

Supporting Information

## Diversity-Oriented Enzymatic Synthesis of Cyclopropane Building Blocks

Bruce J. Wittmann<sup>1,‡</sup>, Anders M. Knight<sup>1,‡</sup>, Julie L. Hofstra<sup>2,†</sup>, Sarah E. Reisman<sup>2</sup>, S. B. Jennifer Kan<sup>2</sup>, and Frances H. Arnold<sup>1,2,\*</sup>

<sup>‡</sup>Equal Contribution

Author affiliation

1: Division of Biology and Bioengineering, California Institute of Technology, 1200 East California Boulevard, MC 210-41, Pasadena, California 91125, United States

2: Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 East California Boulevard, MC 210-41, Pasadena, California 91125, United States

Corresponding Author

\*frances@cheme.caltech.edu

## Contents

Materials and Methods.....	4
Materials and instrumentation.....	4
Site-saturation library construction.....	4
Transformation of genes.....	5
96-Well plate library expression.....	5
96-Well plate reactions and screening.....	5
Medium- and large-scale protein expression .....	6
Lineage validation reactions .....	7
Sonicated lysate preparation and hemochrome assay.....	7
Preparation of lyophilized whole-cell biocatalysts .....	8
Preparative-scale biocatalytic reactions with lyophilized whole cells.....	8
Procedure for column-free purification of trifluoroborated product (4) .....	9
Determination of initial cyclopropanation activity.....	9
Calibration curves for analytical-scale concentration determinations .....	12
Protein Engineering Strategies.....	12
<i>Cis</i> -selective lineage.....	12
<i>Trans</i> -selective lineage.....	14
Protein Crystallography.....	16
Cloning, expression, and purification of for crystallography .....	16
Protein crystallization.....	17
Data collection and refinement .....	18
Small-Molecule Crystal Formation and Data Collection.....	20
Small-molecule crystal formation .....	20
X-Ray refinement details .....	20
Compound Synthesis and Characterization .....	22
Ethyl ( <i>trans</i> )-2-(phenylthio)cyclopropane-1-carboxylate and ethyl ( <i>cis</i> )-2-(phenylthio)cyclopropane-1-carboxylate .....	22
Ethyl ( <i>trans</i> )-2-(( <i>S</i> )-phenylsulfinyl)cyclopropane-1-carboxylate.....	23
Ethyl ( <i>cis</i> )-2-(( <i>S</i> )-phenylsulfinyl)cyclopropane-1-carboxylate .....	24
Ethyl ( <i>cis</i> )-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)cyclopropane-1-carboxylate.....	24
Ethyl ( <i>cis</i> )-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)cyclopropane-1-carboxylate.....	25

Ethyl (1 <i>R</i> ,2 <i>S</i> )-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)cyclopropane-1-carboxylate (enzymatic).....	25
Ethyl (1 <i>R</i> ,2 <i>S</i> )-2-(trifluoro- <i>l</i> -boranyl)cyclopropane-1-carboxylate, potassium salt ((1 <i>R</i> , 2 <i>S</i> )-4) .....	27
Suzuki-Miyaura cross-coupling with 4.....	27
General procedure: Suzuki-Miyaura cross-coupling with <i>cis</i> -3.....	28
Ethyl (1 <i>R</i> ,2 <i>S</i> )-2-(4-methoxyphenyl)cyclopropane-1-carboxylate (5a).....	28
Ethyl (1 <i>R</i> ,2 <i>S</i> )-2-(3,4-difluorophenyl)cyclopropane-1-carboxylate (5b).....	29
Ethyl (1 <i>R</i> ,2 <i>S</i> )-2-(anthracen-9-yl)cyclopropane-1-carboxylate (5c) .....	29
Ethyl (1 <i>R</i> ,2 <i>S</i> )-2-(6-chloropyridin-3-yl)cyclopropane-1-carboxylate (5d).....	30
Suzuki-Miyaura cross-coupling reactions with <i>trans</i> -3 .....	30
Ethyl <i>trans</i> -2-(4-methoxyphenyl)cyclopropane-1-carboxylate .....	31
Ethyl <i>trans</i> -2-(3,4-difluorophenyl)cyclopropane-1-carboxylate .....	31
Chromatographic Data .....	32
Compound separation conditions .....	32
Calibration curves for analytical-scale assays.....	33
Representative chromatographic traces .....	34
DNA and Amino-Acid Sequences .....	40
Nucleotide sequences of proteins of interest .....	40
Amino-acid sequences of proteins of interest. ....	42
Spectroscopic Data.....	44
References .....	65

## Materials and Methods

### Materials and instrumentation

Solvents and reagents were ordered from Sigma Aldrich, TCI, Aurum, or Alfa Aesar and used without further purification. GC-FID data were collected on Shimadzu GC-17A and Agilent 7820A GC systems. GC-MS data were collected on a Shimadzu GCMS-QP2010 SE and Agilent 7820A with 5977B MSD.  $^1\text{H}$  NMR,  $^{11}\text{B}$  NMR, and  $^{13}\text{C}$  NMR spectra were recorded on a Bruker Prodigy 400 MHz instrument;  $^1\text{H}$  NMR spectra were recorded at 400 MHz,  $^{11}\text{B}$  NMR spectra were recorded at 128 MHz, and  $^{13}\text{C}$  NMR spectra were recorded at 101 MHz. For  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR, chemical shifts were normalized to the deuterated solvent's proto impurity ( $\text{CHCl}_3$ :  $^1\text{H}$  NMR 7.26 ppm,  $^{13}\text{C}$  NMR 77.16 ppm;  $\text{DMSO-d}_6$ :  $^1\text{H}$  NMR 2.50 ppm,  $^{13}\text{C}$  NMR 39.52 ppm).  $^{19}\text{F}$  NMR were recorded on a Varian 300 MHz instrument;  $^{19}\text{F}$  NMR spectra were recorded at 282 MHz.

### Site-saturation library construction

Site-saturation mutagenesis was performed using the 22-codon method.<sup>1</sup> Two methods were used for primer design and cloning: kinase-ligase-*DpnI* (KLD) (New England Biolabs) and Gibson assembly. In KLD cloning, primers were designed for blunt-end cloning with the KLD kit, where the degenerate codons were placed at the blunt end. The PCR product was then ligated via the KLD enzyme mix for 1 hour and used to transform electrocompetent cells. In the second approach, primers were designed for Gibson assembly, with the degenerate codons placed within either the forward or reverse primer. Two PCRs were performed for each library, the first containing a mixture of degenerate primers (12:9:1 NDT:VHG:TGG) and the second containing a reverse primer with a complementary overhang to the degenerate primer. The second primer for each PCR is located in the ampicillin cassette to select against off-target PCR products. The two PCR products were gel-purified with the Zymoclean Gel DNA Recovery Kit (Zymo Research Corp, Irvine, CA) and ligated together via Gibson assembly, followed by transformation (*Transformation of genes*).

## **Transformation of genes**

Electrocompetent *Escherichia coli* (*E. coli*, E. cloni EXPRESS BL21(DE3) cells, Lucigen, Middleton, WI) were transformed with Gibson assembly products using a Gene Pulser Xcell (Bio-Rad, Hercules, CA). SOC medium (750  $\mu\text{L}$ ) was added, and the cells were plated on LB-ampicillin (100  $\mu\text{g mL}^{-1}$ ) (LB-amp) agar plates. Overnight cultures (5 mL LB-amp in culture tubes) were grown at 37 °C and 230 rpm for 12–18 hours. Overnight cultures were used to inoculate flask cultures, prepare glycerol stocks, and isolate plasmids. Plasmids were isolated with Qiagen Miniprep kits, and the genes were sequence-verified (T7 promoter / terminator sequencing primers, Laragen, Inc.).

## **96-Well plate library expression**

Single colonies from the LB-amp agar plates were picked using sterile toothpicks and grown in 400  $\mu\text{L}$  LB-amp in 2-mL 96-well deep-well plates at 37 °C, 250 rpm, 80% humidity overnight (12–18 hours). Multi-channel pipettes were used to transfer 50  $\mu\text{L}$  of starter culture into deep-well plates containing 1 mL Hyperbroth (AthenaES) with 100  $\mu\text{g mL}^{-1}$  ampicillin (HB-amp) per well. The deep-well expression culture plate was incubated at 37 °C, 250 rpm, 80% humidity for 2.5 hours. The plate was then chilled on ice for 30 minutes. Expression of the protein of interest was induced with isopropyl- $\beta$ -D-thiogalactoside (IPTG) and cellular heme production was increased with 5-aminolevulinic acid (ALA); an induction mixture containing IPTG and ALA in HB-amp (50  $\mu\text{L}$ ) was added to each well for a final concentration of 0.5 mM IPTG and 1 mM ALA, 1.1 mL culture volume. The plate was incubated at 22 °C and 250 rpm overnight.

## **96-Well plate reactions and screening**

96-Well deep-well plates containing *E. coli* expressing a protein variant were centrifuged at 4000  $\times g$  for 10 minutes at 4 °C. The supernatant was discarded, and pellets were resuspended in nitrogen-free M9 minimal media (M9-N, 380  $\mu\text{L}$ ) and brought into a Coy vinyl anaerobic chamber

(0–20 ppm O<sub>2</sub>). To each well was added 4,4,5,5-tetramethyl-2-vinyl-1,3,2-dioxaborolane (**1**) (10 μL, 400 or 800 mM in ethanol) and EDA (**2**) (10 μL, 400 or 800 mM in ethanol) and sealed with an adhesive foil cover. The reactions were shaken in the Coy chamber at room temperature for 16 hours. The plate was then removed from the Coy chamber and the reactions were worked up: to each reaction internal standard (20 mM 4-methyl anisole in 1:1 hexane:ethyl acetate, 20 μL) and 1:1 hexane:ethyl acetate (480 μL) were added. A silicone plate sealing mat (AWSM1003S, ArcticWhite) was used to cover the plate, and the plate was vortexed. The plate was then centrifuged (4000 × g, 10 minutes) and the organic layer was transferred to GC vial inserts in GC vials. The samples were assayed for product formation via GC.

Wells with improved activity relative to the parent protein were streaked out onto LB-amp agar plates. A single colony from the plate was picked and grown in 5 mL LB-amp overnight (230 rpm, 37 °C). These overnight cultures were used in flask protein expression and small-scale biocatalytic reactions to verify enhanced activity and/or selectivity relative to the parent sequence.

### **Medium- and large-scale protein expression**

Cultures of 50 mL HB-amp (50 mL in 250-mL unbaffled flasks or 1 L in 2.8-L unbaffled flasks) were inoculated 1% (v/v) with saturated overnight cultures and shaken in an Inova 42 shaker at 230 rpm, 37 °C or an NFors shaker at 160 rpm, 37 °C. At an approximate OD<sub>600</sub> of 1.0 – 1.2, cultures were chilled on ice for 30 minutes. Protein expression was induced with 1 M IPTG to a final concentration of 0.5 mM, and heme production was enhanced with supplementation of 1 M ALA to a final concentration of 1 mM. The cultures were shaken in an Inova 42 shaker at 230 rpm (160 rpm for NFors) at 22 °C overnight (18–24 hours). Cells were transferred to 50-mL Falcon tubes or 500-mL centrifuge buckets and pelleted via centrifugation at 4000 × g for 10 minutes at 4 °C.

## Lineage validation reactions

Small-scale reactions were set up in 2-mL GC crimp vials. *E. coli* expressing the appropriate heme protein catalyst (380  $\mu$ L, adjusted to the appropriate optical density with 20 mM MOPS buffer, pH=7) was added to the vials; the vials were then brought into a Coy anaerobic chamber (0 – 20 ppm O<sub>2</sub>). To each vial was added **1** (10  $\mu$ L in ethanol, final concentration 20 mM) followed by **2** (10  $\mu$ L in ethanol final concentration 40 mM). Directly following addition of **2**, the reaction vial was crimped and shaken at 500 rpm at room temperature. Reactions were worked up for GC analysis. For GC analysis, 20  $\mu$ L 100% ethanol and an internal standard (20  $\mu$ L 4-ethyl anisole, 20 mM in 1:1 hexane:ethyl acetate) were added to the vial. A 1:1 mixture of hexane:ethyl acetate (480  $\mu$ L) was added, and the reaction was transferred into 1.7-mL Eppendorf tubes for extraction. The liquid-liquid extraction was carried out with a Retsch MM 301 mixing mill (1 minute, 30 Hz / 1800 rpm). Samples were centrifuged at 20817  $\times$  g for 5 minutes at RT, and the organic layer was used for chromatographic analysis.

## Sonicated lysate preparation and hemochrome assay

An aliquot of cells adjusted to a density of OD<sub>600</sub>= 30 for protein concentration determination was sonicated (QSonica Q500 Sonicator, 1/8 in tip) for 4 minutes, 1 second on, 1 second off at 30% amplitude. Following sonication, a 0.1 volume equivalent Bugbuster 10X protein extraction reagent (EMD Millipore) was added; the lysate-Bugbuster mixture was incubated 15 minutes at room temperature to improve cell lysis. The sonicated lysate was clarified via centrifugation at 20817  $\times$  g and 4 °C for 10 minutes. The concentration of heme-loaded protein was determined with the pyridine hemochromagen (hemochrome) assay.<sup>2</sup> Briefly, sonicated and clarified lysate (500  $\mu$ L) was added to a cuvette. 500  $\mu$ L of solution I (0.2 M NaOH, 40% (v/v) pyridine, 500  $\mu$ M potassium ferricyanide) were added, and the spectrum of this oxidized sample was taken from 350 – 600 nm. Sodium dithionite (10  $\mu$ L of 0.5 M solution in 0.5 M NaOH) was added, and the

reduced spectrum was taken from 350 – 600 nm. The pyridine hemochromagen concentration was determined using its Q bands, with the literature extinction coefficient of  $23.98 \text{ mM}^{-1} \text{ cm}^{-1}$  for ( $557 \text{ nm}_{\text{reduced}} - 540 \text{ nm}_{\text{oxidized}}$ ).

### **Preparation of lyophilized whole-cell biocatalysts**

Cultures of 1 L HB-amp in unbaffled 2.8-L Erlenmeyer flasks were inoculated 1% (v/v) with saturated overnight cultures and shaken in an Infors shaker at 160 rpm, 37 °C. At an approximate  $\text{OD}_{600}$  of 1.0, cultures were chilled on ice for 30 minutes. Protein expression was induced with 1 M IPTG to a final concentration of 0.5 mM, and heme production was enhanced with supplementation of 1 M ALA to a final concentration of 1 mM. The cultures were shaken in an Infors shaker at 230 rpm and 22 °C overnight (18–24 hours). Cells were transferred to 500-mL centrifuge buckets and pelleted via centrifugation at  $4000 \times g$  for 10 minutes at 4 °C. The pellet was transferred into a 50-mL Falcon tube and flash-frozen on powdered dry ice. The lid of the 50-mL Falcon tube was replaced by a Kimwipe held on with a rubber band and put in a lyophilization bottle. Pellets were lyophilized for at least 36 hours on a Labconco 4.5 plus freezezone benchtop lyophilizer (Labconco, Kansas City, MO) to ensure complete lyophilization. Following lyophilization, the cells were transferred into a sealable plastic bag and ground into a fine powder using a mortar and pestle. The lyophilized cell powder was stored at -20°C until use.

### **Preparative-scale biocatalytic reactions with lyophilized whole cells**

To a 250-mL round-bottom flask, lyophilized whole cells harboring *RmaNOD* THRAW (500 mg, 10 mg  $\text{mL}^{-1}$  final loading) were added, and the catalyst was reconstituted in aqueous buffer (KPi 50 mM pH 7.5, 37.6 mL) by stirring for 10 minutes. The round-bottom flask was sealed with a septum and cycled with Argon and vacuum three times. To the reconstituted catalysts **2** (174.8  $\mu\text{L}$  neat substrate, 1 mmol, final concentration 20 mM) was added. Using a syringe pump, EDA



solution (5 mL EtOH, 7 mL dH<sub>2</sub>O, 241.8  $\mu$ L EDA, solution concentration 163 mM, final concentration 40 mM) was added over a period of 2 hours, using a long Luer lock needle such that the needle sat below the solution surface during addition. The reaction was stirred at room temperature overnight (16-hour total reaction time). Upon reaction completion the reaction mixture was stored at -20 °C until workup.

#### **Procedure for column-free purification of trifluoroborated product (4)**

The trifluoroboration procedure was adapted from Bagutski et al.<sup>3</sup> Following completion of the reaction, the product was extracted three times with 1:1 hexane:ethyl acetate (1 volume equivalent per extraction). Deprotected boronic acid was reprotected by stirring the organic layer with pinacol (236 mg, 2 mmol, 2.0 equivalents) in the presence of magnesium sulfate (1 g). The solution was filtered to remove magnesium sulfate and concentrated *in vacuo* to remove solvent and remaining unreacted starting material (**1**). The crude reaction mixture was then resuspended in 1:1 methanol:water (8 mL total). Potassium bifluoride solution (1.67 mL of 3 M solution in water, 5.0 equivalents) was added dropwise, and the reaction was left to proceed overnight. The solution was concentrated *in vacuo* to remove methanol, flash-frozen, and lyophilized to remove water. Acetone (20 mL) was added to the 40-mL scintillation vial containing the lyophilized crude product. The product was resuspended in the sealed scintillation vial via sonication bath. The solution was vacuum filtered, and the vial was washed three times with acetone (10 mL). The acetone was removed *in vacuo*, and the product was dried on high vacuum.

#### **Determination of initial cyclopropanation activity**

A panel of heme proteins previously engineered for carbene or nitrene transfer reactions<sup>4</sup> was tested for their activity in formation of **3**; proteins included in these tests were expressed in 96-

well plates (*96-Well plate library expression*). The cell pellets were resuspended with 400  $\mu$ L 1x-M9-N, and the 96-well deep-well plates were brought into a Coy anaerobic chamber. A substrate solution (50  $\mu$ L of 180 mM **1** and **2**, 20 mM final concentration of each substrate) was added to each well. The plates were sealed with an adhesive foil cover and shaken at 500 rpm and room temperature for 6 hours. The foil covers were removed, 500  $\mu$ L cyclohexane were added to each well, and the plates were sealed with silicone sealing mats and vortexed for 1 minute. The organic layer was transferred to GC vial inserts in GC vials, and formation of **3** was analyzed via GC-MS. The highest initial activity for *cis*-selective cyclopropanation was exhibited by *ApePgb* W59A Y60G F145W and its intermediate variants. The highest initial activity for *trans*-selective cyclopropanation was produced by *RmaNOD* Q52A, *RmaNOD* Q52V, and Mb H64V V68A; *ApePgb* W59A Y60G F145W and *RmaNOD* Q52V were variants engineered for unactivated alkene cyclopropanation.<sup>5</sup> The myoglobin was engineered by the Fasan lab for styrenyl alkene cyclopropanation reactions.<sup>6</sup>

Table S1. Protein scaffolds tested for formation of **3** in the initial-activity screening.

UniProt ID of wild type	Organism	Annotation	Number of variants tested	Formation of <b>3</b>
B3FQS5	<i>Rhodothermus marinus</i>	Cytochrome <i>c</i>	9	Trace detected
P15452	<i>Hydrogenobacter thermophilus</i>	Cytochrome <i>c</i> -552	2	n.d.
P00080	<i>Rhodopila globiformis</i>	Cytochrome <i>c</i> 2	2	n.d.
D0MGT2	<i>Rhodothermus marinus</i>	Nitric oxide dioxygenase	10	Detected
P02185	<i>Physeter catodon</i>	Myoglobin	1	Detected
B3DUZ7	<i>Methylococcus thermophilus</i>	Hell's gate globin	2	Trace detected
O66586	<i>Aquifex aeolicus</i>	Thermoglobin	1	Trace detected
G7VHJ7	<i>Pyrobaculum ferrireducens</i>	Protoglobin	1	Trace detected
Q0PB48	<i>Campylobacter jejuni</i>	Truncated hemoglobin	1	n.d.
Q9YFF4	<i>Aeropyrum pernix</i>	Protoglobin	10	Detected
O31607	<i>Bacillus subtilis</i>	Truncated hemoglobin	1	n.d.
P14779	<i>Bacillus megaterium</i>	Cytochrome P450 <sub>BM3</sub>	50	n.d.

Table S2. Activity of top variants from initial screening. Activity is given as integrated area for the target ions 182 m/z and 195 m/z. \*Variant was found during protein engineering for the referenced project. n.d.: not detected in given ion channel.

Variant	Reference	<i>cis</i> (182 m/z)	<i>trans</i> (182 m/z)	<i>cis</i> (195 m/z)	<i>trans</i> (195 m/z)
<i>ApePgb</i> WT	5	122	n.d.	n.d.	n.d.
<i>ApePgb</i> Y60G	5	1865	n.d.	n.d.	n.d.
<i>ApePgb</i> Y60G F145N	5*	32478	896	2620	1314
<i>ApePgb</i> W59A Y60G	5	5375	n.d.	351	n.d.
<i>ApePgb</i> W59A Y60G F145W	5	30043	n.d.	2513	204
<i>Mb</i> H64V V68A	6	423	1095	n.d.	1437
<i>RmaNOD</i> WT	5	n.d.	576	n.d.	251
<i>RmaNOD</i> Q52V	5	193	1039	n.d.	1074
<i>RmaNOD</i> Q52A	5*	n.d.	1417	n.d.	1767

## Calibration curves for analytical-scale concentration determinations

To generate a calibration curve for analytical-scale reactions, solutions at varying concentrations of the cyclopropane authentic standard were prepared in *E. coli* cell resuspensions (20 mM MOPS, pH = 7.01, 4.8% ethanol, OD<sub>600</sub> = 28 for *cis*-**3**, OD<sub>600</sub> = 30 for *trans*-**3**). Two 420- $\mu$ L aliquots of each solution at each concentration were worked up following the protocols in *Small-scale, whole-cell biocatalytic reaction preparation and work-up*, skipping the addition of 20  $\mu$ L 100% EtOH, and the worked-up sample was analyzed by GC-MS. The y-axis of each calibration curve gives the product:internal standard ratio, calculated as total ion count (TIC) of the GC-MS peak corresponding to the target product divided by the TIC of the GC-MS peak corresponding to the internal standard. The x-axis of each calibration curve is the concentration corresponding to the product:internal standard ratio. The resulting calibration curves reproducibly display non-linearity, in which low concentrations of cyclopropane show a diminished response. While this indicates a product loss event, the standard curve reproducibility shows that it can be utilized to reliably calculate analytical-scale yields after fitting with an exponential function.

## Protein Engineering Strategies

### *Cis*-selective lineage

As *ApePgb* AGW was the most *cis*-diastereoselective variant with good activity, it was selected as the initial parent enzyme for the *cis*-selective lineage. Single-site-saturation libraries were generated for *ApePgb* AGW at active-site residues 63, 86, and 90. Hits (wells identified to have enhanced activity) were identified in screening, but subsequent validation of protein variants identified as hits in screening failed to verify improved variants. At this point, we would typically target additional active-site residues for mutagenesis. In parallel, however, screening *RmaNOD* Q52A for enhanced *trans*-cyclopropylboronate activity yielded variant *RmaNOD* Y32T Q52A with 420 TTN and an inversion of diastereoselectivity from 17:83 *cis:trans* to 90:10 *cis:trans*. We

determined that *RmaNOD* Y32T Q52A produced the same *cis* enantiomer as *ApePgb* AGW with five-fold higher *cis-3* formation, and therefore used *RmaNOD* Y32T Q52A as the parent protein for further evolution on the *cis*-selective lineage. The crystal structure solved for the related protein variant, *RmaNOD* Q52V, was used to select sites for mutagenesis (Figure S1).

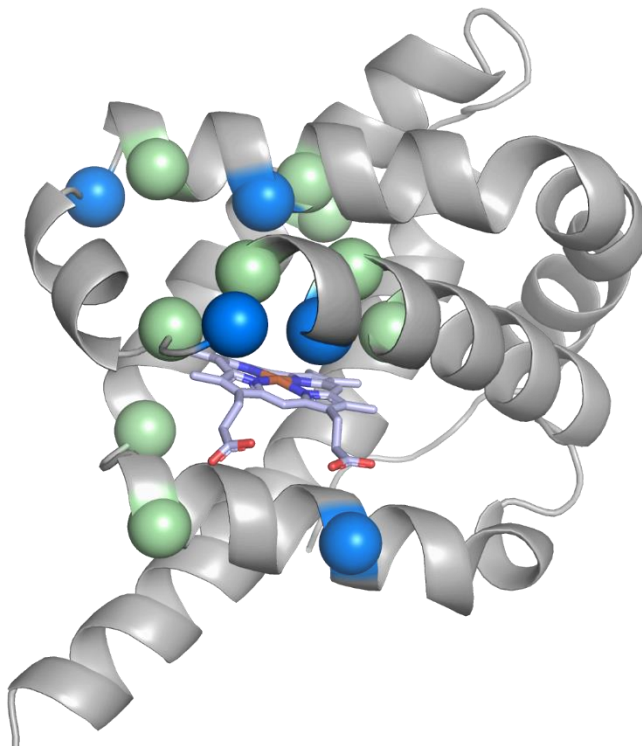


Figure S1. Amino-acid residues targeted in the *cis*-selective *RmaNOD* evolutionary lineage. The  $\alpha$ -carbons of the five fixed mutations relative to *RmaNOD* WT (Y32T, Y39H, L48R, Q52A, and R79W) are shown as blue spheres. The  $\alpha$ -carbons of other residues investigated as part of the *cis*-lineage are shown as green spheres.

