Supplementary Material (ESI) for *Organic and Biomolecular Chemistry*

Synthesis of Optically Active SARS-CoV-2 MPro inhibitor drug Nirmatrelvir (Paxlovid): an Approved Treatment of COVID-19

Arun K. Ghosh*,[a,b] and Monika Yadav[a]

[a]*Department of Chemistry and* [b]*Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, 560 Oval Drive, West Lafayette, IN 47907.*

Table of Contents

Experimental Procedures

General Methods. All chemicals and reagents were purchased from commercial suppliers and used without further purification unless otherwise noted. Tetrahydrofuran was distilled from Na/Benzophenone prior to use. All reactions were performed in oven-dried round-bottom flasks followed by flame-drying under vacuum in the case of moisture-sensitive reactions. The flasks were fitted with rubber septa and kept under a positive pressure of argon. Cannula were used in the transfer of moisture-sensitive liquids. Heated reactions were ran using an oil bath on a hot plate equipped with a temperature probe. TLC analysis was conducted using glass-backed thin-layer silica gel chromatography plates (60 Å, 250 µm thickness, F-254 indicator). Flash

chromatography was done using a 230−400 mesh, a 60 Å pore diameter silica gel. Organic solutions were concentrated at 30−35 °C on rotary evaporators capable of achieving a minimum pressure of ∼25 Torr and further concentrated on a Hi-vacuum pump capable of achieving a minimum pressure of ∼4 Torr. ^{1}H NMR spectra were recorded on 400, 500, and 800 MHz spectrometers. ^{13}C NMR spectra were recorded at 100, 125, and 200 MHz on the respective NMRs. Chemical shifts are reported in parts per million and referenced to the deuterated residual solvent peak ($CDCl_3$, 7.26 ppm for ^{1}H and 77.16 ppm for ^{13}C). NMR data are reported as δ value (chemical shift), J-value (Hz), and integration, where s = singlet, bs = broad singlet, d = doublet, t = triplet, q = quartet, p = quintet, m = multiplet, dd = doublet doublets, and so on. Optical rotations were measured on a Rudolph's AUTOPOL-III automatic digital polarimeter with a sodium lamp and are reported as follows: [α]λ T °C (c = g/100 mL, solvent). Low resolution mass spectra (LRMS) spectra were recorded using a quadrupole LCMS under positive electrospray ionization (ESI+). High-resolution mass spectrometry (HRMS) spectra were recorded at the Purdue University Department of Chemistry Mass Spectrometry Center. These experiments were performed under ESI+ and positive atmospheric pressure chemical ionization (APCI+) conditions using an Orbitrap XL Instrument. HPLC analysis was done an on Agilent 1260 series instrument using a YMC Pack ODS-A (4.6x150mm) C-18 column and eluent: 60% MeOH/ H_2O.

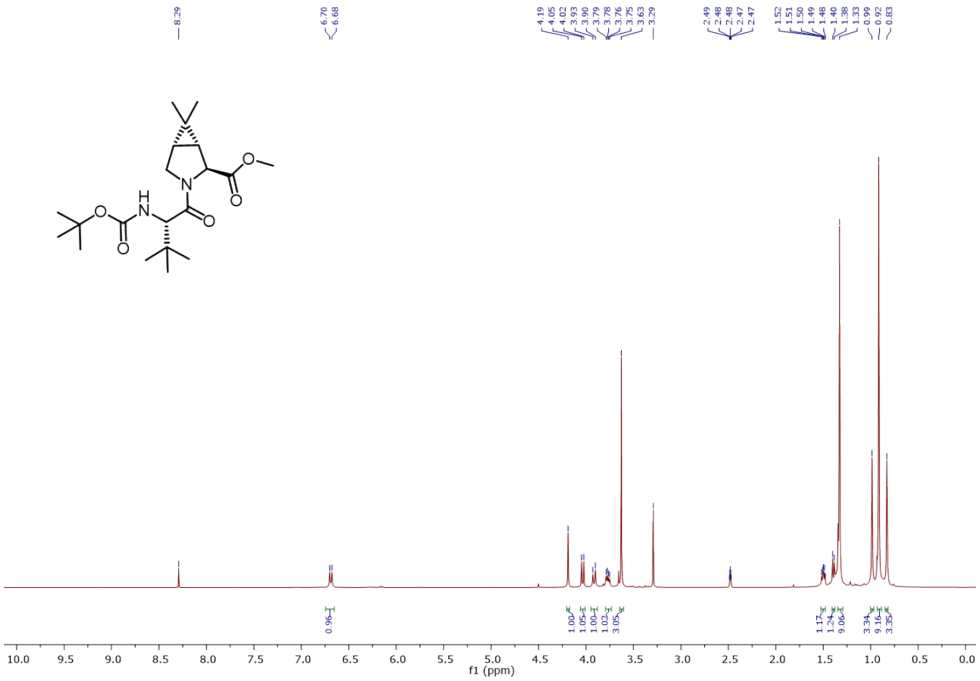


Figure S1. ^{1}H NMR (400 MHz, DMSO-d_6) of amide **10**.

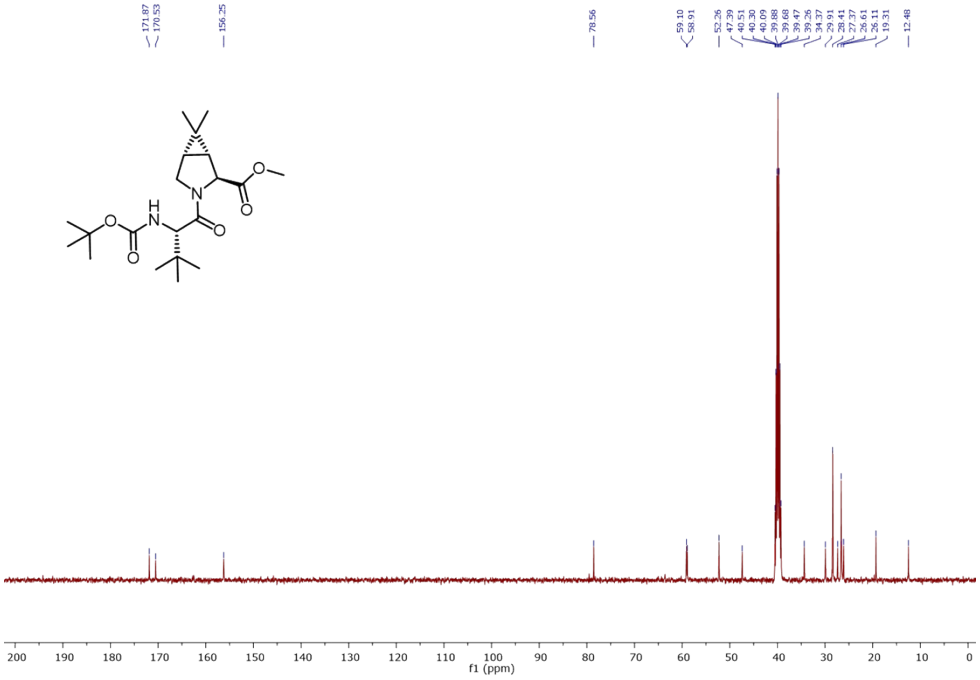


Figure S2. ^{13}C NMR (100 MHz, DMSO-d_6) of amide **10**.

