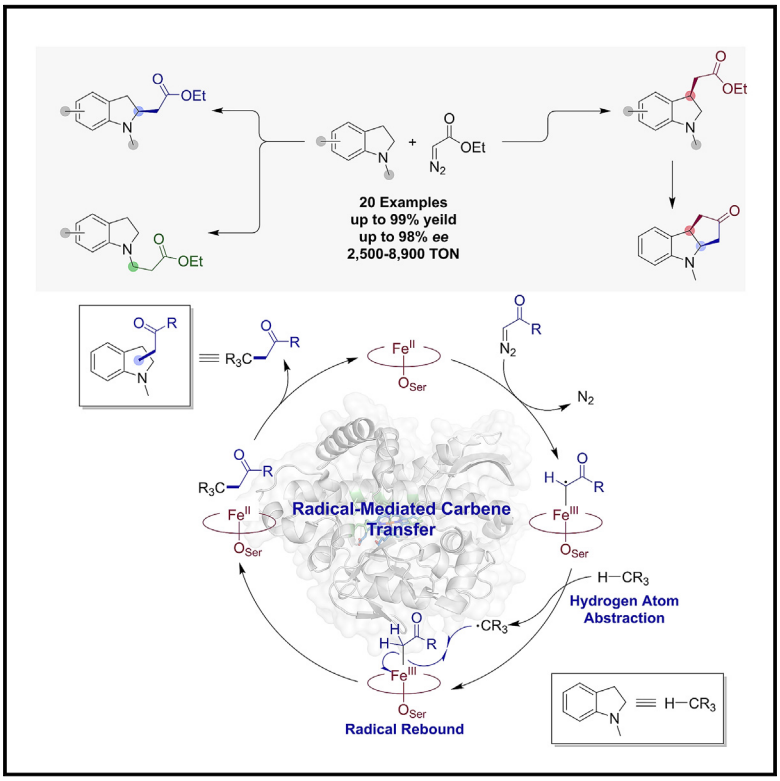
Chem Catalysis

Article

Radical-mediated regiodivergent C(sp^3)–H functionalization of *N*-substituted indolines via enzymatic carbene transfer

Graphical abstract



Highlights

- Biocatalytic approach for stereoselective C(sp^3)–H carbene insertion in indolines
- Tunable regioselectivity toward three distinct C(sp^3)–H sites via protein engineering
- Synthesis of polycyclic indoline-containing building blocks via biocatalytic cascade
- First mechanistic insights into hemoprotein-catalyzed C(sp^3)–H carbene insertion

Authors

Bo M. Couture, Ru Cui, Jia-Min Chu, Zhuofan Shen, Sagar D. Khare, Yong Zhang, Rudi Fasan

Correspondence

yzhang37@stevens.edu (Y.Z.), rudi.fasan@utdallas.edu (R.F.)

In brief

This work reports the development of an efficient biocatalytic strategy for the stereoselective and regiodivergent C(sp^3)–H functionalization of N-substituted indolines, a privileged scaffold in bioactive molecules, using engineered CYP119-based "carbene transferases." A biocatalytic cascade combining two engineered biocatalysts with divergent regioselectivity enabled the stereoselective synthesis of polycyclic indoline-based scaffolds found in natural products. Finally, computational and experimental mechanistic studies offer first-time insights into the mechanism of hemoprotein-catalyzed carbene C(sp^3)–H insertion and protein-controlled site selectivity in this class of transformations.

Couture et al., 2024, Chem Catalysis *4*, 101133
November 21, 2024
https://doi.org/10.1016/j.checat.2024.101133

CellPress

Chem Catalysis

CellPress

Article

Radical-mediated regiodivergent C(sp^3)–H functionalization of *N*-substituted indolines via enzymatic carbene transfer

Bo M. Couture,[1] Ru Cui,[1] Jia-Min Chu,[2] Zhuofan Shen,[3] Sagar D. Khare,[3,4] Yong Zhang,[2,*] and Rudi Fasan[1,5,*]

[1]Department of Chemistry and Biochemistry, University of Texas at Dallas, Richardson, TX 75080, USA

[2]Department of Chemistry and Chemical Biology, Stevens Institute of Technology, Hoboken, NJ 07030, USA

[3]Department of Chemistry and Chemical Biology, Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA

[4]Institute for Quantitative Biomedicine, Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA

[5]Lead contact

*Correspondence: yzhang37@stevens.edu (Y.Z.), rudi.fasan@utdallas.edu (R.F.)

https://doi.org/10.1016/j.checat.2024.101133

THE BIGGER PICTURE Selective C(sp^3)–H functionalization is a powerful approach for the synthesis and diversification of organic molecules. While regiodivergent enzymatic C(sp^3)–H oxidation has been achieved with monooxygenase enzymes, regiodivergent C(sp^3)–H functionalization via "abiological" carbene transfer chemistry has been elusive to biocatalysis and challenging for chemocatalysts. Herein, we report an efficient biocatalytic strategy for the highly stereoselective and regiodivergent C(sp^3)–H carbene insertion in indoline substrates, which are ubiquitous motifs in drugs and natural products. These regiodivergent "carbene transferases" were further leveraged in a biocatalytic cascade to afford polycyclic indoline-based scaffolds. Our studies also provide first-time insight into the mechanism of hemoprotein-catalyzed C(sp^3)–H insertion and protein-mediated regiocontrol in these reactions, which can guide future development of biological catalysts for this important class of transformations.

SUMMARY

Indolines are ubiquitous structural motifs occurring in pharmaceuticals and natural products. Here, we report a strategy for regio- and stereoselective C(sp^3)–H functionalization of *N*-substituted indolines via carbene transfer chemistry mediated by engineered CYP119-based catalysts. These systems offer high enantioselectivity and high catalytic efficiency, as well as regiodivergent selectivity, furnishing an efficient and convenient route for diversification of these important scaffolds via direct C(sp^3)–H functionalization. Selective functionalization of exocyclic C(sp^3)–H bond in *N*-methyl indolines was also achieved, and a biocatalytic cascade combining enzyme-mediated α- and β-C(sp^3)–H functionalization yielded a polycyclic indoline-containing motif found in drugs. Mechanistic and computational studies support a radical-mediated C–H functionalization pathway and provide insights into protein-mediated regiodivergent selectivity. Altogether, this work offers a direct and tunable strategy to access functionalized indolines as key building blocks for medicinal chemistry and natural product synthesis and provides first insights into the mechanism of P450-catalyzed C(sp^3)–H carbene insertion.

INTRODUCTION

Given the privileged nature of functionalized indolines in pharmaceuticals and their ubiquitous presence within biological systems (Figure 1A),[1–4] many efforts have been made toward the synthesis of such scaffolds. Existing methodologies largely rely on either reduction of the indole counterpart[5–7] or catalytic intramolecular cyclization reactions[8–11] to produce functionalized indolines. Comparatively, direct C–H functionalization offers an attractive approach for the synthesis and diversification of indoline-based scaffolds, although methods for their regio- and stereoselective direct C–H functionalization have been elusive.[12]

Biocatalysis has attracted increasing attention as a strategy to address important challenges in chemical synthesis.[13–21] In the area of C–H functionalization, various classes of oxidizing enzymes, including cytochromes P450, unspecific peroxygenases, flavin-dependent halogenases, and non-heme iron-dependent dioxygenases and halogenases, have proved useful for the selective

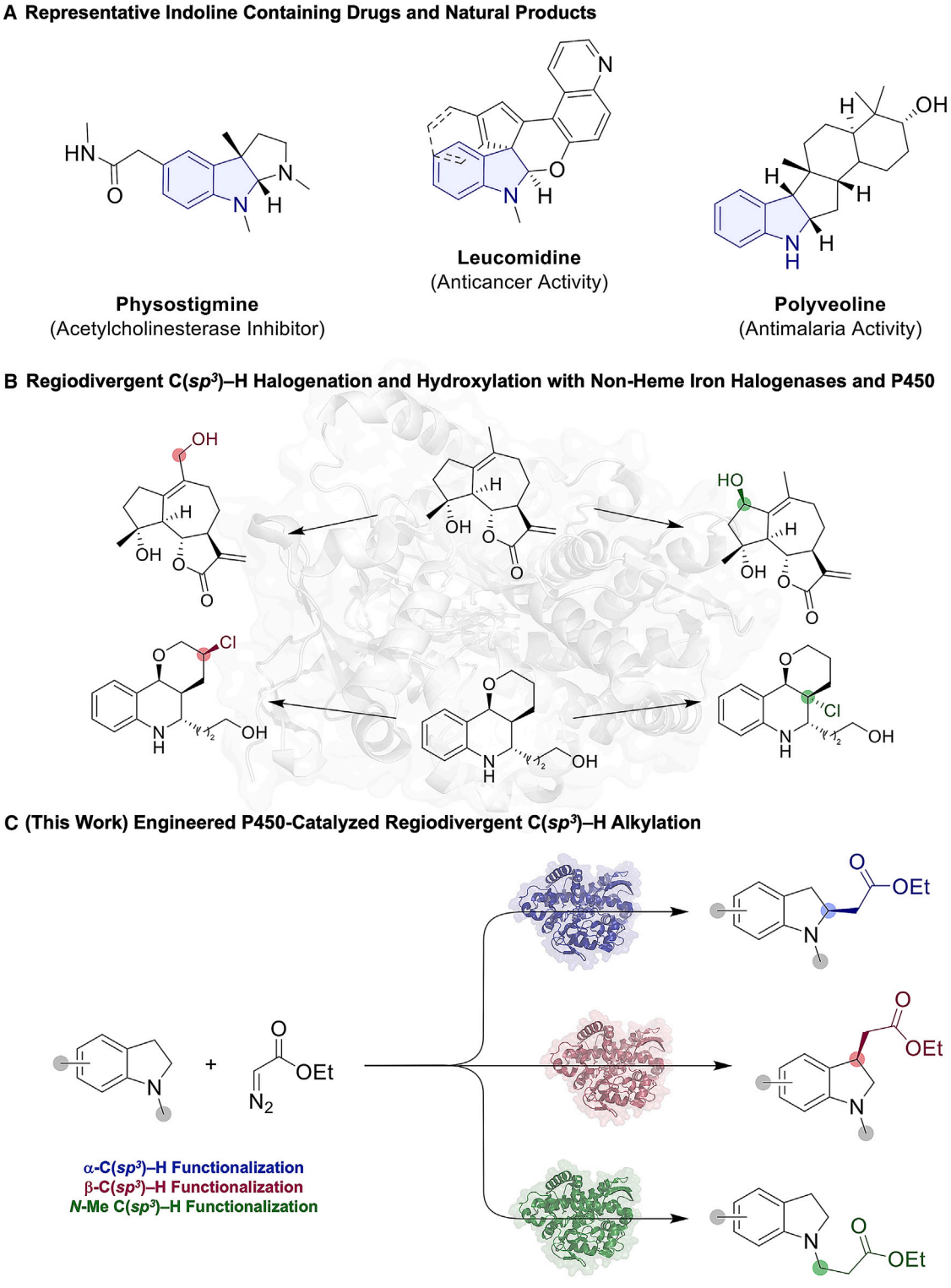


Figure 1. Biological significance of indolines and enzyme-catalyzed regiodivergent C(sp^3)–H functionalization

(A) Representative indoline-containing drugs and natural products.

(B) P450-catalyzed regiodivergent C(sp^3)–H hydroxylation.

(C) Engineered P450-catalyzed regiodivergent C(sp^3)–H alkylation.

functionalization of aromatic and aliphatic C–H bonds in small molecules.[13,14,16,18,21–29] Furthermore, through protein engineering, regiodivergent selectivity has been achieved, for example, for the halogenation of aromatic compounds using evolved flavin-dependent halogenases[22,30,31] or for late-stage C(sp^3)–H hydroxylation of complex natural products using engineered P450 enzymes and halogenases (Figure 1B),[32–39] thus creating new opportunities for the diversification and/or chemoenzymatic synthesis of these molecules. More recently, important progress has also been made in the development of biological catalysts for C–H functionalization via "non-native" chemistry such as carbene transfer catalysis.[40–50] Among them, iron-based cytochrome P450s have constituted attractive systems for achieving selective C(sp^3)–H functionalization in the presence of diazoester-based carbene donor reagents.[46,47,50–52] Despite this progress, and in stark contrast to the scope and regiodivergent selectivity achieved via these enzymes' native reactivity (Figure 1B), these systems and methodologies have been largely limited to a narrow set of substrates and to functionalization of a single C(sp^3)–H site in the target substrate.[45–47,50–52]

Herein, we describe the development of a biocatalytic strategy for the direct C(sp^3)–H functionalization of indolines, a structural subclass found in many bioactive molecules, via P450-catalyzed carbene transfer. In addition to being compatible with a broad range of indoline substrates and diazo reagents, this methodology is shown to allow regiodivergent access to three distinct C(sp^3)–H bonds in the substrate in a highly regio- and enantioselective manner (Figure 1C). As such, this strategy can provide efficient and direct strategy to the synthesis and diversification of these medicinally important scaffolds, whose synthesis is neither straightforward[53] nor readily accessible using metal-catalyzed carbene transfer chemistry. The synthetic utility of the present biocatalysts and methodology is further exemplified through the synthesis of a polycyclic indoline-based core structure akin to that found in many pharmacologically active molecules.[3,54–56] Mechanistic studies provide first-time insights into the mechanism of the present reaction and hemoprotein-catalyzed C(sp^3)–H carbene insertion.

RESULTS AND DISCUSSION

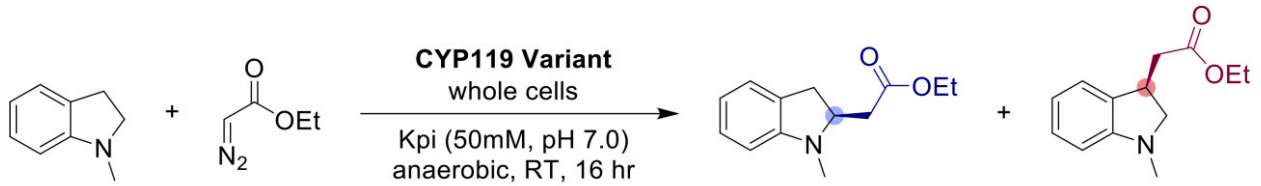
CYP119 biocatalyst for α-C–H functionalization of *N*-methyl indoline with EDA

Motivated by our recent success in developing engineered CYP119 variants for C(sp^3)–H bond functionalization of *N*-aryl-pyrrolidines via carbene transfer with diazoacetone and ethyl diazoacetate,[50] we sought to develop a biocatalytic method for the C(sp^3)–H functionalization of *N*-substituted indolines, which are privileged scaffolds in medicinal chemistry. Of note, this transformation poses a unique challenge in terms of regioselectivity given the presence of two C(sp^3)–H sites with similar reactivities, i.e., at the α and β position with respect to the nitrogen atom.[57] To this end, we began our investigation by screening a small panel (~100) of engineered CYP119 variants derived from our previous work,[50] including CYP119 (F153G, A209G, T213G, V254A, C317S) (called **CYP119-137**). The latter variant features an expanded active site as a result of four space-creating active site mutations and it was previously found to exhibit a pronounced substrate promiscuity toward C–H functionalization of *N*-aryl-pyrrolidines via carbene transfer. These CYP119-based variants also harbor a non-native axial serine ligation, a mutation shown to be beneficial for catalysis of non-native carbene transfer reactions in this[50] and other P450 scaffolds.[58]

The C(sp^3)–H functionalization of *N*-methyl indoline (**1a**) in the presence of ethyl diazoacetate (**2a**, EDA) was investigated as model reaction for this work (Table 1). While wild-type CYP119 shows no activity, the promiscuous variant **CYP119-137** was found capable of converting *N*-methyl indoline (**1a**) in 17% yield (210 turnovers or TON) to yield a mixture of the α-amine C(sp^3)–H functionalization product **3a** and β-C(sp^3)–H functionalization product **4a** in approximately 2:1 ratio (Table 1, entry 4). Along with the desired C–H functionalization products, the **CYP119-137**-catalyzed reaction was accompanied with the formation of various demethylation and desaturation/N–H insertion by-products (Table S5), which have implications with respect to the mechanism of this transformation as discussed later. Importantly, the isolated cofactor hemin along with various organometallic carbene transfer catalysts,[59] such as $Rh_2(OAc)_4$, $Ru(BPY)_2$, Co(TPP), $Cu(OTf)_2$, and Fe(TPP), showed no product formation (Table S1), highlighting the peculiar role of the protein matrix in enhancing the enzyme's reactivity toward this challenging transformation. In addition, the ability of **CYP119-137** to target each of the three distinct C(sp^3)–H bonds in the indoline substrate (i.e., benzylic and the endo and exocyclic α-amino C–H bonds) held promise toward tuning the enzyme's regioselectivity via protein engineering.

Encouraged by these results, we aimed to identify CYP119 catalysts capable of favoring the formation of the α-amine C(sp^3)–H functionalization product **3a** with higher regioselectivity as well as higher chemoselectivity against formation of the undesired demethylation/unsaturation byproducts. To this end, we extended our screening to an in-house library of CYP119-derived variants generated in previous evolution campaigns[50] targeting the partial mutagenesis of active site residues F153, L205, A209, and V254 using (mostly) apolar amino acids of variable size (Ala, Phe, Ile, Leu, Pro, Ser, Thr, Val). The enzyme variants were expressed in *Escherichia coli* C41(DE3) and screened as whole-cell reactions in multi-well plates. These experiments revealed a set of structurally related variants, i.e., CYP119 (F153G, T213A, V254X, C317S, where X is Ile, Leu or Phe) that show a clear beneficial effect of increased steric bulk at the level of position 254 (Phe > Leu > Ile) toward increasing product yield as well as favoring formation of the α-amine C(sp^3)–H functionalization product **3a** over **4a** (Table 1, entries 5–7). Among them, the V254F-containing variant showing the highest levels of catalytic activity and regioselectivity among them (73% yield, 4,180 TON and 87:13 r.r.; Table 1, entry 7; Figure 2B). These variants also exhibited appreciable enantiopreference for formation of the *S*-enantiomer (20%–24% *ee*; Table 1, entry 5–7). Based on this insightful structure-activity data, we chose to investigate the effect of larger aromatic substituents, i.e., tyrosine and tryptophan, at the 254 position. Among them, and in line with the aforementioned trend, the V254W-containing variant CYP119 (F153G, T213A, V254W, C317S), referred to as **CYP119-168**, showed further enhanced catalytic activity (4,180 → 5,270 TON) and regioselectivity (87:13 → 92:8 r.r.) for formation of

Table 1. Intermolecular C–H functionalization of *N*-methyl indoline (1a) with EDA (2a) using hemoprotein CYP119 and variants thereof



1a 2a 3a 4a

Entry	Catalyst[a]	Yield (3a)[b]	3a:4a	TON (3a)[c]	e.r. (3a)[d]
1	hemin	0	0	–	–
2	"empty" *E. coli* cells	0	0	–	–
3	CYP119 (WT)	0	0	–	–
4	[**CYP119-137**] CYP119 (F153A, A209G, T213G, V254A, C317S)	17%	66:34	210	53:47
5	CYP119 (F153G, T213A, V254I, C317S)	26%	70:30	840	58:42
6	CYP119 (F153G, T213A, V254L, C317S)	20%	76:24	1,150	60:40
7	CYP119 (F153G, T213A, V254F, C317S)	73%	87:13	4,180	62:38
8	CYP119 (F153G, T213A, V254Y, C317S)	66%	84:16	3,210	78:12
9	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	92%	92:8	5,270	91:9
10[e]	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	42%	92:8	8,930	91:9
11[f]	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	92%	92:8	460	91:9
12[g]	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	92%	92:8	460	91:9

[a]Standard reaction conditions: protein expressing C41(DE3) *E. coli* cells (OD_{600} = 60), 10 mM **1a**, 80 mM EDA (**2a**), in KPi buffer (50 mM, pH 7), room temperature, 16 hours, anaerobic chamber.
[b]Assay yields as determined by GC using calibration curves with isolated product **3a**.
[c]TON for product **3a** as calculated based on the protein concentration measured in cell lysate using the CO-binding assay.
[d]Enantiomeric ratio (e.r.) of the major product as determined by chiral HPLC.
[e]OD_{600} = 15.
[f]Using 20 μM purified protein and 10 mM $Na_2S_2O_4$.
[g]Using 20 μM lyophilized purified protein and 10 mM $Na_2S_2O_4$.

3a. In addition, **CYP119-168** showed improved enantioselectivity in the reaction (78:12 → 91:9 e.r.) and it catalyzes nearly quantitative conversion of the indoline substrate to the α-C–H functionalized product **3a** with no formation of undesirable byproducts (Table S5).

Additional experiments indicated that up to 8,900 total turnovers (TTN) could be obtained for the **CYP119-168**-catalyzed α-C–H functionalization reaction using whole cells under catalyst limiting conditions (optical density 600 [OD_{600}] = 15) (Table 1, entry 10). Furthermore, this reaction was determined to proceed with equally high yields and enantioselectivity using purified protein at 0.2 mol % as catalyst (Table 1, entry 11) and similar results could be obtained using the same CYP119 variant in lyophilized form (Table 1, entry 12), thus demonstrating the robust nature of this biocatalyst to lyophilization and long-term storage, which are desirable attributes for preparative scale and industrial applications.

Given our prior findings on the C–H carbene insertion reactivity of cobalt-substituted myoglobins using phthalan as the substrate,[51] we investigated the effect of a similar metal substitution in the present CYP119-based biocatalysts, which were prepared via recombinant expression in the presence of Co-ppIX.[51,60] Interestingly, all of these Co-substituted variants showed activity toward the C–H functionalization of *N*-methyl indoline (**1a**) with EDA (**2a**) (Table S2). Although the activity and regioselectivity of these metallo-substituted variants were only comparable or slightly inferior compared to those of the iron-containing counterparts, the functionality of these enzymes as carbene transferases is notable and it could prove useful for other types of non-native transformations.

C–H functionalization of indoline-based substrates

Focusing on the best biocatalyst identified for α-C–H functionalization of **1a**, **CYP119-168**, we next explored the substrate scope of this enzyme using a range of aryl and *N*-substituted indoline compounds (**1a–1o**). Notably, each of these substrates underwent α-C–H functionalization in the presence of EDA with excellent regio- and chemoselectivity, i.e., showing no formation of the potentially competing β-C–H functionalization product and byproducts, respectively. However, these reactions were characterized by variable yields and enantioselectivity (Figure 3C), indicating a certain degree of substrate specificity as observed in other P450-catalyzed native and non-native reactions.[50,61,62] To address this limitation, we employed a substrate versus enzyme library approach, where a set of selected CYP119-derived carbene transferases from the **CYP119-168** lineage and other generations were screened against the substrate panel in a high-throughput manner (~2,000 substrate/enzyme

combinations). From these experiments, a subset of evolved CYP119 variants was shown to catalyze the α-C–H functionalization of each indoline-based substrates with high activity (2,900–6,540 TON) and good to excellent enantioselectivity (up to 96% *ee*; Figure 3A).

In particular, both electron-withdrawing (**3c** and **3f**) and electron-donating (**3d** and **3g**) substitutions on the aryl ring were well tolerated, resulting in high levels of activity and enantioselectivity (85%–94% yield, 5,740–6,310 TON, 83:17–97:3 e.r.). Racemic 3-methyl-substituted indoline derivative **1b** was efficiently converted into the desired product **3b** with both excellent diastereo- and enantioselectivity (99:1 d.r., 98:2 e.r.). This demonstrates the ability of the enzyme to induce kinetic resolution of racemic starting materials, a valuable trait in the context of stereoselective catalysis. In addition to methyl group as *N*-substituent, a variety of linear (**3f–h**, **3j**), branched (**3i**), and cyclic (**3k–n**) alkyl groups were found to be tolerated at this position, with a slight decrease in yield and/or enantioselectivity for bulkier *N*-cycloalkyl substitutions (i.e., **3m–n** vs. **3h–k**). In contrast, *N*-aryl-indolines such as **3o** were converted only with trace activity. Enzymatic conversion of *N*-alkyl indolines was initially affected by the low solubility of these compounds in aqueous media even in the presence of organic cosolvent (10% v/v EtOH). Improved yields for these reactions (e.g., 27% → 44% yield **3n**) could be achieved by raising the reaction temperature to 40°C, a readily applicable condition thanks to the thermostability of the CYP119 variants. Notably, substrate **1j** could be converted to the desired α-C–H functionalization product **3j** with high regio- and chemoselectivity and without affecting the terminal olefinic group, highlighting the ability of the CYP119-based catalyst to favor the more challenging C–H carbene insertion reaction over cyclopropanation, unlike organometallic catalysts used for carbene transfer reactions (Figure 3B).

Whereas different CYP119 variants can be selected for optimal yield and enantioselectivity for each substrate (Figure 3A), each CYP119 variant is capable of catalyzing the desired reaction (Figure 3C). These factors show the generality of our engineered CYP119 variants toward the α-C–H functionalization of indoline-based derivatives. These results also revealed the role of steric bulk in promoting favorable binding conformations to target α-C–H functionalization compared to β-C–H functionalization. In the presence of bulkier *N*-alkyl substitutions, no β-C–H functionalization product is formed (**3f-o**). We also noted an inverse correlation between steric bulk in the substrate and in the catalyst active site, as catalysts that contain bulkier residues within the active site tend to exhibit lower activity toward bulkier substrates. For example, **CYP119-168** and **CYP119-156**, which contain a V254W and A209W mutation, showed basal activity toward substrate **1n**, bearing an *N*-cyclohexyl moiety.

Finally, efficient and selective α-C–H functionalization of *N*-methyl indoline (**1a**) could be achieved using both diazoacetone (**2b**) and diazoacetonitrile (**2c**) as carbene donors, yielding the respective products (**6a** and **7a**) in quantitative yields (97%–99%) and with high catalytic activity (>6,000 TON), regioselectivity (100:0 r.r.) and enantioselectivity (92%–96% *ee*) (Figure 3A). Altogether, these results demonstrate the generality of this CYP119-based methodology for α-C–H functionalization of a broad range of *N*-substituted indolines with high catalytic efficiency and selectivity.

Regiodivergent CYP119 catalysts for indoline C–H functionalization

During initial catalyst development for achieving selective α-C–H functionalization of *N*-methyl-indoline, we noted that early generations of enzymes bearing more open active sites showed appreciable regioselective toward formation of the β-C–H functionalization product (**4a**) (e.g., **CYP119-137**: 66:34 **3a**:**4a**; Table 1, entry 4). Building upon this finding, we sought to develop a regiocomplementary catalyst capable of selectively targeting the β-position (Figure 4A). To this end, we selected variants from the in-house CYP119 library that contained bulky residues (Phe, Tyr, Trp) at positions L205, L69, and A205, which are located on the opposite site of the active site compared to the Val254 residue shown to be instrumental in tuning regioselectivity to the α-C–H functionalization product (Figures 2B and 2C).

Through this approach, we identified CYP119 (F153G, L205W, T213A, V254A, C317S), named **CYP119-235**, as an efficient and selective biocatalyst for functionalization of the β-C–H position in *N*-methyl-indoline (**1a**), producing **4a** in high yield and TON (78% yield, 5,380 TON) as well as high enantio- and regioselectivity (94:6 e.r., 16:84 **3a**:**4a**; Figure 4D). As anticipated, the active site mutations in **CYP119-235** include a bulky substitution (i.e., L205W) (Figure 4C) on the opposite side of the enzyme active site compared to a similar bulky substitution (i.e., V254W) harbored by the α-C–H selective variant **CYP119-168** (Figure 4B).

Given the value of 2,3-difunctionalized indolines in medicinal chemistry (Figure 1) and our success in the kinetic resolution of racemic **1b** with 96% *ee* and 98% *de* (**3b**, Figure 3A), we challenged the β-C−H selective variant **CYP119-235** with racemic *N*-methyl-2-methyl-indoline (**1p**). Importantly, substitution at the α-position was tolerated by the enzyme and the racemic substrate could be converted to the optically active product **4b** in good diastereomeric and enantiomeric excess (67:33 d.r., 79:21 e.r.), albeit with low efficiency (12% yield). Using a different CYP119 variant (**CYP119-137**), this transformation could be carried out with further increased enantio- and diastereoselectivity (92:8 e.r., 81:19 d.r.), illustrating the potential value of this biocatalytic systems for kinetic resolution applications.

Upon exploration of the evolved variant **CYP119-235**, we observed that the regioselectivity of this enzyme could be directed to the functionalization of the *N*-methyl C–H bond in the presence of substitutions at the C_5 (**5e**) and C_6 (**5f**) position on the aryl ring (Figure 4D). These reactions were found to proceed with full regiocontrol and excellent catalytic activity (89%–92% yield, 6,140–6,340 TON), complementing the scope of the α- and β-C–H functionalization reactions catalyzed by the CYP119 catalysts. Altogether, these results demonstrated the capability of the present methodology to target as many as three different C(sp^3)–H sites in a substrate for C–H functionalization via carbene transfer, a feature unprecedented for carbene transferases and rarely achieved with synthetic carbene transfer catalysts.[64,65]

Polycyclic indolines via tandem enzyme-catalyzed C(sp^3)–H carbene insertion

As described above, our investigation of the substrate scope of the engineered CYP119 variants established that both α-C–H and

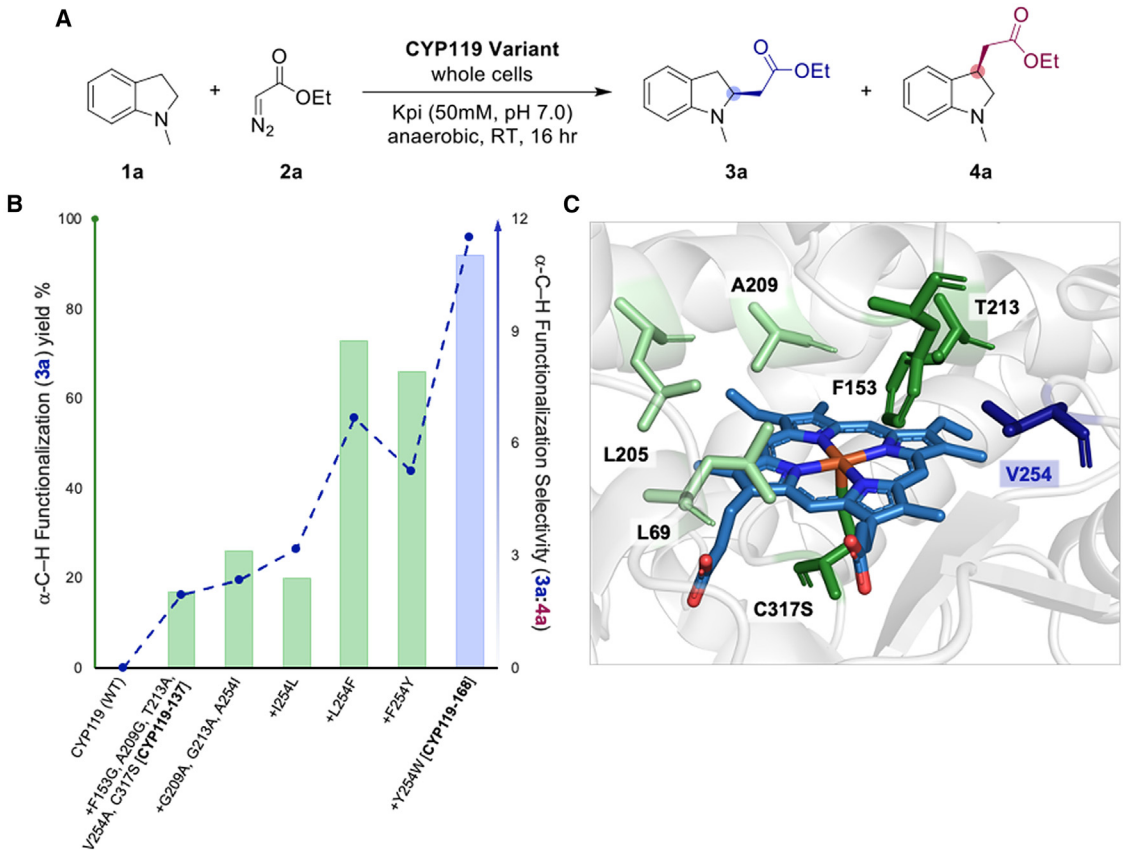


Figure 2. Directed evolution of CYP119 catalysts for regioselective C–H functionalization of *N*-methyl indoline (1a) with EDA (2a)

(A) General scheme for C–H functionalization of *N*-methyl indoline (**1a**) with EDA (**2a**) to form the α-functionalization product **3a** and β-functionalization product **4a**.

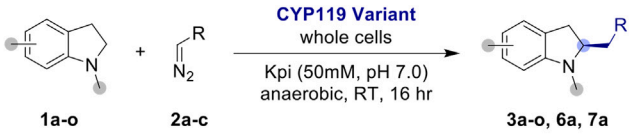
(B) Reconstructed directed evolution of CYP119 catalysts for regioselective C–H functionalization of *N*-methyl indoline (**1a**) with EDA (**2a**). Yields as determined under standard reaction conditions with EDA (Table 1).

(C) X-ray crystal structure of CYP119 from *Sulfolobus solfataricus* (PDB: 1IO7).[63] Active-site residues targeted for mutagenesis are highlighted in dark green, conserved active-site residues are highlighted in light green, active-site residue V254 is highlighted in dark blue, and the heme cofactor is shown in teal.

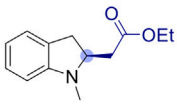
β-C–H selective biocatalysts can tolerate substitutions in the adjacent position in the indoline substrate targeted for functionalization (i.e., **3b** and **4b**). In addition, with product **6a**, we established that diazoketones are accepted as carbene donors by these catalysts. Based on these results, we envisioned the possibility to construct a tricyclic indoline-based scaffold via a tandem CYP119-mediated C–H functionalization strategy, in which the enzymatic β-C–H insertion product **4a** is converted into the corresponding diazo ketone (**8**) and then subjected to an intramolecular C–H insertion to form the desired polycyclic compound **9a** (Figure 4E). The latter is akin to the indoline-based core structure found in various natural products and pharmaceuticals (Figure 1A). Toward this goal, we performed a preparative scale (500 mg, 3.75 mmol) reaction of *N*-methyl indoline (**1a**) in the presence of EDA (**2a**) using the β-C–H selective CYP119 variant **CYP119-235** to afford **4a** in 67% isolated yield. The ester group of the enantioenriched **4a** was then chemically converted to the respective diazoketone **8** in 45% yield over three steps using established chemistry (see supplemental information for details). Gratifyingly, the diazoketone intermediate **8** could be then effectively cyclized by the α-C–H selective variant, **CYP119-168**, to afford the desired polycyclic product **9a** in high yield (64%) and high enantio- and diastereomeric excess (>500:1 d.r.; 97:3 e.r.; Figure 4E). These results highlight utility of the engineered CYP119 library for the synthesis of biologically relevant synthons.

Mechanistic insights into CYP119-catalyzed indoline C–H functionalization

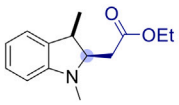
As noted earlier, the C–H functionalization reaction with *N*-methyl-indoline and EDA catalyzed by the early, unoptimized CYP119 catalysts is accompanied by the formation of three distinct by-products, which were determined to correspond to *N*-methyl-indole (**1a-DS**), C_3-functionalized *N*-methyl indole (**4a–DS**), and the *N*-alkylated-indoline **13a** (Table S5, entry 2). While these side-reactions could be suppressed in the presence of the optimized biocatalysts and reaction conditions, they bear important mechanistic implications. Indeed, as summarized in Figure 5, all of these side products can be explained based on desaturation reactions involving substrate-derived radical intermediates, which suggests a radical, stepwise mechanism for the present C–H carbene insertion reaction also in line with the results from computational analyses described further below.



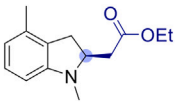
A



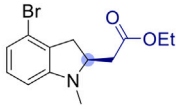
3a, [**CYP119-168**]
92% [83%]
91:9 e.r.
92:8 **3a**:**4a**
5,270 TON



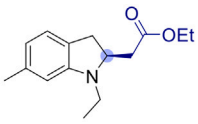
3b, [**CYP119-235**]
42%
98:2 e.r.
99:1 d.r.
2,900 TON



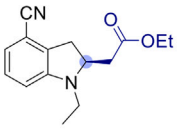
3c, [**CYP119-250**]
85%
83:17 e.r.
6,200 TON



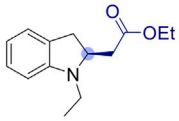
3d, [**CYP119-090**]
94%
97:3 e.r.
6,310 TON



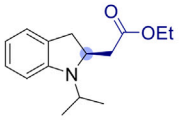
3f, [**CYP119-156**]
93%
84:16 e.r.
5,740 TON



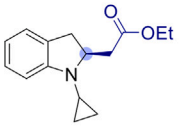
3g, [**CYP119-137**]
89%
93:7 e.r.
5,980 TON



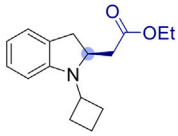
3h, [**CYP119-137**]
97%
84:16 e.r.
6,520 TON



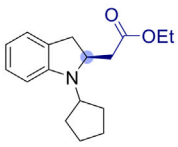
3i, [**CYP119-235**]
84%[a]
96:4 e.r.
5,790 TON



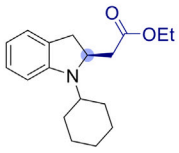
3k, [**CYP119-311**]
77%[a]
90:10 e.r.
6,540 TON



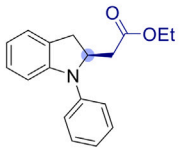
3l, [**CYP119-282**]
67%[a]
66:34 e.r.
4,390 TON



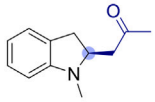
3m, [**CYP119-282**]
52%[a]
60:40 e.r.
3,400 TON



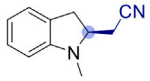
3n, [**CYP119-235**]
44%[a]
87:13 e.r.
3,030 TON



3o, [**CYP119-235**]
Trace[a]



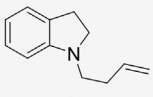
6a, [**CYP119-090**]
97%[b]
98:2 e.r.
6,520 TON



7a, [**CYP119-090**]
99%[b]
96:4 e.r.
6,650 TON

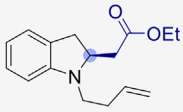
Alternate Carbene Donors

B



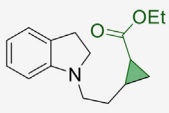
1j

CYP119-235
EDA (**2a**)



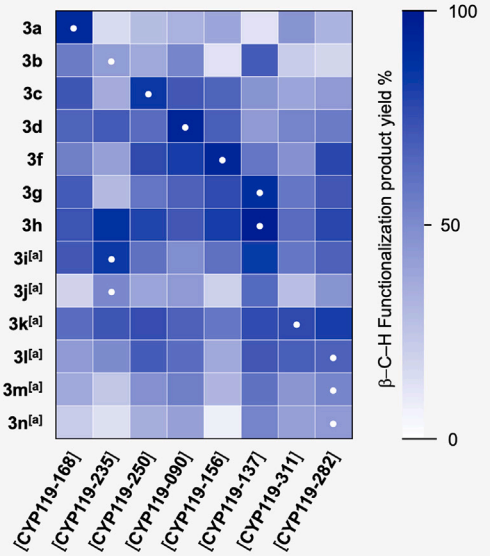
3j, [**CYP119-235**]
51%[a]
80:20 e.r.
3,520 TON

$Rh_2(OAc)_4$
EDA (**2a**)



3j
34%
73:27 d.r.

C



(legend on next page)

For the productive pathway leading to the C–H functionalized products, the reactive heme-carbene species derived from reaction of ferrous protein with EDA is proposed to abstract a hydrogen atom from either the α- or β-C–H site with respect to the indoline N–H, with regioselectivity being controlled by the enzyme's active site configuration (*vide infra*) and resulting in the C-centered radical intermediates **IM$_{I}$** or **IM$_{II}$**, respectively. Radical rebound with the iron-carbenoid species yields the corresponding C–H functionalization products (e.g., **3a** or **4a**). For the side reaction leading to the desaturation product **1a-DS**, we envision the radical intermediate **IM$_{I}$** or **IM$_{II}$** may undergo radical polar crossover (RPC), e.g., via single-electron transfer to the heme cofactor, to give the respective carbocation **IM$_{IV}$** or **IM$_{V}$**, which, upon deprotonation, yield *N*-methyl-indole **1a-*DS*** (Figure 5). Alternatively, the latter can be produced from the radical intermediate **IM$_{I}$** (or **IM$_{II}$**) via a second HAA mediated by the Fe(III)-alkyl species (*vide infra*). Regardless of the nature of this step, the ensuing desaturation product *N*-methyl-indole (**1a-*DS***) can be then converted by the enzyme to **4a-DS** via an indole C(3)–H functionalization reaction with EDA akin to that previously reported by our group[44] and others[66,67] for other engineered hemoproteins.

On the other hand, formation of the *N*-alkylated-indoline product **13a** can be rationalized based on a first step involving enzymatic *N*-demethylation of *N*-methyl-indoline **1a**, followed by N–H carbene insertion of the resulting indoline to give **13a** (Figure 5). Consistent with our results with **5e** and **5f**, hydrogen atom abstraction at the level of *N*-methyl group is also accessible to this biocatalyst, resulting in the carbon-based radical **IM$_{III}$** which, upon RPC, is expected to produce the iminium intermediate (**IM$_{VI}$**). The latter can then undergo hydrolysis to form indoline, which can then give rise to **13a** via N–H carbene insertion with EDA, a known reaction for engineered hemoproteins.[68–70]

Various lines of experimental evidence support the proposed reaction pathways. Since the model substrate *N*-methyl-indoline was found to be susceptible to (slow) desaturation to indole in the presence of air and further modification of the CYP119 product could occur by action of other enzymes in whole-cell reactions, control experiments were first performed to confirm the enzymatic origin of the observed side products. To this end, time-course experiments were carried out in the presence of *N*-cyclopropyl-indole (**1k**), which is stable toward oxidative desaturation, and the unselective variant **CYP119-282** (in purified form) as the catalyst. Under catalytic (air-free) conditions, formation of the C–H insertion product **3k** is accompanied by accumulation of desaturated by-products **1k-*DS*** and **4k-*DS*** in approximately 73:9:18 ratio over 90 min, with corresponding initial formation rates (TOF) of 28, 5.7, and 2.5 turnovers per minute for **3k**, **1k-*DS***, and **4k-*DS***, respectively (Figure S2). In contrast, no formation of either desaturation byproduct was observed in the absence of the enzyme or EDA within the same conditions (Figure S2), clearly indicating that these species are enzymatic products. To further investigate the sequence of reactions leading to the C_3-functionalized *N*-alkyl-indole product (i.e., **4a-DS** from **1a** or **4k-DS** from **1k**), products **4a** and **1a-*DS*** were used as substrates under standard catalytic conditions in the presence of **CYP119-137**. While **4a** was fully preserved in the reaction mixture, **1a-*DS*** was consumed to give rise to the C_3-functionalized product **4a-*DS*** (Figure S8). Although the mechanism of this step was not investigated, previous studies with engineered myoglobin support a stepwise mechanism involving a zwitterionic intermediate.[44,71]

N-Demethylation of *N*-methyl-indoline via the proposed mechanism in Figure 5 implies the release of formaldehyde as byproduct. To test this, the enzymatic reaction mixture was added with the formaldehyde trapping agent *O*-(2,3,4,5,6-pentafluorobenzyl)hydroxylamine (PFBHA, **14**), followed by GC analysis to detect the corresponding formaldehyde adduct (oxime). As anticipated, a detectable amount of the PFBHA-derived oxime was detected in this reaction and this species was found to be proportional to the concentration of the demethylated/*N*-alkylated product **13a** (Figure S7). In separate experiments, we further determined that indoline is readily converted by the enzyme to the *N*-alkylated product **13a** under standard catalytic conditions in the presence of EDA (Figure S8). Collectively, these results support the proposed mechanistic pathways for formation of the experimentally observed products and byproducts in this reaction (Figure 5).

While the application of reducing conditions (i.e., sodium dithionite or intracellular environment) in the current protocol entails the involvement of ferrous hemoprotein as carbene transfer catalyst, we further explored the importance of the redox state of the protein in this reaction, also in light of recent reports on the activity of iron(III)-metalloporphyrins for carbene transfer reactions.[72] In the absence of the reductant, the **CYP119-168**-catalyzed reaction with **1a** and EDA proceeds with reduced catalytic activity (45% → 32% yield; Table S6, entry 1–2) but identical regio- and enantioselectivity as compared to that under reducing conditions. Further experiments were conducted to discern whether the former activity stems from the ferric protein or if the hemoprotein is reduced *in situ* by the diazo compound, as previously observed by our group for certain axial substituted myoglobin-based carbene transferases.[73] Accordingly, the same reactions were carried out in the presence of CO-saturated buffer, with CO being expected to bind with high affinity to only the ferrous form of the protein, thus inhibiting its reactivity. Under these conditions, with or without an external reductant, no carbene transfer activity was observed (Table S6, entry 3–4), indicating that ferrous CYP119 is the catalytic species responsible for formation of the C–H functionalization products.

Figure 3. Activity and selectivity of CYP119 biocatalysts for α-C–H functionalization of indoline-based substrates derivatives (1a–o) with EDA

(A) Yields, TON, and enantioselectivity were determined from whole-cell reactions under standard reaction conditions with EDA as described in Table 1. Analytical yields were determined using GC/LC using calibration curve prepared with isolated products. Isolated yields are indicated in brackets. TON as determined based on P450 concentration in cell lysate. (a) Using 240 mM EDA (**2a**) at 40°C (b) Using 240 mM diazoacetone (**2b**) or diazoacetonitrile (**2c**).

(B) Chemodivergent reactivity of CYP119 **CYP119-253** and $Rh_2(OAc)_4$ (10 mol %) with 1-(but-3-en-1-yl)indoline (**1j**) and EDA (**2a**).

(C) Heatmap depicting intermolecular C–H functionalization of *N*-substituted indolines (**1a–n**) with EDA (**2a**) using hemoprotein CYP119 variants. White dots (•) indicate the best variant for the corresponding substrate. Details about the variants are provided in Figure S1.

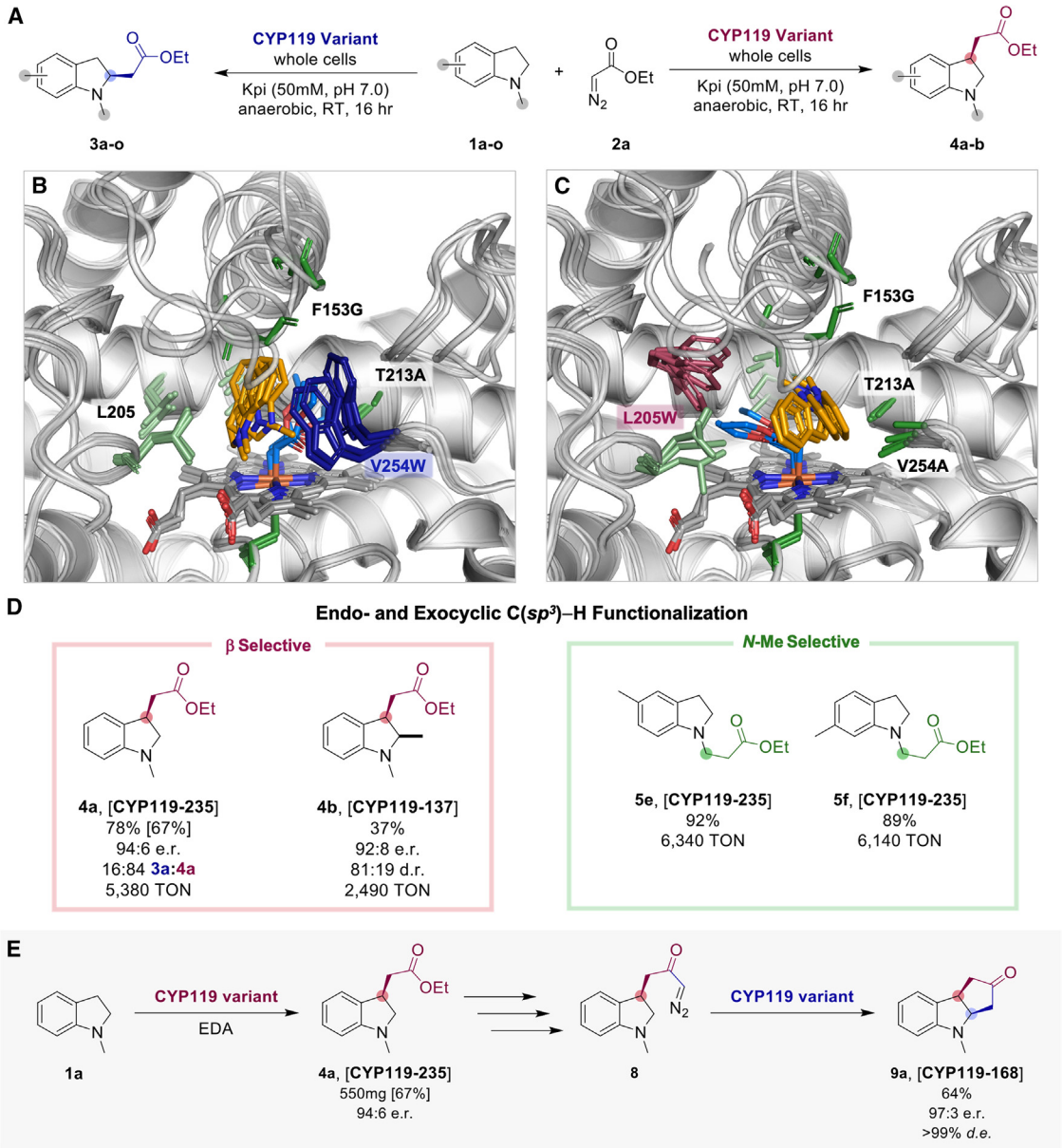


Figure 4. CYP119-catalyzed regiodivergent C–H functionalization of *N*-methyl indoline (1a) with EDA (2a)

(A) Regiodivergent pathway for C–H functionalization of *N*-methyl indoline (**1a**) with EDA to form α-alkylated product **3a** and β-alkylated product **4a**.

(B and C) Active site views of representative Rosetta models for **CYP119-168** in complex with DFT-calculated α-pro(*S*)-C–H transition state (B) and of **CYP119-235** in complex with DFT-calculated β-pro(*R*)-C–H transition state (C). The heme cofactor is shown as a line model, while mutated active site residues (green), *N*-methyl indoline (orange), and heme-bound carbene (blue) are shown as stick models. See Figure S10 for additional data.

(D) Benzylic and *N*-methyl C–H functionalization products. Yields, TON, and enantioselectivity were determined from whole-cell reactions under standard reaction conditions as described in Table 1.

(E) Secondary functionalization pathway for the intramolecular C–H functionalization of diazo tethered *N*-methyl indole (**8**) into product **9a**.

Finally, further mechanistic insights were gained using the deuterium-labeled substrate *N*-methyl-2-D_2-indoline (**1a-*D*$_2$**). In the presence of the α-C–H selective variant **CYP119-168**, enzymatic transformation of **1a-*D*$_2$** in the presence of EDA (**2a**) showed a notable change in the overall product distribution (vs. reaction with **1a**) to favor formation of the β-C–H functionalized product **4a-*D*$_2$**, along with the *N*-methyl-indole byproducts **1a-DS-*D*$_2$** and **4a-DS-*D*$_2$** (92:8 → 68:19:4:9; Figure 6). Insightfully, and in line with the overall mechanism of Figure 5, these results indicate that (1) the H atom abstraction (HAA) step is regioselectivity determining, and (2) that the carbon radical intermediate derived from β-HAA represents a pathway branching point toward formation of the desaturation byproduct (indole). Indeed, as the energy barrier for α-HAA is increased

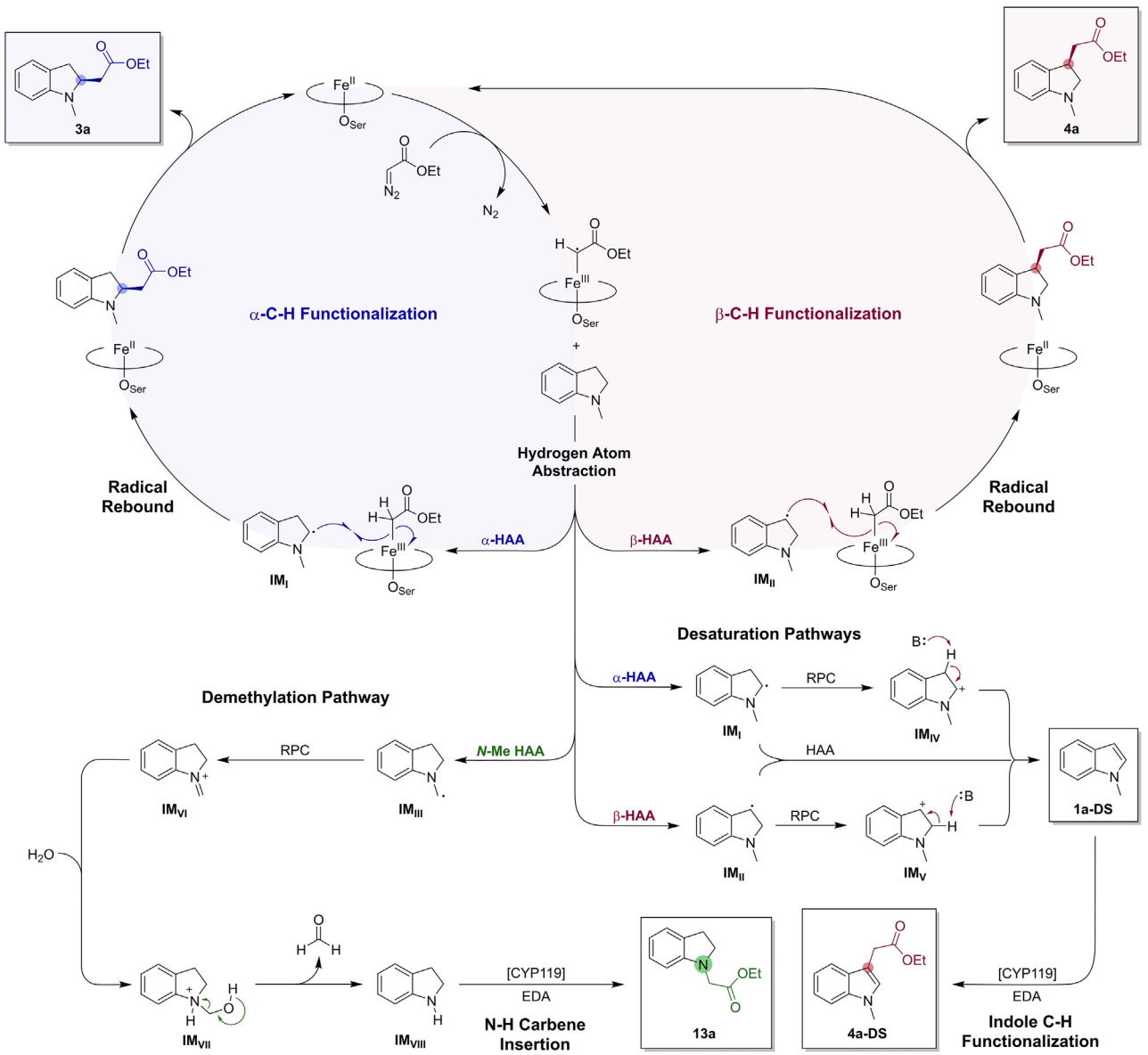


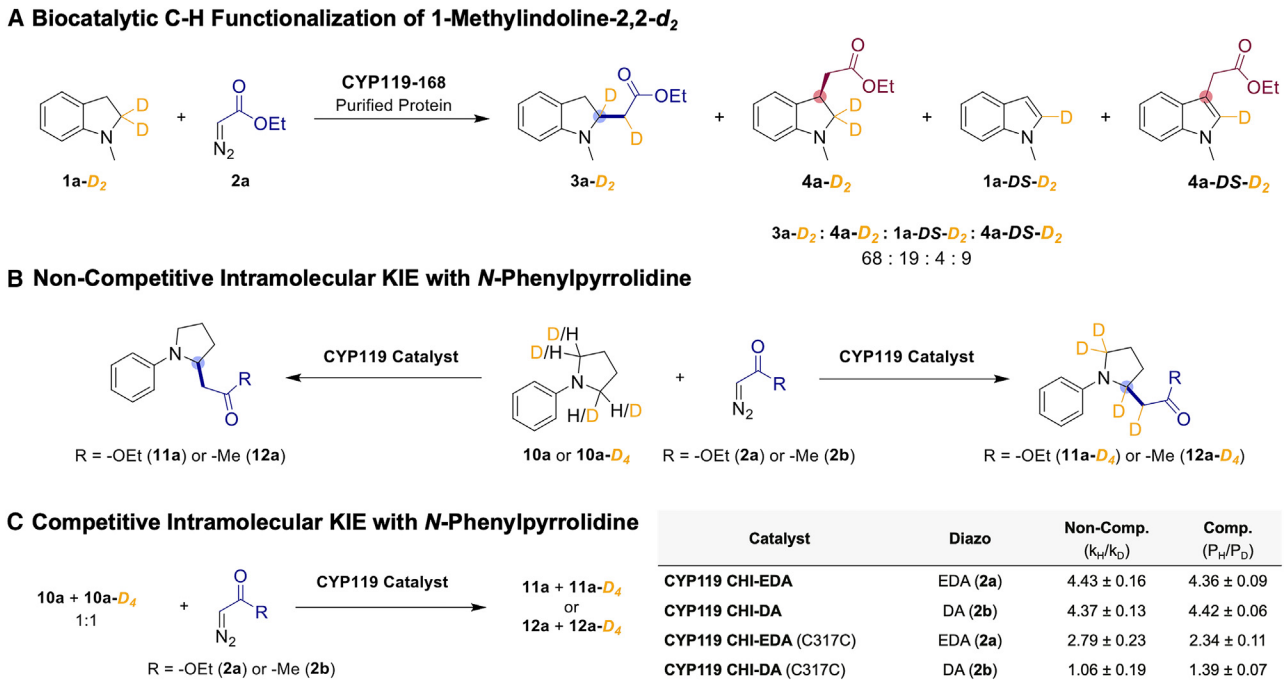
Figure 5. Proposed stepwise radical-mediated mechanism for the C–H functionalization of *N*-methyl indoline (1a) in the presence of engineered CYP119 catalysts

by the H → D substitution, it is conceivable that the reaction pathway is partially diverged to favor the β-HAA manifold, thereby leading to **4a-*D*$_2$** via the productive pathway, or to **1a-DS-*D*$_2$**, via the unproductive side reaction. Since no desaturation byproduct is observed in **CYP119-168**-catalyzed reaction with **1a**, it can be further evinced that the desaturation products **1a-DS-*D*$_2$** (and **4a-DS-*D*$_2$**) primarily derive from the β-HAA pathway (vs. α-HAA pathway), possibly due to a slower radical rebound step and/or more favorable conformation for RPC.

It is worth noting that these desaturation reactions are reminiscent of those described for certain cytochromes P450 in non-classical oxidation reactions under aerobic conditions.[74–76] While the cationic mechanism proposed in Figure 5 is plausible based on computational studies of these analogous reactions,[77] an alternative desaturation mechanism entailing a second HAA on the radical intermediate **IM$_I$** or **IM$_{II}$** by the Fe(III)-alkyl species may be also operative, in analogy to desaturation reactions catalyzed by other metalloenzymes.[78] Further studies will thus be required to elucidate this aspect in more detail.

Mechanistic studies

To further illuminate the mechanism of the present P450-catalyzed C–H carbene insertion, we investigated our previously reported C(sp^3)–H functionalization reaction of aryl pyrrolidines



Catalyst	Diazo	Non-Comp. (k_H/k_D)	Comp. (P_H/P_D)
CYP119 CHI-EDA	EDA (**2a**)	4.43 ± 0.16	4.36 ± 0.09
CYP119 CHI-DA	DA (**2b**)	4.37 ± 0.13	4.42 ± 0.06
CYP119 CHI-EDA (C317C)	EDA (**2a**)	2.79 ± 0.23	2.34 ± 0.11
CYP119 CHI-DA (C317C)	DA (**2b**)	1.06 ± 0.19	1.39 ± 0.07

Figure 6. Kinetic isotope effect experiments

(A) Intermolecular non-competitive KIE experiment with *N*-methylindoline-2,2-d_2 (**1a–*D*$_2$**) and EDA (**2a**).

(B) Intermolecular competitive and non-competitive KIE experiments with *N*-phenylpyrrolidine (**10a**) or *N*-phenylpyrrolidine-2,2,5,5-d_4 (**10-*D*$_4$**) and EDA (**2a**) or DA (**2b**).

(C) Intermolecular competitive and non-competitive KIE experiments with *N*-phenylpyrrolidine (**10a**) or *N*-phenylpyrrolidine-2,2,5,5-d_4 (**10-*D*$_4$**) and EDA (**2a**) or DA (**2b**). Elaborated KIE data in Figures S3 and S4.

with EDA and diazoacetone, as these and the present reaction share the same catalytic system (i.e., serine-ligated CYP119).[50] Importantly, and unlike the indoline reaction, kinetic isotope effect (KIE) values can be more readily measured for the pyrrolidine reactions due to the absence of side products and pathway branching points. Accordingly, KIE values were calculated from both competitive and non-competitive intermolecular KIE experiments using deuterated and non-deuterated *N*-phenyl pyrrolidine (**10a-*D*$_4$** and **10a**) in the presence of both EDA (**2a**) and diazoacetone (**2b**) and in combination with the respective optimized CYP119 catalysts (**CHI-EDA** and **CHI-DA**). In both cases, a relatively large primary KIE value of 4.36–4.43 was measured for both competitive and non-competitive (parallel) kinetic experiments (Figures 6B and 6C). The similar values in both experimental settings indicate that the C–H bond cleavage step is part of the rate-determining step. In addition, the nature of the carbene donor (i.e., diazoester vs. diazoacetone), and thus of the corresponding heme-carbene, has no noticeable effect on the kinetic role of the HAA step. To further investigate the role of the non-native serine axial ligand, the same KIE experiments were also carried out using cysteine-ligated versions of the aforementioned CYP119 variants, i.e., **CHI-EDA** (S317C) and **CHI-DA** (S317C) (Figures 6B and 6C). Interestingly, much smaller KIE values were obtained in both cases compared to the serine-ligated counterparts. For the reaction with EDA, KIE values of 2.34–2.79 were obtained from competitive and non-competitive reactions against a value of 4.36–4.43 measured with **CHI-EDA**. Similarly, the reactions of the cysteine-ligated **CHI-DA** (S317C) enzyme with diazoacetone yielded moderate to no KIE of 1.39 (compet.) and 1.06 (non-compet.) against a much larger KIE of 4.37–4.42 observed for the serine-ligated counterpart. These differences indicate a notable change in the kinetic impact of C–H bond cleavage as a result of the change in heme axial coordination environment, further highlighting the often critical role of the axial ligand for influencing reactivity and the mechanism of hemoprotein-based carbene transferases.[79–81]

A quantum chemical study was performed to further examine the mechanism of C−H carbene insertion catalyzed by the engineered P450s described here. DFT calculations were carried out using the previous method that accurately predicted various experimental structures and reactivities of heme carbenes,[80,82–87] using [Fe(Por)(MeO$^-$)] to mimic the active site core part of the biocatalyst as done previously,[80,86,87] where Por is a non-substituted porphyrin and MeO$^-$ is the model for the Ser ligand.[88–91] Since conformations and spin states may influence the reaction results,[84–86,92] we first conducted a detailed study of these effects involving both the substrate and heme catalyst to select the most favorable conformations and spin states of the species along the reaction pathways, as described in more detail in the supplemental information.

As shown in Figure 7A, both concerted and stepwise mechanisms were investigated. The concerted reaction pathway is

similar to that reported previously on C−H insertions catalyzed by a different iron porphyrin carbene[84] and it features a simultaneous hydride transfer (as shown by a significant negative charge transfer from substrate to carbene, −0.441 e), C−C′ bond formation, and Fe−C bond breaking. In contrast, the stepwise reaction pathway entails a hydrogen atom transfer (HAT) from the substrate to the heme carbene to form a carbon-center radical intermediate, followed by radical rebound to form the new C−C′ bond, while breaking the Fe−C bond (Figure 7B). In the present system, a hydrogen atom transfer feature can be seen from the increase in the C′−H_1 bond length from 1.095 Å in $\mathbf{R_2}$ to 1.354 Å in $\mathbf{TS_1}$ and then to 3.399 Å in **Int**, while C−H_1 bond length decreases from 1.338 Å in $\mathbf{TS_1}$ to 1.086 Å in **Int** (see Table S20). The most favorable spin state calculated for $\mathbf{TS_1}$ is the open-shell singlet, where the radical is equally shared between C (−0.496 e) and C′ (−0.445 e), whereas the hydrogen atom carries spin densities in the opposite direction (0.090 e), indicating partial transfer of the hydrogen atom. In this electronic state, Fe is in the ferric form with spin densities of +0.995 e (Table S22). In the intermediate **Int**, after donating the hydrogen atom, the substrate C′ has the radical with spin densities of 0.995 e, see Table S22). A subsequent radical rebound lead to the formation of the C−C′ bond in $\mathbf{TS_2}$, which has lower energy compared to the hydrogen atom transfer step (−3.77 kcal/mol, see Figure 7B). In this step, again both C and C′ show radical feature, but with opposite spin directions (−0.438 e and 0.641 e respectively), ready for a radical coupling to facilitate the formation of the final product. While partial C−C′ bond formation can be seen by its distance shortening from 4.436 Å in **Int** to 2.543 Å in $\mathbf{TS_2}$, a concomitant Fe−C bond elongation of ∼0.3 Å in this step also indicates partial cleavage of that bond, which proceeds to the final release of the product.

Overall, our calculations indicate that the stepwise reaction pathway exhibits a significantly lower energy barrier (>9 kcal/mol difference) compared to the concerted reaction pathway (see Figure 7B), suggesting that the stepwise radical mechanism is preferred. Furthermore, the calculated KIE values for the concerted C–H carbene insertion step via **TS** and for the first step of stepwise pathway ($\mathbf{TS_1}$) are 2.14 and 4.63, respectively, the latter being closer to the experimentally determined KIE values of 4.3–4.6 for this reaction (Figure 6). To further understand the origin of the reactivity differences between these two mechanisms, the geometric data were examined in more detail. As seen in Table S11, the overall structures of **TS** and $\mathbf{TS_1}$ are similar except that **TS** has a significantly longer Fe−L (1.987 Å) and shorter Fe−C (1.946 Å) bonds compared to $\mathbf{TS_1}$ (Fe−L, 1.912 Å; Fe−C, 1.982 Å). On the one hand, the shorter Fe−L distance in $\mathbf{TS_1}$ results in a stronger *trans* effect that pushes the carbene moiety away from iron center (longer Fe−C bond), thus facilitating attack on the substrate. On the other hand, breaking of the shorter Fe−C bond (1.946 Å) in the concerted **TS** is associated with a higher energy cost compared to cleavage of the Fe−C bond in the radical pathway $\mathbf{TS_2}$, which is longer (2.345 Å). Altogether, these structural features contribute to favor the stepwise radical mechanism for this biocatalytic transformation.

In addition, we computationally studied the rate-determining steps of the C–H insertion reaction with EDA and *N*-methyl indoline **1a** as the substrate, using the most favorable spin states for concerted and radical pathways as derived from the analyses above. As shown in Table S19, these two transition states have barriers similar to those calculated above for the reaction with the pyrrolidine substrate **10a** (within ∼1 kcal/mol differences). Therefore, these results also showed preference for the stepwise radical reaction pathway in the presence of **1a** as the substrate, featuring a lower barrier by 7.44 kcal/mol compared to the concerted pathway. Overall, these data further support a radical reaction mechanism for the C–H functionalization reaction investigated in this work.

