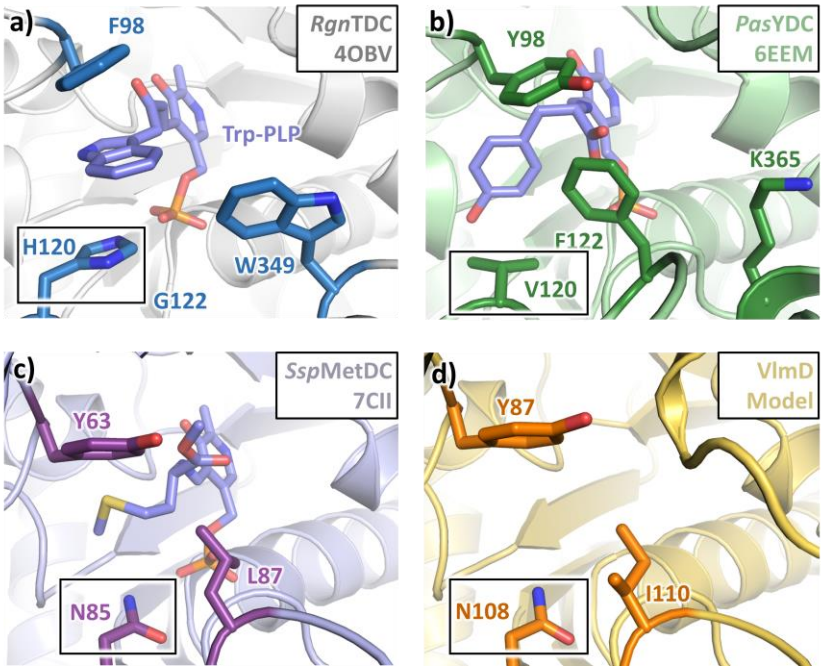


Figure 6. Active-site comparisons between various amino acid decarboxylases. PDB ID's are denoted for each structure (*Rgn*TDC = 4OBV, *Pas*YDC = 6EEM, *Ssp*MetDC = 7CII). The residues in the black box are analogous to the H120 residue of *Rgn*TDC. The structure of VlmD was generated using SwissModel with 7CII serving as the parent structure.

4. 2. 3. Development and application of ObiH – *Rgn*TDC^{NMY} cascade

Upon engineering the *Rgn*TDC^{NMY} variant, which had improved activity on β-OH amino acids without engendering activity on Thr, we decided to return to cascade development by combining both enzymes and optimizing reaction conditions. To this aim, we screened for improved buffer, pH, and cosolvent conditions for both *Rgn*TDC and ObiH activity (Fig 7). From these experiments, we observed that ObiH had maximum activity in Tris buffer, while *Rgn*TDC preferred potassium phosphate (KP_i) buffer. Given the diversity of our intended substrate scope, we envisioned using Tris as a buffer system for substrates challenging for ObiH and using KP_i as the buffer for substrates challenging for *Rgn*TDC.

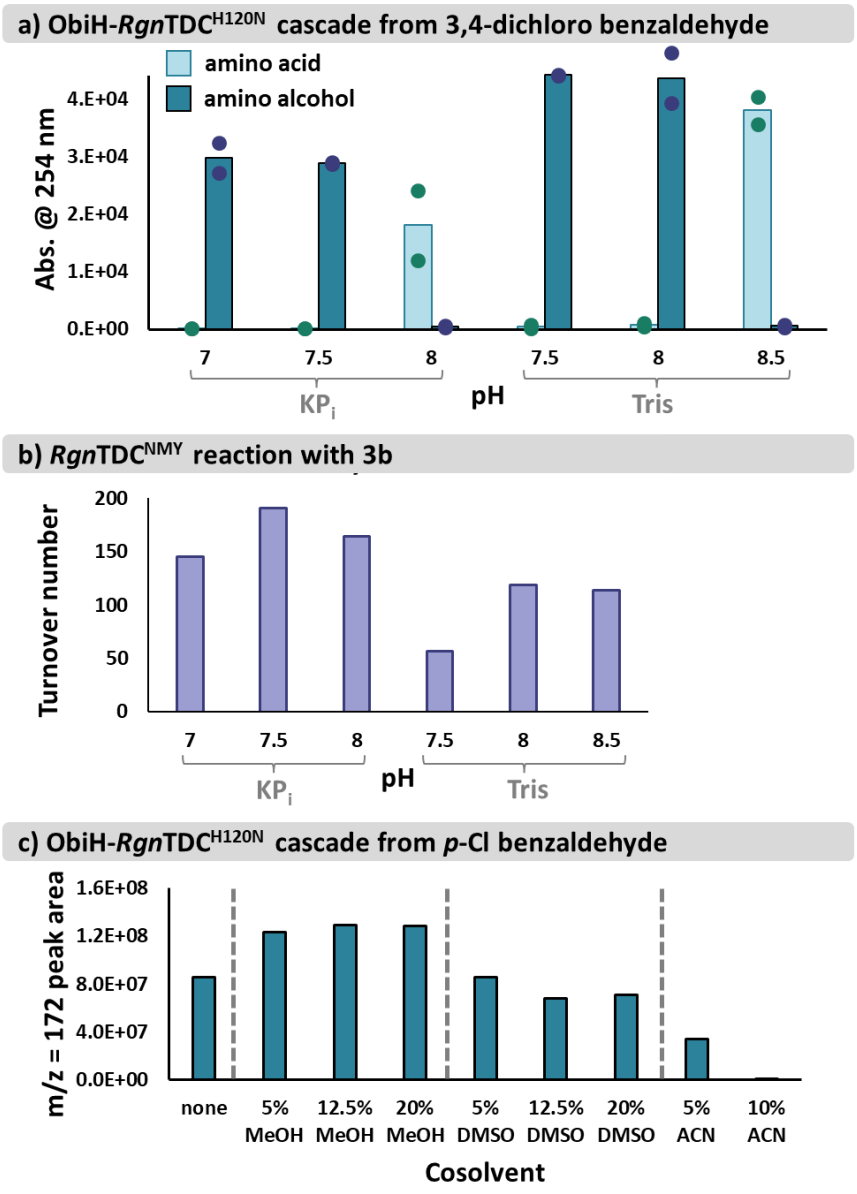


Figure 7. Determination of preferred buffer and pH system for ObiH – *Rgn*TDC cascade. **a)** Buffer screen for ObiH – RgnTDC^{H120N}**.** Reaction conditions: 25 mM 3,4-dichlorobenzaldehyde, 75 mM Thr, 400 µM PLP, 10 µM ObiH, 25 µM *Rgn*TDC^{H120N}, and 50 mM Tris/KPi with 5% MeOH (final volume = 250 µl). Buffer pH's ranged from pH = 7.0 to pH = 8.5. Reactions were conducted for 16 h at 37 °C at 180 RPM's. Product abundance was determined by absorbance at 254 nm. **b)** Buffer screen for *Rgn*TDC^{NMY}. Reaction conditions: 10 mM **3b**, 400 µM PLP, 10 µM *Rgn*TDC^{NMY}, and 50 mM Tris/KPi (final volume = 100 µl). Buffer pH's ranged from pH = 7.0 to pH = 8.5. Reactions were conducted for 16 h at 37 °C at 180 RPM's. Turnover numbers were determined by relative peak areas at 254 nm. **c)** Cosolvent screen for ObiH – *Rgn*TDC^{H120N} cascade. Reaction conditions: 25 mM *p*-chlorobenzaldehyde, 80 mM Thr, 200 µM PLP, 50 mM Tris pH = 8.0, 32 µM ObiH, 12 µM *Rgn*TDC^{H120N} (final volume = 250 µl).

Next, we investigated the diastereoselectivity of *Rgn*TDC^{NMY} for the β-OH moiety. In control experiments with a diastereomeric mixture of **6b**, we found that *Rgn*TDC possessed high activity and little to no selectivity for the β-OH amino acid diastereomers (Fig 8). We postulated that this would result in *Rgn*TDC^{NMY} exerting no stereocontrol over the desired cascade. Instead, the e.r. of the products would reflect the stereoselectivity of the aldol addition catalyzed by ObiH. Previously, it was shown that ObiH forms β-OH amino acids with excellent initial selectivity (>20:1 d.r.), but the d.r. erodes at higher conversion as the products re-enter the catalytic cycle.[26] By coupling the engineered decarboxylase with ObiH, nascent amino acids can be decarboxylated before significant scrambling of the hydroxyl group.

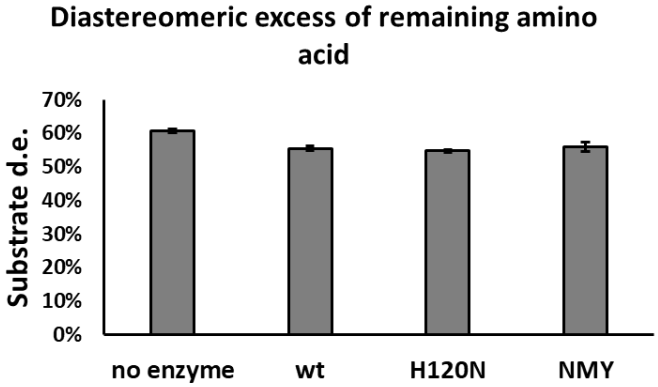


Figure 7: Analysis of *Rgn*TDC variant diastereoselectivity. Reaction conditions: 10 mM *p*-Br, β-OH Phe **(**80:20 d.r.), 400 μM PLP, 10 μM *Rgn*TDC variant, and 50 mM KPi pH = 8.0 (final volume = 100 μl). Substrate d.e. was measured after incubation with *Rgn*TDC variants for 15 min at 37 °C corresponding with ~20% substrate conversion.

With optimized cascade conditions in hand, we sought to investigate the scope of the engineered cascade. Initial reactions were conducted on analytical scale. These assays showed that the two-enzyme combination was indeed promiscuous and capable of producing a wide range of 1,2-amino alcohols (Fig 9a). The cascade produced a variety of aromatic 1,2-amino alcohols with a wide variety in yield. We were initially unsure if *Rgn*TDC would be sufficiently active at the modest catalyst loadings used for these reactions (up to 0.1 – 0.5 mol%) and

maintain high product e.r.'s. However, most aromatic products were produced with high e.r.'s of > 90:10. (**2c**, **6c-10c**), indicating little re-entry into the ObiH active site. The relatively insoluble napthyl (**11a**) and biphenyl (**12a**) substrates were not efficiently carried through the cascade, and **12c** was produced with a notably low d.r. of 75:25. The furanyl aldehyde (**13a**) was not reactive with ObiH in this cascade, resulting in no observed product. Through our engineering efforts, *Rgn*TDC^{NMY} gained activity on numerous aliphatic β-OH amino acids with diverse functionalities, such as **5b** as well as thioether **14b** and the Boc-protected amine **15b**. Analytical scale reactions suggested octanal (**16a**) was accepted by the cascade as well, producing the long-chain 1,2-amino alcohol **16c**. Short-chain aliphatic β-OH amino acids, bearing cyclopentyl, butyl, or isovaleryl groups (**17b-19b**), however, were not decarboxylated by *Rgn*TDC^{NMY} (Fig 17). All aliphatic products were observed with excellent e.r.'s. Since many of these reactions were observed with low overall conversion, we made several changes upon moving to the preparative scale. Reaction times were increased from 16 h to 40 h most substrates to increase yield. Additionally, flask shaking was implemented for poorly soluble aldehyde substrates to increase mass action during the reactions.

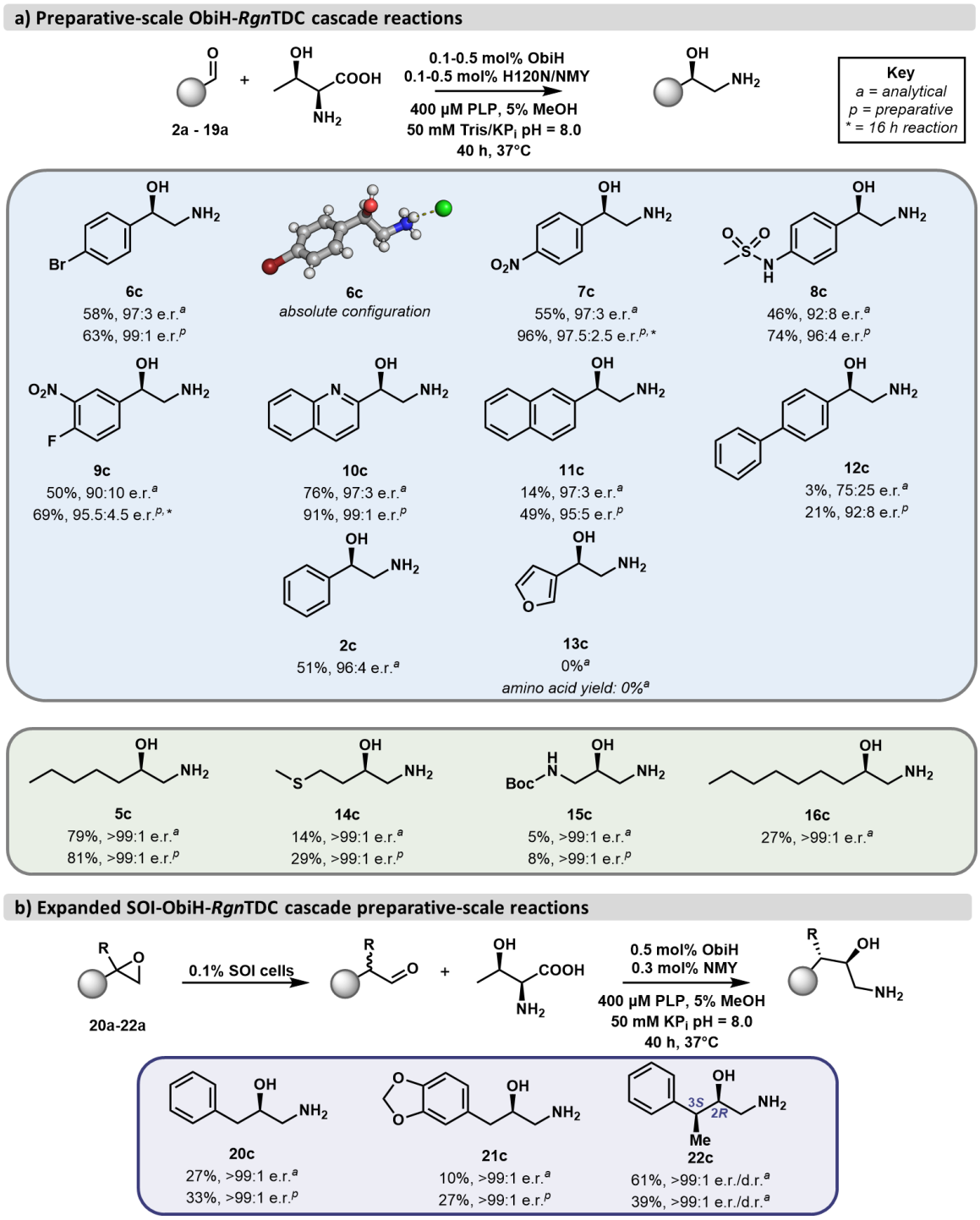


Figure 9. Substrate scope of the engineered ObiH – *Rgn*TDC cascade. a) Analytical and preparative yields of ObiH – *Rgn*TDC cascade. **b)** Analytical (*a*) and preparative (*p*, isolated) yields for expanded SOI – ObiH – *Rgn*TDC cascade. Reactions run for 16 h are denoted by an asterisk (*).

We then used these new conditions with the engineered cascade for the preparative scale syntheses of 1,2-amino alcohols (Fig 9a). Aromatic aldehydes remained the most active substrates for the two-enzyme cascade, giving rise to a series of epinephrine analogs. *para*-Substitutions such as -NO_2, -Br, and -sulfonamido groups were well-tolerated (**6c-8c**), and a large *meta*-substitution such as *m*-NO_2 was accepted as well (**9c**). The *p*-Br analog **6** was crystallized as the HCl salt and the X-ray structure confirmed that the major product was the *R* isomer. Heterocyclic and bicyclic groups such as 2-quinolinyl, naphthyl, and biphenyl (**10c-12c**) were also produced from the cascade. Next, we applied the engineered cascade to synthesize aliphatic amino alcohols. ObiH was previously noted to synthesize aliphatic β-OH amino acids with superb selectivity.[26] Cascade reactions with hexanal proceeded smoothly to synthesize **5c** in good yields and excellent stereochemical purity. Reactions with methional produced **14c** in lower yields but with exceptional e.r. We additionally probed the reactivity of *N*-Boc aminoacetaldehyde, which generated an asymmetrical 1,3-diamino alcohol (**15c**), albeit in poor yield. We highlight these preparative syntheses as examples of products previously unobtainable via the native cascade.

To access benzylic 1,2-amino alcohols with this cascade, α-arylaldehydes are required as substrates. These are a notoriously reactive and unstable class of aldehyde. Therefore, we introduced a third enzyme, styrene oxide isomerase (SOI), to produce these compounds *in situ* from the corresponding epoxide. The cascade reaction of SOI and ObiH was recently reported (CITE), which draws on the strong work of Li et al.[27] With just 0.1% w/v whole cells bearing SOI, the α-arylaldehydes reacted efficiently with the ObiH-*Rgn*TDC^{NMY} cascade on the analytical scale (Fig 8b). We found that *Rgn*TDC^{NMY} has relatively low activity on these challenging substrates; nevertheless, the three-enzyme cascade furnished diverse benzylic 1,2-amino alcohols with excellent e.r.'s, albeit with lower yields (**20c-22c**). The SOI-ObiH cascade was previously characterized to produce **22c** with a 65:35 d.r. with an *anti*-relationship between the

γ-Me and β-OH groups (CITE). To our surprise, *Rgn*TDC^{NMY}, was highly selective for the minor *syn* amino acid, fully consuming the *syn* amino acid and producing exclusively the corresponding *syn* 1,2-amino alcohol (Fig 10).

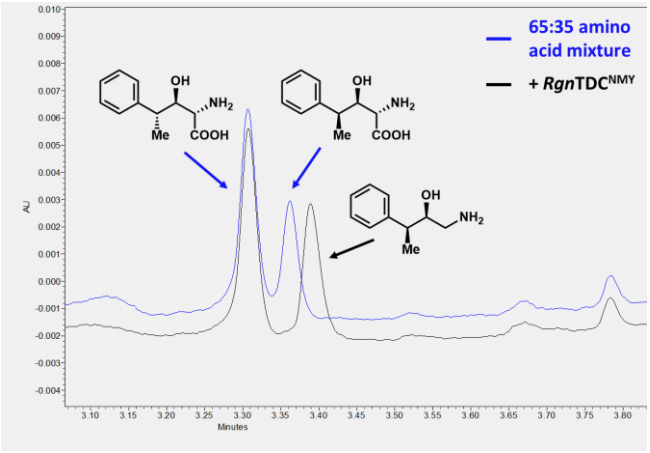


Figure 10. Determination of *Rgn*TDC γ-Me diastereoselectivity. 10 mM **22b** (65:35 *anti*:*syn* for the γ-Me and β-OH groups; prepared as previously described (CITE)), 400 μM PLP, 0 or 10 μM *Rgn*TDC^{NMY}, 50 mM KPi pH = 8.0 (final volume = 100 μl). Reactions were conducted in duplicate for 16 h at 37 °C.

The products of this cascade, and biocatalytic cascades more broadly, seldom bear protecting groups. Given structural similarity of these products to clinically used β-blockers, we were interested in synthesizing several well-known analogues. Such molecules have been clinically used as racemic mixtures, but the biologically more active isomers have been reported as the *R*-isomers.[28] We considered performing the reductive amination reaction as a one-pot reaction directly from the aldehyde starting materials, but opted against such an approach, considering the plethora of reactive amines present in the starting reaction mixture. To this aim, we subjected several 1,2-amino alcohols to reductive amination reactions with acetone (Fig 11). This facile alkylation reaction afforded the stereo-enriched isomers of the clinically studied nifenlol (**7d**) and pronethalol (**11d**) as well as the bioactive isomer (*R*)-sotalol (**8d**) in good yields with retention of enantiopurity. Nevertheless, this chemoenzymatic route is complementary to

previous synthetic approaches to these important compounds, which often rely on formation of the styrene oxide followed by ring-opening with isopropylamine.[29–31]

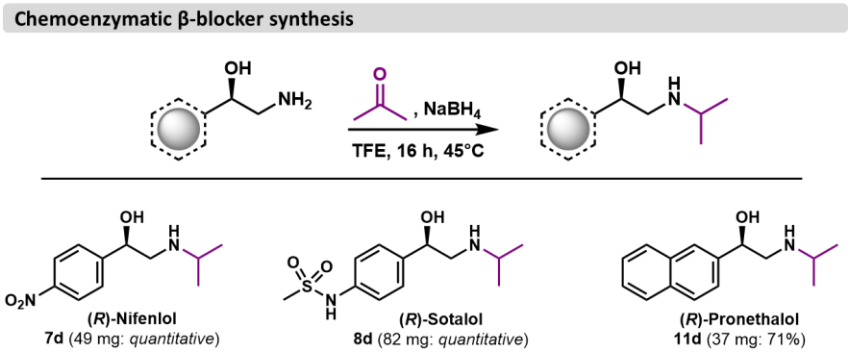


Figure 11. Preparative syntheses of β-adrenergic receptor agonists (β-blockers).

4. 3. Conclusions

This cascade engineering endeavor represents one of the first examples of an *in vitro* directed evolution campaign using a SUMS approach. Directed evolution campaigns using racemic substrates[32,33] or in *in vivo* systems[34] have been described. However, the principles that lead to success for *in vitro* cascade engineering with a diverse substrate pool have not yet been experimentally established. Both the substrate space and the sequence space require consideration. We felt that SUMS as an engineering and screening methodology was well-poised to tackle this challenge of matching substrate scopes for multi-enzyme cascades. SUMS allows many substrates to be assayed simultaneously, provided that the chosen substrates have unique product m/z ratios. This substrate flexibility is especially powerful when searching for gain-of-function mutations, as mutations may only confer activity on a single substrate. Screening on multiple substrates, therefore, decreases the odds of encountering false negatives or evolutionary dead-ends. We leveraged this screening approach to identify gain-of-function on α-aryl and aliphatic β-OH amino acids with *Rgn*TDC while simultaneously monitoring for any variants that might gain activity on Thr. While we did not observe any such *Rgn*TDC variants

with activity on Thr, SUMS would have allowed us to recognize shifts in specificity towards Thr and allowed us to shift course away from such *Rgn*TDC variants. This simultaneous positive and negative selection approach may nevertheless prove valuable in the future as screening for complementary scopes and tailored specificity grows becomes more immediate with increasingly complex biocatalytic cascades.

4. 4. Materials and Methods

Chemicals and reagents were purchased from commercial suppliers (Sigma-Aldrich, VWR, Chem-Impex International, Alfa Aesar, Combi-blocks, Oakwood Products) at the highest quality available and used without further purification unless stated otherwise. β-OH amino acids were obtained from Dr. Tyler Doyon as previously prepared (CITE) or Tony Meza as previously described (CITE). Genes were purchased as gBlocks from Integrated DNA Technologies (IDT). E. coli cells were electroporated with a Bio-Rad MicroPulser at 2500 V. New Brunswick I26R, 120 V/60 Hz shaker incubators (Eppendorf) were used for cell growth. Cell disruption via sonication was performed with a Sonic Dismembrator 550 (Fisher Scientific) sonicator. Optical density measurements were collected using an optical density reader (Amersham Biosciences). Ultra-high pressure liquid chromatography-mass spectrometry (UPLC-MS) data were collected on an Acquity UPLC (Waters) equipped with an Acquity PDA and QDA MS detector using either a BEH C18 column (Waters) or an Intrada Amino Acid column (Imtakt). Preparative column separations were performed on an Isolera One Flash Purification system (Biotage). NMR data were collected on a Bruker 500 MHz spectrometer equipped with a DCH cryoprobe. Signal positions were recorded in ppm with the abbreviations s, d, t, q, dd, and m, denoting singlet, doublet, triplet, quartet, doublet of doublets, and multiplet respectively. All coupling constants J are measured in Hz. High resolution mass spectrometry data were collected with a Q Extractive Plus Orbitrap (NIH 1S10OD020022-1) instrument with samples ionized by ESI.

4. 4. 1. Plasmid and protein information

Protein sequence of *N-His-ObiH* (Uniprot accession code: A0A1X9LWZ7):

MGSSHHHHHHSSMSNVKQQTAQIVDWLSSTLGKDHQYREDSLSLTANENYPSALVRLTSGST
AGAFYHCSFPFEVPAGEWHFPEPGHMNAIADQVRDLGKTLIGAQAFDWRPNGGSTAEQALML
AACKPGEGFVHFAHRDGGHFALESLAQKMGIEIFHLPVNPTSLLIDVAKLDEMVRRNPHIRIVILD
QSFKLRWQPLAEIRSVLPDSCTLTYDMSHDGGLIMGGVFDSPLSCGADIVHGNTHKTIPGPQK
GYIGFKSAQHPLLVDTSLWVCPHLQSNCHAEQLPPMWVAFKEMELFGRDYAAQIVSNAKTLAR
HLHELGLDVTGESFGFTQTHQVHFAVGDLQKALDLCVNSLHAGGIRSTNIEIPGKPGVHGIRLG
VQAMTRRGMKEKDFEVVARFIADLYFKKTEPAKVAQQIKEFLQAFPLAPLAYSFDNYLDEELLA
AVYQGAQR

DNA sequence of *N-His-ObiH*:

ATGTCAAACGTGAAGCAGCAGACGGCGCAAATCGTAGATTGGTTGTCATCAACCTTGGGG
AAAGATCACCAATACCGCGAGGACTCCCTTTCACTTACCGCTAACGAGAACTACCCGTCGG
CATTAGTTCGTTTGACTTCAGGTTCGACCGCCGGCGCATTCTACCATTGTTCTTTCCCCTTC
GAGGTTCCTGCCGGGGAGTGGCACTTCCCGGAGCCCGGTCATATGAATGCAATTGCTGAC
CAGGTTCGTGATTTAGGTAAAACCTTGATTGGTGCCCAGGCATTCGACTGGCGTCCAAATG
GCGGATCAACCGCAGAACAGGCACTTATGCTGGCAGCATGTAAACCGGGAGAGGGGTTC
GTCCATTTTGCTCACCGCGACGGAGGCCATTTCGCTTTAGAATCTCTTGCGCAAAAGATGG
GCATCGAAATTTTCCACTTGCCTGTTAATCCGACCTCTCTGTTAATCGATGTCGCCAAATTG
GATGAAATGGTCCGCCGCAACCCGCATATTCGCATTGTCATTCTTGATCAGAGCTTTAAGC
TGCGCTGGCAACCCCTGGCCGAGATTCGTTCAGTTTTACCAGACTCATGCACGTTGACTTA
TGATATGAGTCATGATGGGGGATTAATTATGGGAGGTGTCTTCGATTCCCCCCTTAGCTGT
GGAGCTGACATCGTCCACGGCAATACTCACAAGACGATTCCTGGACCGCAAAAGGGGTAT
ATCGGTTTCAAGTCCGCGCAACATCCTTTATTAGTCGATACAAGTTTATGGGTATGCCCTCA
CCTTCAAAGTAACTGCCACGCCGAGCAGCTGCCGCCGATGTGGGTTGCCTTCAAGGAAAT

GGAATTATTTGGACGCGATTACGCTGCCCAAATTGTTTCAAACGCAAAAACCTTGGCTCGC
CATCTGCATGAACTGGGATTGGACGTGACCGGAGAATCCTTTGGATTCACACAGACACATC
AGGTCCATTTTGCTGTAGGAGATTTACAGAAAGCGCTTGATCTTTGTGTGAATTCATTACAT
GCTGGAGGTATCCGTTCGACCAATATTGAAATCCCAGGGAAACCAGGAGTACATGGCATTC
GCTTAGGCGTCCAAGCGATGACTCGTCGTGGCATGAAGGAGAAAGACTTTGAGGTGGTCG
CCCGTTTTATCGCCGATCTGTACTTTAAAAAAACGGAACCTGCCAAGGTCGCACAGCAAAT
TAAGGAATTTTTACAGGCGTTTCCGCTTGCACCTTTAGCCTACTCATTTGATAACTATCTTG
ATGAAGAATTATTGGCAGCCGTTTACCAGGGTGCGCAGCGCTGA

Protein sequence of *C-His-RgnTDC* (Uniprot accession code: A7B1V0):

MSQVIKKKRNTFMIGTEYILNSTQLEEAIKSFVHDFCAEKHEIHDQPVVVEAKEHQEDKIKQIKIP
EKGRPVNEVVSEMMNEVYRYRGDANHPRFFSFVPGPASSVSWLGDIMTSAYNIHAGGSKLAP
MVNCIEQEVLKWLAKQVGFTENPGGVFVSGGSMANITALTAARDNKLTDINLHLGTAYISDQTH
SSVAKGLRIIGITDSRIRRIPTNSHFQMDTTKLEEAIETDKKSGYIPFVVIGTAGTTNTGSIDPLTEI
SALCKKHDMWFHIDGAYGASVLLSPKYKSLLTGTGLADSISWDAHKWLFQTYGCAMVLVKDIR
NLFHSFHVNPEYLKDLENDIDNVNTWDIGMELTRPARGLKLWLTLQVLGSDLIGSAIEHGFQLA
VWAEEALNPKKDWEIVSPAQMAMINFRYAPKDLTKEEQDILNEKISHRILESGYAAIFTTVLNGK
TVLRICAIHPEATQEDMQHTIDLLDQYGREIYTEMKKALEHHHHHH

DNA sequence of *C-His-RgnTDC*:

ATGTCGCAGGTCATTAAGAAAAAACGCAATACGTTTATGATTGGAACGGAGTACATCCTTAA
TTCGACACAGTTAGAGGAGGCAATTAAGTCTTTCGTGCACGATTTTTGTGCGGAAAAACAT
GAGATCCATGATCAGCCCGTCGTTGTTGAAGCCAAGGAGCACCAGGAAGATAAAATTAAG
CAGATCAAGATCCCTGAAAAAGGACGCCCAGTAAATGAGGTCGTGAGTGAGATGATGAAT
GAAGTTTACCGCTATCGCGGAGATGCGAACCACCCCCGTTTCTTCTCCTTCGTTCCGGGTC

CAGCTTCGAGCGTCTCCTGGCTTGGAGACATCATGACGAGTGCATATAATATCCATGCCGG
AGGCAGTAAATTGGCTCCCATGGTAAACTGTATTGAGCAAGAAGTGCTGAAGTGGTTGGCA
AAGCAAGTGGGATTTACTGAAAATCCCGGCGGGGTGTTCGTCTCAGGTGGCTCGATGGCG
AACATCACGGCGTTAACAGCAGCCCGTGACAATAAACTTACTGACATTAATTTGCATTTAGG
AACGGCGTATATCAGCGACCAAACACACTCCAGTGTAGCCAAGGGGTTACGTATTATTGGC
ATCACCGACAGCCGTATTCGCCGTATTCCCACTAATTCGCACTTCCAAATGGATACGACCA
AGTTGGAGGAGGCCATTGAAACCGATAAAAAGAGTGGCTATATCCCGTTTGTAGTGATCGG
AACCGCTGGCACGACTAATACAGGATCCATTGACCCATTAACGGAAATTTCTGCATTATGTA
AAAAGCACGATATGTGGTTCCACATCGACGGTGCGTATGGTGCCTCCGTATTGCTTAGTCC
AAAATATAAGTCCCTTTTGACAGGAACAGGATTAGCAGATAGTATTTCTTGGGATGCTCACA
AATGGTTATTCCAGACGTATGGGTGCGCCATGGTATTGGTGAAGGACATCCGCAACCTGTT
CCATTCCTTTCACGTTAACCCCGAATATCTGAAAGACCTTGAGAATGACATTGATAATGTCA
ATACGTGGGATATTGGGATGGAGTTAACACGTCCGGCACGTGGGCTTAAACTTTGGCTGA
CCTTGCAGGTTCTTGGGTCCGACCTTATTGGGTCTGCAATTGAGCACGGTTTCCAATTAGC
GGTATGGGCGGAAGAAGCGCTGAATCCCAAAAAAGATTGGGAAATTGTTAGCCCTGCCCA
GATGGCGATGATTAATTTTCGCTACGCGCCTAAGGATTTAACCAAAGAGGAGCAGGACATC
CTTAATGAAAAGATTTCGCATCGCATCTTGGAATCAGGCTATGCCGCTATTTTTACTACTGT
GCTGAATGGTAAGACAGTGTTACGCATTTGCGCGATTCACCCTGAGGCTACTCAAGAGGAT
ATGCAGCACACCATTGATCTGTTGGACCAATACGGTCGCGAGATCTATACTGAAATGAAAA
AGGCTCTCGAGCACCATCACCATCACCATTGA

DNA sequence of *C-His-RgnTDC* H120N:

ATGTCGCAGGTCATTAAGAAAAAACGCAATACGTTTATGATTGGAACGGAGTACATCCTTAA
TTCGACACAGTTAGAGGAGGCAATTAAGTCTTTCGTGCACGATTTTTGTGCGGAAAAACAT
GAGATCCATGATCAGCCCGTCGTTGTTGAAGCCAAGGAGCACCAGGAAGATAAAATTAAG
CAGATCAAGATCCCTGAAAAAGGACGCCCAGTAAATGAGGTCGTGAGTGAGATGATGAAT

GAAGTTTACCGCTATCGCGGAGATGCGAACCACCCCCGTTTCTTCTCCTTCGTTCCGGGTC
CAGCTTCGAGCGTCTCCTGGCTTGGAGACATCATGACGAGTGCATATAATATCAATGCCGG
AGGCAGTAAAATGGCTCCCATGGTAAACTGTATTGAGCAAGAAGTGCTGAAGTGGTTGGCA
AAGCAAGTGGGATTTACTGAAAATCCCGGCGGGGTGTTCGTCTCAGGTGGCTCGATGGCG
AACATCACGGCGTTAACAGCAGCCCGTGACAATAAACTTACTGACATTAATTTGCATTTAGG
AACGGCGTATATCAGCGACCAAACACACTCCAGTGTAGCCAAGGGGTTACGTATTATTGGC
ATCACCGACAGCCGTATTCGCCGTATTCCCACTAATTCGCACTTCCAAATGGATACGACCA
AGTTGGAGGAGGCCATTGAAACCGATAAAAAGAGTGGCTATATCCCGTTTGTAGTGATCGG
AACCGCTGGCACGACTAATACAGGATCCATTGACCCATTAACGGAAATTTCTGCATTATGTA
AAAAGCACGATATGTGGTTCCACATCGACGGTGCGTATGGTGCCTCCGTATTGCTTAGTCC
AAAATATAAGTCCCTTTTGACAGGAACAGGATTAGCAGATAGTATTTCTTGGGATGCTCACA
AATGGTTATTCCAGACGTATGGGTGCGCCATGGTATTGGTGAAGGACATCCGCAACCTGTT
CCATTCCTTTCACGTTAACCCCGAATATCTGAAAGACCTTGAGAATGACATTGATAATGTCA
ATACGTGGGATATTGGGATGGAGTTAACACGTCCGGCACGTGGGCTTAAACTTTGGCTGA
CCTTGCAGGTTCTTGGGTCCGACCTTATTGGGTCTGCAATTGAGCACGGTTTCCAATTAGC
GGTATGGGCGGAAGAAGCGCTGAATCCCAAAAAAGATTGGGAAATTGTTAGCCCTGCCCA
GATGGCGATGATTAATTTTCGCTACGCGCCTAAGGATTTAACCAAAGAGGAGCAGGACATC
CTTAATGAAAAGATTTCGCATCGCATCTTGGAATCAGGCTATGCCGCTATTTTTACTACTGT
GCTGAATGGTAAGACAGTGTTACGCATTTGCGCGATTCACCCTGAGGCTACTCAAGAGGAT
ATGCAGCACACCATTGATCTGTTGGACCAATACGGTCGCGAGATCTATACTGAAATGAAAA
AGGCTCTCGAGCACCATCACCATCACCATTGA

DNA sequence of *C-His-RgnTDC*^{NMY} (H120N, L126M, W349Y):

ATGTCGCAGGTCATTAAGAAAAAACGCAATACGTTTATGATTGGAACGGAGTACATCCTTAA
TTCGACACAGTTAGAGGAGGCAATTAAGTCTTTCGTGCACGATTTTTGTGCGGAAAAACAT
GAGATCCATGATCAGCCCGTCGTTGTTGAAGCCAAGGAGCACCAGGAAGATAAAATTAAG

CAGATCAAGATCCCTGAAAAAGGACGCCCAGTAAATGAGGTCGTGAGTGAGATGATGAAT
GAAGTTTACCGCTATCGCGGAGATGCGAACCACCCCCGTTTCTTCTCCTTCGTTCCGGGTC
CAGCTTCGAGCGTCTCCTGGCTTGGAGACATCATGACGAGTGCATATAATATCAATGCCGG
AGGCAGTAAATTGGCTCCCATGGTAAACTGTATTGAGCAAGAAGTGCTGAAGTGGTTGGCA
AAGCAAGTGGGATTTACTGAAAATCCCGGCGGGGTGTTCGTCTCAGGTGGCTCGATGGCG
AACATCACGGCGTTAACAGCAGCCCGTGACAATAAACTTACTGACATTAATTTGCATTTAGG
AACGGCGTATATCAGCGACCAAACACACTCCAGTGTAGCCAAGGGGTTACGTATTATTGGC
ATCACCGACAGCCGTATTCGCCGTATTCCCACTAATTCGCACTTCCAAATGGATACGACCA
AGTTGGAGGAGGCCATTGAAACCGATAAAAAGAGTGGCTATATCCCGTTTGTAGTGATCGG
AACCGCTGGCACGACTAATACAGGATCCATTGACCCATTAACGGAAATTTCTGCATTATGTA
AAAAGCACGATATGTGGTTCCACATCGACGGTGCGTATGGTGCCTCCGTATTGCTTAGTCC
AAAATATAAGTCCCTTTTGACAGGAACAGGATTAGCAGATAGTATTTCTTGGGATGCTCACA
AATGGTTATTCCAGACGTATGGGTGCGCCATGGTATTGGTGAAGGACATCCGCAACCTGTT
CCATTCCTTTCACGTTAACCCCGAATATCTGAAAGACCTTGAGAATGACATTGATAATGTCA
ATACGTATGATATTGGGATGGAGTTAACACGTCCGGCACGTGGGCTTAAACTTTGGCTGAC
CTTGCAGGTTCTTGGGTCCGACCTTATTGGGTCTGCAATTGAGCACGGTTTCCAATTAGCG
GTATGGGCGGAAGAAGCGCTGAATCCCAAAAAAGATTGGGAAATTGTTAGCCCTGCCCAG
ATGGCGATGATTAATTTTCGCTACGCGCCTAAGGATTTAACCAAAGAGGAGCAGGACATCC
TTAATGAAAAGATTTCGCATCGCATCTTGGAATCAGGCTATGCCGCTATTTTTACTACTGTG
CTGAATGGTAAGACAGTGTTACGCATTTGCGCGATTCACCCTGAGGCTACTCAAGAGGATA
TGCAGCACACCATTGATCTGTTGGACCAATACGGTCGCGAGATCTATACTGAAATGAAAAA
GGCTCTCGAGCACCATCACCATCACCATTGA

Protein sequence of *C-His-VlmD* (Uniprot accession code: E4N6B4):