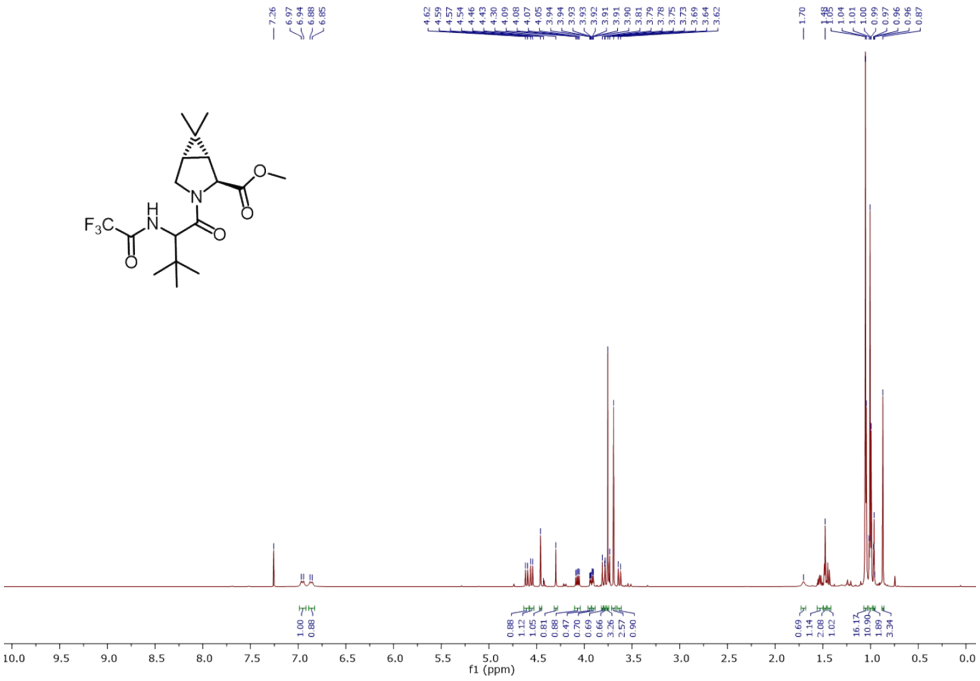


Figure S3. ^{1}H NMR (400 MHz, $CDCl_3$) of trifluoroacetamide **epimer-8**.

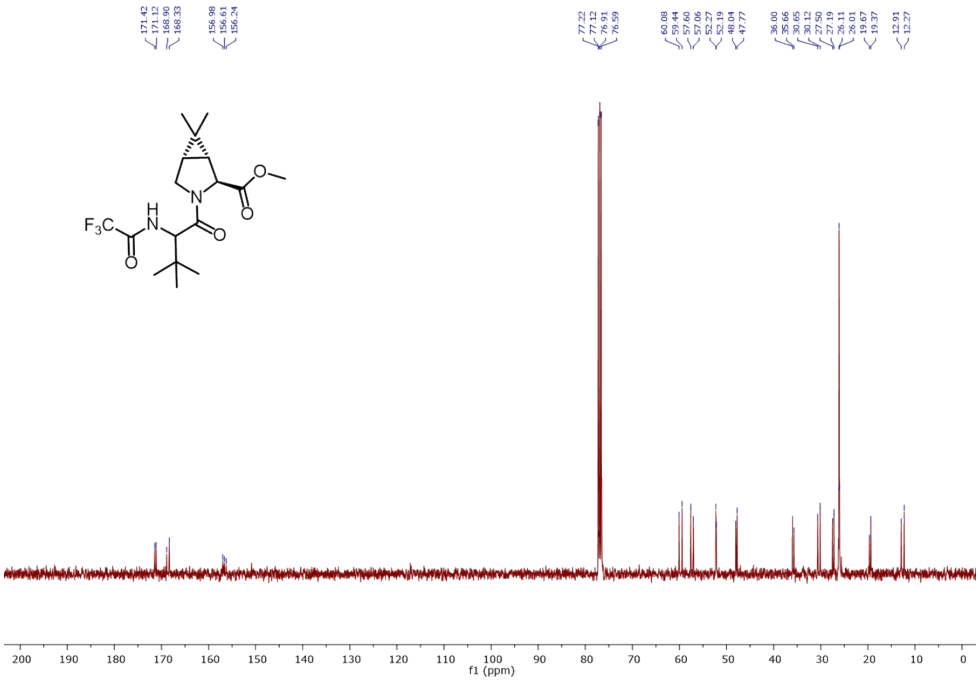


Figure S4. ^{13}C NMR (100 MHz, $CDCl_3$) of trifluoroacetamide **epimer-8**.

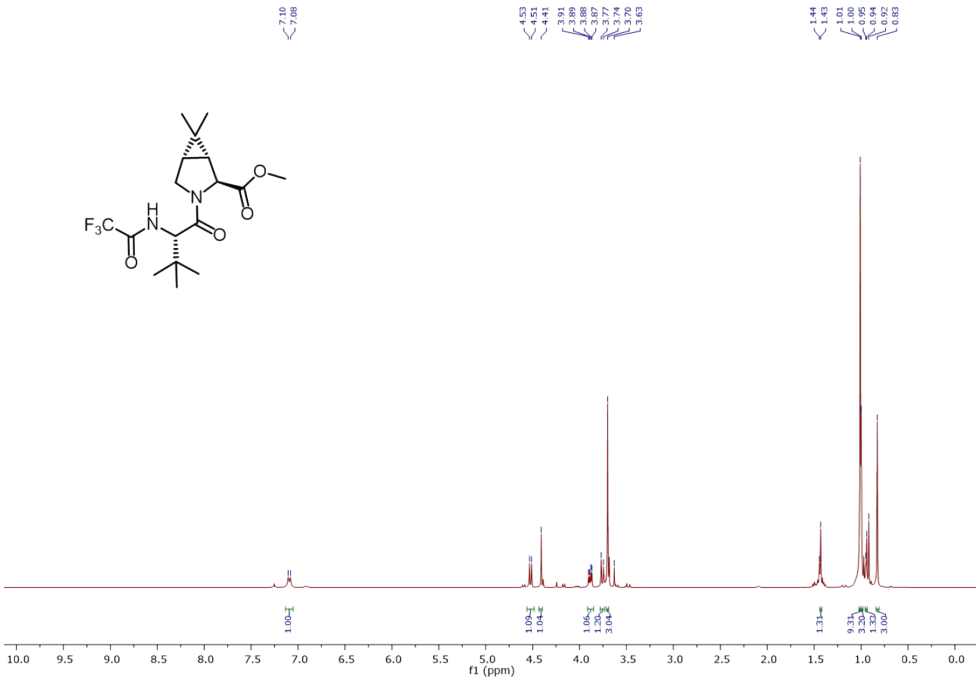


Figure S5. ^{1}H NMR (400 MHz, $CDCl_3$) of trifluoroacetamide **8a**.