Analysis of regio- and stereocontrol in the regiodivergent biocatalysts

Further computational studies were performed to gain insights into the origin of the divergent regioselectivity (i.e., α-C-H vs. β-C–H) of the engineered **CYP119-168** and **CYP119-235** biocatalysts. To this end, the most energetically favorable binding pose of the transition states (TS) for the regio- and stereoselectivity-determining hydrogen atom abstraction (HAA) step were modeled in the two enzymes. Briefly, transition states corresponding to near-attack conformations (NACs) for the insertion of the heme-bound carbene into the pro-(*S*) and pro-(*R*) C–H bond at the α- and β-C–H site of the *N*-methyl indoline substrate were calculated and optimized by DFT (see supplemental information for further details and results). These analyses showed that the α-regio-isomers TS are about 5–9 kcal mol^{-1} lower in energy than the β-regio-isomers TS, indicating a generally higher reactivity of the α-C–H site vs. β-C–H site and no stereoselectivity, in the absence of the protein environment (Table S23). Models of the **CYP119-168** and **CYP119-235** variants were generated using Rosetta[93] based on available crystal structures of this enzyme in both ligand-free form and in complex with imidazole-based ligands (Figure S9). As done previously with other systems,[94] the DFT-optimized TS regio-, stereo-, and conformational isomers were these superimposed onto the **CYP119-168** and **CYP119-235** structures and the resulting protein-TS complexes were optimized using Rosetta to obtain the corresponding energies (Tables S24 and S25; see supplemental information for Rosetta modeling details).

Inspection of most representative, low-energy models of these complexes revealed that mutations F153G and T213A, which are shared by both variants, both contribute to significantly enlarge the active site cavity to accommodate the bulky TS complex. The location and orientation of the *N*-methyl indoline substrate, however, were found to differ drastically in the two enzymes. In the α-C–H selective variant **CYP119-168**, the indoline substrate occupies a space near residue Leu205, with an orientation that favors carbene attack (from the *re*-face) to the α-pro(*S*)-C–H bond (Figure 4B), consistent with the regio and stereoselectivity of this biocatalyst. In **CYP119-235**, this arrangement is disfavored by mutation L205W, which inserts a bulky indole ring into the cavity occupied by the indoline molecule in **CYP119-168** (Figure S10C). Conversely, in the β-C–H selective variant **CYP119-235,** the *N*-methyl-indoline substrate occupies the cavity between Ala213 and Ala254 (Figure 4C), which is greatly enlarged by mutations T213A and V254A, respectively, compared to the wild-type enzyme. In this complex, the *N*-methyl-indoline ring is oriented so that carbene attack (from the *si*-face) is directed toward the

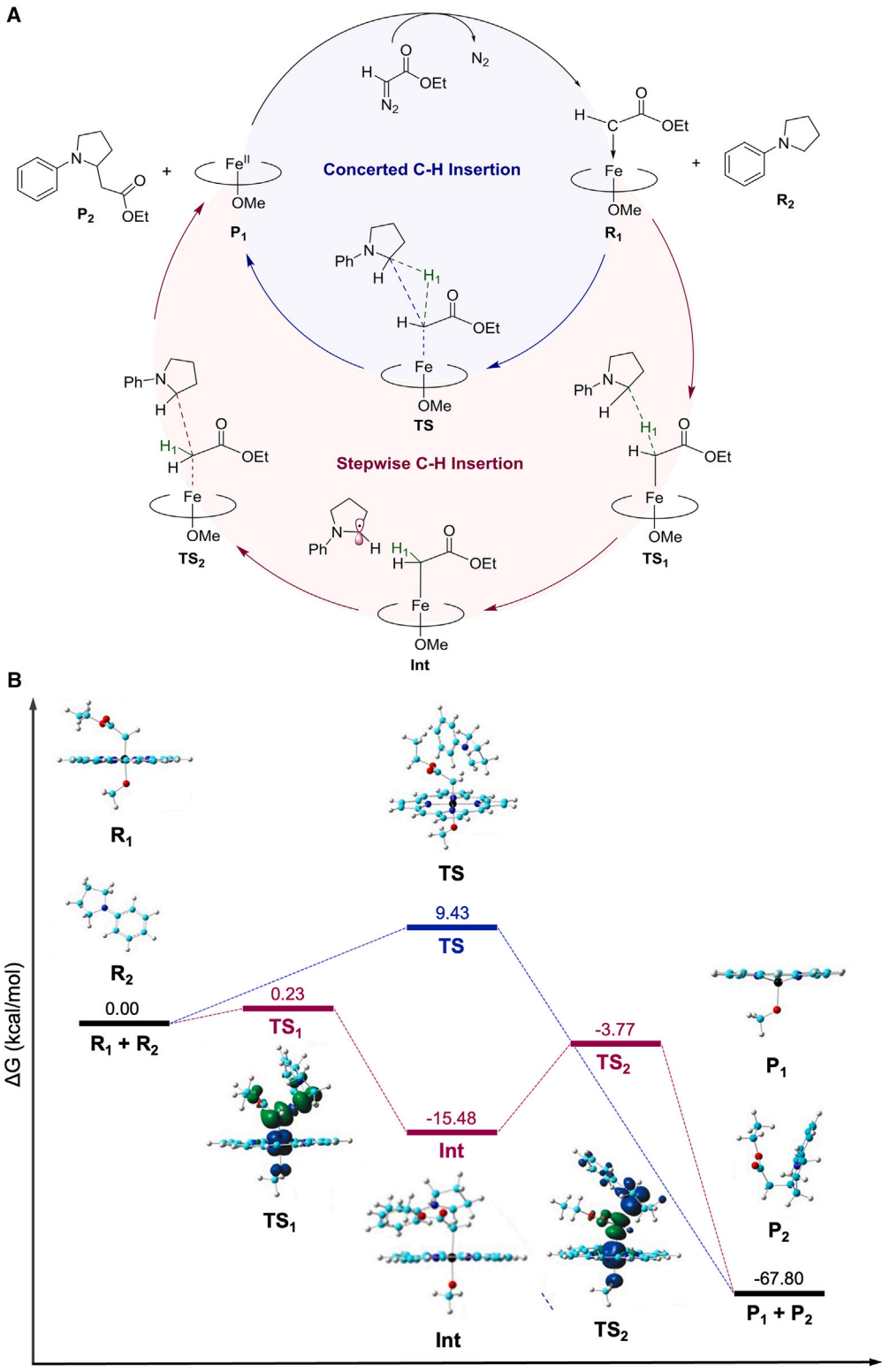


Figure 7. Computational analysis of concerted vs. radical-mediated pathways

(A) Concerted and stepwise pathways for heme catalyzed C−H insertion of *N*-phenylpyrrolidine. Oval represents porphyrin.

(B) Schematic free-energy diagram for heme-catalyzed C−H insertion. Spin densities of the transition states are shown. Atom color scheme: C, cyan; N, blue; O, red; H, grey; Fe, black (spin-density contour value: ±0.004 a.u.).

β-pro(*R*)-C–H bond of the substrate (Figure 4C), in line with the selectivity of this variant. In **CYP119-168**, this arrangement is disfavored by the presence of a bulky Trp residue (vs. Ala in **CYP119-235**) in position 254 (Figure S10B).

Overall, these analyses reveal how mutation of Thr213 and of the ''gating'' residue Phe153 (Figure S9) both contribute to make CYP119 active site more accessible to carbene/indoline complex, thus enhancing its non-native carbene transferase reactivity. On the other hand, modulation of steric encumbrance at the 254 and 205 positions, which are located at opposing sides of the heme pocket, plays a critical role toward dictating the α- vs. β-C–H selectivity in the C–H alkylation reaction catalyzed by the two regiodivergent biocatalysts. These findings also provide a rationale for explaining the experimentally observed beneficial effect of increasing steric bulk in position 254 toward favoring α-C–H regioselectivity and (*S*)-stereoselectivity (Trp > Tyr ≈ Phe ≫ Ile/Leu/Val; Table 1).

Conclusions

In summary, we have developed a new, efficient biocatalytic strategy for the enantioselective C(sp^3)–H functionalization of indoline scaffolds via P450-mediated carbene transfer. This method is amenable to the transformation of variously substituted indolines with diazoacetate and its scope extends to include other carbene donor reagents, such as diazoacetonitrile and diazoketone. Importantly, this work also demonstrates the possibility of achieving regiodivergent selectivity in enzyme-catalyzed C–H carbene insertion through the regio- and enantioselective functionalization of up to three distinct C(sp^3)–H sites in an indoline substrate (Figure 1C). While regiodivergent P450-catalyzed C(sp^3)–H oxidation has been previously achieved through the native monooxygenase reactivity of these enzymes,[32,33,35,37] regiodivergent C(sp^3)–H functionalization via carbene transfer has represented a major challenge[45–47,50–52] and so far has been largely elusive to engineered biological catalysts. On the one hand, this capability can provide rapid access to enantioenriched indoline-derived constitutional isomers, which can be valuable building blocks for medicinal chemistry and whose synthesis would require multi-step syntheses.[53] On the other hand, as demonstrated through the chemoenzymatic synthesis of **9a**, these regiodivergent carbene transferases can be leveraged to afford more complex, polycyclic indoline-based scaffolds akin to those found in various bioactive molecules.[3,54–56] Regiodivergent selectivity is also of particular interest in the construction of compound libraries for drug discovery campaigns.[95,96]

Our mechanistic studies provide valuable, first-time insights into the mechanism of hemoprotein-catalyzed carbene C(sp^3)–H insertion and collectively support the involvement of a radical, stepwise pathway akin to the mechanism of native P450-catalyzed hydroxylation reactions[97] and non-native C–H amination reactions via nitrene transfer catalyzed by engineered P450s and other hemoproteins.[98–100] Our molecular modeling studies clarify the role of key active site mutations in controlling the α- vs. β-C–H selectivity and the enantiopreference of the two regiodivergent biocatalysts developed for this transformation. Overall, this work expands the methodological toolbox for realizing selective C(sp^3)–H functionalization via enzyme-mediated carbene transfer. This advance, along with the insights gained into the productive and non-productive pathways in these reactions, and molecular basis of protein-controlled regioselectivity, are expected to guide future development of new and improved biocatalysts for this important class of transformations.

EXPERIMENTAL PROCEDURES

Details regarding the experimental procedures can be found in the supplemental experimental procedures.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Rudi Fasan (rudi.fasan@utdallas.edu).

Materials availability

All reagents in this study are commercially available or can be prepared as described in the supplemental information.

Data and code availability

There is no dataset or code associated with this publication. All supporting data and findings are available in the manuscript or supplemental information.

ACKNOWLEDGMENTS

This work was supported by the US National Institutes of Health grant GM098628 (R.F.). R.F. acknowledges chair endowment support from the Robert A. Welch Foundation (chair, AT-0051-20221212). S.D.K. acknowledges financial support from the NSF grant CHE-2054897. Y.Z. acknowledges financial support from the NSF grant CHE-2054897 and technical support from Xinyi Zhao and Emily McGuire for preliminary calculations. The authors are grateful to the UTD Center for High-Throughput Reaction Discovery & Synthesis supported by grant RR230018 from the Cancer Prevention and Research Institute of Texas (CPRIT).

AUTHOR CONTRIBUTIONS

R.F. and B.M.C. conceived the project. B.M.C. conducted the bulk of the experiments and experimental mechanistic studies, under guidance and supervision of R.F. R.C. contributed to reaction optimization and product isolation, under guidance and supervision of R.F. J.-M.C. performed DFT analyses of the mechanism under guidance and supervision of Y.Z. Z.S. performed computational analysis of the enzyme regio/stereocontrol under guidance and supervision of S.D.K. B.M.C. and R.F. wrote the manuscript with input and additions from all the other co-authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.checat.2024.101133.

Received: May 7, 2024
Revised: August 12, 2024
Accepted: September 9, 2024
Published: November 21, 2024

REFERENCES

1. Taylor, R.D., MacCoss, M., and Lawson, A.D.G. (2014). Rings in drugs. J. Med. Chem. *57*, 5845–5859. https://doi.org/10.1021/jm4017625.

2. Song, J., Chen, D.F., and Gong, L.Z. (2017). Recent progress in organocatalytic asymmetric total syntheses of complex indole alkaloids. Natl. Sci. Rev. *4*, 381–396. https://doi.org/10.1093/nsr/nwx028.

3. Silva, T.S., Rodrigues, M.T., Santos, H., Zeoly, L.A., Almeida, W.P., Barcelos, R.C., Gomes, R.C., Fernandes, F.S., and Coelho, F. (2019). Recent advances in indoline synthesis. Tetrahedron *75*, 2063–2097. https://doi.org/10.1016/j.tet.2019.02.006.

4. Wei, H., Li, B., Wang, N., Ma, Y., Yu, J., Wang, X., Su, J., and Liu, D. (2023). Development and Application of Indolines in Pharmaceuticals. ChemistryOpen *12*, e202200235. https://doi.org/10.1002/open.202200235.

5. Chandrasekhar, S., Basu, D., and Reddy, C.R. (2007). Palladium-catalyzed reduction of -(-butoxycarbonyl)indoles by polymethylhydrosiloxane. Synthesis-Stuttgart *2007*, 1509–1512. https://doi.org/10.1055/s-2007-966029.

6. Wang, L., Shao, Y., and Liu, Y. (2012). Nucleophilic addition of Grignard reagents to 3-acylindoles: stereoselective synthesis of highly substituted indoline scaffolds. Org. Lett. *14*, 3978–3981. https://doi.org/10.1021/ol301750b.

7. Nieto, M.J., and Lupton, H.K. (2021). Indole and Indoline Scaffolds in Antimicrobials: Overview, Synthesis and Recent Advances in Antimicrobial Research. Curr. Med. Chem. *28*, 4828–4844. https://doi.org/10.2174/0929867327666201102114923.

8. Boger, D.L., and Coleman, R.S. (1984). Intramolecular Diels-Alder Reactions of 1,2-Diazines - General Indoline Synthesis - Studies on the Preparation of the Central and Right-Hand Segments of Cc-1065. J. Org. Chem. *49*, 2240–2245. https://doi.org/10.1021/jo00186a032.

9. Presset, M., Pignon, A., Paul, J., Le Gall, E., Léonel, E., and Martens, T. (2017). Synthesis of Indolines by a Zn-Mediated Mannich Reaction/Pd-Catalyzed Amination Sequence. J. Org. Chem. *82*, 3302–3310. https://doi.org/10.1021/acs.joc.7b00013.

10. Li, Y., Chang, Y., Li, Y., Cao, C., Yang, J., Wang, B., and Liang, D. (2018). Iron-Catalyzed exo-Selective Synthesis of Cyanoalkyl Indolines via Cyanoisopropylarylation of Unactivated Alkenes. Adv. Synth. Catal. *360*, 2488–2492. https://doi.org/10.1002/adsc.201800296.

11. Liang, D., Ge, D., Lv, Y., Huang, W., Wang, B., and Li, W. (2018). Silver-Catalyzed Radical Arylphosphorylation of Unactivated Alkenes: Synthesis of 3-Phosphonoalkyl Indolines. J. Org. Chem. *83*, 4681–4691. https://doi.org/10.1021/acs.joc.8b00450.

12. Guillemard, L., Kaplaneris, N., Ackermann, L., and Johansson, M.J. (2021). Late-stage C-H functionalization offers new opportunities in drug discovery. Nat. Rev. Chem *5*, 522–545. https://doi.org/10.1038/s41570-021-00300-6.

13. Fraley, A.E., and Sherman, D.H. (2018). Halogenase engineering and its utility in medicinal chemistry. Bioorg. Med. Chem. Lett. *28*, 1992–1999. https://doi.org/10.1016/j.bmcl.2018.04.066.

14. Latham, J., Brandenburger, E., Shepherd, S.A., Menon, B.R.K., and Micklefield, J. (2018). Development of Halogenase Enzymes for Use in Synthesis. Chem. Rev. *118*, 232–269. https://doi.org/10.1021/acs.chemrev.7b00032.

15. Zhang, R.K., Huang, X., and Arnold, F.H. (2019). Selective CH bond functionalization with engineered heme proteins: new tools to generate complexity. Curr. Opin. Chem. Biol. *49*, 67–75. https://doi.org/10.1016/j.cbpa.2018.10.004.

16. Zwick, C.R., and Renata, H. (2020). Harnessing the biocatalytic potential of iron- and alpha-ketoglutarate-dependent dioxygenases in natural product total synthesis. Nat. Prod. Rep. *37*, 1065–1079. https://doi.org/10.1039/c9np00075e.

17. Ren, X., and Fasan, R. (2021). Engineered and Artificial Metalloenzymes for Selective C-H Functionalization. Curr. Opin. Green Sustain. Chem. *31*, 100494. https://doi.org/10.1016/j.cogsc.2021.100494.

18. Wang, W., Taber, D.F., and Renata, H. (2021). Practical Enzymatic Production of Carbocycles. Chemistry *27*, 11773–11794. https://doi.org/10.1002/chem.202101232.

19. Pyser, J.B., Chakrabarty, S., Romero, E.O., and Narayan, A.R.H. (2021). State-of-the-Art Biocatalysis. ACS Cent. Sci. *7*, 1105–1116. https://doi.org/10.1021/acscentsci.1c00273.

20. Couture, B., Chattopadhyay, A., and Fasan, R. (2024). Biocatalytic Carbene and Nitrene Transfer Reactions. In Comprehensive Chirality (Elsevier), pp. 111–142.

21. Reetz, M.T., Qu, G., and Sun, Z. (2024). Engineered enzymes for the synthesis of pharmaceuticals and other high-value products. Nat. Synth. *3*, 19–32. https://doi.org/10.1038/s44160-023-00417-0.

22. Hayashi, T., Ligibel, M., Sager, E., Voss, M., Hunziker, J., Schroer, K., Snajdrova, R., and Buller, R. (2019). Evolved Aliphatic Halogenases Enable Regiocomplementary C-H Functionalization of a Pharmaceutically Relevant Compound. Angew. Chem. Int. Ed. Engl. *58*, 18535–18539. https://doi.org/10.1002/anie.201907245.

23. Renata, H., Shimizu, E., and Zwick, C.R. (2021). Regiodivergent Biocatalytic Hydroxylation of L-Glutamine Facilitated by Characterization of Non-Heme Dioxygenases from Non-Ribosomal Peptide Biosyntheses. Tetrahedron *90*, 132190. https://doi.org/10.1016/j.tet.2021.132190.

24. Craven, E.J., Latham, J., Shepherd, S.A., Khan, I., Diaz-Rodriguez, A., Greaney, M.F., and Micklefield, J. (2021). Programmable late-stage C-H bond functionalization enabled by integration of enzymes with chemocatalysis. Nat. Catal. *4*, 385–394. https://doi.org/10.1038/s41929-021-00603-3.

25. Espinoza, R.V., Haatveit, K.C., Grossman, S.W., Tan, J.Y., McGlade, C.A., Khatri, Y., Newmister, S.A., Schmidt, J.J., Garcia-Borràs, M., Montgomery, J., et al. (2021). Engineering P450 TamI as an Iterative Biocatalyst for Selective Late-Stage C-H Functionalization and Epoxidation of Tirandamycin Antibiotics. ACS Catal. *11*, 8304–8316. https://doi.org/10.1021/acscatal.1c01460.

26. Münch, J., Püllmann, P., Zhang, W., and Weissenborn, M.J. (2021). Enzymatic Hydroxylations of sp3-Carbons. ACS Catal. *11*, 9168–9203. https://doi.org/10.1021/acscatal.1c00759.

27. Knorrscheidt, A., Soler, J., Hünecke, N., Püllmann, P., Garcia-Borràs, M., and Weissenborn, M.J. (2021). Accessing Chemo- and Regioselective Benzylic and Aromatic Oxidations by Protein Engineering of an Unspecific Peroxygenase. ACS Catal. *11*, 7327–7338. https://doi.org/10.1021/acscatal.1c00847.

28. Zetzsche, L.E., Chakrabarty, S., and Narayan, A.R.H. (2022). Development of a P450 Fusion Enzyme for Biaryl Coupling in Yeast. ACS Chem. Biol. *17*, 2986–2992. https://doi.org/10.1021/acschembio.2c00690.

29. Monterrey, D.T., Menés-Rubio, A., Keser, M., Gonzalez-Perez, D., and Alcalde, M. (2023). Unspecific peroxygenases: The pot of gold at the end of the oxyfunctionalization rainbow? Curr. Opin. Green Sustainable Chem. *41*, 100786. https://doi.org/10.1016/j.cogsc.2023.100786.

30. Andorfer, M.C., Park, H.J., Vergara-Coll, J., and Lewis, J.C. (2016). Directed Evolution of RebH for Catalyst-Controlled Halogenation of Indole C-H Bonds. Chem. Sci. *7*, 3720–3729. https://doi.org/10.1039/C5SC04680G.

31. Shepherd, S.A., Karthikeyan, C., Latham, J., Struck, A.W., Thompson, M.L., Menon, B.R.K., Styles, M.Q., Levy, C., Leys, D., and Micklefield, J. (2015). Extending the biocatalytic scope of regiocomplementary flavin-dependent halogenase enzymes. Chem. Sci. *6*, 3454–3460. https://doi.org/10.1039/c5sc00913h.

32. Kille, S., Zilly, F.E., Acevedo, J.P., and Reetz, M.T. (2011). Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. Nat. Chem. *3*, 738–743. https://doi.org/10.1038/nchem.1113.

33. Zhang, K., Shafer, B.M., Demars, M.D., second, Stern, H.A., and Fasan, R. (2012). Controlled oxidation of remote sp3 C-H bonds in artemisinin

via P450 catalysts with fine-tuned regio- and stereoselectivity. J. Am. Chem. Soc. *134*, 18695–18704. https://doi.org/10.1021/ja3073462.

34. Greule, A., Stok, J.E., De Voss, J.J., and Cryle, M.J. (2018). Unrivalled diversity: the many roles and reactions of bacterial cytochromes P450 in secondary metabolism. Nat. Prod. Rep. *35*, 757–791. https://doi.org/10.1039/c7np00063d.
35. Alwaseem, H., Frisch, B.J., and Fasan, R. (2018). Anticancer activity profiling of parthenolide analogs generated via P450-mediated chemoenzymatic synthesis. Bioorg. Med. Chem. *26*, 1365–1373. https://doi.org/10.1016/j.bmc.2017.08.009.
36. Lukowski, A.L., Ellinwood, D.C., Hinze, M.E., DeLuca, R.J., Du Bois, J., Hall, S., and Narayan, A.R.H. (2018). C-H Hydroxylation in Paralytic Shellfish Toxin Biosynthesis. J. Am. Chem. Soc. *140*, 11863–11869. https://doi.org/10.1021/jacs.8b08901.
37. Zhang, X., King-Smith, E., Dong, L.B., Yang, L.C., Rudolf, J.D., Shen, B., and Renata, H. (2020). Divergent synthesis of complex diterpenes through a hybrid oxidative approach. Science *369*, 799–806. https://doi.org/10.1126/science.abb8271.
38. Alwaseem, H., Giovani, S., Crotti, M., Welle, K., Jordan, C.T., Ghaemmaghami, S., and Fasan, R. (2021). Comprehensive Structure-Activity Profiling of Micheliolide and its Targeted Proteome in Leukemia Cells via Probe-Guided Late-Stage C-H Functionalization. ACS Cent. Sci. *7*, 841–857. https://doi.org/10.1021/acscentsci.0c01624.
39. Iizaka, Y., Arai, R., Takahashi, A., Ito, M., Sakai, M., Fukumoto, A., Sherman, D.H., and Anzai, Y. (2022). Engineering sequence and selectivity of late-stage C-H oxidation in the MycG iterative cytochrome P450. J. Ind. Microbiol. Biotechnol. *49*, kuab069. https://doi.org/10.1093/jimb/kuab069.
40. Turner, N.J. (2009). Directed evolution drives the next generation of biocatalysts. Nat. Chem. Biol. *5*, 567–573. https://doi.org/10.1038/nchembio.203.
41. Reetz, M.T. (2011). Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions. Angew. Chem. Int. Ed. Engl. *50*, 138–174. https://doi.org/10.1002/anie.201000826.
42. Dydio, P., Key, H.M., Nazarenko, A., Rha, J.Y.E., Seyedkazemi, V., Clark, D.S., and Hartwig, J.F. (2016). An artificial metalloenzyme with the kinetics of native enzymes. Science *354*, 102–106. https://doi.org/10.1126/science.aah4427.
43. Brandenberg, O.F., Fasan, R., and Arnold, F.H. (2017). Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. Curr. Opin. Biotechnol. *47*, 102–111. https://doi.org/10.1016/j.copbio.2017.06.005.
44. Vargas, D.A., Tinoco, A., Tyagi, V., and Fasan, R. (2018). Myoglobin-Catalyzed C-H Functionalization of Unprotected Indoles. Angew. Chem. Int. Ed. Engl. *57*, 9911–9915. https://doi.org/10.1002/anie.201804779.
45. Gu, Y., Natoli, S.N., Liu, Z., Clark, D.S., and Hartwig, J.F. (2019). Site-Selective Functionalization of (sp(3))C-H Bonds Catalyzed by Artificial Metalloenzymes Containing an Iridium-Porphyrin Cofactor. Angew. Chem. Int. Ed. Engl. *58*, 13954–13960. https://doi.org/10.1002/anie.201907460.
46. Zhang, J., Huang, X., Zhang, R.K., and Arnold, F.H. (2019). Enantiodivergent alpha-Amino C-H Fluoroalkylation Catalyzed by Engineered Cytochrome P450s. J. Am. Chem. Soc. *141*, 9798–9802. https://doi.org/10.1021/jacs.9b04344.
47. Zhang, R.K., Chen, K., Huang, X., Wohlschlager, L., Renata, H., and Arnold, F.H. (2019). Enzymatic assembly of carbon-carbon bonds via iron-catalysed sp(3) C-H functionalization. Nature *565*, 67–72. https://doi.org/10.1038/s41586-018-0808-5.
48. Rumo, C., Stein, A., Klehr, J., Tachibana, R., Prescimone, A., Häussinger, D., and Ward, T.R. (2022). An Artificial Metalloenzyme Based on a Copper Heteroscorpionate Enables sp(3) C-H Functionalization via Intramolecular Carbene Insertion. J. Am. Chem. Soc. *144*, 11676–11684. https://doi.org/10.1021/jacs.2c03311.
49. Zhang, J., Maggiolo, A.O., Alfonzo, E., Mao, R., Porter, N.J., Abney, N., and Arnold, F.H. (2023). Chemodivergent C(sp(3))-H and C(sp(2))-H Cyanomethylation Using Engineered Carbene Transferases. Nat. Catal. *6*, 152–160. https://doi.org/10.1038/s41929-022-00908-x.
50. Ren, X., Couture, B.M., Liu, N., Lall, M.S., Kohrt, J.T., and Fasan, R. (2023). Enantioselective Single and Dual alpha-C-H Bond Functionalization of Cyclic Amines via Enzymatic Carbene Transfer. J. Am. Chem. Soc. *145*, 537–550. https://doi.org/10.1021/jacs.2c10775.
51. Sreenilayam, G., Moore, E.J., Steck, V., and Fasan, R. (2017). Metal Substitution Modulates the Reactivity and Extends the Reaction Scope of Myoglobin Carbene Transfer Catalysts. Adv. Synth. Catal. *359*, 2076–2089. https://doi.org/10.1002/adsc.201700202.
52. Zhou, A.Z., Chen, K., and Arnold, F.H. (2020). Enzymatic Lactone-Carbene C-H Insertion to Build Contiguous Chiral Centers. ACS Catal. *10*, 5393–5398. https://doi.org/10.1021/acscatal.0c01349.
53. Chhabra, N., Aseri, M.L., and Padmanabhan, D. (2013). A review of drug isomerism and its significance. Int. J. Appl. Basic Med. Res. *3*, 16–18. https://doi.org/10.4103/2229-516X.112233.
54. Ruiz-Sanchis, P., Savina, S.A., Albericio, F., and Álvarez, M. (2011). Structure, bioactivity and synthesis of natural products with hexahydropyrrolo[2,3-b]indole. Chemistry *17*, 1388–1408. https://doi.org/10.1002/chem.201001451.
55. Zi, W., Zuo, Z., and Ma, D. (2015). Intramolecular dearomative oxidative coupling of indoles: a unified strategy for the total synthesis of indoline alkaloids. Acc. Chem. Res. *48*, 702–711. https://doi.org/10.1021/ar5004303.
56. Wang, X., Griffiths, B., and Burl, J. (2016). Bioinspired Discovery of Chemical Reactions and Biological Probes. Synlett *27*, 2039–2042. https://doi.org/10.1055/s-0035-1561638.
57. Liu, F., and Su, M. (2023). Indole and indoline scaffolds in drug discovery. In Privileged Scaffolds in Drug Discovery, B. Yu, N. Li, and C. Fu, eds. (Academic Press), pp. 147–161.
58. Coelho, P.S., Wang, Z.J., Ener, M.E., Baril, S.A., Kannan, A., Arnold, F.H., and Brustad, E.M. (2013). A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. Nat. Chem. Biol. *9*, 485–487. https://doi.org/10.1038/nchembio.1278.
59. Doyle, M.P., Duffy, R., Ratnikov, M., and Zhou, L. (2010). Catalytic carbene insertion into C-H bonds. Chem. Rev. *110*, 704–724. https://doi.org/10.1021/cr900239n.
60. Weaver, B.R., Perkins, L.J., Fernandez Candelaria, F.O., Burstyn, J.N., and Buller, A.R. (2023). Molecular Determinants of Efficient Cobalt-Substituted Hemoprotein Production in E. coli. ACS Synth. Biol. *12*, 3669–3679. https://doi.org/10.1021/acssynbio.3c00481.
61. Key, H.M., Dydio, P., Liu, Z., Rha, J.Y.E., Nazarenko, A., Seyedkazemi, V., Clark, D.S., and Hartwig, J.F. (2017). Beyond Iron: Iridium-Containing P450 Enzymes for Selective Cyclopropanations of Structurally Diverse Alkenes. ACS Cent. Sci. *3*, 302–308. https://doi.org/10.1021/acscentsci.6b00391.
62. Knight, A.M., Kan, S.B.J., Lewis, R.D., Brandenberg, O.F., Chen, K., and Arnold, F.H. (2018). Diverse Engineered Heme Proteins Enable Stereodivergent Cyclopropanation of Unactivated Alkenes. ACS Cent. Sci. *4*, 372–377. https://doi.org/10.1021/acscentsci.7b00548.
63. Park, S.Y., Yamane, K., Adachi, S.i., Shiro, Y., Weiss, K.E., Maves, S.A., and Sligar, S.G. (2002). Thermophilic cytochrome P450 (CYP119) from Sulfolobus solfataricus: high resolution structure and functional properties. J. Inorg. Biochem. *91*, 491–501. https://doi.org/10.1016/s0162-0134(02)00446-4.
64. Garlets, Z.J., Wertz, B.D., Liu, W., Voight, E.A., and Davies, H.M.L. (2020). Regio- and Stereoselective Rhodium(II)-Catalyzed C-H Functionalization of Cyclobutanes. Chem *6*, 304–313. https://doi.org/10.1016/j.chempr.2019.12.014.
65. Boni, Y.T., Cammarota, R.C., Liao, K., Sigman, M.S., and Davies, H.M.L. (2022). Leveraging Regio- and Stereoselective C(sp(3))-H Functionalization

of Silyl Ethers to Train a Logistic Regression Classification Model for Predicting Site-Selectivity Bias. J. Am. Chem. Soc. *144*, 15549–15561. https://doi.org/10.1021/jacs.2c04383.

66. Hock, K.J., Knorrscheidt, A., Hommelsheim, R., Ho, J., Weissenborn, M.J., and Koenigs, R.M. (2019). Tryptamine Synthesis by Iron Porphyrin Catalyzed C-H Functionalization of Indoles with Diazoacetonitrile. Angew. Chem. Int. Ed. Engl. *58*, 3630–3634. https://doi.org/10.1002/anie.201813631.
67. Brandenberg, O.F., Chen, K., and Arnold, F.H. (2019). Directed Evolution of a Cytochrome P450 Carbene Transferase for Selective Functionalization of Cyclic Compounds. J. Am. Chem. Soc. *141*, 8989–8995. https://doi.org/10.1021/jacs.9b02931.
68. Wang, Z.J., Peck, N.E., Renata, H., and Arnold, F.H. (2014). Cytochrome P450-Catalyzed Insertion of Carbenoids into N-H Bonds. Chem. Sci. *5*, 598–601. https://doi.org/10.1039/C3SC52535J.
69. Sreenilayam, G., and Fasan, R. (2015). Myoglobin-catalyzed intermolecular carbene N-H insertion with arylamine substrates. Chem. Commun. *51*, 1532–1534. https://doi.org/10.1039/c4cc08753d.
70. Steck, V., Sreenilayam, G., and Fasan, R. (2020). Selective Functionalization of Aliphatic Amines via Myoglobin-catalyzed Carbene N-H Insertion. Synlett *31*, 224–229. https://doi.org/10.1055/s-0039-1690007.
71. Vargas, D.A., Khade, R.L., Zhang, Y., and Fasan, R. (2019). Biocatalytic Strategy for Highly Diastereo- and Enantioselective Synthesis of 2,3-Dihydrobenzofuran-Based Tricyclic Scaffolds. Angew. Chem. Int. Ed. Engl. *58*, 10148–10152. https://doi.org/10.1002/anie.201903455.
72. Lee, W.C.C., Wang, D.S., Zhu, Y., and Zhang, X.P. (2023). Iron(III)-based metalloradical catalysis for asymmetric cyclopropanation via a stepwise radical mechanism. Nat. Chem. *15*, 1569–1580. https://doi.org/10.1038/s41557-023-01317-8.
73. Moore, E.J., and Fasan, R. (2019). Effect of proximal ligand substitutions on the carbene and nitrene transferase activity of myoglobin. Tetrahedron *75*, 2357–2363. https://doi.org/10.1016/j.tet.2019.03.009.
74. Rettie, A.E., Boberg, M., Rettenmeier, A.W., and Baillie, T.A. (1988). Cytochrome P-450-catalyzed desaturation of valproic acid in vitro. Species differences, induction effects, and mechanistic studies. J. Biol. Chem. *263*, 13733–13738. https://doi.org/10.1016/S0021-9258(18)68302-4.
75. Guengerich, F.P. (2018). Mechanisms of Cytochrome P450-Catalyzed Oxidations. ACS Catal. *8*, 10964–10976. https://doi.org/10.1021/acscatal.8b03401.
76. Ren, X., O'Hanlon, J.A., Morris, M., Robertson, J., and Wong, L.L. (2016). Synthesis of Imidazolidin-4-ones via a Cytochrome P450-Catalyzed Intramolecular C-H Amination. ACS Catal. *6*, 6833–6837. https://doi.org/10.1021/acscatal.6b02189.
77. Kumar, D., De Visser, S.P., and Shaik, S. (2004). Oxygen economy of cytochrome P450: what is the origin of the mixed functionality as a dehydrogenase-oxidase enzyme compared with its normal function? J. Am. Chem. Soc. *126*, 5072–5073. https://doi.org/10.1021/ja0318737.
78. Cooper, H.L.R., Mishra, G., Huang, X., Pender-Cudlip, M., Austin, R.N., Shanklin, J., and Groves, J.T. (2012). Parallel and competitive pathways for substrate desaturation, hydroxylation, and radical rearrangement by the non-heme diiron hydroxylase AlkB. J. Am. Chem. Soc. *134*, 20365–20375. https://doi.org/10.1021/ja3059149.
79. Chen, K., Zhang, S.Q., Brandenberg, O.F., Hong, X., and Arnold, F.H. (2018). Alternate Heme Ligation Steers Activity and Selectivity in Engineered Cytochrome P450-Catalyzed Carbene-Transfer Reactions. J. Am. Chem. Soc. *140*, 16402–16407. https://doi.org/10.1021/jacs.8b09613.
80. Tinoco, A., Wei, Y., Bacik, J.P., Carminati, D.M., Moore, E.J., Ando, N., Zhang, Y., and Fasan, R. (2019). Origin of high stereocontrol in olefin cyclopropanation catalyzed by an engineered carbene transferase. ACS Catal. *9*, 1514–1524. https://doi.org/10.1021/acscatal.8b04073.
81. Carminati, D.M., and Fasan, R. (2019). Stereoselective Cyclopropanation of Electron-Deficient Olefins with a Cofactor Redesigned Carbene Transferase Featuring Radical Reactivity. ACS Catal. *9*, 9683–9697. https://doi.org/10.1021/acscatal.9b02272.
82. Khade, R.L., Fan, W., Ling, Y., Yang, L., Oldfield, E., and Zhang, Y. (2014). Iron Porphyrin Carbenes as Catalytic Intermediates: Structures, Mossbauer and NMR Spectroscopic Properties, and Bonding. Angew. Chem. Int. Ed. *53*, 7574–7578. https://doi.org/10.1002/anie.201402472.
83. Khade, R.L., and Zhang, Y. (2015). Catalytic and Biocatalytic Iron Porphyrin Carbene Formation: Effects of Binding Mode, Carbene Substituent, Porphyrin Substituent, and Protein Axial Ligand. J. Am. Chem. Soc. *137*, 7560–7563. https://doi.org/10.1021/jacs.5b03437.
84. Khade, R.L., and Zhang, Y. (2017). C-H Insertions by Iron Porphyrin Carbene: Basic Mechanism and Origin of Substrate Selectivity. Chemistry - A Euopean Journal *23*, 17654–17658. https://doi.org/10.1002/chem.201704631.
85. Wei, Y., Tinoco, A., Steck, V., Fasan, R., and Zhang, Y. (2018). Cyclopropanations via Heme Carbenes: Basic Mechanism and Effects of Carbene Substituent, Protein Axial Ligand, and Porphyrin Substitution. J. Am. Chem. Soc. *140*, 1649–1662. https://doi.org/10.1021/jacs.7b09171.
86. Khade, R.,L., Chandgude, A.L., Fasan, R., and Zhang, Y. (2019). Mechanistic Investigation of Biocatalytic Heme Carbenoid Si-H Insertions. ChemCatChem *11*, 3101–3108. https://doi.org/10.1002/cctc.201801755.
87. Vargas, D.A., Khade, R.L., Zhang, Y., and Fasan, R. (2019). Biocatalytic strategy for highly diastereo- and enantioselective synthesis of 2,3-dihydrobenzofuran based tricyclic scaffolds. Angew. Chem. Int. Ed. *58*, 10148–10152. https://doi.org/10.1002/anie.201903455.
88. Li, Z., Burnell, D.J., and Boyd, R.J. (2017). Computational Study of Engineered Cytochrome P450-Catalyzed C–H Amination: The Origin of the Regio- and Stereoselectivity. J. Phys. Chem. B *121*, 10859–10868. https://doi.org/10.1021/acs.jpcb.7b10256.
89. Wang, J., Gao, H., Yang, L., and Gao, Y.Q. (2020). Role of Engineered Iron-haem Enzyme in Reactivity and Stereoselectivity of Intermolecular Benzylic C–H Bond Amination. ACS Catal. *10*, 5318–5327. https://doi.org/10.1021/acscatal.0c00248.
90. Li, X., Dong, L., and Liu, Y. (2020). Theoretical Study of Iron Porphyrin Nitrene: Formation Mechanism, Electronic Nature, and Intermolecular C–H Amination. Inorg. Chem. *59*, 1622–1632. https://doi.org/10.1021/acs.inorgchem.9b02216.
91. Athavale, S.V., Gao, S., Liu, Z., Mallojjala, S.C., Hirschi, J.S., and Arnold, F.H. (2021). Biocatalytic, Intermolecular C−H Bond Functionalization for the Synthesis of Enantioenriched Amides. Angew. Chem. Int. Ed. Engl. *60*, 24864–24869. https://doi.org/10.1002/anie.202110873.
92. Wei, Y., Conklin, M., and Zhang, Y. (2022). Biocatalytic Intramolecular C−H aminations via Engineered Heme Proteins: Full Reaction Pathways and Axial Ligand Effects. Chem. Eur J. *28*, e202202006. https://doi.org/10.1002/chem.202202006.
93. Leaver-Fay, A., Tyka, M., Lewis, S.M., Lange, O.F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P.D., Smith, C.A., Sheffler, W., et al. (2011). Chapter nineteen - Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. In Methods Enzymol., Methods in Enzymology, *487*, M.L. Johnson and L. Brand, eds. (Academic Press), pp. 545–574.
94. Nam, D., Tinoco, A., Shen, Z., Adukure, R.D., Sreenilayam, G., Khare, S.D., and Fasan, R. (2022). Enantioselective Synthesis of alpha-Trifluoromethyl Amines via Biocatalytic N-H Bond Insertion with Acceptor-Acceptor Carbene Donors. J. Am. Chem. Soc. *144*, 2590–2602. https://doi.org/10.1021/jacs.1c10750.
95. Gerry, C.J., Wawer, M.J., Clemons, P.A., and Schreiber, S.L. (2019). DNA Barcoding a Complete Matrix of Stereoisomeric Small Molecules. J. Am. Chem. Soc. *141*, 10225–10235. https://doi.org/10.1021/jacs.9b01203.

96. Bassi, G., Favalli, N., Vuk, M., Catalano, M., Martinelli, A., Trenner, A., Porro, A., Yang, S., Tham, C.L., Moroglu, M., et al. (2020). A Single-Stranded DNA-Encoded Chemical Library Based on a Stereoisomeric Scaffold Enables Ligand Discovery by Modular Assembly of Building Blocks. Adv. Sci. *7*, 2001970. https://doi.org/10.1002/advs.202001970.

97. Denisov, I.G., Makris, T.M., Sligar, S.G., and Schlichting, I. (2005). Structure and chemistry of cytochrome P450. Chem. Rev. *105*, 2253–2277. https://doi.org/10.1021/cr0307143.

98. Singh, R., Kolev, J.N., Sutera, P.A., and Fasan, R. (2015). Enzymatic C(sp(3))-H Amination: P450-Catalyzed Conversion of Carbonazidates into Oxazolidinones. ACS Catal. *5*, 1685–1691. https://doi.org/10.1021/Cs5018612.

99. Yang, Y., Cho, I., Qi, X., Liu, P., and Arnold, F.H. (2019). An Enzymatic Platform for the Asymmetric Amination of Primary, Secondary and Tertiary C(sp(3))—H Bonds. Nat. Chem. *11*, 987–993. https://doi.org/10.1038/s41557-019-0343-5.

100. Roy, S., Vargas, D.A., Ma, P., Sengupta, A., Zhu, L., Houk, K.N., and Fasan, R. (2024). Stereoselective Construction of beta-gamma-and delta-Lactam Rings via Enzymatic C-H Amidation. Nature Catal *7*, 65–76. https://doi.org/10.1038/s41929-023-01068-2.

Chem Catalysis, Volume *4*

Supplemental information

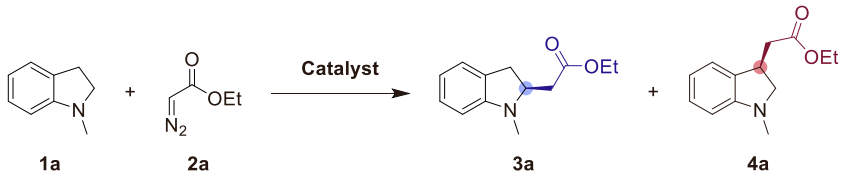
Radical-mediated regiodivergent C(sp^3)–H functionalization of *N*-substituted indolines via enzymatic carbene transfer

Bo M. Couture, Ru Cui, Jia-Min Chu, Zhuofan Shen, Sagar D. Khare, Yong Zhang, and Rudi Fasan

Supplemental Experimental Procedures

Supplementary Tables S1-S4

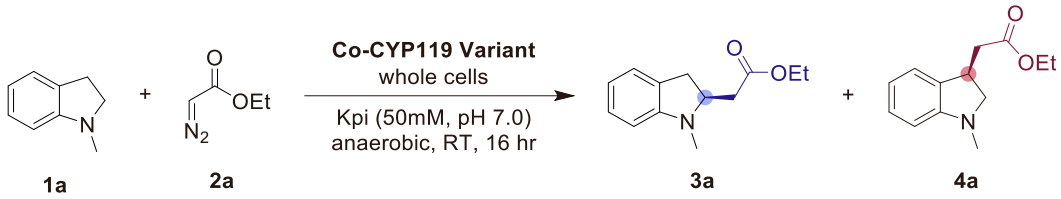
Table S1. Intermolecular C–H functionalization of *N*-methyl Indoline (**1a**) with EDA (**2a**) using representative transition metal catalysts.[a]



No.	Catalyst	Yield[b]	TON
1	Hemin	0	-
2	Fe(TPP)Cl	0	-
3	$Cu(OTf)_2$	0	-
4	Co(TPP)Cl	0	-
5	$Rh_2(OAc)_4$	0	-
6	$Rh_2(TFA)_4$	0	-
7	$[Ru(p\text{-cymene})Cl_2]_2$	0	-
8	$Ru(BPY)_3$	0	-
9	$Cp^*RuCl(PPh_3)_2$	0	-

[a] Standard reaction conditions: 10 mM **1a**, 80 mM EDA (**2a**), 10 mol% chemical catalysts, dry toluene, 70°C, 16 hours. [b] Assay yields as determined by GC using calibration curves with isolated product **3a**/**4a**.

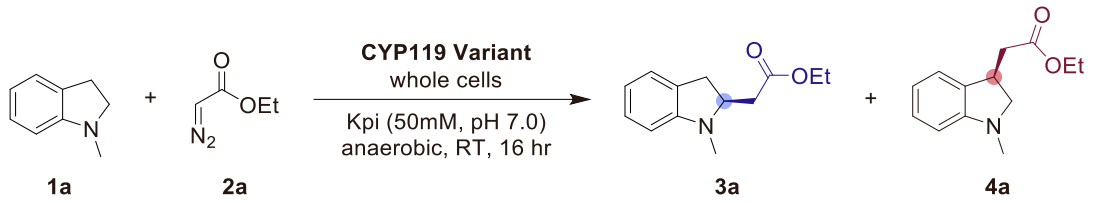
Table S2. Intermolecular C–H functionalization of *N*-methyl Indoline (**1a**) with EDA (**2a**) using Co-substituted CYP119 variants.[a]



No.	Catalyst	Yield (3a)[b]	Yield (4a)[b]	3a:4a
1	Mb[Co]	-	-	-
2	CYP119[Co]	-	-	-
3	[**CYP119-137**] CYP119 [Co] (F153G, A209G, T213G, V254A, C317S)	12	17	41:59
4	[**CYP119-168**] CYP119 [Co] (F153G, T213W, V254W, C317S)	72	15	83:17
5	[**CYP119-235**] CYP119 [Co] (F153G, L205W, T213A, V254A, C317S)	25	46	35:65

[a] Standard reaction conditions: 20 μM Co-CYP119 purified protein, 10 mM **1a**, 80 mM EDA (**2a**), in KPi buffer (50 mM, pH 7), room temperature, 16 hours, in anaerobic chamber. [b] Assay yields as determined by GC using calibration curves with isolated product **3a/4a**.

Table S3. Intermolecular C–H functionalization of *N*-methyl Indoline (**1a**) with EDA (**2a**) using CYP119 variants for the formation of benzylic insertion product **4a**.[a]

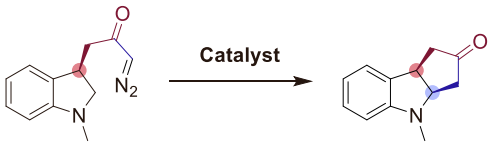


1a **2a** **3a** **4a**

No.	Catalyst	Yield (4a)[b]	3a:4a	TON (4a)[c]	e.r. (4a)[d]
1	CYP119 (WT)	0	0	-	-
2	[**CYP119-137**] CYP119 (F153A, A209G, T213G, V254A, C317S)	17%	66:34	n.d.	n.d.
3	[**CYP119-235**] CYP119 (F153G, L205W, T213A, V254A, C317S)	78%	16:84	5,380	94:6
4[e]	[**CYP119-235**] CYP119 (F153G, L205W, T213A, V254A, C317S)	84%	16:84	500	94:6

[a] Standard reaction conditions: protein expressing C41(DE3) *E. coli* cells (OD_{600} = 60), 10 mM **1a**, 80 mM EDA (**2a**), in KPi buffer (50 mM, pH 7), room temperature, 16 hours, in anaerobic chamber. [b] Assay yields as determined by GC using calibration curves with isolated product **4a**. [c] TON of major product as calculated based on the protein concentration measured from cell lysate. [d] Enantiomeric ratio (e.r.) of major product as determined by chiral HPLC. [e] Using 20 μM purified protein and 10 mM $Na_2S_2O_4$. N.d. = not determined.

Table S4. Intramolecular C–H functionalization of diazo tethered *N*-methyl Indoline (**8**) using representative transition metal catalysts.[a]

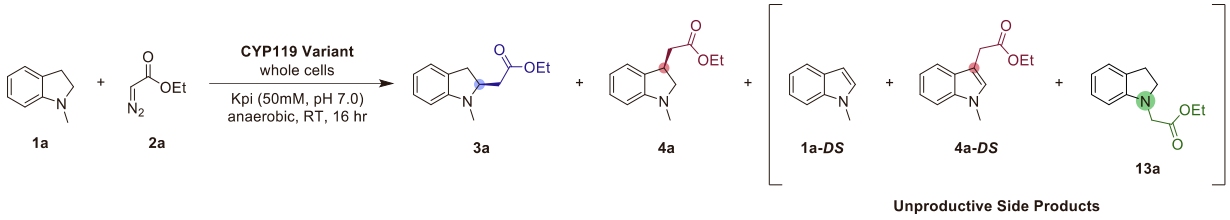


8 **9a**

No.	Catalyst	Yield[b]	TON
1	Hemin	0	-
2	Fe(TPP)Cl	0	-
3	Co(TPP)Cl	0	-
4	$Rh(OAc)_2$	0	-

[a] Standard reaction conditions: 10 mM **8**, 10 mol% chemical catalysts, dry toluene, 70°C, 16 hours. [b] Assay yields as determined by GC using calibration curves with isolated product **9a**.

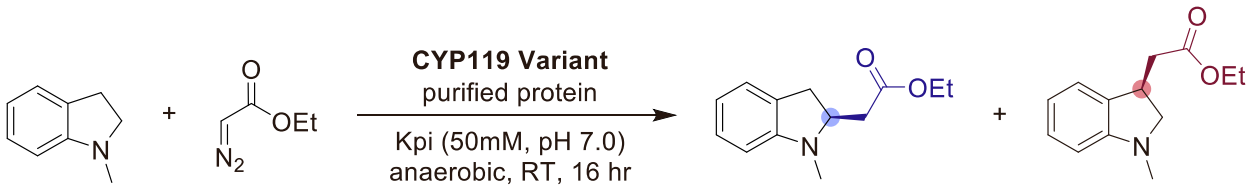
Table S5. Intermolecular C–H functionalization of *N*-methyl-indoline (1a) with EDA (2a) using CYP119 variants for formation of products 3a, 4a, and side products 13a, 14a, and 15a.[a]



No.	Catalyst	Conv.[b]	3a:4a:Other
1	CYP119 (WT)	0	-
2	[**CYP119-137**] CYP119 (F153A, A209G, T213G, V254A, C317S)	26%	42:23:35
3	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	99%	92:8:0
4	[**CYP119-235**] CYP119 (F153G, L205W, T213A, V254A, C317S)	99%	16:84:0

[a] Standard reaction conditions: protein expressing C41(DE3) *E. coli* cells (OD_{600} = 60), 10 mM **1a**, 80 mM EDA (**2a**), in KPi buffer (50 mM, pH 7), room temperature, 16 hours, in anaerobic chamber. [b] Conversion based on the combined areas of products **3a**, **4a**, **1a-*DS***, **4a-*DS***, and **13a**. [c] Product distribution determined by GC using calibration curves with corresponding isolated products.

Table S6. Intermolecular C–H functionalization of *N*-methyl Indoline (1a) with EDA (2a) in the presence or absence of reductant and CO.[a]

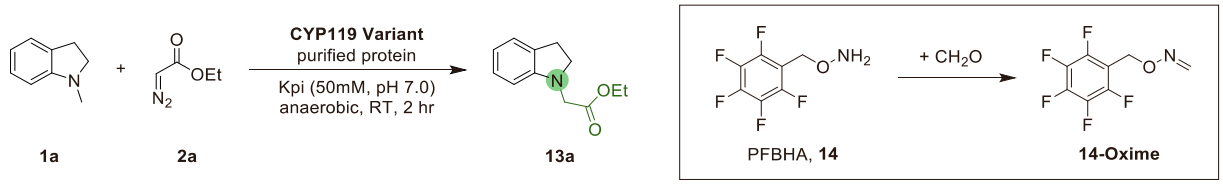


1a **2a** **3a** **4a**

No.	Catalyst	Conditions	Yield (3a)[b]	Yield (4a)[b]	3a:4a
1	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	Std.	45	4	92:8
2	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	No red.	32	3	92:8
3	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	CO	-	-	-
4	[**CYP119-168**] CYP119 (F153G, T213A, V254W, C317S)	No red. + CO	-	-	-

[a] Standard reaction conditions: 5 μM purified protein, 10 mM **1a**, 80 mM EDA (**2a**), 10 mM $Na_2S_2O_4$, in KPi buffer (50 mM, pH 7), room temperature, 16 hours, in anaerobic chamber. [b] Assay yields as determined by GC using calibration curves with isolated product **3a/4a**.

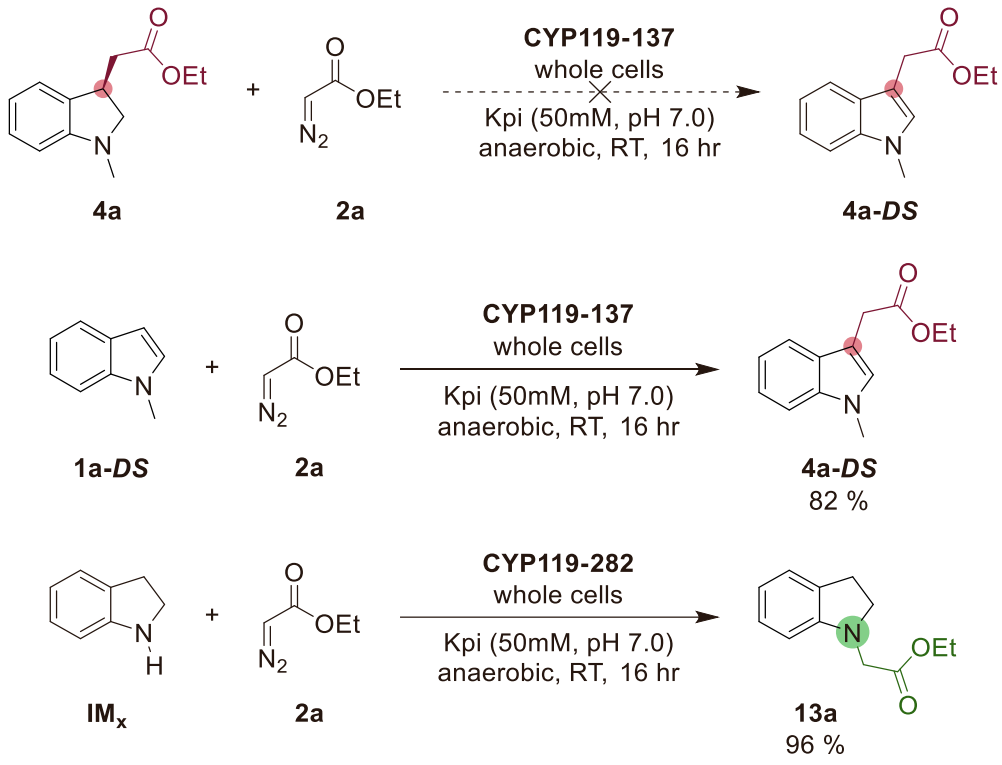
Table S7. Detection of Formaldehyde Formation in the Catalytic Demethylation/ N–H Insertion of *N*-Methyl Indoline with EDA Using PFBHA Formaldehyde Trapping Reagent[a]



No.	Catalyst	Conc. 14-Oxime (mM)[b]	Conc. 13a (mM)[b]	13a:14-Oxime
1	**[CYP119-168]** CYP119 (F153G, T213A, V254W, C317S)	0	0	-
2	**[CYP119-311]** CYP119 (A209W, G210S, T213G, V254A, C317S)	0.025	0.150	6.10
3	**[CYP119-137]** CYP119 (F153G, A209G, T213G, V254A, C317S)	0.080	0.455	5.67
4	**[CYP119-282]** CYP119 (L205Y, A209G, T213G, V254A, C317S)	0.154	0.808	5.26
5	**[CYP119-044]** CYP119 (F153G, T213A, V254A, C317S)	0.168	1.044	6.20

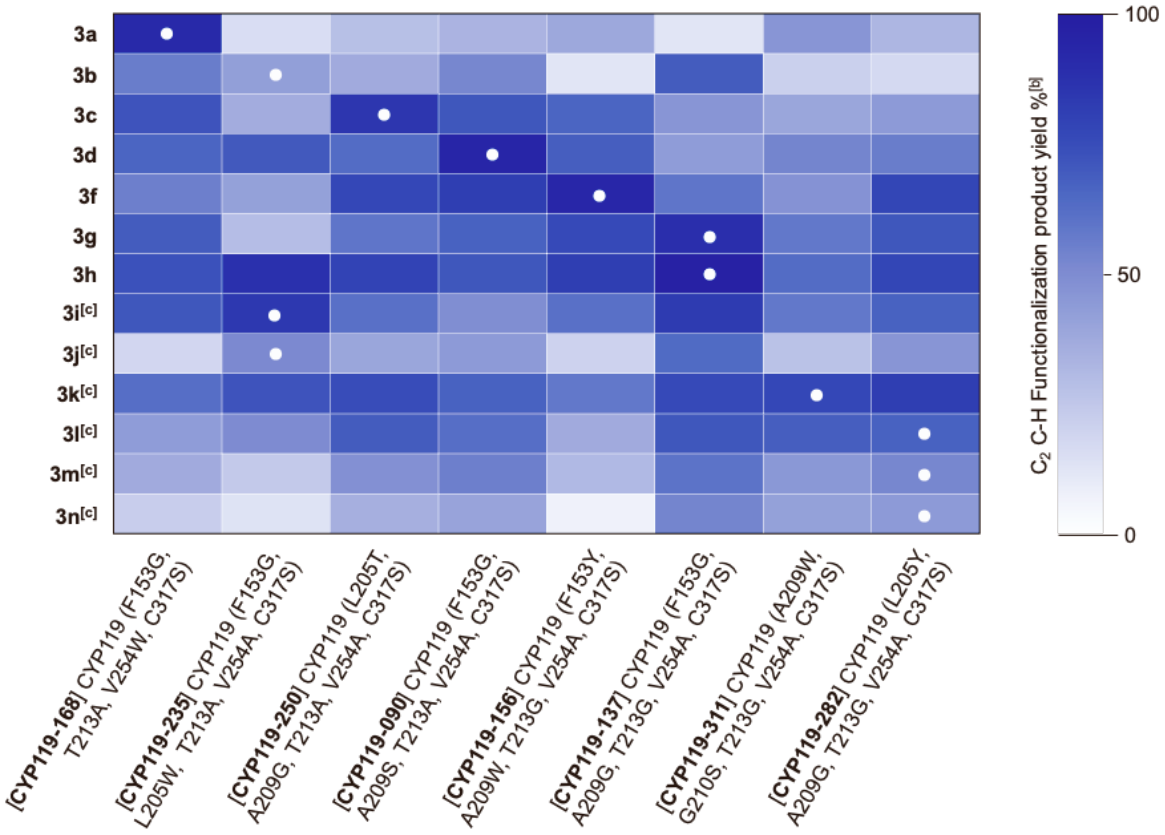
[a] Standard reaction conditions: 20 μM purified protein, 10 mM 1a, 80 mM EDA (**2a**), 10 mM $Na_2S_2O_4$, in KPi buffer (50 mM, pH 7), room temperature, 2 hours, in anaerobic chamber. *O*-(2,3,4,5,6-Pentafluorobenzyl)hydroxylamine Hydrochloride (PFBHA, **14**) was added at the end of the reaction and allowed to incubate for 10 minutes before extraction. [b] Assay yields as determined by GC using calibration curves with isolated product **13a/14-Oxime**.

Table S8. Reactivity of Enzymatic Products/ Intermediates with Representative CYP119 Variants[a]



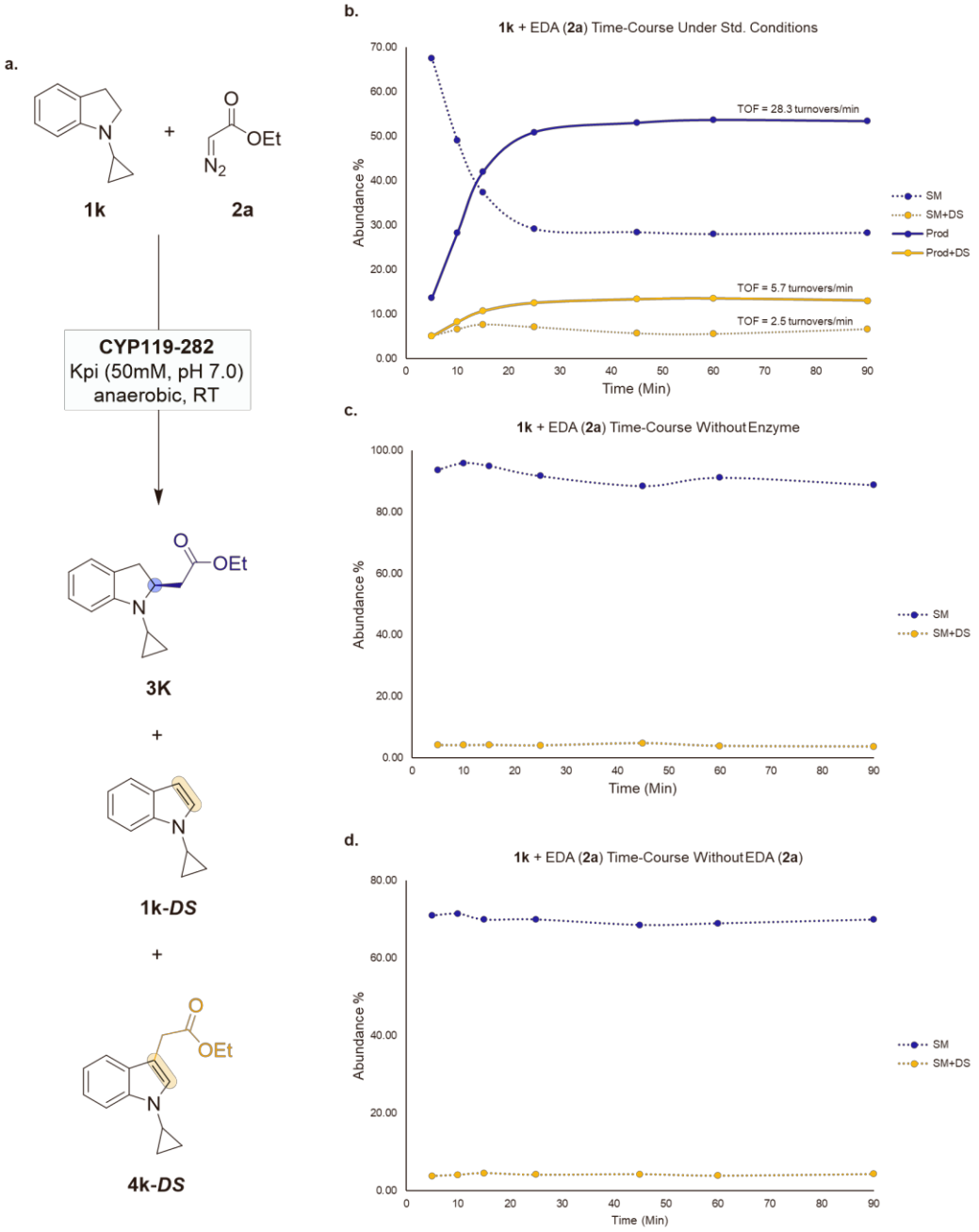
[a] Standard reaction conditions: protein expressing C41(DE3) *E. coli* cells (OD_{600} = 60), 10 mM substrate, 80 mM EDA (**2a**), in KPi buffer (50 mM, pH 7), room temperature, 16 hours, in anaerobic chamber. Conversion based on calibration curves with corresponding isolated products.

Figure S1. Intermolecular C–H functionalization of *N*-substituted Indolines (**1a-n**) with EDA (**2a**) using hemoprotein CYP119 variants.[a]



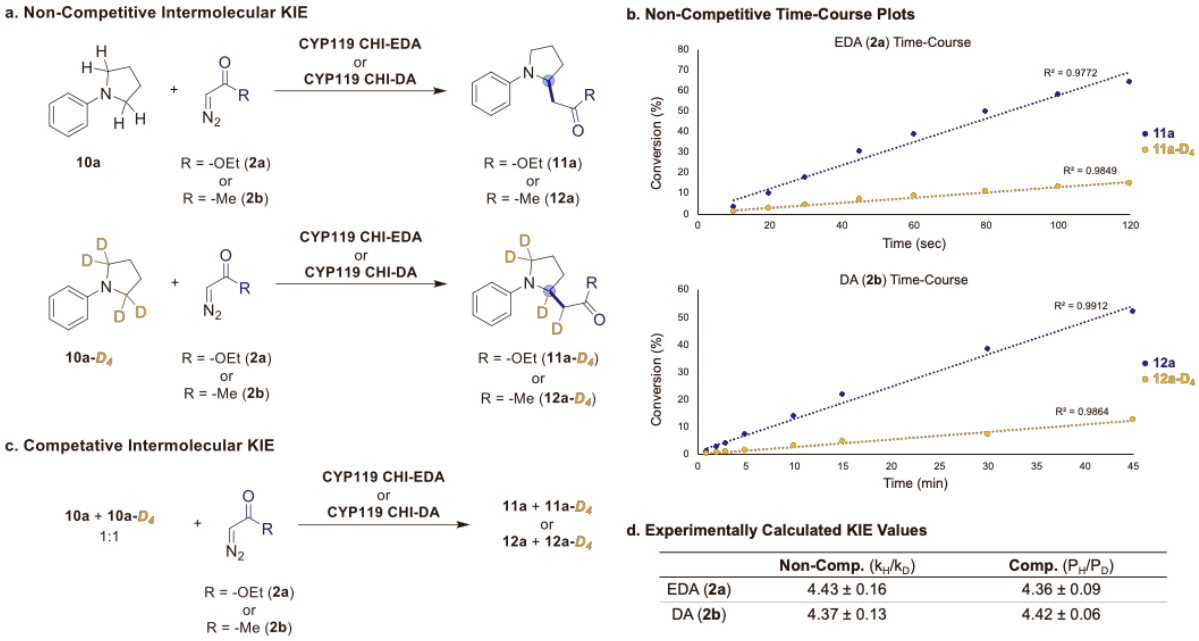
White dots (•) note best variant for the indicated substrate. [a] Standard reaction conditions: protein expressing C41(DE3) *E. coli* cells (OD_{600} = 60), 10 mM **1a**, 80 mM EDA (**2a**), in KPi buffer (50 mM, pH 7), room temperature, 16 hours, in anaerobic chamber. [b] Assay yields as determined by GC using calibration curves with corresponding isolated products. [c] Using 240 mM EDA at 40 °C.

Figure S2. Time-course analysis of CYP119 [**CYP119-282**]-catalyzed intermolecular C–H insertion reactions of N-cyclopropyl indoline (**1k**) with EDA (**2a**).



Conversion was determined by gas chromatography using calibration curves with **3k**, **1k-*DS***, and **4k-*DS***. Reaction conditions: 10 μM CYP119 **[CYP119-282]** purified protein, 10 mM **1k**, 50 mM EDA (**2a**), 50 mM KPi buffer (pH 7.0). (**a**) General scheme showing formation of enzymatic products **3k**, **1k-*DS***, and **4k-*DS***. (**b**) Time-course under standard conditions. (**c**) Time-course in the absence of enzyme. (**d**) Time-course in the absence of EDA (**2a**).

