Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Aldehydes can switch the chemoselectivity of electrophiles in protein labeling

Srinivasa Rao Adusumalli, Dattatraya Gautam Rawale, Vishal Rai*

Organic and Bioconjugate Chemistry Laboratory (OBCL), Department of Chemistry
Indian Institute of Science Education and Research Bhopal, Bhauri, Bhopal Bypass Road, Bhopal 462066, India

vrai@iiserb.ac.in

Contents

1. General

The reagents, resins, proteins, and enzymes were purchased from Merck and Thermo Fisher Scientific. The reagent grade organic solvents were used. Aqueous buffers were prepared freshly using Millipore Grade I water (Resistivity > 5 MΩ cm, Conductivity < 0.2 μS/cm, TOC <30 ppb). Mettler Toledo (FE20) pH meter was used to adjust the final pH. The reaction mixture for the small molecules was stirred (Heidolph magnetic stirrer, 500-600 rpm). Proteins were either vortexed (350 rpm, 25-37 ºC) or incubated (25-37 ºC) in Thermo Fisher Scientific MaxQ 8000 incubator-shaker. Amicon® Ultra-0.5 mL 3-kDa or 10-kDa MWCO centrifugal filters from Merck Millipore were used to remove small molecules from protein mixture, desalting, and buffer exchange. Organic solvents were removed by BUCHI rotavapor R-210/215, whereas aqueous samples were lyophilized by CHRiST ALPHA 2-4 LD plus lyophilizer. Circular Dichroism (CD) measurements were recorded on JASCO J-815 CD spectropolarimeter equipped with a Peltier temperature controller. All the spectra were measured with a scan speed of 50 nm/min and spectral band width of 1 nm using 10 mm path length cuvette at 25 ºC. UV-Vis spectra were recorded in Agilent Carry-100 UV-Vis Spectrophotometer connected with a Peltier temperature controller. Steady-state fluorescence spectra were recorded in HORIBA JOBIN YVON, FLUOROLOG 3-111. The quartz cuvette of 1 cm path length was used to record the fluorescence spectra.

Chromatography: Thin-layer chromatography (TLC) was performed on silica gel coated aluminium TLC plates (Merck, TLC Silica gel 60 F254). The compounds were visualized using a UV lamp (254 nm) and stains such as iodine, ninhydrin, 2,4-diphenylhydrazine. The flash column chromatography was carried out on Combiflash Rf 200 or gravity columns using 230-400 or 100-200 mesh silica gel from Merck.

Nuclear magnetic resonance spectra: ^{1}H, ^{13}C, and ^{19}F NMR spectra were recorded on Bruker Avance III 400 and 500 MHz NMR spectrometer. ^{1}H NMR spectra were referenced to TMS (0 ppm), DMSO-d_6 (2.50 ppm), D_2O (4.79 ppm), and acetone-d_6 (2.05 ppm) whereas ^{13}C NMR spectra were referenced to $CDCl_3$ (77.16 ppm), DMSO-d_6 (39.52 ppm) and acetone-d_6 (29.84 ppm). ^{19}F NMR was referenced to trifluoroacetic acid (-75.70 ppm). Peak multiplicities are designated by the following abbreviations: s, singlet; bs, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublets. All the NMR spectra were recorded at 298 K.

Mass spectrometry: Agilent Technologies 1200 series HPLC paired to Agilent 6130 mass spectrometer (ESI/APCI) was used for ESI-MS data. HRMS data were recorded on Bruker Daltonics MicroTOF-Q-II with electrospray ionization (ESI). Matrix assisted laser desorption/ionisation time of flight mass spectrometry was performed with Bruker Daltonics UltrafleXtreme Software-Flex control version 3.4, using sinapic acid and α-cyano-4-hydroxycinnamic acid (HCCA) matrix. Data analysis was performed using Flex analysis. Peptide mass and fragment ion calculator (http://db.systemsbiology.net:8080/proteomicsToolkit/FragIonServlet.html) were used for peptide mapping and sequencing.

Acetonitrile and H_2O were buffered with 0.01% formic acid and used as the mobile phase. Method A was used for HPLC to record the ESI-MS data for proteins.