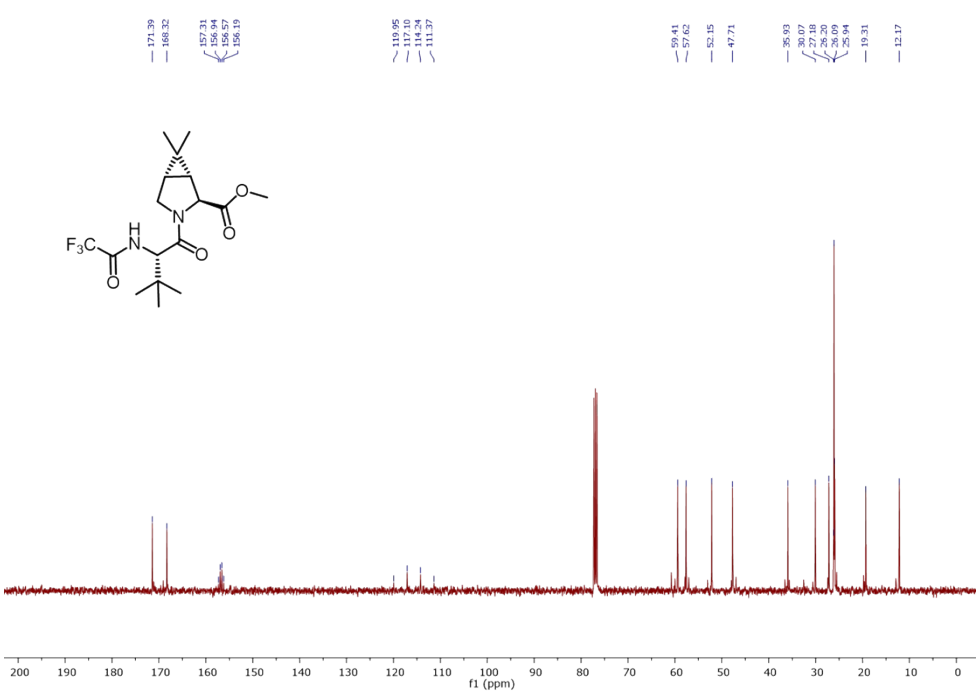


Figure S6. ^{13}C NMR (100 MHz, $CDCl_3$) of trifluoroacetamide **8a**.

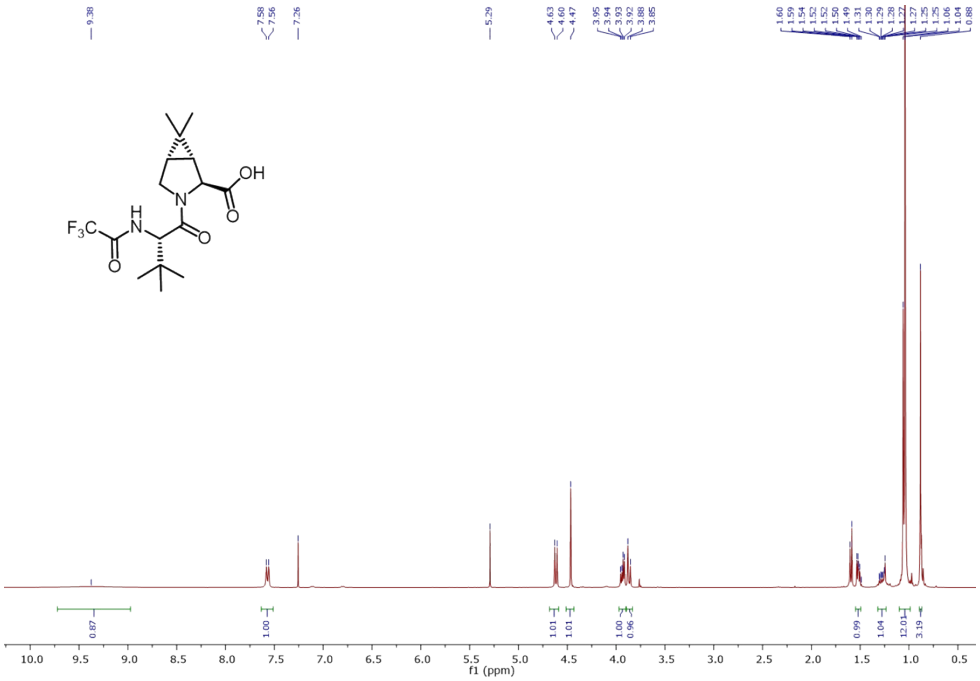


Figure S7. ^{1}H NMR (400 MHz, $CDCl_3$) of acid **2**.

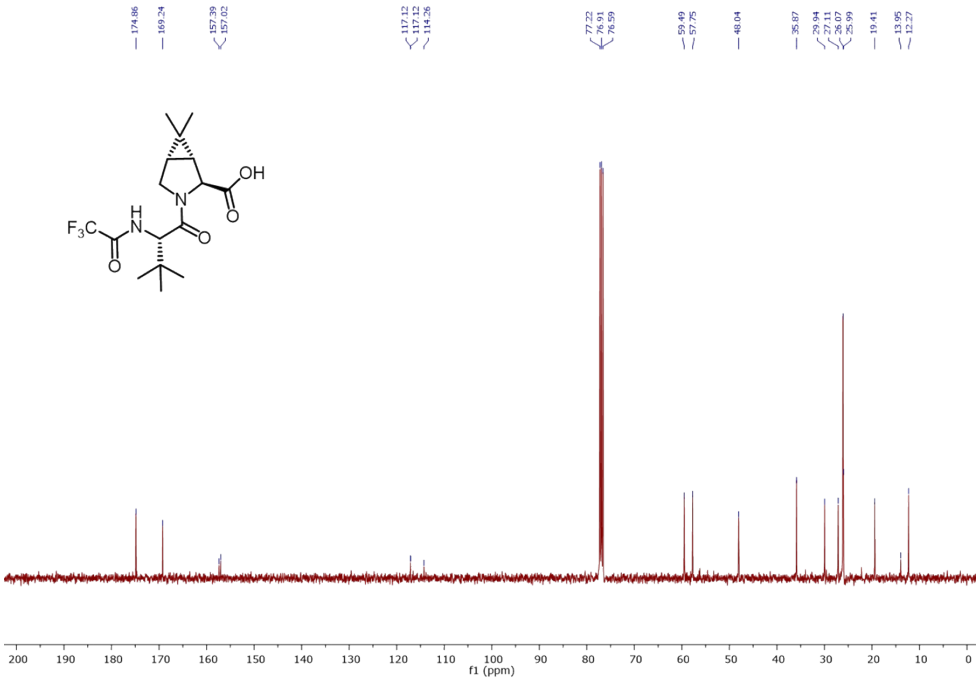


Figure S8. ^{13}C NMR (100 MHz, $CDCl_3$) of acid **2**.