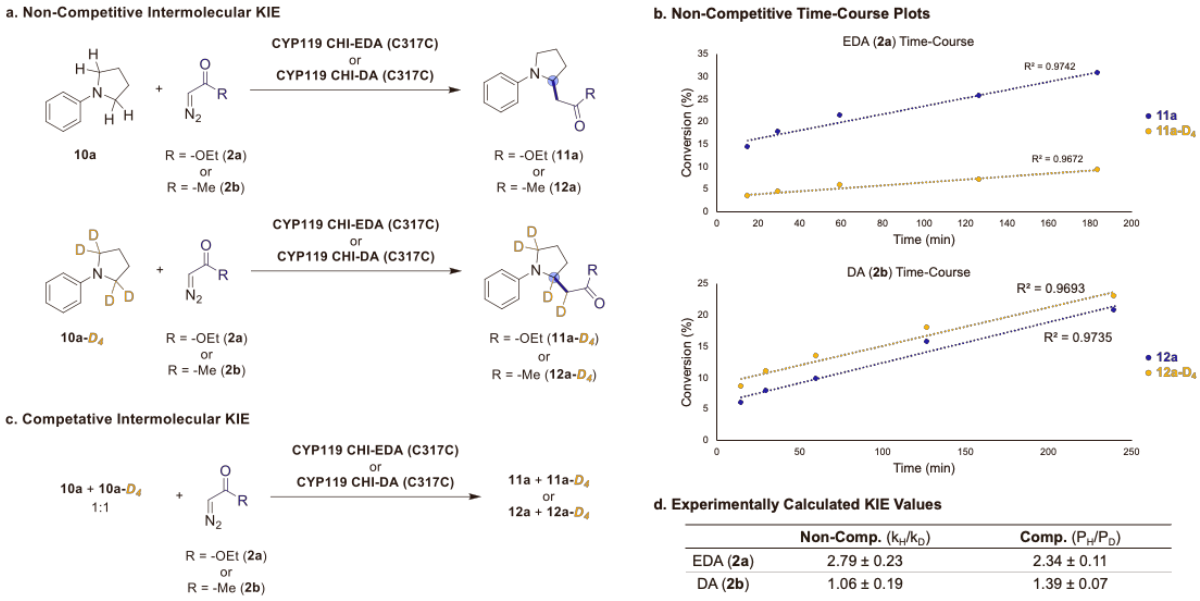
Figure S3. Intermolecular Competitive and Non-Competitive KIE Experiments with *N*-phenylpyrrolidine (**10a**) and EDA (**2a**) or DA (**2b**) Using Evolved CYP119 Variants.



	Non-Comp. (k_H/k_D)	Comp. (P_H/P_D)
EDA (**2a**)	4.43 ± 0.16	4.36 ± 0.09
DA (**2b**)	4.37 ± 0.13	4.42 ± 0.06

(**a**) Non-competitive intermolecular KIE using *N*-phenylpyrrolidine (**10a**) or *N*-phenylpyrrolidine-2,2,5,5-d_4 (**10a-*D*$_4$**) with EDA (**2a**) or DA (**2b**). Reaction conditions: 10 mM substrate, 15 mM diazo compound, 20 μM CYP119 Variant, 10 mM sodium Dithionate, 50 mM phosphate buffer (pH 7.0), anerobic chamber, room temperature. (**b**) Non-competitive time-course plots for both EDA (**2a**) and DA (**2b**). (**c**) Competitive intermolecular KIE using *N*-phenylpyrrolidine (**10a**) and *N*-phenylpyrrolidine-2,2,5,5-d_4 (**10a-*D*$_4$**) (1:1) with EDA (**2a**) or DA (**2b**). Reaction conditions: 10 mM **10a:10a-*D*$_4$** (1:1), 15 mM diazo compound, 20 μM CYP119 Variant, 10 mM sodium Dithionate, 50 mM phosphate buffer (pH 7.0), anerobic chamber, room temperature. (**d**) Summary table of experimentally calculated KIE values.

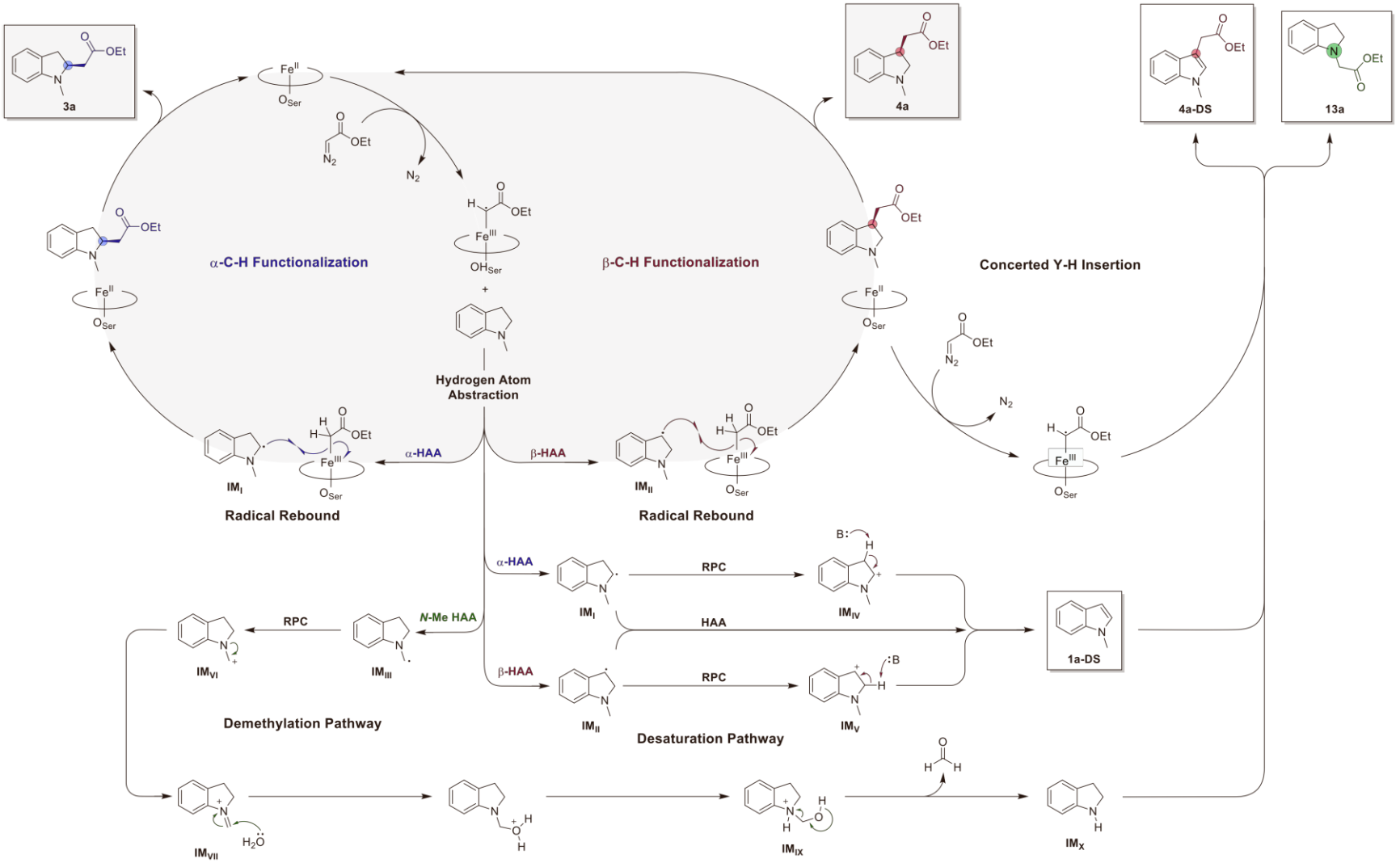
Figure S4. Intermolecular Competitive and Non-Competitive KIE Experiments with *N*-phenylpyrrolidine (**10a**) and EDA (**2a**) or DA (**2b**) Using Evolved Cystine Ligated CYP119 Variants.



	Non-Comp. (k_H/k_D)	Comp. (P_H/P_D)
EDA (**2a**)	2.79 ± 0.23	2.34 ± 0.11
DA (**2b**)	1.06 ± 0.19	1.39 ± 0.07

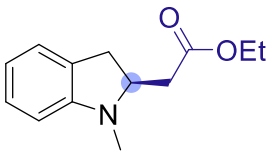
(**a**) Non-competitive intermolecular KIE using *N*-phenylpyrrolidine (**10a**) or *N*-phenylpyrrolidine-2,2,5,5-d_4 (**10a-*D*$_4$**) with EDA (**2a**) or DA (**2b**). Reaction conditions: 10 mM substrate, 15 mM diazo compound, 20 μM CYP119 Variant, 10 mM sodium Dithionate, 50 mM phosphate buffer (pH 7.0), anerobic chamber, room temperature. (**b**) Non-competitive time-course plots for both EDA (**2a**) and DA (**2b**). (**c**) Competitive intermolecular KIE using *N*-phenylpyrrolidine (**10a**) and *N*-phenylpyrrolidine-2,2,5,5-d_4 (**10a-*D*$_4$**) (1:1) with EDA (**2a**) or DA (**2b**). Reaction conditions: 10 mM **10a:10a-*D*$_4$** (1:1), 15 mM diazo compound, 20 μM CYP119 Variant, 10 mM sodium Dithionate, 50 mM phosphate buffer (pH 7.0), anerobic chamber, room temperature. (**d**) Summary table of experimentally calculated KIE values.

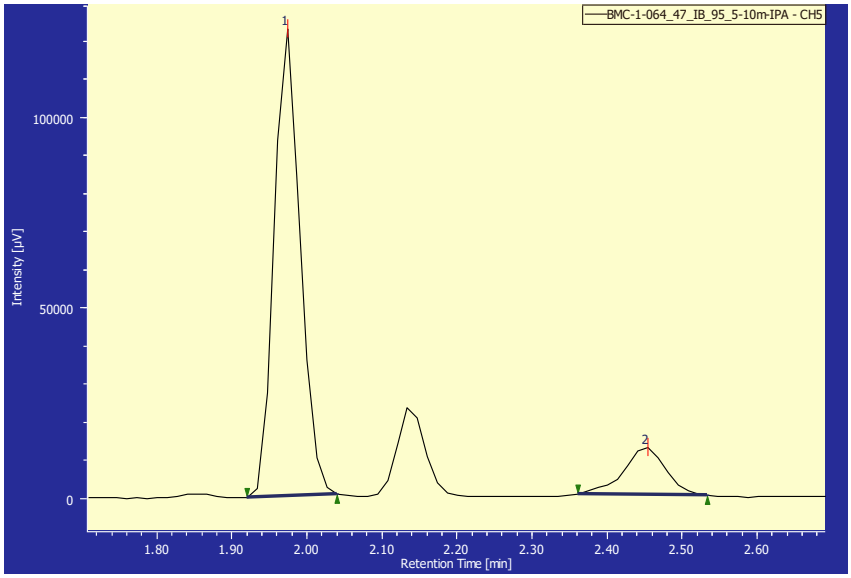
Figure S5. Proposed stepwise radical mediated mechanism for the C–H functionalization of *N*-methyl indoline (**1a**) in the presence of engineered CYP119 catalysts.



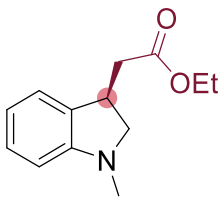
Data S1. Chiral SFC analysis for the determination of an enantiomeric excess in the CYP119-catalyzed intermolecular C–H insertion reactions with EDA (**2a**).

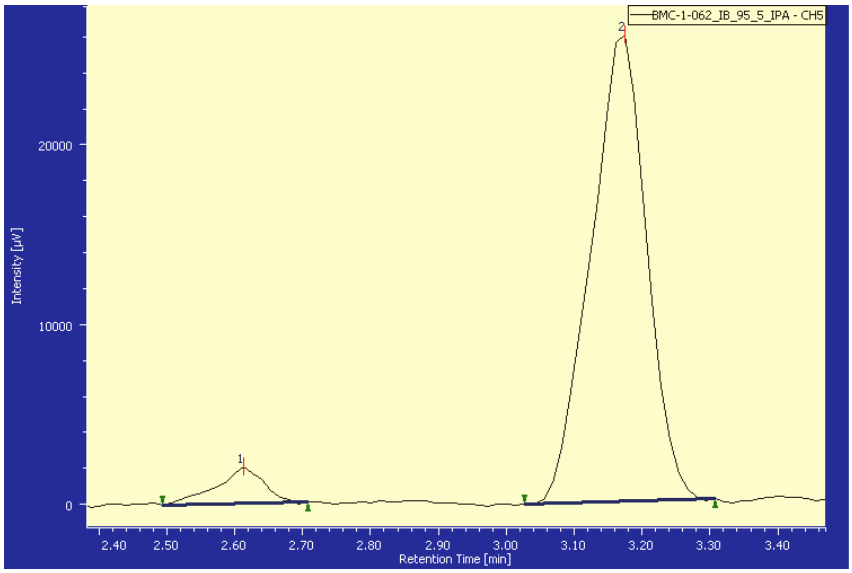
- Chiral SFC analysis of enantioenriched **3a** product using CYP119 (F153G, T213A, V254W, C317S).





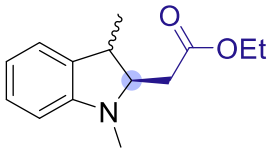
- Chiral SFC analysis of enantioenriched **4a** product using CYP119 (F153G, L205W, T213A, V254A, C317S).

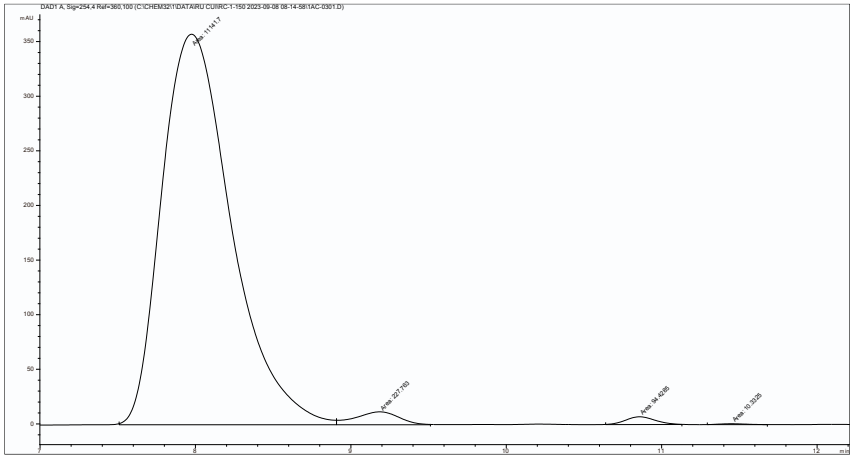




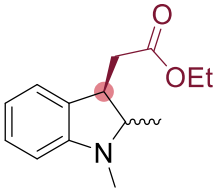
Data S2. Chiral HPLC analysis for the determination of an enantiomeric excess in the CYP119-catalyzed intermolecular C–H insertion reactions with EDA (**2a**).

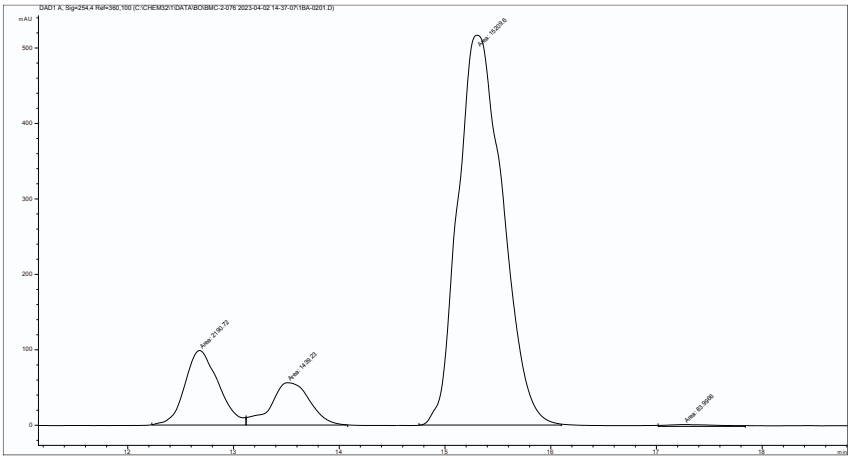
- Chiral HPLC analysis of enantioenriched **3b** product using CYP119 (F153G, L205W, T213A, V254A, C317S).



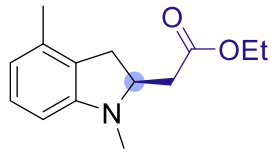


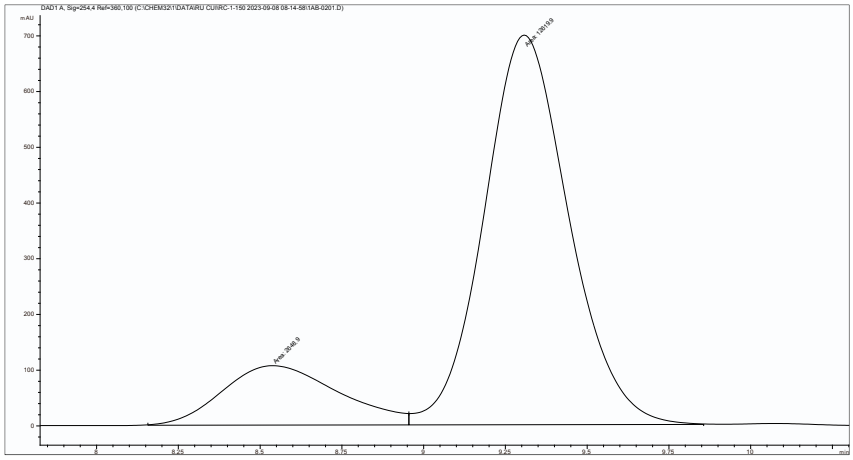
- Chiral HPLC analysis of enantioenriched **4b** product using CYP119 (F153G, A209G, T213G, V254A, C317S).



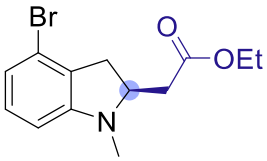


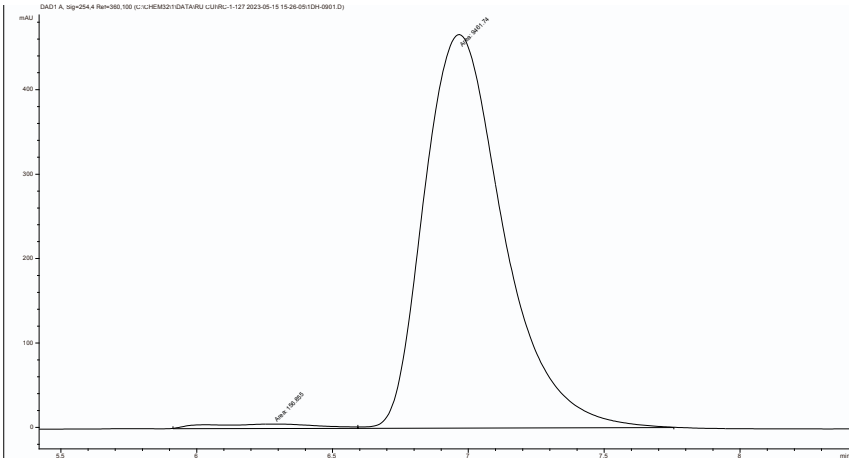
- Chiral HPLC analysis of enantioenriched **3c** product using CYP119 (F153G, A209G, T213G, V254A, C317S).



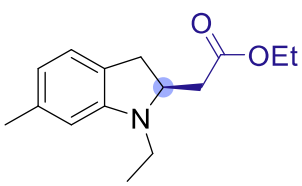


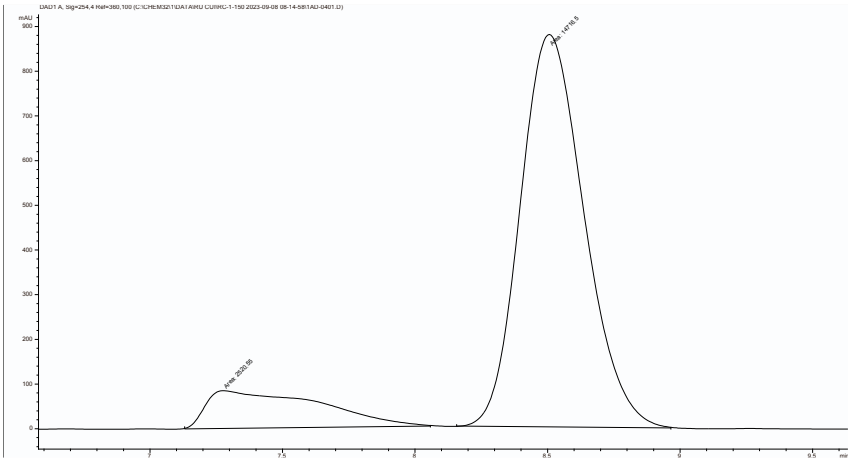
- Chiral HPLC analysis of enantioenriched **3d** product using CYP119 (F153G, A209S, T213A, V254A, C317S).



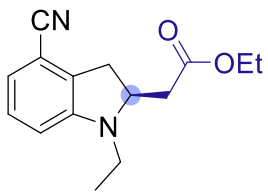


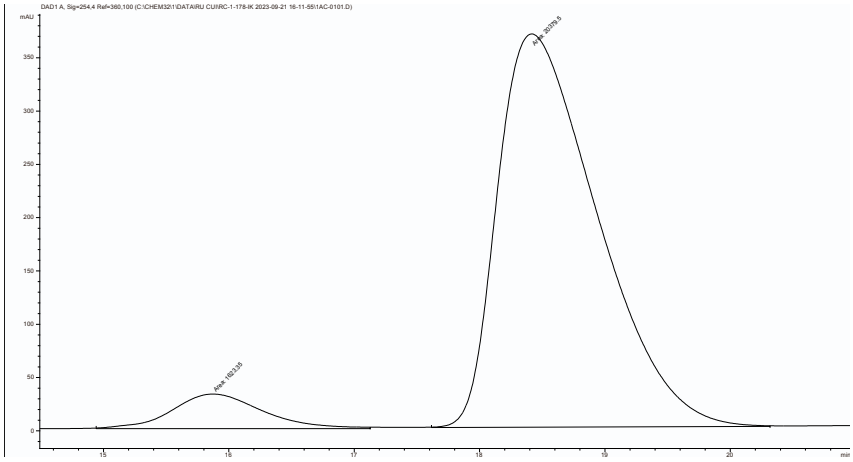
- Chiral HPLC analysis of enantioenriched **3f** product using CYP119 (F153Y, A209W, T213G, V254A, C317S).



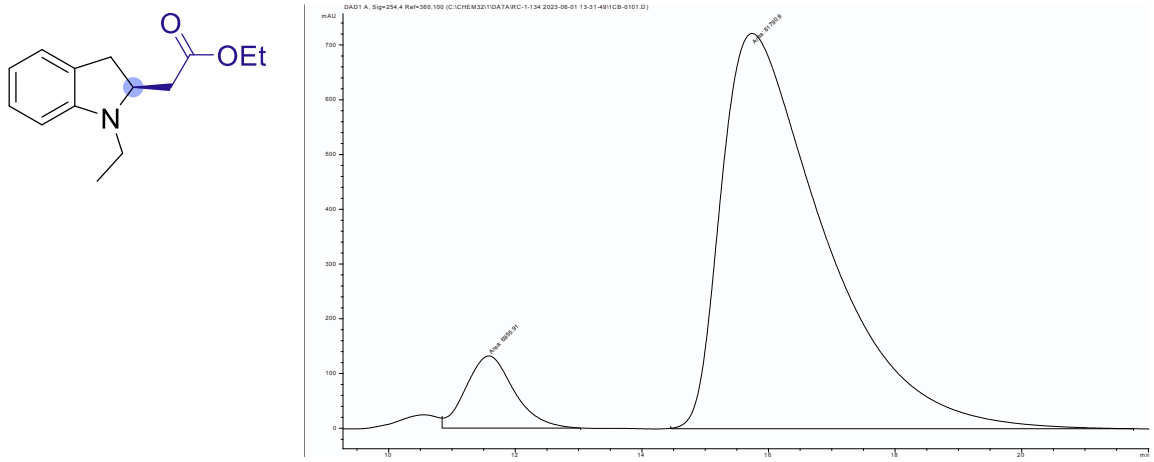


- Chiral HPLC analysis of enantioenriched **3g** product using CYP119 (F153G, A209G, T213G, V254A, C317S).

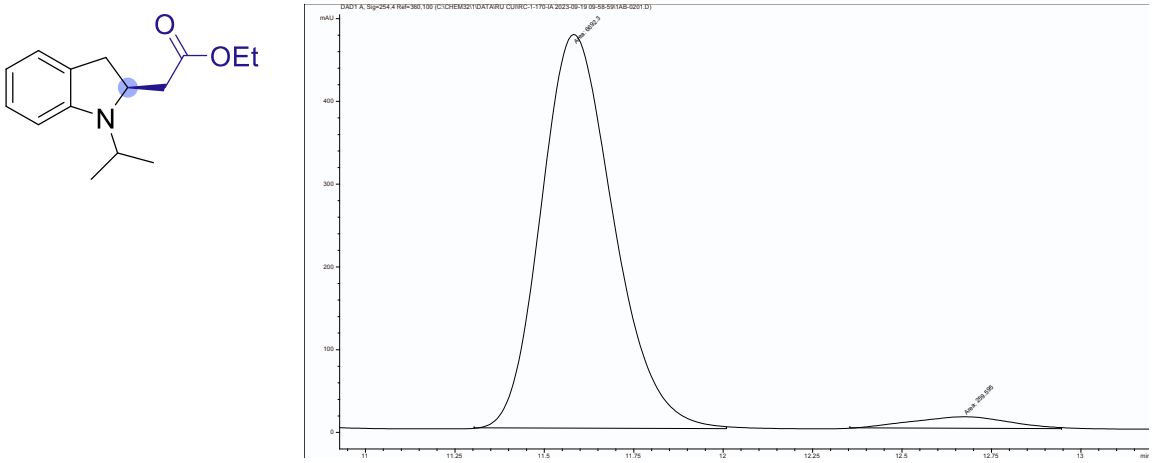




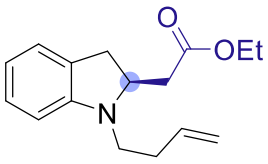
- Chiral HPLC analysis of enantioenriched **3h** product using CYP119 (F153G, A209G, T213G, V254A, C317S).

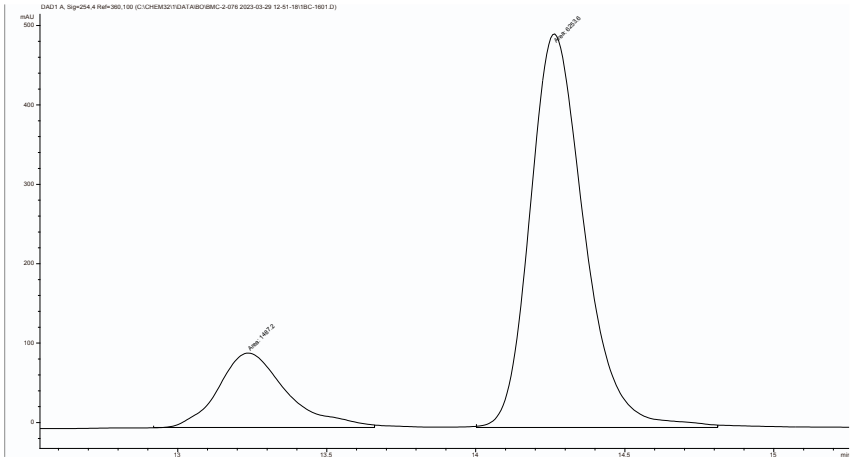


- Chiral HPLC analysis of enantioenriched **3i** product using CYP119 (F153G, L205W, T213A, V254A, C317S).

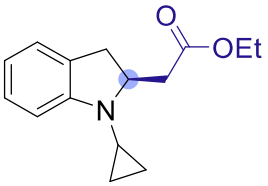


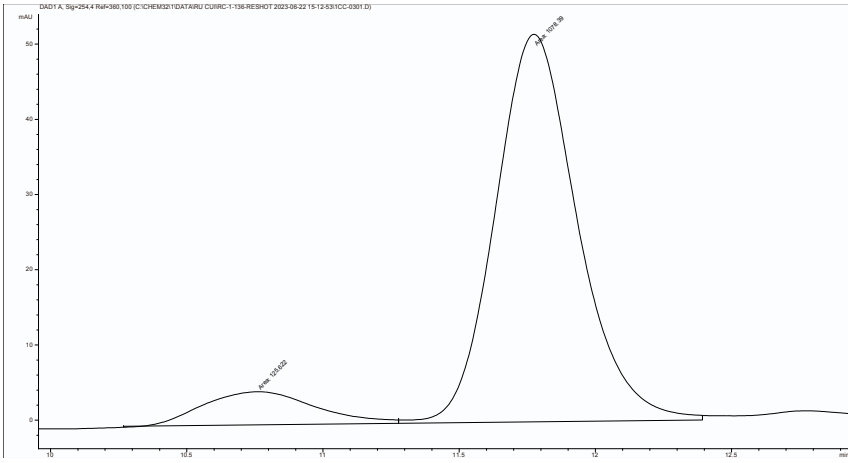
- Chiral HPLC analysis of enantioenriched **3j** product using CYP119 (F153G, L205W, T213A, V254A, C317S).



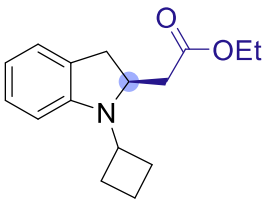


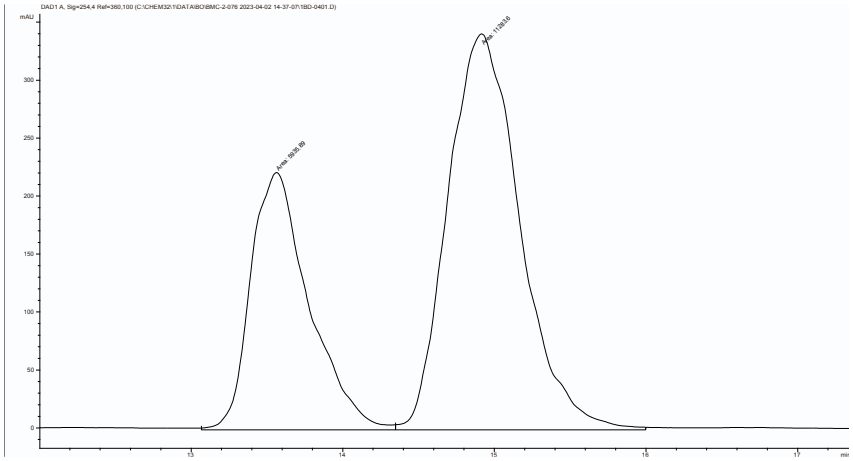
- Chiral HPLC analysis of enantioenriched **3k** product using CYP119 (A209W, G210S, T213G, V254A, C317S).



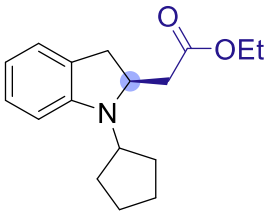


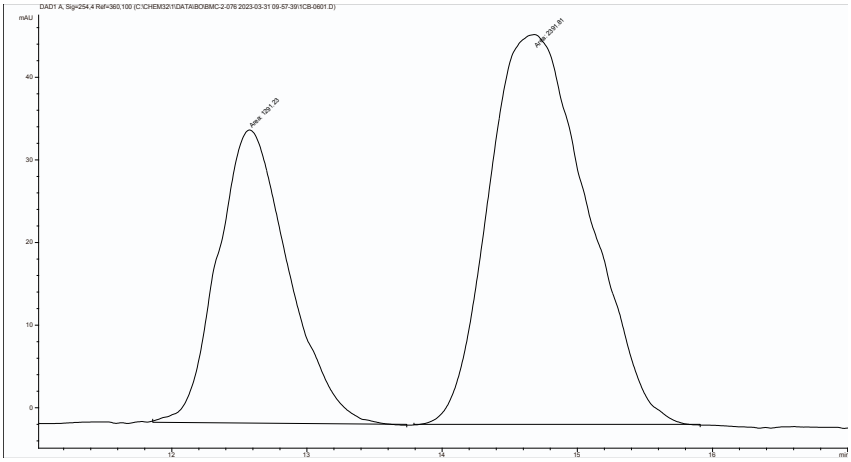
- Chiral HPLC analysis of enantioenriched **3l** product using CYP119 (L205Y, A209G, T213G, V254A, C317S).



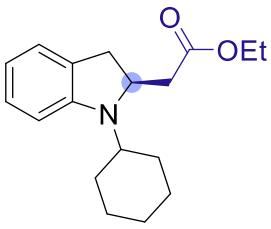


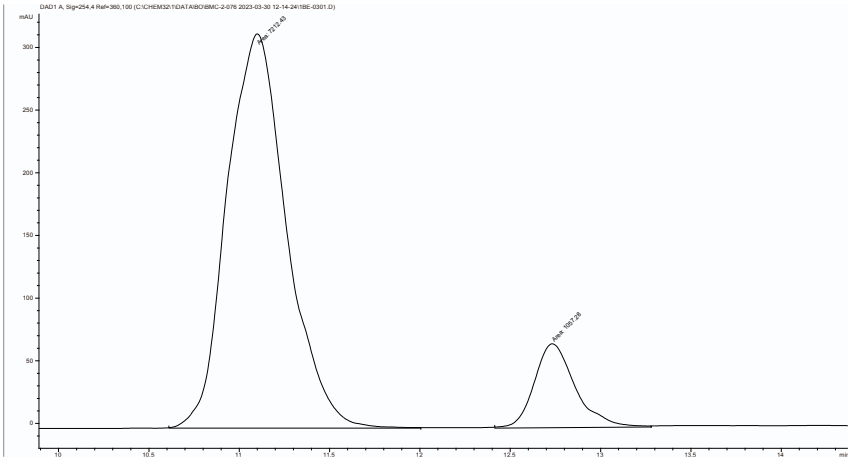
- Chiral HPLC analysis of enantioenriched **3m** product using CYP119 (L205Y, A209G, T213G, V254A, C317S).





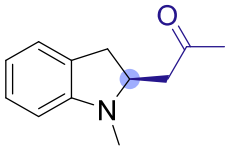
- Chiral HPLC analysis of enantioenriched **3l** product using CYP119 (L205Y, A209G, T213G, V254A, C317S).

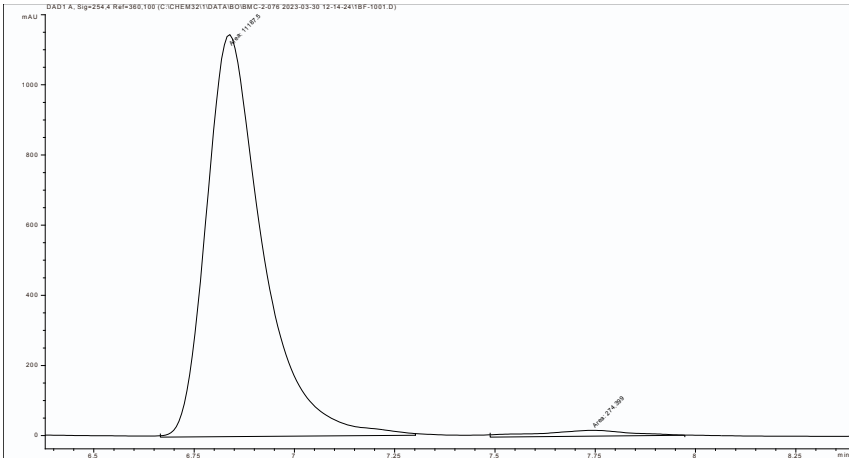




Data S3. Chiral HPLC analysis for the determination of an enantiomeric excess in the CYP119-catalyzed intermolecular C–H insertion reactions with DA (**2b**).

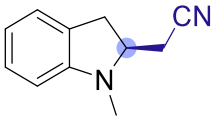
- Chiral HPLC analysis of enantioenriched **6a** using CYP119 (F153G, A209S, T213A, V254A, C317S).

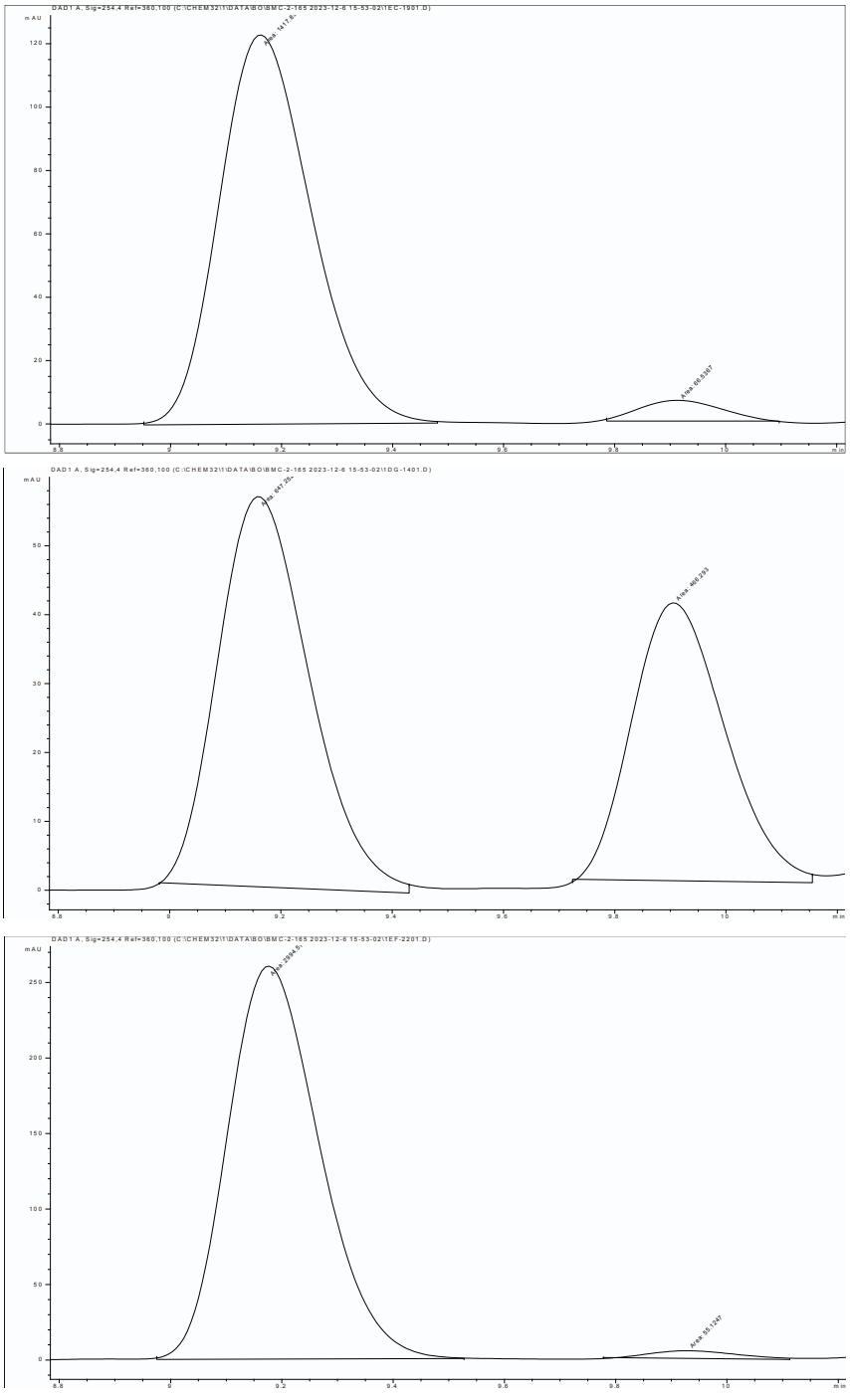




Data S4. Chiral HPLC analysis for the determination of an enantiomeric excess in the CYP119-catalyzed intermolecular C–H insertion reactions with DAN (**2c**).

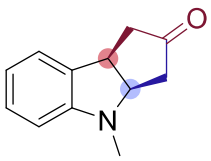
- Chiral HPLC analysis of enantioenriched **7a** using CYP119 (F153G, A209S, T213A, V254A, C317S) (*top*), racemic **7a** using CYP119 (F153G, A209G, T213G, V254A, C317S) (*Middle*), and enantiopure (*S*)-**7a** authentic standard using the described synthetic method (*bottom*).

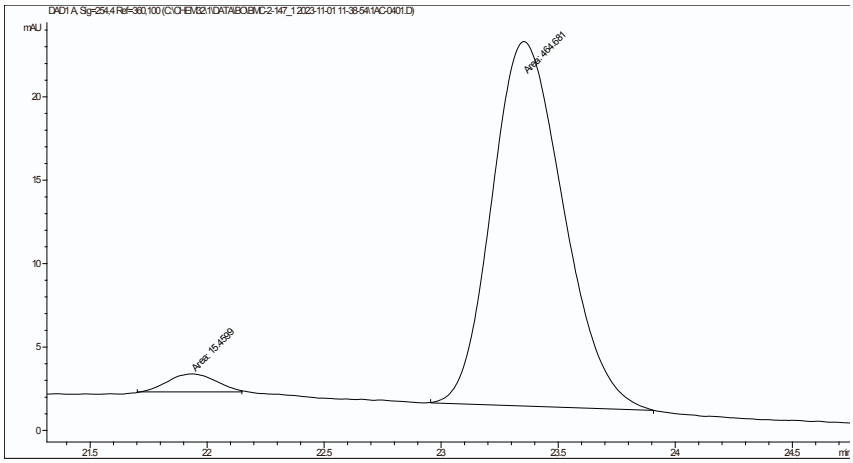




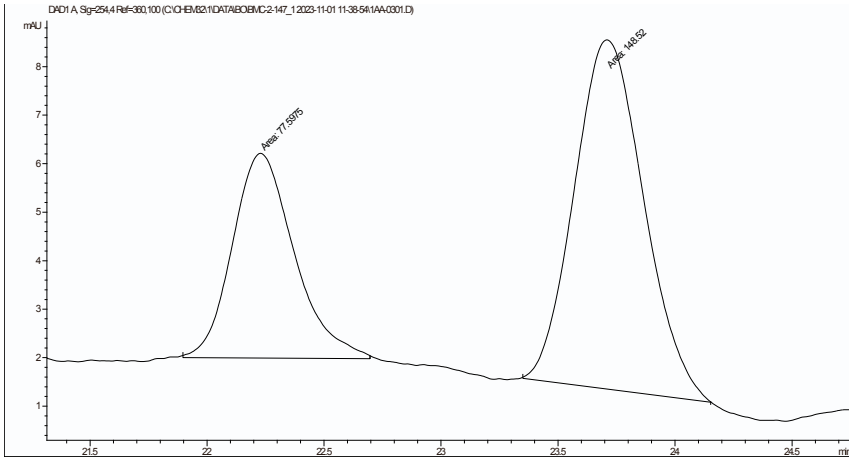
Data S5. Chiral HPLC analysis for the determination of an enantiomeric excess in the CYP119-catalyzed intramolecular C–H insertion reaction.

- Chiral HPLC analysis of enantioenriched **9a** using CYP119 (F153G, T213A, V254W, C317S).





- Chiral HPLC analysis of racemic **9a** using CYP119 (F153G, A209G, T213G, V254A, C317S) from racemic **8**.



Computational Details

All calculations were performed using the program Gaussian 09.[S1] All models investigated in this work were subject to full geometry optimizations without any symmetry constraints using the PCM method[S2] with a dielectric constant of 4.0 to simulate the protein environment effect as done previously.[S3] The frequency analysis was used to verify the nature of the stationary points on respective potential energy surfaces and to provide zero-point energy corrected electronic energies (E_{ZPE}'s), enthalpies (H's), and Gibbs free energies (G's) at 1 atm and experimental reaction temperatures, i.e. room temperature (RT). In addition to the analysis of vibrational mode of the imaginary frequency, intrinsic reaction coordinate calculations as implemented in Gaussian 09 were also used to verify that the nature of the calculated C–H insertion transition states. The atomic charges and spin densities reported here are from the Natural Population Analysis (NPA) and Mulliken schemes respectively, as implemented in Gaussian 09.[S1] All calculations were done using a range-separated hybrid DFT method with dispersion correction, ωB97XD,[S4] based on its excellent performance on heme carbenes and other catalytic systems from previous methodological studies.[S5, S6] Calculations were performed with the heme core part [Fe(Por)(MeO^-)($CHCO_2Et$)] (called $\mathbf{R_1}$). The basis set includes the effective core potential (ECP) basis LanL2DZ[7] for iron and the triple-zeta basis 6-311G(d) for all other elements, which was found to provide accurate predictions of various experimental reaction properties of heme carbenes.[S6, S8-S14] The use of a much larger 6-311++G(2d,2p) basis for all non-metal atoms was found to yield similar results for heme carbene reactions[S9] and thus further support the efficient use of the current basis set here. The use of an ECP basis for metal is common in many reaction studies involving transition metal carbenoids, such as Ir porphyrin carbene,[S15] Ru porphyrin carbene,[S16] and Rh carbene.[S17] The advantage of an ECP basis is the inclusion of relativistic effect basically absent in an all-electron basis set. In addition, it is available for all transition metals, which may allow direct comparisons of effects of a vast amount of metal centers. The alternative use of an all-electron basis for the metal center[S10] was recently found to yield qualitatively same conclusions of geometric, electronic, and energetic features for heme carbene reactions, and therefore supports the use of LanL2DZ basis here.

The KIE calculations with tunneling effect correction (KIE_W) were done using the following formulae reported recently:[S18]

$$KIE_E = e^{((-\Delta G_H + \Delta G_D)/RT)} \quad (1)$$

$$KIE_W = KIE_E \times \frac{Q_{tH}}{Q_{tD}} \quad (2)$$

$$Qt\ (\text{Tunneling correction}) = 1 + \frac{(\frac{h\upsilon}{kT})^2}{24} \quad (3)$$

where h is Planck's constant, υ is the imaginary frequency of the transition state, k is Boltzmann's constant, T is temperature, and subscripts of H and D indicate hydrogen and deuterium respectively. The final KIE_W values were calculated based on experimental 1:1 parallel reactions with and without deuterated species.

Conformation Studies

As seen in **Figure S6**, the final product $\mathbf{P_2}$ can adopt either *cis* or *trans* conformations regarding the phenyl group of the substrate on the same or opposite sides of the carbene substituent with respect to the formed C–C' bond. The energy differences between *cis* and *trans* conformations in $\mathbf{P_2}$ are small (~1 kcal/mol) due to smaller torsional strain of the C–C' single bond. The *cis* conformation has slightly more favorable electronic energy and enthalpy compared to the *trans* conformation (~1-1.5 kcal/mol) due to closer proximity of electronegative oxygens in the carbene substituent and the hydrogens on the 1-phenylpyrrolidine as seen in **Figure S6**: the closest distance between an oxygen atom on the carbene substituent and a hydrogen of the substrate is 2.477 Å in the *cis* conformation, while it is 2.649 Å in the *trans* conformation. The more compact *cis* conformation can also be more favorable in the enclosed active site in the protein environment. So, the *cis* conformation was used in the main text discussion.

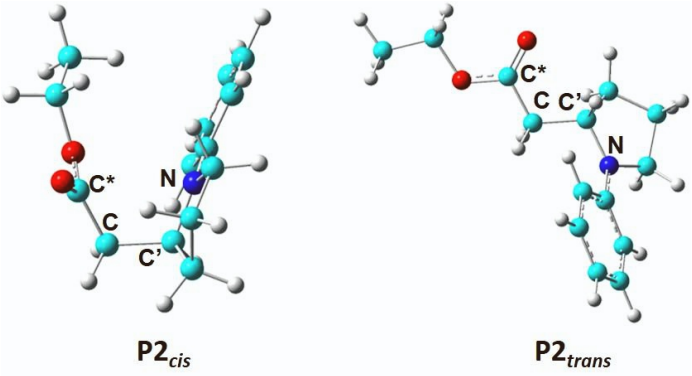


Figure S6. $\mathbf{P_2}$ with *cis* and *trans* conformations. Atom color scheme: C-cyan, N-blue, O- red, H-grey.

Table S9. Key Geometric Results and Relative Energies of $\mathbf{P_2}$ with Different Conformations

System	∠C*CC'N (◦)	ΔE (kcal/mol)	ΔE_{ZPE} (kcal/mol)	ΔH (kcal/mol)	ΔG (kcal/mol)
$\mathbf{P_{2\text{-}trans}}$	173.8	0.00	0.00	0.00	0.00
$\mathbf{P_{2\text{-}cis}}$	-23.6	-1.47	-0.75	-1.00	0.48

We then examined the two different conformations adopted by the species involved in the rate-limiting step. Since **TS** and $\mathbf{TS_1}$ (spin state effects are discussed in the next section, including closed-shell singlet (CSS), open-shell singlet (OSS), triplet, quintet) have similar structural features, in **Figure S7** only the structures of the optimized ^{CSS}**TS** transition states are shown to illustrate the differences. The energy differences between the more favorable *cis* conformations and *trans* conformations for **TS** and $\mathbf{TS_1}$ are approximately 4 kcal/mol (see **Table S10**), which is more significant than the conformation's effect observed in $\mathbf{P_2}$. The *cis* isomer is more stable than the *trans* conformation due to more favorable interactions between the carbonyl O on the carbene substituent with the H's of the two methylene groups in the pyrrolidine substrate. Taking the most favorable $^{OSS}\mathbf{TS_1}$ as an example, these two distances are 3.091 Å and 2.778 Å in the more favorable *cis* isomer, while the corresponding O···H distances in the *trans* isomer are longer, i.e. 3.849 Å and 4.155 Å, respectively. In addition, the shorter C–H_1 bond (by ~0.02 Å) in the optimized *cis* $^{OSS}\mathbf{TS_1}$ vs. *trans* $^{OSS}\mathbf{TS_1}$ structures indicate a larger degree of bond formation and thus stabilization, which is associated with lower barriers (see **Table S2**). After H_1 is transferred, the energy differences between the two conformations in **Int** are not as distinctive, with the most significant difference being for the quintet state ^{5}**Int** (1.35 kcal/mol). Based on the evidently favorable *cis* conformations in the transition states, *cis* conformations were used in the reaction pathway discussion in the next section of spin state study and the main text.

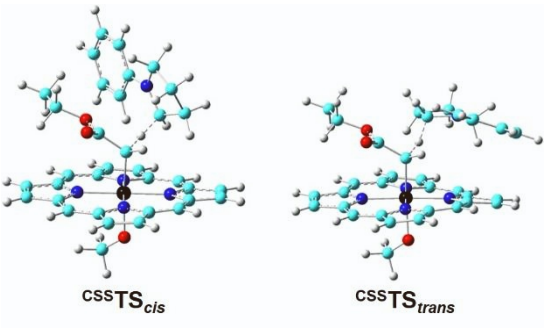


Figure S7. ^{CSS}**TS** in *cis* and *trans* conformations. Atom color scheme: C-cyan, N-blue, O- red, H-grey.

Table S10. Key Geometric Results and Relative Energies of **TS** and $\mathbf{TS_1}$ of Different Conformations

System	R_{CH1} (Å)	$R_{C'H1}$ (Å)	∠FeCC'N (◦)	ΔE (kcal/mol)	ΔE_{ZPE} (kcal/mol)	ΔH (kcal/mol)	ΔG (kcal/mol)
^{CSS}**TS**$_{cis}$	1.330	1.258	139.0	0.00	0.00	0.00	0.00
^{CSS}**TS**$_{trans}$	1.340	1.255	-78.6	4.91	4.80	5.15	3.90
^{OSS}**TS**$_{1\text{-}cis}$	1.338	1.354	139.1	0.00	0.00	0.00	0.00
^{OSS}**TS**$_{1\text{-}trans}$	1.359	1.349	-57.4	3.55	3.65	3.66	3.93
^{3}**TS**$_{1\text{-}cis}$	1.368	1.328	139.7	0.00	0.00	0.00	0.00
^{3}**TS**$_{1\text{-}trans}$	1.392	1.319	-57.0	3.18	3.20	3.20	3.73
^{5}**TS**$_{1\text{-}cis}$	1.376	1.320	139.4	0.00	0.00	0.00	0.00
^{5}**TS**$_{1\text{-}trans}$	1.393	1.316	-59.0	3.59	3.66	3.66	3.51
^{OSS}**Int**$_{cis}$	1.086	3.398	105.2	0.00	0.00	0.00	0.00
^{OSS}**Int**$_{trans}$	1.087	2.595	-63.9	1.83	1.51	2.15	0.16
^{3}**Int**$_{cis}$	1.086	3.399	105.2	0.00	0.00	0.00	0.00
^{3}**Int**$_{trans}$	1.086	2.595	-63.9	1.82	1.46	2.11	0.08
^{5}**Int**$_{cis}$	1.085	3.469	102.9	0.00	0.00	0.00	0.00
^{5}**Int**$_{trans}$	1.085	2.644	-66.5	0.79	0.76	0.76	1.35

Table S11. Geometric Parameters of *cis* and *trans* Conformations of Species in the C–H Insertion Pathways

System	R_{FeL} (Å)	R_{FeC} (Å)	$R_{CC'}$ (Å)	R_{CCO2Et} (Å)	R_{CH} (Å)	R_{CH1} (Å)	$R_{C'H1}$ (Å)	∠FeCC'N (◦)
^{CSS}**TS**$_{cis}$	1.987	1.946	2.503	1.476	1.092	1.330	1.258	139.0
^{CSS}**TS**$_{trans}$	1.982	1.953	2.455	1.478	1.090	1.340	1.255	-78.6
^{OSS}**TS**$_{1\text{-}cis}$	1.912	1.982	2.642	1.457	1.093	1.338	1.354	139.1
^{OSS}**TS**$_{1\text{-}trans}$	1.904	2.004	2.668	1.457	1.092	1.359	1.349	-57.4
^{3}**TS**$_{1\text{-}cis}$	1.891	2.037	2.650	1.451	1.093	1.368	1.328	139.7
^{3}**TS**$_{1\text{-}trans}$	1.884	2.059	2.680	1.454	1.093	1.392	1.319	-57.0
^{5}**TS**$_{1\text{-}cis}$	1.896	2.017	2.652	1.459	1.091	1.376	1.320	139.4
^{5}**TS**$_{1\text{-}trans}$	1.899	2.036	2.673	1.466	1.092	1.393	1.316	-59.0
^{OSS}**Int**$_{cis}$	1.886	2.084	4.436	1.452	1.089	1.086	3.398	105.2
^{OSS}**Int**$_{trans}$	1.891	2.069	3.493	1.453	1.089	1.087	2.595	-63.9
^{3}**Int**$_{cis}$	1.885	2.084	4.436	1.452	1.089	1.086	3.399	105.2
^{3}**Int**$_{trans}$	1.891	2.069	3.492	1.453	1.089	1.086	2.595	-63.9
^{5}**Int**$_{cis}$	1.895	2.099	4.505	1.460	1.088	1.085	3.469	102.9
^{5}**Int**$_{trans}$	1.906	2.089	3.520	1.460	1.087	1.085	2.644	-66.5

Table S12. Atomic Charges of *cis* and *trans* Conformations of Species in the C–H Insertion Pathways (unit: *e*)

System	Q_{Fe}	Q_{L}	Q_{C}	Q_{H}	Q_{CO2Et}	$Q_{C'}$	Q_{H1}	Q_{H2}	Q_{N}	$Q_{R'}$[a)]	Q_{Por}
$^{CSS}TS_{cis}$	0.242	-0.607	-0.323	0.203	-0.064	0.023	0.173	0.245	-0.442	-0.040	-0.853
$^{CSS}TS_{trans}$	0.244	-0.602	-0.318	0.219	-0.068	0.021	0.190	0.204	-0.452	-0.036	-0.853
$^{OSS}TS_{1\text{-}cis}$	0.465	-0.503	-0.536	0.196	-0.139	-0.080	0.224	0.209	-0.456	-0.176	-0.660
$^{OSS}TS_{1\text{-}trans}$	0.476	-0.492	-0.533	0.199	-0.138	-0.081	0.225	0.177	-0.457	-0.176	-0.657
$^{3}TS_{1\text{-}cis}$	0.503	-0.472	-0.554	0.190	-0.171	-0.097	0.226	0.206	-0.456	-0.187	-0.645
$^{3}TS_{1\text{-}trans}$	0.513	-0.463	-0.549	0.193	-0.158	-0.099	0.226	0.174	-0.458	-0.191	-0.645
$^{5}TS_{1\text{-}cis}$	0.879	-0.509	-0.541	0.202	-0.137	-0.098	0.223	0.208	-0.457	-0.182	-1.046
$^{5}TS_{1\text{-}trans}$	0.891	-0.508	-0.532	0.202	-0.120	-0.099	0.224	0.175	-0.458	-0.181	-1.052
$^{OSS}Int_{cis}$	0.527	-0.463	-0.693	0.215	-0.134	-0.030	0.208	0.177	-0.427	-0.150	-0.656
$^{OSS}Int_{trans}$	0.523	-0.473	-0.687	0.212	-0.133	-0.026	0.222	0.190	-0.434	-0.167	-0.660
$^{3}Int_{cis}$	0.527	-0.463	-0.693	0.215	-0.134	-0.030	0.208	0.177	-0.427	-0.150	-0.656
$^{3}Int_{trans}$	0.523	-0.474	-0.686	0.212	-0.133	-0.026	0.222	0.190	-0.434	-0.167	-0.661
$^{5}Int_{cis}$	0.906	-0.505	-0.661	0.218	-0.106	-0.028	0.213	0.178	-0.427	-0.150	-1.063
$^{5}Int_{trans}$	0.908	-0.522	-0.643	0.218	-0.108	-0.027	0.227	0.188	-0.432	-0.162	-1.078

a) R' is the substrate without C', H_1, and H_2.

Table S13. Spin Densities of *cis* and *trans* Conformations of Species in the C–H Insertion Pathways (unit: *e*)

System	$\rho_{\alpha\beta}^{Fe}$	$\rho_{\alpha\beta}^{L}$	$\rho_{\alpha\beta}^{C}$	$\rho_{\alpha\beta}^{H}$	$\rho_{\alpha\beta}^{CO2Et}$	$\rho_{\alpha\beta}^{R'}$	$\rho_{\alpha\beta}^{C'}$	$\rho_{\alpha\beta}^{H1}$	$\rho_{\alpha\beta}^{H2}$	$\rho_{\alpha\beta}^{Por}$
$^{OSS}TS_{1\text{-}cis}$	0.995	0.090	-0.496	0.017	-0.084	-0.091	-0.445	0.090	0.003	-0.077
$^{OSS}TS_{1\text{-}trans}$	0.995	0.088	-0.530	0.014	-0.058	-0.081	-0.449	0.093	0.000	-0.072
$^{3}TS_{1\text{-}cis}$	0.944	0.102	0.480	-0.018	0.113	0.083	0.426	-0.093	-0.001	-0.037
$^{3}TS_{1\text{-}trans}$	0.959	0.101	0.527	-0.014	0.058	0.071	0.429	-0.094	-0.001	-0.039
$^{3}TS_{1'\text{-}trans}$	3.004	0.051	-0.694	0.018	-0.071	-0.088	-0.444	0.097	-0.002	0.130
$^{5}TS_{1\text{-}cis}$	2.982	0.068	0.300	-0.008	0.086	0.084	0.420	-0.089	-0.001	0.159
$^{5}TS_{1\text{-}trans}$	3.012	0.065	0.332	-0.003	0.029	0.073	0.425	-0.089	-0.001	0.156
$^{OSS}Int_{cis}$	1.008	0.102	-0.056	0.006	-0.009	-0.225	-0.794	0.001	0.024	-0.058
$^{OSS}Int_{trans}$	1.012	0.100	-0.063	0.002	-0.010	-0.216	-0.810	0.014	0.029	-0.057
$^{3}Int_{cis}$	1.008	0.103	-0.056	0.006	-0.008	0.225	0.794	0.002	-0.024	-0.049
$^{3}Int_{trans}$	1.019	0.101	-0.056	0.002	-0.009	0.216	0.809	0.001	-0.029	-0.054
$^{5}Int_{cis}$	3.052	0.069	-0.250	0.009	-0.027	0.223	0.795	0.010	-0.024	0.144
$^{5}Int_{trans}$	3.090	0.068	-0.278	0.010	-0.031	0.216	0.806	0.008	-0.028	0.139

a) R' is the substrate without C', H_1, and H_2.

Spin State Analysis

Experimental studies of similar iron porphyrin carbenes (IPCs) showed the favorable spin states of singlet.[S6, S19] Recent computational studies have revealed that the Fe^{II} based CSS electronic structure of these systems, rather than the Fe^{III} based OSS, generate accurate prediction of various experimental spectroscopy properties.[S6, S9] The preferred Fe^{II} over Fe^{III} feature in the singlet of IPC was further supported by additional experimental spectroscopy studies[S19] and high-level multireference quantum chemical calculations.[S20] Based on these previous studies, the CSS state was used for **R_1**. For **TS_1**, as shown in **Table S14**, the OSS state has the lowest energy, while the triplet state is about 2 kcal/mol higher in energy and the quintet has much higher energy of ~ 11 kcal/mol. So, **$^{OSS}TS_1$** was used in the main text for mechanism discussion. For **Int**, quintet still has the highest energy, while the energy difference between antiferromagnetically coupled OSS state and the ferromagnetically coupled triplet state are below 1 kcal/mol (although triplet is slightly more favorable), indicating **Int** can exist in quantum mechanical admixed spin states as observed in other Fe^{III} porphyrins. Although optimizations of all three spin states were attempted for **TS_2**, only the triplet structure was successfully optimized, which has the same spin state as the most favorable **Int**. The results suggested that other spin states may not be stable compared to the triplet state. Thus, **^{3}Int** and **$^3TS_2$** are used in the final reaction pathway discussion. For **P_1**, [Fe(Por)(MeO^-)], several computational results show that $S = 2$ is the ground state[S13, S26, S27] and thus **$^5P_1$** was used in subsequent studies.

Table S14. Key Spin Densities and Relative Energies of Species with Different Spin States

System	$\rho_{\alpha\beta}^{Fe}$ (e)	$\rho_{\alpha\beta}^{C}$ (e)	$\rho_{\alpha\beta}^{C'}$ (e)	ΔE (kcal/mol)	ΔE_{ZPE} (kcal/mol)	ΔH (kcal/mol)	ΔG (kcal/mol)
$^{OSS}TS_1$	0.995	-0.496	-0.445	0.00	0.00	0.00	0.00
$^3TS_1$	0.944	0.480	0.426	3.58	3.38	3.52	2.06
$^5TS_1$	2.982	0.300	0.420	15.05	13.02	13.63	10.77
^{3}Int	1.008	-0.056	0.794	0.00	0.00	0.00	0.00
^{OSS}Int	1.008	-0.056	-0.794	0.00	-0.01	-0.01	0.63
^{5}Int	3.052	-0.250	0.795	12.05	10.28	11.31	7.51

Table S15. Geometric Parameters of Species with Different Spin States

System	R_{FeL} (Å)	R_{FeC} (Å)	$R_{CC'}$ (Å)	R_{CCO2Et} (Å)	R_{CH} (Å)	R_{CH1} (Å)	$R_{C'H1}$ (Å)	∠FeCC'N (°)
$^{OSS}TS_1$	1.912	1.982	2.642	1.457	1.093	1.338	1.354	139.1
$^3TS_1$	1.891	2.037	2.650	1.451	1.093	1.368	1.328	139.7
$^5TS_1$	1.896	2.017	2.652	1.459	1.091	1.376	1.320	139.4
^{3}Int	1.885	2.084	4.436	1.452	1.089	1.086	3.399	105.2
^{OSS}Int	1.886	2.084	4.436	1.452	1.089	1.086	3.398	105.2
^{5}Int	1.895	2.099	4.505	1.460	1.088	1.085	3.469	102.9

Table S16. Atomic Charges of Species with Different Spin States (unit: *e*)

System	Q_{Fe}	Q_L	Q_C	Q_H	Q_{CO2Et}	$Q_{C'}$	Q_{H1}	Q_{H2}	Q_N	$Q_{R'}$[a)]	Q_{Por}
$^{OSS}TS_1$	0.465	-0.503	-0.536	0.196	-0.139	-0.080	0.224	0.209	-0.456	-0.176	-0.660
$^3TS_1$	0.503	-0.472	-0.554	0.190	-0.171	-0.097	0.226	0.206	-0.456	-0.187	-0.645
$^5TS_1$	0.879	-0.509	-0.541	0.202	-0.137	-0.098	0.223	0.208	-0.457	-0.182	-1.046
^{OSS}Int	0.527	-0.463	-0.693	0.215	-0.134	-0.030	0.208	0.177	-0.427	-0.150	-0.656
^{3}Int	0.527	-0.463	-0.693	0.215	-0.134	-0.030	0.208	0.177	-0.427	-0.150	-0.656
^{5}Int	0.906	-0.505	-0.661	0.218	-0.106	-0.028	0.213	0.178	-0.427	-0.150	-1.063

a) R' is the substrate without C', H_1, and H_2.

Table S17. Spin Densities of Species with Different Spin States (unit: *e*)

System	$\rho_{\alpha\beta}^{Fe}$	$\rho_{\alpha\beta}^{L}$	$\rho_{\alpha\beta}^{C}$	$\rho_{\alpha\beta}^{H}$	$\rho_{\alpha\beta}^{CO2Et}$	$\rho_{\alpha\beta}^{R'}$	$\rho_{\alpha\beta}^{C'}$	$\rho_{\alpha\beta}^{H1}$	$\rho_{\alpha\beta}^{H2}$	$\rho_{\alpha\beta}^{Por}$
^{OSS}**TS$_1$**	0.995	0.090	-0.496	0.017	-0.084	-0.091	-0.445	0.090	0.003	-0.077
^{3}**TS$_1$**	0.944	0.102	0.480	-0.018	0.113	0.083	0.426	-0.093	-0.001	-0.037
^{5}**TS$_1$**	2.982	0.068	0.300	-0.008	0.086	0.084	0.420	-0.089	-0.001	0.159
^{OSS}**Int**	1.008	0.102	-0.056	0.006	-0.009	-0.225	-0.794	0.001	0.024	-0.058
^{3}**Int**	1.008	0.103	-0.056	0.006	-0.008	0.225	0.794	0.002	-0.024	-0.049
^{5}**Int**	3.052	0.069	-0.250	0.009	-0.027	0.223	0.795	0.010	-0.024	0.144

a) R' is the substrate without C', H_1, and H_2.