As we had thus far only targeted a small subset of the active-site residues, in order to screen a larger sequence space without increasing the overall screening time with a medium-throughput GC-based screening method, we opted to reduce redundancy by screening 44 clones per library

rather than 88. Screening only 44 clones for a 22-member library still has an 86% probability of full library coverage (compared to 98% probability for 88 clones per library) while allowing for higher sequence space coverage (Figure S2).

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B						Parents	sterile					
C												
D			Library 1							Library 2		
E												
F												
G												
H												

Figure S2. Representative 96-well screening plate layout for screening two libraries per plate.

Using *RmaNOD* Y32T Q52A as parent, in a second round of site-saturation mutagenesis we targeted residues M31, F36, Y39, F46, L48, P49, I53, L56, R79, R86, V89, and L101, screening one half-plate (44 colonies) per library. This round of engineering yielded hits at positions 39, 48, 79, and 86. Residues 39 and 48 are in the distal heme pocket. R79 and R86 coordinate with the heme carboxylate; the underlying cause for the activity enhancement at these positions is unclear. It is noteworthy that the mutations at R79 significantly enhanced activity at the cost of diastereoselectivity, while mutation of residues 39 and 48 enhanced diastereoselectivity with no significant effect on activity. We therefore recombined the mutations at these positions. The combination of mutations Y32T, Y39H, L48R, Q52A, and R79W was found under screening conditions to both improve activity and stereoselectivity.

### ***Trans*-selective lineage**

Site-saturation mutagenesis libraries were generated at positions 31, 36, 46, 49, 53, 56, 79, 89, and 101, using *RmaNOD* Q52A as the parent sequence and the crystal structure of *RmaNOD* Q52V to guide site selection (Figure S3). As with the *cis*-lineage, we screened 44 colonies per library for this and all future site-saturation mutagenesis libraries in the *trans*-lineage. By

screening these libraries, we identified mutation L101N, which improved diastereoselectivity for the *trans*-product without significantly affecting enzyme activity.

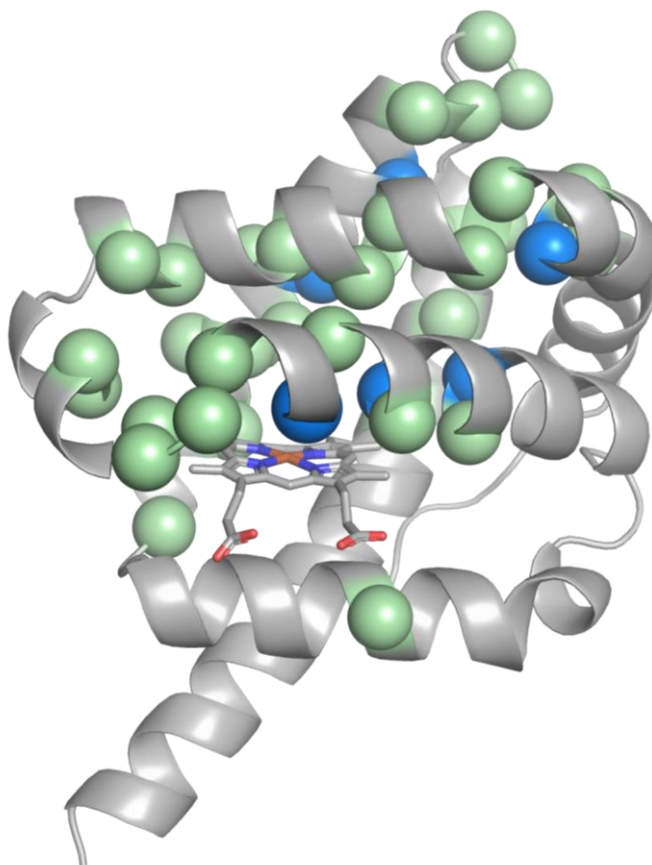


Figure S3. Amino-acid residues targeted in the *trans*-selective *RmaNOD* evolutionary lineage. The  $\alpha$ -carbons of the six fixed mutations relative to *RmaNOD* WT (L20W, Q52A, L56I, L60H, L101N, and I105M) are shown as blue spheres. The  $\alpha$ -carbons of other residues investigated as part of the *trans*- lineage are shown as green spheres.

In the next round of evolution, site-saturation mutagenesis libraries were generated at positions 35, 42, 53, 56, 60, 79, 96, 97, and 125, using *RmaNOD* Q52A L101N as the parent sequence. From this second round of site-saturation mutagenesis, beneficial mutations were identified at positions 56 and 60, with the optimal mutation (defined as boosting activity the most while retaining diastereoselectivity) was identified as L60H.

Because the activity of *RmaNOD* Q52A L60H L101N was still very low, we continued with two further rounds of site-saturation mutagenesis. Using *RmaNOD* Q52A L60H L101N as the parent sequence, in the first round, we targeted positions 27, 28, 43, 47, 48, 55, 59, 94, 105, 111, and 121, identifying beneficial mutations at positions 55 and 105. We decided to move forward with the mutation I105M. In the final round of site-saturation mutagenesis, we targeted positions 16, 19, 20, 23, 24, 27, 28, 31, 36, 46, 49, 56, 79, 98, 100, 108, 109, 112, 117, and 118, now using *RmaNOD* Q52A L60H L101N I105M as the parent sequence. From these libraries, we identified mutations L20W, M31F, L56A, and L56I, all of which boosted activity of the enzyme. By constructing and screening a recombination library of these four mutations, we identified *RmaNOD* L20W Q52A L56I L60H L101N I105M as the combination with the greatest increase in activity, while maintaining near-perfect diastereoselectivity.

## **Protein Crystallography**

### **Cloning, expression, and purification of for crystallography**

DNA encoding *RmaNOD* Q52V was subcloned into a pET22b vector backbone encoding an N-terminally His-tagged SUMO fusion through PCR amplification and Gibson assembly<sup>7</sup> to generate “pSUMO *RmaNOD* Q52V”. BL21(DE3) *E. coli* cells (Lucigen) were transformed with pSUMO *RmaNOD* Q52V. A single colony was used to start an overnight culture in LB-amp, which was grown overnight at 37 °C, 230 rpm. HB-amp (1 L in an unbaffled 2.8-L flask) was inoculated with 1% (v/v) overnight culture and shaken at 37 °C, 160 rpm (NFors shaker). At OD<sub>600</sub> = 1.0 the flask was put on ice for 30 minutes, induced with final concentration of 0.5 mM IPTG, and supplemented with a final concentration of 1 mM 5-ALA. The culture was shaken overnight at 22 °C, 160 rpm. Following expression, the culture was centrifuged at 4000 × g for 10 minutes at 4 °C and frozen at -20 °C.



The pellet was thawed on ice and resuspended in resuspension buffer (25 mM Tris pH 7, 100 mM NaCl). The cells were lysed via sonication (QSonica Q500 Sonicator). The lysate was clarified via centrifugation (15 minutes, 20817 × g) and filtered through 0.22- $\mu$ m filters. The protein was purified on an Akta purifier (GE Healthcare) with a 1-mL HisTrap column running a linear gradient between resuspension buffer (25 mM Tris pH 7, 100 mM NaCl, 20 mM imidazole) and elution buffer (25 mM Tris pH 7, 100 mM NaCl, 500 mM imidazole). Fractions containing protein of interest were pooled and buffer exchanged into 25 mM Tris pH 7, 25 mM NaCl, 5 mM DTT. His-tagged Ulp1 protease (a gift from Christopher Bley, Hoelz lab) was added, and the proteolysis was allowed to proceed overnight. Following proteolysis, the sample was passed through two hand columns with Nickel-NTA agarose bulk resin (Qiagen) to remove the cleaved His-tagged SUMO and His-tagged Ulp1 protease. The flow through containing the protein of interest was buffer exchanged and concentrated in 25 mM Tris pH 7, 25 mM NaCl buffer, and flash-frozen in aliquots.

### **Protein crystallization**

Sparse-matrix screens were set up using 0.2  $\mu$ L protein and 0.2  $\mu$ L well solution using a Gryphon robotic platform (Art Robbins Instruments). Quasicrystals were found in Crystal HT G10 (Hampton Research, 50 mM Cd(II)SO<sub>4</sub>, 0.1 M HEPES pH 7.5, 1.0 M sodium acetate trihydrate). Testing other divalent metal salts resulted in Cu(II)SO<sub>4</sub> as an equivalent replacement for the cadmium salt. Refining from this condition using 24-well sitting-drop well plates (Hampton Research), 2  $\mu$ L protein solution and 2  $\mu$ L well solution yielded crystals in approximately one week in 40 mM Cu(II)SO<sub>4</sub>, 0.1M HEPES pH 7.6, 1.2 M sodium acetate trihydrate. These crystals were cryoprotected with well solution supplemented with glycerol (increased by steps of 5% sequentially up to 25% final concentration glycerol) and flash-frozen in liquid nitrogen.

## Data collection and refinement

Diffraction data were collected on the Stanford Synchrotron Radiation Laboratory Beamline 12-2. As there are no previously reported structures for the *RmaNOD* scaffold, the data were collected as single anomalous diffraction using the iron K edge. Data were processed using XDS<sup>8</sup>, and three data sets were merged to increase anomalous signal redundancy. The data were used with AutoSol<sup>9</sup> for experimental phasing and model building. Following model building and initial refinements in PHENIX,<sup>10</sup> chain A of this model was used as a search model. The highest quality data set was determined in AIMLESS<sup>11</sup> to have a high-resolution shell limit of 2.45 Å. Molecular replacement was performed using Phaser-MR<sup>12</sup> with the merged data set model as the search model. The model was iteratively refined using PHENIX refine and COOT<sup>13</sup> and validated with MolProbity. Crystallographic and model statistics are described in Supplementary table S3.

Table S3. Crystal data and refinement statistics for *RmaNOD Q52V*.

Structure name	Engineered carbene transferase <i>RmaNOD Q52V</i> , putative nitric oxide dioxygenase from <i>Rhodothermus marinus</i>
PDB ID	6WK3
Wavelength	1.74 Å
Resolution range	43.65 - 2.45 (2.538 - 2.45)
Space group	C 1 2 1
Unit cell	a = 124.269, b = 102.472, c = 84.273 $\alpha = 90, \beta = 98.529, \gamma = 90$
Total reflections	248224 (25486)
Unique reflections	36943 (3600)
Multiplicity	6.7 (7.1)
Completeness (%)	95.98 (94.17)
Mean I/sigma(I)	13.34 (1.53)
Wilson B-factor	58.05
R-merge	0.09393 (1.464)
R-meas	0.1022 (1.581)
R-pim	0.03949 (0.5911)
CC1/2	0.997 (0.647)
Reflections used in refinement	36927 (3600)
Reflections used for R-free	1816 (177)
R-work	0.2077 (0.2991)
R-free	0.2467 (0.3343)
CC(work)	0.966 (0.785)
CC(free)	0.959 (0.723)
Number of non-hydrogen atoms	4680
macromolecules	4481
Ligands	198
Solvent	1
Protein residues	578
RMS(bonds)	0.004
RMS(angles)	0.74
Ramachandran favored (%)	99.30
Ramachandran allowed (%)	0.70
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	4.49
Average B-factor	60.86
macromolecules	61.08
Ligands	55.85
Solvent	55.94
Number of TLS groups	1

# Small-Molecule Crystal Formation and Data Collection

## Small-molecule crystal formation

A small sample (~1 mg) of **4** produced from *Rma*NOD THRAW preparative-scale reactions was recrystallized from slow evaporation of acetone to provide crystals suitable for X-ray diffraction.

## X-Ray refinement details

Low-temperature diffraction data ( $\phi$ - and  $\omega$ -scans) were collected on a Bruker AXS D8 VENTURE KAPPA diffractometer coupled to a PHOTON II CPAD detector with either Mo- $K\alpha$  radiation ( $\lambda = 0.71073 \text{ \AA}$ ) from a fine-focus sealed X-ray tube. All diffractometer manipulations, including data collection integration, and scaling were carried out using the Bruker APEXII software.<sup>14</sup> Absorption corrections were applied using SADABS. The structure was solved by intrinsic phasing using SHELXT<sup>15</sup> and refined against  $F^2$  on all data by full-matrix least squares with SHELXL-2014<sup>16</sup> using established refinement techniques.<sup>17</sup> All non-hydrogen atoms were refined anisotropically. All hydrogen atoms were included into the model at geometrically calculated positions and refined using a riding model. The isotropic displacement parameters of all hydrogen atoms were fixed to 1.2 times the U value of the atoms they are linked to (1.5 times for methyl groups). Absolute configuration was determined by anomalous dispersion.<sup>18</sup> Graphical representation of the structure with 50% probability thermal ellipsoids was generated using Mercury visualization software. The crystal data and structure refinement for this compound (noted as D19046\_b) is given in

Table S4. The atomic coordinate positions are provided as a separate supplemental file.

Table S4. Crystal data and structure refinement for D19046\_b.

Identification code	d19046_b
Empirical formula	C <sub>12</sub> H <sub>18</sub> B <sub>2</sub> F <sub>6</sub> K <sub>2</sub> O <sub>4</sub>
Formula weight	440.08
Temperature	100 K
Wavelength	0.71073 Å
Crystal system	Monoclinic
Space group	P 1 21 1
Unit cell dimensions	a = 9.9422(8) Å, α = 90°.
	b = 6.0314(5) Å, β = 103.236(2)°.
	c = 15.4530(12) Å, γ = 90°.
Volume	902.03(13) Å <sup>3</sup>
Z	2
Density (calculated)	1.620 Mg/m <sup>3</sup>
Absorption coefficient	0.599 mm <sup>-1</sup>
F(000)	448
Crystal size	0.39 x 0.28 x 0.02 mm <sup>3</sup>
Theta range for data collection	2.104 to 27.525°.
Index ranges	-12 ≤ h ≤ 12, -7 ≤ k ≤ 7, -20 ≤ l ≤ 20
Reflections collected	30884
Independent reflections	4145 [R(int) = 0.0306]
Completeness to theta = 25.242°	99.9 %
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.7456 and 0.6852
Refinement method	Full-matrix least-squares on F <sup>2</sup>
Data / restraints / parameters	4145 / 1 / 237
Goodness-of-fit on F <sup>2</sup>	1.054
Final R indices [I > 2σ(I)]	R1 = 0.0215, wR2 = 0.0560
R indices (all data)	R1 = 0.0225, wR2 = 0.0565
Absolute structure parameter	0.015(7)
Extinction coefficient	n/a
Largest diff. peak and hole	0.435 and -0.249 e.Å <sup>-3</sup>

The crystal of **4** generated from *cis*-**3** purified from a reaction with *Rma*NOD THRAW was solved as (1*R*, 2*S*). This configuration is in agreement with the absolute configuration assignment by analogy when comparing the enantiomer produced by *Rma*NOD Y32T Q52A and *Ape*Pgb AGW (based on *Ape*Pgb AGW's enantioselective cyclopropanation of 1-octene and EDA).<sup>5</sup>

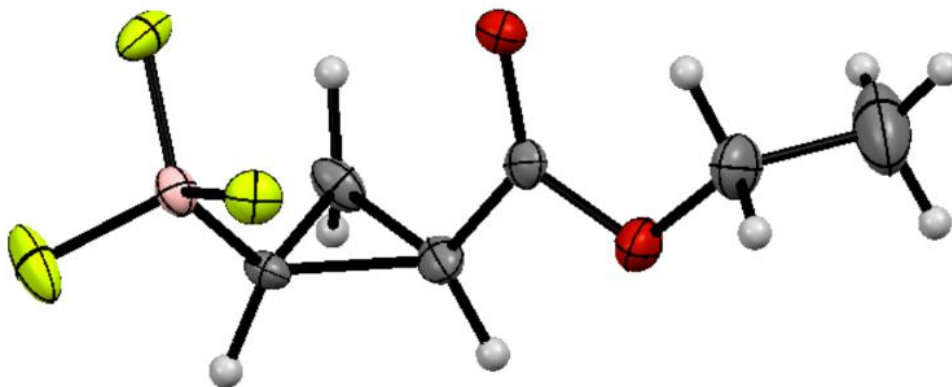
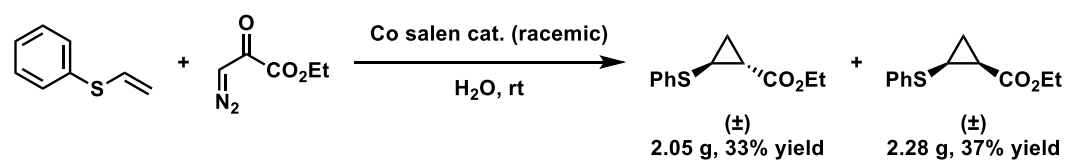


Figure S4. Small-molecule crystal structure of the  $\text{BF}_3\text{K}$ -substituted cyclopropane product **4** generated by the RmaNOD-THRAW-catalyzed reaction of **1** and **2**.

## Compound Synthesis and Characterization

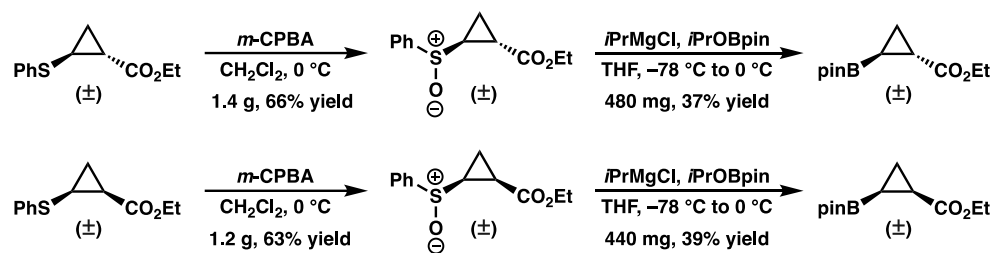
Racemic standards of *cis*-**3** and *trans*-**3** were prepared according to a three-step procedure by Bull and coworkers.<sup>19</sup>

### Ethyl (*trans*)-2-(phenylthio)cyclopropane-1-carboxylate and ethyl (*cis*)-2-(phenylthio)cyclopropane-1-carboxylate



A round-bottom flask containing a racemic mixture of *N,N'*-bis(3,5-di-*tert*-butylsalicylidene)-1,2-cyclohexanediaminocobalt(II) (0.043 equiv, 726 mg total, 363 mg of each enantiomer) was flushed with Ar for 15 minutes, and then 50 mL of degassed water were added (water was degassed by bubbling Ar gas through the liquid for 30 minutes). The ethyl diazoacetate (1.0 equiv, 2.9 mL) and phenyl vinyl sulfide (1.5 equiv, 4.7 mL) were added via syringe, and the reaction was

stirred under a positive pressure of Ar (balloon) at 40 °C for 24 hours. Upon completion, the reaction was cooled to room temperature and 12 mL of isohexane was added. Air was bubbled through the reaction for 30 minutes to oxidize the Co catalyst. The crude mixture was then filtered through a pad of silica gel, which was subsequently washed with CH<sub>2</sub>Cl<sub>2</sub>. The solution was concentrated and purified by column chromatography (silica, 15:1 pentane:Et<sub>2</sub>O) to provide 2.05 g of *cis*-**S1** (33% yield) and 2.28 g of *trans*-**S1** (37% yield).



Each diastereomer of **S1** was monooxidized with *m*-CPBA and subsequently converted to the boronic ester according to the following procedures.

### Ethyl (*trans*)-2-((*S*)-phenylsulfinyl)cyclopropane-1-carboxylate

The cyclopropane *trans*-**S1** (1.97 g) was added to a round-bottom flask, dissolved in 90 mL of CH<sub>2</sub>Cl<sub>2</sub>, and cooled to 0 °C. The *m*-CPBA (99%, 1.63 g) was added in three portions over 30 minutes and then continued to stir for 2 hours at 0 °C. Aqueous KOH (3 M, 35 mL) was added, and the layers were separated. The aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (5 x 35 mL). The combined organic layers were dried with MgSO<sub>4</sub>, filtered, and concentrated. The crude material was purified by column chromatography (silica, 1:2 hexane:Et<sub>2</sub>O) to provide 1.39 g of *trans*-**S2** (66% yield).



### **Ethyl (*cis*)-2-((*S*)-phenylsulfinyl)cyclopropane-1-carboxylate**

The cyclopropane *cis*-**S1** (1.78 g) was added to a round-bottom flask, dissolved in 80 mL of CH<sub>2</sub>Cl<sub>2</sub> and cooled to 0 °C. The *m*-CPBA (99%, 1.44 g) was added in three portions over 30 minutes and then continued to stir for 2 hours at 0 °C. Aqueous KOH (3 M, 30 mL) was added, and the layers were separated. The aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (5 x 30 mL). The combined organic layers were dried with MgSO<sub>4</sub>, filtered, and concentrated. The crude material was purified by column chromatography (silica, 1:2 hexane: diethyl ether) to provide 1.20 g of *cis*-**S2** (63% yield).

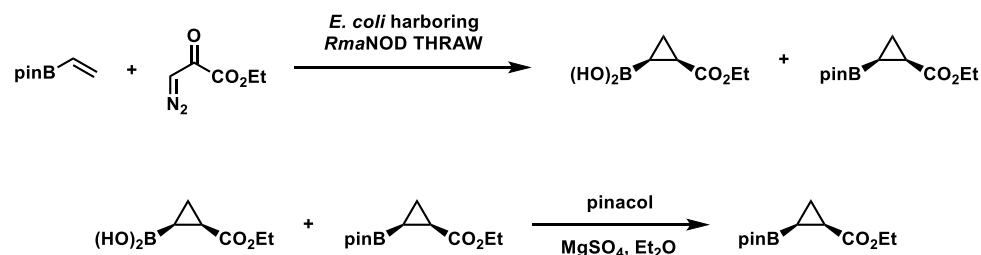
### **Ethyl (*cis*)-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)cyclopropane-1-carboxylate**

The cyclopropane *trans*-**S2** (1.29 g) was added to a round-bottom flask and dissolved in 54 mL of THF. The reaction was cooled to –78 °C, and then the *i*PrMgCl (2 M, 4.4 mL) was added dropwise. The reaction continued to stir for 10 minutes before the addition of 2-isopropoxy-4,4,5,5-tetramethyl-1,3,2-dioxaborolane (2.2 mL). The reaction was warmed to 0 °C and continued to stir for 6 hours before being warmed to room temperature. Methanol (15 mL) was added to quench the reaction and stirred for 5 minutes. Then 150 mL of diethyl ether were added, the mixture was filtered and concentrated. The crude mixture was purified by column chromatography (florisil, 10% to 50% diethyl ether : hexane) to provide 480 mg of *trans*-**3** (37% yield). <sup>1</sup>H NMR (400 MHz, chloroform-*d*) δ 4.19 – 3.98 (m, 2H), 1.83 – 1.73 (m, 1H), 1.26 – 1.21 (m, 12H), 1.12 – 1.04 (m, 1H), 1.04 – 0.96 (m, 1H), 0.43 – 0.31 (m, 1H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 174.5, 83.6, 60.6, 25.0, 25.0, 17.8, 14.4, 11.3. <sup>11</sup>B NMR (128 MHz, CDCl<sub>3</sub>) δ 32.1.

## Ethyl (*cis*)-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)cyclopropane-1-carboxylate

The cyclopropane *cis*-**S2** (1.10 g) was added to a round-bottom flask and dissolved in 46 mL of THF. The reaction was cooled to  $-78\text{ }^{\circ}\text{C}$ , and then the *i*PrMgCl (2 M, 3.5 mL) was added dropwise. The reaction continued to stir for 10 minutes before the addition of 2-isopropoxy-4,4,5,5-tetramethyl-1,3,2-dioxaborolane (1.9 mL). The reaction was warmed to  $0\text{ }^{\circ}\text{C}$  and continued to stir for 6 hours before being warmed to room temperature. Methanol (15 mL) was added to quench the reaction and stirred for 5 minutes. Then 150 mL of diethyl ether were added, the mixture was filtered and concentrated. The crude mixture was purified by column chromatography (florisil, 10% to 50% diethyl ether: hexane) to provide 440 mg of *cis*-**3** (39% yield).  $^1\text{H}$  NMR (400 MHz, chloroform-*d*)  $\delta$  4.18 – 4.04 (m, 2H), 1.79 – 1.69 (m, 1H), 1.21 (d,  $J = 1.7\text{ Hz}$ , 12H), 1.02 – 0.92 (m, 1H), 0.62 – 0.51 (m, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  174.3, 83.5, 60.5, 24.8, 24.7, 18.6, 14.3, 13.0.  $^{11}\text{B}$  NMR (128 MHz,  $\text{CDCl}_3$ )  $\delta$  32.58.