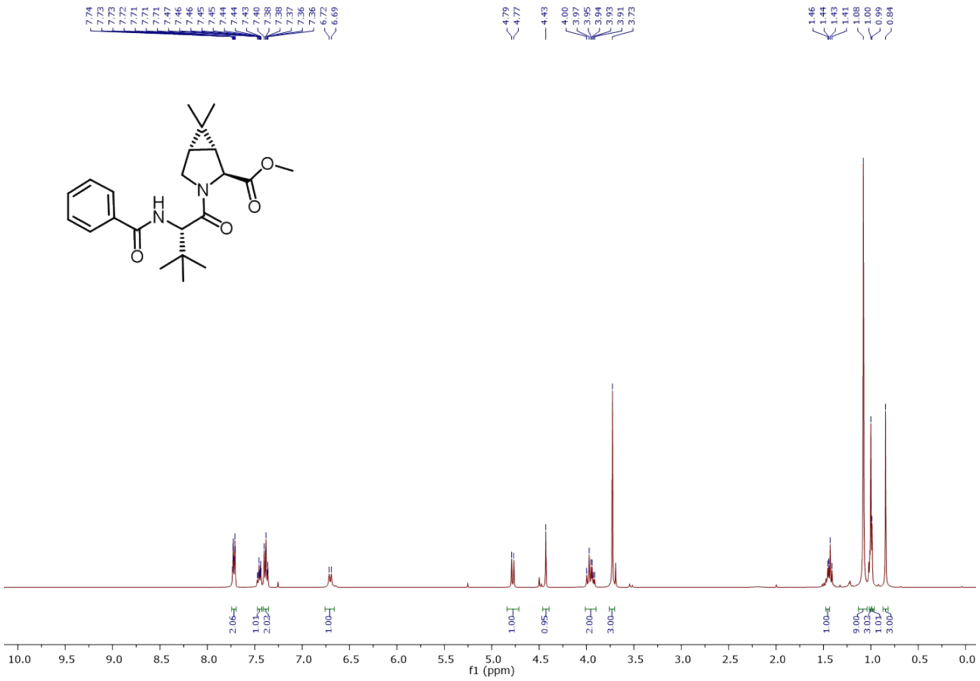


Figure S9. ^{1}H NMR (400 MHz, $CDCl_3$) of benzamide **8b**.

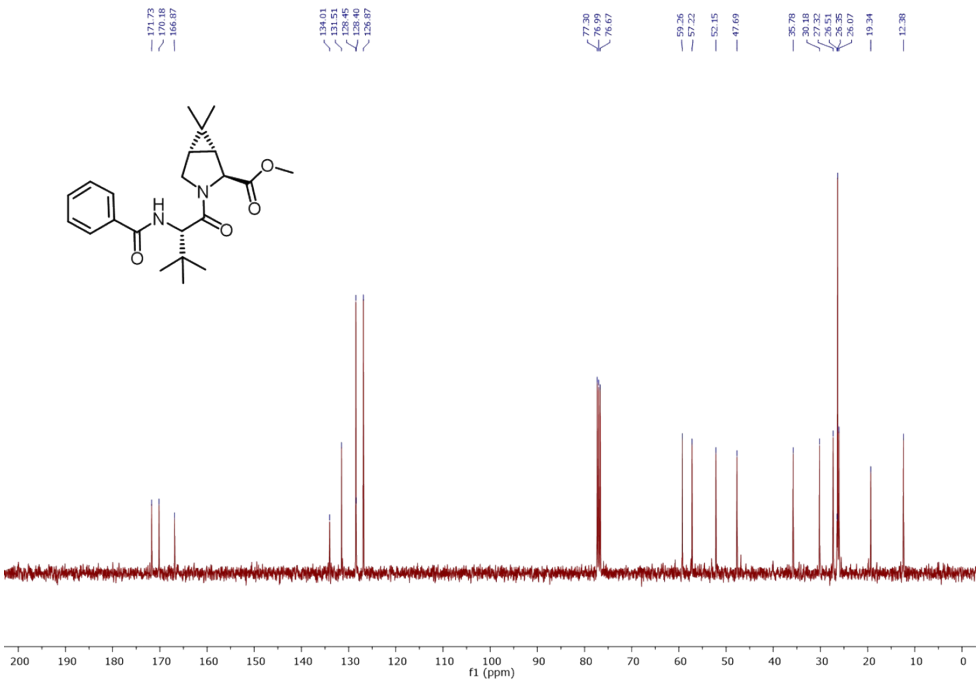


Figure S10. ^{13}C NMR (100 MHz, $CDCl_3$) of benzamide **8b**.