MAALPTGIPFGPDDAAWSTGLDRLRTAGATRVSAPSADPRETYPQLPELPPGRFQLPARGLDA
TEYAQAEDLFRRYVEDHSSRSLGYQLHWSEDFARRLAPYLGLQLNNIGDPYQHGAFMPNSKV

LERAVLDYFASLWNAKWPHRAGDPESYWGYVLTMGASEGNIQALWNARECLSGKPLAGQPR

LPADTAHENPNARHPVVFFSRETHYSLTKAVNLLGLDTFHALGSSRYPDANPLGPGTEWPTEV

PCVGGVDGPGAIDVEKLSLLVRFFVRRGHPVFVNLNYGSTFKGAFDDVPEAARAVHEICAEYG

MAERTFPSDREGTGARPRPGYWIHVDAALGAAYVPYLRMARAAGLVESAPPPFDFRLPQVHS

LTVSAHKWMGAPWPCGVFMTRNGLRMPPPRSSEYIGGTDTTLSGSRNGFSALLMWDYLAHH

SYDDLARQAAECDRLARLAHERLLKLQSGLGVDLLVSRSPWSLAVRFRRPDEAILRRYSLAYET

LLVDGVERPYAHLYVVPHVTEGLIDALLRDLGQPGAFAGGAAALEHHHHHH

DNA sequence of *C-His-vlmD:*

ATGGCAGCTTTGCCCACAGGTATCCCATTTGGTCCCGATGACGCCGCCTGGTCCACTGGA

CTGGATCGCCTGCGCACTGCTGGCGCTACTCGCGTAAGCGCACCGTCAGCTGACCCGCG

CGAAACGTATCCGCAATTGCCCGAGTTGCCGCCGGGCCGCTTCCAACTGCCGGCTCGCG

GTCTGGATGCAACAGAATATGCTCAAGCAGAAGATCTTTTCCGCCGCTACGTGGAAGATCA

CTCTTCTCGTTCTTTGGGCTACCAGTTGCACTGGTCCGAGGACTTTGCCCGTCGCCTGGC

GCCCTATTTAGGATTGCAGCTGAATAATATTGGCGATCCTTATCAACATGGAGCGTTTATGC

CAAATAGCAAGGTATTGGAACGCGCTGTACTGGACTATTTTGCTAGTTTATGGAACGCCAA

ATGGCCACATCGTGCGGGAGATCCCGAAAGTTACTGGGGCTACGTGCTGACAATGGGCG

CGTCCGAAGGAAACATTCAGGCGTTGTGGAATGCACGTGAATGTCTTAGCGGAAAGCCCC

TGGCAGGTCAACCTCGTTTGCCCGCAGACACTGCACATGAGAACCCGAATGCTCGCCATC

CGGTTGTATTCTTTTCCCGTGAGACCCACTATTCACTTACAAAAGCCGTTAACTTATTGGGA

CTGGACACATTCCATGCACTTGGGTCTAGCCGCTATCCAGACGCAAATCCATTAGGTCCCG

GGACAGAGTGGCCGACTGAAGTACCATGTGTGGGAGGGGTCGATGGACCCGGCGCTATT

GACGTAGAGAAGCTTTCCCTGTTAGTTCGTTTCTTTGTACGTCGTGGCCACCCGGTTTTCG

TGAACCTGAATTACGGTAGTACCTTCAAAGGGGCCTTCGACGATGTGCCTGAAGCTGCGC

GTGCTGTTCATGAGATTTGTGCCGAGTACGGGATGGCAGAACGCACTTTCCCATCCGATC

GCGAAGGAACGGGAGCCCGCCCCCGCCCAGGGTACTGGATTCATGTGGACGCGGCTTTG

GGAGCCGCGTACGTGCCGTACTTACGTATGGCACGTGCGGCCGGGTTAGTTGAATCAGCA
CCACCCCCGTTTGACTTTCGCTTACCCCAAGTACATTCTCTGACAGTGTCCGCTCATAAGT
GGATGGGCGCCCCCTGGCCATGCGGTGTATTCATGACTCGCAATGGATTACGTATGCCGC
CCCCTCGCTCTTCCGAATACATCGGAGGAACGGACACCACACTGTCTGGGAGTCGCAATG
GATTTTCCGCGCTGTTAATGTGGGATTATCTTGCGCACCATAGTTATGACGATTTAGCACG
CCAAGCCGCAGAATGCGATCGTTTAGCGCGTTTGGCCCATGAACGTTTATTAAAGTTGCAG
AGTGGATTAGGGGTTGACCTGTTAGTAAGCCGTTCGCCATGGAGCCTTGCCGTACGTTTTC
GCCGCCCGGATGAAGCAATCTTGCGCCGTTATAGTCTTGCCTACGAGACCCTTCTTGTCGA
TGGTGTTGAACGCCCTTACGCTCATCTTTATGTTGTACCTCATGTTACTGAAGGACTTATCG
ATGCCCTTCTTCGCGATTTAGGACAGCCAGGTGCGTTTGCAGGGGGAGCGGCGGCT*CTC*
*GAG*CACCATCACCATCACCATTGA

Protein sequence of *C-His-SOI* (*Pseudomonas sp.* VLB120):

MLHAFERKMAGHGILMIFCTLLFGVGLWMNLVGGFEIIPGYIIEFHVPGSPEGWARAHSGPALN
GMMVIAVAFVLPSLGFADKTARLLGSIIVLDGWSNVGFYLFSNFSPNRGLTFGPNQFGPGDIFS
FLALAPAYLFGVLAMGALAVIGYQALKSTRSRKAVPHAAAELEHHHHHH

DNA sequence of *C-His-soi*:

ATGCTTCATGCGTTCGAGCGTAAGATGGCTGGACATGGAATTTTGATGATTTTTTGCACTTT
ACTGTTTGGGGTAGGTTTGTGGATGAACTTGGTGGGTGGGTTTGAAATCATTCCAGGTTAT
ATTATTGAGTTTCACGTGCCAGGCTCCCCTGAGGGATGGGCACGTGCCCACAGCGGACCC
GCCTTGAATGGCATGATGGTAATCGCAGTTGCATTCGTGTTACCCTCCTTGGGTTTTGCGG
ACAAAACGGCACGTTTGCTGGGCAGTATCATCGTTCTTGATGGGTGGTCAAACGTGGGCTT
TTACTTGTTTTCTAACTTTTCTCCCAATCGTGGTCTGACTTTTGGTCCGAATCAGTTCGGAC
CAGGTGACATCTTTAGTTTCTTGGCTTTGGCGCCTGCGTACCTGTTCGGGGTATTAGCGAT

GGGCGCACTTGCAGTTATTGGTTATCAAGCCCTGAAATCAACACGCTCTCGTAAAGCAGTC
CCTCATGCCGCAGCTGAGCTCGAGCACCATCACCATCACCATTGA

DNA Isolation and Storage: DNA was purified via gel electrophoresis and isolated using a DNA gel extraction kit (Zymo Research). All isolated DNA was stored at -20 ^{o}C.

4. 4. 2. Protein expression methods

Cloning, expression, purification, and storage of ObiH

A codon-optimized copy of the *Pseudomonas fluorescens* ObiH gene was purchased as a gBlock from Integrated DNA Technologies. This DNA fragment was inserted into a pET-28b(+) vector by the Gibson Assembly method.[1] BL21 (DE3) E. coli cells were subsequently transformed with the resulting cyclized DNA product via electroporation. After 45 min of recovery in Terrific Broth II (TB) media containing 0.4% glucose at 37 °C, cells were plated onto Luria-Burtani (LB) plates with 50 µg/mL kanamycin (KAN) and incubated overnight. Single colonies were used to inoculate 5 mL TB + 50 µg/mL KAN (TB-KAN), which were grown overnight at 37 °C, 200 rpm. Expression cultures, typically 1 L of TB-KAN, were inoculated from these starter cultures and shaken (180 rpm) at 37 °C. After 3.5 hours (OD_{600} = > 1.0), the expression cultures were chilled on ice. After 45 min on ice, ObiH expression was induced with 1 mM IPTG, and the cultures were expressed for 16 hours at 23 °C with shaking at 200 rpm. (We found protein yields to be generally invariable to exact cell OD_{600} at the time of induction) Cells were then harvested by centrifugation at 4,300×g at 4 °C for 15 min. Cell pellets were pink in color and were frozen and stored at -20 °C until purification.

To purify ObiH, cell pellets were thawed on ice and then resuspended in lysis buffer (50 mM potassium phosphate buffer (pH = 8.0), 1 mg/ml Hen Egg White Lysozyme (GoldBio), 0.2 mg/ml DNaseI (GoldBio), 1 mM $MgCl_2$, and 200 µM pyridoxal 5′-phosphate (PLP)). A volume of 4 mL

of lysis buffer per gram of wet cell pellet was used. After 45 min of shaking at 37 °C, cells were sonicated with a ½ in. tip for 10 min (1 s on; 1 s off). The resulting lysate was then spun down at 75,000×g to pellet cell debris. Upon successful lysis, the pellet was colorless whereas the supernatant was pink in color. Ni/NTA beads (GoldBio) were added to a gravity column and the lysis supernatant ran over the bead bed for purification by Ni-affinity chromatography. The column was washed with 4 column volumes of 20 mM imidazole, 50 mM potassium phosphate buffer (pH = 8.0). Washing with higher concentrations of imidazole resulted in slow protein elution. ObiH was eluted with 250 mM imidazole, 50 mM potassium phosphate buffer (pH = 8.0). Elution of the desired protein product was monitored by the disappearance of its bright pink color (resulting from the release of ObiH) from the column. The protein product was dialyzed to < 50 μM imidazole in 50 mM Tris-HCl buffer (pH = 8.01) or 50 mM potassium phosphate buffer (pH = 8.08). Purified enzyme was flash frozen in pellet form by pipetting enzyme dropwise into a crystallization dish filled with liquid nitrogen. The enzyme was transferred to a plastic conical and stored at -80 °C until further use. Frozen pellets were thawed at room temperature and centrifuged before use. The concentration of protein was determined by Bradford assay after freeze-thawing using bovine serum albumin for a standard concentration curve. Generally, this procedure yielded > 300 mg per L culture. Protein purity was analyzed by sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel electrophoresis using 12% polyacrylamide gels.

Cloning, expression, purification, and storage of *Rgn*TDC

A codon-optimized copy of the *Ruminococcus gnavus* tryptophan decarboxylase (*Rgn*TDC) gene was purchased as a gBlock from Integrated DNA Technologies. This DNA fragment was inserted into a pET22b vector by the Gibson Assembly method.[1] BL21 (DE3) E. coli cells were subsequently transformed with the resulting cyclized DNA product via electroporation. After 30 min of recovery in LB media at 37 °C, cells were plated onto LB plates with 100 μg/mL ampicillin (AMP) and incubated overnight. Single colonies were used to inoculate 5 mL TB + 100 μg/mL

AMP (TB-AMP), which were grown overnight at 37 °C, 200 rpm. Expression cultures, typically 1 L of TB-AMP were inoculated from these starter cultures and shaken (180 rpm) at 37 °C. After 3.5 to 4 hours (OD_{600} > 1.5), the expression cultures were chilled on ice. After 45 min on ice, expression was induced with 1 mM IPTG, and the cultures were supplemented with 0.5 mM indole. Cultures were expressed overnight at 23 °C with shaking at 180 rpm. Cells were then harvested by centrifugation at 4300xg at 4 °C for 15 min. Cell pellets were frozen and stored at -20 °C until purification.

To purify TDC, cell pellets were thawed on ice and then resuspended in lysis buffer (50 mM potassium phosphate buffer (pH = 8.0), 1 mg/ml Hen Egg White Lysozyme (GoldBio), 0.2 mg/ml DNaseI (GoldBio), 1 mM $MgCl_2$, and 200 µM pyridoxal 5′-phosphate (PLP)). A volume of 4 mL of lysis buffer per gram of wet cell pellet was used. After 45 min of shaking at 37 °C, cells were sonicated with a ½ in. tip for 10 min (1 s on; 1 s off). The resulting lysate was then spun down at 75,000×g to pellet cell debris. Ni/NTA beads (GoldBio) were added to a gravity column and the lysis supernatant ran over the bead bed for purification by Ni-affinity chromatography. The column was washed with 4 column volumes of 20 mM imidazole, 50 mM potassium phosphate buffer (pH = 8.0). Washing with higher concentrations of imidazole resulted in slow protein elution. TDC was eluted with 250 mM imidazole, 50 mM potassium phosphate buffer (pH = 8.0). Elution of the desired protein product was monitored by the disappearance of its bright yellow color (resulting from the release of TDC) from the column. The protein product was dialyzed to < 50 µM imidazole in 50 mM Tris-HCl buffer (pH = 8.01) or 50 mM potassium phosphate buffer (pH = 8.08). Purified enzyme was flash frozen in pellet form by pipetting enzyme dropwise into a crystallization dish filled with liquid nitrogen. The enzyme was transferred to a plastic conical and stored at -80 °C until further use. Frozen pellets were thawed at room temperature and centrifuged before use. The concentration of protein was determined by Bradford assay after freeze-thawing using bovine serum albumin for a standard concentration curve. Generally, this

procedure yielded > 100 mg per L culture. Protein purity was analyzed by sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel electrophoresis using 12% polyacrylamide gels.

Cloning, expression, purification, and storage of VlmD

A codon-optimized copy of the *Kitasatospora setae valine decarboxylase* (VlmD) gene was purchased as a gBlock from Integrated DNA Technologies. This DNA fragment was inserted into a pET22b vector by the Gibson Assembly method.[1] BL21 (DE3) E. coli cells were subsequently transformed with the resulting cyclized DNA product via electroporation. After 30 min of recovery in LB media at 37 °C, cells were plated onto LB plates with 100 µg/mL ampicillin (AMP) and incubated overnight. Single colonies were used to inoculate 5 mL TB + 100 µg/mL AMP (TB-AMP), which were grown overnight at 37 °C, 200 rpm. 0.5 L of TB-AMP was inoculated from these starter cultures and shaken (180 rpm) at 37 °C. After 3 hours (OD_{600} ~ 1.0), the expression culture was chilled on ice. After 1 h on ice, expression was induced with 1 mM IPTG, and the cultures were supplemented with 0.5 mM indole. Cultures were expressed overnight at 23 °C with shaking at 180 rpm. Cells were then harvested by centrifugation at 4300xg at 4 °C for 15 min. Cell pellets were frozen and stored at -20 °C until purification.

To purify VlmD, cell pellets were thawed on ice and then resuspended in lysis buffer (50 mM potassium phosphate buffer (pH = 7.0), 1 mg/ml Hen Egg White Lysozyme (GoldBio), 0.2 mg/ml DNaseI (GoldBio), 1 mM $MgCl_2$, and 200 µM pyridoxal 5′-phosphate (PLP)). A volume of 4 mL of lysis buffer per gram of wet cell pellet was used along with 2.5% BugBuster. After 1 h of shaking at 37 °C, the lysate was spun down at 75,000×g to pellet cell debris. Ni/NTA beads (GoldBio) were added to a gravity column and the lysis supernatant ran over the bead bed for purification by Ni-affinity chromatography. The column was washed with 4 column volumes of 20 mM imidazole, 50 mM potassium phosphate buffer (pH = 7.0). VlmD was eluted with 250 mM imidazole, 50 mM potassium phosphate buffer (pH = 7.0). Elution of the desired protein product was monitored by the disappearance of its bright yellow color (resulting from the release

of VlmD) from the column. The protein product was dialyzed to < 50 μM imidazole in 50 mM potassium phosphate buffer (pH = 7.0). Purified enzyme was flash frozen in pellet form by pipetting enzyme dropwise into a crystallization dish filled with liquid nitrogen. The enzyme was transferred to a plastic conical and stored at -80 °C until further use. Frozen pellets were thawed at room temperature and centrifuged before use. The concentration of protein was determined by Bradford assay after freeze-thawing using bovine serum albumin for a standard concentration curve. This procedure yielded 70 mg per L culture.

Cloning, expression, purification, and storage of SOI

A codon-optimized copy of the *Pseudomonas sp.* VLB120 *StyC* (SOI) gene was purchased as a gBlock from Integrated DNA Technologies. This DNA fragment was inserted into a pET22b vector by the Gibson Assembly method. BL21 (DE3) *E. coli* cells were subsequently transformed with the cyclized DNA product via electroporation. After 45 min of recovery in TB media at 37 °C, cells were plated onto LB plates with 100 μg/mL AMP and incubated overnight. Following initial cloning into the pET22b vector, the gene encoding SOI was transferred to a pBAD vector (p15A origin, KAN resistant).

Single colonies were used to inoculate 10 mL TB-KAN, which was grown overnight at 37 °C, 200 rpm. Expression cultures, typically 1 L of TB-KAN, were inoculated with starter cultures (1% inoculum) and shaken (200 rpm) at 37 °C. After ~3 hours (OD_{600} = ~0.6), the expression cultures were chilled on ice. After 30 min on ice, expression of the protein was induced with 0.2% w/v arabinose. The cultures were expressed for 16-24 hours at 20 °C with shaking at 200 rpm. Cells were then harvested by centrifugation at 4,300×*g* at 4 °C for 10 min. Cell pellets were either freeze dried via lyophilization or extruded from a syringe and flash frozen in liquid nitrogen as small pellets to give flash frozen wet cells.

4. 4. 3. General methods

wt-RgnTDC and wt-VlmD activity assessment

Freeze-thawed *Rgn*TDC and VlmD was spin-filtered, and the supernatant used for the following reaction: 10 mM amino acid, 500 µM PLP, 50 mM KPi pH = 8.0, and 37.5 µM enzyme (final volume = 100 µl). Reactions were allowed to proceed for 16 h at 37 °C. Reactions were quenched via addition of 300 µl ACN followed by 300 µl H_2O, which were then centrifuged to aggregate enzyme, and injected onto UPLC-MS for product detection. Product m/z areas were used to determine relative amounts of product formed (Fig 2).

Initial RgnTDC – ObiH and VlmD – ObiH cascade reactions

Freeze-thawed VlmD and wt-*Rgn*TDC was spin-filtered, and the supernatant used for the following reaction: 10 mM benzaldehyde, 50 mM Thr, 400 µM PLP, 50 mM KPi pH = 8.0, and 10 µM VlmD/*Rgn*TDC (final volume = 100 µl). Reactions were allowed to proceed for 16 h at 37 °C. Reactions were quenched via addition of 300 µl ACN followed by 300 µl H_2O, which were then centrifuged to aggregate enzyme, and injected onto UPLC-MS for product detection. Product m/z areas were used to determine relative amounts of product formed (Fig 12).

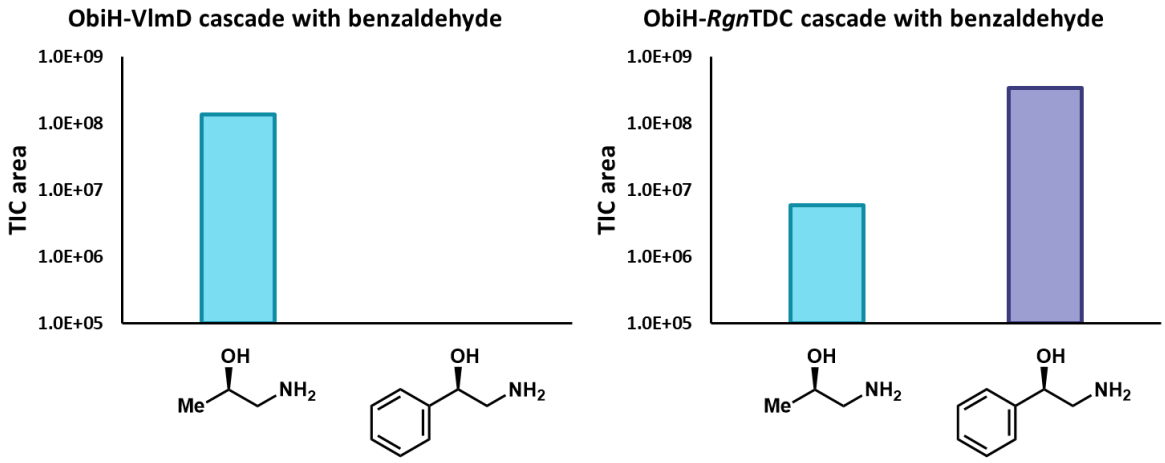


Figure 12. Products observed from cascade reactions with ObiH. 10 mM benzaldehyde, 50 mM Thr, 400 µM PLP, 50 mM KPi pH = 8.0, and 10 µM VlmD/*Rgn*TDC (final volume = 100 µl). Reactions were allowed to proceed for 16 h at 37 °C.

Mutagenesis of wt-RgnTDC

Based on a previously reported structure of an *Rgn*TDC-inhibitor complex (PDB ID: 4OBV), we modelled 3-hydroxyhomophenylalanine (**3b**) bound substrate into the active site (Fig 4a). H120 was chosen for site-saturation mutagenesis (SSM). Primers were purchased from Integrated DNA Technologies. For this site, three primers encoding the degenerate codons NDT, VHG, and TGG at the codon of interest were mixed in a 12:9:1 ratio, respectively.[35] The gene library was amplified first as two separate fragments and then combined via polymerase chain assembly (PCA) to form full-length RgnTDC gene mutagenized at the site of interest.[36] The corresponding gene were then inserted into a pET22b vector as described above and then transformed into BL21(DE3) E. coli cells and plated on LB + 100 µg/mL AMP agar plates.

Mutagenesis of H120N RgnTDC double/triple mutants

Starting from H120N *Rgn*TDC, we next mutagenized further 1st- and 2nd-sphere active-site residues: L126, Y312, and W349 (Fig 4a). Primers were ordered and SSM libraries prepared following the above protocol.

To investigate triple mutants, primers encoding H120N and L126M were used to add these mutations onto W349Y- and W349F-containing *Rgn*TDC gene sequences via PCA. These two gene sequences were transformed and expressed as above. *Rgn*TDC^{NMY} = H120N, L126M, W349Y; *Rgn*TDC^{NMF} = H120N, L126M, W349F.

Screening of RgnTDC libraries

Cell pellets were thawed and then resuspended in lysis buffer: 50 mM potassium phosphate buffer (pH = 8.0), 1 mg/mL Hen Egg White Lysozyme (GoldBio), 0.2 mg/mL DNaseI (GoldBio), 1 mM MgCl2, and 300 µM pyridoxal 5′-phosphate (PLP). A volume of 600 µL lysis buffer per well was used. After 45 min of shaking at 37 °C, the resulting lysate was then spun down at 4000 x*g* to pellet cell debris. Then, 140 µL of the resulting supernatant was added to 60 µL of a substrate mixture in a separate reaction plate. Final substrate concentrations are as follows: 2.5 mM $_{L}$-threonine (**1b**), 5 mM **3b,** and 2.5 mM 3-hydroxyleucine (**4b**), and 2.5 mM 2-amino-3-hydroxyoctanoic acid (**5b**). Reactions were incubated at 37 °C for 16 h, and then 100 µL reaction solution was quenched via addition of 200 µL acetonitrile and centrifuged at 4000 x*g* for 10 min. 200 µL of the quenched reaction mixture supernatant was filtered into a 96-well plate for UPLC-MS analysis. Data were collected on an Acquity UHPLC with an Acquity QDA MS detector (Waters) using a BEH C18 column (Waters). Product m/z ion counts were used to assess product formation from the reaction mixture (Fig 13-16).

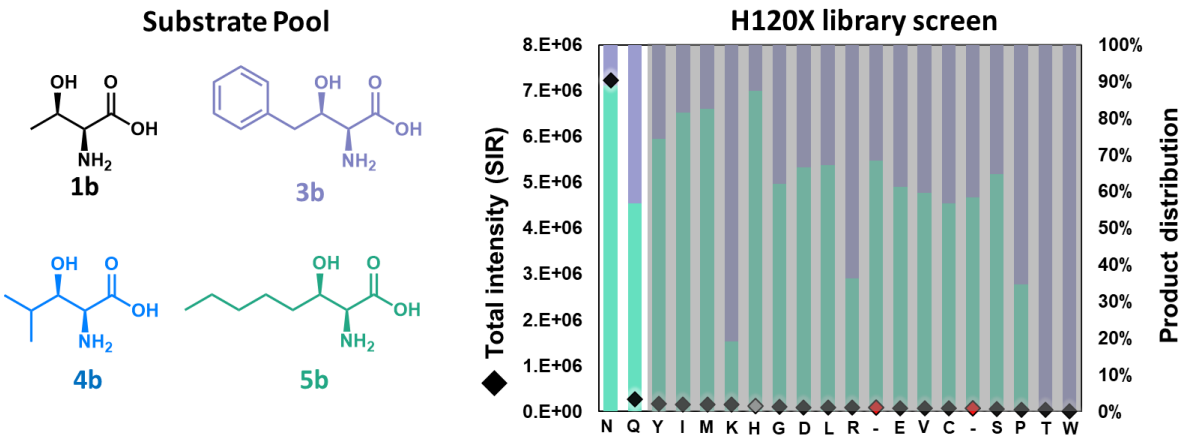


Figure 13. Screening results from the H120X site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention areas. The wild-type sequence is denoted by a grey diamond. Relative product amounts and mM total product were averaged from all wells with the given sequence. The greyed-out section of the graph indicates measurements that were indistinguishable from noise.

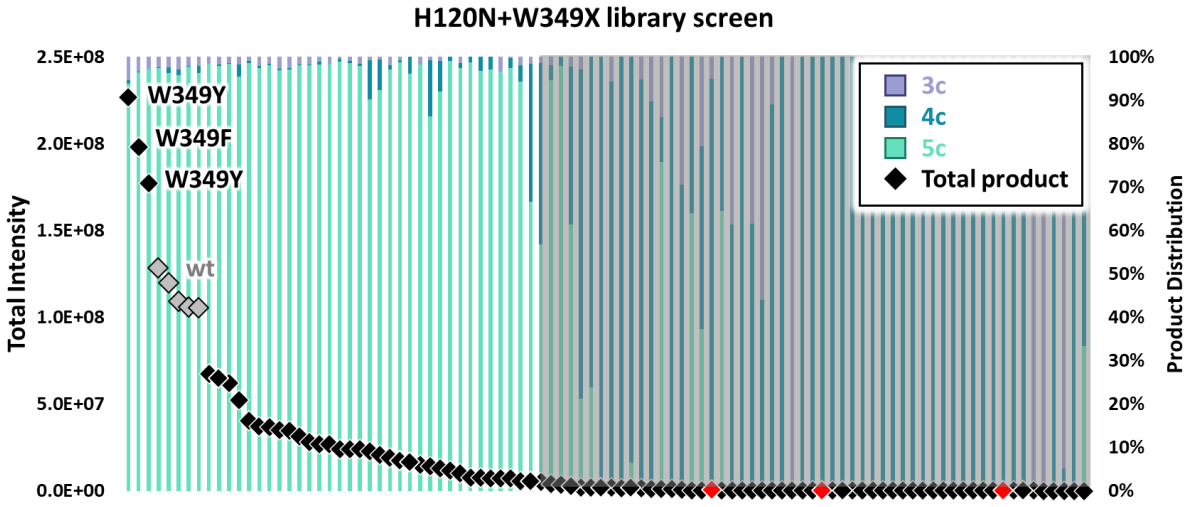


Figure 14. Screening results from the H120N + W349X site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention areas. The wild-type sequence is denoted by a grey diamond, and sequenced wells are labeled. Relative product amounts and mM total product were averaged from all wells with the given sequence. The greyed-out section of the graph indicates measurements that were indistinguishable from noise.

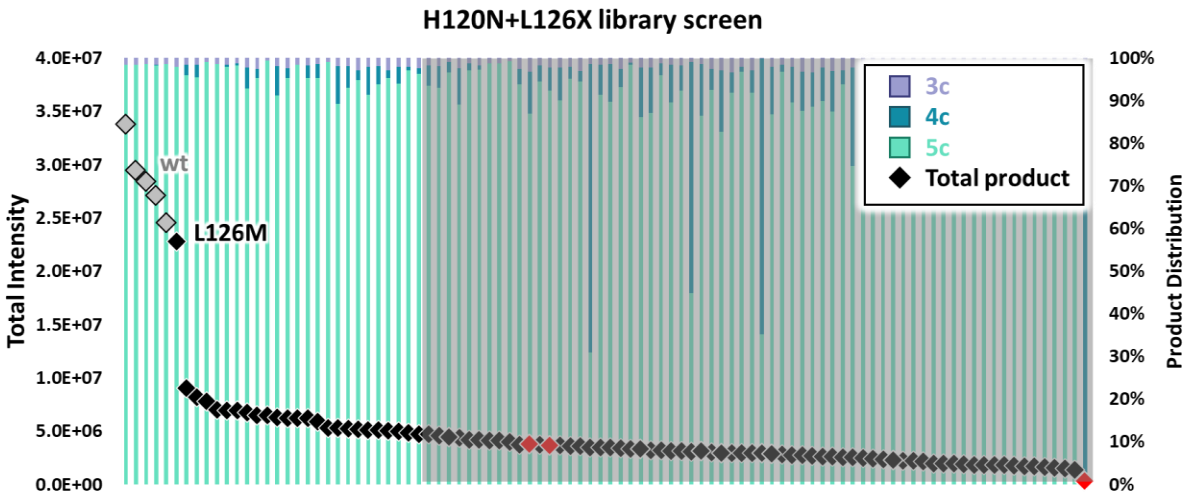


Figure 15. Screening results from the H120N + L126X site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention areas. The wild-type sequence is denoted by a grey diamond, and sequenced wells are labeled. Relative product amounts and mM total product were averaged from all wells with the given sequence. The greyed-out section of the graph indicates measurements that were indistinguishable from noise.

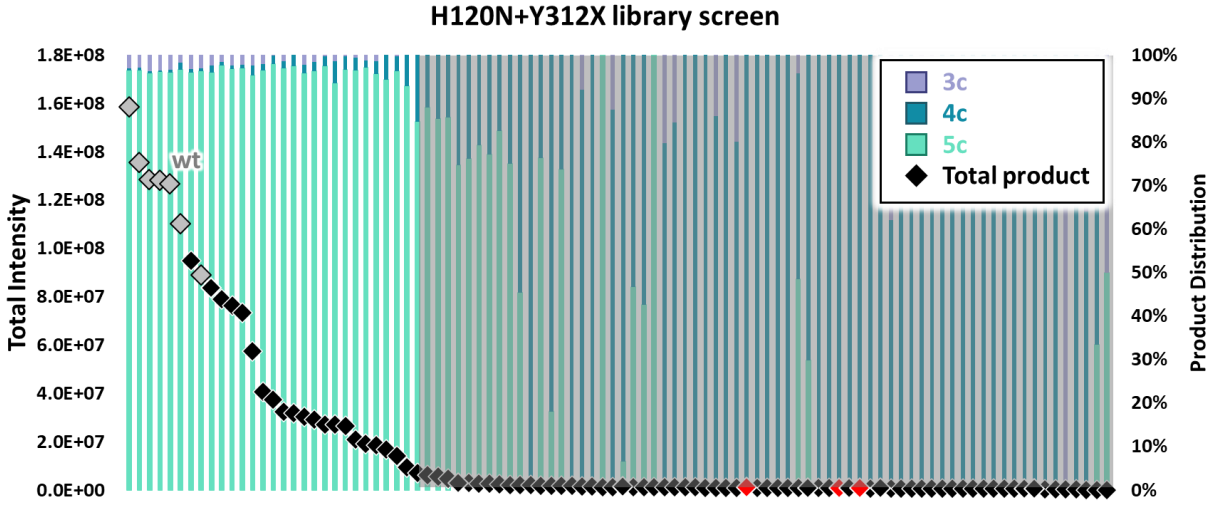


Figure 16. Screening results from the H120N + Y312X site-saturation mutagenesis library. Colored bars represent relative amounts of each product formed, and diamonds represent mM total product produced, as determined by single-ion retention areas. The wild-type sequence is denoted by a grey diamond, and sequenced wells are labeled. Relative product amounts and mM total product were averaged from all wells with the given sequence. The greyed-out section of the graph indicates measurements that were indistinguishable from noise.

Lysate verification of activated RgnTDC variants

Cells from wells containing putatively activated *Rgn*TDC variants were inoculated into 5 mL TB + AMP and grown for 16 h at 37 °C. Each variant strain culture was used to inoculate 500 mL TB + AMP, which was subsequently grown at 37 °C for 4 h, or until OD_{600} ~ 1.5. Cultures were chilled on ice for 45 min, and then induced with 1 mM IPTG and additionally supplemented with 0.5 mM indole. Cultures expressed for 16 h at 23 °C. The next morning, cells were pelleted via centrifugation at 4000 x*g* for 15 min. Supernatants were bleached and discarded and pellets were weighed and stored at -20 °C until lysis.

Cells were thawed and lysed as described above for protein purification. After pelleting of cell debris, however, the lysate supernatant was instead collected and used directly for reaction with **3b** and **5b.**

RgnTDC des-hydroxy amino acid activity comparison

Freeze-thawed *Rgn*TDC variants were spin-filtered, and the supernatants used for the following reaction: 10 mM amino acid, 400 µM PLP, 10 µM *Rgn*TDC (for Leu/homoPhe) or 0.05 µM *Rgn*TDC (for Trp/Phe), and 50 mM KPi pH = 8.0 (final volume = 100 µl). Reactions were allowed to proceed for 16 h at 37 °C. Reactions were quenched via addition of 300 µl ACN followed by 300 µl H_2O, which were then centrifuged to aggregate enzyme, and injected onto UPLC-MS for product detection. Product m/z areas were used to determine relative amounts of product formed (Leu/homoPhe) or UV-vis peak areas at 254 nm (Phe) or 280 nM (Trp) were used to determine relative conversions (Fig 5a).

RgnTDC diastereoselectivity for β-OH moiety

Freeze-thawed *Rgn*TDC variants were spin-filtered, and the supernatants used for the following reaction: 10 mM ***p*-Br, β-OH Phe (6b)** 80:20 d.r., 400 µM PLP, 10 µM *Rgn*TDC variant, and 50 mM KPi pH = 8.0 (final volume = 100 µl). Reactions were allowed to proceed for 15 min at 37

°C. Reactions were quenched via addition of 300 µl ACN followed by 300 µl H_2O, which were then centrifuged to aggregate enzyme, and injected onto UPLC-MS for product detection. UV-vis peak areas at 254 were used to determine relative concentrations of remaining amino acid starting material to observe changes in d.r. over the course of the reaction (Fig 8).

RgnTDC lineage amino acid specificity measurements

Freeze-thawed RgnTDC variants were spin-filtered, and the supernatants used for the following reactions: 2 mM **β-OH homoPhe** (**3b**), 2 mM **β-OH Hex** (**5b**), 0.2 mM ***p*-Br, β-OH Phe** (**6b**), and 2 mM **4-bromotryptophan** 400 µM PLP, 1 µM *Rgn*TDC variant, and 50 mM KPi pH = 7.5 (final volume = 100 µl). Reactions were conducted in duplicate for 16 h at 37 °C. Reactions were quenched using 500 µl ACN, centrifuged to pellet aggregated enzyme, and injected onto UPLC-MS for product detection. Product m/z's were used to quantify relative product abundances (Fig 4d).

For standard amino acid specificity determination, freeze-thawed RgnTDC variants were spin-filtered, and the supernatants used for the following reactions: 2 mM each substrate (**His, Leu, Tyr, Phe, Trp**), 400 µM PLP, 1 µM *Rgn*TDC variant, and 50 mM KPi pH = 8.0 (final volume = 100 µl). Reactions were conducted in duplicate for 10 min at 37 °C. d Reactions were quenched using 500 µl ACN, centrifuged to pellet aggregated enzyme, and injected onto UPLC-MS for product detection. Product m/z's were used to quantify relative product abundances (Fig 5b).

Optimization of ObiH-TDC cascade

Buffer conditions:

Freeze-thawed ObiH and *Rgn*TDC^{H120N} was spin-filtered, and the supernatants used for the following reaction: 25 mM 3,4-dichlorobenzaldehyde, 75 mM Thr, 400 µM PLP, 10 µM ObiH, 25 µM *Rgn*TDC^{H120N}, and 50 mM Tris/KPi with 5% MeOH (final volume = 250 µl). Buffer pH's ranged from pH = 7.0 to pH = 8.5. Reactions were conducted for 16 h at 37 °C at 180 RPM's. Reactions were quenched using 500 µl ACN, centrifuged to pellet aggregated enzyme, and injected onto UPLC-MS for product detection. Product m/z areas were used to quantify relative product abundances. Each reaction condition was conducted in duplicate.

Additionally, *Rgn*TDC activity was assayed independently under different buffer conditions. Freeze-thawed *Rgn*TDC^{NMY} was spin-filtered, and the supernatant used for the following reaction: 10 mM β-OH homoPhe (**3b**), 400 µM PLP, 10 µM *Rgn*TDC^{NMY}, and 50 mM Tris/KPi (final volume = 100 µl). Buffer pH's ranged from pH = 7.0 to pH = 8.5. Reactions were conducted for 16 h at 37 °C at 180 RPM's. Reactions were quenched using 300 µl ACN, followed by 300 µl H_2O, centrifuged to pellet aggregated enzyme, and injected onto UPLC-MS for product detection. Product m/z areas were used to quantify relative product abundances. Each reaction condition was conducted in duplicate (Fig 7a-7b)

Co-solvent conditions:

Freeze-thawed ObiH and H120N *Rgn*TDC was spin-filtered, and the supernatants used for the following reaction: 25 mM 4-chlorobenzaldehyde, 80 mM Thr, 200 µM PLP, 50 mM Tris pH =

8.0, 32 μM ObiH, 12 μM H120N *Rgn*TDC (final volume = 250 μl), with differing v/v % cosolvents (0 – 20% MeOH/DMSO/ACN). Reactions were conducted for 16 h at 37 °C without agitation. Reactions were quenched using 500 μl ACN, centrifuged to pellet aggregated enzyme, and injected onto UPLC-MS for product detection (Fig 7c).

Calculation of UPLC yields using Marfey's derivatization

General Procedure: All reactions were done in triplicate on analytical scale (250 μL). Stocks of Thr were made in either 50 mM Tris pH 8.0 or 50 mM potassium phosphate pH 8.0 and aldehydes were prepared in MeOH. Derivatized amino acid product quantitation was performed by PDA analysis, integrating the area under the product curve at 340 nm and correcting by comparing to the internal standard (arginine) peak area. To calculate product concentrations, a standard curve was generated by subjecting stock solutions of L-Phe in buffer using the identical procedure used to process and derivatize enzymatic reaction solutions, in duplicate. These curves were used to calculate the concentrations of 1,2-amino alcohol product in solution.

In a microcentrifuge tube, 30 μL of quenched reaction mix (1 equiv., 1.0 mM final total amines from unreacted Thr, Tris, unreacted β-hydroxy-α-amino acid, and 1,2-amino alcohol product) was added to a solution of 120 μL of 0.25 mM arginine dissolved in 18.75 mM $NaHCO_3$ (7.5 equiv., 0.1 mM final concentration arginine and 7.5 mM final concentration $NaHCO_3$) followed by addition of 150 μL 5 mM L-FDVA ((*S*)-1-fluoro-2-4-dinitrophenyl-5-L-valine amide) dissolved in ACN (2.5 equiv., 2.5 mM final concentration) to bring the total reaction volume to 300 μL. Each reaction was placed in a dark 37 °C incubator for 12 h, then quenched with 300 μL of 1:1 ACN:60 mM HCl (15 mM post-quench) before analyzing by UPLC-MS.