Method A (Column: Agilent, Poroshell 300 SB-C18 5 µm 2.1 × 75 mm, flow rate 0.4 ml/min)

Time (min)	Acetonitrile (%)	H_2O (%)
0	10	90
1	10	90
8	60	40
12	90	10
15	90	10

Reaction conversion determination for protein labeling

ESI-MS: Conversion for protein labeling was calculated based on the relative peak intensity of native protein and labeled protein in the deconvoluted mass spectrum.

% Conversion = $I_{\text{desired product}} / I_{\text{all relevant species}}$ where $I_{\text{desired product}}$ is the peak intensity of labeled protein, and $I_{\text{all relevant species}}$ is the sum of the peak intensities of native protein and labeled protein in the deconvoluted mass spectra.

MALDI-ToF-MS: Conversion for protein labeling was calculated based on the relative peak intensity of native protein and labeled protein in the mass spectrum.

2. General procedures

2a. Protein labeling

General procedure for screening of chemoselective reagents (2c-2g) for stable Schiff base formation

RNase A **1a** (10 nmol) in phosphate buffer (140 µl, 0.1 M, pH 7.0) was taken in 1.5 ml Eppendorf tube. To this solution, chemoselective reagent **2c-2g** (250 nmol) in DMSO (60 µl) from a freshly prepared stock solution was added and vortexed at 25 ^{o}C. The reaction progress was monitored using MALDI-ToF-MS by taking the aliquots at various time intervals.

General procedure for site-selective labeling of protein

Protein **1** (10 nmol) in phosphate buffer (140 µl, 0.1 M, pH 7.0) was taken in a 2 ml Eppendorf tube. To this solution, reagent (**2a**/**2b**/**2h**-**2j**/**S15**/**S17**, 250 nmol) in DMSO (60 µl) from a freshly prepared stock solution was added and vortexed at 25 ^{o}C or 37 ^{o}C. After 9-48 h, the reaction mixture was diluted with acetonitrile:buffer (10:90, 1800 µl). The use of acetonitrile allows convenient removal of the unreacted reagent. The unreacted reagent and salts were removed by using Amicon® Ultra-0.5 mL 3-kDa MWCO centrifugal filters spin concentrator. The protein mixture was further washed with phosphate buffer (5 × 0.4 ml) and concentrated in phosphate buffer (190 µl, 0.1 M, pH 7.0). To this solution, O-benzylhydroxylamine **4** (2 µmol) in DMSO (10 µl) or hydroxylamine (2 µmol) in Millipore Grade I water (10 µl) from a freshly prepared stock solution was added and vortexed at 25 ^{o}C for 3 h to form the oxime derivative. The excess of reagent and salts were removed by using spin concentrator and the sample was collected in an aqueous medium. Modification of protein was analyzed by ESI-MS or MALDI-ToF-MS. The aqueous sample was concentrated by lyophilization before subjecting it to digestion, peptide mapping, and sequencing by MS-MS.

Note: We examined the deviation in conversion with different batches of RNase A (Sigma-Aldrich, Catalog number R6513). The conversions remain similar (±5%) within the same batch of protein. However, the difference is higher (±11%) with different batches and the conversion ranges between 40-62%. The optimal reaction time ranges between 9-18 h.

Procedure for single-site installation of tags

The labeled protein **15** was prepared according to the procedure mentioned above. The protein **15** was concentrated in phosphate buffer (160 µl, 0.1 M, pH 7.0). To this solution, derivative of O-hydroxylamine **16**/**18**/**20** (2 µmol) in DMSO (40 µl) from a freshly prepared stock solution was added to convert the mono-labeled RNase A **15** to its oxime derivative for 3-6 h. The excess of O-alkoxylamine (**16**/**18**/**20)** and salts were removed by the spin concentrator (0.5 mL, 10-kDa MWCO). The sample was analyzed by ESI-MS.