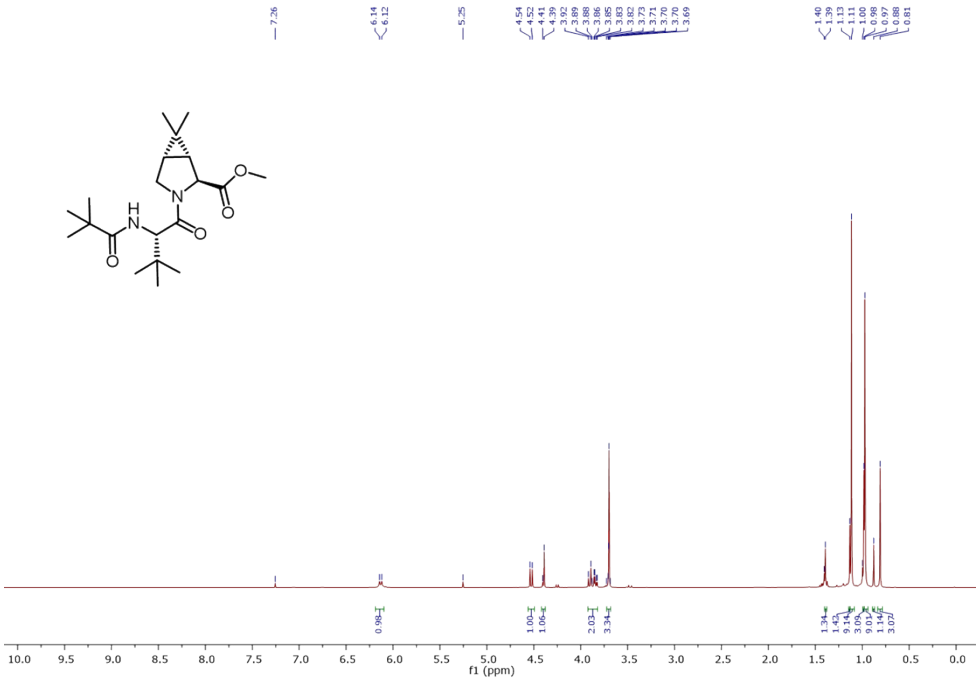


Figure S11. ^{1}H NMR (400 MHz, $CDCl_3$) of *tert*-butyl amide **8c**.

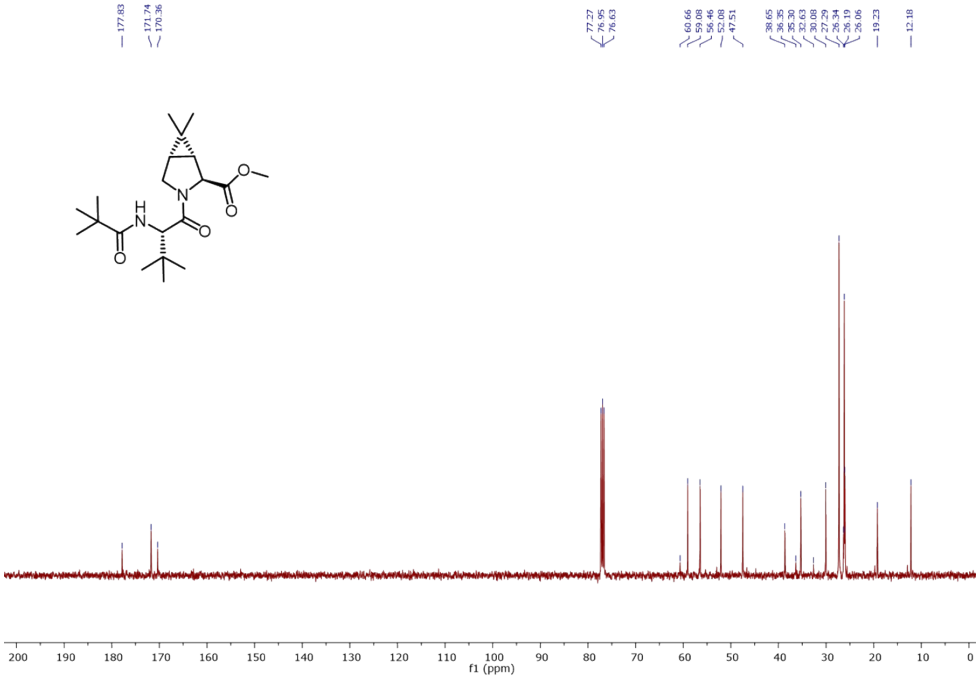


Figure S12. ^{13}C NMR (100 MHz, $CDCl_3$) of *tert*-butyl amide **8c**.

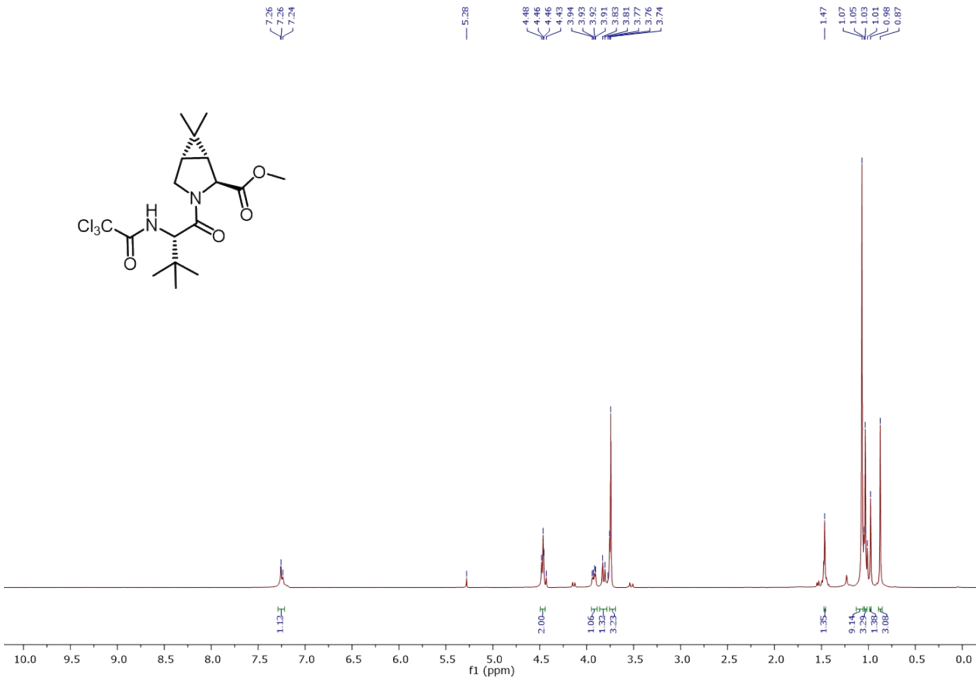


Figure S13. ^{1}H NMR (400 MHz, $CDCl_3$) of trichloroacetamide **8d**.