## Ethyl (1*R*,2*S*)-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)cyclopropane-1-carboxylate (enzymatic)

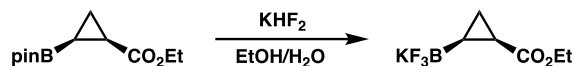


Whole *E. coli* cells were grown fresh as described above (*Medium- and large-scale protein expression*); two 1-L cultures in 2.8-L non-baffled flasks were grown for *RmaNOD THRAW*. The cells were resuspended in 1X M9-N buffer ( $\text{OD}_{600} = 30$ , 664 mL) and transferred into a 1-L round-bottom flask with a stir bar. The round-bottom flask was sealed with a septum and attached to a

Schlenk line; three cycles of Argon purging and vacuum evacuation were used to remove oxygen from the reaction flask. The mixture remained under argon flow throughout the entire reaction. Substrate **(1)** (neat stock, 7.5 mL, 60 mM final concentration) was added via syringe. EDA **(2)** (10.41 mL diluted in 35.9 mL ethanol cosolvent, 120 mM final concentration) was added over 2 hours via syringe pump slow addition. The reaction continued to stir overnight (16 hours) at room temperature. The reactions were then extracted three times with 1:1 hexane:ethyl acetate. The organic fractions were concentrated *in vacuo* and used in the subsequent protection step without purification. To the crude mixture of **3** and the deprotected cyclopropylboronic acid **(6)** were added pinacol (1.18 g), anhydrous magnesium sulfate (15 g), and anhydrous diethyl ether (100 mL). The reaction flask was sealed with a rubber septum and stirred for 2 hours at room temperature. After 2 hours, an aliquot was analyzed by <sup>1</sup>H NMR spectroscopy to ensure complete consumption of **6**. The reaction was filtered over a plug of anhydrous MgSO<sub>4</sub>, concentrated, and purified by distillation.

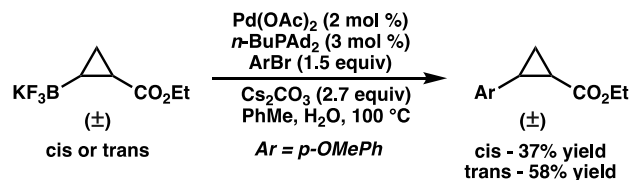
The crude mixture was purified via fractional distillation using a short path distillation setup attached to a small Vigreux column under reduced pressure. The residual alkene starting material was isolated at 10 torr and 40 °C to obtain 820 mg of the reagent as a colorless oil. The Vigreux column was then removed and a mixture of the desired product **3**, diethyl fumarate, and diethyl maleate were co-distilled at 5 torr and 70–100 °C. The obtained mixture was transferred to a round-bottom flask and vigorously stirred while heated under vacuum (40 °C at 1 torr for 24 hours, then 45 °C at 1 torr for 24 hours, then 50 °C at 1 torr for 3.5 hours) to remove the undesired diethyl fumarate and diethyl maleate side products. Product **3** was isolated as a colorless oil (3.7 g, 95% purity, 36% yield, 41% yield BRSM).

## Ethyl (1R,2S)-2-(trifluoro-*l*-boranyl)cyclopropane-1-carboxylate, potassium salt ((1R, 2S)-4)



To a dram vial equipped with a magnetic stir bar were added **3** (50 mg, 0.208 mmol, 1 equiv) and 1 mL of ethanol. A solution of KHF<sub>2</sub> (78 mg, 1 mmol, 5 equiv) in 0.22 mL of water was added to the vial, then sealed with a screw cap, and stirred overnight at room temperature. Upon completion, the solvent was removed *in vacuo* and the BF<sub>3</sub>K salt was extracted with acetone (3 x 4 mL). The combined organic layers were filtered through a plug of cotton and concentrated to provide 35 mg of the desired cyclopropyl BF<sub>3</sub>K salt (76% yield). A small sample (~1 mg) was recrystallized from slow evaporation of acetone to provide crystals suitable for X-ray diffraction. <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ 3.92 (qd, *J* = 7.1, 1.2 Hz, 2H), 1.32 – 1.22 (m, 1H), 1.12 (t, *J* = 7.1 Hz, 3H), 0.75 (ddd, *J* = 8.8, 4.9, 2.5 Hz, 1H), 0.56 (dddt, *J* = 10.8, 8.2, 2.4, 1.1 Hz, 1H), -0.13 (dddt, *J* = 14.7, 10.3, 8.7, 4.3 Hz, 1H). <sup>13</sup>C NMR (101 MHz, DMSO) δ 174.9, 59.1, 40.6, 40.4, 40.2, 40.0, 39.8, 39.6, 39.3, 17.3, 17.3, 17.3, 17.3, 14.7, 10.3, 10.3, 10.2, 10.2. <sup>11</sup>B NMR (128 MHz, DMSO-*d*<sub>6</sub>) δ 3.43, 3.00. <sup>19</sup>F NMR (282 MHz, DMSO-*d*<sub>6</sub>) δ -135.1, -135.2.

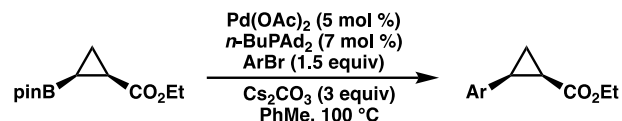
## Suzuki-Miyaura cross-coupling with **4**



In a glovebox under an inert atmosphere, **4** (4.4 mg, 0.02 mmol, 1 equiv) was dissolved in 110 μL of a mixture of 10:1 toluene:degassed water in a half dram vial equipped with a magnetic stir bar. The 4-bromoanisole (5.6 mg, 1.5 equiv), Pd(OAc)<sub>2</sub> (0.09 mg, 0.02 equiv), *n*-BuPAd<sub>2</sub> (0.22 mg, 0.03 equiv), and Cs<sub>2</sub>CO<sub>3</sub> (17.6 mg, 2.7 equivalents) were added sequentially. The vial was

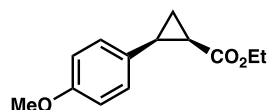
sealed with a Teflon cap, wrapped with electrical tape, and stirred at 100 °C for 20 hours. The reaction was cooled to room temperature, loaded onto a packed silica column, and analyzed by <sup>1</sup>H NMR against an internal standard to determine the product yield.

### General procedure: Suzuki-Miyaura cross-coupling with *cis*-**3**



In a glovebox under an inert atmosphere, *cis*-**3** (48 mg, 0.2 mmol, 1 equiv), aryl bromide (1.5 equiv), Pd(OAc)<sub>2</sub> (2.2 mg, 0.05 equiv), *n*-BuPAd<sub>2</sub> (5 mg, 0.07 equiv), Cs<sub>2</sub>CO<sub>3</sub> (3.0 equiv), and toluene (2 mL) were sequentially added to a 2-dram vial equipped with a magnetic stir bar. The vial was sealed with a Teflon cap, wrapped with electrical tape, and stirred at 100 °C for 36 hours. The reaction was cooled to room temperature, loaded onto a packed silica column, and purified by column chromatography to afford the desired product.

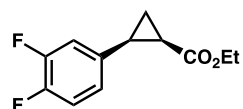
### Ethyl (1*R*,2*S*)-2-(4-methoxyphenyl)cyclopropane-1-carboxylate (**5a**)



Prepared from (1*R*,2*S*)-**3** (48 mg, 1.0 equiv, 0.2 mmol) and 1-bromo-4-methoxybenzene (56 mg, 1.5 equiv, 0.3 mmol). The crude residue was purified by column chromatography (silica, 10% Et<sub>2</sub>O/hexanes) to yield 40.4 mg of **5a** (92% yield). The product was analyzed by chiral SFC and no minor enantiomer peak was detected, confirming the product is formed in >99% ee. This is a known compound, and the spectroscopic data are in agreement with the literature.<sup>6</sup> Chiral SFC: (OJ-H, 2.5 mL/min, 5% IPA in CO<sub>2</sub>, λ = 254 nm): *t*<sub>R</sub> (major) = 3.9 min, *t*<sub>R</sub> (minor) = 5.5 min. <sup>1</sup>H NMR (400 MHz, chloroform-*d*) δ 7.21 – 7.15 (m, 2H), 6.80 (dt, *J* = 8.9, 2.5 Hz, 2H), 3.89 (q, *J* = 7.1 Hz, 2H), 3.77 (s, 3H), 2.52 (q, *J* = 8.7, 8.0 Hz, 1H), 2.03 (ddd, *J* = 9.2, 7.8, 5.6 Hz, 1H), 1.70 – 1.61 (m, 1H), 1.30 (ddd, *J* = 8.7, 7.8, 5.0 Hz, 1H), 1.02

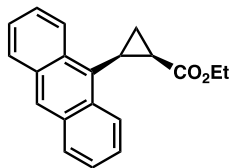
(t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.2, 158.4, 130.4, 128.7, 113.4, 60.3, 55.3, 25.0, 21.8, 14.2, 11.4. HRMS (EI+,  $m/z$ ): calculated for  $\text{C}_{13}\text{H}_{16}\text{O}_3$   $[\text{M}]^+$ : 220.1100, found 220.1095.

### Ethyl (1*R*,2*S*)-2-(3,4-difluorophenyl)cyclopropane-1-carboxylate (**5b**)



Prepared from (1*R*,2*S*)-**3** (48 mg, 1.0 equiv, 0.2 mmol) and 4-bromo-1,2-difluorobenzene (58 mg, 1.5 equiv, 0.3 mmol). The crude residue was purified by column chromatography (silica, 10%  $\text{Et}_2\text{O}$ /hexanes) to yield 29.7 mg of **5b** (66% yield). The product was analyzed by chiral SFC and no minor enantiomer peak was detected, confirming the product is formed in >99% ee. Chiral SFC: (OJ-H, 2.5 mL/min, 0% IPA in  $\text{CO}_2$ ,  $\lambda = 210$  nm):  $t_{\text{R}}$  (major) = 2.9 min,  $t_{\text{R}}$  (minor) = 3.4 min.  $^1\text{H}$  NMR (400 MHz, chloroform- $d$ )  $\delta$  7.13 – 6.93 (m, 3H), 3.92 (q,  $J = 7.1$  Hz, 2H), 2.56 – 2.45 (m, 1H), 2.07 (ddd,  $J = 9.3, 7.9, 5.7$  Hz, 1H), 1.63 (ddd,  $J = 7.4, 5.6, 5.2$  Hz, 1H), 1.34 (ddd,  $J = 8.7, 7.9, 5.3$  Hz, 1H), 1.05 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  170.8, 151.2, 151.1, 150.7, 150.5, 148.7, 148.6, 148.2, 148.1, 133.9, 133.8, 133.8, 133.8, 133.8, 125.5, 125.4, 125.4, 125.4, 118.4, 118.3, 116.8, 116.6, 60.6, 24.7, 24.7, 21.9, 14.2, 11.7.  $^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -138.7 (m), -140.7 (m). HRMS (EI+,  $m/z$ ): calculated for  $\text{C}_{12}\text{H}_{12}\text{F}_2\text{O}_2$   $[\text{M}]^+$ : 226.0805, measured 226.0805.

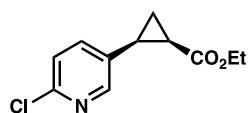
### Ethyl (1*R*,2*S*)-2-(anthracen-9-yl)cyclopropane-1-carboxylate (**5c**)



Prepared from (1*R*,2*S*)-**3** (48 mg, 1.0 equiv, 0.2 mmol) and 9-bromoanthracene (77 mg, 1.5 equiv, 0.3 mmol). The crude residue was purified by column chromatography (silica, 5-15%  $\text{Et}_2\text{O}$ /hexanes) to yield 25.5 mg of **5c** (44% yield). The product was analyzed by chiral SFC and no minor enantiomer peak was detected, confirming the product is formed in >99% ee. Chiral SFC: (OJ-H, 2.5 mL/min, 15% IPA in  $\text{CO}_2$ ,  $\lambda = 254$  nm):  $t_{\text{R}}$  (major) = 7.3 min,  $t_{\text{R}}$  (minor) = 7.9 min.  $^1\text{H}$  NMR (400 MHz,

Chloroform-*d*)  $\delta$  8.63 (d,  $J$  = 8.9 Hz, 1H), 8.50 (d,  $J$  = 8.7 Hz, 1H), 8.36 (s, 1H), 7.98 (t,  $J$  = 8.6 Hz, 2H), 7.51 – 7.42 (m, 4H), 3.55 – 3.34 (m, 2H), 3.09 (q,  $J$  = 8.3, 7.9 Hz, 1H), 2.57 (td,  $J$  = 8.3, 5.6 Hz, 1H), 2.01 (ddd,  $J$  = 8.8, 8.1, 4.6 Hz, 1H), 1.82 (ddd,  $J$  = 7.8, 5.6, 4.7 Hz, 1H), 0.54 (t,  $J$  = 7.1 Hz, 3H).  $^{13}\text{C}$  NMR (126 MHz,  $\text{CDCl}_3$ ): 171.8, 131.9, 131.8, 131.2, 129.7, 129.4, 129.0, 127.0, 125.6, 125.5, 125.3, 125.1, 125.0, 124.8, 60.1, 22.5, 20.9, 16.7, 13.6.

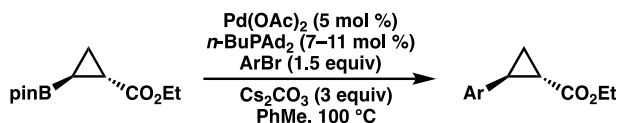
### Ethyl (1*R*,2*S*)-2-(6-chloropyridin-3-yl)cyclopropane-1-carboxylate (**5d**)



Prepared from (1*R*,2*S*)-**3** (48 mg, 1.0 equiv, 0.2 mmol) and 5-bromo-2-chloropyridine (58 mg, 1.5 equiv, 0.3 mmol). The crude residue was purified by column chromatography (silica, 5-40%  $\text{Et}_2\text{O}$ /hexanes) to yield 15 mg of **5d** (33% yield) as a colorless oil. The product was analyzed by chiral SFC, confirming the product is formed in >99% ee. Chiral SFC: (OJ-H, 2.5 mL/min, 10% IPA in  $\text{CO}_2$ ,  $\lambda$  = 254 nm):  $t_R$  (major) = 2.3 min,  $t_R$  (minor) = 2.6 min.

$^1\text{H}$  NMR (400 MHz, chloroform-*d*)  $\delta$  8.30 (dt,  $J$  = 2.5, 0.7 Hz, 1H), 7.53 (ddd,  $J$  = 8.2, 2.5, 0.7 Hz, 1H), 7.25 – 7.19 (m, 1H), 3.94 (qt,  $J$  = 7.1, 3.6 Hz, 2H), 2.48 (q, 1H), 2.14 (ddd,  $J$  = 9.0, 7.9, 5.7 Hz, 1H), 1.67 (dt,  $J$  = 7.5, 5.5 Hz, 1H), 1.42 (ddd,  $J$  = 8.7, 8.0, 5.2 Hz, 1H), 1.06 (t,  $J$  = 7.1 Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  170.6, 150.9, 149.8, 139.5, 131.5, 123.5, 60.8, 22.2, 21.8, 14.3, 11.4. HRMS (FAB,  $m/z$ ): calculated for  $\text{C}_{11}\text{H}_{12}\text{ClNO}_2$   $[\text{M}]^+$ : 225.0557, measured 225.0706.

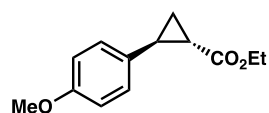
### Suzuki-Miyaura cross-coupling reactions with *trans*-**3**



In a glovebox under an inert atmosphere, *trans*-**3** (4.8 mg, 0.02 mmol, 1 equiv), aryl bromide (1.5 equiv),  $\text{Pd}(\text{OAc})_2$  (0.22 mg, 0.005 equiv), *n*-BuPAD<sub>2</sub> (0.5–0.8 mg, 0.07–0.11 equiv),  $\text{Cs}_2\text{CO}_3$  (3.0

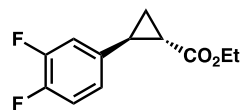
equiv), and toluene (0.2 mL) were sequentially added to a 1/2-dram vial equipped with a magnetic stir bar. The vial was sealed with a Teflon cap, wrapped with electrical tape, and stirred at 100 °C for 24 hours. The reaction was cooled to room temperature, flushed through a pipet over a short pad of silica gel, and concentrated *in vacuo*. The crude reaction was dissolved in CDCl<sub>3</sub> and analyzed by NMR versus an added standard (2,3,5,6-tetrachloronitrobenzene) to obtain an NMR yield of the product.

### Ethyl *trans*-2-(4-methoxyphenyl)cyclopropane-1-carboxylate



Prepared from *trans*-**3** (4.8 mg, 1.0 equiv, 0.02 mmol) and 1-bromo-4-methoxybenzene (5.6 mg, 1.5 equiv, 0.03 mmol). The crude residue was analyzed by NMR spectroscopy against an internal standard (2.0 mg of 2,3,5,6-tetrachloronitrobenzene, singlet at 7.73 ppm) to afford the reaction yield (56%). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 7.06 – 7.01 (m, 2H), 6.84 – 6.80 (m, 2H), 4.17 (q, *J* = 7.1 Hz, 2H), 3.79 (s, 3H), 2.49 (ddd, *J* = 9.2, 6.5, 4.2 Hz, 1H), 1.83 (ddd, *J* = 8.4, 5.2, 4.1 Hz, 1H), 1.56 (ddd, *J* = 9.4, 5.2, 4.4 Hz, 1H), 1.29 (t, *J* = 7.1 Hz, 3H), 1.28 – 1.24 (m, 1H).

### Ethyl *trans*-2-(3,4-difluorophenyl)cyclopropane-1-carboxylate



Prepared from *trans*-**3** (4.8 mg, 1.0 equiv, 0.02 mmol) and 4-bromo-1,2-difluorobenzene (5.8 mg, 1.5 equiv, 0.03 mmol). The crude residue was analyzed by NMR spectroscopy against an internal standard (3.5 mg of 2,3,5,6-tetrachloronitrobenzene, singlet at 7.73 ppm) to afford the reaction yield (61%). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 7.11 (dt, *J* = 10.2, 8.3 Hz, 1H), 6.94 (ddd, *J* = 11.4, 7.5, 2.3 Hz, 1H), 6.89 (ddt, *J* = 8.2, 3.9, 1.8 Hz, 1H), 4.23 (qd, *J* = 7.2, 0.8 Hz, 2H), 2.53 (ddd, *J* = 9.3, 6.4, 4.1 Hz, 1H), 1.91 (ddd, *J* = 8.5, 5.3, 4.2 Hz, 1H), 1.70 – 1.62 (m, 1H), 1.37 – 1.33 (m, 3H), 1.32 – 1.29 (m, 1H).



# Chromatographic Data

## Compound separation conditions

Table S5. Achiral separation conditions for compound **3**.

Compound	Instrument; column	Method program	Retention times
<b>3</b>	GC-FID; HP-5	90 °C isothermal 1 minute 15 °C/min to 110 °C 60 °C/min to 280 °C isothermal 1 minute	Internal standard: 2.9 minutes <i>Cis-3</i> : 4.6 min <i>Trans-3</i> : 4.7 min
<b>3</b>	GC-MS; HP-5ms UI	90 °C isothermal 1 minute 75 °C/min to 325 °C isothermal 1.5 minutes	Internal standard: 2.3 min <i>Cis-3</i> : 3.55 min <i>Trans-3</i> : 3.76 min

Table S6. Chiral separation conditions. Due to the worse separation conditions for *trans-3* enantiomers, a separate method was prepared when only *cis-3* enantioselectivity was to be measured.

Compound	Instrument; column	Method program	Retention times
<i>Cis-3</i>	GC-MS; CP-Chirasil Dex CB (25 m, 0.25 mm ID, 0.25 µm film thickness)	80 °C isothermal 2 minutes 10 °C/min to 130 °C, isothermal 15 minutes	<i>Cis</i> minor: 14.3 min, major: 14.8 min ( <i>trans</i> unresolved): 19.0 min
<i>Trans-3</i>	GC-MS; CP-Chirasil Dex CB (25 m, 0.25 mm ID, 0.25 µm film thickness)	90 °C isothermal 180 minutes	<i>Cis</i> major (1R, 2S): 87.4 min, minor: 98.9 min <i>Trans</i> major: 172.8 min, minor: 177.8 min

## Calibration curves for analytical-scale assays

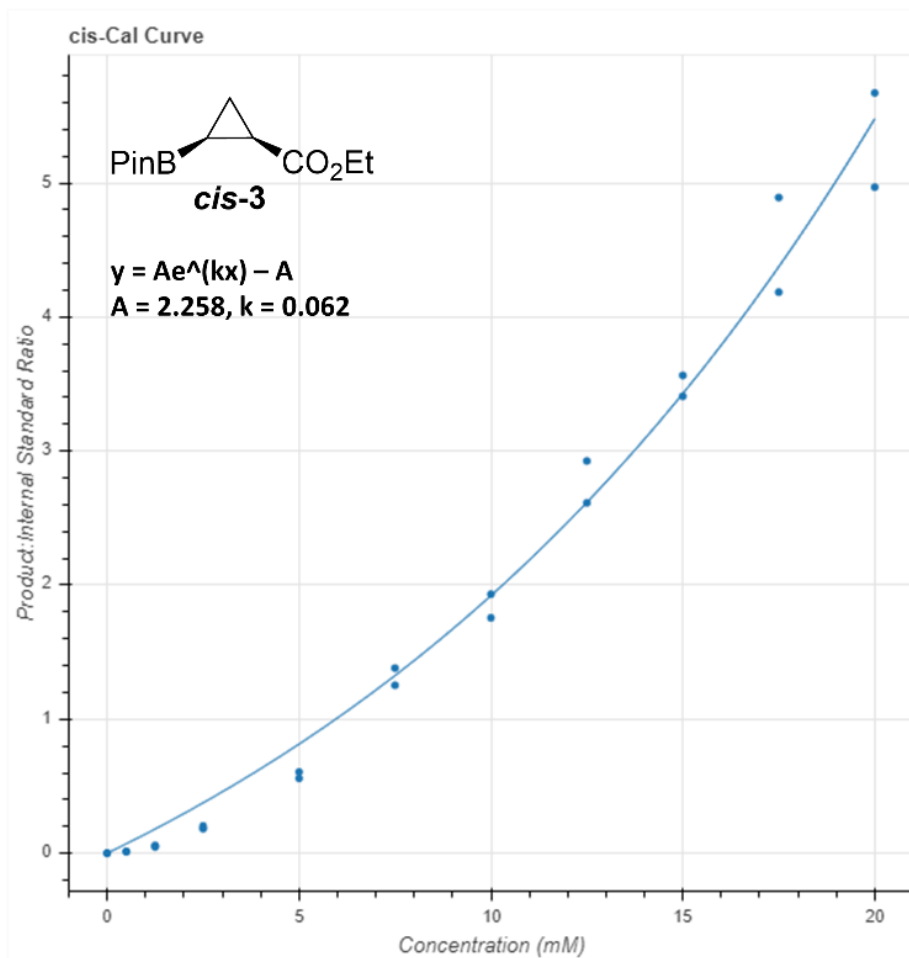


Figure S5. GC-MS calibration curve for *cis-3*.

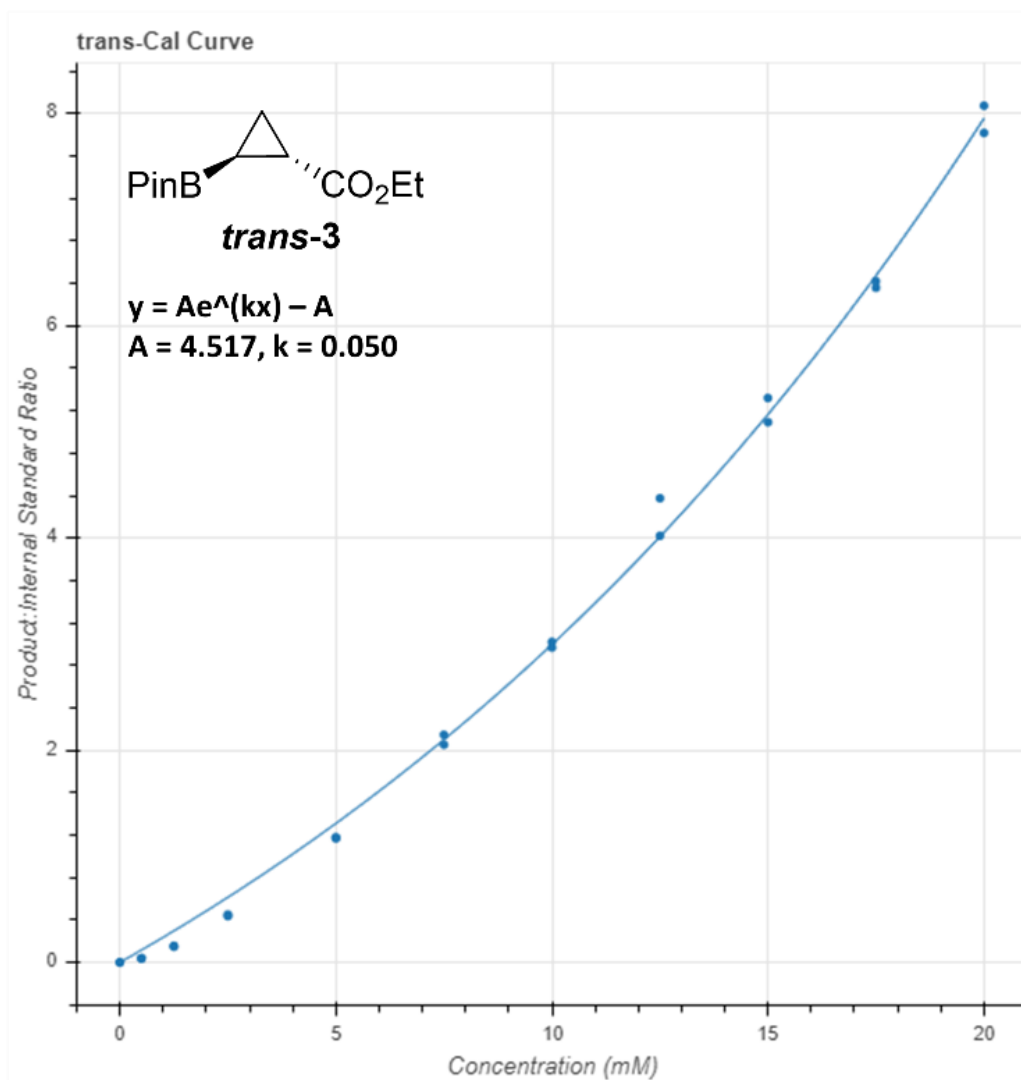


Figure S6. GC-MS calibration curve for *trans-3*.