2b. Purification of labeled protein

Procedure for purification of the labeled protein from reaction mixture

In a 5 ml fritted polypropylene chromatography column with end tip closures, hydrazide beads **22** (200 µL, hydrazide resin loading: 16 µmol/ml) were taken. The beads were washed with phosphate buffer (1 M, pH 7.0, 5 x 1 ml) and re-suspended in phosphate buffer (100 µl, 1 M, pH 7.0). Protein mixture (**1a** and **15**, 100 µM) in phosphate buffer (200 µl, 1 M, pH 7.0) and *p*-phenylenediamine (*p*-PDA, 100 mM) in phosphate buffer (100 µl, 1 M, pH 7.0) were added to the beads followed by end-to-end rotation (30 rpm, rotary mixer) at 25 ^{o}C. The progress of the immobilization of the labeled protein on hydrazide resin was monitored by UV-absorbance of the supernatant. After 8 h, the supernatant was collected and the beads were washed with phosphate buffer (0.3 M, pH 7.3, 4 x 1 ml) and KCl (1 M, 3 x 1 ml) to remove the adsorbed protein from resin. The beads were further washed with the phosphate buffer (0.3 M, pH 7.0, 4 x 1 ml) and re-suspended (phosphate buffer, 200 µl, 0.3 M, pH 7.0). To release the labeled protein from its immobilized derivative **23**, *p*-PDA (100 mM) in phosphate buffer (100 µl, 0.3 M, pH 7.0) and O-hydroxylamine hydrochloride (50 µl, 1 M in phosphate buffer, 0.3 M, pH 7.0) or derivative of O-hydroxylamine (**16**/**18/20**, 50 µl, 150 mM in DMSO) were added. The mixture was subjected to end-to-end rotation at 25 ^{o}C for 4 h. The supernatant was collected while the salts and reagents were removed using the spin concentrator (3 kDa MWCO). The purity of the labeled protein **24a/17/19/21** was confirmed by ESI-MS.

2c. Protein digestion

All solutions were made freshly prior to use.[1]

Procedure for in-solution digestion of the protein

In a 1.5 ml Eppendorf tube, protein **1** (10 nmol) was incubated in 100 mM tris (10 µl, pH 7.8) with urea (6 M) for 30 minutes at 37 °C. To this solution, 1 µl of reducing agent (0.2 M DTT in 0.1 M tris) was added and the mixture incubated for 1 h at 37 °C. Subsequently, the alkylating agent (4 µl, 0.2 M iodoacetamide in 0.1 M tris) was added to the solution and incubated (in dark) for 1 h at 25 °C to block the free sulfhydryl groups. The unreacted iodoacetamide was quenched with the reducing agent (4 µl, 0.2 M DTT in 0.1 M tris) for 1 h at 25 °C. Next, the sample was diluted with grade I water to reduce the urea concentration to 0.6 M. To this solution, 10 µl of enzyme (α-chymotrypsin/trypsin) solution [5 µg, enzyme/protein (1:20); enzyme in 1 mM HCl was dissolved in 0.1 M tris and 0.01 M $CaCl_2$] was added and the mixture was incubated at 37 °C for 18 h. The pH of digested solution was adjusted to < 6 (verified by pH paper) with trifluoroacetic acid (0.5%). Subsequently, the sample was used for peptide mapping by MS and sequencing by MS-MS.

3. Synthesis and characterization data

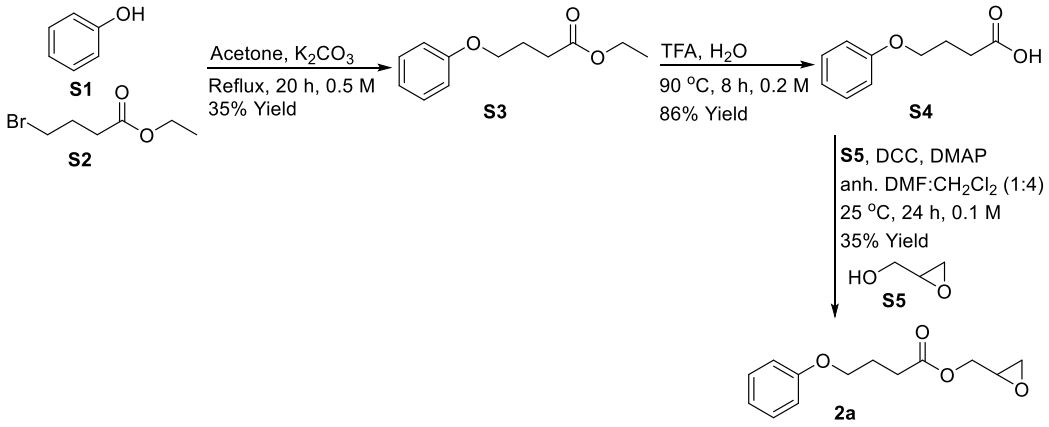
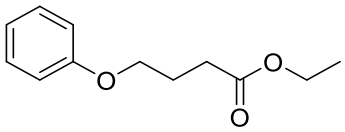


