

Engineered Biosynthesis of β -Alkyl Tryptophan Analogs

Christina E. Boville, Remkes A. Scheele, Philipp Koch, Sabine Brinkmann-Chen, Andrew R. Buller, **Frances H. Arnold**

Submitted date: 13/07/2018 • Posted date: 13/07/2018

Licence: CC BY-NC-ND 4.0

Citation information: Boville, Christina E.; Scheele, Remkes A.; Koch, Philipp; Brinkmann-Chen, Sabine; Buller, Andrew R.; Arnold, Frances H. (2018): Engineered Biosynthesis of β -Alkyl Tryptophan Analogs. ChemRxiv. Preprint.

Non-canonical amino acids (ncAAs) with dual stereocenters at the α and β positions are valuable precursors to natural products and therapeutics. Despite the potential applications of such bioactive β -branched ncAAs, their availability is limited due to the inefficiency of the multi-step methods used to prepare them. Here we report a stereoselective biocatalytic synthesis of β -branched tryptophan analogs using an engineered variant of *Pyrococcus furiosus* tryptophan synthase (PfTrpB), PfTrpB^{7E6}. PfTrpB^{7E6} is the first biocatalyst to synthesize bulky β -branched tryptophan analogs in a single step, with demonstrated access to 27 ncAAs. The molecular basis for the efficient catalysis and broad substrate tolerance of PfTrpB^{7E6} was explored through X-ray crystallography and UV-visible light spectroscopy, which revealed that a combination of active-site and remote mutations increase the abundance and persistence of a key reactive intermediate. PfTrpB^{7E6} provides an operationally simple and environmentally benign platform for preparation of β -branched tryptophan building blocks.

File list (2)

(b) Manuscript ChemRxiv.pdf (1.01 MiB)

[view on ChemRxiv](#) • [download file](#)

(c) Supplemental Information ChemRxiv.pdf (2.45 MiB)

[view on ChemRxiv](#) • [download file](#)

Engineered biosynthesis of β -alkyl tryptophan analogs

Christina E. Boville⁺, Remkes A. Scheele⁺, Philipp Koch, Sabine Brinkmann-Chen, and Andrew R. Buller^{*}, Frances H. Arnold^{*}

Abstract: Non-canonical amino acids (ncAAs) with dual stereocenters at the α and β positions are valuable precursors to natural products and therapeutics. Despite the potential applications of such bioactive β -branched ncAAs, their availability is limited due to the inefficiency of the multi-step methods used to prepare them. Here we report a stereoselective biocatalytic synthesis of β -branched tryptophan analogs using an engineered variant of *Pyrococcus furiosus* tryptophan synthase (*Pf*TrpB)^{7E6}. *Pf*TrpB^{7E6} is the first biocatalyst to synthesize bulky β -branched tryptophan analogs in a single step, with demonstrated access to 27 ncAAs. The molecular basis for the efficient catalysis and broad substrate tolerance of *Pf*TrpB^{7E6} was explored through X-ray crystallography and UV-visible light spectroscopy, which revealed that a combination of active-site and remote mutations increase the abundance and persistence of a key reactive intermediate. *Pf*TrpB^{7E6} provides an operationally simple and environmentally benign platform for preparation of β -branched tryptophan building blocks.

Amino acids are nature's premier synthetic building blocks for bioactive molecules. Alongside the standard proteinogenic amino acids are diverse non-canonical amino acids (ncAAs) that are structurally similar but are not ribosomally incorporated into proteins. Due to the presence of functional groups that confer novel chemical and biological properties,^[1] ncAAs can be found in natural products and 12% of the 200 top-grossing pharmaceuticals.^[2,3] Of interest are β -branched ncAAs, which possess a chiral center at the β -position in addition to the standard chirality at the α -position of an amino acid (**Figure 1a**). The two adjacent stereocenters impose conformational constraints that affect the biochemical properties of both the amino acids themselves and the molecules they compose.^[4–7] These properties make β -branched ncAAs frequent components of useful natural products, biochemical probes, and therapeutics (**Figure 1b**).^[8–13] Despite their broad utility, most β -branched ncAAs are not readily available due to the challenge of forming two adjacent stereocenters while tolerating the reactive functional groups present in amino acids.^[14–18] For example, traditional organic synthesis of (2*S*, 3*S*)- β -methyltryptophan (β -MeTrp) requires multiple steps that incorporate protecting groups, hazardous reagents, and expensive metal catalysts.^[19,20] To take full advantage of these bioactive molecules, an improved methodology is needed to synthesize β -branched ncAAs.

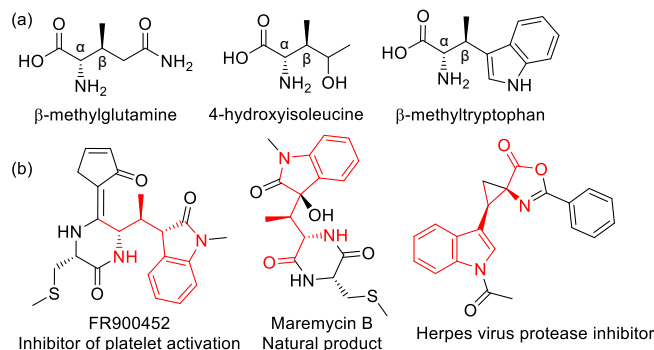


Figure 1. Representative β -branched amino acids. (a) Examples of β -branched ncAAs. (b) Examples of products derived from β -branched tryptophan analogs (red).

Enzymes offer an efficient and sustainable alternative to chemical synthesis and are routinely used to generate enantiopure amino acids from simple materials without the need for protecting groups.^[21] Although several classes of enzymes have been employed in this pursuit, those using the pyridoxal phosphate cofactor (PLP, vitamin B6) are among the most prominent.^[22] The most common biocatalytic route to an amino acid requires a fully assembled carbon skeleton and a PLP-dependent transaminase which is used to set the stereochemistry. However, as with traditional organic methodologies, the enzymatic synthesis of β -branched ncAAs is often confounded by the presence of a second stereocenter. The capacity to incorporate biocatalytic C–C bond-forming steps *en route* to diverse β -branched ncAAs would therefore be a powerful synthetic tool.

Few β -branched ncAA synthases have been reported, and even more rare are enzymes that produce branches larger than a methyl group. We previously engineered the β -subunit of the PLP-dependent enzyme tryptophan synthase from the thermophilic archaeon *Pyrococcus furiosus* (*Pf*TrpB) as a stand-alone ncAA synthase able to generate tryptophan (Trp) analogs from serine (Ser) and the corresponding substituted indole (**Figure 2a**).^[23–25] Further engineering of *Pf*TrpB for improved C–C bond formation with indole analogs and threonine (Thr) led to *Pf*TrpB^{2B9} (eight mutations from wild-type *Pf*TrpB), which exhibited a >1,000-fold improvement in (2*S*, 3*S*)- β -methyltryptophan (β -MeTrp) production relative to wild type (**Figure 2b**).^[26,27] While the reactive amino-acrylate intermediate (E(A-A)) (**Figure 3a**) readily forms with Thr, mechanistic analysis showed that competing hydrolysis of (E(A-A)) resulted in abortive deamination that consumed the amino acid substrate (**Figure 3b**),^[28,29] limiting the enzyme's yield (typically < 50%) with a single equivalent of Thr. Further, *Pf*TrpB^{2B9} accepted only Ser and Thr as substrates since larger β -alkyl substrates were unable to efficiently form E(A-A).

[a] Dr. C. E. Boville, R. A. Scheele, P. Koch, Dr. S. Brinkmann-Chen, Prof. F. H. Arnold
Division of Chemistry and Chemical Engineering 210-41, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, United States
Email: frances@cheme.caltech.edu

[b] Prof. A. R. Buller
Department of Chemistry, University of Wisconsin, 1101 University Avenue, Madison, WI 53706, United States
Email: arbuller@wisc.edu

[+] These authors contributed equally to this work.

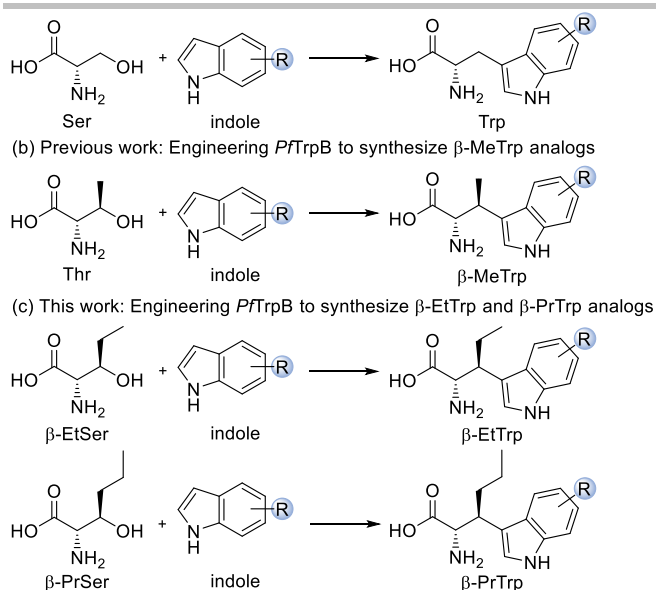


Figure 2. Synthesis of Trp and Trp analogs by *Pf*TrpB.

To surmount these challenges, we sought to identify mutations that would facilitate formation of E(A-A) with the more challenging (2*S*, 3*R*)- β -ethylserine (β -EtSer) and (2*S*, 3*R*)- β -propylserine (β -PrSer) substrates while simultaneously decreasing E(A-A) hydrolysis (**Figure 2c**). The latter is essential, as increasingly bulky alkyl chains are thought to hinder nucleophilic attack. Increased E(A-A) persistence will allow more time for the intrinsically slower addition reaction to occur while reducing the amount of starting material lost to competing hydrolysis (**Figure 3b**).

We chose *Pf*TrpB^{2B9} as our engineering starting point to increase production of β -EtTrp. While an active catalyst with Thr, *Pf*TrpB^{2B9} was sluggish with β -EtSer (80 total turnovers, TTN) and gave too little signal for high-throughput screening.^[23] We speculated that active-site mutations would promote the formation of E(A-A) with larger, sterically demanding β -substituents and used a structure-guided approach to improve activity with β -EtSer. Modeling β -EtSer into the *Pf*TrpB^{2B9} active site as E(A-A) (PDB: 5VM5)^[30] suggested a steric clash with L161 (**Figure 4a**). Hypothesizing that this constraint could be reduced by mutating L161 to a residue with a smaller side chain (**Figure 4b**), we expressed and assayed variants *Pf*TrpB^{2B9} L161V, L161A, and L161G. We found that L161V and L161A increased the TTN 14-fold and 10-fold, respectively, whereas L161G decreased activity by a factor of 2.6 (**Figure 4c**). As our long-term interest is to produce a catalyst that accommodates a wider range of β -alkyl chains, we selected *Pf*TrpB^{2B9} L161A as the parent enzyme for directed evolution, with the rationale that the smaller sidechain of alanine would minimize steric clashes with bulkier substrates.

We then introduced random mutations into the *Pf*TrpB^{2B9} L161A gene and screened for enhanced β -EtTrp synthesis (**Table 1**) at 290 nm under saturating substrate conditions.^[23] Screening made use of starting materials containing a mixture of diastereomers, however only the (2*S*,3*R*) diastereomer underwent a productive reaction. High-throughput screening of 352 variants yielded *Pf*TrpB^{0E3} (L91P), which displayed a 43-fold increase in TTN for β -EtTrp (**Figure 4d**). *Pf*TrpB^{0E3} was then used as the parent for a second round of random mutagenesis, yielding

variant *Pf*TrpB^{8C8} (V173E), which improved β -EtTrp yields by 54-fold relative to *Pf*TrpB^{2B9}. At this juncture, a third round of random mutagenesis failed to yield further improvements after screening 880 variants. Although the accumulated mutations increased activity, we speculated that further improvements were hindered by deleterious mutations that reduced enzyme stability.^[31] We therefore recombined mutations in TrpB^{8C8}, allowing a 50% chance for each residue to retain the mutation or revert to wild type. Recombination included all residues except those which were crucial for starting activity with Ser (T292S), Thr (F95L), and β -EtSer (L161A and L91P) (**Table S1**). Recombination also included F274L, which was previously identified as an activating mutation.^[23] Recombined variants were assayed for β -EtTrp production at 290 nm, which revealed that I68V and T321A were non-essential, but that F274L was beneficial, yielding variant *Pf*TrpB^{7E6}. Though *Pf*TrpB^{7E6} did not show improved stability (**Table S2**), recombination did enhance activity, with a 58-fold

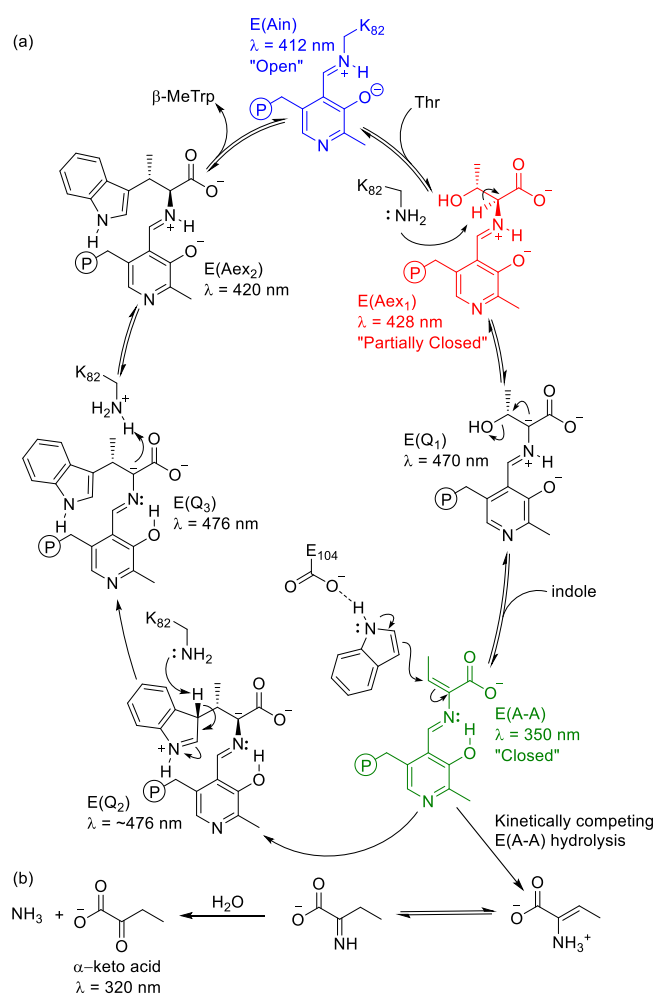


Figure 3. The putative catalytic cycle for *Pf*TrpB synthesizing β -MeTrp. (a) Catalysis initiates as E(Ain) with the mobile COMM domain predominantly in the open conformation (blue). With the addition of Thr, the COMM domain undergoes rigid body motion, transitioning to a partially closed position through E(Aex₁) (red) followed by full closure with formation of the reactive E(A-A) intermediate (green). E(A-A) is then attacked by indole and undergoes an addition reaction to form β -MeTrp. (b) E(A-A) may also undergo a kinetically competing hydrolysis reaction to generate α -keto acids, observable at 320 nm. This deamination reaction consumes an equivalent of the amino acid substrate.

improvement relative to *Pf*TrpB^{2B9} (**Figure 4d**). An additional round of recombination sampled other previously identified activating mutations (Q38R, M139L, N166D, S335N) and allowed for reversion of L91P. This process produced a variant (*Pf*TrpB^{2G8}, see **Table 1**) that lacked the L91P mutation and had only slightly lower activity than *Pf*TrpB^{7E6}. Although subsequent work showed that *Pf*TrpB^{2G8} is also a proficient enzyme (vide infra), the parent *Pf*TrpB^{7E6} was selected for mechanistic characterization as it is a comparatively simple catalyst with excellent activity and more amenable to crystallization.

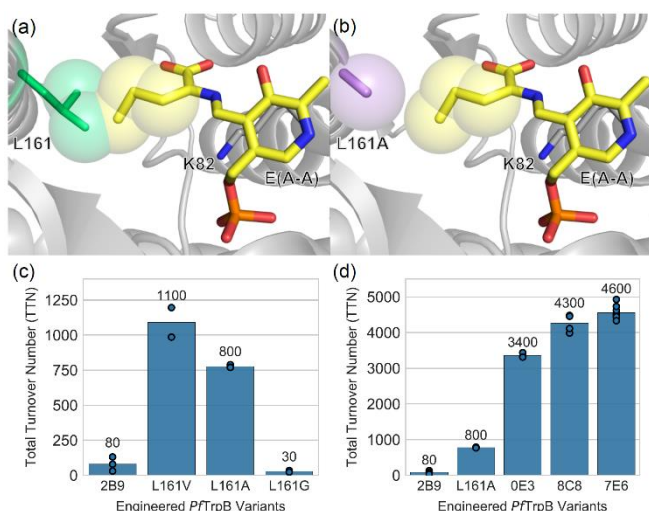


Figure 4. Engineering *Pf*TrpB for β -EtTrp synthesis. (a) β -EtSer as E(A-A) (yellow) modeled in the *Pf*TrpB^{2B9} (PDB: 5VM5, gray) active site. Spheres represent the Van der Waals radii and highlight a clash with L161 (green). (b) As in (a), but with the mutation L161A shown (purple). (c) β -EtTrp production by *Pf*TrpB^{2B9} with L161V, L161A, or L161G mutations. (d) β -EtTrp production by engineered *Pf*TrpB variants. Bars represent the average of all data points, with individual reactions shown as circles. At minimum, reactions were performed in duplicate.

We sought to identify which newly evolved properties of *Pf*TrpB enabled increased TTNs with challenging β -branched substrates. As described above, the activity and substrate scope of the parent enzyme, *Pf*TrpB^{2B9}, were limited by low steady-state population (abundance) and subsequent breakdown (persistence) of the reactive E(A-A) intermediate.^[27] To assess the abundance of E(A-A), we capitalized on the intrinsic spectroscopic properties of the PLP cofactor to visualize the steady-state distribution of intermediates throughout the catalytic cycle (**Figure 3a**).^[32] With the addition of β -EtSer to *Pf*TrpB^{7E6}, the internal aldimine peak (E(Ain), 412 nm) decreased and E(A-A) (350 nm) became the major species (**Figure 5a**). This is a notable change, as when β -EtSer was added to *Pf*TrpB^{2B9}, E(Aex₁) accumulated and no E(A-A) was observed (**Fig 5a**). To assess the persistence of E(A-A), we assayed the deamination rate and coupling efficiency of *Pf*TrpB^{7E6}. In the presence of both Thr and β -EtSer, *Pf*TrpB^{7E6} displayed up to a 4-fold decrease in the deamination reaction relative to *Pf*TrpB^{2B9} (**Table S3**). We then probed the enzyme's coupling efficiency under reaction conditions with high catalyst loading and equimolar substrate equivalents, where product formation is limited only by the consumption of starting material through the competing deamination reaction. We observed an increase in product formation from 5% with *Pf*TrpB^{2B9} to 96% with *Pf*TrpB^{7E6} when β -EtSer was the substrate (**Figure 5b**). Collectively, these data

Table 1. Engineering *Pf*TrpB through directed evolution for improved β -EtTrp production. Engineering began with *Pf*TrpB^{2B9} (*Pf*TrpB L16V, E17G, I68V, F95L, F274S, T292S, T321A, and V384A) with 80 TTN. All reactions were performed in at least duplicate with 0.1% catalyst loading for 24 hours at 75 °C.

Variant	Mutations Added	Mutations Removed	Average TTN
[a] <i>Pf</i> TrpB ^{2B9}	L161A	N/A	800
[b] <i>Pf</i> TrpB ^{0E3}	L91P	N/A	3400
[b] <i>Pf</i> TrpB ^{8C8}	V173E	N/A	4300
[c] <i>Pf</i> TrpB ^{7E6}	F274L	I68V, T321A	4600
[c] <i>Pf</i> TrpB ^{2G8}	M139L, N166D, S335N	L91P	3800

[a] Site-directed mutagenesis. [b] Random mutagenesis. [c] Recombination.

indicate that increased product formation was achieved by incorporating mutations that facilitate the formation of E(A-A) and increase its lifetime in the active site.

During directed evolution, *Pf*TrpB was altered by the introduction of nine mutations. Although *Pf*TrpB^{7E6} has only a single mutation in the active site (**Figure S1**), mutations governing enzyme activity are scattered throughout the protein.^[23,33] Remote mutations may be affecting the enzyme's conformational dynamics, which have been previously shown to be linked to the catalytic cycle of *Pf*TrpB (**Figure 3a**).^[30,33] In its resting state, *Pf*TrpB binds PLP via the catalytic lysine (K82) as E(Ain) with the mobile communication (COMM) domain in a predominantly open conformation. Addition of an amino acid substrate induces formation of the external aldimine (E(Aex₁)), which is accompanied by partial closure of the COMM domain. Dehydration to form the electrophilic E(A-A) species occurs when TrpB populates a fully closed conformation, where it remains until product is formed.^[28,29] To examine the state of the *Pf*TrpB^{7E6} active site and its connection to the COMM domain conformational cycling, we determined the X-ray crystal structures of *Pf*TrpB^{7E6} in the E(Ain) state as well as with β -EtSer bound in the active site as E(A-A).

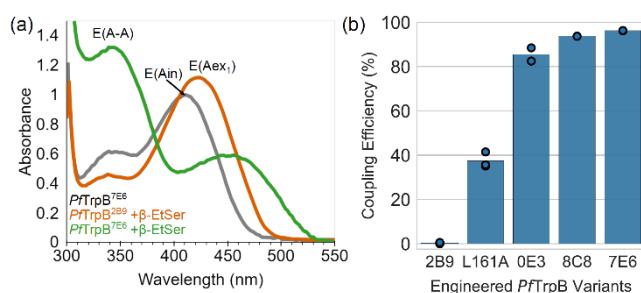


Figure 5. Directed evolution stabilizes E(A-A) and improves coupling efficiency. (a) The steady-state population of *Pf*TrpB as determined by UV-visible light spectroscopy. In the absence of substrate, the predominant population of *Pf*TrpB^{7E6} (black) is E(Ain). β -EtSer-bound *Pf*TrpB^{2B9} (orange) accumulates E(Aex₁) and *Pf*TrpB^{7E6} (green) forms E(A-A). All spectra are normalized to the absorbance value of E(Ain) at 412 nm. (b) Variant coupling efficiency with β -EtSer. Bars represent the average of all data points, with individual reactions shown as circles. At minimum, reactions were performed in duplicate.

Earlier *Pf*TrpB variants, including *Pf*TrpB^{2B9}, were nearly identical to wild-type *Pf*TrpB (PDB: 5DVZ) in the open state. Here, the 2.26-Å structure of *Pf*TrpB^{7E6} (PDB: 6CUV) shows distinct preorganization toward a more closed conformation. Specifically,

in half of the protomers, the COMM domain has shifted into a distinct partially-closed conformation that was previously associated with substrate binding (**Figure 6a**). While many residues may contribute to the stabilization of this state, we hypothesize that the mutation L91P destabilizes open states; this residue lies on an α -helix immediately prior to the COMM domain in the sequence and causes a kink in the helix that shifts the structure toward more closed states (**Figure 6b**).

We next soaked *Pf*TrpB^{7E6} with β -EtSer and obtained a 1.75-Å structure with β -EtSer bound as E(A-A) in two protomers (PDB: 6CUZ). As expected, the COMM domain underwent rigid-body motion to the closed conformation (**Figure 6a**) where the steric complementarity between the longer β -alkyl chain and L161A becomes apparent. Notably, the L161A mutation does not appear to induce significant alterations elsewhere in the active site (**Figure 6c**). When indole is modeled into the active site, there is space to accommodate even longer β -branched substituents as well as a range of indole nucleophiles (**Figure 6d**).

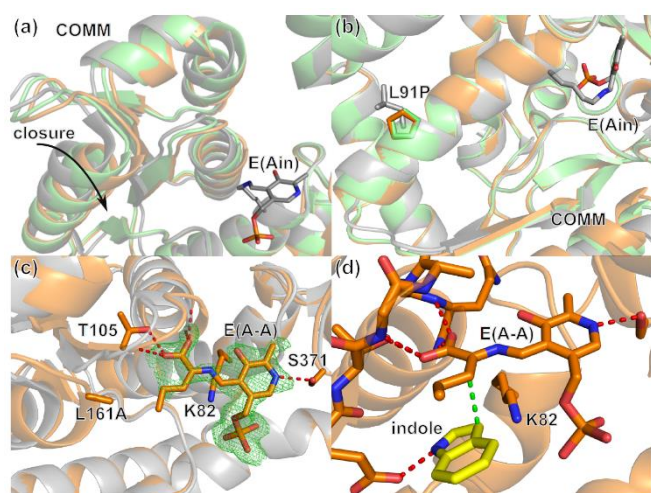


Figure 6. Substrate binding and conformational changes in *Pf*TrpB. (a) The COMM domain of *Pf*TrpB undergoes rigid body motion that is linked to the catalytic cycle. In the absence of substrate, wild-type *Pf*TrpB (PDB: 5DVZ, gray) is in the open conformation, while *Pf*TrpB^{7E6} (PDB: 6CUV, green) assumes a partially closed conformation. When β -EtSer is bound to *Pf*TrpB^{7E6} as E(A-A) (PDB: 6CUV, orange), the COMM domain undergoes a rigid body shift to a closed conformation. (b) The mutation L91P introduces a kink in the α -helix adjacent to the COMM domain. (c) β -EtSer bound to *Pf*TrpB^{7E6} as E(A-A) is shown with $F_o - F_c$ map contoured at 2.0σ (green). The gamma carbon of the amino-acrylate is not well resolved. Hydrogen bonds are shown as red dashes. (d) Indole (yellow) modeled in the active site of *Pf*TrpB^{7E6} with β -EtSer as E(A-A). The green dash represents the bond-forming atoms in indole and β -EtSer.

As our goal was to evolve a versatile β -branched ncAA synthase, we next explored the substrate scope of *Pf*TrpB^{7E6}. We hypothesized that, if improvements in activity came through increased stability of E(A-A), the same mutations should increase activity with multiple amino acid substrates. Indeed, we found that although we screened for β -EtTrp synthesis, the TTN for β -MeTrp and (2*S*, 3*S*)- β -propyltryptophan (β -PrTrp) synthesis were simultaneously improved 3.6-fold and 36-fold, respectively (**Figure 7a**). Consistent with our previous observations, directed evolution improved the enzyme's coupling efficiency (**Figure 7b**) and amino-acrylate persistence (**Figure 7c-d**) with all three acid substrates. Next, we revisited our earlier hypothesis that the L161A mutation would be more beneficial than L161V by reducing steric clashes with larger substrates. We observed that although

*Pf*TrpB^{7E6} L161V is viable for synthesis of β -MeTrp and β -EtTrp, the TTN for β -PrTrp formation was reduced 5-fold (**Figure S2a**). In addition, *Pf*TrpB^{7E6} retained the robust Trp activity that is the hallmark of the wild-type enzyme (**Figure S2b**), demonstrating that the L161A mutation was successful in accommodating bulkier substrates, allowing catalysis with four different amino acid substrates.

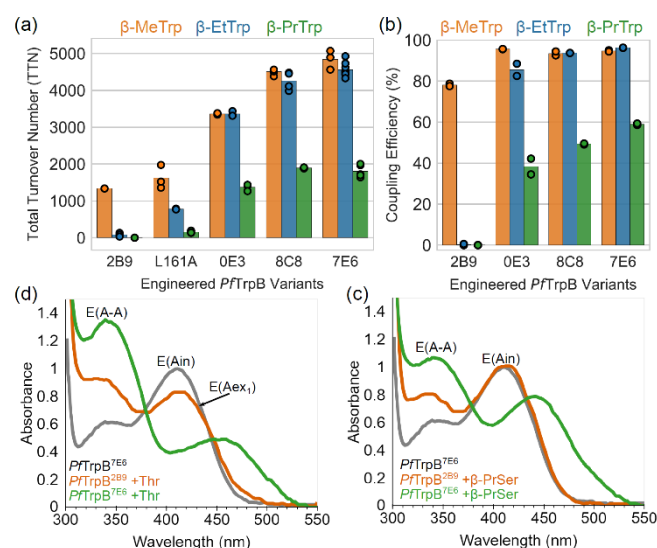
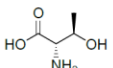













Figure 7. *Pf*TrpB engineering grants access to a range of β -branched tryptophan analogs. (a) TTN of *Pf*TrpB for β -MeTrp (orange), β -EtTrp (blue), and β -PrTrp (green). (b) Variant coupling efficiency with Thr (orange), β -EtSer (blue), and β -PrSer (green). Bars represent the average of all data points, with individual reactions shown as circles. At minimum, reactions were performed in duplicate. (c) The steady-state population of *Pf*TrpB with Thr as determined by UV-visible light spectroscopy. In the absence of substrate, the predominant population of *Pf*TrpB^{7E6} (black) is E(Ain). With the addition of Thr to *Pf*TrpB^{2B9} (orange) has a mixed population of E(Aex₁) and E(A-A), while *Pf*TrpB^{7E6} (green) is predominantly E(A-A). (d) β -PrSer-bound *Pf*TrpB^{2B9} (orange) remains as E(Ain) while *Pf*TrpB^{7E6} (green) predominantly forms E(A-A).

However, activity was not observed with all β -alkyl substrates and reactions with (2*S*)- β -isopropylserine (β -iPrSer) showed only trace activity. To understand why catalysis did not proceed with this bulkier sidechain, we soaked β -iPrSer into *Pf*TrpB^{7E6} crystals and obtained a 1.77-Å structure (PDB: 6CUT), which shows the catalytically unreactive (2*S*, 3*S*) diastereomer of β -iPrSer bound as E(Aex₁) (**Figure S3**). Though (2*S*, 3*S*)- β -iPrSer can form E(Aex₁), dehydration across the C α -C β bond requires a rotameric shift of the side chain that we hypothesize is hindered by steric interactions with an adjacent loop.^[34] Further work is needed to understand whether the poor activity of *Pf*TrpB^{7E6} with (2*S*, 3*R*)- β -iPrSer reflects inhibition by an isomeric analog, increased allylic strain of the amino-acrylate, or a combination of effects.

In addition to acting on multiple amino acid substrates, we hypothesized that *Pf*TrpB^{7E6} would retain the wild-type enzyme's breadth of reactivity with indole analogs.^[23–25] We performed analytical biotransformations with 11 representative nucleophiles with three β -branched amino acid substrates, yielding 27 tryptophan analogs, 20 of which are previously unreported (**Table 2**). Each reaction was analyzed by liquid-chromatography/mass spectrometry (LCMS) and TTN were calculated by comparing product and substrate absorption at the isosbestic wavelength (**Table S4**). Happily, we found that substituted indole analogs

Table 2. β -Branched tryptophan analogs synthesized by *PfTrpB*^{7E6}. Average TTN (10,000 max TTN) are indicated for each combination of amino acid and nucleophilic substrate. At minimum, reactions were performed in duplicate. See Supplemental Information for experimental details.

		Nucleophilic Substrate										
		Indole	2-methylindole	4-methylindole	4-fluoroindole	5-methylindole	5-fluoroindole	5-chloroindole	6-methylindole	7-methylindole	7-azaindole	Indazole
Electrophilic Substrate												
	Thr	4800	4500	1000	4500	800	5800	<100	1400	3000	4800	200
	β -EtSer	4600	3000	700	1600	100	3700	N.D.	700	2800	N.D.	N.D.
β -PrSer	1800	200	100	200	<100	400	N.D.	100	1100	N.D.	N.D.	

remained well-tolerated by *PfTrpB*^{7E6}. Methyl substituents were accepted around the indole ring, though the enzyme demonstrated higher activity with fluoroindoles. We also observed activity with 5-chloroindole and Thr, a reaction that was undetectable for the parent enzyme, *PfTrpB*^{2B9}. In addition, we have abolished the undesirable *N*-alkylation reaction that occurred with *PfTrpB*^{2B9} in the presence of 7-azaindole and 4-fluoroindole.^[27] However, yields with *N*-nucleophilic substrates such as indazole remained low with β -branched substrates relative to their Ser counterparts. Importantly, *PfTrpB*^{7E6} can synthesize these ncAAs using only a single equivalent of the amino acid substrate, whereas *PfTrpB*^{2B9} had required 10 equivalents. This is a testament to the value of improving the stability of the reactive E(A-A) intermediate in the reaction.

All product identities were confirmed by ¹H- and ¹³C-NMR as well as high-resolution mass spectrometry from 100- μ mol preparative reactions using two equivalents of electrophilic substrate. Reactions were conducted at 0.01 to 0.4 mol% catalyst loading, and we found that, under these conditions, *PfTrpB*^{7E6} maintained robust activity: β -MeTrp with 6,600 TTN (88% yield), β -EtTrp with 6,200 TTN (82% yield), and β -PrTrp with 2,100 TTN (84% yield). We also used the recombination variant *PfTrpB*^{2G8} (Table 1) to synthesize and characterize 27 tryptophan analogs on a preparative scale (Table S5). For future applications, reactions may be further optimized by tuning catalyst loading and increasing substrate equivalents (Table S6). In conjunction with the high expression levels of *PfTrpB*^{7E6} (~300 mg enzyme per L culture), a range of β -branched ncAAs are now accessible on a preparative scale. We have developed a new biocatalytic route to (2S, 3S)-tryptophan analogs using the engineered thermostable catalyst, *PfTrpB*^{7E6}. Through directed evolution, we increased the abundance and persistence of the key E(A-A) intermediate by the introduction of active-site and remote mutations. In turn, *PfTrpB*^{7E6} displays improved coupling efficiency with multiple β -branched amino acid substrates.

This work significantly extends previous efforts to engineer *PfTrpB* enzymes, which have proven to be versatile and efficient catalysts for production of tryptophan analogs.

Acknowledgements

The authors thank Dr. David Romney, Patrick Almhjell, Dr. Christopher Prier, Nathaniel Goldberg, and Ella Watkins for their helpful discussions and comments on the manuscript. We thank Dr. Jens Kaiser of the Caltech Molecular Observatory, Dr. Scott Virgil and the Center for Catalysis and Chemical Synthesis, and Dr. Mona Shahgholi and Naseem Torian from the Caltech Mass Spectrometry Laboratory. This work was funded by the Jacobs Institute for Molecular Medicine (Caltech) and the Rothenberg Innovation Initiative. C.E.B. was supported by a postdoctoral fellowship from the Resnick Sustainability Institute. R.A.S. was supported by funding from the University of Groningen.

Keywords: biocatalysis • tryptophan synthase • non-canonical amino acid • β -branched amino acid • directed evolution

- [1] F. Agostini, J.-S. Völler, B. Koksche, C. G. Acevedo-Rocha, V. Kubyshekin, N. Budisa, *Angew. Chem. Int. Ed.* **2017**, *56*, 9680–9703.
- [2] N. A. McGrath, M. Brichacek, J. T. Njardarson, *J. Chem. Educ.* **2010**, *87*, 1348–1349.
- [3] M. A. T. Blaskovich, *Journal of Medicinal Chemistry* **2016**, *59*, 10807–10836.
- [4] S. E. Gibson (née Thomas), N. Guillo, M. J. Tozer, *Tetrahedron* **1999**, *55*, 585–615.
- [5] V. J. Hruby, *J. Med. Chem.* **2003**, *46*, 4215–4231.
- [6] V. W. Cornish, M. I. Kaplan, D. L. Veenstra, P. A. Kollman, P. G. Schultz, *Biochemistry* **1994**, *33*, 12022–12031.
- [7] C. Haskell-Luevano, L. W. Boteju, H. Miwa, C. Dickinson, I. Gantz, T. Yamada, M. E. Hadley, V. J. Hruby, *J. Med. Chem.* **1995**, *38*, 4720–4729.
- [8] C. Milne, A. Powell, J. Jim, M. Al Nakeeb, C. P. Smith, J. Micklefield, *J. Am. Chem. Soc.* **2006**, *128*, 11250–11259.
- [9] J. C. Sheehan, D. Mania, S. Nakamura, J. A. Stock, K. Maeda, *J. Am. Chem. Soc.* **1968**, *90*, 462–470.
- [10] Y. Zou, Q. Fang, H. Yin, Z. Liang, D. Kong, L. Bai, Z. Deng, S. Lin, *Angew. Chem. Int. Ed.* **2013**, *52*, 12951–12955.
- [11] J. Michaux, G. Niel, J.-M. Campagne, *Chem. Soc. Rev.* **2009**, *38*, 2093.
- [12] S. Takase, N. Shigematsu, I. Shima, I. Uchida, M. Hashimoto, T. Tada, S. Koda, Y. Morimoto, *J. Org. Chem.* **1987**, *52*, 3485–3487.
- [13] I. L. Pinto, A. West, C. M. Debouck, A. G. DiLella, J. G. Gorniack, K. C. O'Donnell, D. J. O'Shannessy, A. Patel, R. L. Jarvest, *Bioorganic Med. Chem. Lett.* **1996**, *6*, 2467–2472.
- [14] S. G. Davies, A. M. Fletcher, A. B. Frost, J. A. Lee, P. M. Roberts, J. E. Thomson, *Tetrahedron* **2013**, *69*, 8885–8898.
- [15] S.-Y. Zhang, Q. Li, G. He, W. A. Nack, G. Chen, *J. Am. Chem. Soc.* **2013**, *135*, 12135–12141.
- [16] M. J. O'Donnell, J. T. Cooper, M. M. Mader, *J. Am. Chem. Soc.* **2003**, *125*, 2370–2371.
- [17] C. Xiong, W. Wang, C. Cai, V. J. Hruby, *J. Org. Chem.* **2002**, *67*, 1399–1402.
- [18] S. Lou, G. M. McKenna, S. A. Tymonko, A. Ramirez, T. Benkovics, D. A. Conlon, F. González-Bobes, *Org. Lett.* **2015**, *17*, 5000–5003.
- [19] Y. Sawai, M. Mizuno, T. Ito, J. Kawakami, M. Yamano, *Tetrahedron* **2009**, *65*, 7122–7128.