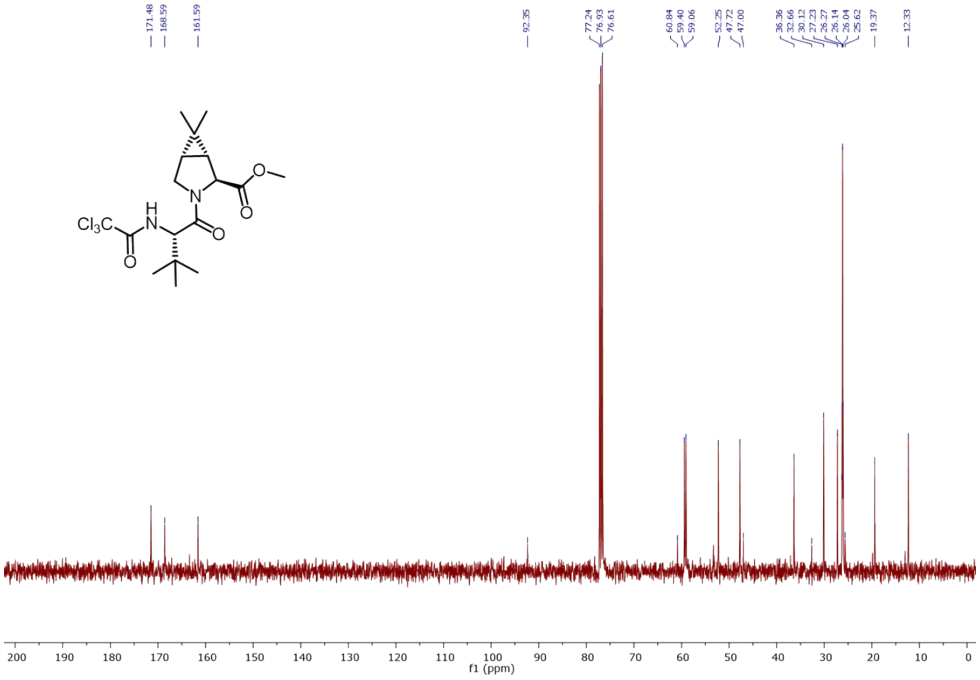


Figure S14. ^{13}C NMR (100 MHz, $CDCl_3$) of trichloroacetamide **8d**.

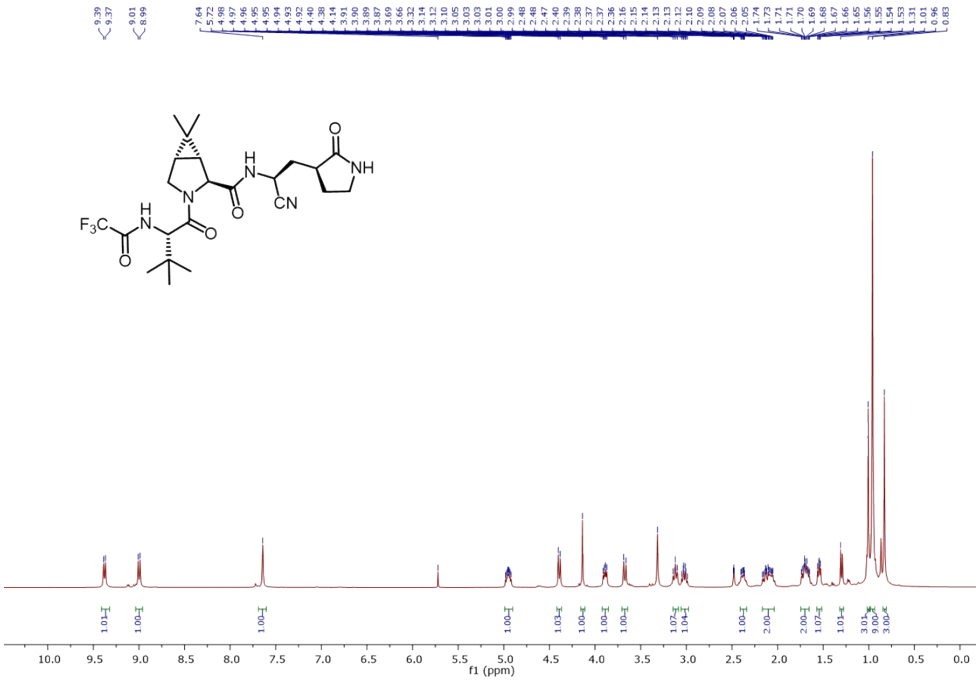


Figure S15. ^{1}H NMR (400 MHz, DMSO-d_6) of nirmatrelvir **1**.

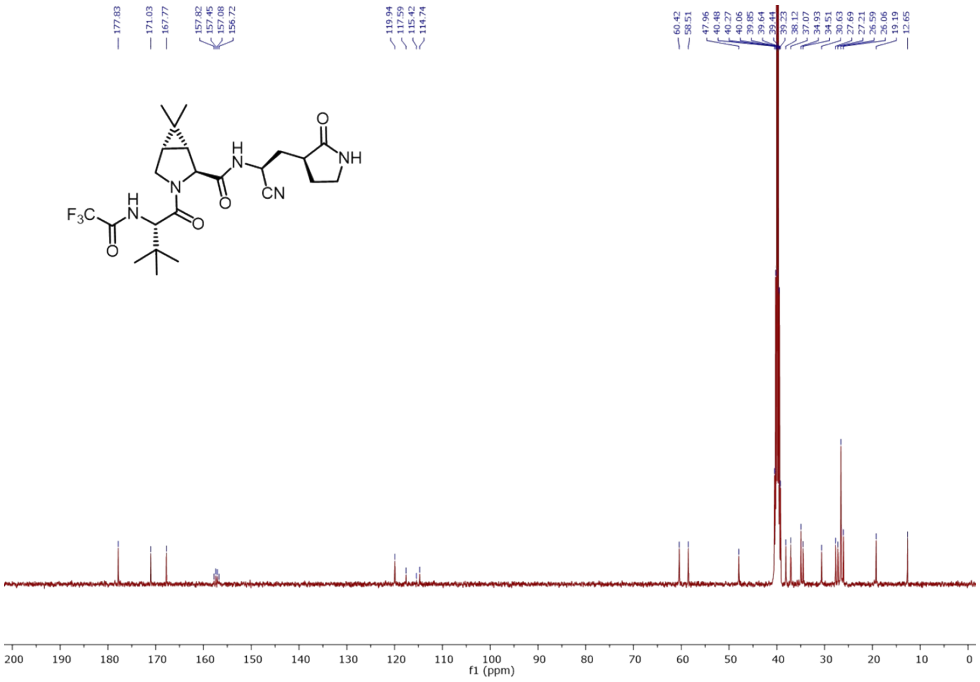


Figure S16. ^{13}C NMR (100 MHz, DMSO-d_6) of nirmatrelvir **1**.

