Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Pentafluorophosphato-Phenylalanines: Amphiphilic Phosphotyrosine Mimetics Displaying Fluorine-Specific Protein Interactions

*M. Accorsi, M. Tiemann, L. Wehrhan, L. M. Finn, R. Cruz, M. Rautenberg, F. Emmerling, J. Heberle, B. G. Keller, J. Rademann**

Table of Contents

General methods

All moisture sensitive reactions were performed in glassware that was previously vacuum heat dried and flushed with Ar or N_2 using Schlenk technology.

Cadmium powder was activated with HCl (1 N) until a metallic shine was observed, washed with H_2O and acetone, dried at high vacuum and stored under inert atmosphere.

NMe_4F was purchased as the tetrahydrate from Sigma Aldrich and dried as described in the literature[1]. For this, the reagent was dissolved in dry MeOH. The solution was concentrated to a syrup at the rotary evaporator, re-dissolved 4 times in dry MeOH and concentrated again. Subsequently, the residue was heated at 130 °C for 3 d at high vacuum. The obtained white powder was stored and handled under inert argon atmosphere.

All other chemicals were purchased from Sigma (Merck), ABCR and Fluka and were used without any further purification.

Dry DMF and ACN were bought as anhydrous and stored over activated molecular sieves 4Å. All other dry solvents were obtained from a column-based solvent system (MBraun, MB-SPS-800).

Removal of volatile components was performed using rotary evaporators from Heidolph with a hot water bath of 40 °C, if not otherwise stated. The high vacuum obtained with an oil pump corresponds to 1 µbar or less. Lyophilized fractions were obtained from Christ Alpha 2-4 LD.

Product isolation was conducted on Biotage, Isolera™ Spektra equipped with KP-Sil or RP-C18 SNAP Cartridges with appropriate HPLC grade solvent mixtures and deionized water, or with HPLC (Agilent Technologies, 1260 series, column Macherey-Nagel, Nucleodur 5 µm C18, 150 x 32 mm, equipped with Agilent 1260 Infinite diode array and multiple wavelength detector and fraction collector).

Melting points were measured with Büchi Melting point apparatus B-545.

Thin layer chromatography analyses were conducted on Merck Aluminum sheets pre-coated with silica gel (Merck, 60 F_{254}). Detection was carried out using 254 nm UV-Light, followed by dipping in ceric ammonium molybdate or ninhydrin stains.

The optical rotation was determined using IBZ Messtechnik Polar LµP (quartz cuvette, optic path 1 cm).

NMR spectra were measured on the following spectrometers: JEOL ECX400 (9.39 T), JEOL ECP500 (11.74 T), JEOL ECZ600 (14.09 T), Bruker Avance III 700 (16.44 T). Chemical shifts (δ) are reported in ppm and coupling constants (J) are given in Hz. ^{1}H and ^{13}C chemical shifts were referenced to the solvent peaks. ^{13}C and ^{31}P NMR spectra were hydrogen decoupled. Chemical shifts are given in ppm relative to the signal of the used deuterated solvent as internal standard.

HPLC chromatograms were recorded with an analytical HPLC system (Agilent Technologies, 1100 Series) equipped with a Luna column (column A), 3 µm C18 100 Å, 4.6 x 100 mm coupled with an ESI single quadrupole mass spectrometer LCMSD (Model# G1956B, Serial# US 44500857) from Agilent and a DAD detector using a gradient of water (A) and 99% ACN/water (B) both with 0.1% formic acid.

Alternatively, for shorter retention times, chromatograms were recorded with an HPLC system (Agilent Technologies, Infinity 1260), using Zorbax Eclipse plus C-18 RRHD column (column B) (2.1 x 50 mm, 1.8 µM, 95 Å) column, coupled with a DAD or mass detector (Agilent Technologies).

ESI high resolution mass spectra were recorded with an equipped with an analytical HPLC system (Agilent Technologies, Infinity II 1290), Zorbax Eclipse plus C-18 RRHD (2.1 x 50 mm, 1.8 µM, 95 Å) column, coupled with an ESI-Q-TOF iFunnel mass spectrometer (Agilent Technologies, 6550).

X-ray crystallographic analysis was performed on single crystals. Single crystals X-ray diffraction was performed on a Bruker D8 Venture system with graphite-monochromatic Mo-Kα radiation (λ = 0.71073 Å). Data reduction and structure solution were conducted as described in the section crystal structure determination below.

Dynamic light scattering experiments were performed on a Nicomp Nano DLS/ZLS system, at 25 °C and a wavelength of 660 nm. The respective viscosity and refraction index for each solvent mixture were taken from the literature.

General method of peptide synthesis

Peptide synthesis was conducted using Fmoc-strategy on Rink amide resin **13** from Merck (loading 0.34 mmol/g, 100-200 mesh, 1% divinyl-benzene/polystyrene). PP-PE syringes equipped with a PE-frit were used as reaction vessels.

Coupling of amino acids
N-Fmoc-protected amino acids with suitable side chain protection (5 equiv. with respect to the loading of the resin) were pre-activated with TBTU (4.9 equiv.) and DIPEA (10 equiv.) in a minimal volume of DMF. The solution was added to the resin pre-swollen in DMF and shaken for 2 h.

The coupling reactions were monitored using the Kaiser test.[2] Coupling effectiveness was quantified via UV-photometric determination of the dibenzofulvene product at 301 nm following to Fmoc cleavage with the following equation:

$$x = \frac{10^5 \cdot E_\lambda}{\varepsilon_\lambda \cdot m_{resin}}$$

x = Loading resin [mmol/g]
E_λ = Extinction
$\varepsilon_{\lambda 301}$ = 7800 l/mol cm
ε_λ = molar extinction coefficient
m_{resin} = mass resin [mg]

The resin was carefully vacuum dried prior to the Fmoc determination.

Washing of the resin
The resin was washed after every coupling and cleavage procedure with 5 syringe volumes of DMF.

Fmoc cleavage
The Fmoc-group was cleaved by using a mixture of piperidine (20%) in DMF. After adding the basic cocktail to the resin, the syringe was shaken for 10 min and then washed with DMF. The cleavage procedure was repeated once.

End capping of peptide
Before capping, the resin was swollen in DMF (2 ml / 100 mg of resin). The peptide N-Terminus was capped with an acetic anhydride/pyridine mixture (1:1, 1 ml/200 mg of resin) for 1 h at RT.

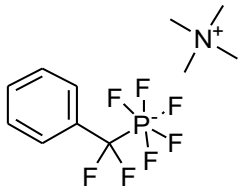
Peptide cleavage
The vacuum-dried resin was treated with Olah´s reagent (pyridinium poly-(hydrogen fluoride), ca. 70% HF and 30% pyridine) + 10% anisole for 90 min at RT. The cleavage mixture was slowly dropped into a saturated solution of $NaHCO_3$. The beads were washed with small portions of THF, followed by THF/ H_2O (1:1) and the washings dropped in the $NaHCO_3$ solution as well. The mixture was then concentrated to a minimum at the rotary evaporator. The residue was purified using an MPLC with a C18 column and a gradient of eluent A (10 mM NH_4HCO_3 in H_2O, pH 7.5) and eluent B (ACN). Collected fractions were analyzed with LCMS and lyophilized.

Scope of cleavage conditions

The cleavage conditions mentioned above (Olahs reagent with 10% anisol for 1h) were shown to successfully deprotect tert-butyl protection groups of commercially available amino acids such as glutamate and aspartate and the trityl protection group of cysteine.

Chemical synthesis

Tetramethylammonium 1-(pentafluorophosphato-difluoromethyl) benzene (2)



Diethyl difluoro-(phenyl)-methyl phosphonate (200 mg, 0.757 mmol, 1 equiv.) was dissolved in a Schlenk flask in dry ACN (5 mL). TMSBr (287 µL, 1.67 mmol 2.2 equiv.) was added dropwise under inert atmosphere. The solution was heated at 60°C for 1 hour. After disappearance of the starting material monitored via LC-MS, the vial was equipped with inert gas inlet and outlet to allow the release of the gaseous components developed after addition of dry DMF (293 µL, 3.78 mmol, 5 equiv.) and $(COCl)_2$ (636 µL, 7.57 mmol, 10 equiv.). After gas development decreased, the reaction was heated in a sealed vessel under inert atmosphere at 40°C with a water bath. After 1.5 hours, the reaction was cooled to 0°C with an ice bath. Previously weighted under inert atmosphere and dried NMe_4F (705 mg, 7.57 mmol, 10 equiv.) was then added slowly under inert atmosphere to the stirred and cooled reaction mixture. After 30 min, the mixture was slowly quenched in a cooled sat. aq. sol. of $NaHCO_3$ (25 mL/mmol starting material) and extracted with DCM (3 x 30 mL). The collected organic layer was then concentrated at the rotary evaporator, redissolved in H_2O/ACN and purified with RP-MPLC (RP c18, ACN / H_2O, 5 to 99%). After purification the product (119 mg, 62%) was isolated as white fluffy solid.

Rf= 0.5 (EtOAc)

Melting point = 107 °C

^{1}H NMR (600 MHz, Acetonitrile- d_3) δ = 7.42 (d, J=7.6, 2H, Ar-H), 7.35 – 7.27 (m, 3H, Ar-H), 3.02 (s, 12H, NMe_4)

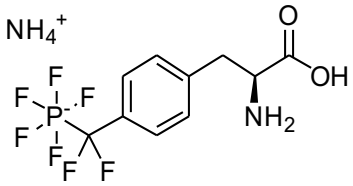
^{13}C NMR (151 MHz, Acetonitrile- d_3) δ = 127.96, 127.32 (Ar-C), 125.53 (t, J=7.8, CF_2), 55.71 – 54.20 (s, NMe4)

^{19}F NMR (565 MHz, Acetonitrile- d_3) δ = -70.07 (dp, J=696.0, 43.5, 1F, F_{ax}), -71.76 (dt, J=855.7, 42.9, 9.1, 2F, F_{eq}), -98.59 (dt, J=119.9.8, 9.5, 2F, F_{eq})

^{31}P NMR (243 MHz, Acetonitrile- d_3) δ = -145.27 (dtquin, J=864.3, 696.3, 125.2)

HRMS (ESI): $[M]^-$ calculated for $C_7H_5F_7P^-$: 253.0023 Da, found: 253.0033 m/z

Ammonium 4-(pentafluorophosphato-difluoromethyl)-L-phenylalanine (3)



The sodium salt of Fmoc-protected amino acid **9** (100 mg, 0.165 mmol, 1 equiv.) was dissolved in ACN (1.8 ml). Piperidine (200 µl, 2.00 mmol, 12 equiv.) was added and the resulting mixture stirred for 8 h at RT. All volatile components were removed under reduced pressure and the obtained crude product purified via MPLC using a C18 reversed phase column and a gradient of eluent A (10 mM NH_4HCO_3 in H_2O, pH 7.5) and eluent B (ACN). Fractions containing the product were concentrated at a rotary evaporator and lyophilized, yielding the product as a pale yellow solid (58 mg, 0.160 mmol, 97%).

^{1}H NMR (700 MHz, D_2O) δ = 7.47 (d, J=7.6, 2H, Ar-H), 7.32 (d, J=8.0, 2H, Ar-H), 3.95 (dd, J=8.5, 4.9, 1H, CHN), 3.30 (dd, J=14.5, 4.8, 1H, $CH_{2\alpha}$-Phe), 3.12lll – 3.04 (m, 1H, $CH_{2\beta}$-Phe)

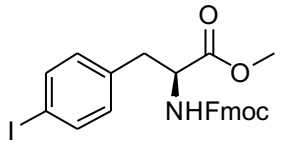
^{13}C NMR (176 MHz, D_2O) δ = 173.7 (O=C), 138.1, 136.3, 129.1, 129.0, 125.8 (6 x Ar-C), 44.5 (CHN)z, 36.1 (CH_2Phe)

^{19}F NMR (376 MHz, D_2O) δ = -68.50 (dp, J=693.6, 43.6, 1F, F_{ax}), -72.14 (ddt, J=864.2, 43.0, 8.5, 4F, F_{eq}), -98.27 (dt, J=126.2, 8.3, 2F, CF_2)

^{31}P NMR (162 MHz, D_2O) δ = -143.00 (pdt, J=864.8, 693.45, 126.3)

HRMS (ESI): $[M]^-$ calculated for $C_{10}H_{10}F_7NO_2P^-$: 340.03429 Da, found: 340.03452 m/z

Methyl N-(fluorenyl-9H-methoxy-carbonyl)-4-iodo-L-phenylalanine (4)



Fmoc-4-I-Phe-OH (from ABCR, 2988 mg, 10.26 mmol, 1 equiv.) was dissolved in dry MeOH (25 ml) in a heat- and vacuum-dried Schlenk flask under Ar atmosphere. 3 drops of dry DMF were added and the solution was cooled to 0 °C. Oxalyl chloride (2.6 ml, 30.795 mmol, 3 equiv.) was added dropwise under stirring and the reaction was allowed to reach RT for 16 h. The amber solution was then concentrated using a rotary evaporator, diluted with EtOAc and washed with H_2O, saturated $NaHCO_3$ solution, and brine. The organic layer was dried over Na_2SO_4, filtrated and concentrated in vacuo. Product **4** (4717 mg, 92%) was obtained as white solid.

R_f = 0.3 (EtOAc/Hex, 20%)

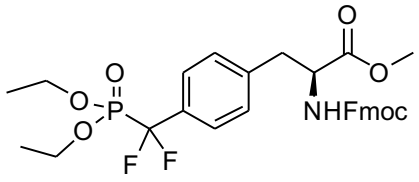
^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.77 (d, J=7.6, 2H, Ar-H), 7.60 (d, J=8.1, 2H, Ar-H), 7.61 – 7.51 (m, 2H, Ar-H), 7.41 (t, J=7.4, 2H, Ar-H), 7.32 (t, J=7.5, 2H, Ar-H), 6.81 (d, J=8.1, 2H, Ar-H), 5.24 (d, J=7.9, 1H, NH), 4.64 (q, J=5.9, 1H, α-H-Phe), 4.47 (dd, J=10.5, 7.2, 1H, $CH_{2\alpha}$-Fmoc), 4.37 (dd, J=10.4, 6.9, 1H, $CH_{2\beta}$-Fmoc), 4.20 (t, J=6.7, 1H, CH-Fmoc), 3.73 (s, 3H, OMe), 3.09 (dd, J=13.9, 5.6, 1H, $CH_{2\alpha}$-Phe), 3.02 (dd, J=13.9, 5.7, 1H, $CH_{2\beta}$-Phe)

^{13}C NMR (126 MHz, $CDCl_3$) δ = 171.7 (O=C-methyl ester), 155.6 (O=C-Fmoc), 143.9, 143.8, 141.5, 141.4, 137.8, 135.5, 131.4, 127.9, 127.2, 120.1 (17 x Ar-C), 92.8 (C-I), 67.0 (CH_2-Fmoc), 54.6 (CHN), 52.6 (OCH_3), 47.3 (CH-Fmoc), 37.8 (CH_2Phe)

HRMS (ESI): $[M+H]^+$ calculated for $C_{25}H_{23}INO_4^+$: 528.06663 Da; found: 528.06571 m/z; $[M+Na]^+$ calculated for $C_{25}H_{22}INNaO_4^+$: 550.04857 Da; found: 550.04790 m/z

Spectral data were consistent with published values.[3]

Methyl N-(fluorenyl-9H-methoxy-carbonyl)-4-(diethoxyphosphoryl-difluoromethyl)-L-phenylalanine (5)



A heat- and vacuum-dried Schlenk flask was charged with cadmium powder (2.578 g, 22.94 mmol, 6 equiv.) activated and dried as described above and dry DMF (3 ml). To the stirred suspension, diethyl bromo-difluoromethyl-phosphonate (2.261 ml, 12.73 mmol, 3.33 equiv.) was added dropwise to the reaction flask at room temperature. The slightly exothermic reaction was stirred for 3 h. In another flask, previously dried Fmoc-(4-I)Phe-OMe **4** (2.016 g, 3.82 mmol, 1

equiv.) was dissolved in dry DMF (1 ml) and CuBr (1.645 g, 11.47 mmol, 3 equiv.) was added. The solution containing the organocadmium reagent was added slowly and dropwise to this stirred mixture under Ar atmosphere and the reaction mixture was stirred for 16 h and monitored via TLC (1:4, EtOAc / Hex). After addition of EtOAc, the precipitate was filtrated off over a bed of Celite and the filtrate washed with a saturated aqueous solution of NH_4Cl (20 ml, 3x), H_2O (20 ml) and brine (20 ml). The organic layer was dried over Na_2SO_4, filtrated and concentrated under reduced pressure. After purification of the crude via column chromatography at MPLC (SiO_2, EtOAc / Hex 1:4 then 1:2), the product (2.256 g, 99%) was isolated as a colorless oil.

R_f = 0.3 (EtOAc/Hex, 50%)

^{1}H NMR (600 MHz, $CDCl_3$) δ = 7.75 (d, J=7.5, 2H, 2 x Ar-H), 7.54 (dd, J=14.7, 7.7, 4H, 2 x Ar-H), 7.38 (t, J=7.4, 2H, 2 x Ar-H), 7.30 (t, J=7.5, 2H, 2 x Ar-H), 7.17 (d, J=7.9, 2H, 2 x Ar-H), 5.31 (d, J=8.2, 1H,NH), 4.66 (q, J=6.0, 1H, α-H-Phe), 4.43 (dd, J=10.6, 7.2, 1H, CH_2-Fmoc), 4.36 (dd, J=10.8, 7.0, 1H, CH_2-Fmoc), 4.23 – 4.05 (m, 5H, 2xCH_2-ethyl, CH-Fmoc), 3.70 (s, 3H, OMe), 3.17 (dd, J=13.9, 5.8, 1H, CH_2-Phe), 3.11 (dd, J=13.9, 6.1, 1H, CH_2-Phe), 1.28 (t, J=7.1, 6H, CH_3-ethyl)

^{13}C NMR (151 MHz, $CDCl_3$) δ = 171.67 (O=C-methyl ester),155.61 (O=C-Fmoc), 143.87, 143.75, 141.41, 138.99, 131.54 (td, J=22.2, 14.0), 129.50, 127.84, 127.16, 126.58 (t, J=7.6), 125.7, 125.14, 120.08 (d, J=3.2, 18 x Ar-C), 117.21 (dd, J=262.5, 218.4, CF_2), 67.02 (CH_2-Fmoc), 64.88 (dd, J=6.8, 1.3, 2 x CH_2 Ethyl), 54.72 (CHN), 52.52 (OCH_3), 47.23 (CH-Fmoc), 38.05 (CH_2Phe), 16.40 (d, J=5.5, 2 x CH_3 Ethyl)

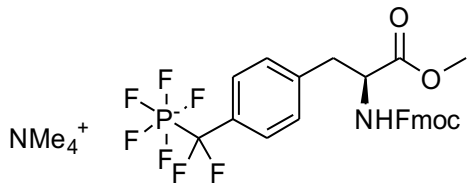
^{19}F NMR (376 MHz, $CDCl_3$) δ = -108.18 (d, J=116.0)

^{31}P NMR (162 MHz, $CDCl_3$) δ = 6.94 (t, J=116.1)

HRMS (ESI): $[M + H]^+$ calculated for $C_{30}H_{33}F_2NO_7P^+$: 588.1957 Da, found: 588.1932 m/z;
$[M+Na]^+$ calculated for $C_{30}H_{32}F_2NNaO_7P^+$: 610.1777 Da, found: 610.1765 m/z;
$[M+K]^+$ calculated for $C_{30}H_{32}F_2KNO_7P^+$: 626.1516, found: 626.1493 m/z.

Spectral data were consistent with published values.[4]

Tetramethylammonium methyl N-(fluorenyl-9H-methoxy-carbonyl)-4-(pentafluorophosphato-difluoromethyl)-L-phenylalanine (8)



Diethyl phosphonic acid ester **5** (828 mg, 1.41 mmol, 1 equiv.) was dissolved in a Schlenk flask in dry ACN (5 ml). TMSBr (930 µL, 7.05 mmol, 5 equiv.) was added dropwise under inert

atmosphere. The solution was heated at 60 °C for 1.5 h. After disappearance of the starting material monitored via LC-MS, the vial was equipped with inert gas inlet and outlet to allow the release of the gaseous components developed after dropwise addition of dry DMF (545 µL, 7.05 mmol, 5 equiv.) followed by $(COCl)_2$ (1.18 ml, 14.09 mmol, 10 equiv.). After gas development ceased, the reaction was heated in a sealed vessel under inert atmosphere at 40 °C with a water bath. After 1.5 h, a small aliquot was taken and MeOH was added. The formation of the dimethyl ester was confirmed via LCMS. The reaction mixture was then cooled to 0 °C with an ice bath. Previously weighted under inert atmosphere and dried as described above, NMe_4F (1050 mg, 14.09 mmol, 10 equiv.) was then added slowly under inert atmosphere to the stirred and cooled reaction mixture. After 1 h, the mixture was quenched by slowly pouring it into an ice cooled saturated aqueous solution of $NaHCO_3$ (30 ml) and extracted with DCM (3 x 30 ml). The collected organic layers were then concentrated at the rotary evaporator, redissolved in H_2O/ACN and purified with RP-MPLC. After purification of the crude via column chromatography at MPLC (RP C18, ACN / H_2O, 5 to 99%), the fractions were analyzed at LCMS and those containing the product were concentrated at the rotary evaporator and lyophilized, yielding 619 mg (68%) of the title compound as white lyophilisate.

Melting point = 140-142 °C

$[\alpha]_D^{20}$ = - 6.9 (c = 1, MeOH)

^{1}H NMR (600 MHz, ACN-d3) δ = 7.80 (d, J=7.6, 2H, Ar-H), 7.65 – 7.53 (m, 2H, Ar-H), 7.39 (t, J=7.5, 2H, Ar-H), 7.38 – 7.23 (m, 4H), 7.15 (d, J=7.8, 2H), 4.39 (q, J=8.2, 1H, CHN), 4.32 – 4.20 (m, 2H, Fmoc CH_2), 4.18 (t, J=6.9, 1H, Fmoc CH), 3.64 (s, 3H, OMe), 3.15 – 3.09 (m, 1H, $CH_{2\alpha}$Phe), 3.02 (s, 12H, NMe_4), 2.97 – 2.88 (m, 1H, $CH_{2\beta}$Phe)

^{13}C NMR (151 MHz, ACN-d3) δ = 172.22 (O=C-methyl ester), 155.93 (O=C-Fmoc), 144.13, 141.19, 136.76, 129.45, 128.24, 127.78, 127.21 (d, J=3.2), 125.71 (t, J=7.0, CF_2), 125.27 (d, J=10.3), 120.05 (18x Ar-C), 66.39 (CH_2-Fmoc) , 55.46 (CHN), 55.24 (NMe_4), 51.89 (OMe), 47.03 (CH-Fmoc), 36.87 (CH_2Phe)

^{19}F NMR (565 MHz, ACN-d3) δ = -69.65 (p, J=42.9, F_{ax}), -70.88 (p, J=42.9, F_{ax}), -71.05 (t, J=8.8, F_{eq}), -71.12 (t, J=8.7, F_{eq}) -98.26 (dt, J=120.3, 9.1, CF_2)

^{31}P NMR (243 MHz, ACN-d3) δ = -143.43 (pdt, J=855.99, 696.35, 120.1)

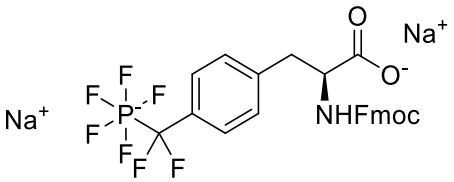
HRMS (ESI): $[M]^-$ calculated for $C_{26}H_{22}F_7NO_4P^-$: 576.11802 Da, found: 576.11809 m/z

Supplementary Table 1: Results of stability tests

Compound	Conditions	PF_5^- integrity (analyzed via NMR)
Phenyl$CF_2PF_5^-$ **2**	HFIP neat, 1 h	50% degradation (to monofluorophosphate)
	HFIP / DCM 1:4, 1 h	stable
Fmoc-(4-$PF_5^-CF_2$)Phe-OH **9**	TFA 100%, 3 h	decomposition
	TFA 95% (aq.), 2 h	decomposition
	0.1 M HCl in HFIP, 2 h	decomposition
	AcOH in DCM (1:9), 1.5 h	decomposition
	5% TFA in DCM, 1 h	decomposition
	Sonication in ACN	stable
	0.1 M HCl (aq.), 1 h	stable
	0.1 M HCl (aq.), 24 h	15% decomposition
	0.01 M HCl (aq.), 1 h	stable
	0.01 M HCl (aq.), 24 h	stable
	0.1 M TFA (aq.), 1 h	10% decomposition
	0.01 M TFA (aq.), 1 h	stable
	0.1% TFA (aq.), 2 h	stable
	0.2% TFA (aq.), 2 h	35% decomposition
	TMSBr (100 equiv.) in ACN, 1 h	decomposition
	SiO_2 (30 µL/mg), in ACN, 1 h	stable
	piperidine (20%) in DMF, 24 h	stable
	pyridine neat, 1 h	stable
	DBU (2%) in DMF, 30 min	stable (analyzed via LCMS)

The compound (3 mg) was added to a vial containing the reagent to be tested and transferred into an NMR tube. After the specified time, a ^{19}F-NMR spectrum was recorded, or in case of the LCMS study, an aliquot was taken and analyzed.

Sodium N-(fluorenyl-9H-methoxy-carbonyl)-4-(pentafluorophosphato-difluoromethyl)-L-phenylalanine (9)



To the methyl ester **8** (60 mg, 0.098 mmol, 1 equiv.) in 50 ml of an aqueous solution of ammonium bicarbonate (50 mM, pH 7.8) were added two spatula tips of *Bacillus licheniformis* protease (from Sigma Aldrich) and 5 ml ACN. The resulting mixture was stirred at RT overnight. All volatile components were removed under reduced pressure and the obtained crude product was purified

via MPLC using a C18 reversed phase column and a gradient of eluent A (10 mM NH_4HCO_3 in H_2O, pH 7.5) and eluent B (ACN). Fractions containing the product were concentrated at a rotary evaporator and lyophilized, yielding the product as an off white solid. Subsequently the pure compound was dissolved in a 1:1 mixture of H_2O/ACN and ion exchanged over Na-loaded Amberlite (IR 120). The exchange was performed in a glass column with a length of 21 cm, an inner diameter of 7 mm and a flowrate of 30 µl/s yielding **9** as a white solid (57 mg, 0.094 mmol, 96%).

$[\alpha]_D^{20}$ = + 12.1 (c = 1, MeOH)

^{1}H NMR (600 MHz, DMSO-D_6) δ = 7.88 (d, *J* = 7.5 Hz, 2H, Ar-H), 7.66 (dd, *J* = 15.9, 7.5 Hz, 2H, Ar-H), 7.44 – 7.37 (m, 2H, Ar-H), 7.37 – 7.27 (m, 2H, Ar-H), 7.21 (d, *J* = 7.9 Hz, 2H, Ar-H), 7.14 (d, *J* = 7.9 Hz, 2H, Ar-H), 6.28 (s, 1H, NH), 4.30 – 4.13 (m, 3H, Fmoc-CH_2, Fmoc-CH), 4.03 – 3.96 (m, 1H, αH), 3.10 – 2.87 (m, 2H, βH).

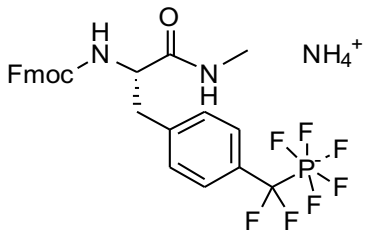
^{13}C NMR (151 MHz, DMSO-D_6) δ = 173.34 (C=O), 155.54 (C=O), 143.9, 140.67, 138.18, 128.94, 128.05, 127.58, 127.08, 125.3, 125.2, 124.9, 121.40, 120.07 (Ar- C), 65.46 (CH_2Fmoc), 56.13 (CHN), 46.62 (CH-Fmoc), 36.60 (CH_2Phe).

^{19}F NMR NMR (565 MHz, DMSO-D_6) δ = -67.57 (dp, *J* = 698.8, 45.3, 44.9 Hz, 1F, F_{ax}), -69.86 (ddt, *J* = 858.5, 44.6, 8.1 Hz, 4F, F_{eq}), -96.65 (d, *J* = 120.9 Hz, 2F, CF_2).

^{31}P NMR (243 MHz, DMSO-D_6) δ = -143.55 (pdt, *J* = 858.5, 698.0, 121.0 Hz).

HRMS (ESI): $[M]^-$ calculated for $C_{25}H_{20}F_7NO_4P^-$: 562.10237 Da, found: 562.10245 m/z

Ammonium N-(fluorenyl-9H-methoxy-carbonyl)-4-(pentafluorophosphato-difluoromethyl)-L-phenylalanyl-N-methylamide (S1)



To **9** (0.17 g, 0.28 mmol, 1 equiv.) in 10 ml dry ACN were added TBTU (0.27 g, 0.84 mmol, 3 equiv.) and diisopropylethylamine (0.24 ml, 1.68 mmol, 6 equiv.). The resulting solution was stirred at RT for 3 min. Methylamine hydrochloride (0.95 mg, 1.40 mmol, 5 equiv.) was added and the reaction mixture stirred for 1 h at RT. The volatile components were removed under reduced pressure and the crude product was purified via MPLC using a C18 reversed phase column and a gradient of eluent A (10 mM NH_4HCO_3 in H_2O, pH 7.5) and eluent B (ACN). Fractions containing the product were concentrated at a rotary evaporator and lyophilized, yielding the product **S1** as an off-white solid (0.16 g, 0.27 mmol, 95%).

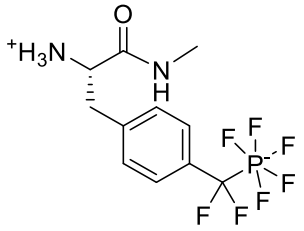
^{1}H NMR (600 MHz, ACN-d_3): δ = 7.83 (d, J = 7.6 Hz, 2H, Ar-H), 7.62 (dd, J = 14.7, 7.5 Hz, 2H, Ar-H), 7.42 (t, J = 7.5 Hz, 2H, Ar-H), 7.34 (dt, J = 15.3, 7.3 Hz, 4H, Ar-H), 7.18 (d, J = 7.8 Hz, 2H, Ar-H), 6.53 (s, 1H, MeNH), 5.91 (d, J = 8.3 Hz, 1H, FmocNH), 4.35 – 4.15 (m, 4H, αH+CH Fmoc+CH_2 Fmoc), 3.13 – 2.83 (m, 2H, βH), 2.64 (d, J = 4.7 Hz, 3H, $NHCH_3$).

^{19}F NMR (565 MHz, ACN -d_3): δ = -70.28 (dp, J = 696.3, 43.1, 42.1 Hz, 1F, F_{ax}), -71.88 (ddt, J = 856.3, 43.0, 9.1 Hz, 4F, F_{eq}), -98.27 (dp, J = 120.7, 9.4, 8.9 Hz, 2F, CF_2).

^{31}P NMR (243 MHz, ACN -d_3): δ = -143.89 (pdt, J = 856.3, 695.5, 119.8 Hz).

ESI-MS (m/z): $[M]^-$ calculated for $([C_{26}H_{23}F_7N_2O_3P])^-$: 575.1 Da; found: 575.0 m/z

4-(Pentafluorophosphato-difluoromethyl)-L-phenylalanyl-N-methylamide (S2)



(S1) (0.15 g, 0.26 mmol, 1 equiv.) was dissolved in ACN (9 ml). Piperidine (1 ml) was added and the resulting solution stirred at RT for 7 h. The volatile components were removed under reduced pressure and the crude product was purified via MPLC using a C18 reversed phase column and a gradient of eluent A (10 mM NH_4HCO_3 in H_2O, pH 7.5) and eluent B (ACN). Fractions containing the product were concentrated at a rotary evaporator and lyophilized, yielding product **S2** as a white solid (0.09 g, 0.25 mmol, 96%.).

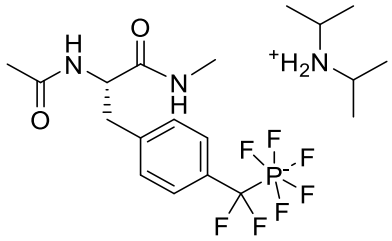
^{1}H NMR (600 MHz, D_2O): δ = 7.51 (d, J = 7.8 Hz, 2H, Ar-H), 7.31 (d, J = 7.9 Hz, 2H, Ar-H), 4.00 (t, J = 7.5 Hz, 1H, αH), 3.19 – 3.08 (m, 2H, βH), 2.63 (s, 3H, $NHCH_3$).

^{19}F NMR (565 MHz, D_2O): δ = -68.48 (dp, J = 692.4, 43.2 Hz, 1F, F_{ax}), -72.12 (ddt, J = 864.4, 43.1, 8.6 Hz, 4F, F_{eq}), -98.30 (ddt, J = 126.2, 17.6, 8.7 Hz, 2F, CF_2).

^{31}P NMR (243 MHz, D_2O): δ = -142.95 (pdt, J = 863.6, 692.9, 126.0 Hz).

ESI-MS (m/z): $[M]^-$ calculated for $([C_{11}H_{13}F_7N_2OP])^-$: 353.0 Da; found: 353.0 m/z

Diisopropylammonium N-Acetyl-4-(pentafluorophosphato-difluoromethyl)-3-(methylamino)-L-phenylalaninyl-amide (10)



To **(S2)** (0.04 g, 0.13 mmol, 1 equiv.) in 10 ml dry ACN were added diisopropylamine (0.13 ml, 0.75 mmol, 6 equiv.) and acetic anhydride (0.06 ml, 0.63 mmol, 5 equiv.). The resulting solution was stirred at RT for 4 h. The volatile components were removed under reduced pressure and the crude product was purified via MPLC using a C18 reversed phase column and a gradient of eluent A (10 mM NH_4HCO_3 in H_2O, pH 7.5) and eluent B (ACN). Fractions containing the product were concentrated at a rotary evaporator and lyophilized, yielding product **10** as a white solid (0.06 g, 0.13 mmol, quant.).

^{1}H NMR (600 MHz, D_2O): δ = 7.48 (d, J = 6.7 Hz, 2H, Ar-H), 7.30 (d, J = 8.0 Hz, 2H, Ar-H), 4.48 (t, 1H, αH), 3.51 (hept, J = 6.5 Hz, 2H, DIPA $NCHCH_3$), 3.16 – 3.00 (m, 2H, βH), 2.63 (s, 3H, $NHCH_3$), 1.95 (s, 3H, $NHCOCH_3$), 1.30 (d, J = 6.5 Hz, 12H, DIPA $NCHCH_3$).

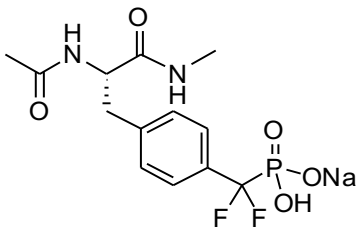
^{19}F NMR (565 MHz, D_2O): δ = -68.38 (dp, J = 692.4, 43.4 Hz, 1F, F_{ax}), -72.11 (ddt, J = 864.8, 43.1, 8.8 Hz, 4F, F_{eq}), -98.10 (dt, J = 126.4, 8.7 Hz, 2F, CF_2).

^{31}P NMR (243 MHz, D_2O): δ = -142.86 (pdt, J = 864.9, 692.4, 127.3 Hz).

^{13}C NMR (151 MHz, D_2O): δ = 174.13 ($NHCOCH_3$), 173.56 ($CONHCH_3$), 137.48 (Ar-$C_{quart.}$), 128.81 (Ar-CH), 125.40 (Ar-CH), 55.34 (αC), 47.27(DIPA $NCHCH_3$), 36.82 (βC), 25.66 ($NHCH_3$), 21.60 ($NHCOCH_3$), 18.29 (DIPA $NCHCH_3$).

ESI-HRMS (m/z): $[M]^-$ calculated for $([C_{13}H_{15}F_7N_2O_2P])^-$: 395.0759 Da; found: 395.0759 m/z

Sodium N-Acetyl-4-(phosphato-difluoromethyl)-3-(methylamino)-L-phenylalaninyl-amide (10a)



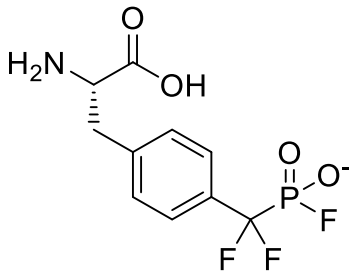
10 (0.01 g, 0.02 mmol, 1 equiv.) was stirred in 2 M HCl for 72 h. The volatile components were removed under reduced pressure and the desired product obtained as an off-white solid (0.01 g, 0.02 mmol, quant.)

^{1}H NMR (600 MHz, D_2O): δ = 7.55 (d, *J* = 5.4 Hz, 1H, Ar-H), 7.34 (d, *J* = 5.1 Hz, 2H, Ar-H), 4.49 (t, *J* = 9.3 Hz, 1H, αH), 3.26 – 2.97 (m, 2H, βH), 2.65 (s, 3H), 1.94 (s, 3H).

^{19}F NMR (565 MHz, D_2O): δ = -108.06 (d, *J* = 106.0 Hz).

^{31}P NMR (243 MHz, D_2O): δ = 4.88 (t, *J* = 105.1 Hz).

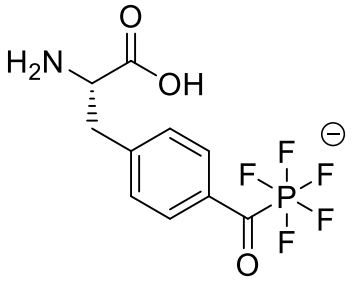
Ammonium 4-(monofluorophosphono-difluoromethyl)-L-phenylalanine (11)



The amino acid **3** (10 mg, 0.03 mmol, 1 equiv.) was stirred in a mixture of glacial acetic acid and deuterated DCM (1:9,) for 90 min. NMR analysis showed full conversion to the product **11**.

^{19}F NMR (376 MHz, DCM): δ = -70.45 (d, J = 1000.3 Hz, 1F), -106.39 (d, J = 97.9 Hz, 2F, CF_2).

Ammonium 4-(pentafluorophosphato-carbonyl)-L-phenylalanine (12)



Amino acid **3** (11 mg, 0.030 mmol, 1 equiv.) was dissolved in 200 µl Olah´s reagent (pyridinium poly-(hydrogen fluoride)) and 2 µl of water. The reaction was stirred for 6 hours and subsequently quenched with a saturated solution of $NaHCO_3$. All volatile components were removed under reduces pressure and the crude product purified via MPLC using an C18 reversed phase column and a gradient of eluent A (10 mM NH_4HCO_3 in H_2O, pH 7.5) and eluent B (ACN). Fractions containing the product were concentrated at a rotary evaporator and lyophilized. The product was obtained as a white solid (9 mg, 0.026, 87%).

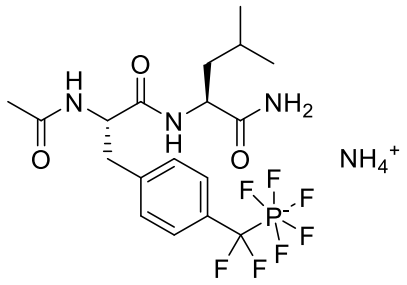
^{1}H NMR (600 MHz, D_2O): δ = 7.97 (d, J = 7.8 Hz, 2H, Ar-H), 7.38 (d, J = 7.7 Hz, 2H, Ar-H), 3.68 – 3.65 (m, 1H, αH), 3.20 – 2.89 (m, 2H, βH).

^{19}F NMR (565 MHz, D_2O): δ = -64.85 (dd, J = 891.8, 44.6 Hz, 4F, F_{eq}), -67.63 (dp, J = 709.7, 44.6, 44.1 Hz, 1F, F_{ax}).

^{31}P NMR (243 MHz, D_2O): δ = -145.71 (pd, J = 892.2, 709.9 Hz).

ESI-MS (m/z): $[M]^-$ calculated for $([C_{10}H_{10}F_5NO_3P])^-$: 318.0319 Da; found: 318.0382 m/z

Ammonium N-acetyl-4-(pentafluorophosphato-difluoromethyl)-L-phenylalaninyl-L-leucyl-amide (14)



The dipeptide **14** was synthesized according to the general method. From 94 mg Rink amide resin, 11.6 mg (75%) of the product were obtained as white lyophilisate.

^{1}H NMR (600 MHz, D_2O): δ = 7.49 (d, J = 7.8 Hz, 2H, Ar-H), 7.33 (d, J = 8.0 Hz, 2H, Ar-H), 4.59 (t, J = 7.6 Hz, 1H, αH-Y*), 4.26 (dd, J = 10.1, 4.1 Hz, 1H, αH-Leu), 3.11 (dd, J = 7.6, 3.2 Hz, 2H, βH-Y*), 1.97 (s, 3H, $NHCOCH_3$), 1.62 – 1.49 (m, 3H, βH+γH-Leu), 0.87 (dd, J = 34.1, 5.8 Hz, 6H, δH-Leu).

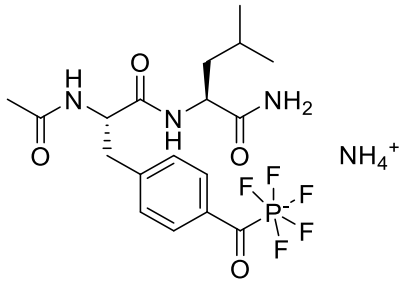
^{13}C NMR (151 MHz, D_2O): δ = 177.10 ($CONH_2$.), 174.16 ($NHCOCH_3$), 173.44 (CO-Y*), 137.70 (Ar-$C_{quart.}$), 137.21 (Ar-$C_{quart.}$), 128.94 (Ar-C), 125.59 (Ar-fC), 55.12 (αC-Y*), 52.20 (αC-Leu), 39.70 (βC-Leu), 36.50 (βC-Y*), 24.14 (γC-Leu), 22.22 (δC-Leu), 21.54 ($NHCOCH_3$), 20.43 (δC-Leu).

^{19}F NMR (565 MHz, D_2O): δ = -68.41 (dp, J = 692.0, 43.1 Hz, 1F, F_{ax}), -72.05 (ddt, J = 864.5, 42.9, 8.5 Hz, 4F, F_{eq}), -97.98 (d, J = 117.9 Hz, 2F, CF_2).

^{31}P NMR (243 MHz, D_2O): δ = -142.91 (pdt, J = 864.7, 691.5, 126.9 Hz).

ESI-HRMS (m/z): $[M]^-$ calculated for $([C_{18}H_{24}F_7N_3O_3P])^-$: 494.1443 Da; found: 494.1445 m/z

Ammonium N-acetyl- 4-(pentafluorophosphato-carbonyl)-L-phenylalaninyl-L-leucyl-amide (15)



The dipeptide **15** was synthesized according to the general method, but cleaved from the resin over 6 h and with the addition of 1% H_2O to the cleavage mixture. From 155 mg Rink amide resin, 16.1 mg (73%) of the product were obtained as white lyophilisate.

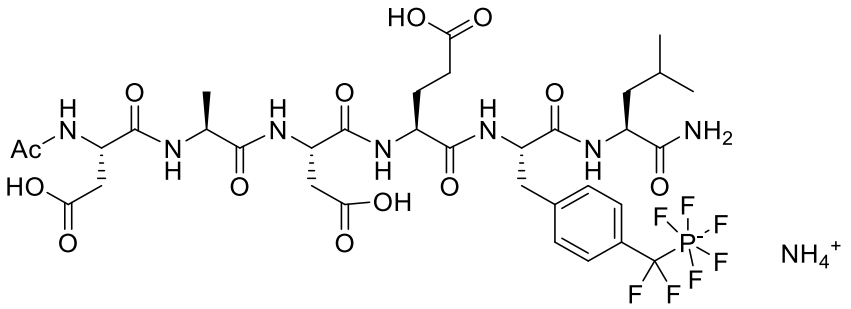
^{1}H NMR (600 MHz, D_2O): δ = 7.97 (d, J = 8.0 Hz, 2H, Ar-H), 7.37 (d, J = 8.2 Hz, 2H, Ar-H), 4.61 (t, J = 7.5 Hz, 1H, αH-Y*), 4.24 (dd, J = 10.0, 4.1 Hz, 1H, αH-Leu), 3.13 (d, J = 7.6 Hz, 2H, βH-Y*), 1.97 (s, 3H, $NHCOCH_3$), 1.64 – 1.40 (m, 3H, βH+γH-Leu), 0.84 (dd, J = 35.2, 5.8 Hz, 6H, δH-Leu).