Determination of turnover numbers for TDC variants via Marfey's derivation:

Freeze-thawed RgnTDC variants were spin-filtered, and the supernatants used for the following reaction: 10 mM amino acid (**3b**, **5b,** or **6b**) 500 µM PLP, 50 µM *Rgn*TDC variant (for **3b** reaction), 20 µM *Rgn*TDC variant (for **5b** reaction), or 0.3 µM *Rgn*TDC variant (for **6b** reaction), and 50 mM KPi pH = 8 (final volume = 100 µl). Reactions were allowed to proceed in a 37 °C incubator for 16 h prior to quenching with 100 µL ACN. The samples were centrifuged to pellet aggregated enzyme and diluted an additional 1:10 with 50 mM KPi pH 8.0 to ensure the total amine concentration for the Marfey's derivatization was low.

In a microcentrifuge tube, 30 µL of diluted reaction mix (1 equiv., 0.05 mM final total amines from unreacted β-OH amino acid substrate, and formed 1,2-amino alcohol product) was added to a solution of 120 µL of 18.75 mM $NaHCO_3$ (150 equiv., 7.5 mM final concentration $NaHCO_3$) followed by addition of 150 µL 5 mM L-FDVA dissolved in ACN (50 equiv., 2.5 mM final concentration) to bring the total reaction volume to 300 µL. Each reaction was placed in a dark 50 °C thermomixer for 8 h, then quenched with 300 µL of 1:1 ACN:60 mM HCl (15 mM post-quench) before analyzing by UPLC-MS. The yield of the TDC reactions were estimated using the ratio of the derivatized amino alcohol product to the sum of the derivatized amino alcohol product and derivatized amino acid starting material. Total turnover numbers were calculated by taking the yield of the TDC reaction and multiplying it by the max TON (Fig 4e).

Determination of analytical substrate scope yields via Marfey's derivation:

For analytical cascade yields, freeze-thawed ObiH and RgnTDC^{NMY} were spin-filtered, and the supernatants used for the following reactions: 10 mM aldehyde/styrene oxide (**2a-22a**), 50 mM Thr, 400 µM PLP, 40 µM ObiH, 40 µM *Rgn*TDC^{NMY}, and 50 mM KPi pH = 8.0 (final volume of 100 µl). Reactions were conducted in triplicate for 16 h at 37 °C. d Reactions were quenched using 200 µl ACN, followed by 100 µl H2O, and centrifuged to pellet aggregated enzyme.

A subsequent Marfey's derivatization reaction was performed to assess UPLC yield of the produced 1,2-amino alcohols. To 8 µl of the quenched reaction solution was added 1 µl 20 mM arginine (final [Arg] = 0.2 mM), 41 µl Na_2CO_3 (final [Na_2CO_3] = 246 mM), and 50 µl 10 mM D-FDVA ((*S*)-1-fluoro-2-4-dinitrophenyl-5-L-valine amide). These reactions proceeded for 16 h at 37 °C. After 16 h, reactions were quenched via addition of 50 µL 1 M HCl and diluted to 1 mL with 1:1 ACN:H_2O. Reaction solutions were analyzed on UPLC-MS by looking at absorbance at 340 nm and comparing relative product peak areas to the derivatized Arg peak area (Fig 9; Fig 17)

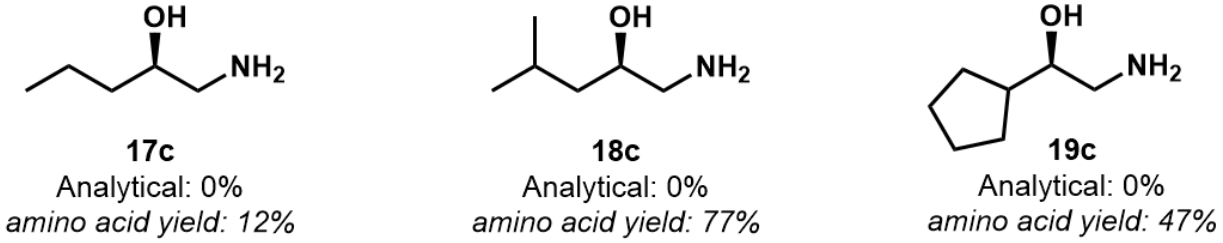


Figure 17. Additional substrates tested in ObiH – *Rgn*TDC^{NMY} cascade.

Calculation of product enantiomeric ratios via Marfey's derivation:

A Marfey's derivatization reaction was performed to assess the enantiomeric ratios of formed 1,2-amino alcohols following the above procedure with purified material. However, the Marfey's reaction was conducted both in the presence of L-FDVA or D-FDVA ((*R*)-1-fluoro-2-4-dinitrophenyl-5-L-valine amide) to ensure enantiomer separation via our UPLC-MS method. Enantiomeric ratio (e.r.) was determined by Marfey-derivatized product peaks integrated at 340 nM.

Stereoisomer determination of 22c:

Freeze-thawed RgnTDC^{NMY} was spin-filtered, and the supernatant used for the following reactions: 10 mM **22b** (65:35 *anti*:*syn* for the γ-Me and β-OH groups; prepared as previously described (CITE)), 400 μM PLP, 0 or 10 μM *Rgn*TDC^{NMY}, 50 mM KPi pH = 8.0 (final volume = 100 μl). Reactions were conducted in duplicate for 16 h at 37 °C. Reactions were quenched using 300 μl ACN, followed by 300 μl H_2O, centrifuged to pellet aggregated enzyme, and injected onto UPLC-MS for product detection. Product m/z's were used to identify substrate and product peaks. Selective amino acid stereoisomer conversion was observed only for the minor *syn* amino acid diastereomer along with appearance of a new product peak, which we attribute to be the *syn* 1,2-amino alcohol (Fig 10).

4. 4. 5 Synthetic methods

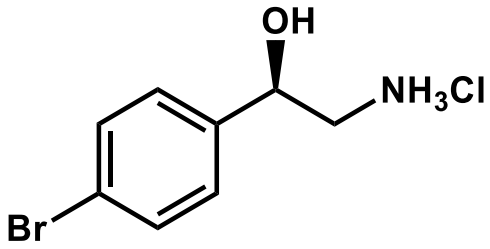
Biocatalytic cascade synthesis of 1,2-amino alcohols

*****We found that purification of the protonated amines was generally easier than purification of the deprotonated amines on C18. We additionally observed high Tris (buffer) retention under basic purification conditions, whereas Tris eluted early as a single peak under acidic conditions. For these reasons, product isolations from C18 were generally performed at pH < 1 via acidification using HCl.

Although acid/base extraction was used for the purification of several of the below compounds, considering the surprising simplicity of C18 purification and high resulting compound purity, we opted for column chromatography purification of many products.

Additionally, since *Rgn*TDC^{H120N} was equally active on phenylserine amino acids, we used this enzyme for many syntheses instead of *Rgn*TDC^{NMY} TDC simply due to catalyst availability.

(*R*)-2-amino-1-(4-bromophenyl)ethan-1-ol (6c)



94.7 mg (0.512 mmol) of 4-bromobenzaldehyde was dissolved in 2.5 mL of MeOH. To this, 440 mg (3.7 mmol) of L-threonine (Thr) was added, and the solution was diluted to ~45 mL with 50 mM Tris-HCl pH = 8.01. 0.5 mL of 20 mM PLP was then added, such that the final concentration was 0.200 µM. Freeze-thawed ObiH and H120N *Rgn*TDC was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 9 µM (0.1 mol% catalyst); [H120N] = 9 µM (0.1 mol% catalyst). Final [aldehyde] = 10 mM; final [Thr] = 74 mM. The reaction was then incubated at 37°C for 40 h.

The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 40 mL acetonitrile (ACN), and centrifuged at 4000 x*g* for 10 min. The pellet was washed with 10 mL 1:1 ACN:H2O. The combined supernatants were concentrated via rotary evaporation down to ~5 mL, and 6 M HCl was added to bring the pH < 1. The resulting suspension was injected onto a 30 g C18 flash chromatography column and the product eluted early at 1% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry.

Later fractions were observed to contain additional product. These fractions were pooled and concentrated via rotary evaporation. This aqueous solution was then basified with 6 M NaOH and extracted twice with 50 mL EtOAc. The desired product was then extracted from the EtOAc layer twice with 50 mL dilute HCl, which were subsequently combined and concentrated via rotary evaporation. This solution was then frozen @ -80°C and lyophilized until dry.

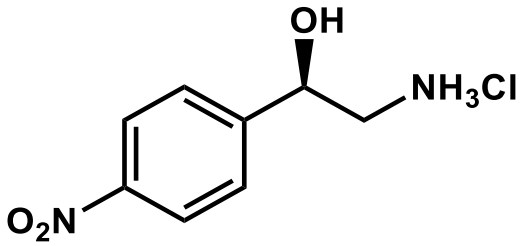
From both purifications, the product was isolated as the HCl salt as a white powder (82 mg: 67 mg from C18, 14 mg from extraction) in 63% yield. Enantiopurity was determined by subsequent Marfey's analysis to be 99:1 e.r. (*R:S*).

^{1}H NMR (500 MHz, MeOD) δ 7.60 – 7.52 (m, 2H), 7.42 – 7.34 (m, 2H), 4.88 (dd, *J* = 9.6, 3.3 Hz, 1H), 3.17 (dd, *J* = 12.8, 3.3 Hz, 1H), 3.00 (dd, *J* = 12.8, 9.6 Hz, 1H).

^{13}C NMR (126 MHz, MeOD) δ 140.34, 131.46, 127.54, 121.65, 68.92, 45.66.

MS/ESI m/z for $[M+H]^+$; C8H10BrNO; calculated 216.0019, observed 216.0018.

(*R*)-2-amino-1-(4-nitrophenyl)ethan-1-ol (7c)



82.8 mg (0.548 mmol) of 4-nitrobenzaldehyde was dissolved in 1.25 mL of methanol (MeOH). To this, 261 mg (2.2 mmol) of L-threonine (Thr) was added, and the solution was diluted to ~45 mL with 50 mM Tris-HCl pH = 8.01. 1 mL of 20 mM PLP was then added, such that the final concentration was 0.400 μM. Freeze-thawed ObiH and H120N *Rgn*TDC was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 40 μM (0.4 mol% catalyst); [H120N] = 10 μM (0.1 mol% catalyst). Final [aldehyde] = 11 mM; final [Thr] = 44 mM. The reaction was then incubated at 37°C for 16 h @ 180 rpm.

The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 50 mL acetonitrile (ACN), and centrifuged at 4000 x*g* for 10 min. The pellet was washed with 10 mL 1:1 ACN:H2O. The combined supernatants were concentrated via rotary

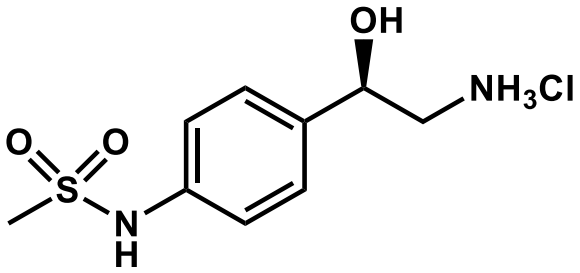
evaporation down to ~5 mL, and 6 M HCl was added to bring the pH < 1. The resulting suspension was injected onto a 30 g C18 flash chromatography column and the product eluted early at 1% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry. The product was isolated as the HCl salt as a white powder (105 mg) in 96% yield. Enantiopurity was determined by subsequent Marfey's analysis to be 97.5:2.5 e.r. (*R:S*).

^{1}H NMR (500 MHz, MeOD) δ 8.31 – 8.24 (m, 2H), 7.73 – 7.67 (m, 2H), 5.03 (dd, *J* = 9.4, 3.3 Hz, 1H), 3.24 (dd, *J* = 12.8, 3.3 Hz, 1H), 3.03 (dd, *J* = 12.8, 9.4 Hz, 1H).

^{13}C NMR (126 MHz, MeOD) δ 149.80, 149.33, 128.18, 124.80, 70.12, 46.83.

MS/ESI m/z for $[M+H]^+$; C8H10N2O3; calculated 183.0764, observed 183.0763

***(R)*-N-(4-(2-amino-1-hydroxyethyl)phenyl)methanesulfonamide (8c)**



96.6 mg (0.485 mmol) of 4-sulfonamidobenzaldehyde was dissolved in 2.5 mL of MeOH. To this, 433 mg (3.6 mmol) of Thr was added, and the solution was diluted to ~45 mL with 50 mM Tris-HCl pH = 8.01. 1 mL of 20 mM PLP was then added, such that the final concentration was 0.400 μM. Freeze-thawed ObiH and NMY *Rgn*TDC was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 50 μM (0.5 mol% catalyst); [NMY] = 14 μM (0.1 mol% catalyst). Final [aldehyde] = 10 mM; final [Thr] = 73 mM. The reaction was then incubated at 37°C for 40 h @ 180 rpm.

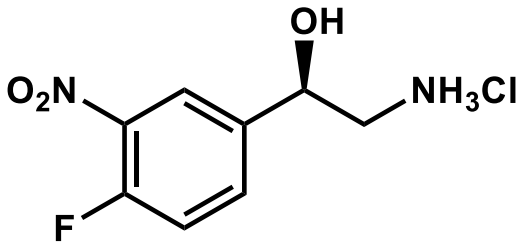
The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 50 mL ACN, and centrifuged at 4000 xg for 10 min. The supernatant was concentrated via rotary evaporation down to ~5 mL, and 6 M HCl was added to bring the pH < 1. The resulting suspension was injected onto a 30 g C18 flash chromatography column and the product eluted at 1% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry. The product was isolated as the HCl salt as a white crystalline powder (96 mg) in 74 % yield. Enantiopurity was determined by subsequent Marfey's analysis to be 96:4 e.r. (*R*).

^{1}H NMR (500 MHz, MeOD) δ 7.47 – 7.39 (m, 2H), 7.33 – 7.27 (m, 2H), 4.89 (dd, *J* = 9.6, 3.4 Hz, 1H), 3.16 (dd, *J* = 12.8, 3.4 Hz, 1H), 3.02 (dd, *J* = 12.7, 9.6 Hz, 1H), 2.97 (s, 3H).

^{13}C NMR (126 MHz, MeOD) δ 138.21, 137.12, 126.72, 120.26, 69.05, 47.09, 45.82, 37.88.

MS/ESI m/z for $[M+H]^+$; $C_9H_{14}N_2O_3S$; calculated 231.0798, observed 231.0795.

(*R*)-2-amino-1-(4-fluoro-3-nitrophenyl)ethan-1-ol (9c)



83.7 mg (0.495 mmol) of 4-fluoro-3-nitrobenzaldehyde was dissolved in 2.5 mL of MeOH. To this, 425 mg (3.6 mmol) of Thr was added, and the solution was diluted to ~45 mL with 50 mM KPi pH = 8.08. 1 mL of 20 mM PLP was then added, such that the final concentration was 0.400 µM. Freeze-thawed ObiH and NMY *Rgn*TDC was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 12 µM (0.1 mol% catalyst); [NMY] = 17 µM (0.2 mol% catalyst). Final [aldehyde] = 10 mM; final [Thr] = 71 mM. The reaction was then incubated at 37°C for 16 h @ 140 rpm.

The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 40 mL ACN, and centrifuged at 4000 xg for 10 min. The supernatant was concentrated via rotary evaporation down to ~5 mL, and 6 M HCl was added to bring the pH < 1. The resulting suspension was injected onto a 30 g C18 flash chromatography column and the product eluted at 1% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry. The product was isolated as the HCl salt as a white crystalline powder (80 mg) in 69 % yield. Enantiopurity was determined by subsequent Marfey's analysis to be 95.5:4.5 e.r. (*R*).

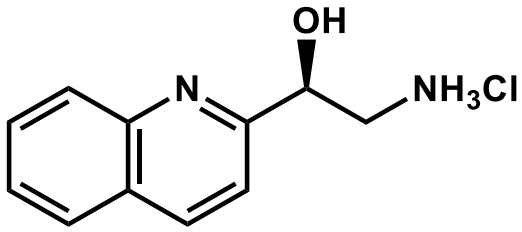
^{1}H NMR (500 MHz, MeOD) δ 8.22 (dd, *J* = 7.1, 2.3 Hz, 1H), 7.83 (ddd, *J* = 8.9, 4.3, 2.3 Hz, 1H), 7.49 (dd, *J* = 10.9, 8.6 Hz, 1H), 5.01 (dd, *J* = 9.2, 3.4 Hz, 1H), 3.25 (dd, *J* = 12.9, 3.4 Hz, 1H), 3.05 (dd, *J* = 12.9, 9.2 Hz, 1H).

^{13}C NMR (126 MHz, MeOD) δ 155.87, 153.78, 138.68, 138.64, 133.02, 132.95, 123.30, 123.28, 118.44, 118.27, 67.93, 45.42.

^{19}F NMR (377 MHz, MeOD) δ -121.80 (ddd, *J* = 11.4, 7.2, 4.2 Hz).

MS/ESI m/z for $[M+H]^+$; C8H9FN2O3; calculated 201.0670, observed 201.0669.

(*R*)-2-amino-1-(quinolin-3-yl)ethan-1-ol (10c)



79.5 mg (0.506 mmol) of quinoline-2-carboxaldehyde was dissolved in 2.5 mL of MeOH. To this, 420 mg (3.5 mmol) of Thr was added, and the solution was diluted to ~45 mL with 50 mM Tris pH = 8.01. 1 mL of 20 mM PLP was then added, such that the final concentration was 0.400 µM. Freeze-thawed ObiH and H120N *Rgn*TDC was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 63 µM (0.6 mol% catalyst); [NMY] = 40 µM (0.4 mol% catalyst). Final [aldehyde] = 10 mM; final [Thr] = 70 mM. The reaction was then incubated at 37°C for 40 h @ 140 rpm.

The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 40 mL ACN, and centrifuged at 4000 xg for 10 min. The supernatant was concentrated via rotary evaporation down to ~3 mL, and 6 M HCl was added to bring the pH < 1. The resulting suspension was injected onto a 30 g C18 flash chromatography column and the product eluted at 1% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry.

Later fractions were observed to contain significant product. These fractions were pooled and concentrated via rotary evaporation. This aqueous solution was then basified with 6 M NaOH and extracted thrice with 50 mL EtOAc. The desired product was then extracted from the EtOAc layer thrice with 50 mL dilute HCl, which were subsequently combined and concentrated via rotary evaporation. This solution was then frozen @ -80°C and lyophilized until dry.

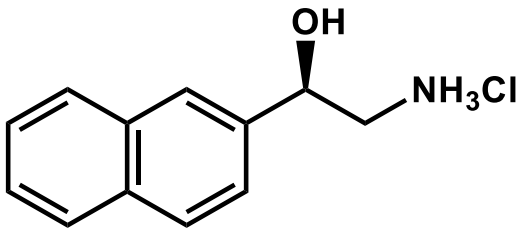
From both purifications, the product was isolated as the HCl salt as a yellow solid (104 mg: 66 mg from C18, 37 mg from extraction) in 91% yield. Enantiopurity was determined by subsequent Marfey's analysis to be 95.5:4.5 e.r. (*R*).

^{1}H NMR (500 MHz, MeOD) δ 8.84 (d, *J* = 8.6 Hz, 1H), 8.30 (d, *J* = 8.5 Hz, 1H), 8.19 (dd, *J* = 8.4, 1.4 Hz, 1H), 8.06 – 7.98 (m, 2H), 7.83 (ddd, *J* = 8.2, 6.9, 1.1 Hz, 1H), 5.42 (dd, *J* =f 8.6, 3.7 Hz, 1H), 3.58 (dd, *J* = 13.0, 3.8 Hz, 1H), 3.37 (dd, *J* = 13.0, 8.6 Hz, 1H).

^{13}C NMR (126 MHz, MeOD) δ 159.30, 142.73, 142.55, 132.57, 128.40, 128.36, 124.25, 119.05, 68.53, 43.99, 19.36.

MS/ESI m/z for [M+H]+ ; $C_{11}H_{12}N_2O$; calculated 189.1022, observed 189.1021.

(*R*)-2-amino-1-(naphthalen-2-yl)ethan-1-ol (11c)



84.4 mg (0.5 mmol) of 2-naphthaldehyde was dissolved in 5 mL of MeOH. To this, 298 mg (2.5 mmol) of Thr was added, and the solution was diluted to ~45 mL with 50 mM Tris-HCl pH = 8.01. 1 mL of 20 mM PLP was then added, such that the final concentration was 0.400 μM.

Freeze-thawed ObiH and H120N *Rgn*TDC was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 40 μM (0.4 mol% catalyst); [H120N] = 10 μM (0.1 mol% catalyst). Final [aldehyde] = 10 mM; final [Thr] = 50 mM. The reaction was then incubated at 37°C for 40 h @ 180 rpm. Insoluble 2-naphthaldehyde was visible at the bottom of the flask throughout the reaction.

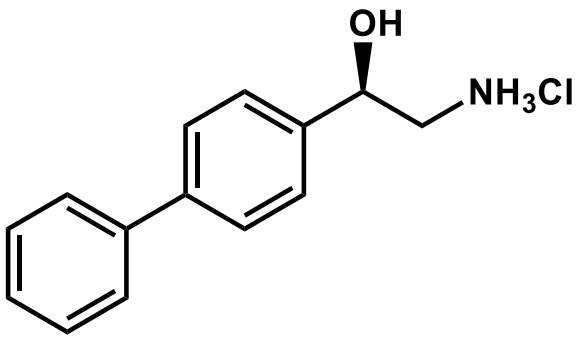
The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 50 mL ACN, and centrifuged at 4000 xg for 10 min. The pellet was washed with 10 mL 1:1 ACN:H2O. The combined supernatants were concentrated via rotary evaporation down to ~5 mL, and 6 M HCl was added to bring the pH < 1. The resulting suspension was injected onto a 30 g C18 flash chromatography column and the product eluted at 20% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry. The product was isolated as the HCl salt as a white powder (55 mg) in 49% yield. Hydration state analysis was performed to account for water mass. Enantiopurity was determined by subsequent Marfey's analysis to be 95:5 e.r. (*R*).

^{1}H NMR (500 MHz, MeOD) δ 7.97 – 7.81 (m, 4H), 7.58 – 7.42 (m, 3H), 5.06 (dd, *J* = 9.4, 3.4 Hz, 1H), 3.25 (dd, *J* = 12.8, 3.4 Hz, 1H), 3.12 (dd, *J* = 12.8, 9.4 Hz, 1H).

^{13}C NMR (126 MHz, MeOD) δ 138.33, 133.43, 133.39, 128.21, 127.61, 127.36, 126.07, 125.92, 124.63, 123.26, 69.66, 45.83.

MS/ESI m/z for $[M+H]^+$; C12H13NO; calculated 188.1070, observed 188.1070.

(*R*)-1-([1,1'-biphenyl]-4-yl)-2-aminoethan-1-ol (12c)



89.3 mg (0.490 mmol) of biphenyl-4-carboxaldehyde was dissolved in 2.5 mL of MeOH. To this, 426 mg (3.6 mmol) of Thr was added, and the solution was diluted to ~45 mL with 50 mM Tris pH = 8.01. 1 mL of 20 mM PLP was then added, such that the final concentration was 0.400 µM. Freeze-thawed ObiH and *Rgn*TDC^{NMY} was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 12 µM (0.1 mol% catalyst); [*Rgn*TDC^{NMY}] = 11 µM (0.1 mol% catalyst). Final [aldehyde] = 10 mM; final [Thr] = 72 mM. The reaction was then incubated at 37°C for 16 h @ 140 rpm.

The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 40 mL ACN, and centrifuged at 4000 xg for 10 min. It was observed that product and remaining starting aldehyde also crashed out at this stage. Therefore, the pellet was extracted with hot MeOH and 1 M HCl, solvent was rotary evaporated, and the remaining solution was injected onto a 30 g C18 flash chromatography column. The product eluted at 70% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry.

The supernatant of the original spin-down was concentrated via rotary evaporation down to ~3 mL, and 6 M HCl was added to bring the pH < 1. The resulting suspension was injected onto a

30 g C18 flash chromatography column and the product eluted at 10% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry.

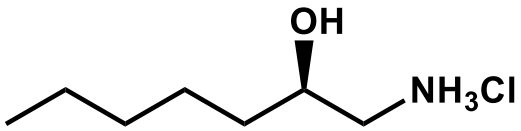
From both purifications, the product was isolated as the HCl salt as a white powder (23 mg: 14 mg from C18, 8 mg from extraction) in 21% yield. Enantiopurity was determined by subsequent Marfey's analysis to be 92:8 e.r. (*R*).

^{1}H NMR (500 MHz, MeOD) δ 7.73 – 7.58 (m, 4H), 7.57 – 7.49 (m, 2H), 7.49 – 7.41 (m, 2H), 7.40 – 7.33 (m, 1H), 4.96 (dd, *J* = 9.5, 3.4 Hz, 1H), 3.21 (dd, *J* = 12.7, 3.4 Hz, 1H), 3.08 (dd, *J* = 12.7, 9.6 Hz, 1H).

^{13}C NMR (126 MHz, MeOD) δ 141.22, 140.47, 139.98, 128.52, 127.15, 126.92, 126.54, 126.09, 69.32, 45.90.

MS/ESI m/z for [M+H]+ ; $C_{14}H_{15}NO$; calculated 214.1226, observed 214.1223.

(*R*)-1-aminoheptan-2-ol (5c)



50.0 mg (0.5 mmol) of hexanal was dissolved in 1 mL of MeOH. To this, 180 mg (1.5 mmol) of Thr was added, and the solution was diluted to ~14 mL with 50 mM KPi pH = 8.01. 0.4 mL of 20 mM PLP was then added, such that the final concentration was 0.400 μM. Freeze-thawed ObiH and *Rgn*TDC^{NMY} was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 33 μM (0.1 mol% catalyst); [*Rgn*TDC^{NMY}] = 33 μM (0.1 mol% catalyst). Final [aldehyde] = 25 mM; final [Thr] = 76 mM. The reaction was then incubated at 37°C for 40 h with no shaking.

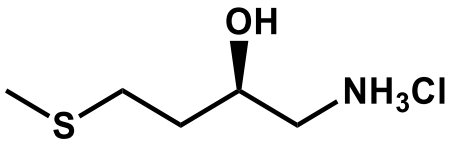
The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 50 mL ACN, and centrifuged at 4000 xg for 10 min. The pellet was washed with 20 mL 1:1 ACN:H2O. The combined supernatants were concentrated via rotary evaporation down to ~20 mL. The pH of the solution was adjusted via addition of 6 M NaOH to > 11 followed by 6x extraction with 50 mL EtOAc. The organic extractions were combined and evaporated to ~50 mL. The product was then extracted from the organic layer via addition of 3x 50 mL dilute HCl (pH ~3). The resulting aqueous layer was concentrated via rotary evaporation and was transferred to a pre-tared flask, frozen @ -80°C, and lyophilized until dry. The product was isolated as the HCl salt as a white powder (68 mg) in 81% yield. Enantiopurity was determined by subsequent Marfey's analysis to be > 99:1 e.r. (*R*).

^{1}H NMR (500 MHz, MeOD) δ 3.76 (dtd, *J* = 9.2, 6.1, 3.1 Hz, 1H), 3.03 (dd, *J* = 12.7, 3.1 Hz, 1H), 2.77 (dd, *J* = 12.7, 9.4 Hz, 1H), 1.56 – 1.27 (m, 8H), 0.98 – 0.91 (m, 3H).

^{13}C NMR (126 MHz, MeOD) δ 67.36, 44.69, 34.59, 31.45, 24.66, 22.20, 12.93.

MS/ESI m/z for $[M+H]^+$; C7H17NO; calculated 132.1383, observed 132.1382.

(*R*)-1-amino-4-(methylthio)butan-2-ol (14c)



52.1 mg (0.500 mmol) of methional was dissolved in 1 mL of MeOH. To this, 181 mg (1.5 mmol) of Thr was added, and the solution was diluted to ~12 mL with 50 mM KPi pH = 8.08. 200 µL of 20 mM PLP was then added, such that the final concentration was 0.200 µM. Freeze-thawed ObiH and *Rgn*TDC^{NMY} was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 25 µM (0.1 mol% catalyst); [*Rgn*TDC^{NMY}] = 50 µM (0.2 mol% catalyst). Final [aldehyde] = 25 mM; final [Thr] = 76 mM. The reaction was then incubated at 37°C for 36 h @ 200 rpm.

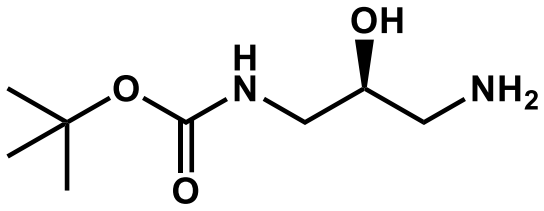
The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 25 mL ACN, and centrifuged at 4000 xg for 10 min. The supernatant was concentrated via rotary evaporation down to ~15 mL. The pH of the solution was adjusted via addition of 6 M NaOH to > 11 followed by 5x extraction with 50 mL EtOAc. The organic extractions were combined and evaporated to ~50 mL. The product was then extracted from the organic layer via addition of 5x 50 mL dilute HCl (pH ~3). The resulting aqueous layer was concentrated via rotary evaporation and was transferred to a pre-tared flask, frozen @ -80°C, and lyophilized until dry. The product was isolated as the HCl salt as a white crystalline powder (25 mg) in 29 % yield. Enantiopurity was determined by subsequent Marfey's analysis to be > 99:1 e.r. (*R*).

^{1}H NMR (500 MHz, MeOD) δ 3.94 (tt, *J* = 7.9, 3.5 Hz, 1H), 3.11 – 3.02 (m, 1H), 2.83 (dd, *J* = 12.6, 9.2 Hz, 1H), 2.72 – 2.57 (m, 2H), 2.12 (s, 3H), 1.77 (hd, *J* = 7.4, 3.7 Hz, 2H).

^{13}C NMR (126 MHz, MeOD) δ 66.13, 44.64, 33.94, 29.30, 13.92.

MS/ESI m/z for [M+H]+ ; $C_5H_{13}NOS$; calculated 136.0791, observed 136.0789.

tert-butyl (*S*)-(3-amino-2-hydroxypropyl)carbamate (15c)



80.6 mg (0.506 mmol) of N-Boc aminoacetaldehyde was dissolved in 1 mL of MeOH. *This substrate was a waxy, sticky solid and did not dissolve easily into solution. To this, 180 mg (1.5 mmol) of Thr was added, and the solution was diluted to ~12 mL with 50 mM Tris pH = 8.01. 400 µL of 20 mM PLP was then added, such that the final concentration was 0.400 µM. Freeze-thawed ObiH and *Rgn*TDC^{NMY} was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 127 µM (0.5 mol% catalyst); [NMY] = 101 µM (0.4 mol% catalyst). Final [aldehyde] = 25 mM; final [Thr] = 76 mM. The reaction was then incubated at 37°C for 36 h @ 140 rpm.

The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 25 mL ACN, and centrifuged at 4000 xg for 10 min. The supernatant was concentrated via rotary evaporation down to ~4 mL. Due to the present Boc group, we opted not to use acid/base extraction, and instead injected directly on a 30 g C18 flash chromatography column. The product eluted at 100% MeOH with a small amount of aldehyde starting material. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and

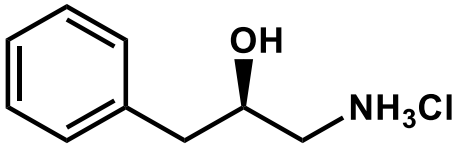
lyophilized until dry. The product was isolated as a tan powder (8.5 mg) in 7 % yield. Enantiopurity was determined by subsequent Marfey's analysis to be > 99:1 e.r. (*R*).

^{1}H NMR (500 MHz, MeOD) δ 3.62 (dtd, *J* = 7.3, 5.9, 4.0 Hz, 1H), 3.17 – 3.05 (m, 2H), 2.72 (dd, *J* = 13.3, 4.1 Hz, 1H), 2.58 (dd, *J* = 13.2, 7.3 Hz, 1H), 1.46 (s, 9H).

^{13}C NMR (126 MHz, MeOD) δ 157.35, 78.80, 70.98, 44.10, 43.39, 27.31.

MS/ESI m/z for $[M+H]^+$; $C_8H_{18}N_2O_3$; calculated 191.1390, observed 191.1387.

(*R*)-1-amino-3-phenylpropan-2-ol (20c)



60.1 mg (0.5 mmol) of styrene oxide was dissolved in 2.5 mL of MeOH. To this, 304 mg (2.55 mmol) of Thr was added, and the solution was diluted to ~45 mL with 50 mM KPi pH = 8.01. 1 mL of 20 mM PLP was then added, such that the final concentration was 0.400 μM. 50 mg frozen SOI cells were then added to the flask. Freeze-thawed ObiH and *Rgn*TDC^{NMY} was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 10 μM (0.1 mol% catalyst); [*Rgn*TDC^{NMY}] = 22 μM (0.2 mol% catalyst). Final [epoxide] = 10 mM; final [Thr] = 51 mM. The reaction was then incubated at 37°C for 40 h @ 180 rpm.

The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 50 mL ACN, and centrifuged at 4000 xg for 10 min. The pellet was washed with 10 mL 1:1 ACN:H2O. The combined supernatants were concentrated via rotary evaporation down to ~40 mL. The pH of the solution was adjusted via addition of 6 M NaOH to > 11 followed by 5x

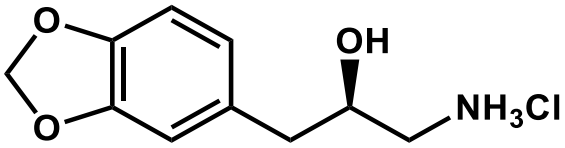
extraction with 75 mL EtOAc. The organic extractions were combined and evaporated to ~50 mL. The product was then extracted from the organic layer via addition of 3x 50 mL dilute HCl (pH ~3). The resulting aqueous layer was concentrated via rotary evaporation and was injected onto a 30 g C18 flash chromatography column. The product eluted at 1% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry. The product was isolated as the HCl salt as a white crystalline solid (25mg) in 33% yield. Enantiopurity was determined by subsequent Marfey's analysis to be > 99:1 e.r. (*R*).

^{1}H NMR (500 MHz, MeOD) δ 7.36 – 7.21 (m, 5H), 4.03 (dtd, *J* = 9.7, 6.6, 2.9 Hz, 1H), 3.01 (dd, *J* = 12.8, 3.0 Hz, 1H), 2.91 – 2.75 (m, 3H).

^{13}C NMR (126 MHz, MeOD) δ 137.24, 129.09, 128.19, 126.31, 68.66, 68.62, 44.07, 41.25.

MS/ESI m/z for $[M+H]^+$; C9H13NO; calculated 188.1070, observed 188.1070.

(*R*)-1-amino-3-(benzo[d][1,3]dioxol-5-yl)propan-2-ol (21c)



55 mg (0.328 mmol) of 5-(oxiran-2-yl)benzo[d][1,3]dioxole (as prepared by Meza *et al*.) was dissolved in 2.5 mL of MeOH. To this, 423 mg (3.55 mmol) of Thr was added, and the solution was diluted to ~45 mL with 50 mM KPi pH = 8.01. 1 mL of 20 mM PLP was then added, such that the final concentration was 0.400 µM. 50 mg frozen SOI cells were then added to the flask. Freeze-thawed ObiH and *Rgn*TDC^{NMY} was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 54 µM (0.6 mol% catalyst); [*Rgn*TDC^{NMY}] = 36 µM (0.5 mol% catalyst). Final [epoxide] = 10 mM; final [Thr] = 71 mM. The reaction was then incubated at 37°C for 40 h @ 140 rpm.

The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 50 mL ACN, and centrifuged at 4000 xg for 10 min. The pellet was washed with 10 mL 1:1 ACN:H2O. The combined supernatants were concentrated via rotary evaporation down to ~3 mL. The remaining liquid was injected onto a 30 g C18 flash chromatography column. The product eluted cleanly at 1% MeOH as well as in a mixture at 30%. The early product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry.

Later fractions were purified via acid/base extraction. These fractions were pooled and concentrated via rotary evaporation to remove MeOH. The pH of the solution was adjusted via addition of 6 M NaOH to > 11 followed by 3x extraction with 50 mL EtOAc. The organic extractions were combined and evaporated to ~50 mL. The product was then extracted from the organic layer via addition of 3x 50 mL dilute HCl (pH ~3). This aqueous solution was

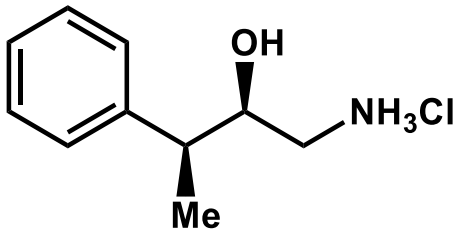
concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry. The product was isolated as the HCl salt as a white solid (20 mg: 4 mg from early flash chromatography fractions, 17 mg from extraction) in 27% yield. Enantiopurity was determined by subsequent Marfey's analysis to be > 99:1 e.r. (*R*).

^{1}H NMR (500 MHz, MeOD) δ 6.79 – 6.76 (m, 2H), 6.72 (dd, *J* = 7.9, 1.7 Hz, 1H), 5.93 (s, 2H), 3.95 (dtd, *J* = 9.7, 6.6, 3.0 Hz, 1H), 2.99 (dd, *J* = 12.7, 3.0 Hz, 1H), 2.81 – 2.68 (m, 3H).

^{13}C NMR (126 MHz, MeOD) δ 147.83, 146.46, 130.83, 122.05, 109.23, 107.78, 100.85, 68.69, 43.96, 40.88.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{13}NO_3$; calculated 196.0968, observed 196.0967.

(2*R*,3*S*)-1-amino-3-phenylbutan-2-ol (22c)



72.5 mg (0.540 mmol) of 2-methyl-2-phenyloxirane was dissolved in 2.5 mL of MeOH. To this, 322 mg (2.7 mmol) of Thr was added, and the solution was diluted to ~45 mL with 50 mM KPi pH = 8.01. 1 mL of 20 mM PLP was then added, such that the final concentration was 0.400 μM. 50 mg frozen SOI cells were then added to the flask. Freeze-thawed ObiH and *Rgn*TDC^{NMY} was spin-filtered, and the supernatants added to the flask with final concentrations of: [ObiH] = 43 μM (0.4 mol% catalyst); [*Rgn*TDC^{NMY}] = 30 μM (0.3 mol% catalyst). Final [epoxide] = 11 mM; final [Thr] = 54 mM. The reaction was then incubated at 37°C for 40 h @ 140 rpm.

The reaction solution was then heat-treated at 75°C for 20 minutes to aggregate protein, diluted with 50 mL ACN, and centrifuged at 4000 xg for 10 min. The combined supernatants were concentrated via rotary evaporation down to ~5 mL. The remaining liquid was injected onto a 30 g C18 flash chromatography column. The product eluted cleanly at 1% MeOH as well as in a mixture at 30%. The early product-containing fractions were pooled and concentrated via rotary evaporation, frozen at -80°C, and lyophilized until dry.