-
- [20] L. Jeannin, M. Boisbrun, C. Nemes, F. Cochard, M. Laronze, E. Dardennes, Á. Kovács-Kulyassa, J. Sapi, J.-Y. Laronze, *Comptes Rendus Chim.* **2003**, *6*, 517–528.
- [21] Y.-P. Xue, C.-H. Cao, Y.-G. Zheng, *Chem. Soc. Rev.* **2018**, *47*, 1516–1561.
- [22] M. L. Di Salvo, N. Budisa, R. Contestabile, *Beilstein Bozen Symp. Mol. Eng. Control* **2012**, 27–66.
- [23] A. R. Buller, S. Brinkmann-Chen, D. K. Romney, M. Herger, J. Murciano-Calles, F. H. Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 14599–14604.
- [24] D. K. Romney, J. Murciano-Calles, J. E. Wehrmüller, F. H. Arnold, *J. Am. Chem. Soc.* **2017**, *139*, 10769–10776.
- [25] J. Murciano-Calles, D. K. Romney, S. Brinkmann-Chen, A. R. Buller, F. H. Arnold, *Angew. Chem. Int. Ed.* **2016**, *55*, 11577–11581.
- [26] A. R. Buller, P. van Roye, J. Murciano-Calles, F. H. Arnold, *Biochemistry* **2016**, *55*, 7043–7046.
- [27] M. Herger, P. van Roye, D. K. Romney, S. Brinkmann-Chen, A. R. Buller, F. H. Arnold, *J. Am. Chem. Soc.* **2016**, *138*, 8388–8391.
- [28] B. G. Caulkins, B. Bastin, C. Yang, T. J. Neubauer, R. P. Young, E. Hilario, Y. M. Huang, C. A. Chang, L. Fan, M. F. Dunn, et al., *J. Am. Chem. Soc.* **2014**, *136*, 12824–12827.
- [29] B. G. Caulkins, R. P. Young, R. A. Kudla, C. Yang, T. J. Bittbauer, B. Bastin, E. Hilario, L. Fan, M. J. Marsella, M. F. Dunn, et al., *J. Am. Chem. Soc.* **2016**, *138*, 15214–15226.
- [30] A. R. Buller, P. van Roye, J. K. B. Cahn, R. A. Scheele, M. Herger, F. H. Arnold, *J. Am. Chem. Soc.* **2018**, *140*, 7256–7266.
- [31] J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 5869–5874.
- [32] W. F. Drewe, M. F. Dunn, *Biochemistry* **1986**, *25*, 2494–2501.
- [33] M. F. Dunn, *Arch. Biochem. Biophys.* **2012**, *519*, 154–166.
- [34] E. W. Miles, D. R. Houck, H. G. Floss, *J. Biol. Chem.* **1982**, *257*, 14203–14210.
-

(b) Manuscript ChemRxiv.pdf (1.01 MiB)

[view on ChemRxiv](#) • [download file](#)

Engineered biosynthesis of β -alkyl tryptophan analogs

Christina E. Boville⁺, Remkes A. Scheele⁺, Philipp Koch, Sabine Brinkmann-Chen, and Andrew R. Buller^{*}, Frances H. Arnold^{*}

Abstract: Non-canonical amino acids (ncAAs) with dual stereocenters at the α and β positions are valuable precursors to natural products and therapeutics. Despite the potential applications of such bioactive β -branched ncAAs, their availability is limited due to the inefficiency of the multi-step methods used to prepare them. Here we report a stereoselective biocatalytic synthesis of β -branched tryptophan analogs using an engineered variant of *Pyrococcus furiosus* tryptophan synthase (*Pf*TrpB), *Pf*TrpB^{7E6}. *Pf*TrpB^{7E6} is the first biocatalyst to synthesize bulky β -branched tryptophan analogs in a single step, with demonstrated access to 27 ncAAs. The molecular basis for the efficient catalysis and broad substrate tolerance of *Pf*TrpB^{7E6} was explored through X-ray crystallography and UV-visible light spectroscopy, which revealed that a combination of active-site and remote mutations increase the abundance and persistence of a key reactive intermediate. *Pf*TrpB^{7E6} provides an operationally simple and environmentally benign platform for preparation of β -branched tryptophan building blocks.

Table of Contents

Results and Discussion	2
Table S1. Summary of the mutations (green) included in the recombination library that led to <i>PfTrpB</i> ^{7E6}	2
Table S2. Thermostability of evolved <i>PfTrpB</i> variants.....	2
Table S3. Enzymatic formation of α -keto acids.....	2
Table S4. Isosbestic point of tryptophan analogs and the corresponding indole analogs.....	3
Table S5. Tryptophan analogs synthesized in preparative reactions.....	4
Table S6. Optimization of reaction yields.....	5
Figure S1. Locations of mutations in <i>PfTrpB</i> ^{7E6}	6
Figure S2. Characterization of <i>PfTrpB</i> ^{7E6}	6
Figure S3. Substrate binding and conformational changes in <i>PfTrpB</i> with iPrSer.....	7
Experimental Procedures	7
Table S7. Primers for random mutagenesis.....	8
Table S8. Summary of the mutations that were subjected to recombination and their variant-of-origin.....	8
Table S9. Primers for cloning recombination libraries.....	8
Table S10. Primers for site-directed and site-saturation mutagenesis.....	8
Table S11. Crystallographic data collection and refinement statistics.....	14
Product Characterization	15

Results and Discussion

Table S1. Summary of the mutations (green) included in the recombination library that led to *PfTrpB*^{7E6}.

Substrate	Serine					Threonine			β -EtSer		
Variant	<i>PfTrpB</i> ^{2G9}	<i>PfTrpB</i> ^{4D11}				<i>PfTrpB</i> ^{4G1}	<i>PfTrpB</i> ^{2B9}		<i>PfTrpB</i> ^{2B9} L161A	<i>PfTrpB</i> ^{0E3}	<i>PfTrpB</i> ^{8C8}
Mutations	T292S	E17G	I68V	F274S	T321A	F95L	I16V	V384A	L161A	L91P	V173E

Table S2. Thermostability of evolved *PfTrpB* variants. Thermostability is reported as the temperature at which half the activity is lost (T_{50}) after 1-hour incubation.

<i>PfTrpB</i> Variant	T_{50} (°C)
<i>PfTrpB</i> ^{2B9}	95.0 \pm 0.2
<i>PfTrpB</i> ^{2B9} L161A	81.3 \pm 0.7
<i>PfTrpB</i> ^{0E3}	86.0 \pm 0.1
<i>PfTrpB</i> ^{8C8}	89.3 \pm 0.8
<i>PfTrpB</i> ^{7E6}	86.6 \pm 0.1

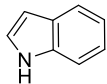
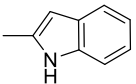
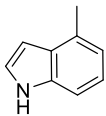
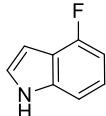
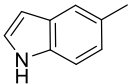
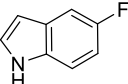
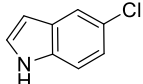
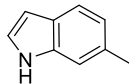
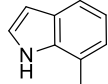
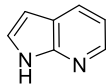
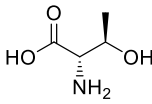
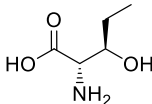
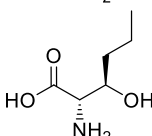
Table S3. Enzymatic formation of α -keto acids. Change in absorption at 320 nm was monitored for 10 min. Deamination rate in units of sec^{-1} is calculated using the extinction coefficient of α -ketobutyrate ($20 \text{ M}^{-1} \text{ cm}^{-1}$). N.D. not detected, E(A-A) was not observed under these reaction conditions.

Enzyme	Substrate deamination (sec^{-1})		
	Thr	β -EtSer	β -PrSer
<i>PfTrpB</i> ^{2B9}	0.4	N.D.	N.D.
<i>PfTrpB</i> ^{8C8}	0.1	0.2	0.2
<i>PfTrpB</i> ^{7E6}	0.1	0.2	0.1

Table S4. Isosbestic point of tryptophan analogs and the corresponding indole analogs.

Nucleophile substrate	Isosbestic point (nm)
Indole	277
2-methylindole	279
4-methylindole	279
4-fluoroindole	267
5-methylindole	280
5-fluoroindole	282
5-chloroindole	260
6-methylindole	273
7-methylindole	272
Indazole	276
7-azaindole	292

Table S5. Tryptophan analogs synthesized in preparative reactions. Preparative reactions were performed with 100 μ mol of nucleophile and two equivalents of electrophile with varied loading of *Pf*TrpB^{2G8}. TTN are reported with yields in parenthesis

		Nucleophilic Substrate									
		Indole	2-methylindole	4-methylindole	4-fluoroindole	5-methylindole	5-fluoroindole	5-chloroindole	6-methylindole	7-methylindole	7-azaindole
											
Electrophilic Substrate	Thr 	5400 (72%) ^a	3700 (92%) ^c	1600 (47%) ^c	3600 (87%) ^c	1800 (45%) ^c	3200 (91%) ^c	100 (20%) ^h	1200 (78%) ^f	3200 (63%) ^b	3900 (77%) ^b
	β -EtSer 	5300 (88%) ^b	2800 (94%) ^c	600 (30%) ^e	1800 (89%) ^e	100 (20%) ^h	2900 (97%) ^c	-	500 (35%) ^f	1900 (97%) ^e	-
	β -PrSer 	1900 (77%) ^d	200 (21%) ^g	100 (23%) ^h	200 (39%) ^h	20 (7%) ⁱ	400 (44%) ^g	-	20 (10%) ⁱ	1100 (56%) ^e	-

Catalyst loading (%): ^a0.01%; ^b0.02%; ^c0.03%; ^d0.04%; ^e0.05%; ^f0.07%; ^g0.1%; ^h0.2%; ⁱ0.4%; - = not tested

Table S6. Optimization of reaction yields. Reaction yields can be improved by increasing the equivalents of electrophilic substrate or increasing catalyst loading. LCMS reactions with PfTrpB2B9 and PfTrpB7E6 were conducted with 20 mM indole, 1 or 10 equivalents of electrophilic substrate, and varied catalyst loading (0.01%–0.1%). Reactions were incubated for 24 h at 75 °C and analyzed by LCMS.

Enzyme	Catalyst Loading (%)	Product	Electrophilic Substrate Equivalents	HPLC yield (%)
<i>PfTrpB^{2B9}</i>	0.01	β -MeTrp	1	13
	0.01	β -MeTrp	10	24
<i>PfTrpB^{7E6}</i>	0.01	β -MeTrp	1	48
	0.01	β -MeTrp	10	97
	0.05	β -MeTrp	1	95
	0.1	β -MeTrp	1	95
	0.01	β -EtTrp	1	46
	0.01	β -EtTrp	10	62
	0.05	β -EtTrp	1	91
	0.1	β -EtTrp	1	96
	0.01	β -PrTrp	1	18
	0.01	β -PrTrp	10	14
	0.05	β -PrTrp	1	52
	0.1	β -PrTrp	1	59

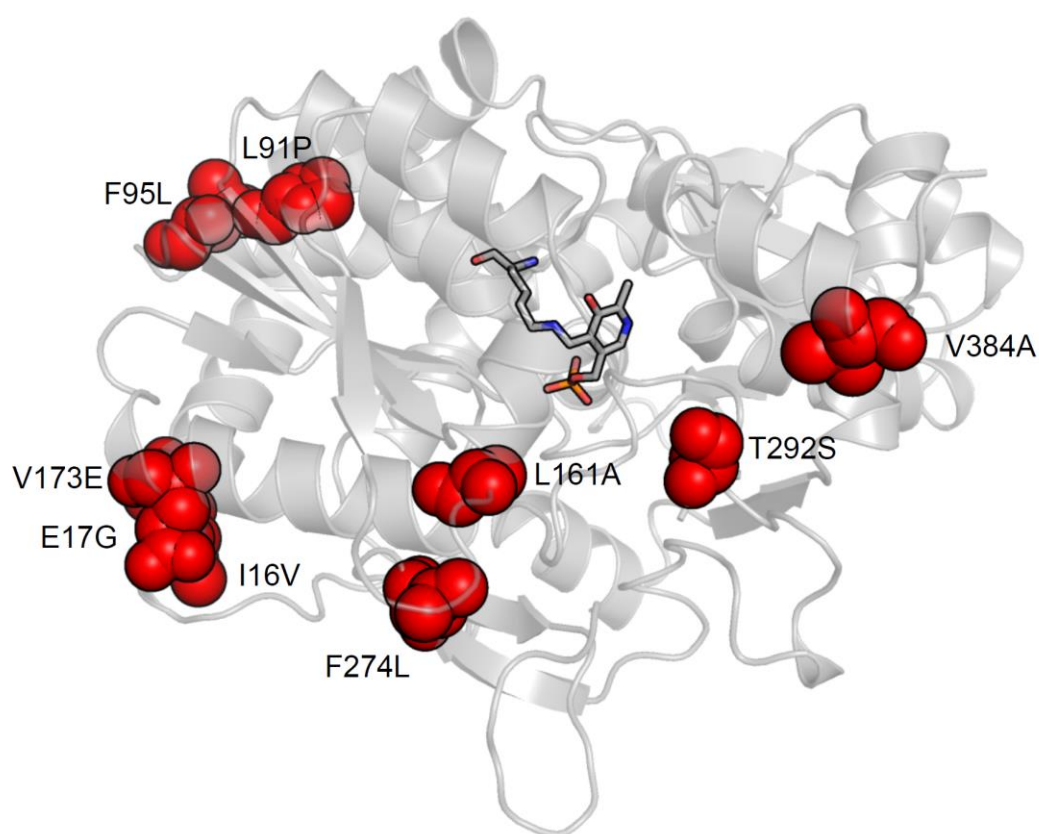


Figure S1. Locations of mutations in *PfTrpB*^{7E6}. *PfTrpB*^{7E6} contains nine mutations relative to wild-type *PfTrpB*. Mutations are distributed throughout the enzyme and are indicated in red. L161A is an active site mutation.

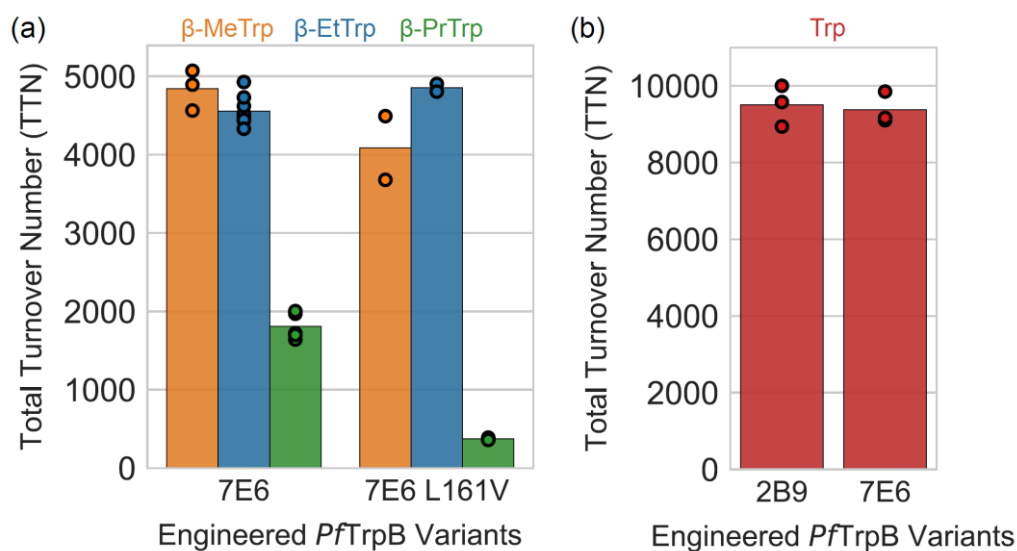


Figure S2. Characterization of *PfTrpB*^{7E6}. (a) Comparison of TTN values with *PfTrpB*^{7E6} and *PfTrpB*^{7E6} L161V. (b) Comparison of *PfTrpB*^{7E6} TTN with serine and indole. Bars represent the average of all data points, with individual reactions shown as circles. At minimum, reactions were performed in duplicate.

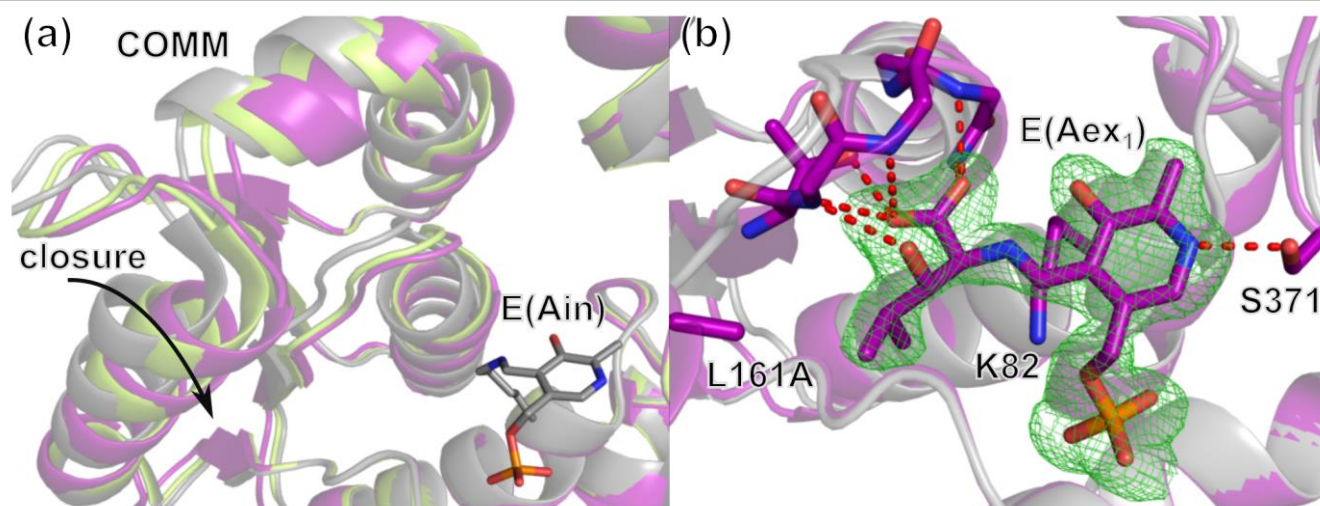


Figure S3. Substrate binding and conformational changes in *PflTrpB* with iPrSer. (a) (2S, 3S)-β-iPrSer is bound to *PflTrpB*^{7E6} as E(Ain) (PDB: 6CUT, purple). The *PflTrpB*^{7E6} COMM domain assumes a more closed conformation when compared to wild-type *PflTrpB* without substrate (PDB: 5DVZ, gray) or with Ser as E(Aex₁) (PDB 5DW0, lime). (b) iPrSer shown with F_o-F_c map contoured at 2.5σ (green). Hydrogen bonds are shown as red dashes.

Experimental Procedures

General experimental methods. Chemicals and reagents were purchased from commercial sources and used without further purification. Proton and carbon NMR spectra were recorded on a Bruker 400 MHz (100 MHz) spectrometer equipped with a cryogenic probe. Proton chemical shifts are reported in ppm (δ) relative to tetramethylsilane and calibrated using the residual solvent resonance (DMSO, δ 2.50 ppm). Data are reported as follows: chemical shift (multiplicity [singlet (s), doublet (d), doublet of doublets (dd), doublet of doublets of doublets (ddd), triplet (t), triplet of doubles (td), multiplet (m)], coupling constants [Hz], integration). Carbon NMR spectra were recorded with complete proton decoupling. Carbon chemical shifts are reported in ppm relative to tetramethylsilane and calibrated using the residual solvent proton resonance as an absolute reference. All NMR spectra were recorded at ambient temperature (about 25 °C). Preparative reversed-phase chromatography was performed on a Biotage Isolera One purification system, using C-18 silica as the stationary phase, with CH₃OH as the strong solvent and H₂O (0.1% HCl by weight) as the weak solvent. Liquid chromatography/mass spectrometry (LCMS) was performed on an Agilent 1290 UPLC-LCMS equipped with a C-18 silica column (1.8 μ m, 2.1 × 50 mm) using CH₃CN/H₂O (0.1% acetic acid by volume): 5% to 95% CH₃CN over 4 min; 1 mL/min.

Cloning. *PflTrpB*^{WT} (UNIPROT ID Q8U093) was previously codon optimized for expression in *Escherichia coli*, and cloned into pET-22b(+) with a C-terminal 6x His tag.^[1] Parent variant *PflTrpB*^{2B9} (E17G, I68V, T292S, F274S, T321A, F95L, I16V, V384A) was cloned and expressed as described previously.^[2]

Substrate modeling in the active site. β-EtSer was modeled as E(A-A) in the crystal structure of *PflTrpB*^{2B9} (PDB: 5VM5) by following the coordinates of the E(A-A) formed from Ser and following the known *E*- stereochemistry of the intermediate. Each staggered rotamer for the methyl group was sampled manually in Coot. The single rotamer lacking a steric clash was selected and the entire active site subject to 10 cycles of structure idealization in Refmac5.

Construction of random mutagenesis libraries. Random mutagenesis libraries were generated with the appropriate *PflTrpB* gene as template by the addition of 200–400 μ M MnCl₂ to a *Taq* PCR reaction as reported previously.^[1] PCR fragments were treated with DpnI for 2 h at 37 °C, purified by gel extraction, and then inserted into a pET-22b(+) vector via Gibson assembly.^[3] BL21(DE3) E. cloni[®] Express cells were transformed with the Gibson assembly product.

CTGCGTGATTGGGTGGCTACTTTTGAATACACCCACTACCTAATCGGTTCCGTGGTCCGATCCGTATCCGACCATCGTTCTGTATTTTCAGTCTGTTATCGGTCGTGAGGCTAAAGCGCAGATCCTGGAGGCTGAAGGTCAGCTGCCAGATGTAATCGTTGCTTGTGTTGGTGGTGGCTCTAACGCGATGGGTATCTTTTACCCGTTTCGTGAACGACAAAAAAGTTAAGCTGGTTGGCGTTGAGGCTGTGGTAAAGGCCTGGAATCTGGTAAGCATTCCGCTAGCCTGAACGCAGGTGAGGTTGGTGTGTCCCATGGCATGCTGTCCTACTTTCTGCAGGACGAAGAAGGTCAGATCAAACCAAGCCACTCCATCGCACCAGGTCTGGATTATCCAGGTGTTGGTCCAGAACACGCTTACCTGAAAAAAATTCAGCGTGCTGAATACGTGGCTGTAACCGATGAAGAAGCACTGAAAGCGTTCCATGAACTGAGCCGTACCGAAGGTATCATCCAGCTCTGGAATCTGCGCATGCTGTGGCTTACGCTATGAACTGGCTAAGGAAATGTCTCGTGATGAGATCATCATCGTAAACCTGTCTGGTCGTGGTGACAAAGACCTGGATATTGTCCTGAAAGCGTCTGGCAACGTGCTCGAGCACCACCAACACCACCACTGA

PfTrpB^{2B9} L161G

ATGTGGTTCGGTGAATTTGGTGGTCAGTACGTGCCAGAAACGCTGGTTGGACCCCTGAAAGAGCTGGAAAAAGCTTACAAACGTTCAAAGATGACGAAGAATTCAATCGTCAGCTGAATTACTACCTGAAAACCTGGGCAGGTCTGCCAACCCCACTGTACTACGCAAACGCCTGACTGAAAAAATCGGTGGTGCTAAAGTCTACCTGAAACGTGAAGACCTGGTTCACGGTGGTGCACACAAGACCAACAACGCCATCGGTGAGGCACTGCTGGCAAAGCTCATGGGTAAACTCGTCTGATCGCTGAGACCGGTGCTGGTCAGCACGGCGTAGCGACTGCAATGGCTGGTGCCTGCTGGGCATGAAAGTGGACATTTACATGGGTGCTGAGGACGTAGAACGTCAGAAAAATGAACGTATTCGGTATGAAGCTGCTGGGTGCAAACGTAATTCAGTTAACTCCGTTCTCGCACCGGGAAAGACGCAATCAACGAGGCTGTGCGTGATTGGGTGGCTACTTTTGAATACACCCACTACCTAATCGGTTCCGTGGTCCGATCCGTATCCGACCATCGTTCTGTGATTTTCAGTCTGTTATCGGTCGTGAGGCTAAAGCGCAGATCCTGGAGGCTGAAGGTCAGCTGCCAGATGTAATCGTTGCTTGTGTTGGTGGTGGCTCTAACGCGATGGGTATCTTTTACCCGTTTCGTGAACGACAAAAAAGTTAAGCTGGTTGGCGTTGAGGCTGTGGTAAAGGCCTGGAATCTGGTAAGCATTCCGCTAGCCTGAACGCAGGTGAGGTTGGTGTGTCCCATGGCATGCTGTCCTACTTTCTGCAGGACGAAGAAGGTCAGATCAAACCAAGCCACTCCATCGCACCAGGTCTGGATTATCCAGGTGTTGGTCCAGAACACGCTTACCTGAAAAAAATTCAGCGTGCTGAATACGTGGCTGTAACCGATGAAGAAGCACTGAAAGCGTTCCATGAACTGAGCCGTACCGAAGGTATCATCCAGCTCTGGAATCTGCGCATGCTGTGGCTTACGCTATGAACTGGCTAAGGAAATGTCTCGTGATGAGATCATCATCGTAAACCTGTCTGGTCGTGGTGACAAAGACCTGGATATTGTCCTGAAAGCGTCTGGCAACGTGCTCGAGCACCACCAACACCACCACTGA

PfTrpB^{2B9} L161V

ATGTGGTTCGGTGAATTTGGTGGTCAGTACGTGCCAGAAACGCTGGTTGGACCCCTGAAAGAGCTGGAAAAAGCTTACAAACGTTCAAAGATGACGAAGAATTCAATCGTCAGCTGAATTACTACCTGAAAACCTGGGCAGGTCTGCCAACCCCACTGTACTACGCAAACGCCTGACTGAAAAAATCGGTGGTGCTAAAGTCTACCTGAAACGTGAAGACCTGGTTCACGGTGGTGCACACAAGACCAACAACGCCATCGGTGAGGCACTGCTGGCAAAGCTCATGGGTAAACTCGTCTGATCGCTGAGACCGGTGCTGGTCAGCACGGCGTAGCGACTGCAATGGCTGGTGCCTGCTGGGCATGAAAGTGGACATTTACATGGGTGCTGAGGACGTAGAACGTCAGAAAAATGAACGTATTCGGTATGAAGCTGCTGGGTGCAAACGTAATTCAGTTAACTCCGTTCTCGCACCGGTGAAAGACGCAATCAACGAGGCTCTGCGTGATTGGGTGGCTACTTTTGAATACACCCACTACCTAATCGGTTCCGTGGTCCGATCCGTATCCGACCATCGTTCTGTGATTTTCAGTCTGTTATCGGTCGTGAGGCTAAAGCGCAGATCCTGGAGGCTGAAGGTCAGCTGCCAGATGTAATCGTTGCTTGTGTTGGTGGTGGCTCTAACGCGATGGGTATCTTTTACCCGTTTCGTGAACGACAAAAAAGTTAAGCTGGTTGGCGTTGAGGCTGTGGTAAAGGCCTGGAATCTGGTAAGCATTCCGCTAGCCTGAACGCAGGTGAGGTTGGTGTGTCCCATGGCATGCTGTCCTACTTTCTGCAGGACGAAGAAGGTCAGATCAAACCAAGCCACTCCATCGCACCAGGTCTGGATTATCCAGGTGTTGGTCCAGAACACGCTTACCTGAAAAAAATTCAGCGTGCTGAATACGTGGCTGTAACCGATGAAGAAGCACTGAAAGCGTTCCATGAACTGAGCCGTACCGAAGGTATCATCCAGCTCTGGAATCTGCGCATGCTGTGGCTTACGCTATGAACTGGCTAAGGAAATGTCTCGTGATGAGATCATCATCGTAAACCTGTCTGGTCGTGGTGACAAAGACCTGGATATTGTCCTGAAAGCGTCTGGCAACGTGCTCGAGCACCACCAACACCACCACTGA

PfTrpB^{0E3}

ATGTGGTTCGGTGAATTTGGTGGTCAGTACGTGCCAGAAACGCTGGTTGGACCCCTGAAAGAGCTGGAAAAAGCTTACAAACGTTCAAAGATGACGAAGAATTCAATCGTCAGCTGAATTACTACCTGAAAACCTGGGCAGGTCTGCCAACCCCACTGTACTACGCAAACGCCTGACTGAAAAAATCGGTGGTGCTAAAGTCTACCTGAAACGTGAAGACCTGGTTCACGGTGGTGCACACAAGACCAACAACGCCATCGGTGAGGCACTGCTGGCAAAGCTCATGGGTAAACTCGTCTGATCGCTGAGACCGGTGCTGGTCAGCACGGCGTAGCGACTGCAATGGCTGGTGCCTGCTGGGCATGAAAGTGGACATTTACATGGGTGCTGAGGACGTAGAACGTCAGAAAAATGAACGTATTCGGTATGAAGCTGCTGGGTGCAAACGTAATTCAGTTAACTCCGTTCTCGCACCGCGAAAGACGCAATCAACGAGGCTCTGCGTGATTGGGTGGCTACTTTTGAATACACCCACTACCTAATCGGTTCCGTGGTCCGATCCGTATCCGACCATCGTTCTGTGATTTTCAGTCTGTTATCGGTCGTGAGGCTAAAGCGCAGATCCTGGAGGCTGAAGGTCAGCTGCCAGATGTAATCGTTGCTTGTGTTGGTGGTGGCTCTAACGCGATGGGTATCTTTTACCCGTTTCGTGAACGACAAAAAAGTTAAGCTGGTTGGCGTTGAGGCTGTGGTAAAGGCCTGGAATCTGGTAAGCATTCCGCTAGCCTGAACGCAGGTGAGGTTGGTGTGTCCCATGGCATGCTGTCCTACTTTCTGCAGGACGAAGAAGGTCAGATCAAACCAAGCCACTCCATCGCACCAGGTCTGGATTATCCAGGTGTTGGTCCAGAACACGCTTACCTGAAAAAAATTCAGCGTGCTGAATACGTGGCTGTAACCGATGAAGAAGCACTGAAAGCGTTCCATGAACTGAGCCGTACCGAAGGTATCATCCAGCTCTGGAATCTGCGCATGCTGTGGCTTACGCTATGAACTGGCTAAGGAAATGTCTCGTGATGAGATCATCATCGTAAACCTGTCTGGTCGTGGTGACAAAGACCTGGATATTGTCCTGAAAGCGTCTGGCAACGTGCTCGAGCACCACCAACCACCACCACTGA

PfTrpB^{8C8}

ATGTGGTTCGGTGAATTTGGTGGTCAGTACGTGCCAGAAACGCTGGTTGGACCCCTGAAAGAGCTGGAAAAAGCTTACAAACGTTCAAAGATGACGAAGAGTTCAATCGTCAGCTGAATTACTACCTGAAAACCTGGGCAGGTCTGCCAACCCCACTGTACTACGCAAACGCCTGACTGAAAAAATCGGTGGTGCTAAAGTCTACCTGAAACGTGAAGACCTGGTTCACGGTGGTGCACACAAGACCAACAACGCCATCGGTGAGGCACTGCTGGCAAAGCTCATGGGTAAACTCGTCTGATCGCTGAGACCGGTGCTGGTCAGCACGGCGTAGCGACTGCAATGGCTGGTGCCTGCTGGGCATGAAAGTGGACATTTACATGGGTGCTGAGGACGTAGAACGTCAGAAAAATGAACGTATTCGGTATGAAGCTGCTGGGTGCAAACGTAATTCAGTTAACTCCGTTCTCGCACCGCGAAAGACGCAATCAACGAGGCTCTGCGTGATTGGGAGGCTACTTTTGAATACACCCACTACCTAATCGGTTCCGTGGTCCGATCCGTATCCGACCATCGTTCTGTGATTTTCAGTCTGTTATCGGTCGTGAGGCTAAAGCGCAGATCCTGGAGGCTGAAGGTCAGCTGCCAGATGTAATCGTTGCTT

GTGTTGGTGGTGGCTCTAACGCGATGGGTATCTTTTACCCGTTTCGTGAACGACAAAAAGTTAAGCTGGTTGGCGTTGAGGCTG
 GTGGTAAAGGCCTGGAATCTGGTAAGCATTCCGCTAGCCTGAACGCAGGTCAGGTTGGTGTGTCCCATGGCATGCTGTCCTACT
 TTCTGCAGGACGAAGAAGGTCAGATCAAACCAAGCCACTCCATCGCACCAGGTCGGATTATCCAGGTGTTGGTCCAGAACACG
 CTTACCTGAAAAAATTCAGCGTGTGTAATACGTGGCTGTAACCGACGAAGAAGCACTGAAAGCGTTCCATGAACTGAGCCGTA
 CCGAAGGTATCATCCCAGCTCTGGAATCTGCGCATGCTGTGGCTTACGCTATGAAACTGGCTAAGGAAATGTCTCGTGATGAGA
 TCATCATCGTAAACCTGTCTGGTCGTGGTGACAAAGACCTGGATATTGTCTGAAAGCGTCTGGCAACGTGCTCGAGCACCACC
 ACCACCACCACTGA

PfTrpB^{7E6}

ATGTGGTTCGGTGAATTTGGTGGTCAGTACGTGCCAGAAACGCTGGTAGGACCCCTGAAAGAGCTGGAAAAAGCTTACAAACGT
 TTCAAAGATGACGAAGAGTTCAATCGTCAGCTGAATTACTACCTGAAACCTGGGCAGGTCGTCCAACCCCACTGTACTACGCAA
 AACGCCTGACTGAAAAAATCGGTGGTGCTAAAAATATACCTGAAACGTGAAGACCTGGTTCACGGTGGTGCACACAAGACCAACA
 ACGCCATCGGTGAGGCACCGCTGGCAAAGCTCATGGGTAAAACTCGTCTGATCGCTGAGACCGGTGCTGGTCAGCACGGCGTA
 GCGACTGCAATGGCTGGTGCCTGCTGGGCATGAAAGTGGACATTTACATGGGTGCTGAGGACGTAGAACGTGAGAAATGAAC
 GTATTCCGTATGAAGCTGCTGGGTGCAAACGTAATTCAGTTAACTCCGTTCTCGCACCAGGTCAGGATTATCCAGGTGTTGGTCCAGAACACG
 CTGCGTGATTGGGAAGCTACTTTTGAATACACCCACTACCTAATCGGTTCCGTGGTCCACATCCGTATCCGACCATCGTTC
 GTGATTTTCAGTCTGTTATCGGTCTGAGGCTAAAGCGCAGATCCTGGAGGCTGAAGGTCAGCTGCCAGATGTAATCGTTGCTT
 GTGTTGGTGGTGGCTCTAACGCGATGGGTATCTTTTACCCGTTTCGTGAACGACAAAAAGTTAAGCTGGTTGGCGTTGAGGCTG
 GTGGTAAAGGCCTGGAATCTGGTAAGCATTCCGCTAGCCTGAACGCAGGTCAGGTTGGTGTGTTGCATGGCATGCTGTCCTACT
 TTCTGCAGGACGAAGAAGGTCAGATCAAACCAAGCCACTCCATCGCACCAGGTCGGATTATCCAGGTGTTGGTCCAGAACACG
 CTTACCTGAAAAAATTCAGCGTGTGTAATACGTGACAGTAACCGATGAAGAAGCACTGAAAGCGTTCCATGAACTGAGCCGTAC
 CGAAGGTATCATCCCAGCTCTGGAATCTGCGCATGCTGTGGCTTACGCTATGAAACTGGCTAAGGAAATGTCTCGTGATGAGAT
 CATCATCGTAAACCTGTCTGGTCGTGGTGACAAAGACCTGGATATTGTCTGAAAGCATCTGGCAACGTGCTCGAGCACCACCA
 CCACCACCACTGA

PfTrpB^{7E6} L161V

ATGTGGTTCGGTGAATTTGGTGGTCAGTACGTGCCAGAAACGCTGGTAGGACCCCTGAAAGAGCTGGAAAAAGCTTACAAACGT
 TTCAAAGATGACGAAGAGTTCAATCGTCAGCTGAATTACTACCTGAAACCTGGGCAGGTCGTCCAACCCCACTGTACTACGCAA
 AACGCCTGACTGAAAAAATCGGTGGTGCTAAAAATATACCTGAAACGTGAAGACCTGGTTCACGGTGGTGCACACAAGACCAACA
 ACGCCATCGGTGAGGCACCGCTGGCAAAGCTCATGGGTAAAACTCGTCTGATCGCTGAGACCGGTGCTGGTCAGCACGGCGTA
 GCGACTGCAATGGCTGGTGCCTGCTGGGCATGAAAGTGGACATTTACATGGGTGCTGAGGACGTAGAACGTGAGAAATGAAC
 GTATTCCGTATGAAGCTGCTGGGTGCAAACGTAATTCAGTTAACTCCGTTCTCGCACCAGTGAAGACGCAATCAACGAGGCT
 CTGCGTGATTGGGAAGCTACTTTTGAATACACCCACTACCTAATCGGTTCCGTGGTCCACATCCGTATCCGACCATCGTTC
 GTGATTTTCAGTCTGTTATCGGTCTGAGGCTAAAGCGCAGATCCTGGAGGCTGAAGGTCAGCTGCCAGATGTAATCGTTGCTT
 GTGTTGGTGGTGGCTCTAACGCGATGGGTATCTTTTACCCGTTTCGTGAACGACAAAAAGTTAAGCTGGTTGGCGTTGAGGCTG
 GTGGTAAAGGCCTGGAATCTGGTAAGCATTCCGCTAGCCTGAACGCAGGTCAGGTTGGTGTGTTGCATGGCATGCTGTCCTACT
 TTCTGCAGGACGAAGAAGGTCAGATCAAACCAAGCCACTCCATCGCACCAGGTCGGATTATCCAGGTGTTGGTCCAGAACACG
 CTTACCTGAAAAAATTCAGCGTGTGTAATACGTGACAGTAACCGATGAAGAAGCACTGAAAGCGTTCCATGAACTGAGCCGTAC
 CGAAGGTATCATCCCAGCTCTGGAATCTGCGCATGCTGTGGCTTACGCTATGAAACTGGCTAAGGAAATGTCTCGTGATGAGAT
 CATCATCGTAAACCTGTCTGGTCGTGGTGACAAAGACCTGGATATTGTCTGAAAGCATCTGGCAACGTGCTCGAGCACCACCA
 CCACCACCACTGA

PfTrpB^{2G8}

ATGTGGTTCGGTGAATTTGGTGGTCAGTACGTGCCAGAAACGCTGGTAGGACCCCTGAAAGAGTTGAAAAAGCTTACAAACGT
 TTCAAAGATGACGAAGAGTTCAATCGTCAGCTGAATTACTACCTGAAACCTGGGCAGGTCGTCCAACCCCACTGTACTACGCAA
 AACGCCTGACTGAAAAAATCGGTGGTGCTAAAAATATACCTGAAACGTGAAGACCTGGTTCACGGTGGTGCACACAAGACCAACA
 ACGCCATCGGTGAGGCACCTGCTGGCAAAGCTCATGGGTAAAACTCGTCTGATCGCTGAGACCGGTGCTGGTCAGCACGGCGTA
 GCGACTGCAATGGCTGGTGCCTGCTGGGCATGAAAGTGGACATTTACATGGGTGCTGAGGACGTAGAACGTGAGAAATGAAC
 CGTATTCCGTATGAAGCTGCTGGGTGCAAACGTAATTCAGTTAACTCCGTTCTCGCACCAGTGAAGACGCAATCAACGAGGCT
 TCTGCGTGATTGGGAAGCTACTTTTGAATACACCCACTACCTAATCGGTTCCGTGGTCCACATCCGTATCCGACCATCGTTC
 CGTGATTTTCAGTCTGTTATCGGTCTGAGGCTAAAGCGCAGATCCTGGAGGCTGAAGGTCAGCTGCCAGATGTAATCGTTGCT
 TGTGTTGGTGGTGGCTCTAACGCGATGGGTATCTTTTACCCGTTTCGTGAACGACAAAAAGTTAAGCTGGTTGGCGTTGAGGCT
 GGTGGTAAAGGCCTGGAATCTGGTAAGCATTCCGCTAGCCTGAACGCAGGTCAGGTTGGTGTGTTGCATGGCATGCTGTCCTAC
 TTTCTGCAGGACGAAGAAGGTCAGATCAAACCAAGCCACTCCATCGCACCAGGTCGGATTATCCAGGTGTTGGTCCAGAACAC
 GCTTACCTGAAAAAATTCAGCGTGTGTAATACGTGACAGTAACCGATGAAGAAGCACTGAAAGCGTTCCATGAACTGAACCGTA
 CCGAAGGTATCATCCCAGCTCTGGAATCTGCGCATGCTGTGGCTTACGCTATGAAACTGGCTAAGGAAATGTCTCGTGATGAGA
 TCATCATCGTAAACCTGTCTGGTCGTGGTGACAAAGACCTGGATATTGTCTGAAAGCATCTGGCAACGTGCTCGAGCACCACC
 ACCACCACCACTGA

Protein expression and purification. A single colony containing the appropriate *PfTrpB* gene was used to inoculate 5 mL Terrific Broth supplemented with 100 µg/mL ampicillin (TB_{amp}) and incubated overnight at 37 °C and 230 rpm. For expression, 2.5 mL of overnight culture were used to inoculate 250 mL TB_{amp} in a 1-L flask and incubated at 37 °C and 250 rpm for 3 h to reach OD₆₀₀ 0.6 to 0.8. Cultures were chilled on ice for 20 min and expression was induced with a final concentration of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). Expression proceeded at 25 °C and 250 rpm for approximately 20 h. Cells were harvested by centrifugation at 5,000g for 5 min at 4 °C, and then the supernatant was decanted. The pellet was stored at -20 °C until further use.