^{19}F NMR (565 MHz, D_2O): δ = -64.76 (dd, J = 891.9, 44.8 Hz, 4F, F_{eq}), -67.65 (dp, J = 710.5, 44.5 Hz, 1F, F_{ax}).

^{31}P NMR (243 MHz, D_2O): δ = -145.76 (pd, J = 892.1, 710.1 Hz).

ESI-HRMS (m/z): $[M]^-$ calculated for $([C_{18}H_{24}F_5N_3O_4P])^-$: 472.1424 Da; found: 472.1421 m/z

N-Acetyl-L-aspartyl-L-alaninyl-L-aspartyl-L-glutamyl-4-(pentafluorophosphato-difluoromethyl)-L-phenylalaninyl-L-leucyl-amide (16)



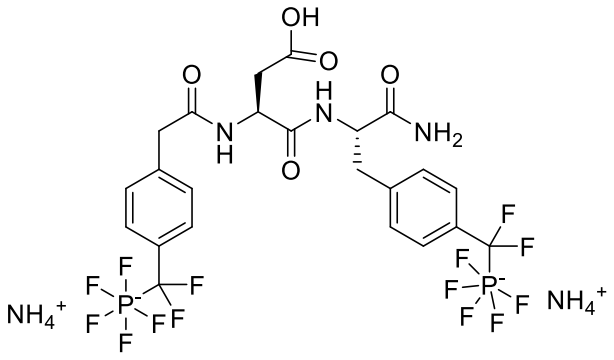
The dipeptide **16** was synthesized according to the general method. From 94 mg Rink amide resin, 15.3 mg (65%) of the product were obtained as white lyophilisate.

^{1}H NMR (600 MHz, D_2O) δ = 7.43 – 7.26 (m, 2H, Ar-H), 7.25 – 7.15 (m, 2H, Ar-H), 4.52 – 4.11 (m, 6H, CHN), 3.15 – 2.89 (m, 2H, CH_2Phe), 2.64 (dd, J=63.7, 15.9, 4H, 2xCH_2Asp), 2.17 (m, 2H, CH_2Glu), 1.91 (s, 3H, Ac), 1.81 (dd, J=29.9, 7.5, 2H, CH_2Glu), 1.56 – 1.37 (m, 3H, CH_2CH) 1.27 (m, 3H, CH_3Ala), 0.76 (d, J=36.1, 6H, 2xCH_3Leu)

^{19}F NMR (565 MHz, D_2O) δ = -67.75 (p, J=42.9, F_{ax}), -68.97 (p, J=44.0, 42.4, F_{ax}), -70.69 – -71.76 (d, J=42.9, 2F, F_{eq}), -72.78 (d, J=42.9, 2F, F_{eq}), -97.98 (dd, J=126.4, 8.7, 2F, CF_2)

HRMS (ESI): $[M]^-$ calculated for $C_{34}H_{46}F_7N_7O_{13}P^-$: 924.2785 Da, found: 924.2795 m/z

Bis-ammonium 4-(pentafluorophosphato-difluoromethyl)-phenylacetamidyl-L-aspartyl- 4-(pentafluorophosphato-difluoromethyl)-L-phenylalaninyl-amide (17)



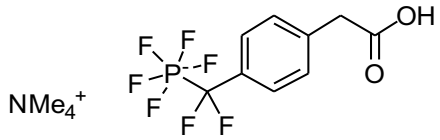
Compound **17** was synthesized applying the general methods for peptide synthesis starting from resin **13** (0.2 g, loading 0.34 mmol/g). Following to the final Fmoc cleavage, compound **18** (5 equiv.) was pre-activated with TBTU (4.9 equiv.) and DIPEA (10 equiv.) in DMF and then added to the N-unprotected dipeptide resin for 2 h. Cleavage and purification was conducted again as described in the general methods. From 159 mg of resin, 11.5 mg (46%) of the product **17** were obtained as white lyophilisate.

^{1}H NMR (600 MHz, D_2O) δ = 7.37 (t, J=9.5, 4H, Ar-H), 7.17 (d, J=7.8, 4H, Ar-H), 4.52 – 4.43 (m, 2H, CHN), 3.45 (s, 2H, CH_2Phe), 3.10 (d, J=13.3, 1H, CH_2Phe), 2.86 (dd, J=13.2, 10.2, 1H, CH_2Phe), 2.43 (ddd, J=96.3, 15.9, 7.2, 2H, CH_2Asp)

^{19}F NMR (565 MHz, D_2O) δ = -67.53 – -68.13 (m), -69.00 (h, J=43.2, 42.7, 2F, F_{ax}), -72.04 (dddd, J=864.7, 42.9, 25.4, 7.9, 8F, F_{eq}), -98.00 (dd, J=126.4, 8.7, 4F, 2xCF_2)

HRMS (ESI): $[M]^-$ calculated for $C_{23}H_{21}F_{14}N_3O_5P_2^{2-}$: 373.5372 Da, found: 373.5387 m/z

Tetramethyl 4-(pentafluorophosphato-difluoromethyl)-phenylacetic acid (18)



The methyl ester **21** (320 mg, 0.985 mmol, 1 equiv.) was added to B*acillus licheniformis* protease (from Sigma Aldrich) (20 mg, 20 mg /mmol) in aqueous buffer (NH_4HCO_3, 50 mM, pH = 7.8, 30 ml) and stirred at 50 °C overnight. The reaction mixture was concentrated under reduced pressure. The residue was purified via column chromatography at MPLC (RP C18, ACN / H_2O + 10 mM NH_4HCO_3, 5 to 99%). Fractions containing the product were concentrated at rotary evaporator and lyophilized, yielding the product **18** (255 mg, 83%) as a white lyophilisate.

^{1}H NMR (500 MHz, MeOD-d_4) δ = 7.41 (d, J=7.7, 2H, Ar-H), 7.26 (d, J=7.9, 2H, Ar-H), 3.57 (s, 2H, CH_2Phe), 3.09 (s, 12H, NMe_4^+)

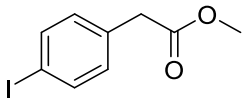
^{13}C NMR (126 MHz, MeOD-d_4) δ = 176.15 (s, C=O), 135.86, 127.86, 125.55 (s, 6 x Ar-H), 54.53 (s, $NMe_4^{+)}$, 42.27 (s, CH_2Phe)

^{19}F NMR (471 MHz, MeOD-d_4) δ = -71.18 (dp, J=696.1, 43.6, 1F, F_{eq}), -72.30 (d, J=43.6, 2F, F_{ax}), -74.13 (d, J=41.4, 2F, F_{ax}), -99.54 (d, J=124.3, 2F, CF_2)

^{31}P NMR (202 MHz, MeOD-d_4) δ = -143.37 (pdt, J=861.4, 694.0, 122.9)

HRMS (ESI): $[M]^-$ calculated for $C_9H_7F_7O_2P^-$: 311.00774 Da, found: 311.00795 m/z

Methyl 4-iodophenylacetate (19)



4-Iodophenylacetic acid (from ABCR, 5 g, 19.1 mmol, 1 equiv.) was dissolved in dry MeOH (25 ml) in a heat- and vacuum-dried Schlenk flask under Ar atmosphere. 3 drops of dry DMF were added and the solution was cooled to 0 °C. Oxalyl chloride (3.27 ml, 38.2 mmol, 2 equiv.) was added dropwise under stirring and the reaction was allowed to reach RT for 16 h. The amber solution was then concentrated using a rotary evaporator, diluted with EtOAc and washed with H_2O, saturated $NaHCO_3$ solution, and brine. The organic layer was dried over Na_2SO_4, filtrated and concentrated in vacuo. The crude was purified via column chromatography at MPLC (SiO_2, EtOAc / Hex, 5 to 100%), product **19** (3.7 g, 70%) was obtained as a colorless oil.

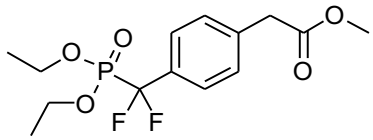
^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.64 (d, J=8.4, 2H, Ar-H), 7.02 (d, J=8.5, 2H, Ar-H), 3.69 (s, 3H, OMe), 3.56 (s, 2H, CH_2)

^{13}C NMR (126 MHz, $CDCl_3$) δ = 171.54 (C=O), 137.76, 133.66, 131.37, (Ar-C), 92.73 (C-I), 52.25 (OMe), 40.73 (CH_2Phe)

HRMS (ESI): $[M+H]^+$ calculated for $C_9H_{10}IO_2^+$: 276.97200 Da, found: 276.97089 m/z; $[M+Na]^+$ calculated for $C_9H_9INaO_2^+$: 298.95394 Da, found: 298.95356 m/z

Spectral data were consistent with published values.[5]

Methyl 4-(diethyl-phosphonato-difluoromethyl)-phenylacetate (20)



A heat- and vacuum-dried Schlenk flask was charged with metallic cadmium (1832 mg, 16.9 mmol, 6 equiv.) activated and dried as previously described and dry DMF (8 ml). To the stirred suspension, diethyl bromophosphonate (1.595 ml, 8.965 mmol, 3.33 equiv.) was added dropwise at RT. The slightly exothermic reaction was stirred for 3 h. In another flask, previously dried 4-iodophenylacetic acid methylester **19** (750 mg, 2.72 mmol, 1 equiv.) was dissolved in dry DMF (1 ml) and CuBr (1169 mg, 8.15 mmol, 3 equiv.) was added. The solution containing the organocadmium was added slowly and dropwise to this stirred mixture under Ar atmosphere and the reaction mixture was stirred for 16 h and monitored via TLC (1:3, EtOAc / Hex). After addition of EtOAc, the precipitate was filtrated off over a bed of Celite and the filtrate washed with sat. aq. sol. NH_4Cl (20 ml, 3x), H_2O (20 ml) and brine (20 ml). The organic layer was dried over Na_2SO_4, filtrated and concentrated under reduced pressure. After purification of the crude via chromatography (SiO_2, EtOAc / Hex 5 to 100%), product **20** (836 mg, 99%) was isolated as a colorless oil.

R_f = 0.4 (50% EtOAc/hexane)

^{1}H NMR (500 MHz, $CDCl_3$) δ = 7.57 (d, J=7.9, 2H, Ar-H), 7.37 (d, J=8.3, 2H, Ar-H), 4.21 – 4.09 (m, 4H, OCH_2CH_3), 3.69 (s, 3H, OMe), 3.66 (s, 2H, CH_2), 1.30 (t, J=7.1, 6H, CH_2CH_3)

^{13}C NMR (126 MHz, $CDCl_3$) δ = 171.43 (C=O), 136.91 (Ar-C), 131.64 (Ar-C), 129.48 (Ar-C), 129.47 (Ar-C), 126.58 (Ar-C), 118.97 (CF_2), 64.88 (2x CH_2 ethyl), 52.25 (OMe), 41.01 (CH_2), 16.43 (2x CH_3 ethyl)

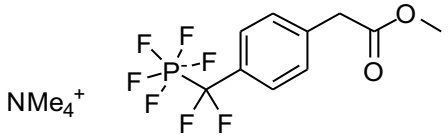
^{19}F NMR (471 MHz, $CDCl_3$) δ = -108.14 (d, J=116.5)

^{31}P NMR (202 MHz, $CDCl_3$) δ = 6.92 (d, J=234.3)

HRMS (ESI): $[M+H]^+$ calculated for $C_{14}H_{20}F_2O_5P^+$: 337.10109 Da, found: 337.10247 m/z $[M+Na]^+$ calculated for $C_{14}H_{19}F_2NaO_5P^+$: 359.08304 Da, found: 359.08386 m/z

Spectral data were consistent with published values.[5]

Tetramethylammonium O-methyl-4-(pentafluorophosphato-difluoromethyl)-phenylacetate (21)



Diethyl phosphonic acid ester **20** (670 mg, 1.99 mmol, 1 equiv.) was dissolved in a Schlenk flask in dry ACN (10 ml). TMSBr (1.314 ml, 9.96 mmol, 5 equiv.) was added dropwise under inert atmosphere. The solution was heated at 60°C for 1.5 h. After disappearance of the starting material monitored via LC-MS, the vial was equipped with inert gas inlet and outlet to allow the release of the gaseous components developed after dropwise addition of dry DMF (766 µL, 9.96 mmol, 5 equiv.) and $(COCl)_2$ (1.67 ml, 19.92 mmol, 10 equiv.). After gas development ceased, the reaction was heated in a sealed vessel under inert atmosphere at 40 °C with a water bath. After 1.5 h, a small aliquot was taken and MeOH was added. The formation of the dimethyl ester was confirmed via LCMS. The reaction mixture was then cooled to 0°C with an ice bath. Previously weighted under inert atmosphere and dried as previously described NMe_4F (1856 mg, 19.92 mmol, 10 equiv.) was then added slowly under inert atmosphere to the stirred and cooled reaction mixture. After 1 h, the mixture was slowly quenched in a cooled sat. aq. sol. of $NaHCO_3$ (30 ml) and extracted with DCM (3 x 30 ml). The collected organic layers were then concentrated at the rotary evaporator, redissolved in H_2O/ACN and purified with RP-MPLC. After purification of the crude via chromatography (RP-C18, ACN / H_2O, 5 to 99%), the fractions were analyzed at LCMS and the one containing the product were concentrated at rotary evaporator and lyophilized, yielding the product **21** (320 mg, 49%) as white lyophilisate.

^{1}H NMR (600 MHz, ACN-d3) δ = 7.40 – 7.32 (m, 2H, Ar-H), 7.21 (d, J=8.1, 2H, Ar-H), 3.62 (s, 5H, CH_2, OMe), 3.02 (s, 12H, NMe_4)

^{13}C NMR (151 MHz, ACN-d3) δ = 172.01 (O=C), 134.40 (d, J=2.1), 129.69 (d, J=1.2), 128.41 (Ar-C), 125.72 (t, J=7.5, CF_2), 55.20 (NMe_4), 51.57 (OMe), 40.14 (CH_2Phe)

^{19}F NMR (565 MHz, ACN-d3) δ = -70.09 (dp, J=696.7, 43.4, 42.6, F_{ax}), -71.75 (ddt, J=855.5, 43.1, 9.4, 4F, F_{eq}), -98.39 (dp, J=120.6, 9.3, 2F, CF_2)

^{31}P NMR (243 MHz, ACN-d3) δ = -143.92 (pdt, J=856.2, 696.6, 120.0)

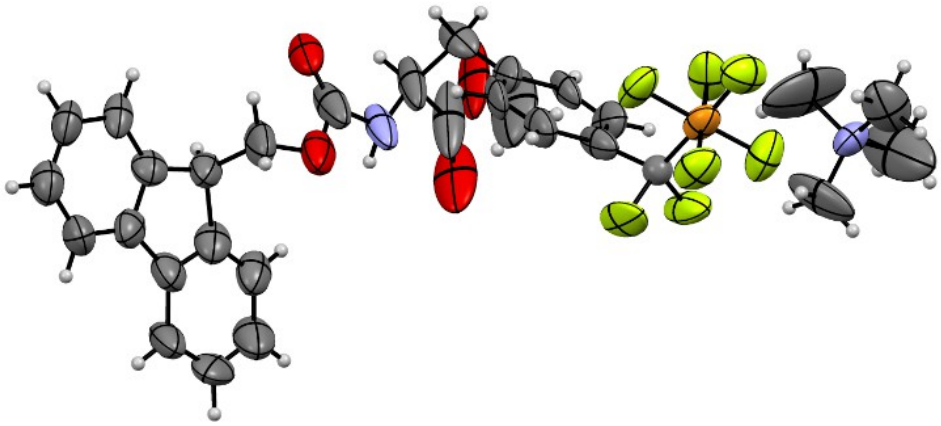
HRMS (ESI): $[M]^-$ calculated for $C_{10}H_9F_7O_2P^-$: 325.0234 Da, found: 325.0241 m/z

Crystal structure determination

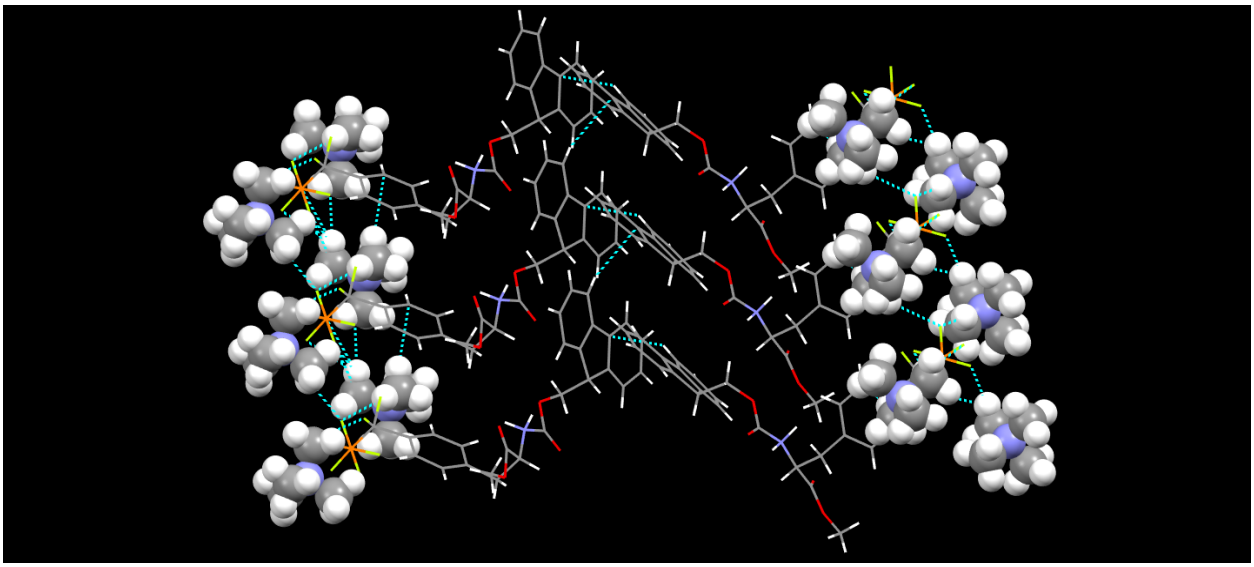
To obtain single crystals of **8**, the compound was dissolved in a 1:1 mixture of water and methanol, transferred into a beaker and sealed with parafilm. Several small holes were punctured into the film to allow slow evaporation of the methanol. After four weeks at room temperature single crystals had grown. Single crystals X-ray diffraction was performed on a Bruker D8 Venture system with graphite-monochromatic Mo-Kα radiation (λ = 0.71073 Å). Data reduction was performed with Bruker AXS SAINT[6] and SADABS[7] packages. The structure was solved by SHELXS 2018[8] using direct methods and followed by successive Fourier and difference Fourier synthesis. Full matrix least-squares refinements were performed on F2 using SHELXL 2018[8] with anisotropic displacement parameters for all non-hydrogen atoms. All other calculations were carried out using SHELXS 2018[8], SHELXL 2018[7] and WinGX (Ver-1.80)[9]. Mercury 2020.1[10] and Diamond 4.6.5[11] were used for structure visualization. Data collection, structure refinement parameters and crystallographic data of compounds **8** are summarized in Supplementary Table 2.

Supplementary Table 2. Crystallographic data of compound **8**

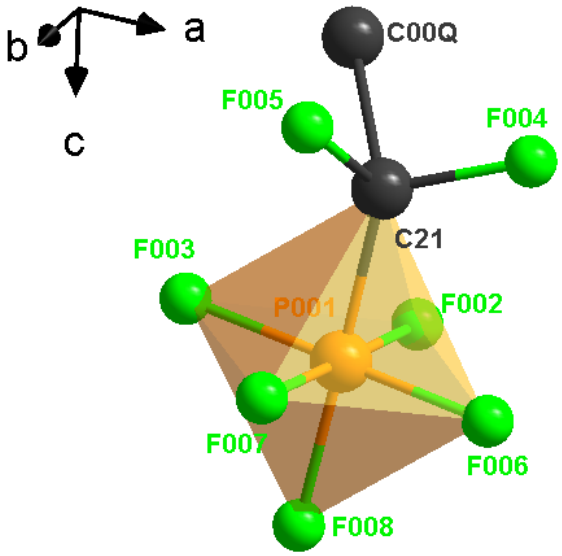
Identification code	**8**	
Empirical formula	C30 H35 F7 N2 O4 P	
Formula weight	651.57	
Temperature	293(2) K	
Wavelength	0.71073 Å	
Crystal system	Orthorhombic	
Space group	P 21 21 21	
Unit cell dimensions	a = 6.6335(10) Å	a= 90°.
	b = 13.663(3) Å	b= 90°.
	c = 33.745(5) Å	g = 90°.
Volume	3058.4(9) $Å^3$	
Z	4	
Density (calculated)	1.415 Mg/m^3	
Absorption coefficient	0.171 mm^{-1}	
F(000)	1356	
Crystal size	0.18 x 0.22 x 0.29 mm^3	
Theta range for data collection	2.345 to 23.277°.	
Index ranges	-6<=h<=7, -15<=k<=15, -37<=l<=37	
Reflections collected	24008	
Independent reflections	4391 [R(int) = 0.1564]	
Completeness to theta = 25.242°	80.5 %	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	4391 / 0 / 406	
Goodness-of-fit on F^2	1.119	
Final R indices [I>2sigma(I)]	R1 = 0.1063, wR2 = 0.2783	
R indices (all data)	R1 = 0.1932, wR2 = 0.3354	
Absolute structure parameter	0.1(6)	
Extinction coefficient	0.003(3)	
Largest diff. peak and hole	0.634 and -0.403 $e.Å^{-3}$	



Supplementary Figure 1. Crystal structure of compound **8**



Supplementary Figure 2. Crystal structure of compound **8**



Bond lengths:

P001	F006	1.5485(101)
	F007	1.5774(92)
	F002	1.6109(93)
	F003	1.6191(90)
	F008	1.6237(115)
	C21	1.6833(115)

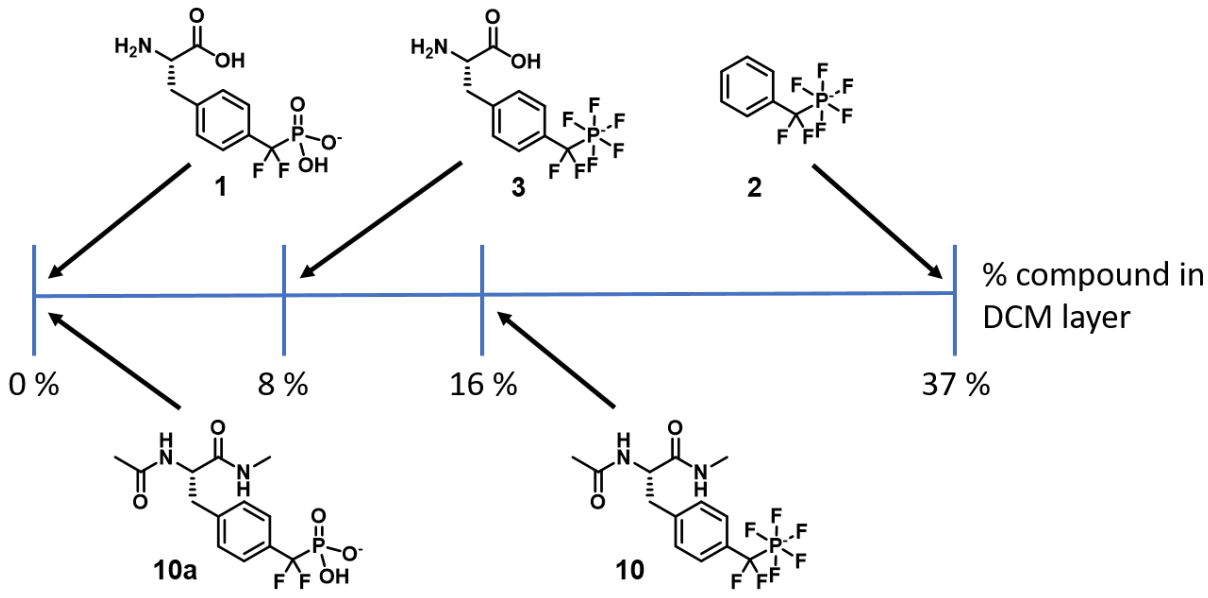
Bond angles:

P001	F006	F007	89.656(548)
	F006	F002	91.873(529)
	F006	F003	178.168(531)
	F006	F008	90.256(574)
	F006	C21	91.517(615)
	F007	F002	178.380(514)
	F007	F003	88.599(535)
	F007	F008	89.025(584)
	F007	C21	92.528(566)
	F002	F003	89.877(516)
	F002	F008	90.432(570)
	F002	C21	87.967(565)
	F003	F008	90.273(568)
	F003	C21	88.002(594)
	F008	C21	177.649(619)

Supplementary Figure 3: Bond angles and lengths of the R-CF_2-PF_5 group of **8**

Determination of partition coefficients

Compounds were weighted and added to 50 ml round bottom flasks. 10 ml DCM and 10 ml water (MilliQ, 0.055 µS) were added and the mixture stirred for 24 h at room temperature. The phases were separated, evaporated and the remaining solids weighted to determine the concentration ratio between the two phases and thereof the logarithmic partition coefficient log P.



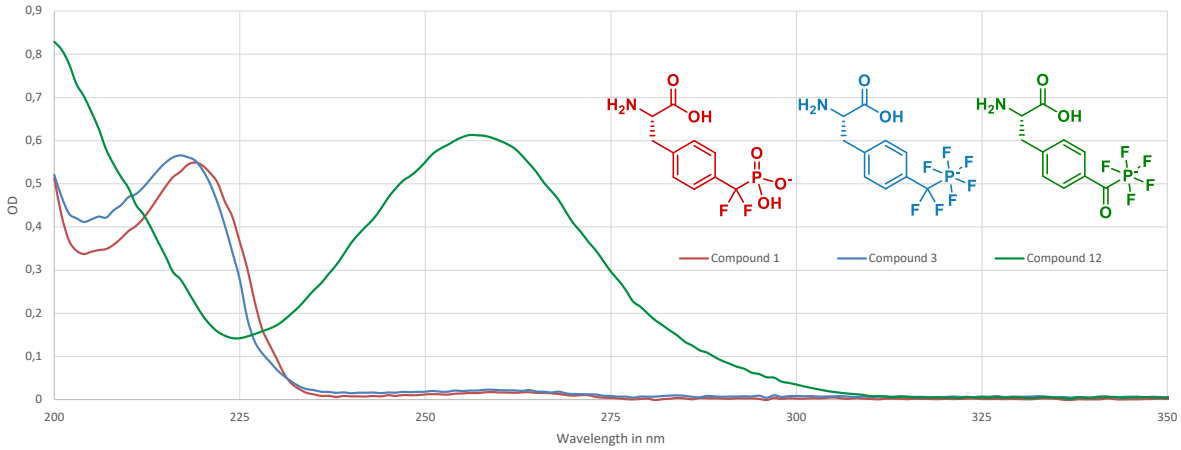
Supplementary Figure 4. Visual representation of the partition of compounds **1**, **2**, **3**, **10** and **10a** between water and DCM

Supplementary Table 3. Partition of compounds **1**, **2**, **3**, **10**, **10a** between water and DCM

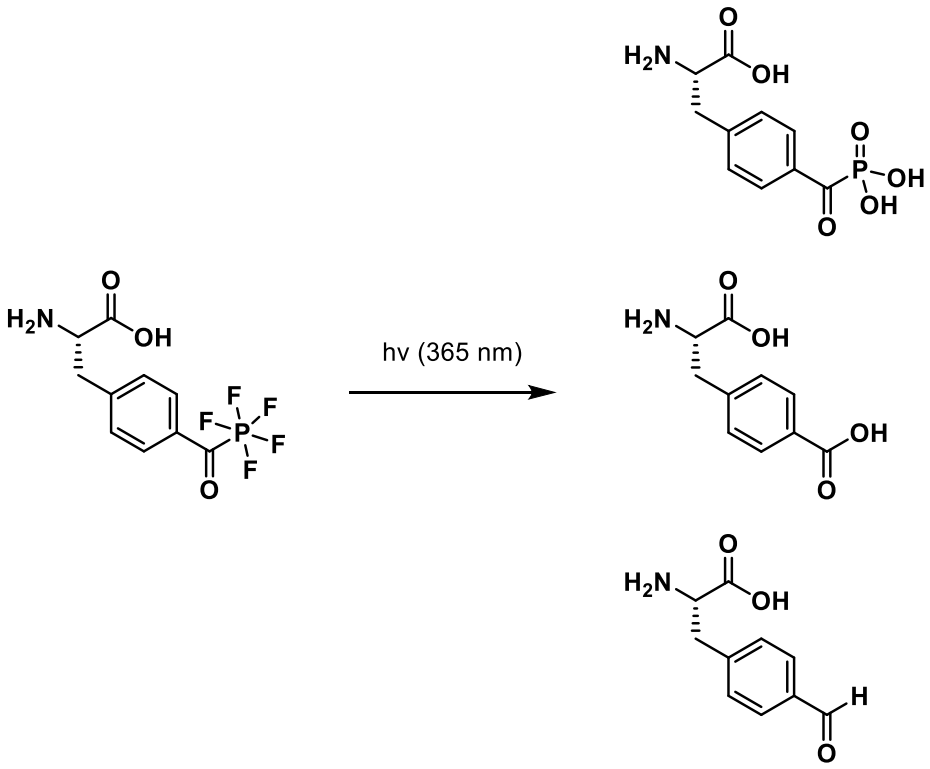
Compound	Percent compound in DCM	Log P (DCM/water)
1	< 1 %	-
2	37 %	-0.23
3	8.2 %	-1.05
10	16 %	-0.81
10a	< 1 %	-

UV spectroscopy and irradiation experiments

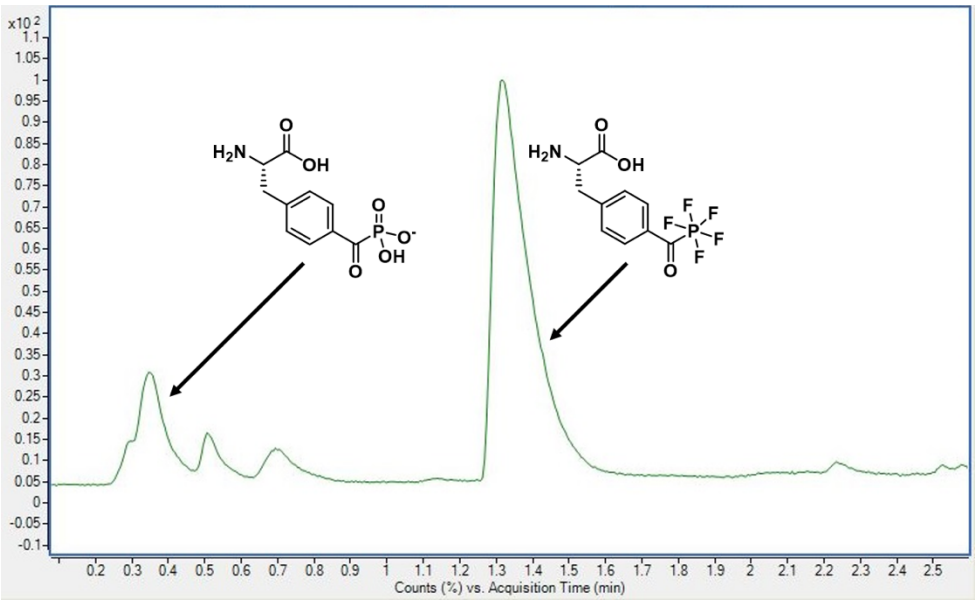
The three amino acids **1**, **3**, and **12** were dissolved in water (MilliQ, 0.055 µS) at a concentration of 0.5 mM. 50 µl of these solutions were added into a UV-star 96-well plate and an absorbance scan was performed. Compounds **1** and **3** displayed a similar absorbance pattern, with **1** giving a distinct maximum at 219 nm (ε = 17259 $M^{-1}cm^{-1}$) and **3** at 219 nm (ε = 17788 $M^{-1}cm^{-1}$). Compound **12** showed a maximum at 256 nm (ε = 19271 $M^{-1}cm^{-1}$) and a second peak at 343 nm (ε = 192 $M^{-1}cm^{-1}$) (see Supplementary Figure 71). Thus, substitution of the CF_2 group by a CO group leads to a shift of the maximum by 39 nm. The second peak can be attributed to the n-π* transition of the ketophosphonate. Accordingly, irradiation at 365 nm for 2 h led to the photoconversion of the amino acid **12**.



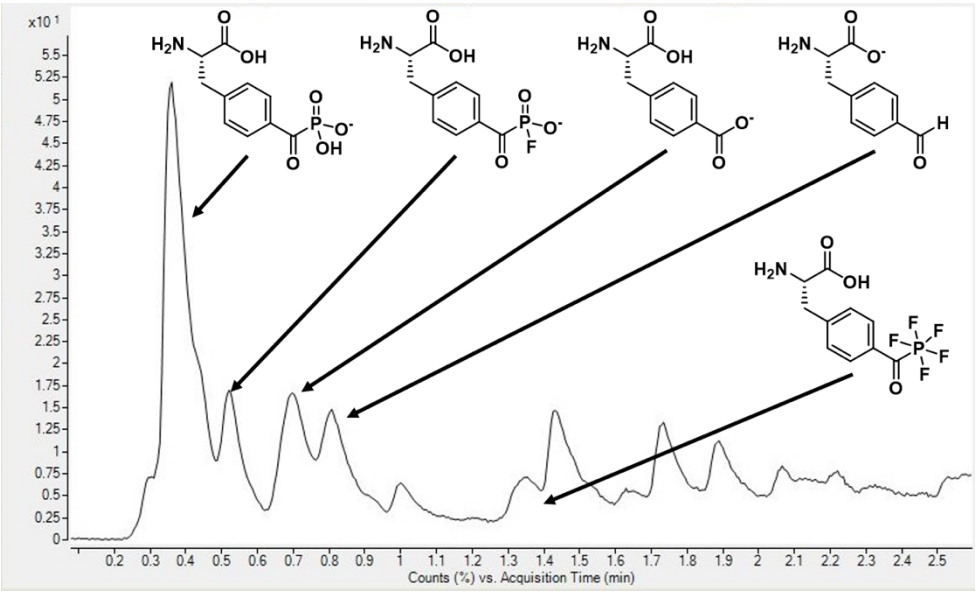
Supplementary Figure 5: UV spectra of 0.5 mM solutions of compounds **1**, **3** and **12** in water



Supplementary Figure 6: After irradiation at 365 nm for 2 h at room temperature in 70/30 iPrOH/H_2O compound **12** (7.5 mM) several derivatives of compound **12** were identified, namely the keto-phosphonate, carboxylic acid and aldehyde.



Supplementary Figure 7: HPLC chromatogram of a 7.5 mM solution of compound **12** in 70/30 iPrOH/H_2O prior to irradiation. Column B, eluent 5-95% ACN in 8 min, detector: total ion current (TIC).



Supplementary Figure 8: HPLC chromatogram of a 7.5 mM solution of compound **12** in 70/30 iPrOH/H_2O after irradiation at 365 nm for 2 h at room temperature. Column B, eluent 5-95% ACN in 8 min, detector: total ion current (TIC).

IR spectroscopy experiments

Methods

FTIR spectra were recorded on a Bruker Vertex 70 IR spectrometer, using an attenuated total reflection (ATR) element from IRubis and a custom-made PTFE cell. A background spectrum was first recorded with 300 µl of milliQ water, which were then replaced by 300 µl of a 10 mM aqueous sample solution of **1a** or **2**. The difference absorbance spectrum of the sample was recorded against the pure water background. All spectra were measured using a DTGS detector by averaging 4096 spectra at a resolution of 2 cm^{-1}. Data were baseline-corrected applying a 4^{th} degree polynomial.

Band assignment

The theoretical IR spectra of the tetramethylammonium pentafluorophosphato-difluoromethyl-benzene **2** and the sodium phosphono-difluoromethyl-benzene fragment **1a** were calculated by density functional theory (DFT) with Gaussian 16 software using the B3LYP/6-31++G(d,p) level of theory. Spectra of **2** and **1a** samples between 4000 and 400 cm^{-1} were recorded in water solution and in transmission using KBr pellets (not shown), the latter to improve the signal to noise ratio and to avoid interference from water bands during the band assignment.

Bands between 1380 and 810 cm^{-1} were assigned to the stretching vibrations of the CF_2 spacer and are observed in both compounds. The beating vibrations of the phenyl group appear at ~1630 cm^{-1} in both spectra, overlapped by the O-H deformation mode from water (solvent). In both DFT models, CF_2 stretching modes were strongly coupled with the a1 and b2 beating vibrations of the phenyl group. As a result, multiple CF_2 stretching bands can be observed (Supporting Table 3) instead of the two expected from a single CF_2 group (symmetric and asymmetric stretching).

In **1a**, the stretching vibrations from the CF_2 appear in the experimental spectra between 1354 cm^{-1} and 824 cm^{-1} overlapped with the P-O stretching and P-O-H deformation vibrations. In the DFT simulation, CF_2 stretching modes were strongly coupled to these vibrations. Simulation was performed for different protonation degrees of the PO_3 headgroup and the monoprotonated form was found to reproduce better the experimental bands (Supplementary Table 3). The band at 1198 cm^{-1}, assigned to one of the P-O stretching and P-O-H deformation modes, did not appear in the monosodium derivative but could be observed in the disodium salt (deprotonated) at 1125 cm^{-1}. The experimental spectrum therefore suggests that the monoprotonated (monoanion) and deprotonated (dianion) forms of the phosphonate group coexist in solution under these conditions.

Supplementary Table 4. Band assignment of **2** and **1a** fragments from comparison with DFT calculations. Abbreviations: v=stretching, δ=deformation, asym=asymmetric, sym=symmetric, //=parallel to.

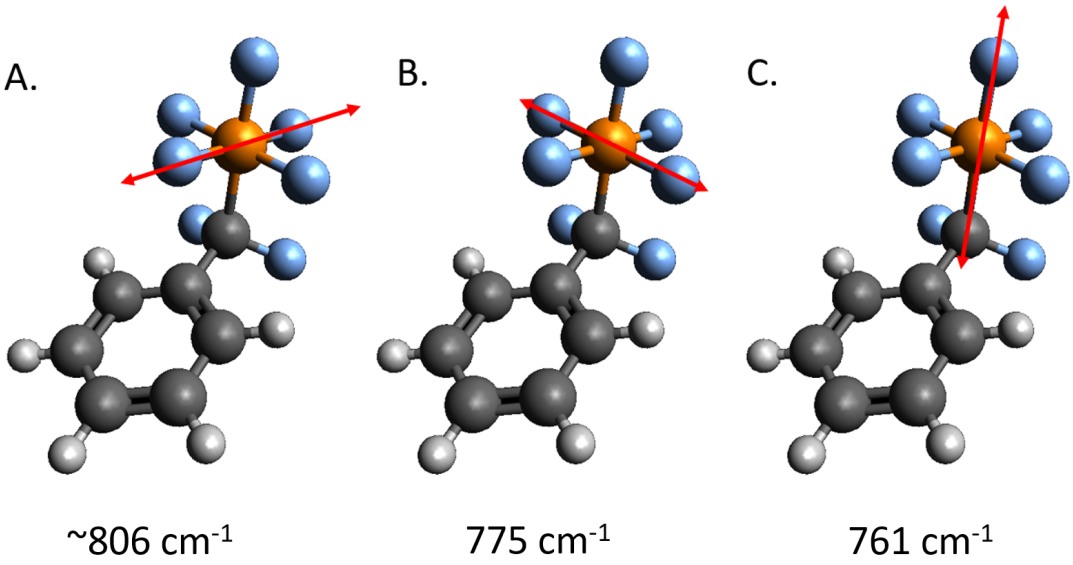
2

DFT	**Exp.**	**Assignment**
wavenumber / cm^{-1}	**wavenumber / cm^{-1}**	
776.5	761	v(PF) axial // b1
808.1	775	v(PF) in plane // b2
814.7	~806	v(PF) in plane // a1
1037.1	1035	$v(CF_2)_{asym}$
1090.1	1103	$v(CF_2)_{sym}$
1254.1	1249	$v(CF_2)_{sym}$

1a

DFT (monoNa)	**Exp.**	**Assignment**
wavenumber / cm^{-1}	**wavenumber / cm^{-1}**	
1036.4	1033	$v(CF2)_{asym}$, v(PO), δ(POH)
1059.6	1069	$v(CF2)_{sym}$, v(PO), δ(POH)
1090	1100	$v(CF2)_{sym}$, v(PO)
1116.7	1133	$v(CF2)_{asym}$
*	1198	v(PO), δ(POH)
1286.1, 1251.3	1251	v(PO), δ(POH), $v(CF2)_{sym}$

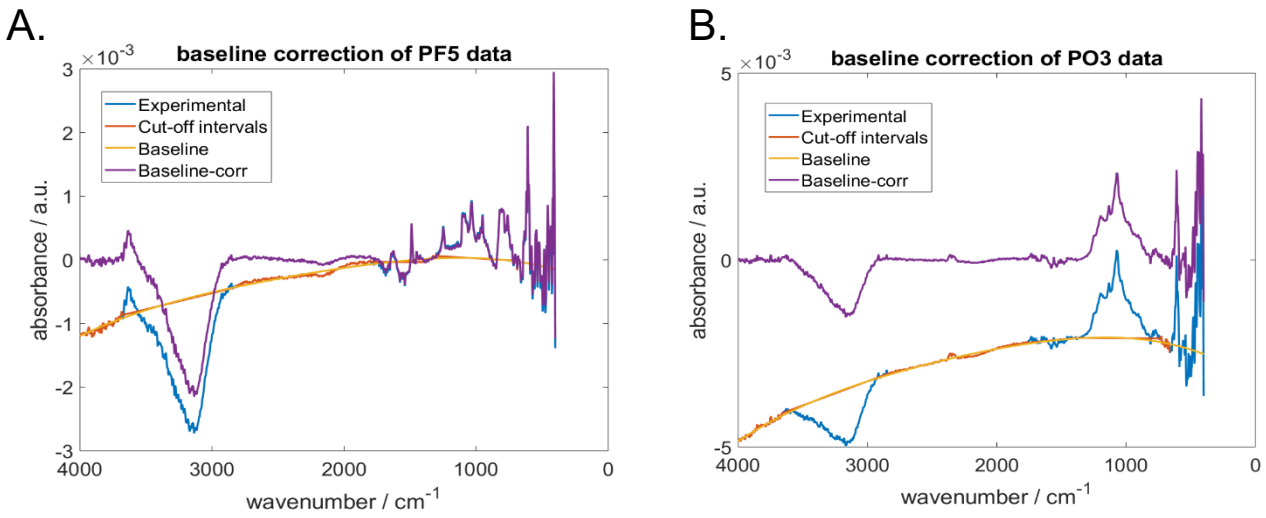
The characteristic PF stretching vibrational bands of the fragment **2** appear at lower frequencies between 900 and 710 cm^{-1}. In both, experimental and theoretical spectra, 3 peaks could be distinguished at ~806, 775 and 761 cm^{-1} (aqueous solution) and at 815, 808 and 777 cm^{-1} (DFT). These were assigned to the in-plane PF stretching vibration along the a1 direction of the phenyl group (Supplementary Figure 3, A), to the in-plane PF stretching along the b2 direction of the phenyl group (Supplementary Figure 3, B), and to the axial (terminal) PF stretching (Supplementary Figure 3, C) along the b1 direction of the phenyl group. Fluorination of the phosphorous also decouples the CF_2 stretching modes from those of the headgroup, which can be distinguished clearly in the **2** experimental spectra.