Later fractions were purified via acid/base extraction. These fractions were pooled and concentrated via rotary evaporation to remove MeOH. The pH of the solution was adjusted via addition of 6 M NaOH to > 11 followed by 4x extraction with 75 mL EtOAc. The organic extractions were combined and evaporated to ~75 mL. The product was then extracted from the organic layer via addition of 4x 75 mL dilute HCl (pH ~3). This aqueous solution was concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry. The product was isolated as the HCl salt as a white solid (43 mg: 25 mg from early flash chromatography fractions, 18 mg from extraction) in 39% yield. Only one diastereomer was observable via ^{1}H NMR. The methyl stereocenter was assigned by extension, as *Rgn*TDC^{NMY} was found to be active only on the corresponding (2*S*, 3R, 4S) amino acid stereoisomer (Fig 9). Enantiopurity was determined by subsequent Marfey's analysis to be > 99:1 e.r. and > 99:1 d.r. (2*R,* 3*S*).

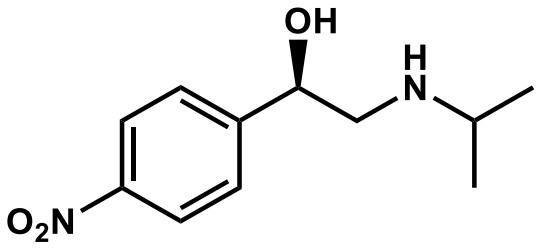
^{1}H NMR (500 MHz, MeOD) δ 7.37 – 7.31 (m, 2H), 7.26 (td, *J* = 6.8, 1.6 Hz, 3H), 3.81 (td, *J* = 9.2, 3.6 Hz, 1H), 2.78 – 2.57 (m, 3H), 1.40 (d, *J* = 7.0 Hz, 3H).

^{13}C NMR (126 MHz, MeOD) δ 143.09, 128.51, 127.17, 126.65, 72.26, 45.09, 43.40, 17.16.

MS/ESI m/z for $[M+H]^+$; $C_{10}H_{15}NO$; calculated 166.1226, observed 166.1226.

Chemoenzymatic syntheses

(*R*)-2-(isopropylamino)-1-(4-nitrophenyl)ethan-1-ol (7d)



Reductive amination was conducted following the procedure of Tajbakhsh et al. 2011 with slight modifications. 45.4 mg (0.207 mmol) of the (*R*)-2-amino-1-(4-nitrophenyl)ethan-1-ol HCl salt was added to a round bottom along with 0.80 mL acetone (1.1 mmol) and 2 mL trifluoroethanol (TFE). The round bottom was heated to 45°C upon which 113.49 mg $NaBH_4$ (3 mmol) was added. The solution was then stirred at 45°C for 16 h.

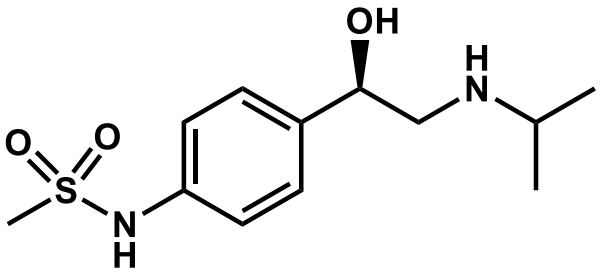
After the reaction, remaining insoluble $NaBH_4$ was filtered away via washing with MeOH and the filtrate was collected and evaporated via rotary evaporation to ~ 2 mL. The resulting solution was diluted with 3 mL H2O and injected onto a 30 g C18 flash chromatography column. The product eluted at 80% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry. The product was isolated as a white crystalline powder (49 mg) in quantitative yield. Hydration state analysis was performed to account for water mass.

^{1}H NMR (500 MHz, CD_3CN) δ 8.23 – 8.18 (m, 2H), 7.66 – 7.59 (m, 2H), 4.75 (dd, *J* = 8.5, 4.0 Hz, 1H), 2.94 (dd, *J* = 12.2, 4.1 Hz, 1H), 2.81 (hept, *J* = 6.2 Hz, 1H), 2.61 (dd, *J* = 12.2, 8.5 Hz, 1H), 1.05 (dd, *J* = 6.3, 1.8 Hz, 6H).

^{13}C NMR (126 MHz, CD_3CN) δ 152.31, 147.75, 127.40, 123.83, 71.76, 54.84, 48.84, 23.08, 22.76.

MS/ESI m/z for $[M+H]^+$; C11H16N2O3; calculated 225.1234 observed 225.1231.

(*R*)-N-(4-(1-hydroxy-2-(isopropylamino)ethyl)phenyl)methanesulfonamide (8d)



Reductive amination was conducted following the procedure of Tajbakhsh et al. 2011 with slight modifications. 77.5 mg (0.290 mmol) of the (*R*)-N-(4-(2-amino-1-hydroxyethyl)phenyl)-methanesulfonamide HCl salt was added to a round bottom along with 1.05 mL acetone (1.4 mmol) and 3 mL trifluoroethanol (TFE). The round bottom was heated to 45°C upon which 119.2 mg $NaBH_4$ (3 mmol) was added. The solution was then stirred at 45°C for 16 h.

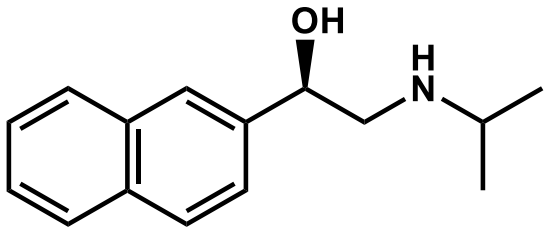
After the reaction, remaining insoluble $NaBH_4$ was filtered away via washing with MeOH and the filtrate was collected and evaporated via rotary evaporation to ~ 2 mL. The resulting solution was diluted with 3 mL H2O and injected onto a 30 g C18 flash chromatography column. The product eluted at 70% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen @ -80°C, and lyophilized until dry. The product was isolated as a white crystalline powder 82 mg) in quantitative yield.

^{1}H NMR (500 MHz, MeOD) δ 7.18 (d, *J* = 8.3 Hz, 2H), 7.09 (d, *J* = 8.5 Hz, 2H), 4.65 (t, *J* = 6.7 Hz, 1H), 2.89 – 2.83 (m, 1H), 2.82 (s, 3H), 2.77 – 2.74 (m, 2H), 1.09 (dd, *J* = 16.3, 6.3 Hz, 6H).

^{13}C NMR (126 MHz, MeOD) δ 147.91, 147.88, 133.33, 133.31, 126.11, 121.14, 72.15, 54.19, 37.54, 21.31, 21.07.

MS/ESI m/z for $[M+H]^+$; $C_{12}H_{20}N_2O_3S$; calculated 273.1267, observed 273.1264.

(*R*)-2-(isopropylamino)-1-(naphthalen-2-yl)ethan-1-ol (11d)



Reductive amination was conducted following the procedure of Tajbakhsh et al. 2011 with slight modifications. 50.0 mg (0.224 mmol) of the (*R*)-2-amino-1-(naphthalen-2-yl)ethan-1-ol HCl salt was added to a round bottom along with 0.22 mL acetone (3 mmol) and 3 mL trifluoroethanol (TFE). The round bottom was heated to 45°C upon which 127.8 mg $NaBH_4$ (3 mmol) was added. The solution was then stirred at 45°C for 16 h.

After the reaction, remaining insoluble $NaBH_4$ was filtered away via washing with MeOH and the filtrate was collected and evaporated via rotary evaporation to ~ 0.5 mL. The resulting solution was diluted with 2 mL H2O and injected onto a 30 g C18 flash chromatography column. The product eluted at 100% MeOH. Product-containing fractions were pooled and concentrated via rotary evaporation, frozen at -80°C, and lyophilized until dry. The product was isolated as a white crystalline powder (36.5 mg) in 71% yield.

^{1}H NMR (500 MHz, CD_3CN) δ 7.93 – 7.85 (m, 4H), 7.57 – 7.47 (m, 3H), 4.78 (dd, *J* = 8.7, 4.0 Hz, 1H), 2.97 (dd, *J* = 12.1, 4.0 Hz, 1H), 2.83 (hept, *J* = 6.3 Hz, 1H), 2.70 (dd, *J* = 12.1, 8.7 Hz, 1H), 1.06 (d, *J* = 6.2 Hz, 6H).

^{13}C NMR (126 MHz, CD_3CN) δ 142.11, 133.90, 133.40, 128.33, 128.25, 128.12, 126.62, 126.21, 125.02, 124.91, 72.61, 55.21, 48.88, 23.14, 22.85.

MS/ESI m/z for $[M+H]^+$; $C_{15}H_{19}NO$; calculated 230.1537, observed 230.1536.

Spectroscopic Data and NMR spectra

Small molecule crystallography

Data Collection

A colorless crystal with approximate dimensions 0.057 x 0.036 x 0.011 mm^3 was selected under oil under ambient conditions and attached to the tip of a MiTeGen MicroMount©. The crystal was mounted in a stream of cold nitrogen at 100(1) K and centered in the X-ray beam by using a video camera. The crystal evaluation and data collection were performed on a Bruker D8 VENTURE PhotonIII four-circle diffractometer with Cu Kα (λ = 1.54178 Å) radiation and the detector to crystal distance of 4.0 cm.[3] The initial cell constants were obtained from a 180° φ scan conducted at a 2θ = 50° angle with the exposure time of 1 second per frame. The reflections were successfully indexed by an automated indexing routine built in the APEX3 program. The final cell constants were calculated from a set of 8521 strong reflections from the actual data collection. The data were collected by using the full sphere data collection routine to survey the reciprocal space to the extent of a full sphere to a resolution of 0.78 Å. A total of 10187 data were harvested by collecting 15 sets of frames with 0.5° scans in ω and φ with an exposure time 10– 20 sec per frame. These highly redundant datasets were corrected for Lorentz and polarization effects. The absorption correction was based on fitting a function to the empirical transmission surface as sampled by multiple equivalent measurements.[4]

Structure Solution and Refinement

The systematic absences in the diffraction data were consistent for the space groups $P2_1$ and $P2_1/m$. The E-statistics strongly suggested the non-centrosymmetric space group $P2_1$ that yielded chemically reasonable and computationally stable results of refinement [5-8]. A successful solution by intrinsic phasing provided most non-hydrogen atoms from the E-map. The remaining non-hydrogen atoms were located with an alternating series of least-squares

cycles and difference Fourier maps. All nonhydrogen atoms were refined with anisotropic displacement coefficients. All hydrogen atoms (except those bound to non-C atoms) were included in the structure factor calculation at idealized positions and were allowed to ride on the neighboring atoms with relative isotropic displacement coefficients. The absolute configuration of the C2 chiral atom is *R*. The final least-squares refinement of 121 parameters against 2083 data resulted in residuals *R* (based on F^2 for $I \geq 2\sigma$) and *wR* (based on F^2 for all data) of 0.0350 and 0.0960, respectively. The final difference Fourier map was featureless.

Summary

Crystal Data for $[C_8H_{11}BrNO]^+Cl^-$ (*M* = 252.54 g/mol): monoclinic, space group $P2_1$ (no. 4), *a* = 5.823(2) Å, *b* = 7.573(2) Å, *c* = 11.404(4) Å, *β* = 102.507(16)°, *V* = 490.9(3) $Å^3$, *Z* = 2, *T* = 100.0 K, μ(Cu Kα) = 7.840 mm^{-1}, *Dcalc* = 1.708 g/cm^3, 7685 reflections measured (7.94° ≤ 2Θ ≤ 164.192°), 2083 unique (R_{int} = 0.0433, R_{sigma} = 0.0376) which were used in all calculations. The final R_1 was 0.0350 ($I > 2\sigma(I)$) and wR_2 was 0.0960 (all data).

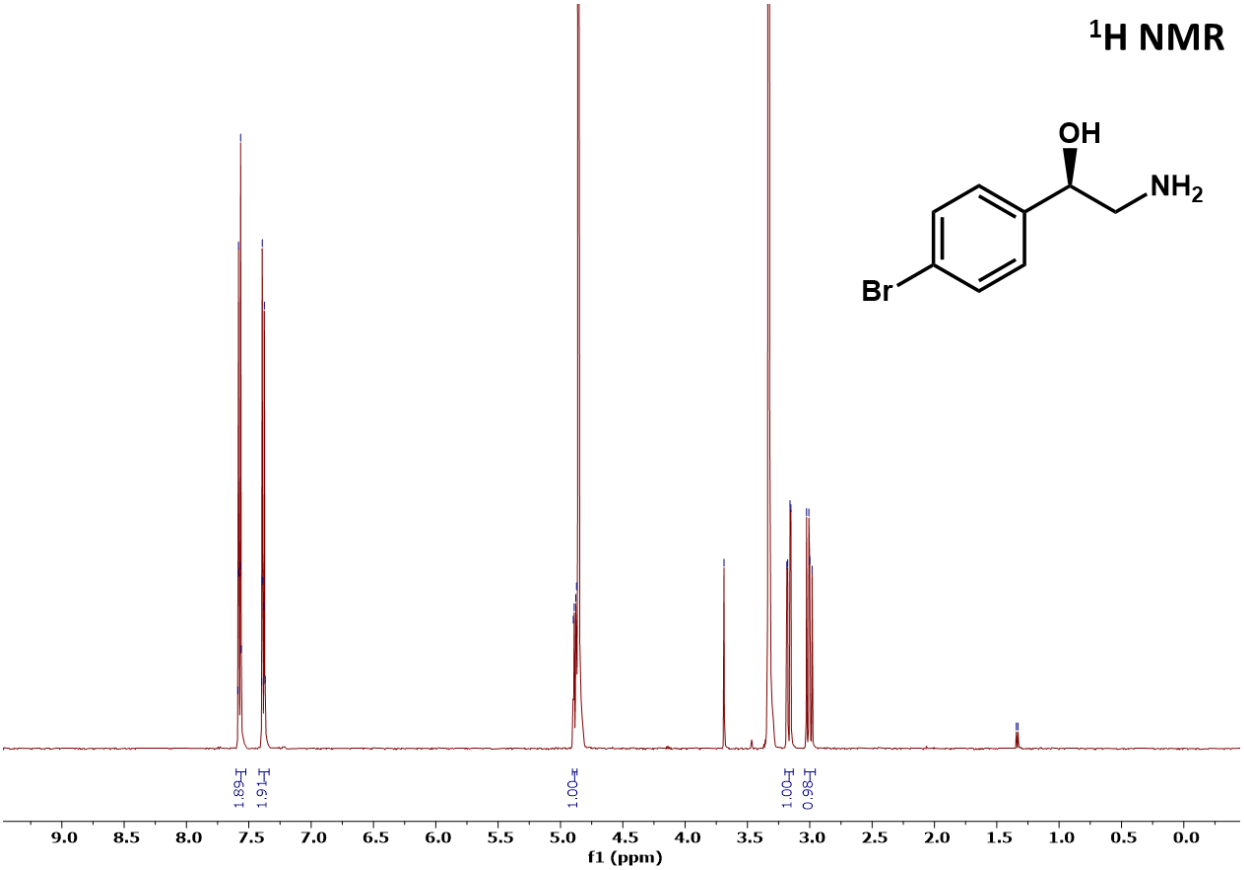


Figure 18. ^{1}H NMR spectrum of (*R*)-2-amino-1-(4-bromophenyl)ethan-1-ol.

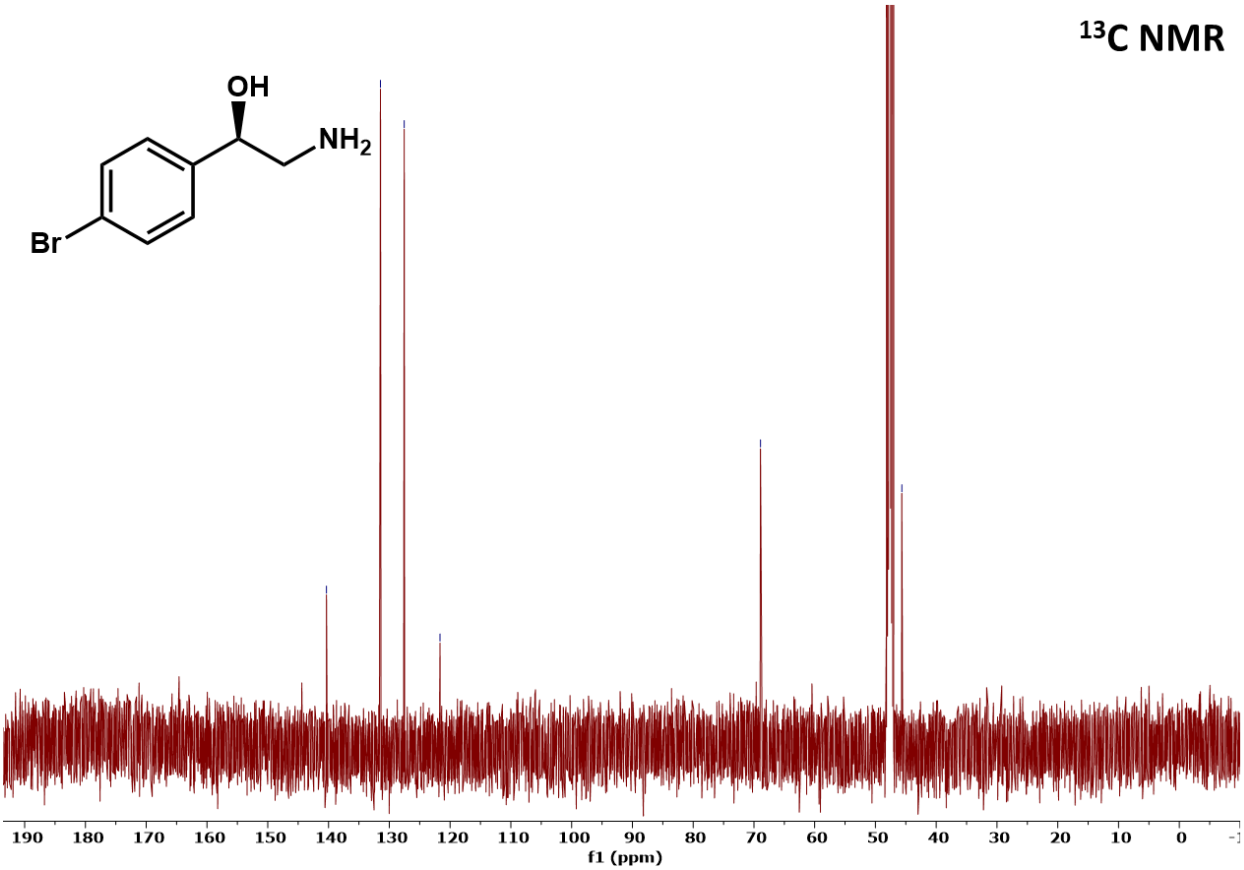


Figure 19. ^{13}C NMR spectrum of (*R*)-2-amino-1-(4-bromophenyl)ethan-1-ol.

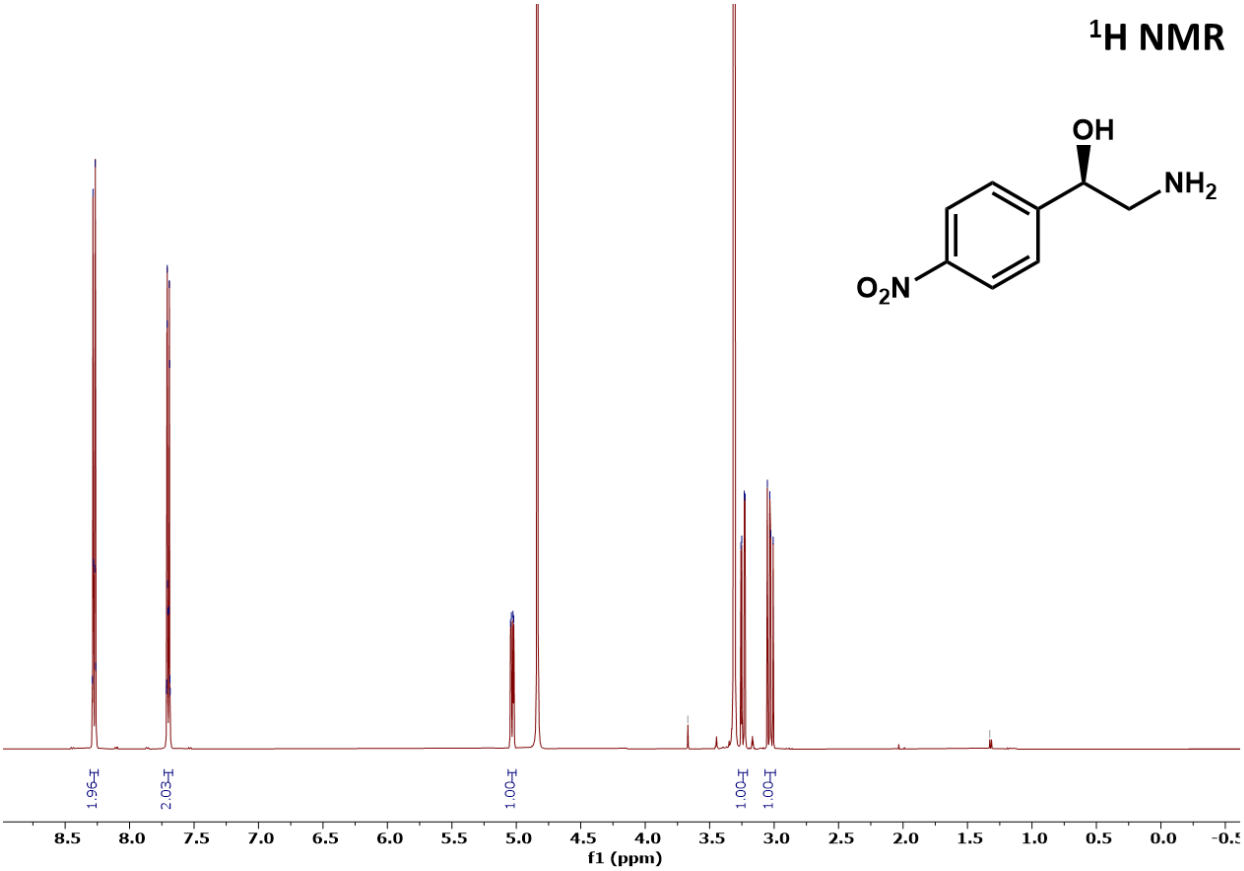


Figure 20. ^{1}H NMR spectrum of (*R*)-2-amino-1-(4-nitrophenyl)ethan-1-ol.

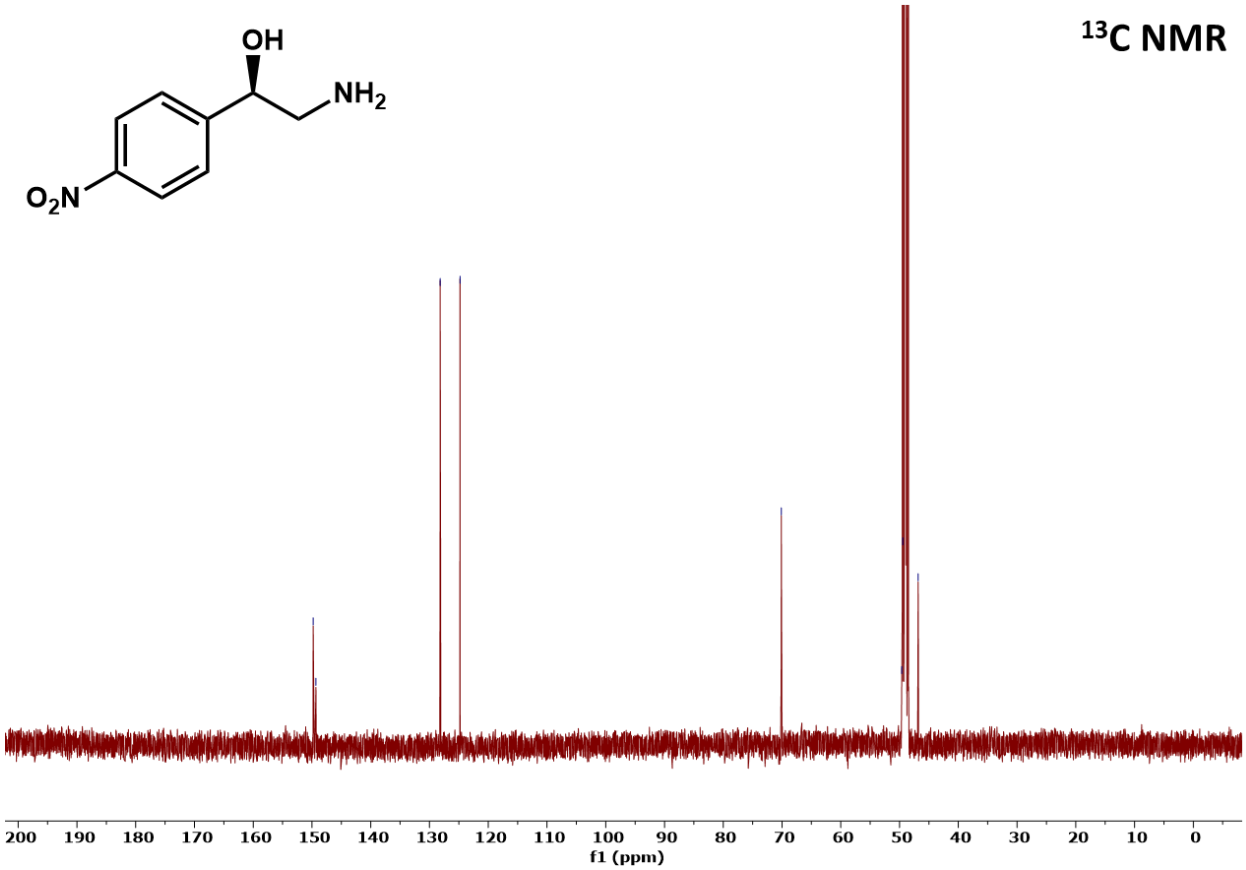


Figure 21. ^{13}C NMR spectrum of (*R*)-2-amino-1-(4-nitrophenyl)ethan-1-ol.

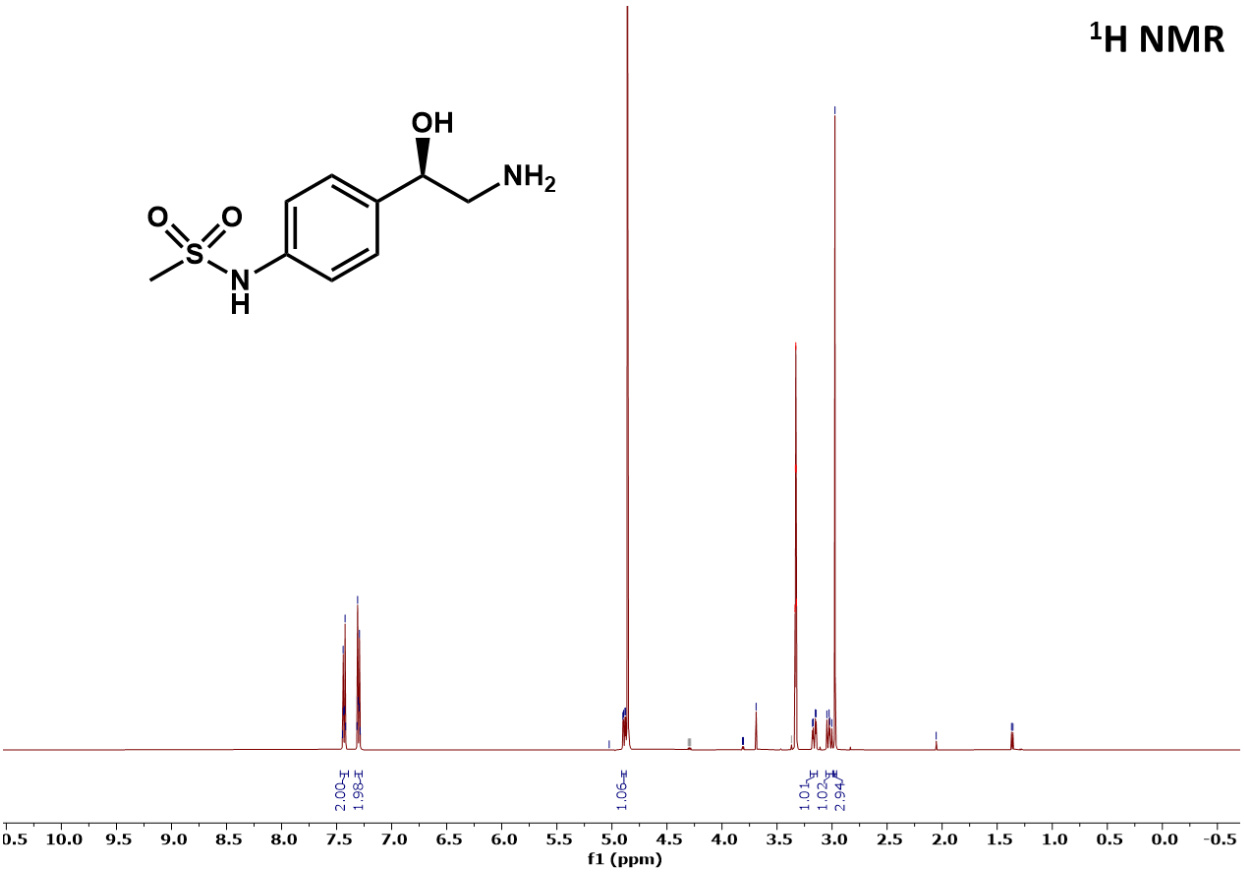


Figure 22. ^{1}H NMR spectrum of (*R*)-N-(4-(2-amino-1-hydroxyethyl)phenyl)methanesulfonamide.

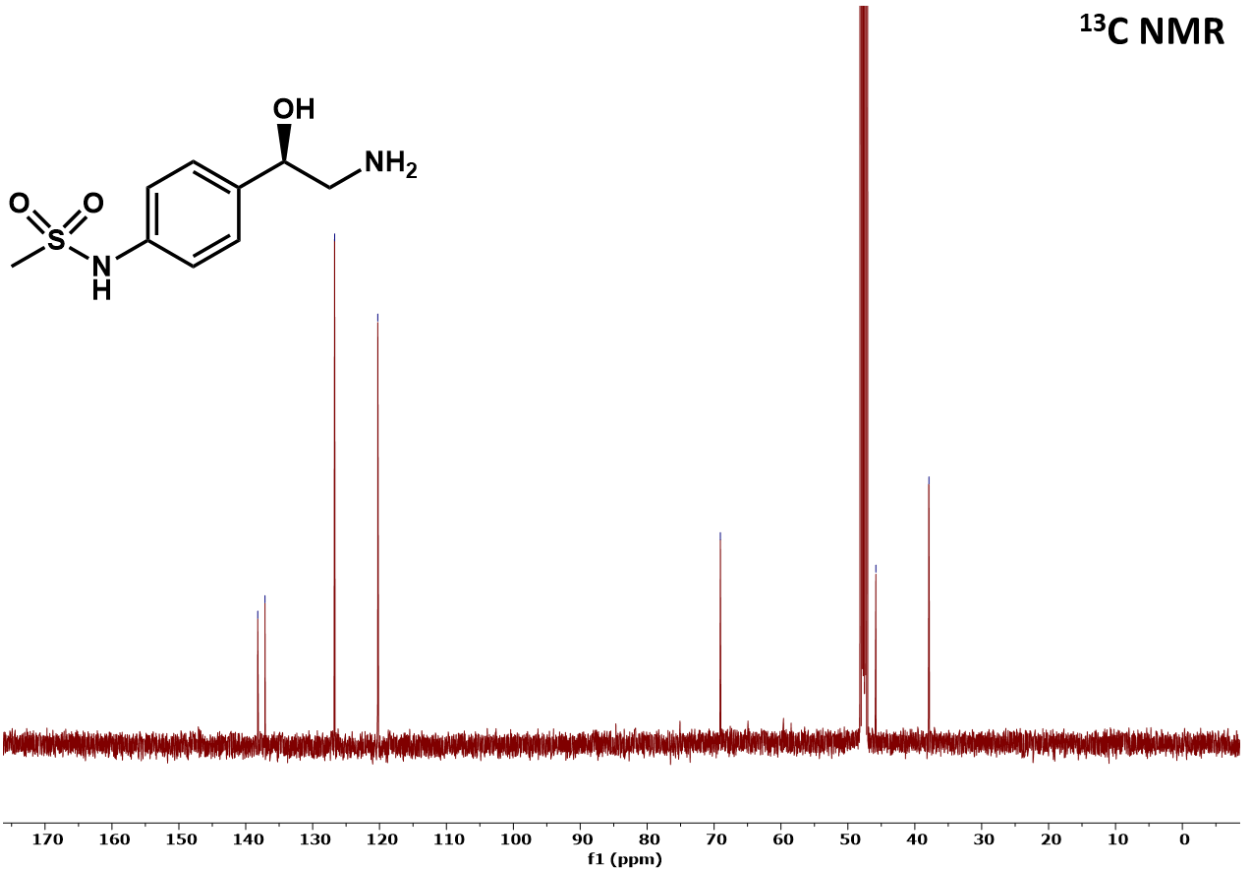


Figure 23. ^{13}C NMR spectrum of (*R*)-N-(4-(2-amino-1-hydroxyethyl)phenyl)methanesulfonamide.

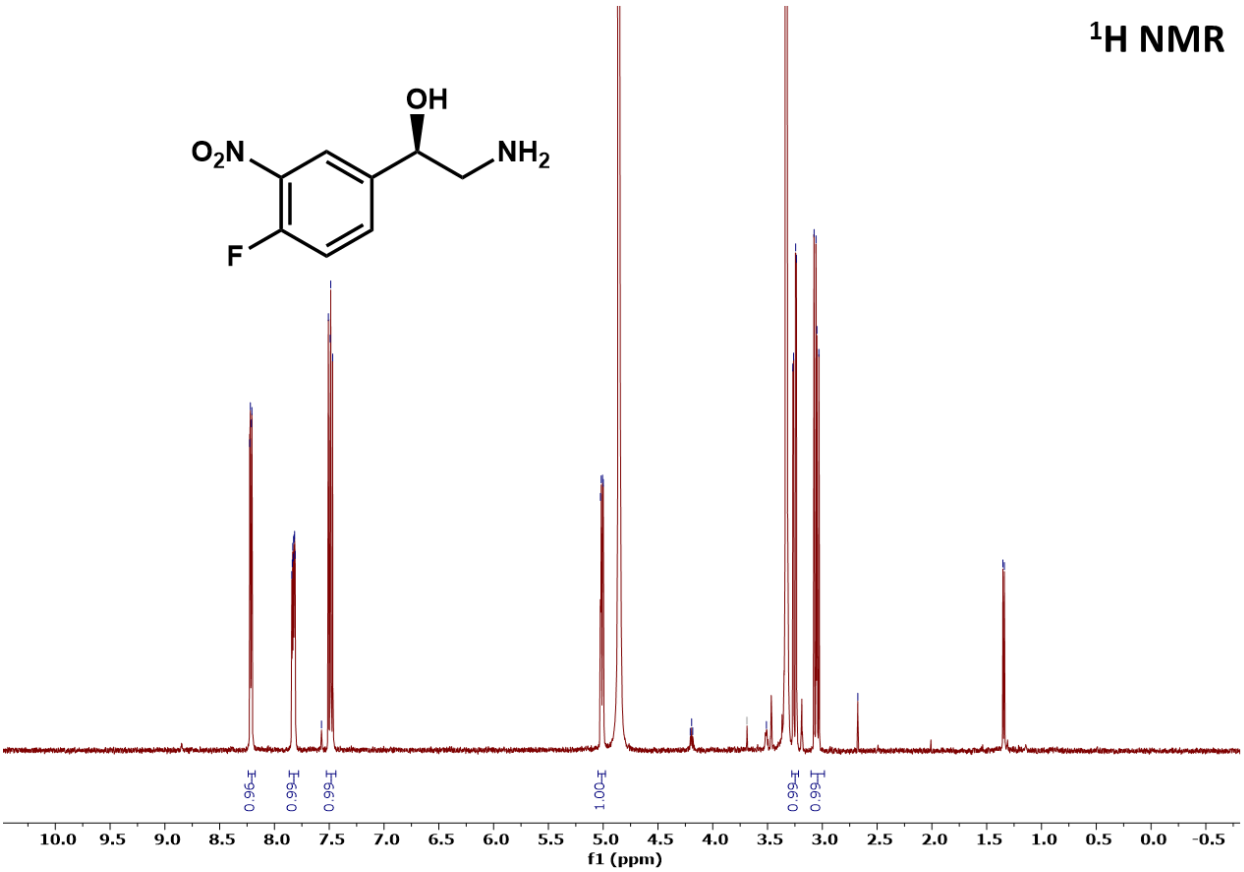


Figure 24. ^{1}H NMR spectrum of (*R*)-2-amino-1-(4-fluoro-3-nitrophenyl)ethan-1-ol.

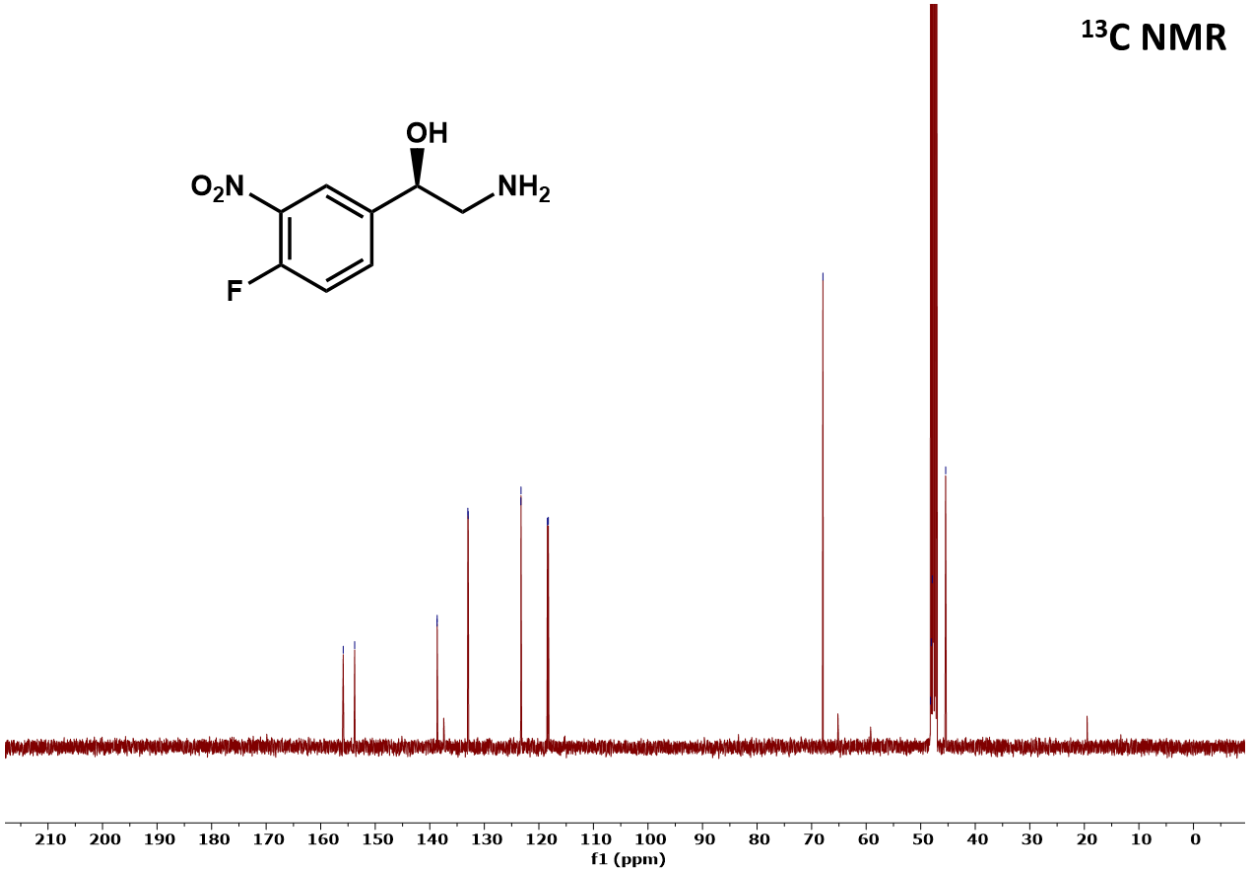


Figure 25. ^{13}C NMR spectrum of (*R*)-2-amino-1-(4-fluoro-3-nitrophenyl)ethan-1-ol.