### Representative chromatographic traces

Chromatographic separation conditions for each compound are reported in the *Compound separation conditions* section above. The chiral SFC method information is detailed in compound synthesis and characterization for the title compound.

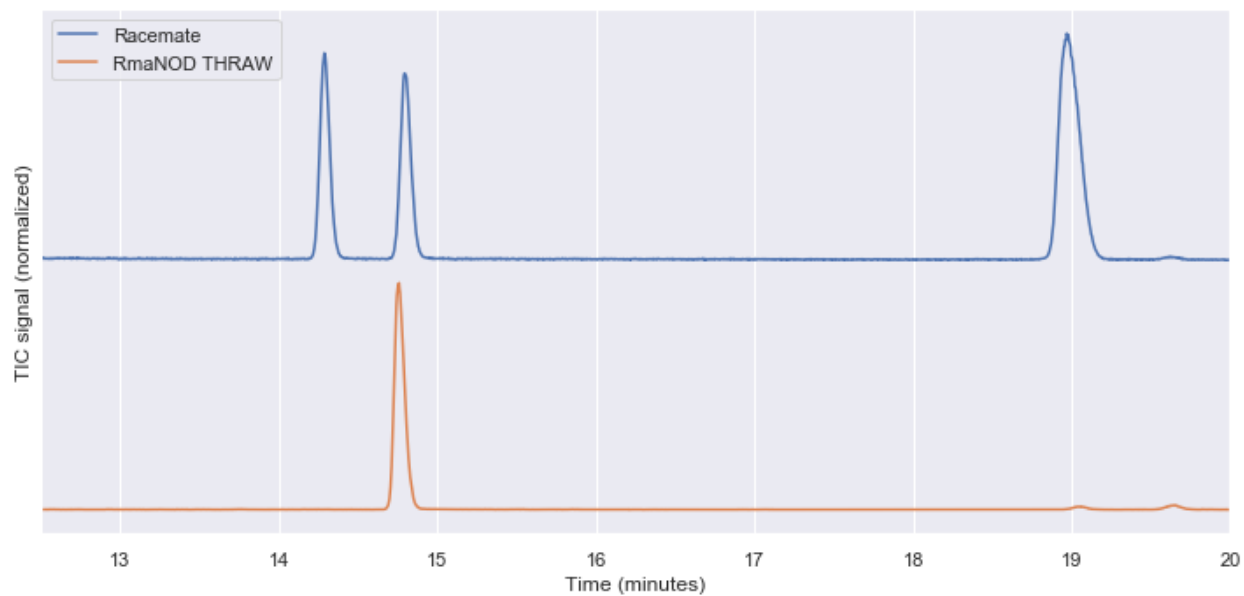


Figure S7. Chiral GC-MS trace of racemic authentic standard **3** on method to determine *cis*-**3** enantioselectivity and *Rma*NOD THRAW catalyzed production of **3**.

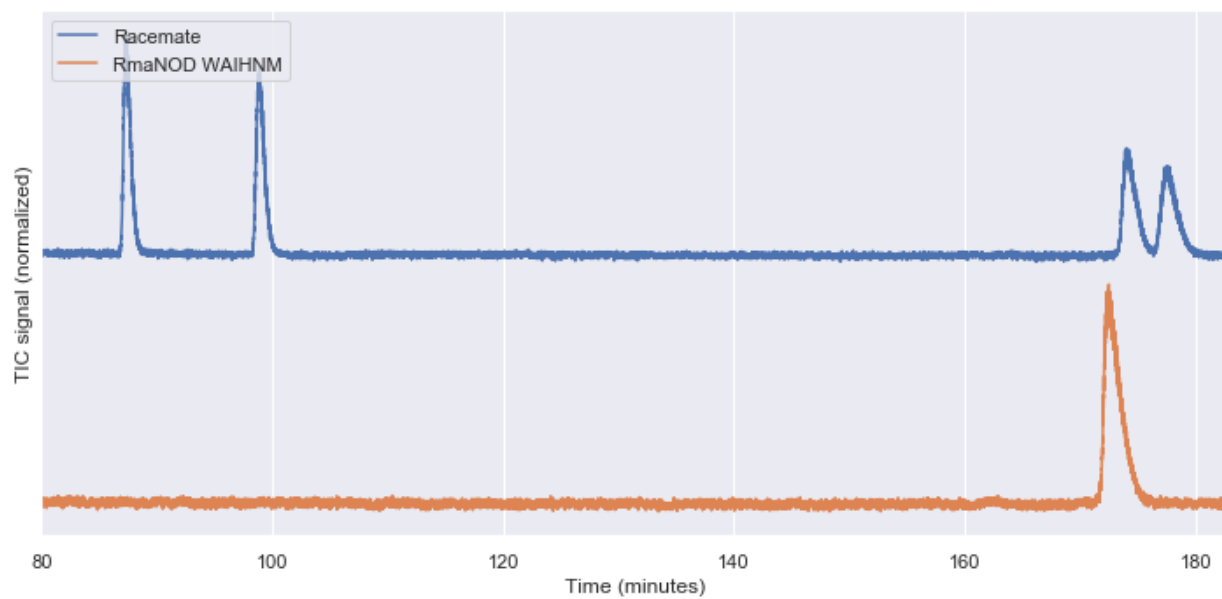
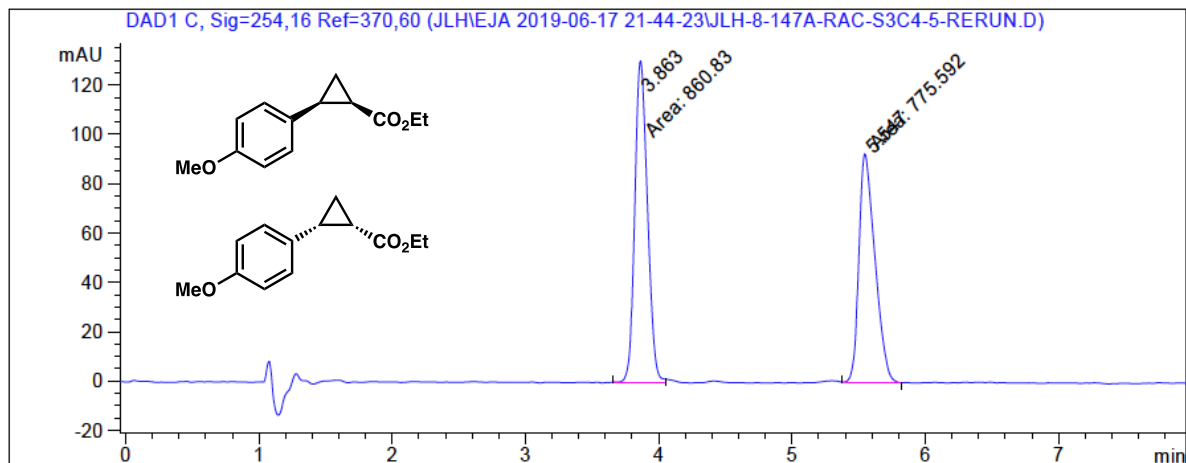
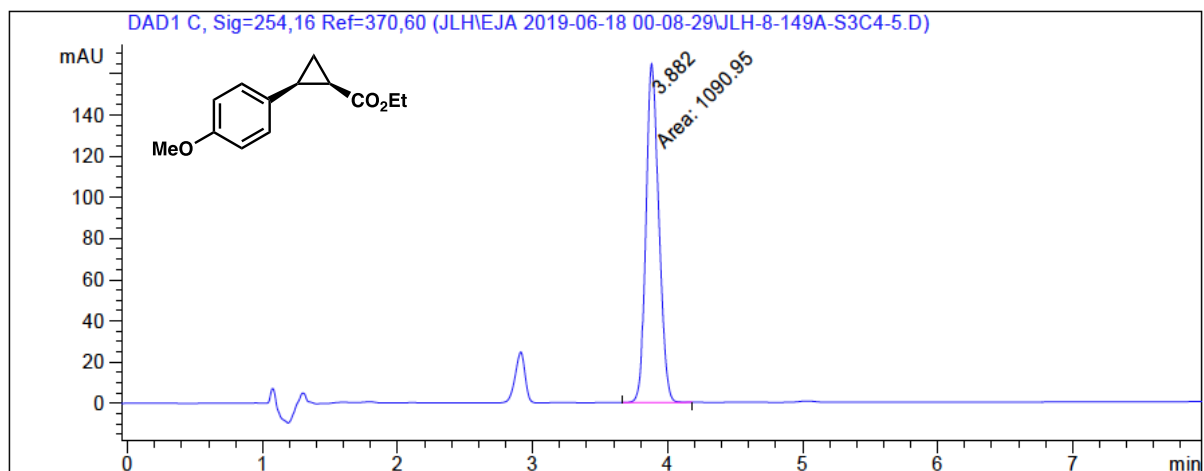


Figure S8. Chiral GC-MS trace of racemic authentic standard **3** on method to determine *trans*-**3** enantioselectivity and *Rma*NOD WAIHNM catalyzed production of **3**.



Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	3.863	MM	0.1094	860.82959	131.20370	52.6044
2	5.547	MM	0.1392	775.59241	92.85722	47.3956

Figure S9. Chiral SFC trace of *rac*-**5a**.



Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	3.882	MM	0.1099	1090.94519	165.46260	100.0000

Figure S10. Chiral SFC trace of chemoenzymatic product **5a**. The enantioenriched sample shows no minor stereoisomer peak.

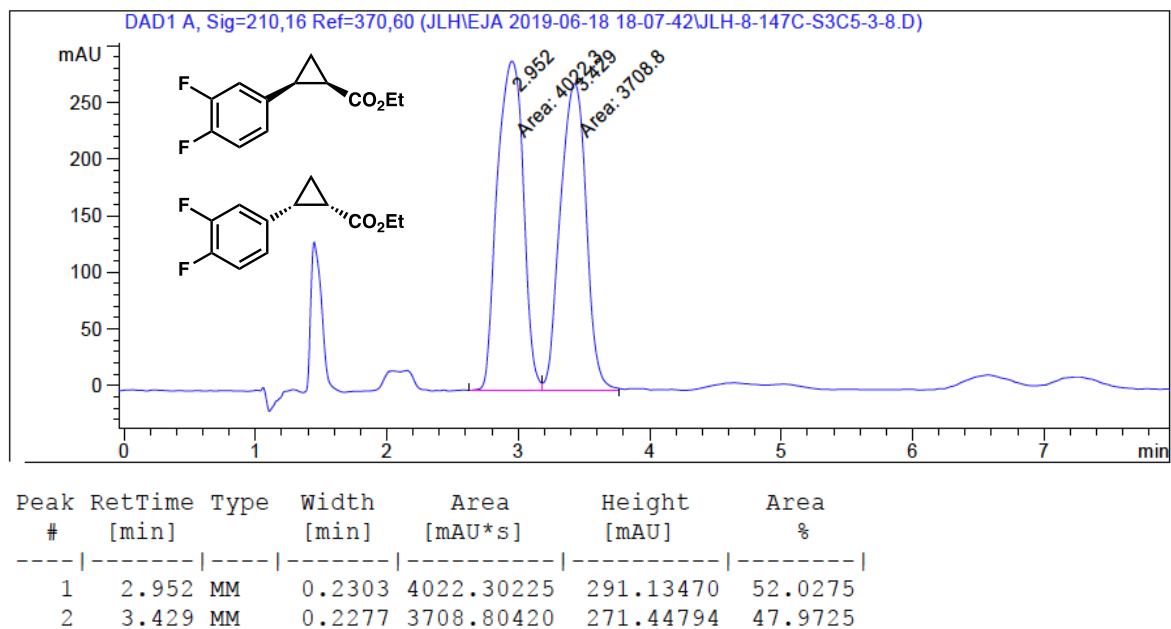


Figure S11. Chiral SFC trace of *rac*-**5b**.

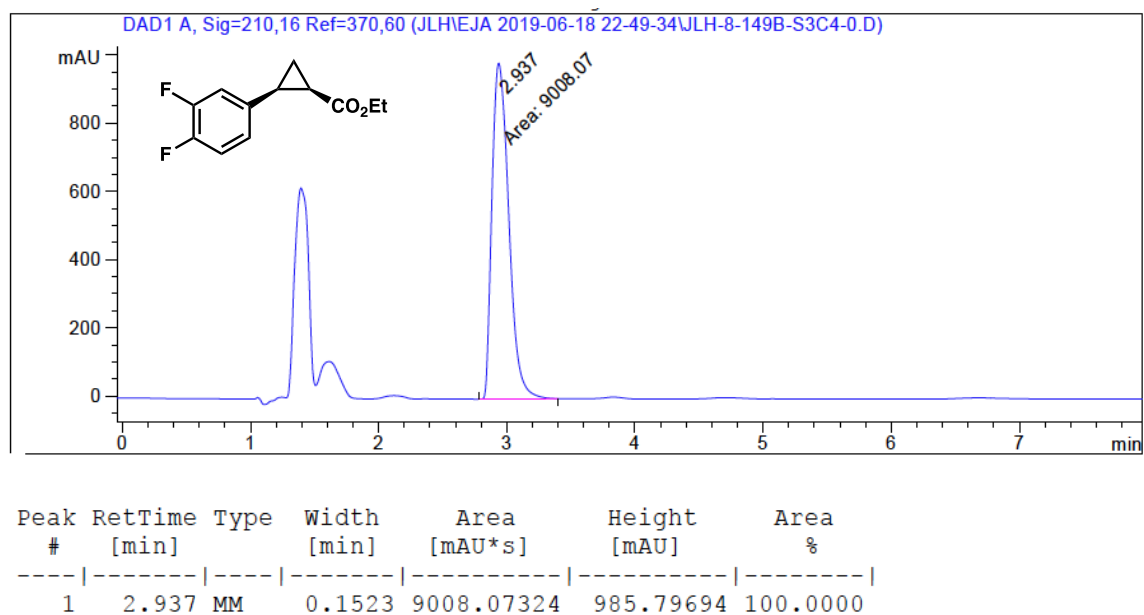


Figure S12. Chiral SFC trace of chemoenzymatic product **5b**. The enantioenriched sample shows no minor stereoisomer peak.

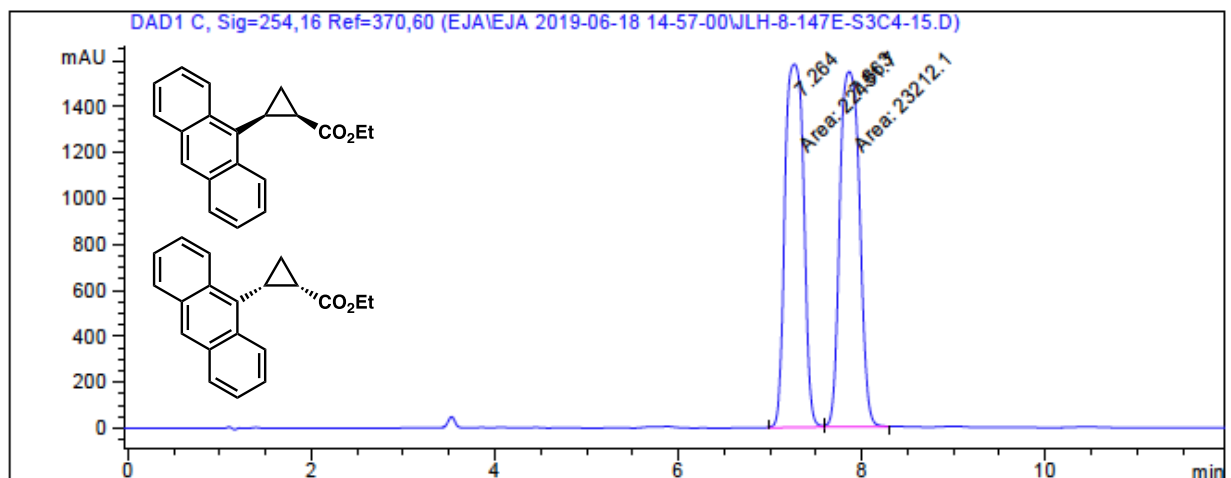


Figure S13. Chiral SFC trace of *rac*-**5c**.

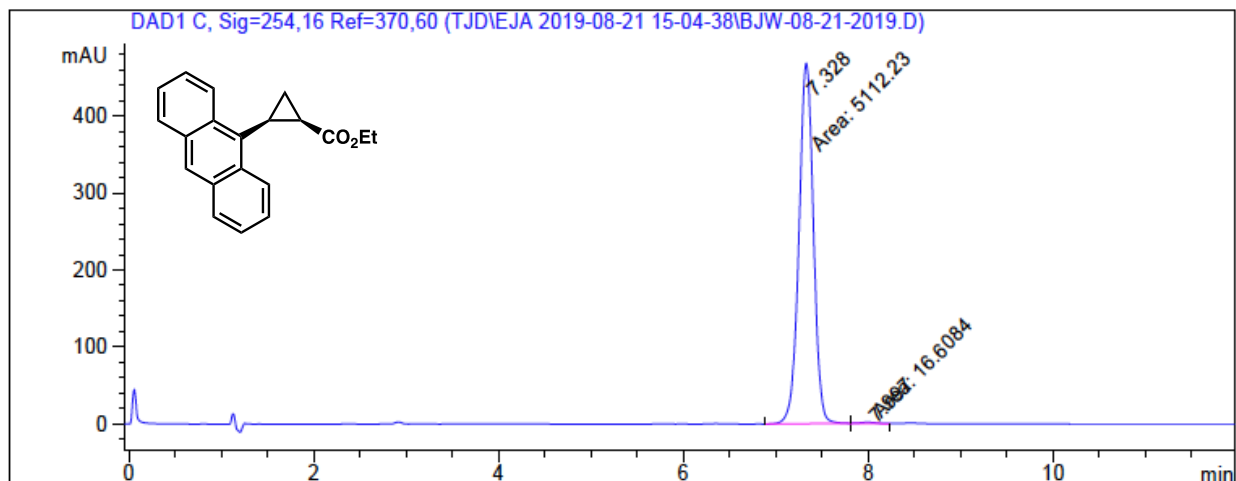
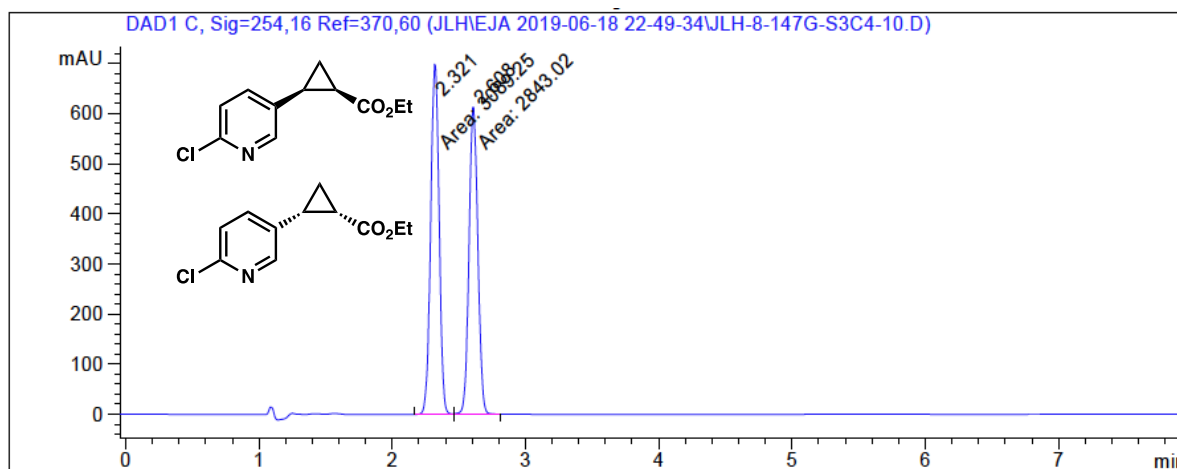
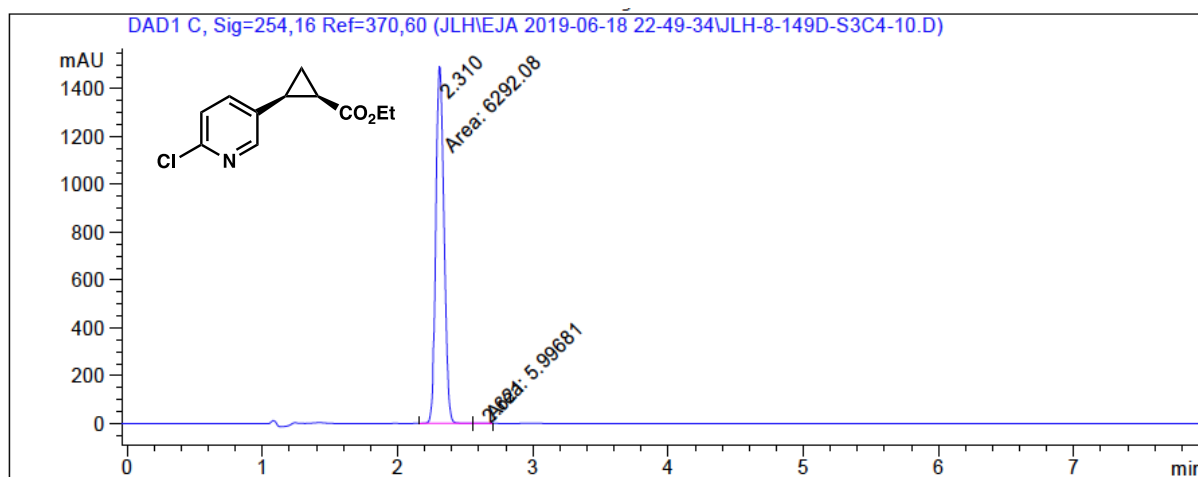


Figure S14. Chiral SFC trace of chemoenzymatic product **5c**. The enantioenriched sample shows only trace of the minor stereoisomer peak.



Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.321	MM	0.0732	3089.25415	703.42719	52.0754
2	2.608	MM	0.0766	2843.01636	618.20642	47.9246

Figure S15. Chiral SFC trace of *rac*-5d.



Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.310	MM	0.0697	6292.08105	1504.73425	99.9048
2	2.621	MM	0.0661	5.99681	1.51112	0.0952

Figure S16. Chiral SFC trace of chemoenzymatic product **5d**. The enantioenriched sample shows only trace of the minor stereoisomer peak.



## DNA and Amino-Acid Sequences

The DNA and amino-acid sequences for variants used in this work, as well as primer sequences for mutagenesis and cloning, are also provided as a separate supplemental Excel file.

### Nucleotide sequences of proteins of interest

Table S7. Nucleotide sequences of proteins of interest.