Thawed cell pellets were resuspended in 9 mL of lysis buffer containing 25 mM potassium phosphate buffer, pH 8.0 (KPi buffer) with 100 mM NaCl, 20 mM imidazole, 1 mg/mL hen egg white lysozyme (HEWL), 200 μ M pyridoxal phosphate (PLP), 2 mM MgCl_2 , 0.02 mg/mL DNase I. Pellets were completely resuspended and then lysed with 1 mL BugBuster[®] according to manufacturer's recommendations. Lysate was heat treated at 75 °C for 15 min. The supernatant was collected from clarified lysate following centrifugation for 15 min at 15,000g and 4 °C. Purification was performed with an AKTA purifier FPLC system (GE Healthcare) and a 1-mL Ni-NTA column. Protein was eluted by applying a linear gradient of 100 mM to 500 mM imidazole in 25 mM KPi buffer, pH 8.0 and 100 mM NaCl. Fractions containing purified protein were dialyzed into 50 mM KPi buffer, pH 8.0, flash frozen in liquid nitrogen, and stored at -80 °C. Protein concentrations were determined using the Bio-Rad Quick Start[™] Bradford Protein Assay.

Library expression and screening. Single colonies from libraries containing the appropriate *PfTrpB* variant genes were expressed in 96-well deep-well plates containing 300 μ L of TB_{amp} and incubated overnight (approximately 20 h) at 25 °C and 250 rpm with 80% humidity. For expression, 20 μ L of overnight culture were transferred into 630 μ L TB_{amp} and incubated for 3 h at 37 °C and 250 rpm with 80% humidity. Cells were then chilled on ice for 20 min and induced with 50 μ L of IPTG in TB_{amp} (0.5 mM–1 mM final concentration), followed by overnight incubation at 37 °C and 250 rpm. Cells were harvested by centrifugation at 4 °C and 4,000g for 15 min and then stored at -20 °C for at least 24 h. Cell plates were thawed and resuspended in 400 μ L/well 50 mM KPi buffer, pH 8.0 with 1 mg/mL HEWL, 100 μ M PLP, 2 mM MgCl_2 , and 0.02 mg/mL DNase. Cells were lysed by a 30–60-min incubation at 37 °C and heat treatment in a 75 °C water bath for 20 min. Lysate was clarified by centrifugation at 5,000g for 10 min.

Reactions were performed in a UV-transparent 96-well assay plate with a total volume of 200 μ L/well comprised of 20–40 μ L heat-treated lysate, 500 μ M indole, and 5 mM β -DL-ethylserine in 50 mM KPi buffer, pH 8.0. Due to the racemic nature of the substrate, the effective concentration of β -L-ethylserine is 2.5 mM. Reactions proceeded in a 75 °C water bath and were assessed for product formation at multiple time points (0.5–4 h). Prior to being measured, plates were cooled on ice and centrifuged briefly to collect condensation and assayed by measuring absorption at 290 nm.

UV-Vis Spectroscopy. Spectra were collected on a Shimadzu UV1800 spectrophotometer in a quartz cuvette with a 1 cm path length at 75 °C.

Steady-state distribution of catalytic intermediates. Spectra were collected between 250 nm and 500 nm immediately following substrate addition. Samples were prepared in a total volume of 400 μ L with 20 μ M purified enzyme and 20 mM substrate (threonine, β -L-ethylserine, β -L-propylserine) in 50–200 mM KPi buffer, pH 8.0. Data were baseline subtracted and normalized to the E(Ain) peak at 412 nm. Catalytic intermediates were assigned at the following wavelengths: E(Ain) at 412 nm, E(Aex₁) at 428 nm, and E(A-A) at 350 nm.

Deamination of the amino-acrylate. Spectra were collected between 250–550 nm immediately following substrate addition, and then once per min for ten min. Samples were prepared in a total volume of 400 μ L with 20 μ M purified enzyme and 20 mM substrate (Threonine, β -L-ethylserine, β -L-propylserine) in 50–200 mM KPi buffer, pH 8.0. Data were baseline subtracted and plotted as absorbance over time where α -keto acid formation is represented by the slope. Deamination is described in AU/min as the extinction coefficient is unknown for β -L-ethylserine and β -L-propylserine.

Isosbestic points. Spectra were collected between 250 nm and 550 nm immediately following Ser addition, and then once per min for ten min. Samples were prepared in a total volume of 400 μ L with 1 μ M of purified enzyme and 100 μ M–1 mM nucleophile

substrate in 50 mM KPi buffer, pH 8.0. The isosbestic point was defined as the overlapped position of the starting material and product UV peaks. The isosbestic point of some nucleophiles have been reported previously.^[2]

Small-scale analytical reactions. All analytical reactions were performed in 2-mL glass HPLC vials charged with nucleophile substrate, followed by addition of amino acid substrate and purified enzyme in 50 mM KPi buffer, pH 8.0 to a final volume of 150 μ L. Reactions were incubated in a 75 °C water bath for 24 h. The reaction was then diluted with 850 μ L of 1:1 1-M aq. HCl/CH₃CN and vortexed thoroughly. The reaction mixture was then subjected to centrifugation at >20,000g for 10 min and the supernatant analyzed by HPLC. Yields were determined at the relevant isosbestic point (**Table S4**) and calculated as area of the product peak divided by the sum of the integrated product and substrate peaks. All reactions were performed at least in duplicate.

TTN determination. A 2-mL glass HPLC vial was charged with 20 mM nucleophile substrate as 6 μ L of a 500-mM solution in DMSO. Next, 20 mM amino acid substrate and then 2 μ M purified enzyme (0.01% catalyst loading, 10,000 max TTN) were added as solutions in 50 mM KPi buffer, pH 8.0. The reactions were worked up and analyzed as described above. TTN were determined as yield times max TTN.

Coupling efficiency. A 2-mL glass HPLC vial was charged with 20 mM nucleophile substrate as 6 μ L of a 500-mM solution in DMSO. Next, 20 mM amino acid substrate and 20 μ M purified enzyme (0.1% catalyst loading, 1,000 max TTN) were added as a solution in 50 mM KPi buffer, pH 8.0. Coupling efficiency was described as the yield under reaction conditions with high catalyst loading and equimolar substrate equivalents.

Synthesis and characterization of tryptophan analogs. Preparative reactions were carried out by adding 100 μ mol of nucleophile substrate and 200 μ mol L-amino acid substrate to a 40-mL reaction vial. Following substrate addition, 10 mL of 50 mM KPi buffer, pH 8.0 containing purified *Pf*TrpB^{7E6} or *Pf*TrpB^{2G8} at 0.01–0.4% catalyst loading. The reaction mixture was incubated in a 75 °C water bath for 24 h, frozen on dry ice, and then the water was removed by lyophilization. Approximately 4 mL of 1:1 CH₃CN/1 M aq. HCl were added to the remaining solid and the volume was reduced *in vacuo*. The sample was resuspended in water and loaded onto a 12g C-18 column equilibrated with 1% methanol/water (0.1% HCl by mass) on a Biotage Isolera One purification system. The column was washed with three column volumes (CV) of 1% methanol/water mixture. The product was eluted with a gradient from 1% to 100% methanol over 10 CV. The fractions containing the UV-active product were combined and the volume reduced *in vacuo*. The product was then suspended in water (0.1% HCl by mass) and transferred to a tared vial before being frozen on dry ice and lyophilized. Yields were determined by product mass following lyophilization relative to theoretical yield with indole analog as the limiting reagent. Products were obtained as hydrochloride salts and product identities were confirmed by ¹H- and ¹³C-NMR and high-resolution mass spectrometry.

Determination of Optical Purity. Product optical purity was estimated by derivatization with FDNP-alanamide. Approximately 0.5 μ mol of purified β -MeTrp, β -MeTrp, or β -MeTrp were added to a 2-mL vial. The product was resuspended in 100 μ L of 1 M aq. NaHCO₃. FDNP-alanamide (10 μ L of a 33-mM solution in acetone, 0.33 μ mol) was added to each vial, followed by a two-hour incubation at 37 °C and 230 rpm. The reaction mixture was then cooled to room temperature and diluted with 1:1 CH₃CN/1-M aq. HCl (600 μ L). The resulting solution was analyzed directly by LCMS at 330 nm. Each amino acid was derivatized with both racemic and enantiopure FDNP-alanamide for comparison. Absolute stereochemistry was inferred by analogy to L-tryptophan. All products were >99% ee.

Determination of T_{50} values. A solution of 1 μ M purified enzyme in 50 mM KPi buffer, pH 8.0 was aliquoted into 12 PCR tubes with a volume of 95 μ L/tube. Ten of these samples were incubated in a thermocycler for 60 min with a temperature gradient from 75 °C to 95 °C, while the two remaining samples were incubated at room temperature as controls. All 12 tubes were centrifuged for three min to pellet precipitated enzyme, and then 75 μ L of the supernatant were transferred from each tube to a UV-transparent 96-well assay plate. Enzyme activity was determined by adding an additional 75 μ L of 50 mM KPi buffer, pH 8.0 containing 1 mM indole and 1 mM serine to each well. Reactions were incubated for 10 min at 75 °C and then briefly centrifuged to collect condensation. Activity was determined by measuring product formation at 290 nm. Activity was correlated to incubation temperature, and thermostability is reported as the temperature at which half of the activity is lost (T_{50}) after 1-hour incubation. Measurements were conducted in duplicate.

Crystallography. Seed stocks of wild-type *PfTrpB* crystals were used to seed crystallization of *PfTrpB*^{7E6}. The wild-type *PfTrpB* crystal was obtained by sitting drop vapor diffusion against a 1-mL reservoir containing 24% PEG3350 and 50 mM Na HEPES, pH 7.85. The seed stock was prepared according to the Seed Bead method (Hampton Research) using 24% PEG3350 and 50 mM Na HEPES, pH 7.85 as stabilization buffer. The seed stock was diluted 2,000x in stabilization buffer before use. *PfTrpB*^{7E6} crystals were grown in sitting drops against a 1-mL reservoir of 14% PEG3350 and 0.1 M Na HEPES (pH 7.85) with mother liquor comprised of 1.5 μ L of 18.8 mg/mL *PfTrpB*^{7E6} and 1.5 μ L of 2,000x diluted seed stock.

Ligand-bound structures were determined by soaking *PfTrpB*^{7E6} crystals with the compound of interest. From a 50/50% (v/v) mixture containing 0.5 M β -DL-ethylserine in 0.2 M KPi buffer, pH 8.0 and stabilization buffer, 0.5 μ L were added to the sitting drop and incubated for 2 h. (2S)- β -isopropylserine was soaked into *PfTrpB*^{7E6} crystals by adding powdered substrate directly to the sitting drop, mixing gently, and incubating for one hour.

Crystals were cryoprotected through oil immersion in Fomblin Y (Sigma) and flash-frozen in liquid nitrogen until diffraction. Diffraction data were collected remotely at the Stanford Synchrotron Radiation Laboratories on beamline 12-2. Crystals routinely diffracted at or below 2.25 Å, and the data were integrated and scaled using XDS^[5] and AIMLESS.^[6] A resolution cutoff of $CC_{1/2} > 0.3$ was applied along the strongest axis of diffraction.^[6,7] These data were later judged to contribute to model quality based on R_{free} in the final bin < 0.4 . Structures were solved using molecular replacement with PHASER, as implemented in CCP4.^[8,9] The search model comprised a single monomer of *PfTrpB*^{2B9} (*holo* and (2S, 3R)- β -EtSer, PDB: 5VM5) or *PfTrpB*^{4D11} ((2S, 3S)- β -iPrSer) and was prepared by removing all ligand, trimmed mutated residues to C β , and then subjected to ten cycles of geometric idealization in REFMAC5.^[9] Model-building was performed in Coot^[10] beginning with data processed at 2.4 Å, followed by subsequent inclusion of increasingly higher-resolution shells of data with relaxed geometric constraints. This procedure was particularly important for the structures of β -L-ethylserine and β -L-isopropylserine-bound *PfTrpB*^{7E6}, which contained a large rigid body motion of the COMM domain relative to the search model. Refinement was performed using REFMAC5⁹. The MolProbity server was used to identify rotamer flips and to identify clashes.^[11] After the protein, ligand, and solvent atoms were built, TLS operators were added to refinement, which resulted in substantial improvements in R_{free} for the models. Despite these improvements, R_{free} were still high compared to structures of similar resolution. This has proved to be a general feature of TrpB crystals in the spacegroup, and we hypothesize the high R factors are caused by the considerable domain-level conformational heterogeneity that is difficult to parameterize with a single structure.^[12] This heterogeneity was increased when ligands were soaked into *PfTrpB*^{7E6} and bound with incomplete occupancy. We observed much weaker density in the COMM domain of several protamers, which led to an uncommonly

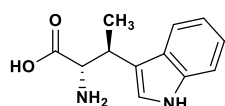
high number of residues flagged as density outliers that were nonetheless well-determined in structure. Crystallographic and refinement statistics are reported in **Table S11**. Coordinates are deposited in the Protein Data Bank with ID codes 6CUV (holo *PfTrpB*^{7E6}), 6CUZ ((2*S*, 3*R*)- β -ethylserine-bound *PfTrpB*^{7E6}), and 6CUT ((2*S*, 3*S*)- β -isopropylserine-bound *PfTrpB*^{7E6}).

Table S11. Crystallographic data collection and refinement statistics.

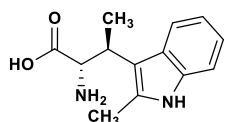
Protein	<i>PfTrpB</i> ^{7E6}	<i>PfTrpB</i> ^{7E6}	<i>PfTrpB</i> ^{7E6}
PDB ID Code	6CUV	6CUZ	6CUT
Ligand	None	(2 <i>S</i> ,)- β -ethylserine	(2 <i>S</i> , 3 <i>S</i>)- β -isopropylserine
Space Group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions, Å	a,b,c = 83.6, 108.6, 159.3	a,b,c = 84.2, 109.3, 159.9	a,b,c = 82.2, 107.4, 159.3
Cell angles	$\alpha = \beta = \gamma = 90^\circ$	$\alpha = \beta = \gamma = 90^\circ$	$\alpha = \beta = \gamma = 90^\circ$
Data Collection			
Wavelength, Å	1.19499	1.19499	0.97946
Beamline	SSRL 12.2	SSRL 12.2	SSRL 12.2
Resolution, Å	40-2.26	40-1.75	40-1.77
Last bin (Å)	2.31-2.26	1.78-1.75	1.80-1.77
No. observations	422,578	610,237	920,320
Completeness (%)	100.0 (100.0)	100.0 (100.0)	99.9 (99.9)
R _{pim} (%)	0.058 (0.719)	0.050 (0.613)	0.030 (1.25)
CC(1/2)	0.990 (0.655)	0.981 (0.753)	0.998 (0.452)
I/ σ I	8.9 (1.0)	8.2 (0.8)	12.0 (0.6)
Redundancy	6.2 (6.2)	4.1 (4.1)	6.7 (6.7)
Refinement			
Total no. of reflections	63,878	141,404	130,162
Total no. of atoms	11,687	11,996	11,972
Final bin (Å)	2.32-2.26	1.80-1.75	1.82-1.77
R _{work} (%)	21.1(36.3)	23.5 (36.7)	19.3 (39.5)
R _{free} (%)	25.9 (38.3)	26.1 (38.0)	22.5 (40.1)
Average B factor, Å ²	26.1	14.8	24.7
Ramachandran plot favored, %	97	98	98
Allowed, %	99.8	99.9	99.8
Outliers, %	0.2	0.1	0.2

Values in parenthesis are for the highest resolution shell. R_{merge} is $\Sigma||I_o - \bar{I}| / \Sigma I_o$, where I_o is the intensity of an individual reflection, and \bar{I} is the mean intensity for multiply recorded reflections. R_{work} is $\Sigma||F_o - F_c| / F_o$, where F_o is an observed amplitude and F_c a calculated amplitude; R_{free} is the same statistic calculated over a 5% subset of the data that has not been included. Ramachandran statistics calculated by the MolProbity server.

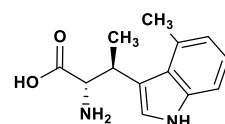
Product Characterization



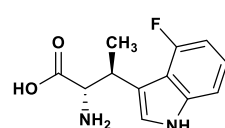
β -methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.66 (dt, J = 8.0, 0.9 Hz, 1H), 7.49 (dt, J = 8.2, 0.9 Hz, 1H), 7.32 (s, 1H), 7.23 (ddd, J = 8.3, 6.8, 1.1 Hz, 1H), 7.13 (ddt, J = 7.9, 7.0, 0.8 Hz, 1H), 4.23 (d, J = 5.5 Hz, 1H), 3.85 (qd, J = 7.3, 5.4 Hz, 1H), 1.53 (d, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.87, 136.35, 125.56, 124.20, 122.28, 119.49, 118.67, 112.08, 57.91, 32.26, 17.29. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{12}\text{H}_{15}\text{N}_2\text{O}_2$ requires 219.1134, observed 219.1113.



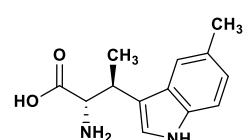
β -methyl-2-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.61 (dt, J = 7.8, 1.0 Hz, 1H), 7.39 (dt, J = 8.1, 0.9 Hz, 1H), 7.14 (ddd, J = 8.1, 7.0, 1.2 Hz, 1H), 7.07 (ddd, J = 8.1, 7.0, 1.2 Hz, 1H), 4.28 (d, J = 9.4 Hz, 1H), 3.49 (dq, J = 9.4, 7.1 Hz, 1H), 2.35 (s, 3H), 1.51 (d, J = 7.2 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.89, 135.64, 134.86, 125.73, 121.22, 119.31, 118.27, 111.41, 107.25, 57.63, 33.00, 16.77, 10.88. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{13}\text{H}_{17}\text{N}_2\text{O}_2$ requires 233.1290, observed 233.1278.



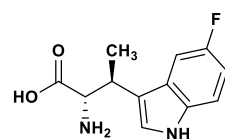
β -methyl-4-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.33 (d, J = 7.3 Hz, 2H), 7.10 (dd, J = 8.2, 7.1 Hz, 1H), 6.89 (dt, J = 7.1, 1.0 Hz, 1H), 4.24 (d, J = 7.0 Hz, 1H), 4.08 – 3.98 (m, 1H), 2.64 (s, 3H), 1.44 (d, J = 7.1 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.29, 136.31, 130.16, 124.17, 123.11, 122.22, 121.29, 114.44, 109.96, 58.50, 32.59, 19.56, 18.53. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{13}\text{H}_{17}\text{N}_2\text{O}_2$ requires 233.1290, observed 233.1297.



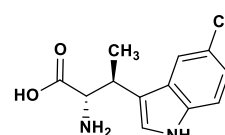
β -methyl-4-fluorotryptophan. ^1H NMR (400 MHz, D_2O) δ 7.24 (d, J = 8.1 Hz, 2H), 7.11 (td, J = 8.0, 5.2 Hz, 1H), 6.78 (ddd, J = 12.0, 7.9, 0.7 Hz, 1H), 4.21 (d, J = 6.7 Hz, 1H), 3.71 (p, J = 7.0 Hz, 1H), 1.47 – 1.40 (m, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.57, 157.25, 154.84, 139.42, 139.30, 124.56, 122.70, 122.62, 114.00, 113.80, 111.18, 111.15, 108.28, 108.24, 104.53, 104.33, 58.14, 58.11, 33.21, 16.85, 16.83. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{12}\text{H}_{14}\text{FN}_2\text{O}_2$ requires 237.1039, observed 237.1011.



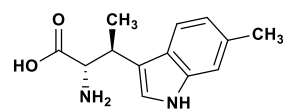
β -methyl-5-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.40 (dt, J = 1.8, 0.9 Hz, 1H), 7.32 (d, J = 8.3 Hz, 1H), 7.19 (s, 1H), 7.01 (dd, J = 8.3, 1.5 Hz, 1H), 4.11 (d, J = 6.1 Hz, 1H), 3.42 (dt, J = 10.1, 5.9 Hz, 1H), 2.31 (s, 3H), 1.93 – 1.75 (m, 2H), 0.75 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.63, 134.73, 129.20, 126.51, 124.96, 123.75, 118.11, 111.90, 109.50, 57.49, 39.87, 24.60, 20.48, 11.41. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{13}\text{H}_{17}\text{N}_2\text{O}_2$ requires 233.1290, observed 233.1291.



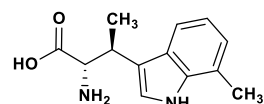
β -methyl-5-fluorotryptophan. ^1H NMR (400 MHz, D_2O) δ 7.41 – 7.29 (m, 2H), 7.25 (dd, J = 10.3, 2.5 Hz, 1H), 6.95 (td, J = 9.3, 2.5 Hz, 1H), 4.18 (d, J = 5.2 Hz, 1H), 3.72 (qd, J = 7.3, 5.0 Hz, 1H), 1.44 (d, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.62, 158.54, 156.24, 132.85, 125.83, 125.74, 112.82, 112.72, 112.07, 112.03, 110.52, 110.26, 103.33, 103.09, 57.69, 32.05, 17.06. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{12}\text{H}_{13}\text{FN}_2\text{O}_2$ requires 237.1039, observed 237.1031.



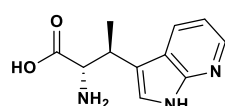
β -methyl-5-chlorotryptophan. ^1H NMR (400 MHz, D_2O) δ 7.66 (d, J = 1.9 Hz, 1H), 7.43 (d, J = 8.7 Hz, 1H), 7.35 (s, 1H), 7.19 (dd, J = 8.7, 1.9 Hz, 1H), 4.17 (d, J = 5.5 Hz, 1H), 3.85 – 3.73 (m, 1H), 1.52 (d, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.11, 134.79, 126.68, 125.55, 124.56, 122.22, 117.95, 113.18, 112.04, 58.15, 32.20, 17.22. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{12}\text{H}_{14}\text{ClN}_2\text{O}_2$ requires 253.0744, observed 253.0740.



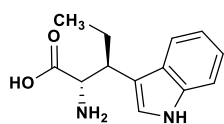
β -methyl-6-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.53 (d, J = 8.2 Hz, 1H), 7.29 (dt, J = 1.6, 0.8 Hz, 1H), 7.23 (s, 1H), 6.98 (dd, J = 8.2, 1.4 Hz, 1H), 4.21 (d, J = 5.5 Hz, 1H), 3.80 (qd, J = 7.2, 5.3 Hz, 1H), 2.39 (s, 3H), 1.50 (d, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.80, 136.87, 132.59, 123.57, 123.41, 121.14, 118.49, 111.89, 111.71, 57.83, 32.31, 20.58, 17.29. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{13}\text{H}_{17}\text{N}_2\text{O}_2$ requires 233.1290, observed 233.1283.



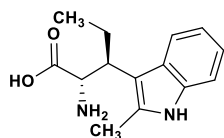
β -methyl-7-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.50 – 7.43 (m, 1H), 7.31 (s, 1H), 7.07 – 6.98 (m, 2H), 4.22 (d, J = 5.4 Hz, 1H), 3.86 – 3.74 (m, 1H), 2.43 (d, J = 0.9 Hz, 3H), 1.49 (d, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.63, 135.84, 125.23, 123.98, 122.45, 122.06, 119.79, 116.26, 112.47, 57.74, 32.28, 17.21, 15.85. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{13}\text{H}_{17}\text{N}_2\text{O}_2$ requires 233.1290, observed 233.1281.



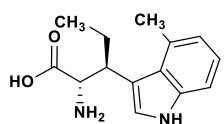
β -methyl-7-azatryptophan. ^1H NMR (400 MHz, D_2O) δ 8.68 (dd, J = 8.1, 1.2 Hz, 1H), 8.35 (dd, J = 6.1, 1.1 Hz, 1H), 7.68 (s, 1H), 7.55 (dd, J = 8.1, 6.0 Hz, 1H), 4.32 (d, J = 4.8 Hz, 1H), 4.06 – 3.94 (m, 1H), 1.56 (d, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.07, 138.47, 136.99, 132.90, 127.86, 124.76, 115.43, 114.35, 57.55, 31.33, 16.30. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{11}\text{H}_{11}\text{N}_3\text{O}_2$ requires 221.1133, observed 221.1144.



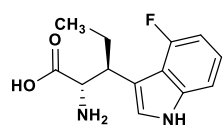
β -ethyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.66 (dt, J = 8.0, 1.0 Hz, 1H), 7.50 (dt, J = 8.2, 0.9 Hz, 1H), 7.31 (s, 1H), 7.23 (ddd, J = 8.2, 7.0, 1.2 Hz, 1H), 7.13 (ddd, J = 8.1, 7.1, 1.1 Hz, 1H), 4.28 (d, J = 5.5 Hz, 1H), 3.61 (dt, J = 9.8, 5.8 Hz, 1H), 2.03 – 1.83 (m, 2H), 0.84 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.00, 136.36, 126.28, 124.75, 122.27, 119.49, 118.63, 112.07, 109.76, 56.92, 39.52, 24.61, 11.38. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{13}\text{H}_{17}\text{N}_2\text{O}_2$ requires 233.1290, observed 233.1293.



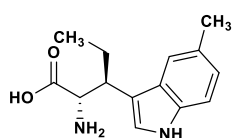
β -ethyl-2-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.62 – 7.56 (m, 1H), 7.41 (dt, J = 8.1, 0.9 Hz, 1H), 7.14 (ddd, J = 8.1, 7.0, 1.1 Hz, 1H), 7.07 (ddd, J = 8.2, 7.1, 1.2 Hz, 1H), 4.32 (d, J = 9.2 Hz, 1H), 3.31 – 3.20 (m, 1H), 2.36 (s, 3H), 2.10 – 1.96 (m, 1H), 1.96 – 1.84 (m, 1H), 0.65 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.18, 136.36, 135.67, 125.90, 121.19, 119.30, 118.14, 111.40, 104.83, 57.21, 123.61, 23.45, 11.21, 10.94. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{19}\text{N}_2\text{O}_2$ requires 247.1447, observed 247.1445.



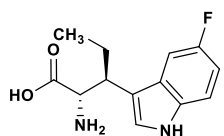
β -ethyl-4-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.40 – 7.29 (m, 2H), 7.12 (dd, J = 8.2, 7.1 Hz, 1H), 6.90 (dt, J = 7.1, 1.0 Hz, 1H), 4.18 (d, J = 6.7 Hz, 1H), 3.91 (s, 1H), 2.66 (s, 3H), 1.99 – 1.86 (m, 1H), 1.86 – 1.71 (m, 1H), 0.85 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.08, 136.22, 130.38, 125.44, 123.61, 122.16, 121.50, 111.88, 110.10, 58.37, 39.19, 26.73, 20.00, 10.92. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{19}\text{N}_2\text{O}_2$ requires 247.1447, observed 247.1448.



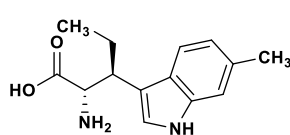
β -ethyl-4-fluorotryptophan. ^1H NMR (400 MHz, D_2O) δ 7.30 – 7.23 (m, 2H), 7.11 (td, J = 8.0, 5.2 Hz, 1H), 6.78 (ddd, J = 12.0, 7.9, 0.8 Hz, 1H), 4.21 (d, J = 7.1 Hz, 1H), 3.52 – 3.41 (m, 1H), 1.90 – 1.75 (m, 2H), 0.71 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.95, 157.32, 154.91, 139.54, 139.42, 125.66, 122.63, 122.55, 114.47, 114.27, 108.42, 108.38, 108.31, 108.28, 104.59, 104.39, 57.50, 57.47, 40.66, 24.50, 24.47, 11.31. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{13}\text{H}_{16}\text{FN}_2\text{O}_2$ requires 251.1196, observed 251.1186.



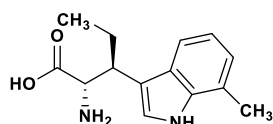
β -ethyl-5-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.40 (dt, J = 1.8, 0.9 Hz, 1H), 7.32 (d, J = 8.3 Hz, 1H), 7.19 (s, 1H), 7.01 (dd, J = 8.3, 1.5 Hz, 1H), 4.11 (d, J = 6.1 Hz, 1H), 3.42 (dt, J = 10.1, 5.9 Hz, 1H), 2.31 (s, 3H), 1.93 – 1.75 (m, 2H), 0.75 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.63, 134.73, 129.20, 126.51, 124.96, 123.75, 118.11, 111.90, 109.50, 57.49, 39.87, 24.60, 20.48, 11.41. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{19}\text{N}_2\text{O}_2$ requires 247.1447, observed 247.1451.



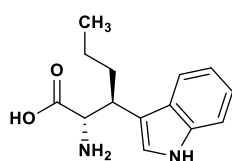
β -ethyl-5-fluorotryptophan. ^1H NMR (400 MHz, D_2O) δ 7.40 (dd, J = 8.9, 4.6 Hz, 1H), 7.34 – 7.24 (m, 2H), 6.97 (td, J = 9.3, 2.5 Hz, 1H), 4.24 (d, J = 5.3 Hz, 1H), 3.51 (td, J = 7.9, 5.3 Hz, 1H), 1.87 (p, J = 7.4 Hz, 2H), 0.80 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.94, 158.60, 156.30, 132.90, 126.59, 126.49, 126.31, 112.83, 112.73, 110.54, 110.28, 109.94, 109.89, 103.35, 103.11, 56.78, 39.37, 24.47, 11.32. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{13}\text{H}_{16}\text{FN}_2\text{O}_2$ requires 251.1196, observed 251.1186.



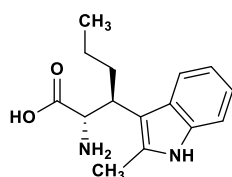
β -ethyl-6-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.54 (d, J = 8.2 Hz, 1H), 7.31 (s, 1H), 7.22 (s, 1H), 6.99 (d, J = 8.2 Hz, 1H), 4.23 (d, J = 5.7 Hz, 1H), 3.54 (dt, J = 9.7, 5.9 Hz, 1H), 2.40 (s, 3H), 2.02 – 1.82 (m, J = 6.7 Hz, 2H), 0.83 (t, J = 7.2 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.23, 136.89, 132.57, 124.13, 124.10, 121.13, 118.50, 111.71, 109.74, 57.11, 39.70, 24.61, 20.58, 11.40. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{19}\text{N}_2\text{O}_2$ requires 247.1447, observed 247.1444.



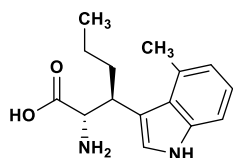
β -ethyl-7-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.49 – 7.39 (m, 1H), 7.28 (d, J = 4.4 Hz, 1H), 7.05 – 6.93 (m, 2H), 4.23 (t, J = 4.0 Hz, 1H), 3.52 (td, J = 8.7, 4.1 Hz, 1H), 2.42 (d, J = 5.7 Hz, 3H), 1.86 (dtd, J = 13.5, 7.8, 5.4 Hz, 2H), 0.82 – 0.71 (m, 3H). ^{13}C NMR (101 MHz, D_2O) δ 171.90, 135.85, 125.96, 124.48, 122.43, 121.99, 119.77, 116.25, 110.19, 56.85, 39.60, 24.60, 15.87, 11.38. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{19}\text{N}_2\text{O}_2$ requires 247.1447, observed 247.1448.



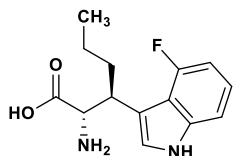
β -propyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.65 (dt, J = 8.1, 1.0 Hz, 1H), 7.49 (dt, J = 8.0, 0.8 Hz, 1H), 7.29 (s, 1H), 7.22 (ddd, J = 8.1, 6.9, 1.1 Hz, 1H), 7.12 (ddd, J = 8.1, 7.1, 1.1 Hz, 1H), 4.23 (d, J = 5.7 Hz, 1H), 3.69 (dt, J = 10.8, 5.4 Hz, 1H), 2.00 – 1.87 (m, 1H), 1.86 – 1.76 (m, 1H), 1.26 – 1.14 (m, 2H), 0.81 (t, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.05, 136.35, 126.21, 124.74, 122.24, 119.48, 118.63, 112.07, 109.92, 57.26, 37.36, 33.31, 20.01, 12.84. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{19}\text{N}_2\text{O}_2$ requires 247.1447, observed 247.1456.