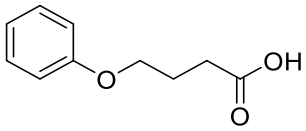
Figure S1. Synthesis of oxiran-2-ylmethyl 4-phenoxybutanoate **2a.**

Synthesis of ethyl 4-phenoxybutanoate S3[2]



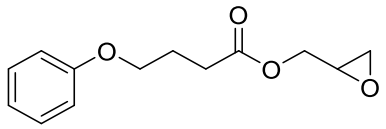
In a 25 ml round bottom flask, phenol **S1** (282 mg, 3 mmol), potassium carbonate (828 mg, 6 mmol) and ethyl 4-bromobutanoate **S2** (501 µl, 3.6 mmol) were dissolved in acetone (6 ml) and refluxed for 20 h. The progress of the reaction was monitored by TLC. The reaction mixture was concentrated using rotary evaporator. The purification of crude mixture was performed by silica gel column chromatography using ethyl acetate:n-hexane (2:98) to isolate ethyl 4-phenoxybutanoate **S3** (219 mg, 35% yield; R_f 0.59, ethyl acetate:n-hexane 10:90; colorless liquid). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.31-7.23 (m, 2H), 6.97-6.91 (m, 1H), 6.91-6.85 (m, 2H), 4.14 (q, *J* = 7.1 Hz, 2H), 4.00 (t, *J* = 6.1 Hz, 2H), 2.51 (t, *J* = 7.3 Hz, 2H), 2.14-2.07 (m, 2H), 1.25 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 173.4, 159.0, 129.6, 120.8, 114.6, 66.7, 60.6, 31.0, 24.8, 14.4. MS (ESI) $[M+H]^+$ calcd. $C_{12}H_{17}O_3$ 209.1, found 209.1.

Synthesis of 4-phenoxybutanoic acid S4[3]



In a 10 ml round bottom flask, 4-phenoxybutanoate **S3** (208 mg, 1 mmol) and trifluoroacetic acid (307 µl, 4 mmol) were dissolved in water (5 ml). The reaction mixture was stirred at 90 ^{o}C for 8 h. The reaction mixture turns into a clear solution while heating. The progress of the reaction was followed by TLC. The purification of crude reaction mixture was performed by silica gel flash column chromatography using ethyl acetate:n-hexane (5:95) to isolate gave 4-phenoxybutanoic acid **S4** (155 mg, 86% yield; R_f 0.30, ethyl acetate:n-hexane 30:70; colorless liquid). ^{1}H NMR (500 MHz, $CDCl_3$) δ 11.52 (bs, 1H), 7.31-7.24 (m, 2H), 6.97-6.92 (m, 1H), 6.91-6.86 (m, 2H), 4.02 (t, *J* = 6.1 Hz, 2H), 2.59 (t, *J* = 7.3 Hz, 2H), 2.16-2.08 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 179.0, 158.9, 129.6, 121.0, 114.6, 66.5, 30.6, 24.5. MS (ESI) $[M+H]^+$ calcd. $C_{10}H_{13}O_3$ 181.1, found 181.1.

Synthesis of oxiran-2-ylmethyl 4-phenoxybutanoate 2a



In a 25 ml round bottom flask, 4-phenoxybutanoic acid **S4** (224 mg, 1 mmol), DCC (227 mg, 1.1 mmol) and DMAP (37 mg, 0.3 mmol) were dissolved in anh. DMF:CH_2Cl_2 (1:4, 10 ml) at 0 ^oC. After 10 minutes, glycidol **S5** (133 µl, 2 mmol) was added to the reaction mixture and the stirring was continued at room temperature for 24 h. The reaction mixture was concentrated in vacuo, diluted with ethyl acetate, and filtered to remove DCU. The filtrate was washed with brine, dried over anh. Na_2SO_4, and concentrated in vacuo. The crude mixture was purified by silica gel flash column chromatography using ethyl acetate:n-hexane (10:90) to yield oxiran-2-ylmethyl 4-phenoxybutanoate **2a** (83 mg, 35% yield; R_f 0.60, ethyl acetate:n-hexane 30:70; colorless viscous liquid). ^1H NMR (400 MHz, $CDCl_3$) δ 7.32-7.22 (m, 2H), 7.00-6.84 (m, 3H), 4.43 (dd, *J* = 12.3, 3.0 Hz, 1H), 4.01 (t, *J* = 6.0 Hz, 2H), 3.94 (dd, *J* = 12.3, 6.3 Hz, 1H), 3.25-3.16 (m, 1H), 2.88-2.80 (m, 1H), 2.64 (dd, *J* = 4.8, 2.6 Hz, 1H), 2.59 (t, *J* = 7.3 Hz, 2H), 2.19-2.08 (m, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.1, 158.9, 129.6, 120.9, 114.6, 66.6, 65.1, 49.5, 44.8, 30.7, 24.7 ppm. HRMS (ESI) $[M+Na]^+$ calcd. $C_{13}H_{16}NaO_4$ 259.0946, found 259.0963.