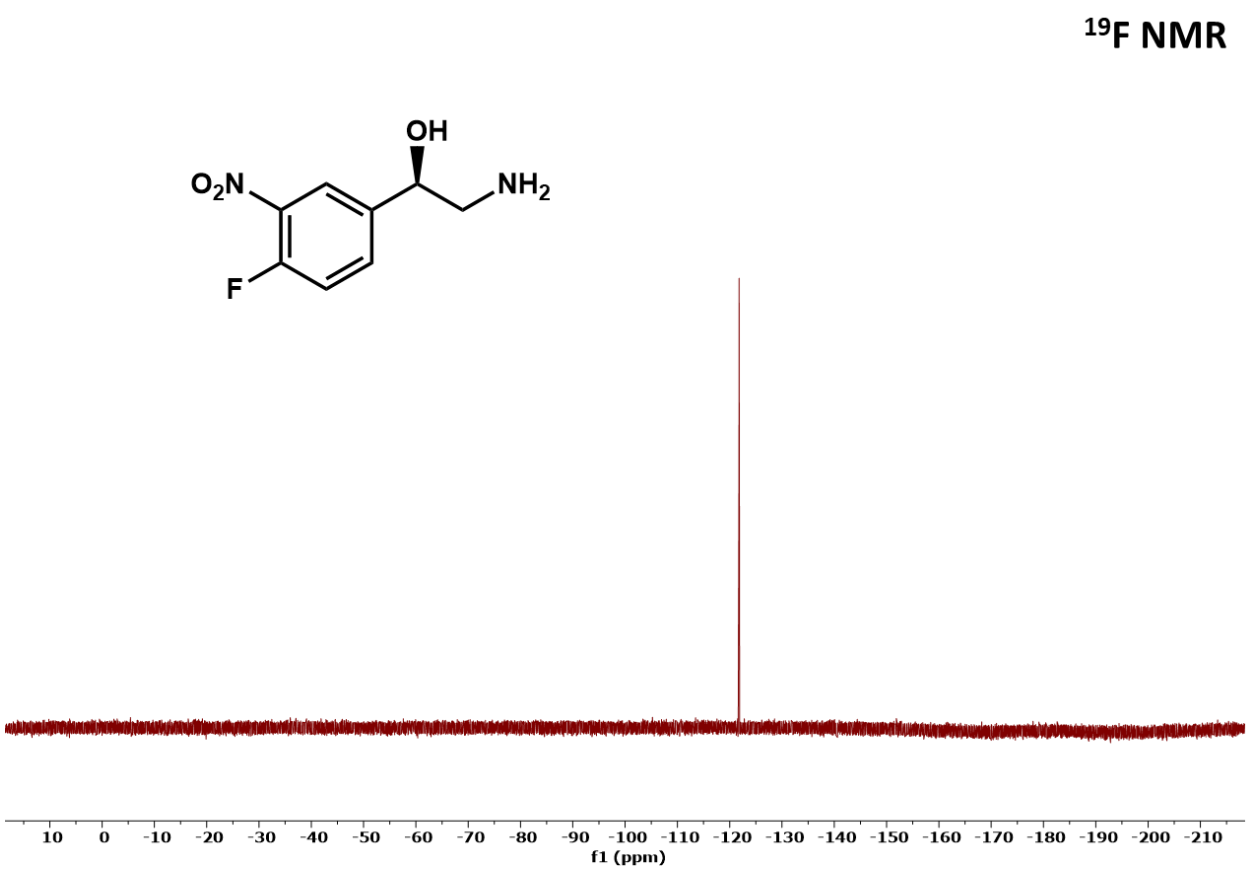


Figure 26. ^{19}F NMR spectrum of (*R*)-2-amino-1-(4-fluoro-3-nitrophenyl)ethan-1-ol.

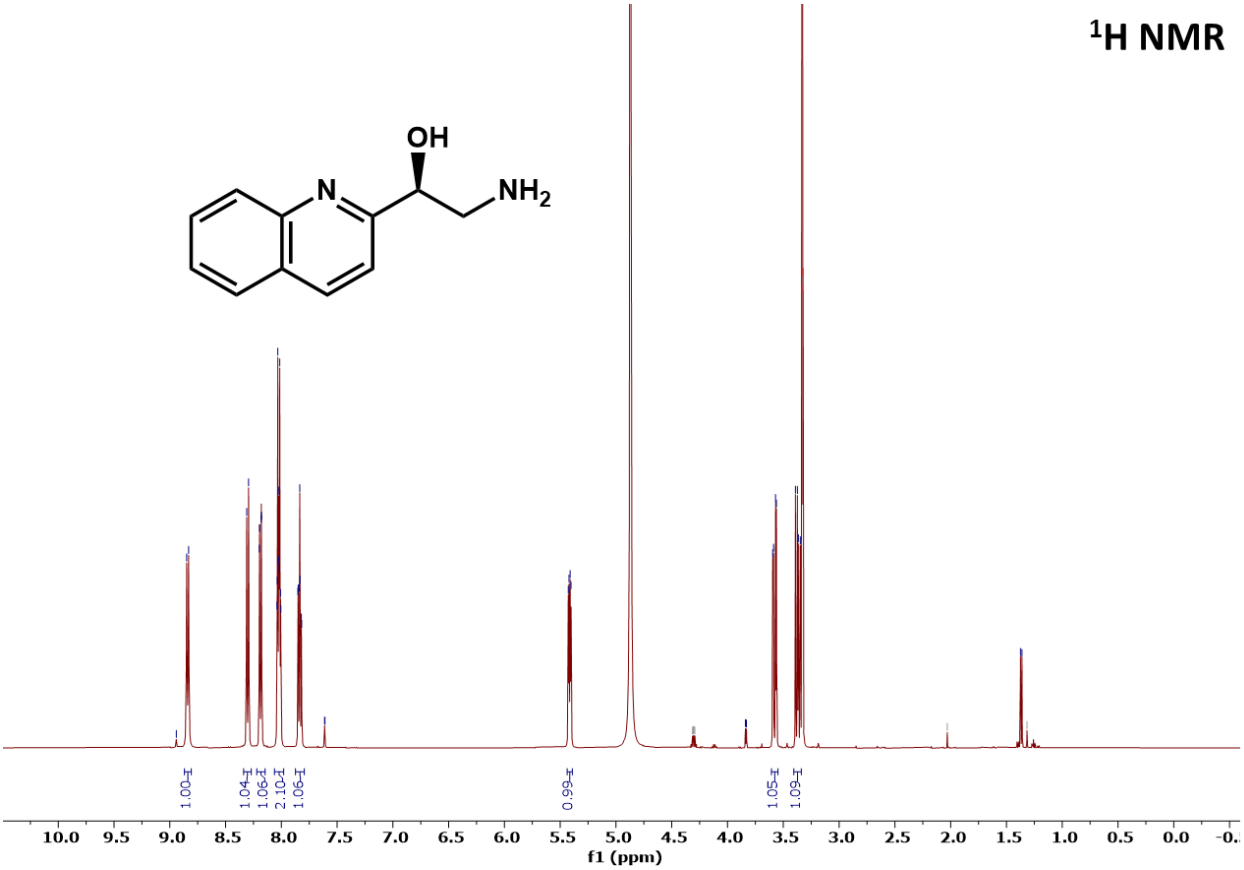


Figure 27. ^{1}H NMR spectrum of (*R*)-2-amino-1-(quinolin-3-yl)ethan-1-ol.

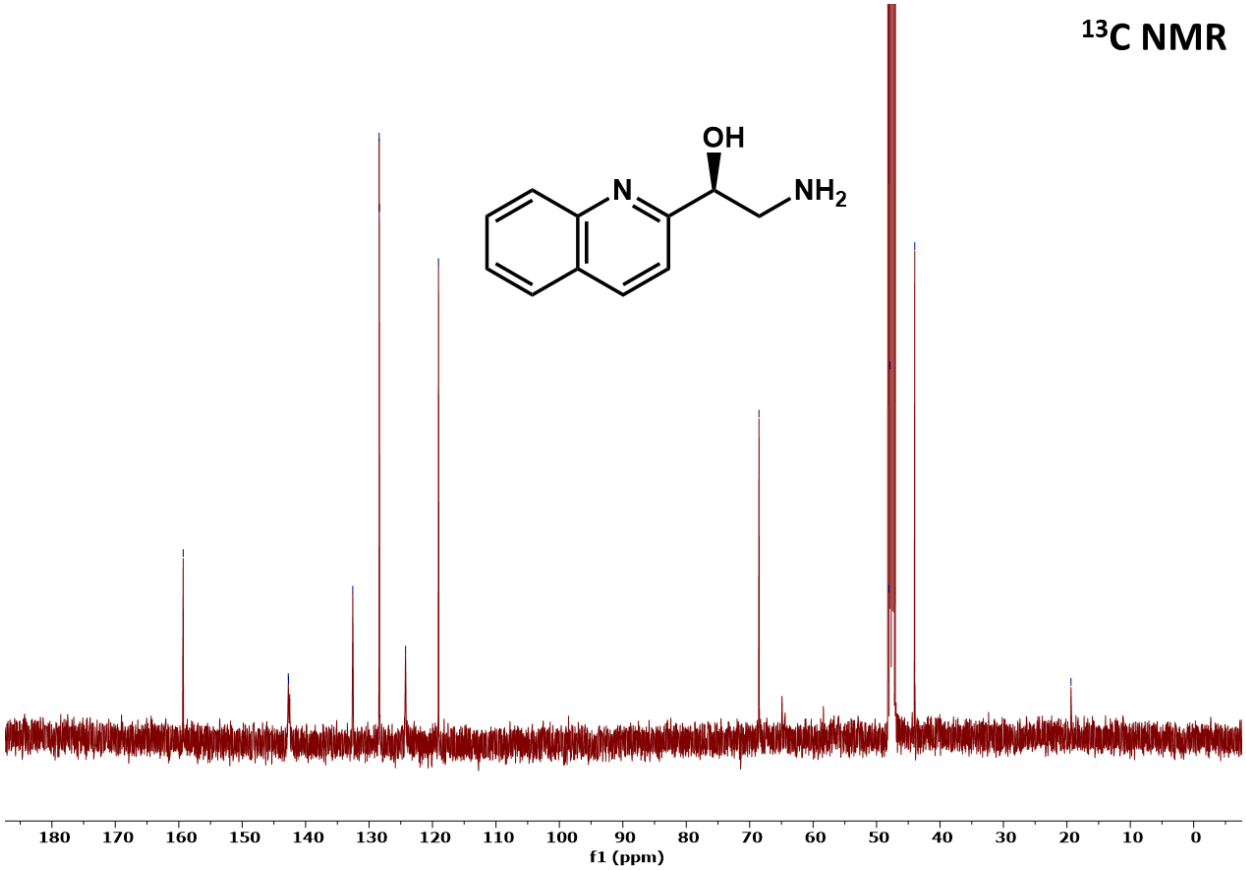


Figure 28. ^{13}C NMR spectrum of (*R*)-2-amino-1-(quinolin-3-yl)ethan-1-ol.

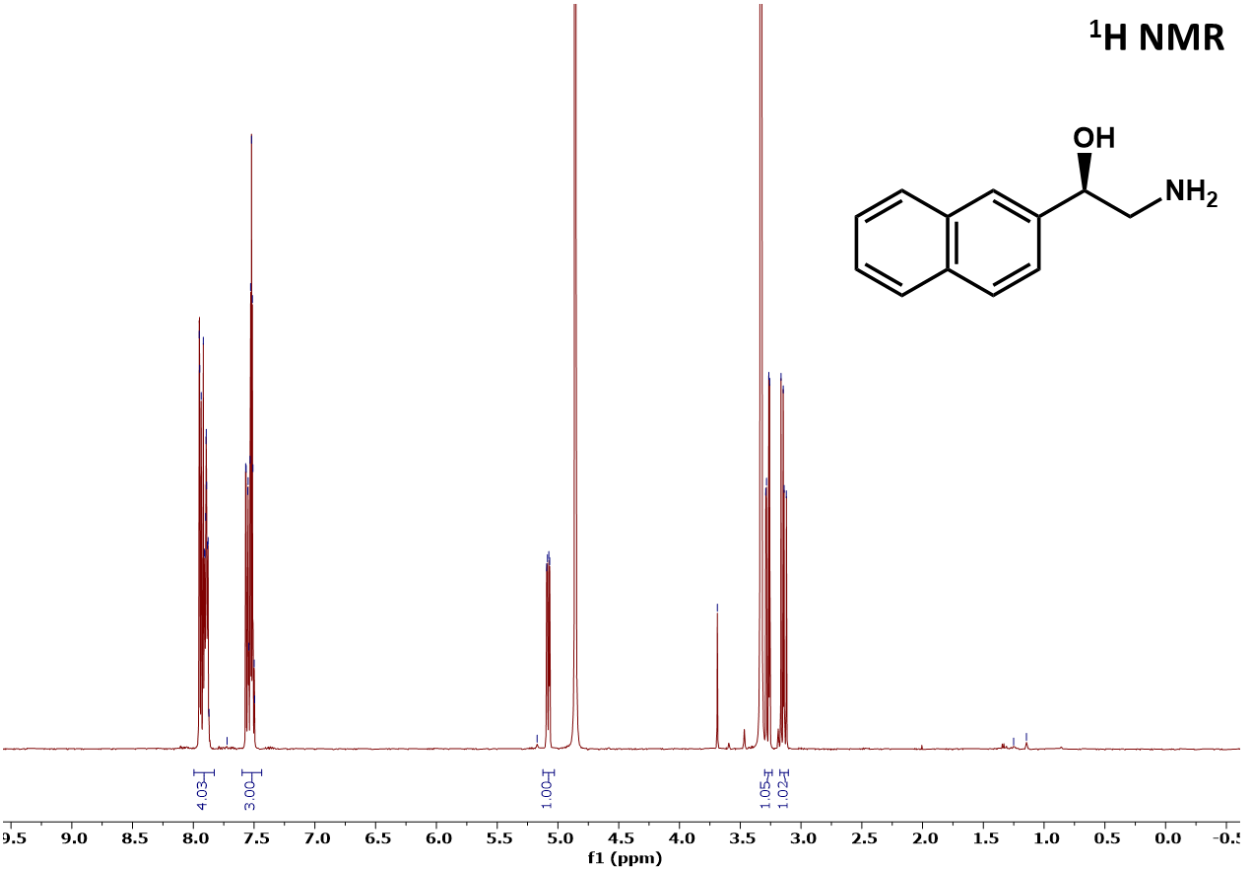


Figure 29. ^{1}H NMR spectrum of (*R*)-2-amino-1-(naphthalen-2-yl)ethan-1-ol.

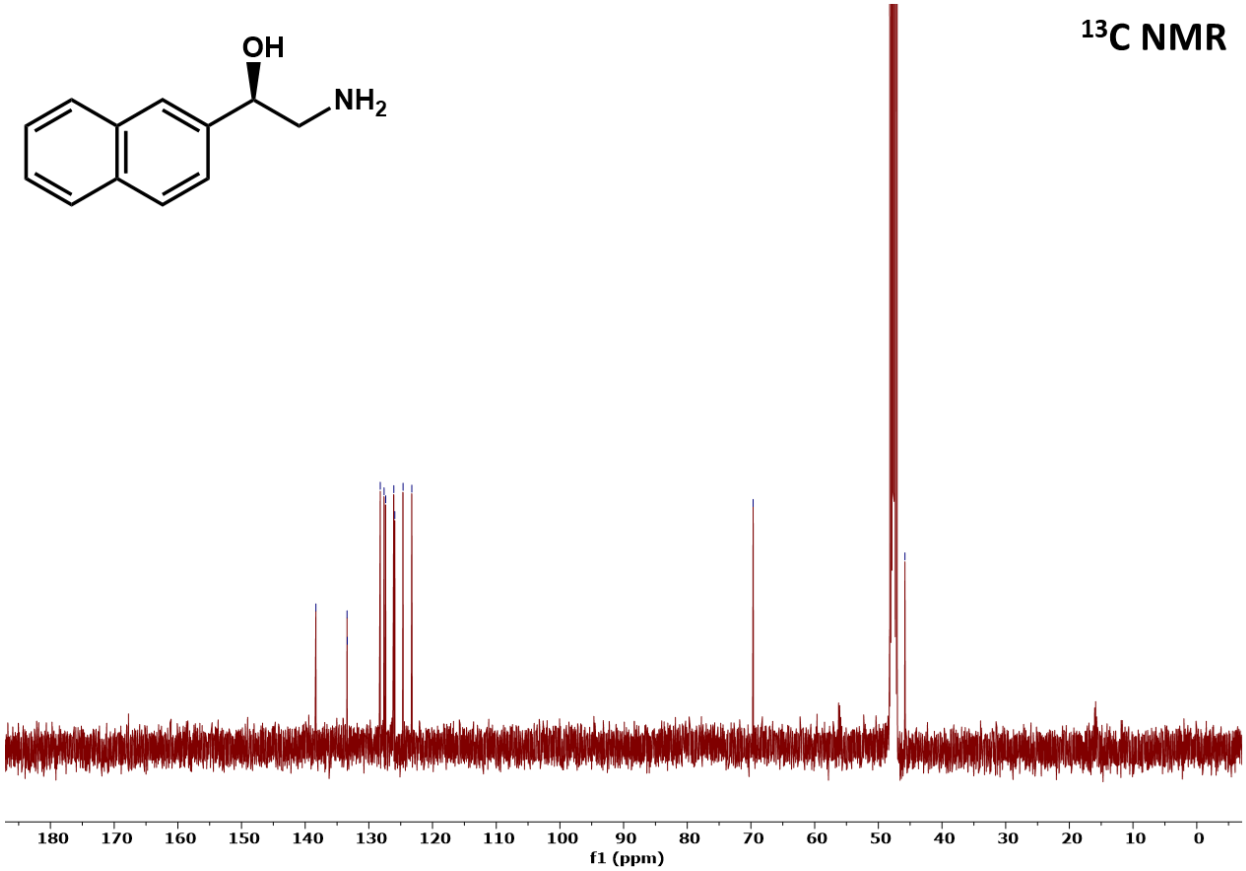


Figure 30. ^{13}C NMR spectrum of (*R*)-2-amino-1-(naphthalen-2-yl)ethan-1-ol.

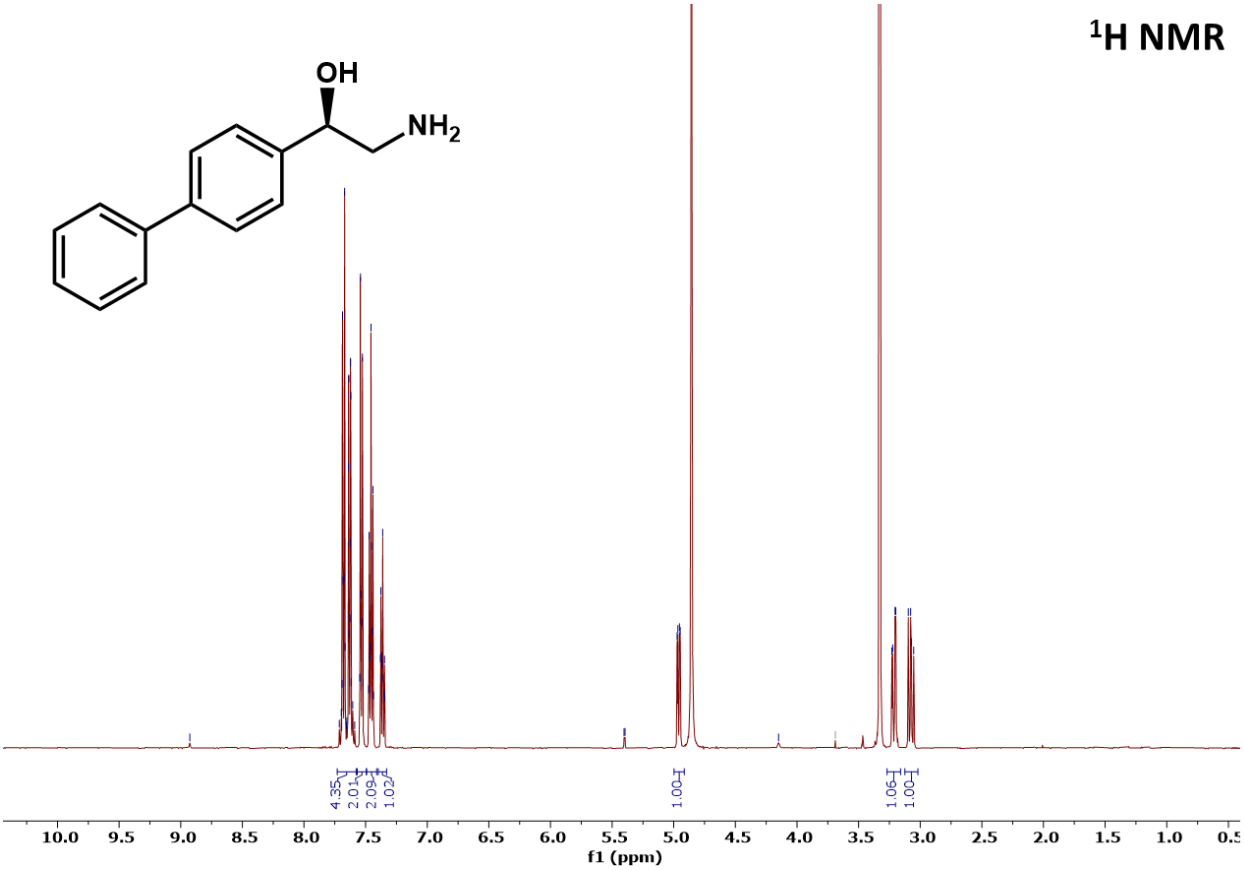


Figure 31. ^{1}H NMR spectrum of (*R*)-1-([1,1'-biphenyl]-4-yl)-2-aminoethan-1-ol.

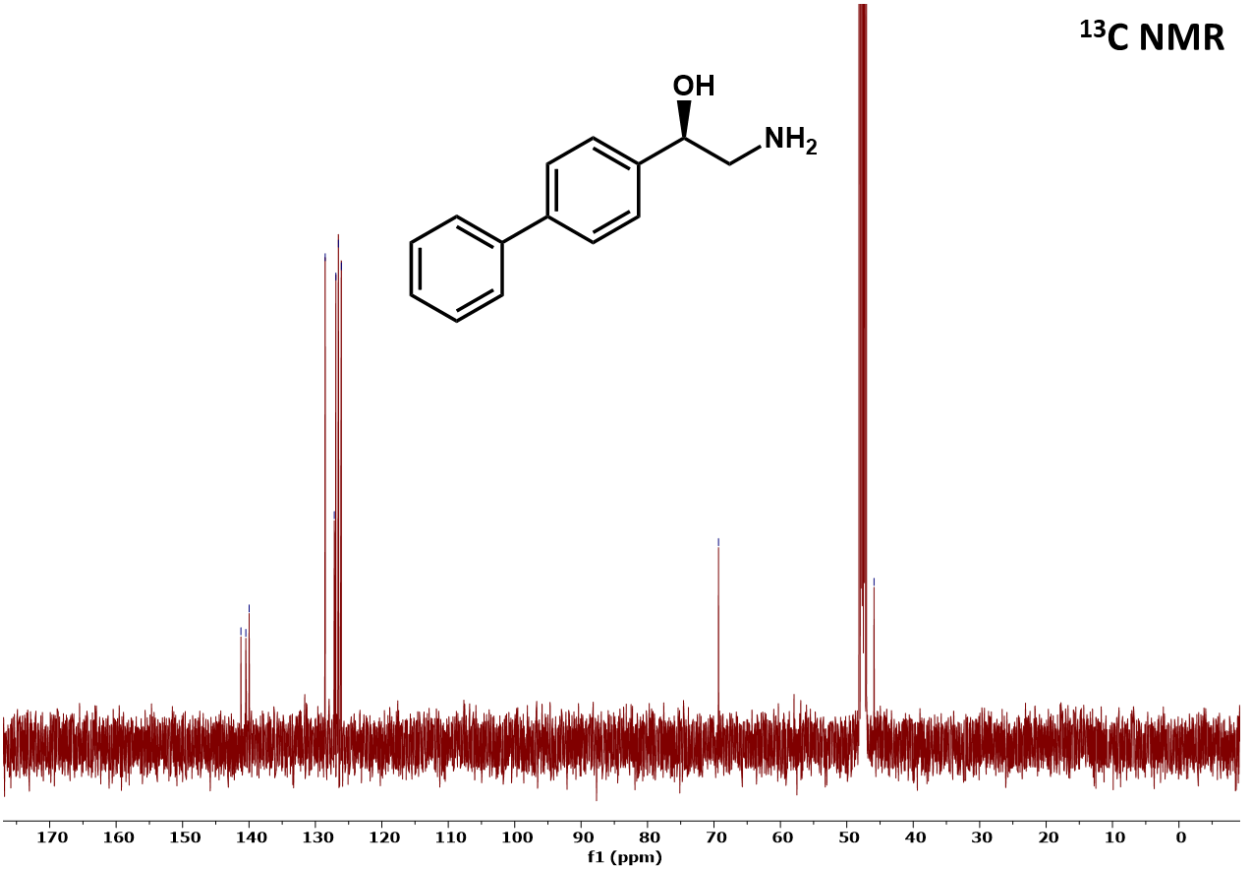


Figure 32. ^{13}C NMR spectrum of (*R*)-1-([1,1'-biphenyl]-4-yl)-2-aminoethan-1-ol.

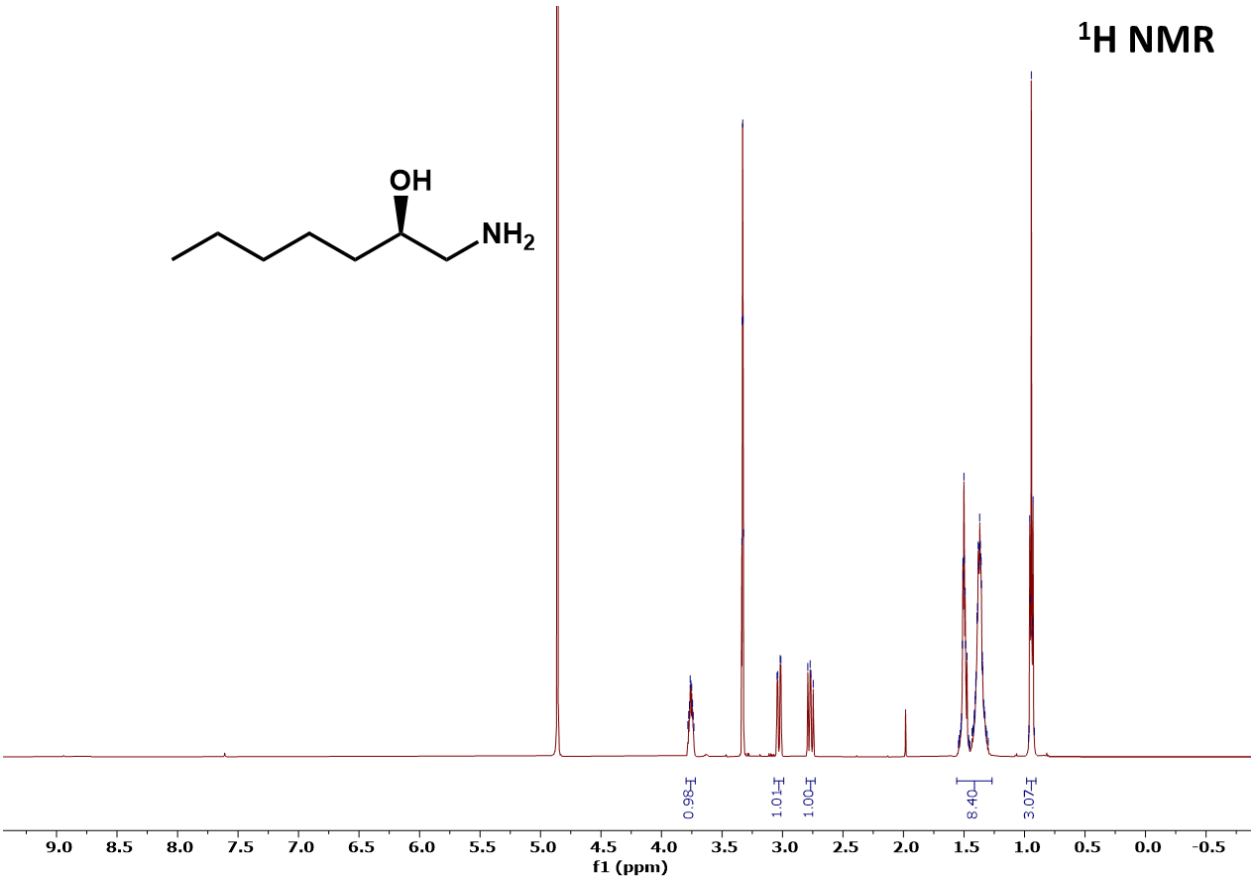


Figure 33. ^{1}H NMR spectrum of (*R*)-1-aminoheptan-2-ol.

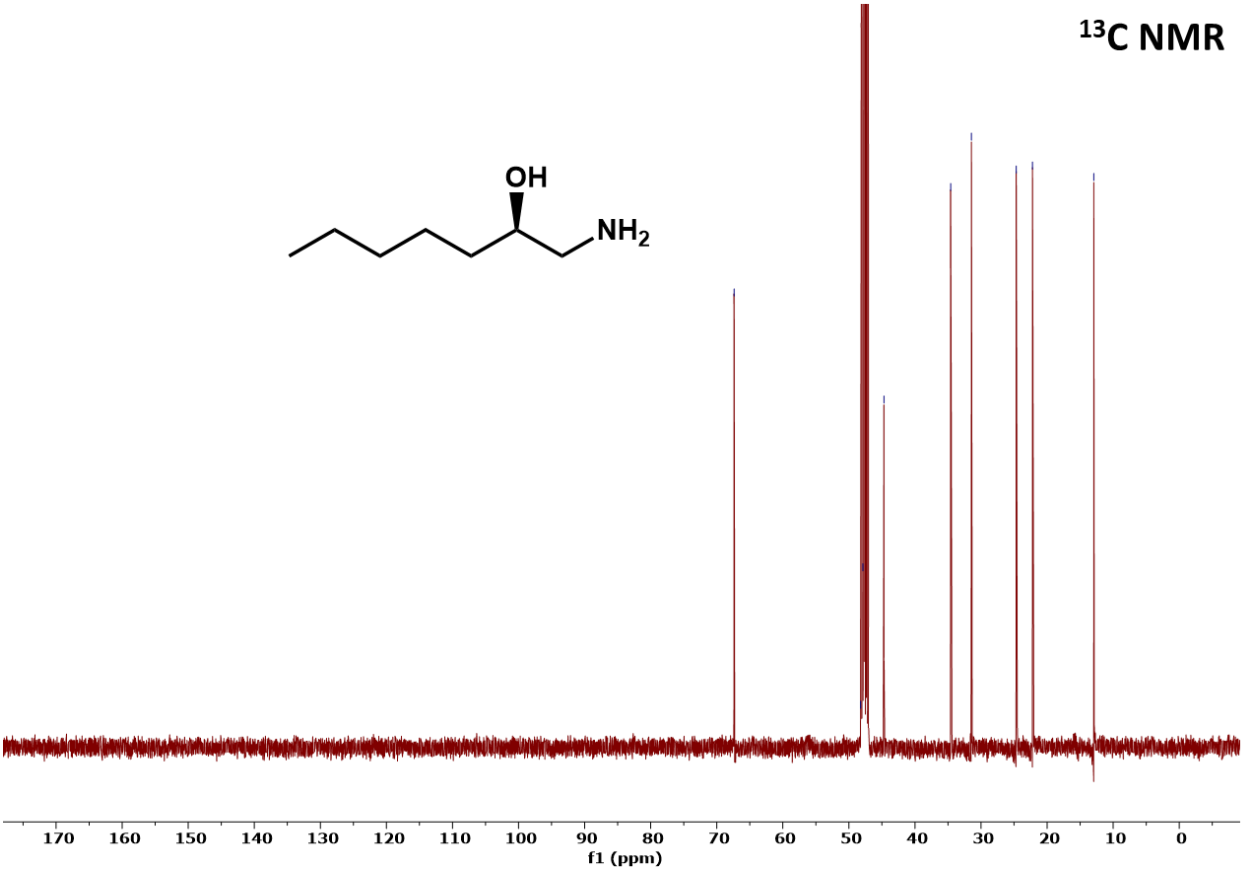


Figure 34. ^{13}C NMR spectrum of (*R*)-1-aminoheptan-2-ol.

Figure 35. ^{1}H NMR spectrum of (*R*)-1-amino-4-(methylthio)butan-2-ol.

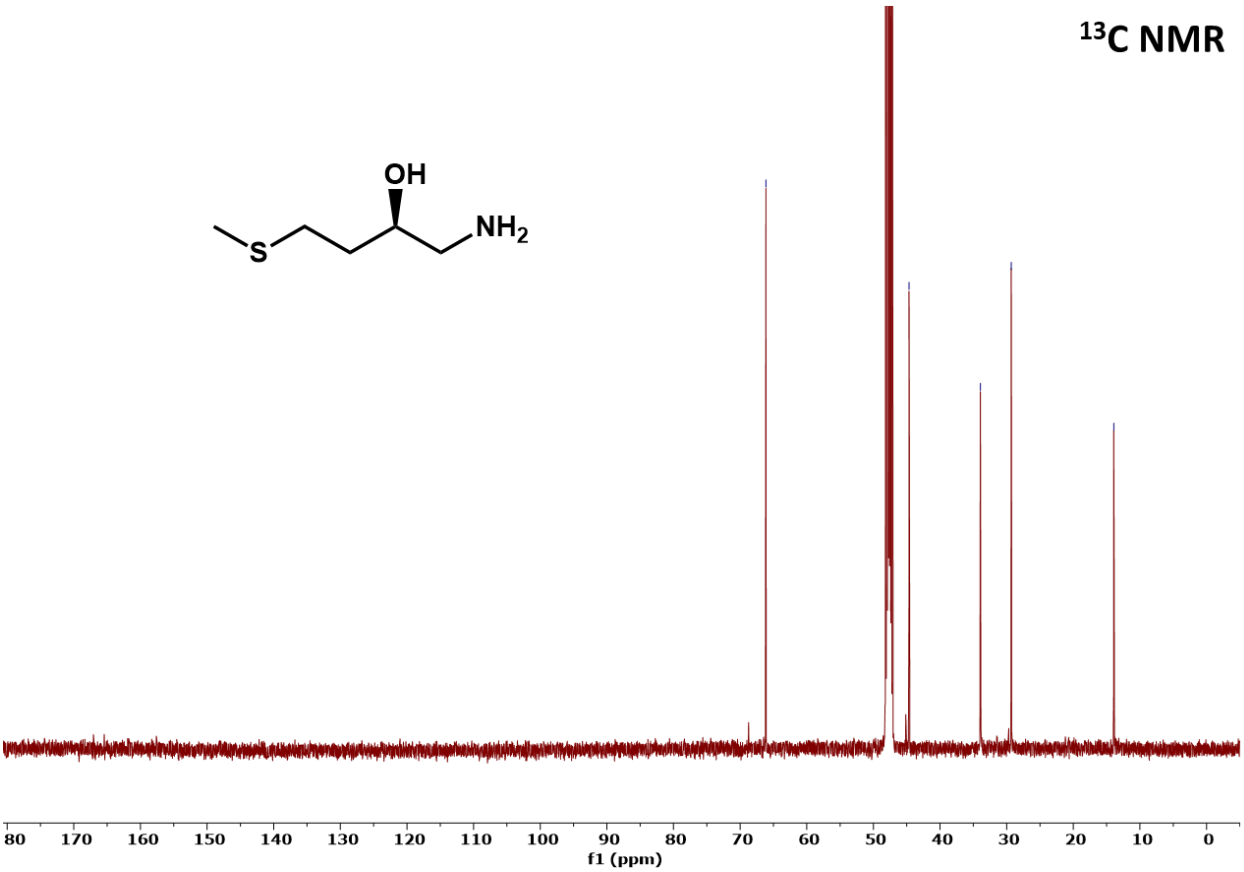


Figure 36. ^{13}C NMR spectrum of (*R*)-1-amino-4-(methylthio)butan-2-ol.

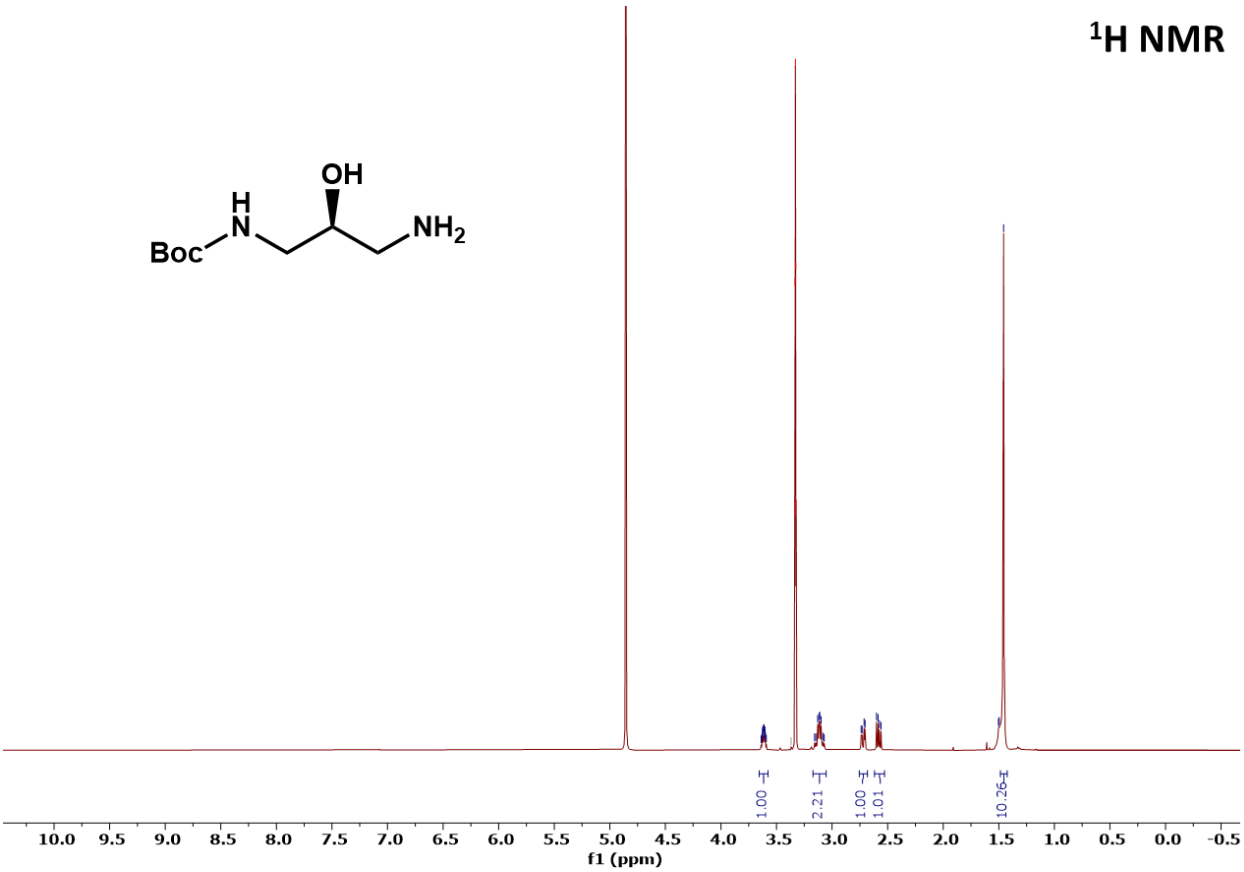


Figure 37. ^{1}H NMR spectrum of tert-butyl (*S*)-(3-amino-2-hydroxypropyl)carbamate.