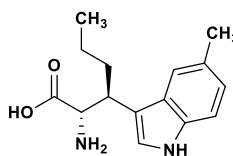
β -propyl-2-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.65 (d, J = 7.9 Hz, 1H), 7.42 (dt, J = 8.1, 1.0 Hz, 1H), 7.17 (ddd, J = 8.2, 7.1, 1.2 Hz, 1H), 7.10 (ddd, J = 8.1, 7.1, 1.2 Hz, 1H), 4.27 (d, J = 9.2 Hz, 1H), 3.34 (d, J = 15.7 Hz, 1H), 2.38 (s, 3H), 2.11 (tt, J = 13.2, 7.0 Hz, 1H), 1.80 (dtd, J = 12.6, 8.0, 4.4 Hz, 1H), 1.07 (h, J = 7.4 Hz, 2H), 0.77 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.66, 135.67, 121.21, 119.30, 118.23, 111.38, 105.31, 57.76, 38.31, 32.23, 30.23, 30.23, 20.04, 12.89, 10.82. HRMS (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{15}\text{H}_{21}\text{N}_2\text{O}_2$ requires 261.1603, observed 261.1611.



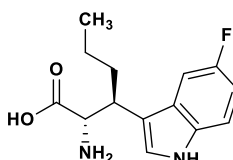
β -propyl-4-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.39 – 7.30 (m, 2H), 7.11 (dd, J = 8.2, 7.1 Hz, 1H), 6.89 (d, J = 7.0 Hz, 1H), 4.13 (d, J = 6.4 Hz, 1H), 3.94 (d, J = 37.8 Hz, 1H), 2.66 (s, 3H), 1.89 – 1.72 (m, 2H), 1.25 (p, J = 7.3 Hz, 2H), 0.82 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.15, 136.13, 130.27, 125.22, 123.58, 122.04, 121.42, 112.06, 110.00, 58.66, 37.26, 35.62, 19.75, 13.03, 12.76. **HRMS** (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{15}\text{H}_{16}\text{N}_2\text{O}_2\text{H}_3$ requires 262.1640, observed 262.1635.



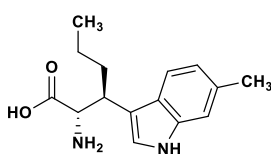
β -propyl-4-fluorotryptophan. ^1H NMR (400 MHz, D_2O) δ 7.34 – 7.26 (m, 2H), 7.15 (td, J = 8.0, 5.2 Hz, 1H), 6.83 (dd, J = 12.1, 7.8 Hz, 1H), 4.20 (d, J = 7.3 Hz, 1H), 3.59 (dd, J = 10.6, 5.6 Hz, 1H), 1.99 – 1.87 (m, 1H), 1.84 – 1.73 (m, 1H), 1.23 – 1.07 (m, 2H), 0.80 (t, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.19, 157.32, 154.90, 139.55, 139.43, 125.67, 122.68, 122.60, 108.76, 108.73, 108.32, 108.29, 104.63, 104.43, 94.96, 57.93, 38.54, 38.18, 33.24, 19.97, 12.78. **HRMS** (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{15}\text{FN}_2\text{O}_2\text{H}_3$ requires 268.1541, observed 268.1531.



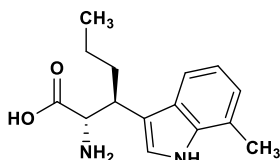
β -propyl-5-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.50 (dt, J = 1.7, 0.9 Hz, 1H), 7.47 – 7.37 (m, 1H), 7.27 (s, 1H), 7.09 (dd, J = 8.3, 1.6 Hz, 1H), 4.13 (d, J = 6.4 Hz, 1H), 3.57 (dt, J = 11.2, 5.9 Hz, 1H), 2.40 (s, 3H), 1.97 (dddd, J = 13.8, 11.0, 8.7, 5.5 Hz, 1H), 1.80 (td, J = 13.4, 7.7 Hz, 1H), 1.27 – 1.14 (m, 2H), 0.82 (t, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.93, 134.74, 129.18, 126.43, 124.96, 123.72, 118.16, 111.91, 109.80, 58.17, 37.76, 33.36, 20.48, 20.03, 12.85. **HRMS** (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{15}\text{H}_{18}\text{N}_2\text{O}_2\text{H}_3$ requires 264.1791, observed 264.1799.



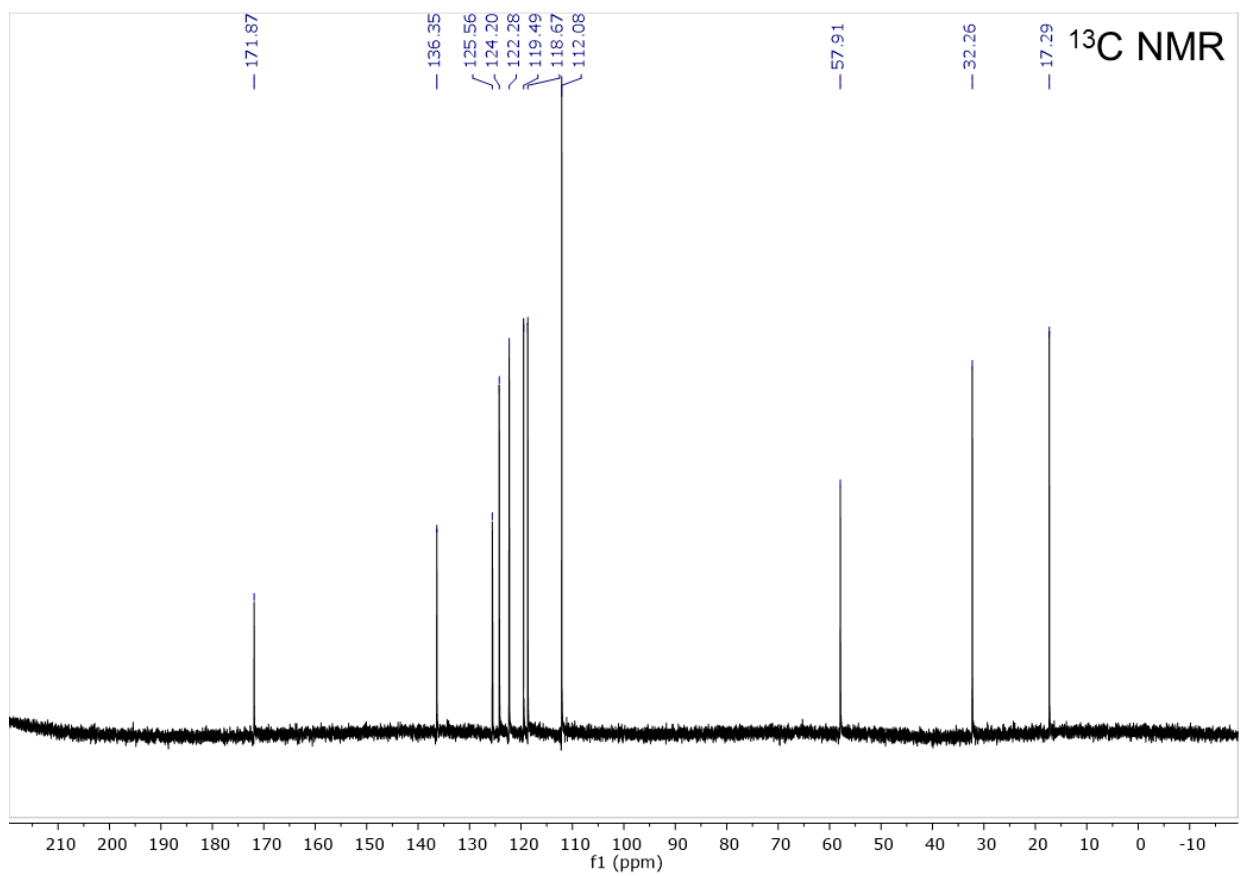
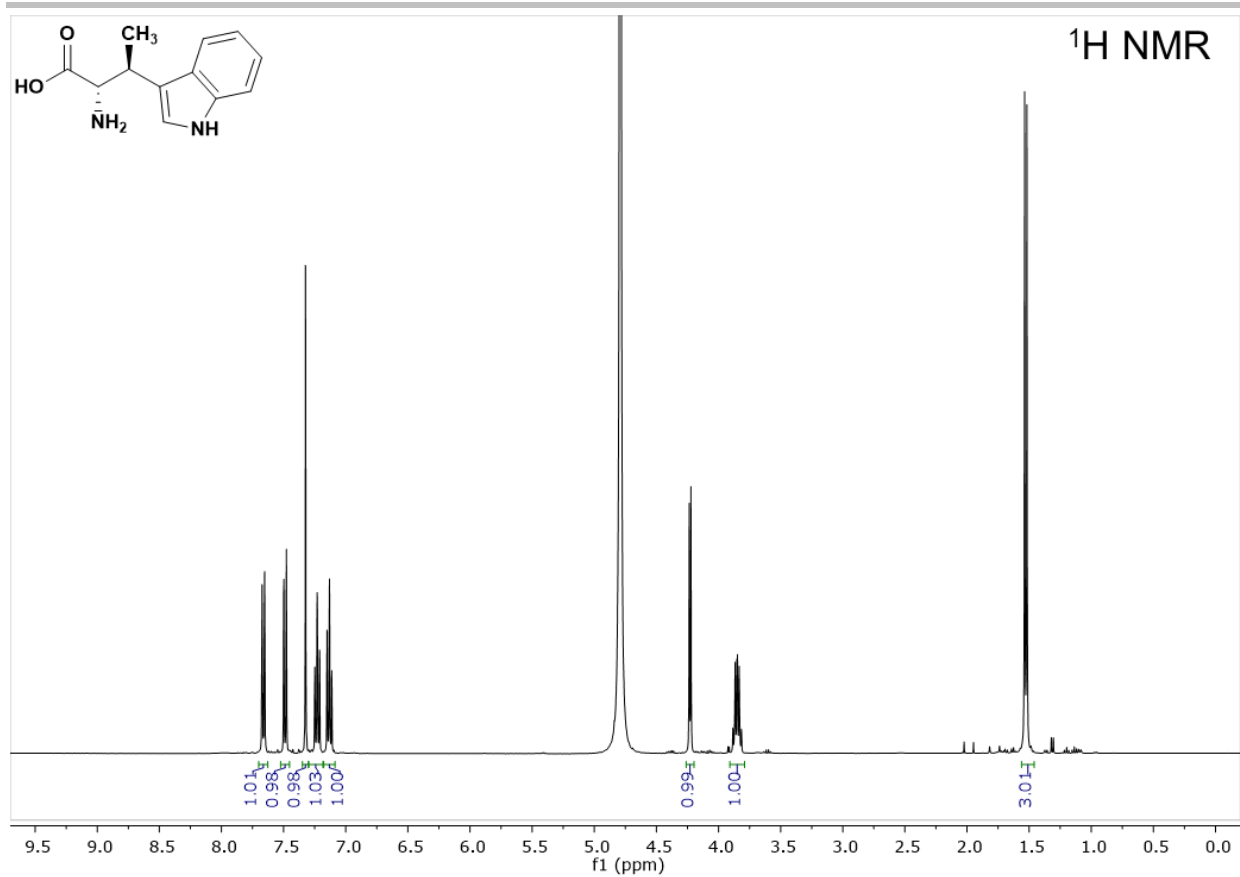
β -propyl-5-fluorotryptophan. ^1H NMR (400 MHz, D_2O) δ 7.42 (dd, J = 8.9, 4.6 Hz, 1H), 7.33 (d, J = 10.4 Hz, 2H), 7.00 (td, J = 9.3, 2.5 Hz, 1H), 4.21 (d, J = 5.6 Hz, 1H), 3.64 (dt, J = 10.8, 5.5 Hz, 1H), 2.00 – 1.71 (m, 2H), 1.22 (dddt, J = 13.6, 11.7, 9.4, 7.0 Hz, 2H), 0.82 (t, J = 7.3 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.13, 158.62, 156.31, 132.91, 126.51, 126.41, 126.37, 112.85, 112.75, 110.56, 110.30, 110.18, 110.14, 103.39, 103.15, 57.28, 37.30, 33.20, 25.06, 24.68, 21.95, 21.53, 19.98, 12.83. **HRMS** (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{13}\text{FN}_2\text{O}_2\text{H}_3$ requires 266.1390, observed 266.1384.

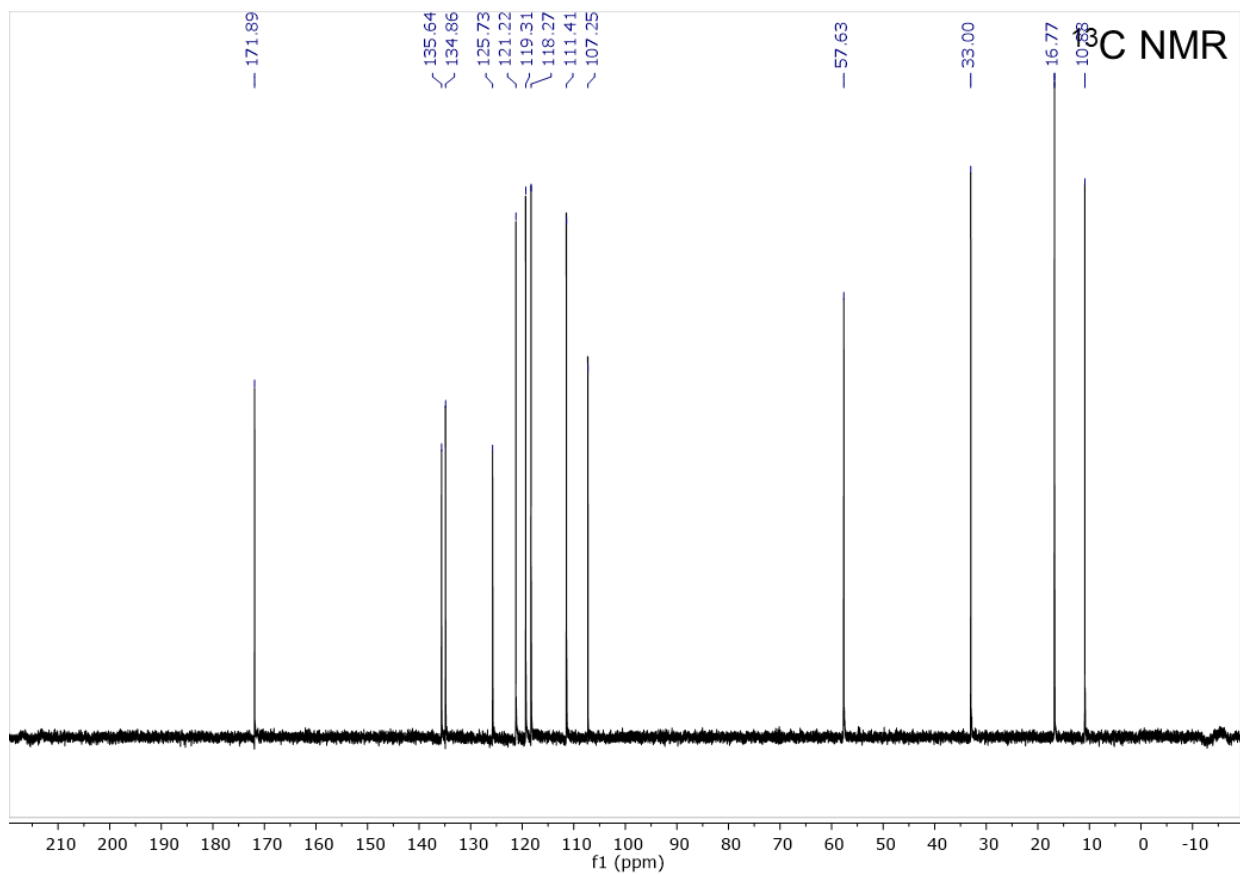
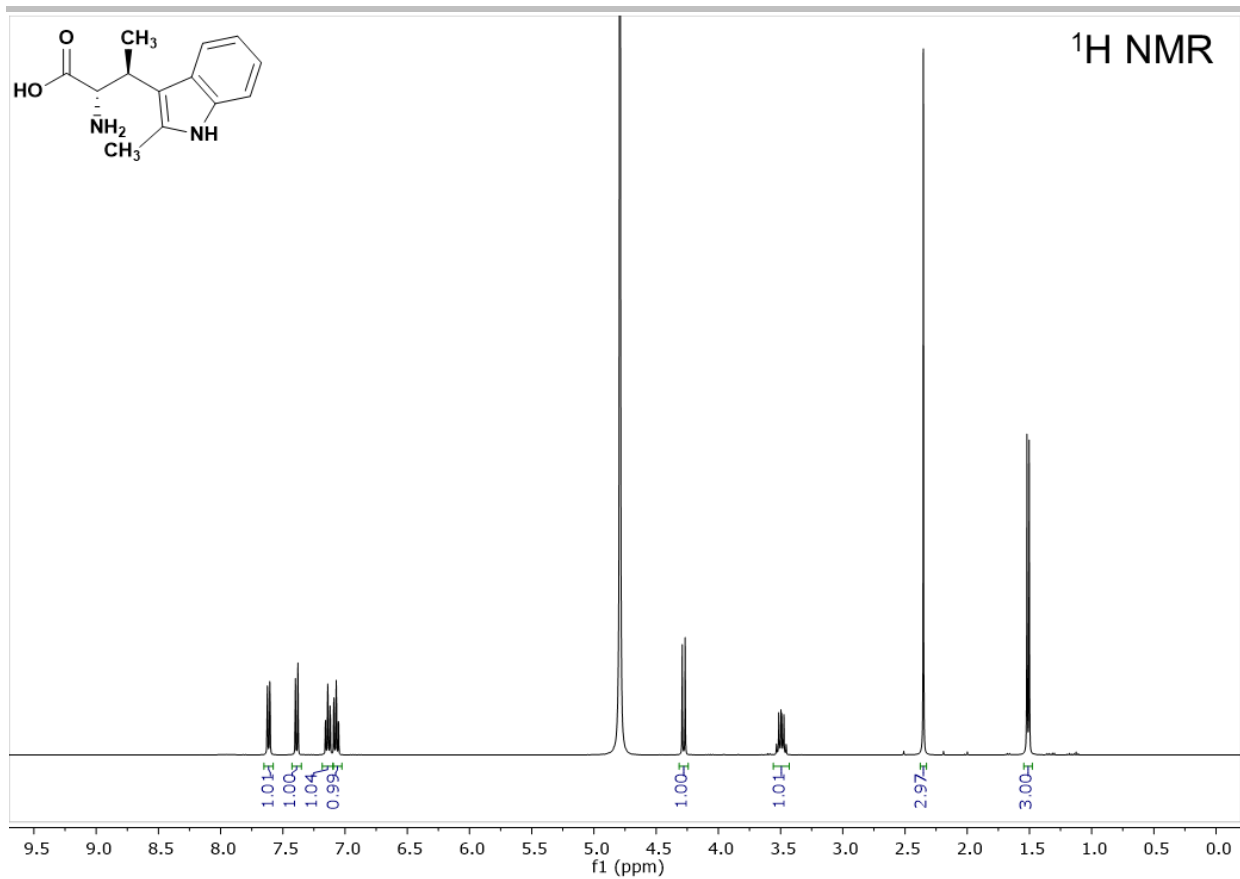


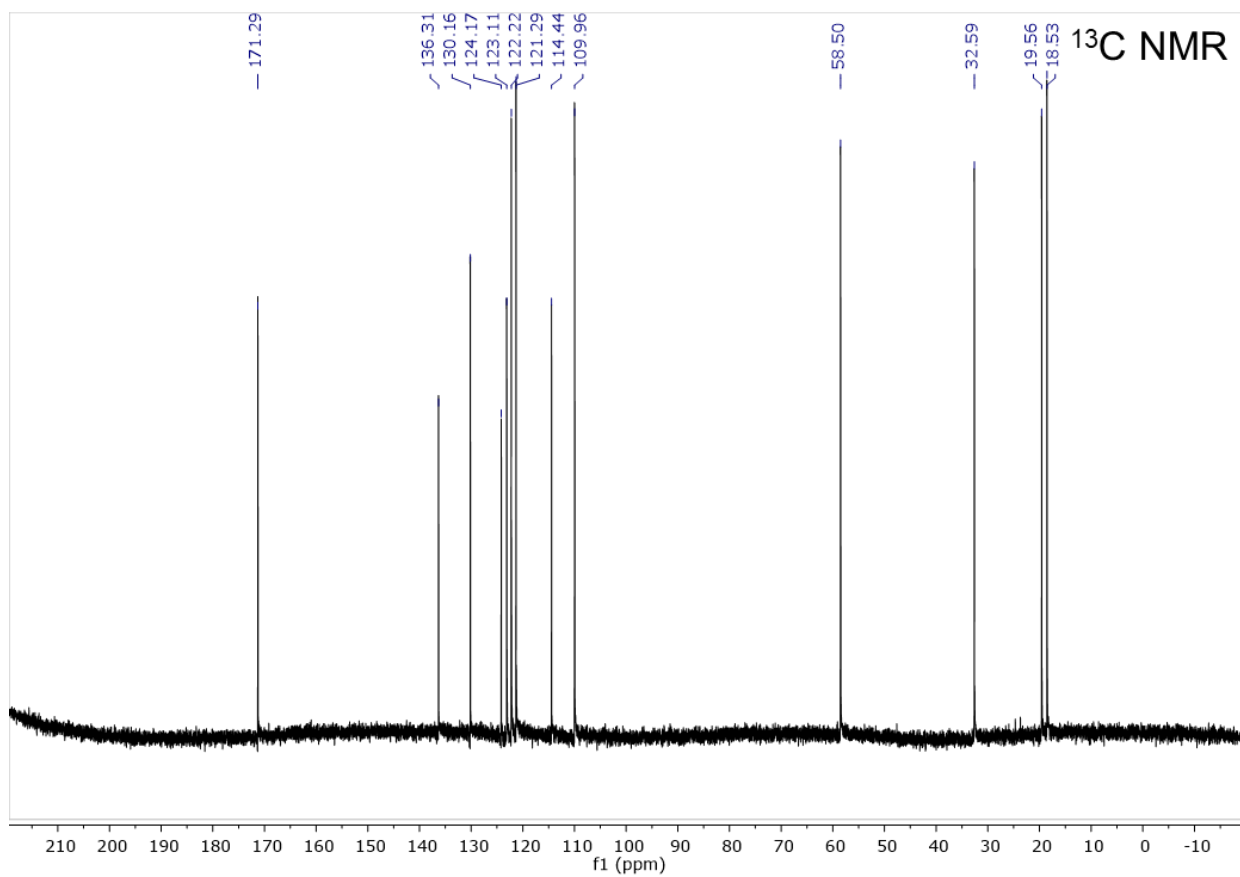
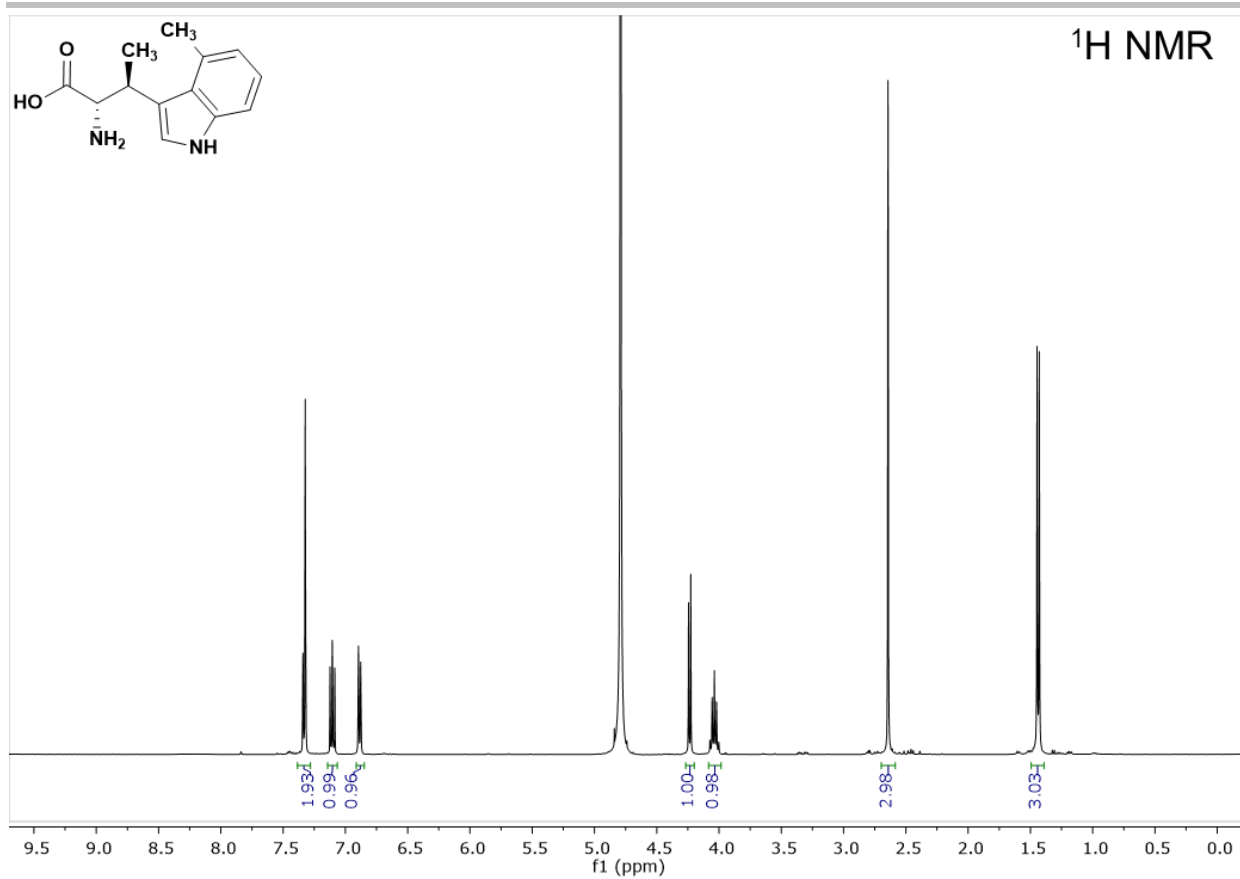
β -propyl-6-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.58 (d, J = 8.2 Hz, 1H), 7.32 (td, J = 1.4, 0.7 Hz, 1H), 7.23 (s, 1H), 7.00 (dd, J = 8.3, 1.4 Hz, 1H), 4.14 (dd, J = 6.2, 1.7 Hz, 1H), 3.70 – 3.56 (m, 1H), 2.41 (s, 3H), 1.96 (dddd, J = 13.9, 10.8, 8.6, 5.6 Hz, 1H), 1.87 – 1.74 (m, 1H), 1.31 – 1.11 (m, 2H), 0.96 – 0.79 (m, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.80, 136.84, 132.46, 124.06, 123.94, 121.03, 118.53, 111.63, 110.10, 57.89, 37.63, 33.33, 20.50, 19.95, 12.78. **HRMS** (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{15}\text{H}_{18}\text{N}_2\text{O}_2\text{H}_3$ requires 264.1791, observed 264.1800.

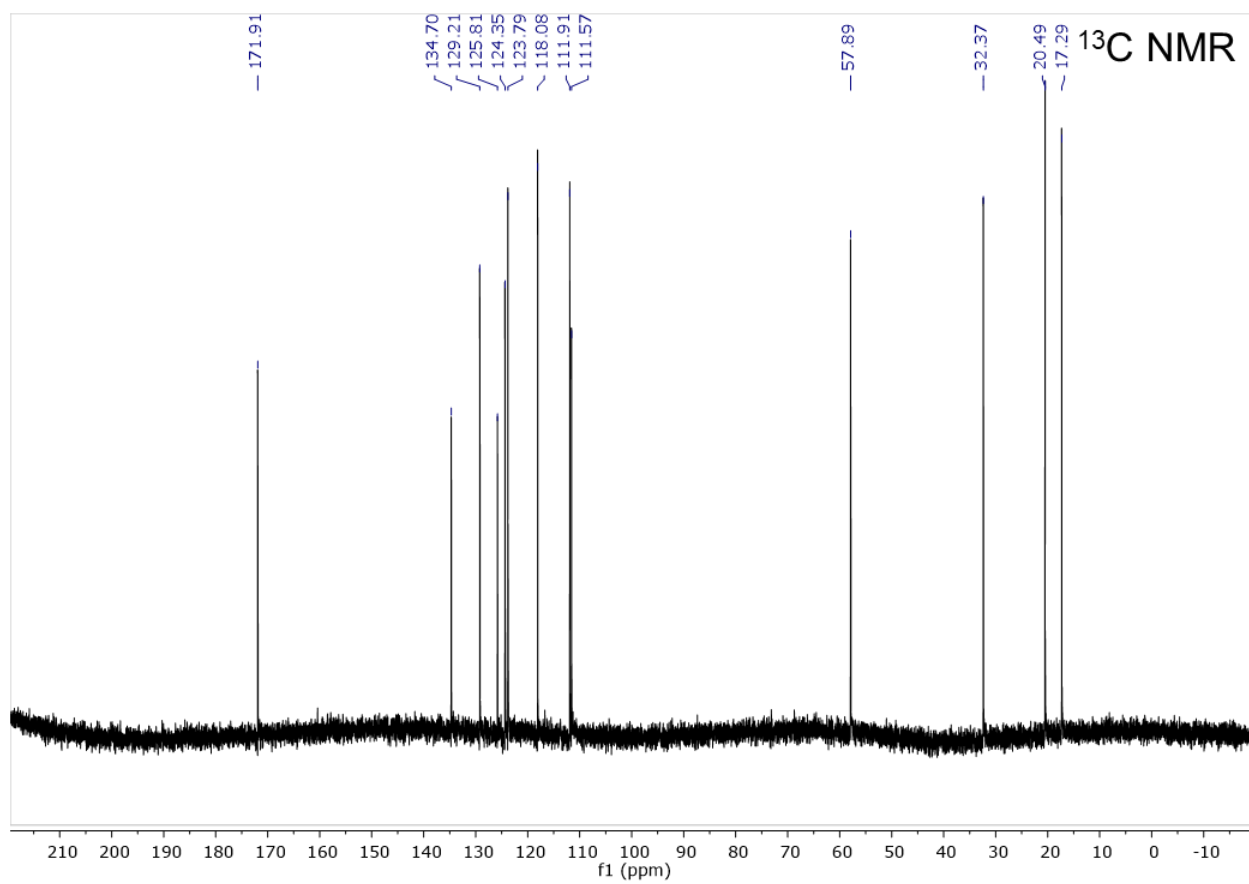
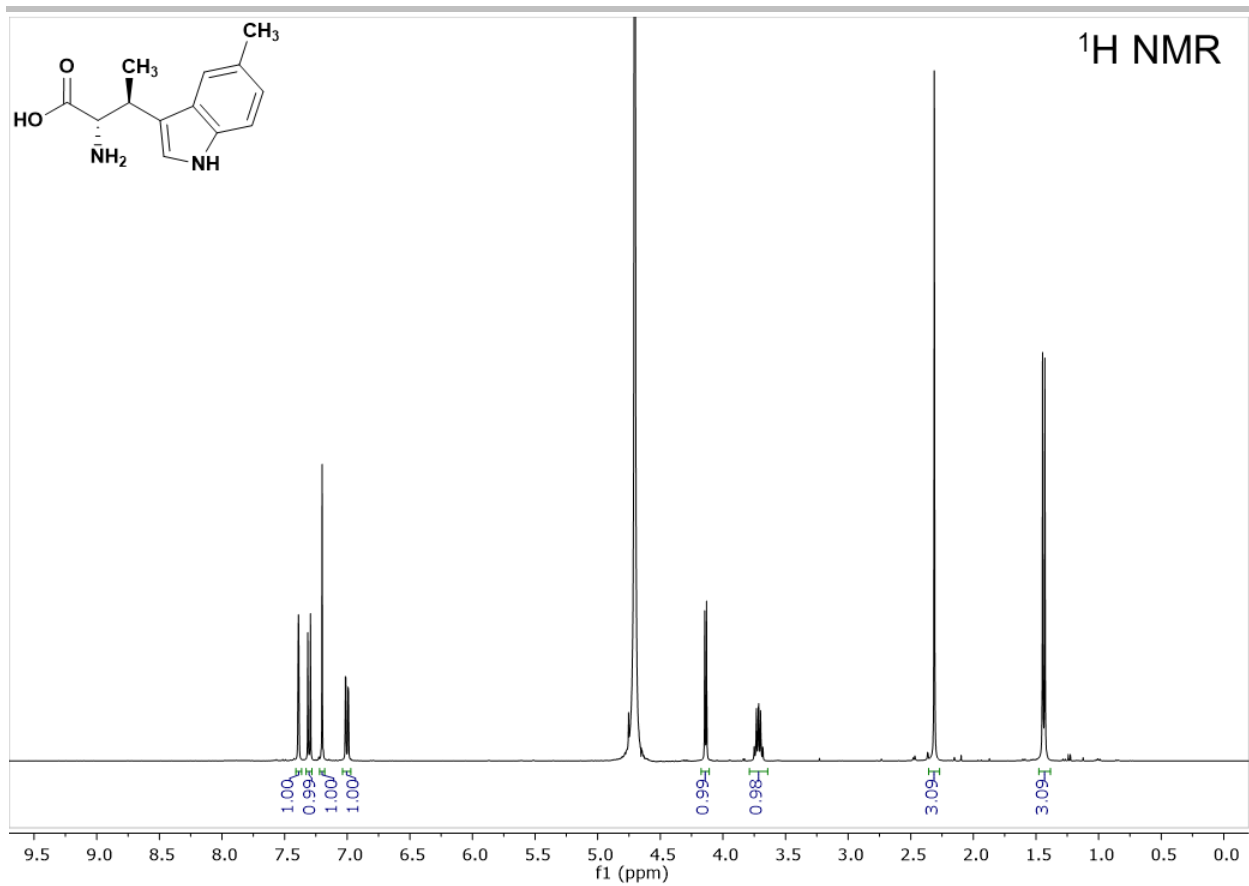


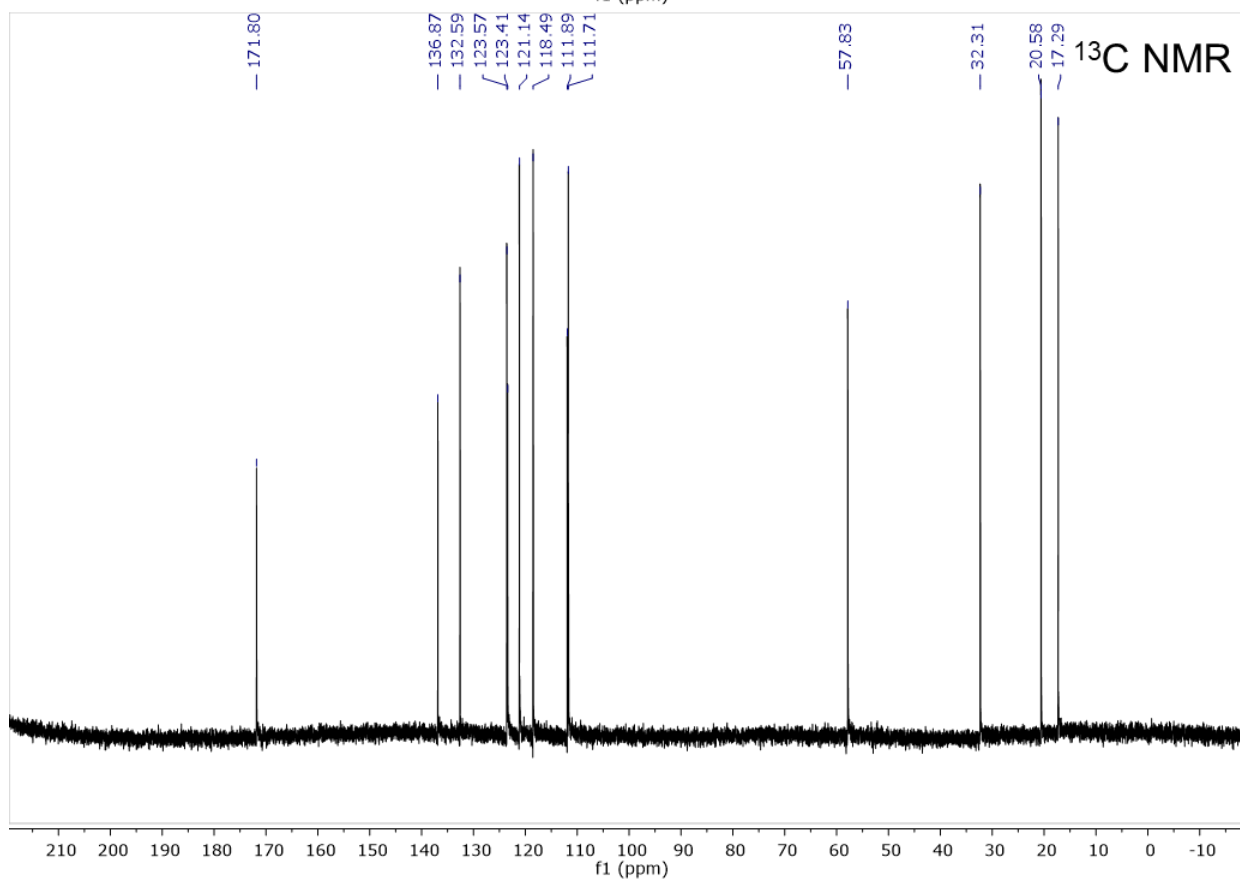
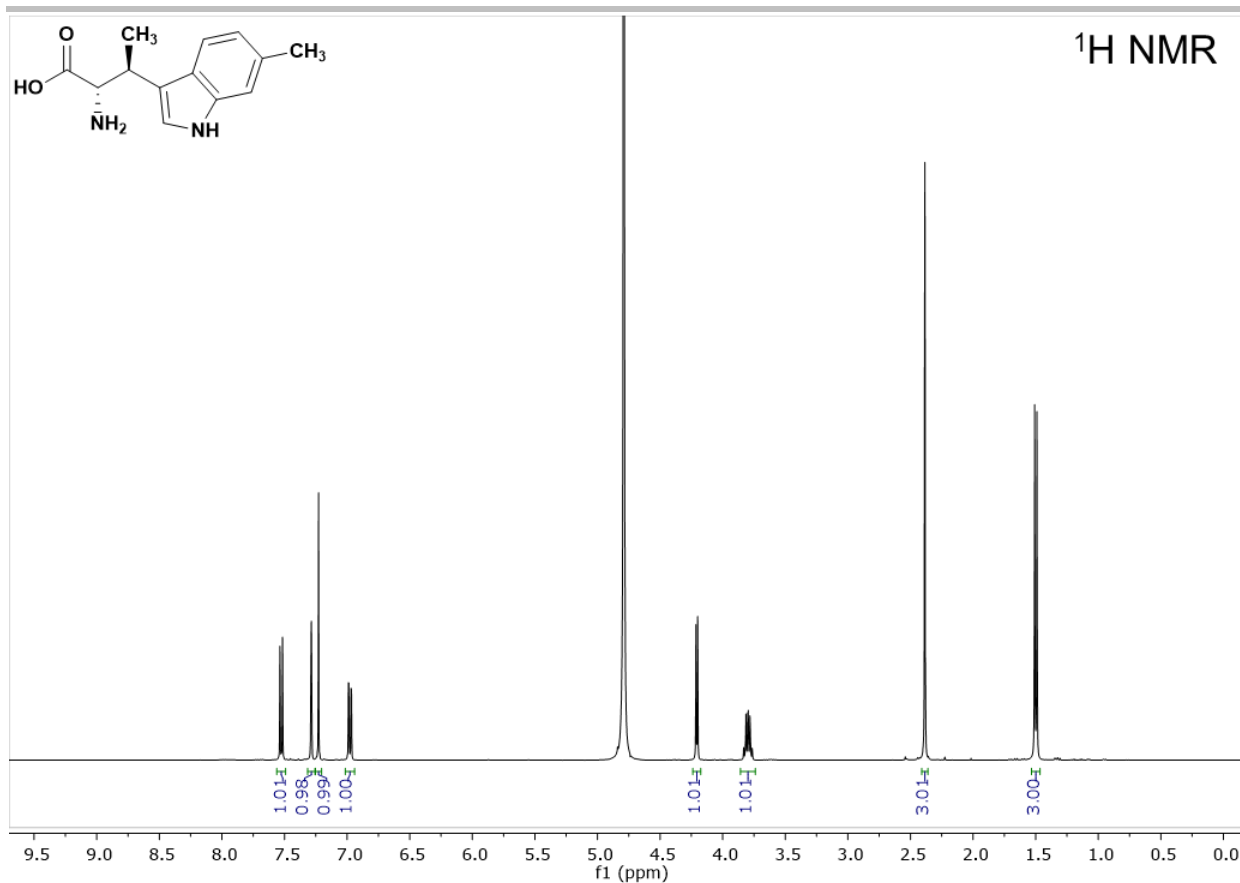
β -propyl-7-methyltryptophan. ^1H NMR (400 MHz, D_2O) δ 7.54 – 7.45 (m, 1H), 7.31 (s, 1H), 7.09 – 6.99 (m, 2H), 4.20 (d, J = 5.8 Hz, 1H), 3.65 (dt, J = 10.8, 5.4 Hz, 1H), 2.45 (s, 3H), 1.93 (dddd, J = 13.8, 10.7, 8.6, 5.6 Hz, 1H), 1.86 – 1.72 (m, 1H), 1.27 – 1.09 (m, 2H), 0.79 (t, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.18, 135.85, 125.90, 124.54, 122.43, 122.05, 119.80, 116.30, 110.47, 57.42, 37.50, 33.34, 20.00, 15.86, 12.83. **HRMS** (FAB+) (m/z) for $[\text{M}+\text{H}]^+$ $\text{C}_{15}\text{H}_{19}\text{N}_2\text{O}_2\text{H}_2$ requires 263.1729, observed 263.1723.