Protein name	Nucleotide sequence
<i>RmaNOD</i> wild type (WT)	ATGGCGCCGACCCTGTCGGAACAGACCCGTCAGTTGGTACGTGCGTCTG TGCCTGCACTGCAGAAACTCAGTCGCTATTAGCGCCACGATGTATCGG CTGCTTTTTCGAACGGTATCCCGAAACGCGGAGCTTGTTTGAACCTCCTGA GAGACAGATACACAAGCTTGCGTCGGCCCTGTTGGCCTACGCCCGTAGT ATCGACAACCCATCGGCGTTACAGGCGGCCATCCGCCGCATGGTGCTTT CCCACGCACGCGCAGGAGTGCAGGCCGTCCATTATCCGCTGGTTTGGGA ATGTTTGAGAGACGCTATAAAAGAAGTCCTGGGCCCGGATGCCACCGAG ACCCTTCTGCAGGCGTGGAAGGAAGCCTATGATTTTTTAGCTCATTACTG TCTACCAAGGAAGCGCAAGTCTACGCTGTGTTAGCTGAACTCGAGCATCA CCATCACCATCACTGA
<i>RmaNOD</i> Q52V	ATGGCGCCGACCCTGTCGGAACAGACCCGTCAGTTGGTACGTGCGTCTG TGCCTGCACTGCAGAAACTCAGTCGCTATTAGCGCCACGATGTATCGG CTGCTTTTTCGAACGGTATCCCGAAACGCGGAGCTTGTTTGAACCTCCTGA GAGAGTTATACACAAGCTTGCGTCGGCCCTGTTGGCCTACGCCCGTAGTA TCGACAACCCATCGGCGTTACAGGCGGCCATCCGCCGCATGGTGCTTTC CCACGCACGCGCAGGAGTGCAGGCCGTCCATTATCCGCTGGTTTGGGAA TGTTTGAGAGACGCTATAAAAGAAGTCCTGGGCCCGGATGCCACCGAGA CCCTTCTGCAGGCGTGGAAGGAAGCCTATGATTTTTTAGCTCATTACTGT CTACCAAGGAAGCGCAAGTCTACGCTGTGTTAGCTGAACTCGAGCACCAC CACCACCACCTGA
<i>RmaNOD</i> Q52V with cleavable N-terminal His – SUMO construct (underline denotes cleaved region)	<u>ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGC</u> <u>GCGGCAGCCATATGGCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGC</u> <u>TAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAA</u> <u>AGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACT</u> <u>CCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAAT</u> <u>GGACTCCTTAAGATTCTTGACGACGGTATTAGAATTCAAGCTGATCAGAC</u> <u>CCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCACAGAG</u> <u>AACAGATTGGTGGATCCATGGCGCCGACCCTGTCGGAACAGACCCGTCA</u> <u>GTTGGTACGTGCGTCTGTGCCTGCACTGCAGAAACTCAGTCGCTATTA</u> <u>GCGCCACGATGTATCGGCTGCTTTTTCGAACGGTATCCCGAAACGCGGAG</u> <u>CTTGTTTGAACCTCCTGAGAGAGTTATACACAAGCTTGCGTCGGCCCTGT</u> <u>TGGCCTACGCCCGTAGTATCGACAACCCATCGGCGTTACAGGCGGCCAT</u> <u>CCGCCGCATGGTGCTTTCCACGCACGCGCAGGAGTGCAGGCCGTCCAT</u> <u>TATCCGCTGGTTTGGGAATGTTTGAGAGACGCTATAAAAGAAGTCCTGGG</u> <u>CCCGGATGCCACCGAGACCCTTCTGCAGGCGTGGAAGGAAGCCTATGAT</u>

	TTTTTAGCTCATTACTGTCTACCAAGGAAGCGCAAGTCTACGCTGTGTTA GCTGAATAATGA
<i>RmaNOD</i> Q52A	ATGGCGCCGACCCTGTCGGAACAGACCCGTCAGTTGGTACGTGCGTCTG TGCCTGCACTGCAGAAACTCAGTCGCTATTAGCGCCACGATGTATCGG CTGCTTTTTCGAACGGTATCCCGAAACGCGGAGCTTGTTTGAACCTCCTGA GAGAGCGATACACAAGCTTGCGTCGGCCCTGTTGGCCTACGCCCGTAGT ATCGACAACCCATCGGCGTTACAGGCGGCCATCCGCCGCATGGTGCTTT CCCACGCACGCGCAGGAGTGCAGGCCGTCCATTATCCGCTGGTTTGGGA ATGTTTGAGAGACGCTATAAAGAAGTCCTGGGCCCGGATGCCACCGAG ACCCTTCTGCAGGCGTGGAAGGAAGCCTATGATTTTTTAGCTCATTACTG TCTACCAAGGAAGCGCAAGTCTACGCTGTGTTAGCTGAACTCGAGCACC CCACCACCACCTGA
<i>RmaNOD</i> Y32T Q52A	ATGGCGCCGACCCTGTCGGAACAGACCCGTCAGTTGGTACGTGCGTCTG TGCCTGCACTGCAGAAACTCAGTCGCTATTAGCGCCACGATGACGCG GCTGCTTTTTCGAACGGTATCCCGAAACGCGGAGCTTGTTTGAACCTCCTG AGAGAGCGATACACAAGCTTGCGTCGGCCCTGTTGGCCTACGCCCGTAG TATCGACAACCCATCGGCGTTACAGGCGGCCATCCGCCGCATGGTGCTT TCCCACGCACGCGCAGGAGTGCAGGCCGTCCATTATCCGCTGGTTTGGG AATGTTTGAGAGACGCTATAAAGAAGTCCTGGGCCCGGATGCCACCGA GACCCTTCTGCAGGCGTGGAAGGAAGCCTATGATTTTTTAGCTCATTACT GTCTACCAAGGAAGCGCAAGTCTACGCTGTGTTAGCTGAACTCGAGCACC ACCACCACCACCTGA
<i>RmaNOD</i> Y32T Y39H L48R Q52A R79W ("THRAW")	ATGGCGCCGACCCTGTCGGAACAGACCCGTCAGTTGGTACGTGCGTCTG TGCCTGCACTGCAGAAACTCAGTCGCTATTAGCGCCACGATGACGCG GCTGCTTTTTCGAACGGCATCCCGAAACGCGGAGCTTGTTTGAACGTCCTG AGAGAGCGATACACAAGCTTGCGTCGGCCCTGTTGGCCTACGCCCGTAG TATCGACAACCCATCGGCGTTACAGGCGGCCATCCGCTGGATGGTGCTT TCCCACGCACGCGCAGGAGTGCAGGCCGTCCATTATCCGCTGGTTTGGG AATGTTTGAGAGACGCTATAAAGAAGTCCTGGGCCCGGATGCCACCGA GACCCTTCTGCAGGCGTGGAAGGAAGCCTATGATTTTTTAGCTCATTACT GTCTACCAAGGAAGCGCAAGTCTACGCTGTGTTAGCTGAACTCGAGCACC ACCACCACCACCTGA
<i>RmaNOD</i> Q52A L101N	ATGGCGCCGACCCTGTCGGAACAGACCCGTCAGTTGGTACGTGCGTCTG TGCCTGCACTGCAGAAACTCAGTCGCTATTAGCGCCACGATGTATCGG CTGCTTTTTCGAACGGTATCCCGAAACGCGGAGCTTGTTTGAACCTCCTGA GAGAGCGATACACAAGCTTGCGTCGGCCCTGTTGGCCTACGCCCGTAGT ATCGACAACCCATCGGCGTTACAGGCGGCCATCCGCCGCATGGTGCTTT CCCACGCACGCGCAGGAGTGCAGGCCGTCCATTATCCGCTGGTTTGGGA ATGTAATAGAGACGCTATAAAGAAGTCCTGGGCCCGGATGCCACCGAG ACCCTTCTGCAGGCGTGGAAGGAAGCCTATGATTTTTTAGCTCATTACTG TCTACCAAGGAAGCGCAAGTCTACGCTGTGTTAGCTGAACTCGAGCACC CCACCACCACCTGA
<i>RmaNOD</i> Q52A 60H L101N	ATGGCGCCGACCCTGTCGGAACAGACCCGTCAGTTGGTACGTGCGTCTG TGCCTGCACTGCAGAAACTCAGTCGCTATTAGCGCCACGATGTATCGG CTGCTTTTTCGAACGGTATCCCGAAACGCGGAGCTTGTTTGAACCTCCTGA GAGAGCGATACACAAGCTTGCGTCGGCCATTTGGCCTACGCCCGTAGT ATCGACAACCCATCGGCGTTACAGGCGGCCATCCGCCGCATGGTGCTTT CCCACGCACGCGCAGGAGTGCAGGCCGTCCATTATCCGCTGGTTTGGGA ATGTAATAGAGACGCTATAAAGAAGTCCTGGGCCCGGATGCCACCGAG ACCCTTCTGCAGGCGTGGAAGGAAGCCTATGATTTTTTAGCTCATTACTG

	TCTACCAAGGAAGCGCAAGTCTACGCTGTGTTAGCTGAACTCGAGCACCA CCACCACCACCACTGA
<i>RmaNOD</i> Q52A 60H L101N I105M	ATGGCGCCGACCCTGTCGGAACAGACCCGTCAGTTGGTACGTGCGTCTG TGCCTGCACTGCAGAAACTCAGTCGCTATTAGCGCCACGATGTATCGG CTGCTTTTTCGAACGGTATCCCGAAACGCGGAGCTTGTGTTGAACTTCCTGA GAGAGCGATACACAAGCTTGCCTCGGCCATTTGGCCTACGCCCGTAGT ATCGACAACCCATCGGCGTTACAGGCGGCCATCCGCCGCATGGTGCTTT CCCACGCACGCGCAGGAGTGCAGGCCGTCCATTATCCGCTGGTTTGGGA ATGTAATAGAGACGCTATGAAAGAAGTCCTGGGCCCGGATGCCACCGAG ACCCTTCTGCAGGCGTGGAAGGAAGCCTATGATTTTTTAGCTCATTACTG TCTACCAAGGAAGCGCAAGTCTACGCTGTGTTAGCTGAACTCGAGCACCA CCACCACCACCACTGA
<i>RmaNOD</i> L20W Q52A L56I 60H L101N I105M (WAIHNM)	ATGGCGCCGACCCTGTCGGAACAGACCCGTCAGTTGGTACGTGCGTCTG TGCCTGCATGGCAGAAACTCAGTCGCTATTAGCGCCACGATGTATCGG CTGCTTTTTCGAACGGTATCCCGAAACGCGGAGCTTGTGTTGAACTTCCTGA GAGAGCGATACACAAGATTGCCTCGGCCATTTGGCCTACGCCCGTAGT ATCGACAACCCATCGGCGTTACAGGCGGCCATCCGCCGCATGGTGCTTT CCCACGCACGCGCAGGAGTGCAGGCCGTCCATTATCCGCTGGTTTGGGA ATGTAATAGAGACGCTATGAAAGAAGTCCTGGGCCCGGATGCCACCGAG ACCCTTCTGCAGGCGTGGAAGGAAGCCTATGATTTTTTAGCTCATTACTG TCTACCAAGGAAGCGCAAGTCTACGCTGTGTTAGCTGAACTCGAGCACCA CCACCACCACCACTGA

### Amino-acid sequences of proteins of interest.

Table S8. Amino-acid sequences of proteins of interest.

Protein name	Amino-acid sequence
<i>RmaNOD</i> wild type (WT)	MAPTLSEQTRQLVRASVPALQKHSVAISATMYRLLFERYPETRSL FELPERQIHKLASALLAYARSIDNPSALQAAIRRMVLSHARAGVQA VHYPLVWECLRDAIKEVLGPDATETLLQAWKEAYDFLAHLLSTKE AQVYAVLAELEHHHHHH*
<i>RmaNOD</i> Q52V	MAPTLSEQTRQLVRASVPALQKHSVAISATMYRLLFERYPETRSL FELPERVIHKLASALLAYARSIDNPSALQAAIRRMVLSHARAGVQA VHYPLVWECLRDAIKEVLGPDATETLLQAWKEAYDFLAHLLSTKE AQVYAVLAELEHHHHHH*
<i>RmaNOD</i> Q52V with cleavable N-terminal His – SUMO construct (underline denotes cleaved region)	<u>MGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKP</u> <u>ETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFKRQGKEMDSLRF</u> <u>YDGIRIQADQTPEDLDMEDNDIIEAHREQIGGSMAPTLSEQTRQLV</u> RASVPALQKHSVAISATMYRLLFERYPETRSLFELPERVIHKLASA LLAYARSIDNPSALQAAIRRMVLSHARAGVQAVHYPLVWECLRDA IKEVLGPDATETLLQAWKEAYDFLAHLLSTKEAQVYAVLAE**
<i>RmaNOD</i> Q52A	MAPTLSEQTRQLVRASVPALQKHSVAISATMYRLLFERYPETRSL FELPERAIHKLASALLAYARSIDNPSALQAAIRRMVLSHARAGVQA VHYPLVWECLRDAIKEVLGPDATETLLQAWKEAYDFLAHLLSTKE AQVYAVLAELEHHHHHH*

<i>Rma</i> NOD Y32T Q52A	MAPTLSEQTRQLVRASVPALQKHSVAISATMTRLLFERYPETRSL FELPERAIHKLASALLAYARSIDNPSALQAAIRRMVLSHARAGVQA VHYPLVWECLRDAIKEVLGPDATETLLQAWKEAYDFLAHLLSTKE AQVYAVLAELEHHHHHH*
<i>Rma</i> NOD Y32T Y39H L48R Q52A R79W ("THRAW")	MAPTLSEQTRQLVRASVPALQKHSVAISATMTRLLFERHPETRSL FERPERAIHKLASALLAYARSIDNPSALQAAIRWMVLSHARAGVQ AVHYPLVWECLRDAIKEVLGPDATETLLQAWKEAYDFLAHLLSTK EAQVYAVLAELEHHHHHH*
<i>Rma</i> NOD Q52A L101N	MAPTLSEQTRQLVRASVPALQKHSVAISATMYRLLFERYPETRSL FELPERAIHKLASALLAYARSIDNPSALQAAIRRMVLSHARAGVQA VHYPLVWECNRDAIKEVLGPDATETLLQAWKEAYDFLAHLLSTKE AQVYAVLAELEHHHHHH*
<i>Rma</i> NOD Q52A 60H L101N	MAPTLSEQTRQLVRASVPALQKHSVAISATMYRLLFERYPETRSL FELPERAIHKLASAHLAYARSIDNPSALQAAIRRMVLSHARAGVQA VHYPLVWECNRDAIKEVLGPDATETLLQAWKEAYDFLAHLLSTKE AQVYAVLAELEHHHHHH*
<i>Rma</i> NOD Q52A 60H L101N I105M	MAPTLSEQTRQLVRASVPALQKHSVAISATMYRLLFERYPETRSL FELPERAIHKLASAHLAYARSIDNPSALQAAIRRMVLSHARAGVQA VHYPLVWECNRDAMKEVLGPDATETLLQAWKEAYDFLAHLLSTK EAQVYAVLAELEHHHHHH*
<i>Rma</i> NOD L20W Q52A L56I 60H L101N I105M	MAPTLSEQTRQLVRASVPAWQKHSVAISATMYRLLFERYPETRSL FELPERAIHKIASAHLAYARSIDNPSALQAAIRRMVLSHARAGVQA VHYPLVWECNRDAMKEVLGPDATETLLQAWKEAYDFLAHLLSTK EAQVYAVLAELEHHHHHH*

# Spectroscopic Data

20190718\_cis.1.fid

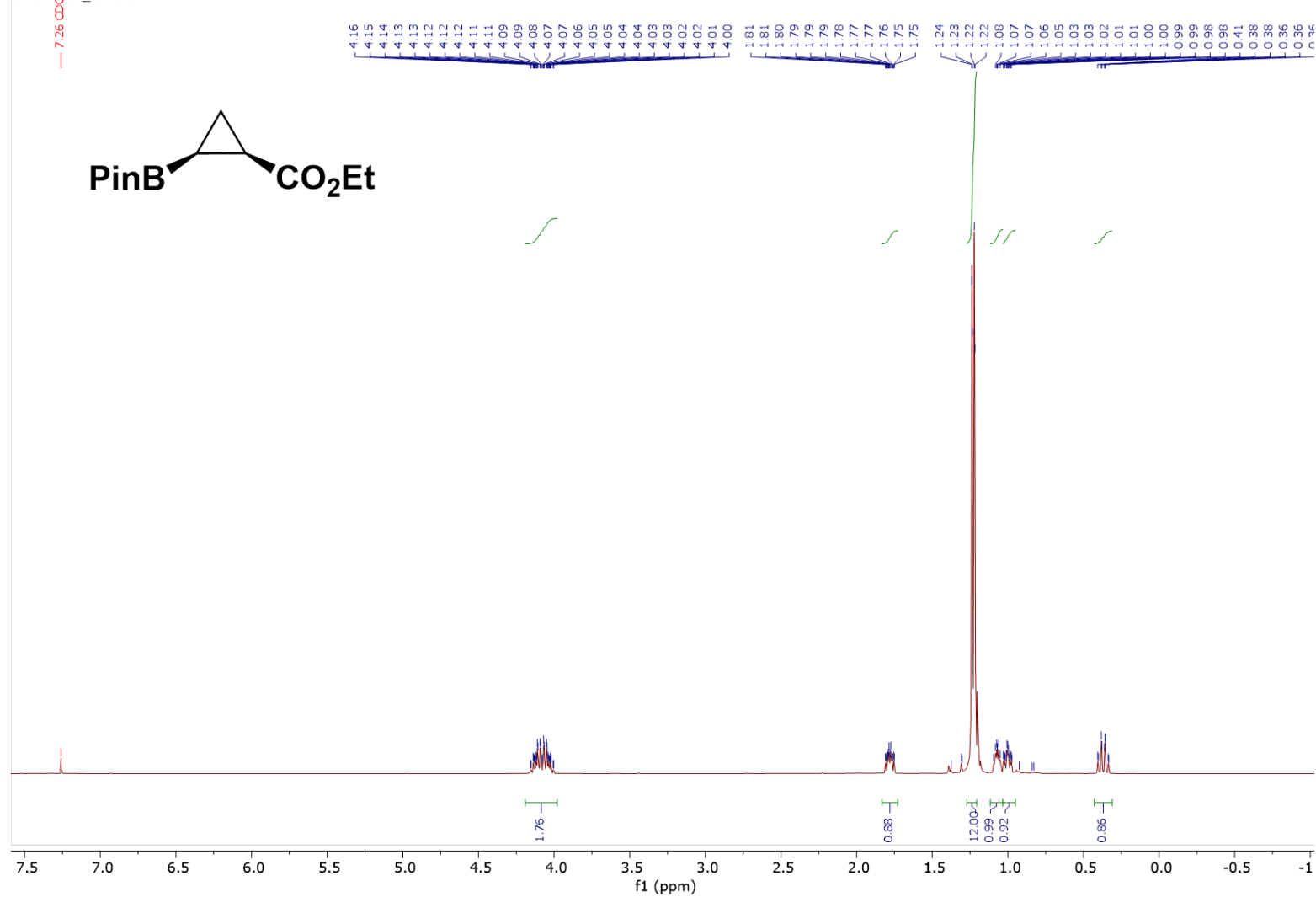


Figure S17. <sup>1</sup>H NMR of *cis*-3.

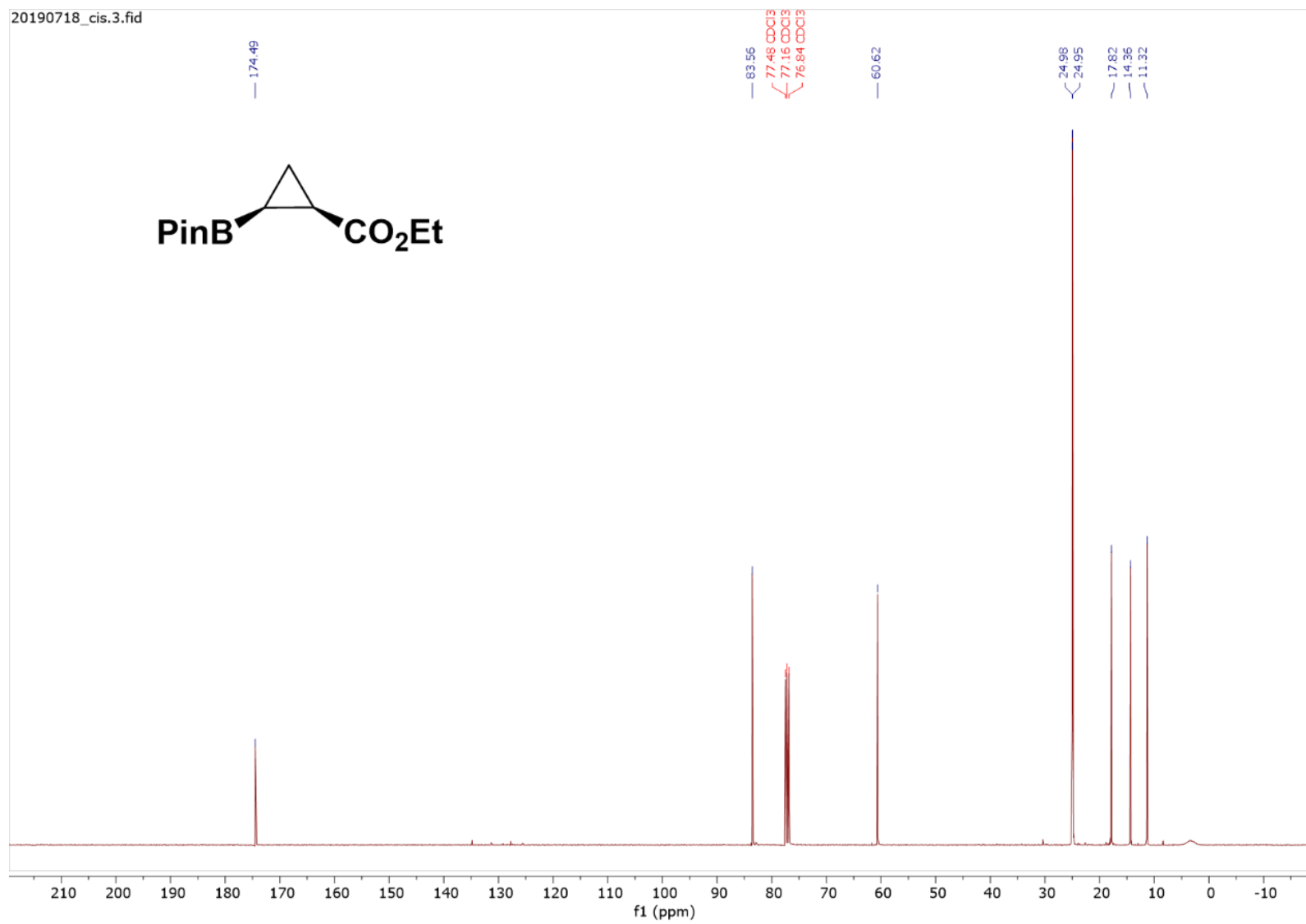


Figure S18.  $^{13}\text{C}$  NMR of cis-3.

20190718\_cis.2.fid

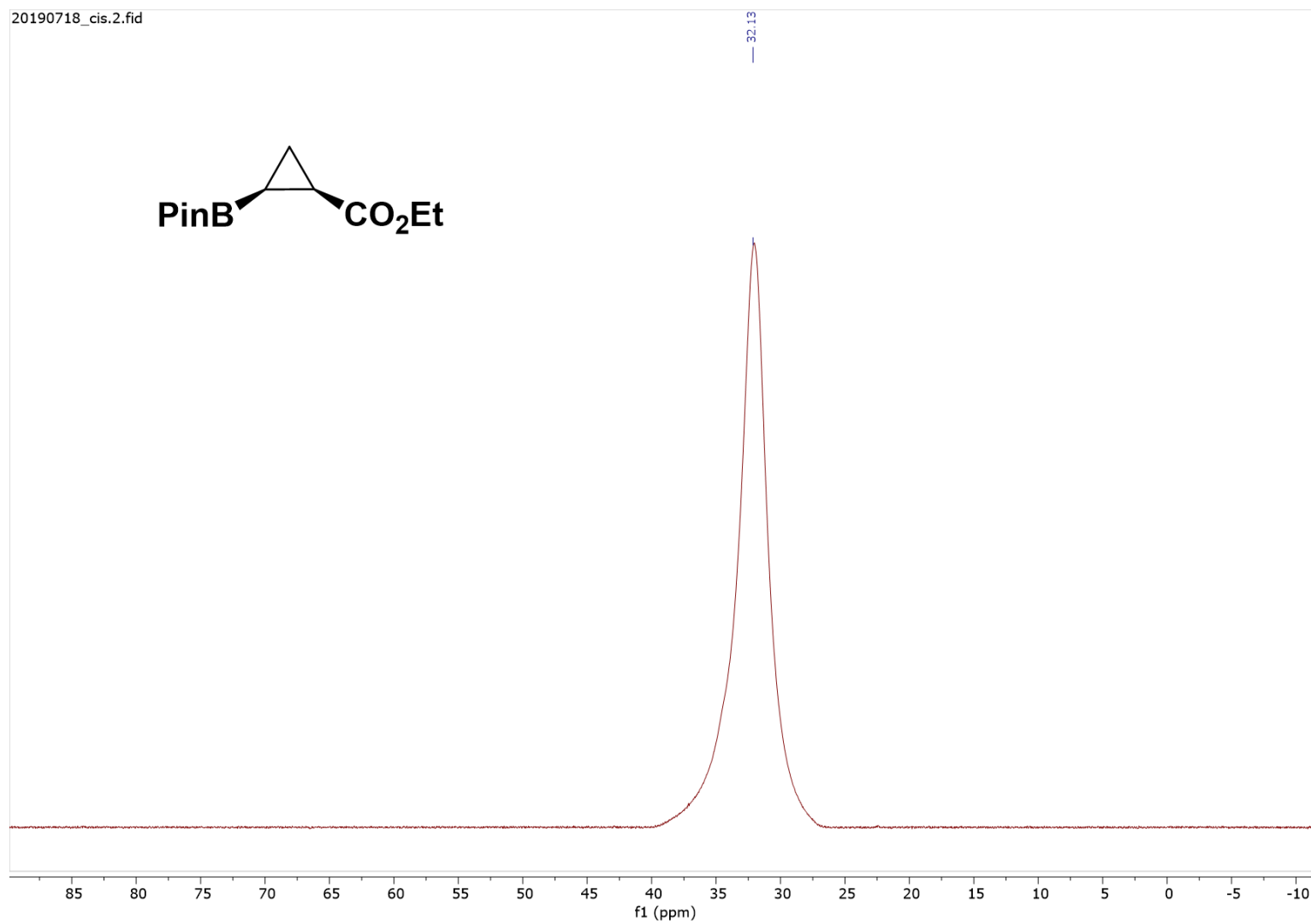
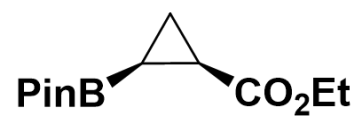


Figure S19.  $^{11}\text{B}$  NMR of cis-3.