Supplementary Figure 9. The three PF stretching vibrations with experimental values

Baseline correction

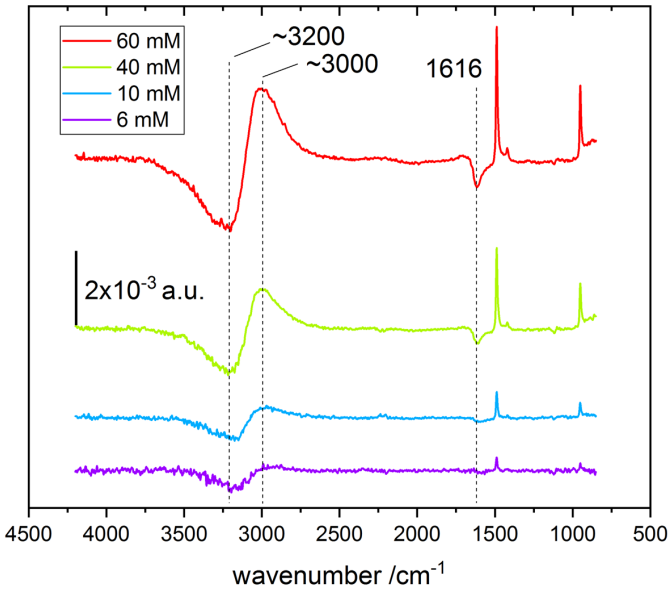
Experimental spectra were baseline corrected by manually selecting those intervals where no specific bands neither from the compound nor from water were observed. The ends of these intervals were connected by straight lines and data were averaged using a sliding window with a length of 200 data points (~1/20 of total spectral interval) over the complete spectral range. The resulting trace was fitted by a 4th degree polynomial to obtain the baseline, which was then subtracted from the experimental data. A visual representation of all steps can be found in Supplementary Figure 4.



Supplementary Figure 10. Baseline correction of experimental data from A.) **2** and B.) **1a** in 10 mM aqueous solution.

Hydration shell of tetramethylammonium counterion

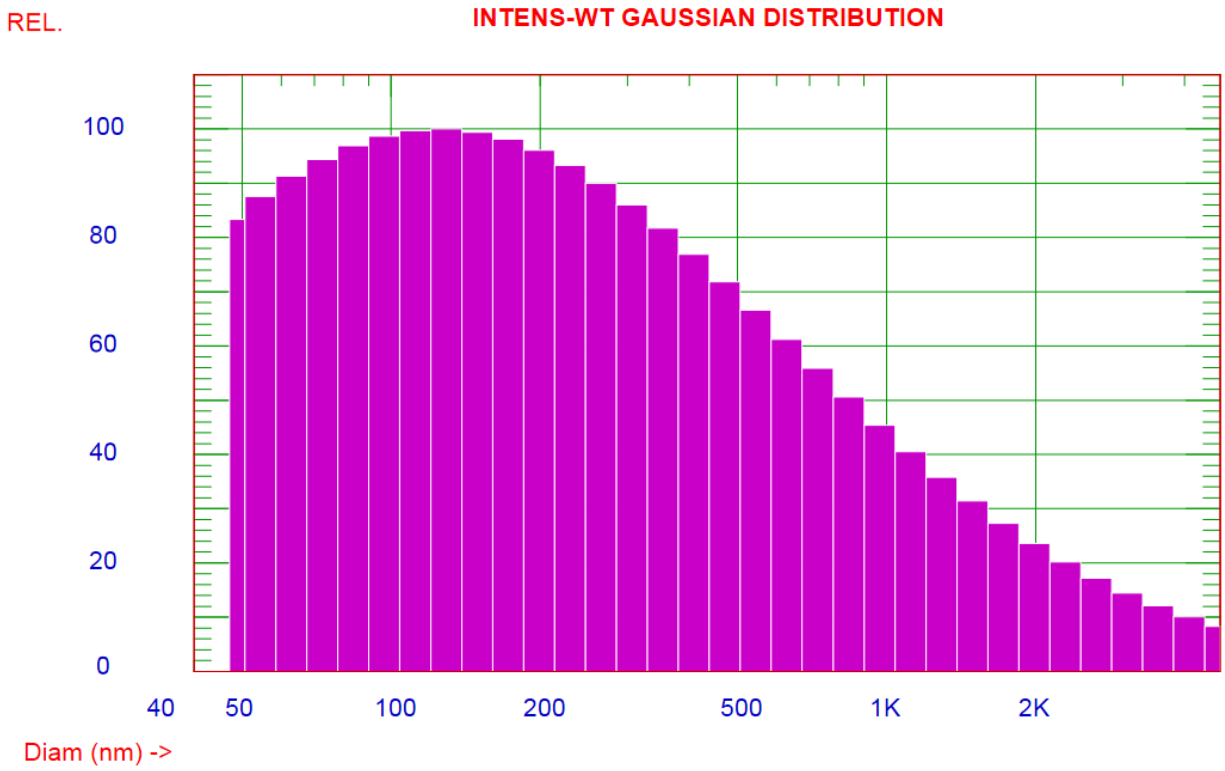
The hydration shell of the tetramethylammonium counterion was evaluated in an independent experiment to discard it as the cause of the observed dangling-water specific bands. The difference spectrum of a TMA fluoride aqueous solution against a pure-water background was recorded at different concentrations (Supplementary Figure 5). The spectra showed two negative peaks at ~3200 cm^{-1} and at ~1620 cm^{-1}, caused by the reduction in water concentration, and linked to water molecules fully exposed to hydrogen bonding. No dangling-water specific bands (~3630 cm^{-1}) were observed at any concentration.



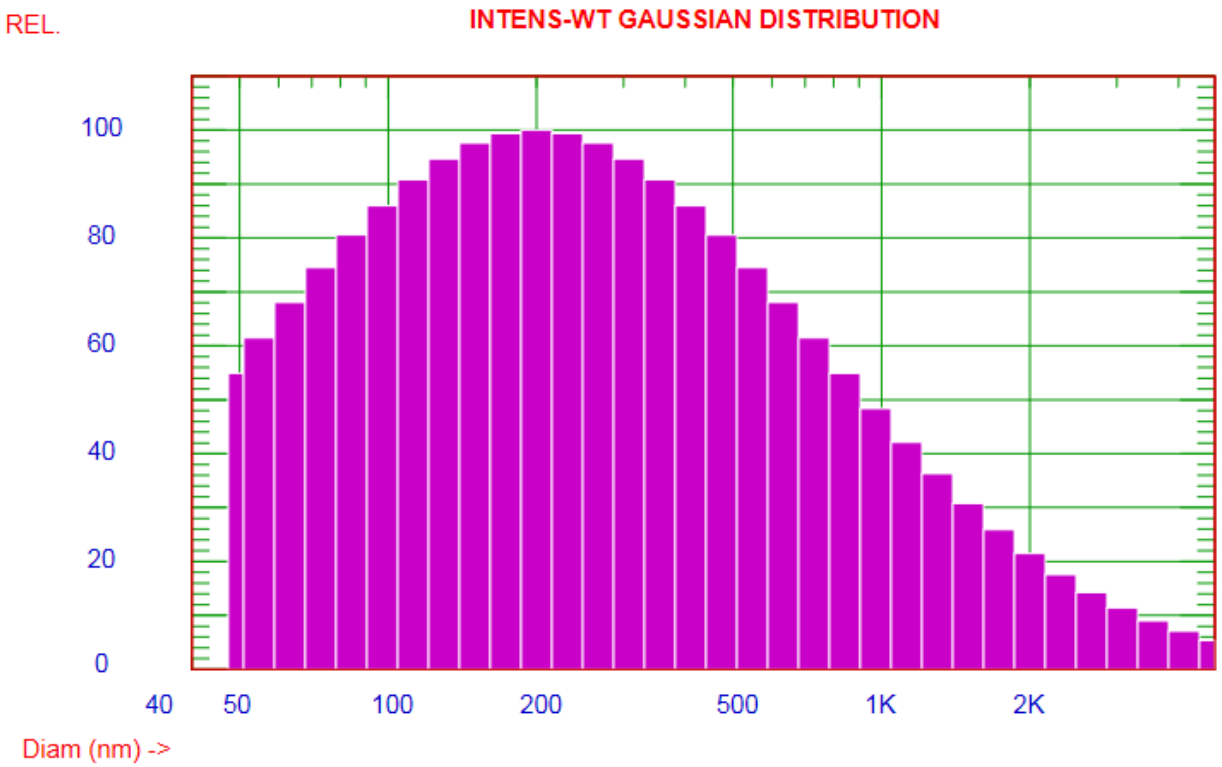
Supplementary Figure 11. Spectra of tetramethylammonium fluoride at different concentrations versus a pure water background. Spectra recorded with 512 co-additions with a LN-MCT detector

Dynamic Light Scattering

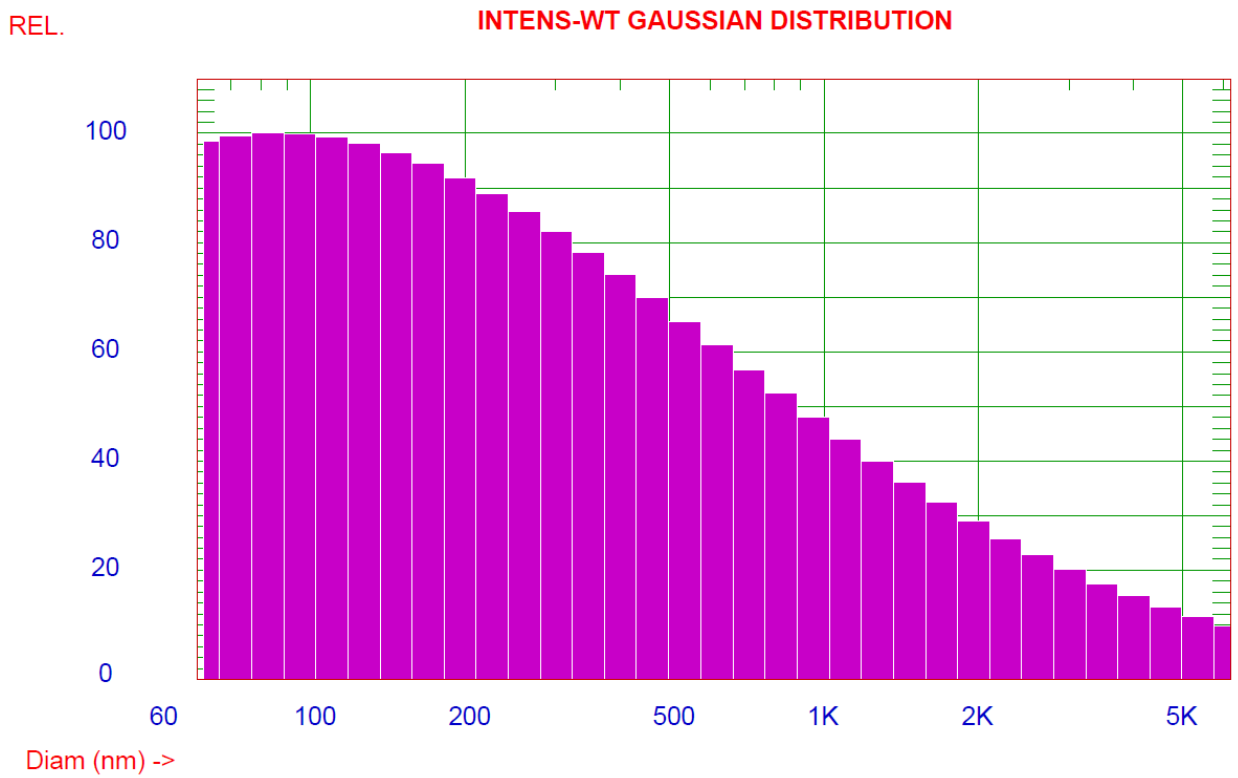
A clear solution of compound **16** in DMSO / buffer 1:1 (650 µM) was subjected to dynamic light scattering indicated the formation of nanoparticular aggregates with a mean diameter of 427 nm (median ca. 120 nm). Dilution of this sample with 1:1 buffer/DMSO to a concentration of 325 µM led to a slight change in particle size distribution, with a mean diameter of 474 nm (median 180 nm).
Furthermore, two samples of **16** in pure DMSO were diluted with buffer to contain 5% DMSO (assay conditions). Here, at concentrations of 250 µM and 125 µM, nano particles with a mean diameter of 619 nm and 776 nm, respectively, were observed. These results indicate that **16** was not fully dissolved under assay conditions.



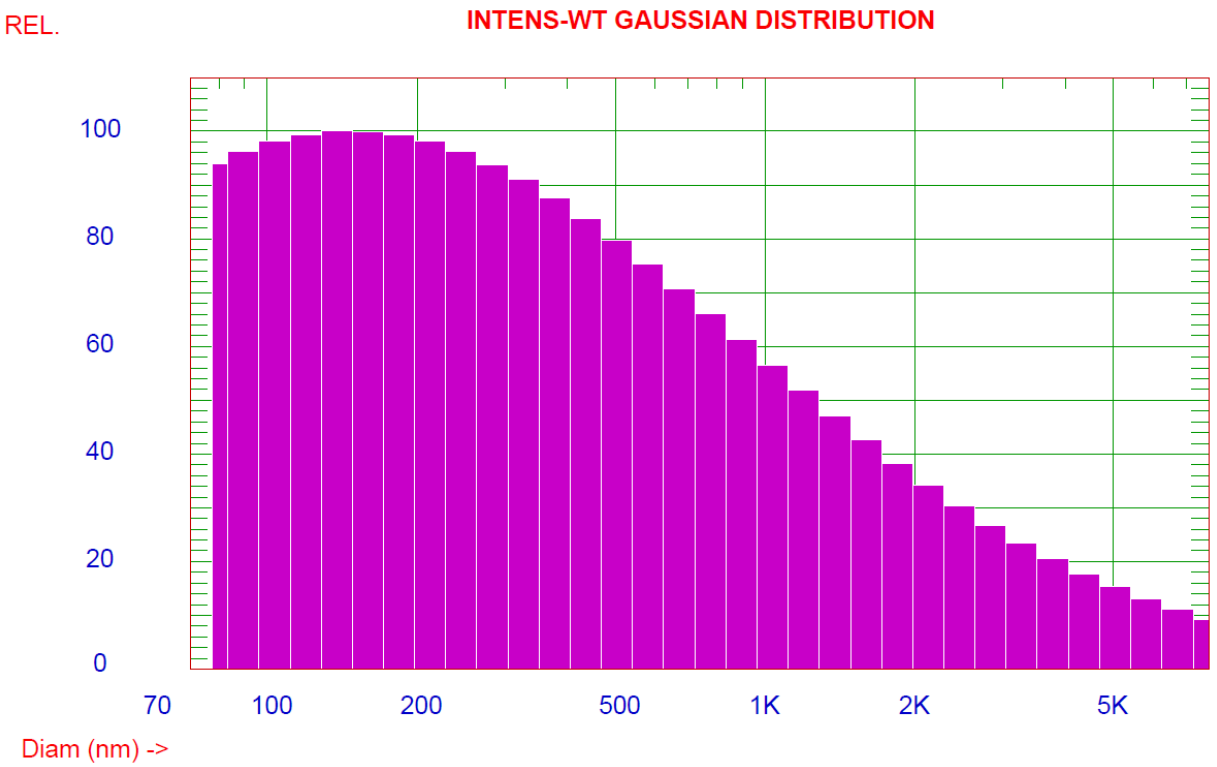
Supplementary Figure 12. DLS of compound **16** at an apparent concentration of 650 µM (50% DMSO) shows undissolved compound in the form of nano particles. These have a mean diameter of 471.7 nm.



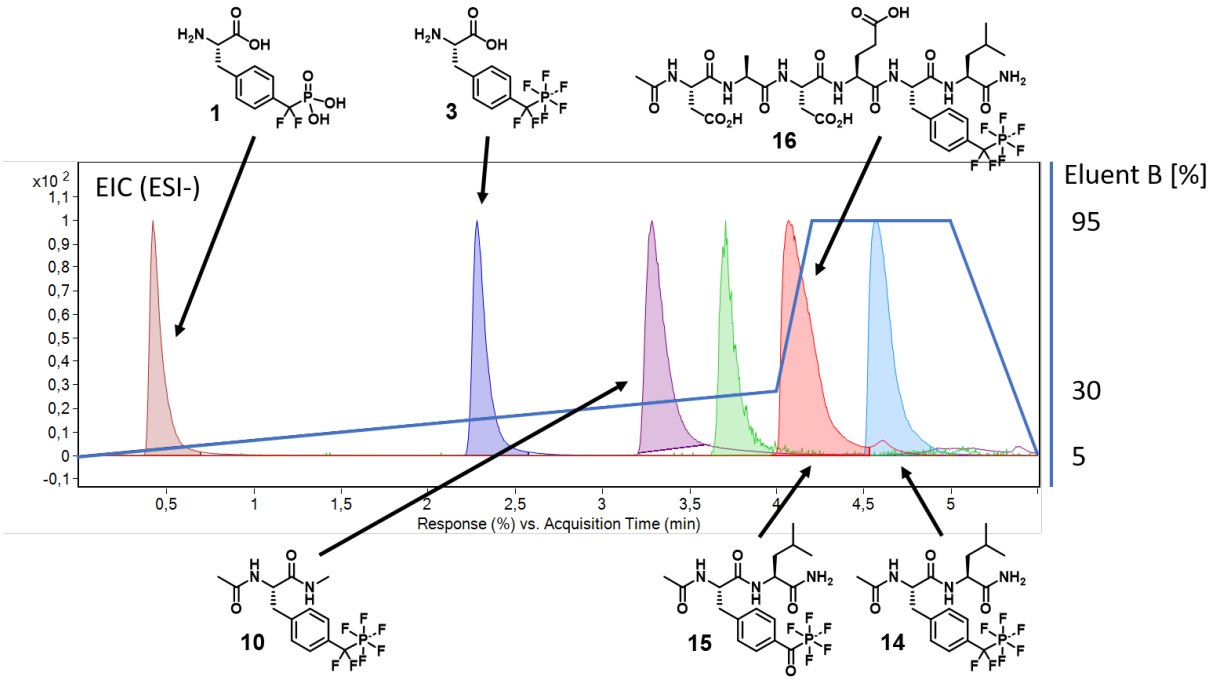
Supplementary Figure 13. DLS of compound **16** at an apparent concentration of 325 μM (50% DMSO) shows undissolved compound in the form of nano particles. These have a mean diameter of 473.7 nm.



Supplementary Figure 14. DLS of compound **16** at an apparent concentration of 250 µM (5% DMSO) shows undissolved compound in the form of nano particles. These have a mean diameter of 618.5 nm.


REL.
INTENS-WT GAUSSIAN DISTRIBUTION
100
80
60
40
20
0
70
100
200
500
1K
2K
5K
Diam (nm) ->
Supplementary Figure 15. DLS of compound 16 at an apparent concentration of 125 µM (5% DMSO) shows undissolved compound in the form of nano particles. These have a mean diameter of 775.5 nm.

HPLC retention times



Supplementary Figure 16. Overlayed HPLC chromatograms of compounds **1**, **3**, **10**, **14**, **15** and **16**, column B, eluent ACN, gradient as shown, EIC detector.

Biochemical evaluation of synthesized compounds

PTP1B

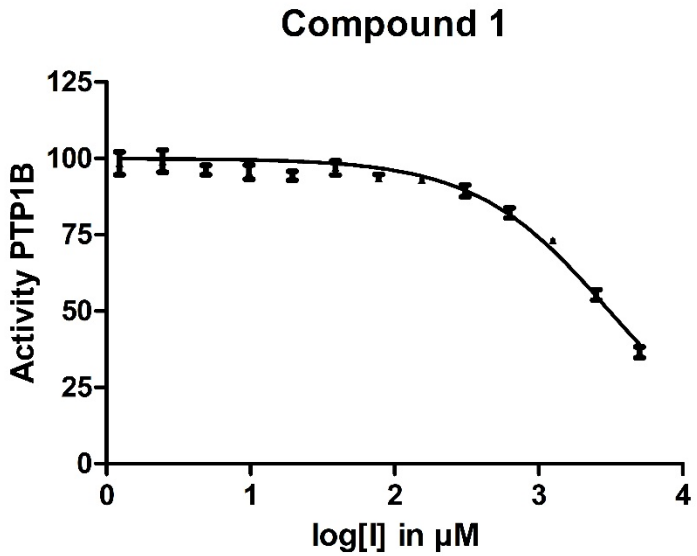
Recombinant human PTP1B was obtained from Abcam (ab51277) at a concentration of 100 µM and used as received, without further purification.

Enzymatic DiFMUP assay:

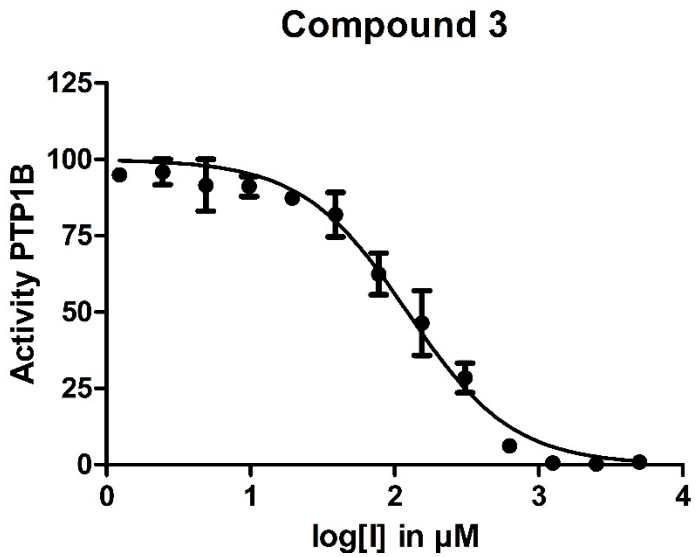
An enzyme assay with 6,8-difluoro-4-methylumbelliferyl phosphate (DiFMUP) as substrate was used to determine the activity of inhibitors toward PTP1B. In Method A, test compounds were dissolved and serially diluted in buffer (0% DMSO). In Method B, test compounds were dissolved in a 1:1 mixture of DMSO and buffer (20 mM stock) and serially diluted with the same mixture resulting in a final DMSO concentration of 2.5% in the assay. In Method C, compounds were dissolved and diluted in DMSO to a final DMSO concentration of 5% in the assay, and in method D, compounds were dissolved in DMSO and diluted with buffer to a DMSO concentration of 5%, serially diluted with 5% DMSO in buffer resulting in a final DMSO concentration of 2.5% DMSO in the buffer.

The assay buffer contained 50 mM MOPSO (pH 6.5), 200 mM NaCl, 0.03% Tween-20, 50 µM tris-(2-carboxyethyl)-phosphin (TCEP) (freshly added prior to each measurement) and 1.5 nM PTP1B (final concentration). The final assay volume was 20 µL. Enzyme and test compound in buffer solution were incubated for 30 min at RT. The reaction was started by adding DiFMUP to a final concentration of 67 µM. This substrate concentration matches the experimentally determined K_M value of the enzyme. Measurements were performed on a Genius Pro Reader (SAFIRE II, instrument serial number: 512000014) with the following settings: measurement mode: Fluorescence Top; λ_{ex}: 360 nm (bandwidth 20 nm); λ_{em}: 460 nm (bandwidth 20 nm) ; gain (manual): 60; number of scans: 8; FlashMode: high sensitivity; integration time: 40 µs; lag time: 0 µs; Z-position (manual): 13900 µM; number of kinetic cycles 10; kinetic interval: 60 s; total kinetic run time 10 min. Measurements were performed in triplicate. IC_{50} values were calculated with Prism 5 (for Windows, Version 5.01, GraphPad Software Inc.) and were converted into the corresponding K_I values applying the Cheng Prusoff equation.[12]

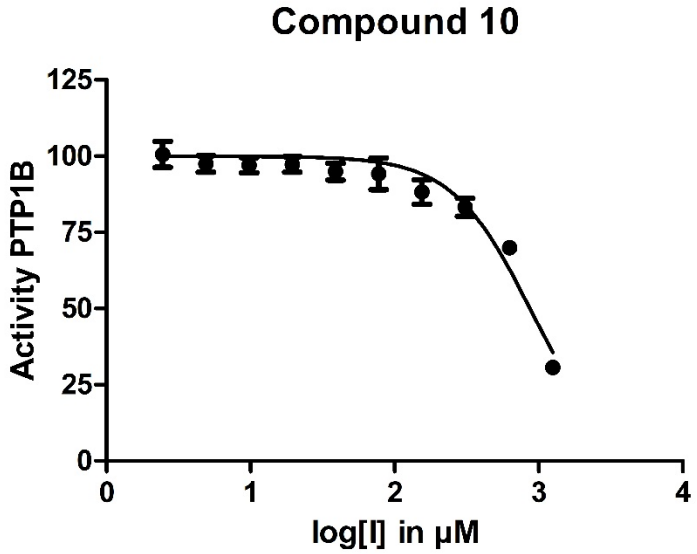
Inhibition of PTP1B



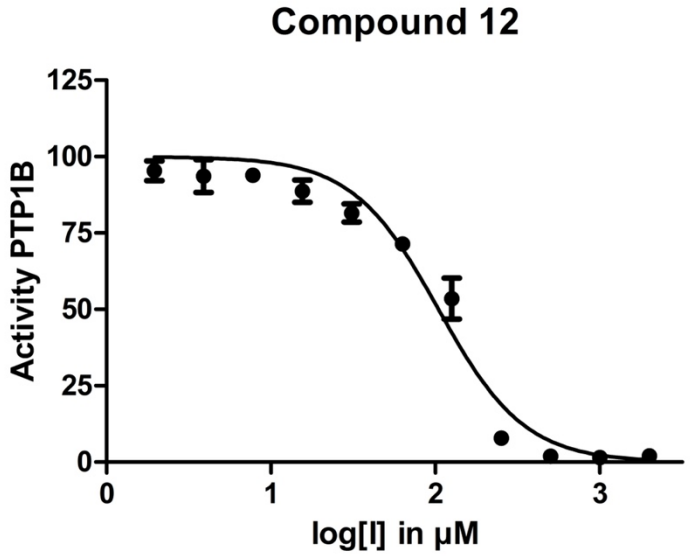
Supplementary Figure 17. Inhibition of PTP1B by compound **1**, IC_{50}: 3.1 ± 0.37 mM (Method A)



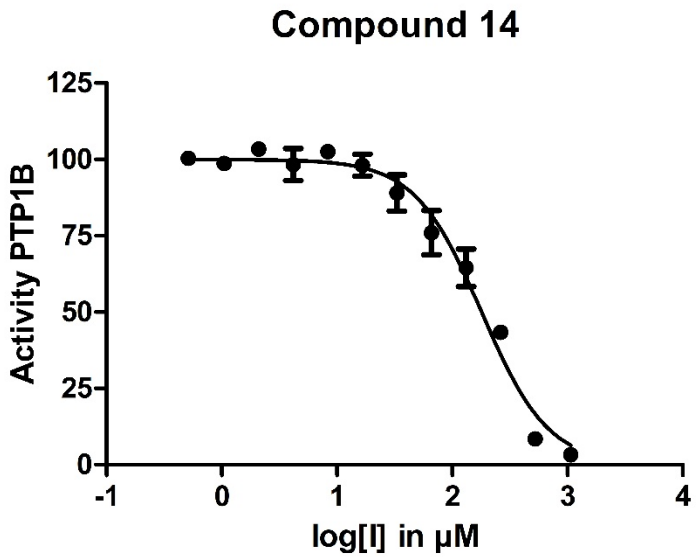
Supplementary Figure 18. Inhibition of PTP1B by compound **3**, IC_{50}: 122 ± 16 µM (Method B)



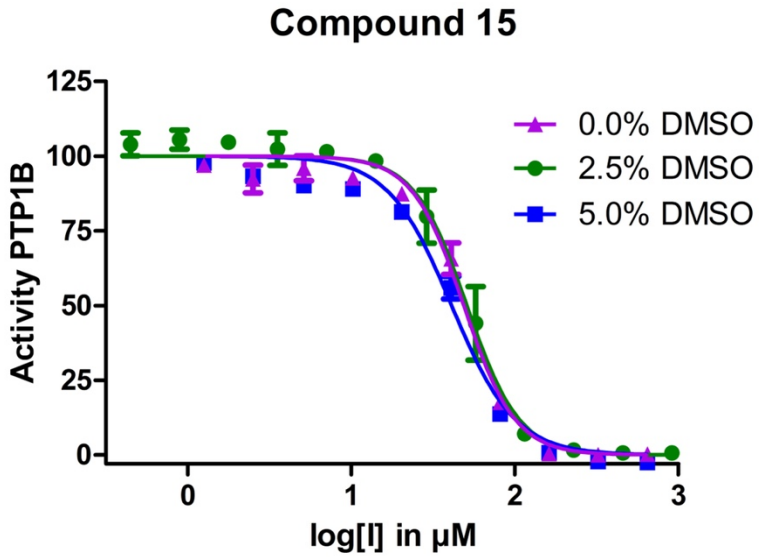
Supplementary Figure 19. Inhibition of PTP1B by compound **10**, IC_{50}: 872 ± 95 µM (Method B)



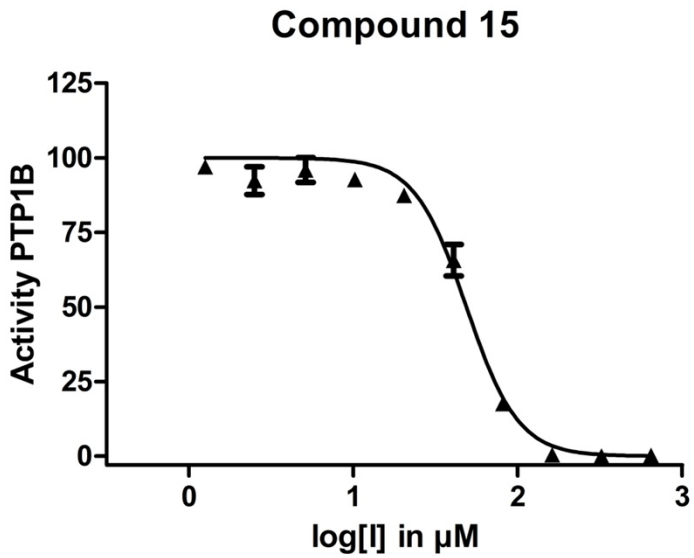
Supplementary Figure 20. Inhibition of PTP1B by compound **12**, IC_{50}: 104 ± 14 µM (Method B)



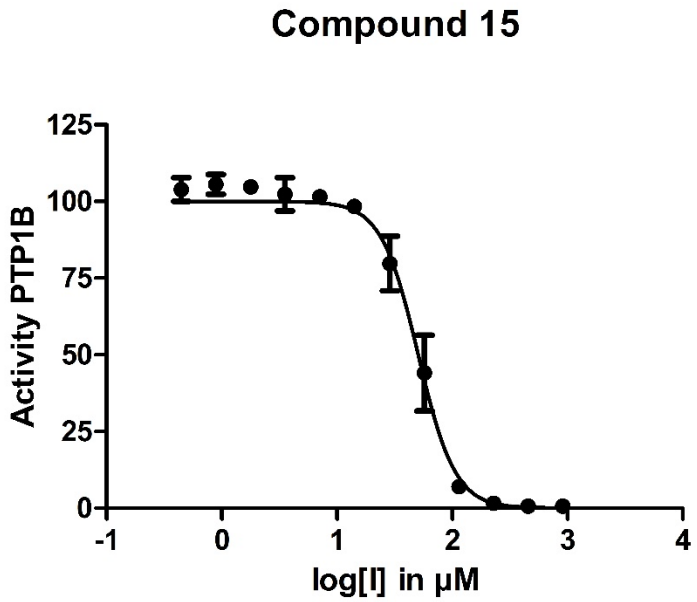
Supplementary Figure 21. Inhibition of PTP1B by compound **14**, IC_{50} = 180 ± 20 µM (Method B)



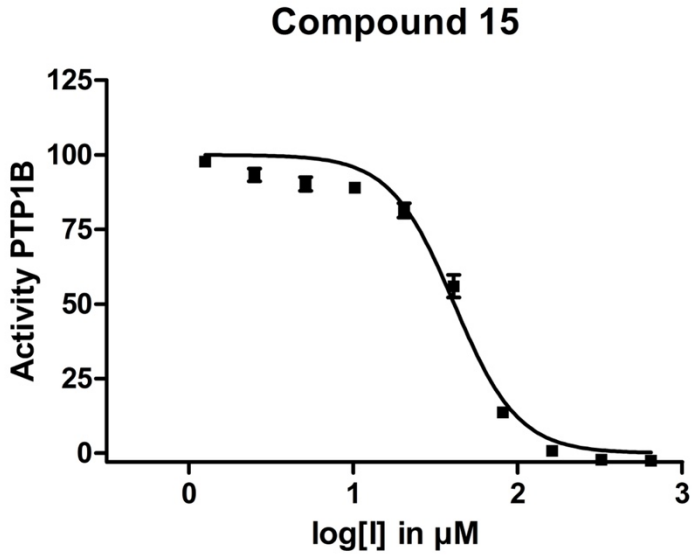
Supplementary Figure 22. Inhibition of PTP1B by compound **15** under different assay conditions. Method A (magenta): final DMSO concentration of 0%. Method B (green): final DMSO concentration of 2.5%. Method C (blue): final DMSO concentration of 5%.



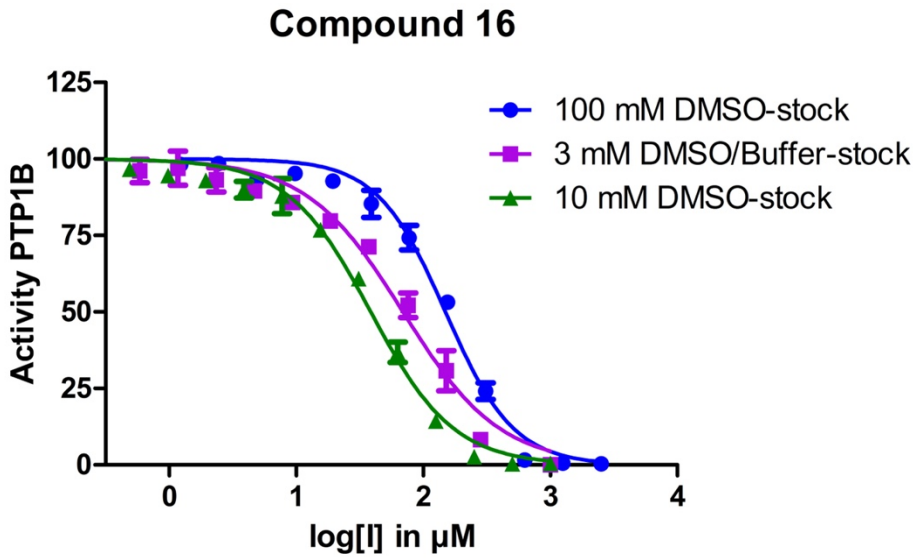
Supplementary Figure 23. Inhibition of PTP1B by compound **15**, IC_{50} = 48 ± 8 µM (Method A)



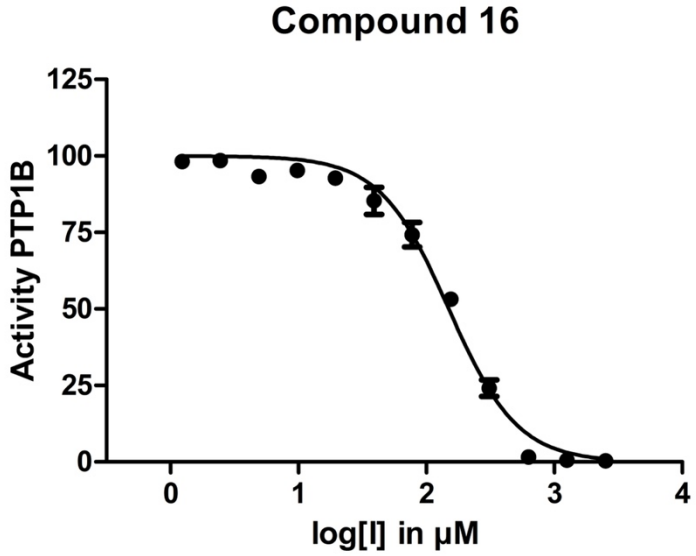
Supplementary Figure 24. Inhibition of PTP1B by compound **15**, IC_{50} = 50 ± 8 µM (Method B)



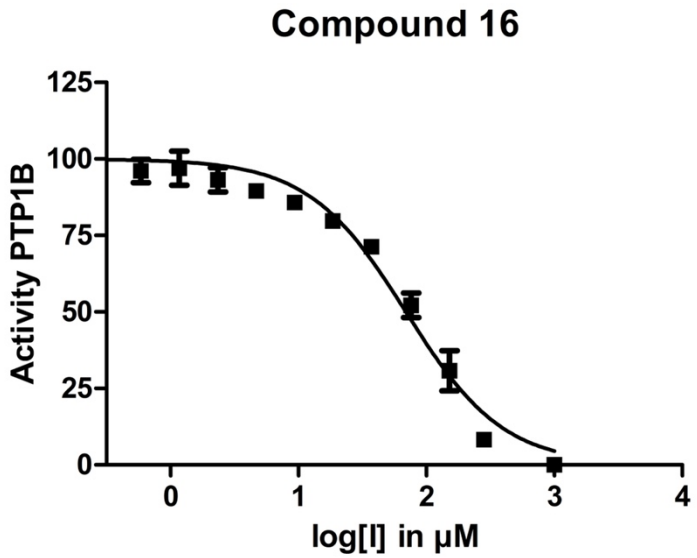
Supplementary Figure 25. Inhibition of PTP1B by compound **15**, IC_{50} = 41 ± 6 µM (Method C)



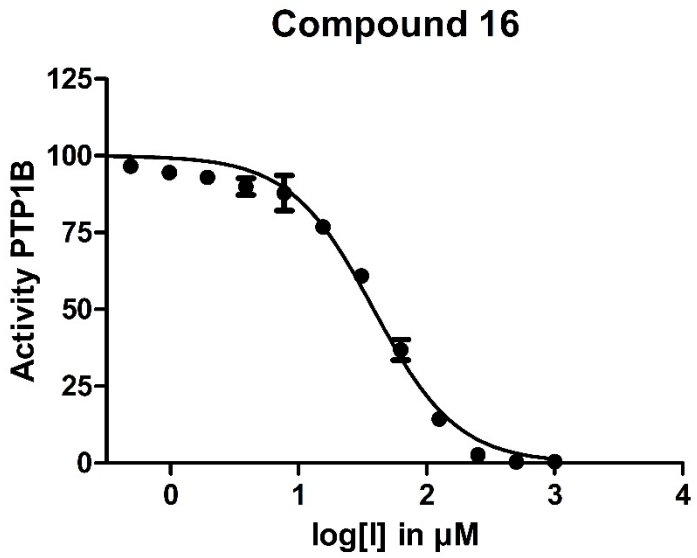
Supplementary Figure 26. Inhibition of PTP1B by compound **16** under different assay conditions. Method D (blue): final DMSO concentration of 2.5%. Method B (magenta): final DMSO concentration of 2.5%. Method C (green): final DMSO concentration of 5%.



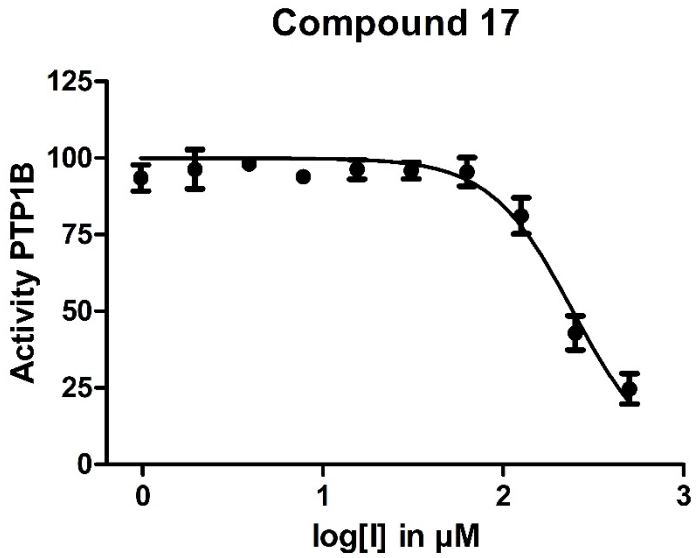
Supplementary Figure 27. Inhibition of PTP1B by compound **16**, IC_{50} = 149 ± 26 µM (method D)



Supplementary Figure 28. Inhibition of PTP1B by compound **16**, IC_{50} = 67 ± 10 µM (method B)



Supplementary Figure 29. Inhibition of PTP1B by compound **16**, IC_{50} = 38 ± 6 µM (method C)



Supplementary Figure 30. Inhibition of PTP1B by compound **17**, IC_{50} = 243 ± 23 µM (method B)

Computational methods

Protein preparation for docking

The protein X-ray diffraction crystal structure of PTP1B (PDB code: 4Y14)[13] was prepared for docking and simulations with Schrödinger's Protein Preparation Wizard.[14] The protonation states of amino acid sidechains were assigned with PROPKA at pH 7.0. Small molecules, crystal water and the A Chain of the dimer were deleted. The hydrogen-bond network was optimized and a brief molecular mechanics minimization using the OPLS4 force field[7] was run.

Ligand docking

The structures of 1 and 3 were docked to the binding pocket of PTP1B using Schrödinger's Glide [15] and OPLS4 force field[16]. A receptor grid was generated using the default setting with OH- and SH- groups within the binding pocket allowed to rotate. Ligand docking was performed with the XP protocol, which applies sampling based on anchors and refined growth as well as a scoring function which scores the docking poses based on physico-chemical descriptors. Non-planar amide conformations were penalized and halogens were included as weak noncovalent interaction acceptors of hydrogen bond type.

Molecular dynamics simulation

Molecular dynamics simulations with PTP1B without ligand and with ligands 1 and 3 were performed with GROMACS 2019-4[17] and our amended version of the AMBER14SB force field.[18] The protein was prepared in the same way as for docking and placed into a dodecahedric box of TIP3P[19] water with 1.1 nm distance of the box edges to the solute. The starting structure was energy minimized and equilibrated for 100 ps in the NVT ensemble, followed by a 1 ns equilibration in the NPT ensemble. Production run had a length of 100 ns, the integration timestep was 2 fs and a snapshot was saved every 1 ps. Periodic boundary conditions were used in all three directions. Covalent bonds to hydrogen atoms were treated as constraints. The applied thermostat was a velocity rescaling scheme[20] and the applied barostat the Parrinello-Rahman barostat[21]. The cut-off for Lennard-Jones and Coulomb interactions was set to 1.0 nm. For Coulomb interactions, the PME method[22] was used. The Verlet cut-off scheme was used to generate neighbor lists.

Molecular dynamics simulation analysis

Molecular dynamics simulations were analysed using the mdtraj[23] python package. Distances and RMSF were calculated using its built-in methods. For the detection of π-π interactions, the centroid of the two respective aromatic rings were calculated and their distance was measured. To be considered for a π-π interaction, the distance had to be less than 4.4 Å. Additionally, the angle between the unit vectors orthogonal to the ring planes had to be less than 30 degrees to be counted as π-π interaction.