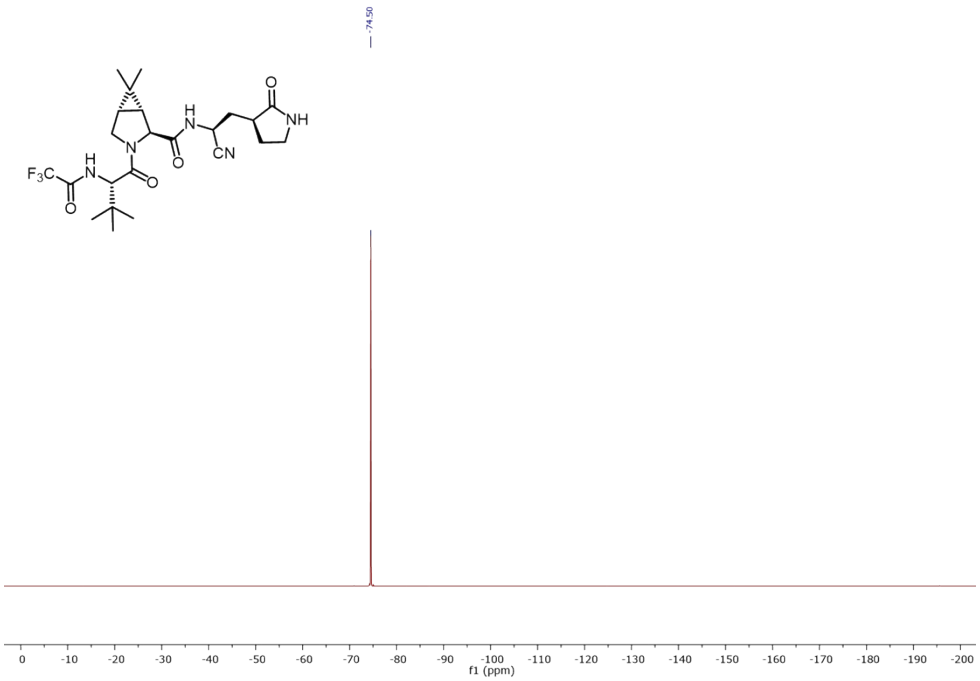
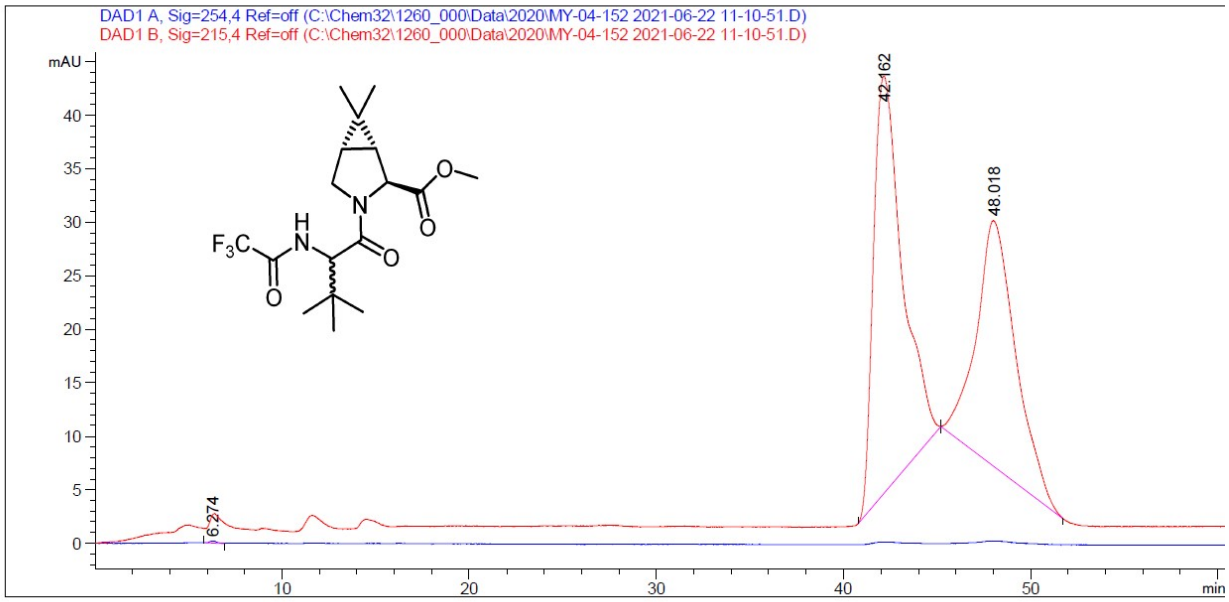


Figure S17. (376 MHz, DMSO-d_6) of nirmatrelvir **1.**

Reverse Phase HPLC analysis of epimer 8

```
Sample Info     : YMC PAK ODS-A
                  60% MeOH
                  5 uL inj.
                  0.5  mL/min
                  254 nm
                  215 nm
                  1 mg/mL
```



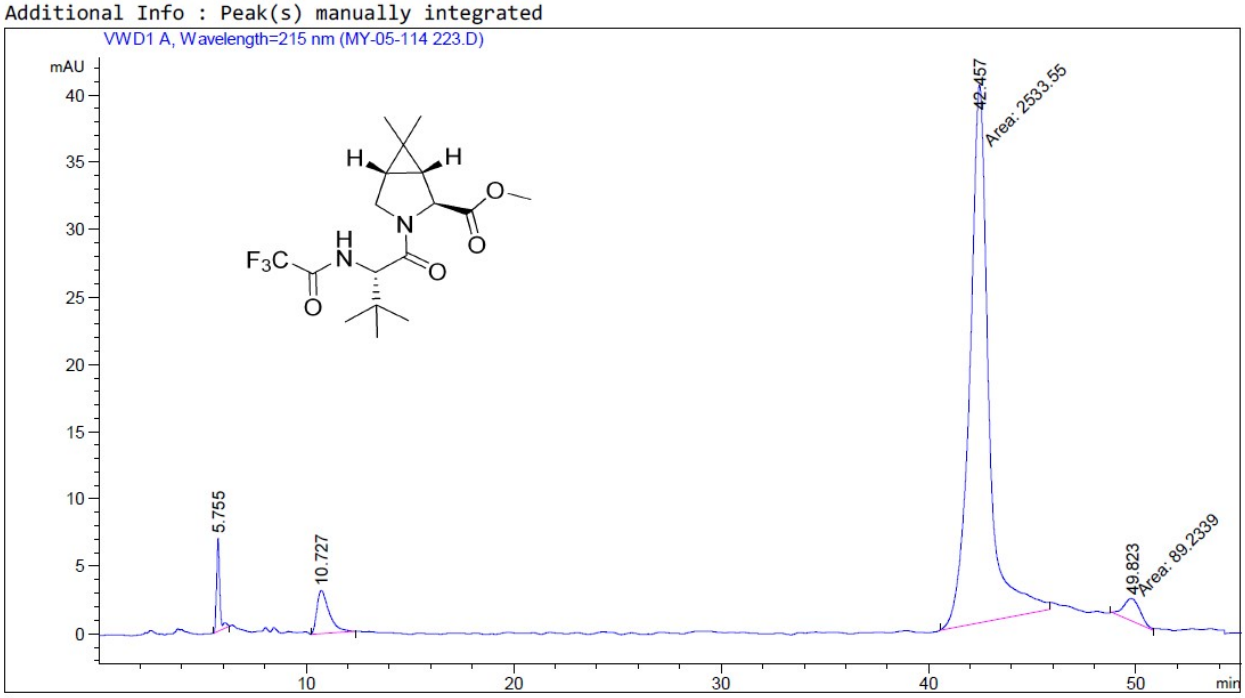

```
Signal 2: DAD1 B, Sig=215,4 Ref=off

Peak RetTime Type  Width     Area      Height     Area
  #   [min]       [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1  42.162 BB    1.5991 4164.03906   38.95371  55.1795
   2  48.018 BB    2.1042 3382.31860   22.93942  44.8205
```

Note: Sample has no chromophore and hence it's detected only in 215nM wavelength. Retention time at 6.274 min corresponds to solvent impurities.