20190730\_Trans.1.fid

7.26 CDCl<sub>3</sub>

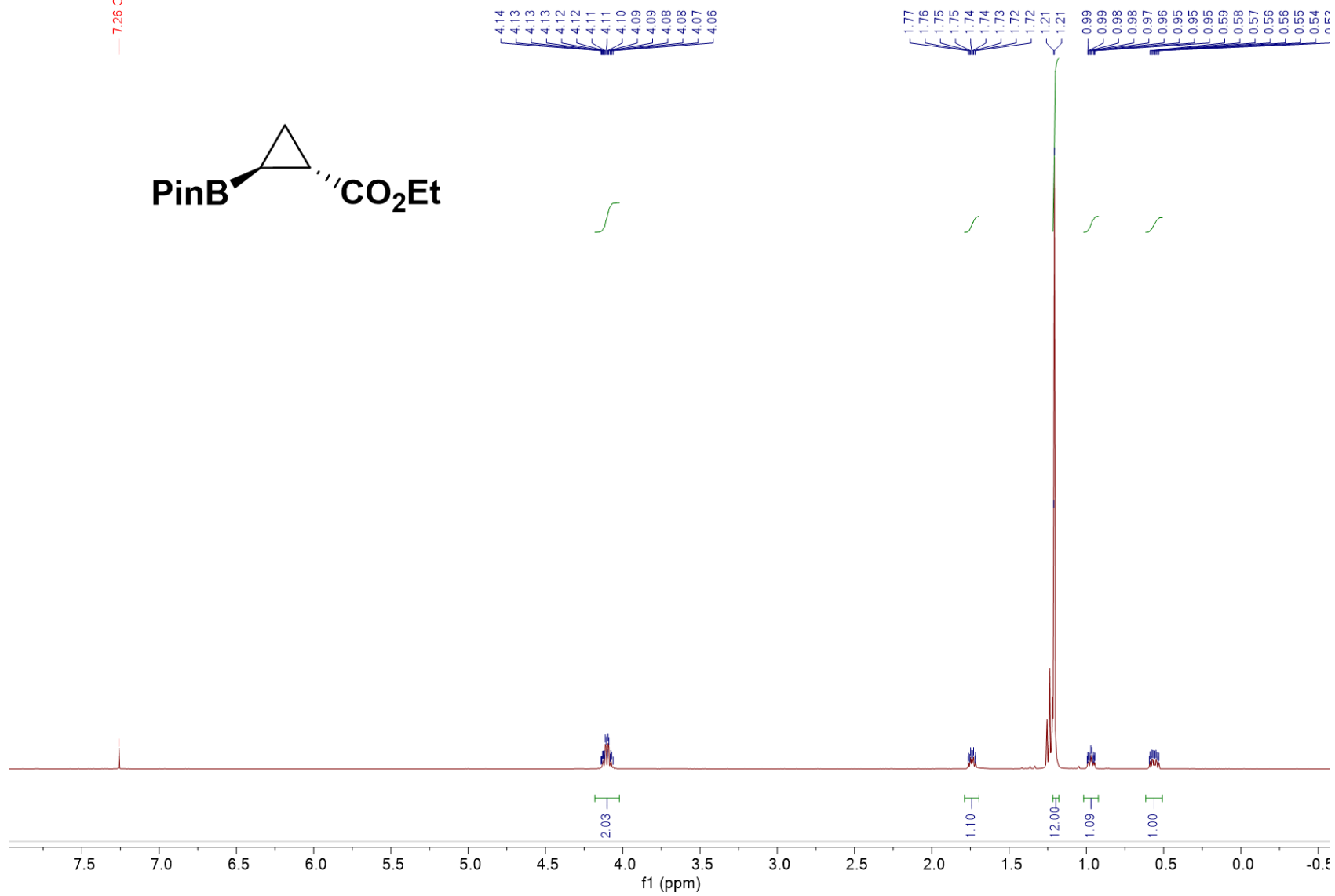


Figure S 20.. <sup>1</sup>H NMR of *trans*-3.



20190718\_trans.3.fid

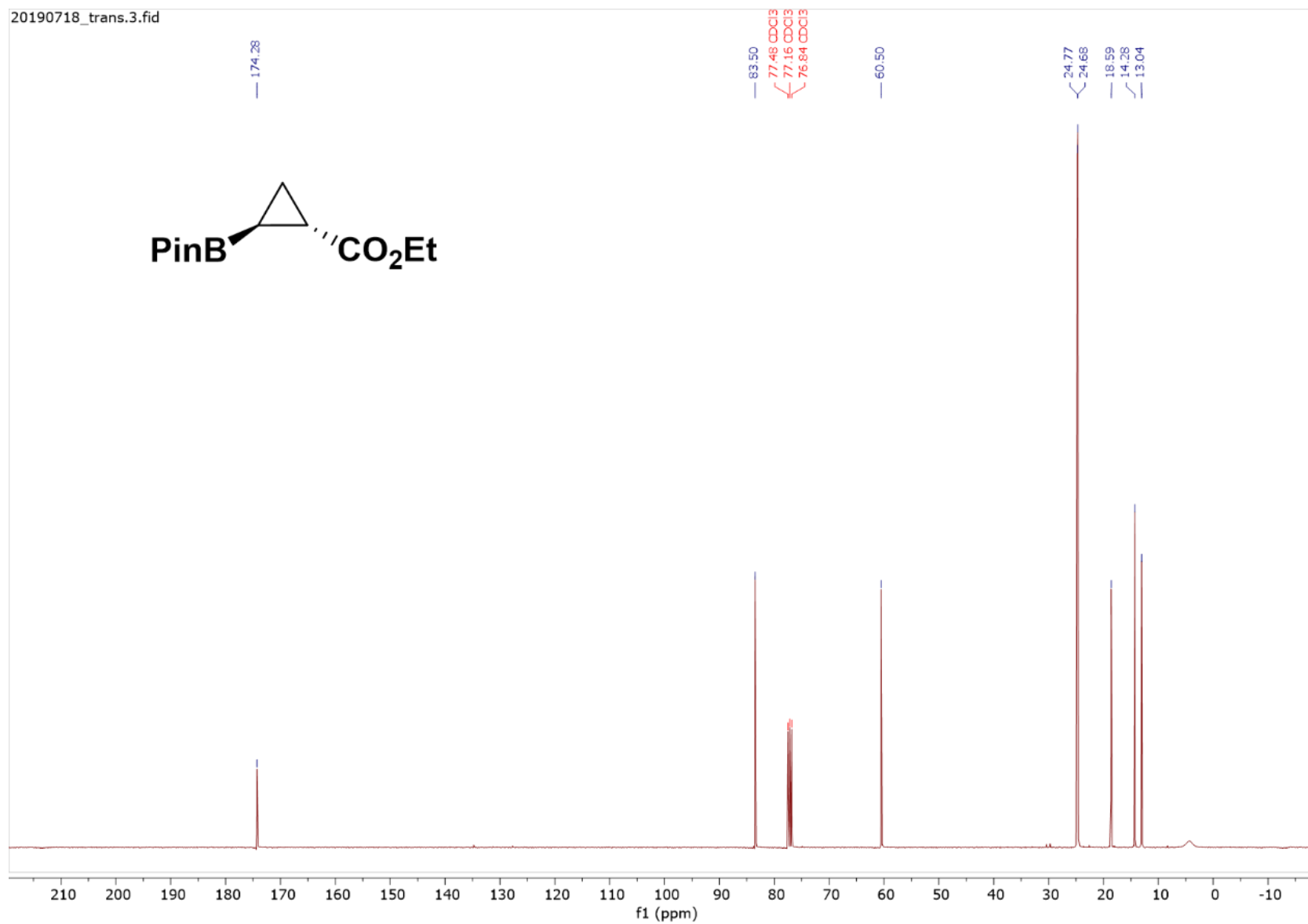


Figure S21. <sup>13</sup>C NMR of *trans*-3.

20190718\_trans.2.fid

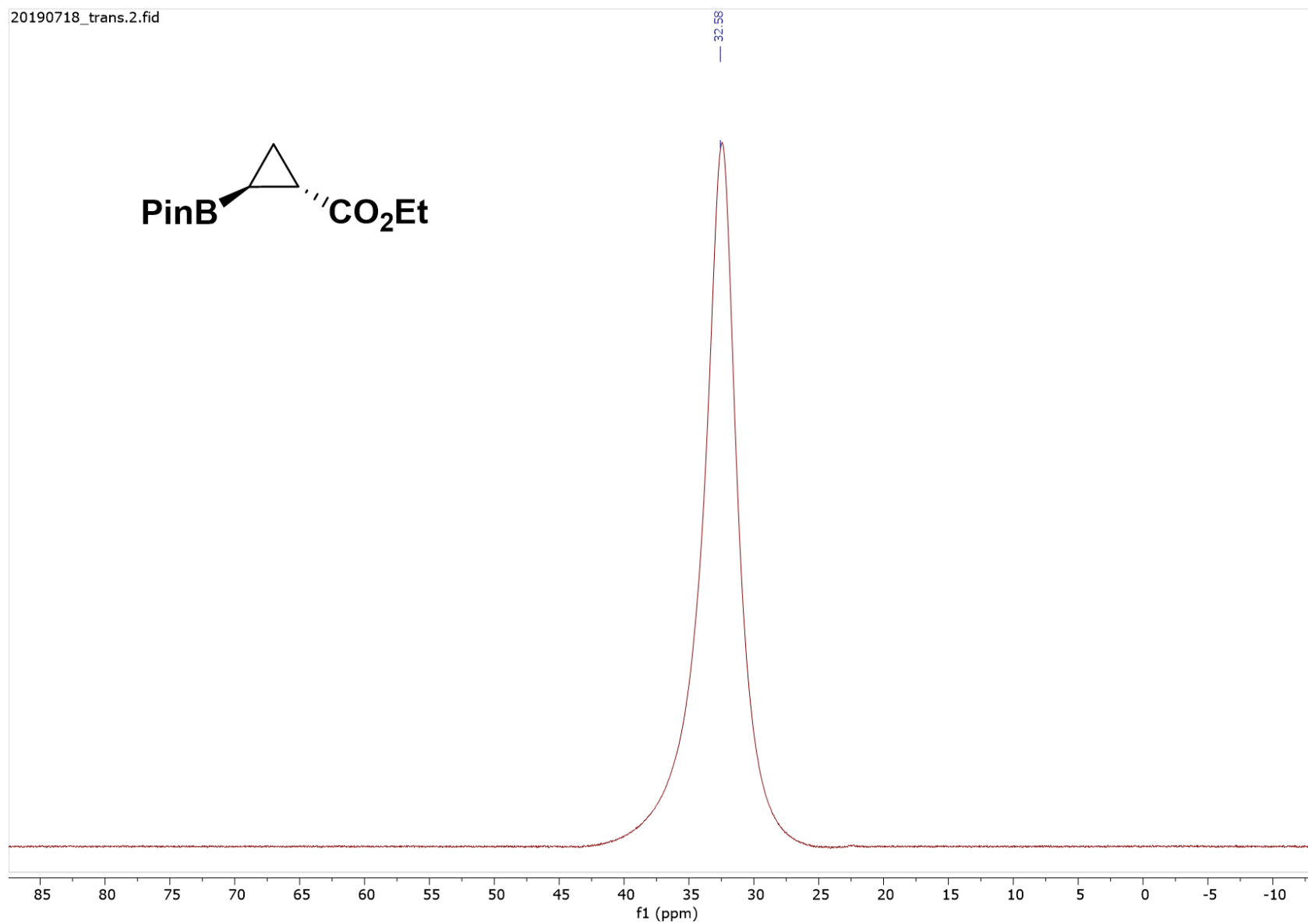


Figure S22.  $^{11}\text{B}$  NMR of *trans*-3.

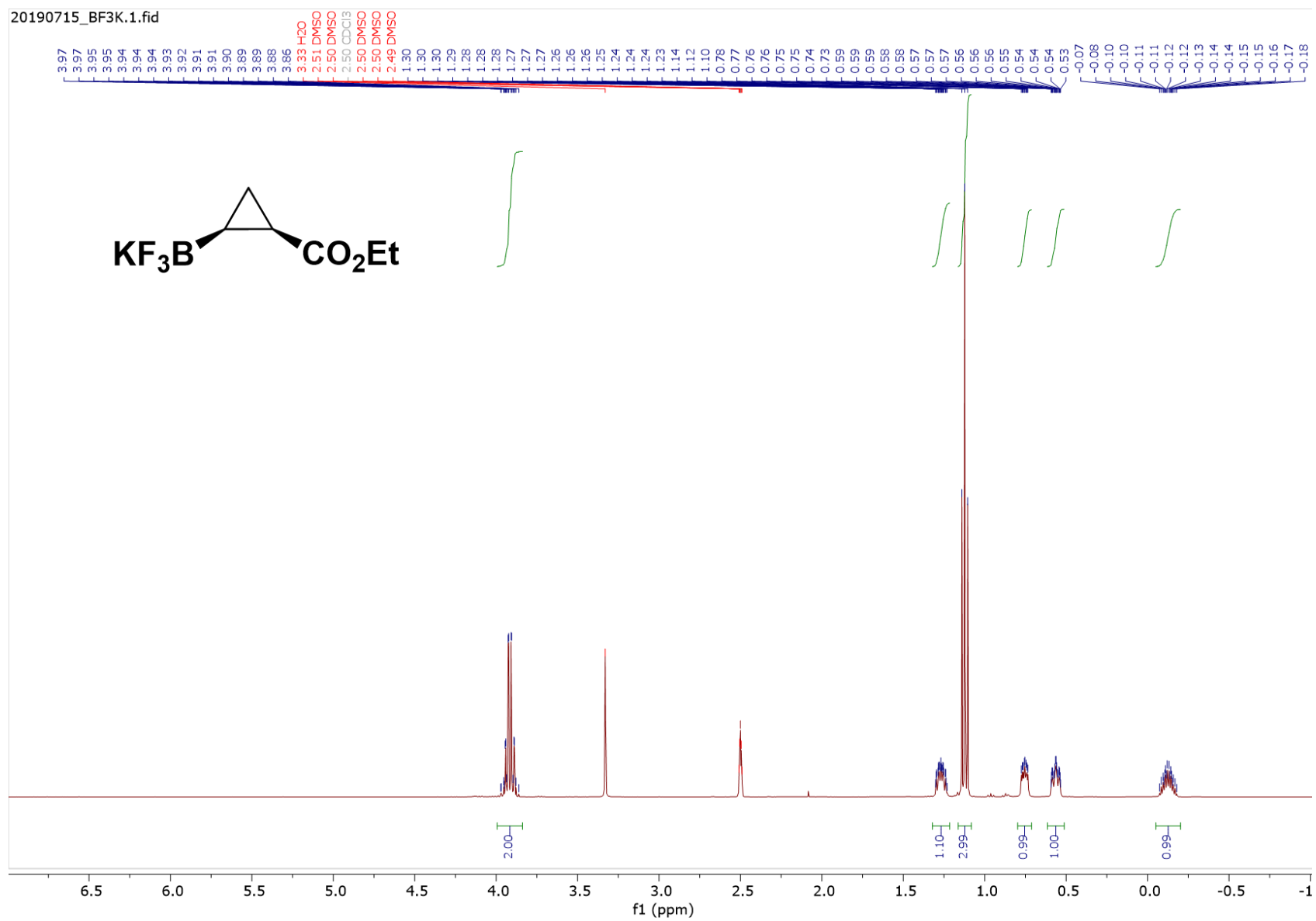


Figure S23. <sup>1</sup>H NMR of *cis*-4.

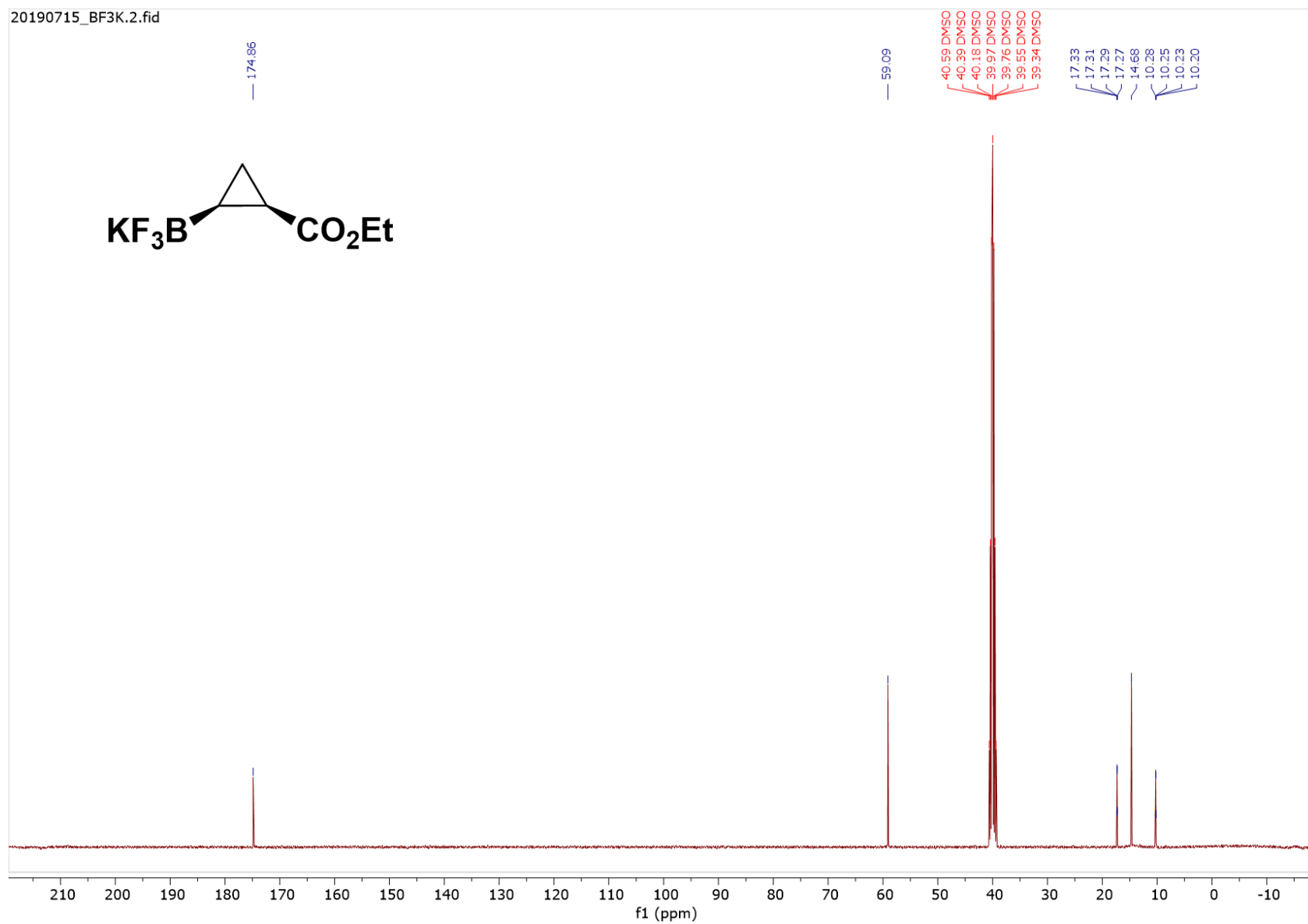


Figure S24.  $^{13}\text{C}$  NMR of *cis*-4.

20190715\_BF3K.3.fid

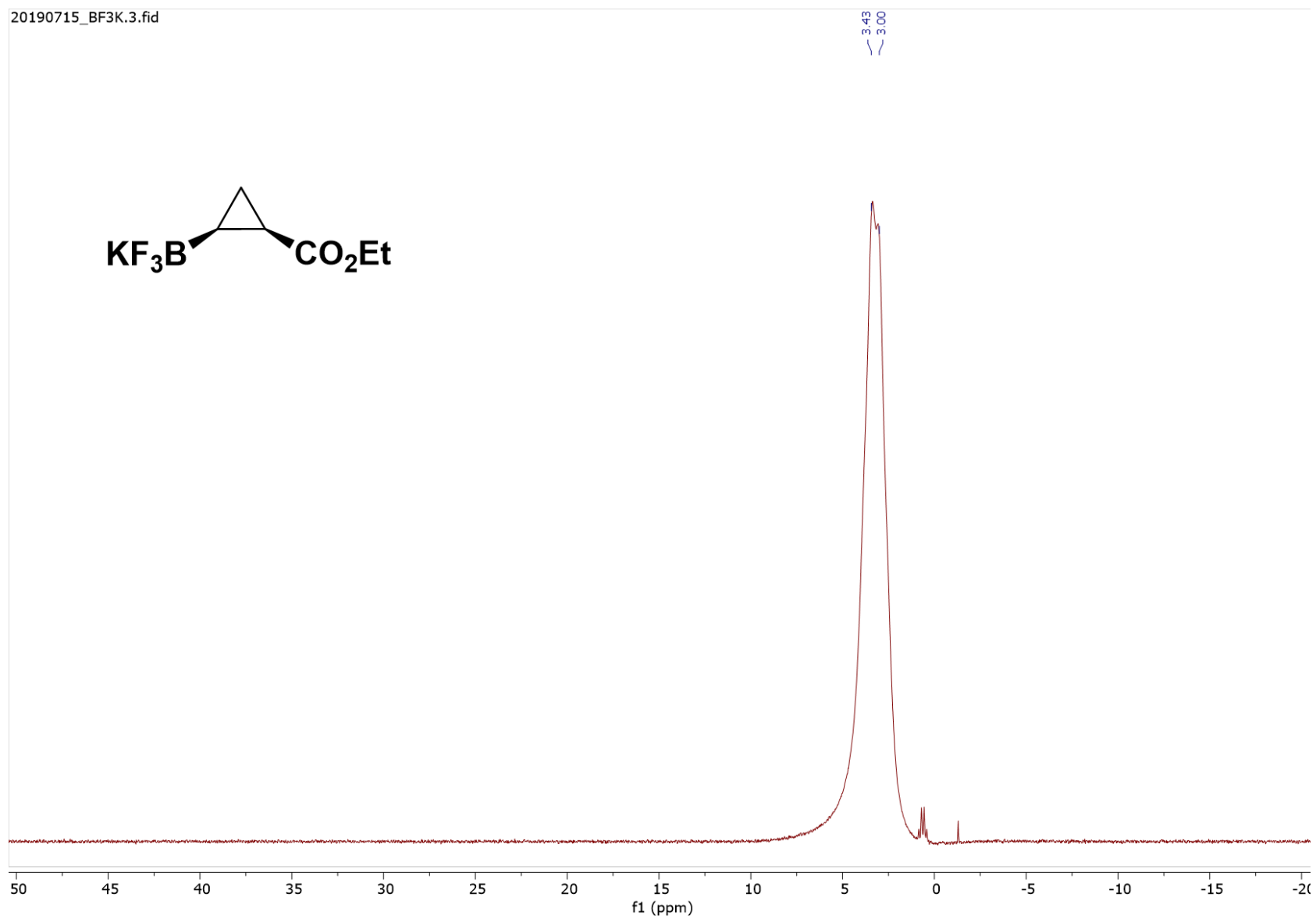


Figure S25.  $^{11}\text{B}$  NMR of *cis*-4.