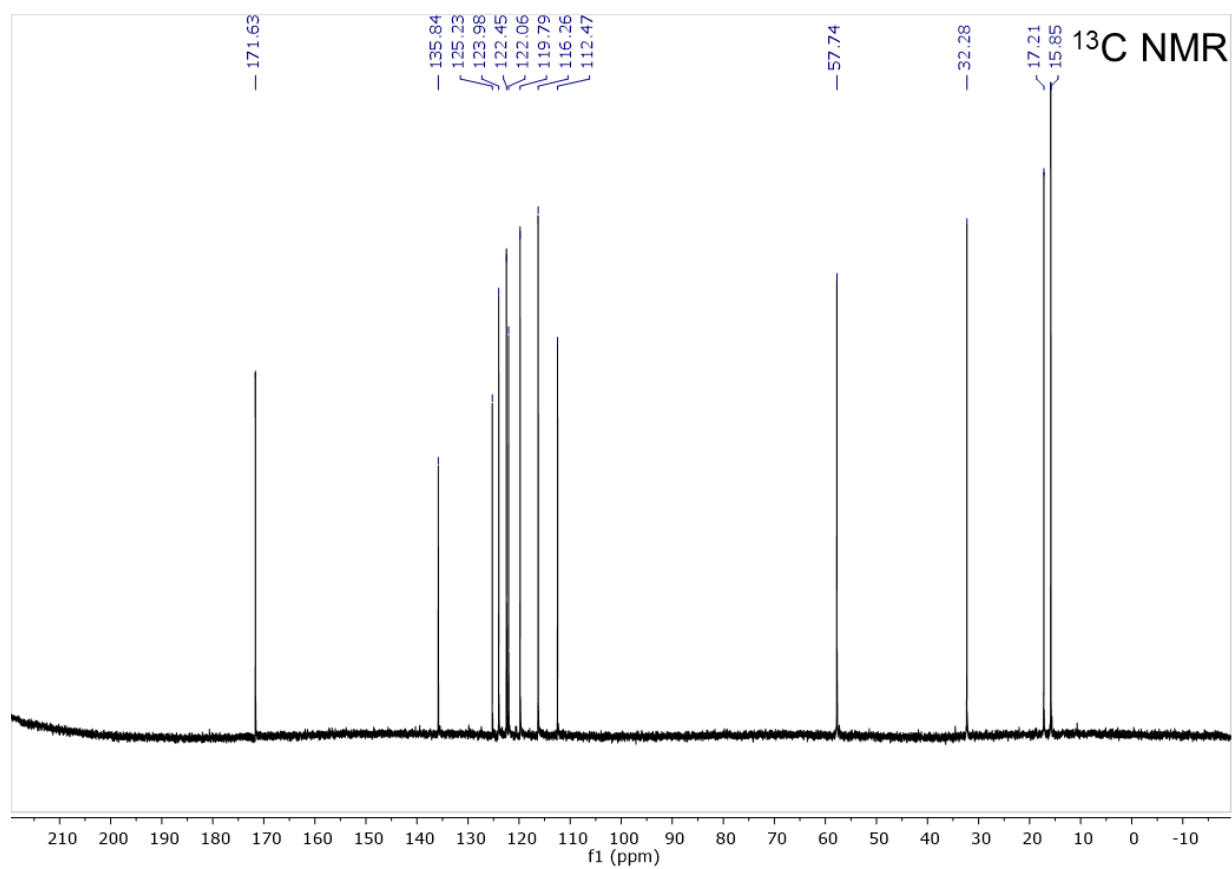
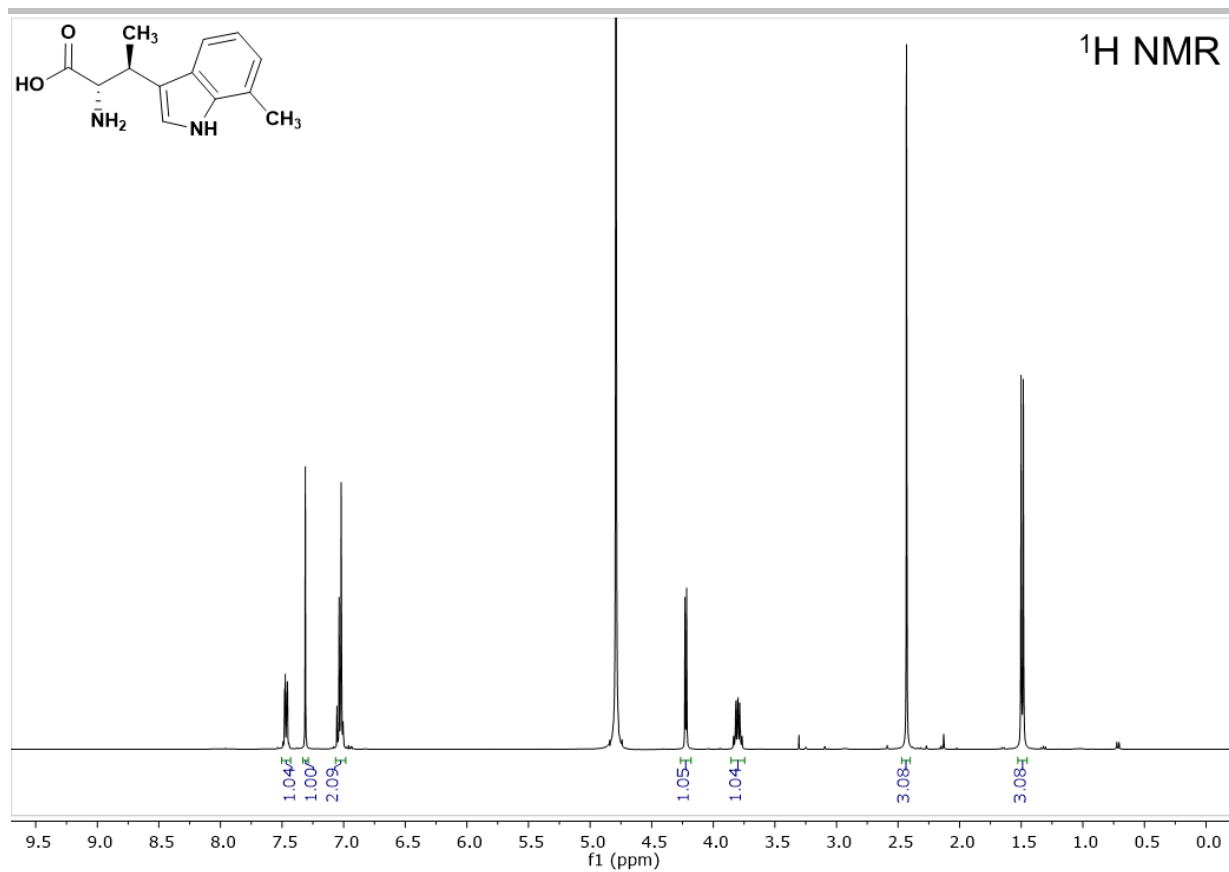


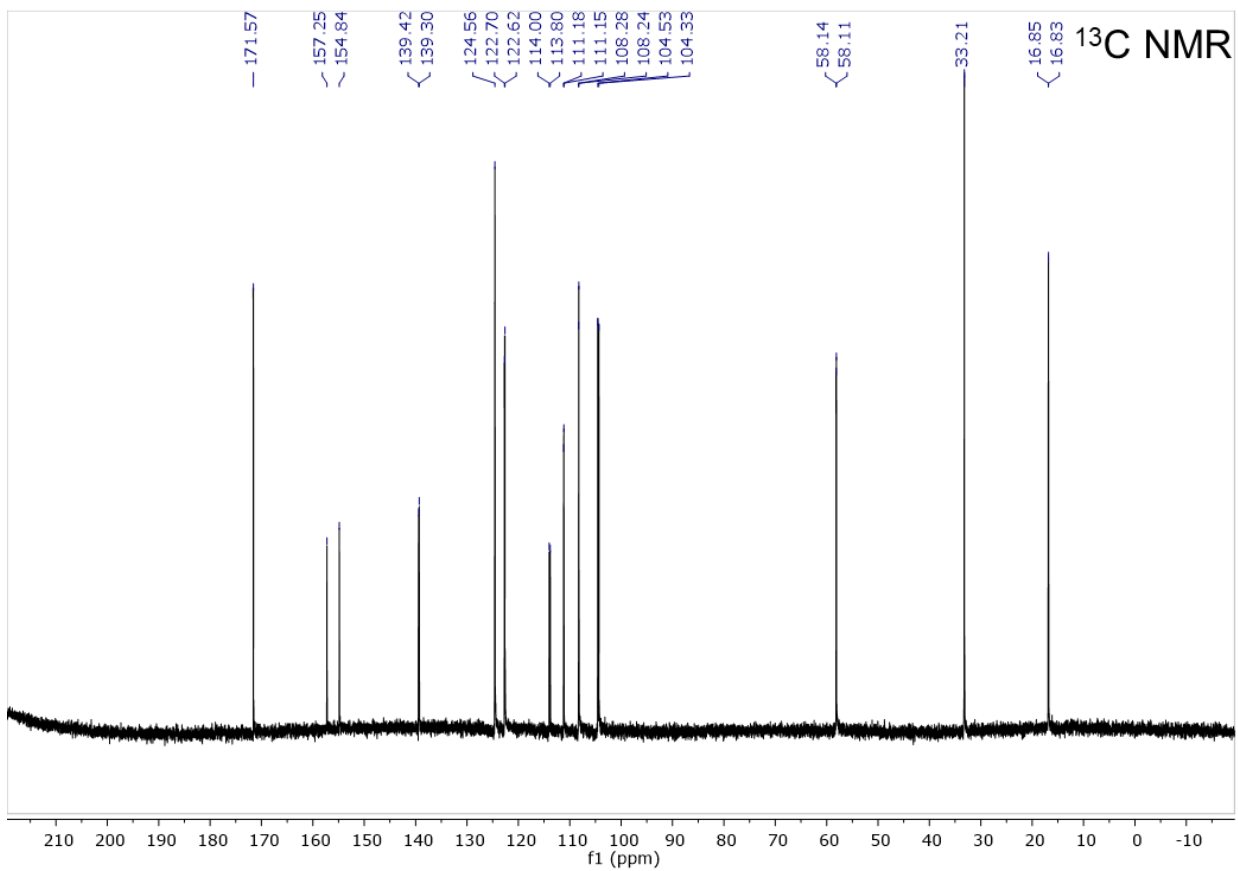
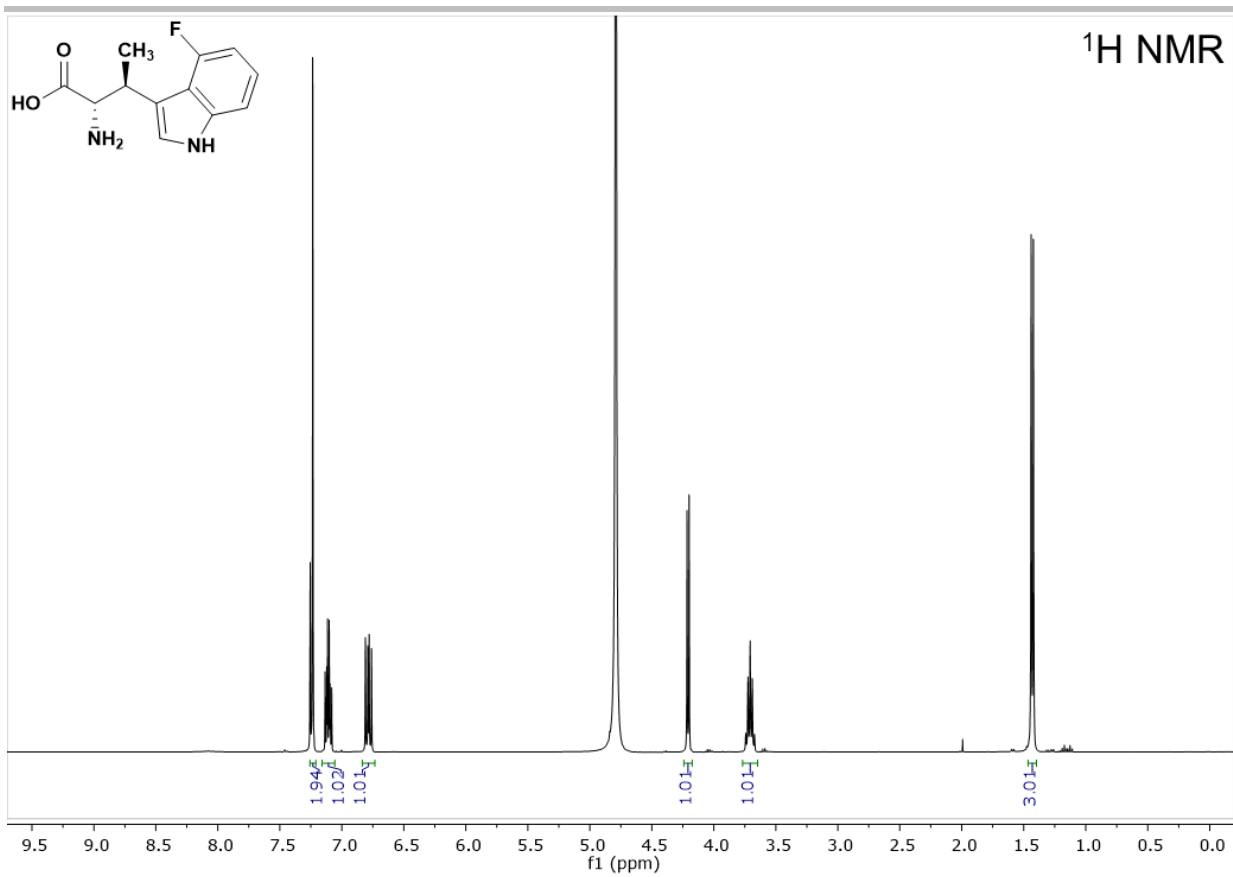


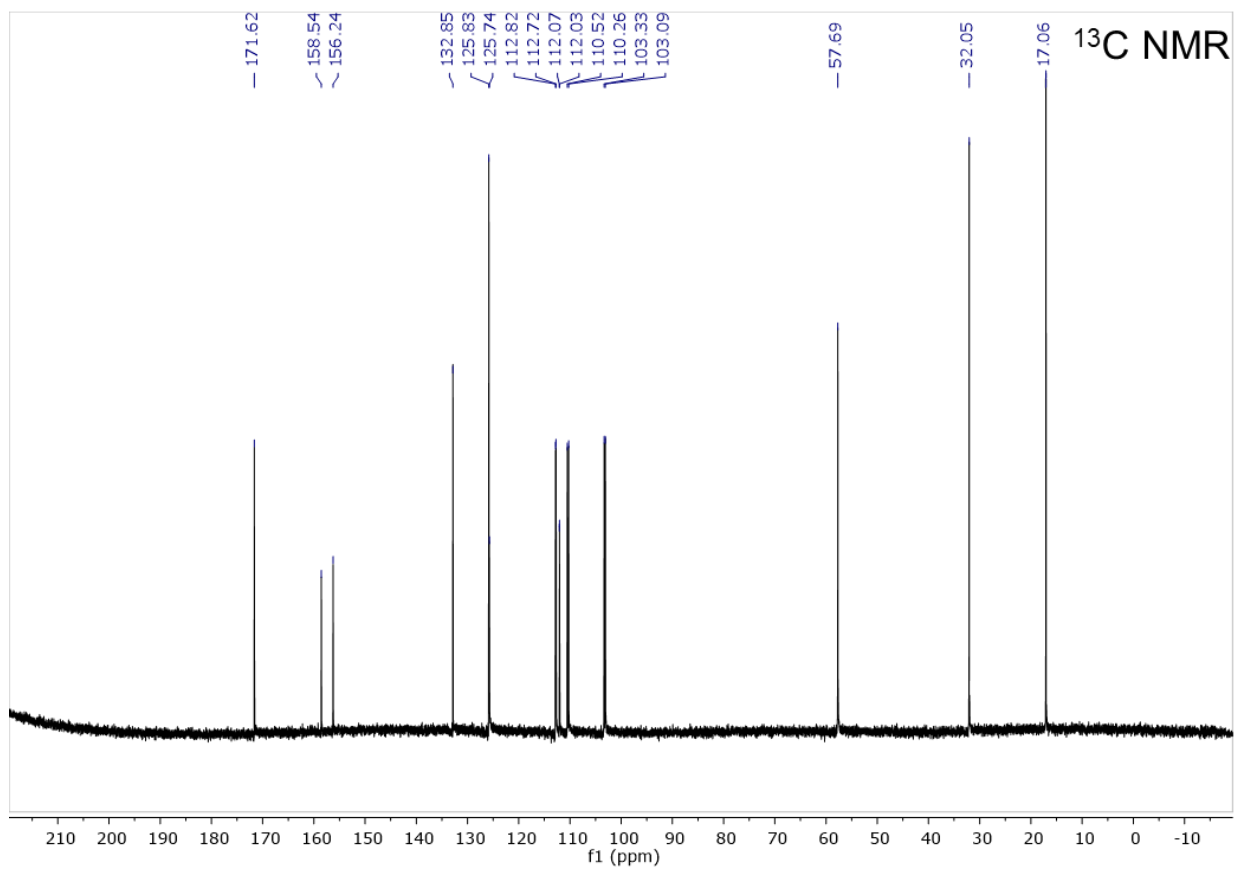
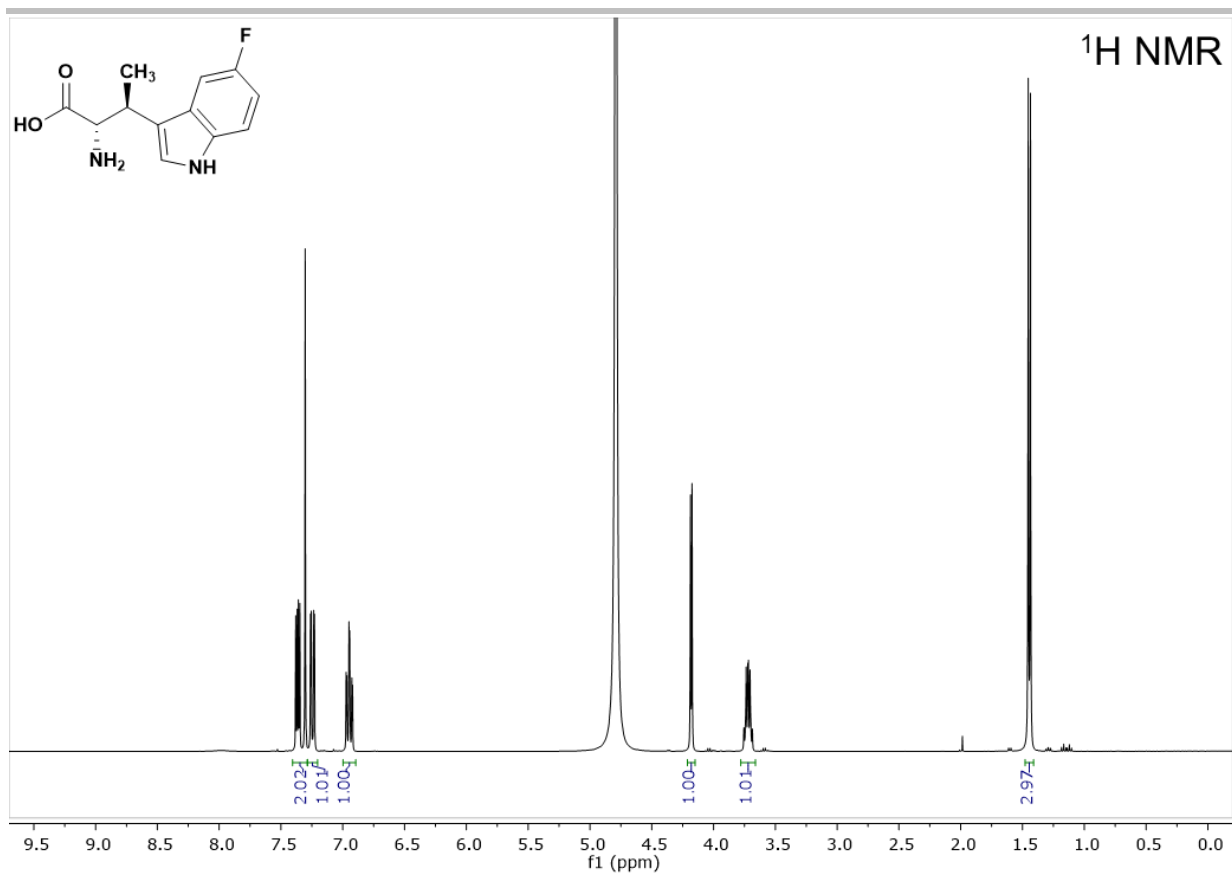


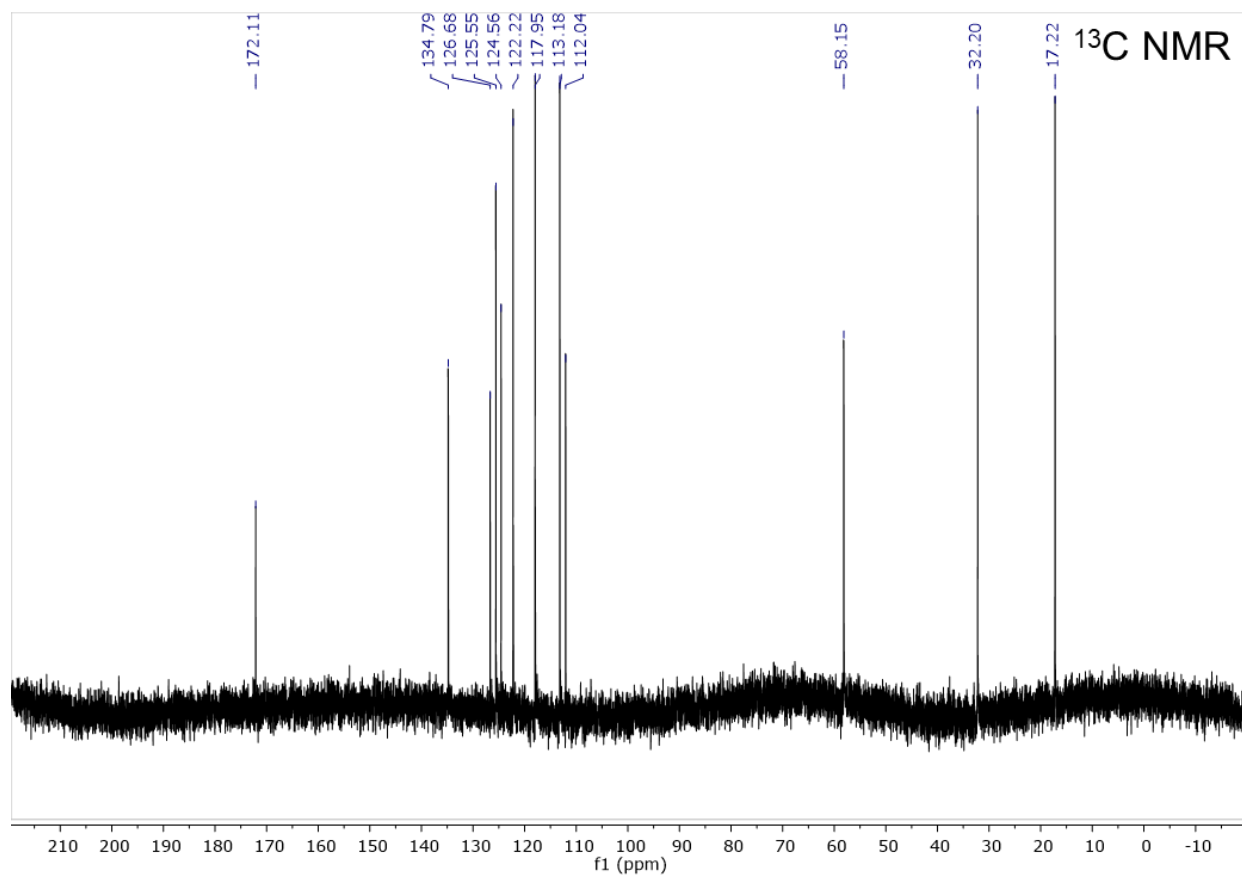
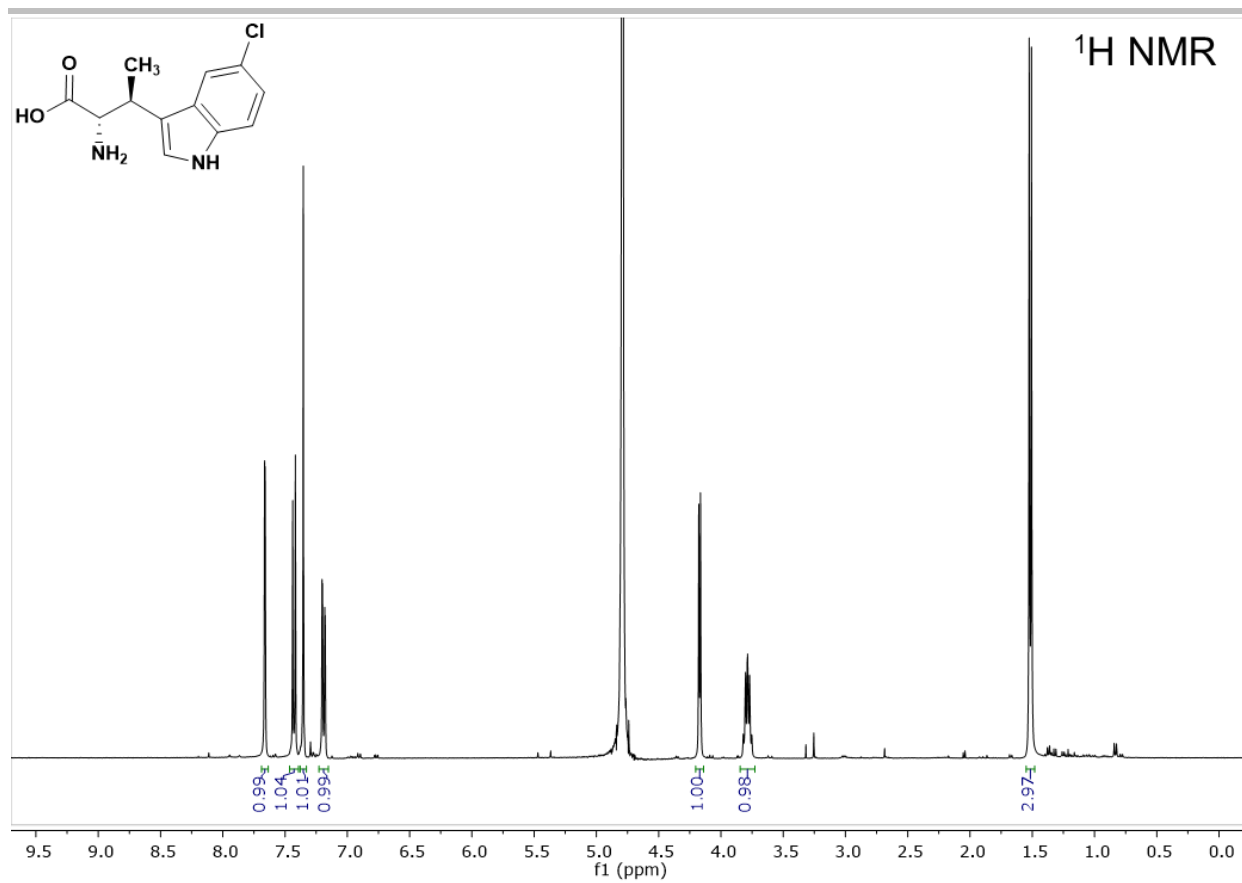


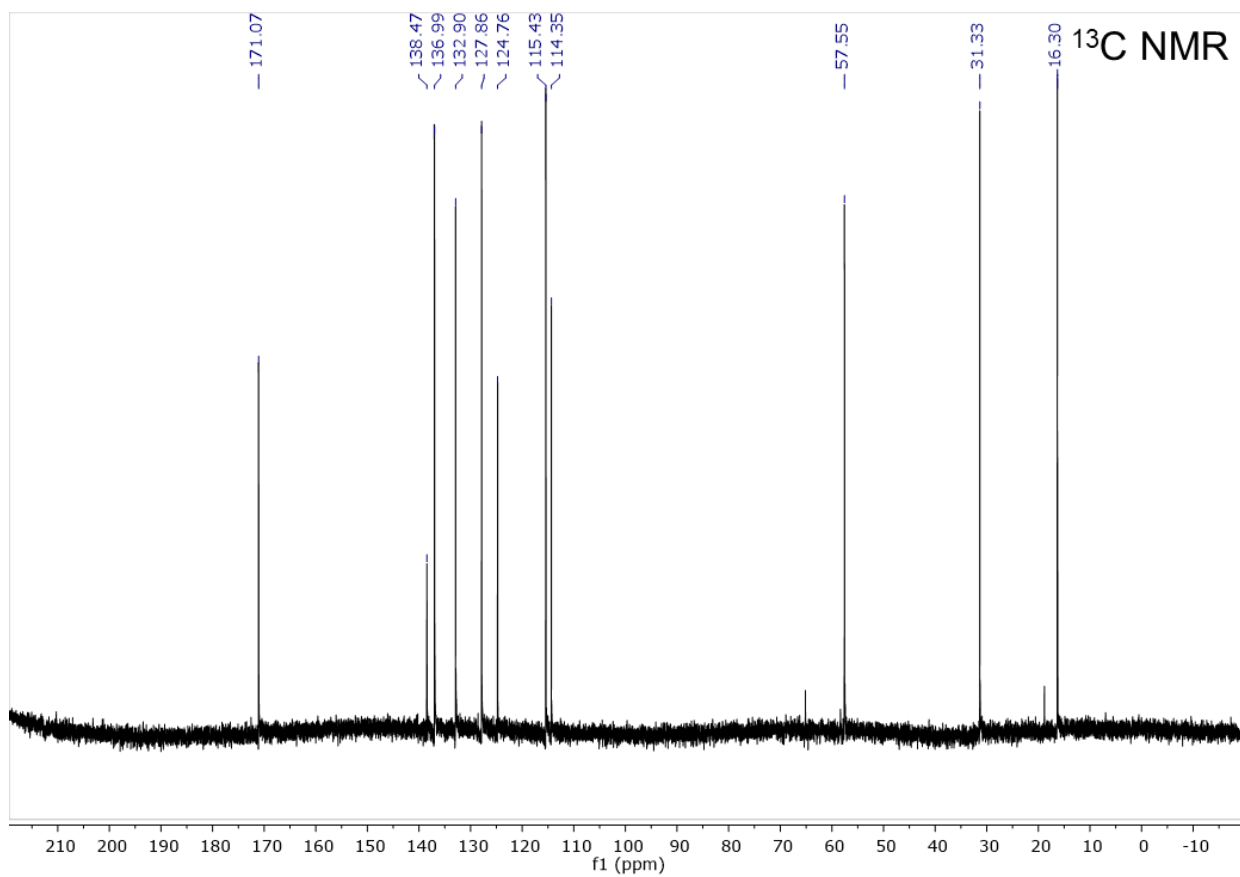
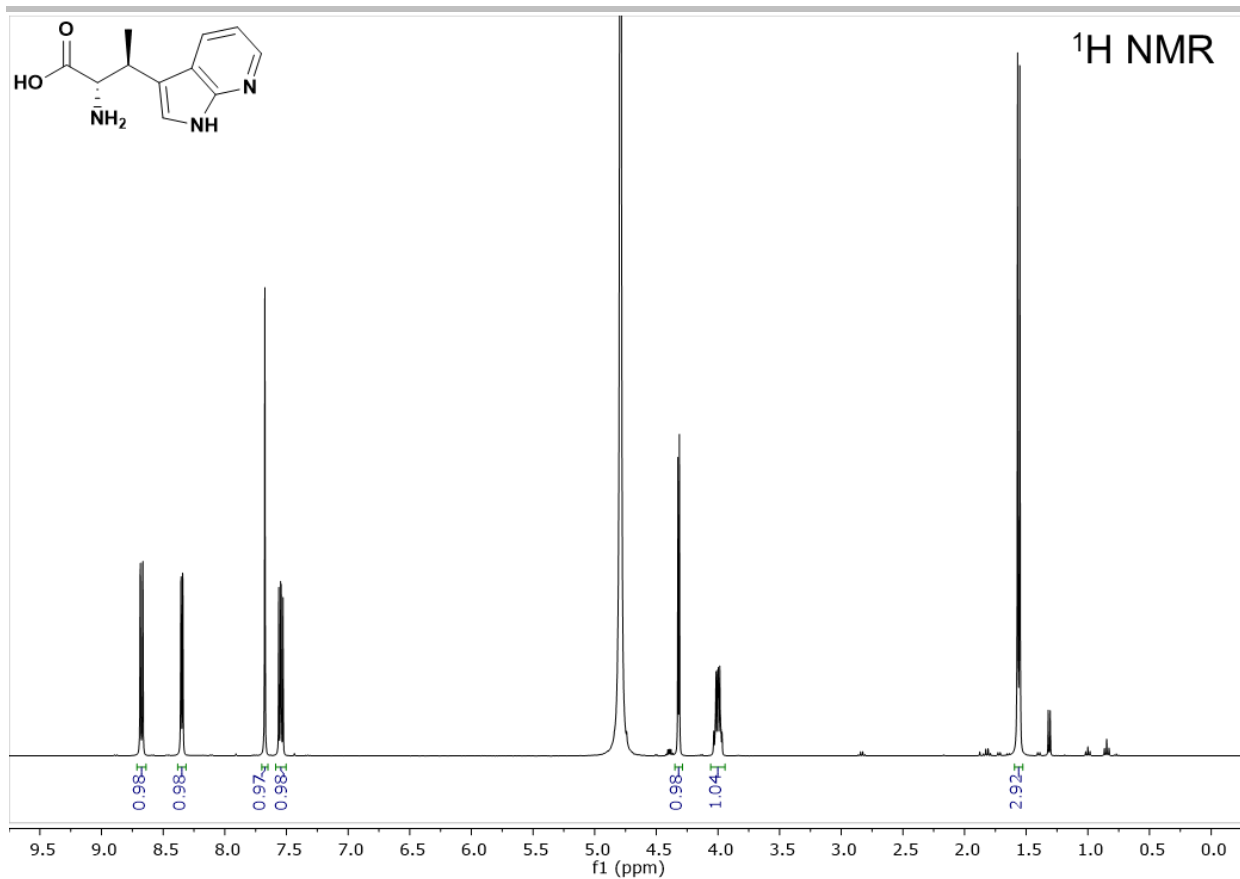