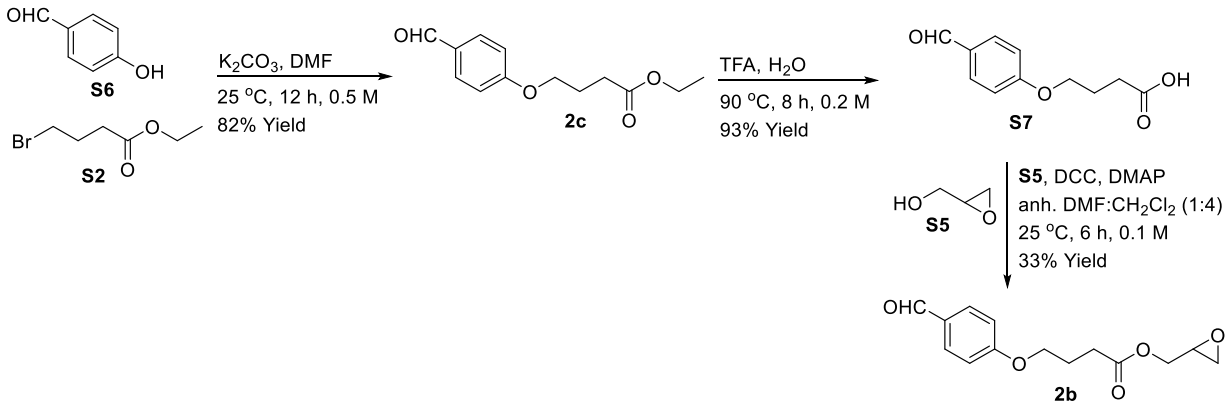
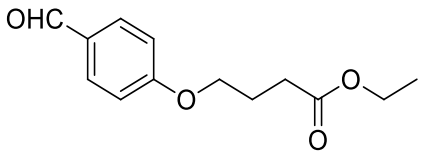


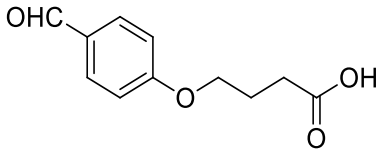
Figure S2. Synthesis of oxiran-2-ylmethyl 4-(4-formyl-3-hydroxyphenoxy)butanoate **2b**.

Ethyl 4-(4-formylphenoxy)butanoate 2c[4]



In a 25 ml round bottom flask, p-hydroxybenzaldehyde **S6** (611 mg, 5 mmol), ethyl 4-bromobutyrate **S2** (859 µl, 6 mmol) and K_2CO_3 (1.037 g, 7.5 mmol) were dissolved in DMF (10 ml). The reaction mixture was stirred at room temperature for 12 h followed by filtration to remove potassium carbonate. The filtrate was extracted using ethyl acetate:n-hexane (4:6, 3 × 100 ml) and the combined organic layers were washed with brine solution. The organic layer was dried over anh. Na_2SO_4, concentrated, and the residue was purified using silica gel flash column chromatography (ethyl acetate:n-hexane 3:97; R_f 0.59, ethyl acetate:n-hexane 35:65; colorless liquid) to give **2c** (969 mg, 82% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.88 (s, 1H), 7.83 (d, *J* = 8.7 Hz, 2H), 6.99 (d, *J* = 8.7 Hz, 2H), 4.15 (q, *J* = 7.2 Hz, 2H), 4.11 (t, *J* = 6.2 Hz, 2H), 2.53 (t, *J* = 7.2 Hz, 2H), 2.21-2.10 (m, 2H), 1.26 (t, *J* = 7.1 Hz, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 190.9, 173.1, 164.0, 132.1, 130.1, 114.8, 67.2, 60.6, 30.7, 24.5, 14.3 ppm. MS (ESI) $[M+H]^+$ calcd. $C_{13}H_{17}O_4$ 237.1, found 237.1.