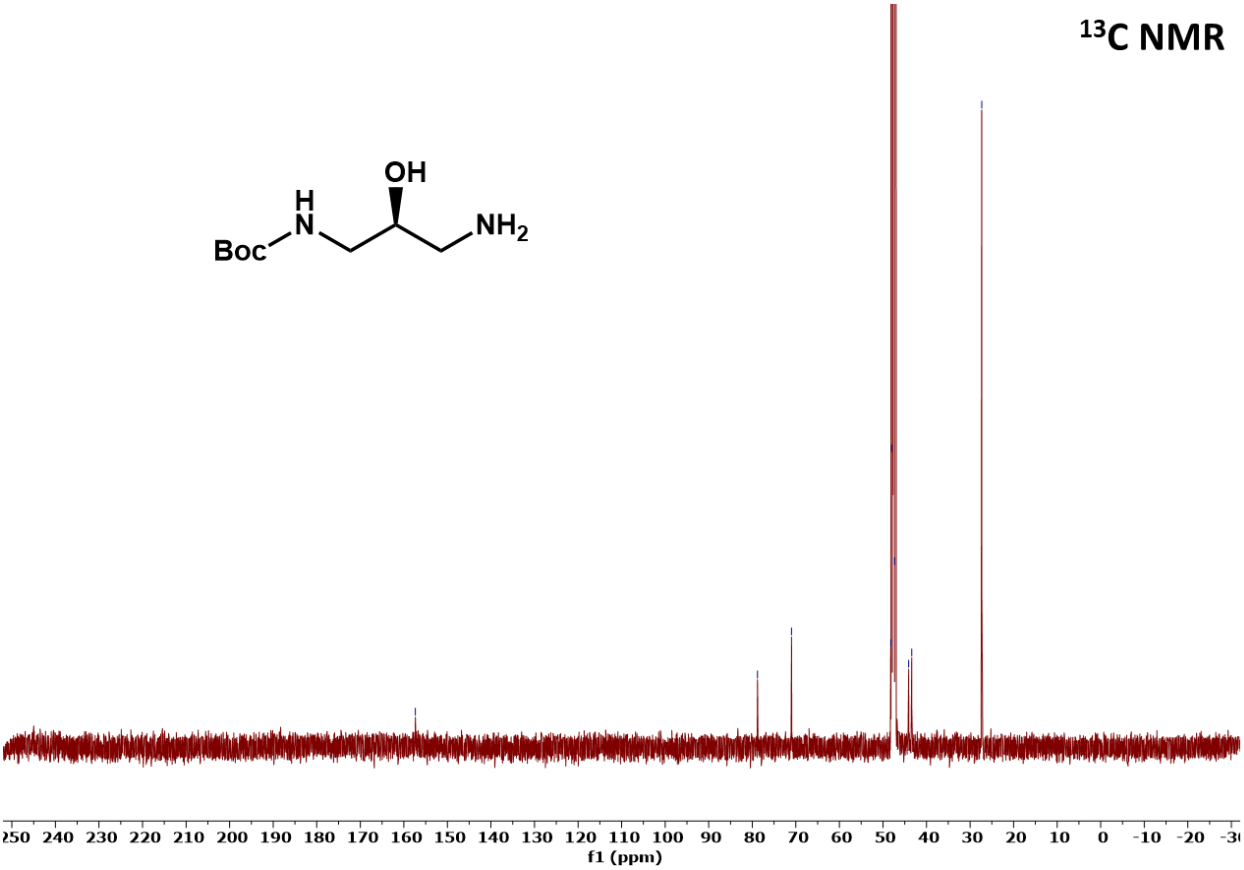


Figure 38. ^{13}C NMR spectrum of tert-butyl (*S*)-(3-amino-2-hydroxypropyl)carbamate.

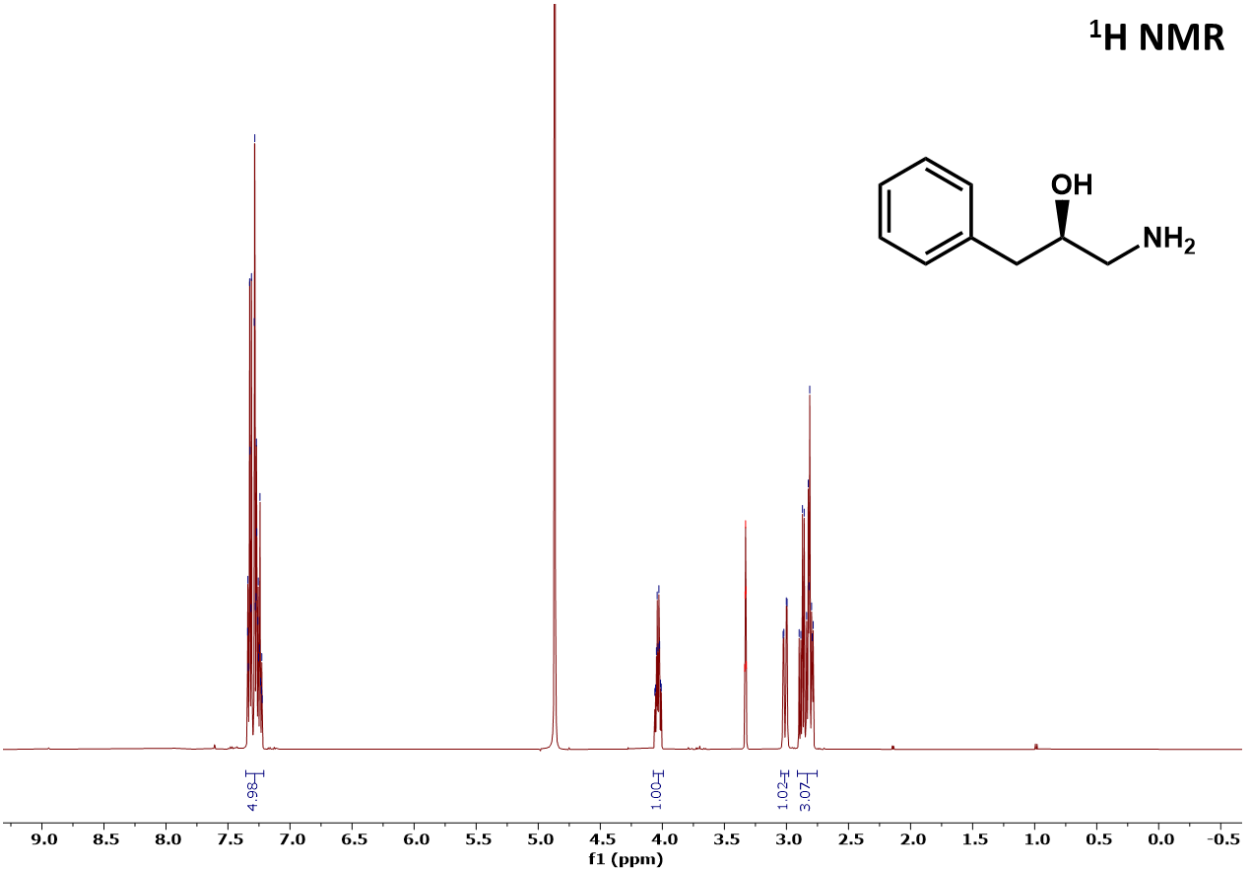


Figure 39. ^{1}H NMR spectrum of (*R*)-1-amino-3-phenylpropan-2-ol.

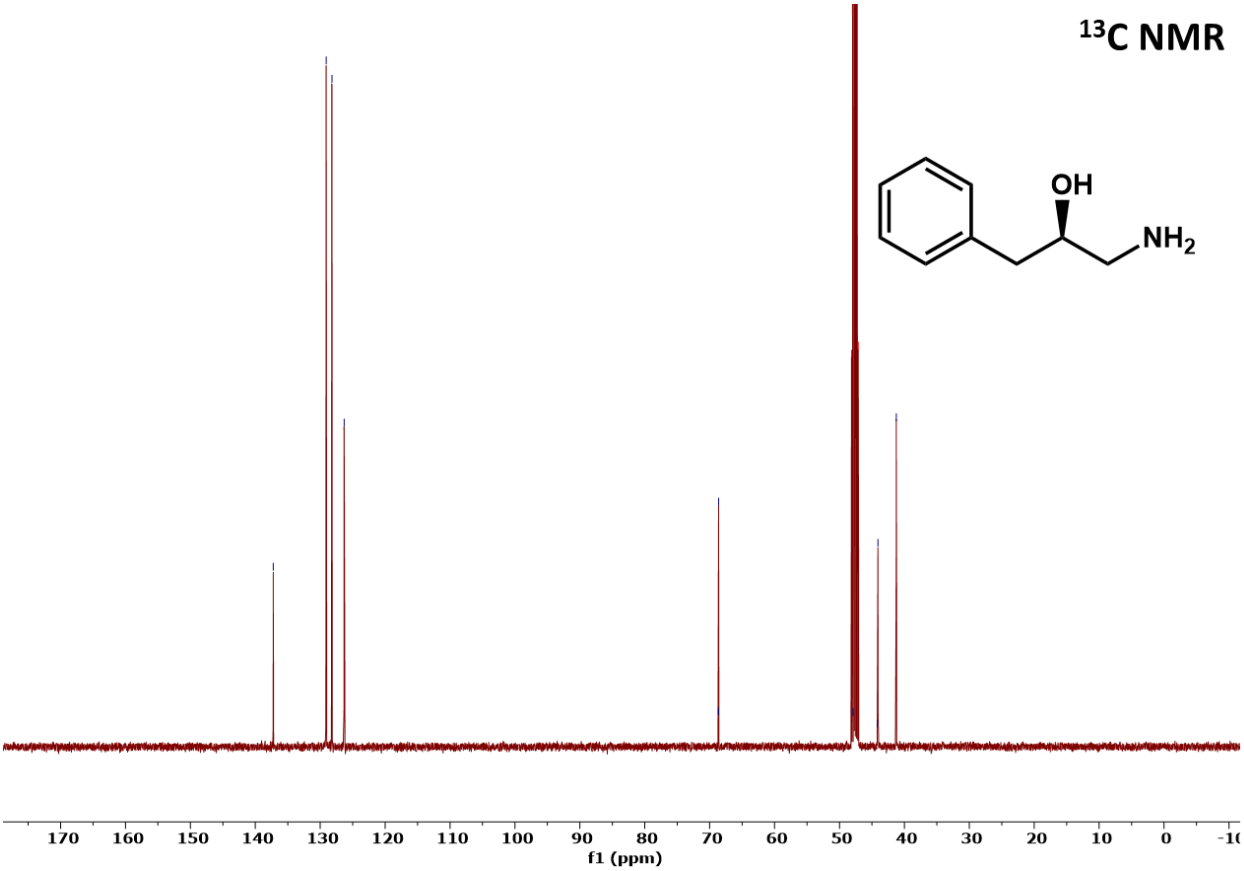


Figure 40. ^{13}C NMR spectrum of (*R*)-1-amino-3-phenylpropan-2-ol.

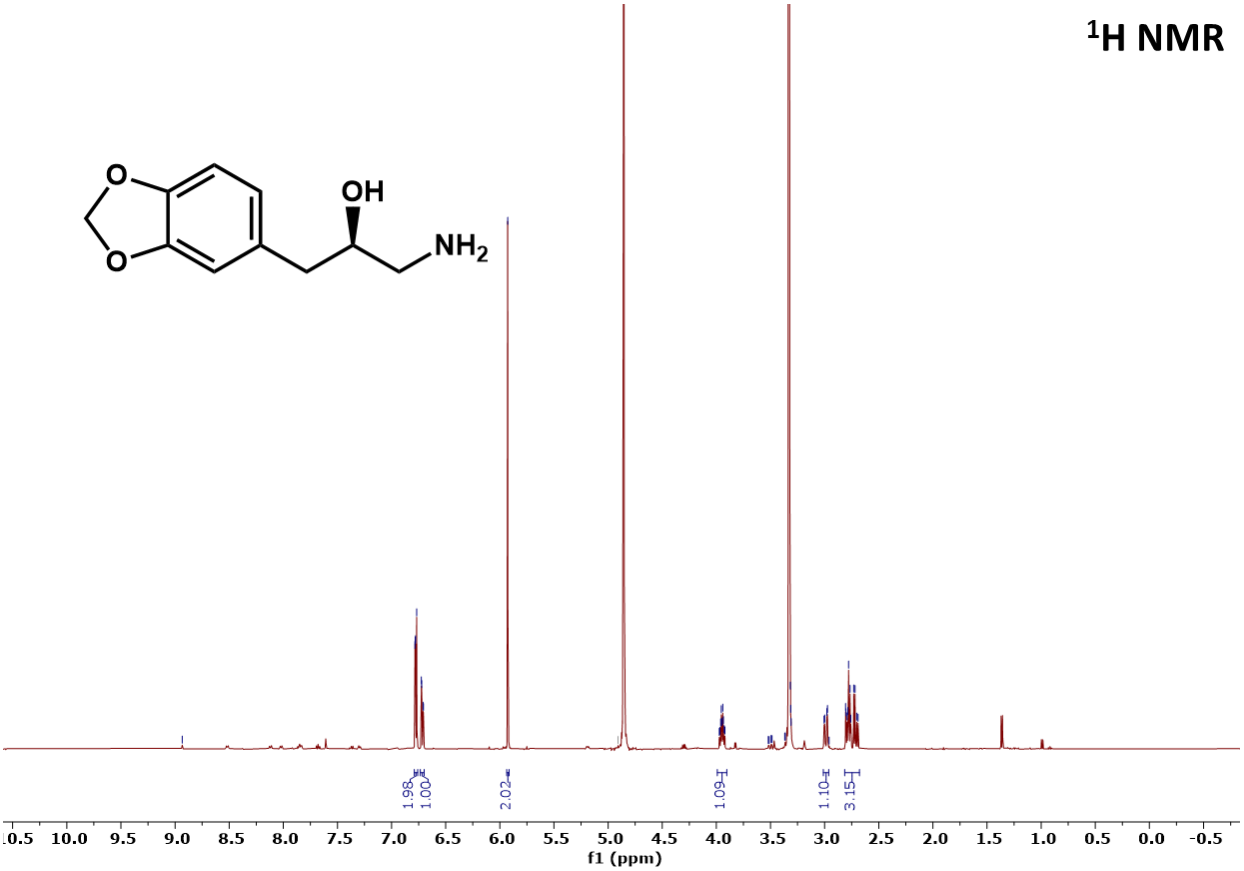


Figure 41. ^{1}H NMR spectrum of (*R*)-1-amino-3-(benzo[d][1,3]dioxol-5-yl)propan-2-ol.

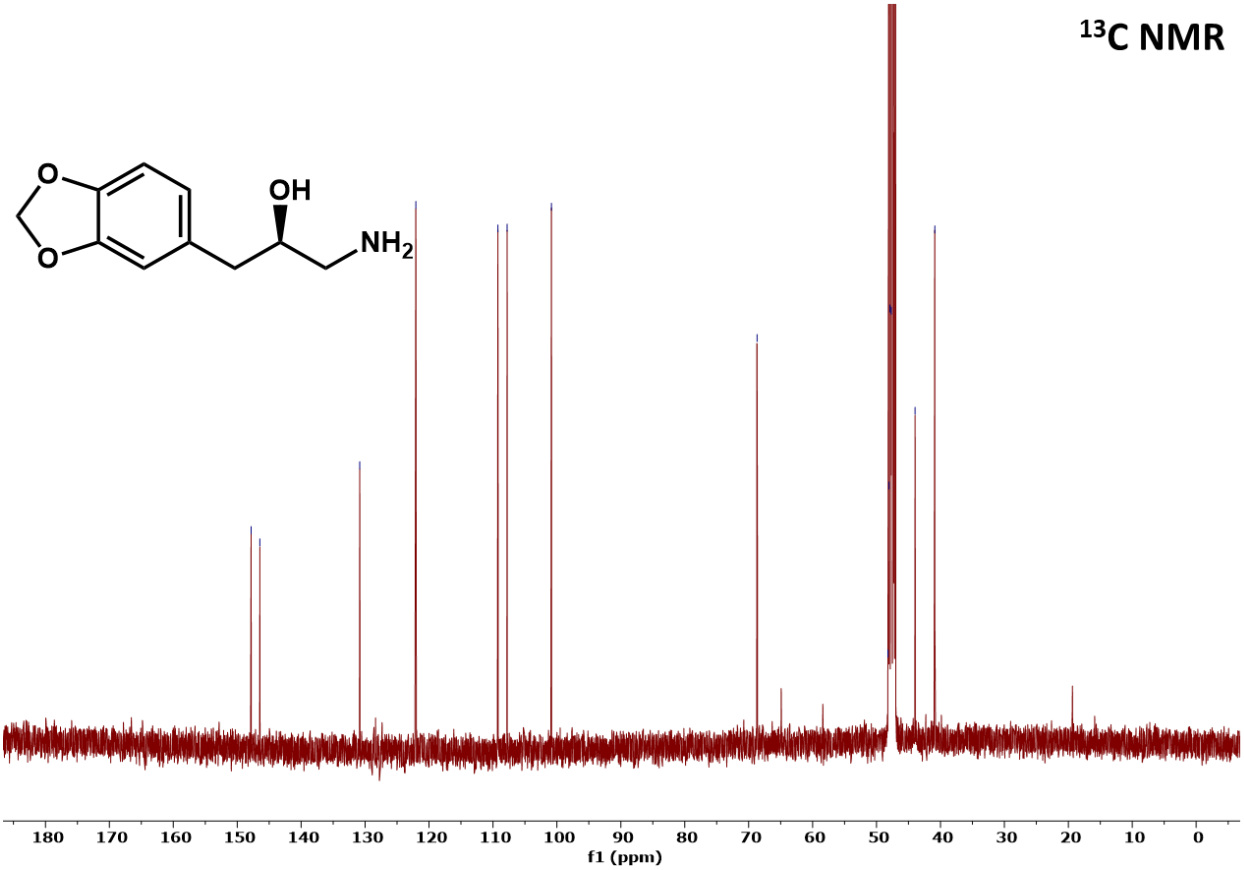


Figure 42. ^{13}C NMR spectrum of (*R*)-1-amino-3-(benzo[d][1,3]dioxol-5-yl)propan-2-ol.

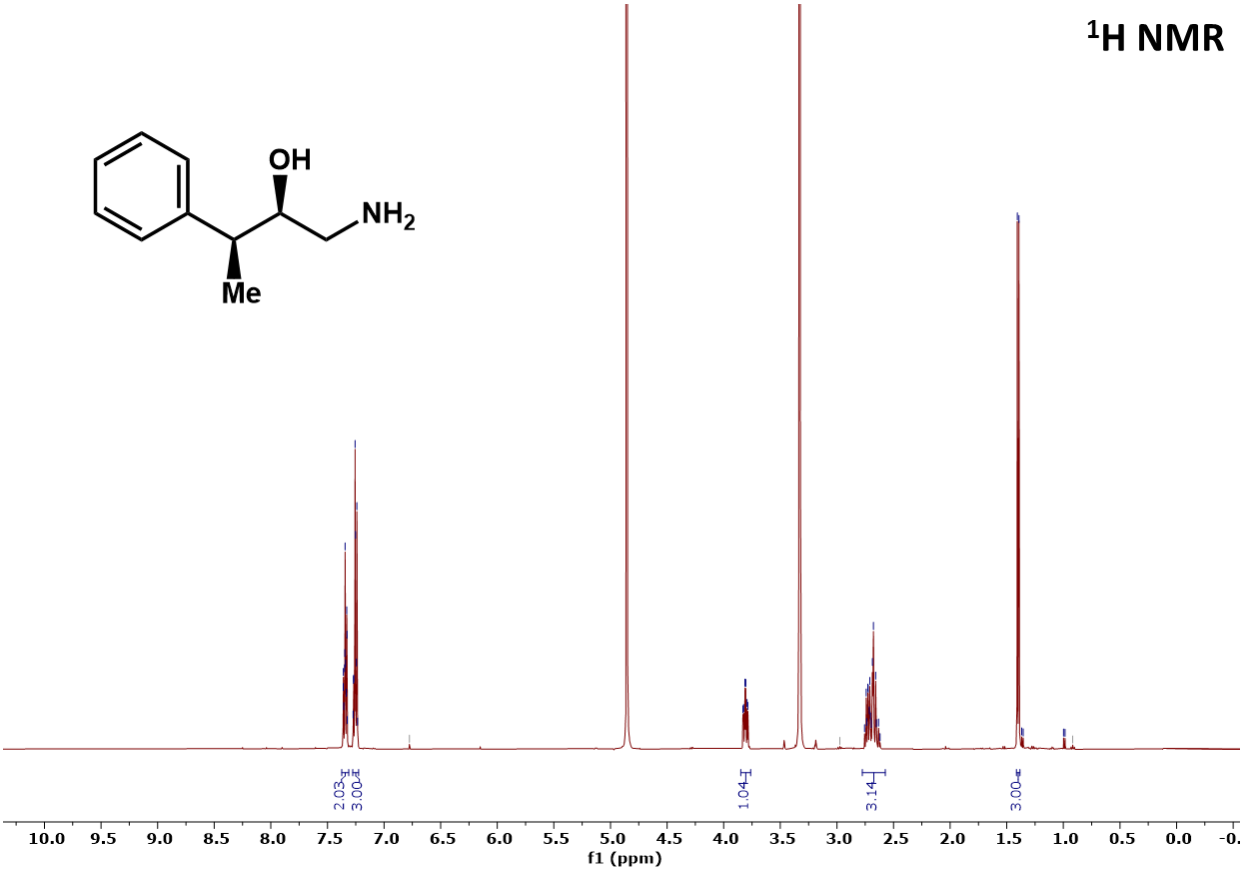


Figure 43. ^{1}H NMR spectrum of (2*R*)-1-amino-3-phenylbutan-2-ol.

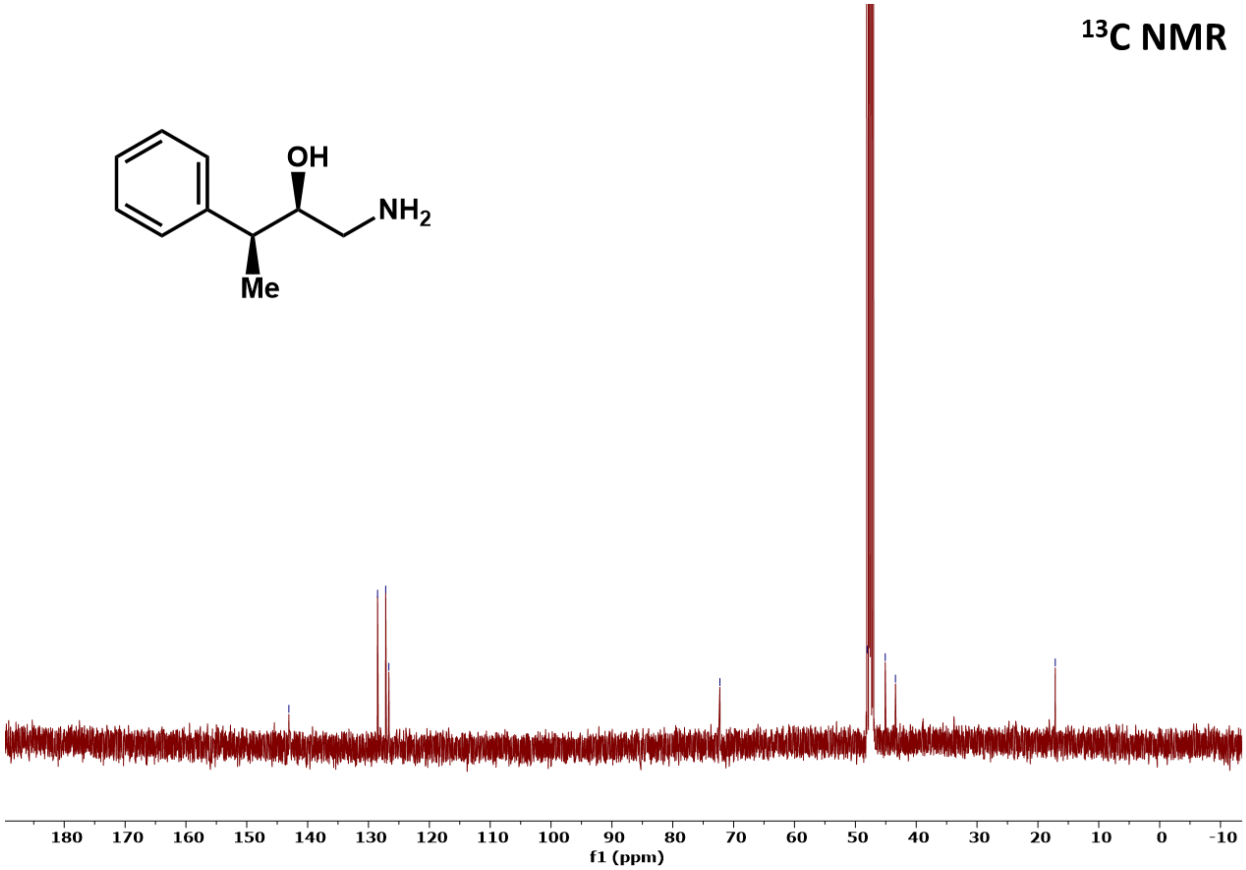


Figure 44. ^{13}C NMR spectrum of (2*R*)-1-amino-3-phenylbutan-2-ol.

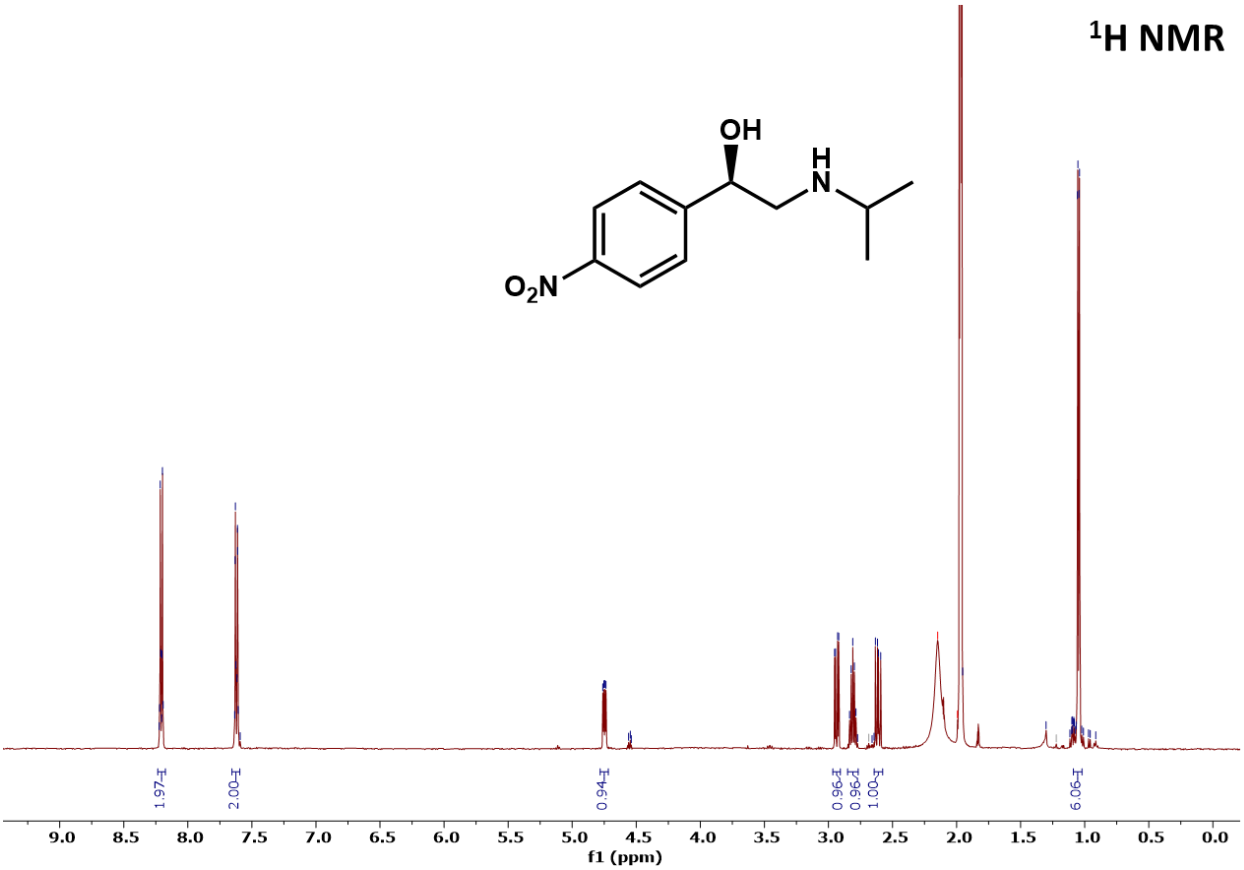


Figure 45. ^{1}H NMR spectrum of (*R*)-2-(isopropylamino)-1-(4-nitrophenyl)ethan-1-ol.

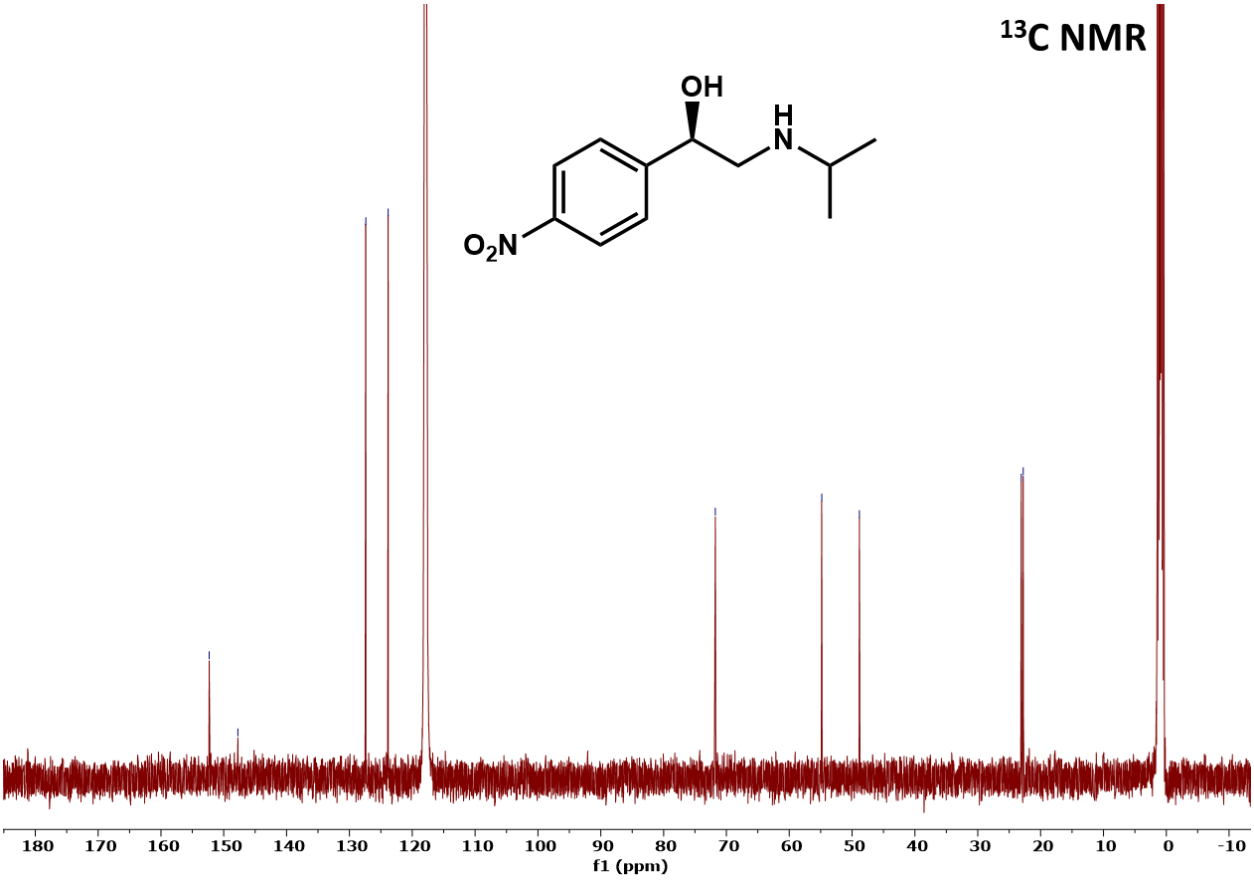


Figure 46. ^{13}C NMR spectrum of (*R*)-2-(isopropylamino)-1-(4-nitrophenyl)ethan-1-ol.

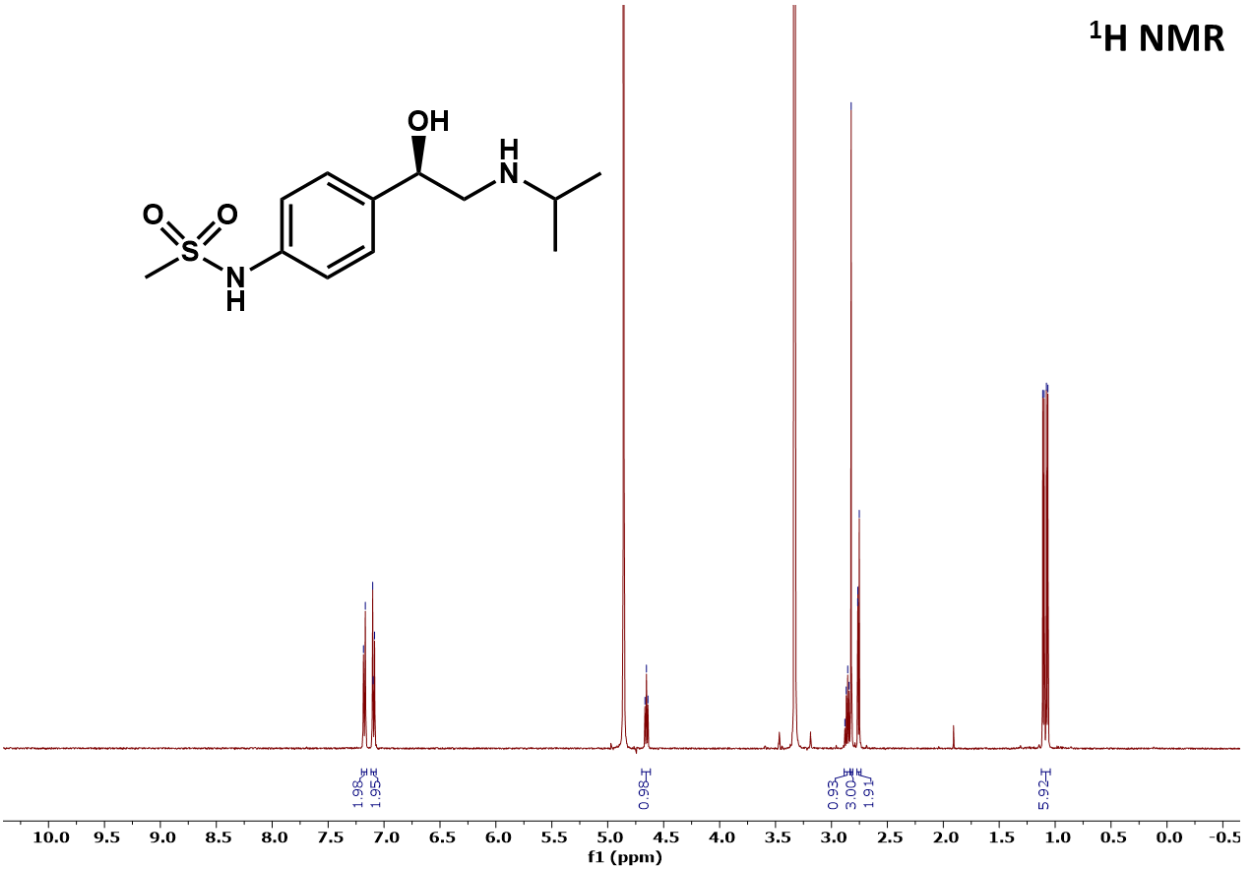


Figure 47. ^{1}H NMR spectrum of (*R*)-N-(4-(1-hydroxy-2-(isopropylamino)ethyl)phenyl)methanesulfonamide.

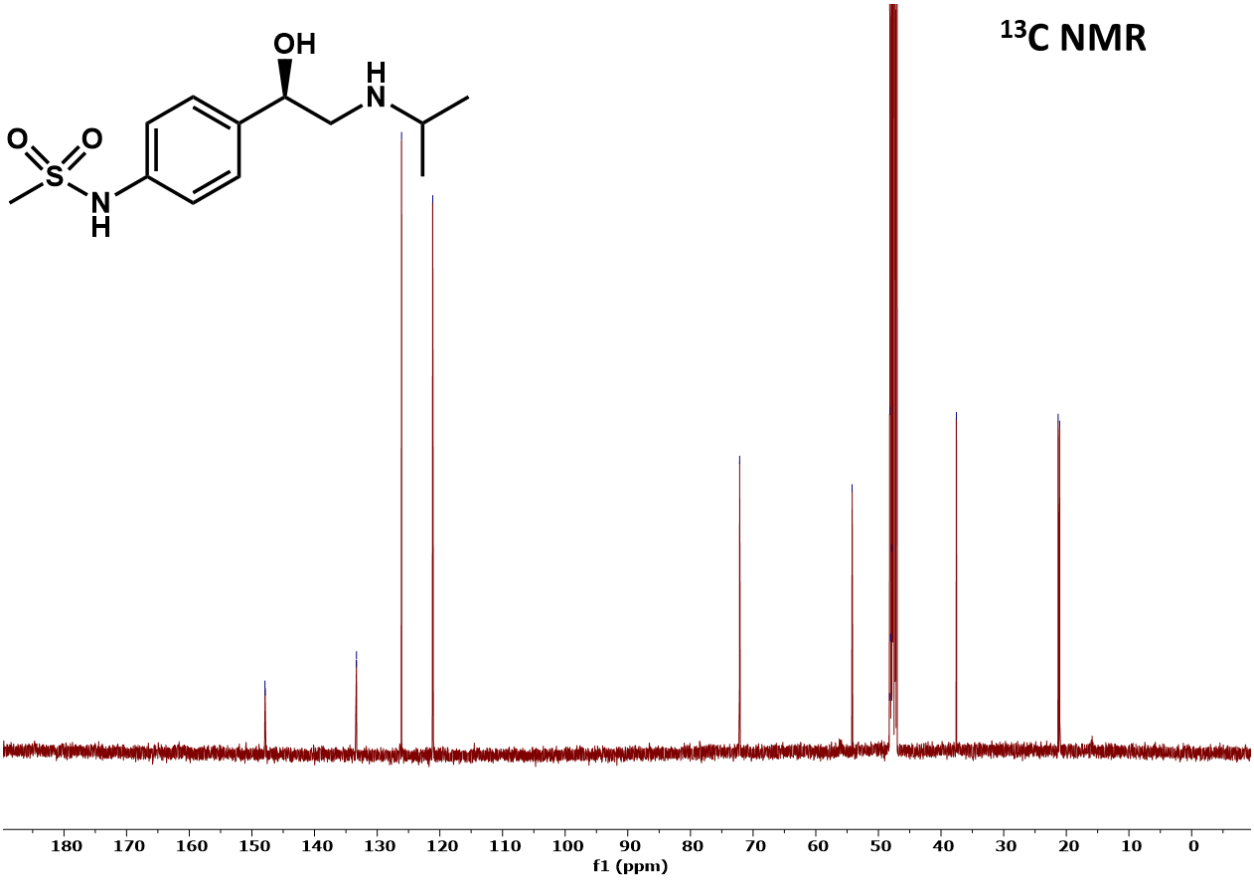


Figure 48. ^{13}C NMR spectrum of (*R*)-N-(4-(1-hydroxy-2-(isopropylamino)ethyl)phenyl)methanesulfonamide.