Energies, Geometric Parameters, Charges, and Spin Densities Results of Optimized Structure of the Most Favorable Species

Table S18. Absolute Energies of the Most Favorable Species in the C–H Insertion Pathway (unit: *Hartrees*)

Substrate	System	E	E_{ZPE}	H	G
10a	$^{1}R_1$	-1533.34355	-1532.92759	-1532.89930	-1532.98281
	$^{1}R_2$	-443.58942	-443.37724	-443.36704	-443.41237
	^{CSS}TS	-1976.94327	-1976.31566	-1976.27817	-1976.38015
	$^{OSS}TS_1$	-1976.95532	-1976.32942	-1976.29147	-1976.39482
	^{3}Int	-1976.98299	-1976.35303	-1976.31494	-1976.41984
	$^{3}TS_2$	-1976.96128	-1976.33286	-1976.29435	-1976.40119
	$^{5}P_1$	-1226.95490	-1226.64029	-1226.61800	-1226.69085
	$^{1}P_2$	-750.08373	-749.76925	-749.75272	-749.81237
1a	$^{1}R_2$	-404.28021	-404.09765	-404.08857	-404.13026
	^{CSS}TS	-1937.62951	-1937.03314	-1936.99602	-1937.09870
	$^{OSS}TS_1$	-1937.64105	-1937.04566	-1937.00847	-1937.11055

Table S19. Relative Energies of Species in C–H Insertion (unit: kcal/mol)

Substrate	System	ΔE	ΔE_{ZPE}	ΔH	ΔG
10a	$^{1}R_1$+$^{1}R_2$	0.00	0.00	0.00	0.00
	^{CSS}TS	-6.46	-6.79	-7.42	9.43
	$^{OSS}TS_1$	-14.02	-15.43	-15.77	0.23
	^{3}Int	-31.39	-30.24	-30.50	-15.48
	$^{3}TS_2$	-17.76	-17.59	-17.58	-3.77
	$^{5}P_1$+$^{1}P_2$	-66.30	-65.70	-65.50	-67.80
1a	$^{1}R1$+$^{1}R2$	0.00	0.00	0.00	0.00
	^{CSS}TS	-3.61	-4.96	-5.11	9.02
	$^{OSS}TS1$	-10.85	-12.81	-12.93	1.58

Table S20. Geometric Parameters of the Most Favorable Species in the C–H Insertion Pathway

Substrate	System	R_{FePor} (Å)	R_{FeL} (Å)	R_{FeC} (Å)	$R_{CC'}$ (Å)	R_{CCO2Et} (Å)	R_{CH} (Å)	R_{CH1} (Å)	$R_{C'H1}$ (Å)	∠FeCC'N (◦)
10a	$^{1}R_1$	2.020	1.967	1.773	/	1.475	1.098	/	/	/
	$^{1}R_2$	/	/	/	/	/	/	/	1.095	/
	^{CSS}TS	2.021	1.987	1.946	2.503	1.476	1.092	1.330	1.258	139.0
	$^{OSS}TS_1$	2.017	1.912	1.982	2.642	1.457	1.093	1.338	1.354	139.1
	^{3}Int	2.014	1.885	2.084	4.436	1.452	1.089	1.086	3.399	105.2
	$^{3}TS_2$	2.023	1.993	2.345	2.543	1.441	1.072	1.079	2.744	-132.6
	$^{5}P_1$	2.119	1.973	/	/	/	/	/	/	/
	$^{1}P_2$		/	/	1.562	1.508	1.091	1.09	2.193	/
1a	$^{1}R_2$	/	/	/	/	/	/	/	1.092	/
	^{CSS}TS	2.020	1.985	1.941	2.531	1.476	1.093	1.346	1.255	109.9
	$^{OSS}TS_1$	2.017	1.910	1.986	2.663	1.458	1.092	1.357	1.346	107.3

Table S21. Atomic Charge of the Most Favorable Species Involved in the C–H Insertion Pathway (unit: *e*)

Sub.	Sys.	Q_{Fe}	Q_L	Q_C	Q_H	Q_{CO2Et}	$Q_{C'}$	Q_{H1}	Q_{H2}	Q_N	$Q_{R'}$[a]	Q_{Por}
10a	^{1}**R$_1$**	0.138	-0.571	-0.011	0.183	-0.072	/	/	/	/	/	-0.666
	^{1}**R$_2$**	/	/	/	/	/	-0.172	0.192	0.176	-0.470	-0.196	/
	^{CSS}**TS**	0.242	-0.607	-0.322	0.204	-0.063	0.018	0.177	0.244	-0.441	-0.043	-0.851
	^{OSS}**TS$_1$**	0.465	-0.503	-0.536	0.196	-0.139	-0.080	0.224	0.209	-0.456	-0.176	-0.660
	^{3}**Int**	0.527	-0.463	-0.693	0.215	-0.134	-0.030	0.208	0.177	-0.427	-0.150	-0.656
	^{3}**TS$_2$**	0.503	-0.472	-0.554	0.19	-0.171	-0.097	0.226	0.206	-0.456	-0.187	-0.645
	^{5}**P$_1$**	1.082	-0.696	/	/	/	/	/	/	/	/	-1.386
	^{1}**P$_2$**	/	/	-0.490	0.223	0.007	0.014	0.220	0.203	-0.494	-0.061	/
1a	^{1}**R2**	/	/	/	/	/	-0.151	0.193	0.166	-0.487	-0.208	/
	^{CSS}**TS**	0.239	-0.608	-0.315	0.203	-0.064	0.029	0.175	0.24	-0.434	-0.052	-0.848
	^{OSS}**TS1**	0.464	-0.502	-0.533	0.197	-0.13	-0.067	0.216	0.204	-0.462	-0.182	-0.667

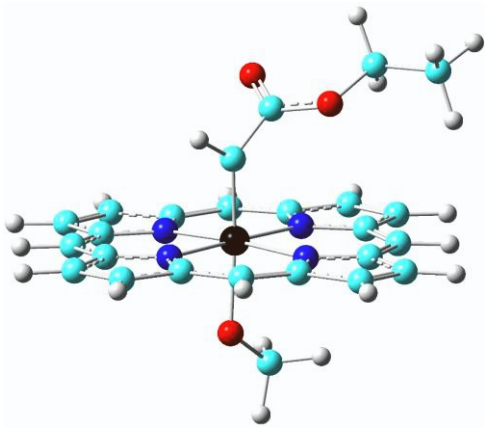
a) R' is the substrate without C', H_1, and H_2.

Table S22. Spin Densities of the Most Favorable Species Involved in the C–H Insertion Pathway (unit: *e*)

Substrate	System	$\rho_{\alpha\beta}^{Fe}$	$\rho_{\alpha\beta}^{L}$	$\rho_{\alpha\beta}^{C}$	$\rho_{\alpha\beta}^{H}$	$\rho_{\alpha\beta}^{CO2Et}$	$\rho_{\alpha\beta}^{R'}$	$\rho_{\alpha\beta}^{C'}$	$\rho_{\alpha\beta}^{H1}$	$\rho_{\alpha\beta}^{H2}$	$\rho_{\alpha\beta}^{Por}$
10a	^{OSS}**TS$_1$**	0.995	0.090	-0.496	0.017	-0.084	-0.091	-0.445	0.090	0.003	-0.077
	^{3}**Int**	1.008	0.103	-0.056	0.006	-0.008	0.225	0.794	0.002	-0.024	-0.049
	^{3}**TS$_2$**	1.654	0.118	-0.438	0.021	-0.071	0.163	0.641	0.021	-0.018	-0.091
	^{5}**P$_1$**	3.716	0.127	/	/	/	/	/	/	/	0.157
1a	^{OSS}**TS1**	0.983	0.086	-0.511	0.018	-0.060	-0.097	-0.435	0.091	0.005	-0.079

Data S6. Cartesian Coordinates and 3d Structures of Optimized Structures of the Most Favorable Conformations and Spin States

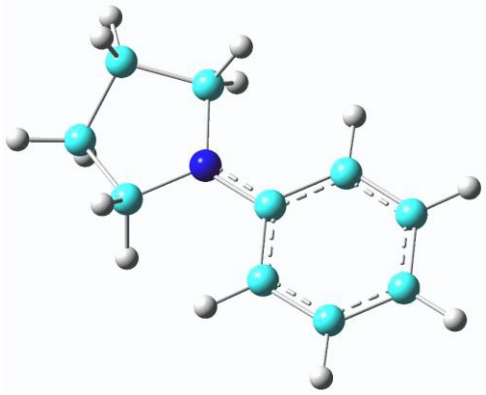
Cartesian Coordinates and 3d Structures of Optimized Structures of $^1R_1$



$^1R_1$

Fe	0.08964500	-0.00853300	-0.18697000
N	-1.68587600	0.98505900	-0.19643000
C	-2.91193500	0.47076200	-0.47840900
C	-1.88095100	2.29829400	0.09445800
C	-3.92006400	1.49898400	-0.38719400
C	-3.27738800	2.63812600	-0.03260100
H	-4.97700000	1.34926400	-0.56511600
H	-3.69524400	3.62103400	0.14092200
N	1.01406100	1.66152800	0.44749300
C	2.34047200	1.80332100	0.69818100
C	0.44559300	2.87546300	0.66461400
C	2.62683500	3.16008400	1.10826300
C	1.44975000	3.82607800	1.08724900
H	3.60740200	3.53268100	1.37375000
H	1.25735200	4.86265400	1.33038900
N	1.83664000	-1.00486300	-0.14009800
C	2.02611700	-2.31627600	-0.42727400
C	3.05264700	-0.49915500	0.18320000
C	3.42396100	-2.66070300	-0.29049500
C	4.06293700	-1.52945500	0.09001100
H	3.84014500	-3.64583000	-0.45589800
H	5.11459600	-1.38919200	0.30291400
N	-0.88908000	-1.70542400	-0.66623000
C	-0.32358000	-2.92322300	-0.86877100
C	-2.22293300	-1.85897800	-0.86826200
C	-1.34025100	-3.89355100	-1.20847000
C	-2.52070800	-3.23279000	-1.20508800
H	-1.15400000	-4.93944100	-1.41395800
H	-3.51059200	-3.61992900	-1.40802700
C	-3.17365300	-0.85498600	-0.78426600
C	1.02799300	-3.21513600	-0.77482400
C	-0.89437100	3.18519600	0.49464000
C	3.29831200	0.80954600	0.57136000
H	1.32816400	-4.23972400	-0.97163800
H	-4.20709200	-1.13021400	-0.97145400
H	-1.19691700	4.20708000	0.70035700
H	4.32431200	1.07602200	0.80554200
O	-0.07041100	-0.54629500	1.69941500
C	-1.18845200	-0.26205700	2.43592500
H	-2.12463000	-0.69102800	2.01799400
H	-1.10287600	-0.66609900	3.46518400
H	-1.38487400	0.82633600	2.55060900
C	0.39130300	0.37847800	-1.88906000
H	1.33137600	0.10292200	-2.38435900
C	-0.50717400	1.14896600	-2.76684100
O	-1.47396100	0.39922900	-3.30497000
O	-0.39595500	2.34002300	-2.95701300
C	-2.54105100	1.10206300	-3.95484300
H	-2.96861700	1.81208700	-3.24184500
H	-2.14219300	1.66673200	-4.80130600
C	-3.55597400	0.07130900	-4.38973500
H	-4.40272000	0.56173000	-4.87690000
H	-3.11535000	-0.63732700	-5.09496900
H	-3.92314800	-0.48562700	-3.52523900

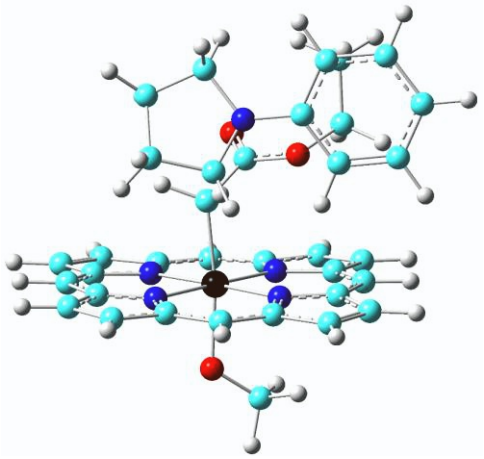
Cartesian Coordinates and 3d Structures of Optimized Structures of $^1\mathbf{R}_2$



$^1\mathbf{R}_2$

C	-0.80975500	-0.99665600	0.05367300
C	-1.21961800	0.07218900	-0.76816500
C	-0.65210100	1.33037400	-0.62945300
C	0.33596400	1.57469500	0.31856300
C	0.74721600	0.52523500	1.13347400
C	0.19050000	-0.73938100	1.01231200
H	-1.97709100	-0.08591200	-1.52705400
H	-0.98850500	2.13126800	-1.28112000
H	0.77622100	2.56042200	0.42013200
H	1.51442200	0.69044100	1.88401600
H	0.52491100	-1.52924700	1.67478300
C	-2.48991300	-2.53621400	-0.94529900
C	-0.88987300	-3.41083000	0.64868000
C	-2.91630000	-3.94037600	-0.51673500
H	-2.19327000	-2.51241800	-2.00408100
H	-3.29342000	-1.80411800	-0.81243700
C	-1.60691600	-4.57174300	-0.04024400
H	-1.14985600	-3.35258500	1.71567300
H	0.19904400	-3.50393800	0.57926600
H	-3.62629700	-3.87565800	0.31325000
H	-3.39398100	-4.49523000	-1.32574100
H	-1.75515400	-5.42346200	0.62543300
H	-1.02021300	-4.91182900	-0.89892100
N	-1.36348000	-2.24182500	-0.07456700

Cartesian Coordinates and 3d Structures of Optimized Structures of $^{CSS}TS_{cis}$

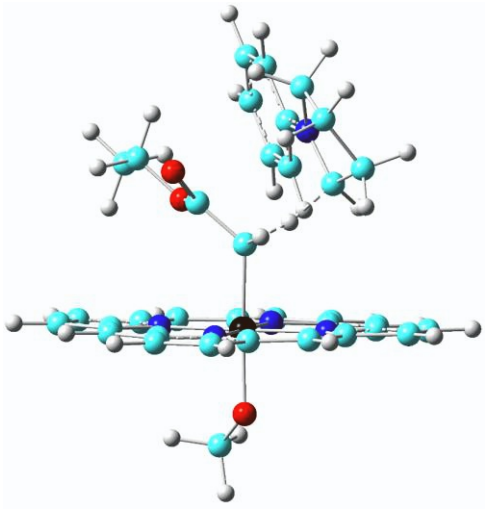


$^{CSS}TS_{cis}$

Fe	0.12304300	-0.26854400	-0.52006700	H	4.40582500	1.06158700	-0.72096900
N	-1.66911800	0.66728200	-0.44509300	O	0.25577300	-0.28890300	1.46256400
C	-2.89606400	0.08646300	-0.44531400	C	-0.86025700	-0.40852300	2.24033900
C	-1.87655300	2.00550800	-0.34607800	H	-1.48255200	-1.30053000	2.00570900
C	-3.92810600	1.09552300	-0.35880900	H	-0.59871400	-0.50341000	3.31570600
C	-3.29357700	2.28973100	-0.29063700	H	-1.55838900	0.45771800	2.18072400
H	-4.99204400	0.89710300	-0.33982300	C	0.20957100	-0.07488200	-2.45720300
H	-3.72731400	3.27780700	-0.20828200	H	1.14826900	0.33154500	-2.84210100
N	1.06763200	1.51962300	-0.48057900	C	-0.93285500	0.46694200	-3.21780200
C	2.40465500	1.73517000	-0.55651700	O	-1.02180500	1.62955700	-3.55646500
C	0.48215000	2.73973700	-0.38319300	H	0.30723100	-1.32471300	-2.89069500
C	2.68593500	3.15500500	-0.50478600	C	-1.70347700	-3.21817300	-6.26322600
C	1.49062300	3.77881200	-0.38909600	C	-1.88572900	-3.39096500	-3.86547400
H	3.67345300	3.59654400	-0.54302400	C	-2.93237100	-3.85479800	-6.39457400
H	1.29045600	4.83993100	-0.31561900	H	-1.17376600	-2.90266100	-7.15439100
N	1.90630900	-1.21074800	-0.56784900	C	-3.10513100	-4.02867500	-4.01256200
C	2.12092000	-2.55108100	-0.53477400	H	-1.52381800	-3.16497000	-2.87204800
C	3.12945800	-0.62616200	-0.64772600	C	-3.64405200	-4.26552000	-5.27507800
C	3.53775000	-2.83382800	-0.61270600	H	-3.33460400	-4.02535400	-7.38829600
C	4.16493900	-1.63599800	-0.68948500	H	-3.64461000	-4.32668200	-3.11907400
H	3.97739000	-3.82286800	-0.60229200	H	-4.60495100	-4.75703900	-5.38253300
H	5.22687700	-1.43646300	-0.75395000	C	-1.16571100	-2.97562900	-4.99305100
N	-0.82351100	-2.05343400	-0.50176400	N	0.05375400	-2.32434500	-4.85520000
C	-0.23776500	-3.27895400	-0.42532400	C	0.76231000	-2.20909700	-3.66456000
C	-2.16625700	-2.26794600	-0.46718900	C	0.74408900	-1.69007600	-5.97655900
C	-1.24607200	-4.31221600	-0.34608900	C	2.17202900	-1.82203500	-4.04665600
C	-2.44543000	-3.68296700	-0.37407300	H	0.64436400	-3.03330000	-2.96222900
H	-1.04503600	-5.37321100	-0.27461300	C	1.94824900	-1.00599200	-5.32349000
H	-3.43422200	-4.12035700	-0.32353400	H	0.08368200	-0.98550200	-6.49052800
C	-3.13798700	-1.27870500	-0.47301700	H	1.05828300	-2.45056600	-6.70220600
C	1.13004800	-3.52060800	-0.45126300	H	2.69019100	-1.28665300	-3.25186200
C	-0.88349100	2.97366900	-0.31205000	H	2.73630300	-2.73614000	-4.26015600
C	3.36949300	0.74212400	-0.65961200	H	1.69949500	0.02717700	-5.06997300
H	1.45111200	-4.55723900	-0.40734100	H	2.81780000	-0.99048100	-5.98122300
H	-4.17500600	-1.60083900	-0.45170100	O	-1.90271800	-0.43425800	-3.41079000
H	-1.20424100	4.00834700	-0.23531100	C	-3.15845200	0.03088600	-3.90779700

H	-3.87486100	-0.71458400	-3.56084000
H	-3.39024100	0.98669400	-3.43474500
C	-3.16586200	0.13636200	-5.41988200
H	-4.15362300	0.45266300	-5.76773100
H	-2.43251300	0.87156200	-5.75652700
H	-2.93629900	-0.83074300	-5.87203700

Cartesian Coordinates and 3d Structures of Optimized Structures of ^{OSS}**TS**$_{1\text{-}cis}$

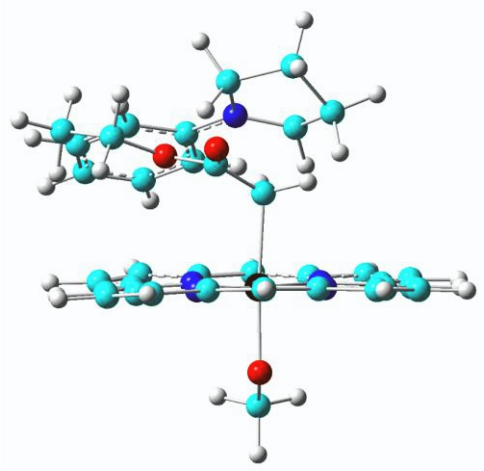


^{OSS}**TS**$_{1\text{-}cis}$

Fe	-0.00455400	-0.21149100	-0.47369200	H	-4.31498000	-1.49244500	-0.70653200
N	-1.77589600	0.73540600	-0.58743200	H	-1.27406300	4.07920700	-0.61323000
C	-3.00926400	0.17295400	-0.66094100	H	4.30954000	1.07142000	-0.28342400
C	-1.96192700	2.07929000	-0.62807700	O	-0.04384000	-0.20930300	1.43755000
C	-4.02096900	1.19777100	-0.75216100	C	-1.18291800	-0.45262000	2.16981200
C	-3.36992900	2.38426200	-0.73661700	H	-1.61522500	-1.45564000	1.98728100
H	-5.08491600	1.01366000	-0.82220500	H	-0.96852100	-0.39490100	3.25192100
H	-3.78591700	3.38157600	-0.78951200	H	-1.99580800	0.27192200	1.97126200
N	0.96750500	1.55931500	-0.40139400	C	0.28337700	-0.09697700	-2.43147900
C	2.30771200	1.76058600	-0.33636500	H	1.27616100	0.31683700	-2.62334700
C	0.39320800	2.78803400	-0.45560500	C	-0.64345900	0.47564600	-3.39825100
C	2.59906500	3.17615200	-0.32691600	O	-0.40443600	1.41210600	-4.15091300
C	1.40881600	3.81474300	-0.40211700	H	0.45022000	-1.35042200	-2.86988500
H	3.59089100	3.60520200	-0.27387800	C	-1.76957200	-2.77221000	-6.32639000
H	1.21565500	4.87899600	-0.42608000	C	-1.79880600	-3.47075700	-4.01980600
N	1.77432600	-1.16320000	-0.38815200	C	-3.01120300	-3.36685300	-6.51097000
C	1.96446100	-2.50730000	-0.38067000	H	-1.29652700	-2.26895500	-7.16151300
C	3.00832700	-0.59802900	-0.35079500	C	-3.03107300	-4.06621100	-4.22466000
C	3.37641100	-2.81079900	-0.34616000	H	-1.37254200	-3.47084400	-3.02608900
C	4.02561000	-1.62279500	-0.32977900	C	-3.65234100	-4.03031300	-5.47099300
H	3.79710400	-3.80757000	-0.33238700	H	-3.48087800	-3.30715000	-7.48845700
H	5.09126100	-1.43811400	-0.29757000	H	-3.52002800	-4.54683700	-3.38227700
N	-0.97286800	-1.97575200	-0.54418600	H	-4.62194500	-4.49256600	-5.62205200
C	-0.39920100	-3.20549700	-0.45606300	C	-1.12726000	-2.82738300	-5.07590600
C	-2.31626800	-2.17998900	-0.59004900	N	0.12103800	-2.28422900	-4.89445800
C	-1.41672800	-4.22899000	-0.44861400	C	0.86334400	-2.34636600	-3.68909600
C	-2.60867300	-3.59055600	-0.53240700	C	0.81786100	-1.53263500	-5.93141000
H	-1.22501000	-5.29207200	-0.38814500	C	2.28121800	-1.97967400	-4.06159400
H	-3.60172200	-4.01918100	-0.55334700	H	0.72035600	-3.25478600	-3.10754300
C	-3.27570500	-1.18442600	-0.65500600	C	2.07481300	-1.00662800	-5.22952900
C	0.96172700	-3.46419800	-0.39486500	H	0.19421300	-0.71687700	-6.31104800
C	-0.96452400	3.04005700	-0.56628300	H	1.06530700	-2.19046800	-6.77601400
C	3.26957200	0.76333200	-0.32255300	H	2.81942200	-1.53594900	-3.22095100
H	1.26663400	-4.50528400	-0.35668000	H	2.83572000	-2.86854700	-4.38777800

H	1.88216200	-0.00243300	-4.84809400
H	2.93020700	-0.95718300	-5.90560800
O	-1.84127000	-0.14862100	-3.43441900
C	-2.77772300	0.33389300	-4.38870800
H	-2.94074500	1.40394200	-4.22958100
H	-2.36437000	0.20554400	-5.39496700
C	-4.05646800	-0.45304100	-4.22297400
H	-4.79589700	-0.12203600	-4.95834600
H	-3.87520400	-1.51829300	-4.36945700
H	-4.47039400	-0.30478000	-3.22344400

Cartesian Coordinates and 3d Structures of Optimized Structures of $^{3}\mathbf{Int}_{cis}$

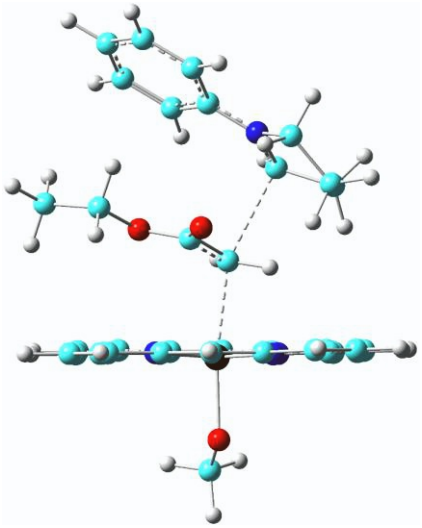


$^{3}\mathbf{Int}_{cis}$

Fe	-0.12040300	0.06266300	-0.33958000	H	4.28886200	0.70950500	0.07415700
N	-1.75191400	1.21838200	-0.08818000	O	-0.20646000	-0.37263600	1.49295800
C	-3.04412100	0.81338000	-0.18113100	C	-0.04939900	-1.67047400	1.93630300
C	-1.78444600	2.53499500	0.24593400	H	0.94282500	-2.09549700	1.69755600
C	-3.93452800	1.91557200	0.10157400	H	-0.16044900	-1.71427800	3.03290300
C	-3.15073200	2.98792900	0.35765500	H	-0.79529900	-2.36676600	1.51101700
H	-5.01484000	1.85552100	0.09362500	C	0.10440700	0.56407200	-2.35019800
H	-3.44866200	3.99785900	0.60594700	H	1.00290900	1.17803800	-2.39691100
N	1.03822900	1.64375200	0.13410600	C	-1.04315300	1.26220900	-2.90231700
C	2.39479200	1.63838100	0.22228400	O	-1.23945200	2.46182700	-2.96466300
C	0.63597000	2.90199700	0.43887800	H	0.25169400	-0.40611500	-2.81558500
C	2.87199400	2.94640600	0.60474400	C	-0.22109400	-5.13172500	-4.51632200
C	1.77793700	3.73354100	0.73531400	C	-1.95649500	-3.53454300	-5.02700400
H	3.91195800	3.20828400	0.74844200	C	-1.18507300	-6.12057900	-4.38951500
H	1.72817200	4.77938200	1.00733800	H	0.82377800	-5.38774700	-4.39015600
N	1.51114600	-1.08377200	-0.63625100	C	-2.90444800	-4.53633200	-4.90219600
C	1.54710900	-2.36755500	-1.07869600	H	-2.28138700	-2.52293100	-5.23272800
C	2.80168900	-0.70393900	-0.44095900	C	-2.53501200	-5.83919700	-4.57968300
C	2.91013100	-2.82818600	-1.15681600	H	-0.87198300	-7.13131400	-4.14523300
C	3.69082900	-1.79488000	-0.75641200	H	-3.95190400	-4.28816400	-5.04421800
H	3.21301400	-3.81588100	-1.47839000	H	-3.28272400	-6.61826500	-4.47884000
H	4.76944000	-1.75634800	-0.68070800	C	-0.59077800	-3.81023100	-4.83199900
N	-1.28076600	-1.48462900	-0.87705600	N	0.34826700	-2.80988700	-4.94534400
C	-0.87423100	-2.70669900	-1.31826600	C	1.69010200	-2.90378200	-4.61627600
C	-2.63775100	-1.50905700	-0.88078900	C	-0.00801500	-1.46343800	-5.39972200
C	-2.01616200	-3.53327500	-1.62523400	C	2.27878200	-1.52929000	-4.58570300
C	-3.11173400	-2.79325300	-1.34054300	H	1.99433400	-3.67902700	-3.92477900
H	-1.96225500	-4.53621300	-2.02692000	C	1.33504500	-0.73248400	-5.50039400
H	-4.15355500	-3.06618100	-1.44567000	H	-0.68158900	-0.99289700	-4.67938200
C	-3.46859000	-0.45762700	-0.53169100	H	-0.52466900	-1.51462100	-6.36348100
C	0.43836900	-3.12704300	-1.42181800	H	2.26700300	-1.12243100	-3.56459500
C	-0.68114000	3.33417600	0.48152800	H	3.31680000	-1.49995300	-4.92869100
C	3.22010400	0.55472500	-0.03557100	H	1.24206900	0.31174200	-5.19925400
H	0.60912500	-4.13320000	-1.78974700	H	1.70058600	-0.76027200	-6.53049700
H	-4.53890600	-0.63197900	-0.57692700	O	-1.98671100	0.38468300	-3.36296100
H	-0.85978600	4.37317300	0.73844100	C	-3.24353000	0.92988300	-3.73544100

H	-3.68607100	1.43039200	-2.86896700
H	-3.10945200	1.68186200	-4.51853600
C	-4.10847600	-0.21894600	-4.20680000
H	-5.12082200	0.13123100	-4.42659500
H	-3.69956900	-0.66856800	-5.11587100
H	-4.16181700	-0.99236600	-3.43703200

Cartesian Coordinates and 3d Structures of Optimized Structures of $^{3}TS_{2\text{-}cis}$

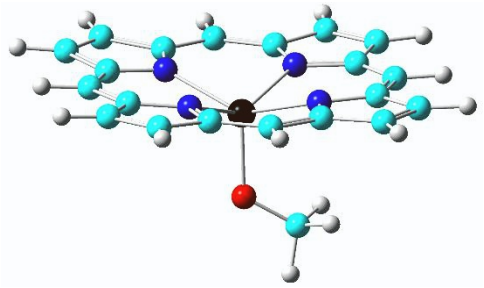


$^{3}TS_{2\text{-}cis}$

Fe	-0.14041300	-0.19477000	-0.60174300	H	-4.29810400	-1.75173900	0.03701000
N	-1.98783700	0.63342500	-0.59743900	H	-1.78031400	3.97093800	-1.09414400
C	-3.15717600	-0.01197300	-0.35015300	H	3.98276000	1.33082100	-1.47884200
C	-2.29540900	1.95096900	-0.73466400	O	0.12648400	0.14179700	1.34489300
C	-4.25245600	0.92824400	-0.33900700	C	0.37700500	-0.90174100	2.19836000
C	-3.71617000	2.14882900	-0.57967200	H	1.28764800	-1.48174300	1.93339600
H	-5.28967700	0.67168300	-0.16758900	H	0.53253900	-0.55266500	3.23969900
H	-4.22028300	3.10356200	-0.64925900	H	-0.44460300	-1.64811300	2.24651100
N	0.64395000	1.61717300	-1.01774400	C	-0.12166100	-0.54525600	-2.92004600
C	1.95493200	1.88736900	-1.25609400	H	0.75324200	0.07375800	-2.94562700
C	-0.01906700	2.79582000	-1.13188100	C	-1.35695400	0.02673500	-3.39260600
C	2.13251400	3.29733300	-1.52417800	O	-1.51806600	1.16814500	-3.78908000
C	0.90568900	3.86253200	-1.44327200	H	-0.02533000	-1.61649300	-2.83552800
H	3.07966900	3.77369200	-1.74054600	C	-2.05076600	0.23105100	-7.51696500
H	0.63277500	4.90001300	-1.58414300	C	-1.69746900	-1.82561900	-6.30804000
N	1.68863200	-1.05828900	-0.80034100	C	-3.24049600	-0.27779200	-8.01771400
C	1.99034600	-2.37852700	-0.70547800	H	-1.74373900	1.23076300	-7.79944500
C	2.85423600	-0.40912700	-1.05637900	C	-2.88478300	-2.32005900	-6.82269400
C	3.40638600	-2.57921500	-0.89436400	H	-1.14019400	-2.42365700	-5.59949500
C	3.94386800	-1.35348900	-1.11377900	C	-3.67133800	-1.55647300	-7.68011200
H	3.90864600	-3.53710200	-0.86224100	H	-3.83885000	0.34005100	-8.68034100
H	4.97936700	-1.09851300	-1.29760900	H	-3.20689500	-3.31548300	-6.53243500
N	-0.95066300	-2.02430300	-0.35855600	H	-4.60284100	-1.94903300	-8.07323400
C	-0.29602700	-3.21774100	-0.32408200	C	-1.25498700	-0.53488200	-6.64805200
C	-2.26663700	-2.30373700	-0.16830300	N	-0.07452400	-0.02807100	-6.14003100
C	-1.23261200	-4.29481000	-0.09655200	C	0.78439700	-0.69844500	-5.29111400
C	-2.45623500	-3.72599600	0.00326200	C	0.31041000	1.37708300	-6.29441300
H	-0.96965900	-5.34208700	-0.02577000	C	2.01108600	0.14277200	-5.11456400
H	-3.41012700	-4.20817600	0.17185400	H	0.83066500	-1.77768200	-5.35317200
C	-3.29578500	-1.37544700	-0.14318900	C	1.44623200	1.55848600	-5.28405100
C	1.06874700	-3.39398900	-0.48485000	H	-0.53624000	2.02743900	-6.07074400
C	-1.38695600	2.96508600	-0.98722300	H	0.63811600	1.56088800	-7.32564300
C	2.98461900	0.95709800	-1.27149100	H	2.50213000	-0.01169100	-4.15077100
H	1.44986500	-4.40925300	-0.43389700	H	2.75092100	-0.07955800	-5.89839800

H	1.02545800	1.90409500	-4.33821000
H	2.18878700	2.28347600	-5.62096400
O	-2.37411100	-0.86682800	-3.37094300
C	-3.63660400	-0.38775900	-3.82067800
H	-3.91647000	0.49224100	-3.23520300
H	-3.55604800	-0.08863000	-4.87062800
C	-4.63388100	-1.50924100	-3.64350900
H	-5.62143600	-1.19379500	-3.99195100
H	-4.33099000	-2.38762100	-4.21879000
H	-4.70816100	-1.79258100	-2.59101300

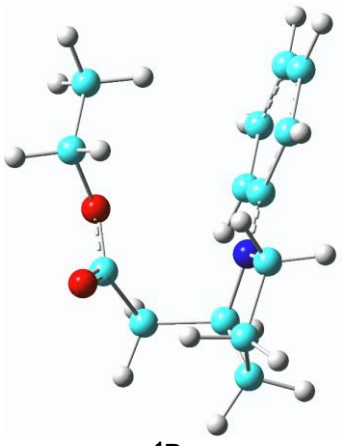
Cartesian Coordinates and 3d Structures of Optimized Structures of $^5P_1$



$^5P_1$

Fe	-0.04091870	-0.30199379	0.70453985	N	-0.92256061	-1.84333129	-0.50353169
N	-1.74637563	0.86822503	0.12848219	C	-0.33451634	-3.02466247	-0.82531407
C	-2.94971756	0.37590575	-0.26228789	C	-2.24105594	-1.95369570	-0.80882753
C	-1.92074677	2.18740380	0.40083265	C	-1.32989447	-3.94077476	-1.34297627
C	-3.94610086	1.42634910	-0.21896859	C	-2.51296230	-3.27638695	-1.33240787
C	-3.30761057	2.55034223	0.19305156	H	-1.13632932	-4.95157666	-1.67823613
H	-4.98994313	1.31056871	-0.48095581	H	-3.48141524	-3.63418535	-1.65752242
H	-3.72396714	3.53921192	0.33623797	C	-3.18529565	-0.93466131	-0.67868810
N	1.01354241	1.57301919	0.71683074	C	1.02825170	-3.30376116	-0.71814423
C	2.35030635	1.71489856	0.90783887	C	-0.91391318	3.07251064	0.78746495
C	0.44584212	2.78862723	0.92156091	C	3.29700809	0.70081832	0.75748525
C	2.65026962	3.08695502	1.26226247	H	1.34355685	-4.30028515	-1.01471259
C	1.46810248	3.75357279	1.27059998	H	-4.20487025	-1.17791433	-0.96442409
H	3.63808323	3.47942518	1.46709743	H	-1.21640326	4.09877751	0.97630928
H	1.29513209	4.80059820	1.48374532	H	4.33277914	0.97270122	0.94130307
N	1.83909689	-1.14254712	0.08382989	O	-0.23469539	-0.76916439	2.55124671
C	2.03470828	-2.42556951	-0.31399456	C	0.78221498	-0.56156619	3.46044233
C	3.05999981	-0.61428137	0.35511015	H	1.69334496	-1.14895531	3.23219719
C	3.44894215	-2.73659549	-0.28534623	H	1.10856379	0.49567468	3.51290402
C	4.08535324	-1.61261309	0.13135386	H	0.47656288	-0.84563811	4.48433651
H	3.88668680	-3.68847248	-0.55744248				
H	5.14845543	-1.46041711	0.26711638				

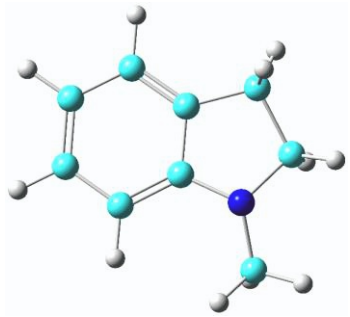
Cartesian Coordinates and 3d Structures of Optimized Structures of $^1\mathbf{P_2}$



$^1\mathbf{P_2}$

C	0.25661400	-1.07038300	-2.69585100
H	1.08145900	-0.62386500	-2.13889100
C	-0.48508200	0.01830000	-3.43041800
O	0.02079800	0.98715700	-3.94332800
H	-0.42436800	-1.53553400	-1.98238300
C	-1.73467000	-2.92829800	-6.38580500
C	-1.82150900	-3.40036200	-4.02759200
C	-2.95922600	-3.56277300	-6.54787500
H	-1.24563700	-2.50406100	-7.25414600
C	-3.04023000	-4.03408900	-4.20754200
H	-1.41296300	-3.32929600	-3.02751600
C	-3.62760200	-4.12244200	-5.46583000
H	-3.39368900	-3.61852900	-7.54131500
H	-3.54159800	-4.45924000	-3.34347700
H	-4.58468100	-4.61447600	-5.59823200
C	-1.13727800	-2.82771300	-5.11584600
N	0.06411300	-2.17684800	-4.95175100
C	0.80492900	-2.13865100	-3.69481200
C	0.83821800	-1.67711900	-6.08173100
C	2.24540000	-1.88193000	-4.15948700
H	0.74822900	-3.11199700	-3.19828800
C	2.08310000	-1.06654900	-5.44142800
H	0.26939500	-0.93606700	-6.65244000
H	1.09813900	-2.49745000	-6.76603500
H	2.84783900	-1.38086800	-3.40044400
H	2.72020400	-2.84196400	-4.38304400
H	1.90307700	-0.01778800	-5.20148300
H	2.95295200	-1.12611100	-6.09720000
O	-1.78824300	-0.25182000	-3.48739000
C	-2.58828100	0.54349100	-4.37637700
H	-2.71070800	1.53948900	-3.94354600
H	-2.05555100	0.64906200	-5.32388300
C	-3.90765200	-0.17067900	-4.55355500
H	-4.55462700	0.40857200	-5.21694600
H	-3.75352700	-1.15927900	-4.99160200
H	-4.41830900	-0.29093800	-3.59545400

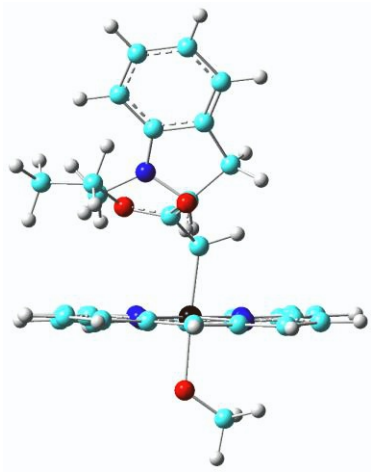
Cartesian Coordinates and 3d Structures of Optimized Structures of $^{1}R_{2}$-1[a]



$^{1}R_{2}$-1[a]

C	1.92582900	-7.07118300	-1.27777100	H	3.75830900	-5.92108800	-1.22559600
C	0.52677500	-7.03343000	-1.19172700	H	1.35604700	-9.48663200	0.02297100
C	-0.13309400	-5.83263300	-1.03461500	H	1.32092100	-10.20809000	-1.60195500
C	0.60888400	-4.64944400	-0.94654000	N	2.37134200	-8.37786100	-1.46499200
C	1.99526400	-4.69471300	-1.01982100	C	3.70186100	-8.74486800	-1.04262700
C	2.67478700	-5.90398500	-1.18079200	H	4.44819600	-8.14967900	-1.57221800
C	1.27852500	-9.26434500	-1.05445600	H	3.85128200	-8.61000700	0.04067400
C	0.01124600	-8.44440100	-1.33837200	H	3.88044300	-9.79300600	-1.28871100
H	-1.21691500	-5.80534900	-0.96740500	H	-0.34034900	-8.61698800	-2.36192200
H	0.10276700	-3.69935200	-0.81392100	H	-0.80573500	-8.68895300	-0.65768400
H	2.56629000	-3.77492100	-0.94045600				

Cartesian Coordinates and 3d Structures of Optimized Structures of $^{CSS}TS_{cis}$-1[a]

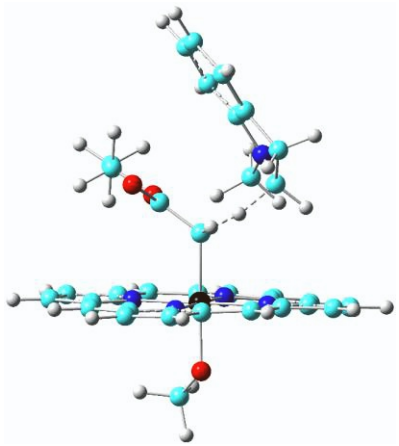


$^{CSS}TS_{cis}$-1[a]

Fe	-0.11772100	-0.27872900	-0.58361600	H	-4.34034600	-1.36949700	-1.64706900
N	-1.81352500	0.75800900	-0.94431600	H	-1.25771900	4.06443900	-0.51648100
C	-3.02097700	0.24601000	-1.28647100	H	4.14057300	0.82359600	0.33266300
C	-1.97346200	2.10464800	-0.87527400	O	-0.70129500	-0.35843800	1.31175400
C	-3.98531700	1.30836800	-1.47064700	C	0.22746100	-0.41218200	2.31173900
C	-3.33356600	2.46657500	-1.21041900	H	0.98337500	0.40368800	2.27494600
H	-5.02271400	1.16623400	-1.74537900	H	-0.24789000	-0.33289100	3.31242100
H	-3.72174900	3.47668100	-1.23232300	H	0.82024500	-1.35580100	2.33733300
N	0.86779200	1.45300500	-0.22619900	C	0.49673100	-0.04168800	-2.40982000
C	2.19251900	1.59923500	0.02914200	H	1.54857800	0.24071500	-2.50127300
C	0.33557100	2.70253100	-0.23735100	C	-0.29693100	0.67857600	-3.42489300
C	2.51766400	2.99896900	0.20181700	O	-0.10315100	1.83948300	-3.71988400
C	1.36094100	3.68379700	0.04426700	H	0.59057800	-1.31066600	-2.84847600
H	3.50453600	3.38710800	0.41847500	N	0.19266400	-2.57878400	-4.56103500
H	1.19888800	4.75235000	0.10072200	C	1.09348800	-2.22057600	-3.55191700
N	1.57648600	-1.31419400	-0.20958100	C	0.46779200	-1.82326100	-5.69751900
C	1.70611300	-2.66385300	-0.13626000	C	2.32725100	-1.69447500	-4.26461400
C	2.80798800	-0.79375900	0.03618100	H	1.22102700	-2.95958500	-2.76283500
C	3.07710500	-3.02005500	0.15734400	C	1.73328700	-1.23445200	-5.57184000
C	3.76400100	-1.85666700	0.25506800	H	2.85529900	-0.92582000	-3.70255000
H	3.44693100	-4.03063300	0.27413600	H	3.02940900	-2.52089800	-4.43211500
H	4.81480500	-1.71241100	0.47103000	O	-1.32200900	-0.03127200	-3.91881000
N	-1.11020000	-2.00594800	-0.87723400	C	-2.22377300	0.67729300	-4.76973900
C	-0.63172900	-3.25966800	-0.66550800	H	-2.55747300	1.57878700	-4.25123100
C	-2.42886500	-2.14577600	-1.17602200	H	-1.69356600	0.97819100	-5.67836700
C	-1.68229000	-4.23721700	-0.85117200	C	-3.38637900	-0.23805600	-5.07545000
C	-2.79907500	-3.54384200	-1.17534500	H	-4.11735500	0.28526200	-5.69743800
H	-1.56270600	-5.30671200	-0.73510500	H	-3.05831900	-1.13107200	-5.61161600
H	-3.79153700	-3.92471300	-1.37933300	H	-3.87493300	-0.55003500	-4.15002400
C	-3.31529700	-1.10507700	-1.40317000	C	-1.09879300	-3.14572300	-4.25228900
C	0.67785800	-3.57769000	-0.32654400	H	-0.97885600	-3.89891700	-3.47340100
C	-0.98286000	3.01403100	-0.53329400	H	-1.79027400	-2.38258800	-3.88794300
C	3.10745000	0.55899000	0.12710200	H	-1.50971800	-3.62891600	-5.14078000
H	0.91761800	-4.63002800	-0.20532100	C	-0.31249400	-1.61814900	-6.82627300

C	2.22726500	-0.42311200	-6.57063200	H	1.81653400	0.45141500	-8.49694400
C	0.19976600	-0.79594400	-7.83079300	H	3.20172200	0.04413500	-6.46883300
C	1.44850700	-0.19725900	-7.70959700	H	-1.29521500	-2.06434400	-6.92606800
H	-0.39853400	-0.61172200	-8.71745400				

Cartesian Coordinates and 3d Structures of Optimized Structures of $^{OSS}TS_{1\text{-}cis}\text{-}1^{a}$



$^{OSS}TS_{1\text{-}cis}\text{-}1^{a}$

Fe	0.05614400	-0.25422100	-0.51547500	H	4.30592300	0.98390300	0.31052700
N	-1.68504900	0.71436200	-0.83582600	O	-0.25924300	-0.34353700	1.36658900
C	-2.89330900	0.16814500	-1.12632600	C	-1.50751700	-0.54156600	1.91078600
C	-1.86907200	2.05950500	-0.79897400	H	-1.99187700	-1.47488000	1.56480000
C	-3.88340400	1.20417200	-1.29231200	H	-1.44236700	-0.61031300	3.01157200
C	-3.24643100	2.38178400	-1.08878500	H	-2.22009800	0.27544200	1.68649800
H	-4.92545700	1.03363400	-1.52872000	C	0.61410600	-0.03364800	-2.40853800
H	-3.65428500	3.38317000	-1.12582900	H	1.60368300	0.42769300	-2.44362000
N	1.00634300	1.50072000	-0.20779200	C	-0.23803700	0.62962600	-3.38743000
C	2.32714800	1.68966900	0.03687700	O	-0.05275600	1.73459100	-3.87133900
C	0.43815500	2.73271000	-0.24630400	H	0.90728900	-1.26385300	-2.90018500
C	2.61072300	3.09950300	0.17823500	N	0.49900800	-2.67810900	-4.60792100
C	1.43551700	3.74757800	0.00637700	C	1.46522700	-2.22590000	-3.65894400
H	3.58734900	3.51897700	0.38018800	C	0.55208100	-1.84246200	-5.71127800
H	1.24280400	4.81171700	0.03631100	C	2.53254800	-1.51339400	-4.47605500
N	1.79851500	-1.22350200	-0.20693600	H	1.77962800	-2.98648000	-2.94606300
C	1.97184500	-2.56884400	-0.17954100	C	1.74175900	-1.09689200	-5.69323200
C	3.01915000	-0.67132100	0.01456000	H	2.98079400	-0.67765800	-3.93509100
C	3.36088200	-2.88793000	0.05927200	H	3.34353900	-2.19856200	-4.76059500
C	4.01350500	-1.70720300	0.17265400	O	-1.32700900	-0.10142800	-3.73289600
H	3.76461800	-3.88928400	0.13113700	C	-2.24445000	0.51878000	-4.62398500
H	5.06514300	-1.53393000	0.35894500	H	-2.60133100	1.45383300	-4.18149800
N	-0.89814900	-2.00216700	-0.79735900	H	-1.73314400	0.76086800	-5.55986400
C	-0.35593600	-3.23848500	-0.63781800	C	-3.38554600	-0.44663200	-4.85377900
C	-2.21955600	-2.19147000	-1.05749400	H	-4.11813100	-0.00865700	-5.53776300
C	-1.37028300	-4.25109600	-0.80759400	H	-3.02241200	-1.37904900	-5.29185300
C	-2.52701600	-3.60034400	-1.07625300	H	-3.88497500	-0.68212200	-3.91169800
H	-1.19976800	-5.31670400	-0.72829000	C	-0.73620000	-3.26464800	-4.14576000
H	-3.50787000	-4.01862600	-1.25942100	H	-1.39879000	-2.51103500	-3.70887300
C	-3.15598200	-1.18715700	-1.22977100	H	-1.24314500	-3.76783700	-4.97284500
C	0.97408600	-3.51396900	-0.35824300	H	-0.50971200	-4.01058300	-3.38391700
C	-0.89334800	3.00278400	-0.51727100	C	-0.37560600	-1.68000100	-6.73375500
C	3.27936600	0.68585600	0.12262500	C	2.00567200	-0.18169800	-6.68915300
H	1.25755700	-4.55888400	-0.27953000	C	-0.09281200	-0.75264900	-7.73762100
H	-4.17836200	-1.48529400	-1.43932800	C	1.07826400	-0.00535200	-7.72303900
H	-1.19505200	4.04522000	-0.52560100	H	-0.81524100	-0.60374100	-8.53467400

H	1.26766800	0.72401400	-8.50341900	H	-1.30683200	-2.23523000	-6.74657000
H	2.91672100	0.40911600	-6.66364500				

Computational analysis of regio- and stereocontrol in the engineered enzymes

Density Functional Theory (DFT) Methods

Modeling studies were carried out using DFT-calculated transition states (TS) corresponding to the regio- and stereoselectivity-determining HAA step in the reaction with the model substrate *N*-methyl indoline. Regarding the transition states, we focused on near-attack conformations (NACs) for the reaction of *N*-methyl indoline with the heme-bound carbene, including NACs for carbene insertion into the pro-(*S*) and pro-(*R*) C–H bond at both α- and β-C–H site of the substrate, and two possible orientations of the heme-bound carbene with respect to the substrate (denoted herein as pseudo-*re* and pseudo-*si* conformers). Additional conformational variations include rotation of *N*-methyl-indoline substrate relative to the heme-bound carbene (along the cleaving C–H bond axis) and rotation of the TS relative to the heme along the Fe–C bond. The corresponding TS structures were optimized by DFT calculations using the Gaussian 16 software package.[S1] Geometry optimization and frequency calculations were performed using the ω-B97XD hybrid functional.[S28] The pseudopotential LANL2DZ[S7] was applied to the iron atom and the basis set 6-311G**[S29] was applied to other atoms. An initial broken-symmetry guess and the unrestricted (U)ω-B97XD method were set up to obtain an unrestricted solution and the spin multiplicity was confirmed at the end of the calculation to converge to the open-shell singlet spin state. Conductor-like polarizable continuum model (CPCM) (diethyl ether)[S30] to mimic the dielectric permittivity in the enzyme active site.

Crystal structures of CYP119 (PDB IDs: 1IO7, 1IO8, 1IO9, 4TT5, 4TUV, 4WQJ, 4WPD, 1F4T, 1F4U) were used as initial structures for Rosetta-based homology modeling. After stripping solvent and ligand molecules from the crystal structures, the Rosetta FastRelax mover[S31] and the "beta_nov16_cst" energy function was applied to introduce the mutations and relax the structures to generate unbound **CYP119-168** and **CYP119-235** structures. DFT-generated TS models were then superimposed onto to the heme cofactor of the pre-relaxed **CYP119-168** and **CYP119-235** structures. The binding poses of TS as well as the protein binding pocket conformations were further optimized using FastRelax. All the single bonds were set to be rotatable in the modeling, with appropriate dihedral constraints applied to the carbenoid Fe–C bond and the ester (C=O)–O bond. Since hydrogen atoms were only allowed to form one covalent bond in the Rosetta params file, for the hydrogen atom transferring from the *N*-methyl indoline to the carbenoid, we only kept the breaking α- or β-C–H bond and did not consider the forming C–H bond on the carbenoid carbon. We created a pseudo-bond between the *N*-methyl indoline α- or β-carbon and the carbenoid carbon to hold fixed the two substrates in TS conformation. The rotation of this pseudo-bond was also sampled with appropriate dihedral constraint to mimic the rotation of *N*-methyl indoline relative to the carbenoid along the cleaving α- or β-C–H bond axis. During both of the two FastRelax optimization stages, coordinate constraints were applied to the protein backbone atoms except residues 151-165 to reflect the higher flexibility of the F and G helices connection part. Final energy scores reported in **Table S24-S25** were the total scores subtracting the coordinate constraint scores.

Table S23. DFT-calculated relative thermodynamic quantities (unit: kcal·mol^{-1}) of different regio- and conformational isomers of the transition states for the hydrogen atom abstraction step.

Step	ΔE(0)	ΔU(T)	ΔH(T)	ΔG(T)
α-(*S*)-*si*	0	0	0	0
α-(*S*)-*re*	-0.21	-0.77	-0.77	1.46
β-(*S*)-*si*-indoline conformer 1	6.69	6.86	6.86	6.32

β-(*S*)-*si*-indoline conformer 2	8.34	8.46	8.46	8.75
β-(*S*)-*re*-indoline conformer 1	5.62	5.85	5.85	5.37
β-(*S*)-*re*-indoline conformer 2	5.68	5.85	5.85	5.28

Table S24. Rosetta-calculated energies (unit: REU) of CYP119-168 complexed with hydrogen atom transfer TS modeled using different starting crystal structures. The lowest energy α-regio-isomer is highlighted in bold text while the lowest energy β-regio-isomer is underscored. "α-(*S*)-*si*" refers to the TS for attack of a-pro(*S*)-C–H bond in the N-methyl indoline substrate from the pseudo-*si* face of the heme-bound carbene; "b-(*R*)-*re*" refers to the TS for attack of b-pro(*R*)-C–H bond in the N-methyl indoline substrate from the pseudo-*re* face of the heme-bound carbene, and so on.

PDB ID	unbound	α-(*S*)-*si*	α-(*S*)-*re*	α-(*R*)-*si*	α-(*R*)-*re*		β-(*S*)-*re*	β-(*S*)-*si*	β-(*R*)-*re*	β-(*R*)-*si*
1IO7	-997.47	-938.83	**-940.96**	-913.45	-929.68		-945.81	-953.74	-930.89	-946.10
1IO8	-940.92	-907.55	**-908.75**	-883.42	-890.45		-894.17	-906.28	-892.87	-896.32
1IO9	-967.85	**-908.81**	-903.52	-892.29	-896.48		-919.42	-916.76	-924.11	-929.72
4TT5	-1041.94	-1022.19	**-1034.60**	-1024.55	-1025.95		-1041.06	-1034.28	-1034.34	-1036.25
4TUV	-992.90	-982.22	**-989.42**	-981.35	-980.87		-988.00	-985.36	-989.04	-988.20
4WQJ	-1017.20	**-1015.90**	-1013.38	-1014.89	-1009.01		-1010.34	-1017.41	-1012.01	-1019.47
4WPD	-1063.24	-1045.55	-1049.09	**-1059.95**	-1048.24		-1048.70	-1061.59	-1048.59	-1062.14
1F4T	-1031.45	-1010.29	-1012.76	**-1024.16**	-1020.29		-1002.13	-1019.57	-1011.37	-1032.79
1F4U	-826.52	**-789.63**	-781.27	-786.02	-780.72		-777.10	-796.80	-777.94	-794.27

Table S25. Rosetta-calculated energies (unit: REU) of CYP119-235 complexed with C–H insertion TS modeled using different starting crystal structures. The lowest energy β-regio-isomer is highlighted in bold text while the lowest energy of α- regio-isomer is underscored. "α-(*S*)-*si"* refers to the TS for attack of a-pro(*S*)-C–H bond in the N-methyl indoline substrate from the pseudo-*si* face of the heme-bound carbene; "b-(*R*)-*re"* refers to the TS for attack of b-pro(*R*)-C–H bond in the N-methyl indoline substrate from the pseudo-*re* face of the heme-bound carbene, and so on.

PDB ID	unbound	α-(*S*)-*si*	α-(*S*)-*re*	α-(*R*)-*si*	α-(*R*)-*re*		β-(*S*)-*re*	β-(*S*)-*si*	β-(*R*)-*re*	β-(*R*)-*si*
1IO7	-989.27	-954.51	-932.34	-944.43	-931.12		-944.34	**-946.42**	-935.20	-934.37
1IO8	-932.91	-906.41	-883.59	-896.98	-885.56		-894.84	-906.63	-908.96	**-915.77**
1IO9	-952.54	-917.61	-913.83	-901.79	-894.13		-906.83	-909.25	-923.50	**-930.51**
4TT5	-1033.93	-1028.27	-1025.32	-1014.58	-1022.24		**-1031.82**	-1025.78	-1024.73	-1030.95
4TUV	-983.36	-971.51	-973.87	-971.85	-978.08		**-980.74**	-978.35	-978.44	-980.38
4WQJ	-1008.9	-1006.80	-1002.36	-1006.95	-1004.24		-1008.26	-1011.33	-1002.06	**-1011.79**
4WPD	-1047.77	-1028.82	-1034.27	-1042.75	-1027.43		-1034.80	**-1045.83**	-1037.11	-1045.40
1F4T	-1012.33	-997.96	-992.63	-1004.60	-985.03		-987.36	-1000.66	-985.71	**-1001.69**
1F4U	-817.33	-778.84	-785.36	-795.62	-773.49		-791.43	-786.61	-788.34	**-800.11**

Figure S8. DFT-optimized transition states for the hydrogen atom abstraction step. Different regio- and conformational isomers are shown.

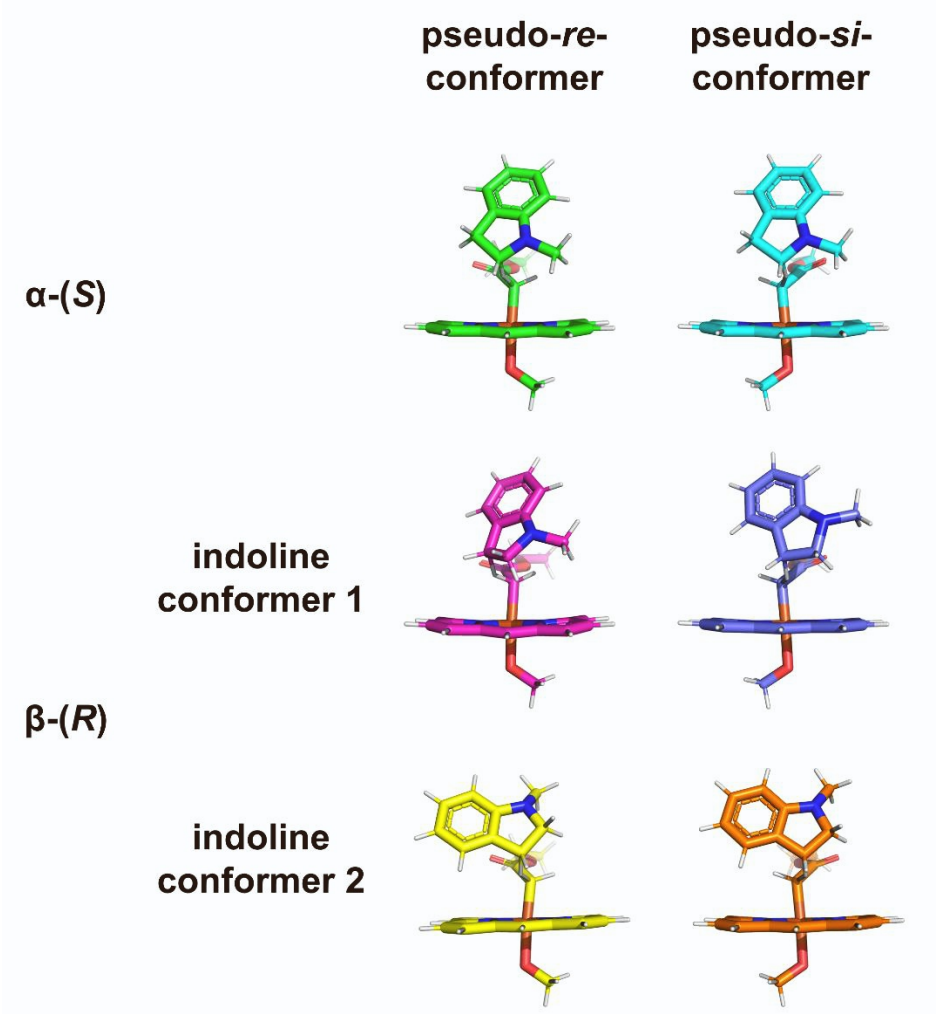


Figure S9. Crystal structures of CYP119 in (**A**) ligand-free, (**B-C**) 4-phenylimidazole-bound, and (**D**) imidazole-bound form. In the ligand-free enzyme, Phe153 sidechain points toward the heme pocket. In the 4-phenylimidazole bound structures, the protein adopts two alternative conformations (**B** and **C**), where the F and G helices undergo extensive secondary and tertiary structure rearrangements and Phe153 points away from the heme center to accommodate the ligand. In the imidazole-bound form, Phe153 is rotated away from the heme (as in **B-C**) but Leu155 sidechain points toward the heme cofactor, filling the active site space. these crystal structures denoted a high degree of conformational flexibility of the F and G helices upon ligand binding, as well as a 'gating' role of Phe153 side chain, which rotates away from the heme pocket to accommodate binding of the imidazole ligand to the heme center.

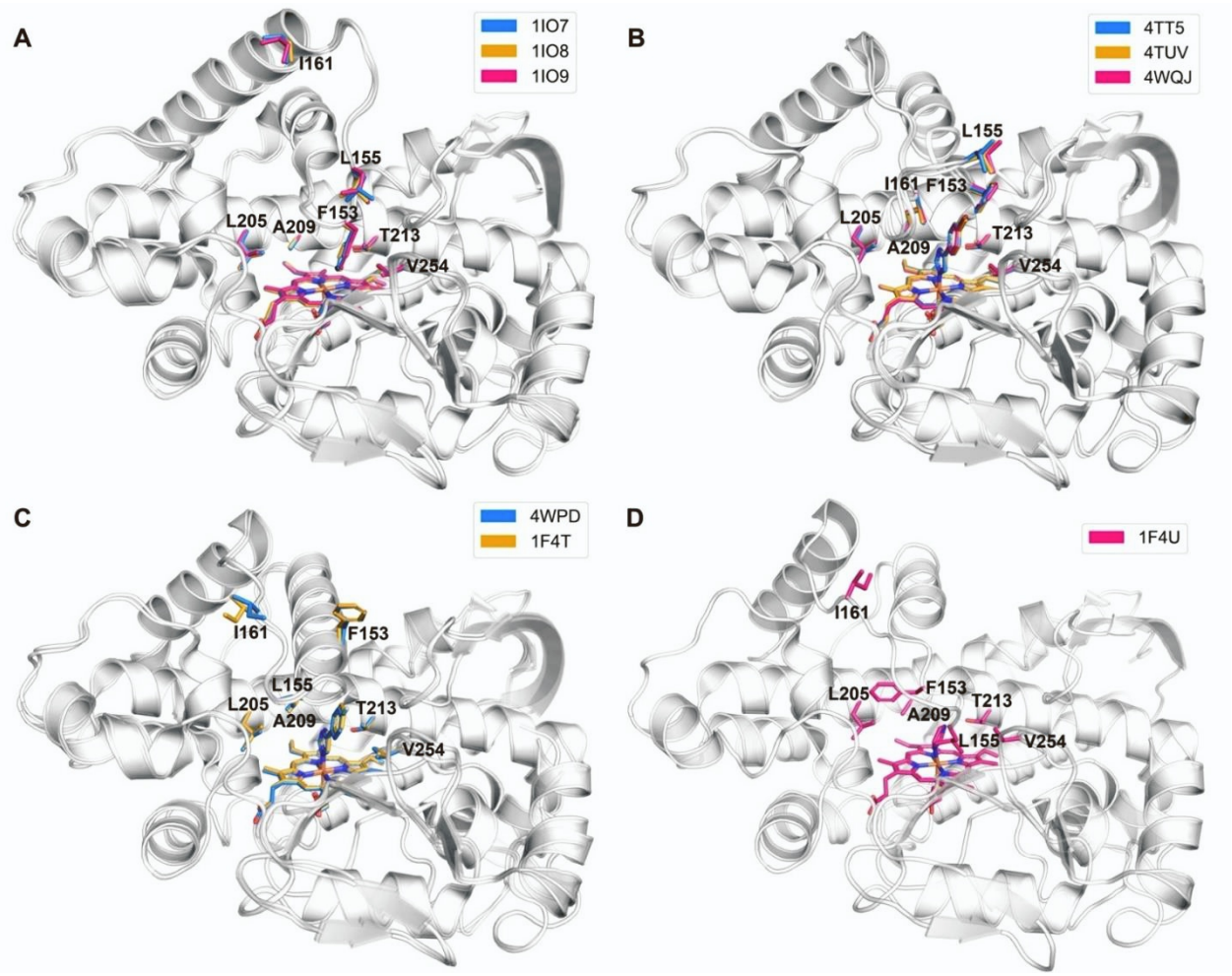
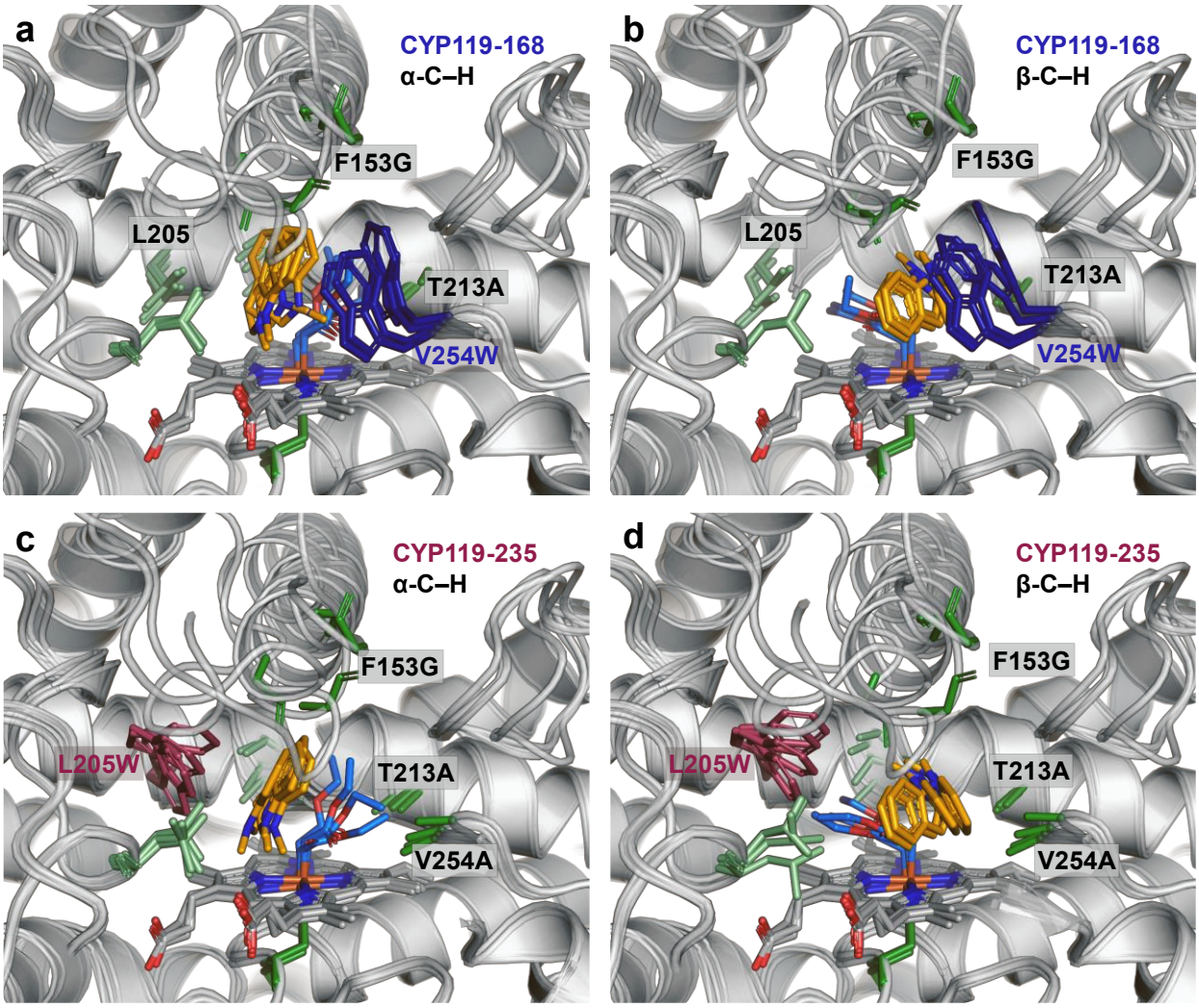


Figure S10. Analysis of enzyme regioselectivity via molecular modeling. (**a-b**) Active site views of representative Rosetta models for **CYP119-168** complexed with (**a**) α-pro(*S*)-C–H TS and (**b**) β-pro(*R*)-C–H TS. (**c-d**) Active site views of representative Rosetta models for **CYP119-235** complexed with (**c**) α-pro(*S*)-C-H TS and (**d**) β-pro(*R*)-C–H TS. The heme cofactor is shown as line model, while mutated active site residues (green), N-methyl indoline (orange) and heme-bound carbene (blue) are shown as stick models.