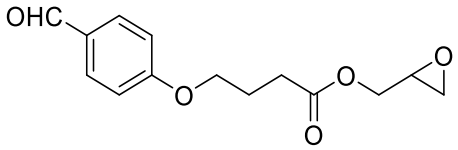
4-(4-formylphenoxy)butanoic acid S7[5]



This compound was synthesized according to the procedure for synthesis of **S4**.

Yield 93%; R_f 0.39, ethyl acetate/hexane 50:50; white solid. ^{1}H NMR (400 MHz, DMSO-d_6) δ 9.86 (s, 1H), 7.85 (d, *J* = 8.7 Hz, 2H), 7.11 (d, *J* = 8.6 Hz, 2H), 4.10 (t, *J* = 6.4 Hz, 2H), 2.39 (t, *J* = 7.3 Hz, 2H), 2.04-1.91 (m, 2H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6) δ 191.3, 174.1, 163.6, 131.9, 129.6, 115.0, 67.2, 30.1, 24.1 ppm. MS (ESI) $[M+H]^+$ calcd. $C_{11}H_{13}O_4$ 209.1, found 209.0.

Oxiran-2-ylmethyl 4-(4-formylphenoxy)butanoate 2b



This compound was synthesized according to the procedure for synthesis of **2a**.

Yield 33%; R_f 0.43, ethyl acetate:n-hexane 30:70; colorless viscous liquid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.88 (s, 1H), 7.83 (d, *J* = 8.8 Hz, 2H), 6.99 (d, *J* = 8.7 Hz, 2H), 4.46 (dd, *J* = 12.3, 2.9 Hz, 1H), 4.11 (t, *J* = 6.1 Hz, 2H), 3.93 (dd, *J* = 12.3, 6.4 Hz, 1H), 3.25-3.17 (m, 1H), 2.88-2.81 (m, 1H), 2.63 (dd, *J* = 4.9, 2.6 Hz, 1H), 2.58 (t, *J* = 7.2 Hz, 2H), 2.20-2.10 (m, 2H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 190.9, 172.8, 163.9, 132.1, 130.1, 114.8, 67.1, 65.2, 49.4, 44.7, 30.5, 24.4 ppm. HRMS (ESI) $[M+H]^+$ calcd. $C_{14}H_{17}O_5$ 265.1076, found 265.1084.

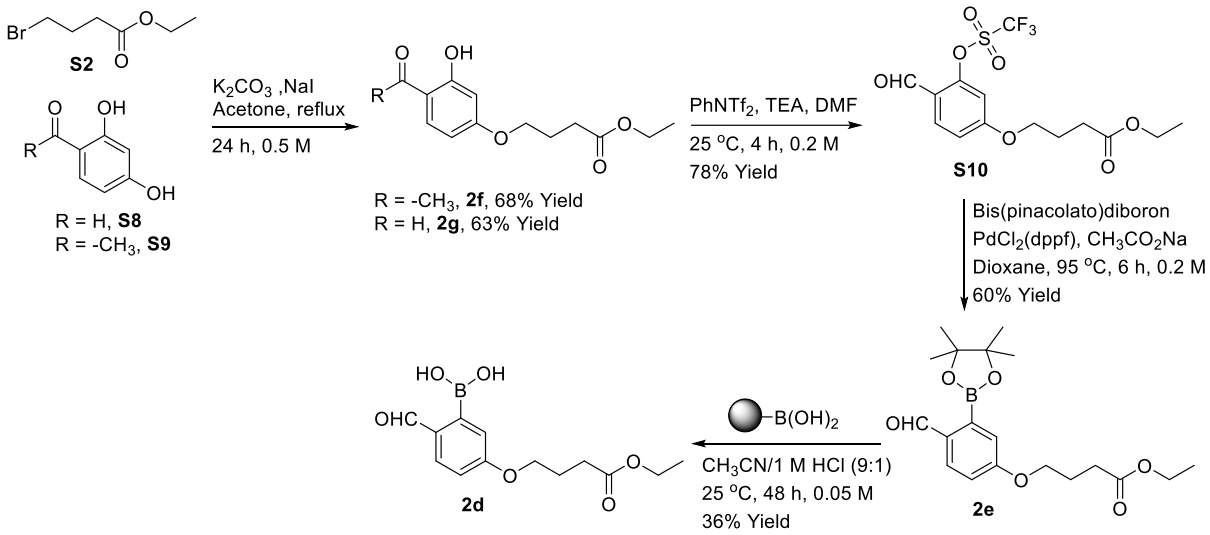
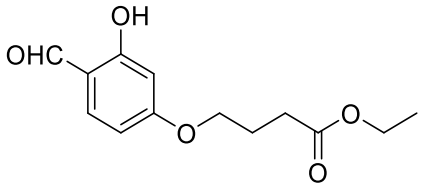