Force field parametrization

Parameters were retained from the AMBER14SB force field where possible, however supplementary parameters were required to simulate the nonstandard amino acid structures.[18]

Missing bonded parameters were provided by the general Amber force field (GAFF) for organic molecules,[24] with the help of the acpype tool.[25] Although GAFF can describe an extensive variety of organic molecules, it does not provide a sufficient model for the sp^3d^2-hybridized phosphorous atom in **3**, necessitating the determination of additional bonded parameters using ab initio methods. The bond parameters and the angular force constant of GAFF atom type "p5" (phosphorous with four substituents) were reappropriated for phosphorous with six substituents. The geometry was approximated by an octahedral shape. This assumption was made based on a density functional theory (DFT) geometry optimization. The angle parameters 90° and 180° were chosen to enforce octahedral geometry. Missing torsional parameters were obtained via relaxed dihedral scans in 72 steps of 5° intervals at the HF/6-31G* level of theory. The torsional parameters k_ϕ, n and ϕ_s, used for proper dihedrals in the Amber force field family, were obtained by optimizing the function, $V_d(\phi_{ijkl}) = k_\phi(1 + \cos(n\phi - \phi_s)$ to fit the relaxed scan data using the SciPy module scipy.optimize.curve_fit.[26] Re-optimized Lennard Jones parameters, which better model hydrophobic properties, were used as nonbonding parameters for fluorine in place of GAFF parameters.[27]

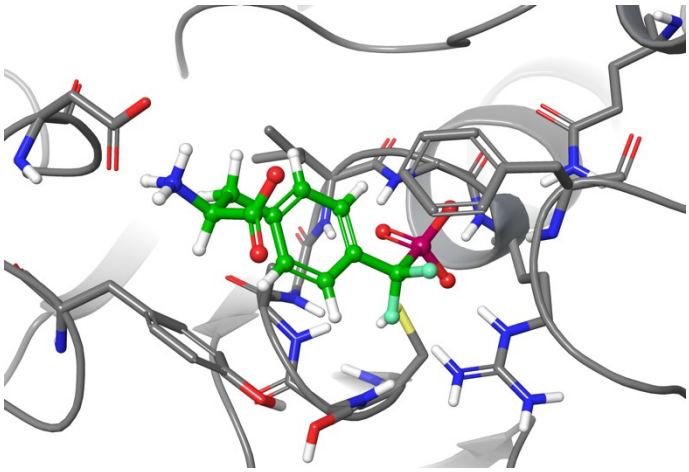
Partial atomic charges for the amino acids **1** and **3** were determined using a charge fitting procedure adapted from Robalo et al.[27] The method involves two iterations of the two stage restrained electrostatic potential (RESP) protocol,[28] in which the first iteration relies on a single conformation and the second iteration averages RESP-fitted charges over multiple conformations. The initial RESP-fitted charges were applied to simulate the free amino acid in TIP3P water for a production run of 100 ns in the NPT ensemble. Conformations at 1 ns intervals were extracted from the simulation and submitted to conventional two stage RESP fitting, such that the final partial atomic charges are based on the average values from 101 conformations. The RESP methodology was implemented using the antechamber program in the Ambertools package and ab initio calculations were performed in Gaussian 16.[29]

Docking studies

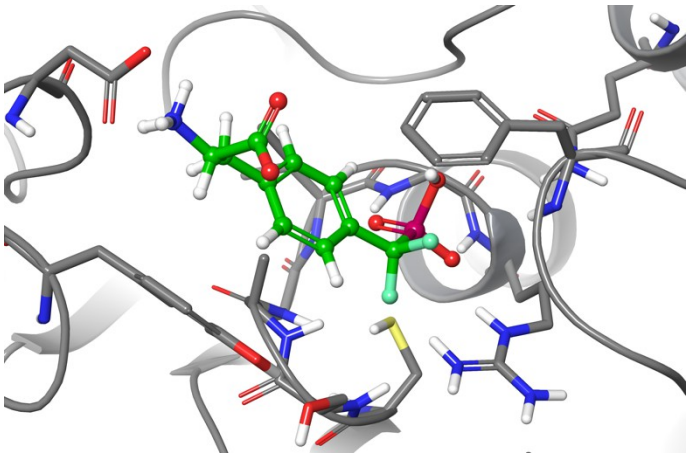
Structure and druggability of the PTP1B binding pocket

The binding site was assessed using Schrödinger's SiteMap[30], which applies a grid-based algorithm to detect and score binding pockets suitable for drug-like ligand binding based on electrostatic and geometric properties. A positively charged main pocket and a negatively charged side pocket were identified. The main pocket consists of the backbone NH-groups of residues 215-220. Also the positively charged sidechain of Arg221 can be found there. The negative charge in the side pocket can be allocated to the sidechain of Asp48 and the backbone carbonyl and sidechain C=O of Asn262. Between the main- and side pocket there are two aromatic rings of Tyr46 and Phe182. The pocket was evaluated by SiteMap's Dscore. The Dscore of the binding pocket is calculated to be 0.6, which rates it as an "undruggable" pocket, meaning it is difficult to address by drug-like ligands.[31]

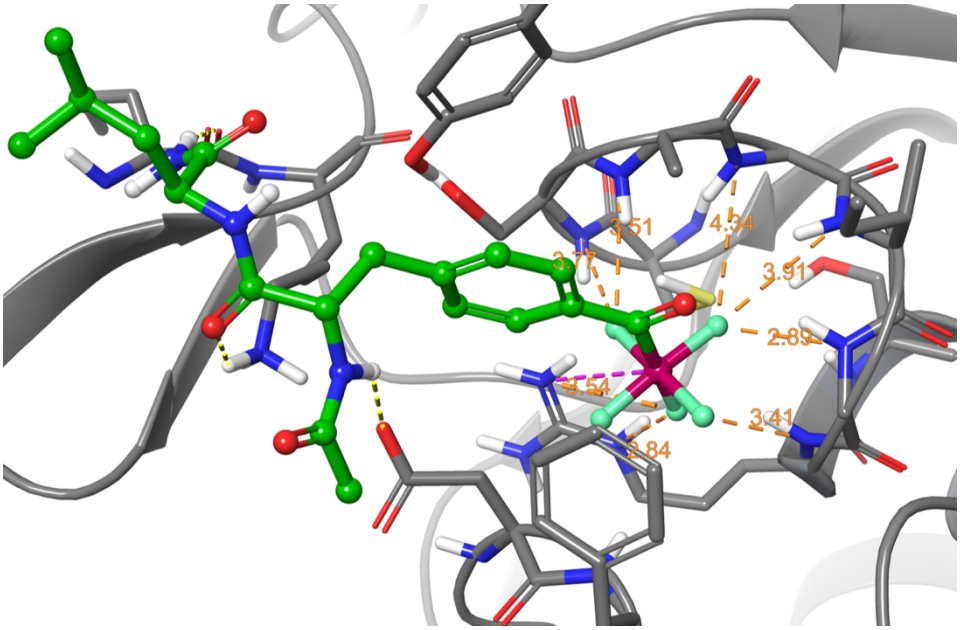
Docking poses



Supplementary Figure 31. Ligand **1** (di-anion) final docking pose in PTP1B binding pocket



Supplementary Figure 32. Protonated ligand **1** (mono-anion) final docking pose in PTP1B binding pocket



Supplementary Figure 33. Ligand **15**, final docking pose in PTP1B binding pocket

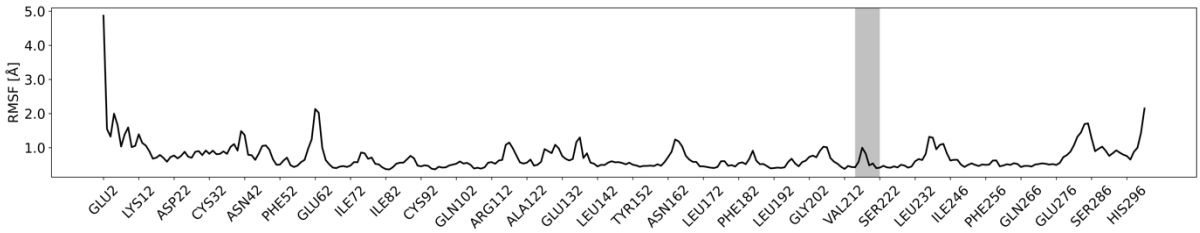
Docking score

The final docking poses of **1** and **3** were evaluated for their docking score, as seen in Table 4. Amino acid **3** scores better (more negative) than amino acid **1**. This result is expected as a better inhibition is measured for amino acid **3** as well, however the lower docking score does not necessarily imply a higher binding affinity of compound **1** as approximations in the Glide score "omit essential thermodynamics of the free energy of binding" and "accurately estimating ligand-protein affinities remains beyond the capabilities of docking scoring functions".[32]

Supplementary Table 6. Glide docking scores of amino acids **1** and **3**

Compound	Glide Docking Score
1 di-anion	-7.494
1 mono-anion	-7.793
3	-10.333

Flexibility of the PTP1B protein backbone

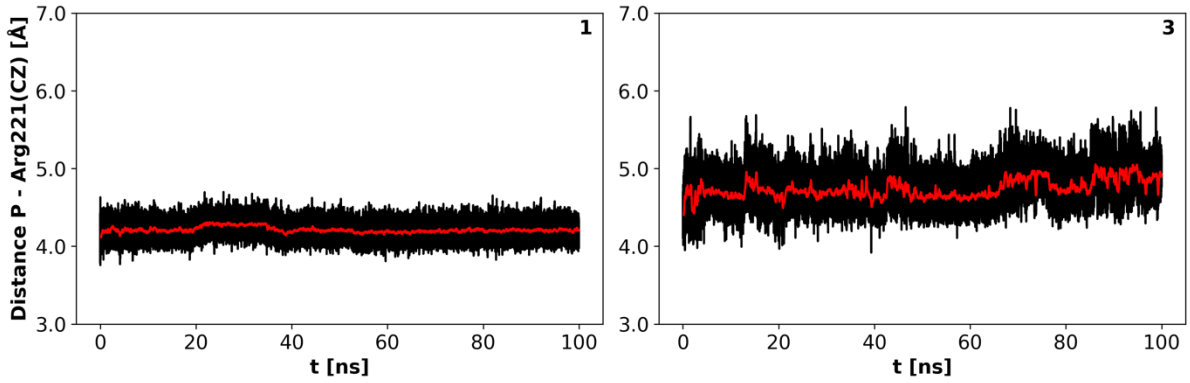


Supplementary Figure 34. RMSF of C_α–atoms of each residue of PTP1B throughout a 10 ns MD simulation of the apo protein. The binding pocket area of residues 215-221 is highlighted in grey

The RMSF of C_α-atoms of PTP1B were evaluated in an MD simulation to get a measure of the flexibility of the protein backbone, especially at the binding site region (Supplementary Figure 15). The backbone of the binding site region (residues 215-221) shows a little spike of flexibility but is in one of the least flexible areas of the protein.

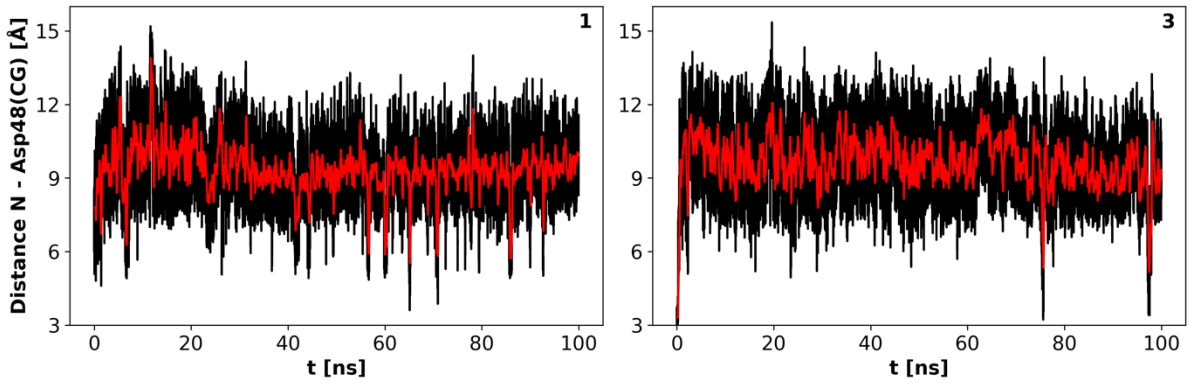
Ligand-protein key interactions throughout MD simulation

Both ligands remain stable inside the pocket throughout the 100 ns of simulation. Ligand-protein interactions found through docking were evaluated during molecular dynamics simulations of ligands **1** and **3** in complex with PTP1B.



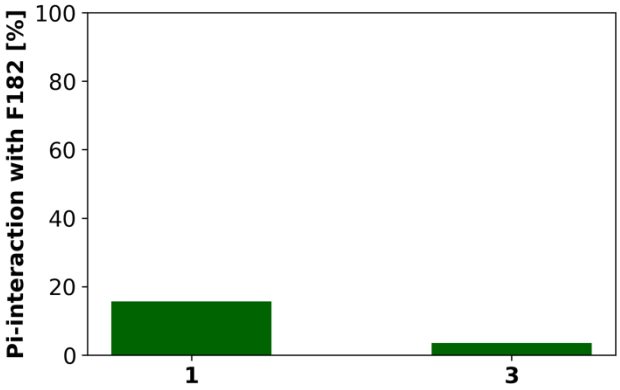
Supplementary Figure 35. Distance between phosphorus of ligand headgroup of **1** (left) and **3** (right). The moving average with a sliding window of 200 frames is shown in red

For both complexes it can be seen that the headgroup phosphorus remains close to the sidechain of Arg221, here indicated by sidechain carbon CZ. The distance is smaller and shows less fluctuation for ligand **1** than for ligand **3**.



Supplementary Figure 36. Distance between backbone nitrogen of ligands **1** (left) and **3** (right). The moving average with a sliding window of 200 frames is shown in red

The distance between the ligands' positively charged backbone nitrogen and the Asp48 sidechain is too large for an intact salt bridge. This is already the case at the beginning of the simulation, after equilibration. It can be explained by the backbone ammonium ion turning outward and thereby preferring solvent exposure over the formation of the salt bridge.

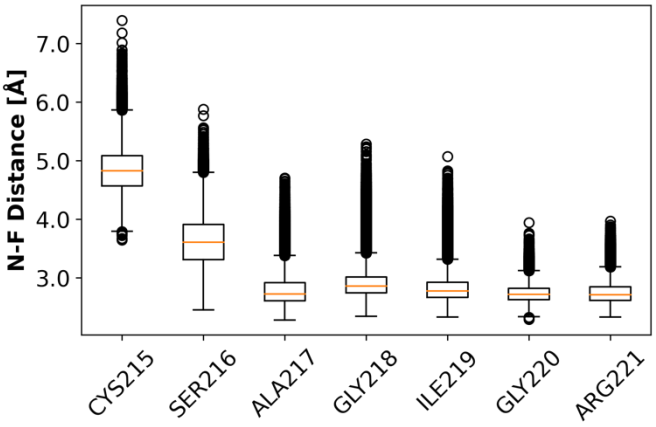


Supplementary Figure 37. Percentage of frames in which a pi-interaction between the aromatic rings of the ligand and Phe182 of PTP1B is present

We do not observe π-π interaction of any ligand with Tyr46. We observe significantly more π-π interaction with Phe182 for **1** compared to **3**, but for both ligands this interaction occurs rarely.

The backbone NH groups remain close to the fluorines of the PF_5 moiety, which implies the presence of N-H..F interactions. Supplementary Figure 19 shows the distribution of N..F distances

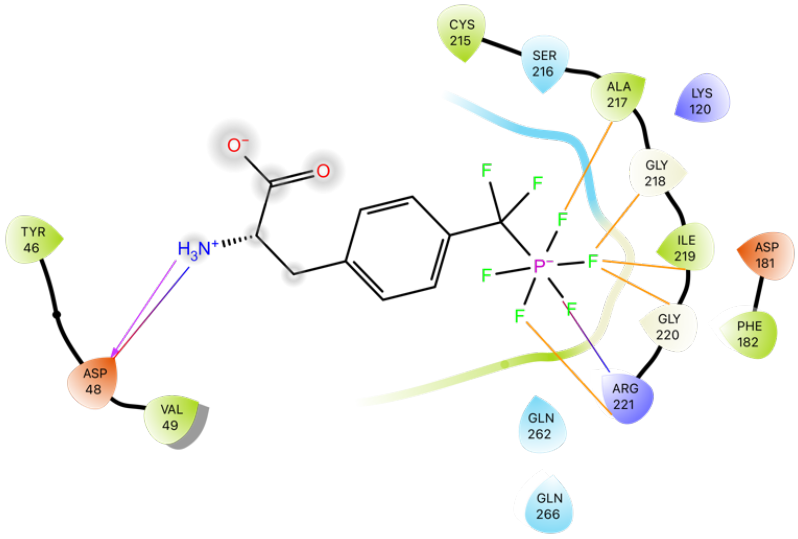
inside the binding pocket. Specifically, the figure contains the distance of the respective backbone nitrogen to the closest fluorine of the PF_5 moiety. For residues 217-221, the closest N..F distance can be stated to be generally below 3 Å.



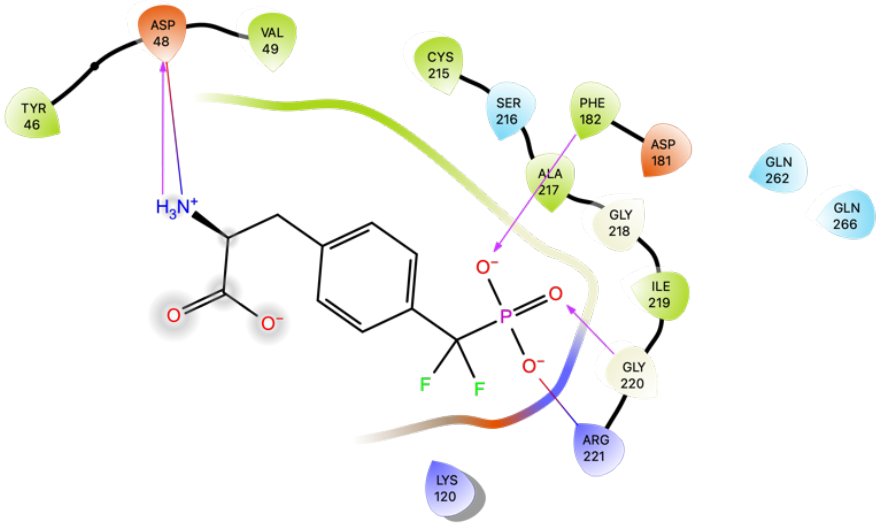
Supplementary Figure 38. Distance of backbone nitrogen of respective residue to the closest fluorine of the PF_5 moiety

Ligand-protein interaction in the PTP1B binding pocket

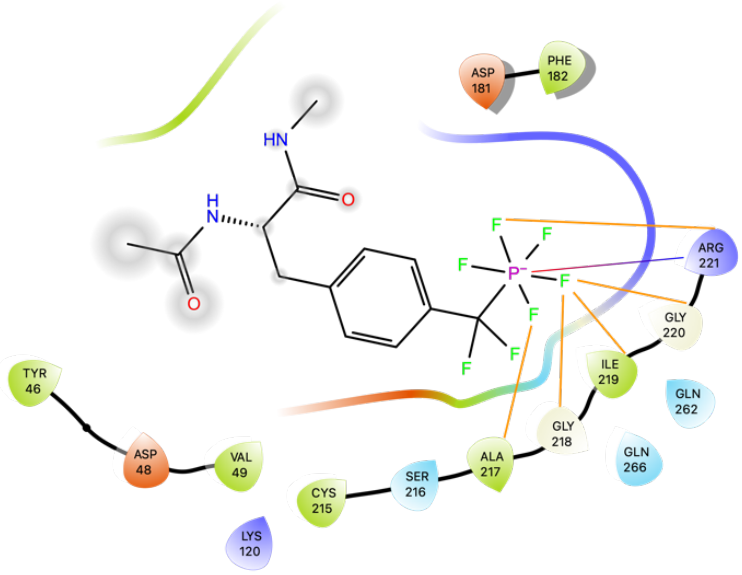
Figures 20-23 show the interactions of **1** and **3** as well as of their ACE/NME capped counterparts when bound to the pocket of PTP1B. The interaction was analysed using the final docking poses. For all ligands salt bridges are identified between phosphorus head group and Arg221, and between backbone nitrogen and Asp48. Hydrogen bonding, indicated by pink arrows, is detected to the phosphate group of **1** and between the backbone amide of the uncapped amino acids and Asp48. The capped derivative of **1** also shows hydrogen bonds between its backbone amide and Asp48.



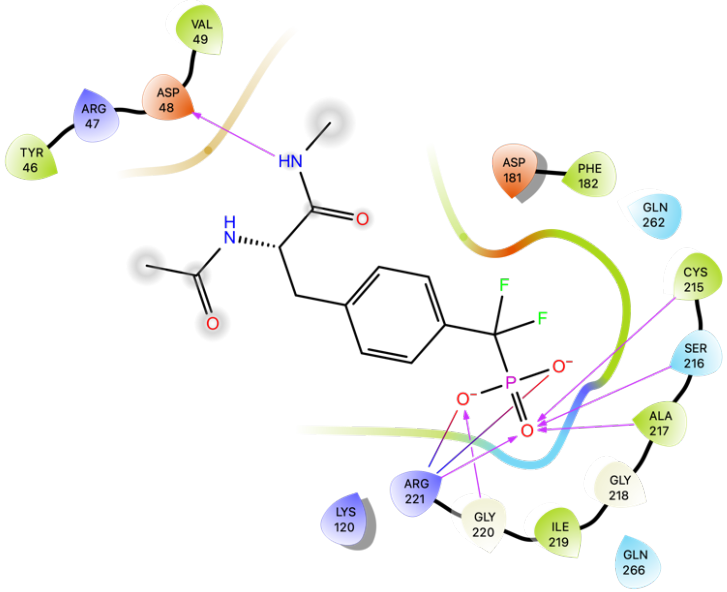
Supplementary Figure 39. Ligand interaction diagram of **3** in PTP1B binding pocket. Salt bridges are indicated by a straight line, hydrogen bonds by a pink arrow. The orange lines indicate contacts between backbone NH and F of the PF_5 moiety. Grey shaded atoms are solvent exposed. The binding pocket is represented by a line around the ligand with the color matching the closest amino acid



Supplementary Figure 40.2 Ligand interaction diagram of **1** in PTP1B binding pocket. Salt bridges are indicated by a straight line, hydrogen bonds by a pink arrow. Grey shaded atoms are solvent exposed. The binding pocket is represented by a line around the ligand with the color matching the closest amino acid

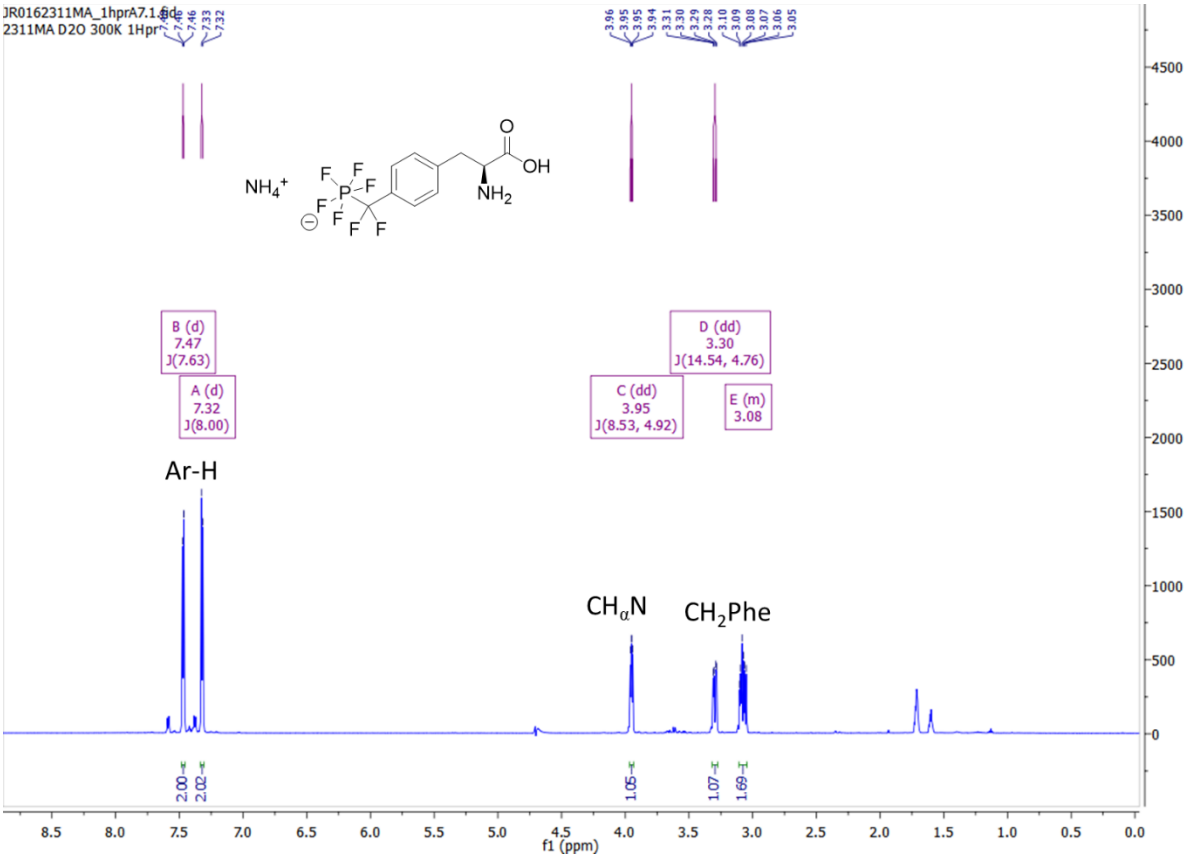


Supplementary Figure 41. Ligand interaction diagram of capped **10** in PTP1B binding pocket. Salt bridges are indicated by a straight line, hydrogen bonds by a pink arrow. The orange lines indicate contacts between backbone NH and F of the PF_5 moiety. Grey shaded atoms are solvent exposed. The binding pocket is represented by a line around the ligand with the color matching the closest amino acid

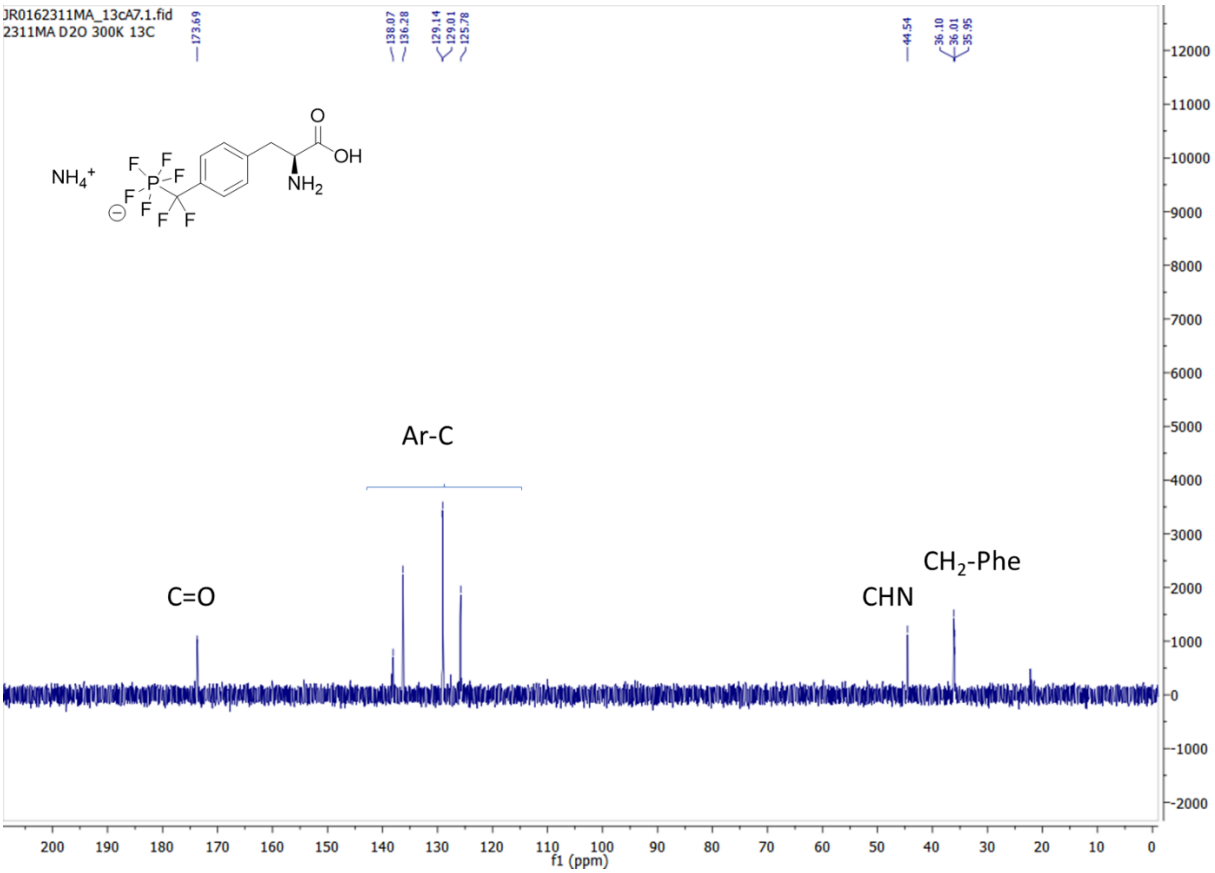


Supplementary Figure 42. Ligand interaction diagram of capped **1** in PTP1B binding pocket. Salt bridges are indicated by a straight line, hydrogen bonds by a pink arrow. Grey shaded atoms are solvent exposed. The binding pocket is represented by a line around the ligand with the color matching the closest amino acid

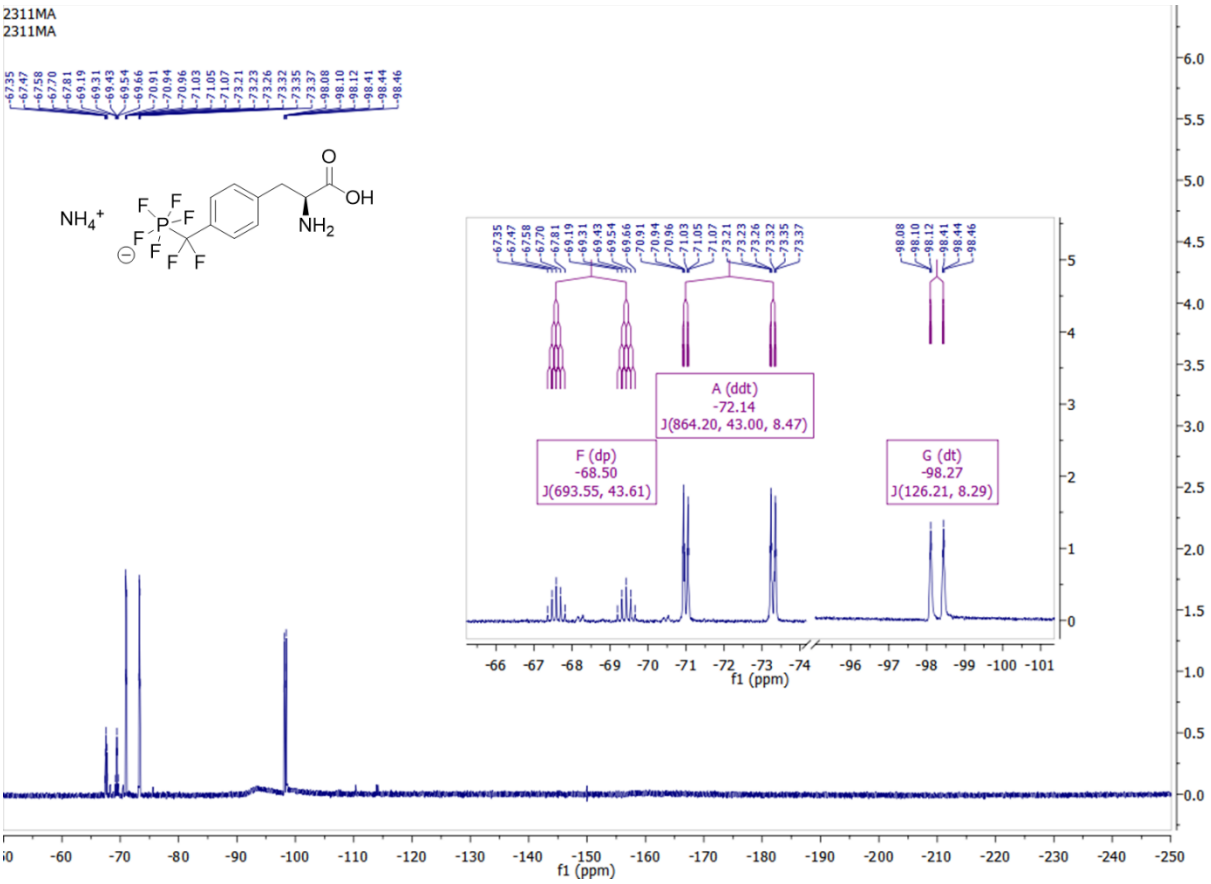
NMR and UV/vis spectra, HPLC chromatograms



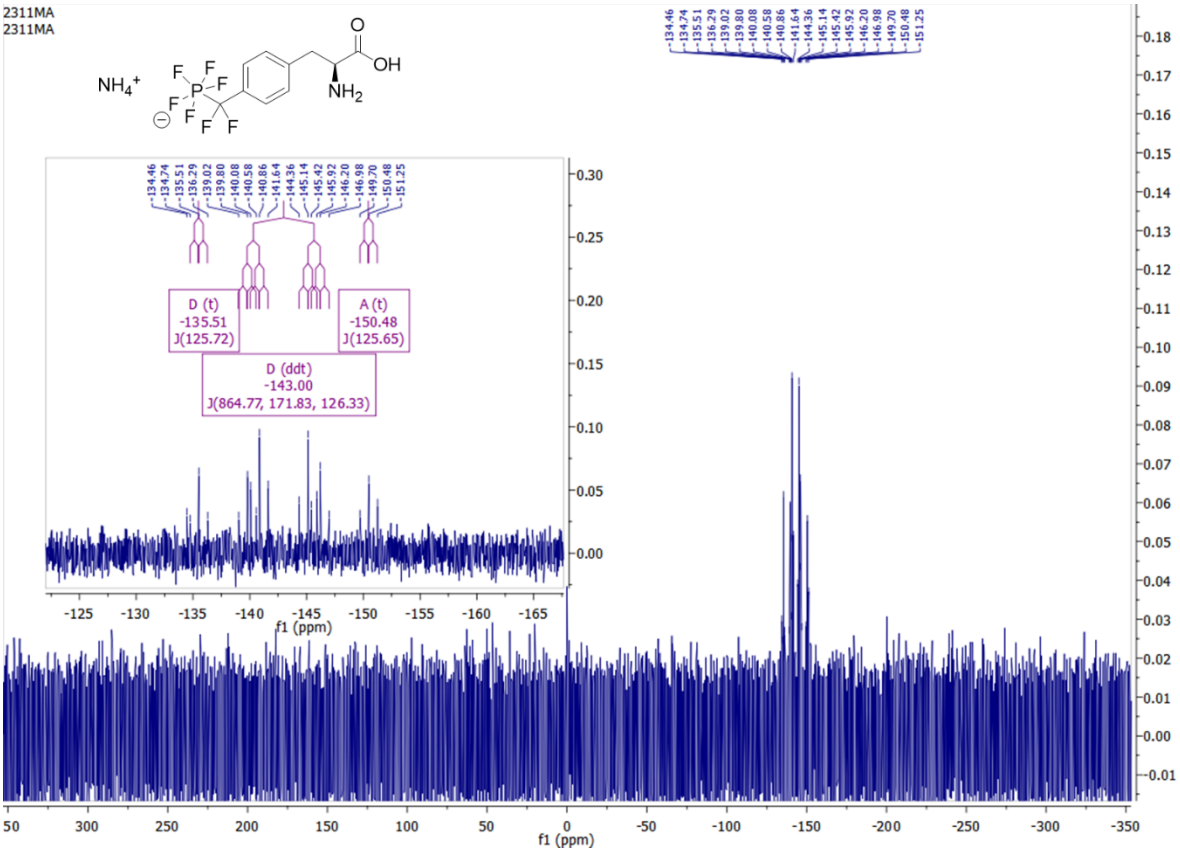
Supplementary Figure 43. ^{1}H NMR spectrum (700 MHz, D_2O) of **3**



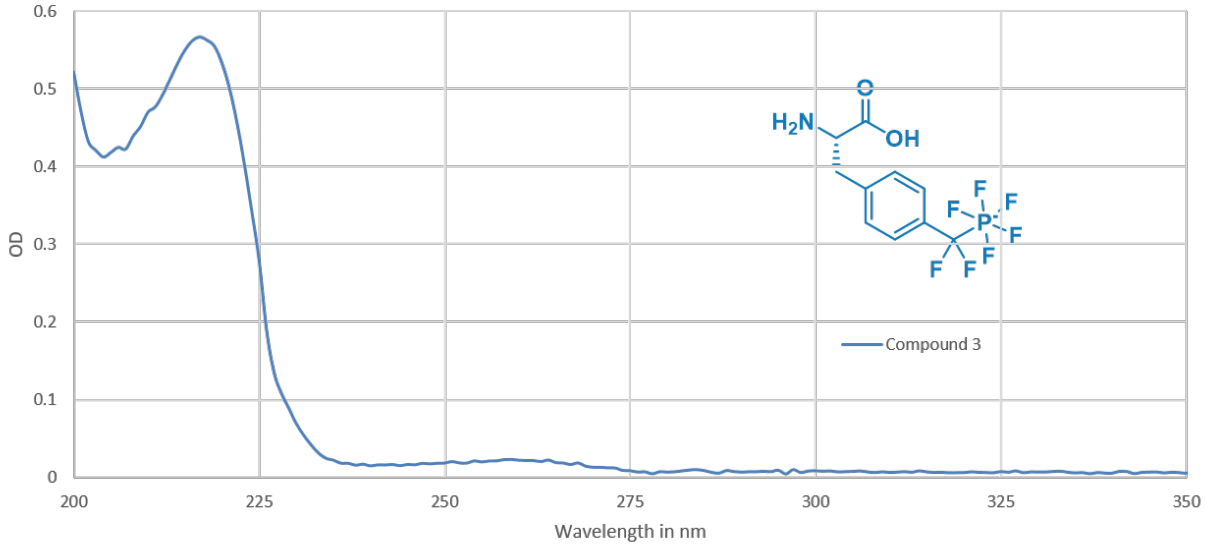
Supplementary Figure 44. ^{13}C NMR spectrum (176 MHz, D_2O) of **3**



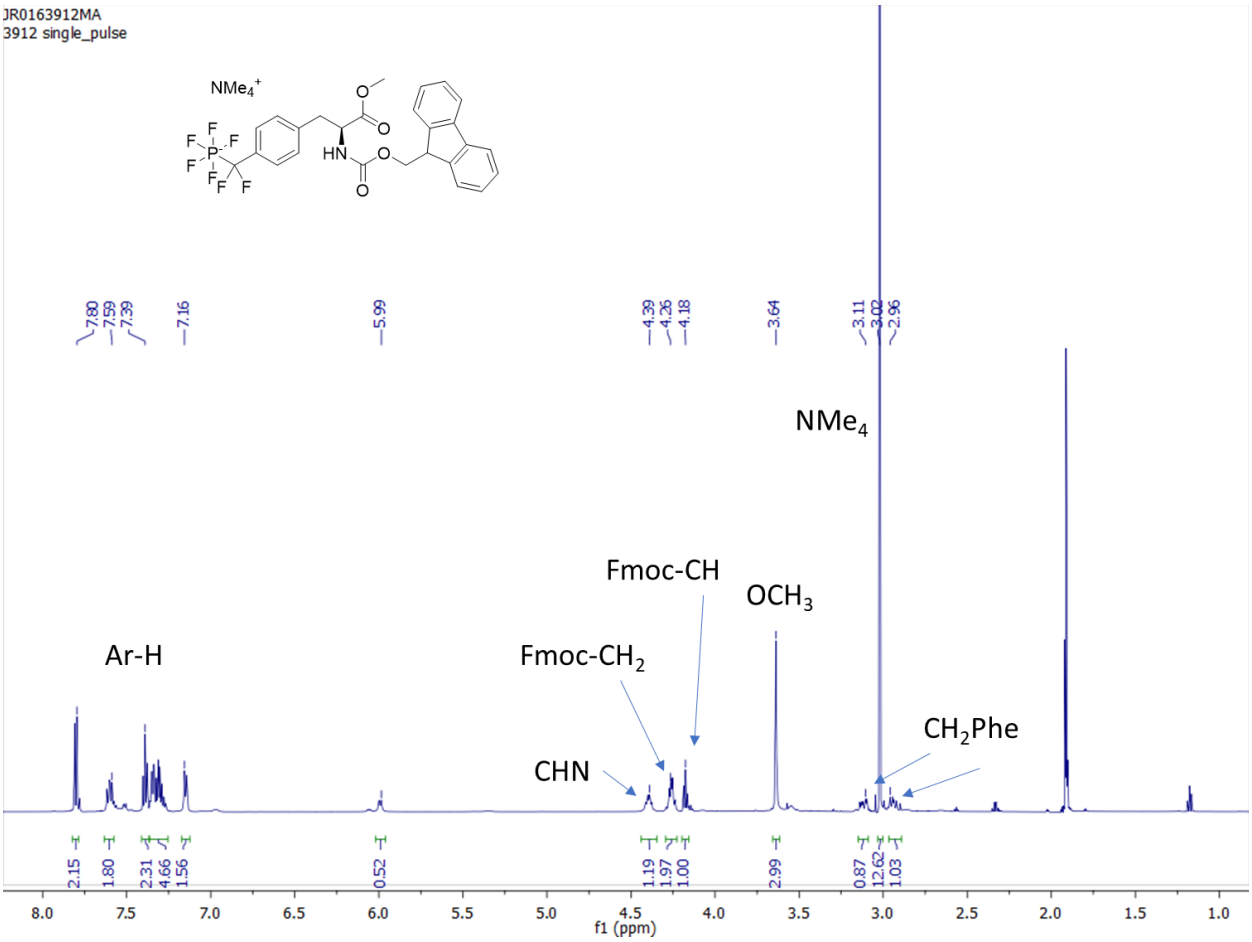
Supplementary Figure 45. ^{19}F NMR spectrum (376 MHz, D_2O) of **3**



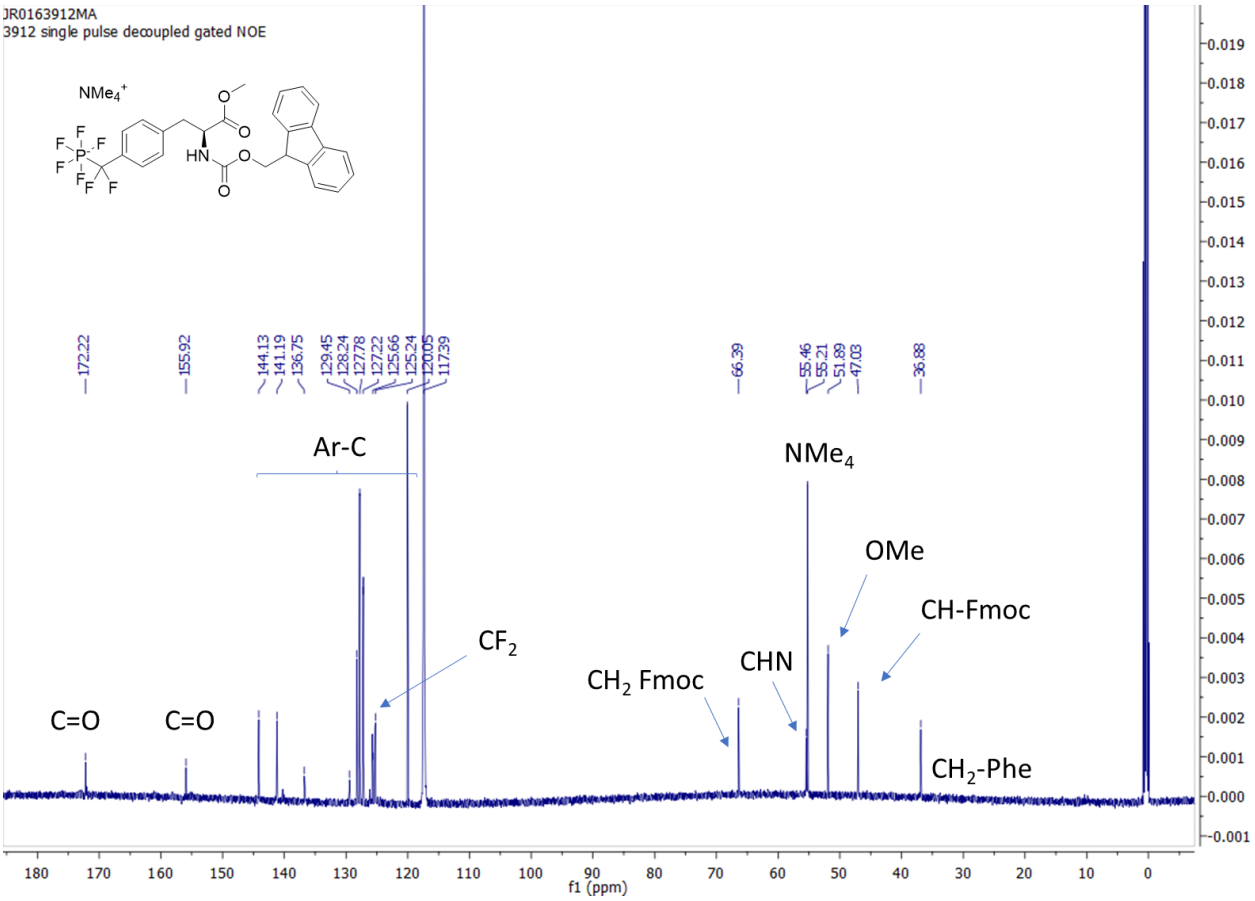
Supplementary Figure 46. ^{31}P NMR spectrum (161 MHz, D_2O) of **3**