Data S7. Cartesian Coordinates of DFT-Optimized TS Structures

Cartesian Coordinates of **α-(*S*)-*si***

Fe	-1.344095233999	-0.389809825985	0.112214072190
N	-1.731824590858	0.458446725811	-1.683058339520
N	-1.870008233836	1.322273426023	1.039936326596
N	-0.867924560368	-2.116656387557	-0.815457861965
N	-1.004847860022	-1.249864714493	1.895722706378
C	-1.525012175311	-0.092474788077	-2.907454019588
C	-1.818115439363	1.576034224696	2.370555236765
C	-2.222343012179	1.705314463257	-1.897484336406
C	-2.351604510326	2.447991797049	0.452626873154
C	-0.815272438049	-2.333948940007	-2.155799134023
C	-1.050616323694	-0.646717856362	3.113154884939
C	-0.615394720853	-3.312001162058	-0.221989940986
C	-0.699025890253	-2.553775566691	2.124965045612
C	-1.894036240443	0.845341432867	-3.945784607799
C	-2.276370072262	2.920439095947	2.645351174244
C	-2.337867504343	1.959238381383	-3.317340550182
C	-2.616954106983	3.461081201793	1.451848290749
C	-0.507708484788	-3.719874763939	-2.426414330991
C	-0.733153368465	-1.598013042291	4.155763193016
C	-0.397705364960	-4.332456183756	-1.222881464135
C	-0.524624353359	-2.786839964226	3.542296322151
H	-1.827294338656	0.652534867160	-5.007478478973
H	-2.331655761872	3.368069184703	3.627882151915
H	-2.707589500075	2.876702104064	-3.753865403001
H	-3.008189139378	4.447620351550	1.245452779146
H	-0.410054065404	-4.151973465537	-3.412654102697
H	-0.697303176923	-1.370086113893	5.211965779646
H	-0.183256411207	-5.371159907717	-1.012972542084
H	-0.276139363957	-3.740221793079	3.987482356838
C	-1.078025111032	-1.385859990419	-3.136706159932
C	-1.410827647680	0.673051382358	3.342512091551
C	-0.538991243920	-3.526881026893	1.147766752557
C	-2.535368496718	2.628512027041	-0.910763602696
H	-0.978360459503	-1.701133472594	-4.170028880256
H	-1.411778583579	1.015915450229	4.371754668817
H	-0.312196591053	-4.533959139446	1.481541396302
H	-2.923575060184	3.588228270912	-1.235710166462
O	-3.251650602936	-0.911827860723	0.260148429644
C	-3.836372935933	-1.707370168855	-0.689344061464
H	-3.424761802053	-2.741220450746	-0.739344761365
H	-3.749073665367	-1.311931786474	-1.725265906243
H	-4.924105552628	-1.828723818369	-0.505331693979
C	0.449699224270	0.241078258696	-0.151292628797
H	0.929277562291	-0.031275321970	-1.097572772540
C	0.866467840725	1.599806253834	0.266205478052
O	1.178201712504	1.982029654925	1.372584349756
O	0.770001486944	2.434630222620	-0.784603630866
C	0.806608750830	3.836959924156	-0.492352197462
H	1.788306398266	4.093322950078	-0.085295736703
H	0.051753753719	4.054706760488	0.268226935002
C	0.522502614140	4.570984816513	-1.782399951723
H	1.278487469510	4.338511460009	-2.536081314620

H	0.529967206505	5.649313183118	-1.605544839700
H	-0.457010547359	4.285444297819	-2.171625440048
H	1.231739256535	-0.608342235599	0.646652574514
C	2.230714376261	-1.232590725332	0.983933729569
H	1.750379822738	-2.103226019384	1.428902313407
C	3.061954787206	-0.330131215615	1.883518131540
H	3.596187184808	-0.938005090330	2.623260900584
H	2.450977453774	0.407524124185	2.400165867895
C	4.022953294402	0.272336149427	0.889930670728
C	3.972806769850	-0.495697508797	-0.281682534848
N	3.054330591373	-1.527131943642	-0.125283221288
C	2.616619163900	-2.362141560811	-1.217713170704
H	2.029427451657	-1.808396331436	-1.961086016878
H	1.990355518556	-3.162095407191	-0.824833721210
H	3.483514099305	-2.806827462782	-1.711685567009
C	4.865263735329	1.360256419654	0.969630275801
H	4.897597841706	1.964678418182	1.869653567948
C	4.755672111861	-0.187056135077	-1.385379301745
H	4.701708634995	-0.769277481007	-2.297540446308
C	5.665627183517	1.683893038894	-0.132336626586
H	6.326509375105	2.541652156427	-0.086412384450
C	5.607977805803	0.915120269726	-1.288375821187
H	6.225102502902	1.180577186177	-2.13978285533

Cartesian Coordinates of **α-(*S*)-*re***

Fe	1.385148246701	-0.245948813362	0.109970617349
N	2.033990968921	1.324604094483	-0.986546387668
N	1.357441713674	0.913326344553	1.765354730113
N	1.445705374388	-1.411781336066	-1.533940851663
N	0.791709158077	-1.820620505707	1.209513128548
C	2.215974629298	1.358843953976	-2.331360672114
C	0.962748387521	0.549410904510	3.010972595621
C	2.339426937647	2.563453838710	-0.521663686445
C	1.763948490802	2.206277842592	1.853218851861
C	1.705719639559	-1.010523409517	-2.806205900752
C	0.481102612584	-1.821319870125	2.533553806112
C	1.247429991914	-2.754254099170	-1.585306204610
C	0.699845576283	-3.110928839067	0.794914304093
C	2.658945392533	2.673572617959	-2.741781425817
C	1.104530353817	1.658778869153	3.927739481696
C	2.746323798350	3.419284583265	-1.614811919518
C	1.611573478325	2.688559964661	3.208318073977
C	1.652248501783	-2.141189278232	-3.706685348470
C	0.160336863666	-3.161500680158	2.973065142836
C	1.379549301610	-3.227852069909	-2.945647383492
C	0.306366634505	-3.965966170834	1.894290775066
H	2.878590951226	2.964044695832	-3.759746392717
H	0.855888825193	1.630097989377	4.979472735827
H	3.047139210221	4.452729182223	-1.513142842738
H	1.863062917252	3.685324330657	3.543256218719
H	1.820522615719	-2.092440675958	-4.773570560965
H	-0.122836294576	-3.431055997433	3.981011000090
H	1.270576336710	-4.257452067728	-3.256908981126
H	0.164806100430	-5.035713192376	1.827016300576
C	2.037584608997	0.281753726461	-3.190256565878
C	0.530685125522	-0.719254279008	3.373034857338
C	0.924090569433	-3.559011857632	-0.500774797579
C	2.238312024221	2.974346326077	0.799185216412
H	2.219688886375	0.453652661980	-4.245906241777
H	0.252445060822	-0.872625735143	4.410371379326
H	0.806839750267	-4.621398770950	-0.686718468095
H	2.526147250530	3.996533235442	1.022316453002
O	3.260482287252	-0.677516870764	0.586535194910
C	4.195517735215	-0.895847282900	-0.390203871771
H	4.305509865213	-0.053043287205	-1.108035257891
H	3.996244776481	-1.789856045022	-1.023862425754
H	5.205686414596	-1.055327637070	0.040987392159
C	-0.359707020060	0.345215013775	-0.392770965680
H	-0.484429532208	0.644226818494	-1.438338660456
C	-1.204423829104	1.168074822624	0.497386618793
O	-1.916180089662	0.791445427627	1.401444316887
O	-1.018800329179	2.470513237607	0.215896544309
C	-1.479866422175	3.404662326350	1.200733161950
H	-2.568835502145	3.341699587386	1.273460423108
H	-1.053225340357	3.129140345414	2.169007486530
C	-1.022791295521	4.777906599874	0.766590502421
H	-1.442081426274	5.037537152449	-0.208294460354
H	-1.349674496479	5.524833289990	1.494185286836
H	0.066535541431	4.808403385417	0.697932027598

H	-1.103328177978	-0.861378727759	-0.493949981985
C	-1.934980221111	-1.616104743527	-0.954554817340
H	-1.345241278046	-2.530185805323	-1.006049220279
C	-2.467037608522	-1.047085364443	-2.262218135755
H	-2.773872801254	-1.867478213858	-2.921959340995
H	-1.728766446685	-0.451037793909	-2.798454912560
C	-3.673940019413	-0.281645177574	-1.777883053633
C	-3.975757597789	-0.731702966377	-0.483864170977
N	-3.049793548252	-1.689818019046	-0.092928483265
C	-2.912676213143	-2.174396261786	1.262309825773
H	-2.473235309481	-1.409775304393	1.909962139033
H	-2.259052956669	-3.047849222777	1.257629467171
H	-3.890348886046	-2.474572741933	1.645445330851
C	-4.461980224706	0.673250778282	-2.385128024968
H	-4.227172683628	1.025779182423	-3.383998956974
C	-5.063916942020	-0.229703158353	0.216778878606
H	-5.281705010962	-0.554947813217	1.226641546069
C	-5.562159836909	1.189138000192	-1.691727034108
H	-6.184596691234	1.947778556669	-2.151694265101
C	-5.853045166533	0.735246472290	-0.409955479381
H	-6.702466208984	1.148092422676	0.123481348436

Cartesian Coordinates of **β-(*R*)-*si*-indoline conformer 1**

Fe	-1.290819943681	-0.399474418259	0.005213328793
N	-1.553549899189	0.895067155302	-1.521262639983
N	-1.911833589328	0.997721503106	1.308178535859
N	-0.711134501749	-1.818512427336	-1.309779920941
N	-1.058831801818	-1.708753133067	1.507823568178
C	-1.222773809227	0.698051007803	-2.823855173962
C	-1.921438099514	0.899590096896	2.659971152883
C	-2.083229772162	2.141466757089	-1.440496431656
C	-2.382241781372	2.234664207493	1.008623493524
C	-0.508252459258	-1.657677098450	-2.643207713676
C	-1.173916485365	-1.445785245880	2.837172220125
C	-0.533509498306	-3.141076011895	-1.050820533215
C	-0.802117063647	-3.040159741243	1.399604469887
C	-1.558254049990	1.868093329127	-3.606809175642
C	-2.418596857735	2.124525002552	3.250065761967
C	-2.108273945453	2.759271713334	-2.748891130785
C	-2.715863374723	2.951939101601	2.221456543284
C	-0.159780942489	-2.921461703247	-3.253004765174
C	-0.941580937604	-2.648547414082	3.606668348174
C	-0.191727829002	-3.847745914529	-2.264728105198
C	-0.726431319385	-3.642752151130	2.712148654307
H	-1.401240284479	1.967557287017	-4.671851359611
H	-2.527462754330	2.302673381974	4.310854339048
H	-2.492728649294	3.747590991321	-2.960188496624
H	-3.116361237864	3.955546299040	2.258151319858
H	0.055884299099	-3.065205006079	-4.302694194900
H	-0.968229556683	-2.708731774331	4.685817038128
H	0.000242287153	-4.909687460426	-2.331290105320
H	-0.531504423905	-4.688884714417	2.903404045493
C	-0.699248218258	-0.475404736718	-3.348385788600
C	-1.546063961171	-0.224621393723	3.383272070084
C	-0.598704687781	-3.724345080360	0.208698401487
C	-2.488449947738	2.766611355023	-0.269879714967
H	-0.491695006335	-0.493991726952	-4.413183734425
H	-1.603355109623	-0.162011420353	4.464922422892
H	-0.416704856079	-4.791953963574	0.272859241196
H	-2.884687677795	3.773041628495	-0.356871916298
O	-3.186547777842	-0.960034099623	-0.151196462808
C	-3.673909717626	-1.475313223054	-1.322798224341
H	-3.255674207581	-2.470527195861	-1.597929417926
H	-3.491255308681	-0.829783214826	-2.209973835221
H	-4.773636785185	-1.617899702708	-1.277033305234
C	0.571878596749	0.327937383052	0.075932739345
H	1.188575914341	0.070827500598	-0.789762975592
C	0.835813158026	1.706058003099	0.542069662945
O	0.959420313768	2.071395495849	1.690601227639
O	0.880942798397	2.554233498321	-0.497625715690
C	0.878918710468	3.950655635481	-0.177261552939
H	1.767649726915	4.182840489974	0.415374710477
H	-0.004127799817	4.166837027304	0.430126541233
C	0.857795620464	4.712495935510	-1.481920540652
H	1.742510067940	4.481167793704	-2.079844001052
H	0.845713846271	5.787004966429	-1.283373830052
H	-0.030948836918	4.450800749166	-2.059681065094

H	1.036783857375	-0.378466030349	1.023411038303
C	2.364921631336	-2.072953165087	0.543686423496
H	1.448088184130	-2.370490121845	0.038514270012
H	2.693414200553	-2.911755450179	1.170030156644
C	2.173858407447	-0.833201826098	1.391645334866
H	1.861778114792	-0.975558349293	2.425295672856
C	3.351826141926	0.000947378155	1.102278687381
C	3.969214220739	-0.517054291421	-0.053219365620
N	3.402508573485	-1.715169165923	-0.422747974450
C	3.260062431349	-2.132943607256	-1.800323067837
H	4.197752433843	-1.983064515838	-2.337535444880
H	3.020162795558	-3.196976046657	-1.832663295630
H	2.457022870940	-1.588708986766	-2.318198389728
C	3.831189994704	1.151910523773	1.709684220264
H	3.341173802992	1.542533458029	2.592218731943
C	5.026598017900	0.167816525326	-0.651983418167
H	5.493677253981	-0.198884716023	-1.558027537096
C	4.911576308155	1.816427693456	1.136207117277
H	5.298982754468	2.718974802621	1.593465133898
C	5.483073381144	1.330955531502	-0.040525950150
H	6.311926049573	1.867958036336	-0.489526352607

Cartesian Coordinates of **β-(*R*)-*si*-indoline conformer 2**

Fe	1.069907580691	0.764292361470	0.453691008617
N	2.584274047462	-0.183424217190	-0.499990594712
N	1.099890360219	-0.641561082761	1.898501652996
N	1.028750911460	2.163949756699	-1.002870286388
N	-0.416884032997	1.727477800751	1.413659624950
C	3.118550923515	0.145046201456	-1.704286875298
C	0.286115966154	-0.712157236960	2.981118007572
C	3.233232395412	-1.300892940374	-0.081825109766
C	1.943677046063	-1.702062523188	1.984755622649
C	1.774910177898	2.173520773610	-2.139077617089
C	-0.989093466262	1.364886221163	2.593078483370
C	0.254340028394	3.280102539119	-1.041995718061
C	-0.969423167568	2.921640603939	1.072513587841
C	4.158177996806	-0.795152048408	-2.064256377712
C	0.610390768603	-1.872705528084	3.781683188290
C	4.234881674051	-1.689793052254	-1.050872322346
C	1.646170036746	-2.487400302233	3.163639054447
C	1.456223208655	3.337266529465	-2.937188765616
C	-1.944112147871	2.364544031923	3.016497338090
C	0.515695804886	4.029514188043	-2.251638112127
C	-1.930329020211	3.335228568820	2.071639851528
H	4.743463130637	-0.754750607546	-2.972442090894
H	0.107193909487	-2.155827170208	4.695712345463
H	4.893218515565	-2.541964767674	-0.953150261017
H	2.173649533734	-3.383073730503	3.461263440914
H	1.910927120036	3.580604632120	-3.887507646527
H	-2.534285417603	2.313451959430	3.920877777975
H	0.033064975446	4.958478319165	-2.521958163209
H	-2.506578836030	4.249325700116	2.036423348739
C	2.734447178812	1.230034539760	-2.481854689080
C	-0.691716411987	0.216308443846	3.311218446572
C	-0.668814396674	3.649189584712	-0.072842124745
C	2.950094241340	-2.006296115201	1.077730750679
H	3.250291080622	1.369876421206	-3.426148890328
H	-1.249672862957	0.045695047956	4.225900243322
H	-1.204443317875	4.580390127213	-0.224574339782
H	3.549217250233	-2.888090504184	1.280380317090
O	2.373977332786	1.779118676278	1.543724960132
C	3.297578757023	2.586478558140	0.934952610340
H	2.861579279968	3.462977892620	0.403087732474
H	3.918959428629	2.067350066854	0.172211845098
H	4.015087734632	3.009778474580	1.668587128549
C	-0.087649264613	-0.364534739744	-0.714893208700
H	0.234493254895	-0.330013929952	-1.756039380196
C	-0.511144007331	-1.714205511322	-0.315671371004
O	-1.398122233485	-2.018731027791	0.449363187220
O	0.301455244655	-2.634130257663	-0.883718152266
C	0.271959327329	-3.945875109843	-0.309661640614
H	-0.705891859141	-4.399520271085	-0.493999191409
H	0.408786213319	-3.857553296150	0.771625270850
C	1.389906423658	-4.738794821425	-0.946714697054
H	1.245144533363	-4.818101037585	-2.026884708663
H	1.417638170786	-5.747149866393	-0.526514548313
H	2.349959485715	-4.253293109819	-0.758991105310

H	-1.106395011374	0.381653749863	-0.692384516612
C	-2.258134411672	-0.048778261264	-2.742634340164
H	-1.296658977176	-0.334476677469	-3.168166815158
H	-2.742228858485	0.637162661218	-3.451041363590
C	-2.167562338921	0.634513392438	-1.391528449717
H	-1.983336617319	1.707881087944	-1.390150117446
C	-3.292067721769	0.085352198269	-0.632389294073
C	-3.737737432809	-1.071653135859	-1.294888798102
N	-3.113206563465	-1.213903083714	-2.515447469379
C	-2.758432117818	-2.496015124173	-3.081624108179
H	-3.565345546527	-3.210249500565	-2.915778162908
H	-2.618893068907	-2.397489013658	-4.160110925906
H	-1.835128897342	-2.900293764771	-2.642403603384
C	-3.826862885877	0.461901724250	0.591162573014
H	-3.458440591420	1.346386692992	1.098150601320
C	-4.711508848951	-1.883151299523	-0.712895203468
H	-5.056833979527	-2.787787993287	-1.198621030630
C	-4.814534313206	-0.330148732131	1.161265524012
H	-5.247619515967	-0.056150234486	2.115923534884
C	-5.236628981746	-1.494413054530	0.513004975050
H	-5.998386290394	-2.112705456314	0.976098425601

Cartesian Coordinates of **β-(*R*)-*re*-indoline conformer 1**

Fe	1.319877192596	-0.508485586318	0.280211627599
N	2.264261904094	0.374446939816	-1.270353696458
N	1.578538093693	1.158272805739	1.391744671099
N	1.054835660024	-2.174838417146	-0.832402374430
N	0.385076595209	-1.394833911815	1.828894078841
C	2.433692085803	-0.134592978940	-2.517139630901
C	1.144476095150	1.368537880960	2.660838902623
C	2.852193328567	1.598889960897	-1.269423939404
C	2.261858763414	2.274134738415	1.031094215355
C	1.369366410922	-2.332887480019	-2.145226239755
C	0.111498764512	-0.840552746305	3.040444347776
C	0.498994485458	-3.347088769963	-0.425924191909
C	-0.070842439203	-2.674467517282	1.881687937252
C	3.161663867794	0.802944656466	-3.346045236754
C	1.551558664110	2.677575544561	3.121742943677
C	3.430619888045	1.876654301594	-2.566860623293
C	2.254460463412	3.239134674690	2.108897114344
C	0.995970841116	-3.656465406226	-2.593539981483
C	-0.555482602522	-1.800960784699	3.891735453351
C	0.462784018783	-4.290769681589	-1.522061483399
C	-0.662578932226	-2.944012091447	3.173294413074
H	3.428208582810	0.638586367233	-4.380912886405
H	1.329502019478	3.088967941475	4.096609947670
H	3.960336455602	2.782324124479	-2.827958310126
H	2.728122412339	4.210514285263	2.074804006062
H	1.141014869656	-4.035483725005	-3.595711702591
H	-0.880742502391	-1.609990415658	4.904988853686
H	0.073069630723	-5.297427725023	-1.461219884486
H	-1.098739150687	-3.887746001481	3.470066907377
C	1.996470963589	-1.382864796276	-2.940804326684
C	0.444327721788	0.448859258954	3.428990078689
C	-0.013994477860	-3.593392361355	0.840399439210
C	2.872767954166	2.479575595612	-0.199122848723
H	2.193043192311	-1.652997255618	-3.973112813989
H	0.160146716108	0.752920012941	4.431007157783
H	-0.428752018572	-4.578697839725	1.025495944236
H	3.388356281557	3.423377433705	-0.344274633300
O	3.035795323353	-1.174305863299	1.014520999393
C	4.014538802265	-1.633154748458	0.174003543127
H	4.468282049929	-0.847918852966	-0.472621225231
H	3.673875669613	-2.424615580834	-0.530089490218
H	4.860930477522	-2.078188805926	0.737958324168
C	-0.359549171858	0.225510398525	-0.539424627422
H	-0.397764074863	0.140440863424	-1.626296384793
C	-0.941589662531	1.480554721356	-0.040789171641
O	-1.678680007758	1.623367139871	0.911161199656
O	-0.470147231340	2.514044435640	-0.756439978678
C	-0.650220826644	3.816238882670	-0.183961600973
H	-1.718399390277	4.044897550401	-0.142282378243
H	-0.259613001999	3.804968976642	0.837009587596
C	0.101213629991	4.798067560980	-1.052428941353
H	-0.286300926505	4.790428919670	-2.073886786641
H	-0.006941105150	5.807660113382	-0.648486200594
H	1.162414425597	4.542684997777	-1.078572794271

H	-1.069788157044	-0.707263068764	-0.125523038503
C	-2.905918536511	-1.182843587655	0.983004166008
H	-2.253142780666	-0.764081711215	1.751866114611
H	-3.157303354748	-2.209459468707	1.268339789510
C	-2.238627736657	-1.160740335982	-0.387012499528
H	-1.854628260235	-2.107656951833	-0.763646846309
C	-3.149466654510	-0.388715214129	-1.245690807596
C	-4.181156232769	0.125208376369	-0.442134363182
N	-4.114383960303	-0.368889105773	0.837651199249
C	-4.567341767413	0.395863930072	1.978179979237
H	-5.584003829011	0.755620023828	1.809997696515
H	-4.588036194045	-0.251302477097	2.856956249416
H	-3.906278544612	1.248522881617	2.176901550914
C	-3.085759900935	-0.099814254079	-2.600401339497
H	-2.283828872116	-0.507458387297	-3.206996455233
C	-5.123605840279	0.998407622544	-0.990966690622
H	-5.908139052202	1.432749469552	-0.383433629876
C	-4.040776522059	0.744075644345	-3.157414559750
H	-4.004253071069	0.988230644214	-4.212132430385
C	-5.036719563886	1.292365638715	-2.346788612911
H	-5.766317819249	1.965538062114	-2.784304375053

Cartesian Coordinates of **β-(*R*)-*re*-indoline conformer 2**

Fe	1.163419516272	-0.738975571550	0.011113689941
N	1.571745972495	0.021888167122	-1.815393220450
N	2.309850914745	0.679219331576	0.866200680872
N	0.011848213494	-2.163787381631	-0.840841736196
N	0.755580249609	-1.508172893014	1.828851623730
C	1.046395738471	-0.378880510359	-3.002378780436
C	2.502235653509	0.880760958864	2.192503484136
C	2.409104994485	1.052759107554	-2.092427891019
C	3.051507440282	1.618569323689	0.223378803299
C	-0.279678516304	-2.286853528181	-2.163350111117
C	1.151146486509	-1.007082983268	3.029997520081
C	-0.587156711303	-3.211150352135	-0.216314062722
C	0.052320410640	-2.640314296805	2.099405046647
C	1.569040554305	0.432439111958	-4.080219784253
C	3.395685408871	1.998953592314	2.408324767683
C	2.425256355723	1.316294545485	-3.514938943156
C	3.744312449893	2.453641769081	1.181986193042
C	-1.099564725376	-3.455906907227	-2.388383163288
C	0.663577224352	-1.841944267066	4.106689978520
C	-1.284415684945	-4.036109799257	-1.177574311876
C	-0.012119242925	-2.862361292202	3.527932649014
H	1.311552251560	0.317914366674	-5.123977609514
H	3.708696373088	2.364961065497	3.376298927141
H	3.015152061308	2.084127573773	-3.996215685374
H	4.401832348578	3.273817341676	0.928934543535
H	-1.467909536890	-3.778797643702	-3.352232884544
H	0.838043976444	-1.661278949279	5.158275869873
H	-1.839038099227	-4.932866854229	-0.938718441618
H	-0.513839762277	-3.693447011436	4.003881931762
C	0.168735523490	-1.441604503793	-3.169492259671
C	1.948822569138	0.113253842641	3.209126115627
C	-0.562702881091	-3.449205150090	1.152895330970
C	3.119104442669	1.786089440166	-1.151870918565
H	-0.161079895160	-1.657869322615	-4.180270203605
H	2.188459769227	0.393380051198	4.229531443035
H	-1.090332042161	-4.325366335109	1.514745695695
H	3.753384195111	2.584698381462	-1.522771954632
O	2.796479266875	-1.851988045843	-0.084468920657
C	2.946110879159	-2.773332784491	-1.086845891096
H	2.835388928309	-2.349969757426	-2.109421864346
H	2.222300311813	-3.619119165672	-1.045053289633
H	3.952468749530	-3.240922948910	-1.062760542276
C	-0.425096517258	0.482712710103	-0.036122178107
H	-1.023058430011	0.389270872588	-0.943398816492
C	-0.311601313551	1.879701904861	0.402897452438
O	-0.349224317081	2.315425111377	1.537000721411
O	-0.066922911804	2.666348635091	-0.660309263603
C	0.419766542849	3.985409616409	-0.384554053470
H	-0.357057002704	4.556730716082	0.130636607094
H	1.285488876806	3.906977661051	0.278625555400
C	0.789481140512	4.612264398899	-1.709098779695
H	-0.081954814407	4.676470197237	-2.364927116470
H	1.176156005737	5.621390446720	-1.547004947141
H	1.557263142084	4.017391045417	-2.207729840723

H	-1.094776507702	-0.168187904760	0.808989664556
C	-2.717331469144	0.628776970340	2.235712045343
H	-1.856543965267	1.170478067971	2.631875792781
H	-3.219590224808	0.105042290252	3.059797163372
C	-2.326181306898	-0.365347310682	1.165726305889
H	-2.097644368449	-1.381805684150	1.482148218129
C	-3.288211669194	-0.161392661366	0.078564338453
C	-3.967699125571	1.047528000629	0.327607619324
N	-3.643241186815	1.538024529311	1.565555103752
C	-3.716628879423	2.923754909937	1.956887923736
H	-4.541779607386	3.414848277031	1.440430453117
H	-3.905868335283	2.996211552940	3.030740612905
H	-2.782575362716	3.454246497548	1.730697614464
C	-3.500222372915	-0.881197988972	-1.086839283351
H	-2.955056051838	-1.802421914974	-1.264541379693
C	-4.841904339828	1.564582788136	-0.630975653295
H	-5.360626793200	2.502954311565	-0.476503970112
C	-4.392937529210	-0.381552517028	-2.030476540288
H	-4.573182283060	-0.924721685584	-2.950316548822
C	-5.040068347362	0.833484332926	-1.798197115943
H	-5.721637777396	1.222686647653	-2.547271049519

Analytical Methods

Chiral SFC Methods. Stereoisomer resolution for compounds **3a** and **4a** was performed by Supercritical Fluid Chromatography (SFC) analysis, using a JASCO Analytical and Semi-Preparative SFC instrument equipped with a column oven (35 °C), photodiode array detector, a backpressure regulator (~14.0 MPa), a carbon dioxide pump and a sample injection volume of 3 µL. Daicel Chiralpak IA (0.46 cm ID × 25 cm L) was used for separation of enantiomers. All samples were eluted using an isocratic solvent system with the indicated modifier (see table below) in liquid CO_2 at an elution rate of 4 mL/min and detected at λ = 220 nm. Total run time was 10.2 min. Modifier solvent percentages and retention times (t_R) for chiral SFC analyses of samples **3a** and **4a** are reported in the tables below:

Product	Column	Modifier Solvent	t_R - 1st isomer (min)	t_R - 2nd isomer (min)
3a	IA	5% IPA	1.98	2.46
4a	IA	2% IPA	2.62	3.18

Chiral HPLC Methods. Stereoisomer resolution for compound **3b-3n**, **4b**, **6a**, **7a**, and **9a** was performed by High Performance Liquid Chromatography (HPLC) analysis, using Agilent 1200 series, equipped with column oven (25 °C), UV/PDA detector and a sample injection volume of 5 µL. Daicel Chiralpak IA, IF, IG, or IK column (0.46 cm ID × 15 cm L) was used for separation of enantiomers. All samples were eluted using a constant gradient system with the indicated modifier (see table below) in hexanes at an elution rate of 1 mL/min and detected at 254 nm.

Product	Column	Modifier	t_R - 1st isomer (min)	t_R - 2nd isomer (min)
3c	IA	25 min, 1-3% IPA	8.5	9.3
3d	IK	25 min, 1-3% IPA	6.3	7.0
3f	IA	25 min, 1-3% IPA	7.4	8.6
3g	IK	25 min, 1-3% IPA	15.9	18.6
3h	IA	35min, 0.5-1% IPA	11.6	15.9
3i	IA	25 min, 1-3% IPA	11.6	12.6
3j	IG	25 min, 1-3% IPA	13.3	14.3
3k	IG	25 min, 1-3% IPA	10.7	11.8
3l	IK	35min, 0.5-1% IPA	13.6	15.0
3m	IG	35min, 0.5-1% IPA	12.6	14.8
3n	IG	25 min, 1-3% IPA	11.1	12.7
6a	IA	25 min, 1-3% IPA	6.8	7.8
7a	IB	25 min, 10% IPA	9.2	9.9
9a	IF	25 min, 1-8% IPA	21.9	23.4

Product	Column	Modifier	t_R - 1st isomer (min)	t_R - 2nd isomer (min)	t_R - 3rd isomer (min)	t_R - 4th isomer (min)

3b	IA	25 min 1-3% IPA	8.1	9.2	10.9	11.5
4b	IA	25 min 1-3% IPA	12.7	13.5	15.4	17.4

GC Methods. Gas chromatography (GC) analysis was carried out using a Shimadzu GC-2010 gas chromatograph equipped with a FID detector, and a Cyclosil-B column (30 m x 0.25 mm x 0.25 µm film). The following GC methods were used for TON analysis: 1 µL injection, injector temp.: 200 ºC, detector temp: 300 ºC. Method: column temperature set at 80 ºC for 2 min, then to 245 ºC at 40 ºC/min, then held at 245 ºC for 6.5 min. Total run time was 12.63 min.

Experimental Procedures

General information. All chemicals were purchased from commercial sources and used without further purification. Substrates **1a-1n** were synthesized according to the reported procedure. Unless stated otherwise, all column purification were performed on a Biotage Selekt Flash Chromatography, eluted with ethyl acetate: hexane, 0-15% gradient, detected by UV absorption at 254nm. Thin Layer Chromatography (TLC) was carried out using Merck Millipore TLC silica gel 60 F254 glass plates. 1H, 13C, and 19F NMR spectra were measured on a Bruker AVANCE III™ 500 Spectrometer with Cryo Probe (operating at 500 MHz for 1H and 125 MHz for 13C).

Protein Expression. Wild-type and engineered CYP119 variants were expressed in *E. coli* C41(DE3) cells as follows. After transformation, cells were grown in LB medium (ampicillin, 100 mg/L) at 37°C (200 rpm) until OD_{600} reached 1 to 1.2. Cells were then induced with 0.25 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and 0.3 mM δ-aminolevulinic acid (δ-ALA) (the addition of 1.0 mM $CoCl_2$ was used to form Co-substituted CYP119 variants). After induction, cultures were shaken at 180 rpm and 37°C and harvested after 40 hours by centrifugation at 4,000 rpm at 4°C. The cells were resuspended in 20 mL of Ni-NTA Lysis Buffer (50 mM KPi, 250 mM NaCl, 10 mM histidine, pH 8.0). Resuspended cells were frozen and stored at -80°C until purification. Cell suspensions were thawed at room temperature, lysed by sonication, and clarified by centrifugation (14,000 rpm, 50 min, 4°C). The clarified lysate was transferred to a Ni-NTA column equilibrated with Ni-NTA Lysis Buffer. The resin was washed with 50 mL of Ni-NTA Lysis Buffer and then 50 mL of Ni-NTA Wash Buffer (50 mM KPi, 250 mM NaCl, 20 mM histidine, pH 8.0). Proteins were eluted with Ni-NTA Elution Buffer (50 mM KPi, 250 mM NaCl, 250 mM histidine, pH 7.0). After elution from the Ni-NTA column, the protein was buffer exchanged against 50 mM KPi buffer (pH 7.0) using 30 kDa Centricon filters. The concentration of the CYP119 variants was determined from CO-binding assays (difference spectra) using ε_{406} = 130 $mM^{-1}cm^{-1}$ as the extinction coefficient.

Protein Engineering. Mutagenesis was conducted using the Quickchange method.[S32] Designed forward primers containing desired mutation and a universal reverse primer was used and the PCR products were transformed into *E. coli* DH5a cells after digestion with DpnI restriction enzyme. The colonies were collected in LB medium (ampicillin, 100 mg L^{-1}) and plasmids were extracted by QIAprep Spin Miniprep Kit (Cat No.27104). Incorporation was then assessed by DNA sequencing to confirm the incorporation of desired mutations. The CYP119 variant was then transformed into *E. coli* DH5α cells and the protein expressed under the conditions described above. After expression, the cells were pelleted by centrifugation and resuspended in KPi buffer (50 mM, pH 7).

Representative Forward Primers:

Name	**Sequence**
C317S_f	GAATACATCTGTTAGGTGCTCC
T213A_f	GGTAATGAGGCTACAACTAAC
F153G_f	CTTAGTCGCAGGCAGGTTGGGTAAG
V254W_f	GTATTCTCCTCCTTGGATGAGGACTG
V254A_f	GTATTCTCCTCCTGCCATGAGGACTG
L209W_f	GATACATTATTTGGCTTCTCATAG

*Mutated regions highlighted in red.

Universal CYP119 Reverse Primer:

Name	**Sequence**
CYP119u_r	AGCCGGATCTCAGTGGTGGTG

Enzymatic Reactions. Analytical scale enzymatic reactions with purified proteins were carried out at a 500 µL-scale using the CYP119 variant (or other protein), indoline substrate, ethyl diazoacetate (or otherwise stated) and sodium dithionite $Na_2S_2O_4$ at the concentrations indicated in the Tables and legends. In a typical procedure, a solution of $Na_2S_2O_4$ in potassium phosphate buffer (50 mM, pH 7.0) was degassed by bubbling argon into the mixture for 3 min in a sealed vial. A buffered solution containing the CYP119 variant was carefully degassed in a similar manner in a separate vial. The two solutions were then mixed via cannula

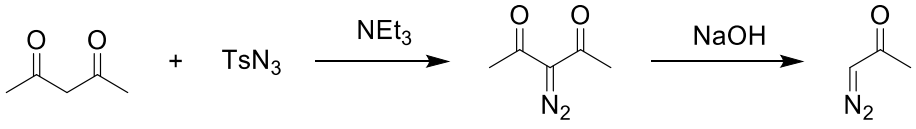
transfer. Reactions were initiated by addition of indoline derivative (from a 0.5 M stock solution in ethanol), followed by the addition of diazo compound (from a 0.5-5 M stock solution in ethanol) with a syringe, and the reaction mixture was stirred for 16 hours at room temperature, under positive argon pressure. For whole cell experiments, reactions were carried out at a 500 µL-scale using *E. coli.* whole cells expressing the CYP119 variant, indoline substrate, and ethyl diazoacetate (or otherwise stated) at the concentrations indicated in the Tables and legends. In a typical procedure, a sealed vial containing whole cells was degassed with argon for 3 min. The reactions were initiated by addition of indoline derivative (from a 0.5 M stock solution in ethanol), followed by the addition of diazo compound (from a 0.5-5 M stock solution in ethanol) with a syringe. The reaction mixture was stirred for 16 hours at room temperature under positive argon pressure. The TON for the whole-cell reactions were calculated based on CYP119 concentration in the reaction mixture as measured via UV-vis spectroscopy after cell lysis.

Hemoprotein Extinction Coefficient Calculation. Serine ligated CYP119 variant extinction coefficients were found using a reported pyridine hemochromagen assay.[S33-S35] The extinction coefficient of the CO-bound serine ligated CYP119 variants was found to be ca. ε_{406} = 130 $mM^{-1}cm^{-1}$.

Reaction Analysis. The reactions were analyzed by adding 25 µL of internal standard (benzodioxole, 50 mM in methanol) to a 500 µL aliquot of the reaction mixture, followed by extraction with 500 µL dichloromethane (DCM) and centrifugation at 14,000 rpm. The organic layer was collected and analyzed by GC for yield, and chiral HPLC or SFC for enantioselectivity. The TON for the whole-cell reactions were calculated based on CYP119 concentration in the reaction mixture as measured via UV-vis using the CO-binding assay (ε_{406} = 130 $mM^{-1}cm^{-1}$) after cell lysis. Calibration curves of the different products were constructed using authentic standards from the whole cell reactions (**General Procedure A**). Enantioselectivity was determined by HPLC or SFC using a chiral column as described in analytical methods section.

General Procedure A: Whole-cell biocatalytic reactions for C–H functionalization on a preparative scale. These reactions were carried out on a 50 mL-scale using C41(DE3) *E. coli* cells expressing the CYP119 variant, 10 mM indoline substrate, 80 mM diazo reagent. In a typical procedure, the substrate (0.4 mmol in 1 mL of ethanol) was added slowly to a 125 mL Erlenmeyer flask containing a stirring suspension of CYP119-expressing cells (OD_{600} = 60 in KPi, pH 7) in an anaerobic chamber. After stirring for 5 minutes, 640 µL of 5 M diazo solution (8 equiv.) in ethanol were added dropwise into the Erlenmeyer flask, then sealed with a rubber septum. Reaction mixture stirred at room temperature overnight. The reaction mixtures were extracted with ethyl acetate (100 mL x 3) and the combined organic layers were dried over $MgSO_4$ and concentrated under reduced pressure. The TON for the whole-cell reactions were calculated based on CYP119 concentration in the reaction mixture as measured via UV-vis spectroscopy using the CO-binding assay (ε_{406} = 130 $mM^{-1}cm^{-1}$) after cell lysis. The crude product was purified by flash column •purified product was characterized by NMR, GC-MS, and chiral HPLC or SFC for stereoselectivity determination and they were used as authentic standards for the construction of the calibration curves (TON and % conversion determination).

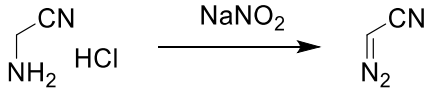
Synthesis of diazoacetone (2b)



2b

Diazoacetone (**2b**) was prepared according to a reported procedure.[S36] Isolated yield: 81%. NMR data matched the reported data.

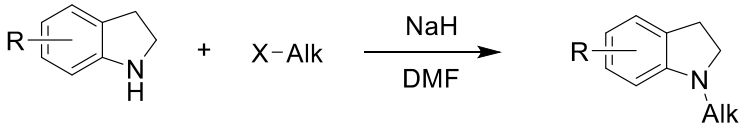
Synthesis of diazoacetonitrile (2c)



2c

Aminoacetonitrile hydrochloride (25 mmol, 1.0 equiv.) in water (2 mL) was added slowly to a solution of $NaNO_2$ (25 mmol, 1.0 equiv.) in ethanol (4 mL) over 1 hr via syringe pump at -10 °C (brine/ ice bath). After addition, reaction was stirred at r.t. for 1 hr or until completion. The concentration of the diazoacetonitrile (**2c**) solution was measured by ^1H NMR using DMS as a standard, diluted to the final concentration, and used directly for enzymatic reactions.

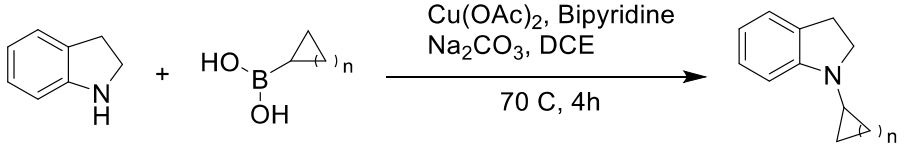
Synthesis of Indoline derivatives 1a-1j



1a-1j

N-substituted indolines **1a**-**1j** were prepared from the respective alkyl halide and substituted indole as follows. Sodium hydride (7.5 mmol, 1.25 equiv.) was added to a solution of indoline (5 mmol, 1.0 equiv.) in dry THF (10 mL) cooled to 0 °C. The solution was brought to r.t. and stirred for 15 minutes. Alkyl halide (7.5 mmol, 1.25 equiv.) was then added dropwise, and the reaction was stirred at r.t. for 4 hrs. Once completed, the reaction was quenched with a saturated $NaHCO_3$ and extracted 3 times with DCM (25 mL). The combined organic layers were dried over sodium sulfate, concentrated, and purified by flash chromatography on silica gel (EtoAc:Hex 1:4) to afford the desired product **1a**-**1j** in 48-93% yield.

Synthesis of Indoline derivatives 1k and 1l

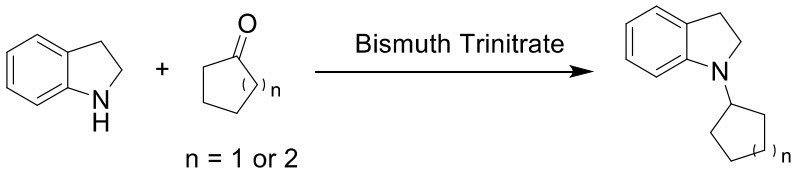


n = 1 or 2

1k/1l

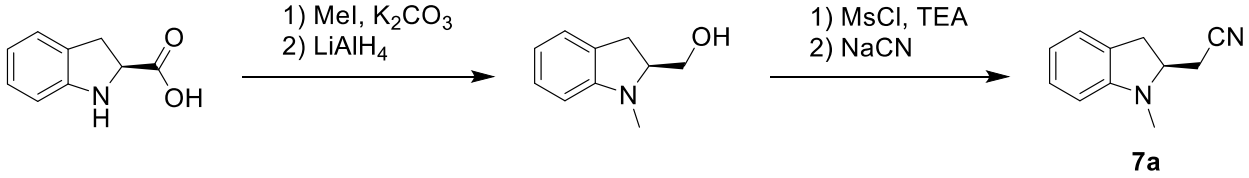
N-substituted indolines **1k** and **1l** were prepared according to a reported procedure.[S37] To a suspension of $Cu(OAc)_2$ (5.0 mmol, 1.0 equiv.) and bipyridine (5.0 mmol, 1.0 equiv.) in hot dichloroethane (40 mL), a suspension of cycloalkyl boronic acid (10 mmol, 2.0 equiv.) and Na_2CO_3 (10 mmol, 2.0 equiv.) in dichloroethane (10 mL) was added portionwise while stirring. The mixture was heated to 70 °C and stirred for 2-5 hours or until completion (monitored by TLC). Once completed, the reaction was cooled and a solution of 25% aqueous NH_4OH (40 mL) was added. The aqueous layer was extracted 3 times with DCM (40 mL), washed with brine and dried over sodium sulfate. After concentration, crude reactions were purified by flash chromatography on silica gel to afford 1-cyclopropylindoline (**1k**) in 85% yield (r.f. 0.78, EtOAc:Hex 1:4) or 1-cyclobutylindoline (**1l**) in 62% yield (r.f. 0.67, EtOAc:Hex 1:4).

Synthesis of Indoline derivatives 1m and 1n



N-substituted indolines **1m** and **1n** were prepared according to a reported procedure.[S38] Indoline (5.0 mmol, 1.0 equiv.), cyclic aliphatic ketone (2.5 mmol, 0.5 equiv.) and $Bi(NO_3)_3 \cdot H_2O$ (0.5 mmol, 0.1 equiv.) were stirred at 120 °C for 1 hr. After completion, reaction was diluted with ethyl acetate (30 mL), washed 3 times with water (30 mL), and dried over sodium sulfate. After concentration, crude reactions were purified by flash chromatography on silica gel to afford 1-cyclopentylindoline (**1m**) in 67% yield (r.f. 0.73, EtOAc:Hex 1:4) or 1-cyclohexylindoline (**1n**) in 72% yield (r.f. 0.73, EtOAc:Hex 1:4).

Synthesis of *S*-Isomer Authentic Standard (7a)



Authentic standard **7a** was prepared in a multistep sequence according to a reported procedure.[S39]

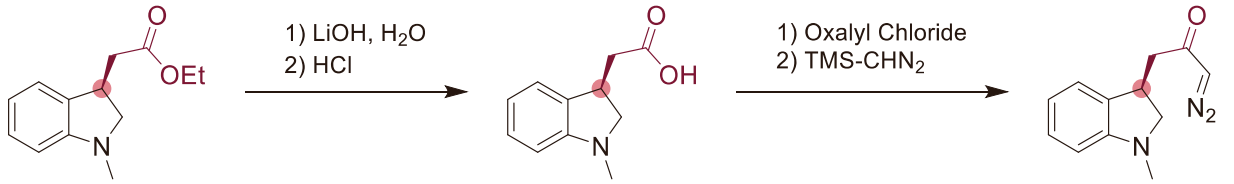
Step 1: (*S*)-indoline-2-carboxylic acid (4 mg, 2.5 mmol, 1.0 equiv.) and K_2CO_3 (7 mg, 5 mmol, 2.0 equiv.) were added to a round bottom flask and purged with argon. Acetone (8 mL) was added and iodomethane (530 mg, 3.75 mmol, 1.5 equiv.) were added dropwise. The reaction was heated to 40 °C and monitored by TLC. Once the reaction completed (~ 8 hours), the reaction was cooled and 1:1 EtOAc:Hex (15 mL) was added. The mixture was filtered, concentrated, taken up in DCM (15 mL), and subsequently washed with water and brine. The crude mixture was dried and concentrated for use in the next step.

Step 2: To a dry round bottom flask, the crude mixture from **Step 1** and dry THF (15 mL) were added. The reaction was cooled to 0 °C and $LiAlH_4$ (300 mg, 7.5 mmol, 3 equiv.) was added portionwise, after which the reaction was warmed to r.t. and stirred for 1 hr. Once completed, the reaction was cooled to 0 °C and quenched via the addition of acetone, followed by the addition of a saturated solution of Rochelle's salts (12 mL). The suspension was stirred until the solution became clear and extracted with Ether (3 x 20 mL). The combined organic layers were dried over sodium sulfate, concentrated, and purified by flash chromatography on silica gel to afford the desired product (*S*)-(1-methylindolin-2-yl)methanol in 61% yield.

Step 3: To a dry round bottom flask purged with argon, (*S*)-(1-methylindolin-2-yl)methanol (160 mg, 1.0 mmol, 1.0 equiv.) and dry DCM (5 mL) were added via syringe. The reaction was cooled to 0 °C and triethylamine (TEA) (280 µL, 2.0 mmol, 2.0 equiv.), then methanesulfonyl chloride (MsCl) (120 µL, 1.5 mmol, 1.5 equiv.) were added dropwise. The reaction was then warmed to r.t. and stirred for 1 hr. Once completed, the reaction was washed with saturated $NaHCO_3$ and brine, then subsequently dried over sodium sulfate and concentrated for use in the next step.

Step 4: To a dry round bottom flask purged with argon, the crude oil from **Step 3**, dry DMF (5 mL), and NaCN (150 mg, 3.0 mmol, 3.0 equiv.) were added. The sealed reaction was then heated to 60 °C under argon for 3 hr. After completion, the reaction was cooled and diluted with ethyl acetate (5 mL) and water (5 mL). The organic layer was washed with brine (3 x 10 mL), dried over sodium sulfate, concentrated, and purified by flash chromatography on silica gel to afford the desired product (*S*)-2-(1-methylindolin-2-yl)acetonitrile (**7a**) in 73% yield.

Synthesis of diazo tethered Indoline (8)

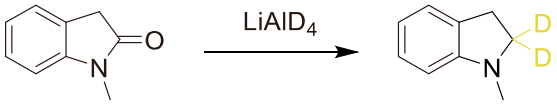


4a **8**

Diazo tethered indoline **8** was synthesized starting from the enzymatically obtained enantioenriched **4a** (general procedure A, 500mg scale). Ethyl (*R*)-2-(1-methylindolin-3-yl)acetate (**4a**) (2 mmol, 1.0 equiv.) and LiOH (10 mmol, 5.0 equiv.) in THF (10 mL) and water (10mL), were stirred at r.t. for 16 hrs. The reaction was quenched with 1 M NaOH (5 mL) and extracted with EtOAc (15 mL). The aqueous layer was then acidified with 1 M HCl to pH 3-4, extracted 3 times with EtOAc (15 mL), and dried over sodium sulfate. After concentration, the product was pure enough for subsequent reactions (97% yield, r.f. 0.44, EtOAc:Hex 1:1).

(*R*)-2-(1-methylindolin-3-yl)acetic acid previously obtained (1.94 mmol, 1.0 equiv.) and oxalyl chloride (2.91 mmol, 1.5 eq) in DCM (10 mL) were cooled to 0 °C. To the cooled reaction, THF (0.1 mmol, 0.05 equiv.) was added, the reaction was warmed to r.t. and stirred for 2 hrs. Once completed, the reaction was concentrated, and the crude mixture was used in the following reaction. The crude acid chloride (1.0 equiv.) was taken up in acetonitrile (10 mL) and cooled to 0 °C. To the reaction, 1.8-2.4 M TMS-diazomethane in hexanes (9.7 mmol, 5.0 equiv.) was cooled and added dropwise over 20 minutes via syringe pump. Once addition was completed, the reaction was slowly brought to r.t. and stirred overnight. The reaction was then concentrated under reduced pressure while cooling, taken up in ether (20 mL) and washed with brine (20 mL). After concentration, the crude reaction was purified by flash chromatography on silica gel to afford (*R*)-1-diazo-3-(1-methylindolin-3-yl)propan-2-one (**8**) in 46% yield (r.f. 0.61, EtOAc:Hex 1:1).

Synthesis of a-deuterated substrates (1a-*D*$_2$ and 10a-*D*$_4$)



1a-*D*$_2$

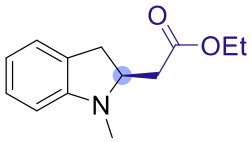
A flame dried round bottom charged with $LiAlD_4$ (2.0 mmol, 2 equiv.) and a stir bar were purged and backfilled with argon (x3). Dry ether (3 mL) was added, and the flask was cooled to 0 °C. 1-methylindolin-2-one (1.0 mmol, 1.0 eq) in dry THF (3 mL) was added dropwise, and the reaction was stirred at r.t. for 3 hrs. Due to the stability of the desired deuterated product in air, the reaction mixture was quenched with degassed sat. K_2CO_3 (3 mL) and degassed water (3 mL). The organic layer was extracted via canula to an argon purged Schlenk flask and reduced under vacuum. After concentration, the crude reaction was backfilled with argon and quickly loaded onto silica to be purified by flash chromatography to afford 1-methylindoline-2,2-d_2 (**1a-*D*$_2$**) in 62% yield (r.f. 0.63, EtOAc:Hex 1:4).

1-phenylpyrrolidine-2,2,5,5-d_4 (**10a-*D*$_4$**) was synthesized in a similar manner from 1-phenylpyrrolidine-2,5-dione using 4.0 equiv. of $LiAlD_2$ (compared to 2 equiv.). No need for using degassed solvents or purged flasks was observed, and the crude reaction was purified by flash chromatography to afford 1-phenylpyrrolidine-2,2,5,5-d_4 (**10a-*D*$_4$**) in 96% yield (r.f. 0.77, EtOAc:Hex 1:4).

Compound Characterization

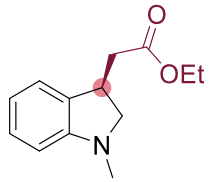
Absolute configuration of enzymatic products **3a-3n**, **6a**, and **7a** were determined by relation to the synthetically derived *S*-isomer authentic standard of **7a**, assuming facial selectivity of the enzymatic products is retained. Similarly, the absolute configuration of enzymatic products **4a**, **4b**, and **9a** were determined assuming retained facial selectivity from **7a** for the secondary enzymatic insertion to form *cis*-**9a** selectively, displaying a retained facial selectivity of the enzyme for both the C_2 and C_3 insertion products.

Ethyl (*S*)-2-(1-methylindolin-2-yl)acetate (3a)



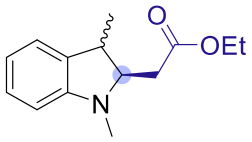
Ethyl (*S*)-2-(1-methylindolin-2-yl)acetate (**3a**) was prepared according to the general procedure A (r.f. 0.49, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.1-7.0 (m, 2H), 6.7 (td, *J* = 7.4, 1.0 Hz, 1H), 6.5 (d, *J* = 7.8 Hz, 1H), 4.2 (q, *J* = 7.2 Hz, 2H), 3.7 (qd, *J* = 8.9, 4.3 Hz, 1H), 3.3-3.2 (m, 1H), 2.8 (dd, *J* = 15.1, 4.3 Hz, 1H), 2.8-2.7 (m, 1H), 2.7 (s, 2H), 2.5 (dd, *J* = 15.1, 8.9 Hz, 1H), 1.3-1.2 (m, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.7, 152.8, 128.5, 127.5, 124.2, 118.1, 107.3, 63.6, 60.6, 38.7, 35.5, 29.7, 14.2. **HRMS** (ESI): Calculated for $C_{13}H_{18}NO_2$ $[M+H]^+$ 220.1338; found 220.1324.

Ethyl (*R*)-2-(1-methylindolin-3-yl)acetate (4a)



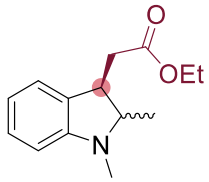
Ethyl (*R*)-2-(1-methylindolin-3-yl)acetate (**4a**)was prepared according to the general procedure A (r.f. 0.48, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.1 (dd, *J* = 7.8, 1.4 Hz, 1H), 7.1-7.0 (m, 1H), 6.7 (td, *J* = 7.4, 1.0 Hz, 1H), 6.5 (d, *J* = 7.9 Hz, 1H), 4.2-4.1 (m, 2H), 3.6 (ddd, *J* = 15.0, 8.7, 6.0 Hz, 1H), 3.5 (t, *J* = 8.6 Hz, 1H), 3.0 (dd, *J* = 8.8, 6.7 Hz, 1H), 2.8-2.7 (m, 4H), 2.5 (dd, *J* = 16.0, 9.2 Hz, 1H), 1.3 (t, *J* = 7.3 Hz, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.4, 153.0, 132.2, 128.1, 123.5, 117.9, 107.4, 62.0, 60.5, 38.8, 37.3, 35.9, 14.3. **HRMS** (ESI): Calculated for $C_{13}H_{18}NO_2$ $[M+H]^+$ 220.1338; found 220.1321.

Ethyl 2-((2*R*)-1,3-dimethylindolin-2-yl)acetate (3b)



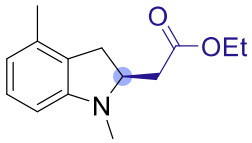
Ethyl 2-((2*R*)-1,3-dimethylindolin-2-yl)acetate (**3b**) was prepared according to the general procedure A (r.f. 0.54, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.1-7.0 (m, 2H), 6.6-6.6 (m, 1H), 6.4 (d, *J* = 7.8 Hz, 1H), 4.1 (q, *J* = 7.1 Hz, 2H), 3.7 (ddq, *J* = 9.7, 8.5, 6.1 Hz, 1H), 3.5 (dq, *J* = 21.2, 7.5 Hz, 2H), 3.1 (dd, *J* = 15.4, 8.5 Hz, 1H), 2.7 (s, 1H), 2.6-2.5 (m, 3H), 1.3-1.2 (m, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 173.2, 152.2, 128.1, 124.9, 118.2, 107.2, 61.1, 60.4, 42.6, 38.1, 32.8, 20.1, 14.9, 14.9. **HRMS** (ESI): Calculated for $C_{14}H_{20}NO_2$ $[M+H]^+$ 234.1494; found 234.1484.

Ethyl 2-((3*R*)-1,2-dimethylindolin-3-yl)acetate (4b)



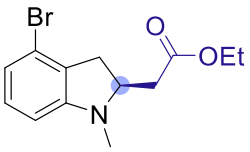
Ethyl 2-((3*R*)-1,2-dimethylindolin-3-yl)acetate (**4b**) was prepared according to the general procedure A (r.f. 0.61, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, CDCl3) δ 7.1 (t, J = 7.7 Hz, 1H), 7.0 (dd, J = 21.9, 7.3 Hz, 1H), 6.7 (td, J = 7.5, 2.8 Hz, 1H), 6.5 (dd, J = 10.6, 7.8 Hz, 1H), 4.2-4.1 (m, 2H), 3.6 (q, J = 7.7 Hz, 1H), 3.5 (p, J = 6.8 Hz, 1H), 2.7 (dd, J = 16.7, 1.5 Hz, 3H), 2.6 (dd, J = 15.8, 7.2 Hz, 1H), 2.5 (dd, J = 15.7, 7.1 Hz, 1H), 2.5-2.4 (m, 1H), 1.3-1.3 (m, 3H), 1.3 (d, 1H), 1.2 (d, J = 6.4 Hz, 2H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 128.0, 124.0, 123.3, 118.1, 117.8, 107.1, 68.3, 64.6, 60.6, 60.5, 45.1, 40.9, 38.5, 34.4, 33.5, 33.2, 17.7, 14.3, 14.2, 12.7. **HRMS** (ESI): Calculated for $C_{14}H_{20}NO_2$ $[M+H]^+$ 234.1494; found 234.1486.

Ethyl (*S*)-2-(1,4-dimethylindolin-2-yl)acetate (3c)



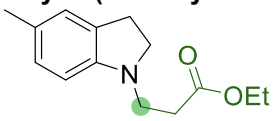
Ethyl (*S*)-2-(1,4-dimethylindolin-2-yl)acetate (**3c**) was prepared according to the general procedure A (r.f. 0.20, EtOAc:Hex 1:12). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.0 (t, *J* = 7.7 Hz, 1H), 6.5 (d, *J* = 7.6 Hz, 1H), 6.3 (d, *J* = 7.8 Hz, 1H), 4.3-4.1 (m, 4H), 3.8-3.6 (m, 1H), 3.2 (dd, *J* = 15.7, 8.7 Hz, 1H), 2.8 (dd, *J* = 15.2, 4.2 Hz, 1H), 2.7 (s, 3H), 2.6 (dd, *J* = 15.7, 9.4 Hz, 2H), 2.5 (dd, *J* = 15.2, 9.1 Hz, 1H), 1.3-1.2 (m, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 166.0, 153.3, 130.5, 128.3, 120.3, 105.6, 64.1, 62.0, 61.3, 39.7, 35.2, 35.1, 19.2, 15.0. **HRMS** (ESI): Calculated for $C_{14}H_{20}NO_2$ $[M+H]^+$ 234.1494; found 234.1481.

Ethyl (*S*)-2-(4-bromo-1-methylindolin-2-yl)acetate (3d)



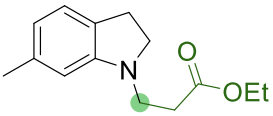
Ethyl (*S*)-2-(4-bromo-1-methylindolin-2-yl)acetate (**3d**) was prepared according to the general procedure A (r.f. 0.20, EtOAc:Hex 1:8). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 6.9 (t, *J* = 7.8 Hz, 1H), 6.8 (d, *J* = 8.0 Hz, 1H), 6.3 (d, *J* = 7.8 Hz, 1H), 4.2 (q, *J* = 7.1 Hz, 2H), 3.8 (td, *J* = 8.8, 4.3 Hz, 1H), 3.3 (dd, *J* = 16.3, 9.0 Hz, 1H), 2.8-2.7 (m, 5H), 2.5 (dd, *J* = 15.3, 8.7 Hz, 1H), 1.3-1.2 (m, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.1, 154.4, 129.9, 121.4, 106.2, 78.0, 77.7, 77.5, 63.2, 61.4, 39.5, 37.4, 32.3, 23.4, 14.9. **HRMS** (ESI): Calculated for $C_{13}H_{17}BrNO_2$ $[M+H]^+$ 298.0443; found 298.0425.

Ethyl 3-(5-methylindolin-1-yl)propanoate (5e)



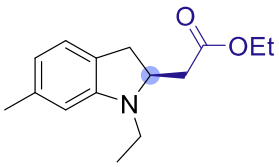
Ethyl 3-(5-methylindolin-1-yl)propanoate (**5e**) was prepared according to the general procedure A (r.f. 0.20, EtOAc:Hex 1:12). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 6.9 (dd, *J* = 18.9, 10.8 Hz, 2H), 6.4 (dd, *J* = 10.8, 7.6 Hz, 1H), 4.2-4.1 (m, 1H), 4.1-4.1 (m, 1H), 3.4-3.2 (m, 4H), 2.9 (q, *J* = 9.0 Hz, 2H), 2.6-2.5 (m, 2H), 2.2 (d, *J* = 9.7 Hz, 3H), 1.2 (dt, *J* = 19.0, 7.3 Hz, 3H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 173.1, 150.5, 131.0, 128.2, 128.0, 126.1, 107.8, 61.3, 54.2, 46.3, 33.4, 29.4, 21.4, 15.0. **HRMS** (ESI): Calculated for $C_{14}H_{20}NO_2$ $[M+H]^+$ 234.1494; found 234.1484.

Ethyl 3-(6-methylindolin-1-yl)propanoate (5f)



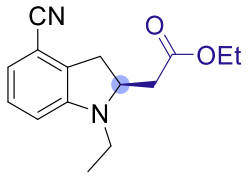
Ethyl 3-(6-methylindolin-1-yl)propanoate (**5f**) was prepared according to the general procedure A (r.f. 0.20, EtOAc:Hex 1:12). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 6.9 (d, *J* = 7.3 Hz, 1H), 6.5 (d, *J* = 7.3 Hz, 1H), 6.3 (s, 1H), 4.1 (q, *J* = 7.2 Hz, 2H), 3.4 (t, *J* = 7.1 Hz, 2H), 3.3 (t, *J* = 8.2 Hz, 2H), 2.9 (t, *J* = 8.3 Hz, 2H), 2.6 (t, *J* = 7.1 Hz, 2H), 2.3 (s, 3H), 1.2 (t, *J* = 7.1 Hz, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.3, 152.1, 137.2, 127.1, 124.1, 118.4, 108.0, 60.6, 53.4, 45.1, 32.6, 28.2, 21.7, 14.2. **HRMS** (ESI): Calculated for $C_{14}H_{20}NO_2$ $[M+H]^+$ 234.1494; found 234.1485.

Ethyl (*S*)-2-(1-ethyl-6-methylindolin-2-yl)acetate (3f)



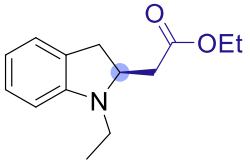
Ethyl (*S*)-2-(1-ethyl-6-methylindolin-2-yl)acetate (**3f**) was prepared according to the general procedure A (r.f. 0.40, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 6.9-6.9 (m, 1H), 6.4 (ddd, *J* = 7.3, 1.6, 0.8 Hz, 1H), 6.3-6.3 (m, 1H), 4.3-4.2 (m, 1H), 4.2-4.1 (m, 2H), 4.1 (d, *J* = 7.1 Hz, 1H), 3.4-3.3 (m, 2H), 2.9 (t, *J* = 8.3 Hz, 2H), 2.6 (dd, *J* = 14.4, 6.5 Hz, 1H), 2.4 (dd, *J* = 14.4, 8.1 Hz, 1H), 2.3 (d, *J* = 1.0 Hz, 3H), 1.3-1.2 (m, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.1, 150.8, 137.1, 127.2, 124.1, 118.0, 108.2, 60.5, 48.0, 46.2, 38.4, 27.8, 21.8, 16.0, 14.2. **HRMS** (ESI): Calculated for $C_{15}H_{22}NO_2$ $[M+H]^+$ 248.1651; found 248.1637.

Ethyl (*S*)-2-(4-cyano-1-ethylindolin-2-yl)acetate (3g)



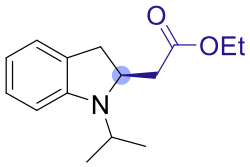
Ethyl (*S*)-2-(4-cyano-1-ethylindolin-2-yl)acetate (**3g**) was prepared according to the general procedure A (r.f. 0.30, EtOAc:Hex 1:10). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.1 (t, *J* = 7.9 Hz, 1H), 6.8 (d, *J* = 7.7 Hz, 1H), 6.6 (d, *J* = 8.0 Hz, 1H), 4.1 (q, *J* = 7.0 Hz, 1H), 4.1 (q, *J* = 7.1 Hz, 2H), 3.5-3.4 (m, 2H), 3.2-3.0 (m, 2H), 2.6 (dd, *J* = 14.6, 7.2 Hz, 1H), 2.4 (dd, *J* = 14.6, 7.4 Hz, 1H), 1.2 (dt, *J* = 7.2, 3.8 Hz, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.3, 152.0, 129.2, 120.3, 118.7, 111.0, 109.3, 61.4, 48.8, 46.4, 39.2, 28.4, 17.0, 14.9. **HRMS** (ESI): Calculated for $C_{15}H_{19}N_2O_2$ $[M+H]^+$ 259.1447; found 259.1432.

Ethyl (*S*)-2-(1-ethylindolin-2-yl)acetate (3h)



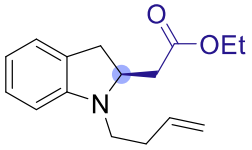
Ethyl (*S*)-2-(1-ethylindolin-2-yl)acetate (**3h**) was prepared according to the general procedure A (r.f. 0.52, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.1-7.0 (m, 2H), 6.6 (td, *J* = 7.3, 1.0 Hz, 1H), 6.5-6.5 (m, 1H), 4.3-4.1 (m, 3H), 3.4-3.3 (m, 2H), 3.0-2.8 (m, 2H), 2.6 (dd, *J* = 14.4, 6.5 Hz, 1H), 2.4 (dd, *J* = 14.4, 8.1 Hz, 1H), 1.3-1.1 (m, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.8, 151.4, 130.8, 128.1, 125.1, 118.1, 107.9, 61.2, 48.8, 46.7, 39.1, 28.9, 16.8, 14.9. **HRMS** (ESI): Calculated for $C_{14}H_{20}NO_2$ $[M+H]^+$ 234.1494; found 234.1482.

Ethyl (*S*)-2-(1-isopropylindolin-2-yl)acetate (3i)



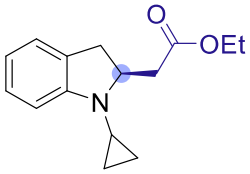
Ethyl (*S*)-2-(1-isopropylindolin-2-yl)acetate (**3i**) was prepared according to the general procedure A (r.f. 0.51, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.1-7.0 (m, 2H), 6.6 (t, *J* = 7.3 Hz, 1H), 6.5 (d, *J* = 7.8 Hz, 1H), 4.2 (q, *J* = 7.2 Hz, 2H), 4.1 (dtd, *J* = 9.7, 5.6, 2.8 Hz, 1H), 3.6 (p, *J* = 6.8 Hz, 1H), 3.3 (dd, *J* = 16.1, 9.6 Hz, 1H), 2.8-2.7 (m, 2H), 2.5 (dd, *J* = 15.2, 9.7 Hz, 1H), 1.3-1.2 (m, 9H). **^{13}C NMR** (126 MHz, CDCl3) δ 171.8, 150.4, 128.9, 127.2, 124.5, 117.3, 108.6, 77.3, 77.0, 76.8, 60.4, 56.8, 48.5, 41.9, 35.9, 19.9, 19.6, 14.2. **HRMS** (ESI): Calculated for $C_{15}H_{22}NO_2$ $[M+H]^+$ 248.1651; found 248.1636.