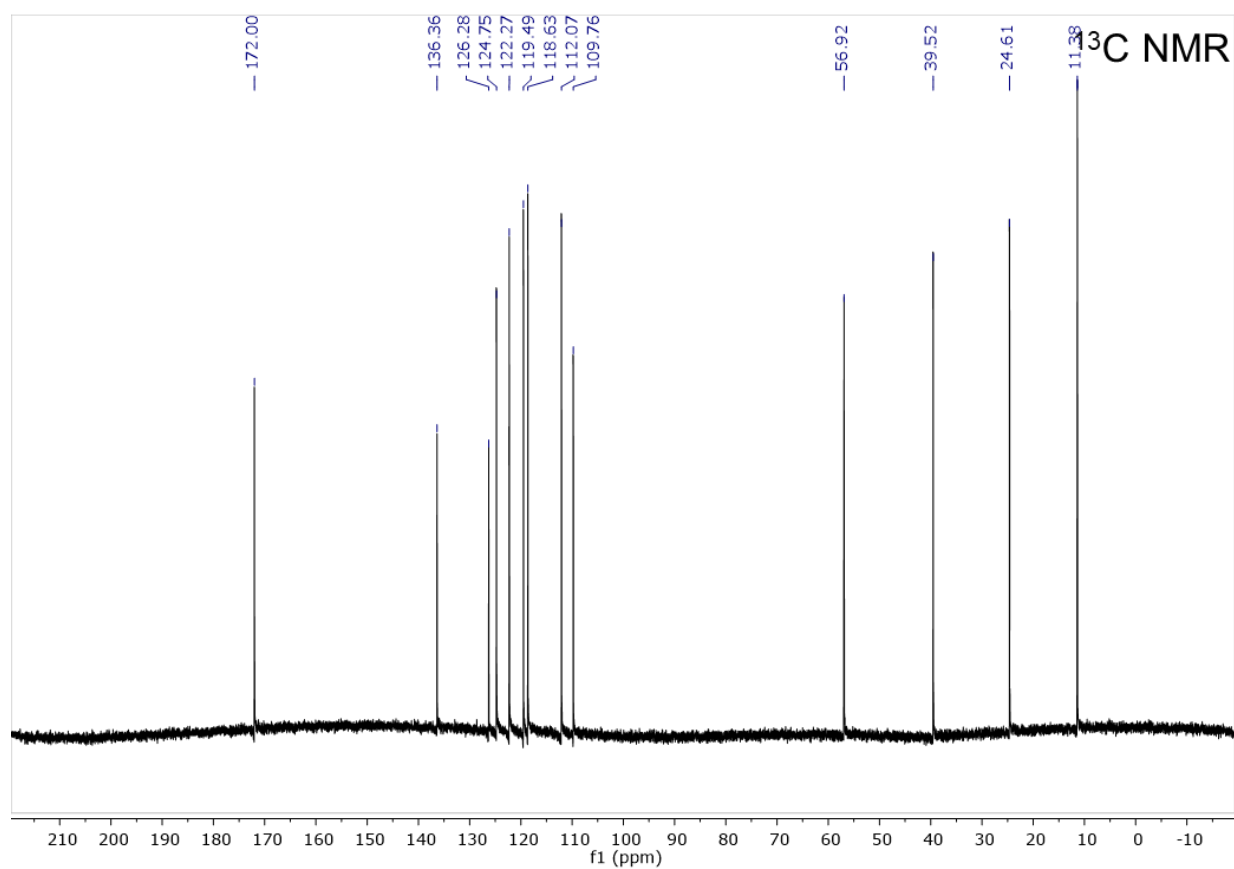
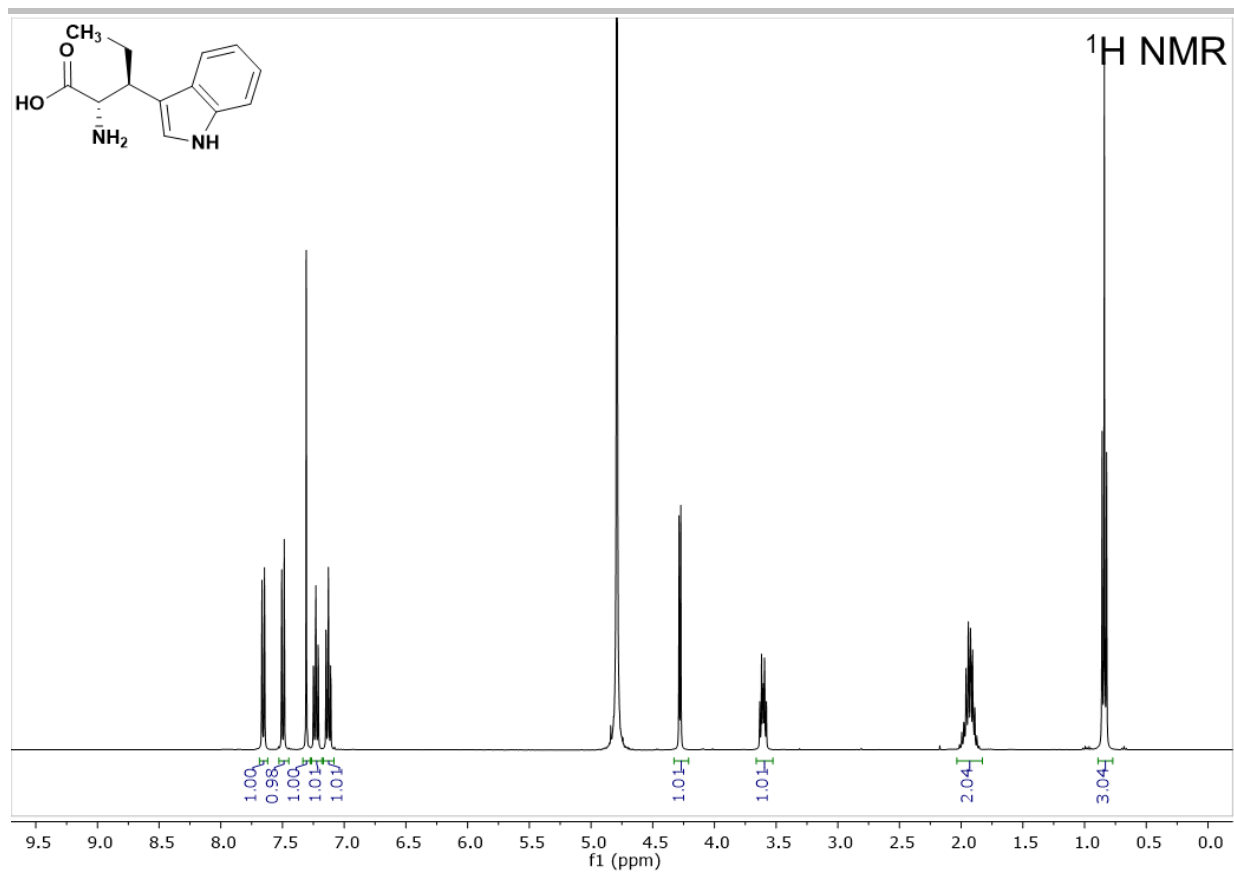


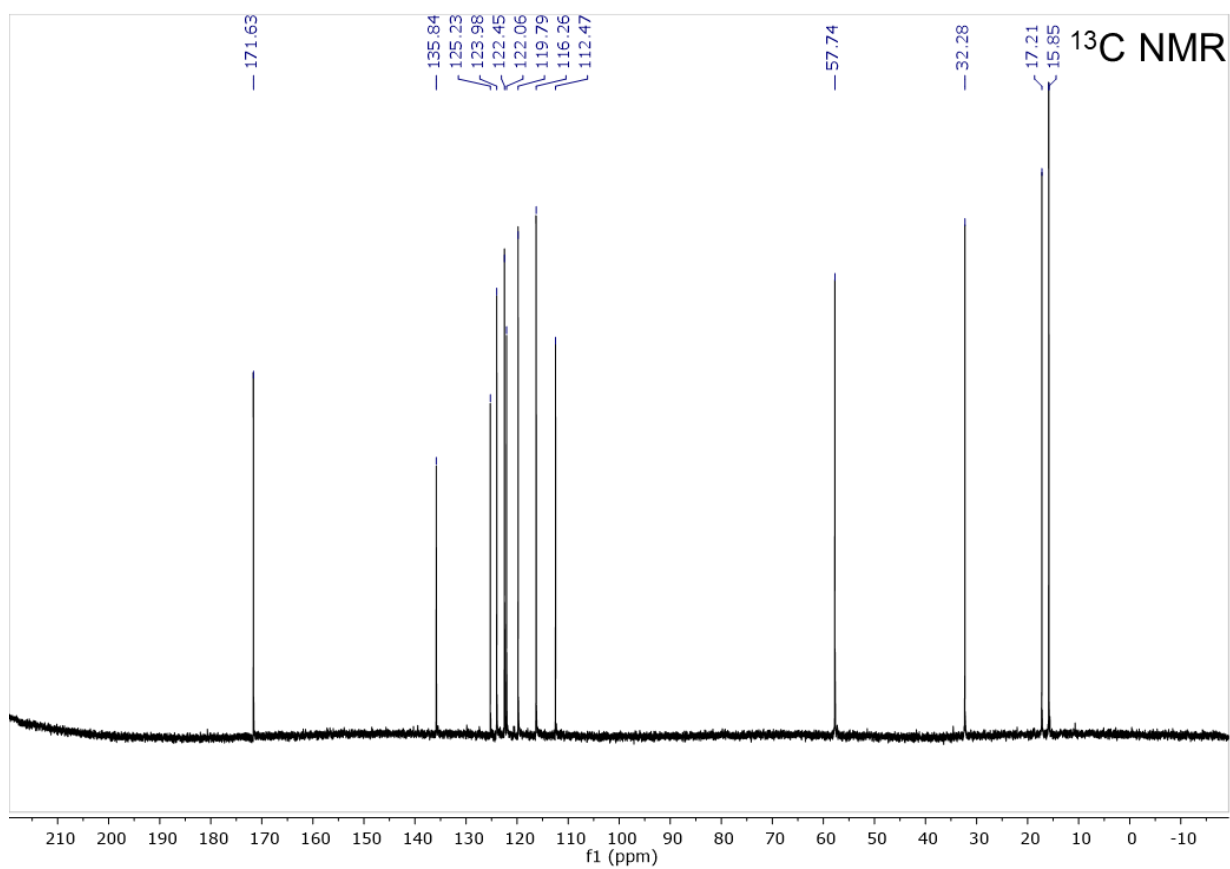
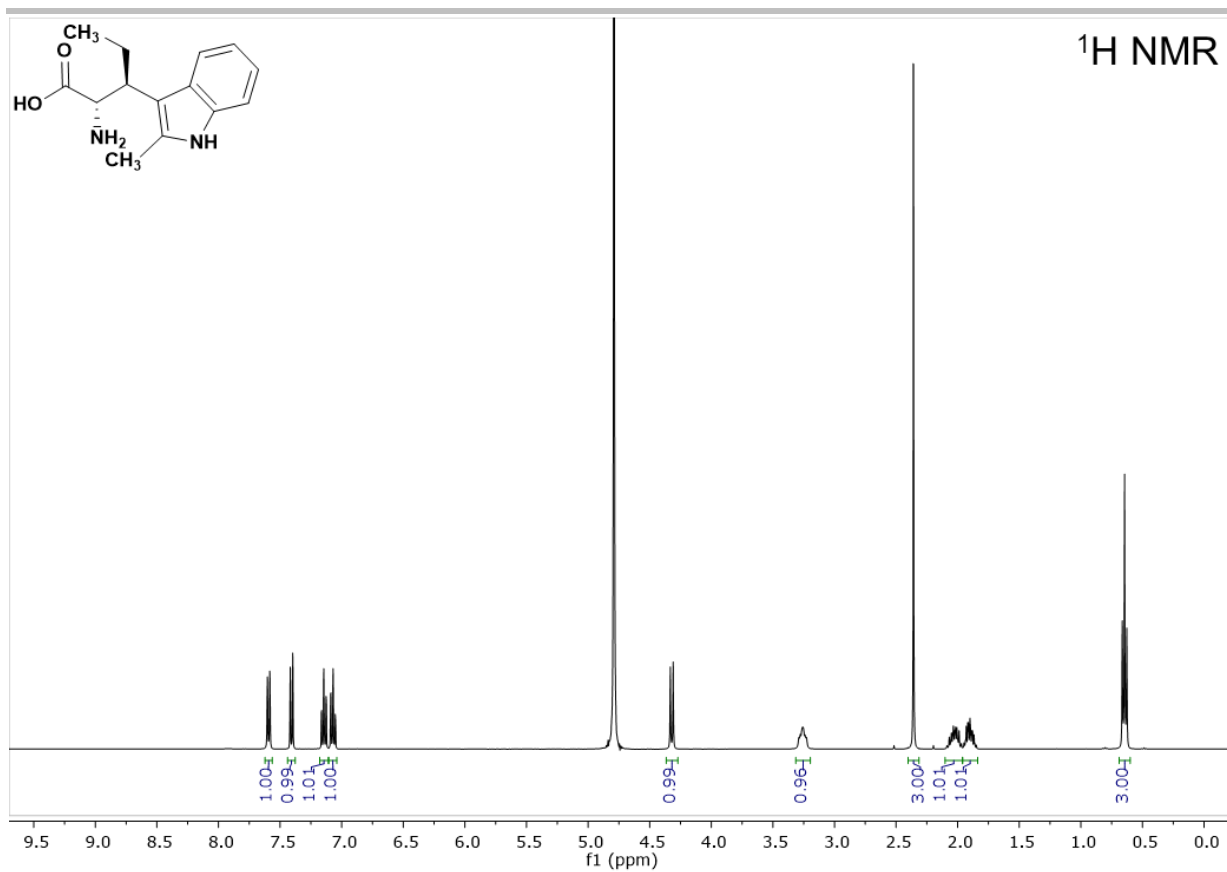


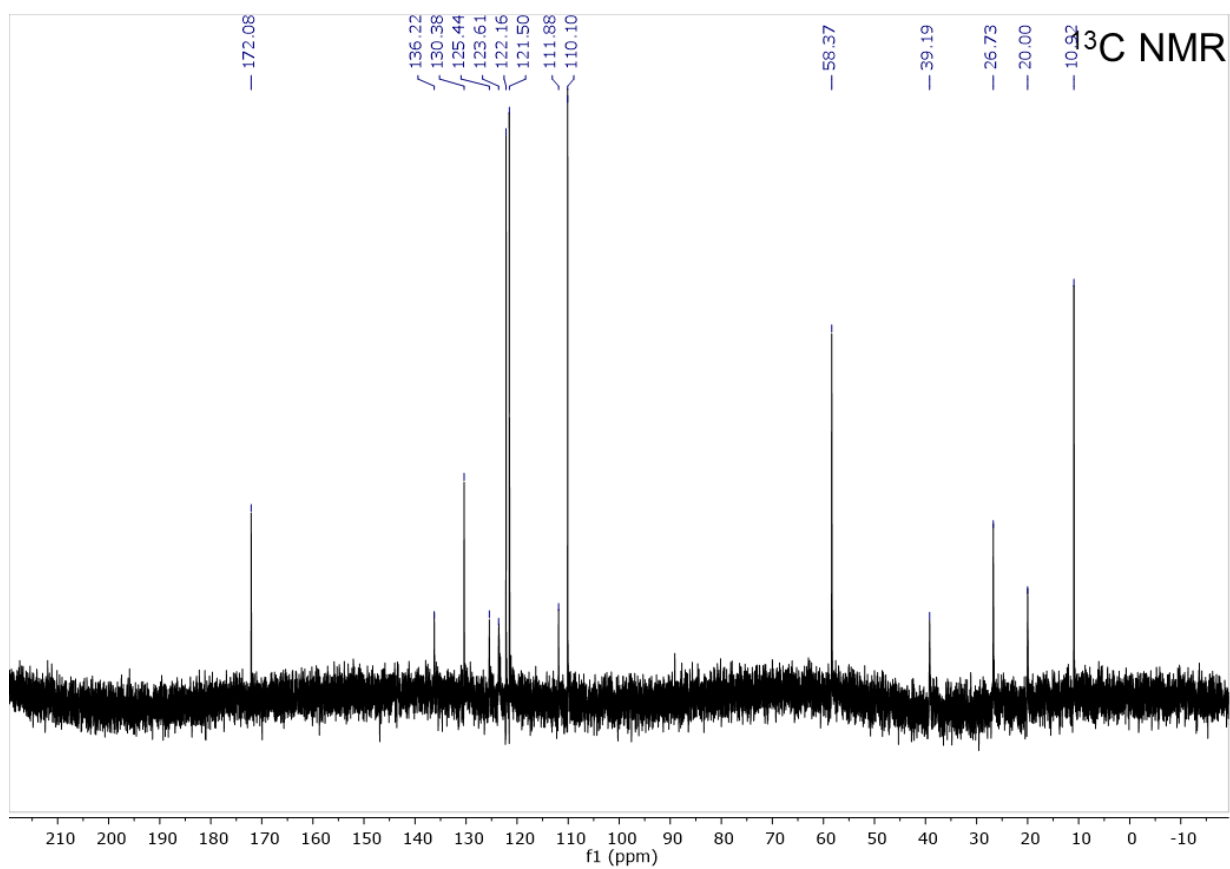
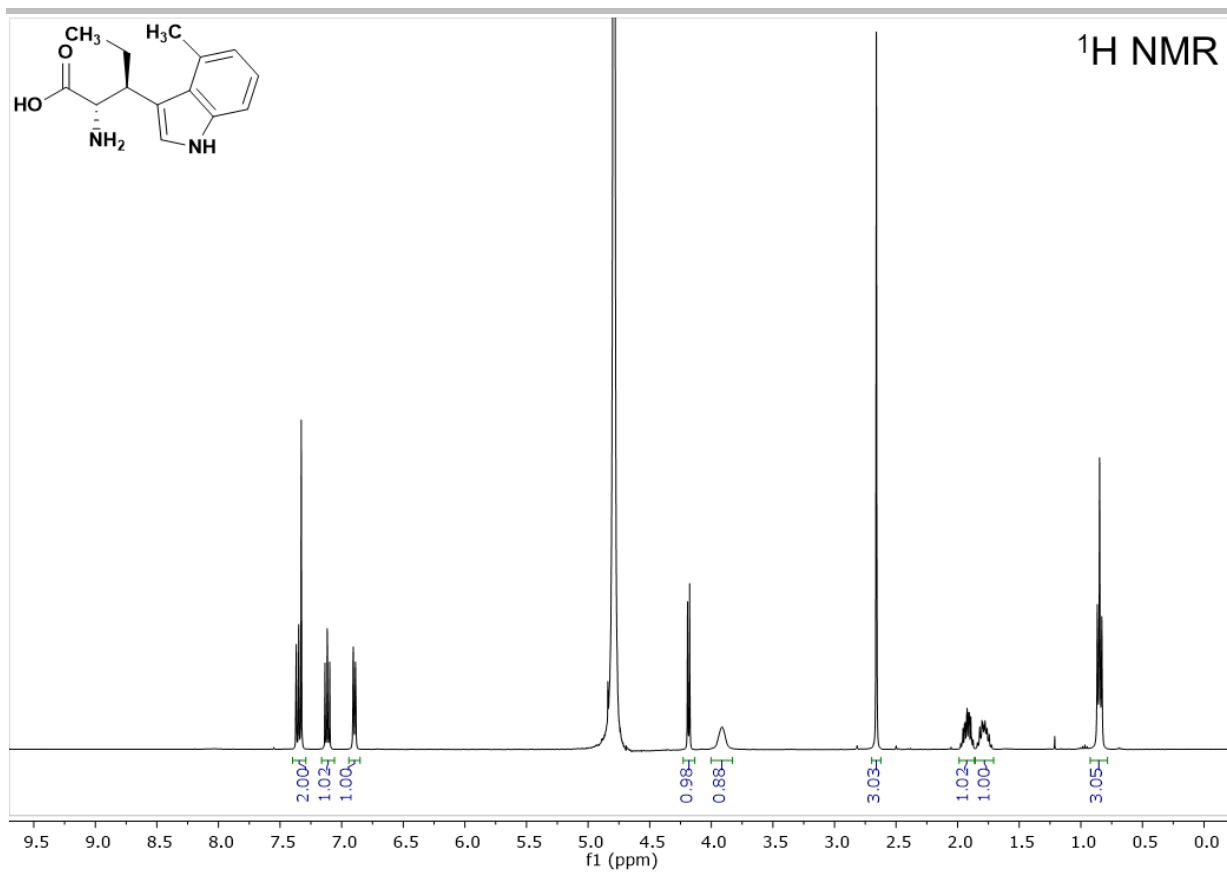


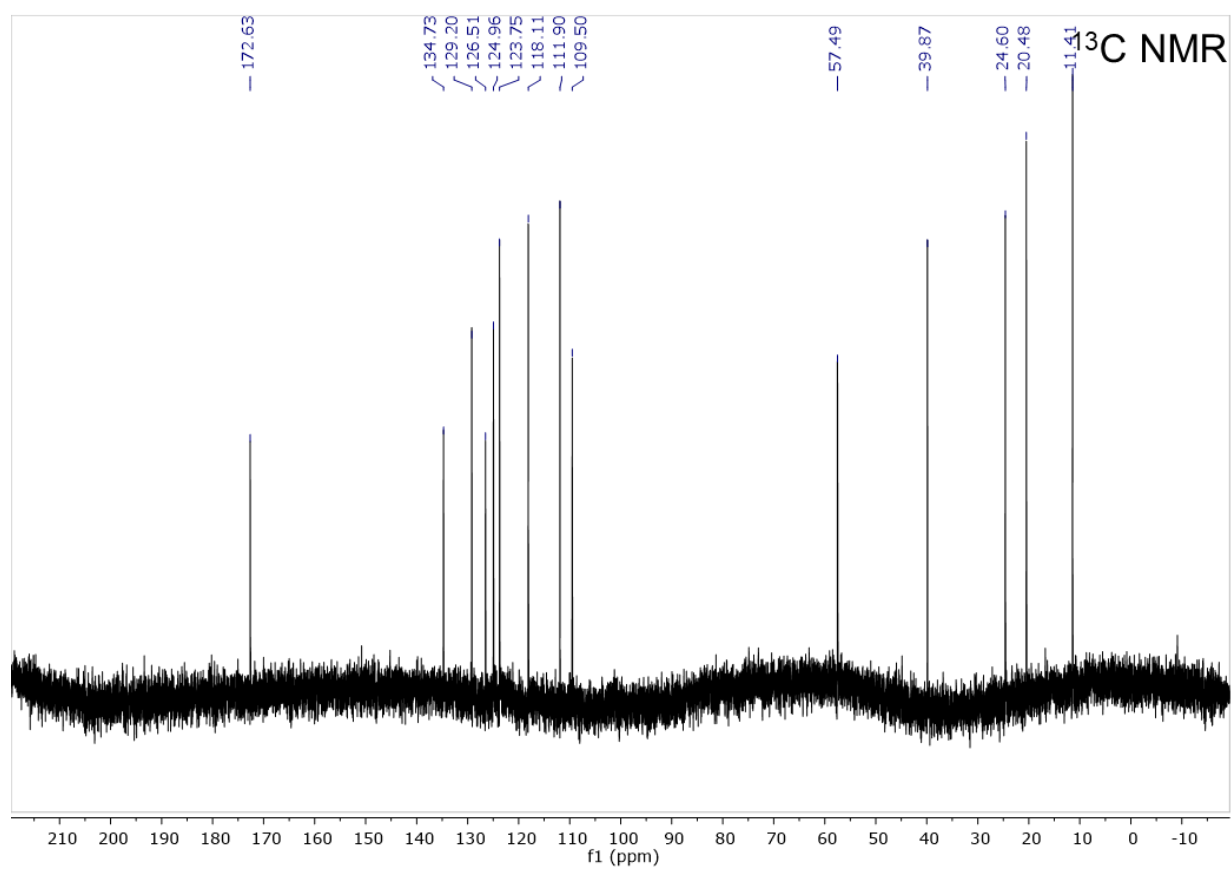
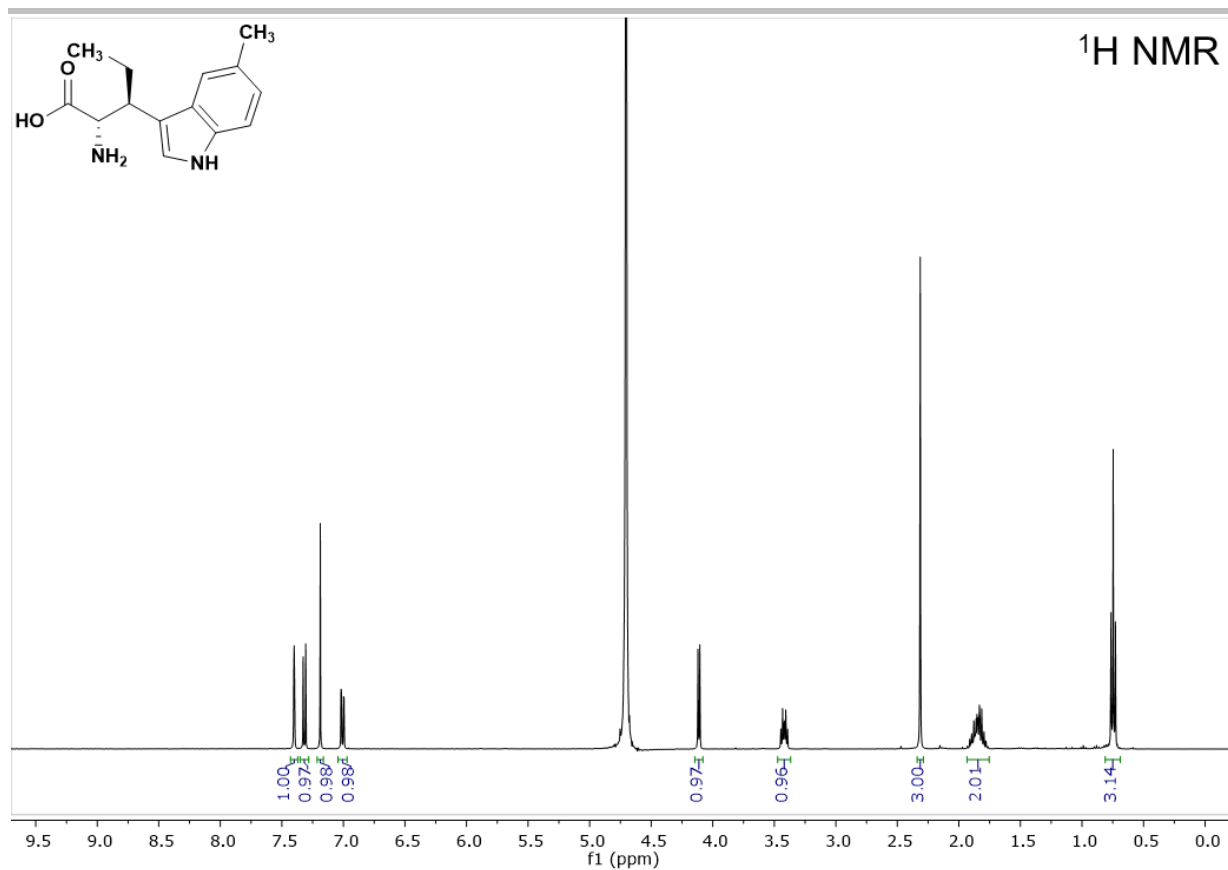


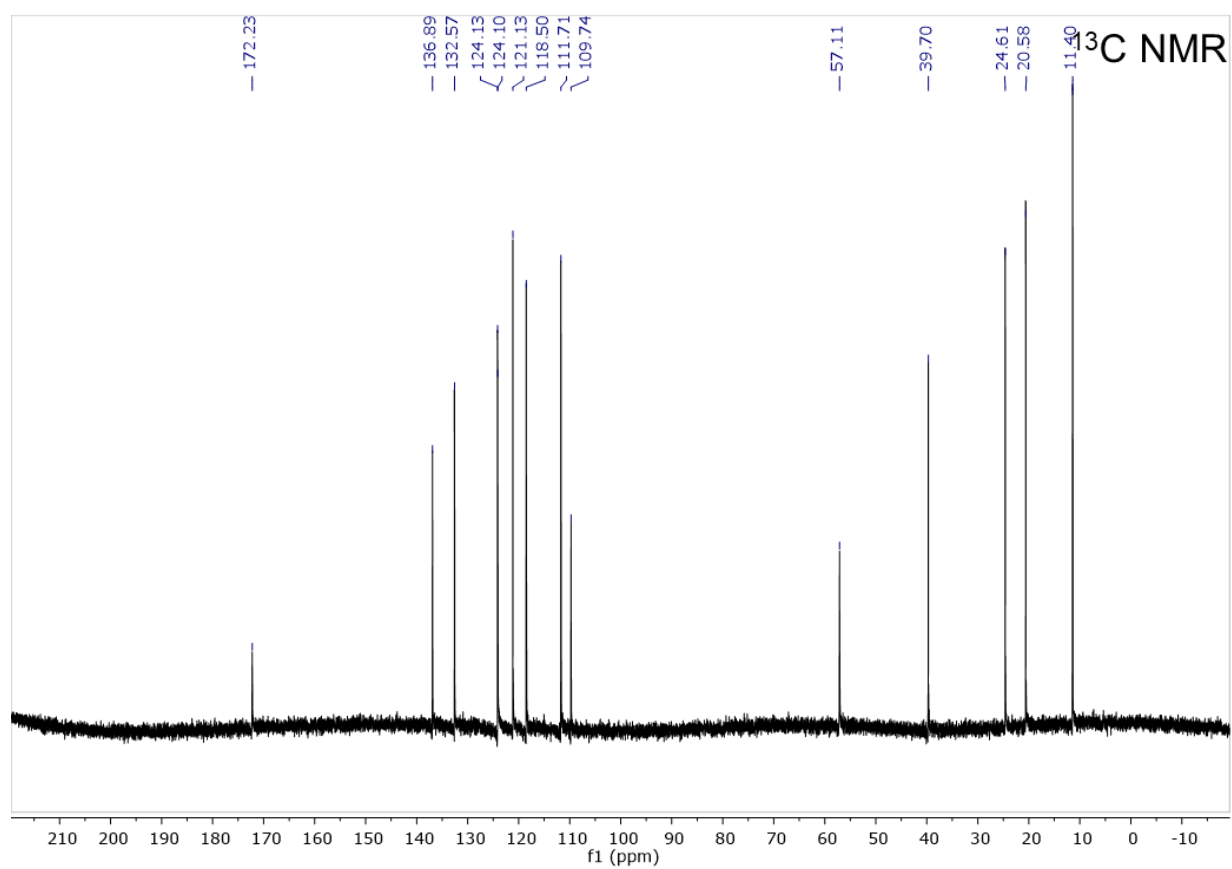
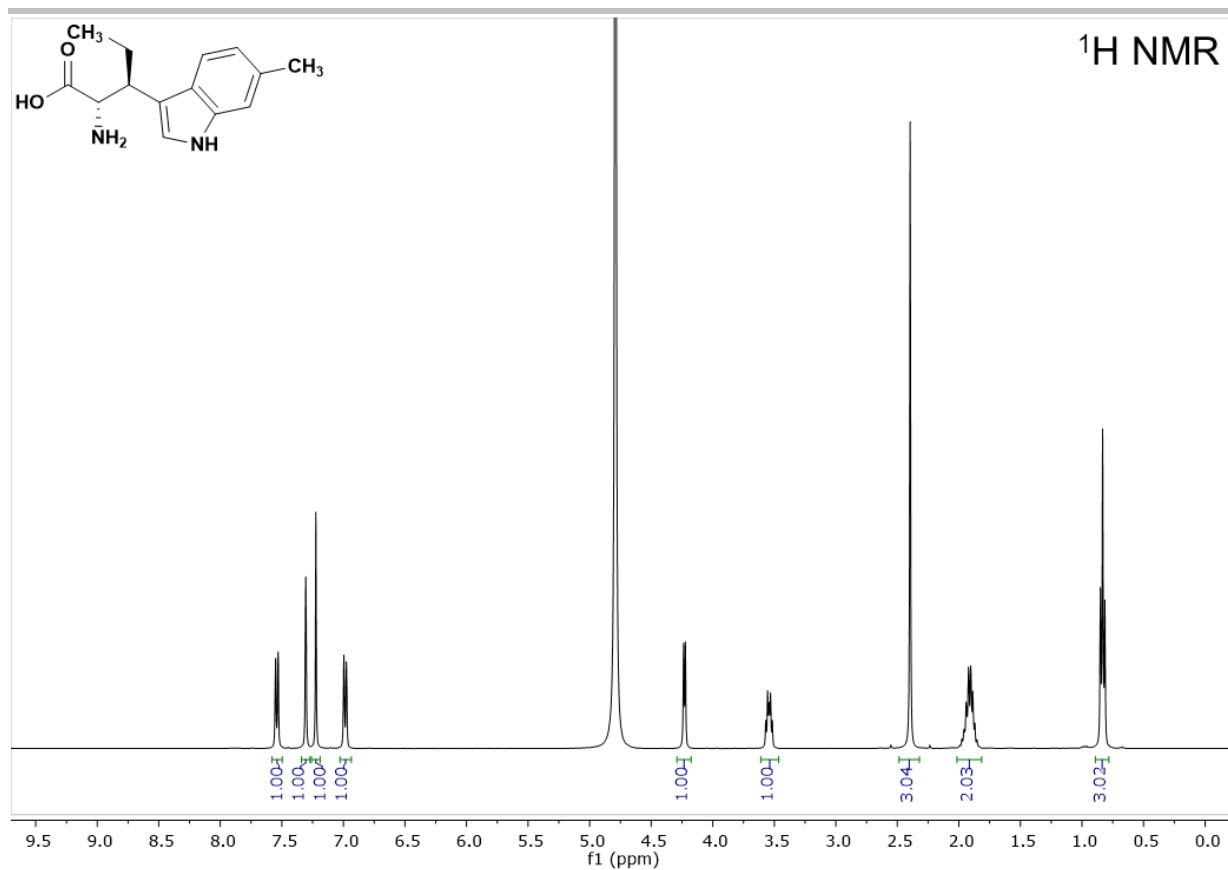