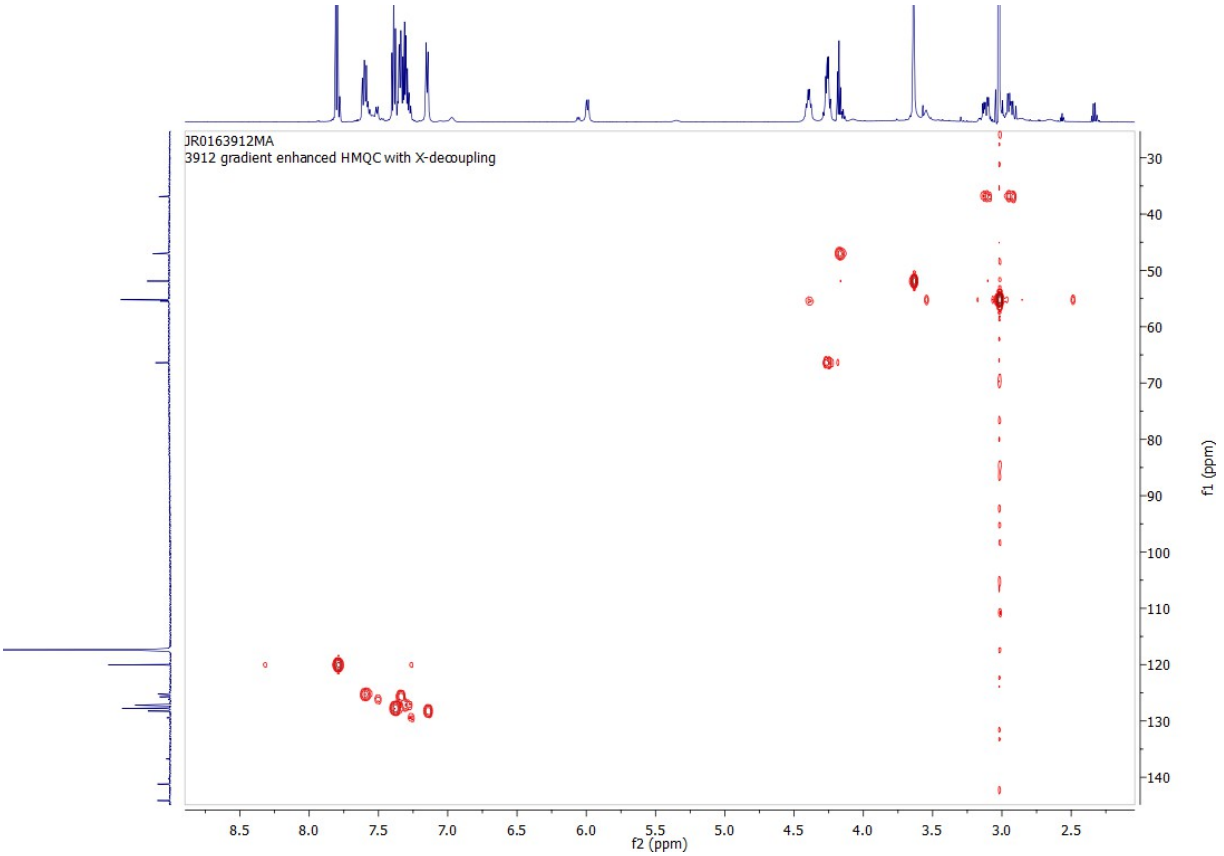
Supplementary Figure 47: UV spectrum of 0.5 mM solutions of compound **12** in water



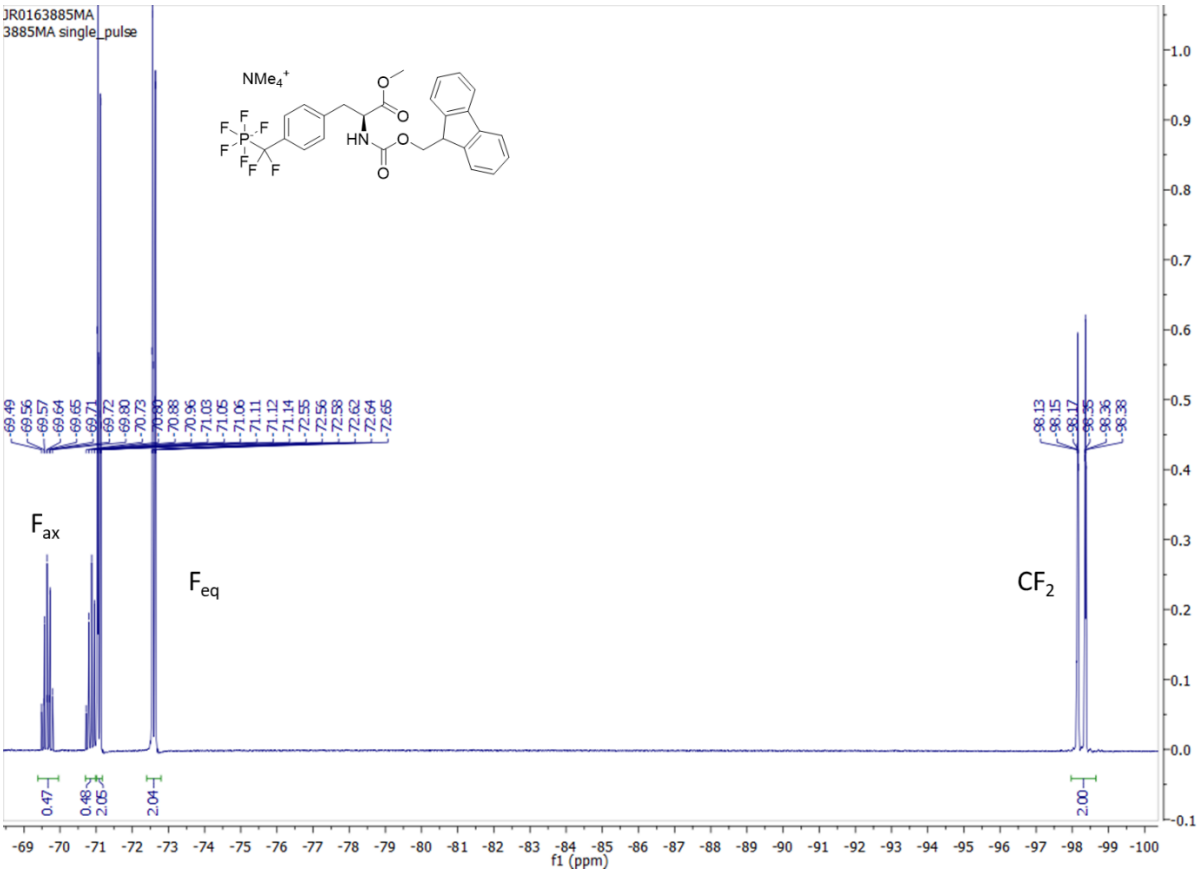
Supplementary Figure 48. ^{1}H NMR spectrum (600 MHz, ACN-d_3) of **8**



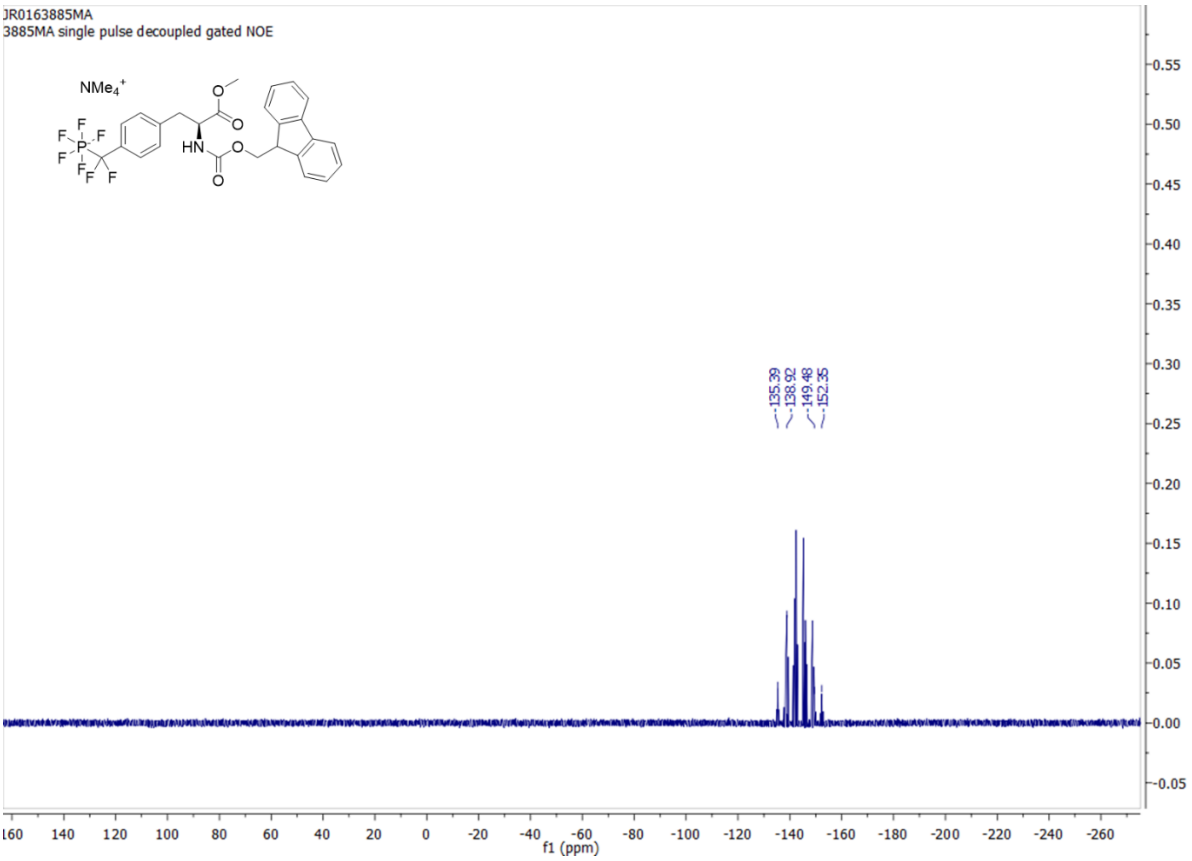
Supplementary Figure 49. ^{13}C NMR spectrum (151 MHz, ACN-d_3) of **8**



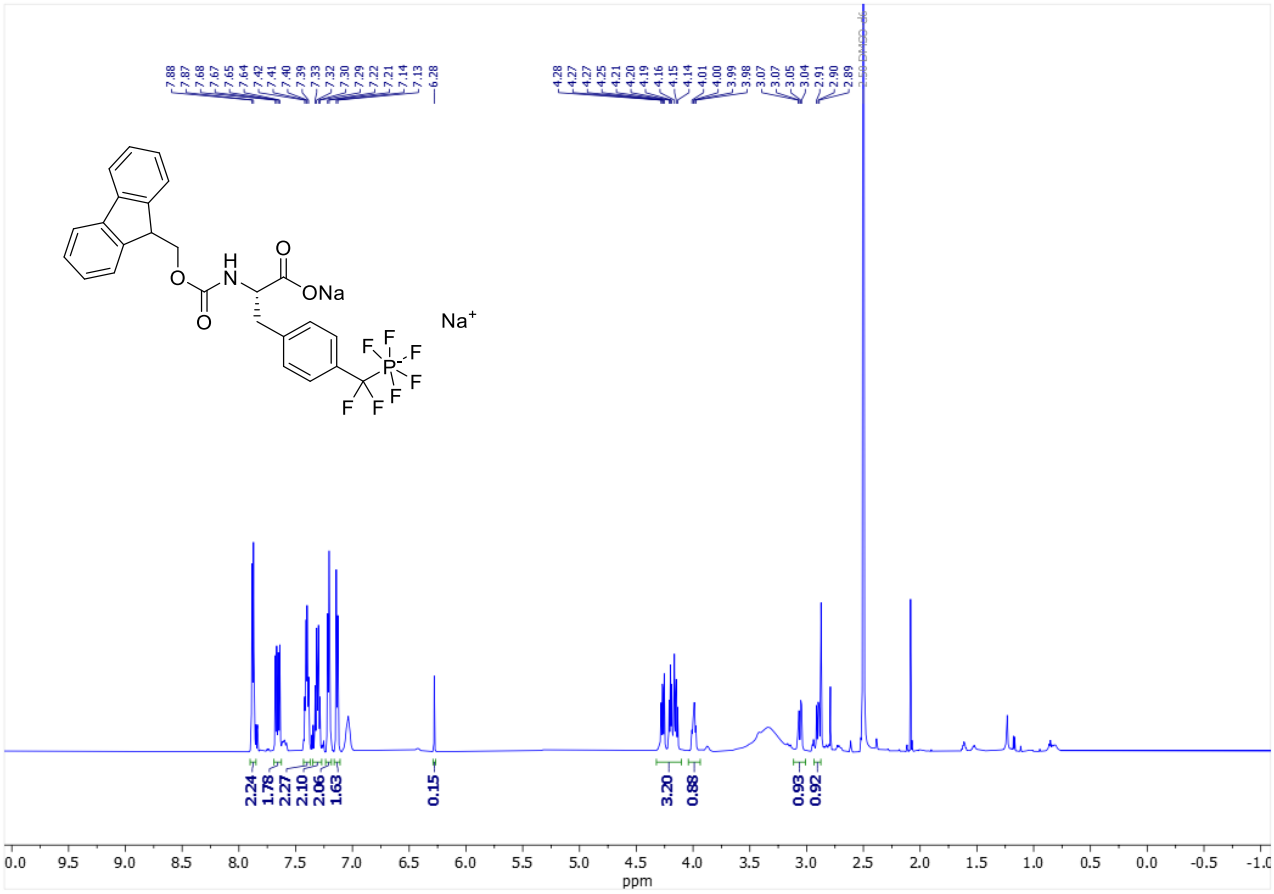
Supplementary Figure 50: HMQC spectrum (1H, 13C) in ACN-d_3 of **8**



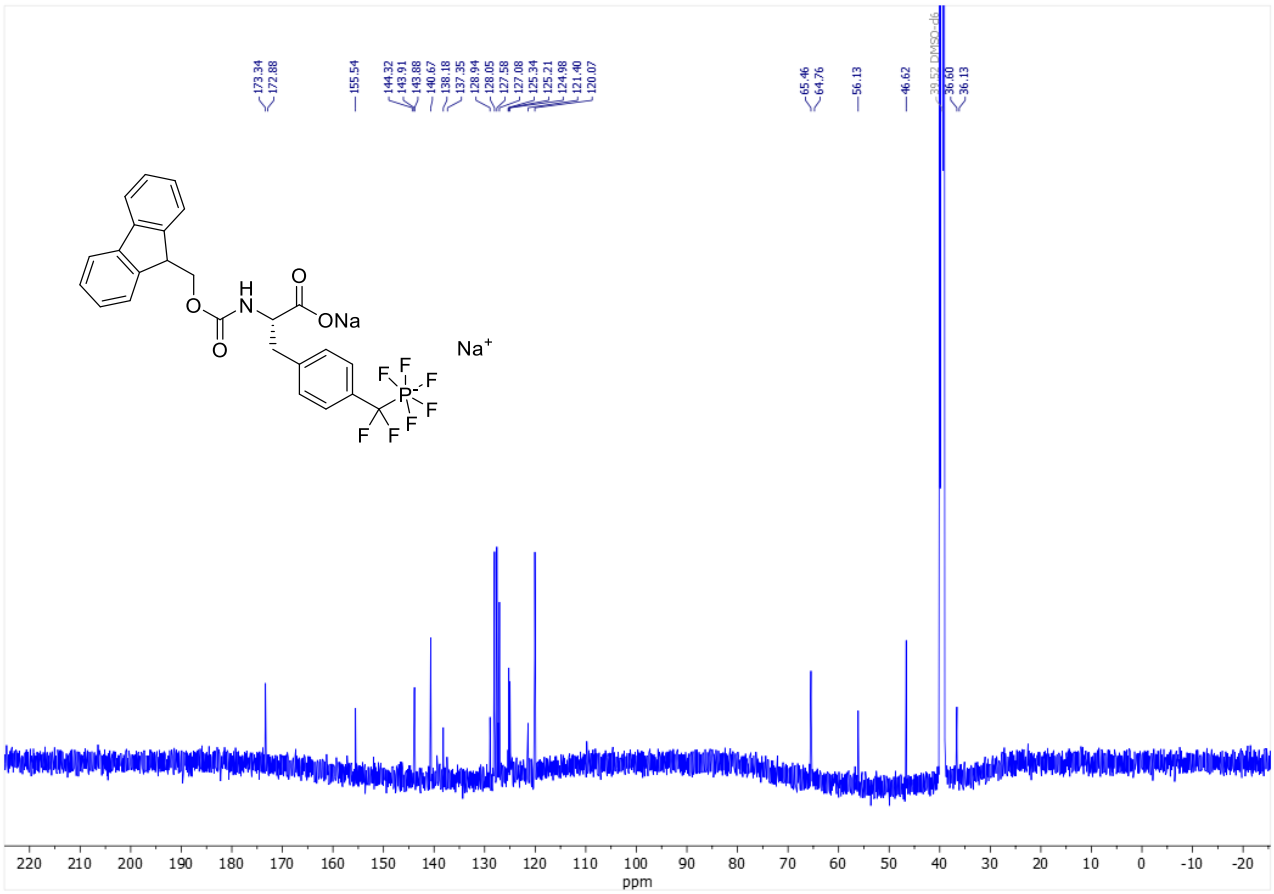
Supplementary Figure 51. ^{19}F NMR spectrum (565 MHz, ACN-d3) of **8**



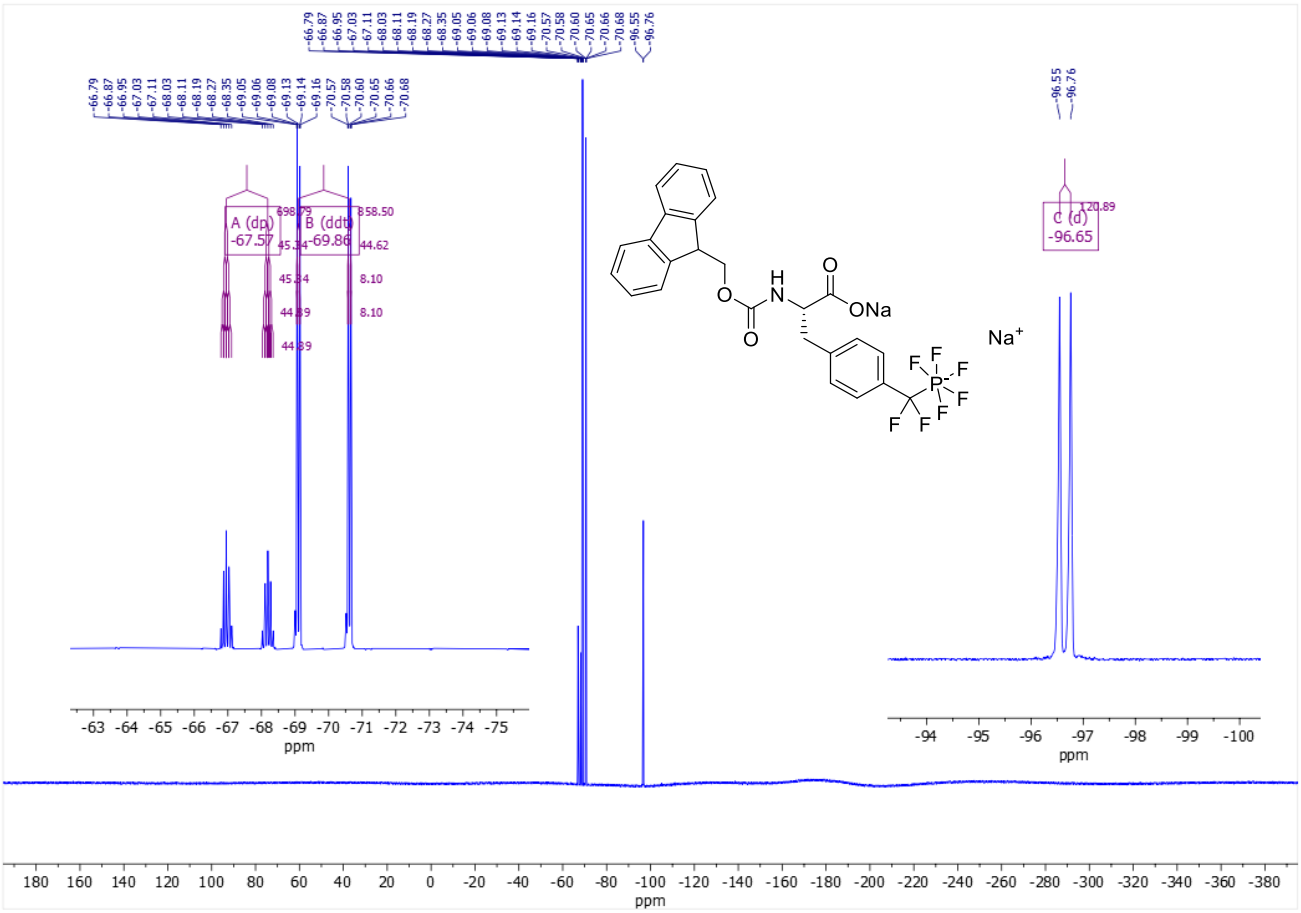
Supplementary Figure 52.3 ^{31}P NMR spectrum (243 MHz, ACN-d3) of **8**



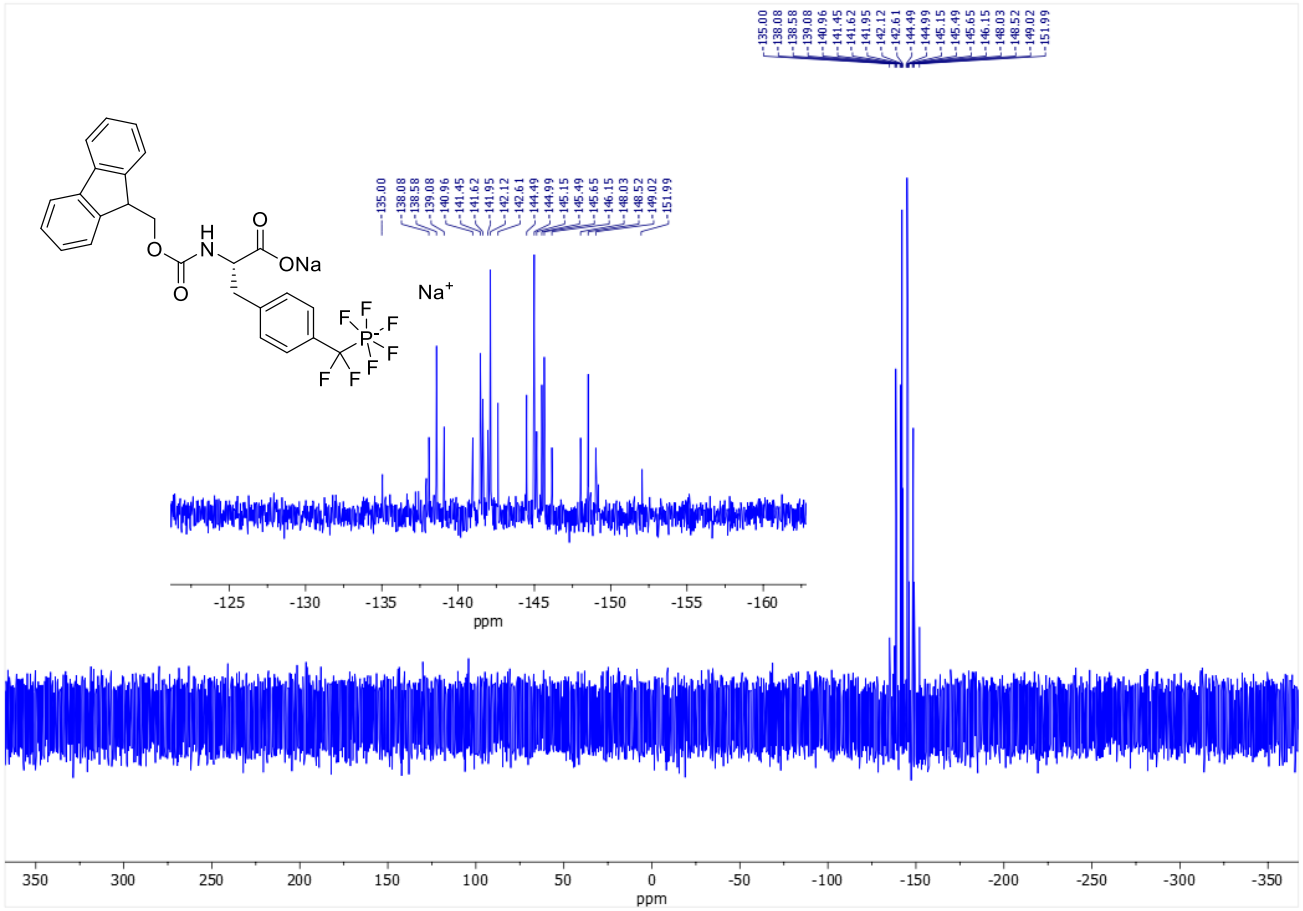
Supplementary Figure 53. ^{1}H NMR spectrum (600 MHz, DMSO-d6) of **9**



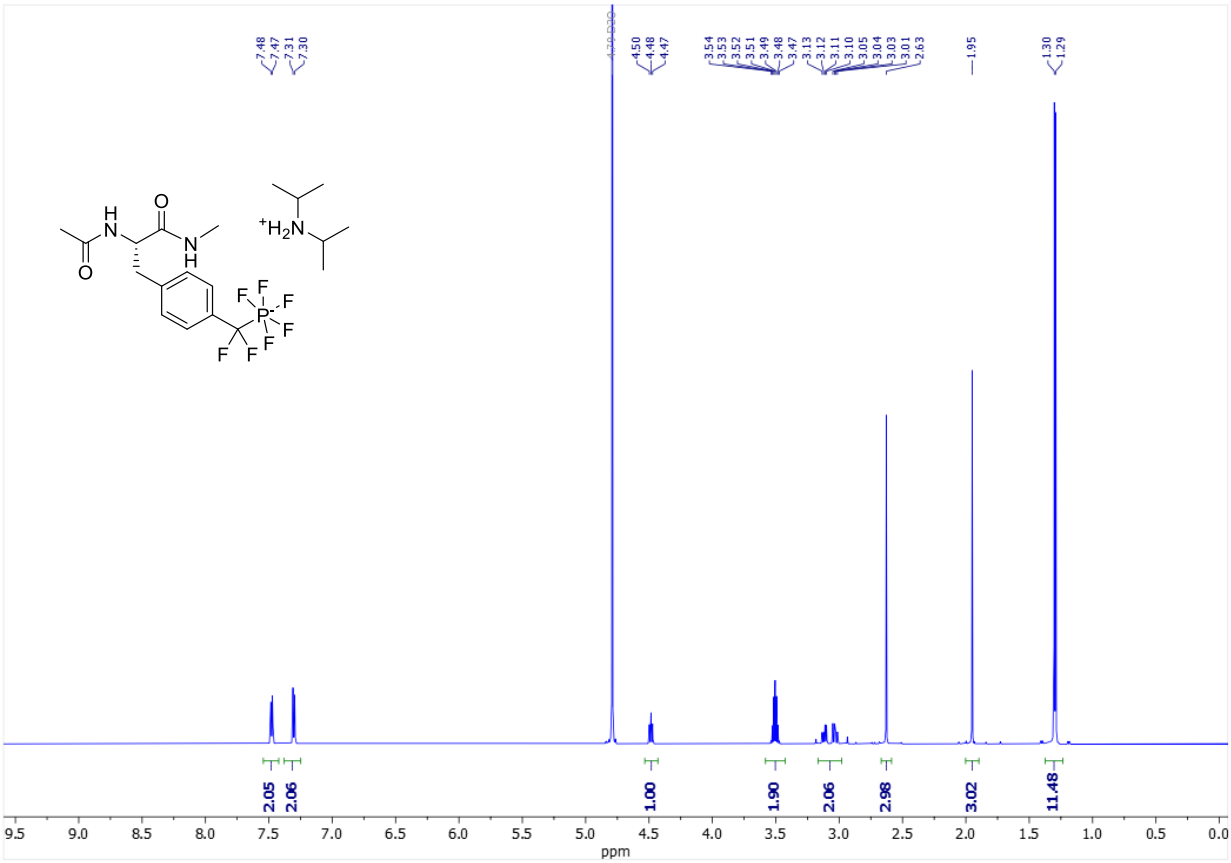
Supplementary Figure 54. ^{13}C NMR spectrum (151 MHz, DMSO-d6) of **9**



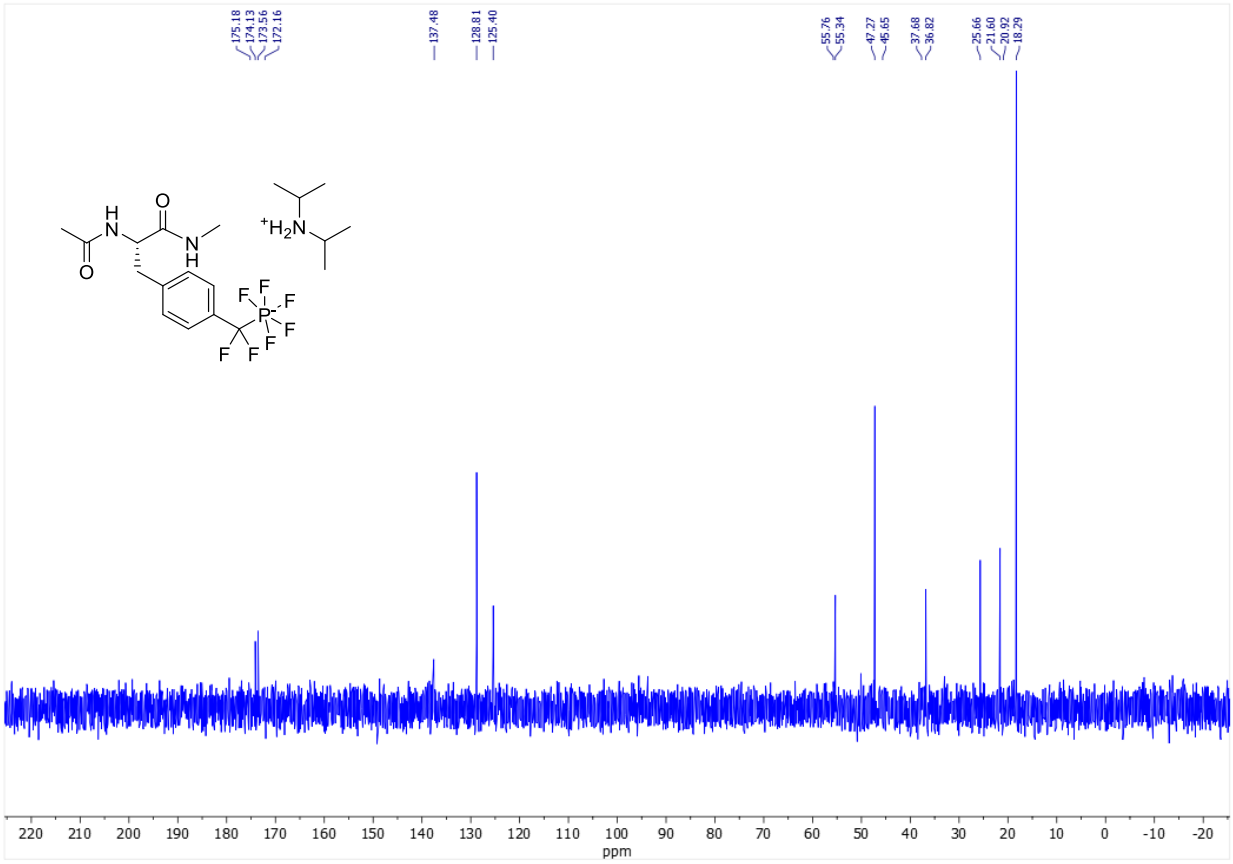
Supplementary Figure 55. ^{19}F NMR spectrum (565 MHz, DMSO-d6) of **9**



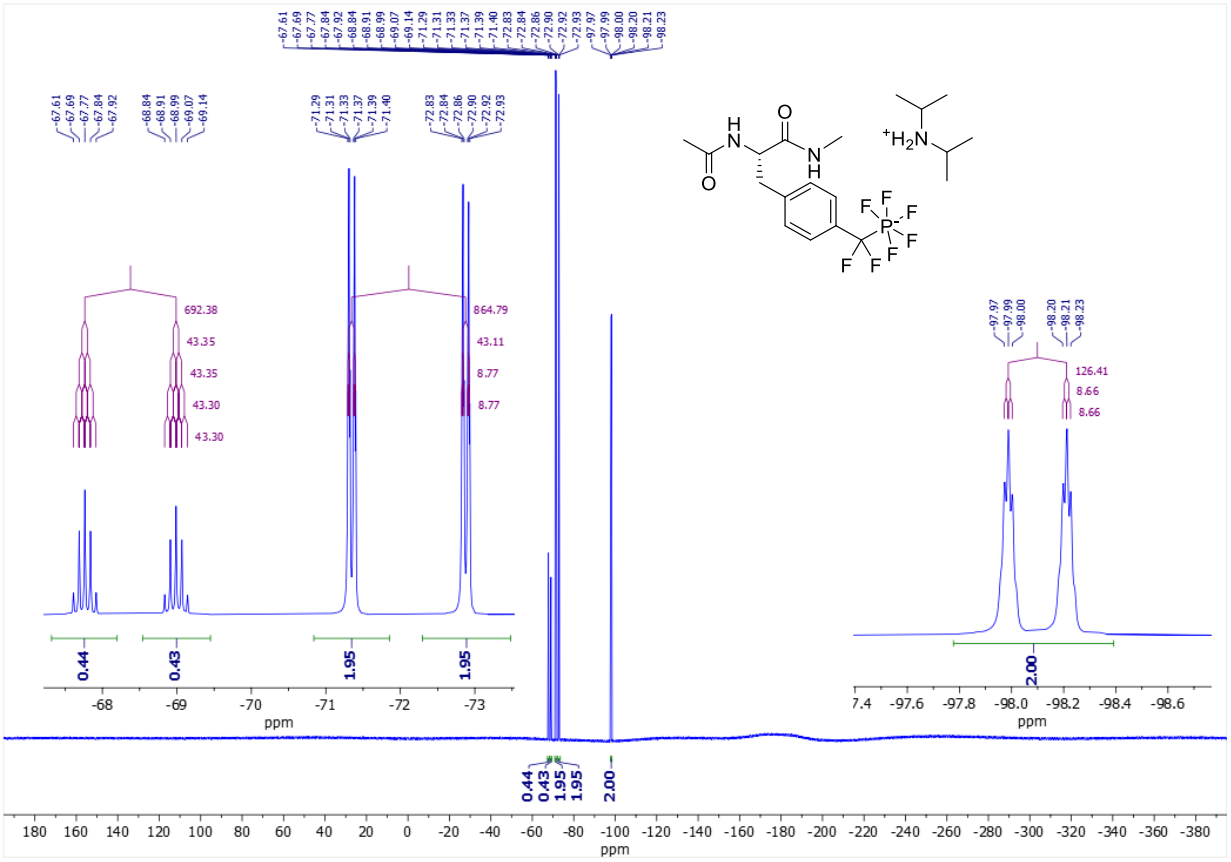
Supplementary Figure 56. ^{31}P NMR spectrum (243 MHz, DMSO-d6) of **9**



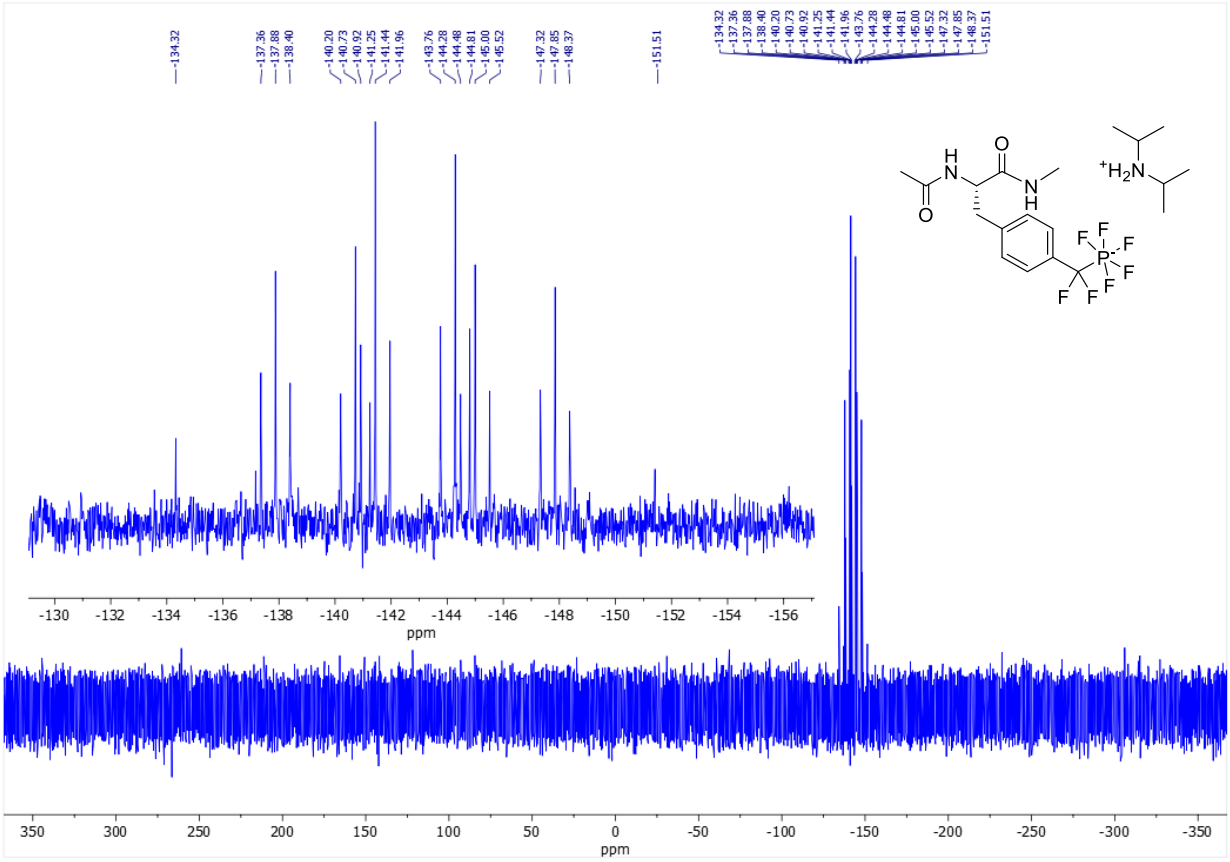
Supplementary Figure 57. ^{1}H NMR spectrum (600 MHz, D_2O) of **10**



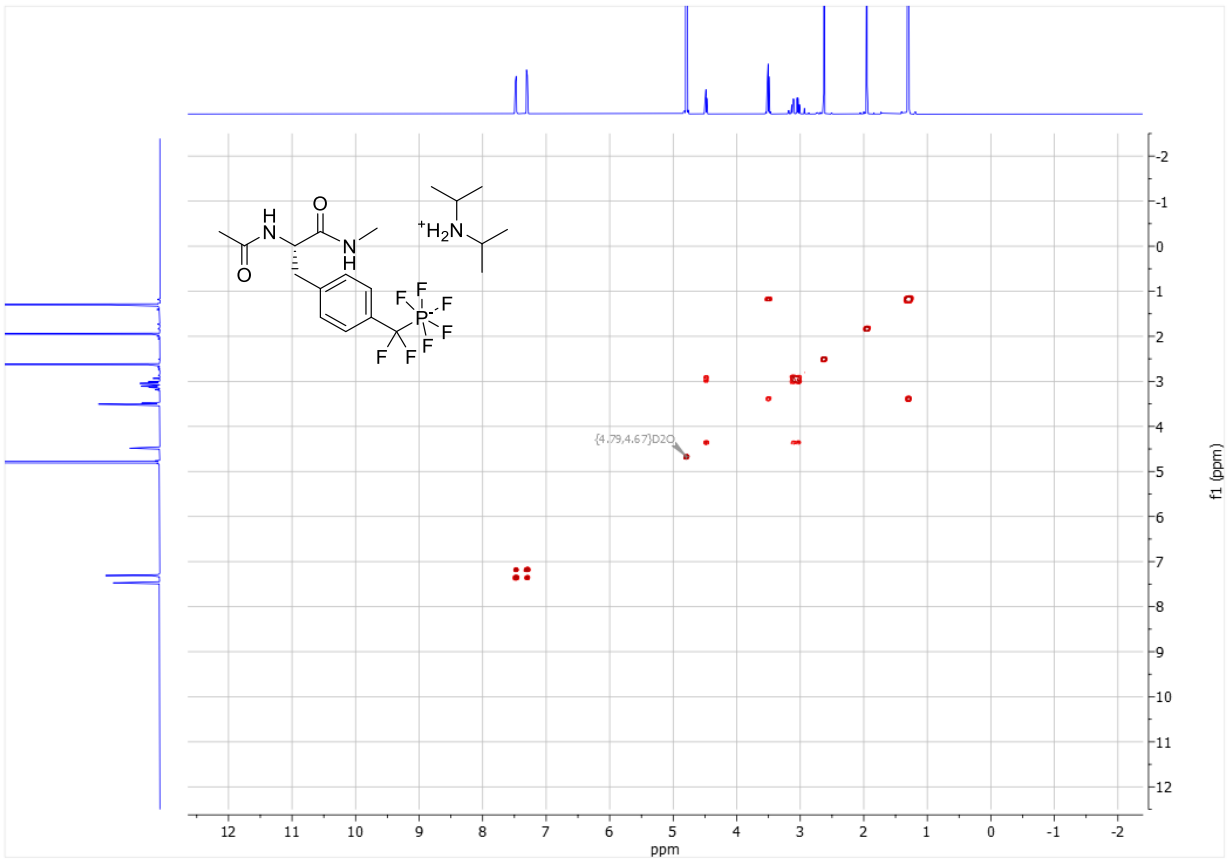
Supplementary Figure 58. ^{13}C NMR spectrum (151 MHz, ACN-d_3) of **10**