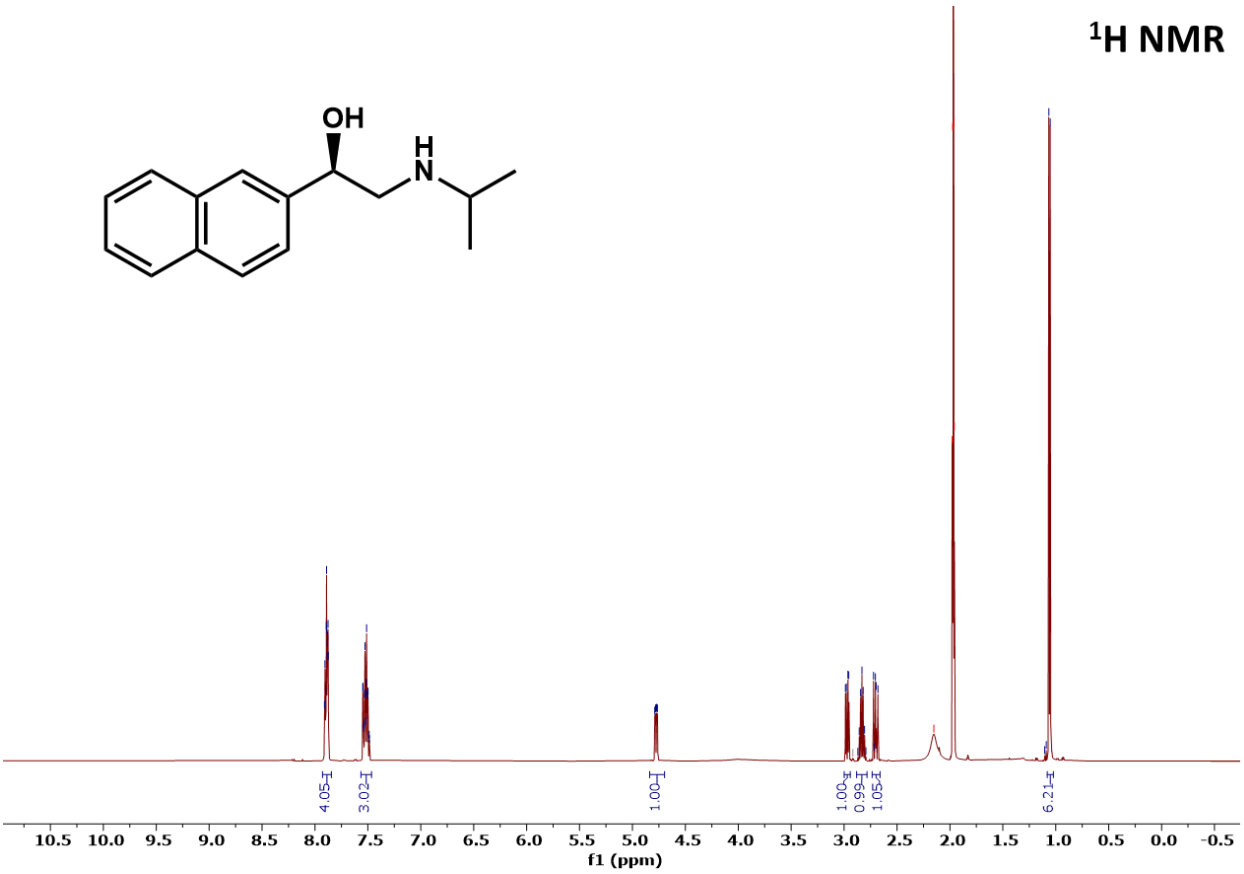


Figure 49. ^{1}H NMR spectrum of (*R*)-2-(isopropylamino)-1-(naphthalen-2-yl)ethan-1-ol.

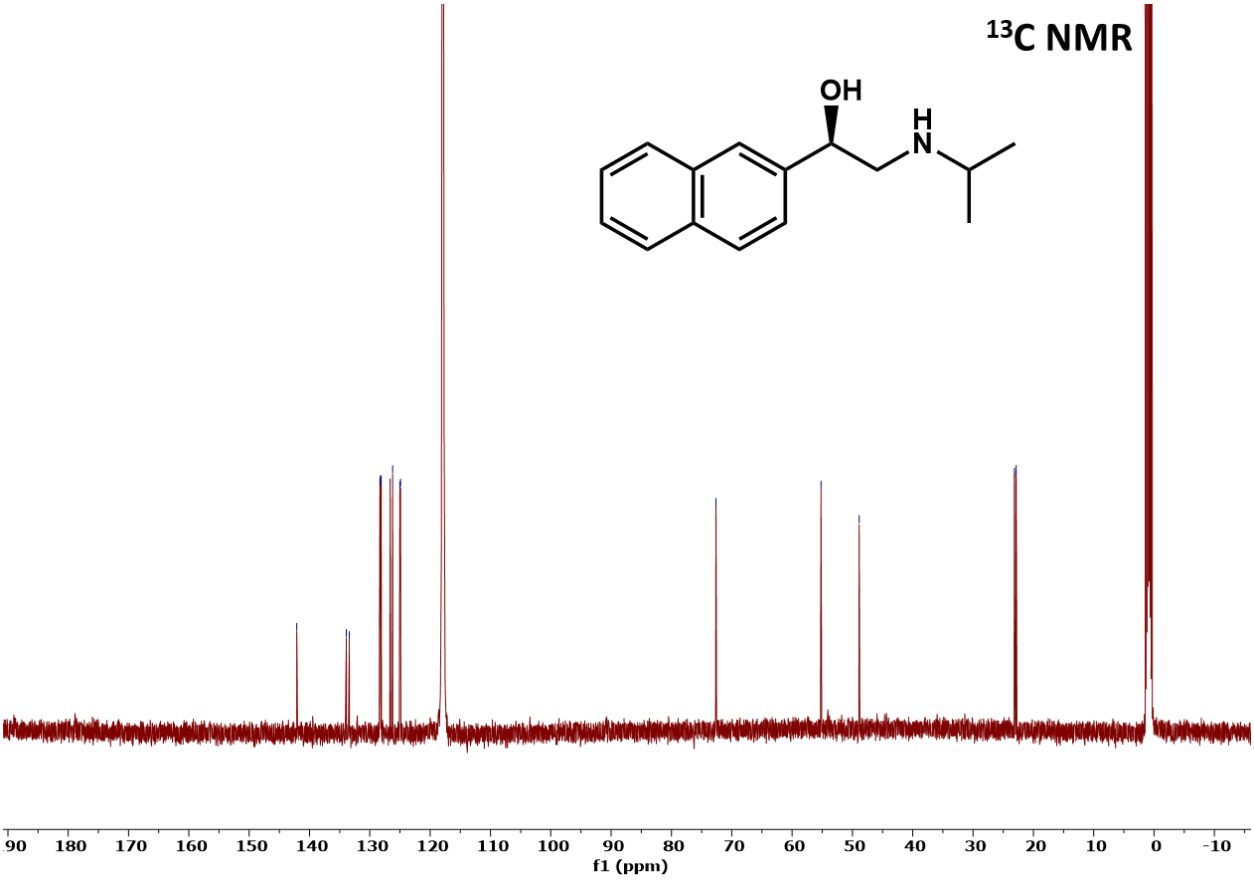


Figure 50. ^{13}C NMR spectrum of (*R*)-2-(isopropylamino)-1-(naphthalen-2-yl)ethan-1-ol.

4. 5. References

1. Schrittwieser, J. H., Velikogne, S., Hall, M. & Kroutil, W. Artificial Biocatalytic Linear Cascades for Preparation of Organic Molecules. *Chemical Reviews* vol. 118 270–348 (2018).
2. Muschiol, J., Peters, C., Oberleitner, N. & Mihovilovic, M. D. Cascade catalysis – strategies and challenges en route to preparative synthetic biology. *Chem. Commun.* **51**, 5798–5811 (2015).
3. McIntosh, J. A. *et al.* Engineered Ribosyl-1-Kinase Enables Concise Synthesis of Molnupiravir, an Antiviral for COVID-19. *ACS Cent. Sci.* **7**, 1980–1985 (2021).
4. Savile, C. K. *et al.* Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science (80-.).* **329**, 305–310 (2010).
5. Huffman, M. A. *et al.* Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science (80-.).* **366**, 1255–1259 (2019).
6. Wang, H. *et al.* Data mining of amine dehydrogenases for the synthesis of enantiopure amino alcohols. *Catal. Sci. Technol.* **10**, 5945–5952 (2020).
7. Zhang, J. *et al.* Enantioselective Cascade Biocatalysis for Deracemization of Racemic β - Amino Alcohols to Enantiopure (S) - β -Amino Alcohols by Employing Cyclohexylamine Oxidase and ω -Transaminase. *ChemBioChem* **22**, 124–128 (2021).
8. Liu, S. *et al.* Designing of a Cofactor Self-Sufficient Whole-Cell Biocatalyst System for Production of 1 ,2-Amino Alcohols from Epoxides. *ACS Syn* **8**, 734–743 (2019).
9. Corrado, M. L., Knaus, T., Schwaneberg, U. & Mutti, F. G. High-Yield Synthesis of Enantiopure 1,2-Amino Alcohols from l -Phenylalanine via Linear and Divergent Enzymatic Cascades . *Org. Process Res. Dev.* (2022) doi:10.1021/acs.oprd.1c00490.
10. Schrittwieser, J. H. & Hollmann, F. One-pot combination of enzyme and Pd nanoparticle catalysis for the synthesis of enantiomerically pure 1,2-amino alcohols. *Green Chem.* **15**, 3318–3331 (2013).
11. Busto, E., Simon, R. C. & Kroutil, W. Vinylation of Unprotected Phenols Using a Biocatalytic System. *Angew. Chemie* **54**, 10899–10902 (2015).
12. Fischereder, E.-M., Pressnitz, D. & Kroutil, W. Stereoselective Cascade to C3-Methylated Strictosidine Derivatives Employing Transaminases and Strictosidine Synthases. *ACS Catal.* **6**, 23–30 (2016).
13. Sehl, T. *et al.* Two Steps in One Pot : Enzyme Cascade for the Synthesis of Nor(pseudo)ephedrine from Inexpensive Starting Materials. *Angew. Chemie* **52**, 6772–6775 (2013).
14. Liu, S. *et al.* 1 ,2-Amino oxygenation of alkenes with hydrogen evolution reaction. *Nat. Commun.* **13**, (2022).

15. Hana, S. & Lange, A. B. Cloning and Functional Characterization of Octβ2-Receptor and Tyr1-Receptor in the Chagas Disease Vector, Rhodnius prolixus. *Front. Physiol.* **8**, 1–15 (2017).

16. Cizmarikova, R., Habala, L., Valentova, J. & Markuliak, M. Survey of Pharmacological Activity and Pharmacokinetics of Selected β -Adrenergic Blockers in Regard to Their Stereochemistry. *Appl. Sci.* **9**, 625 (2019).

17. Pereira, R. B. *et al.* Amino Alcohols from Eugenol as Potential Semisynthetic Insecticides : Chemical , Biological , and Computational Insights. *Molecules* **26**, 1–20 (2021).

18. Baker, J. G., Hill, S. J. & Summers, R. J. Evolution of b blockers: from anti-anginal drugs to ligand-directed signalling. *Trends Pharmacol. Sci.* **32**, 227–234 (2011).

19. Erdmann, V. *et al.* Cascade Reactions Enzymatic and Chemoenzymatic Three-Step Cascades for the Synthesis of Stereochemically Complementary Trisubstituted Tetrahydroisoquinolines. *Angew. Chemie* **56**, 12503–12507 (2017).

20. Li, Y. *et al.* Asymmetric Synthesis of N-Substituted 1,2-Amino Alcohols from Simple Aldehydes and Amines by One-Pot Sequential Enzymatic Hydroxymethylation and Asymmetric Reductive Amination. *Angew. Chemie* (2022) doi:10.1002/anie.202116344.

21. Xu, F., Wang, J., Liu, B., Wu, Q. & Lin, X. Enzymatic synthesis of optical pure B-nitroalcohols by combining D-aminoacylase-catalyzed nitroaldol reaction and immobilized lipase PS-catalyzed kinetic resolution. *Green Chem.* **13**, 2359–2361 (2011).

22. Steinreiber, J. *et al.* Overcoming Thermodynamic and Kinetic Limitations of Aldolase-Catalyzed Reactions by Applying Multienzymatic Dynamic Kinetic Asymmetric Transformations. *Angew. Chem. Int. Ed.* 1624–1626 (2007) doi:10.1002/anie.200604142.

23. Steinreiber, J. *et al.* Synthesis of Aromatic 1 ,2-Amino Alcohols Utilizing a Bienzymatic Dynamic Kinetic Asymmetric Transformation. *Adv. Synth. Catal.* **349**, 1379–1386 (2007).

24. Scott, T. A., Heine, D., Qin, Z. & Wilkinson, B. An L -threonine transaldolase is required for L-threo-b-hydroxy-a-amino acid assembly during obafluorin biosynthesis. *Nat. Commun.* **8**, 1–11 (2017).

25. Kumar, P. *et al.* Transaldolase Activity Is Enabled by a Persistent Catalytic Intermediate. *ACS ch* **16**, 86–95 (2021).

26. Doyon, T. J. *et al.* Scalable and Selective β -Hydroxy- α -Amino Acid Synthesis Catalyzed by Promiscuous l -Threonine Transaldolase ObiH. *ChemBioChem* **22**, 1–10 (2021).

27. Wu, S., Liu, J. & Li, Z. Biocatalytic Formal Anti-Markovnikov Hydroamination and Hydration of Aryl Alkenes. *ACS Catal.* **7**, 5225–5233 (2017).

28. Mcconathy, J. & Owens, M. J. Stereochemistry in Drug Action. *Prim. Care Companion J Clin Psychiatry* **5**, 70–73 (2003).

29. Yadav, J. S., Reddy, A. R., Narsaiah, A. V. & Reddy, B. V. S. An efficient protocol for regioselective ring opening of epoxides using samarium triflate: Synthesis of propranolol, atenolol and RO363. *J. Mol. Catal. A Chem.* **261**, 207–212 (2007).

30. Kamble, V. T. & Joshi, N. S. Synthesis of β-amino alcohols by ring opening of epoxides with amines catalyzed by cyanuric chloride under mild and solvent-free conditions. *Green Chem. Lett. Rev.* **3**, 275–281 (2010).

31. Saddique, F. A. *et al.* Synthetic approaches towards the synthesis of beta-blockers (betaxolol, metoprolol, sotalol, and timolol). *Turkish J. Chem.* **40**, 193–224 (2016).

32. Ma, F. *et al.* Efficient molecular evolution to generate enantioselective enzymes using a dual-channel microfluidic droplet screening platform. *Nat. Commun.* **9**, 1–8 (2018).

33. Engstrom, K., Nyhlen, J., Sandstrom, A. G. & Bäckvall, J.-E. Directed Evolution of an Enantioselective Lipase with Broad Substrate Scope for Hydrolysis of r -Substituted Esters. *J. Am. Chem. Soc.* **132**, 7038–7042 (2010).

34. Jones, K. A., Snodgrass, H. M., Belsare, K., Dickinson, B. C. & Lewis, J. C. Phage-Assisted Continuous Evolution and Selection of Enzymes for Chemical Synthesis. *ACS Cent. Sci.* **7**, 1581–1590 (2021).

35. Kille, S. *et al.* Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2**, 83–92 (2013).

36. TerMaat, J. R., Pienaar, E., Whitney, S. E., Mamedov, T. G. & Subramanian, A. Gene synthesis by integrated polymerase chain assembly and PCR amplification using a high-speed thermocycler. *J. Microbiol. Methods* **79**, 295–300 (2009).

Chapter 5

Leveraging substrate multiplexed screening to map active regions of sequence space

Jonathan Ellis has been instrumental in developing this project and has been a central contributor to the intellectual development of this work.

Chapter 5: Leveraging substrate multiplexed screening to map active regions of sequence space

5. 1. Introduction

Sequence space is be defined as the possible amino acid sequence combinations that are some number of permutations away from a parent enzyme sequence. Through convergent and divergent evolution, Nature regularly samples wide portions of sequence space for any given enzymatic function. Since enzymatic activity in Nature is improved until 'good enough', most enzymes are only moderately efficient at their native reaction and the local space around the native sequence can be rich in diverse activities.[1,2] One way scientists can take advantage of this natural sequence space diversity is by screening a diverse panel of homologs for a non-native activity of interest. Using sequence similarity networks, such an approach was used to find selective FAD-dependent monooxygenases for azaphilone biosynthesis[3] as well as to find flavin-dependent halogenases for selective indole halogenation.[4] Since all investigated sequences yield native enzymes, most studied sequences will result in properly folded and active enzymes. This traversal of Nature's sequence space is therefore a rich and comparately low-risk way to investigate native function.

Directed evolution methodologies have additionally proven powerful for investigation of sequence space. By directing sequence space traversal towards specific, potentially non-native functions, protein engineering can access different regions of sequence space than are naturally available.[5] For example, a key early example of directed evolution from Frances Arnold engineered subtilisin E for activity in high concentrations of dimethyl formamide, a commonly used organic solvent.[6] The field of biocatalysis has since leveraged the utility of directed evolution to expand our knowledge of sequence-function relationships specifically for enzymatic transformations. Many well-studied enzyme systems, such as halogenases,[7] cytochrome P450's,[8] lipases,[9] transaminases,[10] and ene-reductases[11] have been heavily engineered for wide-ranging activities on diverse, non-native substrates. During these efforts, active-site

mutagenesis is a universally successful step at altering the scope of an enzyme. One recent, well-explored model system for sequence space traversal is the tryptophan synthase β-subunit (TrpB).[12–14] Through years of laboratory evolution, the sequence space around the native TrpB sequence has been tested for activity with many diverse indole analogs as well as non-indole substrates. This large substrate space additionally means that many TrpB variants have been tested with many, diverse substrates. Due to this knowledge of variants' substrate scopes, the curated TrpB sequence space has proven useful for further engineering. Sampling from an already diverse and active set of variants increases the chance of finding activity with new substrates.

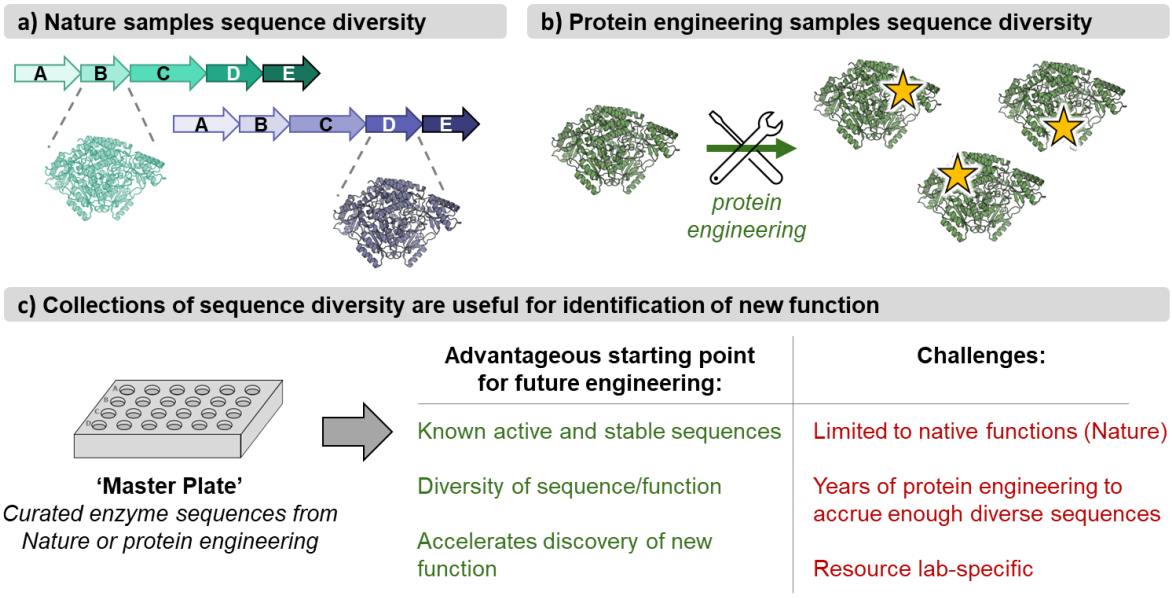


Figure 1. Sequence diversity can be curated through natural or engineered sequences.

The advantage of undertaking multiple evolutionary campaigns from the same enzyme (either iteratively or successively) is that a wide collection of active sequence spaces is accrued. Such collections of enzymes along evolutionary lineages are often dubbed 'master plates'. These proprietary troves of rich sequence diversity accelerate future engineering campaigns, as master plates can be screened to determine good starting points for gaining new function. Such

master plates have seen recurring use in P450 and TrpB engineering (Arnold),[15] and also for RebH engineering for olefin halogenation (Lewis)[16] and ene-reductases for new cross-coupling activity (Hyster).[17] These master plates consist of enzymes that were all engineered for specific reactions, and the result is a collection of diverse active site sequences that can be panned for new, untested reactions. Such in-house panels of biocatalysts thus confer large advantages to individual labs but are difficult or impossible to access for the general scientific community. Moreover, despite the success of this approach, sparse effort has been applied to the development of sequence diversity via protein engineering to create master plates from scratch.[16]

Part of the difficulty in accumulating sequence diversity stems from the challenge of studying sequence-activity relationships. It is not always clear how even single point mutations can impact enzyme activity or enzyme stability.[18] Such problems are amplified when large swathes of sequence space are sampled, as combinatorial effects from multiple mutations are famously difficult to predict.[19] This difficulty is in part due to the large impact consecutive mutations have on protein stability. Individual mutations can be highly activating on their own or even in pairs, but iteratively decrease the stability of the catalyst until the 'folding robustness threshold' of the protein is crossed.[20] Then, addition of new mutations, even if they are highly activating on their own, leads to unfolded and non-functional enzymes.[21] Such destabilized sequences are called 'holey sequences,' since the potential of a true sequence-activity relationship is undermined by a lack of stability.[22,23] A holey sequence landscape, therefore, hinders our ability to accurately model sequence-activity relationships and gain useful information from all variants screened in the sequence space.[22] A common problem of holey protein landscapes is distinguishing between destabilized sequences and stable sequences that are simply inactive on the target substrate, since most engineering campaigns are undertaken using single substrate screening. Engineering strategies that could more accurately map out

sequence-activity relationships in holey sequence spaces, therefore, would have utility in accruement of sequence diversity as starting points for protein engineering.

We successfully used substrate multiplexed screening (SUMS) to engineer *Rgn*TDC for improved activity on non-standard tryptophan (Trp) analogs as well as β-OH amino acids. However, we envisioned that the depth of information provided by SUMS could be useful not just for producing practical biocatalysts, but as a model system for exploring active site sequence space. Specifically, the rich activity information simultaneously obtained for multiple substrates would be useful for mapping out holey sequence spaces from active-site recombination. By investigating activity on more than one substrate, destabilized sequences will be more easily distinguished from sequences with low activity on just one or a few of the substrates in the screen (Fig XX). In this chapter, we describe our efforts, both past and ongoing, to intentionally seek out active regions of sequence space using *Rgn*TDC as a model system and provide a general workflow for the curation of sequence diversity.

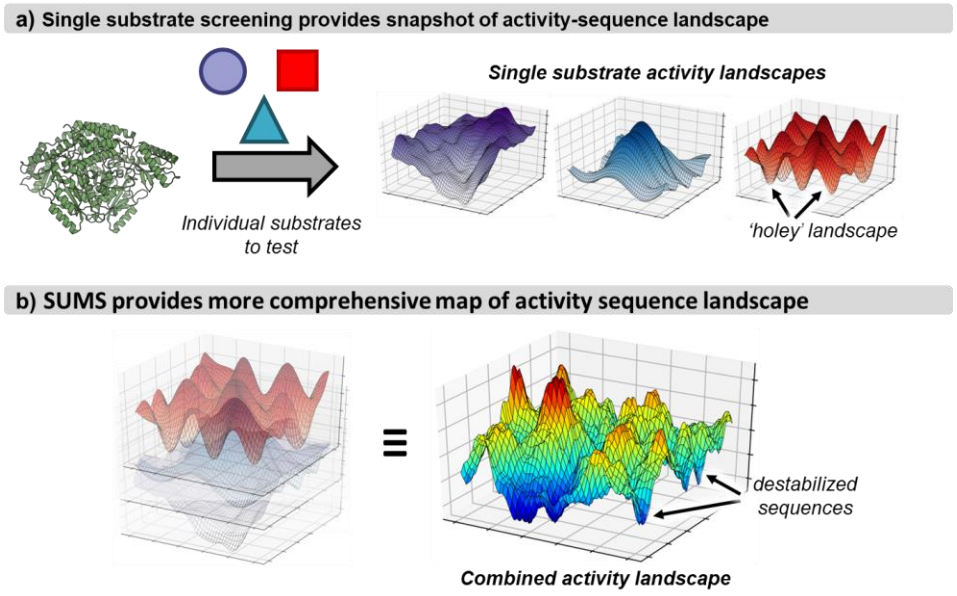


Figure 2. Sequence-activity landscapes are substrate specific. *3D surfaces created by Jon Ellis using Geogebra.*

5. 2. Results and Discussion

5. 2. 1. Initial recombination of active-site mutations in *Rgn*TDC

We previously engineered *Rgn*TDC for improved activity on non-native Trp analogs. During the engineering campaign, many point mutations within the active-site were found to positively impact activity on at least one substrate, specifically the F98, V99, L339, W349, and L355 sites (Fig 3). To access a more diverse sequence space, we set out to recombine only neutral-to-activating point mutations across these sites (Table 1). Because nearly all mutations to W349 had increased activity with 5-substituted Trp analogs, we decided to simply include all possible residues for the recombination.

Table 1. Recombined library sequence space.

Site	Included residues	# possible residues	Sequence space
F98	Ala, Arg, Cys, Gly, Ile, Leu, Met, Phe, Ser, Thr, Trp, Val	12	**28,800**
V99	Ala, Cys, Gly, Ile, Phe, Ser, Thr, Val	8	
L339	Leu, Met, Val	3	
W349	**All**	20	
L355	Ala, Leu, Met, Ser, Thr	5	

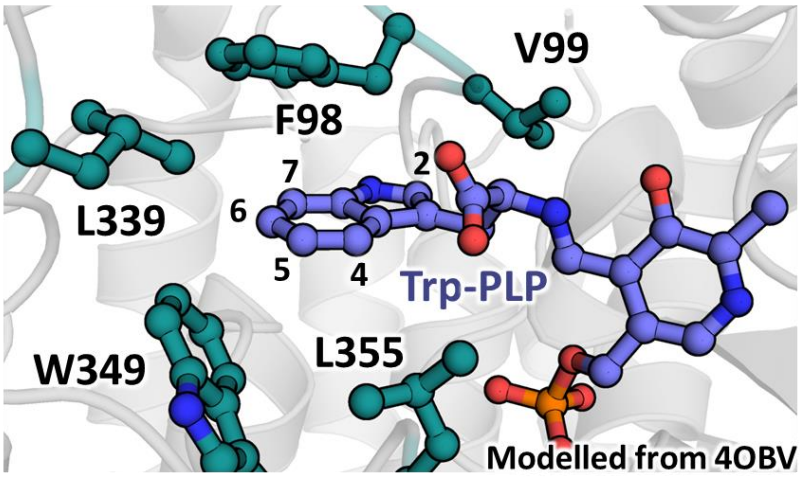


Figure 3. Active-site model of *Rgn*TDC (*PDB ID: 4OBV*).

Next, we investigated possible substrate mixtures on which to screen this recombination library. Specifically, we sought to use a substrate pool of substrates on which wt-*Rgn*TDC has relatively low activity, such that our screen would be highly sensitive both to shifts in relative specificity as well as boosts in activity. Because *Rgn*TDC was highly active with 6- and 7-substituted Trp analogs, these substrates would outcompete the less active substrates in competition and were therefore excluded from consideration at this stage. Previously we had observed that activity on the similar Trp substituents 5-OMe and 5-OEt was not well-correlated, so we sought pairs of 4- and 5-substituted Trp analogs. This reasoning led to five substrates, with which wt-*Rgn*TDC has < 100 TON: β-Me-Trp, 4-CN-Trp, 4-OMe-Trp, 5-OEt-Trp, and 5-NO2-Trp (Fig 4). Because screening itself would eventually take place in cell lysate, we also reasoned that native activity could be assessed in a qualitative fashion by measuring accumulation of tryptamine.

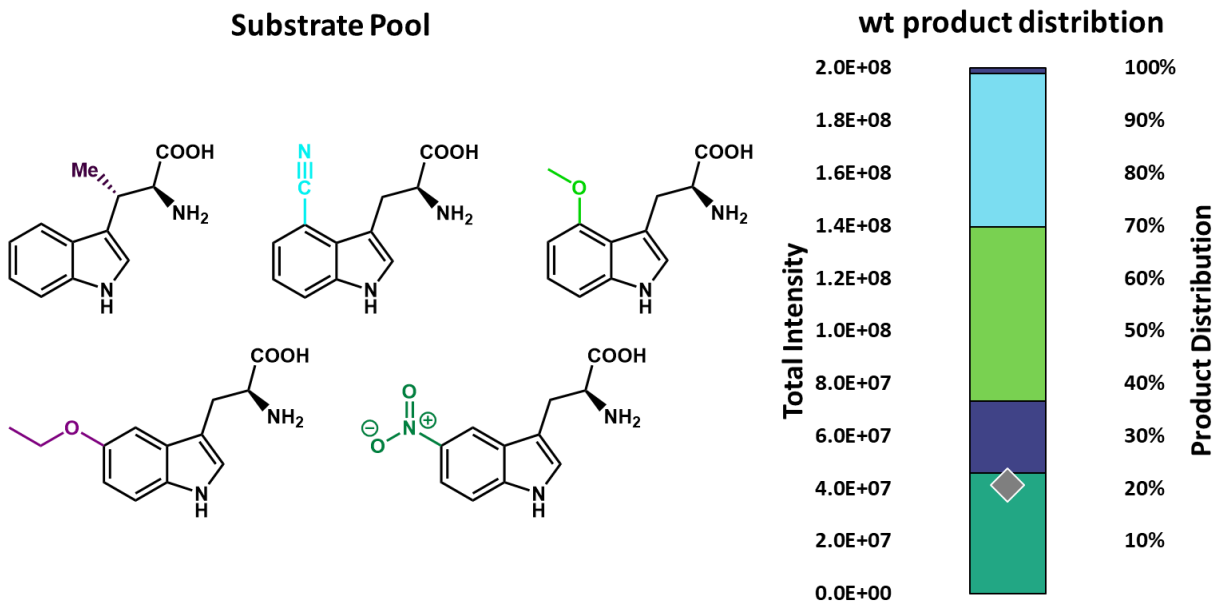


Figure 4. SUMS substrate pool and wt-*Rgn*TDC product profile. Conditions for multiplexed reaction: 2 mM each sub, 2 µM *Rgn*TDC, 500 uM PLP, 50 mM KPi pH = 8.0; 4 h @ 37 °C.

We screened and sequenced ~150 variants from the active site recombination library with this substrate mixture (Fig 5a). We observed that 23% of the library displayed any catalytic activity, and only a handful of variants had improvements over wt-*Rgn*TDC. Most active sequences were double and triple mutants, with diverse mutations observed for all sites of mutagenesis. Different variants in the library displayed varied product profiles (Fig 5b). Since no single mutation in this space was a deactivating mutation, we hypothesized that combinatorial effects, either cooperative effects on catalysis or on thermal stability, are leading to inactive enzymes, leading to a holey sequence space. Such occurrences are well-documented in the literature.[18,19] If few variants of the sequence space display activity, the onus of screening thousands of variants becomes the main route to finding desired activity. Therefore, even with a 'smart' library construction consisting of carefully chosen sites and mutations, much of the observed sequence space was inactive.

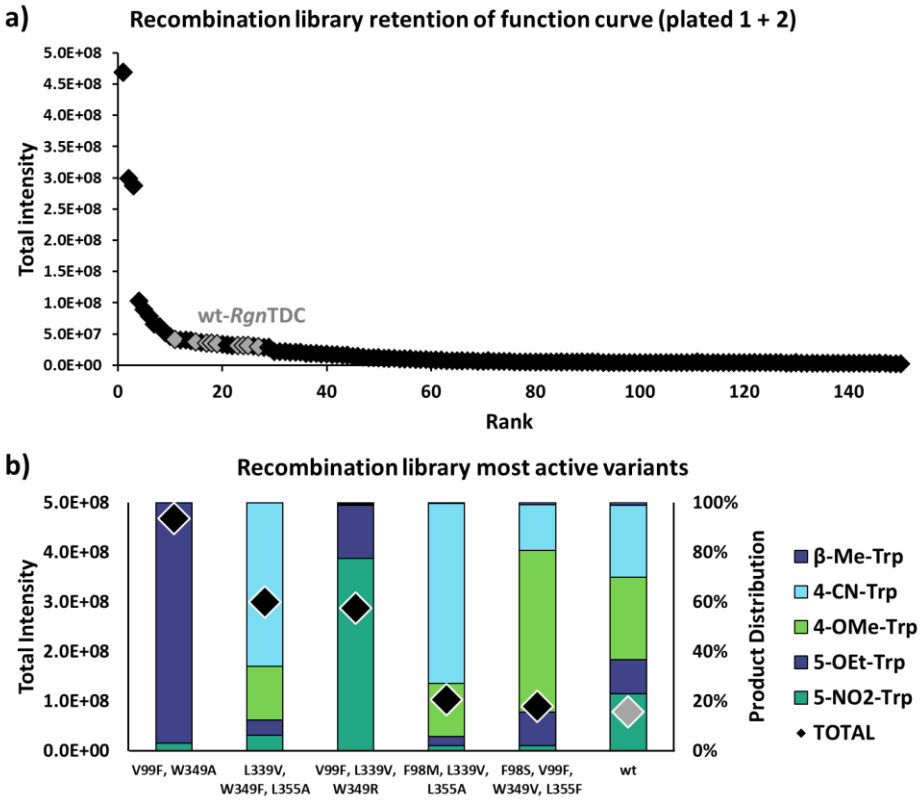


Figure 5. Screening results from initial recombination library. a) SUMS reaction: 2 mM β-Me-, 4-OMe-, 4-CN-, 5-NO2-, 5-OEt-Trp, 180 ul TDC lysate.; 50 mM KPi pH = 8.0, 4 h @ 37 °C. Wild-type sequences are denoted in gray diamonds. **b)** SUMS product profile of the most active variants from library.

5. 2. 2. Wild-type primer doping selectively tunes sequence space

Because the recombination sequence space was generally inactive, we decided to refocus the region of sequence space we intended to screen. We calculated the likelihood of finding variants with given numbers of simultaneous mutations for this original library, and found the library enriched in quadruple and quintuple mutation variants (Fig 6 left), with an average mutational rate of 4.4 mutations per variant. To focus this sequence space to variants with fewer mutations, we implemented a strategy we dub ‘wt primer doping.’ By including additional copies of the wt sequence during library generation, we increase the likelihood that a site resamples the wt residue. Consequently, the sequence space sampled during screening will be enriched in double and triple mutant variants (Fig 6 right). Extensive review of the literature has found that

this mutational approach has been applied to decrease mutual mutational abundance in the generation of SELEX libraries, but we found few examples of this approach being used for protein engineering.[24]

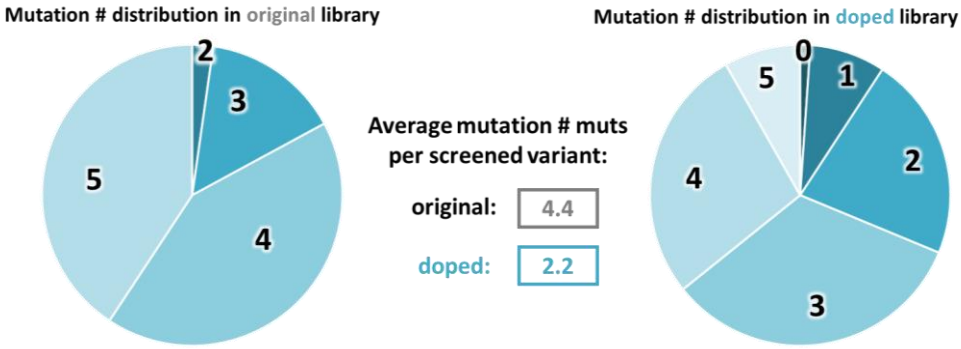


Figure 6. Calculated mutation distribution in original (left) or doped (right) recombination libraries.

During library generation, we doped in wt primers such that we expect an average mutational rate of 2.5 mutations per variant. With this newly generated library, we screened and sequenced ~300 variants comprising 200 unique sequences, with an average mutational rate of 2.2 mutations per variant. We observed nearly twice as many active sequences than from the original library, with ~45% of variants in the library showing activity above baseline (Fig 7a). Many variants were significantly more active than the parent sequence, with varying specificities for the substrates in the screen (Fig 7b). Variants that were specialists for either 4-substitued or 5-substited Trp analogs were common, although generally improved variants were also observed.