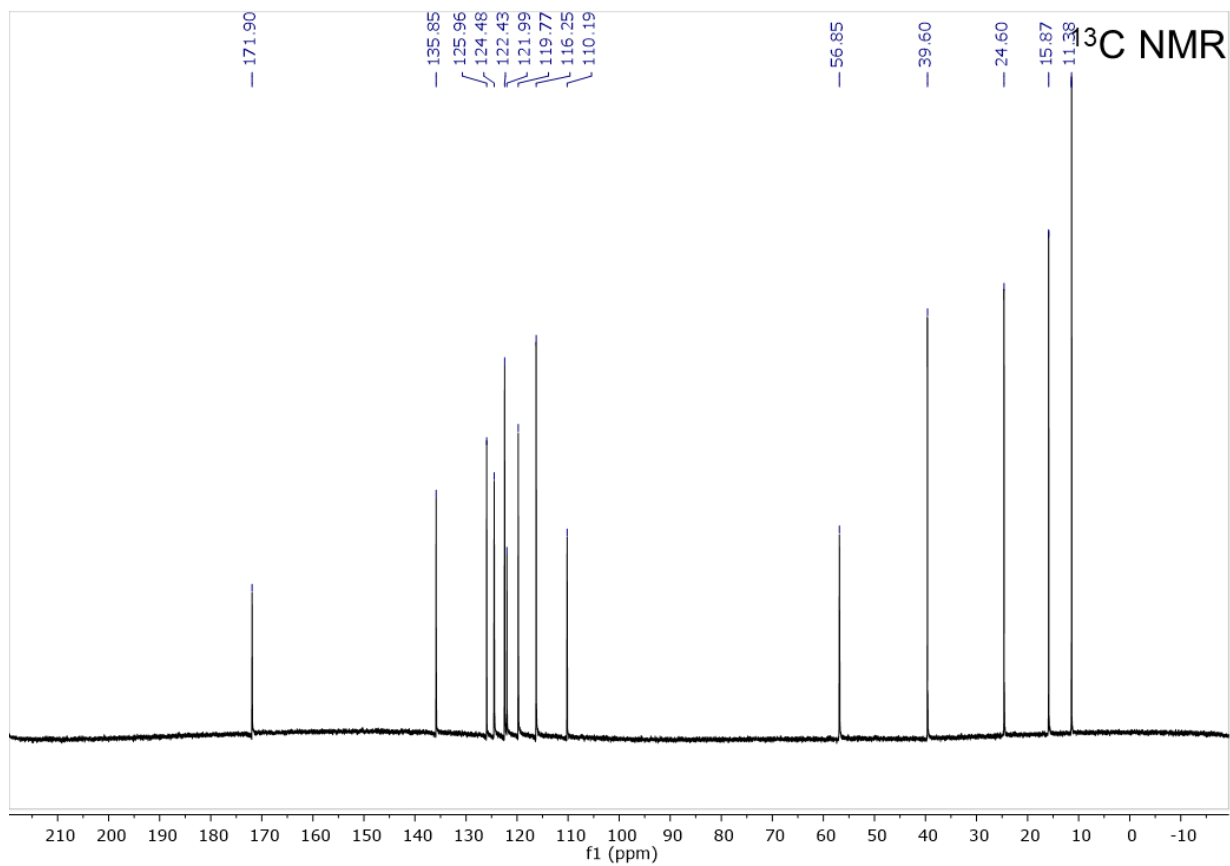
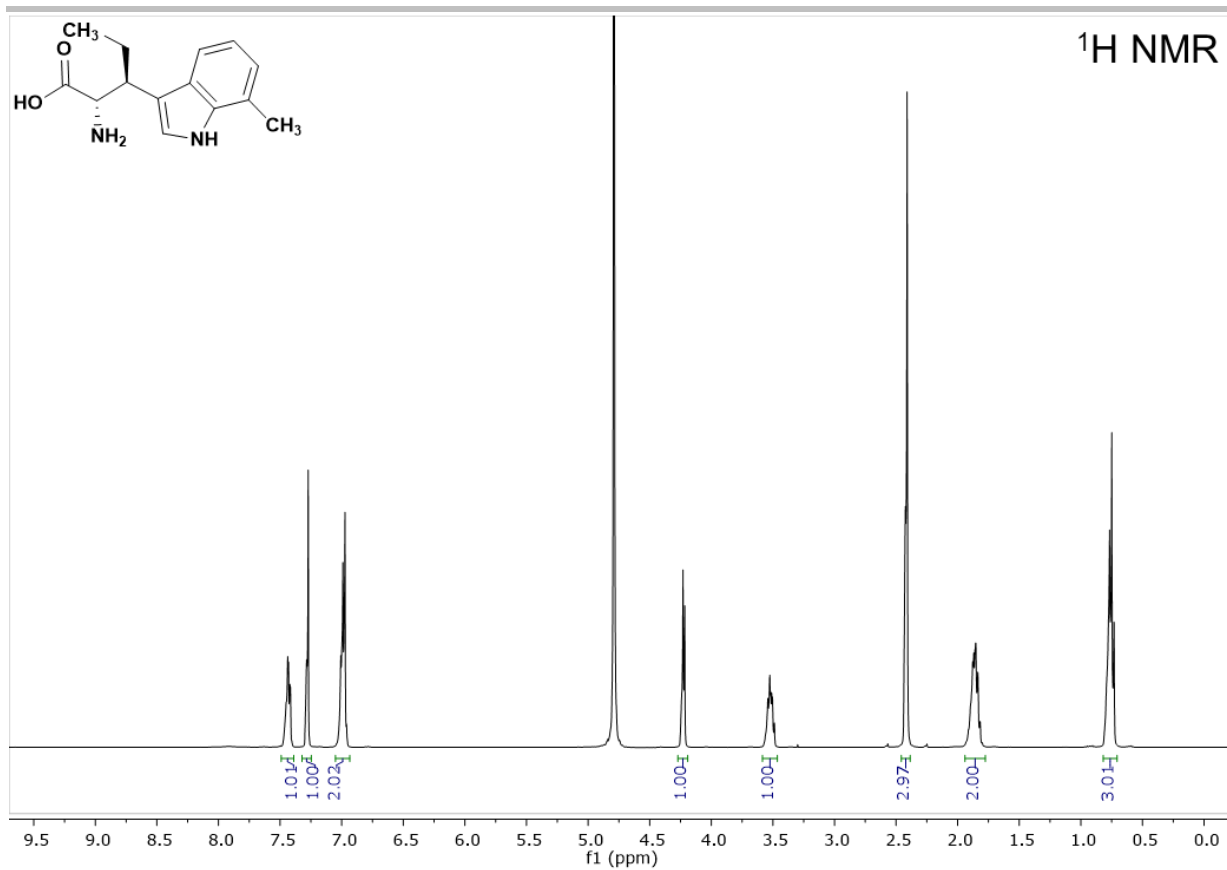


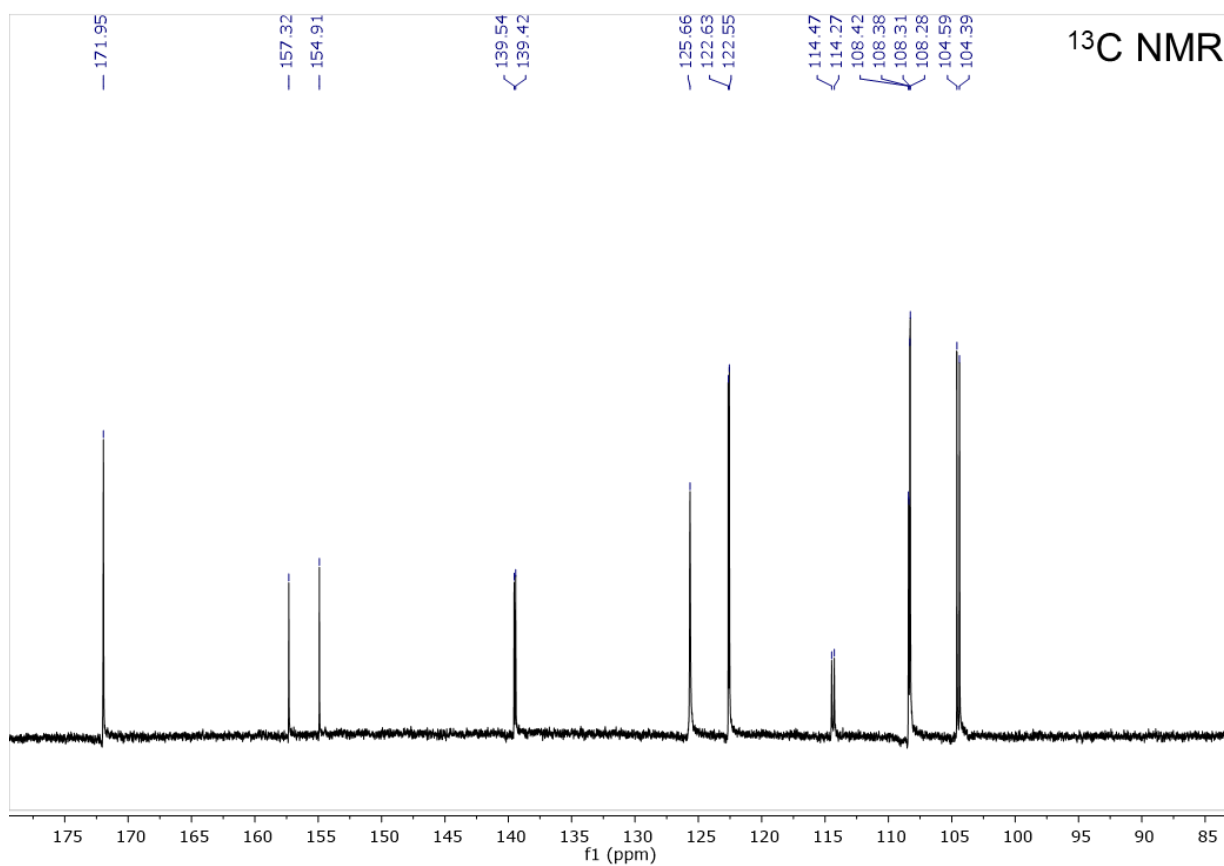
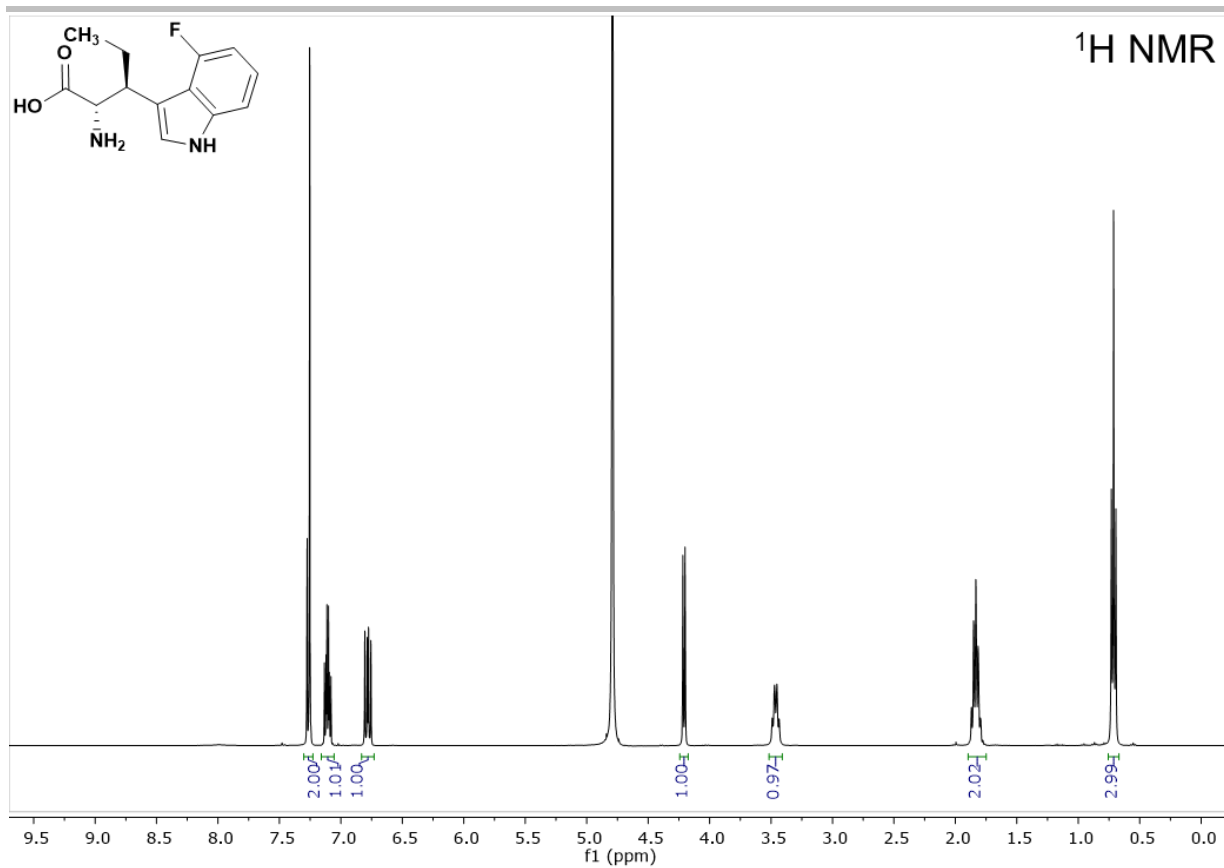


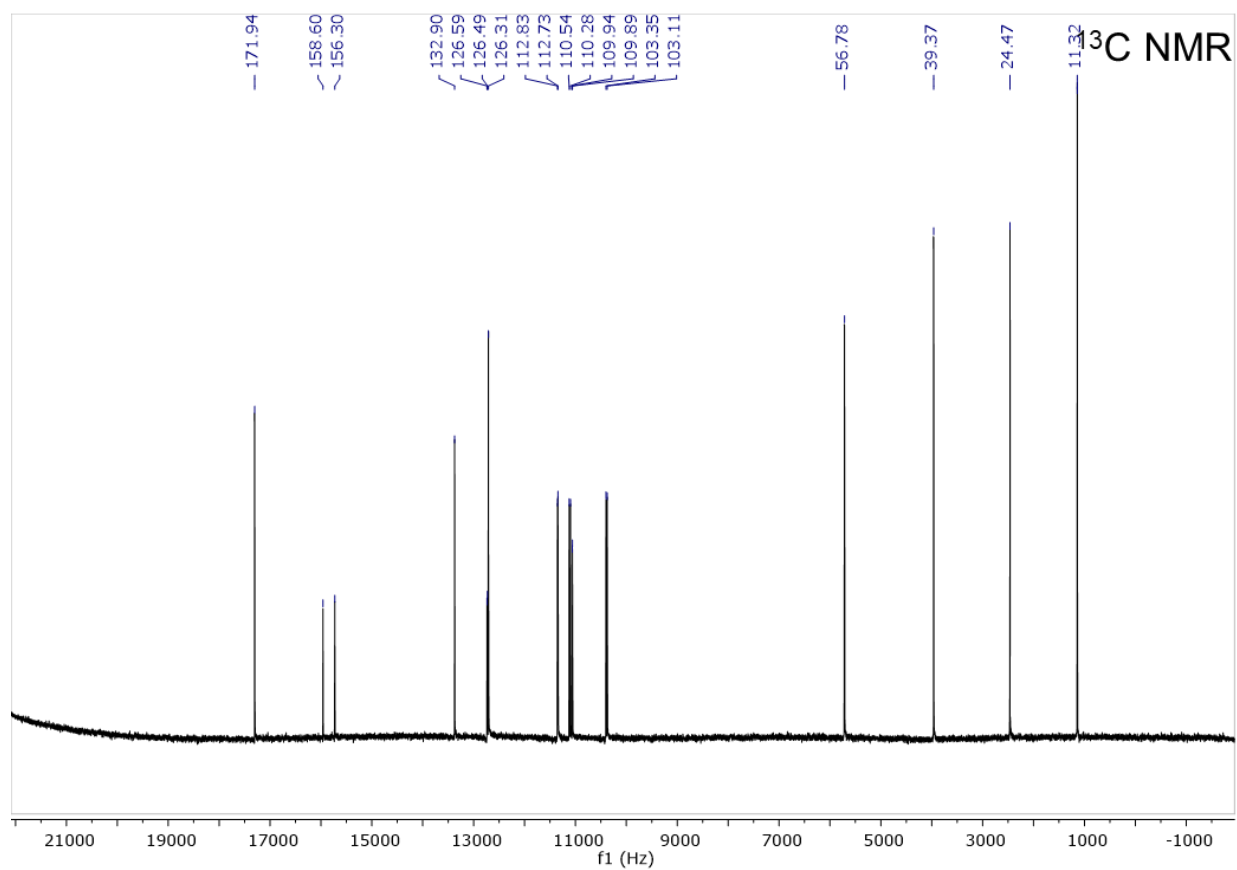
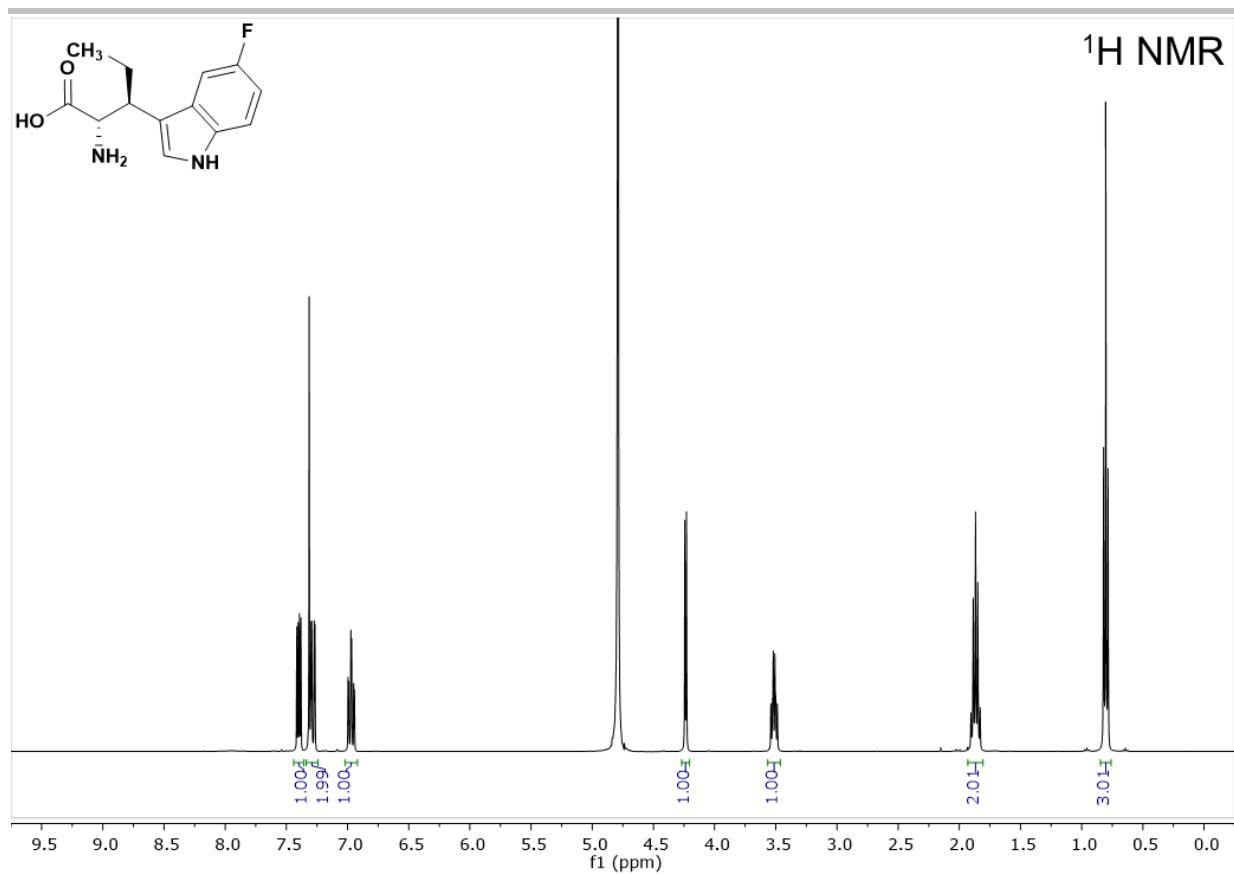


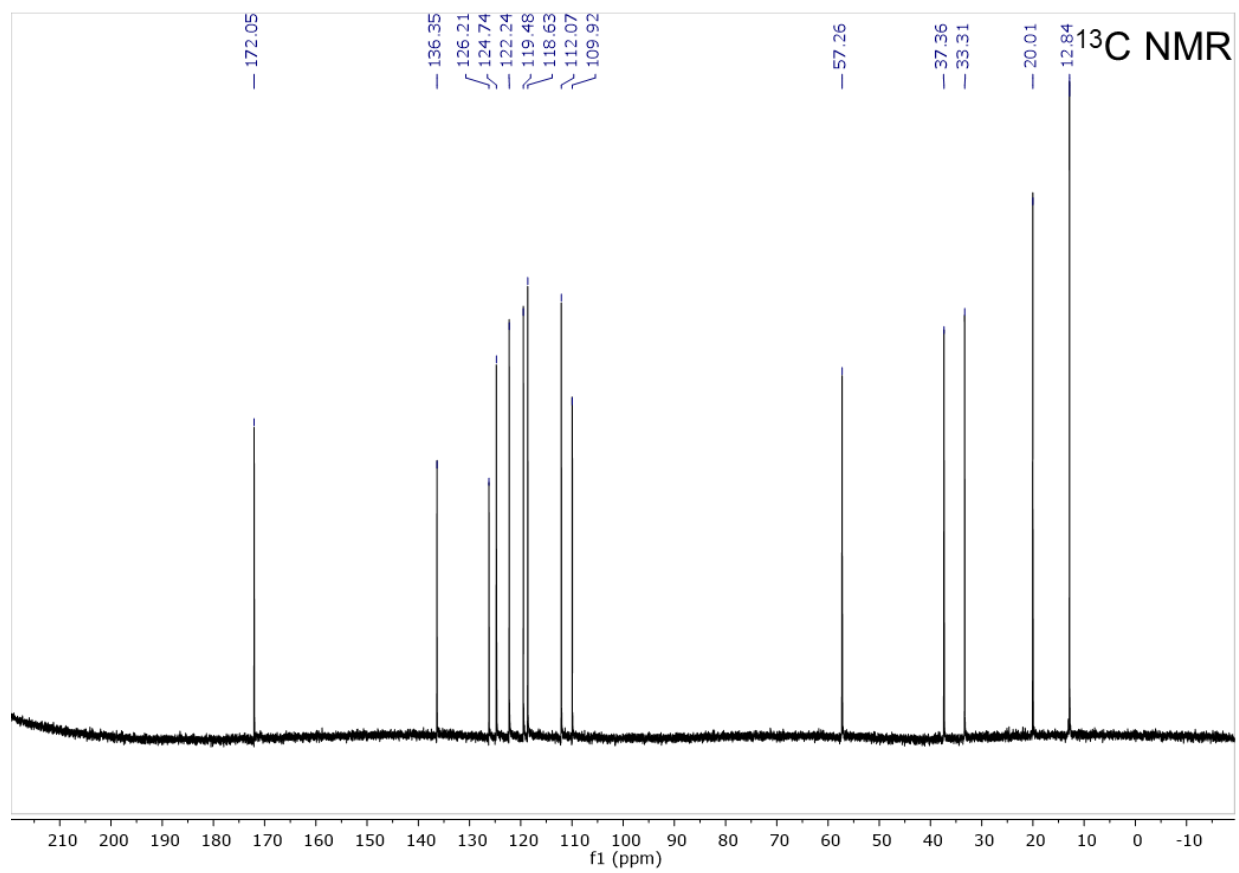
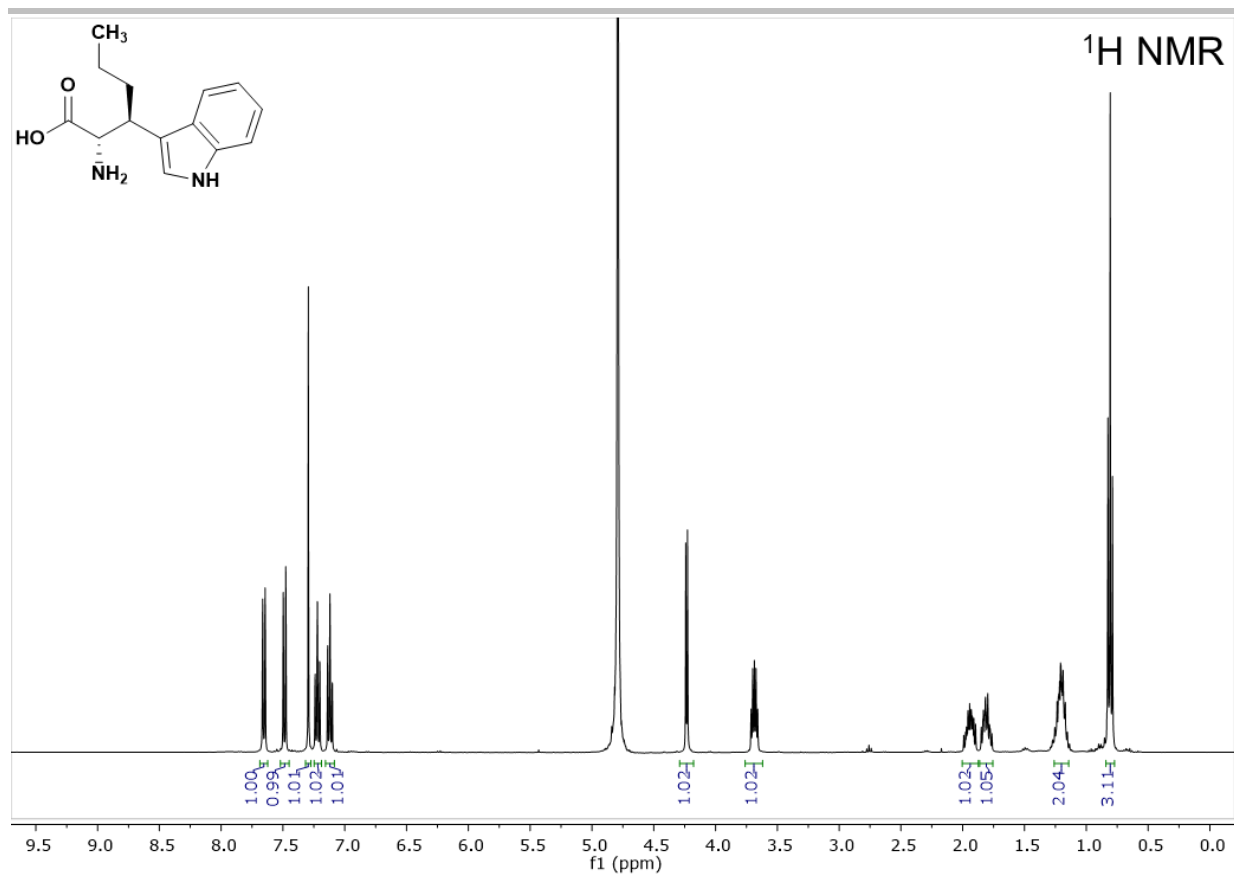


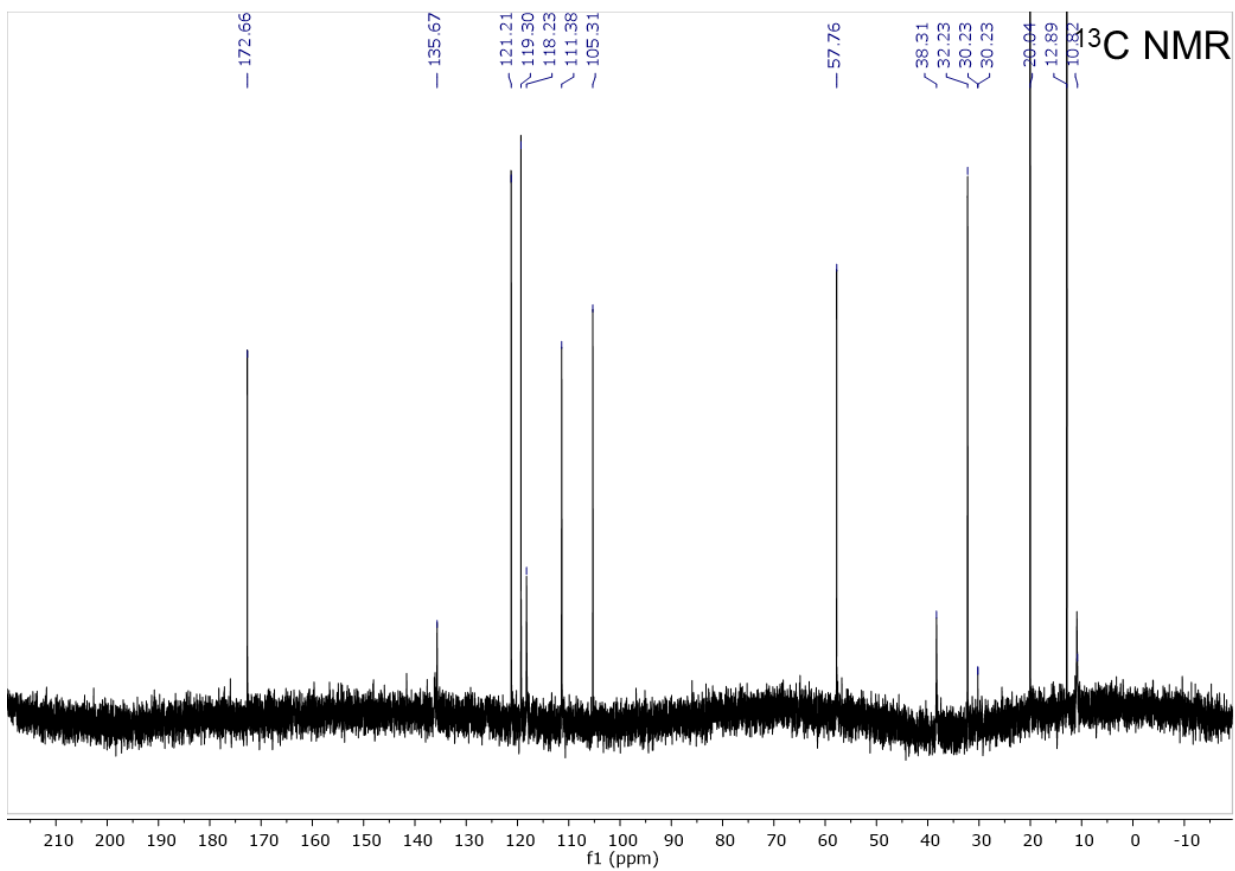
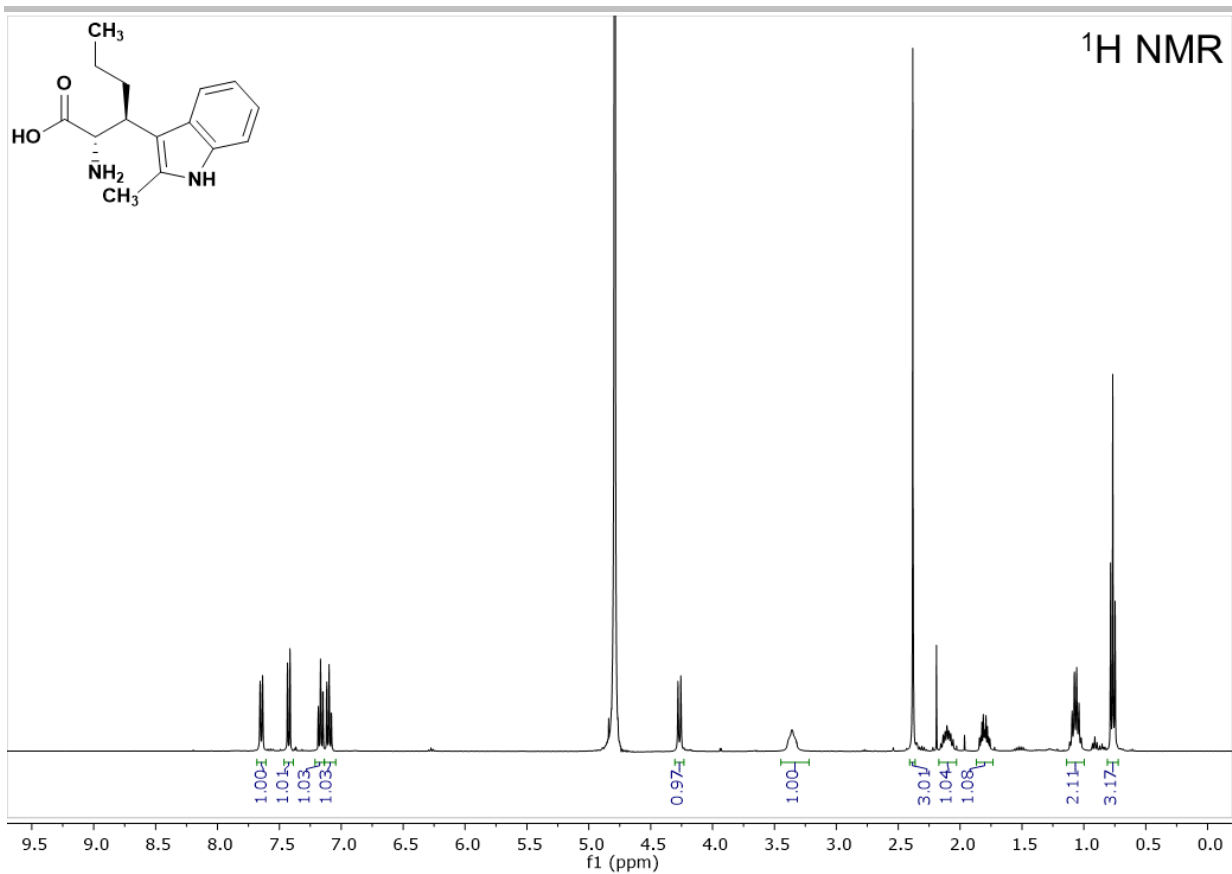


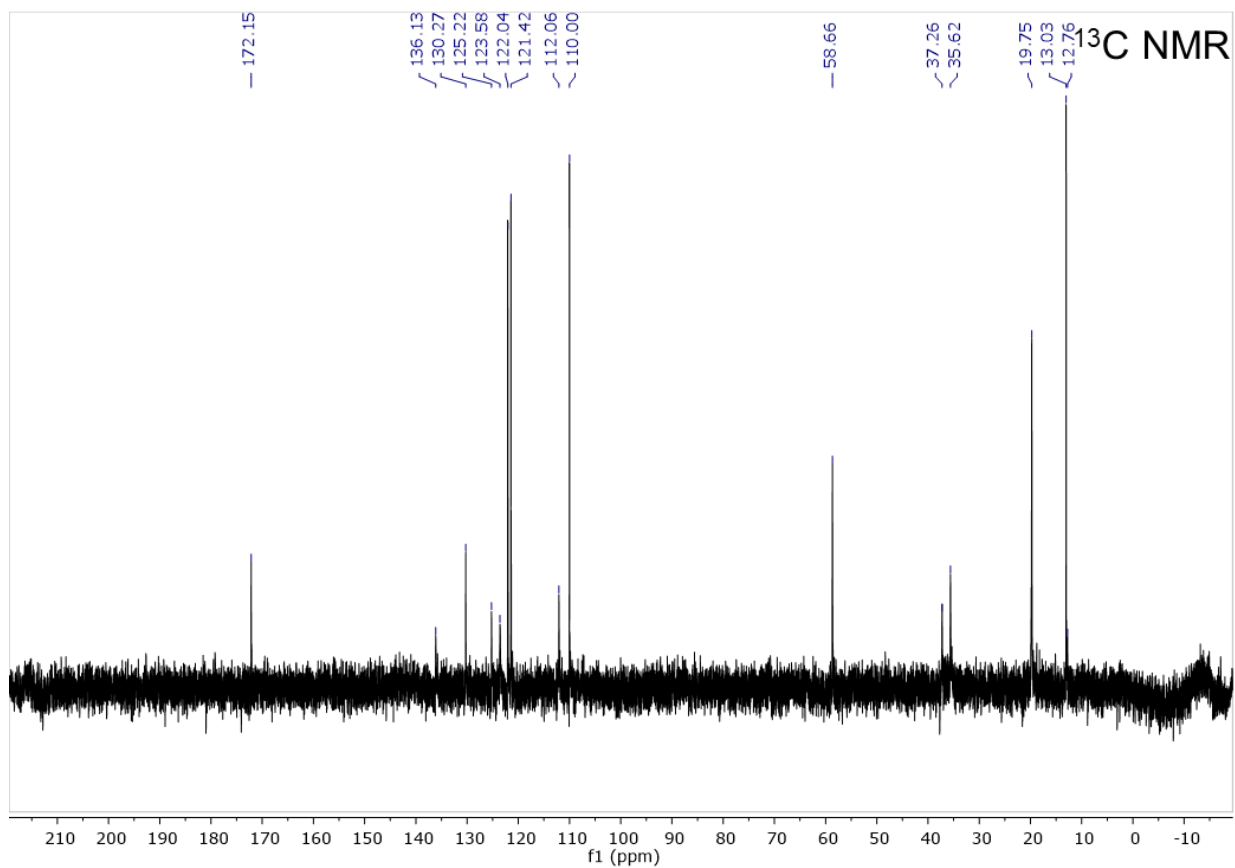
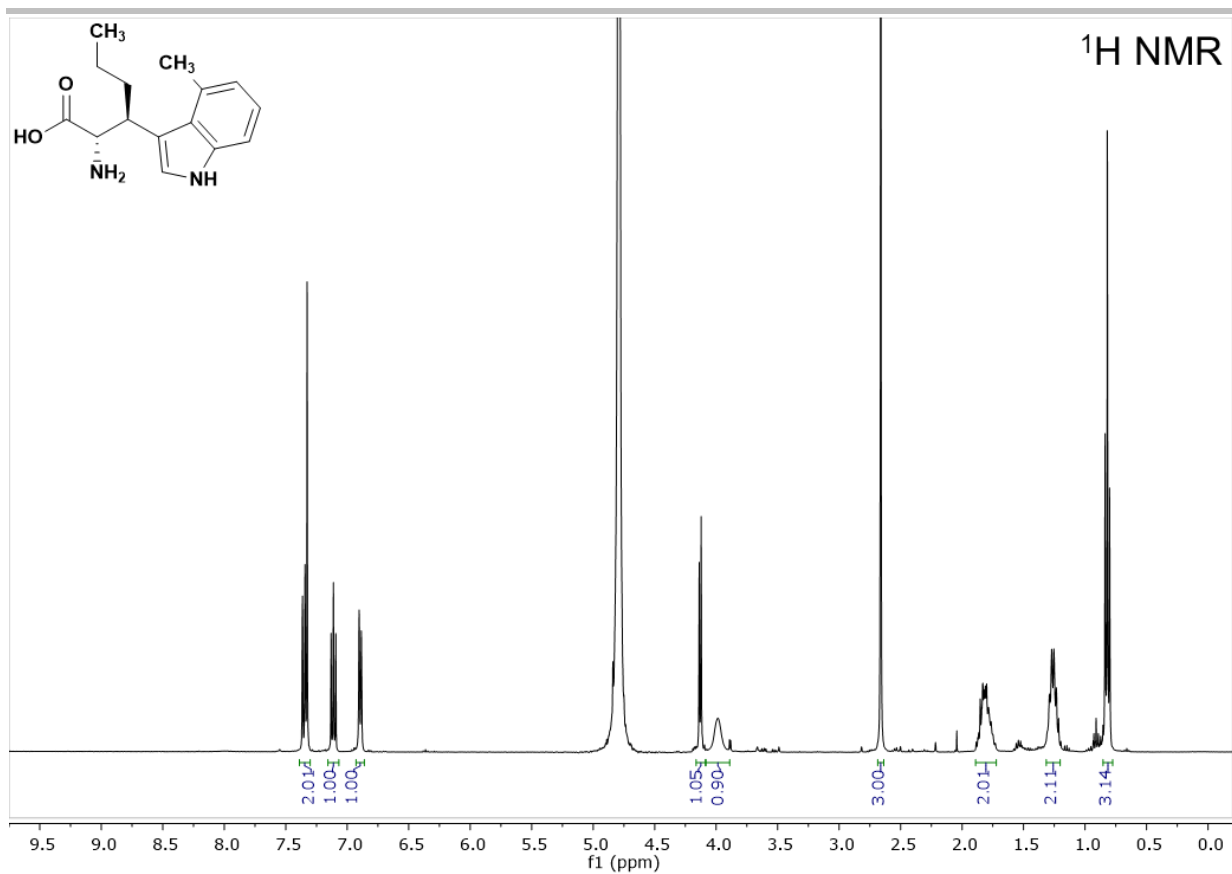