FLUORINE01  
190715\_BJW\_salt

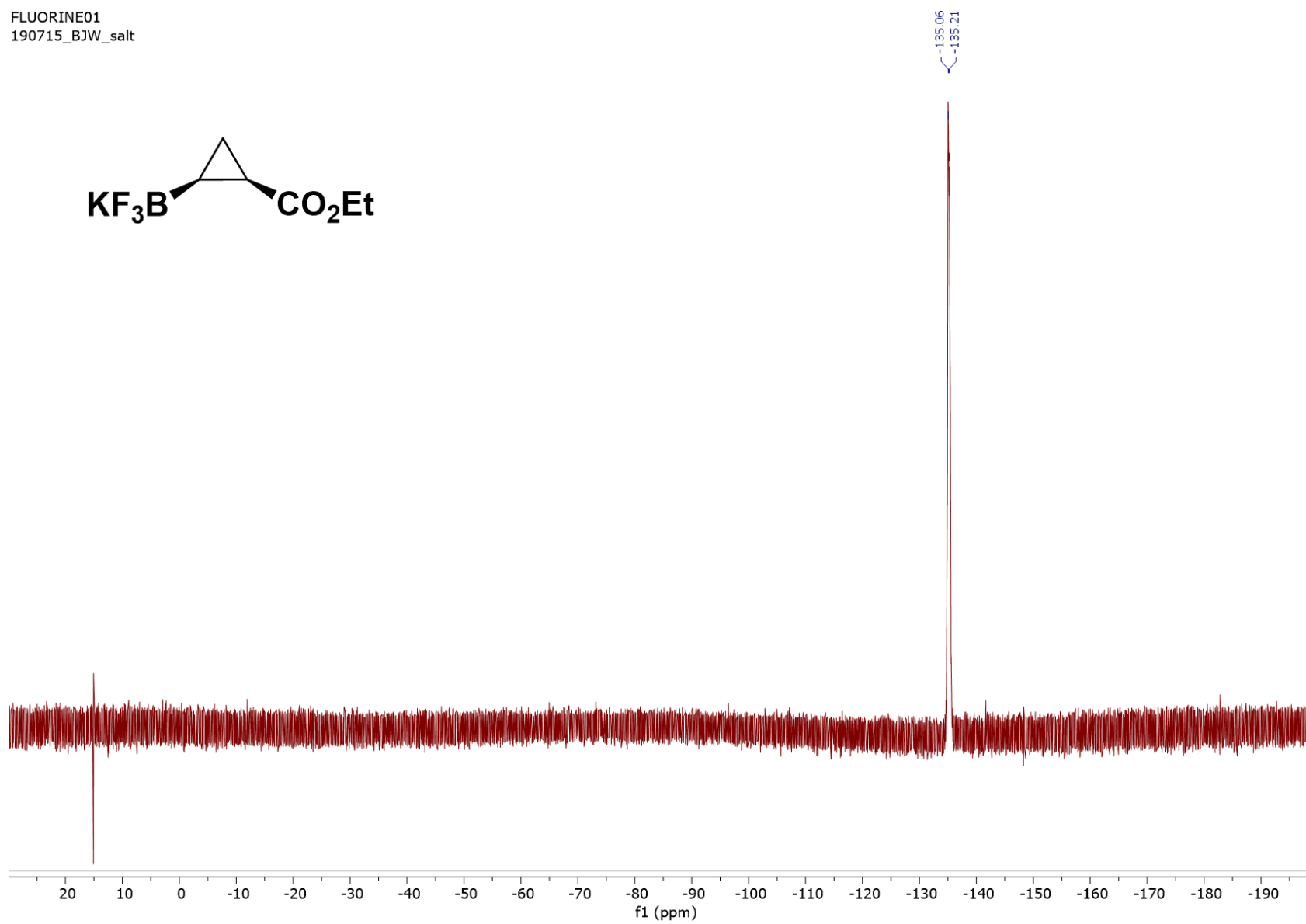


Figure S26. <sup>19</sup>F NMR of *cis*-4.

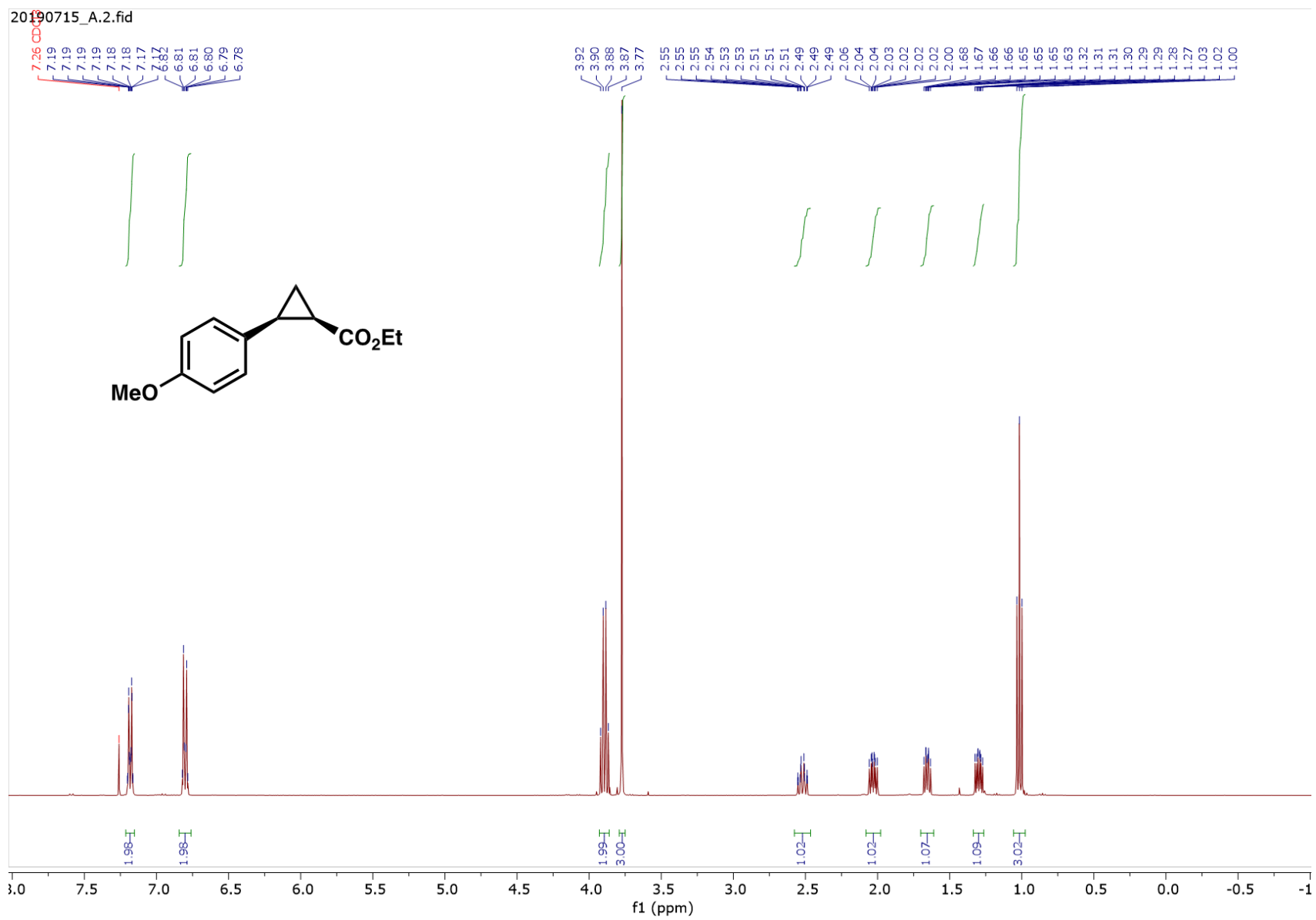


Figure S27.  $^1\text{H}$  NMR of *cis*-5a.

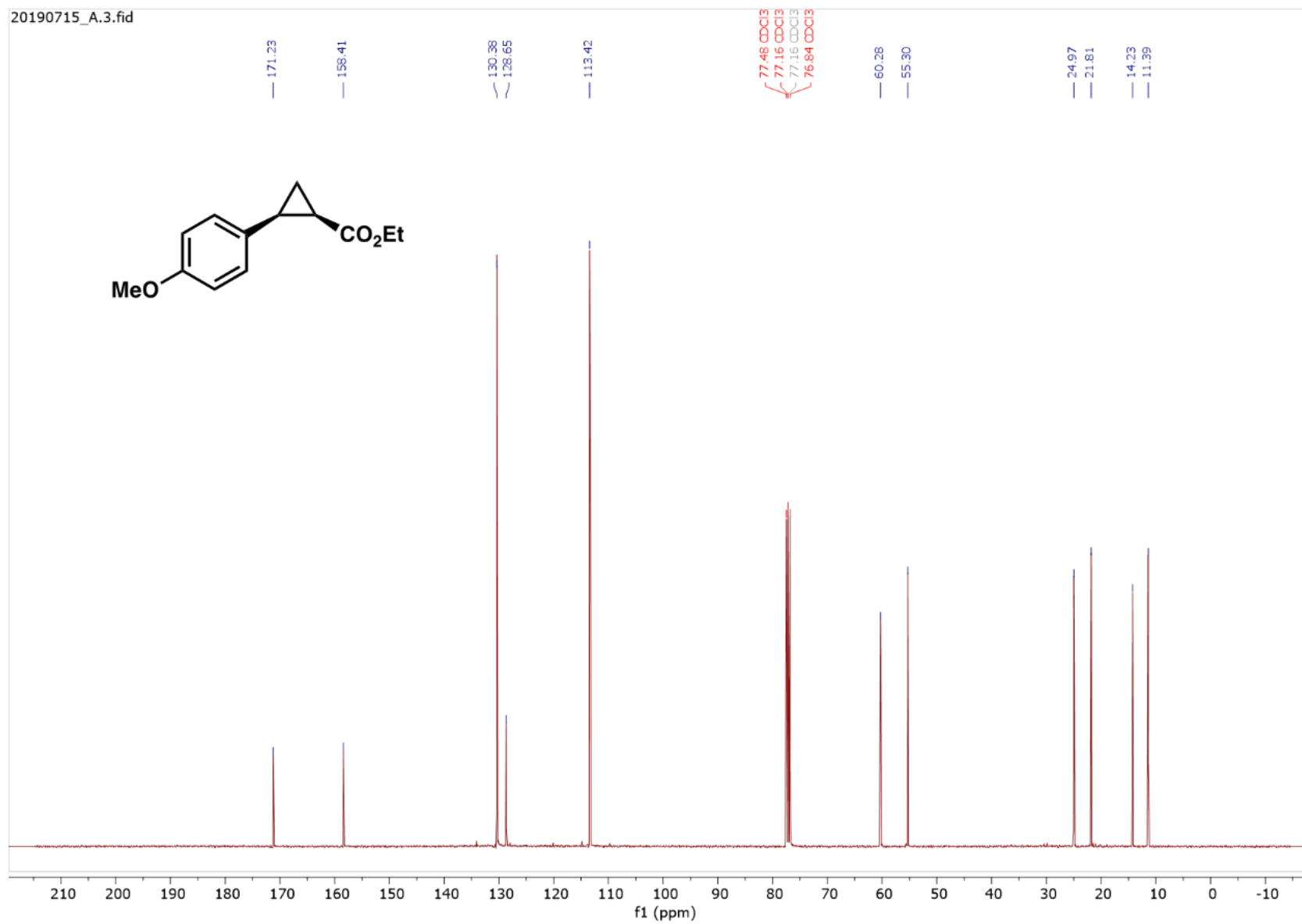
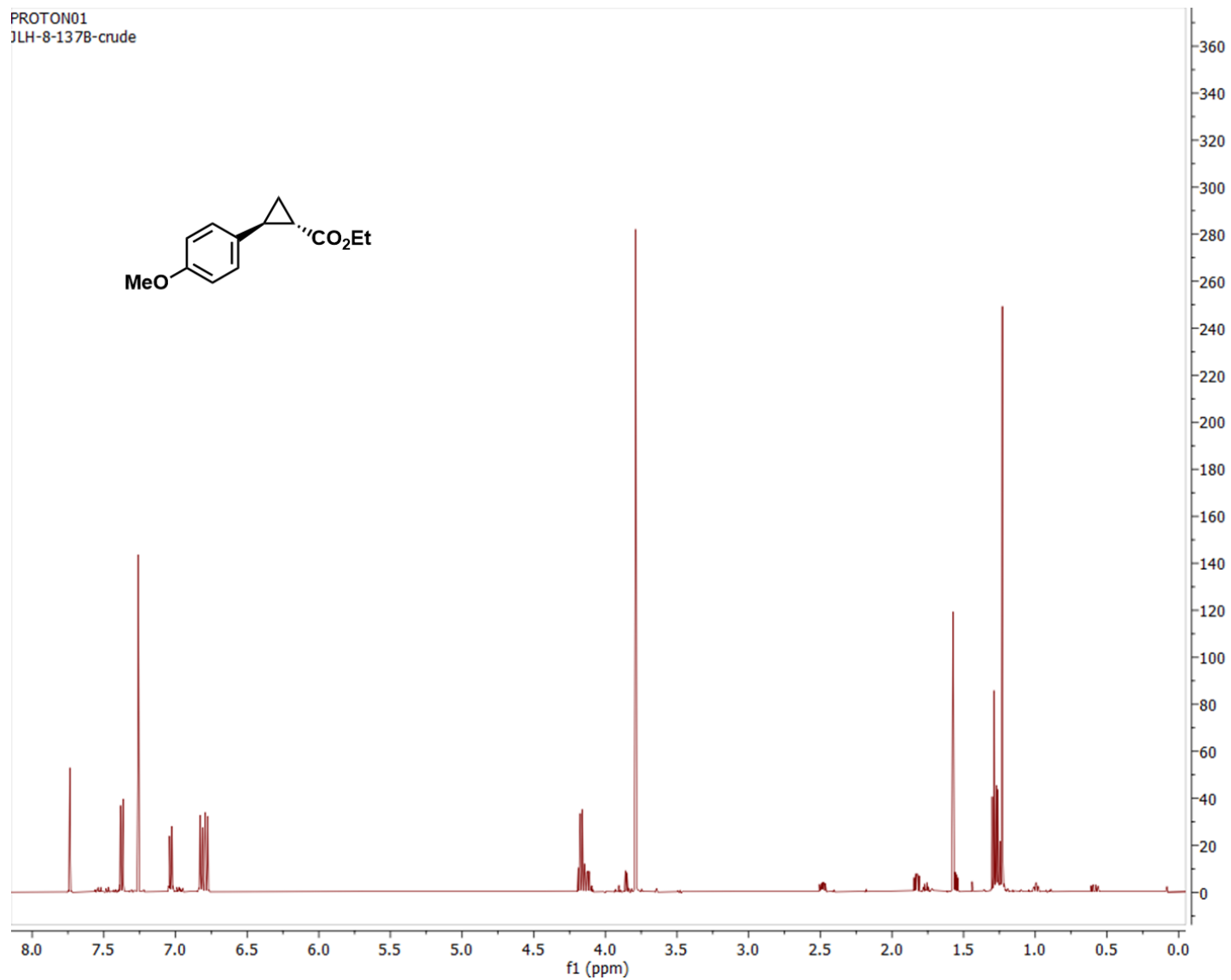


Figure S28. <sup>13</sup>C NMR of *cis*-5a.





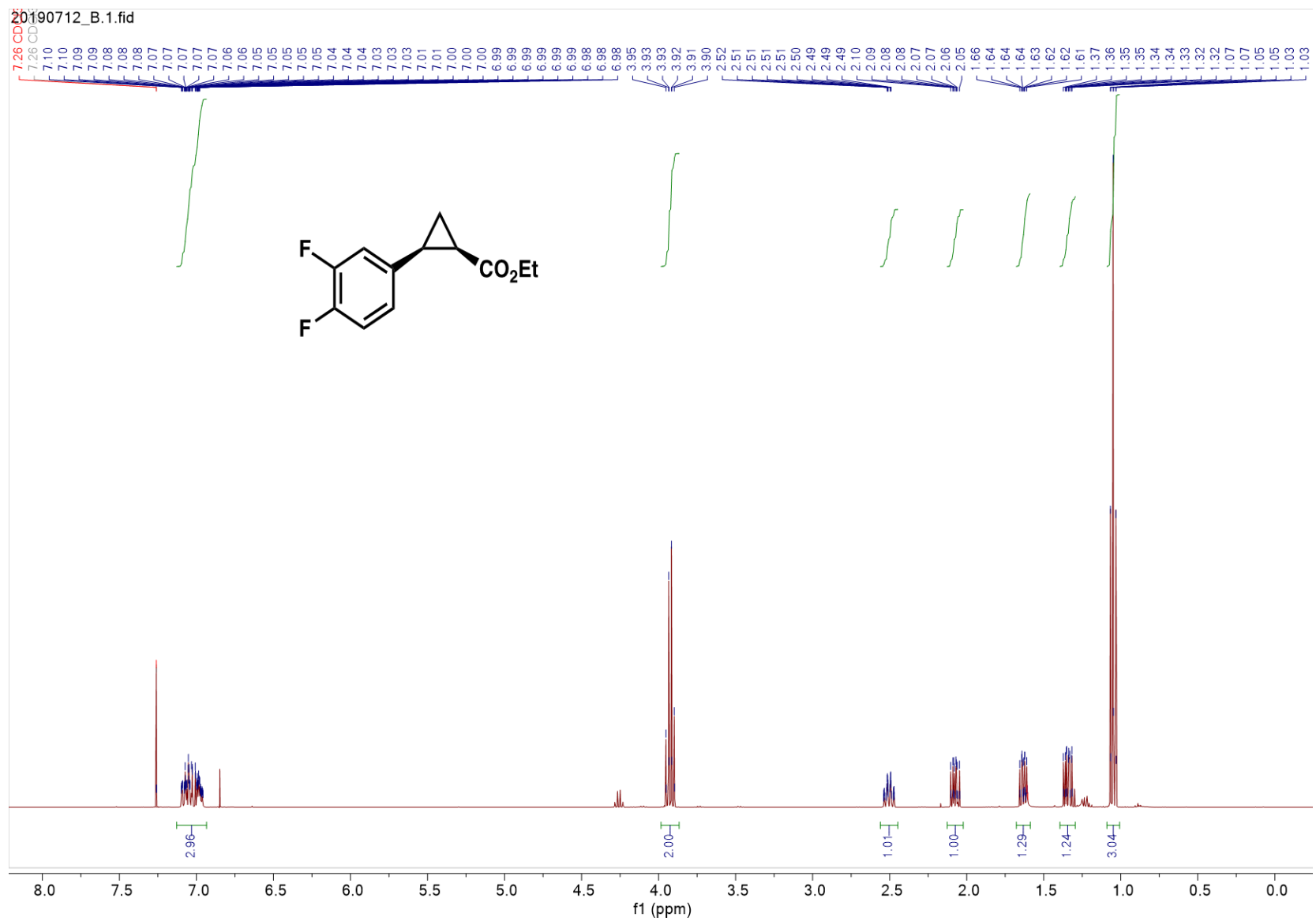


Figure S30.  $^1\text{H}$  NMR of *cis*-5b.

20190712\_B.2.fid

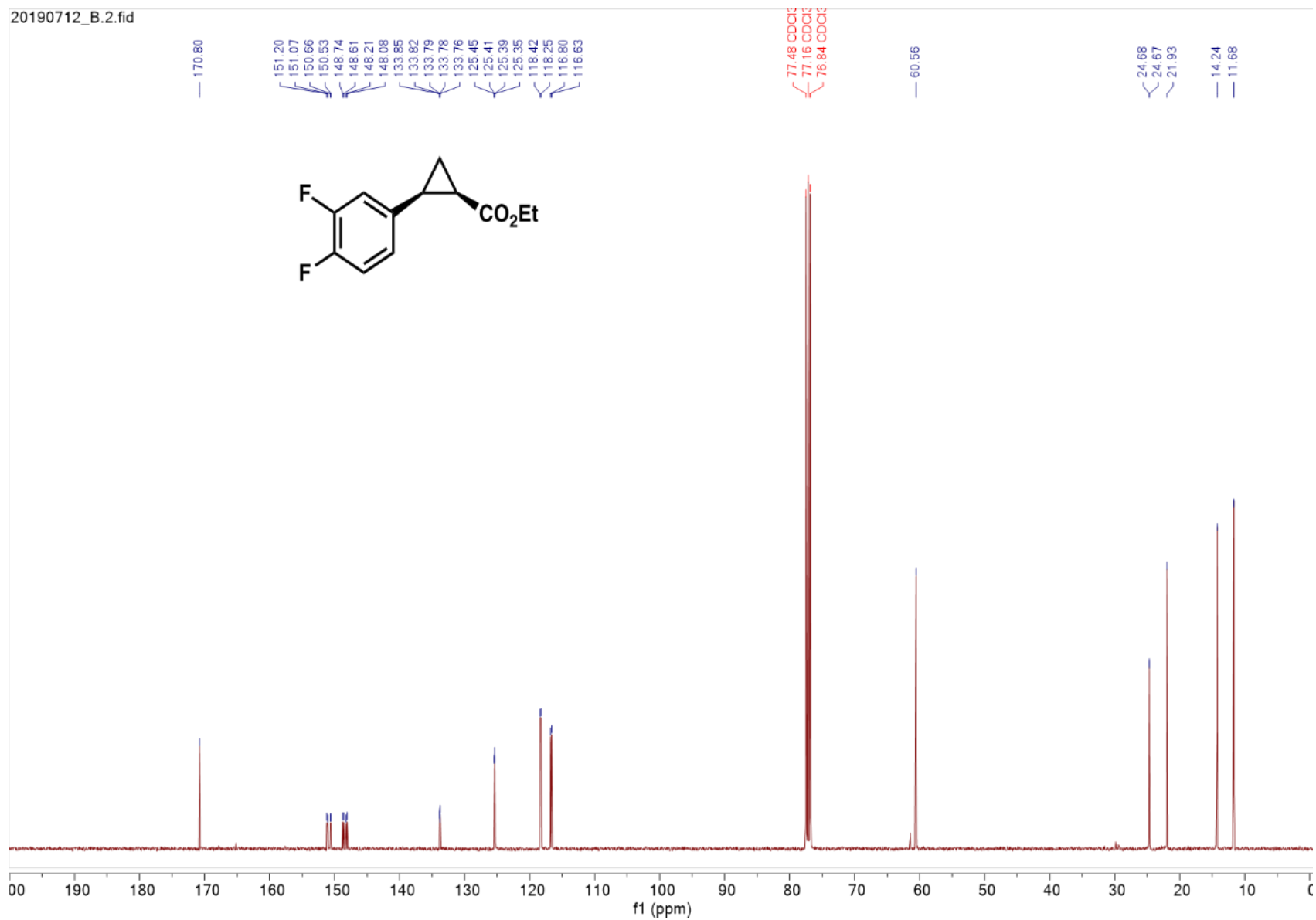


Figure S31. <sup>13</sup>C NMR of *cis*-5b.

FLUORINE01  
190715-BJW-B

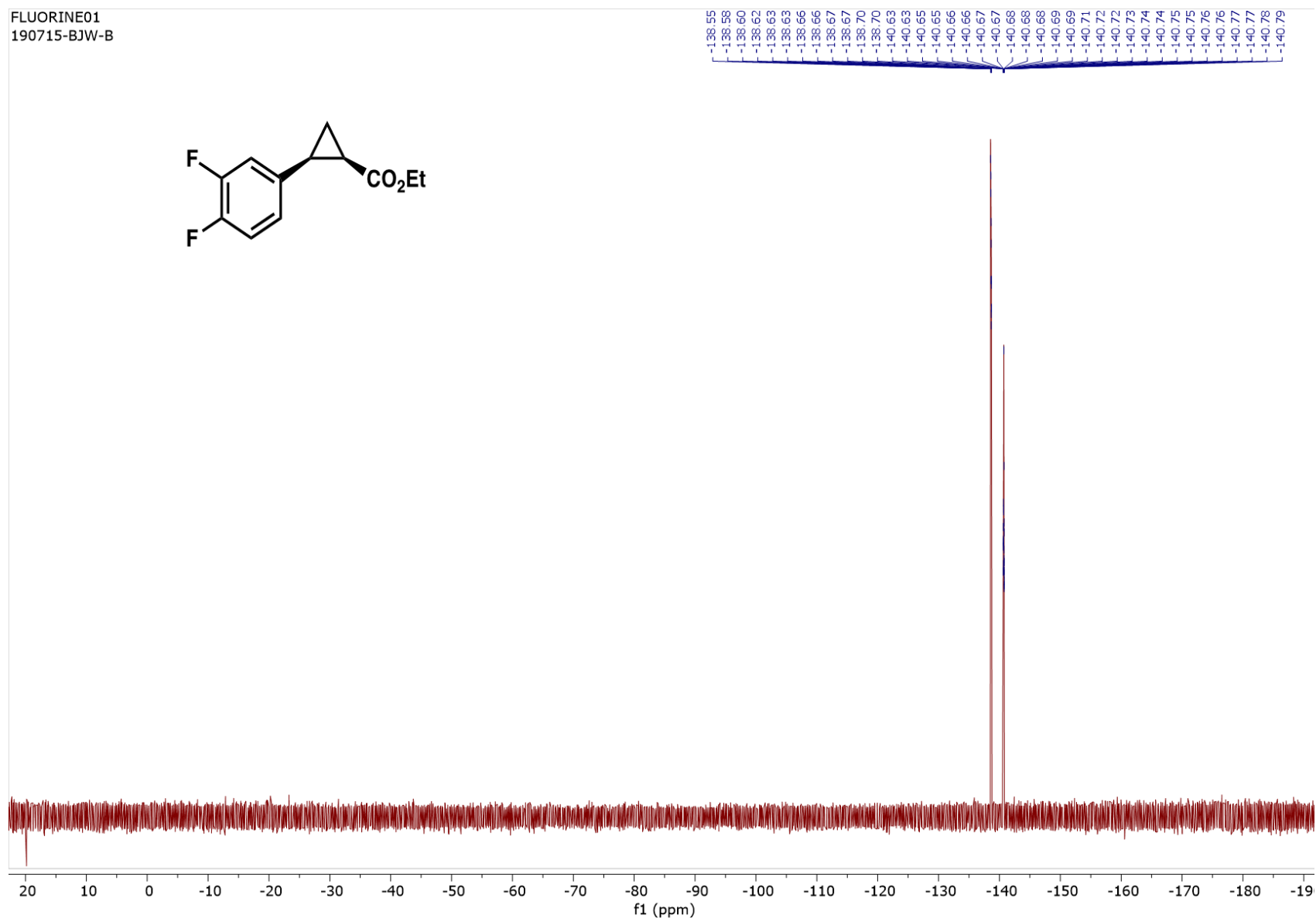
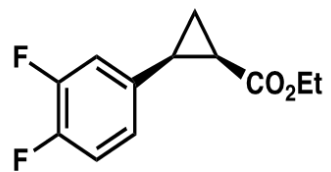


Figure S32.  $^{19}\text{F}$  NMR of *cis*-5b.