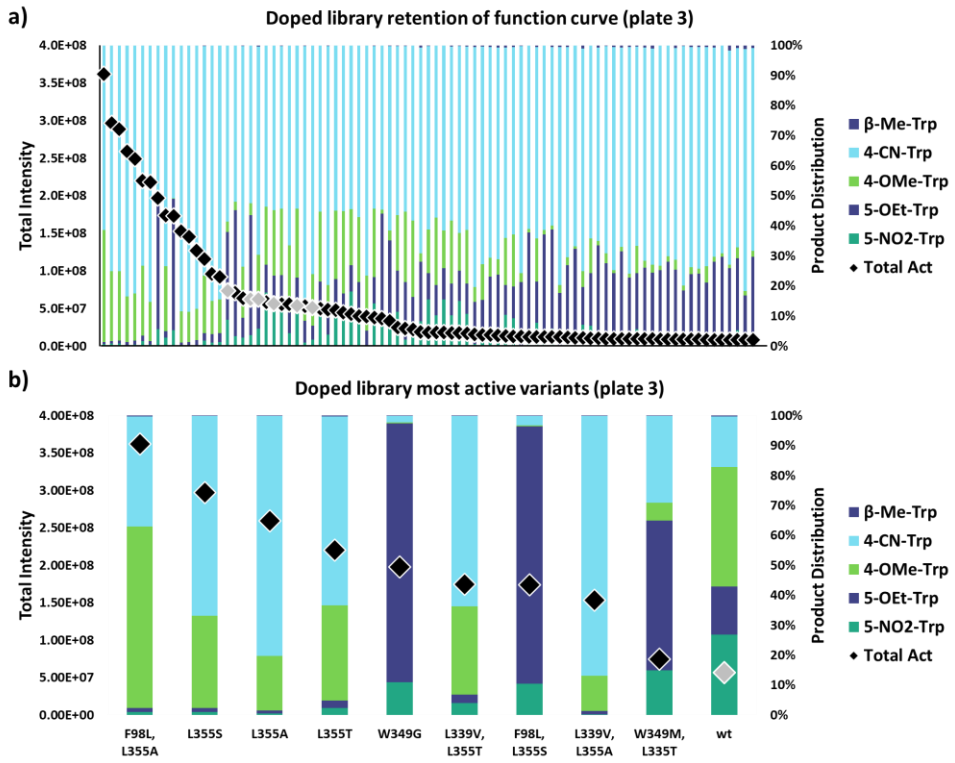


Figure 7. Example retention of function curve from doped recombination library. a) Full retention of function curve for plate 3. Reaction conditions: 2 mM β-Me-, 4-OMe-, 4-CN-, 5-NO2-, 5-OEt-Trp, 180 ul TDC lysate.; 50 mM KPi pH = 8.0, 4 h @ 37 °C. Wild-type sequences are denoted in gray diamonds. **b)** Zoomed in retention of function curve of select wells with sequence information.

We then analyzed the sequence space of this doped library. We observed a large share of single and double mutants with fewer variants with 3+ mutations (Fig 8a). Additionally, our hypothesis that variants with more mutations were less likely to be highly active was held true for this sampling of sequence space (Fig 8b). We did, however, observe some triple and quadruple mutation variants with activity, showing that even large perturbations to the active site can be tolerated. However, even though more than half of the sequence space is inactive, due to wt primer doping, the mutational rate is sufficiently low such that we resample the individual point mutants. Our stated goal for this sequence space analysis was to investigate recombination of mutations, so resampling point mutants is undesirable.

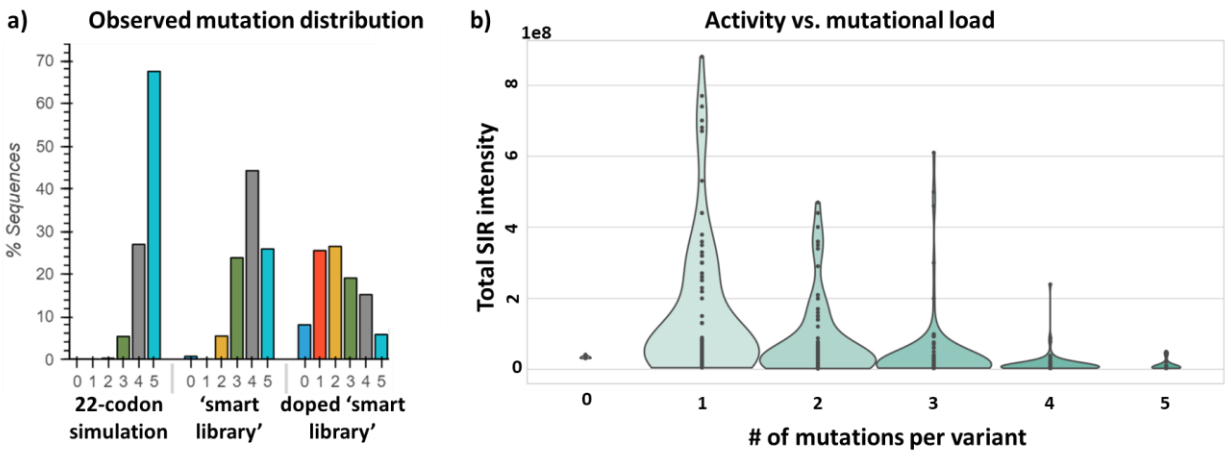


Figure 8. Analysis of wt primer doped library screening results. a) Mutation distribution for: 20 residue recombination at all five sites of interest (left, calculated), original recombination library (middle, observed), and doped recombination library (right, observed). *Figure adapted from Jon Ellis.* **b)** Violin plot of observed activity vs. # mutations for *Rgn*TDC variants in the doped recombination library. Each point represents data from one well.

5. 2. 3. Removal of negative cooperativity mutations increases activity of sequence space

The general lack of active triple and quadruple variants prompted us to again reconsider the sequence space being screened. We classified mutations based on their prevalence in high activity or low activity variants. The code for this analysis was written by Jon Ellis. By looking at this enrichment in either high activity or low activity variants, mutations were ranked on their ability to engage in positive cooperativity with other mutations in the active site. An example depiction of this type of analysis is shown in Figure 7. Note that for the high activity sequences, the L355 site not conserved, with many possible mutations (Fig 9a). The F98, L339, and W349 sites are also amenable to mutation, but to a lesser degree. The V99 site, however, is highly conserved for these sequences, indicating that mutation to this site is deleterious to overall activity. Conversely, among inactive sequences V99 is frequently mutated, indicating that such mutations are present in the library, but are generally inactive (Fig 9b). We emphasize the unpredictable nature of this finding, as mutation at V99 was activating for 2-Me-Trp and 4-Br-Trp (Chapter 3). Additionally, we noted the serendipitous mutation I343N had appeared during

library generation and was present in ~60 variants. This mutation was enriched in active variants, and so was included in future library designs.

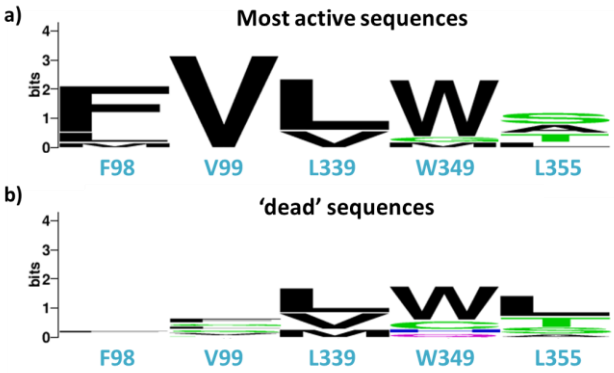


Figure 9. Most observed mutations at sites of interest as WebLogos. a) Mutations for the twelve most active sequences. **b)** Mutations for twelve inactive sequences. Figure created using WebLogo website at https://weblogo.berkeley.edu/logo.cgi.

By leveraging such activity-sequence information, we redesigned the library sequence space. We excluded the nine most destabilizing mutations, which includes any mutation to V99, W349E, and W349Q. Notably, these mutations to W349 are highly activating on their own for 5-substituted Trps. This edit shrinks the sequence space to 4384 unique sequences. We again used wt primer doping to generate recombination libraries but reasoned that by removing the 'bad actors' we could increase the mutational load to and aimed for 3.0 mutations per variant, which would significantly reduce the probability of re-sampling point mutations. We screened ~180 sequences and were pleased to observe that, despite the a higher mutational rate, this library had ~45% active sequences, (Fig 10a). We sequenced four of the most active variants and found that all of these variants contained 3+ mutations (Fig 10b). These variants also display unique specificities for the substrate panel, suggestive of their divergent active-site geometries. The remaining sequences of this library will also be sequenced in the near future.

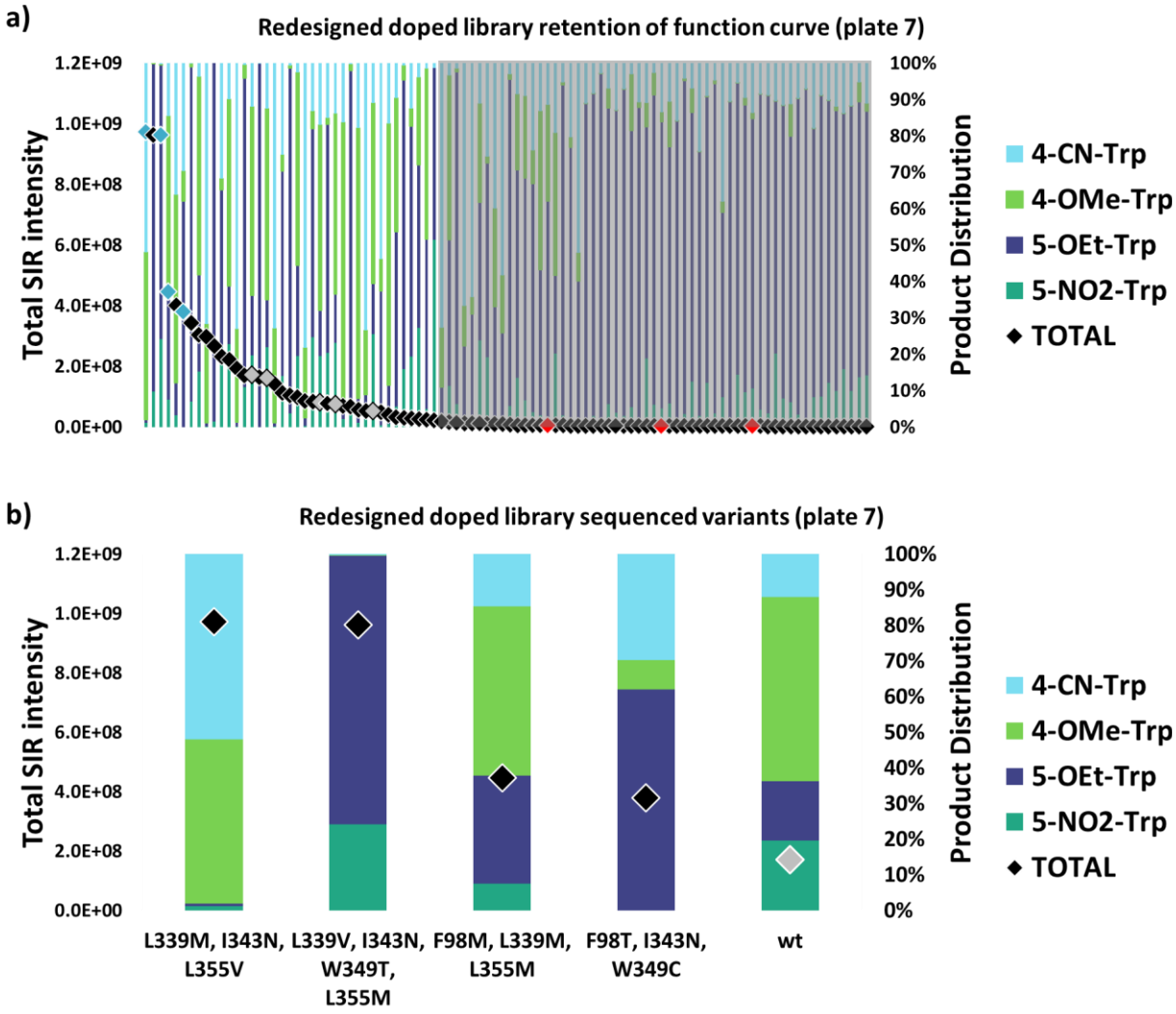


Figure 10. Retention of function curve from redesigned doped recombination library. **a)** Full retention of function curve for plate 8. Reaction conditions: 2 mM β-Me-, 4-OMe-, 4-CN-, 5-NO2-, 5-OEt-Trp, 180 ul TDC lysate.; 50 mM KPi pH = 8.0, 4 h @ 37 °C. Wild-type sequences are denoted in gray diamonds, sequenced wells are denoted with cyan diamonds, and negative controls are denoted with red diamonds. **b)** Specificity of sequenced variants compared to the wt-*Rgn*TDC specificity.

This refined library will subsequently be screened with new substrate mixes to investigate the generality (or lack thereof) of the focused sequence space. 6- and 7-substitued Trp analogs will be tested as well as non-Trp amino acids. Understanding the importance of adding more diversity to the initial screening stages will better inform users of this methodology to best traverse the sequence space for a given enzyme.

5. 3. Conclusions

Efficiently exploring protein sequence space is difficult and finding active areas of sequence space is an ongoing challenge for biocatalysis. SUMS provides advantages over single-substrate screening to seek out active sequences due to simultaneously reading out on activity from multiple, diverse substrates. This additional information better distinguishes ‘dead’ sequences from inactive sequences, allowing researchers to focus in on active sequence space for a desired transformation. By leveraging SUMS information, we were able to identify a sequence space enriched with highly active triple and quadruple mutant variants with activities on substrates not included in the original substrate pool. This approach represents the targeted sampling of just NN% of the original 5-site recombination sequence space. This methodology represents a streamlined and efficient route to quickly create a library of diverse, active sequences for biocatalysts of interest.

5. 4. Materials and Methods

Chemicals and reagents were purchased from commercial suppliers (Sigma-Aldrich, VWR, Chem-Impex International, Alfa Aesar, Combi-blocks, Oakwood Products) at the highest quality available and used without further purification unless stated otherwise. Genes were purchased as gBlocks from Integrated DNA Technologies (IDT). E. coli cells were electroporated with a Bio-Rad MicroPulser at 2500 V. New Brunswick I26R, 120 V/60 Hz shaker incubators (Eppendorf) were used for cell growth. Cell disruption via sonication was performed with a Sonic Dismembrator 550 (Fisher Scientific) sonicator. Optical density measurements were collected using an optical density reader (Amersham Biosciences). Ultra-high pressure liquid chromatography-mass spectrometry (UPLC-MS) data were collected on an Acquity UPLC (Waters) equipped with an Acquity PDA and QDA MS detector using either a BEH C18 column (Waters) or an Intrada Amino Acid column (Imtakt). Preparative column separations were

performed on an Isolera One Flash Purification system (Biotage). NMR data were collected on a Bruker 500 MHz spectrometer equipped with a DCH cryoprobe. Signal positions were recorded in ppm with the abbreviations s, d, t, q, dd, and m, denoting singlet, doublet, triplet, quartet, doublet of doublets, and multiplet respectively. All coupling constants J are measured in Hz. High resolution mass spectrometry data were collected with a Q Extractive Plus Orbitrap (NIH 1S10OD020022-1) instrument with samples ionized by ESI.

5. 4. 1. Plasmid and protein information

Protein sequence of *C-His-RgnTDC* (Uniprot accession code: A7B1V0):

MSQVIKKKRNTFMIGTEYILNSTQLEEAIKSFVHDFCAEKHEIHDQPVVVEAKEHQEDKIKQIKIP
EKGRPVNEVVSEMMNEVYRYRGDANHPRFFSFVPGPASSVSWLGDIMTSAYNIHAGGSKLAP
MVNCIEQEVLKWLAKQVGFTENPGGVFVSGGSMANITALTAARDNKLTDINLHLGTAYISDQTH
SSVAKGLRIIGITDSRIRRIPTNSHFQMDTTKLEEAIETDKKSGYIPFVVIGTAGTTNTGSIDPLTEI
SALCKKHDMWFHIDGAYGASVLLSPKYKSLLTGTGLADSISWDAHKWLFQTYGCAMVLVKDIR
NLFHSFHVNPEYLKDLENDIDNVNTWDIGMELTRPARGLKLWLTLQVLGSDLIGSAIEHGFQLA
VWAEEALNPKKDWEIVSPAQMAMINFRYAPKDLTKEEQDILNEKISHRILESGYAAIFTTVLNGK
TVLRICAIHPEATQEDMQHTIDLLDQYGREIYTEMKKALEHHHHHH

DNA sequence of *C-His-RgnTDC*:

ATGTCGCAGGTCATTAAGAAAAAACGCAATACGTTTATGATTGGAACGGAGTACATCCTTAA
TTCGACACAGTTAGAGGAGGCAATTAAGTCTTTCGTGCACGATTTTTGTGCGGAAAAACAT
GAGATCCATGATCAGCCCGTCGTTGTTGAAGCCAAGGAGCACCAGGAAGATAAAATTAAG
CAGATCAAGATCCCTGAAAAAGGACGCCCAGTAAATGAGGTCGTGAGTGAGATGATGAAT
GAAGTTTACCGCTATCGCGGAGATGCGAACCACCCCCGTTTCTTCTCCTTCGTTCCGGGTC
CAGCTTCGAGCGTCTCCTGGCTTGGAGACATCATGACGAGTGCATATAATATCCATGCCGG
AGGCAGTAAATTGGCTCCCATGGTAAACTGTATTGAGCAAGAAGTGCTGAAGTGGTTGGCA

AAGCAAGTGGGATTTACTGAAAATCCCGGCGGGGTGTTCGTCTCAGGTGGCTCGATGGCG
AACATCACGGCGTTAACAGCAGCCCGTGACAATAAACTTACTGACATTAATTTGCATTTAGG
AACGGCGTATATCAGCGACCAAACACACTCCAGTGTAGCCAAGGGGTTACGTATTATTGGC
ATCACCGACAGCCGTATTCGCCGTATTCCCACTAATTCGCACTTCCAAATGGATACGACCA
AGTTGGAGGAGGCCATTGAAACCGATAAAAAGAGTGGCTATATCCCGTTTGTAGTGATCGG
AACCGCTGGCACGACTAATACAGGATCCATTGACCCATTAACGGAAATTTCTGCATTATGTA
AAAAGCACGATATGTGGTTCCACATCGACGGTGCGTATGGTGCCTCCGTATTGCTTAGTCC
AAAATATAAGTCCCTTTTGACAGGAACAGGATTAGCAGATAGTATTTCTTGGGATGCTCACA
AATGGTTATTCCAGACGTATGGGTGCGCCATGGTATTGGTGAAGGACATCCGCAACCTGTT
CCATTCCTTTCACGTTAACCCCGAATATCTGAAAGACCTTGAGAATGACATTGATAATGTCA
ATACGTGGGATATTGGGATGGAGTTAACACGTCCGGCACGTGGGCTTAAACTTTGGCTGA
CCTTGCAGGTTCTTGGGTCCGACCTTATTGGGTCTGCAATTGAGCACGGTTTCCAATTAGC
GGTATGGGCGGAAGAAGCGCTGAATCCCAAAAAAGATTGGGAAATTGTTAGCCCTGCCCA
GATGGCGATGATTAATTTTCGCTACGCGCCTAAGGATTTAACCAAAGAGGAGCAGGACATC
CTTAATGAAAAGATTTCGCATCGCATCTTGGAATCAGGCTATGCCGCTATTTTTACTACTGT
GCTGAATGGTAAGACAGTGTTACGCATTTGCGCGATTCACCCTGAGGCTACTCAAGAGGAT
ATGCAGCACACCATTGATCTGTTGGACCAATACGGTCGCGAGATCTATACTGAAATGAAAA
AGGCTCTCGAGCACCATCACCATCACCATTGA

Cloning, expression, purification, and storage of RgnTDC variants

A codon-optimized copy of the *Ruminococcus gnavus* tryptophan decarboxylase (*Rgn*TDC) gene was purchased as a gBlock from Integrated DNA Technologies. This DNA fragment was inserted into a pET22b vector by the Gibson Assembly method.[1] BL21 (DE3) E. coli cells were subsequently transformed with the resulting cyclized DNA product via electroporation. After 30 min of recovery in LB media at 37 °C, cells were plated onto LB plates with 100 µg/mL ampicillin (AMP) and incubated overnight. Single colonies were used to inoculate 5 mL TB + 100 µg/mL

AMP (TB-AMP), which were grown overnight at 37 °C, 200 rpm. Expression cultures, typically 1 L of TB-AMP were inoculated from these starter cultures and shaken (180 rpm) at 37 °C. After 3.5 to 4 hours (OD_{600} > 1.5), the expression cultures were chilled on ice. After 45 min on ice, expression was induced with 1 mM IPTG, and the cultures were supplemented with 0.5 mM indole. Cultures were expressed overnight at 23 °C with shaking at 180 rpm. Cells were then harvested by centrifugation at 4300xg at 4 °C for 15 min. Cell pellets were frozen and stored at -20 °C until purification.

To purify TDC, cell pellets were thawed on ice and then resuspended in lysis buffer (50 mM potassium phosphate buffer (pH = 8.0), 1 mg/ml Hen Egg White Lysozyme (GoldBio), 0.2 mg/ml DNaseI (GoldBio), 1 mM $MgCl_2$, and 200 µM pyridoxal 5′-phosphate (PLP)). A volume of 4 mL of lysis buffer per gram of wet cell pellet was used. After 45 min of shaking at 37 °C, cells were sonicated with a ½ in. tip for 10 min (1 s on; 1 s off). The resulting lysate was then spun down at 75,000×g to pellet cell debris. Ni/NTA beads (GoldBio) were added to a gravity column and the lysis supernatant ran over the bead bed for purification by Ni-affinity chromatography. The column was washed with 4 column volumes of 20 mM imidazole, 50 mM potassium phosphate buffer (pH = 8.0). Washing with higher concentrations of imidazole resulted in slow protein elution. TDC was eluted with 250 mM imidazole, 50 mM potassium phosphate buffer (pH = 8.0). Elution of the desired protein product was monitored by the disappearance of its bright yellow color (resulting from the release of TDC) from the column. The protein product was dialyzed to < 50 µM imidazole in 50 mM Tris-HCl buffer (pH = 8.01) or 50 mM potassium phosphate buffer (pH = 8.08). Purified enzyme was flash frozen in pellet form by pipetting enzyme dropwise into a crystallization dish filled with liquid nitrogen. The enzyme was transferred to a plastic conical and stored at -80 °C until further use. Frozen pellets were thawed at room temperature and centrifuged before use. The concentration of protein was determined by Bradford assay after freeze-thawing using bovine serum albumin for a standard concentration curve. Generally, this

procedure yielded > 100 mg per L culture. Protein purity was analyzed by sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel electrophoresis using 12% polyacrylamide gels.

5. 4. 2. General methods

Generation of original recombination library

Primers were purchased from Integrated DNA Technologies. Five active-site residues were targeted for mutagenesis, based on site-saturation mutagenesis library results from Chapter 2. The following degenerate codons were used for library construction: **F98**: DBK; **V99**: DBC; **L339**: DTG; **W349**: NDT + VHG + TGG; **L355**: DCA + ATG + TTG. Each gene library was amplified first as two separate fragments and then combined via polymerase chain assembly (PCA) to form full-length RgnTDC gene mutagenized at the sites of interest.[25] The corresponding genes were then inserted into a pET22b vector as described above and then transformed into BL21(DE3) E. coli cells and plated on LB + 100 µg/mL AMP agar plates.

Generation of wt primer doped recombination library

Wild-type codon primers for the five active-site positions were purchased from Integrated DNA Technologies. The following degenerate codons were used for library construction: **F98**: DBK; **V99**: DBC; **L339**: DTG; **W349**: NDT + VHG + TGG; **L355**: DCA + ATG + TTG. Each gene library was amplified first as two separate fragments, with inclusion of wt primers along with the degenerate codon-containing mutagenesis primers. Ratio of mutagenesis primers to wild-type primer for each amplification was 3:2, with an expected mutational rate of 2.5. The PCR-generated fragments were then combined via polymerase chain assembly (PCA) to form full-length RgnTDC gene mutagenized at the sites of interest.[25] The corresponding genes were then inserted into a pET22b vector as described above and then transformed into BL21(DE3) E. coli cells and plated on LB + 100 µg/mL AMP agar plates.

Generation of redesigned wt primer doped recombination library

Wild-type codon primers for the five active-site positions were purchased from Integrated DNA Technologies. The following degenerate codons were used for library construction: **F98**: BKC + AYS + TGG + GCG; **L339**: DTG; **I343N:** AAT**; W349**: NDT + VYS + AAA + TGG; **L355**: DCA + ATG + TTG. Each gene library was amplified first as two separate fragments, with inclusion of wt primers along with the degenerate codon-containing mutagenesis primers. Ratio of mutagenesis primers to wild-type primer for each amplification was 7:3, with an expected mutational rate of 3.0. The PCR-generated fragments were then combined via polymerase chain assembly (PCA) to form full-length RgnTDC gene mutagenized at the sites of interest.[25] The corresponding genes were then inserted into a pET22b vector as described above and then transformed into BL21(DE3) E. coli cells and plated on LB + 100 µg/mL AMP agar plates.

Screening of RgnTDC libraries

Cell pellets were thawed and then resuspended in lysis buffer: 50 mM potassium phosphate buffer (pH = 8.0), 1 mg/mL Hen Egg White Lysozyme (GoldBio), 0.2 mg/mL DNaseI (GoldBio), 1 mM MgCl2, and 300 µM pyridoxal 5′-phosphate (PLP). A volume of 600 µL lysis buffer per well was used. After 1 h of shaking at 37 °C, the resulting lysate was then spun down at 4000 x*g* to pellet cell debris. Then, 180 µL of the resulting supernatant was added to 20 µL of a substrate mixture in a separate reaction plate. The substrate master mix contained 20 mM each Trp substrate dissolved in 50% MeOH, 40% H_2O, and 10% 1 M HCl. Final substrate concentrations are as follows: 2 mM **β-methyltryptophan**, 2 mM **4-cyanotryptophan**, 2 mM **4-methoxytryptophan**, 2 mM **5-ethoxytryptophan**, and 2 mM **5-nitrotryptophan**. Reactions were incubated at 37 °C for 4 h, and then 100 µL reaction solution was quenched via addition of 200 µL acetonitrile and centrifuged at 4000 x*g* for 10 min. 200 µL of the quenched reaction mixture supernatant was filtered into a 96-well plate for UPLC-MS analysis. Data were collectesd on an Acquity UHPLC with an Acquity QDA MS detector (Waters) using an Intrada

column (Imtakt). Product m/z ion counts were used to assess product formation from the reaction mixture.

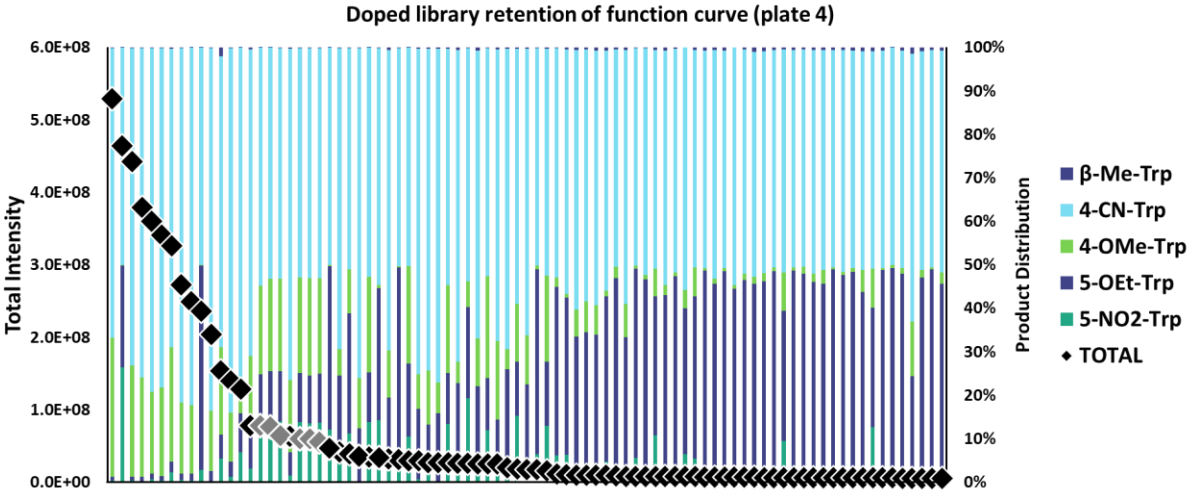


Figure 11. Retention of function curve from doped recombination library (plate 4). Reaction conditions: 2 mM β-Me-, 4-OMe-, 4-CN-, 5-NO2-, 5-OEt-Trp, 180 ul TDC lysate.; 50 mM KPi pH = 8.0, 4 h @ 37 °C. Wild-type sequences are denoted in gray diamonds.

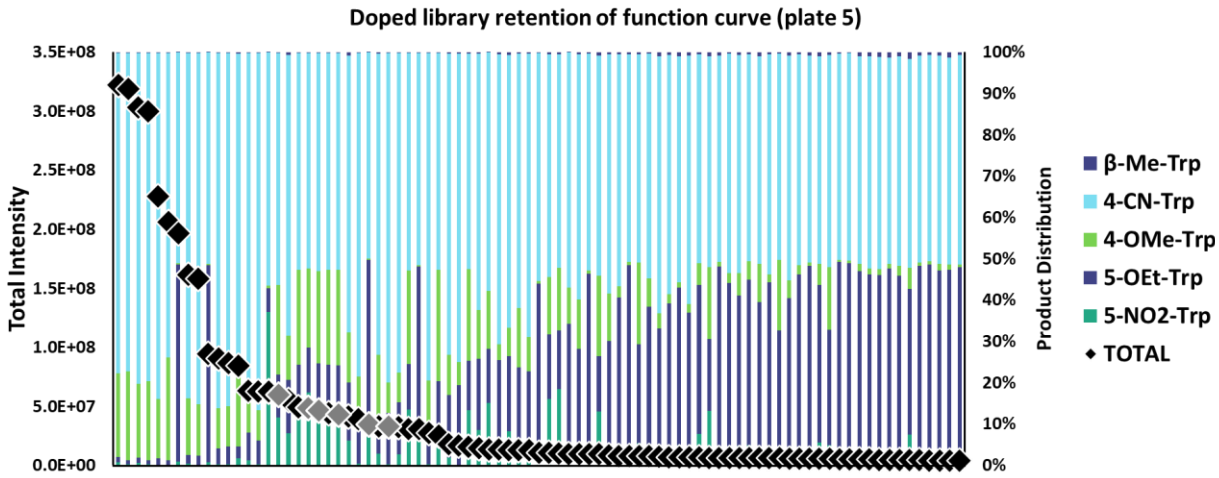


Figure 12. Retention of function curve from doped recombination library (plate 5). Reaction conditions: 2 mM β-Me-, 4-OMe-, 4-CN-, 5-NO2-, 5-OEt-Trp, 180 ul TDC lysate.; 50 mM KPi pH = 8.0, 4 h @ 37 °C. Wild-type sequences are denoted in gray diamonds.

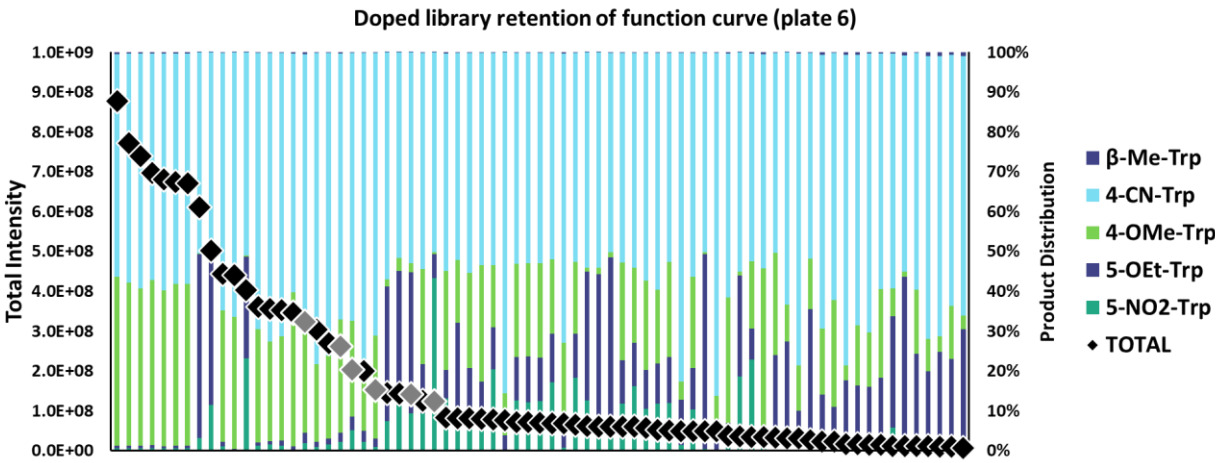


Figure 13. Retention of function curve from doped recombination library (plate 6). Reaction conditions: 2 mM β-Me-, 4-OMe-, 4-CN-, 5-NO2-, 5-OEt-Trp, 180 ul TDC lysate.; 50 mM KPi pH = 8.0, 4 h @ 37 °C. Wild-type sequences are denoted in gray diamonds.

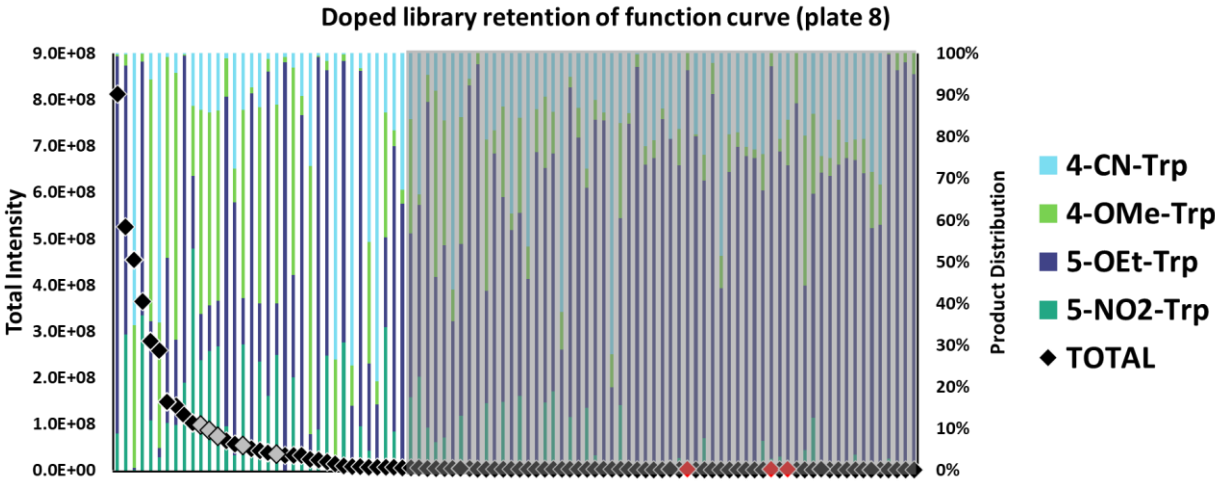


Figure 14. Retention of function curve from redesigned doped recombination library (plate 8). Reaction conditions: 2 mM β-Me-, 4-OMe-, 4-CN-, 5-NO2-, 5-OEt-Trp, 180 ul TDC lysate.; 50 mM KPi pH = 8.0, 4 h @ 37 °C. Wild-type sequences are denoted in gray diamonds.

5. 5. References

1. Bar-Even, A. *et al.* The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410 (2011).
2. Khersonsky, O. & Tawfik, D. S. Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. *http://dx.doi.org/10.1146/annurev-biochem-030409-143718* **79**, 471–505 (2010).
3. Pyser, J. B. *et al.* Stereodivergent, Chemoenzymatic Synthesis of Azaphilone Natural Products. *J. Am. Chem. Soc.* **141**, 18551–18559 (2019).
4. Fisher, B. F., Snodgrass, H. M., Jones, K. A., Andorfer, M. C. & Lewis, J. C. Site-Selective C-H Halogenation Using Flavin-Dependent Halogenases Identified via Family-Wide Activity Profiling. *ACS Cent. Sci.* **5**, 1844–1856 (2019).
5. Romero, P. A. & Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nature Reviews Molecular Cell Biology* vol. 10 866–876 (2009).
6. Chen, K. & Arnold, F. H. Enzyme engineering for nonaqueous solvents: random mutagenesis to enhance activity of subtilisin E in polar organic media. *Bio/Technology* **9**, 1073–1077 (1991).
7. Andorfer, M. C. *et al.* Understanding Flavin-Dependent Halogenase Reactivity via Substrate Activity Profiling. *ACS Catal.* **7**, 1897–1904 (2017).
8. Jung, S. T., Lauchli, R. & Arnold, F. H. Cytochrome P450: Taming a wild type enzyme. *Curr. Opin. Biotechnol.* **22**, 809–817 (2011).
9. Dasetty, S., Blenner, M. A. & Sarupria, S. Engineering Lipases: Walking the fine line between activity and stability. *Mater. Res. Express* **4**, (2017).
10. Guo, F. & Berglund, P. Transaminase biocatalysis: optimization and application. *Green Chem.* **19**, 333–360 (2017).
11. Kumar Roy, T., Sreedharan, R., Ghosh, P., Gandhi, T. & Maiti, D. Ene-Reductase: A Multifaceted Biocatalyst in Organic Synthesis. *Chem. – A Eur. J.* **28**, e202103949 (2022).
12. 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. *J. Am. Chem. Soc.* **139**, 10769–10776 (2017).
13. Watkins-Dulaney, E., Straathof, S. & Arnold, F. Tryptophan Synthase: Biocatalyst Extraordinaire. *ChemBioChem* vol. 22 5–16 (2021).
14. Almhjell, P. J., Boville, C. E. & Arnold, F. H. Engineering enzymes for noncanonical amino acid synthesis. *Chemical Society Reviews* vol. 47 8980–8997 (2018).
15. Romney, D. K., Sarai, N. S. & Arnold, F. H. Nitroalkanes as Versatile Nucleophiles for Enzymatic Synthesis of Noncanonical Amino Acids. *ACS Catal.* **9**, 8726–8730 (2019).

16. Jiang, Y., Mondal, D. & Lewis, J. C. Expanding the Reactivity of Flavin Dependent Halogenases Toward Olefins via Enantioselective Intramolecular Haloetherification and Chemoenzymatic Oxidative Rearrangements Yuhua. *ChemRxiv* 1–6 (2022).

17. Fu, H. *et al.* An Asymmetric sp 3-sp 3 Cross-Electrophile Coupling Using Biocatalysis. *Chem RXIV* 1–17 (2022).

18. Bloom, J. D. & Glassman, M. J. Inferring Stabilizing Mutations from Protein Phylogenies : Application to Influenza Hemagglutinin. *PLOS Comput. Biol.* **5**, (2009).

19. Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N. & Tawfik, D. S. Robustness – epistasis link shapes the fitness landscape of a randomly drifting protein. *Nat. Lett.* **444**, 929–932 (2006).

20. Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N. & Tawfik, D. S. Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature* **444**, 929–932 (2006).

21. Bloom, J. D., Labthavikul, S. T., Otey, C. R. & Arnold, F. H. Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5869–5874 (2006).

22. Romero, P. A., Krause, A. & Arnold, F. H. Navigating the protein fitness landscape with Gaussian processes. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E193–E201 (2013).

23. Gavrilets, S. Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.* **12**, 307–312 (1997).

24. Bartel, D. P., Zapp, M. L., Green, M. R. & Szostak, J. W. HIV-1 rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **67**, 529–536 (1991).

25. TerMaat, J. R., Pienaar, E., Whitney, S. E., Mamedov, T. G. & Subramanian, A. Gene synthesis by integrated polymerase chain assembly and PCR amplification using a high-speed thermocycler. *J. Microbiol. Methods* **79**, 295–300 (2009).