Reverse Phase HPLC analysis of 8a

Sample Info : 60% MeOH/H2O
0.5 mL/min
215 nm
5 uL inj.
1 mg/mL
YMCPAK ODS-A

Additional Info : Peak(s) manually integrated



Signal 1: VWD1 A, Wavelength=215 nm

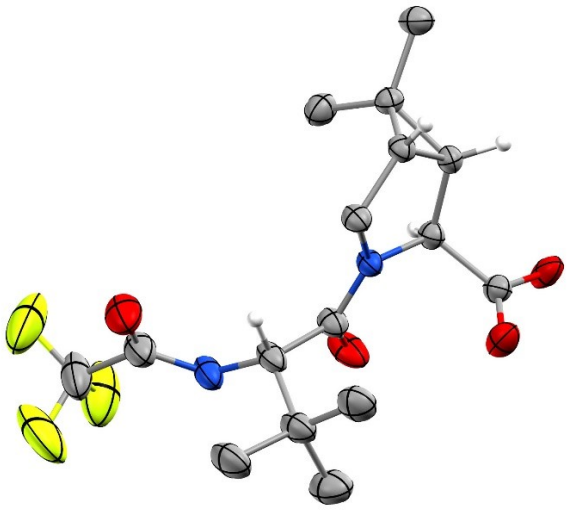
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	5.755	BV R	0.1573	73.88456	6.93667	2.6129
2	10.727	BB	0.5976	131.00914	3.21387	4.6331
3	42.457	MM	1.0583	2533.55444	39.89836	89.5983
4	49.823	MM	0.9024	89.23389	1.64804	3.1557

Note: Sample has no chromophore and hence it's detected only in 215nM wavelength. Retention times at 5.7 min and 10.72 correspond to solvent impurities.

X-Ray Diffraction Data for acid 2

Single crystals of compound **2** were formed by slow evaporation of the solution of the compound **2** in 1 mL THF and 4 mL hexanes over a week. Examination and data collection were performed with Cu Kα radiation (λ =1.54178 Å) at 150 K. Data were collected, reflections were indexed and processed, and the files scaled and corrected for absorption using APEX3[1]. The space groups were assigned, and the structures were solved by direct methods using XPREP within the SHELXTL suite of programs[2,3] and refined by full matrix least squares against F2 with all reflections using Shelxl2018[4] using the graphical interface Shelxle. If not specified otherwise H atoms attached to carbon were positioned geometrically and constrained to ride on their parent atoms. C-H bond distances were constrained to 0.95 Å for aromatic and alkene C-H and CH2 and alkyne C-H moieties, and to 1.00, 0.99 and 0.98 Å for aliphatic C-H, CH2 and CH3 moieties, respectively. Methyl H atoms were allowed to rotate but not to tip to best fit the experimental electron density. U_{iso}(H) values were set to a multiple of U_{eq}(C) with 1.5 for CH3, and 1.2 for C-H and CH2 units, respectively. Additional data collection and refinement details can be found in the Supporting Information. Complete crystallographic data, in CIF format, have been deposited with the Cambridge Crystallographic Data Centre. **CCDC 2181895** contains the supplementary crystallographic data for compound **2**. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif. X-ray diffraction data were collected using an instrument funded by the NSF (CHE-1625543).

Table 1. Crystal data and structure refinement for carboxylic acid **2**.



Crystal data	
Chemical formula	C16H23F3N2O4
Mr	364.36
Crystal system, space group	Trigonal, R3:H
Temperature (K)	150
a, c (Å)	23.5990 (6), 10.4104 (4)
V (Å3)	5020.9 (3)
Z	9
F(000)	1728
Radiation type	Cu Kα
No. of reflections for cell measurement	9772
θ range (°) for cell measurement	3.8–72.5
μ (mm-1)	0.81
Crystal shape	Needle
Colour	Colorless
Crystal size (mm)	
Data collection	
Diffractometer	Bruker AXS D8 Quest
diffractometer with PhotonIII_C14 charge-integrating and photon counting pixel array detector	
Radiation source	I-mu-S microsource X-ray tube
Monochromator	Laterally graded multilayer (Goebel) mirror
Scan method	ω and phi scans
Absorption correction	Multi-scan
SADABS 2016/2: Krause, L., Herbst-Irmer, R., Sheldrick G.M. & Stalke D. (2015). J. Appl. Cryst. 48, 3-10.	
Tmin, Tmax	0.635, 0.754

No. of measured, independent and	
observed [I > 2σ (I)] reflections	
Rint	0.046
θ values (°)	θ_{max} = 72.5, θ_{min} = 4.8
(sin θ/λ)$_{max}$ (Å-1)	0.619
Range of h, k, l	h = -29→29, k = -29→29, l = -9→12
Refinement	
Refinement on	F2
R[$F^2 > 2\ \sigma(F^2)$], *wR*(F^2), *S*	0.026, 0.070, 1.06
No. of reflections	3995
No. of parameters	272
No. of restraints	160
H-atom treatment	H atoms treated by a mixture of independent and constrained refinement
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.0429P)^2 + 0.6317P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$	0.009
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.10, -0.10
Extinction method	*SHELXL2018*/3 (Sheldrick 2018), $Fc^*=kFc[1+0.001xFc^2\lambda^3/\sin(2\theta)]^{-1/4}$
Extinction coefficient	0.00030 (6)
Absolute structure	Flack x determined using 1676 quotients [(I+)-(I-)]/[(I+)+(I-)] (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259).
Absolute structure parameter	-0.09 (6)

References

1. Bruker (2016). Apex3 v2016.9-0, Saint V8.34A, SAINT V8.37A, Bruker AXS Inc.: Madison (WI), USA, 2013/2014.
2. SHELXTL suite of programs, Version 6.14, 2000-2003, Bruker Advanced X-ray Solutions, Bruker AXS Inc., Madison, Wisconsin: USA.
3. Sheldrick, G. *Acta Crystallogr A* **64** (1), 112 (2008).
4. Sheldrick, G. *Acta. Crystallogr. Sect. C. Struct. Chem*. **71** (1), 3 (2015).