PROTON01  
JLH-8-146B

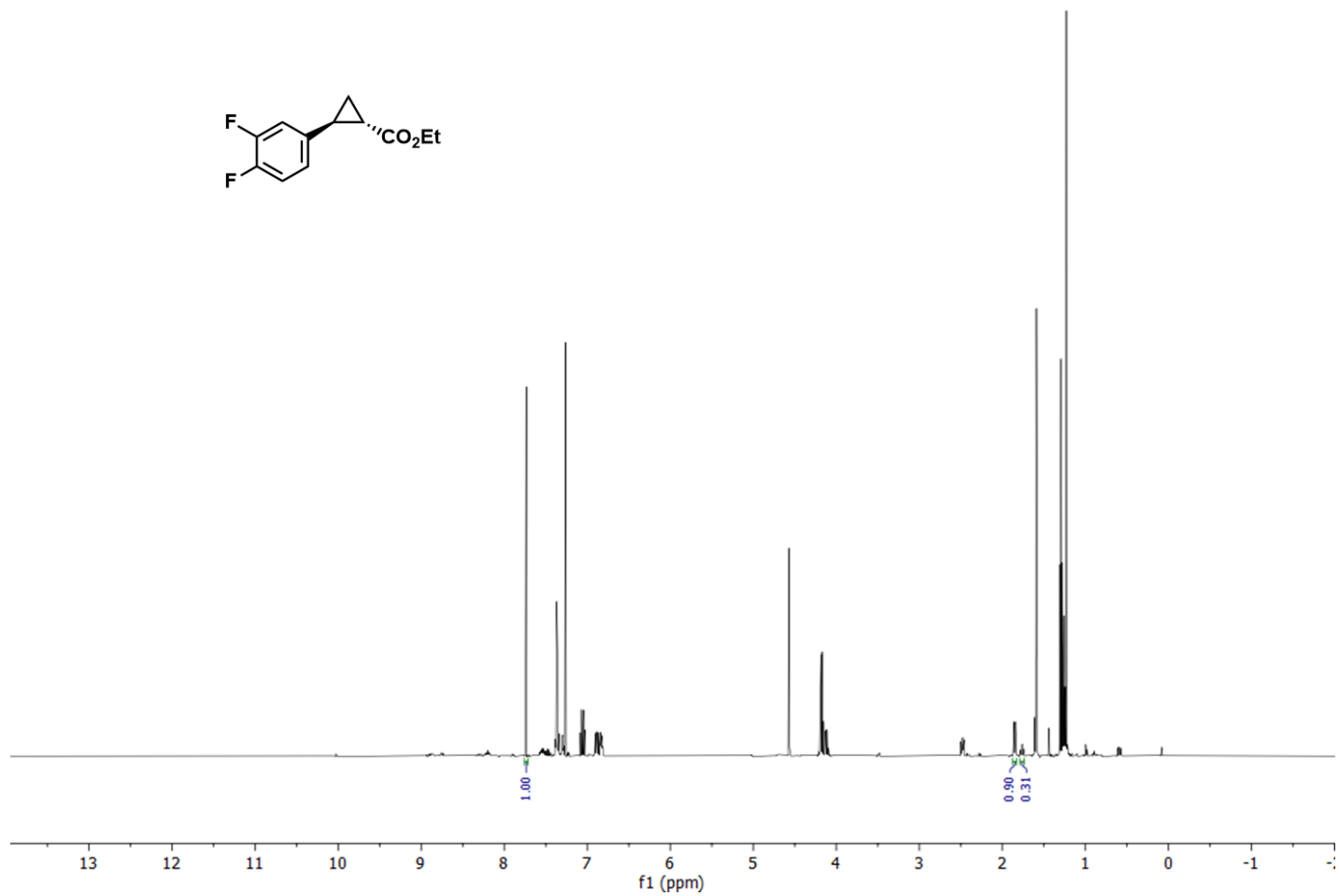
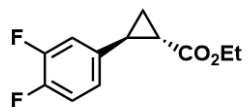


Figure S33. <sup>1</sup>H NMR of crude *trans*-5b.

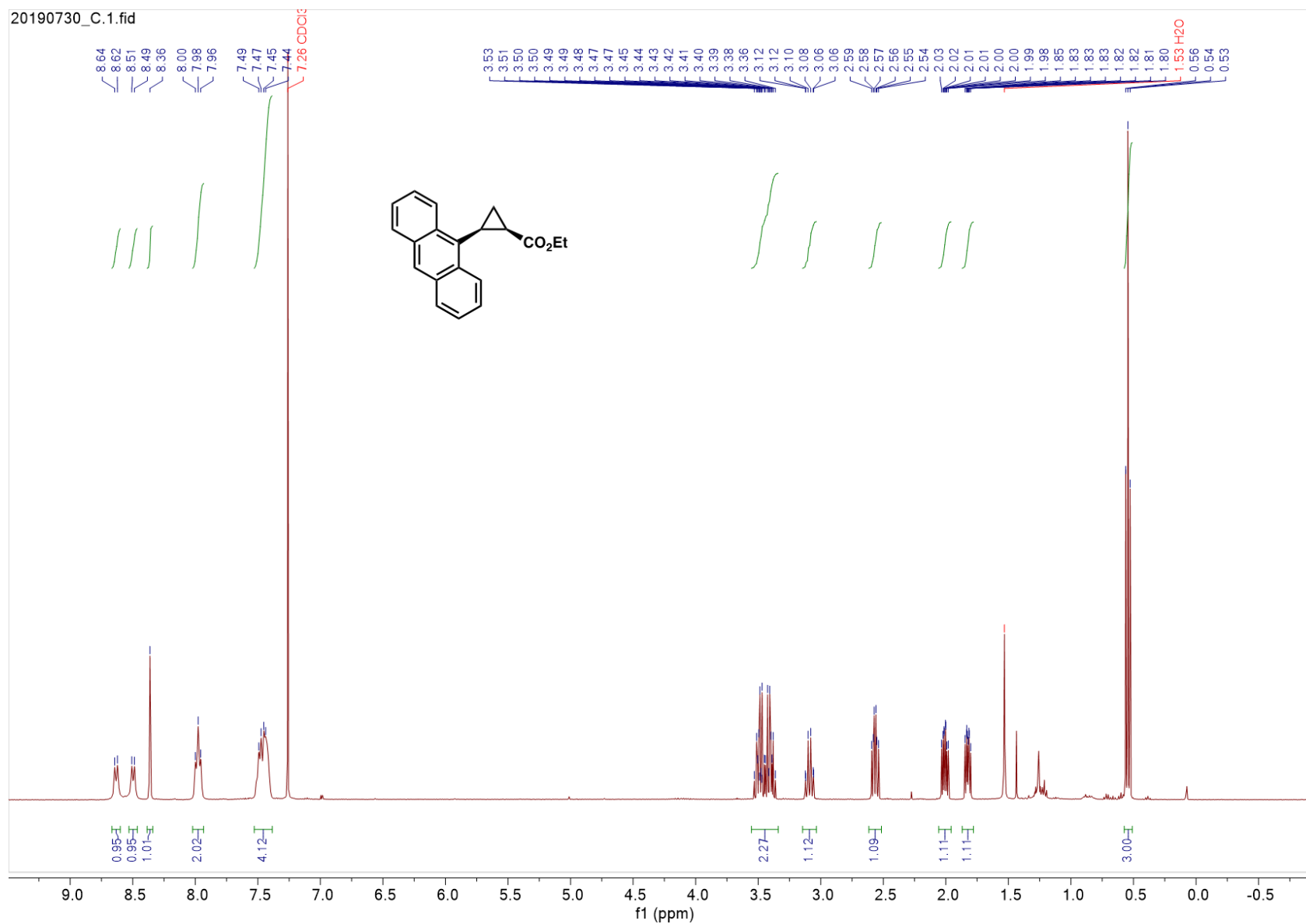


Figure S34. <sup>1</sup>H NMR of *cis*-5c.

20200303\_C.1.fid

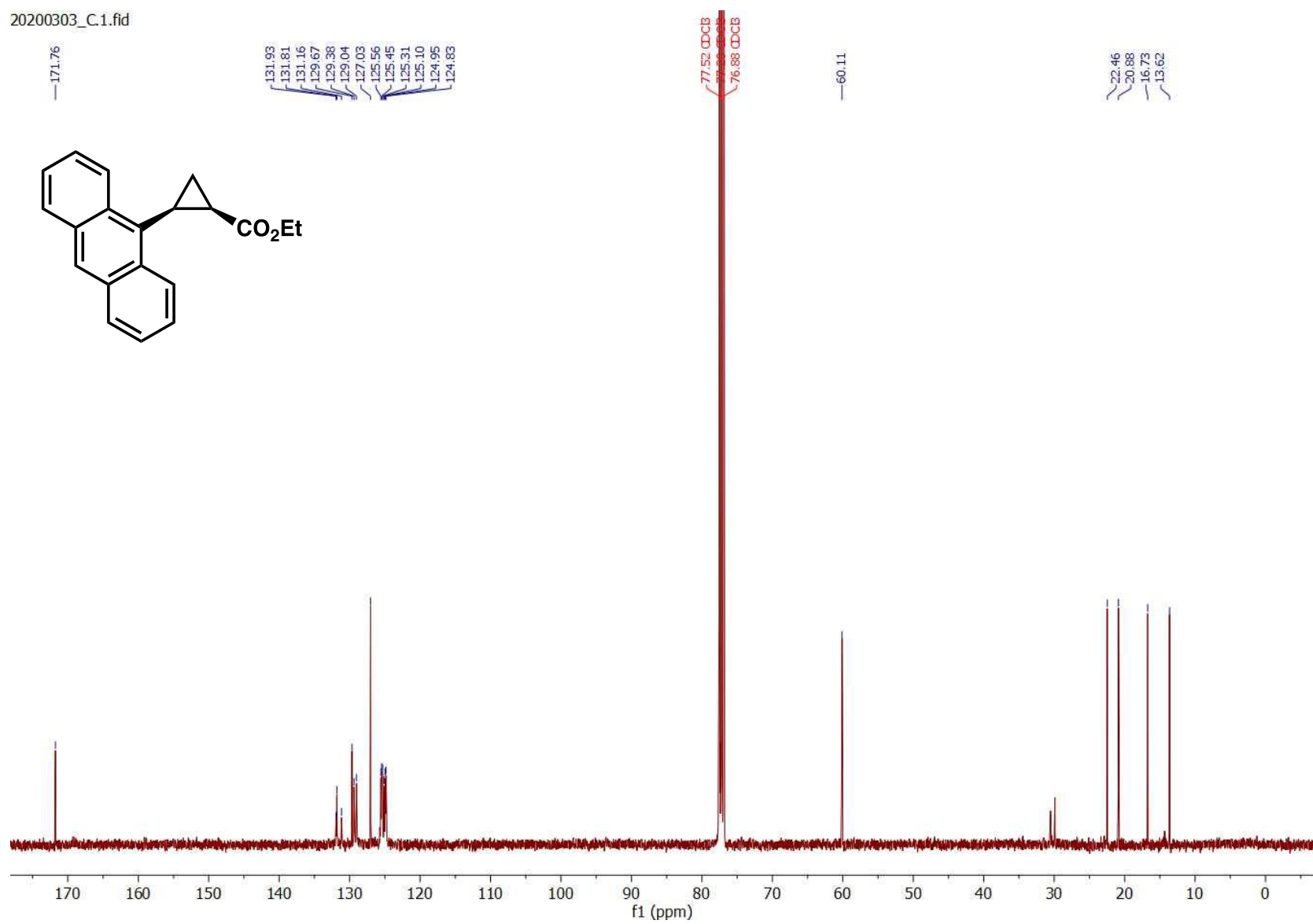


Figure S35. <sup>13</sup>C NMR of *cis*-5c.

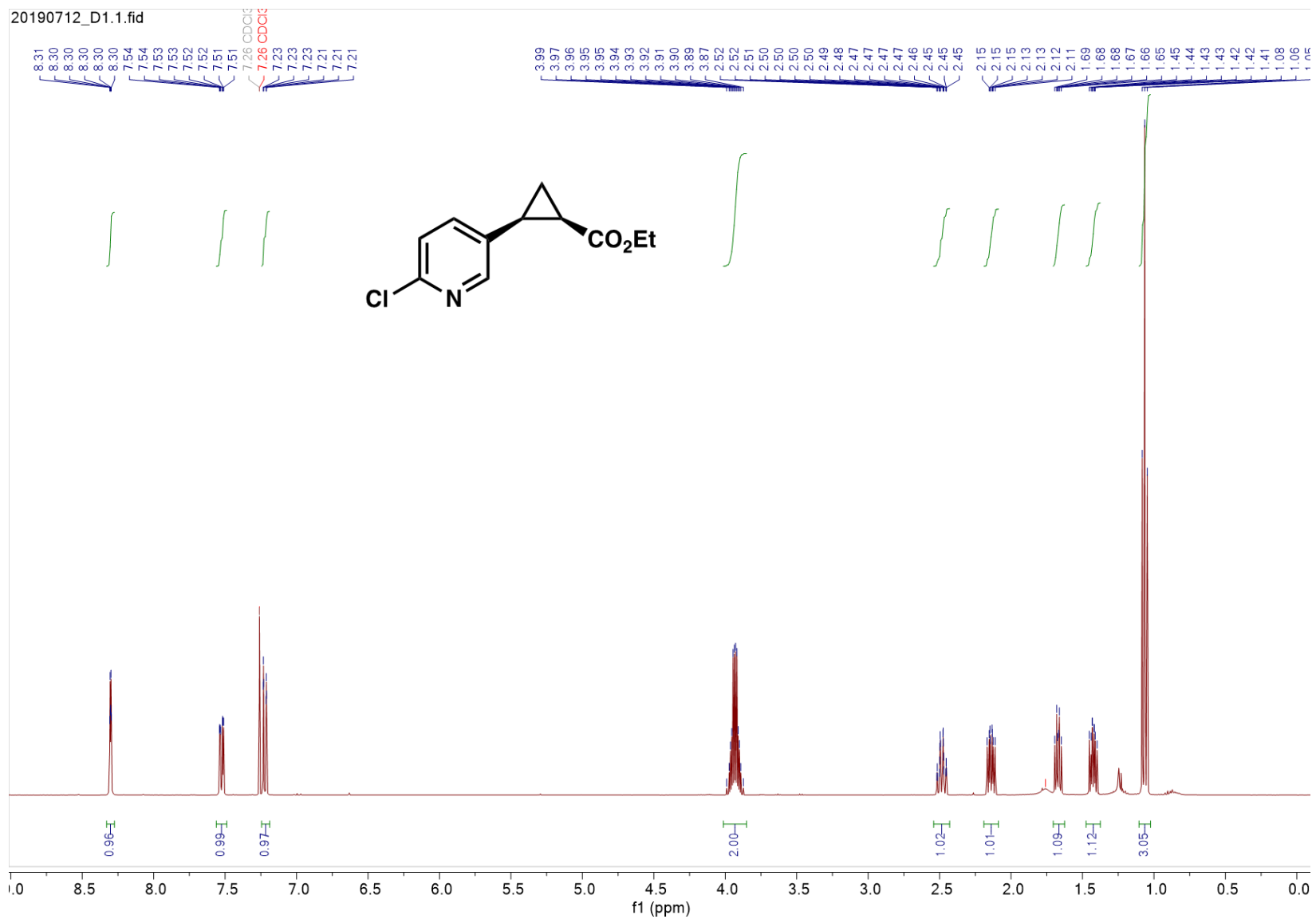


Figure S36. <sup>1</sup>H NMR of *cis*-5d.



20190712\_D1.2.fid

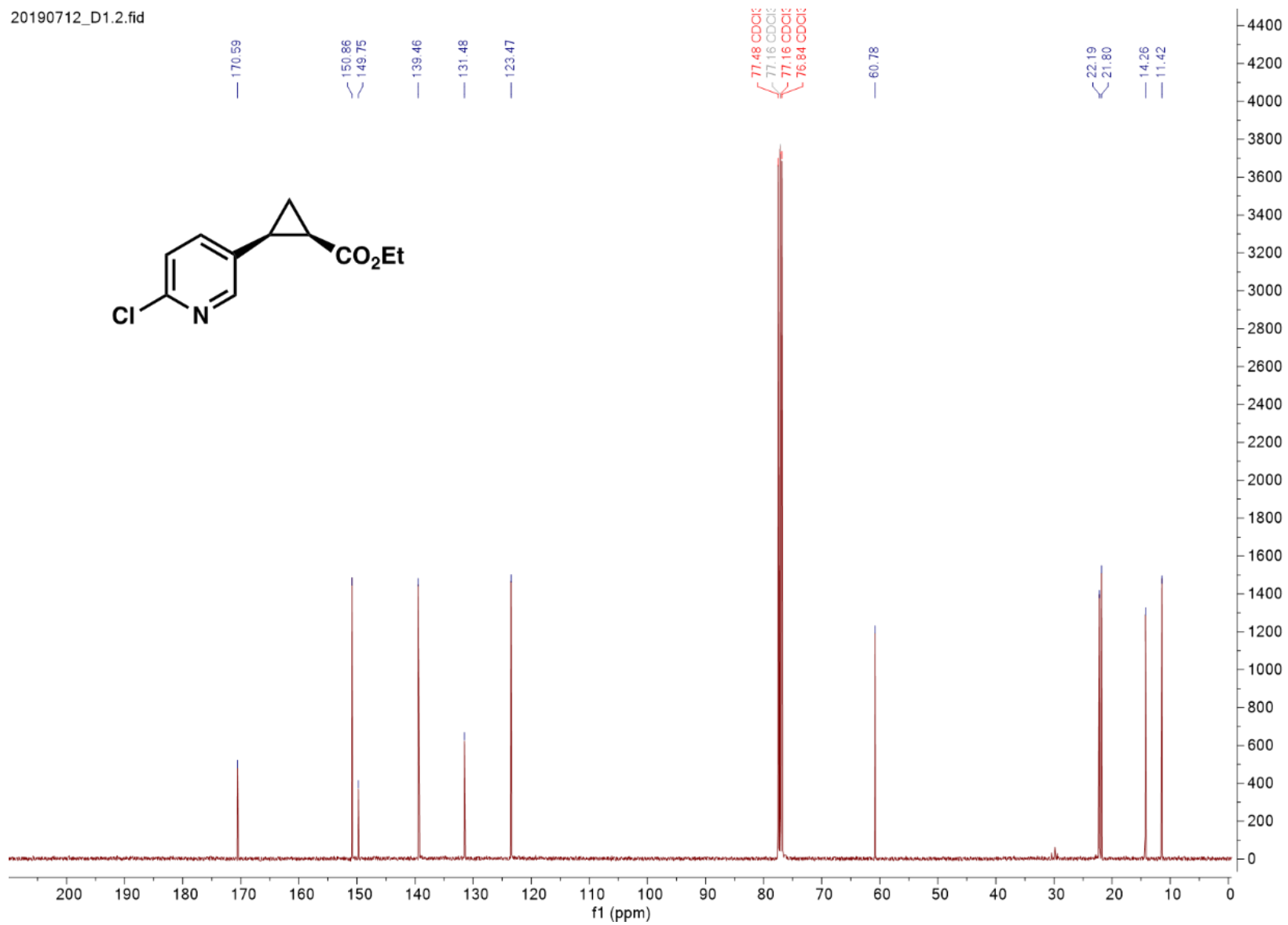


Figure S37. <sup>13</sup>C NMR of *cis*-5d.

## References

- (1) Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. Reducing Codon Redundancy and Screening Effort of Combinatorial Protein Libraries Created by Saturation Mutagenesis. *ACS Synth. Biol.* **2013**, *2*, 83–92. <https://doi.org/10.1021/sb300037w>.
- (2) Barr, I.; Guo, F. Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *Bio-Protocol* **2015**, *5*. <https://doi.org/10.21769/bioprotoc.1594>.
- (3) Bagutski, V.; Ros, A.; Aggarwal, V. K. Improved Method for the Conversion of Pinacolboronic Esters into Trifluoroborate Salts: Facile Synthesis of Chiral Secondary and Tertiary Trifluoroborates. *Tetrahedron* **2009**, *65*, 9956–9960. <https://doi.org/10.1016/j.tet.2009.10.002>.
- (4) Brandenburg, O. F.; Fasan, R.; Arnold, F. H. Exploiting and Engineering Hemoproteins for Abiological Carbene and Nitrene Transfer Reactions. *Curr. Opin. Biotechnol.* **2017**, *47*, 102–111. <https://doi.org/10.1016/j.copbio.2017.06.005>.
- (5) Knight, A. M.; Kan, S. B. J.; Lewis, R. D.; Brandenburg, O. F.; Chen, K.; Arnold, F. H. Diverse Engineered Heme Proteins Enable Stereodivergent Cyclopropanation of Unactivated Alkenes. *ACS Cent. Sci.* **2018**, *4*, 372–377. <https://doi.org/10.1021/acscentsci.7b00548>.
- (6) Bordeaux, M.; Tyagi, V.; Fasan, R. Highly Diastereoselective and Enantioselective Olefin Cyclopropanation Using Engineered Myoglobin-Based Catalysts. *Angew. Chem. Int. Ed.* **2015**, *54*, 1744–1748. <https://doi.org/10.1002/anie.201409928>.
- (7) Gibson, D. G.; Young, L.; Chuang, R.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6*, 343–345. <https://doi.org/10.1038/nmeth.1318>.
- (8) Kabsch, W. Xds. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125–132. <https://doi.org/10.1107/S0907444909047337>.
- (9) Terwilliger, T. C.; Adams, P. D.; Read, R. J.; McCoy, A. J.; Moriarty, N. W.; Grosse-Kunstleve, R. W.; Afonine, P. V.; Zwart, P. H.; Hung, L. W. Decision-Making in Structure Solution Using Bayesian Estimates of Map Quality: The PHENIX AutoSol Wizard. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *65*, 582–601. <https://doi.org/10.1107/S0907444909012098>.
- (10) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213–221. <https://doi.org/10.1107/S0907444909052925>.
- (11) Evans, P. R.; Murshudov, G. N. How Good Are My Data and What Is the Resolution? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204–1214. <https://doi.org/10.1107/S0907444913000061>.

- (12) Bunkóczi, G.; Echols, N.; McCoy, A. J.; Oeffner, R. D.; Adams, P. D.; Read, R. J. Phaser.MRage: Automated Molecular Replacement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 2276–2286. <https://doi.org/10.1107/S0907444913022750>.
- (13) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486–501. <https://doi.org/10.1107/S0907444910007493>.
- (14) *APEX2, Version 2 User Manual, M86-E01078, Bruker Analytical X-Ray Systems*; Madison, WI, 2006.
- (15) Sheldrick, G. M. A Short History of SHELX. *Acta Crystallogr. A* **2008**, *A64*, 112–122. <https://doi.org/10.1107/S0108767307043930>.
- (16) Sheldrick, G. M. Crystal Structure Refinement with SHELXL. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **2015**, *71*, 3–8. <https://doi.org/10.1107/S2053229614024218>.
- (17) Müller, P. Practical Suggestions for Better Crystal Structures. *Crystallogr. Rev.* **2009**, *15*, 57–83. <https://doi.org/10.1080/08893110802547240>.
- (18) Parsons, S.; Flack, H. D.; Wagner, T. Use of Intensity Quotients and Differences in Absolute Structure Refinement. *Acta Crystallogr. B* **2013**, *B69*, 249–259. <https://doi.org/10.1107/S2052519213010014>.
- (19) Chawner, S. J.; Cases-Thomas, M. J.; Bull, J. A. Divergent Synthesis of Cyclopropane-Containing Lead-Like Compounds, Fragments and Building Blocks through a Cobalt Catalyzed Cyclopropanation of Phenyl Vinyl Sulfide. *Eur. J. Org. Chem.* **2017**, *34*, 5015–5024. <https://doi.org/10.1002/ejoc.201701030>.