Ethyl (*S*)-2-(1-(but-3-en-1-yl)indolin-2-yl)acetate (3j)



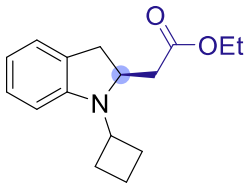
Ethyl (*S*)-2-(1-(but-3-en-1-yl)indolin-2-yl)acetate (**3j**) was prepared according to the general procedure A (r.f. 0.62, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.0 (t, *J* = 8.2 Hz, 2H), 6.6 (t, *J* = 7.3 Hz, 1H), 6.5 (d, *J* = 7.7 Hz, 1H), 5.8 (ddt, *J* = 17.0, 10.2, 6.9 Hz, 1H), 5.1 (dt, *J* = 17.1, 1.7 Hz, 1H), 5.0 (dd, *J* = 10.1, 1.7 Hz, 1H), 4.1 (dq, *J* = 19.5, 7.2 Hz, 3H), 3.4 (dt, *J* = 17.1, 8.4 Hz, 2H), 2.9 (t, *J* = 8.4 Hz, 2H), 2.5 (qd, *J* = 14.6, 7.2 Hz, 2H), 2.4-2.3 (m, 1H), 2.3 (dt, *J* = 14.2, 7.2 Hz, 1H), 1.3-1.2 (m, 3H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.5, 150.9, 129.6, 127.9, 125.3, 117.8, 108.8, 61.2, 59.7, 59.3, 40.5, 36.0, 29.6, 25.0, 24.7, 15.0. **HRMS** (ESI): Calculated for $C_{16}H_{22}NO_2$ $[M+H]^+$ 260.16505; found 260.1636.

Ethyl (*S*)-2-(1-cyclopropylindolin-2-yl)acetate (3k)



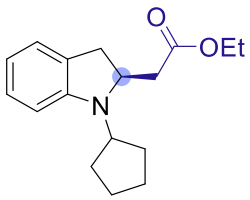
Ethyl (*S*)-2-(1-cyclopropylindolin-2-yl)acetate (**3k**) was prepared according to the general procedure A (r.f. 0.63, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.1-7.0 (m, 2H), 6.8 (d, *J* = 7.7 Hz, 1H), 6.7 (td, *J* = 7.4, 1.0 Hz, 1H), 4.1 (qd, *J* = 7.1, 2.1 Hz, 2H), 4.0 (tdd, *J* = 8.7, 6.3, 4.6 Hz, 1H), 3.1 (dd, *J* = 15.6, 8.5 Hz, 1H), 2.9 (dd, *J* = 15.0, 4.5 Hz, 1H), 2.7 (dd, *J* = 15.6, 6.3 Hz, 1H), 2.5 (dd, *J* = 15.0, 9.0 Hz, 1H), 2.2 (ddd, *J* = 8.6, 6.6, 3.7 Hz, 1H), 1.2 (t, *J* = 7.2 Hz, 3H), 0.8-0.6 (m, 5H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.7, 152.8, 129.0, 128.0, 125.2, 119.1, 109.0, 62.8, 61.2, 38.6, 35.7, 28.1, 15.0, 7.7, 5.7. **HRMS** (ESI): Calculated for $C_{15}H_{20}NO_2$ $[M+H]^+$ 246.1494; found 246.1480.

Ethyl (*S*)-2-(1-cyclobutylindolin-2-yl)acetate (3l)



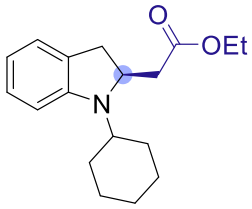
Ethyl (*S*)-2-(1-cyclobutylindolin-2-yl)acetate (**3l**) was prepared according to the general procedure A (r.f. 0.68, EtOAc:Hex 1:4). **^{1}H NMR (500 MHz, $CDCl_3$)** δ 7.0-7.0 (m, 2H), 6.6 (t, *J* = 7.3 Hz, 1H), 6.4 (d, *J* = 7.8 Hz, 1H), 4.2 (tt, *J* = 9.3, 3.4 Hz, 1H), 4.1 (qd, *J* = 7.1, 1.5 Hz, 2H), 3.9-3.8 (m, 1H), 3.2 (dd, *J* = 15.9, 8.9 Hz, 1H), 2.7 (dd, *J* = 15.9, 3.4 Hz, 1H), 2.5 (dd, *J* = 14.9, 3.6 Hz, 1H), 2.4-2.4 (m, 1H), 2.4-2.2 (m, 3H), 2.2 (p, *J* = 10.0 Hz, 1H), 1.9-1.8 (m, 2H), 1.2 (s, 3H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.5, 150.9, 129.3, 128.0, 125.4, 118.3, 108.5, 61.2, 58.9, 52.6, 39.1, 35.6, 29.9, 28.8, 16.5, 14.9. **HRMS** (ESI): Calculated for $C_{16}H_{22}NO_2$ $[M+H]^+$ 260.1651; found 260.1635.

Ethyl (*S*)-2-(1-cyclopentylindolin-2-yl)acetate (3m)



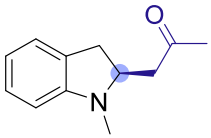
Ethyl (*S*)-2-(1-cyclopentylindolin-2-yl)acetate (**3m**) was prepared according to the general procedure A (r.f. 0.65, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.0 (t, *J* = 7.7 Hz, 2H), 6.6 (td, *J* = 7.4, 0.9 Hz, 1H), 6.5 (d, *J* = 7.8 Hz, 1H), 4.2-4.1 (m, 3H), 3.7 (h, *J* = 7.5 Hz, 1H), 3.3 (dd, *J* = 15.9, 9.2 Hz, 1H), 2.7 (dd, *J* = 15.9, 4.5 Hz, 1H), 2.6 (dd, *J* = 15.2, 3.4 Hz, 1H), 2.4 (dd, *J* = 15.1, 9.8 Hz, 1H), 1.9 (qd, *J* = 7.3, 3.3 Hz, 3H), 1.8 (qd, *J* = 6.3, 2.5 Hz, 5H), 1.6-1.5 (m, 2H), 1.2 (t, *J* = 7.1 Hz, 3H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.5, 150.9, 129.6, 127.9, 125.3, 117.8, 108.8, 61.2, 59.7, 59.3, 40.5, 36.0, 29.6, 29.5, 25.0, 24.7, 15. **HRMS** (ESI): Calculated for $C_{17}H_{24}NO_2$ $[M+H]^+$ 274.1807; found 274.1787.

Ethyl (*S*)-2-(1-cyclohexylindolin-2-yl)acetate (3n)



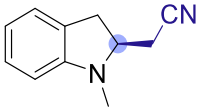
Ethyl (*S*)-2-(1-cyclohexylindolin-2-yl)acetate (**3n**) was prepared according to the general procedure A (r.f. 0.65, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.0-7.0 (m, 2H), 6.6 (td, *J* = 7.3, 1.0 Hz, 1H), 6.5 (d, *J* = 7.8 Hz, 1H), 4.2-4.1 (m, 3H), 3.3 (dd, *J* = 16.2, 9.8 Hz, 1H), 3.2 (tt, *J* = 12.1, 3.7 Hz, 1H), 2.7-2.7 (m, 2H), 2.5 (ddd, *J* = 15.3, 9.8, 1.2 Hz, 1H), 1.9-1.7 (m, 4H), 1.7-1.6 (m, 1H), 1.6-1.4 (m, 3H), 1.4-1.2 (m, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.5, 151.2, 129.5, 128.0, 125.2, 117.9, 109.2, 61.1, 58.4, 57.5, 43.0, 36.7, 31.6, 30.7, 27.1, 27.1, 26.7, 15.0. **HRMS** (ESI): Calculated for $C_{18}H_{26}NO_2$ $[M+H]^+$ 288.1964; found 288.1946.

(*S*)-1-(1-methylindolin-2-yl)propan-2-one (6a)



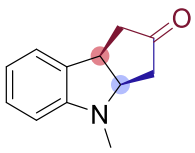
(*S*)-1-(1-methylindolin-2-yl)propan-2-one (**6a**) was prepared according to the general procedure A (r.f. 0.62, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.0 (dt, *J* = 15.0, 7.5 Hz, 3H), 6.7 (d, *J* = 7.3 Hz, 1H), 6.4 (d, *J* = 7.7 Hz, 1H), 3.8-3.7 (m, 1H), 3.3-3.2 (m, 1H), 3.1-2.9 (m, 2H), 2.7 (s, 4H), 2.6 (dd, *J* = 15.7, 9.3 Hz, 1H), 2.2 (s, 3H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 208.1, 166.0, 144.2, 129.5, 128.2, 125.0, 119.0, 107.9, 78.0, 77.7, 77.5, 63.5, 48.4, 36.5, 35.2, 31.7. **HRMS** (ESI): Calculated for $C_{12}H_{16}NO$ $[M+H]^+$ 190.1232; found 190.1228.

(*S*)-2-(1-methylindolin-2-yl)acetonitrile (7a)



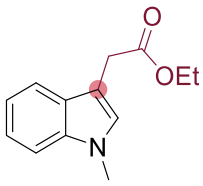
(*S*)-2-(1-methylindolin-2-yl)acetonitrile (**7a**) was prepared according to the general procedure A (r.f. 0.36, EtOAc:Hex 1:4), product characterization data matched reported values.[S40] **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.15-7.04 (m, 2H), 6.68 (m, 1H), 6.48 (d, *J* = 7.4 Hz, 1H), 3.69 (m, 1H), 3.29 (dd, *J* = 15.7, 8.8 Hz, 1H), 2.88 (dd, *J* = 15.6, 9.0 Hz, 1H), 2.79 (s, 3H), 2.75-2.62 (m, 2H). **^{13}C NMR** (125 MHz, $CDCl_3$) δ 152.3, 128.0, 127.4, 124.5, 119.0, 117.6, 107.9, 63.0, 35.2, 34.5, 22.1. **HRMS** (ESI): Calculated for $C_{11}H_{12}N_2$ $[M+H]^+$ 173.1079; found 173.1077.

(3a*R*,8b*R*)-4-methyl-3,3a,4,8b-tetrahydrocyclopenta[b]indol-2(1H)-one (9a)



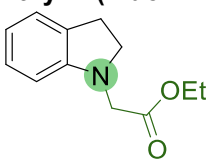
(3a*R*,8b*R*)-4-methyl-3,3a,4,8b-tetrahydrocyclopenta[b]indol-2(1H)-one (**9a**) was prepared according to the general procedure A (r.f. 0.60, EtOAc:Hex 1:1). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.2-7.1 (m, 2H), 6.7 (td, *J* = 7.4, 1.0 Hz, 1H), 6.5 (d, *J* = 7.8 Hz, 1H), 4.1 (ddd, *J* = 8.0, 6.1, 2.0 Hz, 1H), 3.8-3.8 (m, 1H), 2.7-2.6 (m, 5H), 2.5 (ddd, *J* = 18.7, 6.1, 1.5 Hz, 1H), 2.3 (ddd, *J* = 19.2, 7.6, 1.5 Hz, 1H), 2.2 (d, *J* = 3.3 Hz, 1H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 217.5, 152.2, 128.2, 124.1, 118.5, 107.7, 68.8, 43.1, 42.3, 41.9, 33.5, 29.7. **HRMS** (ESI): Calculated for $C_{12}H_{14}NO$ $[M+H]^+$ 188.1075; found 188.1064.

Ethyl 2-(1-methyl-1H-indol-3-yl)acetate (4a-*DS*)



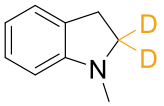
Ethyl 2-(1-methyl-1H-indol-3-yl)acetate (**4a-*DS***) was prepared according to the general procedure A (r.f. 0.0.39, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, CDCl3) δ 7.6 (d, J = 8.0 Hz, 1H), 7.2-7.2 (m, 2H), 7.1 (ddd, J = 8.0, 6.4, 1.5 Hz, 1H), 6.9 (s, 1H), 4.1 (q, J = 7.1 Hz, 2H), 3.7 (d, J = 1.4 Hz, 2H), 3.6 (s, 3H), 1.2 (t, J = 7.2 Hz, 3H). **^{13}C NMR** (126 MHz, CDCl3) δ 172.2, 137.0, 127.9, 121.8, 119.2, 119.2, 109.4, 107.0, 60.8, 32.7, 31.5, 29.9, 14.4. **HRMS** (ESI): Calculated for $C_{13}H_{16}NO_2$ $[M+H]^+$ 218.1181; found 218.1179.

Ethyl 2-(indolin-1-yl)acetate (13a)



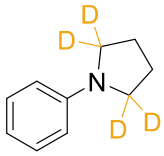
Ethyl 2-(indolin-1-yl)acetate (**13a**) was prepared according to the general procedure A (r.f. 0.52, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, CDCl3) δ 7.1-7.0 (m, 2H), 6.7 (td, J = 7.4, 1.0 Hz, 1H), 6.4 (d, J = 7.9 Hz, 1H), 4.2 (q, J = 7.2 Hz, 2H), 3.8 (s, 2H), 3.5 (t, J = 8.4 Hz, 2H), 3.0 (t, J = 8.4 Hz, 2H), 1.2 (t, J = 7.2 Hz, 3H). **^{13}C NMR** (126 MHz, CDCl3) δ 170.3, 151.4, 129.7, 127.3, 124.6, 118.2, 106.7, 60.9, 53.8, 50.6, 28.7, 14.3. **HRMS** (ESI): Calculated for $C_{12}H_{15}NO_2$ $[M+H]^+$ 205.1103; found 205.1096.

1-methylindoline-2,2-*d*$_2$ (1a-*D*$_2$)



1-methylindoline-2,2-*d*$_2$ (**1a-*D*$_2$**) was prepared according to the described procedure under the experimental section (r.f. 0.66, EtOAc:Hex 1:4). **^{1}H NMR (500 MHz, $CDCl_3$)** δ 7.0-7.0 (m, 2H), 6.6 (t, *J* = 7.3 Hz, 1H), 6.4 (d, *J* = 8.0 Hz, 1H), 2.9 (s, 2H), 2.7 (s, 3H). **HRMS** (ESI): Calculated for $C_9H_9D_2N$ $[M+H]^+$ 136.1095; found 136.1091.

1-phenylpyrrolidine-2,2,5,5-*d*$_4$ (10a-*D*$_4$)

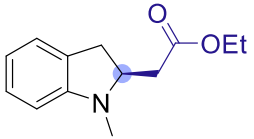


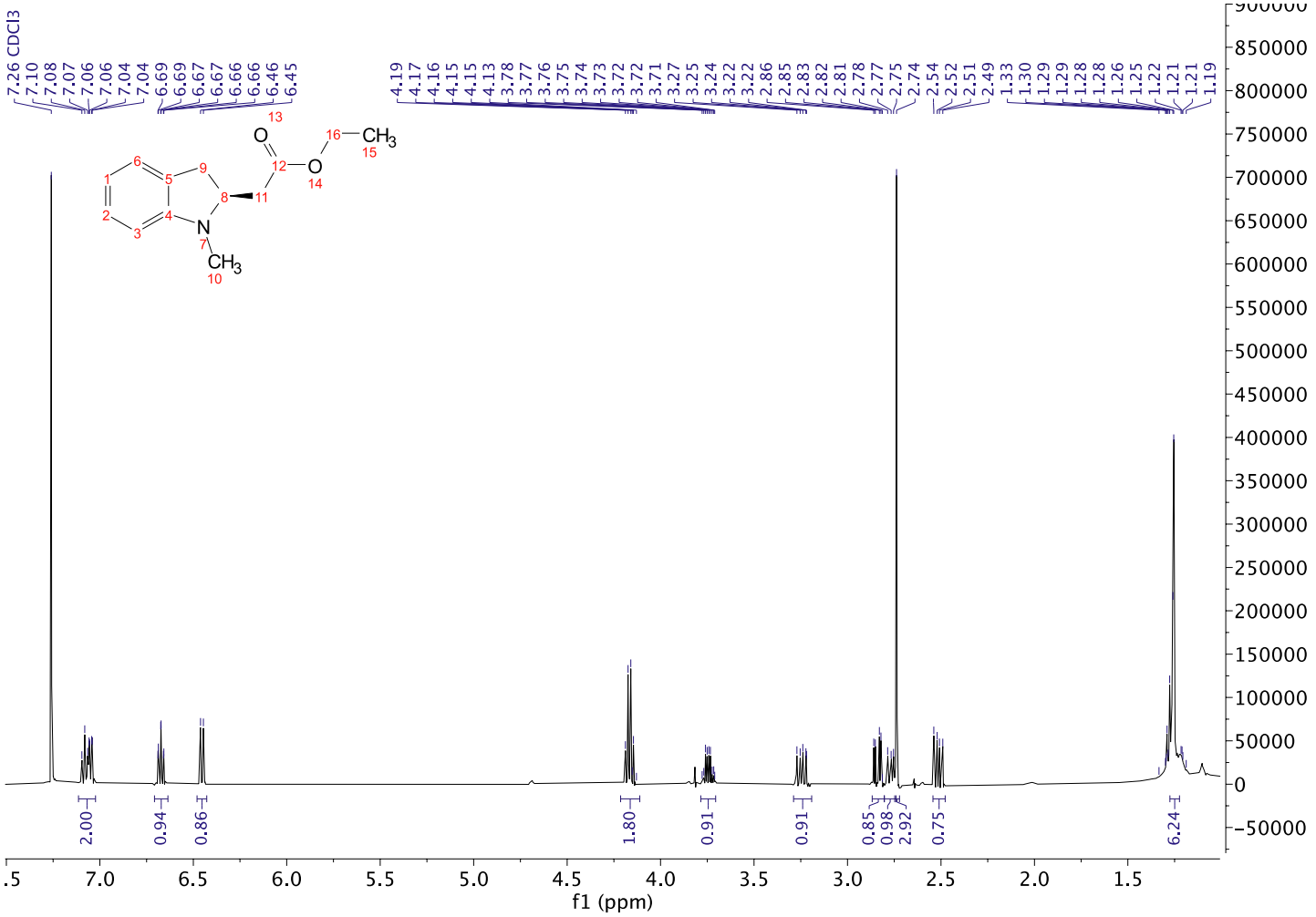
1-phenylpyrrolidine-2,2,5,5-*d*$_4$ (**10a-*D*$_4$**) was prepared according to the described procedure under the experimental section (r.f. 0.77, EtOAc:Hex 1:4). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.3-7.3 (m, 2H), 6.7 (tt, *J* = 7.2, 1.1 Hz, 1H), 6.7-6.6 (m, 2H), 2.0 (s, 4H). **HRMS** (ESI): Calculated for $C_{10}H_9D_4N$ $[M+H]^+$ 152.1377; found 152.1372.

Data S8. NMR Data

Ethyl (*S*)-2-(1-methylindolin-2-yl)acetate (3a)

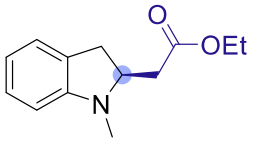
^{1}H NMR (500 MHz, $CDCl_3$)

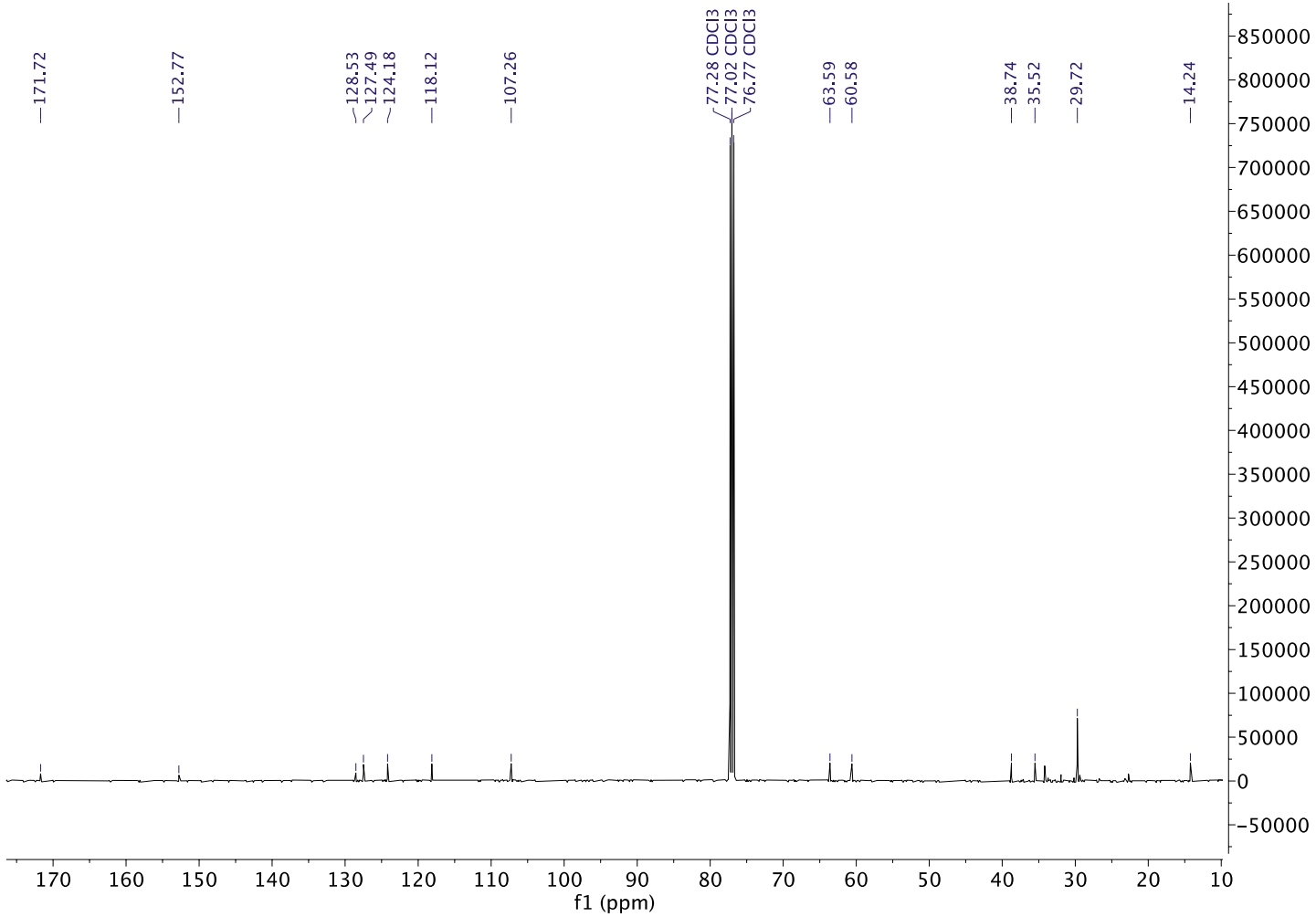




Ethyl (*S*)-2-(1-methylindolin-2-yl)acetate (3a)

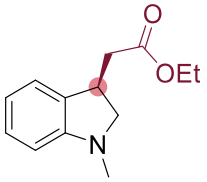
^{13}C NMR (126 MHz, $CDCl_3$)

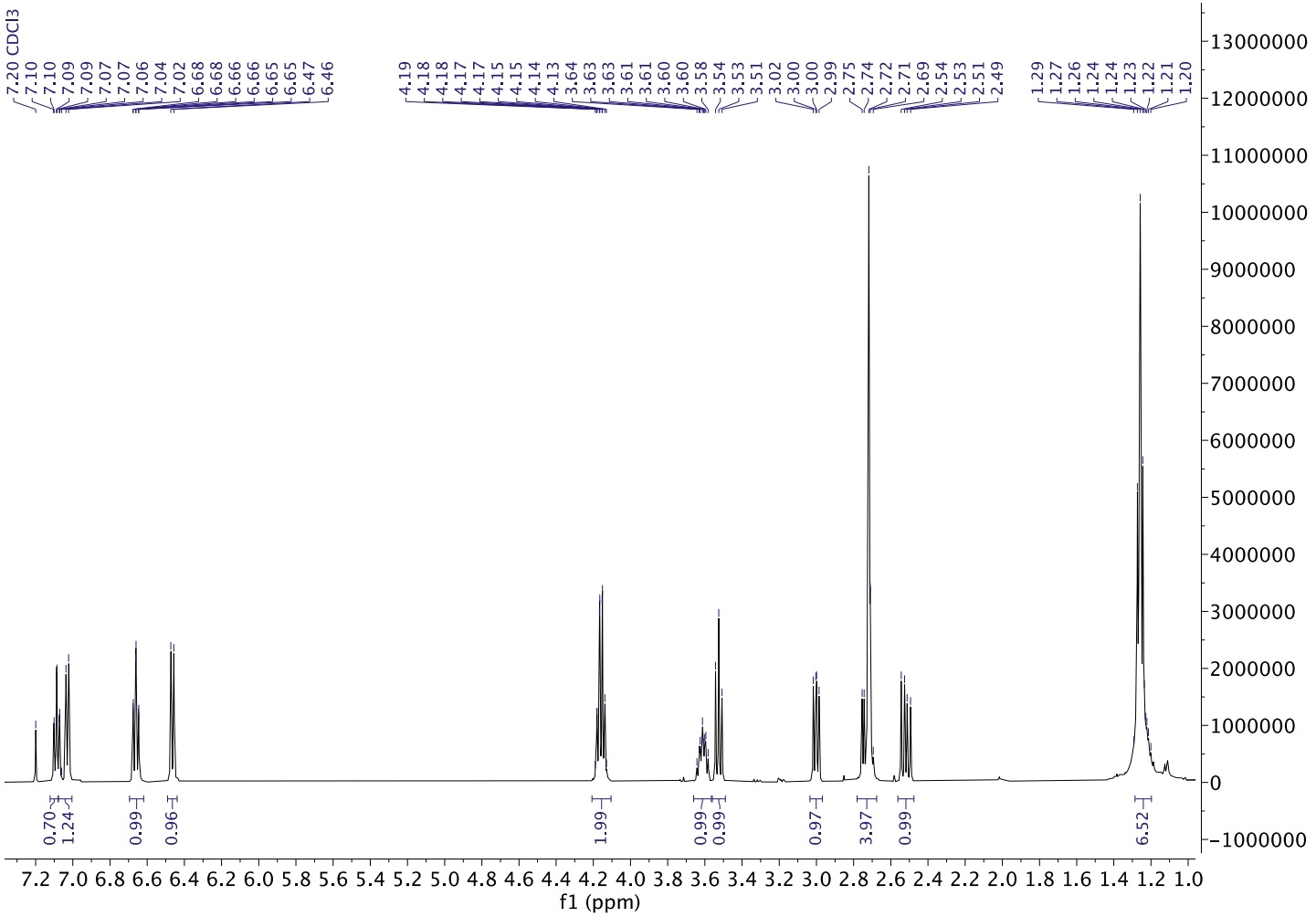




Ethyl (*R*)-2-(1-methylindolin-3-yl)acetate (4a)

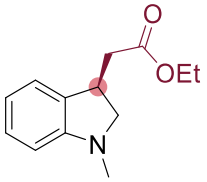
^{1}H NMR (500 MHz, $CDCl_3$)

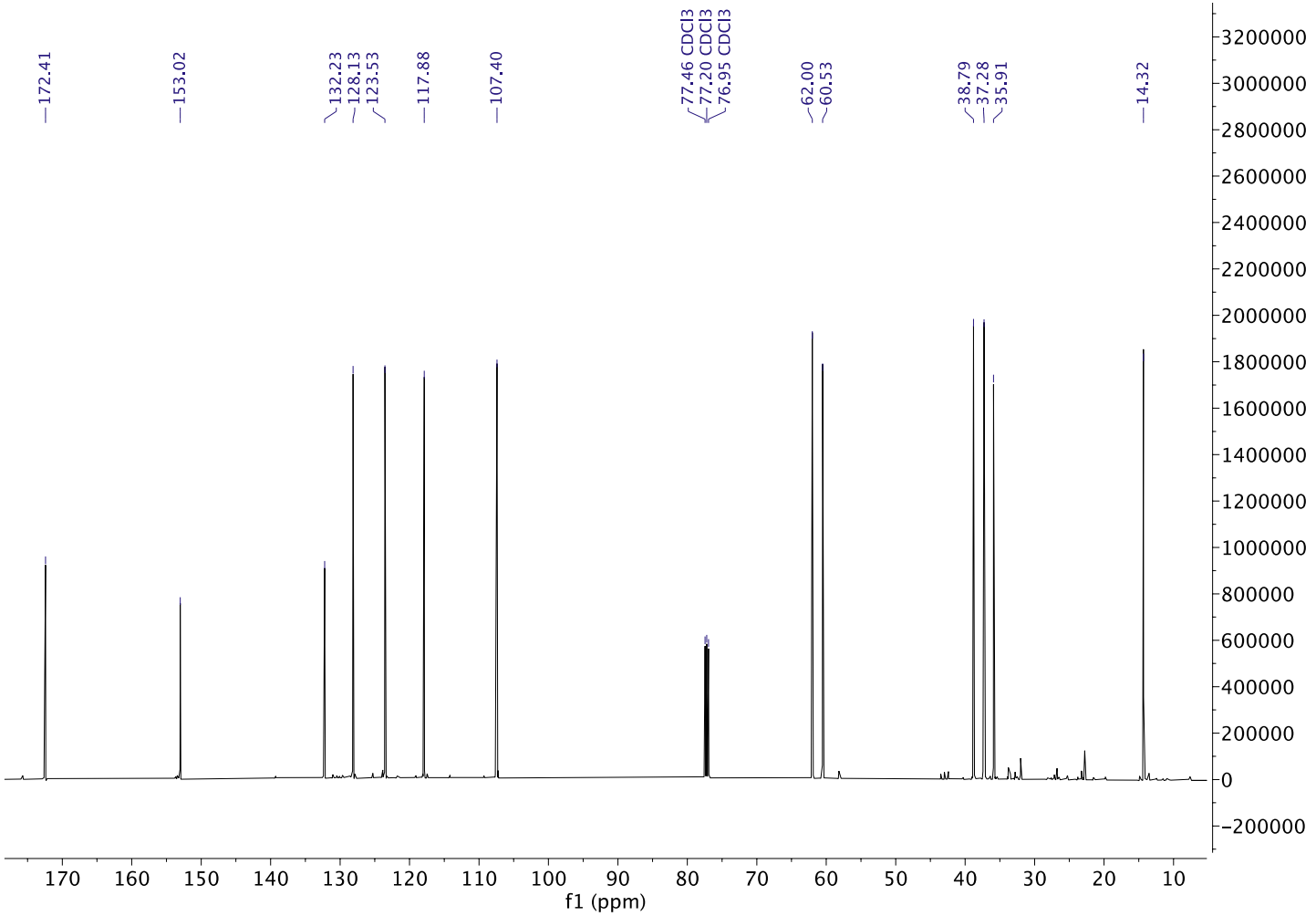




Ethyl (*R*)-2-(1-methylindolin-3-yl)acetate (4a)

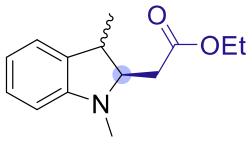
^{13}C NMR (126 MHz, $CDCl_3$)

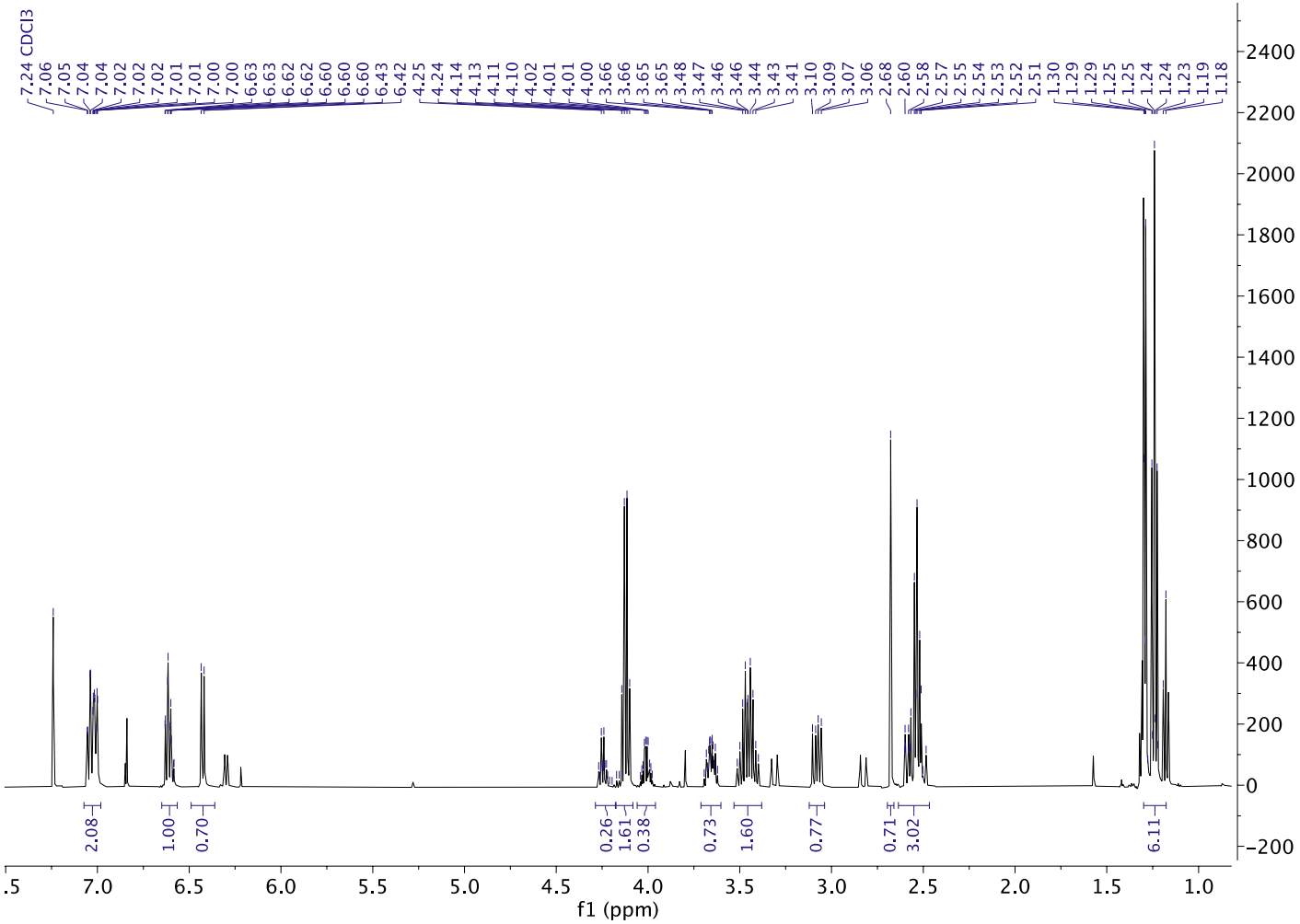




Ethyl 2-((2*R*)-1,3-dimethylindolin-2-yl)acetate (3b)

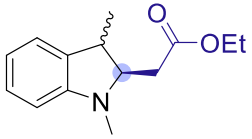
^{1}H NMR (500 MHz, $CDCl_3$)

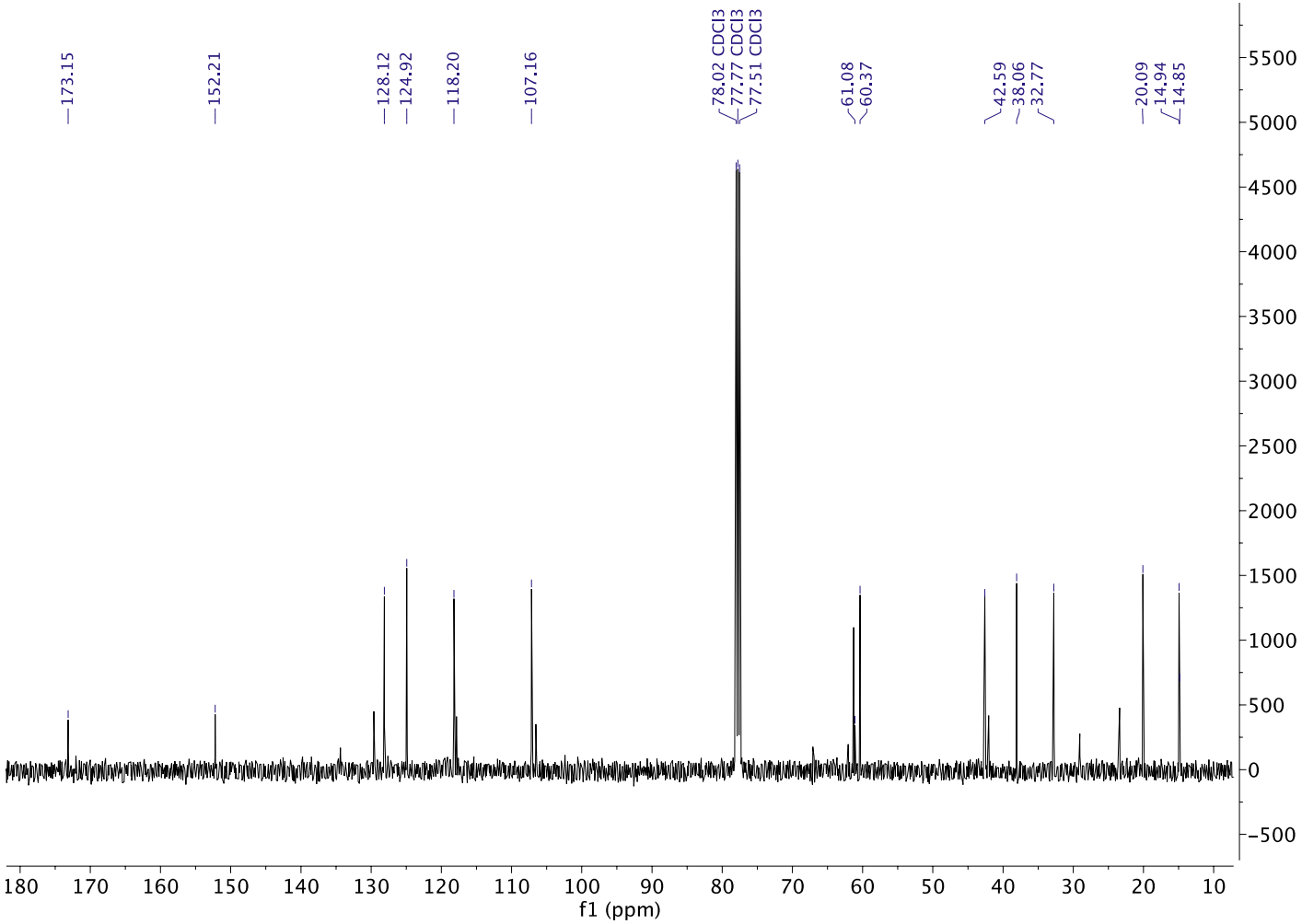




Ethyl 2-((2*R*)-1,3-dimethylindolin-2-yl)acetate (3b)

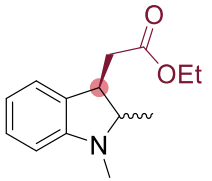
^{13}C NMR (126 MHz, $CDCl_3$)

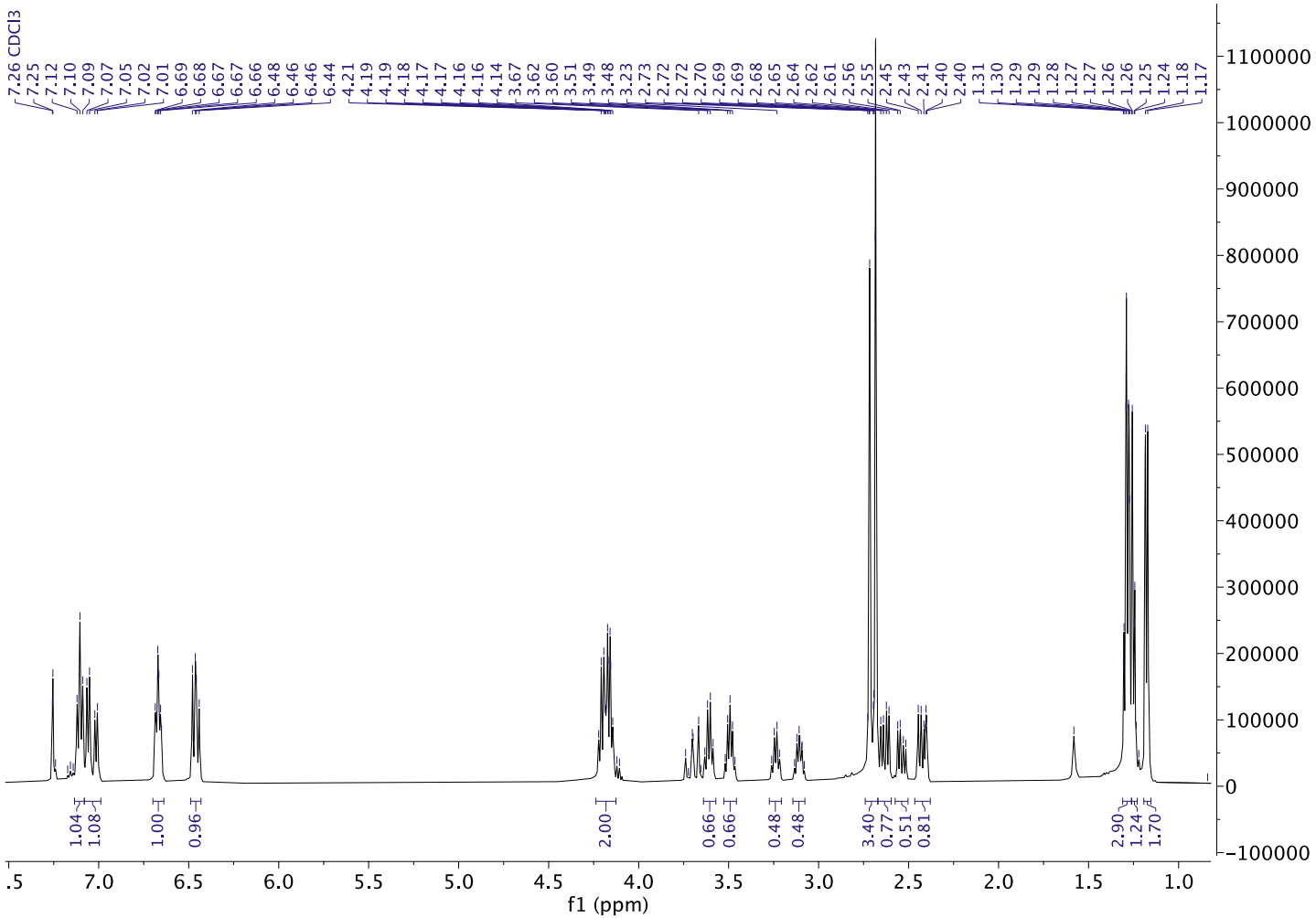




Ethyl 2-((3*R*)-1,2-dimethylindolin-3-yl)acetate (4b)

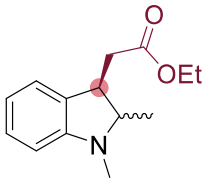
^{1}H NMR (500 MHz, $CDCl_3$)

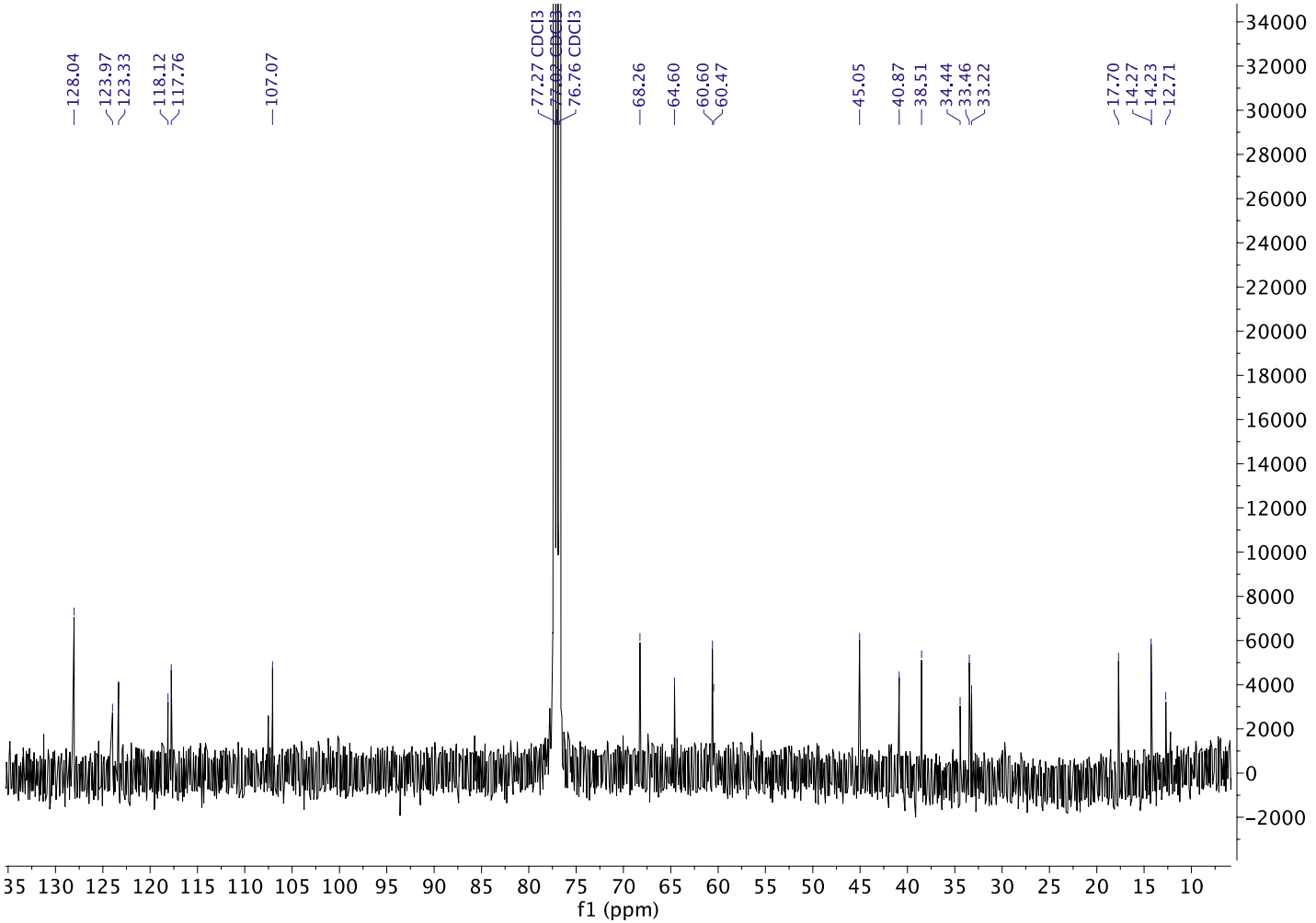




Ethyl 2-((3*R*)-1,2-dimethylindolin-3-yl)acetate (4b)

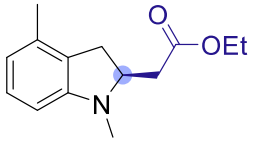
^{13}C NMR (126 MHz, $CDCl_3$)

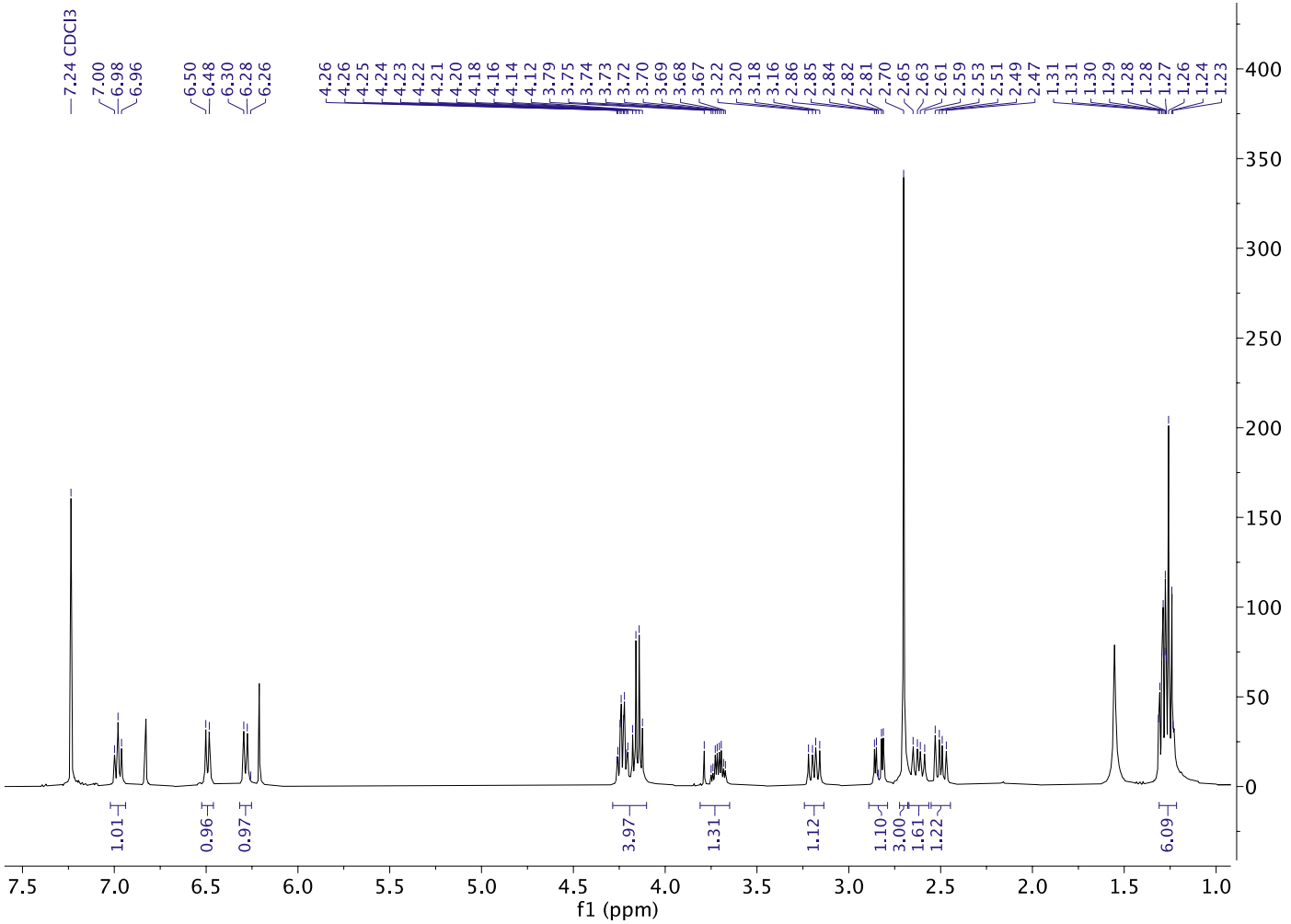




Ethyl (*S*)-2-(1,4-dimethylindolin-2-yl)acetate (3c)

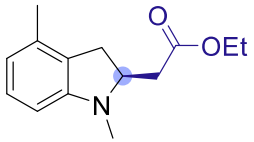
^{1}H NMR (500 MHz, $CDCl_3$)

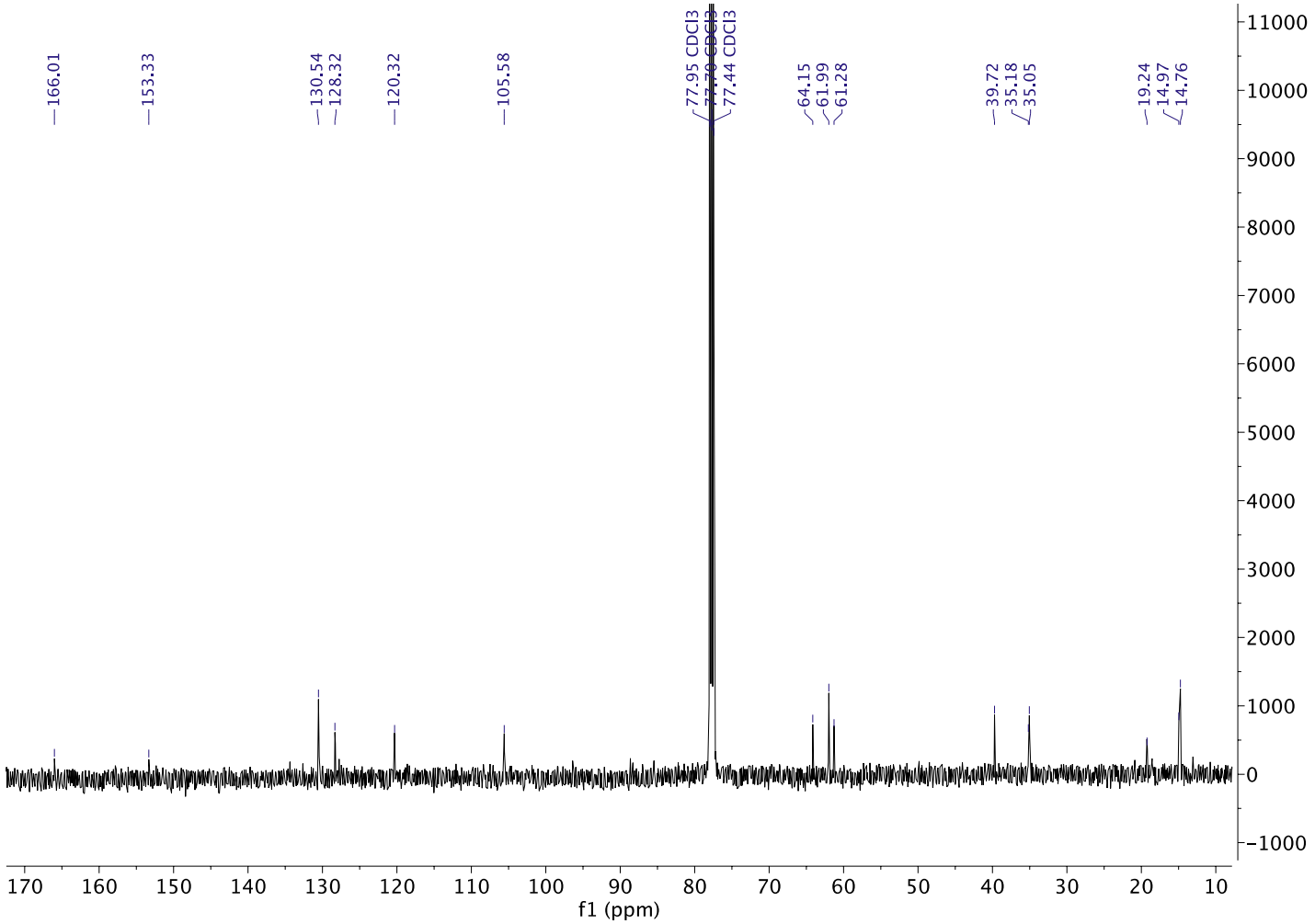




Ethyl (*S*)-2-(1,4-dimethylindolin-2-yl)acetate (3c)

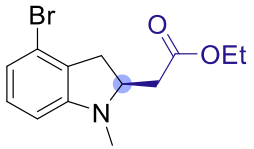
^{13}C NMR (126 MHz, $CDCl_3$)

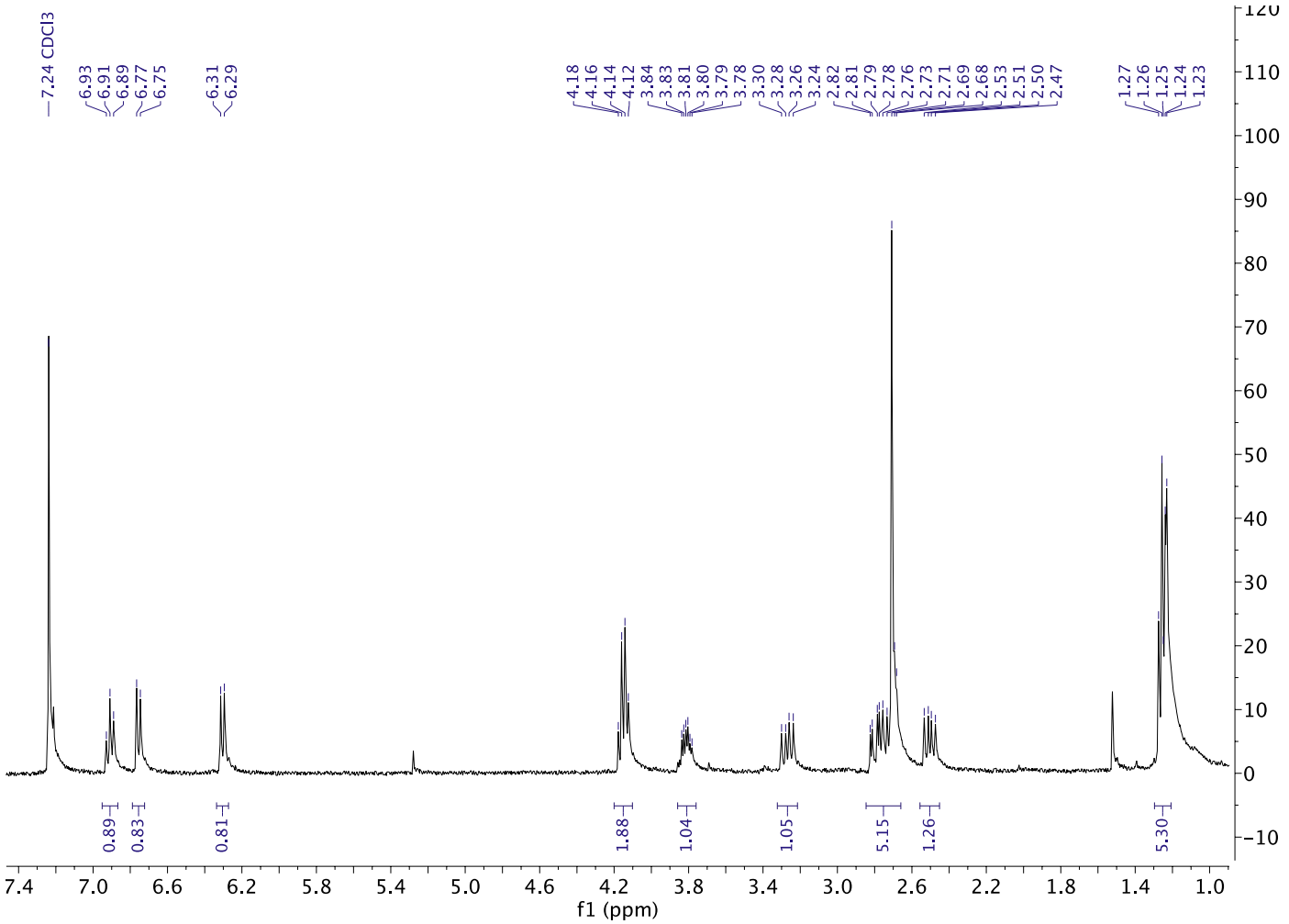




Ethyl (*S*)-2-(4-bromo-1-methylindolin-2-yl)acetate (3d)

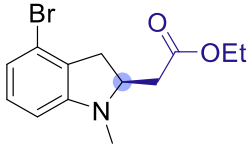
^{1}H NMR (500 MHz, $CDCl_3$)

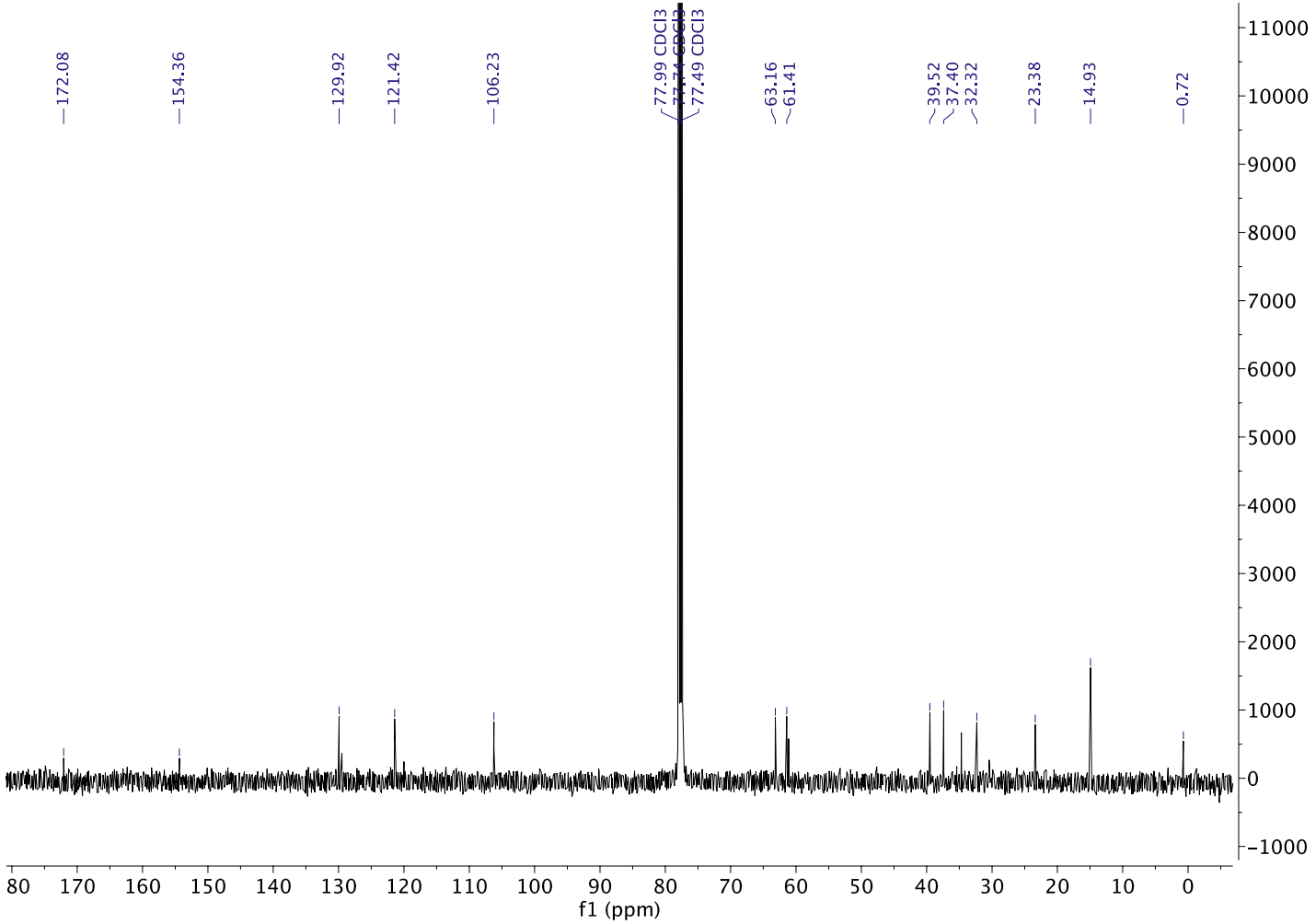




Ethyl (*S*)-2-(4-bromo-1-methylindolin-2-yl)acetate (3d)

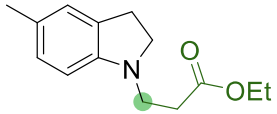
^{13}C NMR (126 MHz, $CDCl_3$)

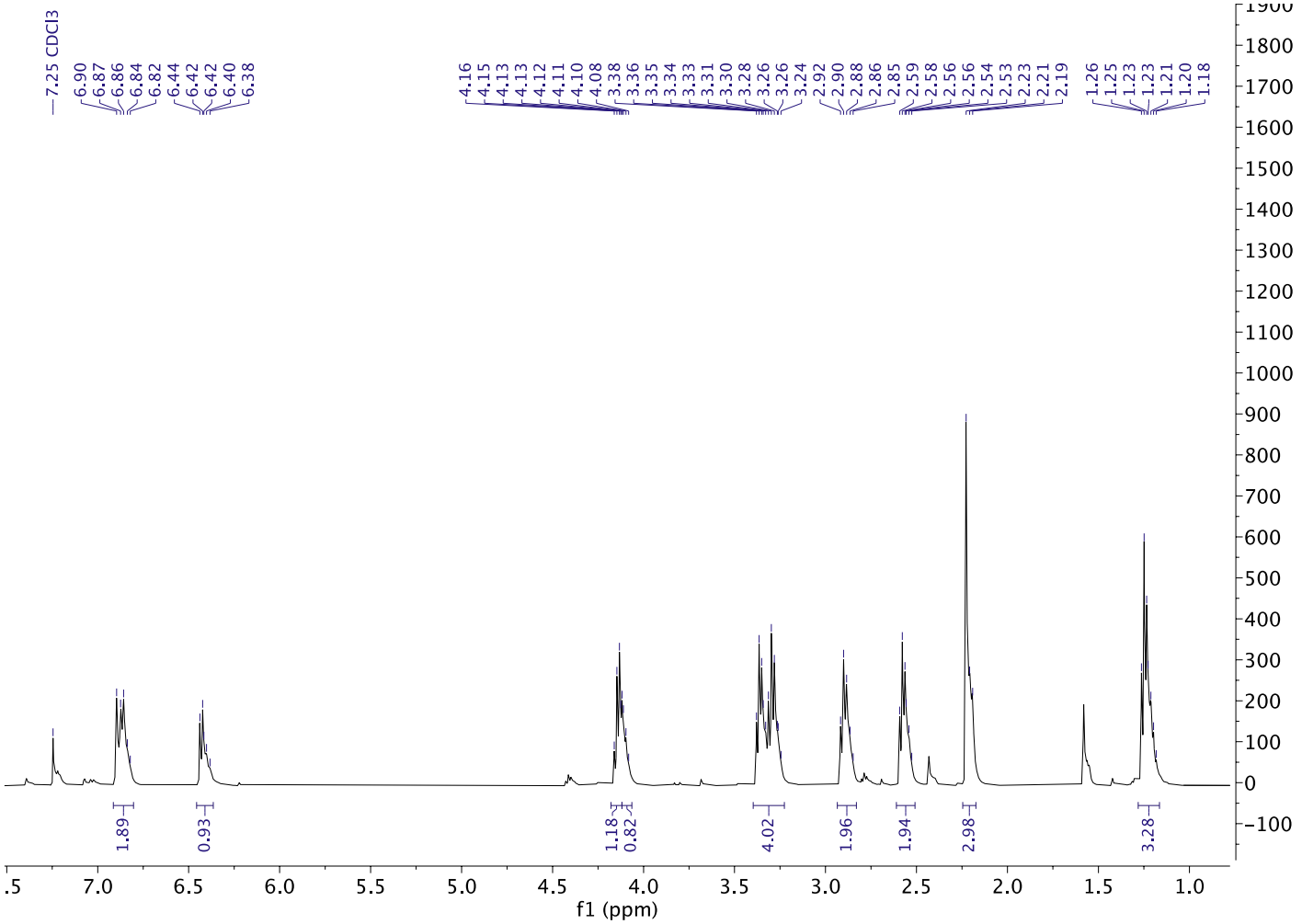




Ethyl 3-(5-methylindolin-1-yl)propanoate (5e)

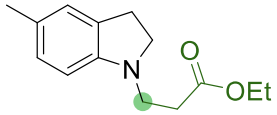
^{1}H NMR (500 MHz, $CDCl_3$)