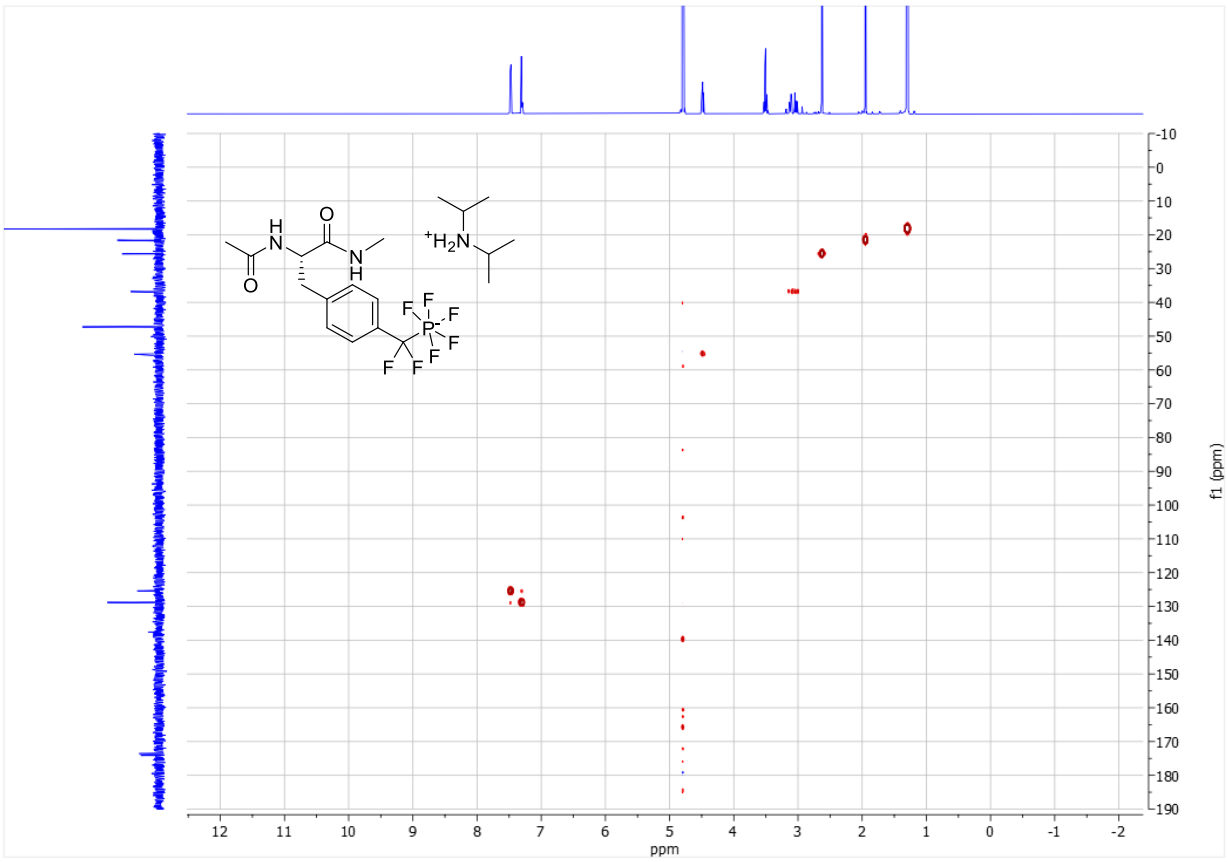
Supplementary Figure 59. ^{19}F NMR spectrum (565 MHz, D_2O) of **10**



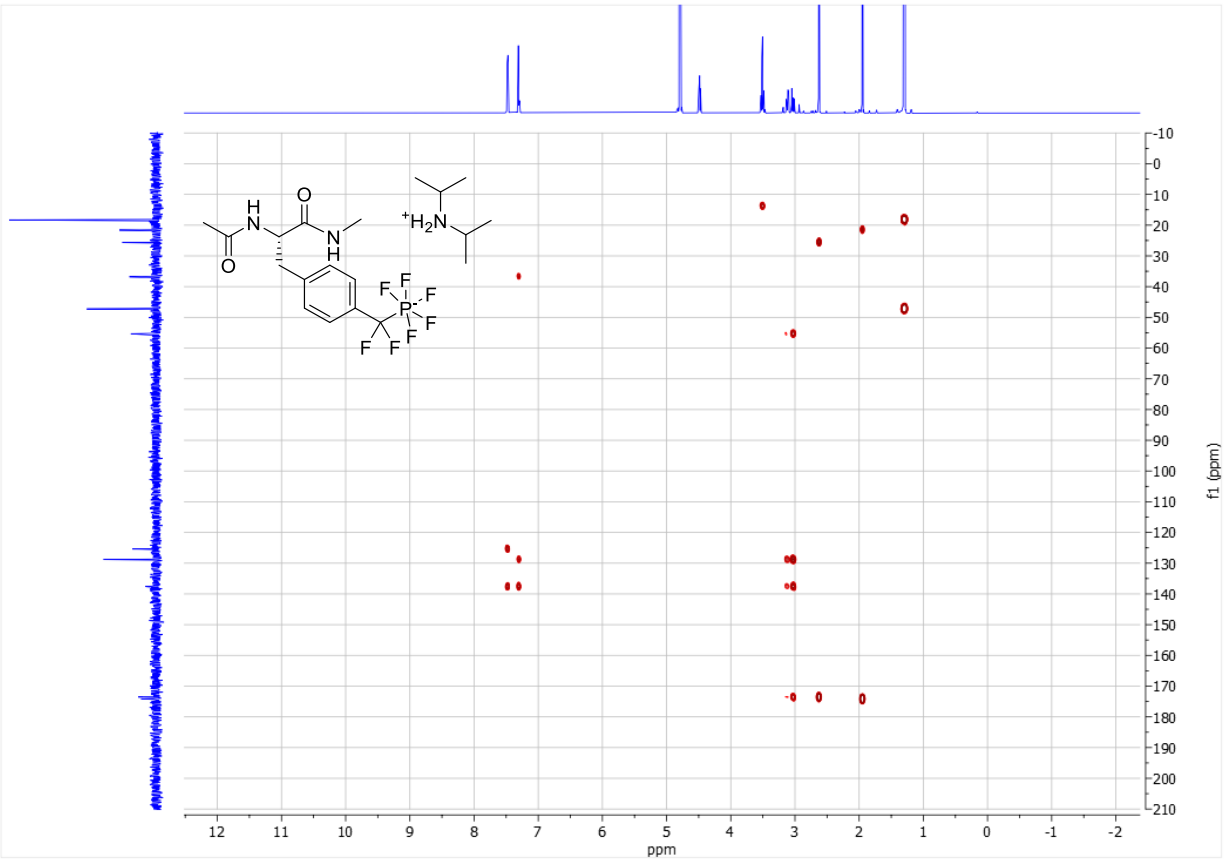
Supplementary Figure 60. ^{31}P NMR spectrum (243 MHz, D_2O) of **10**



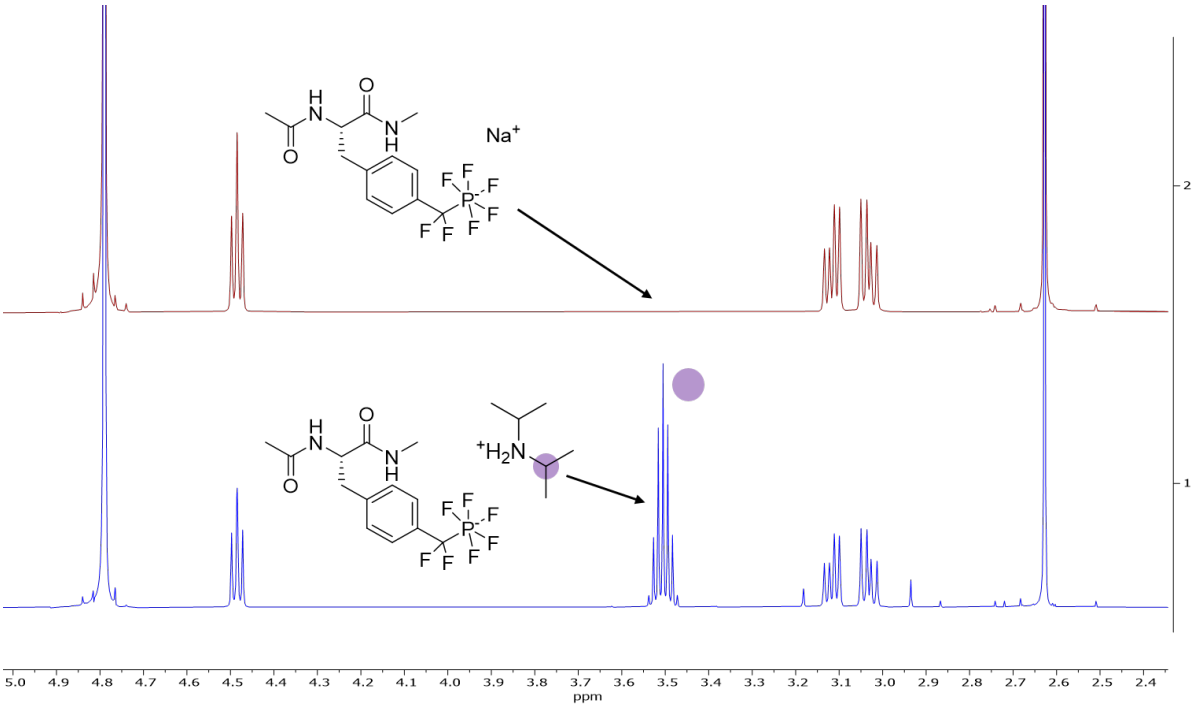
Supplementary Figure 61. H,H-COSY-NMR spectrum (600 MHz, D_2O) of **10**



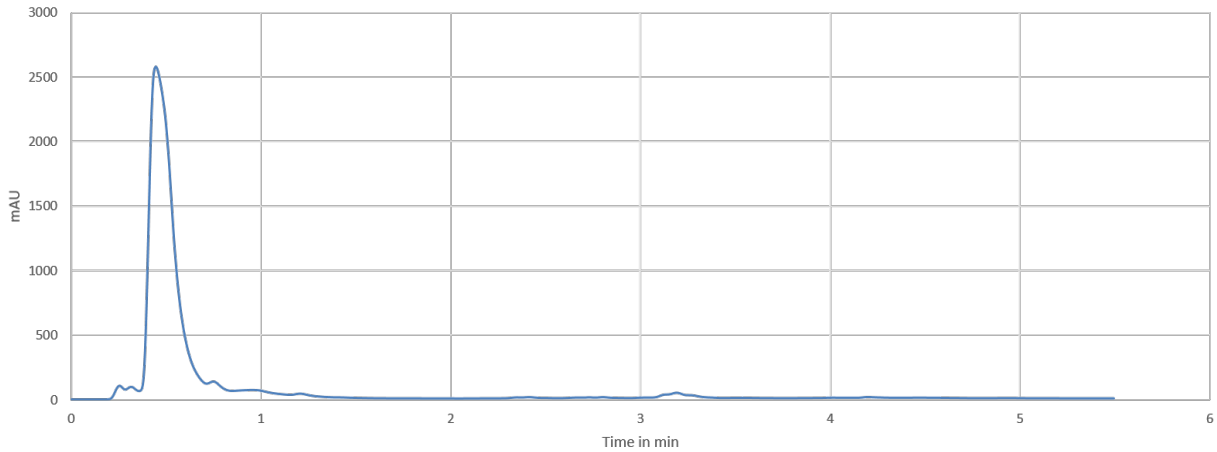
Supplementary Figure 62. HMQC spectrum (600 MHz, D_2O) of **10**



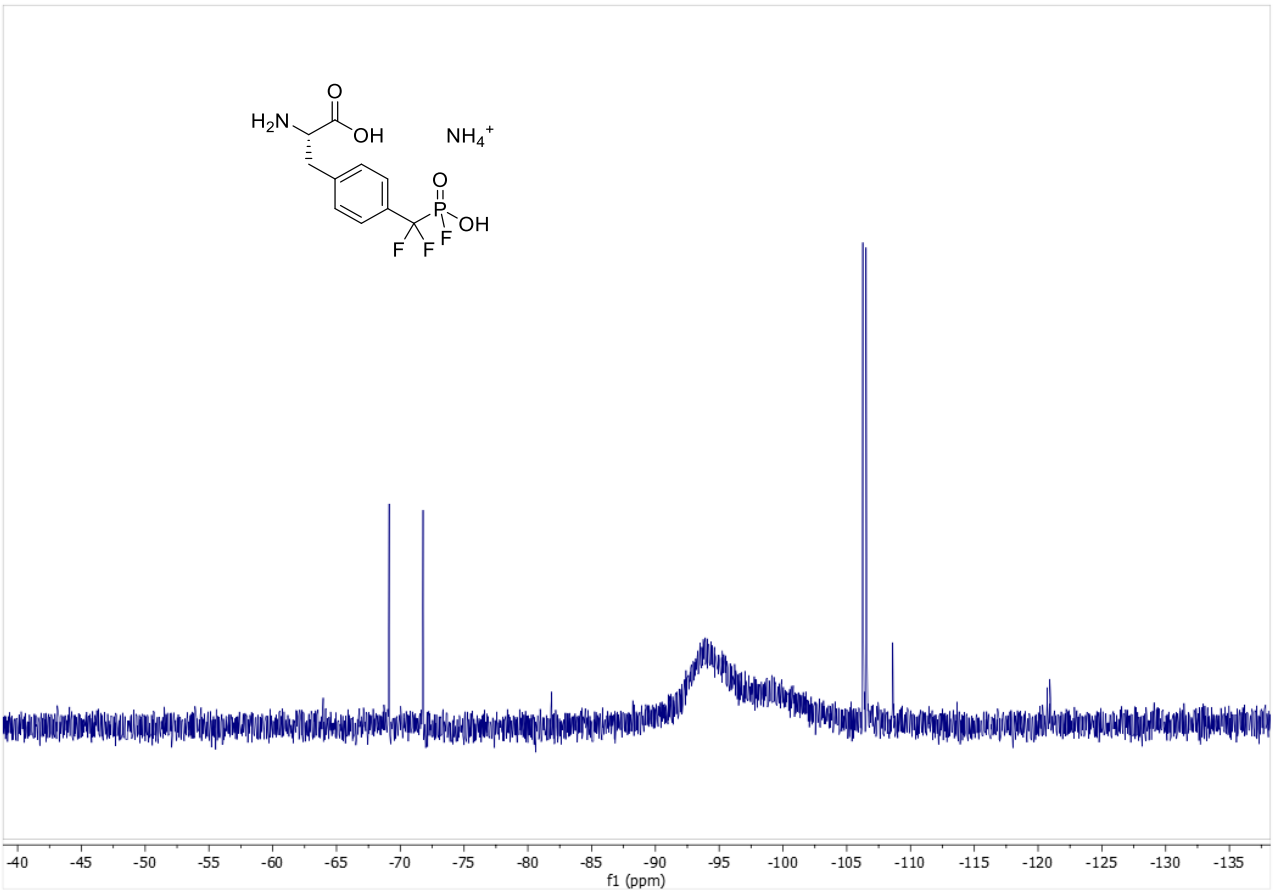
Supplementary Figure 63. HMBC (600 MHz, D_2O) of **10**



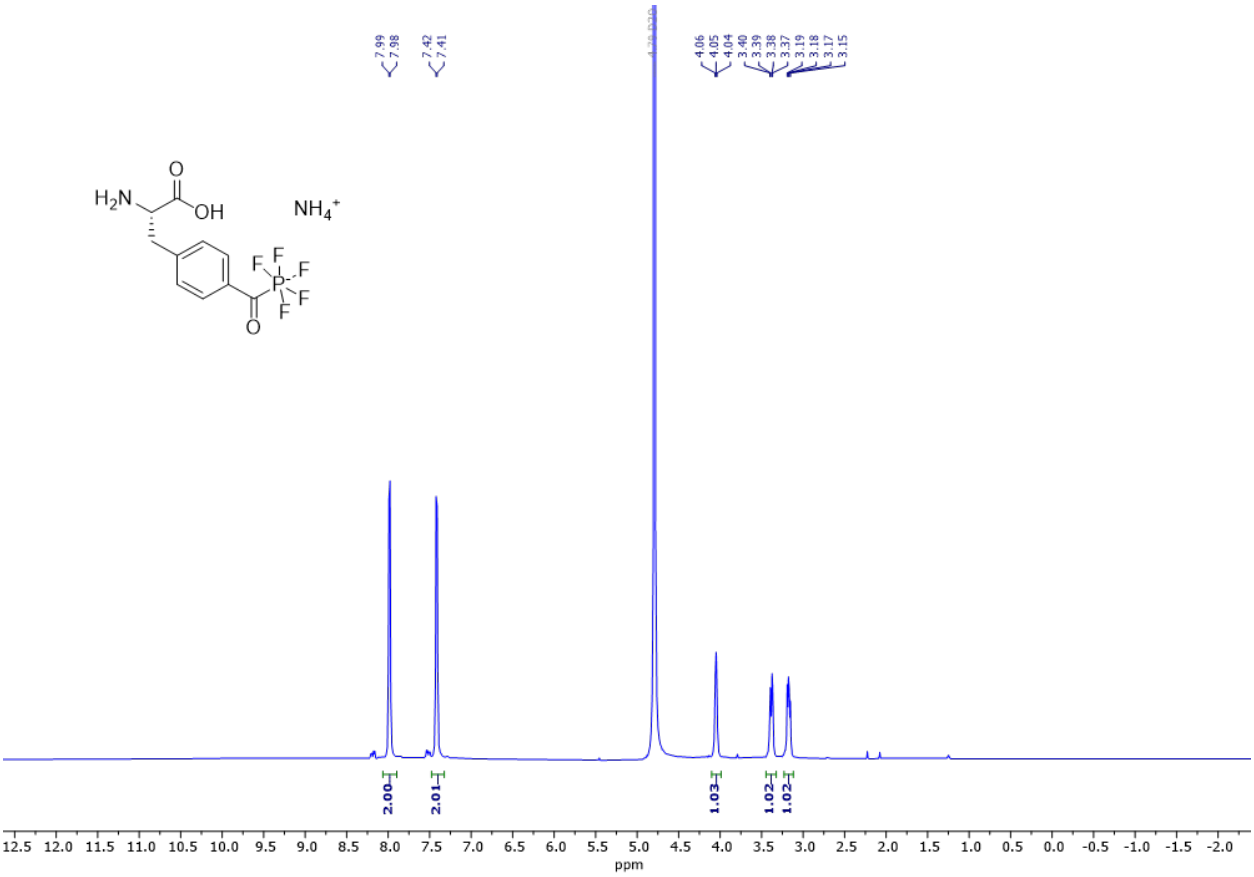
Supplementary Figure 64. ^{1}H NMR (600 MHz, D_2O) of **10**, before and after ion exchange



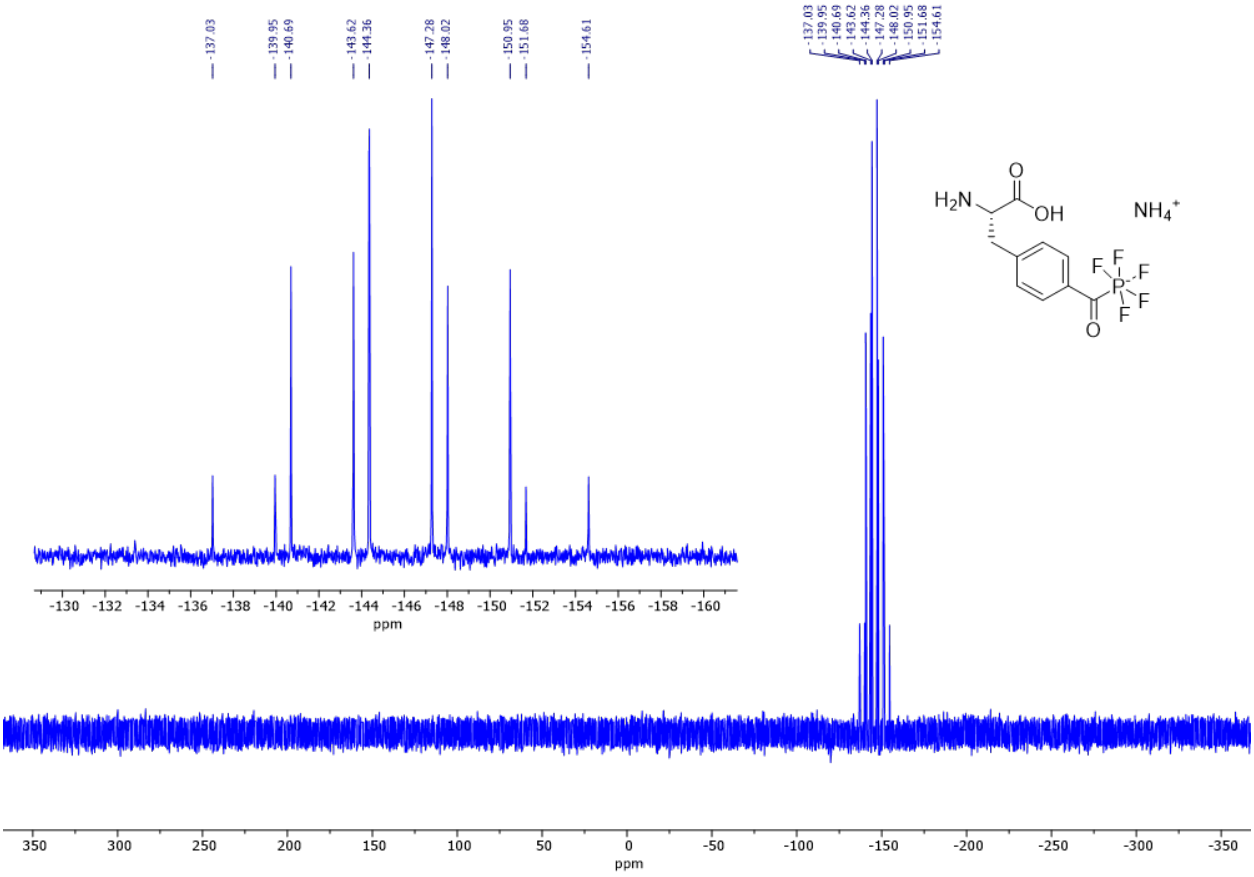
Supplementary Figure 65. HPLC chromatogram of compound **10**. Column B, eluent 15-95% ACN in 5 min, 210 nm.



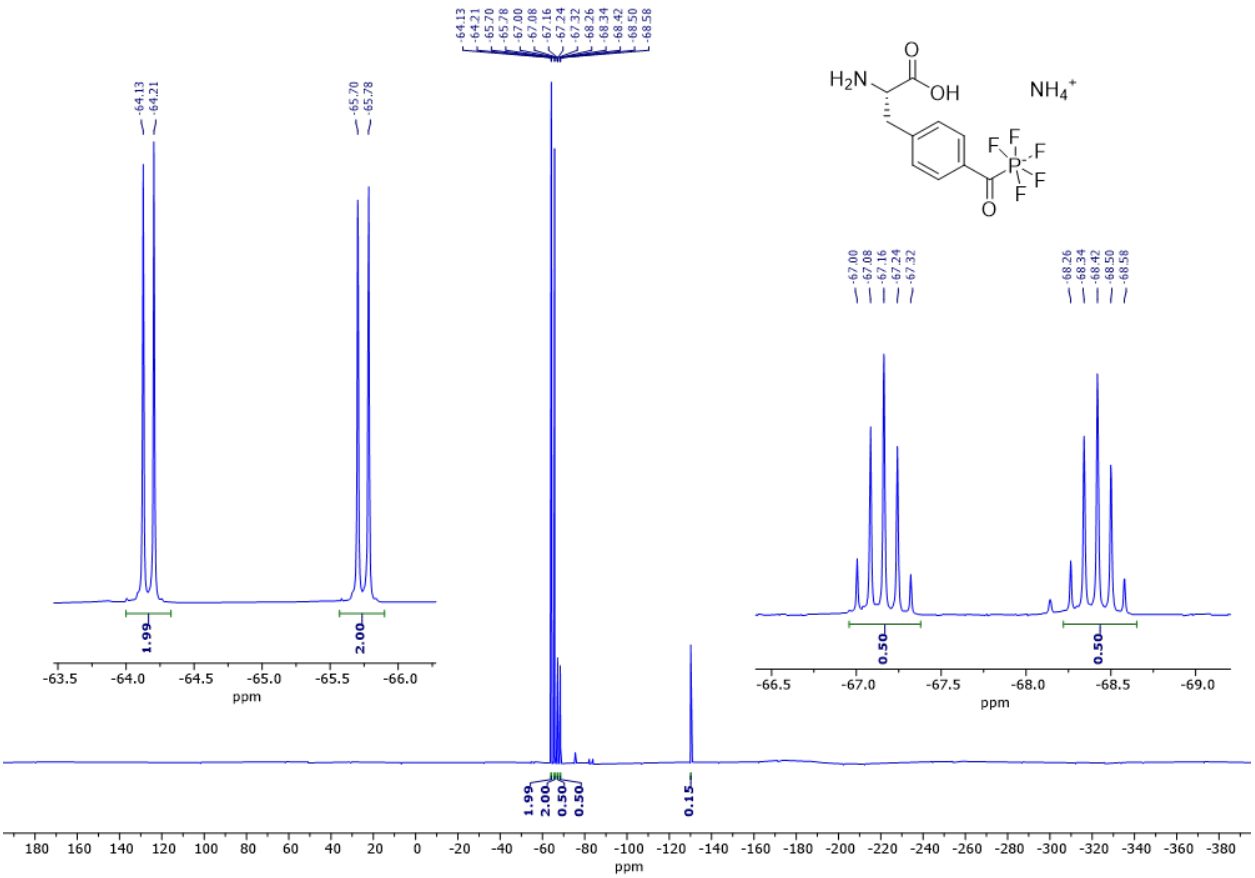
Supplementary Figure 66. ^{19}F NMR spectrum (565 MHz, D_2O) of **11**



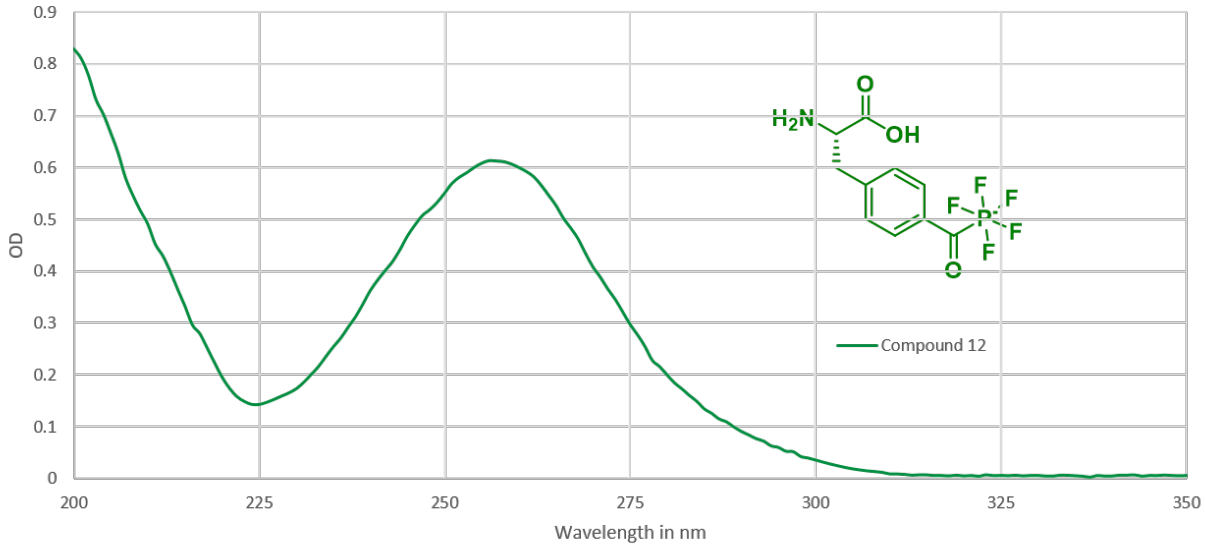
Supplementary Figure 67. ^{1}H NMR spectrum (600 MHz, D_2O) of **12**



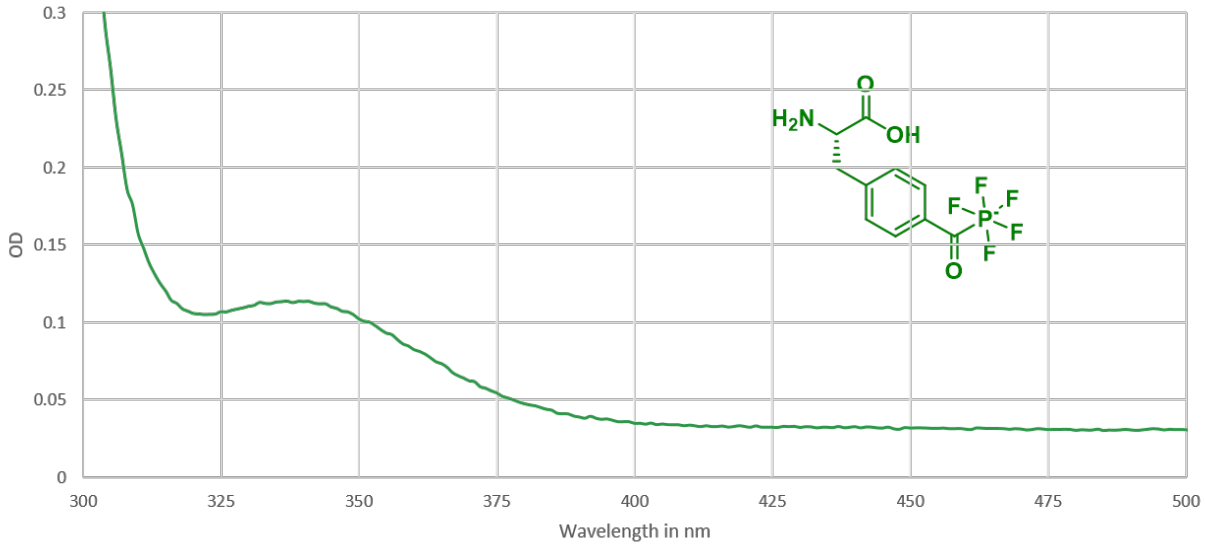
Supplementary Figure 68. ^{31}P NMR spectrum (243 MHz, D_2O) of **12**



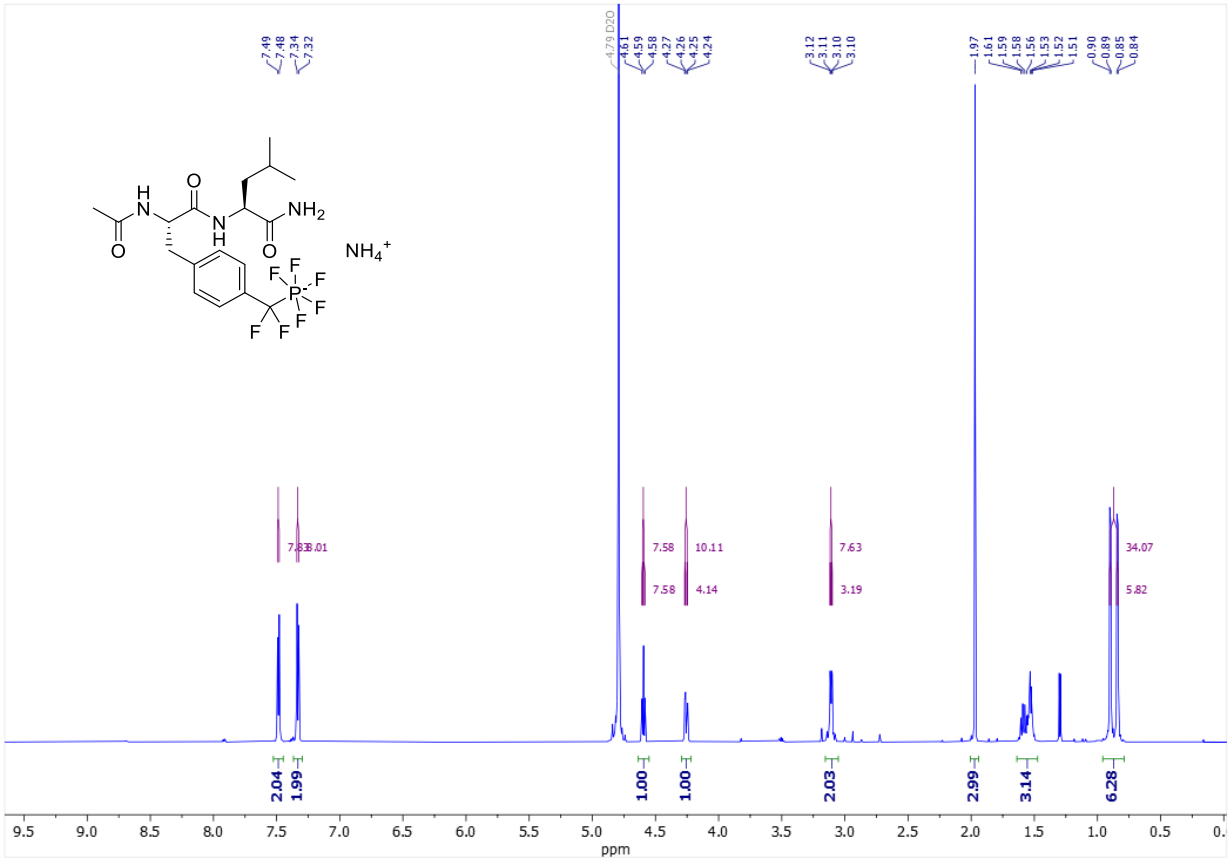
Supplementary Figure 69. ^{19}F NMR spectrum (565 MHz, D_2O) of **12**



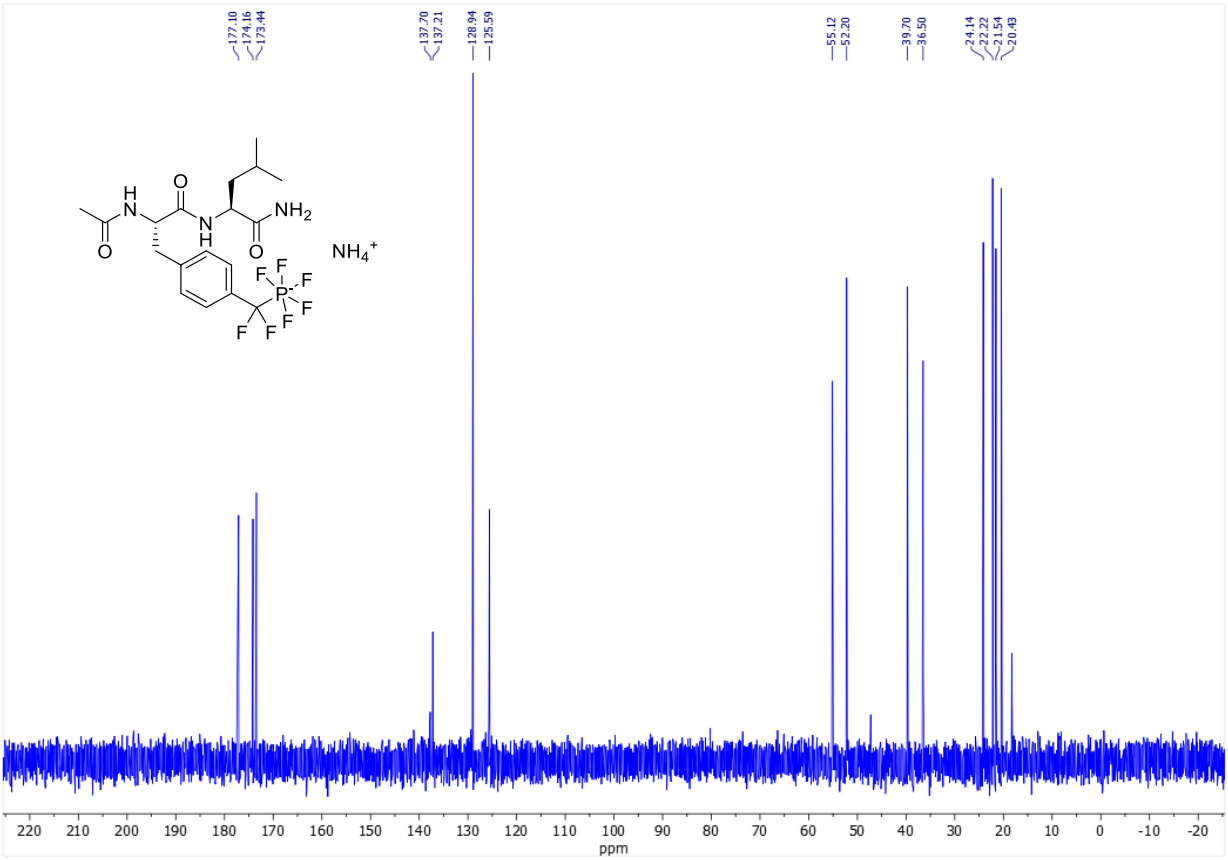
Supplementary Figure 70: UV spectrum of 0.5 mM solution of compounds **12** in water



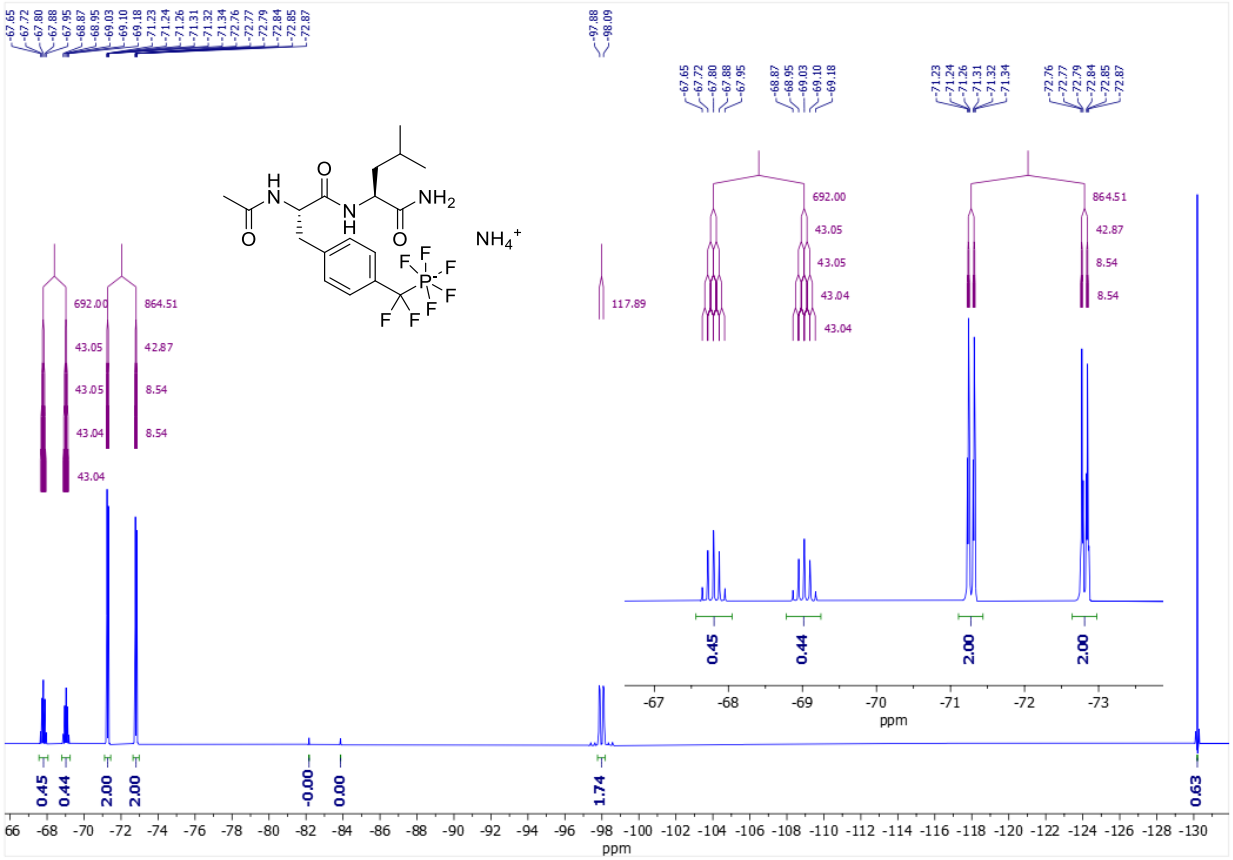
Supplementary Figure 71: UV spectrum of **12** in water displaying an n-π* transition



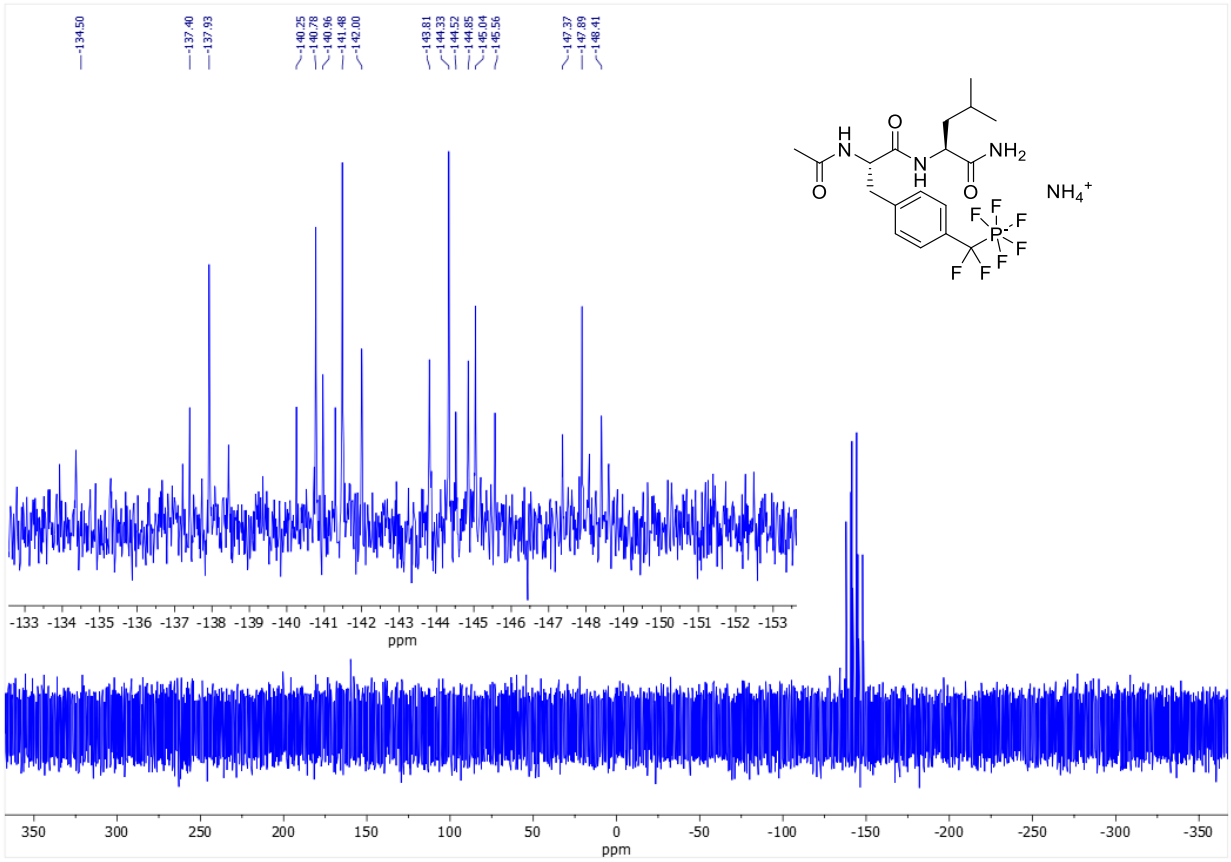
Supplementary Figure 72. ^{1}H NMR spectrum (600 MHz, D_2O) of **14**