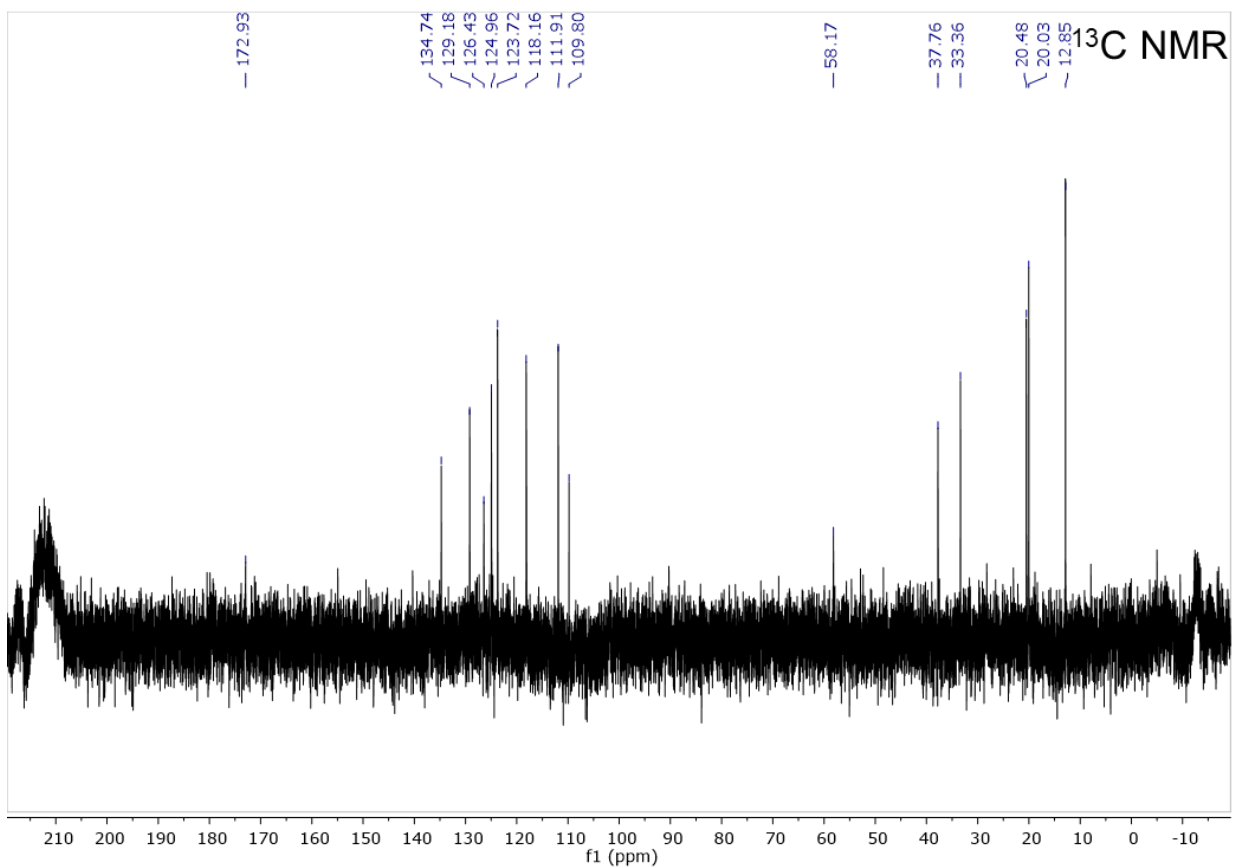
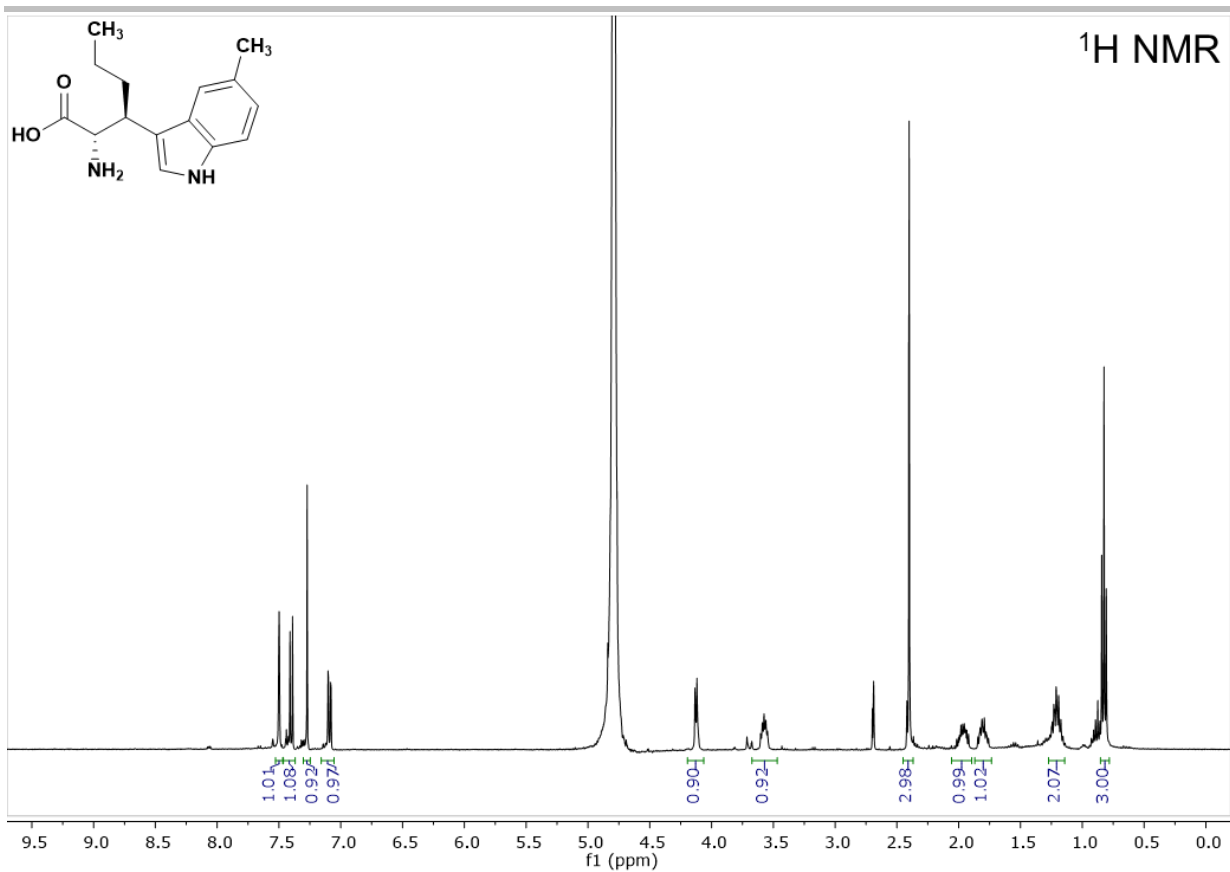


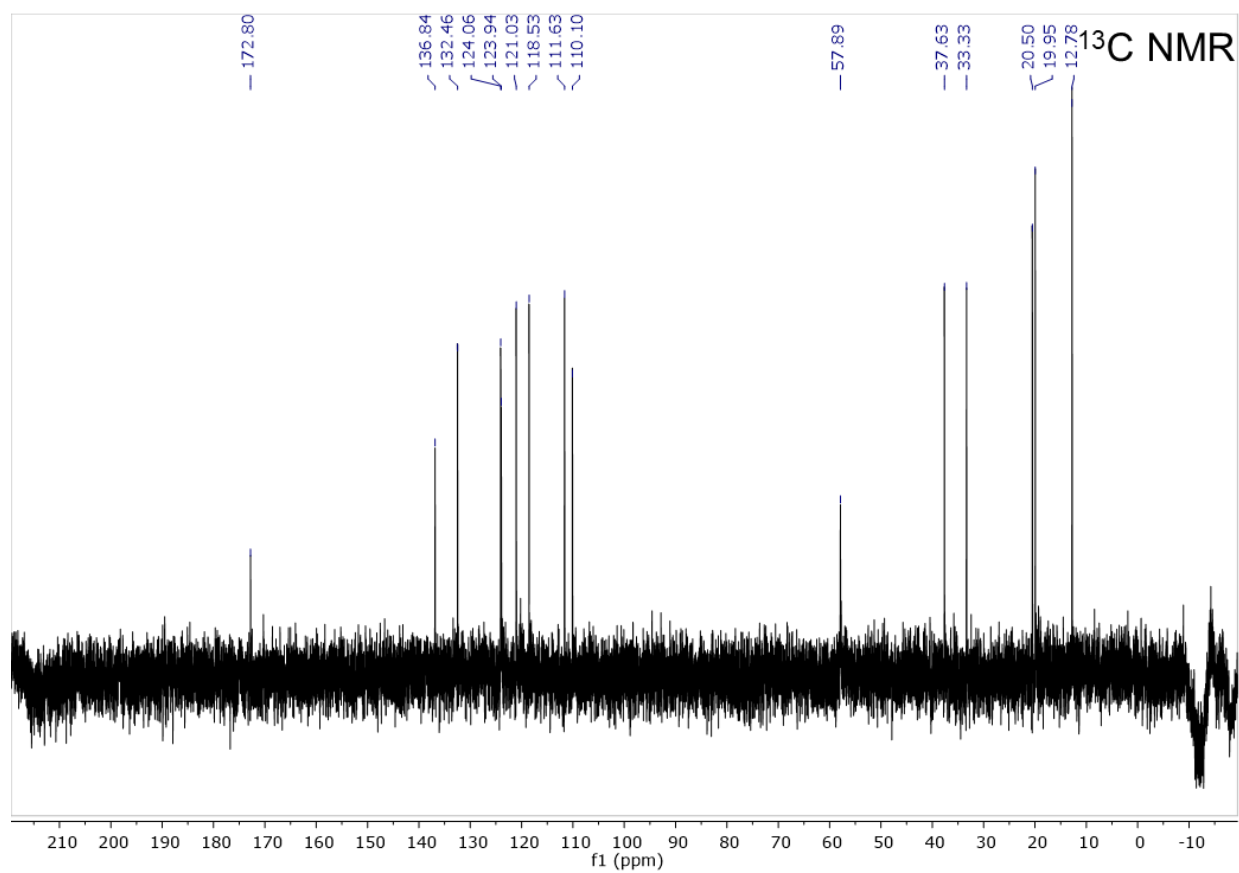
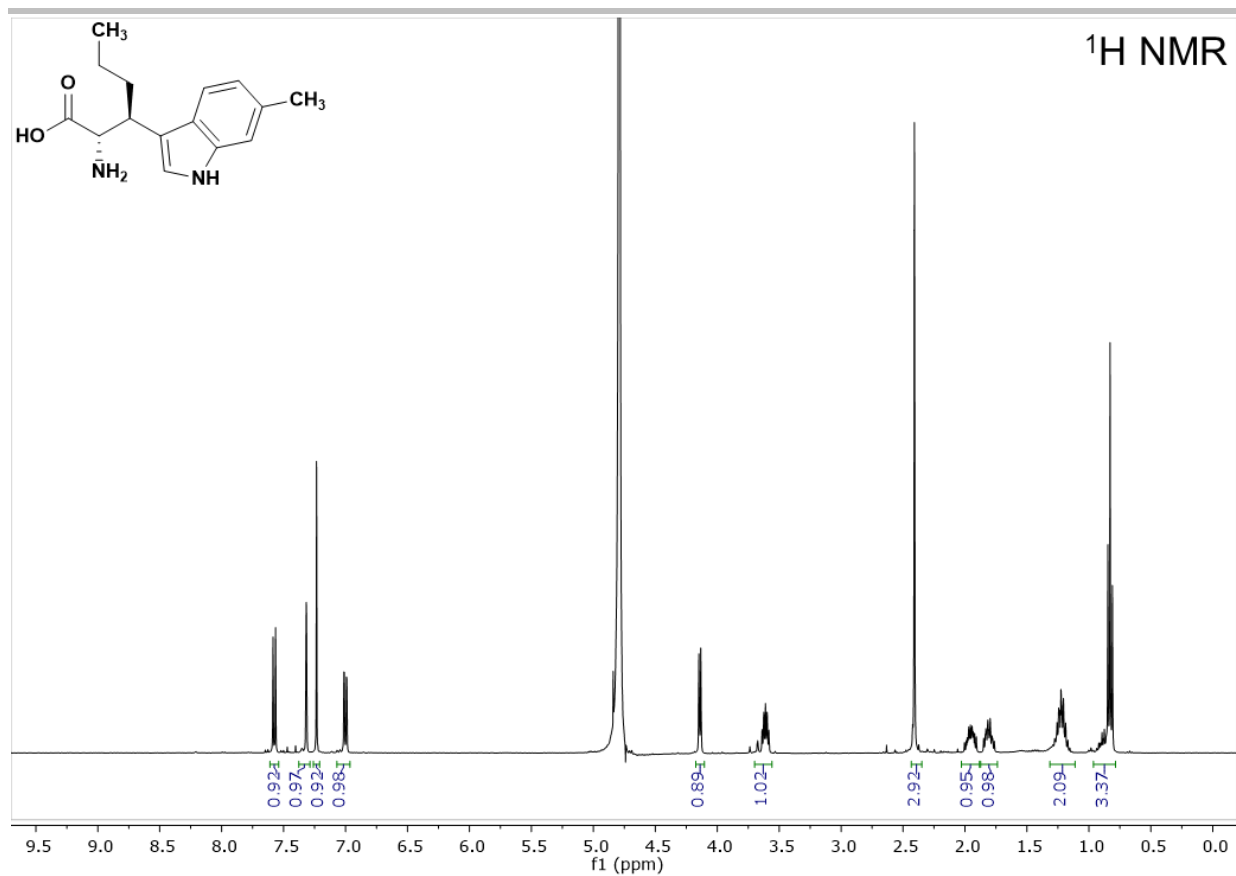


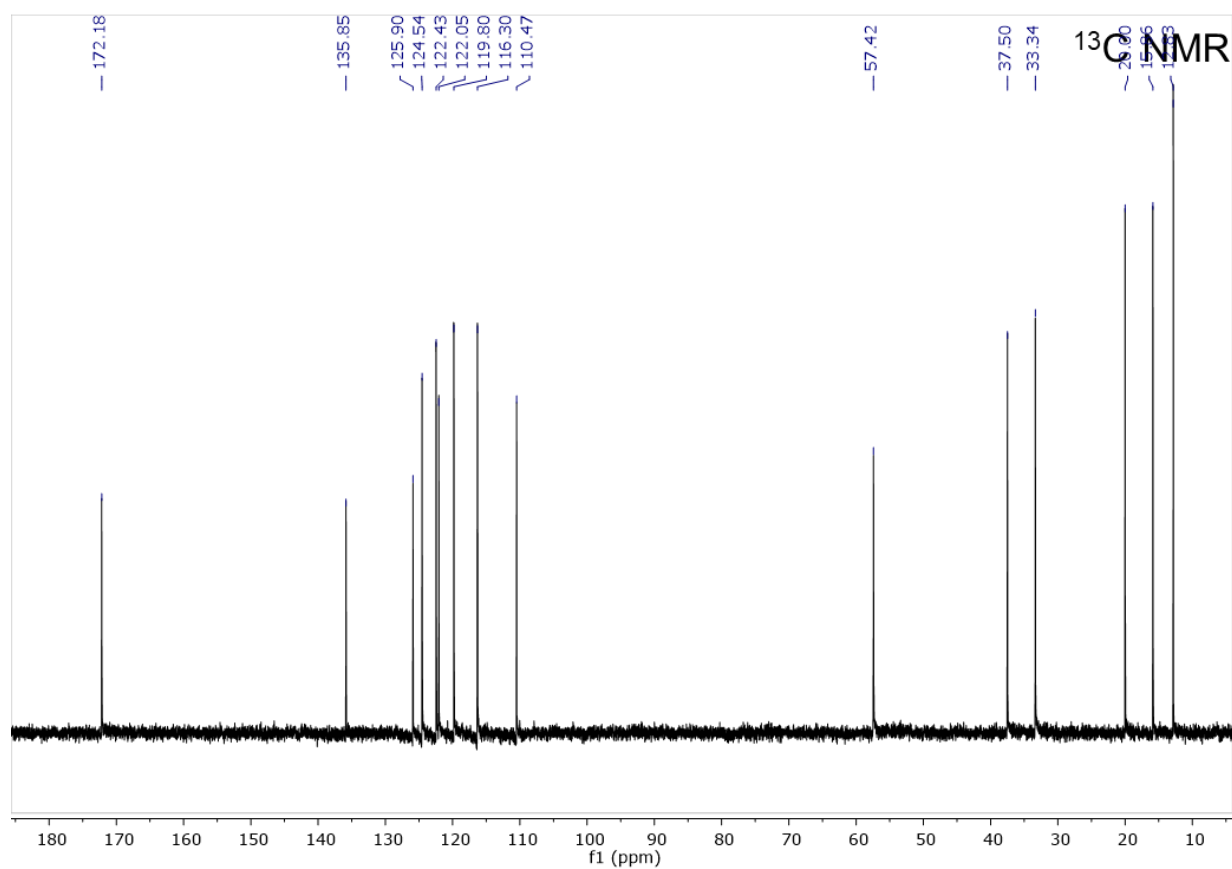
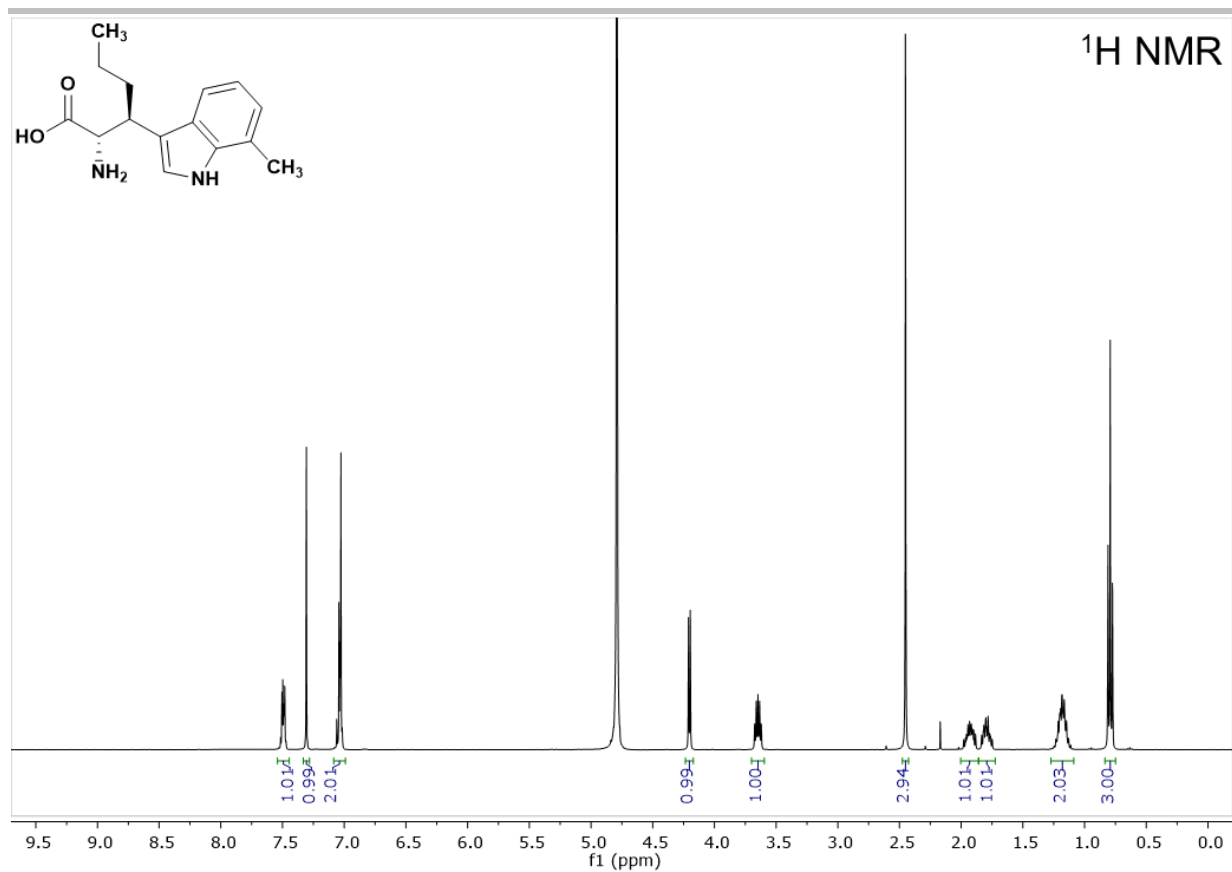


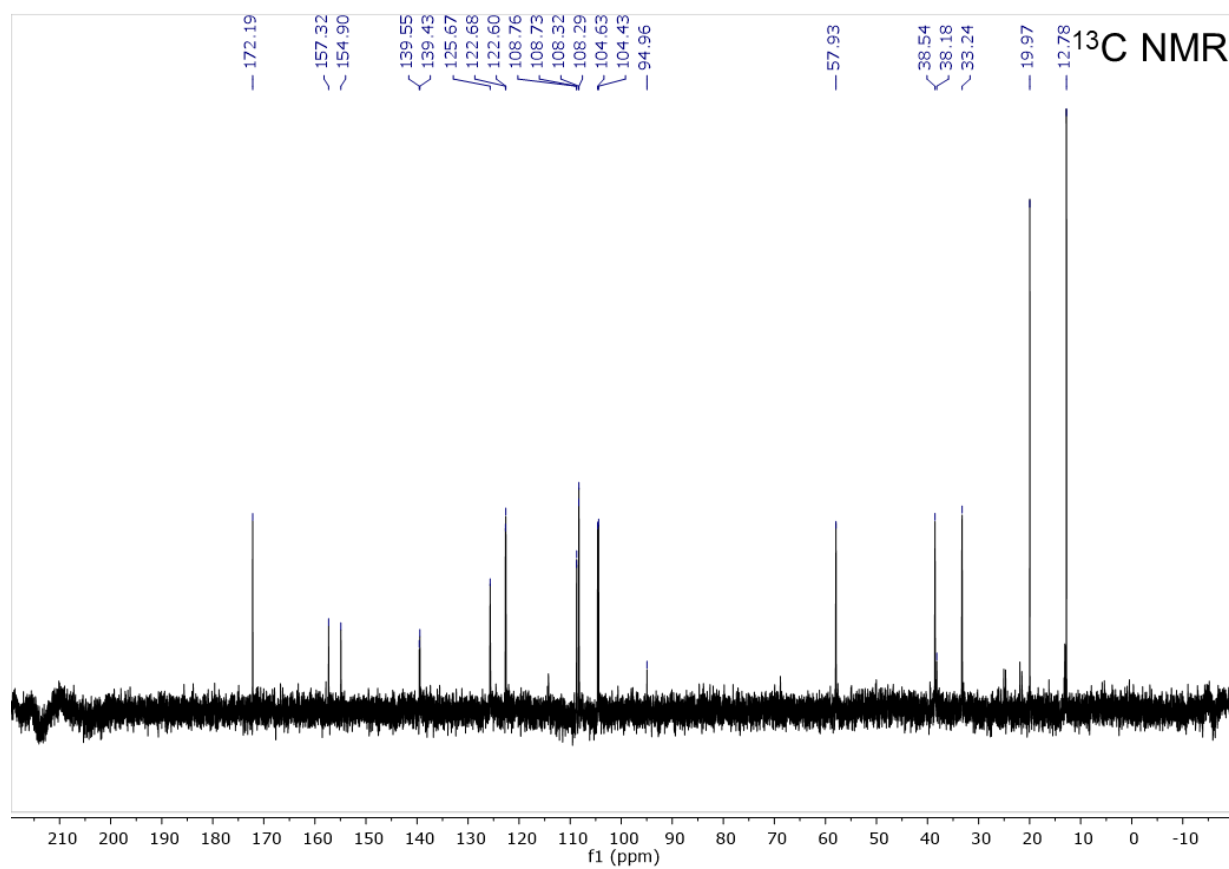
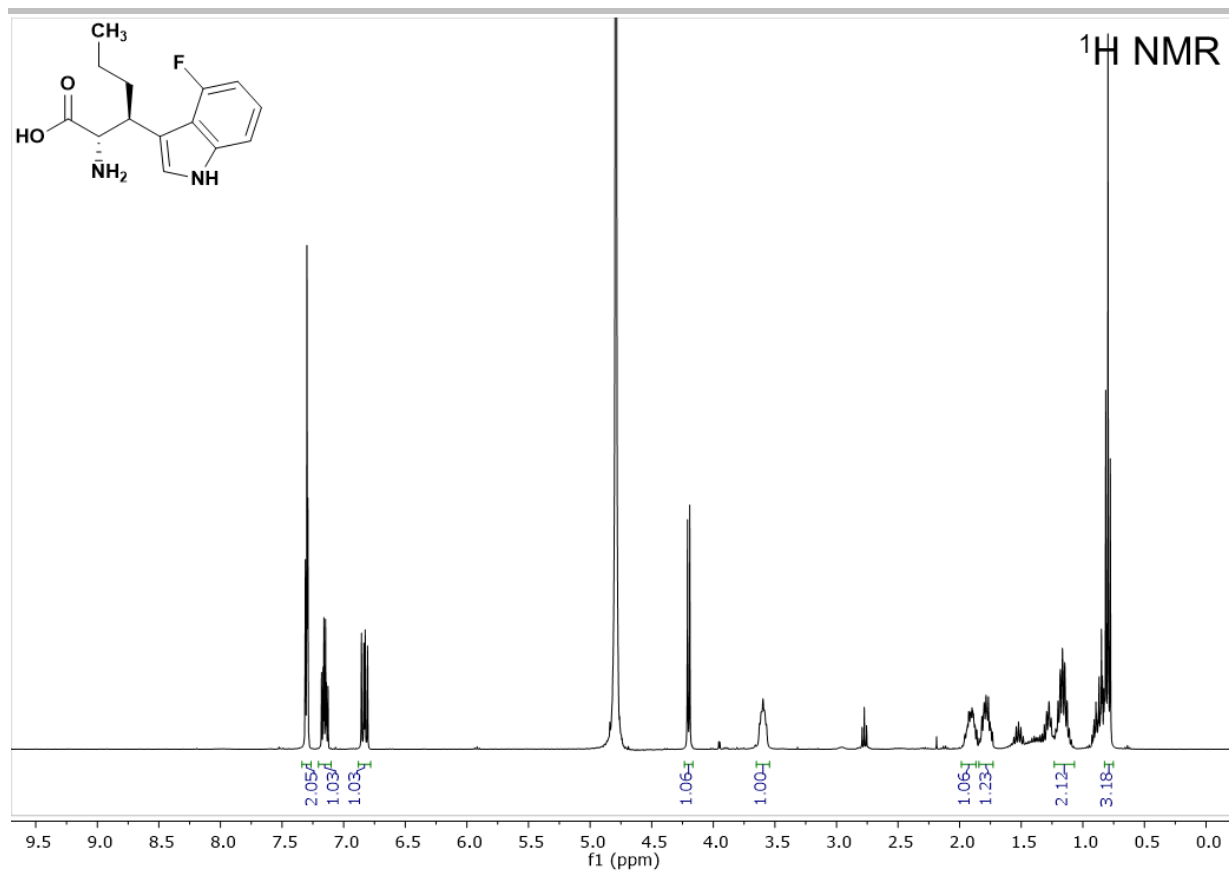


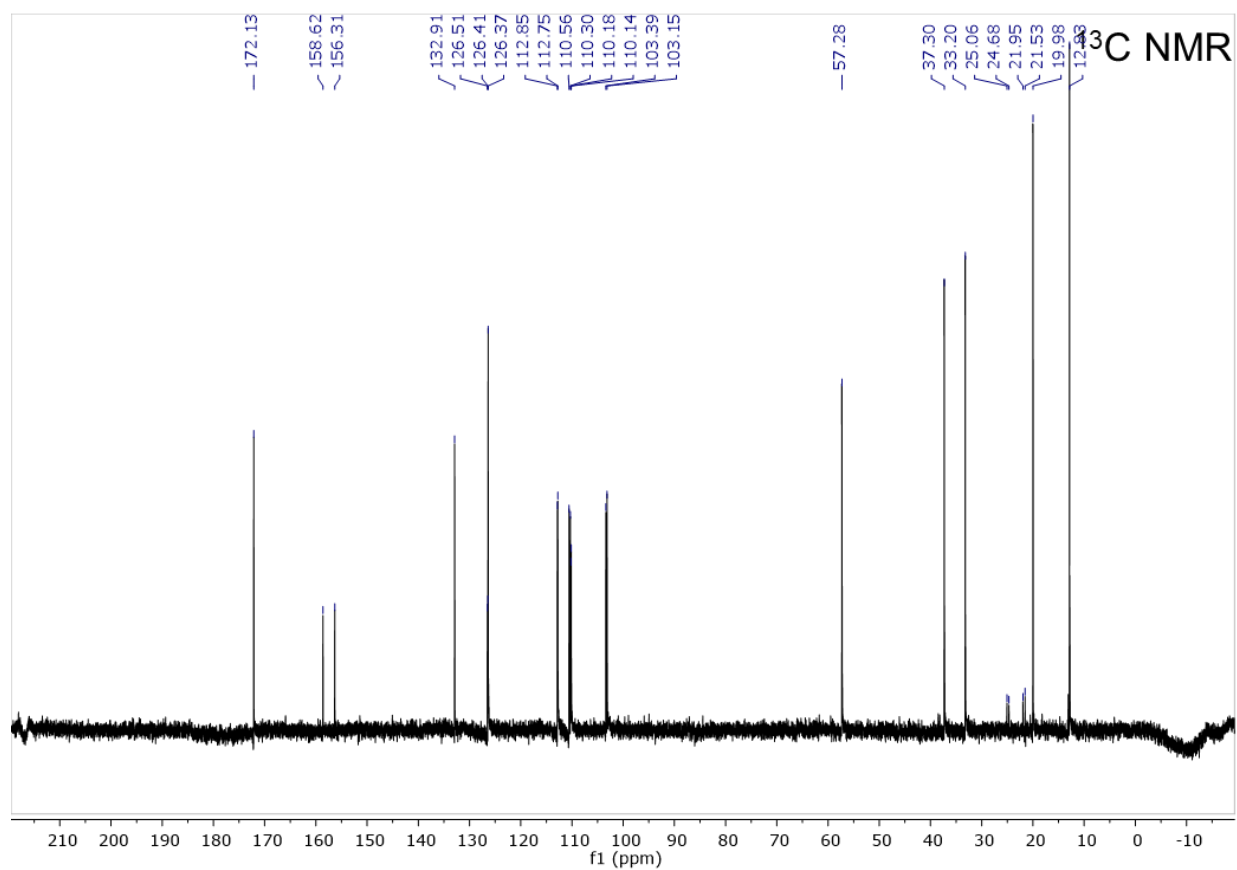
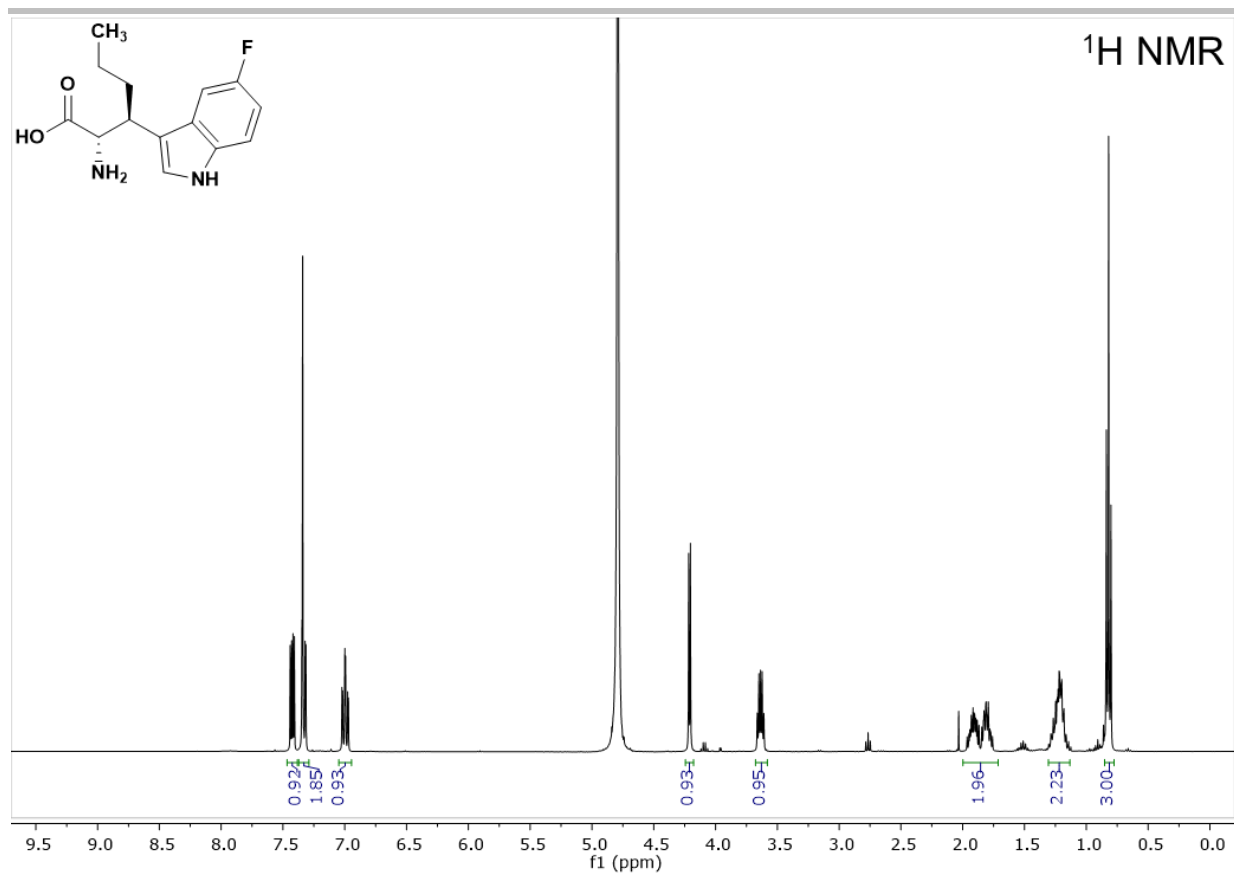












References

- [1] A. R. Buller, S. Brinkmann-Chen, D. K. Romney, M. Herger, J. Murciano-Calles, F. H. Arnold, *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 14599–14604.
- [2] M. Herger, P. van Roye, D. K. Romney, S. Brinkmann-Chen, A. R. Buller, F. H. Arnold, *J. Am. Chem. Soc.* **2016**, *138*, 8388–8391.
- [3] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, *Nat. Methods* **2009**, *6*, 343–345.
- [4] S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z.-G. Zhang, D. J. Opperman, M. T. Reetz, J. P. Acevedo, *ACS Synth. Biol.* **2013**, *2*, 83–92.
- [5] W. Kabsch, *Acta Crystallogr. D* **2010**, *66*, 133–144.
- [6] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. D* **2013**, *69*, 1204–1214.
- [7] P. A. Karplus, K. Diederichs, *Science* **2012**, *336*, 1030–1033.
- [8] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Cryst.* **2007**, *40*, 658–674.
- [9] M. D. Winn, G. N. Murshudov, M. Z. Papiz, *Meth. Enzymol.* **2003**, *374*, 300–321.
- [10] P. Emsley, K. Cowtan, *Acta Crystallogr. D* **2004**, *60*, 2126–2132.
- [11] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. D* **2010**, *66*, 12–21.
- [12] A. R. Buller, P. van Roye, J. K. B. Cahn, R. A. Scheele, M. Herger, F. H. Arnold, *J. Am. Chem. Soc.* **2018**, *140*, 7256–7266.

Author Contributions

Author contributions are noted using the CRediT taxonomy. Conceptualization: S.B.-C., A.R.B.; Methodology: C.E.B., R.A.S., P.K., S.B.-C., A.R.B.; Validation: C.E.B., R.A.S., P.K., S.B.-C.; Formal Analysis: C.E.B., R.A.S., P.K., S.B.-C.; Investigation: C.E.B., R.A.S., P.K., S.B.-C.; Writing-Original Draft: C.E.B.; Writing-Review & Editing: C.E.B., R.A.S., P.K., S.B.-C., A.R.B., F.H.A.; Visualization: C.E.B.; Supervision: A.R.B.; Funding Acquisition: C.E.B., R.A.S., A.R.B., F.H.A.

(c) Supplemental Information ChemRxiv.pdf (2.45 MiB)

[view on ChemRxiv](#) • [download file](#)