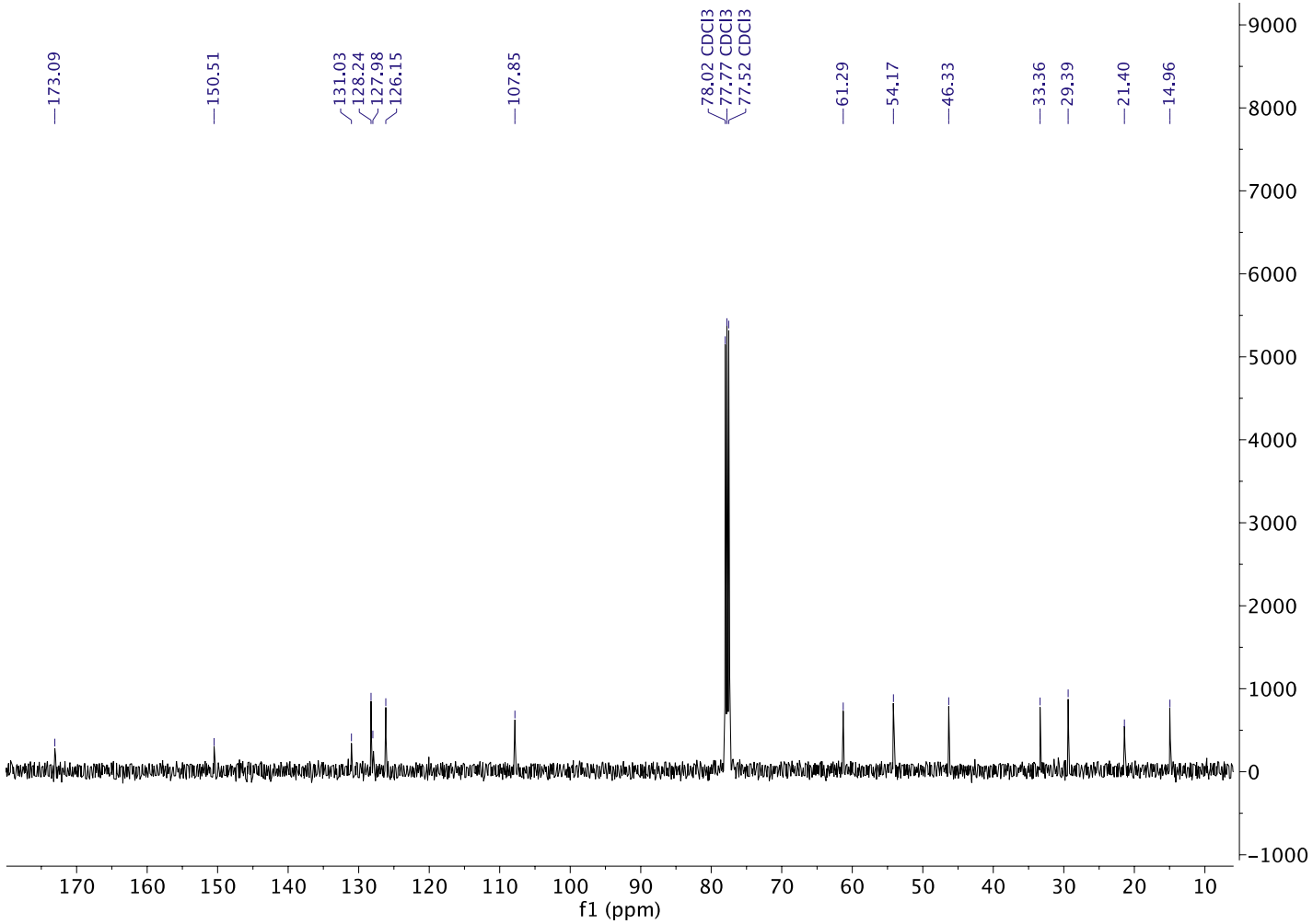




Ethyl 3-(5-methylindolin-1-yl)propanoate (5e)

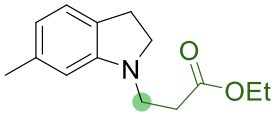
^{13}C NMR (126 MHz, $CDCl_3$)

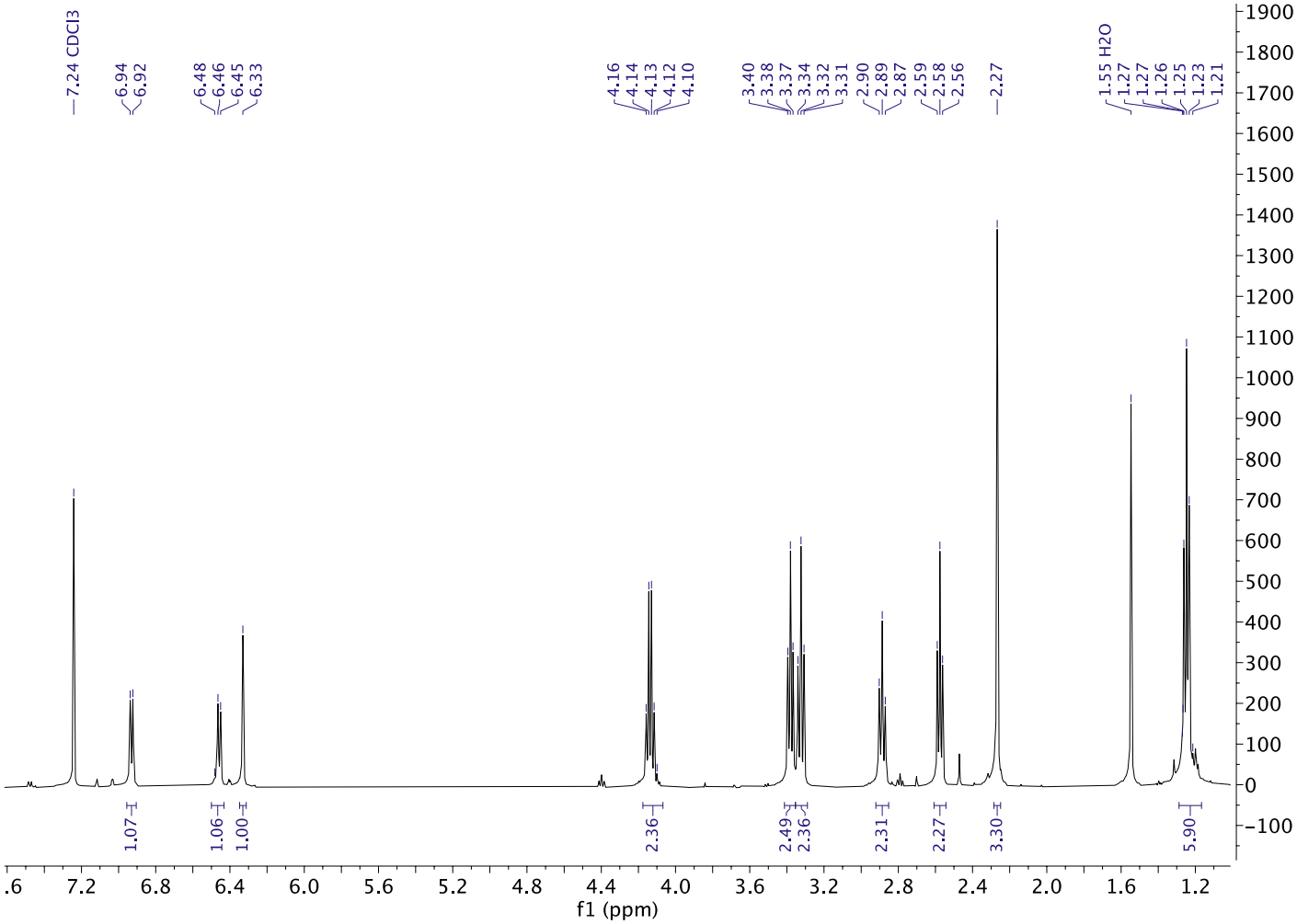




Ethyl 3-(6-methylindolin-1-yl)propanoate (5f)

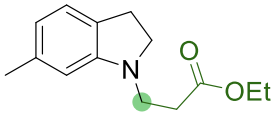
^{1}H NMR (500 MHz, $CDCl_3$)

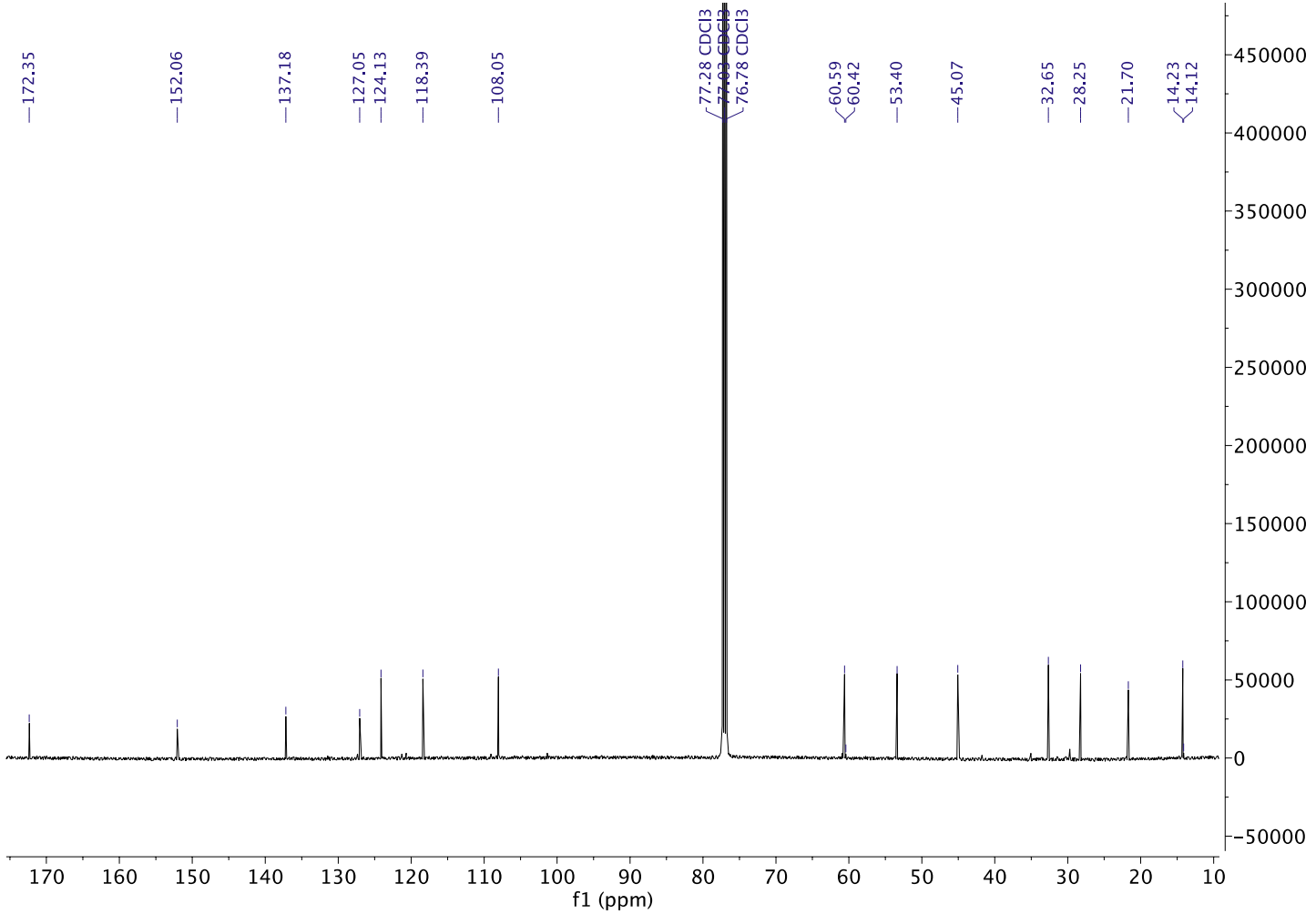




Ethyl 3-(6-methylindolin-1-yl)propanoate (5f)

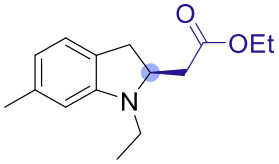
^{13}C NMR (126 MHz, $CDCl_3$)

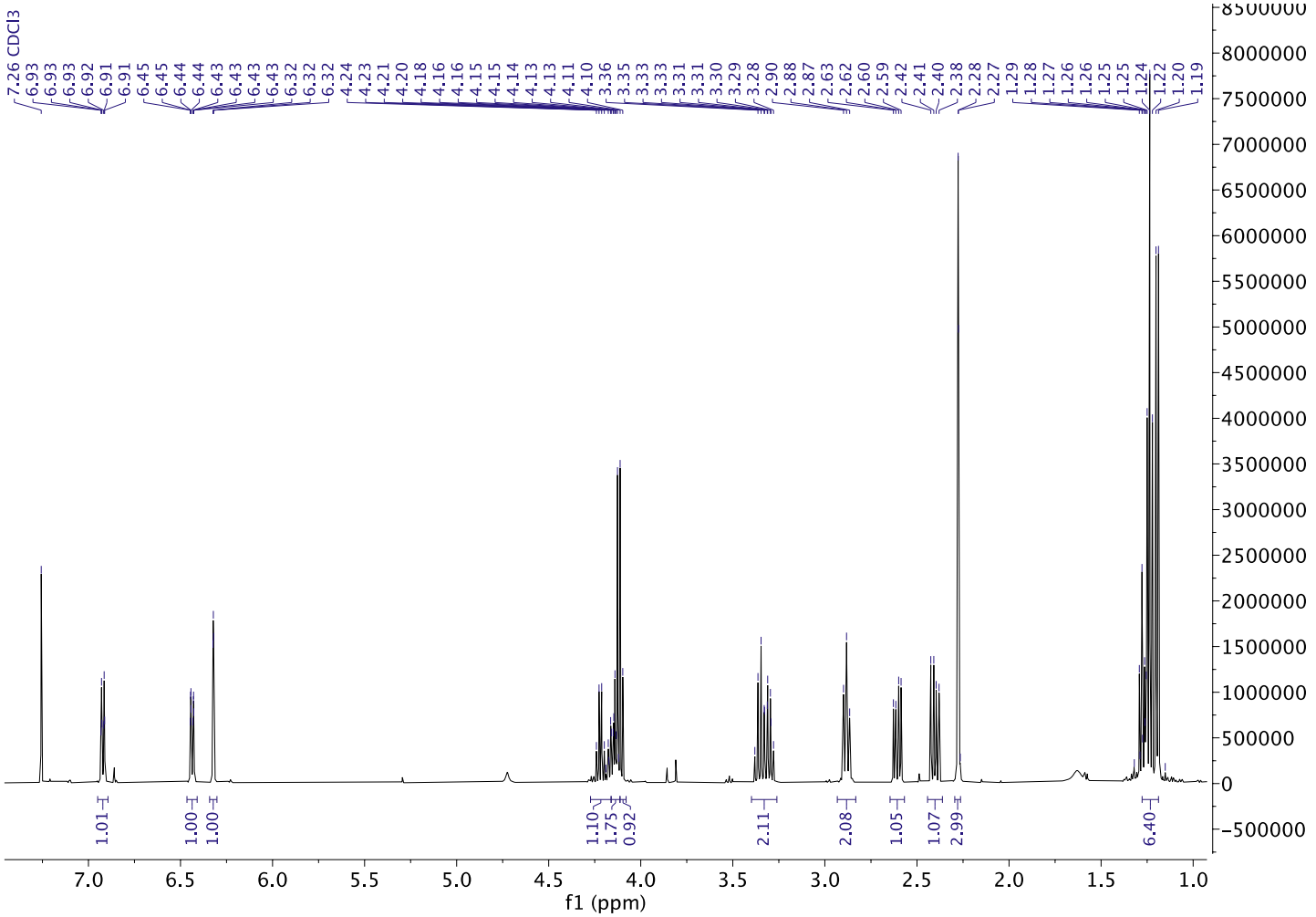




Ethyl (*S*)-2-(1-ethyl-6-methylindolin-2-yl)acetate (3f)

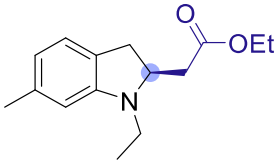
^{1}H NMR (500 MHz, $CDCl_3$)

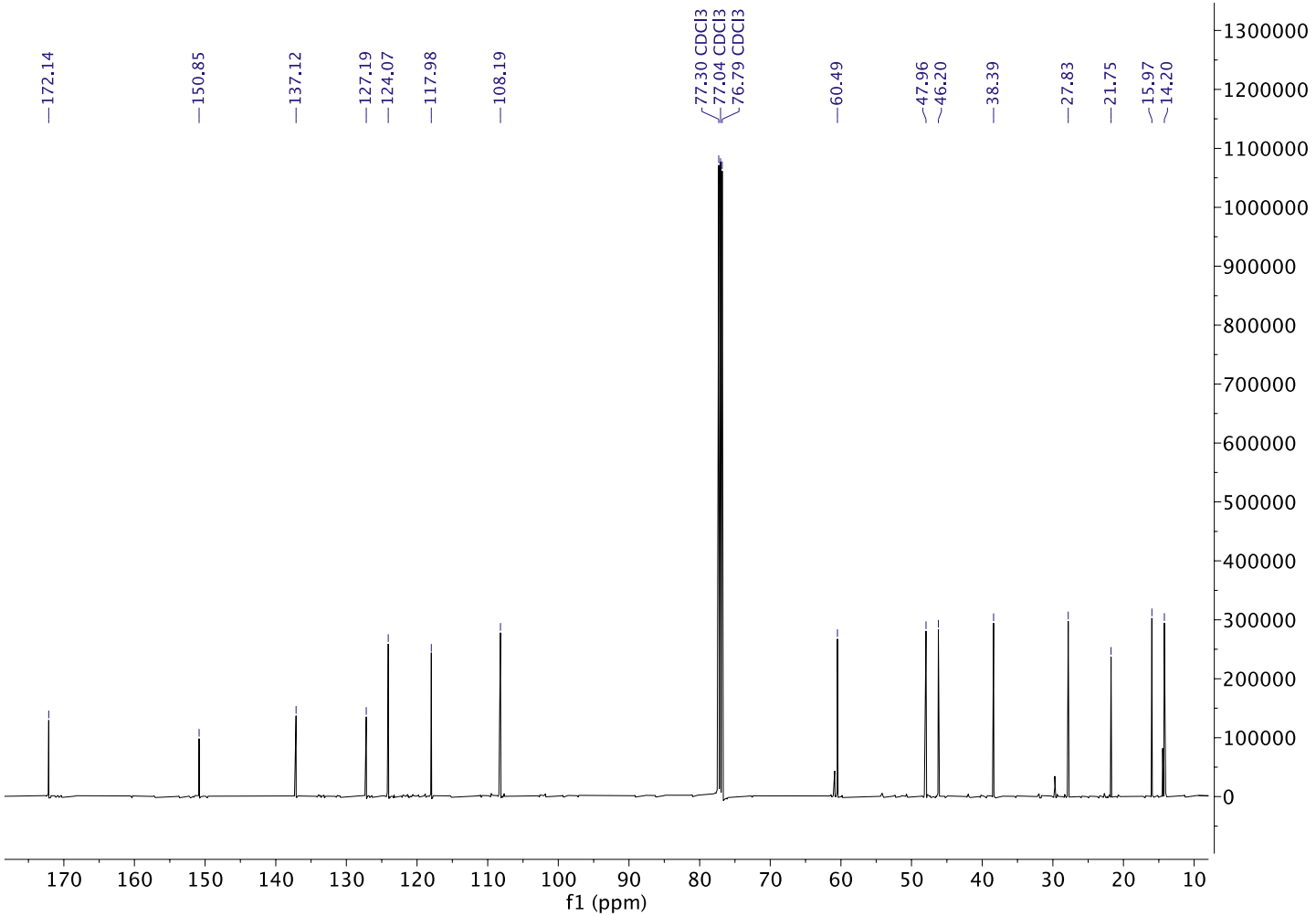




Ethyl (*S*)-2-(1-ethyl-6-methylindolin-2-yl)acetate (3f)

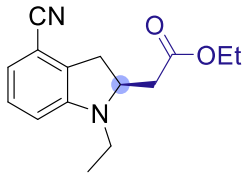
^{13}C NMR (126 MHz, $CDCl_3$)

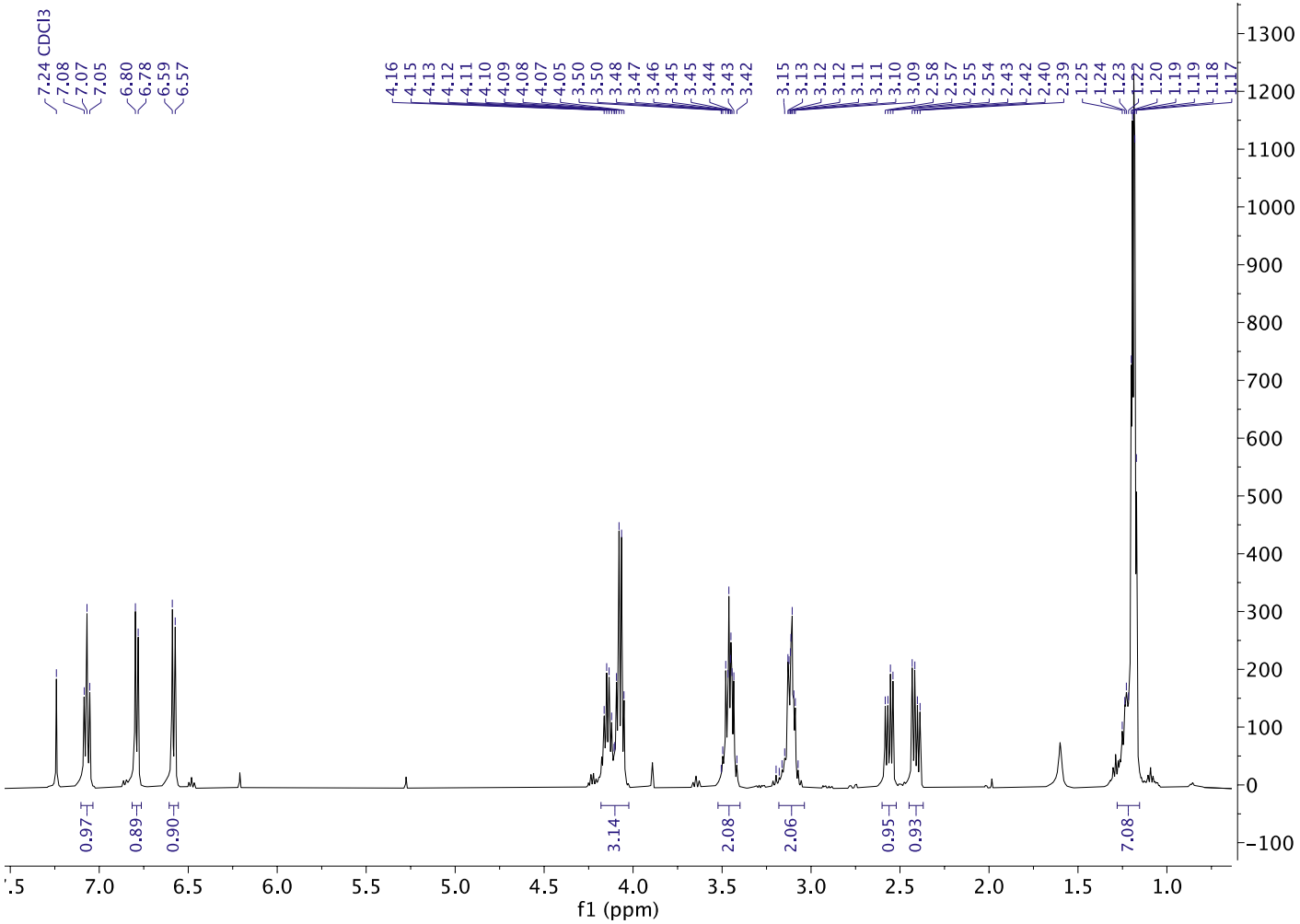




Ethyl (*S*)-2-(4-cyano-1-ethylindolin-2-yl)acetate (3g)

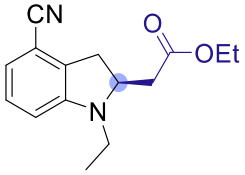
^{1}H NMR (500 MHz, $CDCl_3$)

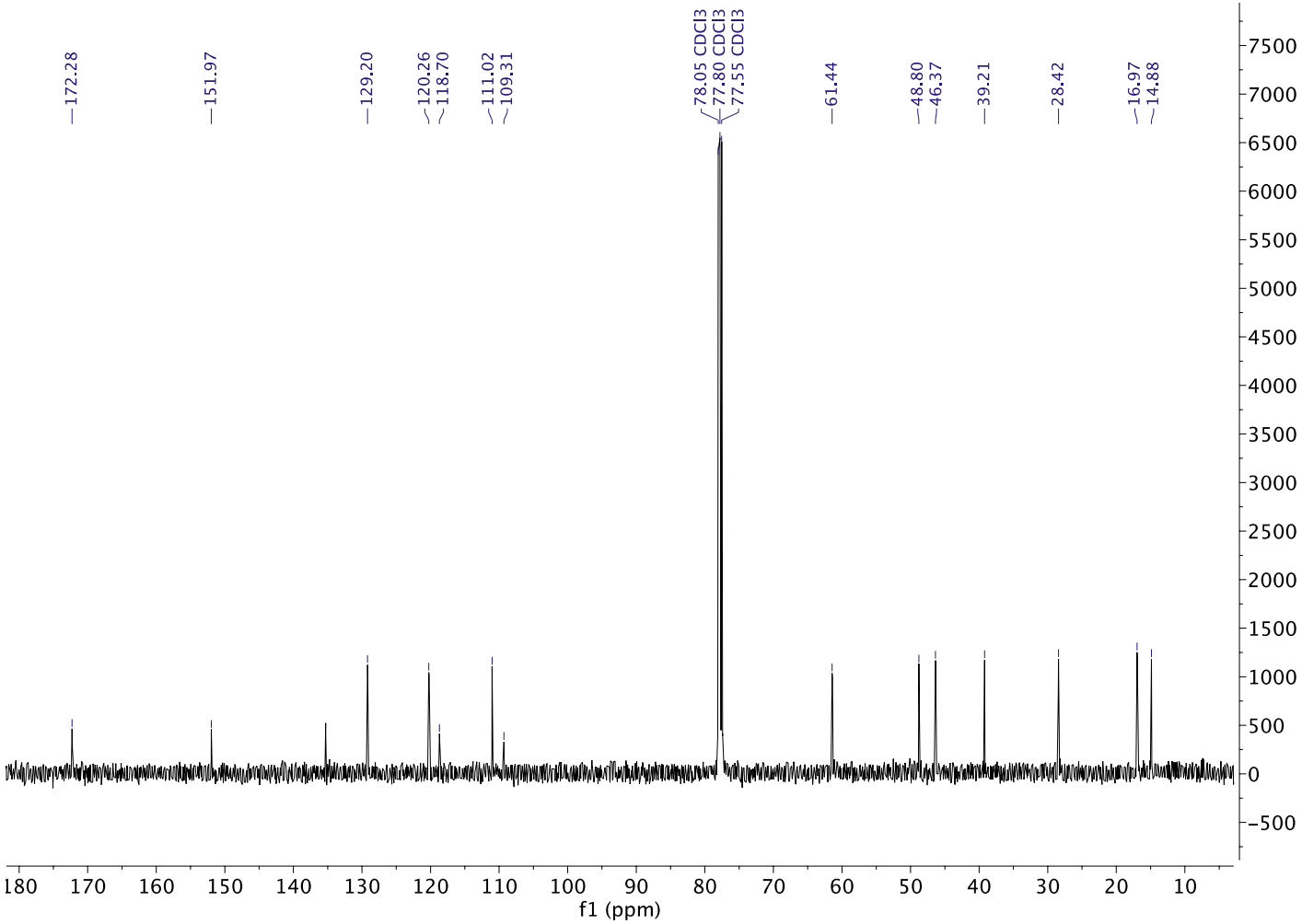




Ethyl (*S*)-2-(4-cyano-1-ethylindolin-2-yl)acetate (3g)

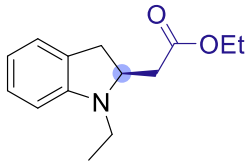
^{13}C NMR (126 MHz, $CDCl_3$)

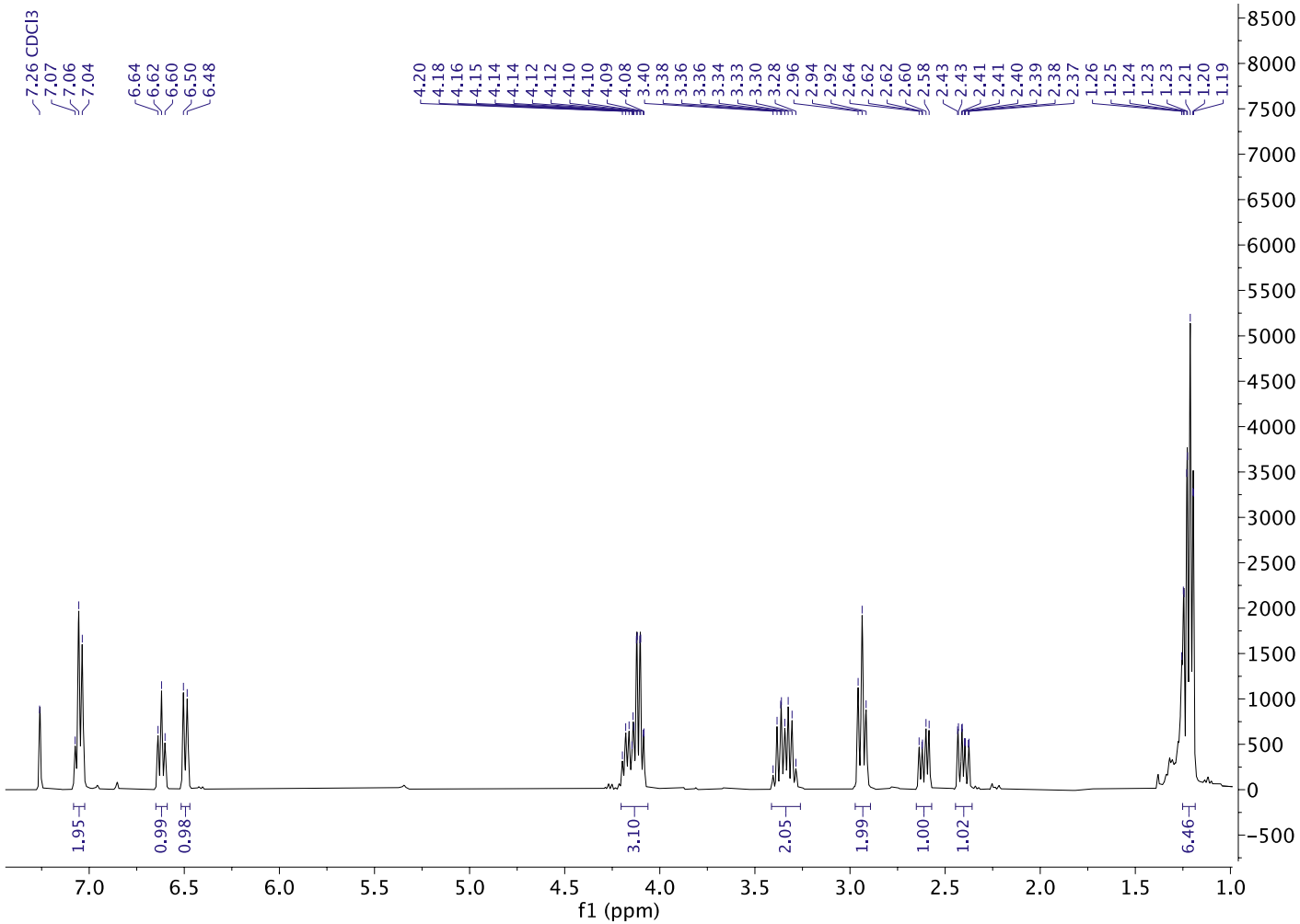




Ethyl (*S*)-2-(1-ethylindolin-2-yl)acetate (3h)

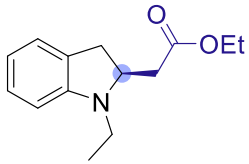
^{1}H NMR (500 MHz, $CDCl_3$)

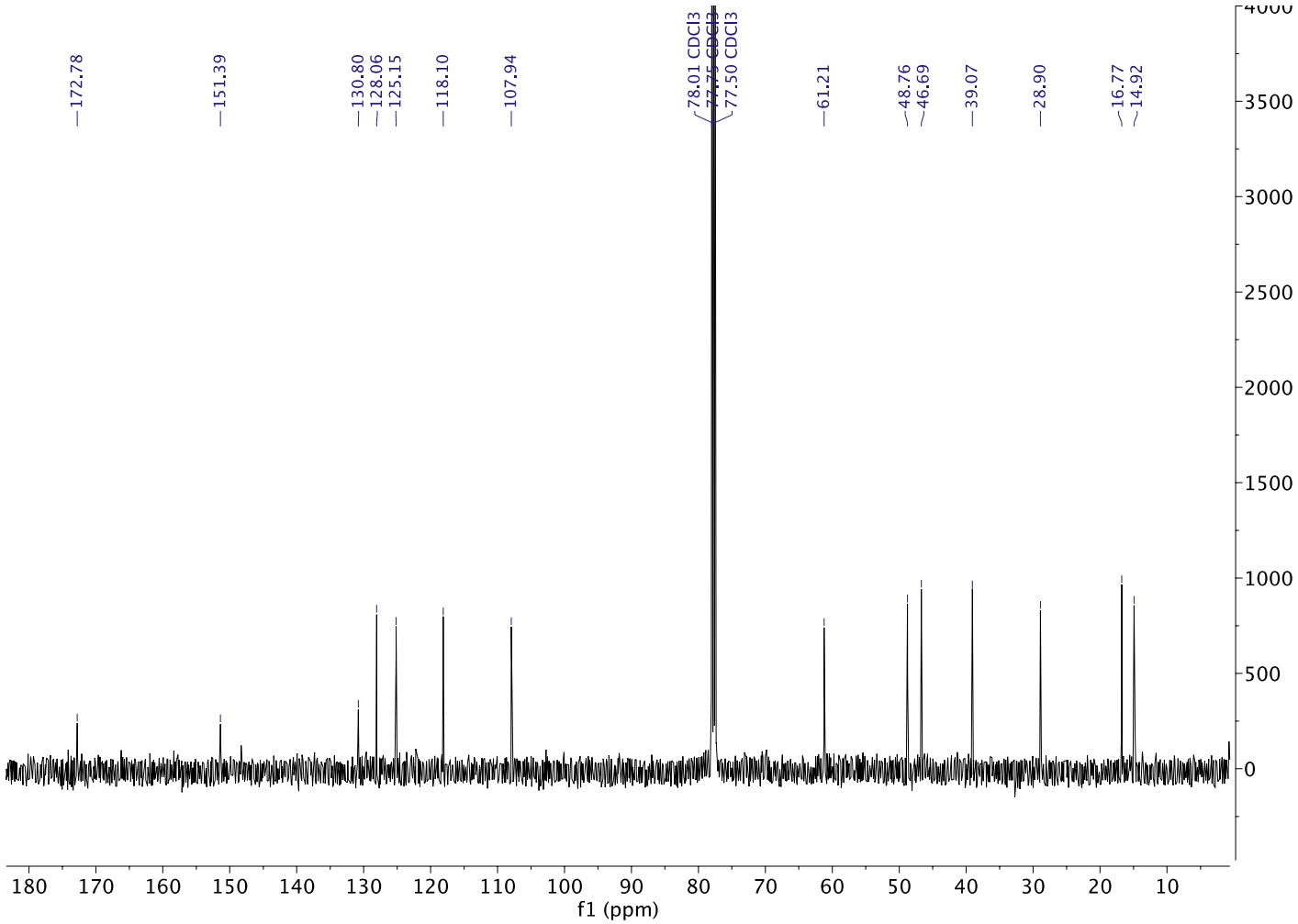




Ethyl (*S*)-2-(1-ethylindolin-2-yl)acetate (3h)

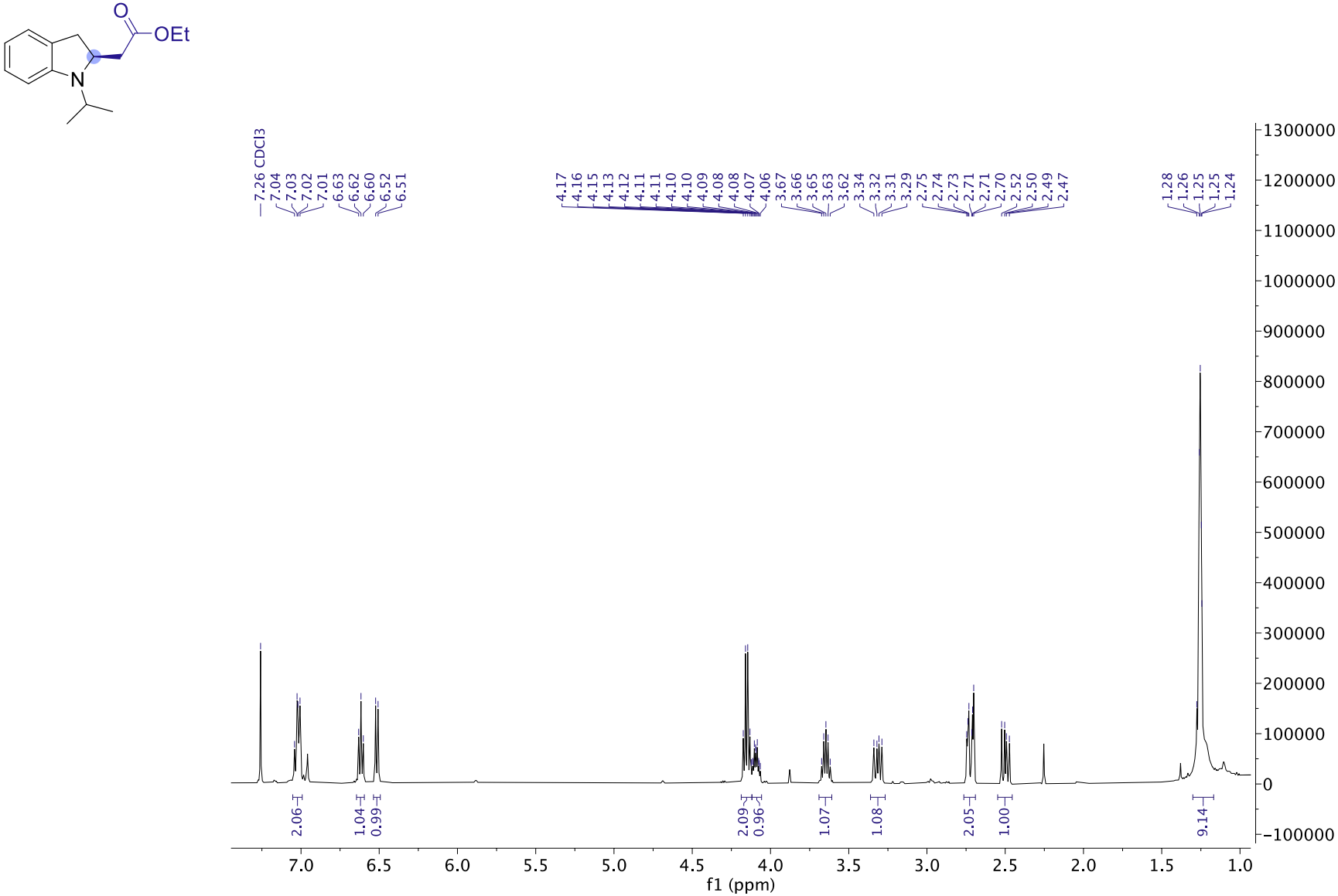
^{13}C NMR (126 MHz, $CDCl_3$)





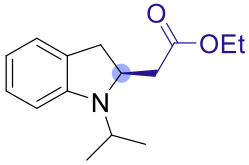
Ethyl (*S*)-2-(1-isopropylindolin-2-yl)acetate (3i)

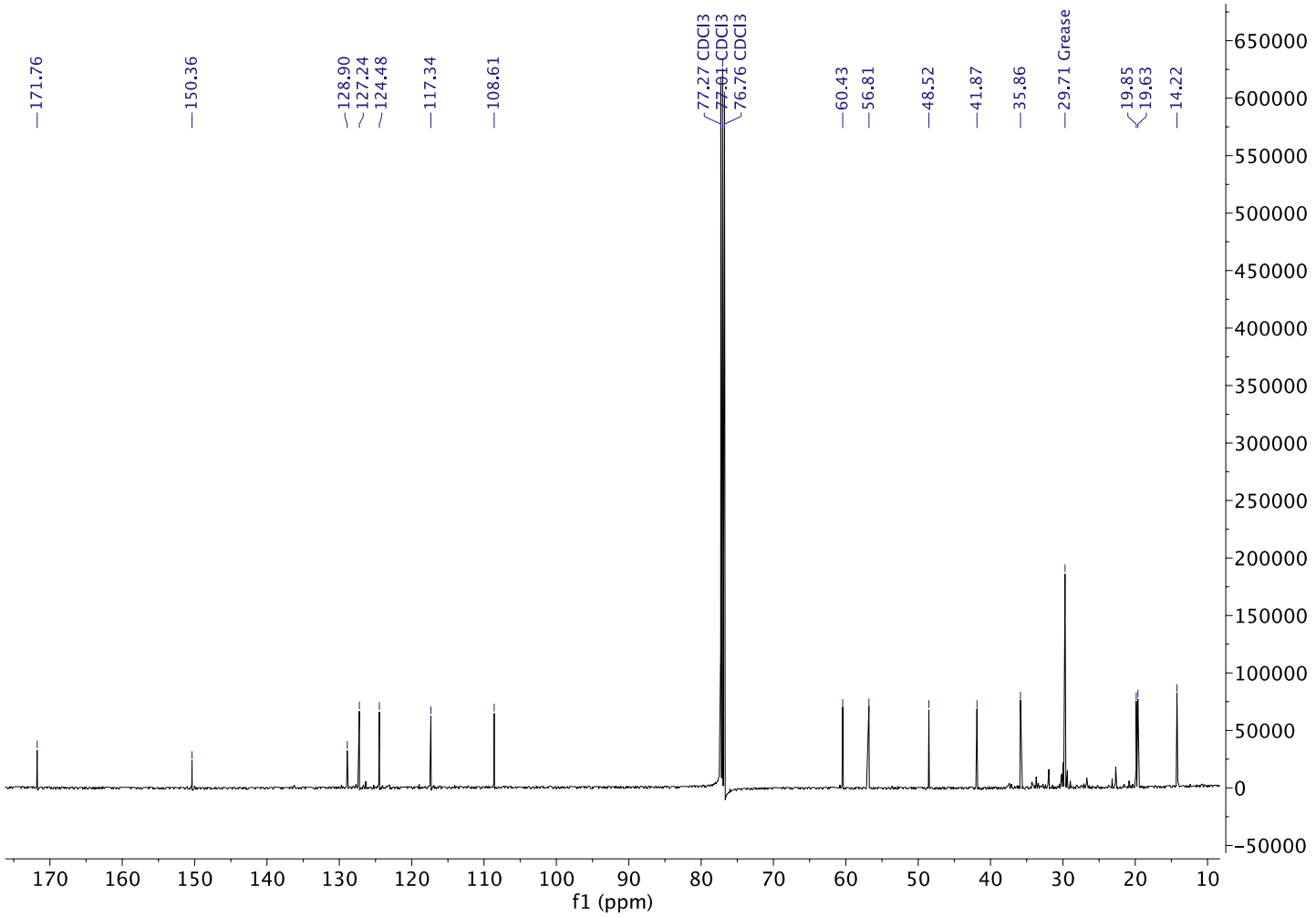
^{1}H NMR (500 MHz, $CDCl_3$)



Ethyl (*S*)-2-(1-isopropylindolin-2-yl)acetate (3i)

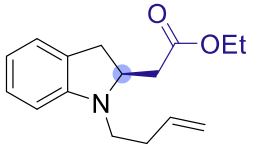
^{13}C NMR (126 MHz, $CDCl_3$)

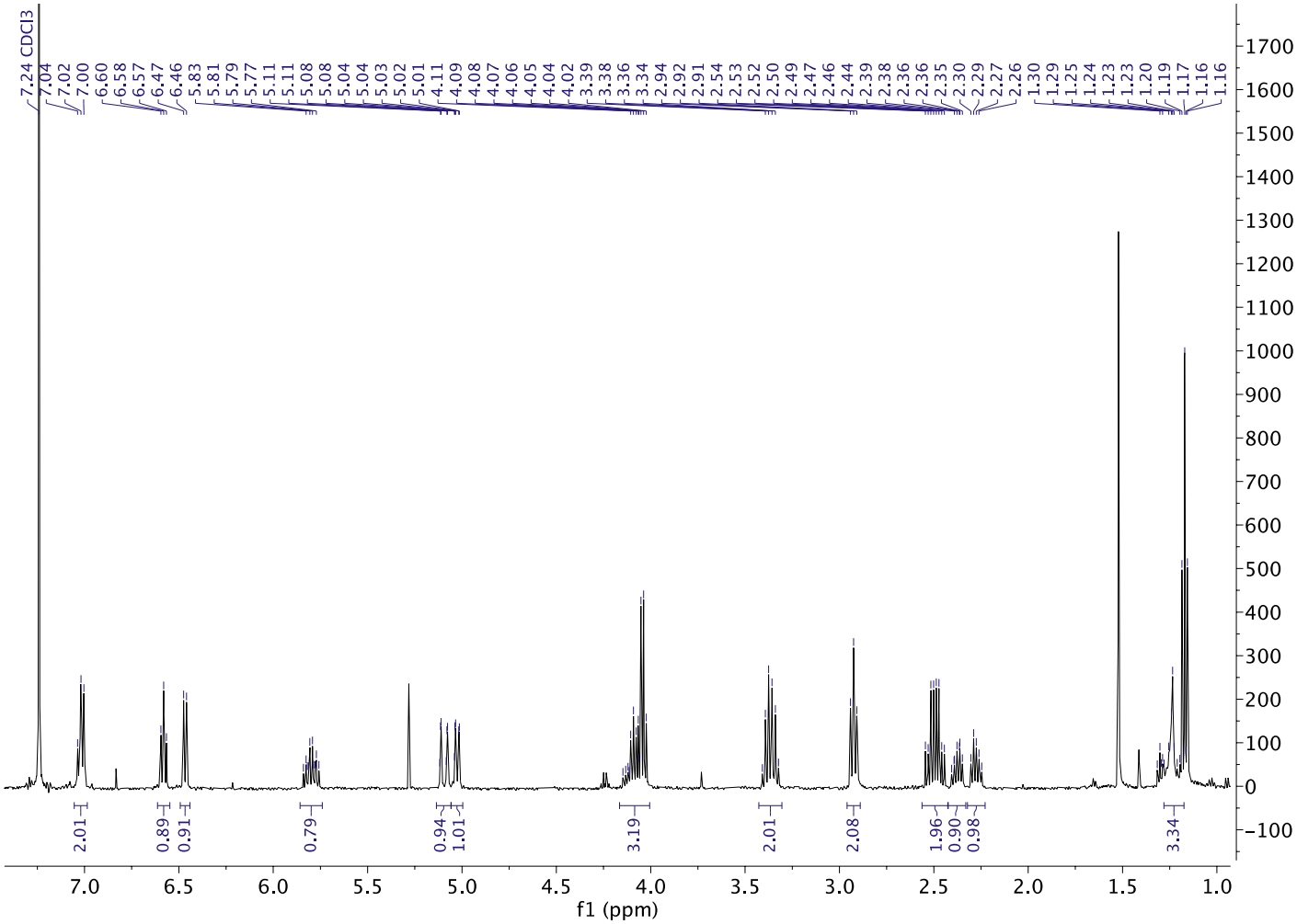




Ethyl (*S*)-2-(1-(but-3-en-1-yl)indolin-2-yl)acetate (3j)

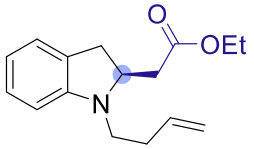
^{1}H NMR (500 MHz, $CDCl_3$)

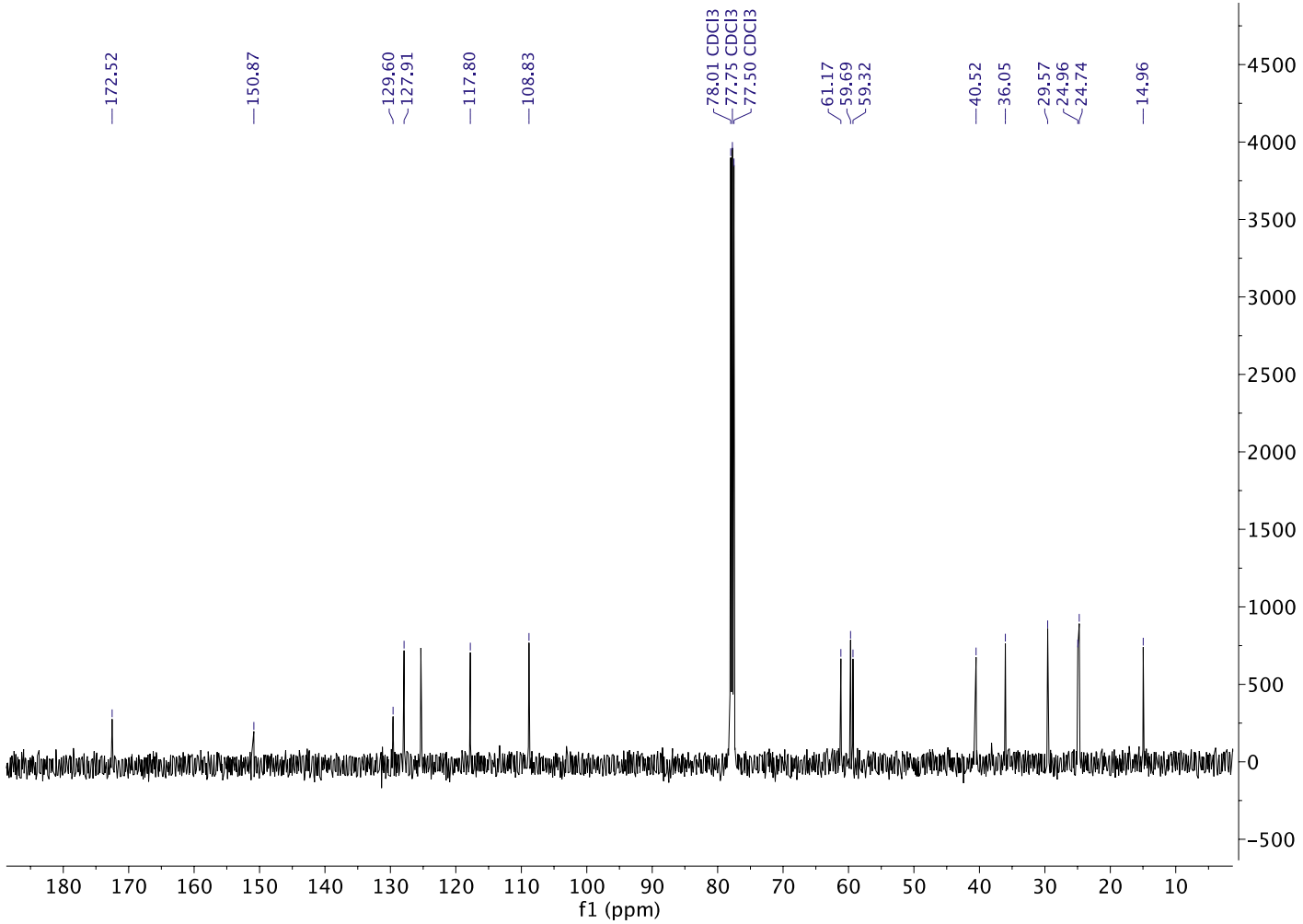




Ethyl (*S*)-2-(1-(but-3-en-1-yl)indolin-2-yl)acetate (3j)

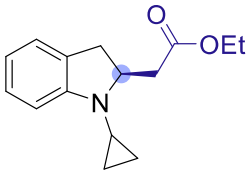
^{13}C NMR (126 MHz, $CDCl_3$)

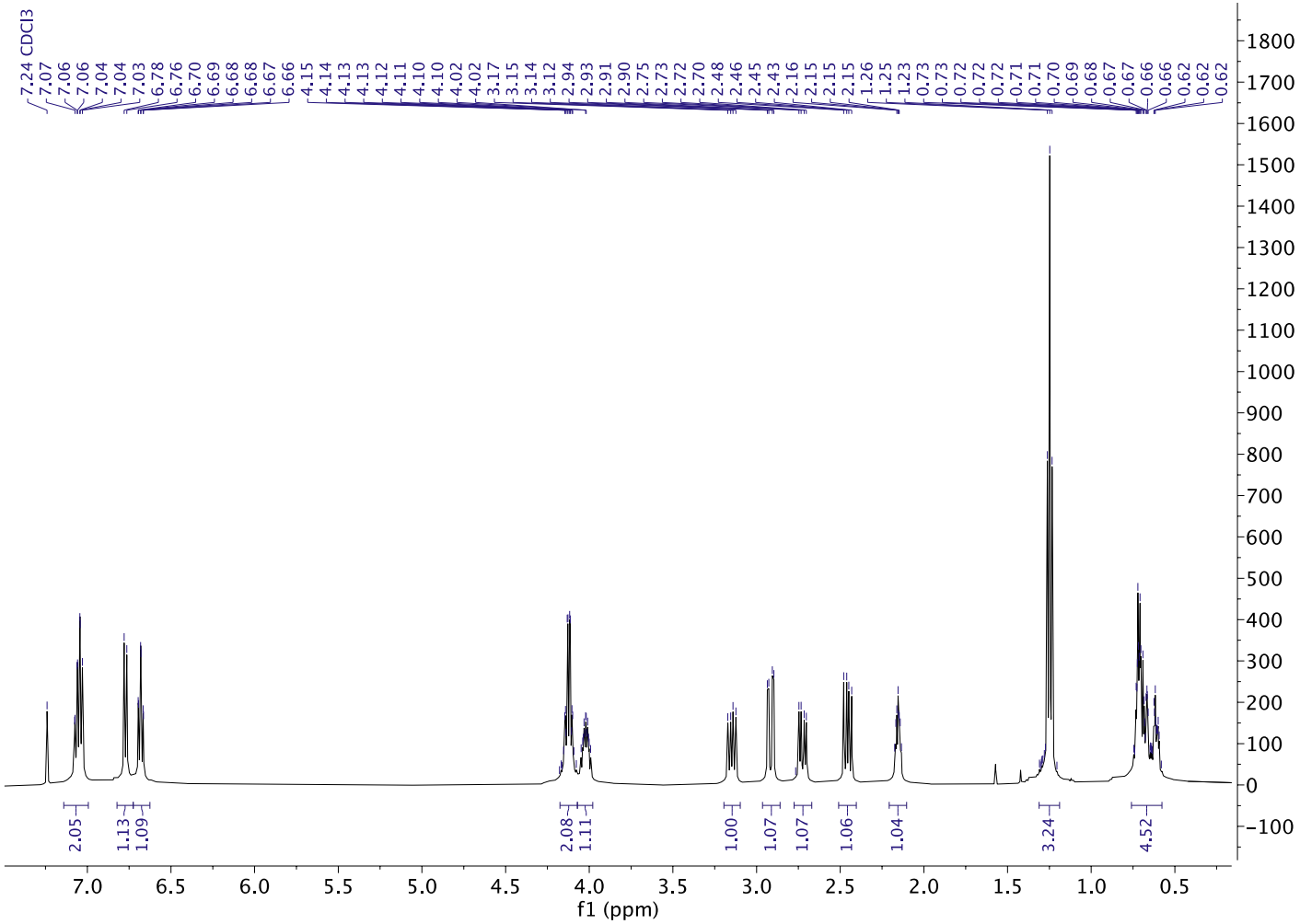




Ethyl (*S*)-2-(1-cyclopropylindolin-2-yl)acetate (3k)

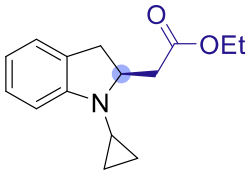
^{1}H NMR (500 MHz, $CDCl_3$)

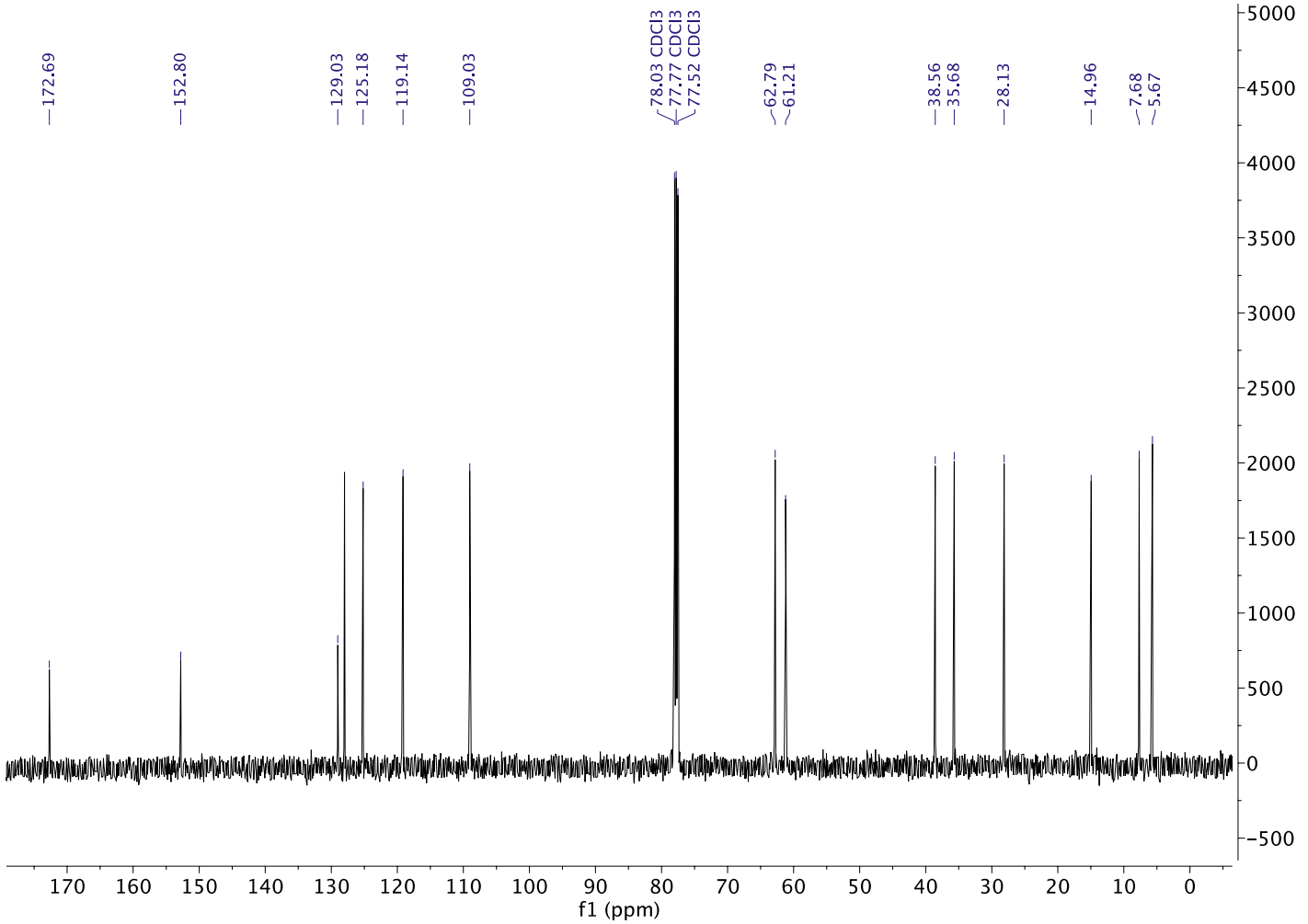




Ethyl (*S*)-2-(1-cyclopropylindolin-2-yl)acetate (3k)

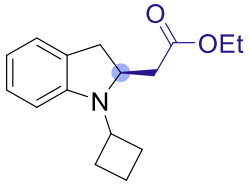
^{13}C NMR (126 MHz, $CDCl_3$)

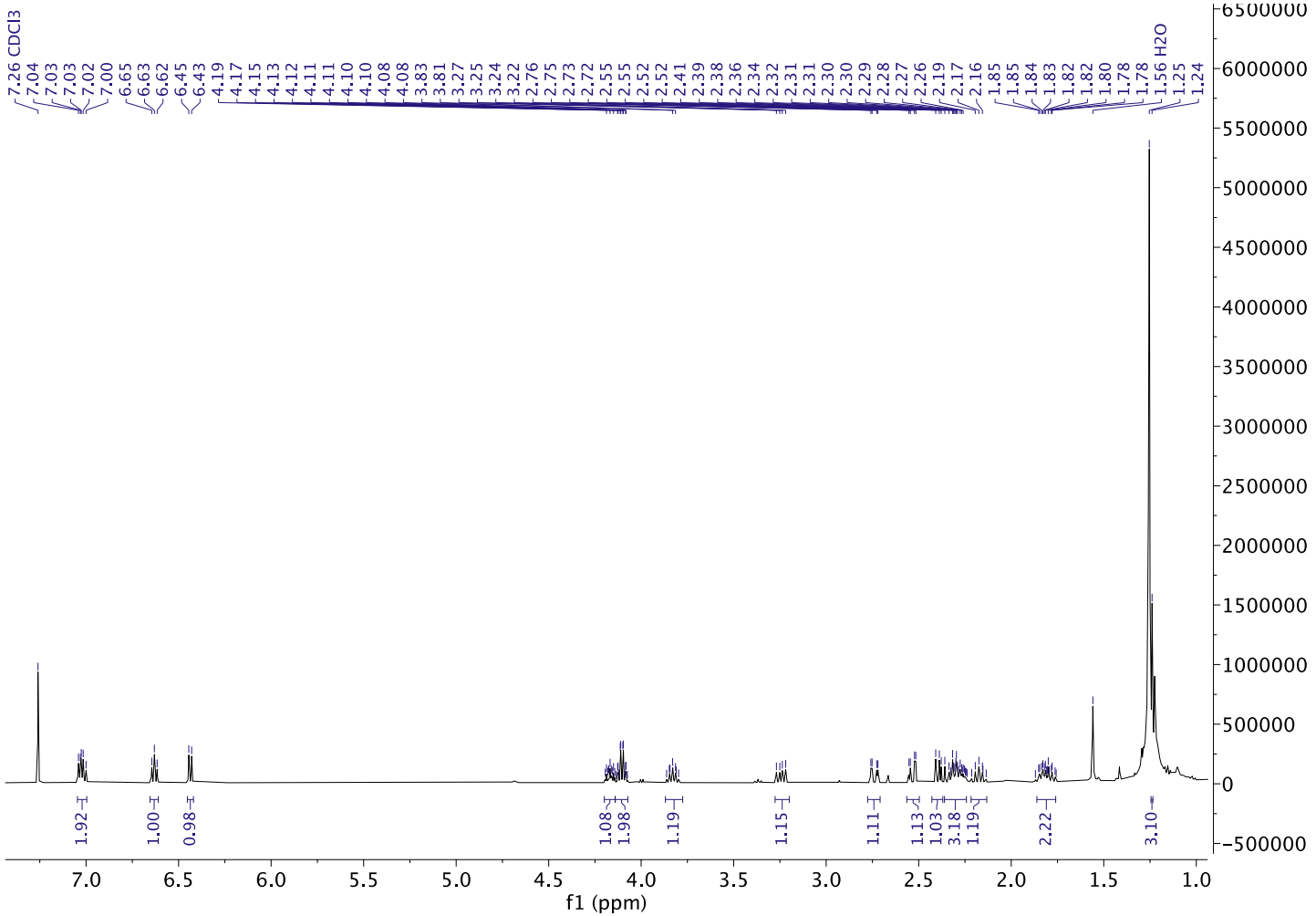




Ethyl (*S*)-2-(1-cyclobutylindolin-2-yl)acetate (3l)

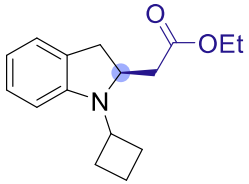
^{1}H NMR (500 MHz, $CDCl_3$)

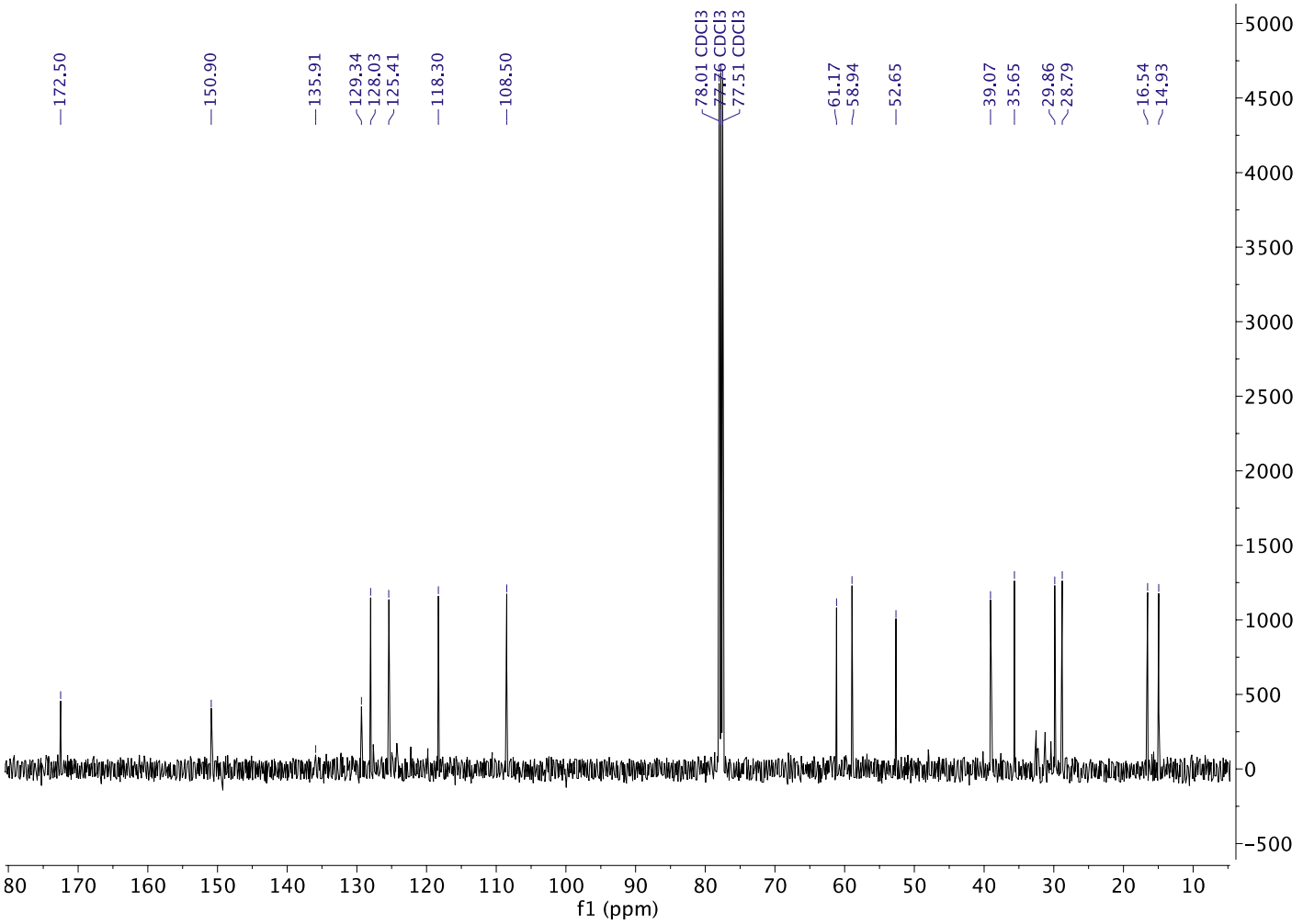




Ethyl (*S*)-2-(1-cyclobutylindolin-2-yl)acetate (3l)

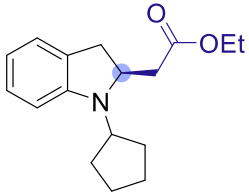
^{13}C NMR (126 MHz, $CDCl_3$)