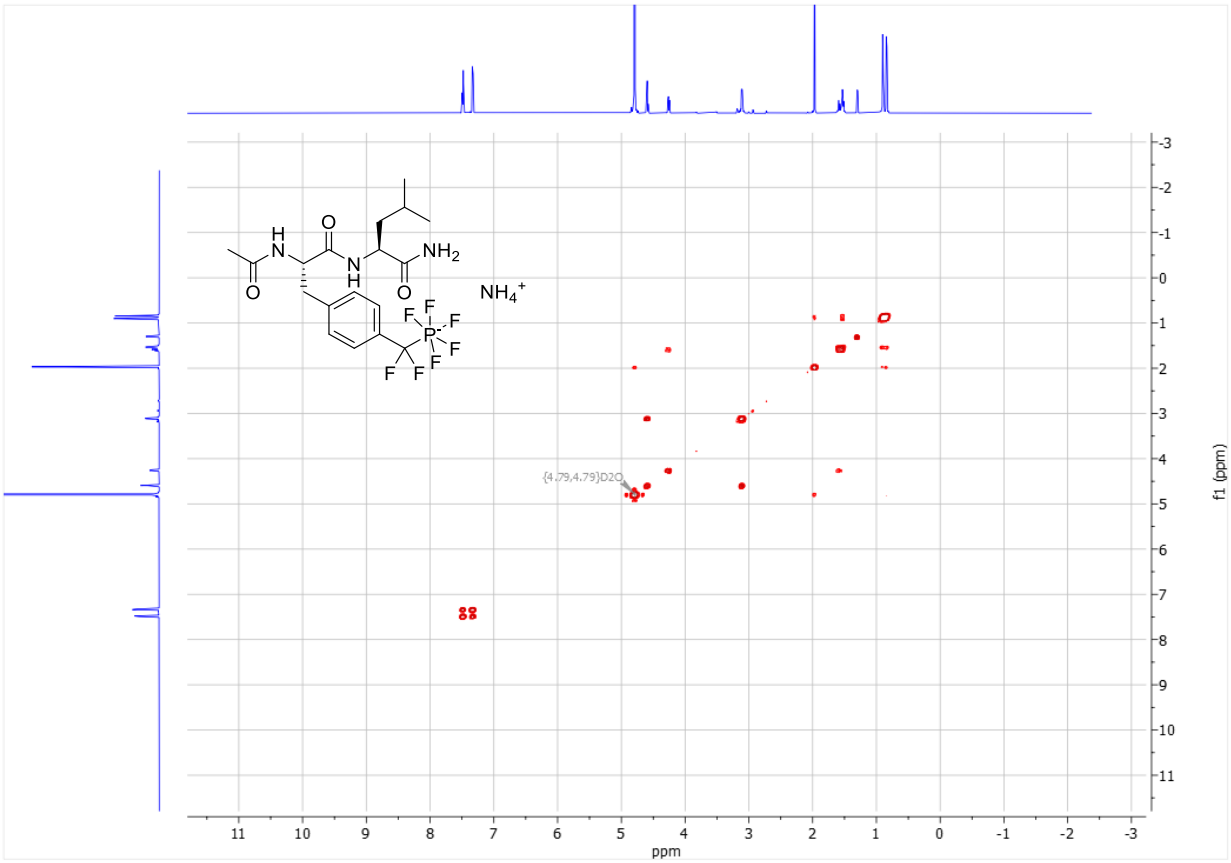
Supplementary Figure 73. ^{13}C NMR spectrum (151 MHz, D_2O) of **14**



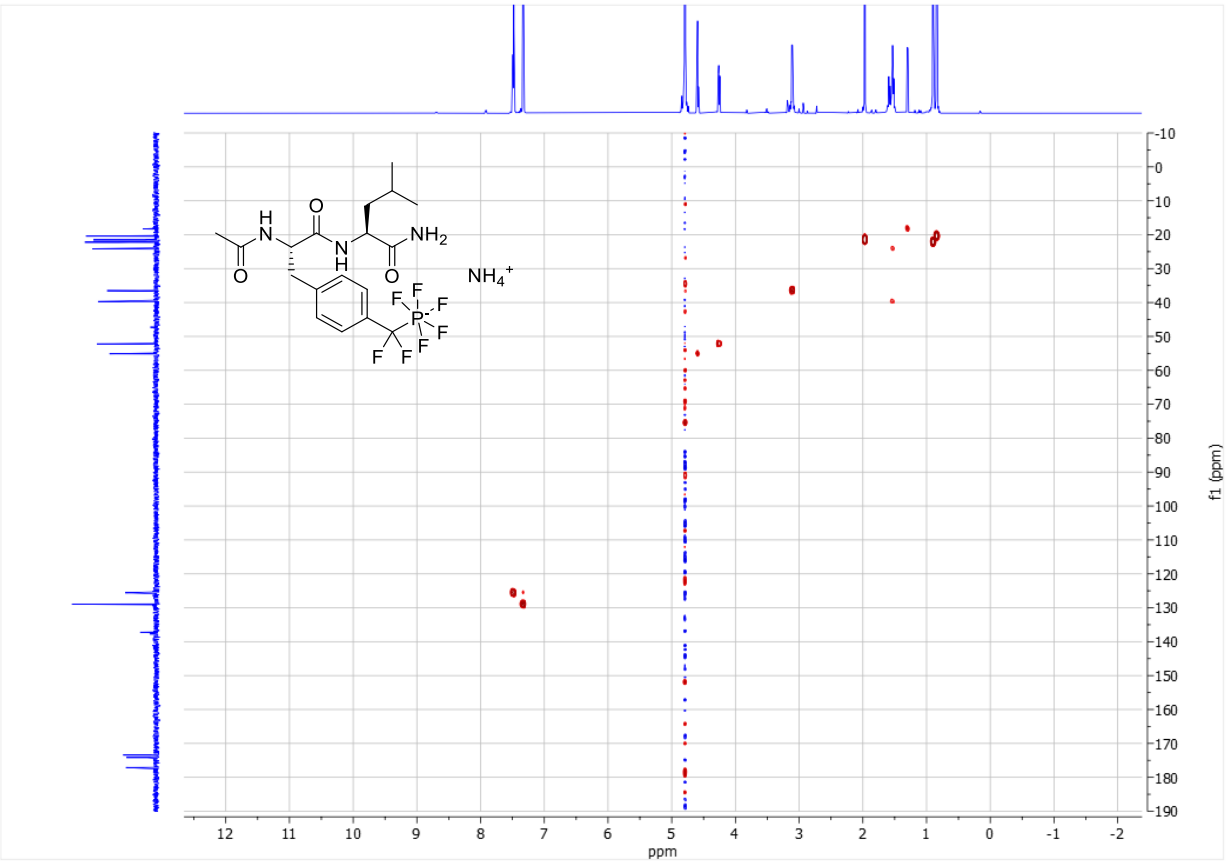
Supplementary Figure 74. ^{19}F NMR spectrum (565 MHz, D_2O) of **14**



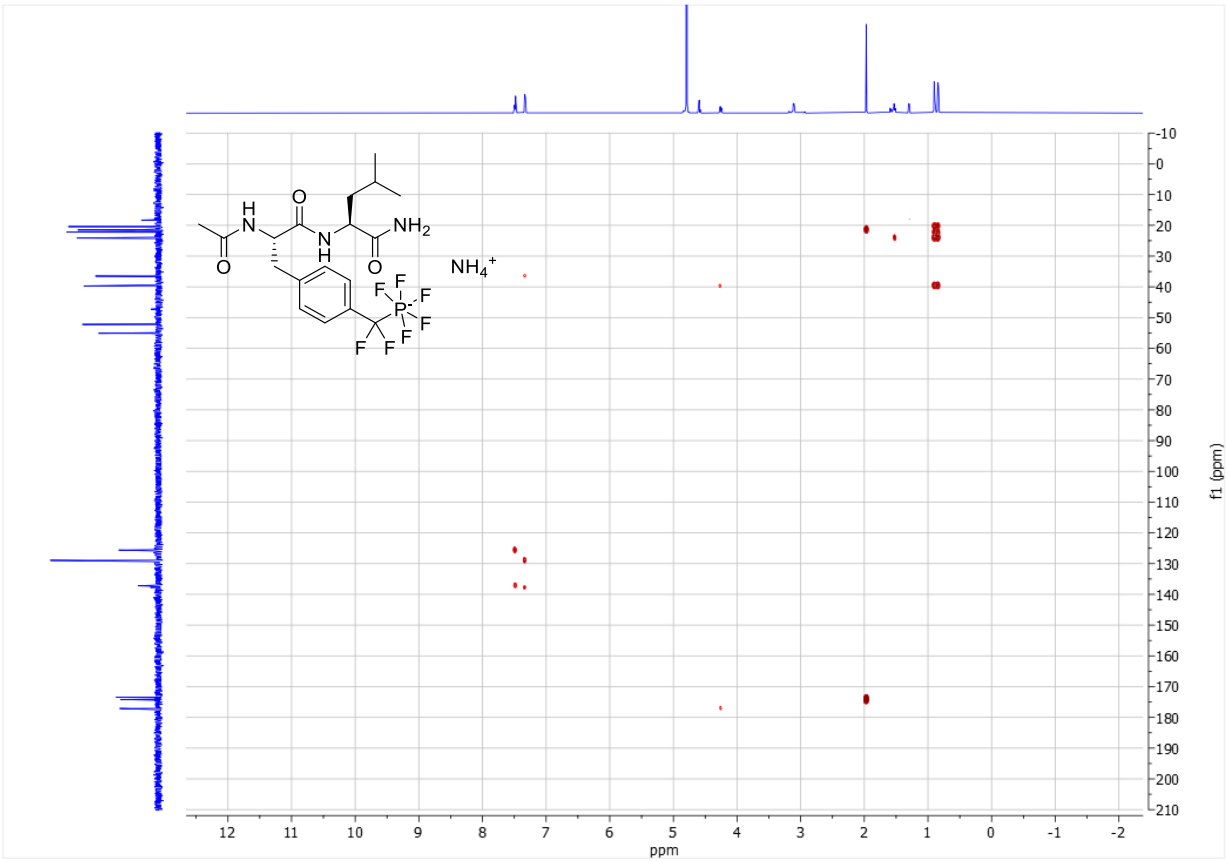
Supplementary Figure 75. ^{31}P NMR spectrum (243 MHz, D_2O) of **14**



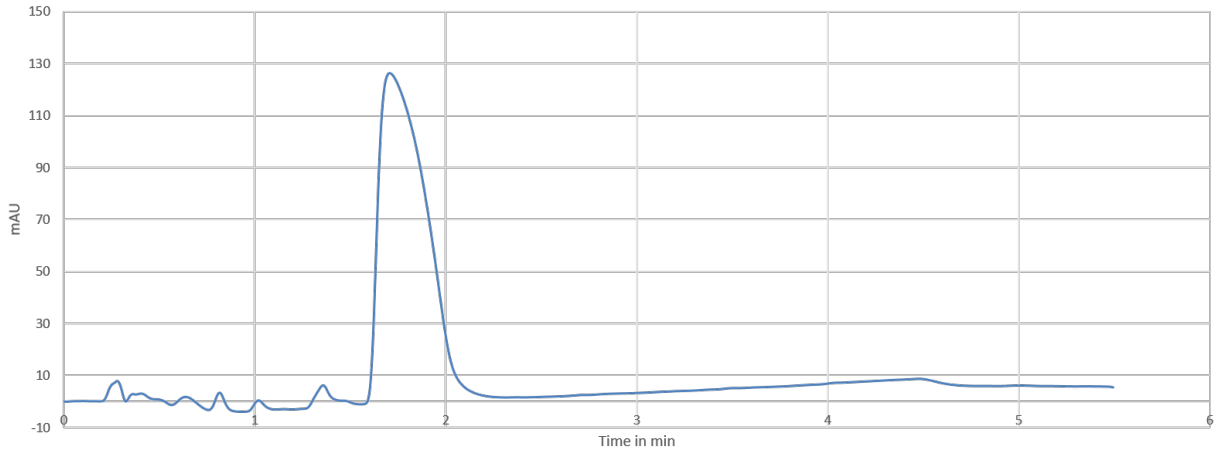
Supplementary Figure 76. H,H-COSY-NMR NMR spectrum (600 MHz, D_2O) of **14**



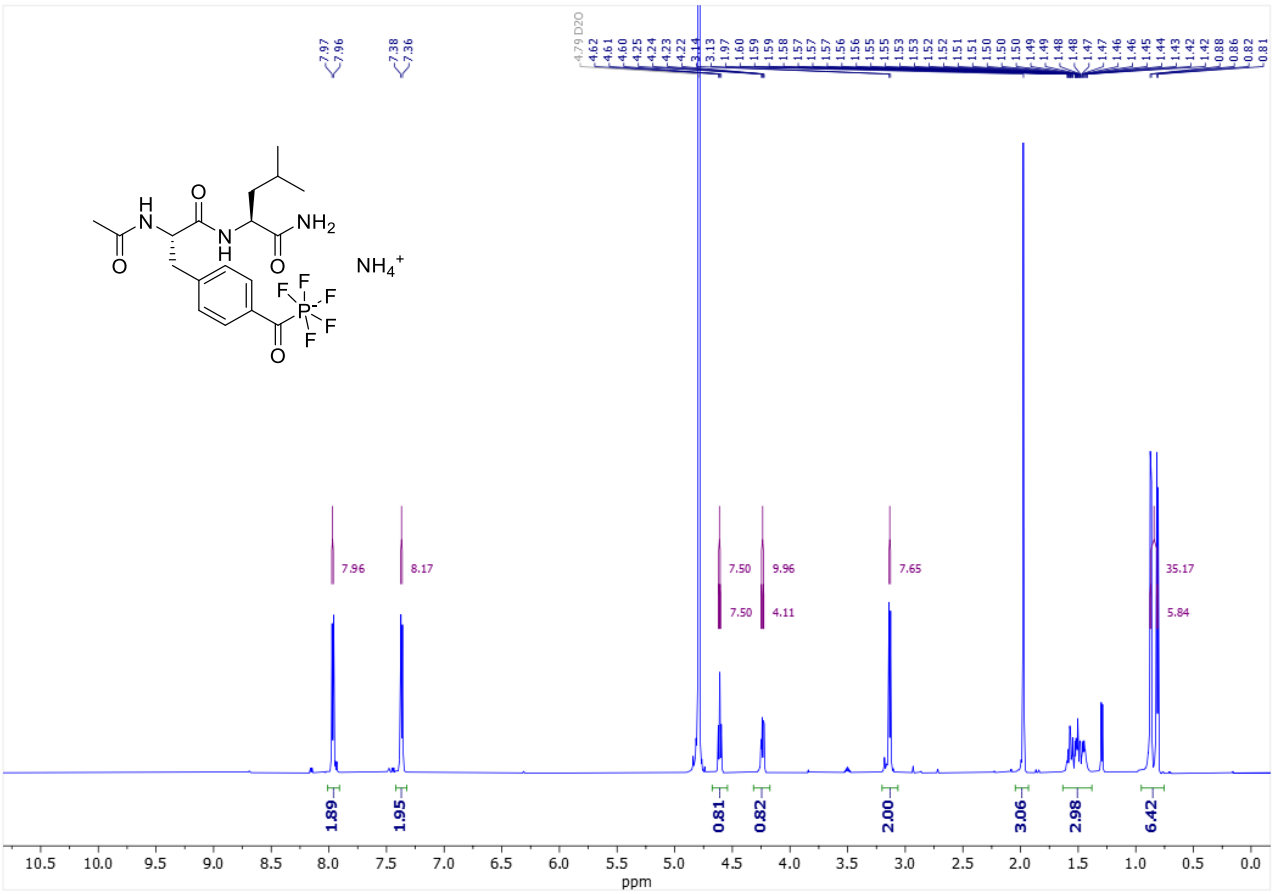
Supplementary Figure 77. HMBC NMR spectrum (600 MHz, D_2O) of **14**



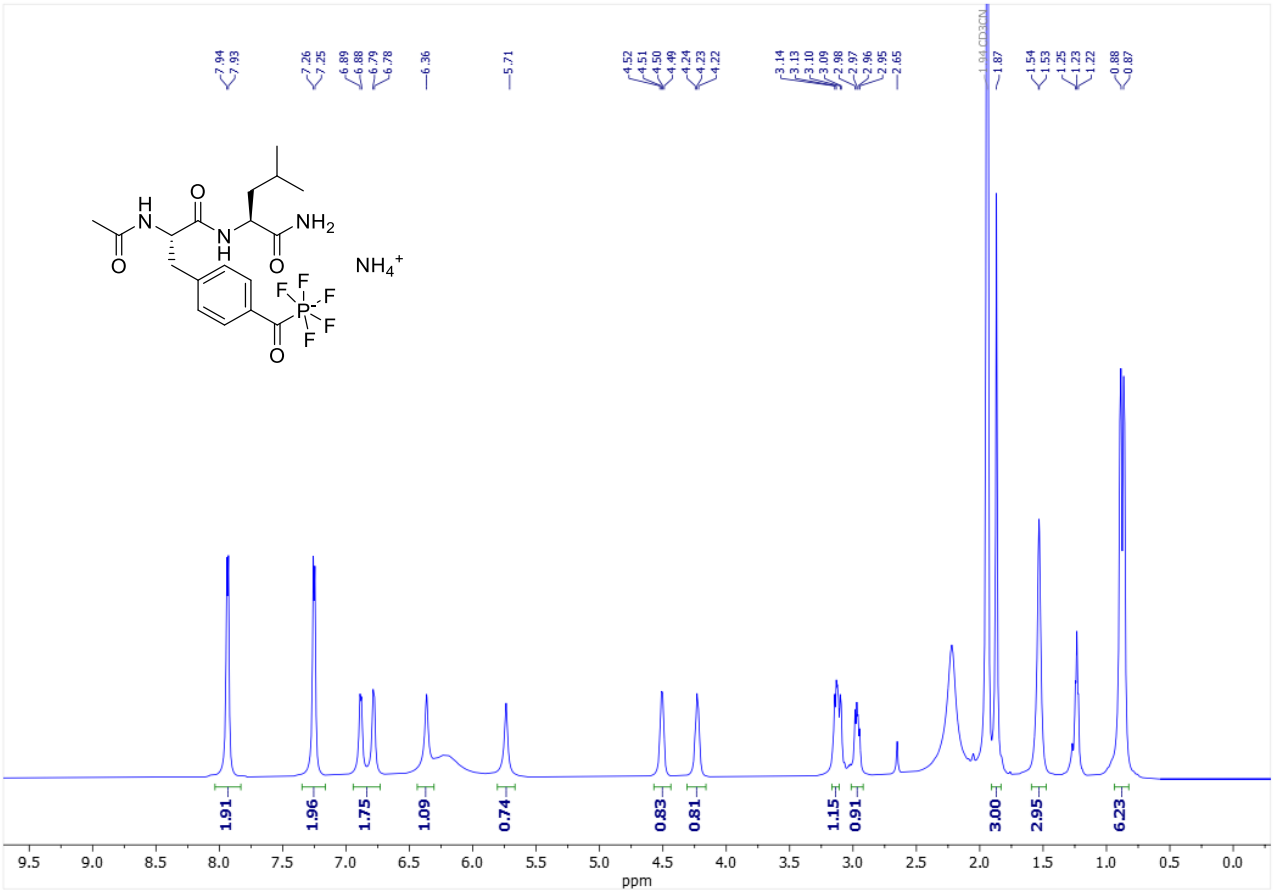
Supplementary Figure 78. HMBC NMR spectrum (600 MHz, D_2O) of **14**



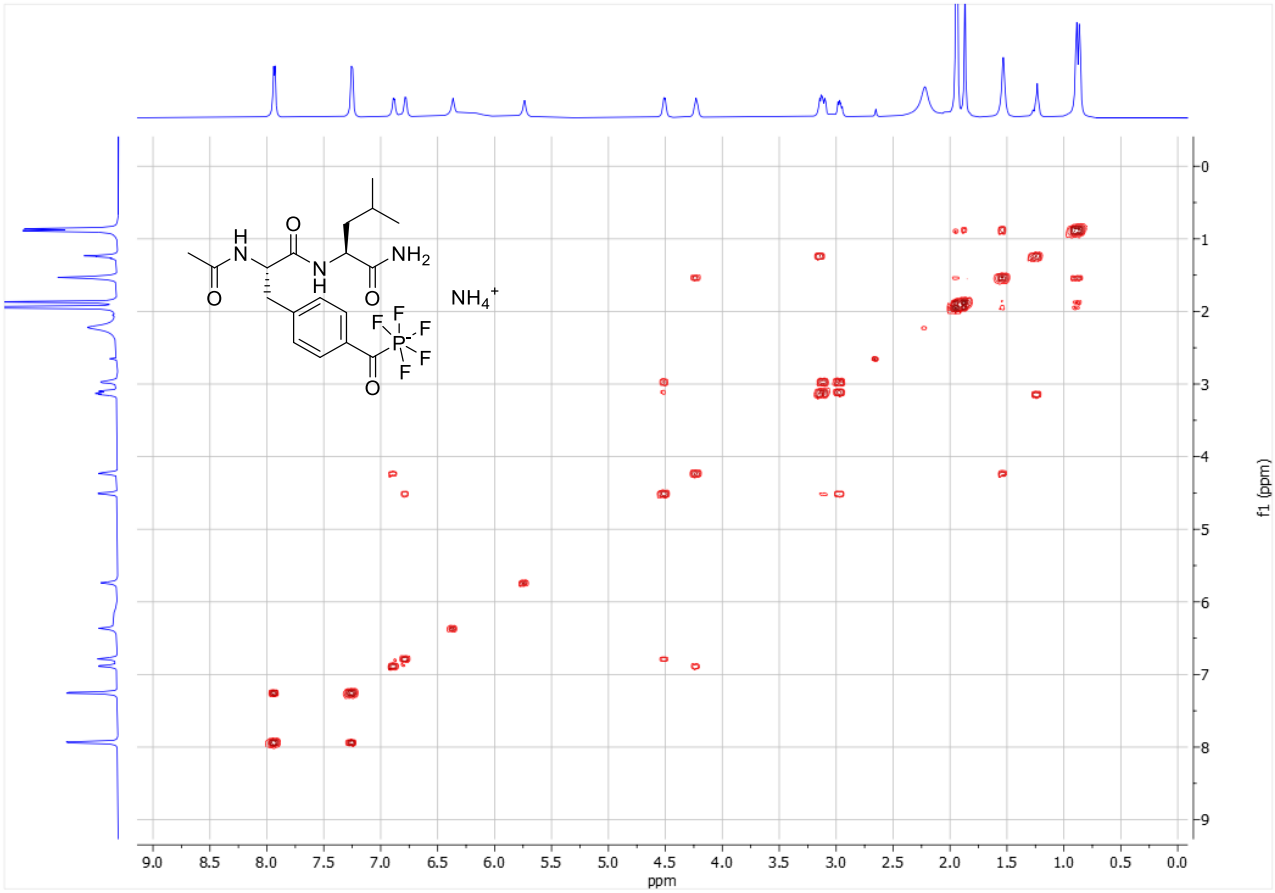
Supplementary Figure 79. HPLC chromatogram of compound **14**. Column B, eluent 15-95% ACN in 5 min, DAD 210 nm.



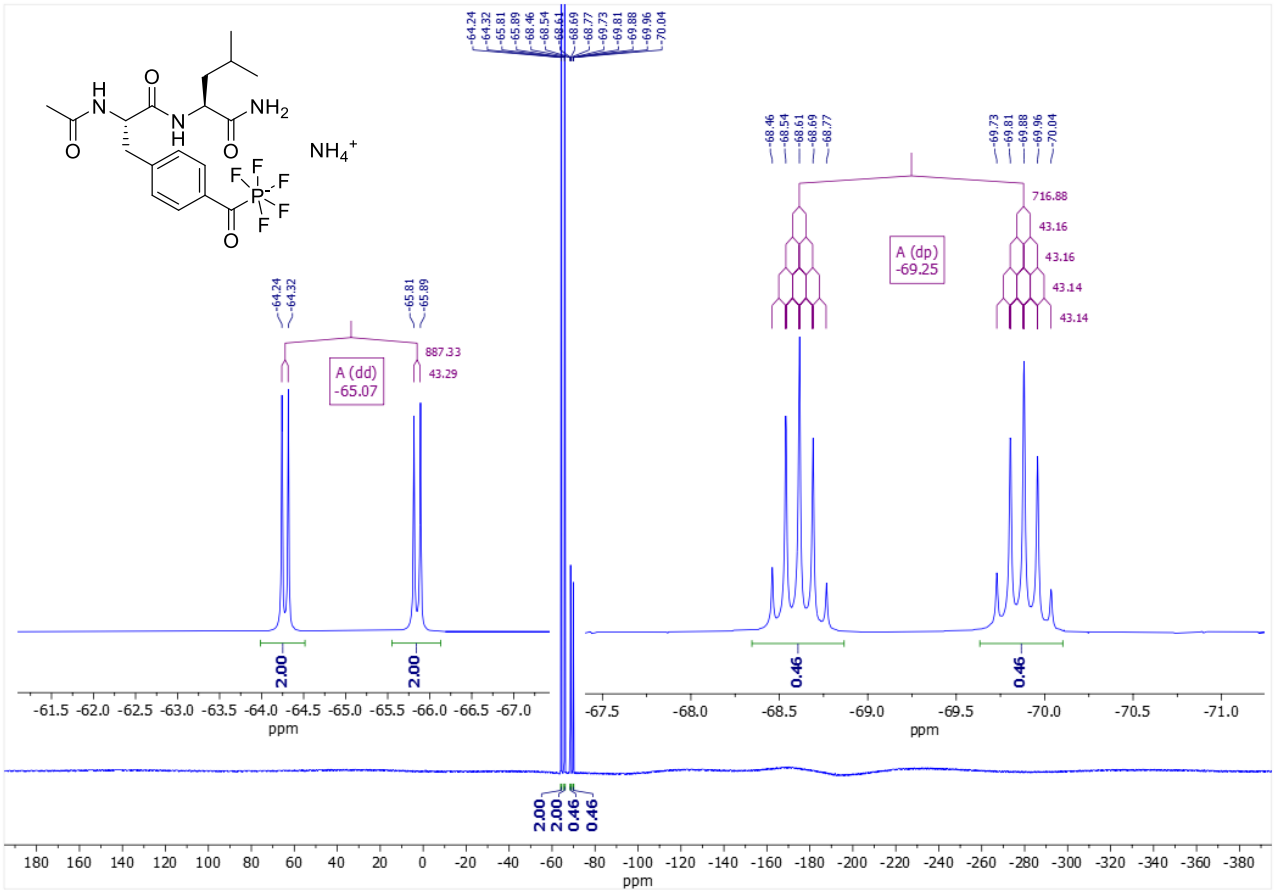
Supplementary Figure 80. ^{1}H NMR spectrum (600 MHz, D_2O) of **15**



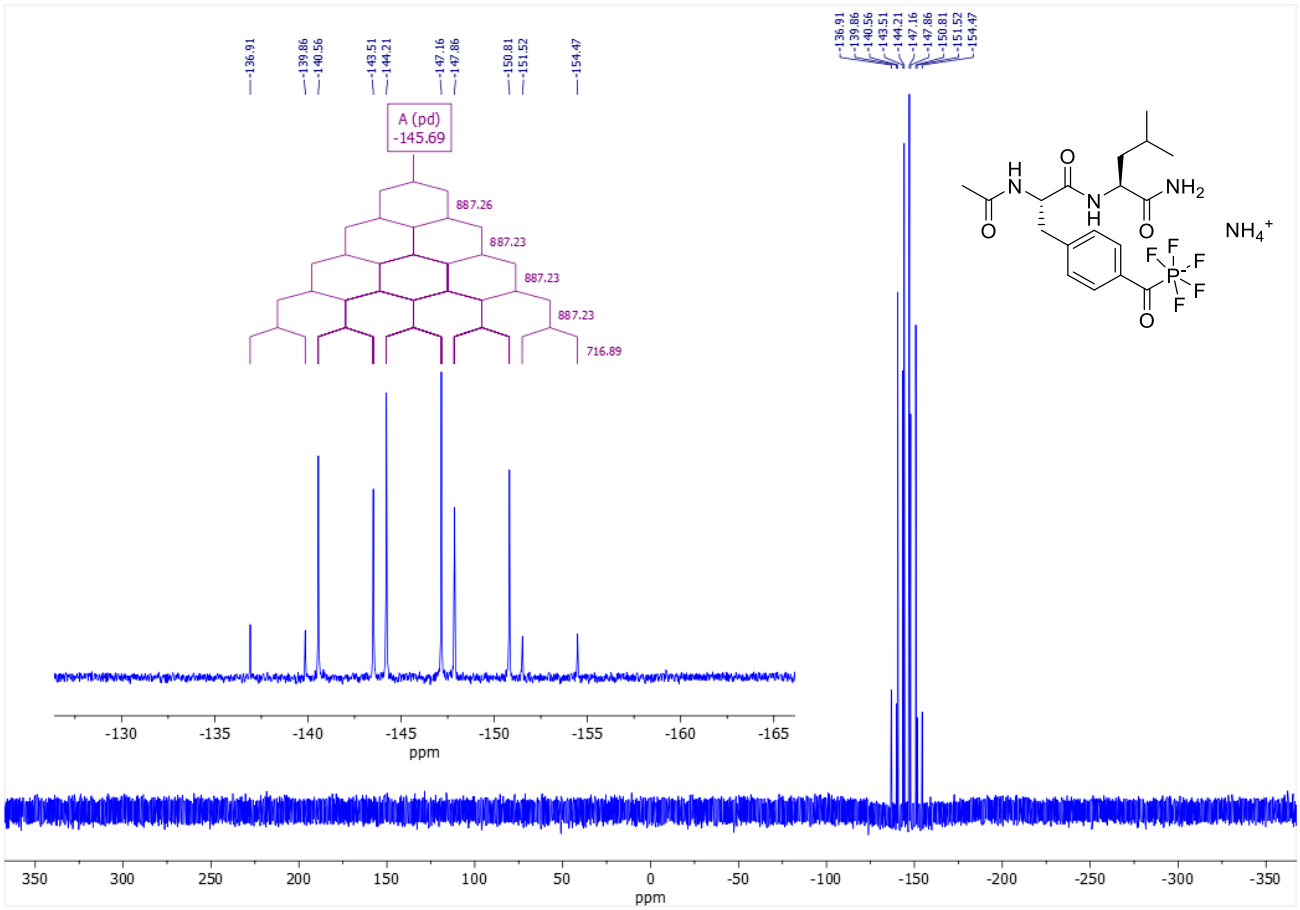
Supplementary Figure 81. ^{1}H NMR spectrum (600 MHz, ACN-D3) of **15**



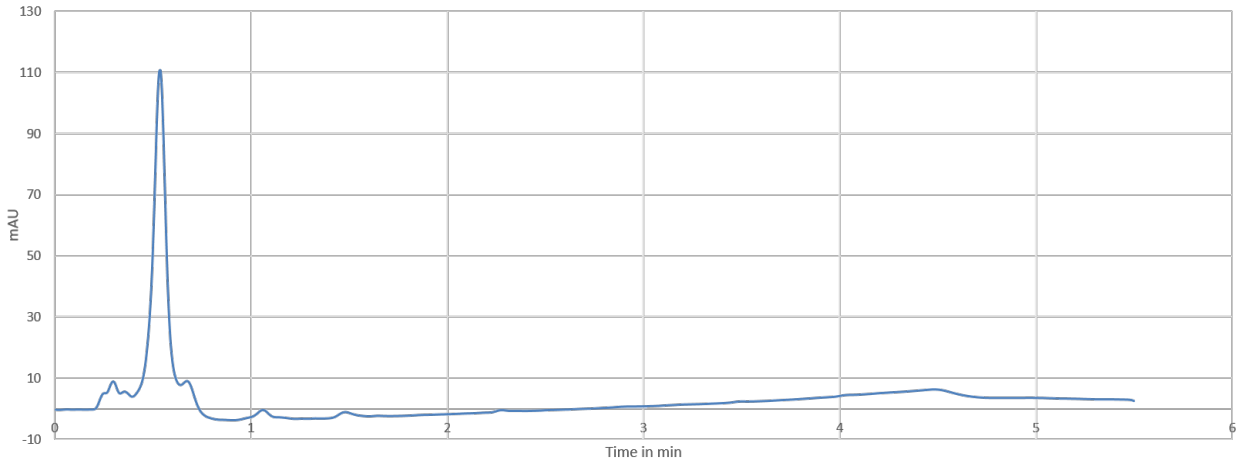
Supplementary Figure 82. H,H-COSY-NMR spectrum (600 MHz, ACN-D3) of **15**



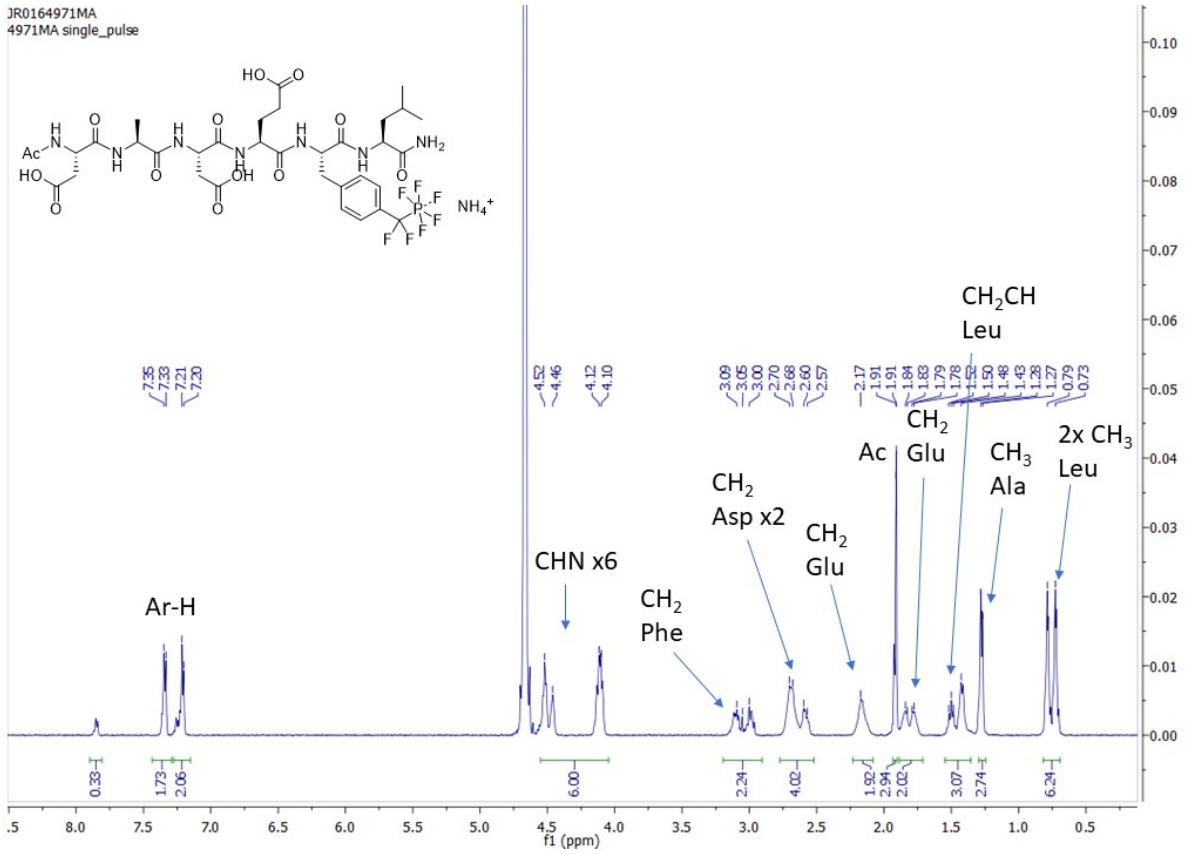
Supplementary Figure 83. ^{19}F NMR spectrum (565 MHz, ACN-D3) of **15**



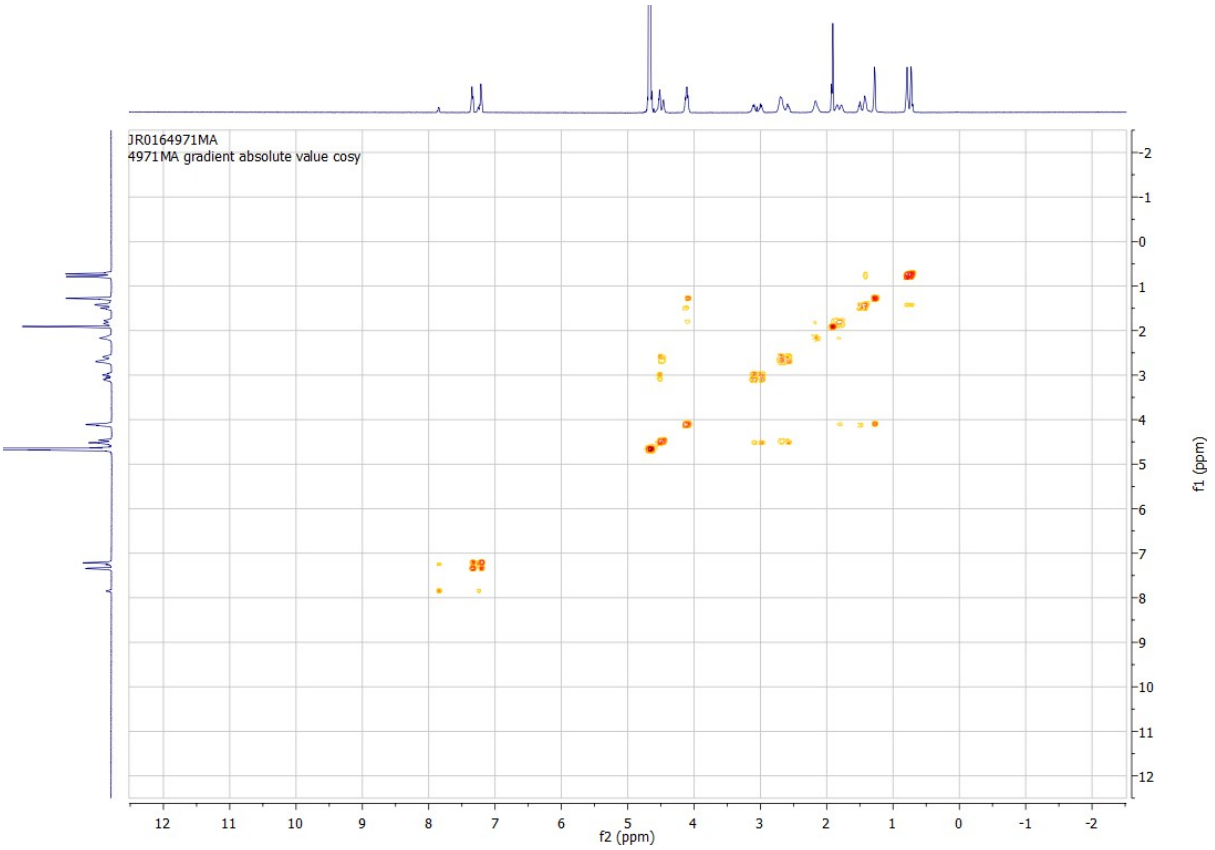
Supplementary Figure 84. ^{31}P NMR spectrum (243 MHz, ACN-D3) of **15**



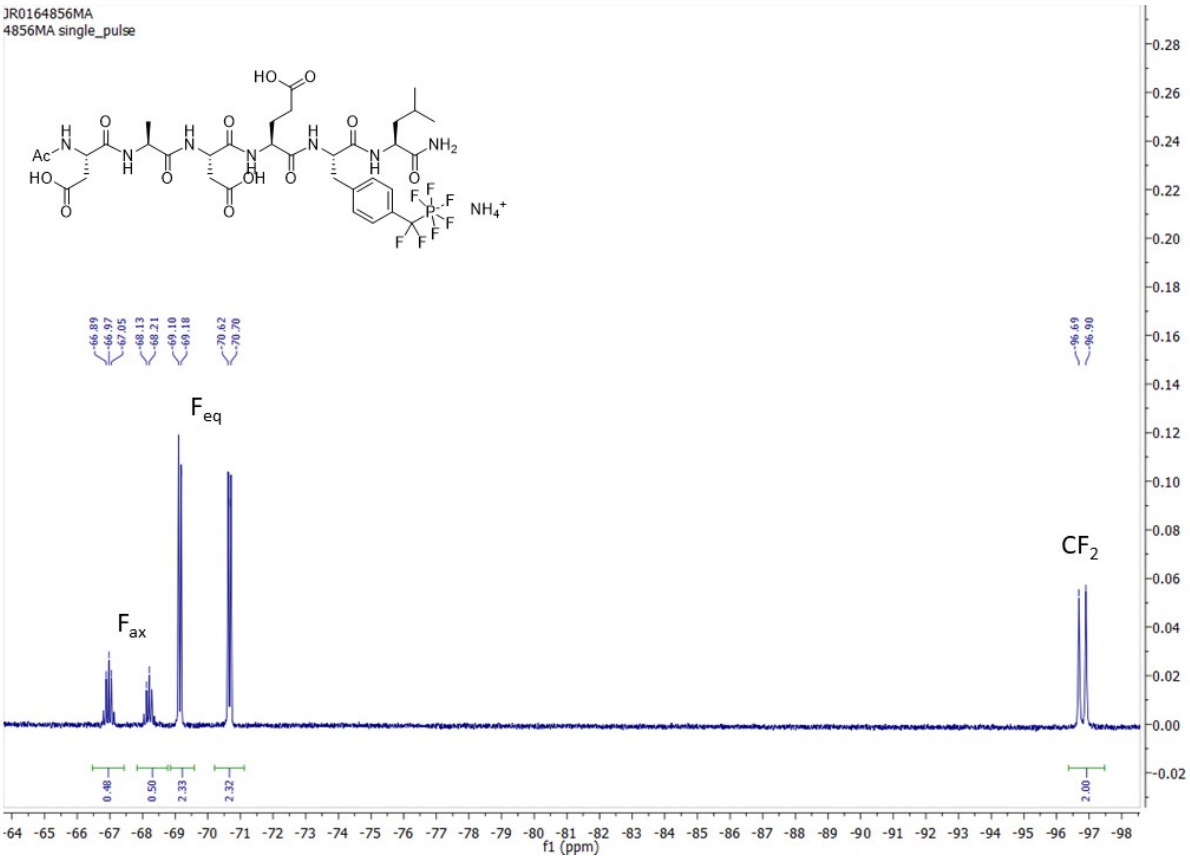
Supplementary Figure 85. HPLC chromatogram of compound **15** with UV detection at 220 nm. Column B, eluent 15-95% ACN in 5 min, DAD 210 nm.



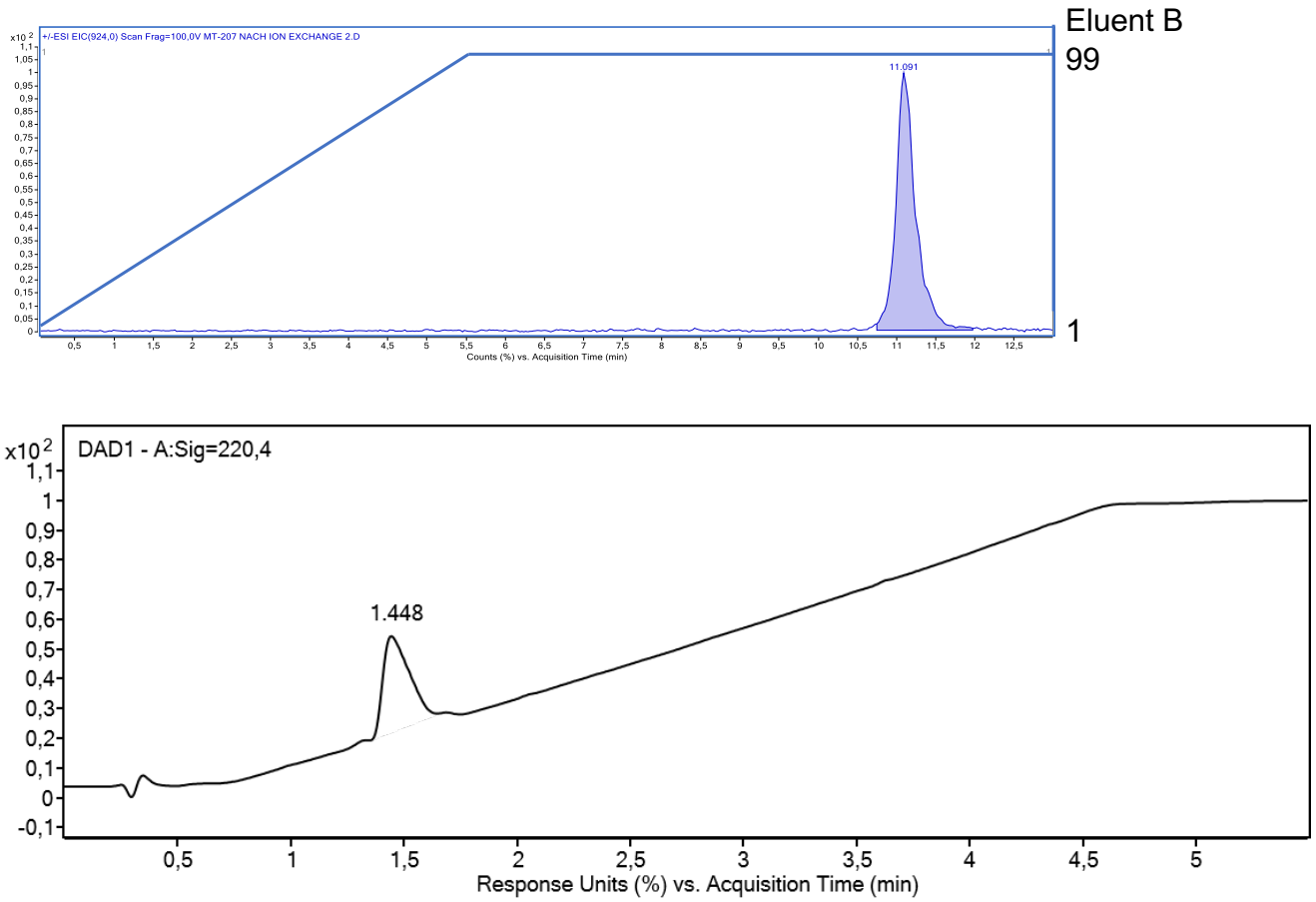
Supplementary Figure 86. ^{1}H NMR spectrum (600 MHz, D_2O) of **16**



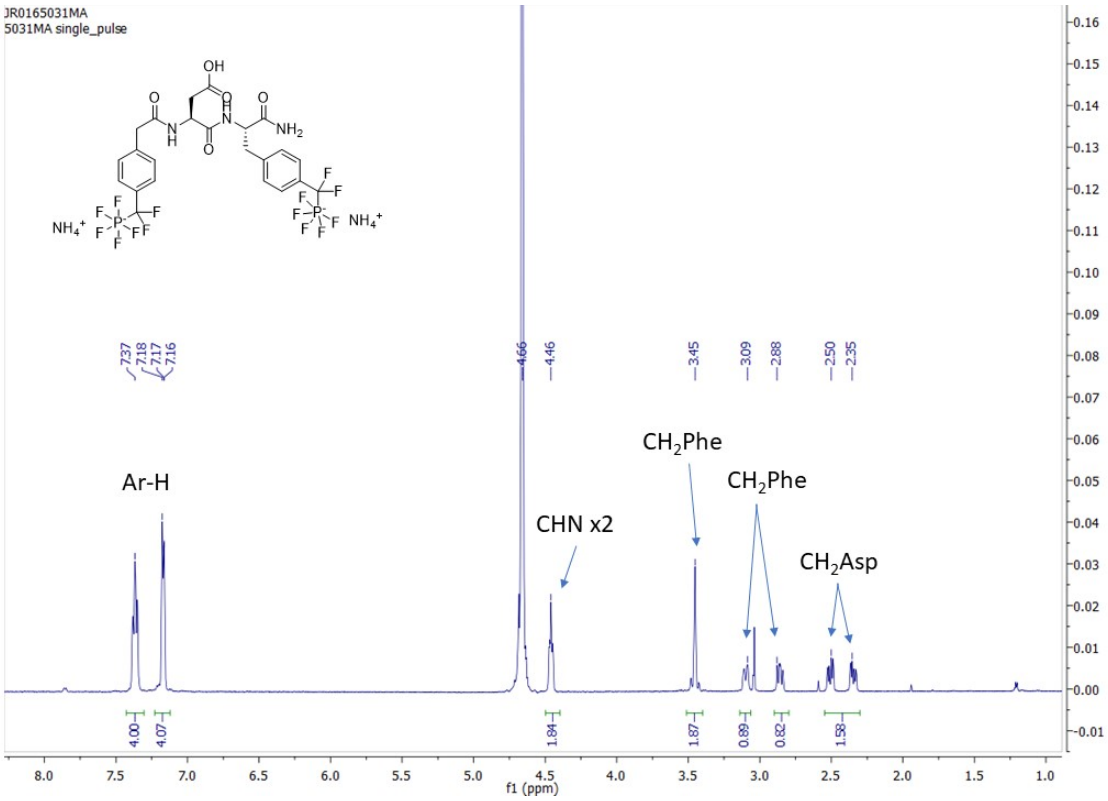
Supplementary Figure 87. H,H-COSY-NMR spectrum (600 MHz, D_2O) of **16**



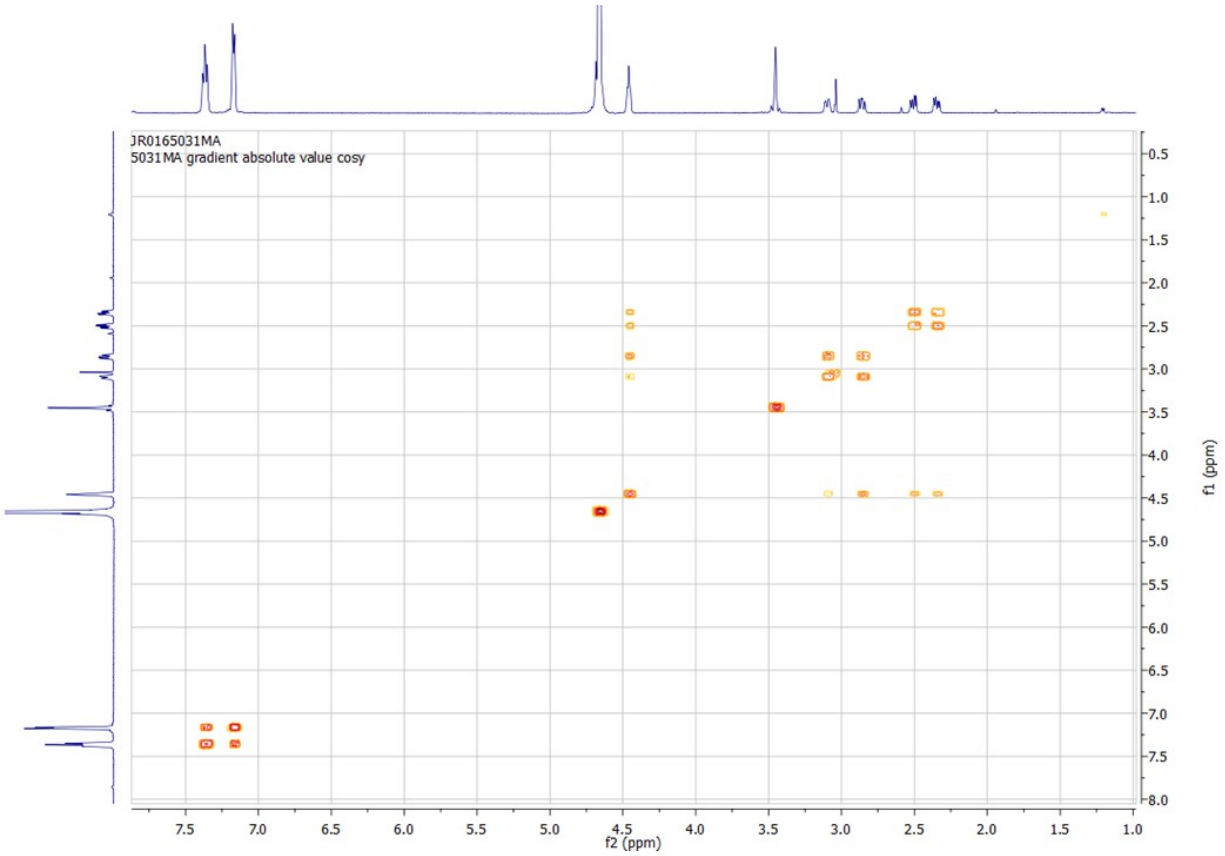
Supplementary Figure 88. ^{19}F NMR spectrum (565 MHz, D_2O) of **16**



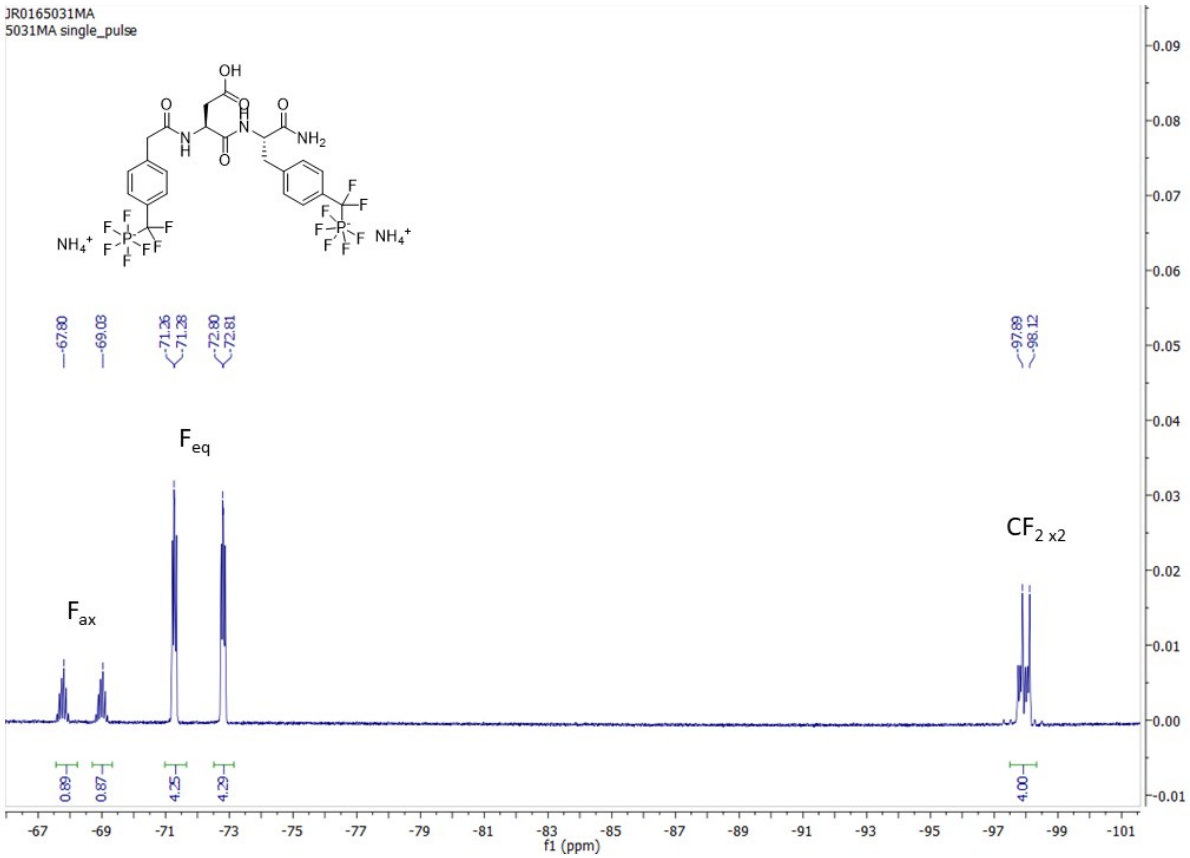
Supplementary Figure 89. HPLC chromatogram of compound **16**. Top: Column A, 1-99% eluent ACN in 5.5 min. Bottom: Column B, eluent 15-95% ACN in 4.5 min, 220 nm DAD.



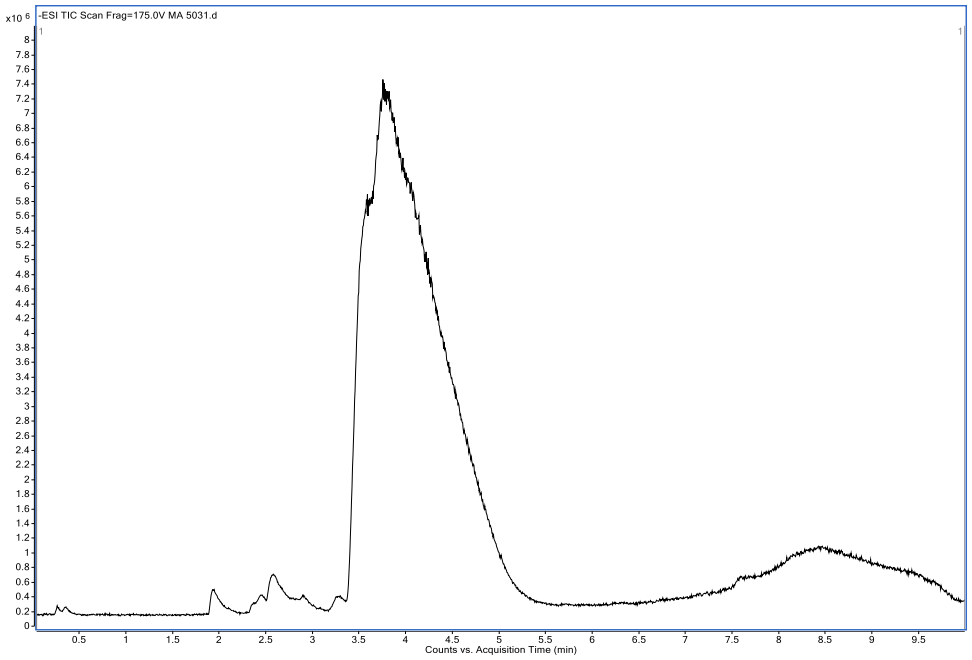
Supplementary Figure 90. ^{1}H NMR spectrum (600 MHz, D_2O) of **17**



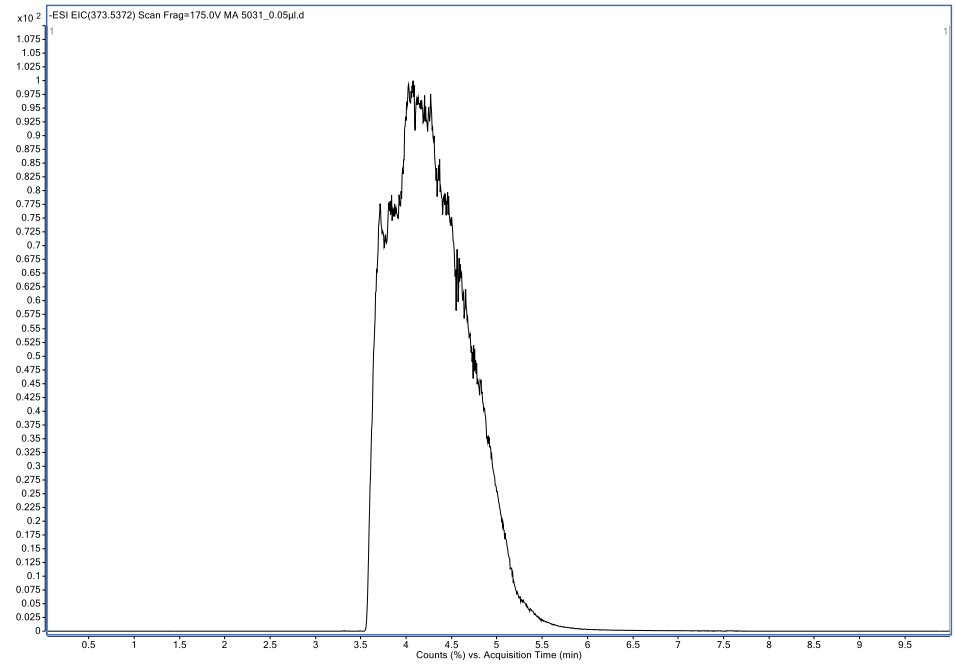
Supplementary Figure 91. H,H-COSY-NMR spectrum (600 MHz, D_2O) of **17**



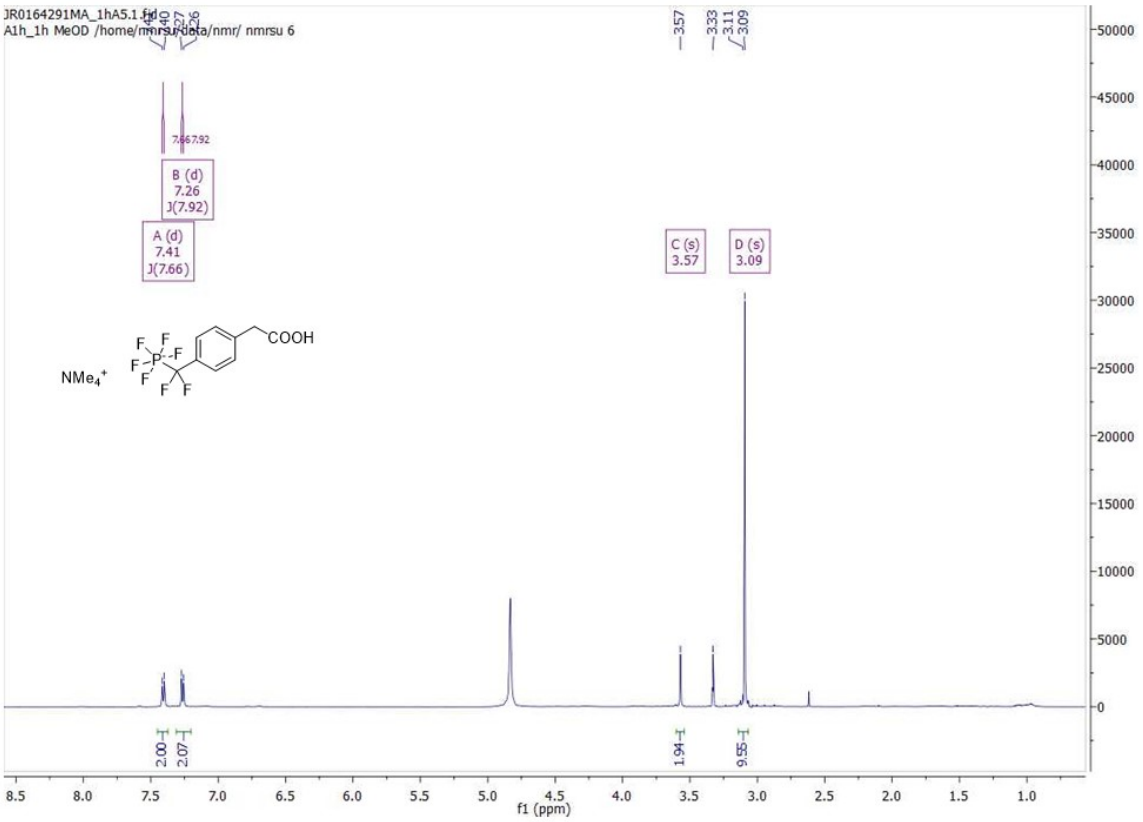
Supplementary Figure 92. ^{19}F NMR spectrum (565 MHz, D_2O) of **17**



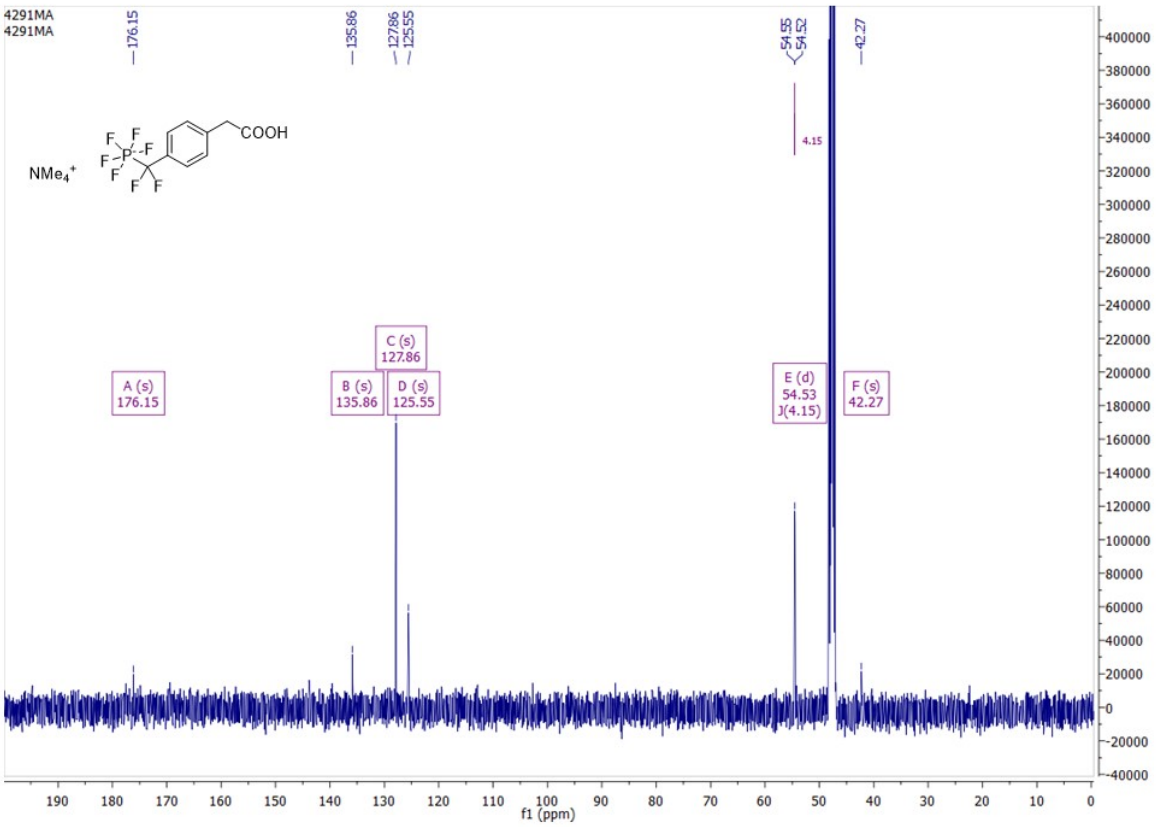
Supplementary Figure 93. Total ion chromatogram of compound **17**. Column B, eluent 15-95% ACN in 8 min



Supplementary Figure 94. Extracted ion chromatogram of compound **17**. Column B, eluent 15-95% ACN in 8 min

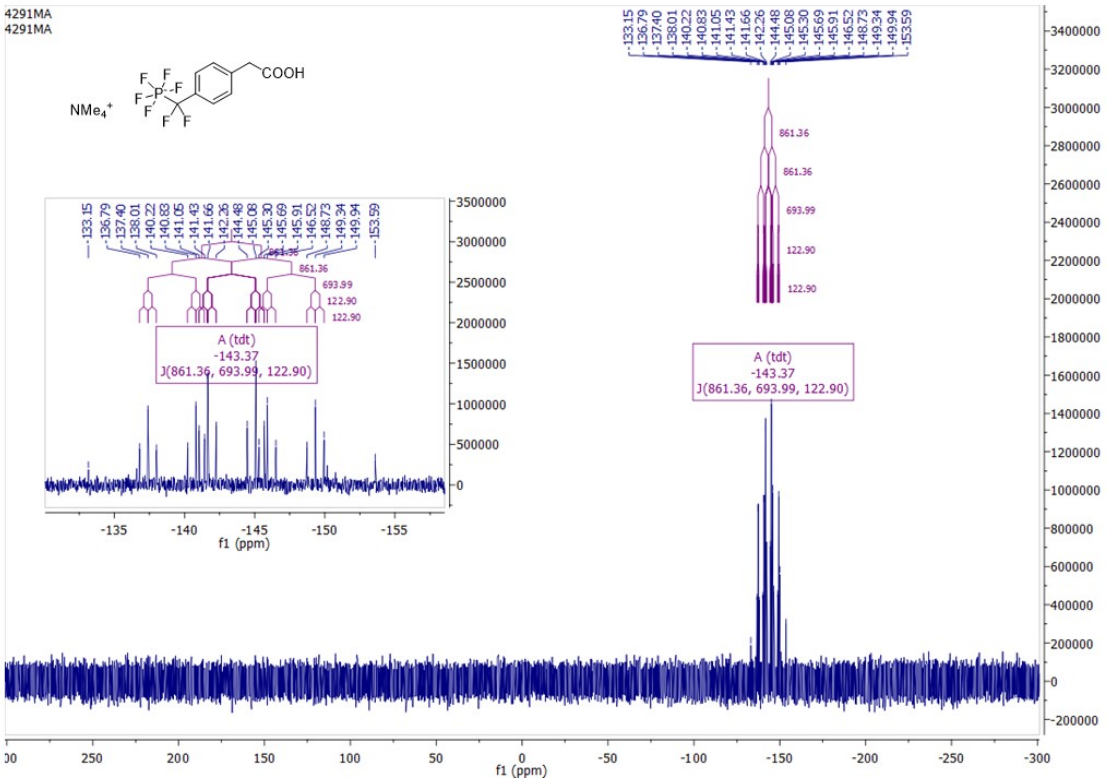


Supplementary Figure 95. ^{1}H NMR spectrum (500 MHz, MeOD) of **18**

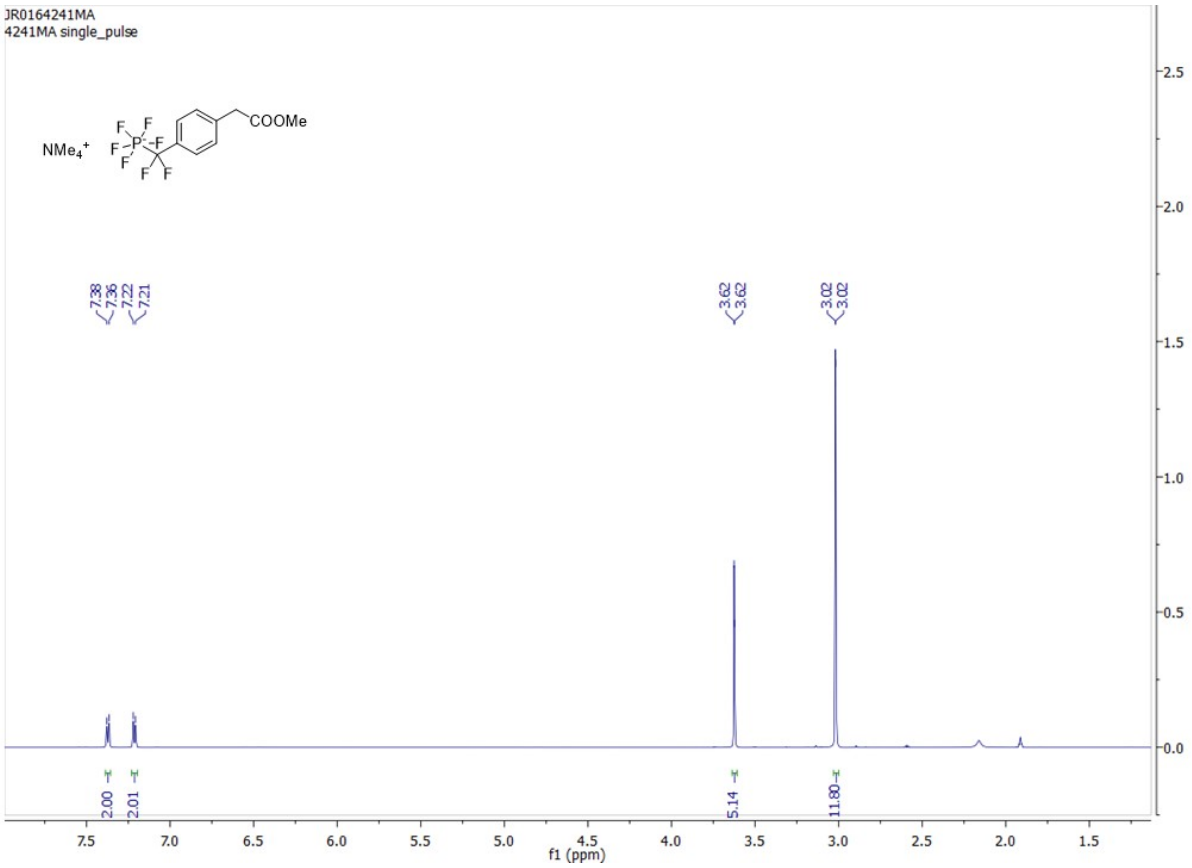


Supplementary Figure 96. ^{13}C NMR spectrum (125 MHz, MeOD) of **18**

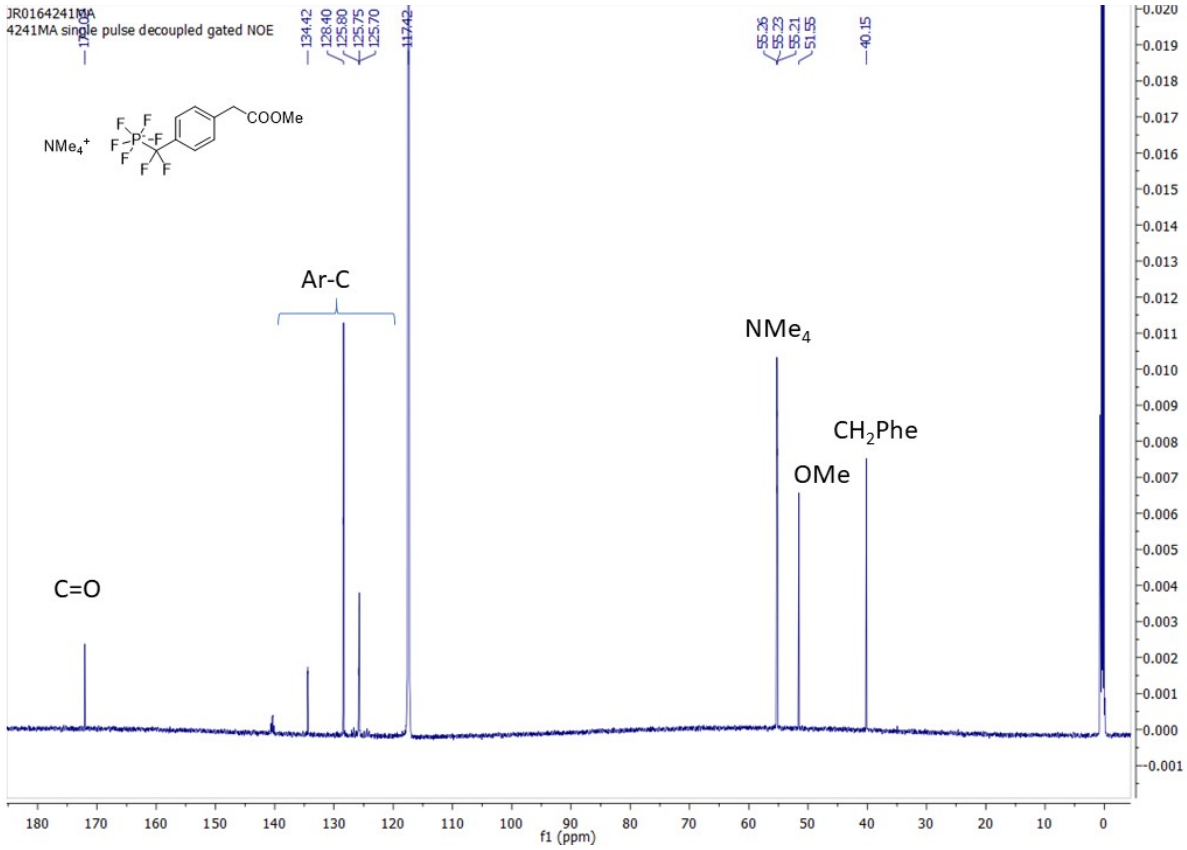
Supplementary Figure 97. ^{19}F-NMR spectrum (470 MHz, MeOD) of **18**



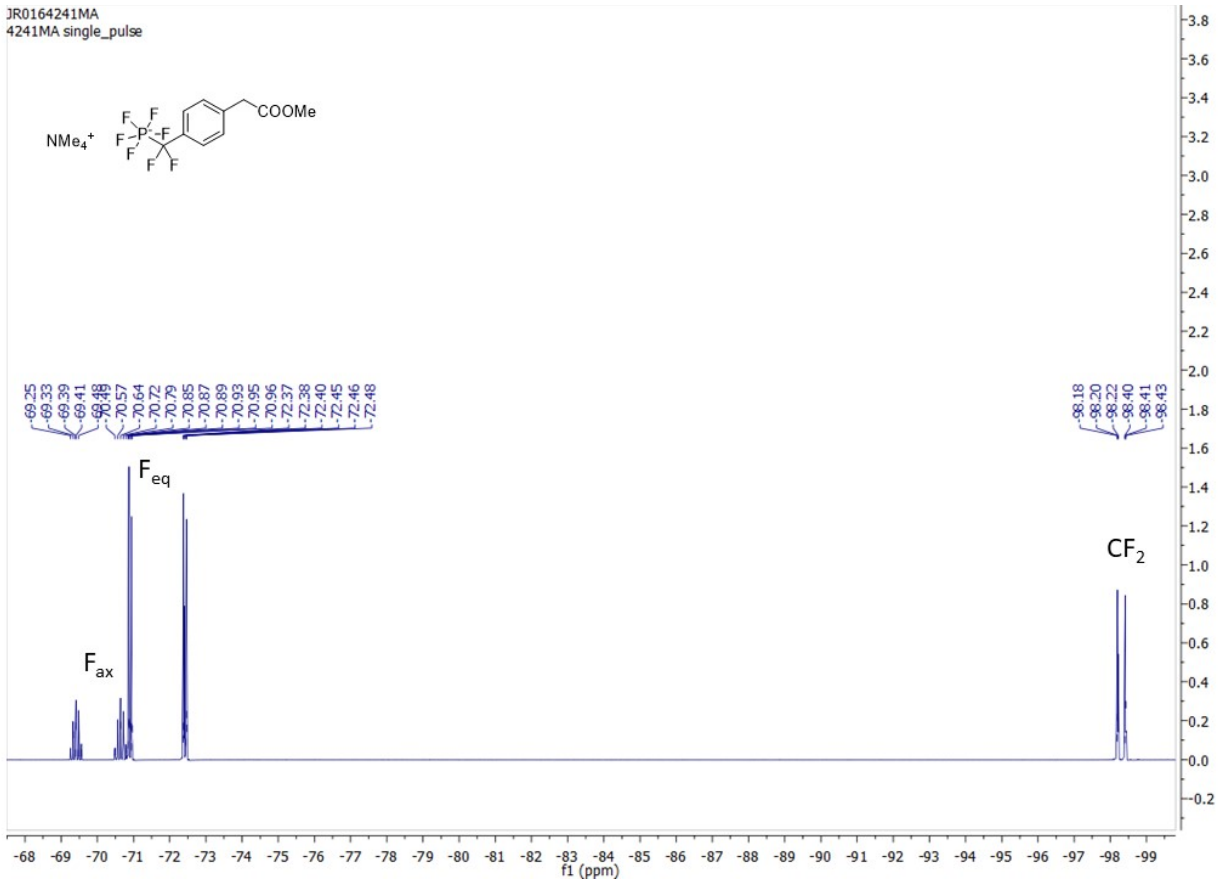
Supplementary Figure 98. ^{31}P-NMR spectrum (202 MHz, MeOD) of **18**



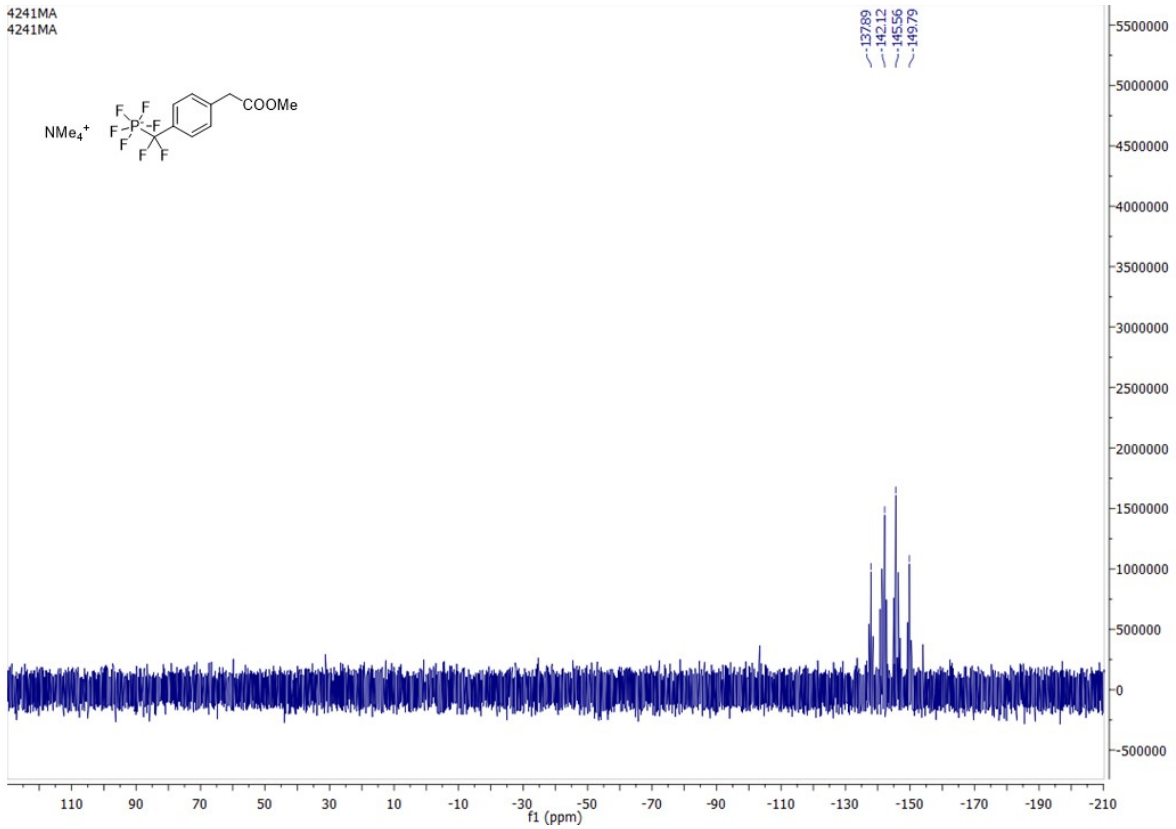
Supplementary Figure 99. ^{1}H NMR spectrum (600 MHz, ACN-d3) of **21**



Supplementary Figure 100. ^{13}C NMR spectrum (151 MHz, ACN-d3) of **21**



Supplementary Figure 101. ^{19}F NMR spectrum (565 MHz, ACN-d3) of **21**



Supplementary Figure 102. ^{31}P NMR spectrum (243 MHz, ACN-d3) of **21**

References

[1] K. O. Christe, W. W. Wilson, R. D. Wilson, R. Bau, J. A. Feng, *J. Am. Chem. Soc.* **1990**, *112*, 7619-7625.
[2] E. Kaiser, R. L. Colescott, C. D. Bossinger, P. I. Cook, *Anal. Biochem.* **1970**, *34*, 595-598.
[3] J. Ge, H. Wu, S. Q. Yao, *Chem. Commun.* **2010**, *46*, 2980-2982.
[4] M. N. Qabar, J. Urban, M. Kahn, *Tetrahedron* **1997**, *53*, 11171-11178.
[5] I. G. Boutselis, X. Yu, Z. Y. Zhang, R. F. Borch, *J. Med. Chem.* **2007**, *50*, 856-864.
[6] a) A. S. Bruker, AXS Inc., **2004** Madison, WI; b) G. M. Sheldrick, *Acta Crystallogr. A* **2008**, *64*, 112-122.
[7] G. M. Sheldrick, SADABS (Version 2.03), **2002** University of Göttingen. Göttingen, Germany.
[8] G. M. Sheldrick, *Acta Crystallogr. A Found Adv.* **2015**, *71*, 3-8.
[9] L. J. Farrugia, *Journal of Applied Crystallography* **1999**, *32*, 837-838.
[10] C. F. Macrae, I. Sovago, S. J. Cottrell, P. T. A. Galek, P. McCabe, E. Pidcock, M. Platings, G. P. Shields, J. S. Stevens, M. Towler, P. A. Wood, *J. Appl. Crystallogr.* **2020**, *53*, 226-235.
[11] a) K. Brandenburg, H. Putz, Crystal Impact 2.6.4. **2021** GbR, Bonn, Germany; b) G. Bergerhoff, M. Berndt, K. Brandenburg, *J. Res. Natl. Inst. Stand. Technol.* **1996**, *101*, 221-225.
[12] C. Yung-Chi, W. H. Prusoff, *Biochemical Pharmacology* **1973**, *22*, 3099-3108.
[13] N. Krishnan, K. Krishnan, C. R. Connors, M. S. Choy, R. Page, W. Peti, L. Van Aelst, S. D. Shea, N. K. Tonks, *J. Clin. Invest.* **2015**, *125*, 3163-3177.
[14] G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, *J. Comput. Aided Mol. Des.* **2013**, *27*, 221-234.
[15] a) R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.* **2006**, *49*, 6177-6196; b) R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739-1749; c) T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750-1759.
[16] C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, R. Abel, R. A. Friesner, E. D. Harder, *J. Chem. Theory Comput.* **2021**, *17*, 4291-4300.
[17] a) E. Lindahl, B. Hess, D. van der Spoel, *J. Mol. Model.* **2001**, *7*, 306-317; b) S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, *Bioinformatics* **2013**, *29*, 845-854; c) M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1-2*, 19-25.
[18] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696-3713.
[19] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926-935.
[20] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014101.
[21] M. Parrinello, A. Rahman, *Journal of Applied Physics* **1981**, *52*, 7182-7190.
[22] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577-8593.

[23] R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernandez, C. R. Schwantes, L. P. Wang, T. J. Lane, V. S. Pande, *Biophys. J.* **2015**, *109*, 1528-1532.
[24] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157-1174.
[25] A. W. Sousa da Silva, W. F. Vranken, *BMC Res. Notes* **2012**, *5*, 367.
[26] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, I. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, C. SciPy, *Nat. Methods* **2020**, *17*, 261-272.
[27] J. R. Robalo, S. Huhmann, B. Koksch, A. Vila Verde, *Chem* **2017**, *3*, 881-897.
[28] a) C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, *The Journal of Physical Chemistry* **1993**, *97*, 10269-10280; b) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
[29] a) D. A. Case, T. E. Cheatham, 3rd, T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*, 1668-1688; b) D. A. Case, K. Belfon, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, *AMBER 2020,* **2020**; c) M. J. Frisch, G. W. Trucks, H. B. Schlegel, **2016**, Gaussian 16, Revision C.01.
[30] a) T. A. Halgren, *J. Chem. Inf. Model.* **2009**, *49*, 377-389; b) T. Halgren, *Chem. Biol. Drug Des.* **2007**, *69*, 146-148.
[31] A. C. Cheng, R. G. Coleman, K. T. Smyth, Q. Cao, P. Soulard, D. R. Caffrey, A. C. Salzberg, E. S. Huang, *Nat. Biotechnol.* **2007**, *25*, 71-75.
[32] Schrödinger Knowledge Base. Available at: https://www.schrodinger.com/kb/144. (Accessed: 122nd September 2021).