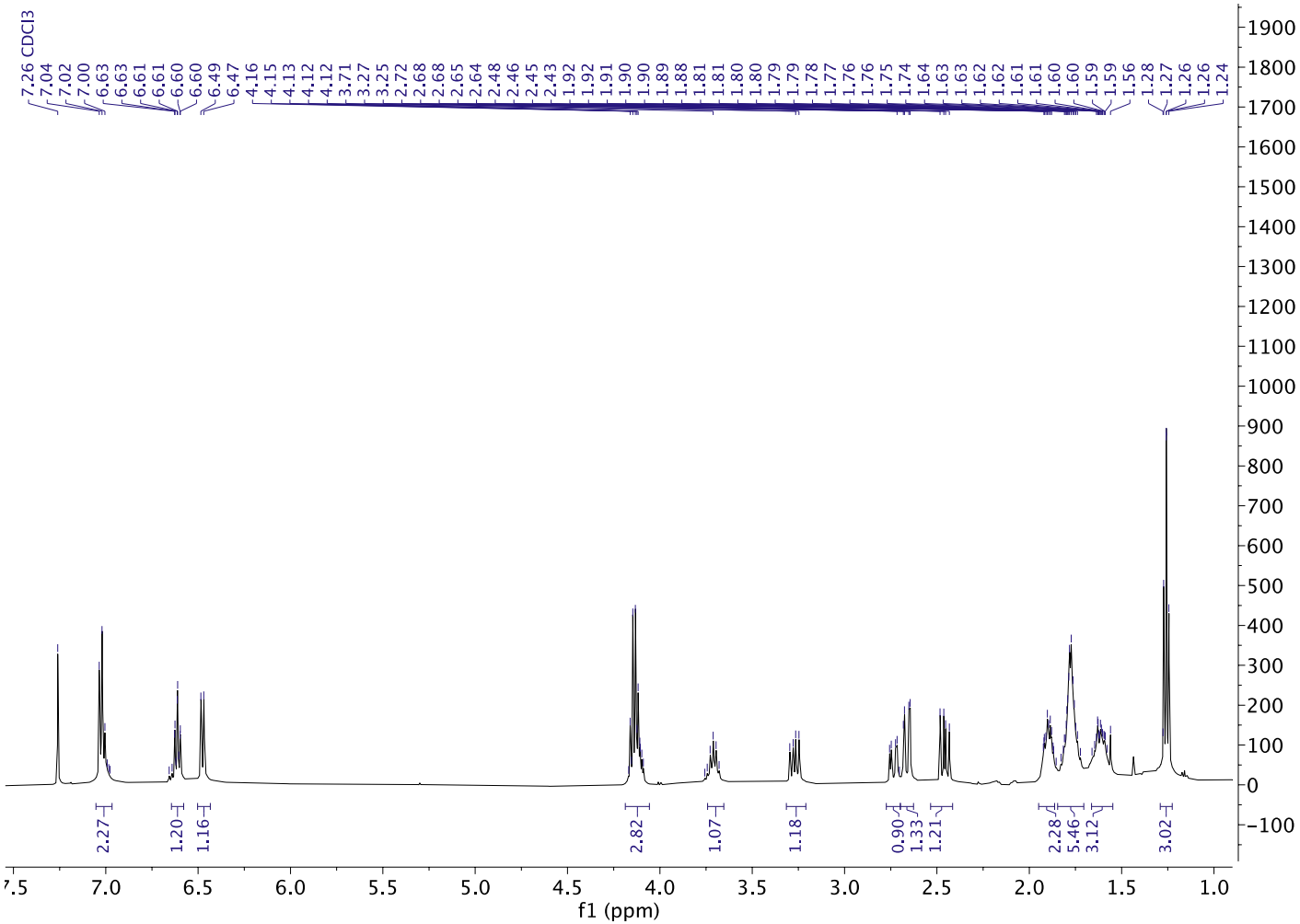




Ethyl (*S*)-2-(1-cyclopentylindolin-2-yl)acetate (3m)

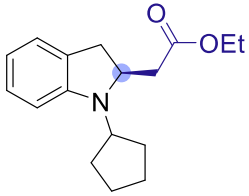
^{1}H NMR (500 MHz, $CDCl_3$)

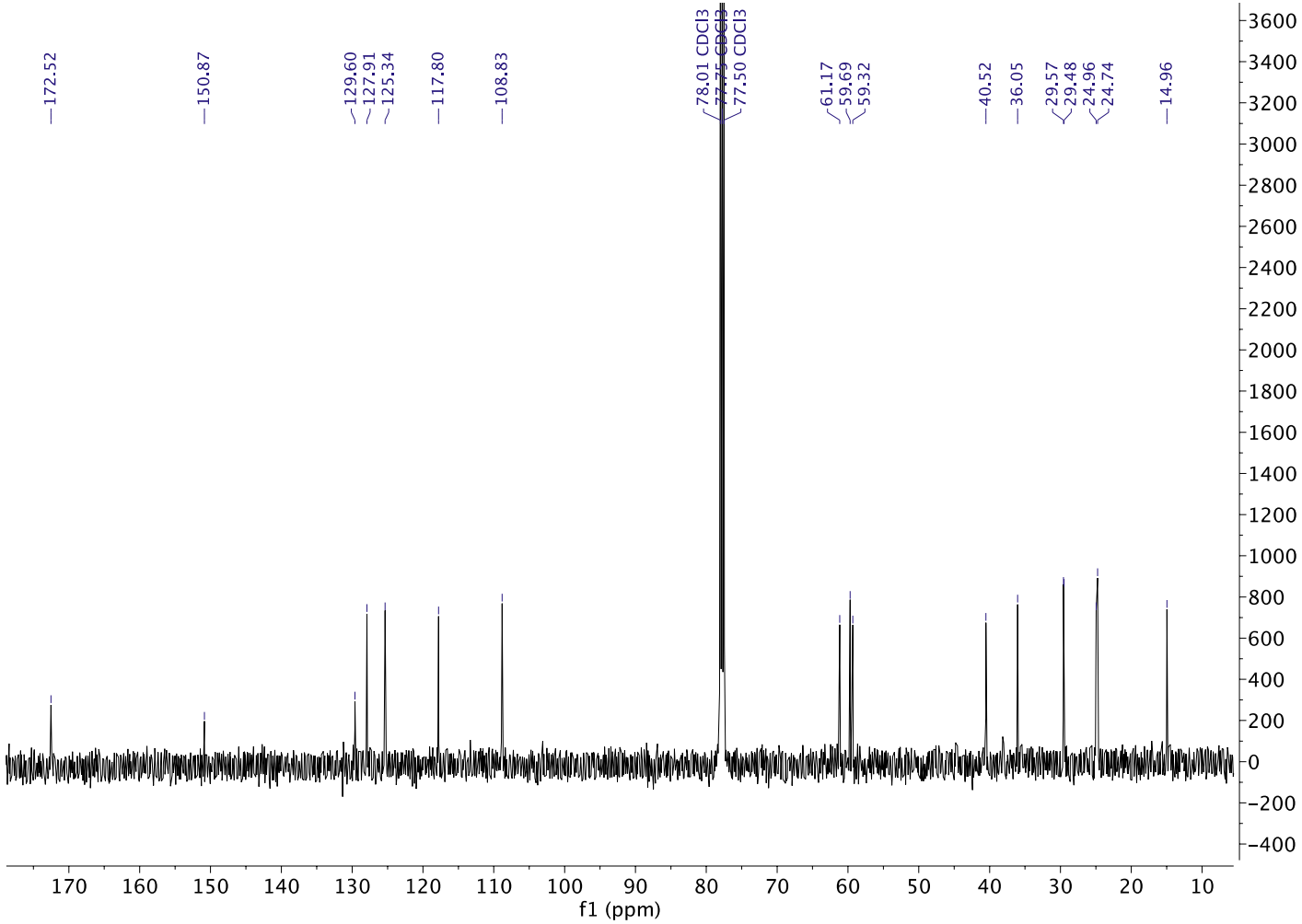




Ethyl (*S*)-2-(1-cyclopentylindolin-2-yl)acetate (3m)

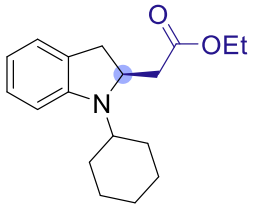
^{13}C NMR (126 MHz, $CDCl_3$)

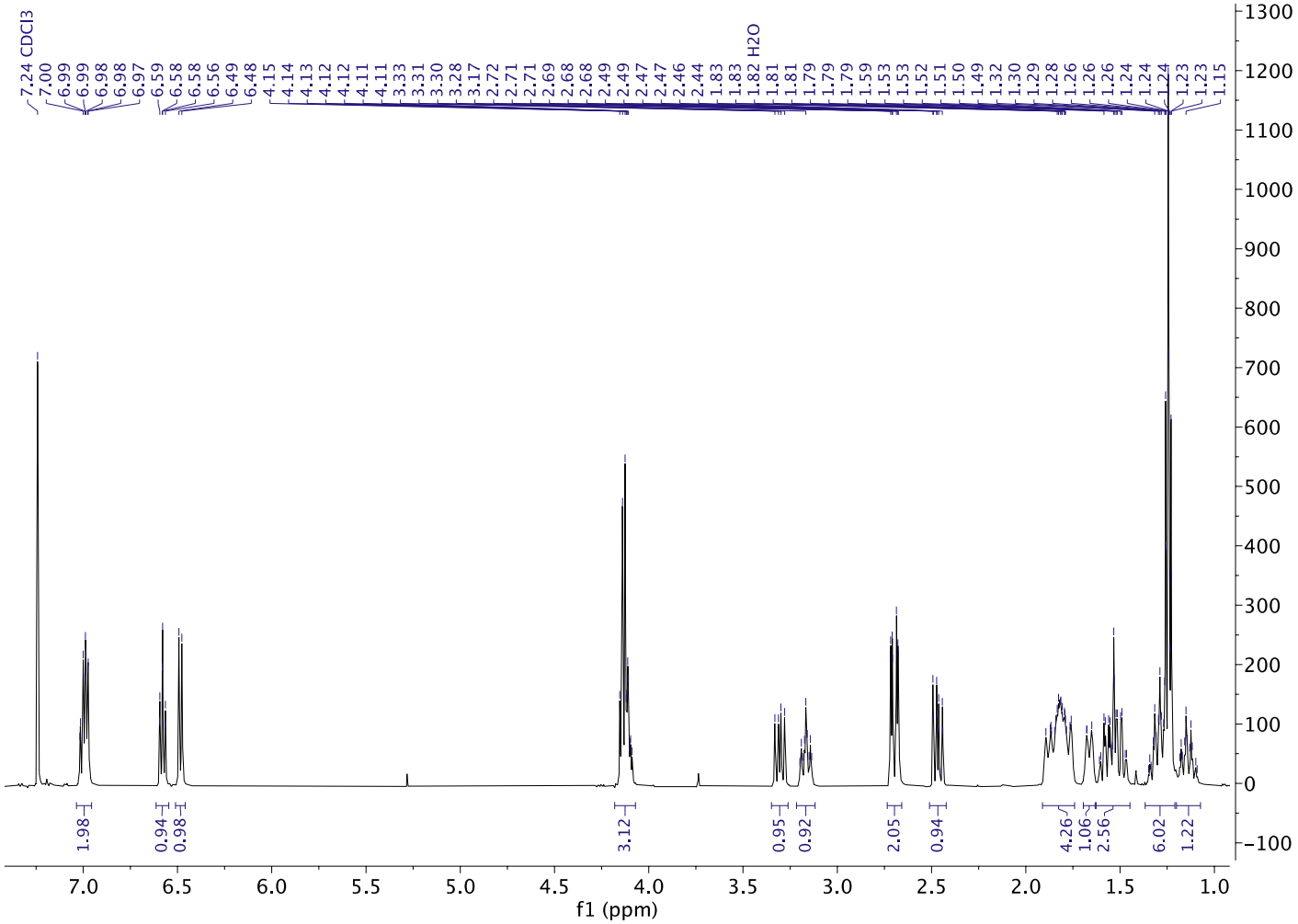




Ethyl (*S*)-2-(1-cyclohexylindolin-2-yl)acetate (3n)

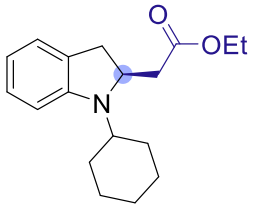
^{1}H NMR (500 MHz, $CDCl_3$)

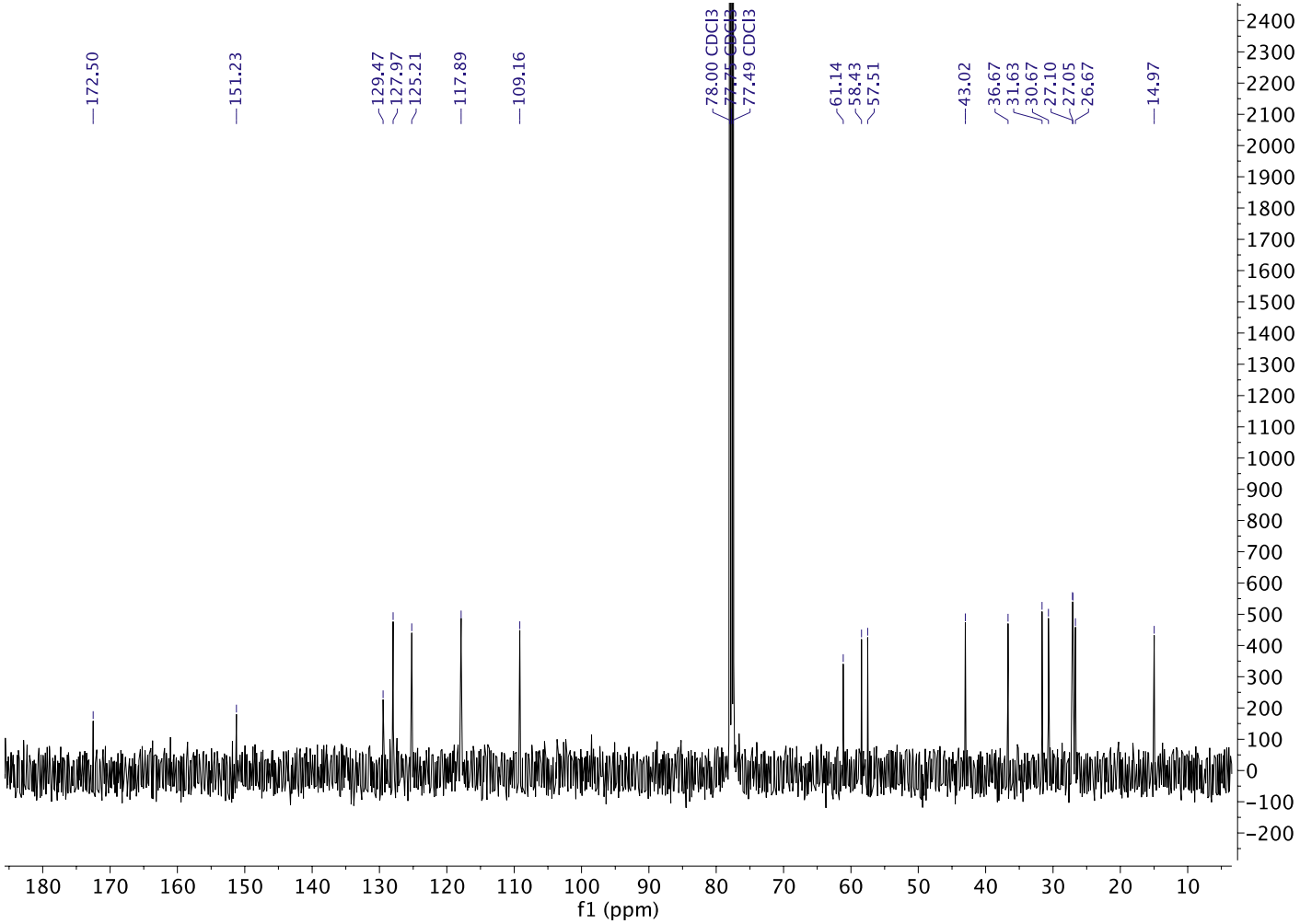




Ethyl (*S*)-2-(1-cyclohexylindolin-2-yl)acetate (3n)

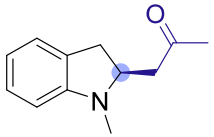
^{13}C NMR (126 MHz, $CDCl_3$)

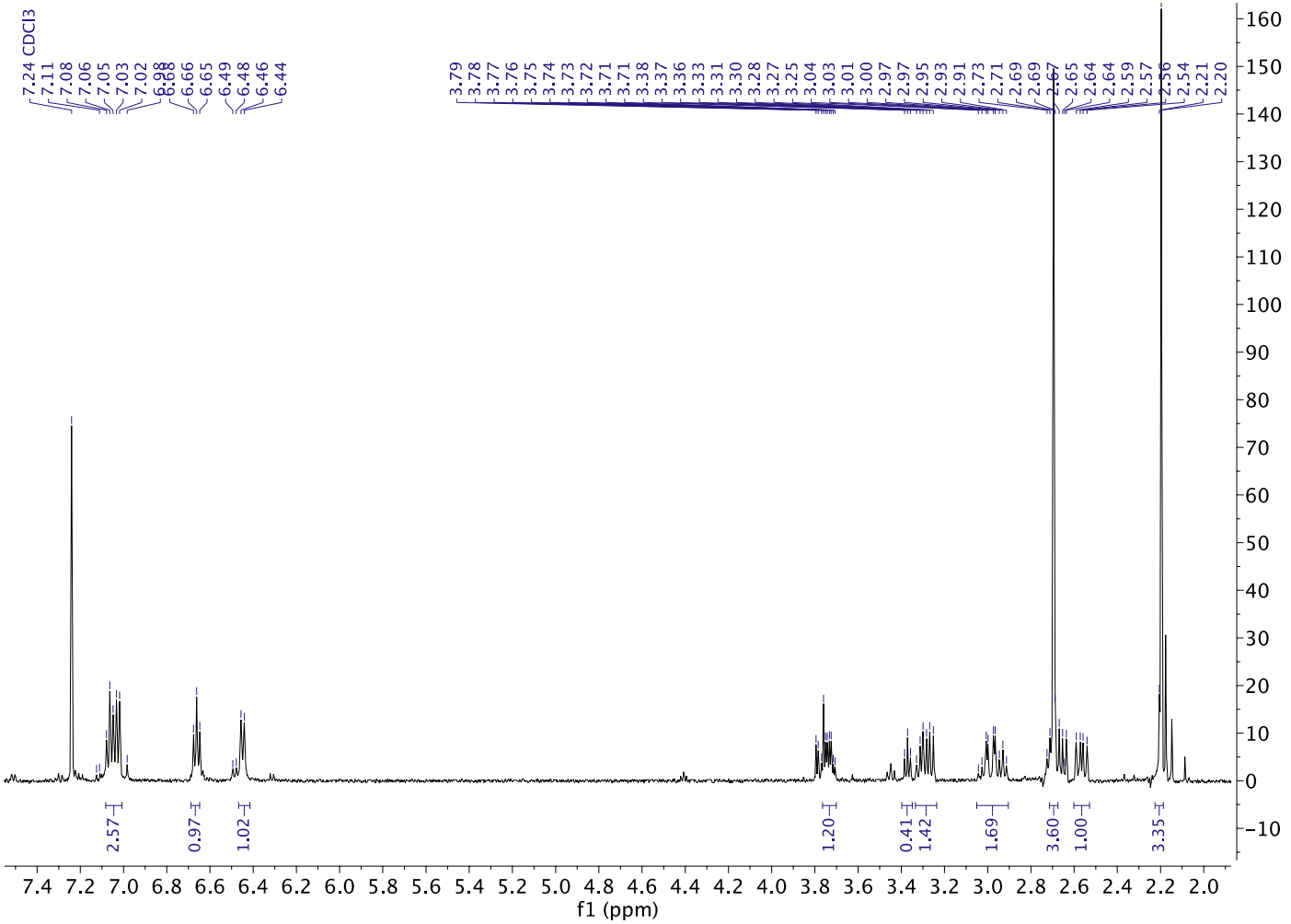




(*S*)-1-(1-methylindolin-2-yl)propan-2-one (6a)

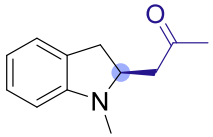
^{1}H NMR (500 MHz, $CDCl_3$)

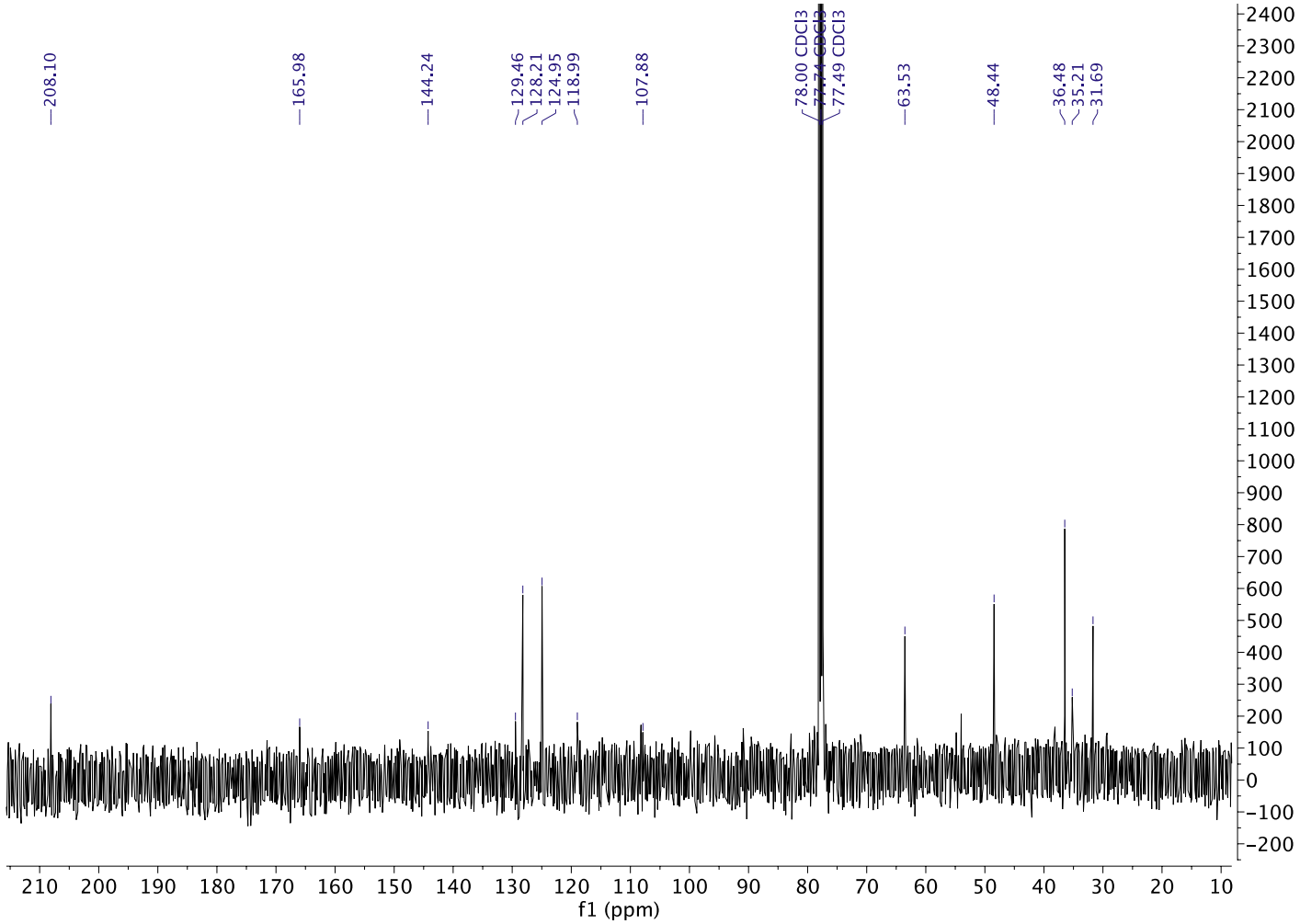




(*S*)-1-(1-methylindolin-2-yl)propan-2-one (6a)

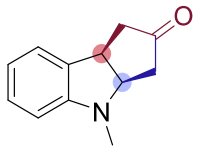
^{13}C NMR (126 MHz, $CDCl_3$)

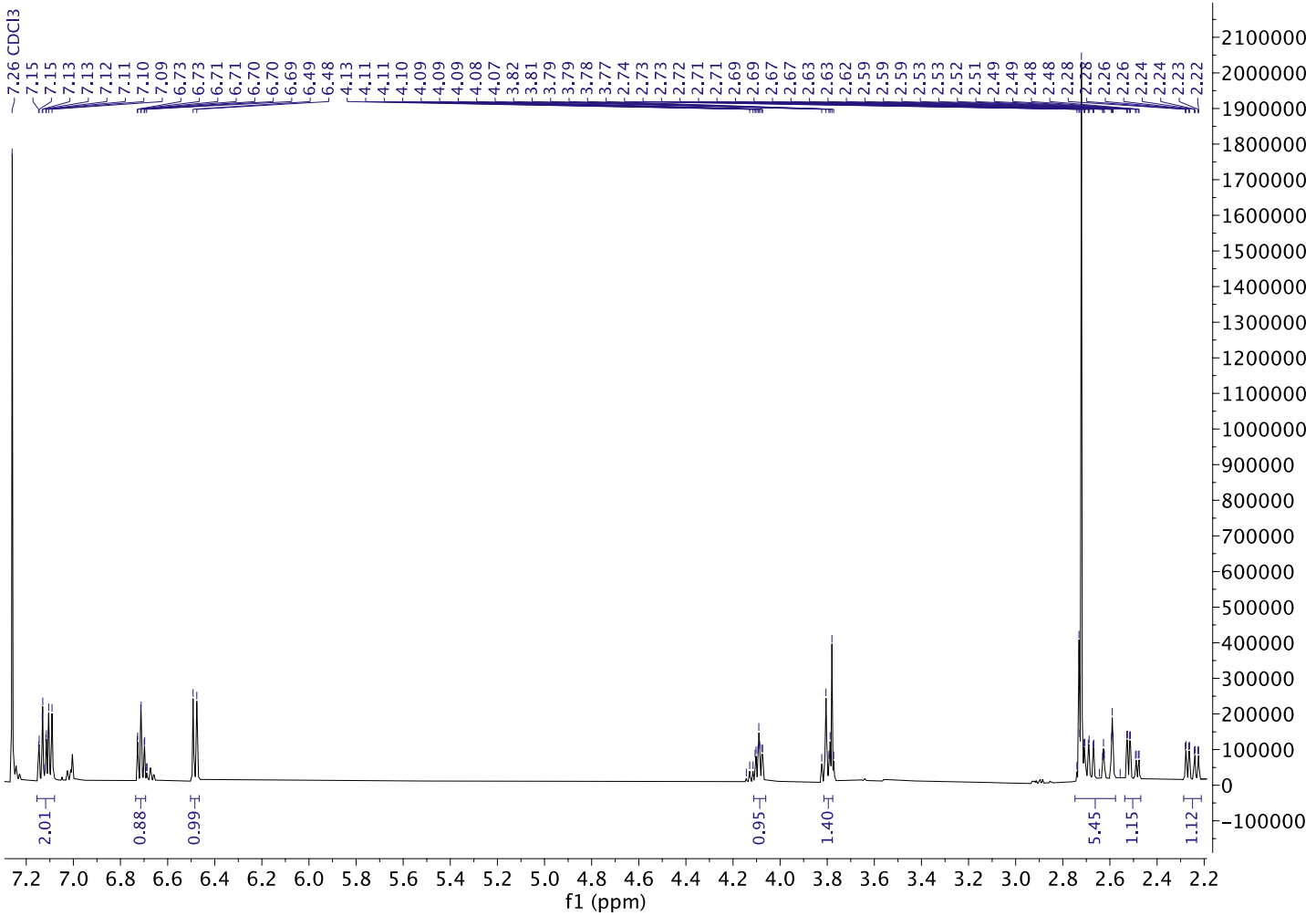




(3a*R*,8b*R*)-4-methyl-3,3a,4,8b-tetrahydrocyclopenta[b]indol-2(1H)-one (9a)

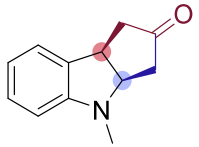
^{1}H NMR (500 MHz, $CDCl_3$)

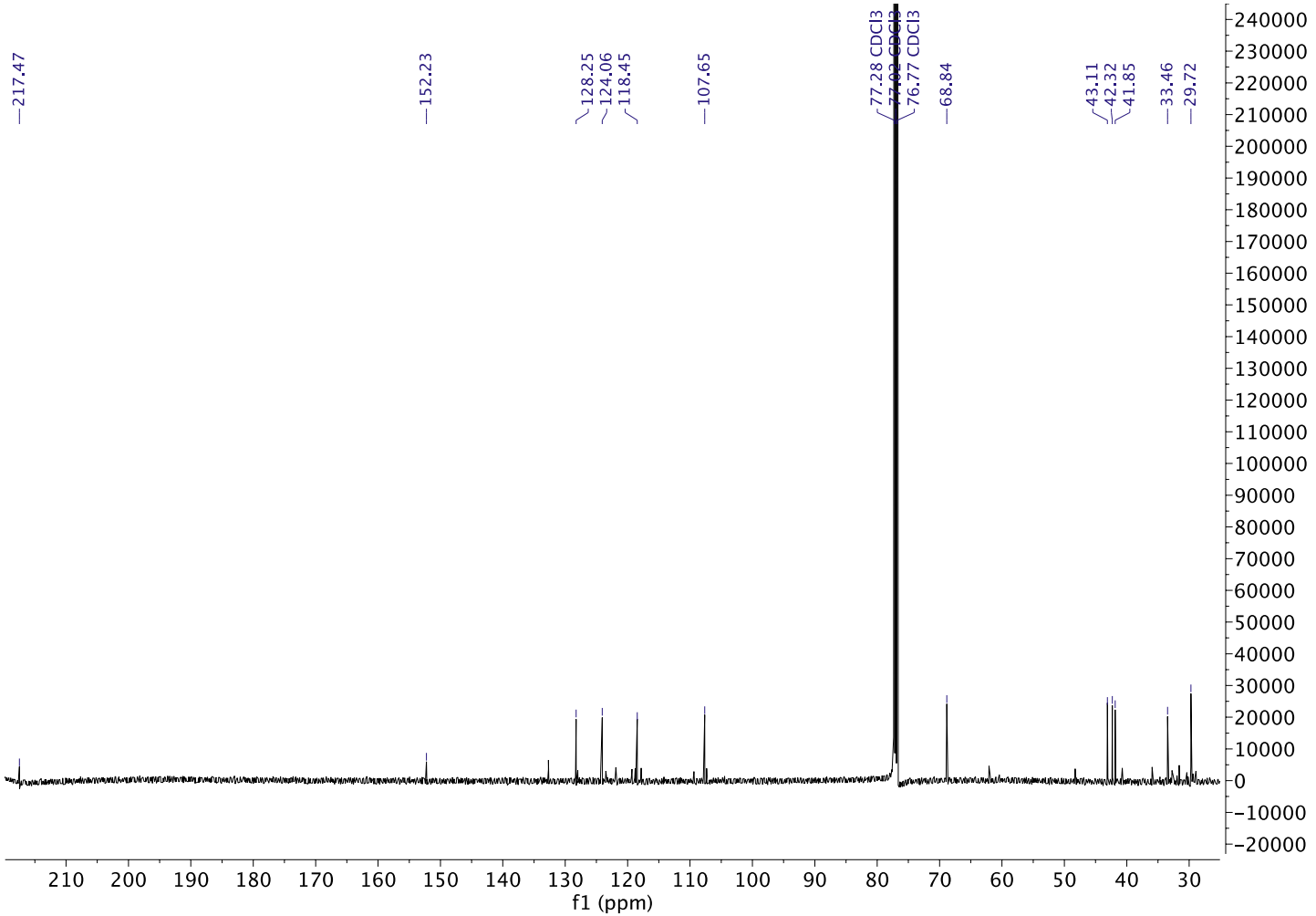




(3a*R*,8b*R*)-4-methyl-3,3a,4,8b-tetrahydrocyclopenta[b]indol-2(1H)-one (9a)

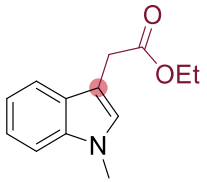
^{13}C NMR (126 MHz, $CDCl_3$)

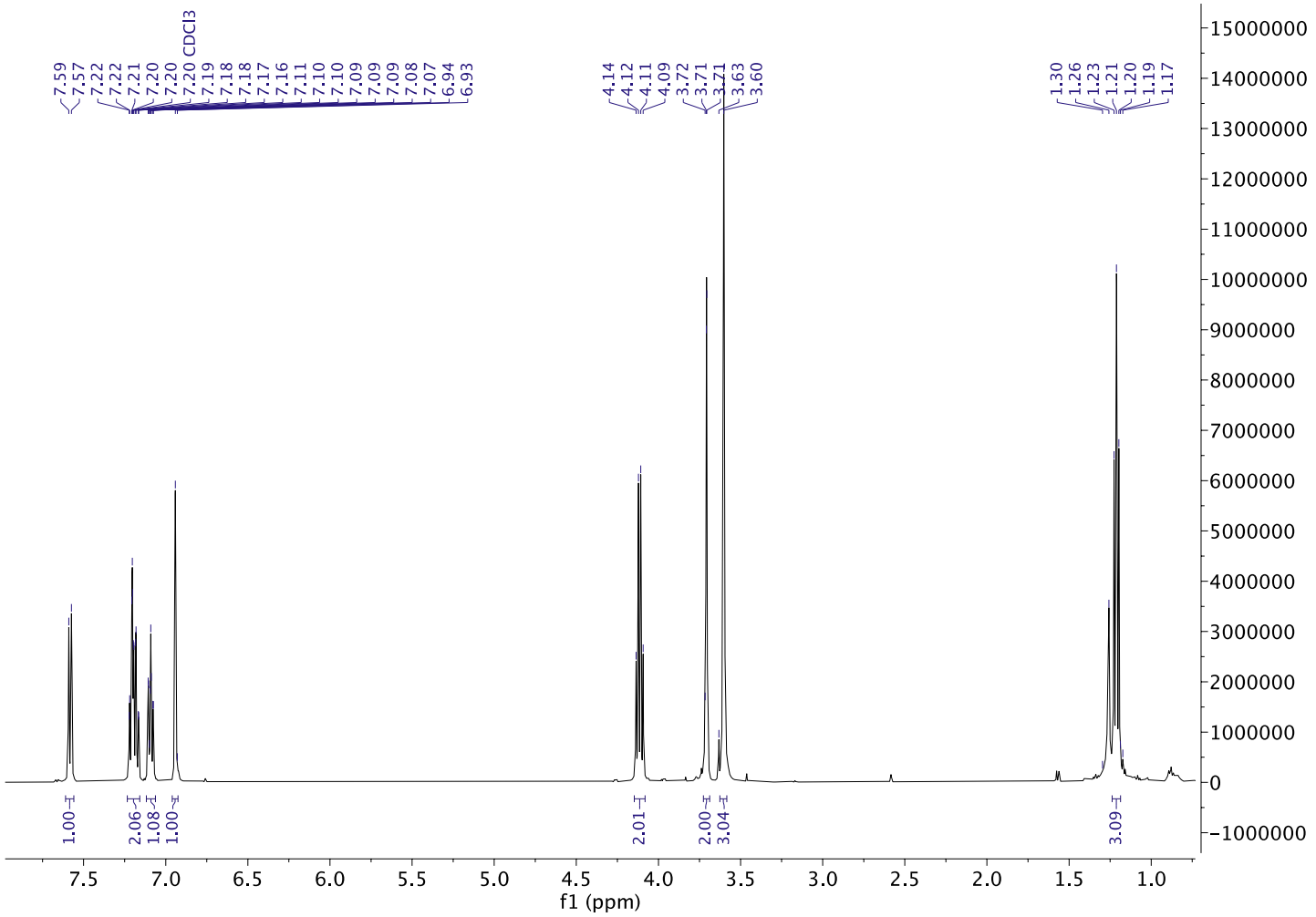




Ethyl 2-(1-methyl-1H-indol-3-yl)acetate (4a-*DS*)

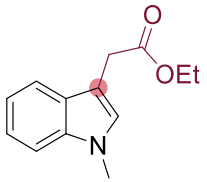
^{1}H NMR (500 MHz, $CDCl_3$)

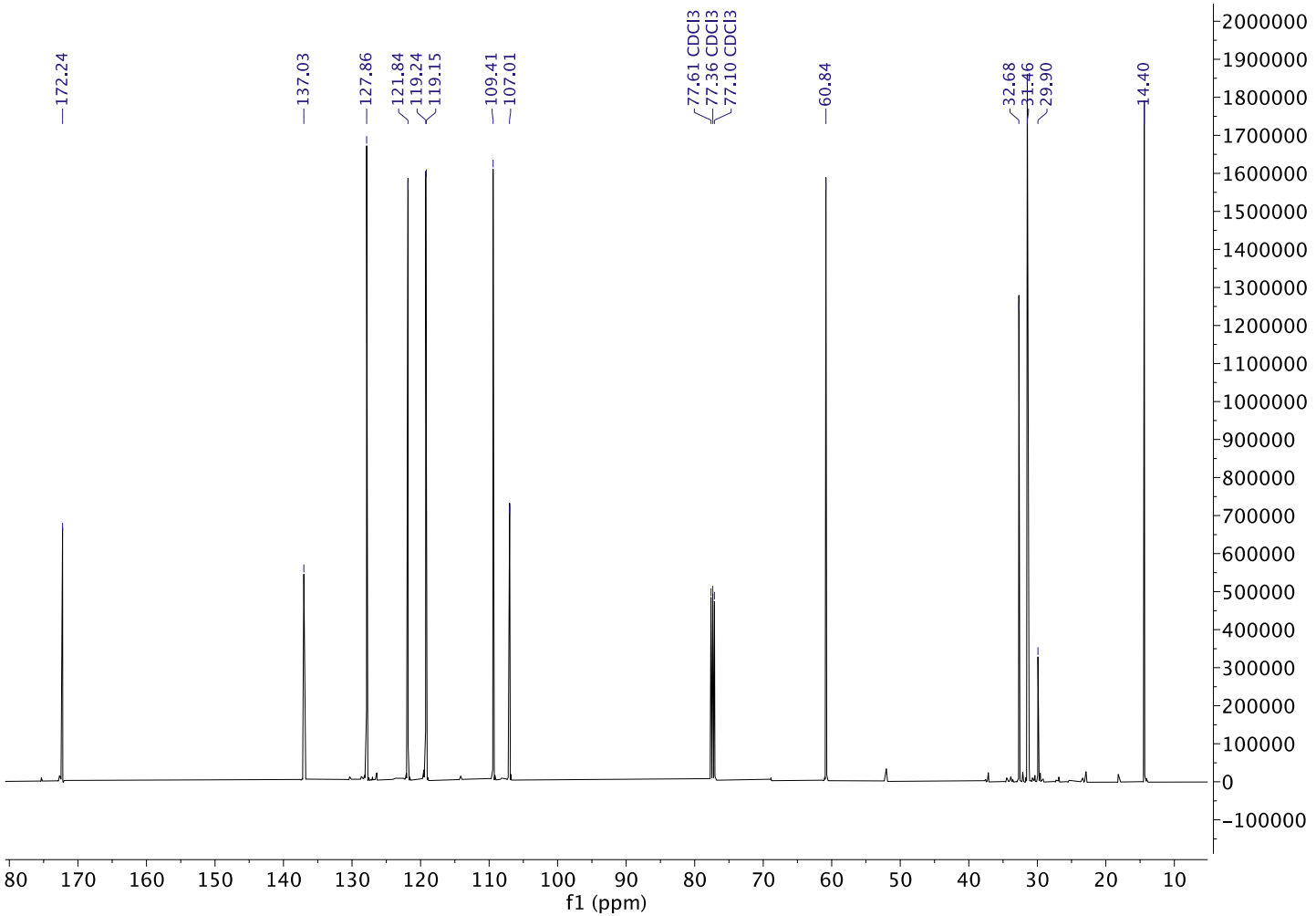




Ethyl 2-(1-methyl-1H-indol-3-yl)acetate (4a-*DS*)

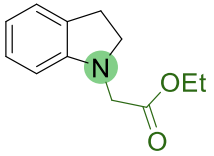
^{13}C NMR (126 MHz, $CDCl_3$)

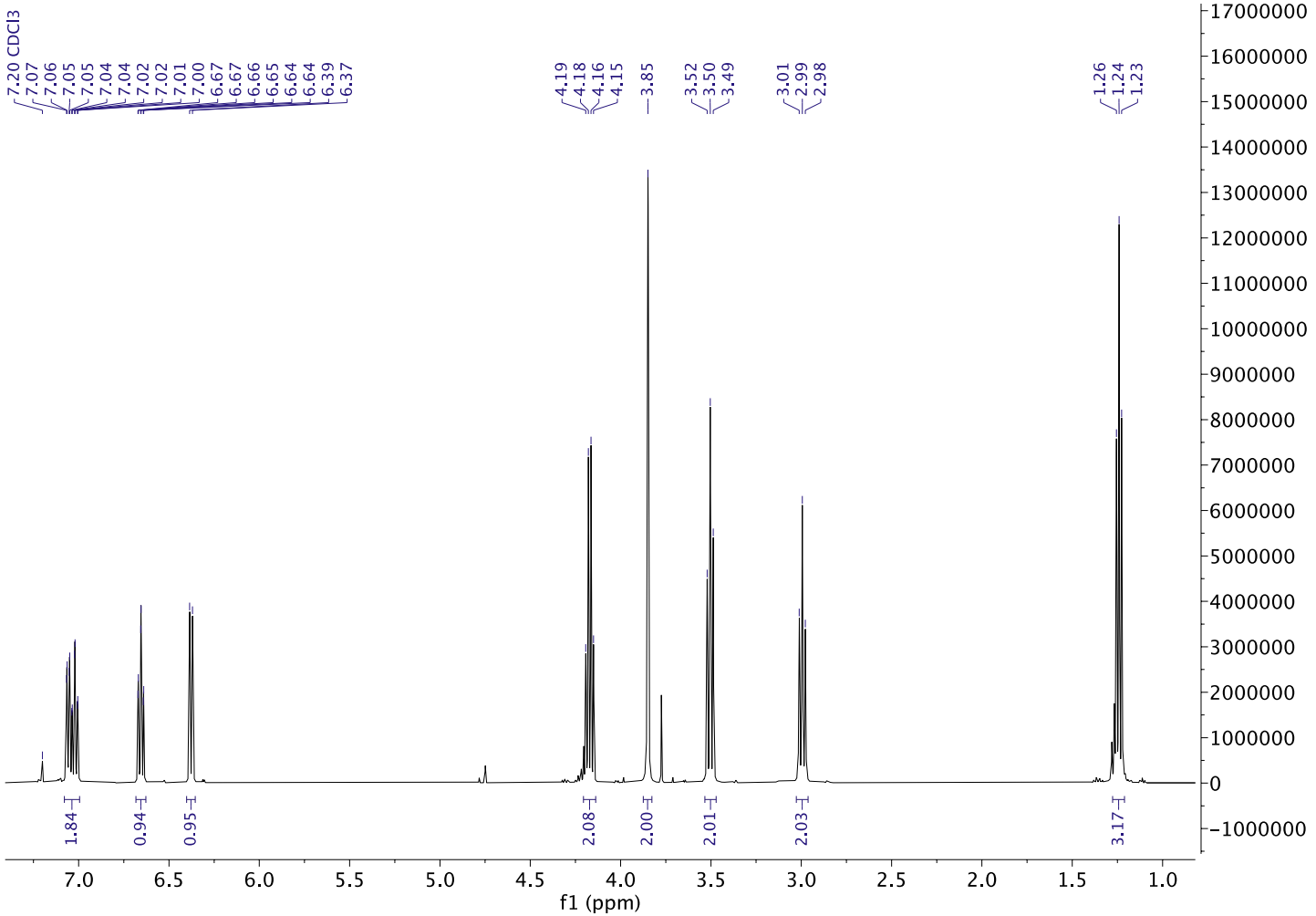




Ethyl 2-(indolin-1-yl)acetate (13a)

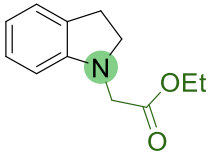
^{1}H NMR (500 MHz, $CDCl_3$)

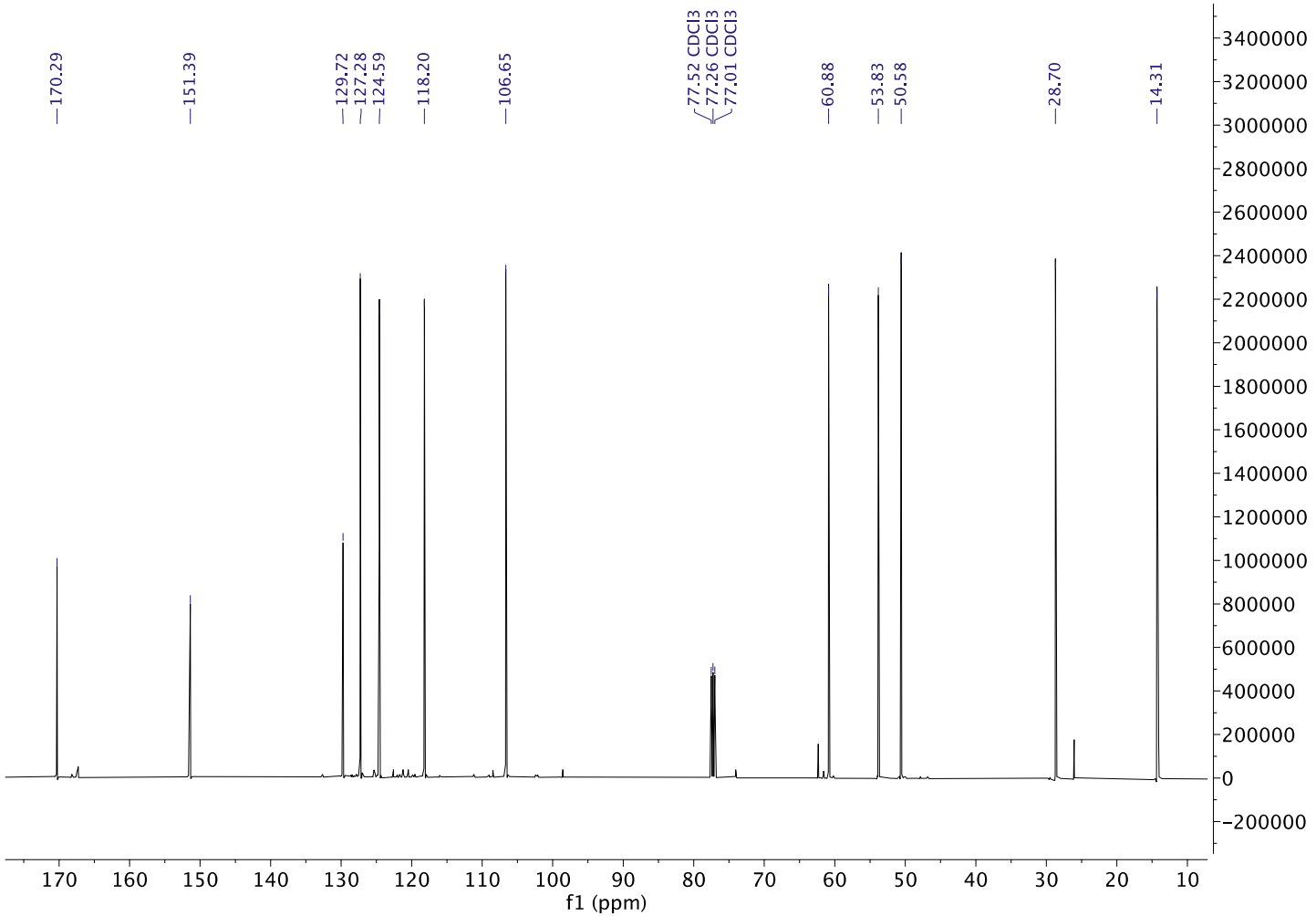




Ethyl 2-(indolin-1-yl)acetate (13a)

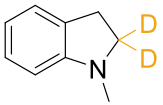
^{13}C NMR (126 MHz, $CDCl_3$)

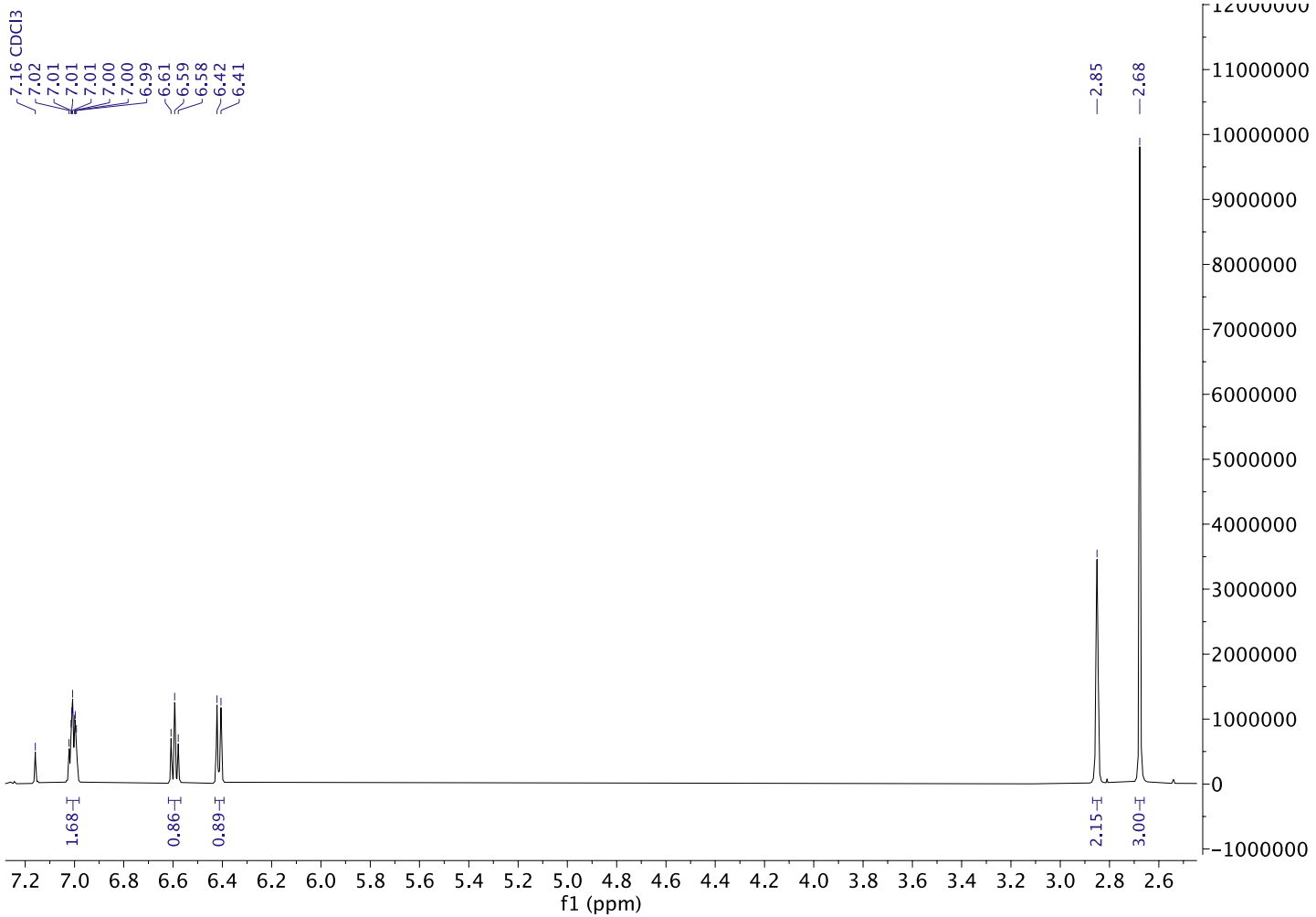




1-methylindoline-2,2-d$_2$ (1a-D$_2$)

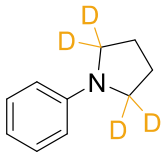
^{1}H NMR (500 MHz, $CDCl_3$)

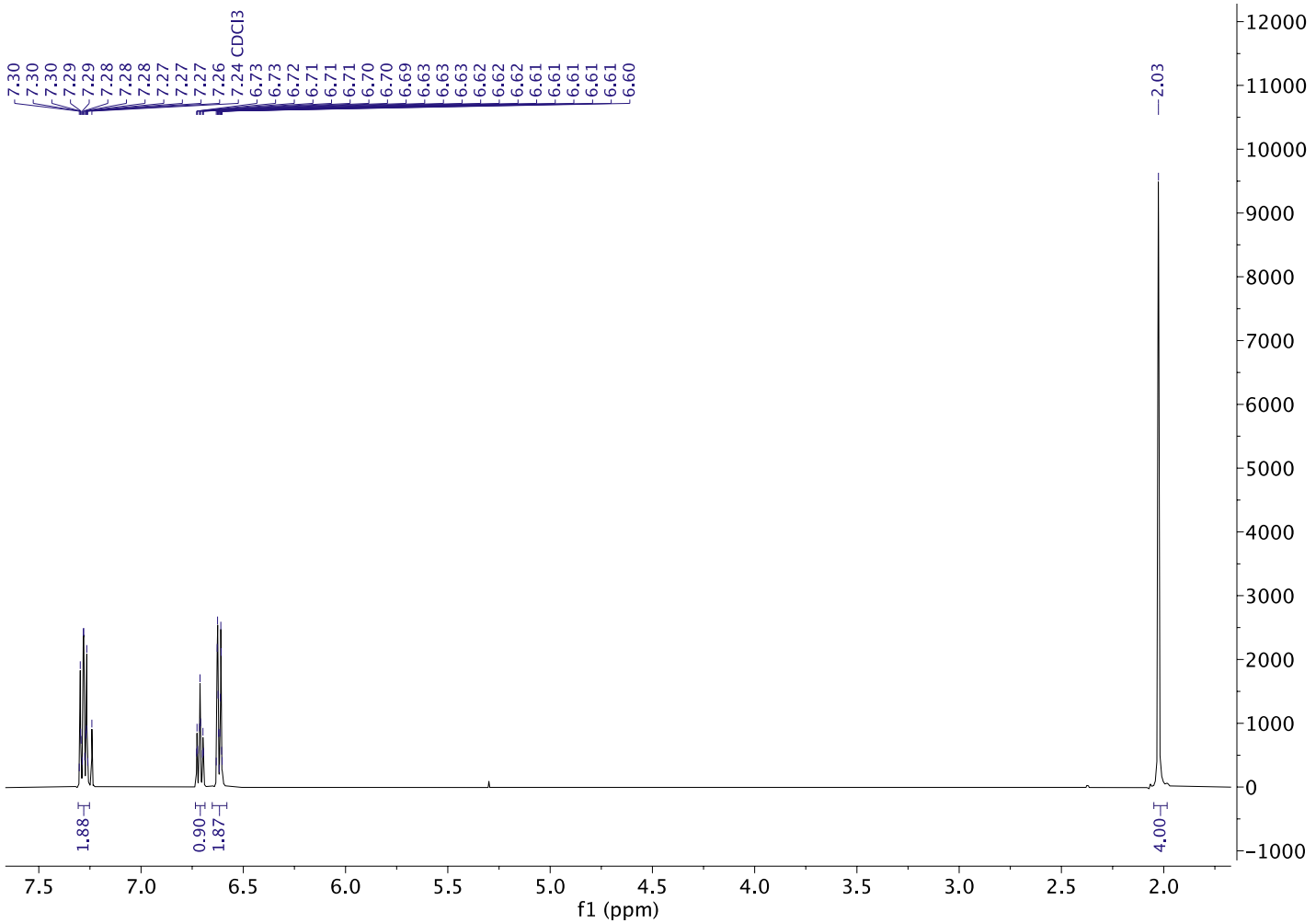




1-phenylpyrrolidine-2,2,5,5-d$_4$ (10a-D$_4$)

^{1}H NMR (500 MHz, $CDCl_3$)





Supplementary References

(S1) Gaussian 09, Revision D.01, M. J. Frisch, G. W. T., H. B. Schlegel, G. E.; Scuseria, M. A. R., J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.; A. Petersson, H. N., M. Caricato, X. Li, H. P. Hratchian, A. F.; Izmaylov, J. B., G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota,; R. Fukuda, J. H., M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai,; T. Vreven, J. A. M., Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J.; Heyd, E. B., K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J.; Normand, K. R., A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi,; M. Cossi, N. R., J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken,; C. Adamo, J. J., R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin. Gaussian, Inc. Wallingford CT, 2013.

(S2) Mennucci, B.; Tomasi, J. (1997). Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries. *J. Chem. Phys.* 106 (12), 5151-5158. 10.1063/1.473558

(S3) Torres, R. A.; Lovell, T.; Noodleman, L.; Case, D. A. (2003). Density Functional and Reduction Potential Calculations of Fe4S4 Clusters. *J. Am. Chem. Soc.* 125 (7), 1923-1936. 10.1021/ja0211104

(S4) Chai, J.-D.; Head-Gordon, M. (2008). Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Physical Chemistry Chemical Physics* 10, 6615-6620. 10.1039/B810189B

(S5) Yang, K.; Zheng, J.; Zhao, Y.; Truhlar, D. G. (2010). Tests of the RPBE, revPBE, tHCTHhyb, wB97X-D, and MOHLYP density functional approximations and 29 others against representative databases for diverse bond energies and barrier heights in catalysis. *J. Chem. Phys.* 132, 164117. 10.1063/1.3382342

(S6) Khade, R. L.; Fan, W.; Ling, Y.; Yang, L.; Oldfield, E.; Zhang, Y. (2014). Iron Porphyrin Carbenes as Catalytic Intermediates: Structures, Mossbauer and NMR Spectroscopic Properties, and Bonding. *Angew. Chem. Int. Ed.* 53, 7574-7578. 10.1002/anie.201402472

(S7) Hay, P. J.; Wadt, W. R. (1985). Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *The Journal of chemical physics* 82 (1), 270-283. 10.1063/1.448799

(S8) Khade, R. L.; Zhang, Y. (2015). Catalytic and Biocatalytic Iron Porphyrin Carbene Formation: Effects of Binding Mode, Carbene Substituent, Porphyrin Substituent, and Protein Axial Ligand. *J. Am. Chem. Soc.* 137 (24), 7560-7563. 10.1021/jacs.5b03437

(S9) Khade, R. L.; Zhang, Y. (2017). C-H Insertions by Iron Porphyrin Carbene: Basic Mechanism and Origin of Substrate Selectivity. *Chemistry - A Euopean Journal* 23, 17654-17658. 10.1002/chem.201704631

(S10) Wei, Y.; Tinoco, A.; Steck, V.; Fasan, R.; Zhang, Y. (2018). Cyclopropanations via Heme Carbenes: Basic Mechanism and Effects of Carbene Substituent, Protein Axial Ligand, and Porphyrin Substitution. *J. Am. Chem. Soc.* 140, 1649-1662. 10.1021/jacs.7b09171

(S11) Vargas, D. A.; Khade, R. L.; Zhang, Y.; Fasan, R. (2019). Biocatalytic strategy for highly diastereo- and enantioselective synthesis of 2,3-dihydrobenzofuran based tricyclic scaffolds. *Angew. Chem. Int. Ed.* 58, 10148-10152. 10.1002/anie.201903455

(S12) Khade, R., L.; Chandgude, A. L.; Fasan, R.; Zhang, Y. (2019). Mechanistic Investigation of Biocatalytic Heme Carbenoid Si-H Insertions. *ChemCatChem* 11, 3101-3108. 10.1002/cctc.201801755

(S13) Wei, Y.; Conklin, M.; Zhang, Y. (2022). Biocatalytic Intramolecular C-H aminations via Engineered Heme Proteins: Full Reaction Pathways and Axial Ligand Effects. *Chemistry* 28 (59), e202202006. 10.1002/chem.202202006

(S14) Tinoco, A.; Wei, Y.; Bacik, J. P.; Moore, E. J.; Ando, N.; Zhang, Y.; Fasan, R. (2019). Origin of high stereocontrol in olefin cyclopropanation catalyzed by an engineered carbene transferase. *ACS Catal.* 9, 1514-1524 10.1021/acscatal.8b04073

(S15) Wang, J.-C.; Xu, Z.-J.; Guo, Z.; Deng, Q.-H.; Zhou, C.-Y.; Wan, X.-L.; Che, C.-M. (2012). Highly enantioselective intermolecular carbene insertion to C–H and Si–H bonds catalyzed by a chiral iridium(III) complex of a D4-symmetric Halterman porphyrin ligand. *Chem. Commun.* 48, 4299-4301. 10.1039/C2CC30441D

(S16) Chan, K.-H.; Guan, X.; Lo, V. K.-Y.; Che, C.-M. (2014). Elevated Catalytic Activity of Ruthenium(II)–Porphyrin-Catalyzed Carbene/Nitrene Transfer and Insertion Reactions with N-Heterocyclic Carbene Ligands. *Angewandte Chemie-International Edition* 53, 2982-2987. 10.1002/ange.201309888
(S17) Nakamura, E.; Yoshikai, N.; Yamanaka, M. (2002). Mechanism of C-H bond activation/C-C bond formation reaction between diazo compound and alkane catalyzed by dirhodium tetracarboxylate. *J. Am. Chem. Soc.* 124 (24), 7181-7192. 10.1021/ja017823o
(S18) Tahsini, L.; Bagherzadeh, M.; Nam, W.; de Visser, S. P. (2009). Fundamental Differences of Substrate Hydroxylation by High-Valent Iron(IV)-Oxo Models of Cytochrome P450. *Inorg. Chem.* 48, 6661-6669. 10.1021/jp060908x
(S19) Liu, Y.; Xu, W.; Zhang, J.; Fuller, W.; Schultz, C. E.; Li, J. (2017). Electronic Configuration and Ligand Nature of Five-Coordinate Iron Porphyrin Carbene Complexes: An Experimental Study. *J. Am. Chem. Soc.* 139, 5023-5026. 10.1021/jacs.7b01722
(S20) Stroscio, G. D.; Srnec, M.; Hadt, R. G. (2020). Multireference Ground and Excited State Electronic Structures of Free- versus Iron Porphyrin-Carbenes. *Inorg. Chem.* 59 (13), 8707-8715. 10.1021/acs.inorgchem.0c00249
(S21) Abucayon, E. G.; Chu, J. M.; Ayala, M.; Khade, R. L.; Zhang, Y.; Richter-Addo, G. B. (2021). Insight into the preferential N-binding versus O-binding of nitrosoarenes to ferrous and ferric heme centers. *Dalton Trans.* 50 (10), 3487-3498. 10.1039/d0dt03604h
(S22) Sainna, M. A.; Sil, D.; Sahoo, D.; Martin, B.; Rath, S. P.; Comba, P.; de Visser, S. P. (2015). Spin-State Ordering in Hydroxo-Bridged Diiron(III)bisporphyrin Complexes. *Inorg. Chem.* 54 (4), 1919-1930. 10.1021/ic502803b
(S23) Nakamura, M.; Ohgo, Y.; Ikezaki, A. Electronic and Magnetic Structures of Iron Porphyrin Complexes. In *Handbook of Porphyrin Science*, Handbook of Porphyrin Science, Vol. Volume 7; World Scientific Publishing Company, 2010; pp 1-146.
(S24) Weiss, R.; Gold, A.; Terner, J. (2006). Cytochromes c ': Biological models for the S = (3)/(2),(5)/(2) spin-state admixture? *Chem. Rev.* 106 (6), 2550-2579. 10.1021/cr040416l
(S25) Reed, C. A.; Mashiko, T.; Bentley, S. P.; Kastner, M. E.; Scheidt, W. R.; Spartalian, K.; Lang, G. (1979). The missing heme spin state and a model for cytochrome c'. The mixed S = 3/2, 5/2 intermediate spin ferric porphyrin: perchlorato(meso-tetraphenylporphinato)iron(III). *J. Am. Chem. Soc.* 101 (11), 2948-2958. 10.1021/ja00505a023
(S26) Li, X.; Dong, L.; Liu, Y. (2020). Theoretical Study of Iron Porphyrin Nitrene: Formation Mechanism, Electronic Nature, and Intermolecular C–H Amination. *Inorg. Chem.* 59 (3), 1622-1632. 10.1021/acs.inorgchem.9b02216
(S27) Li, Z.; Burnell, D. J.; Boyd, R. J. (2017). Computational Study of Engineered Cytochrome P450-Catalyzed C–H Amination: The Origin of the Regio- and Stereoselectivity. *The Journal of Physical Chemistry Part B* 121 (48), 10859-10868. 10.1021/acs.jpcb.7b10256
(S28) Chai, J. D.; Head-Gordon, M. (2008). Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys Chem Chem Phys* 10 (44), 6615-6620. 10.1039/b810189b
(S29) Hehre, W. J.; Ditchfield, R.; Pople, J. A. (1972). Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *The Journal of Chemical Physics* 56 (5), 2257-2261. 10.1063/1.1677527
(S30) Barone, V.; Cossi, M. (1998). Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *Journal of Physical Chemistry A* 102 (11), 1995-2001. 10.1021/jp9716997
(S31) Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K. W.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; et al. Chapter nineteen - Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. In *Methods in Enzymology*, Johnson, M. L., Brand, L. Eds.; Vol. 487; Academic Press, 2011; pp 545-574.

(S32) Weiner, M. P.; Costa, G. L.; Schoettlin, W.; Cline, J.; Mathur, E.; Bauer, J. C. (1994). Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. *Gene* 151 (1-2), 119-123. 10.1016/0378-1119(94)90641-6
(S33) Berry, E. A.; Trumpower, B. L. (1987). Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. *Anal Biochem* 161 (1), 1-15. 10.1016/0003-2697(87)90643-9
(S34) Ellis-Guardiola, K.; Soule, J.; Clubb, R. T. (2021). Methods for the Extraction of Heme Prosthetic Groups from Hemoproteins. *Bio Protoc* 11 (18), e4156. 10.21769/BioProtoc.4156
(S35) Barr, I.; Guo, F. (2015). Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *Bio Protoc* 5 (18), e1594-e1594. 10.21769/bioprotoc.1594
(S36) Abid, I.; Gosselin, P.; Mathe-Allainmat, M.; Abid, S.; Dujardin, G.; Gaulon-Nourry, C. (2015). TBAF-Triggered Aldol-Type Addition of alpha-Triethylsilyl-alpha-diazoacetone. *J Org Chem* 80 (20), 9980-9988. 10.1021/acs.joc.5b01554
(S37) Benard, S.; Neuville, L.; Zhu, J. (2010). Copper-promoted N-cyclopropylation of anilines and amines by cyclopropylboronic acid. *Chem Commun (Camb)* 46 (19), 3393-3395. 10.1039/b925499d
(S38) Bayindir, S.; Erdogan, E.; Kilic, H.; Aydin, O.; Saracoglu, N. (2015). Synthesis of N-Alkylated Indolines and Indoles from Indoline and Aliphatic Ketones. *Journal of Heterocyclic Chemistry* 52 (5), 1589-1594. 10.1002/jhet.2337
(S39) Torisu, K.; Kobayashi, K.; Iwahashi, M.; Nakai, Y.; Onoda, T.; Nagase, T.; Sugimoto, I.; Okada, Y.; Matsumoto, R.; Nanbu, F.; et al. (2004). Discovery of a new class of potent, selective, and orally active prostaglandin D2 receptor antagonists. *Bioorg Med Chem* 12 (20), 5361-5378. 10.1016/j.bmc.2004.07.048
(S40) Shen, H.; Deng, Q. F.; Liu, R. J.; Feng, Y. Y.; Zheng, C. K.; Xiong, Y. (2017). Intramolecular aminocyanation of alkenes promoted by hypervalent iodine. *Organic Chemistry Frontiers* 4 (9), 1806-1811. 10.1039/c7qo00214a