Figure S3. Synthesis of (5-(4-ethoxy-4-oxobutoxy)-2-formylphenyl)boronic acid **2d.**[6]

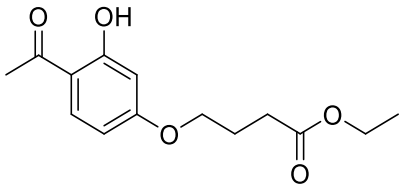
Synthesis of ethyl 4-(4-formyl-3-hydroxyphenoxy)butanoate 2g[7]



In a 25 ml round bottom flask 1 (RB1), 2,4-dihydroxybenzaldehyde **S8** (691 mg, 5 mmol) and K_2CO_3 (691 mg, 5 mmol) were dissolved in acetone (5 ml) and refluxed for 2 h. Simultaneously

in another 10 ml round bottom flask (RB2), ethyl 4-bromobutyrate **S2** (716 µl, 5 mmol) and sodium iodide (825 mg, 5.5 mmol) in acetone (5 ml) were stirred at room temperature for 2 h. The solution in RB2 was transferred to RB1 and refluxed for another 24 h. Acetone was removed using rotary evaporator and the crude mixture was re-suspended in ethyl acetate and water. After solvent-solvent extraction, the combined organic layers were dried with anh. sodium sulfate, filtered, and concentrated under vacuum. The crude mixture was purified by silica gel flash column chromatography using ethyl acetate:n-hexane (3:97) to afford ethyl 4-(4-formyl-3-hydroxyphenoxy)butanoate **2g** (795 mg, 63% yield; R_f 0.44, ethyl acetate:n-hexane 10:90; white solid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.46 (s, 1H), 9.71 (s, 1H), 7.42 (d, *J* = 8.7 Hz, 1H), 6.52 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.41 (d, *J* = 2.3 Hz, 1H), 4.15 (q, *J* = 7.1 Hz, 2H), 4.06 (t, *J* = 6.2 Hz, 2H), 2.50 (t, *J* = 7.2 Hz, 2H), 2.18-2.08 (m, 2H), 1.26 (t, *J* = 7.1 Hz, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 194.5, 173.0, 166.2, 164.6, 135.4, 115.3, 108.7, 101.3, 67.4, 60.7, 30.7, 24.4, 14.3 ppm. MS (ESI) $[M+H]^+$ calcd. $C_{13}H_{17}O_5$ 253.1, found 253.1.

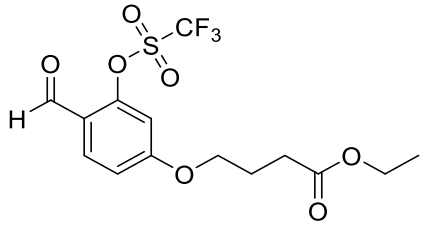
Synthesis of ethyl 4-(4-acetyl-3-hydroxyphenoxy)butanoate 2f [6]



This compound was synthesized according to the procedure for synthesis of **2g**.

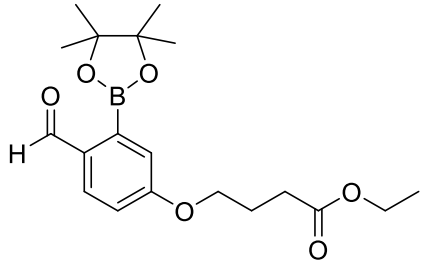
Yield 68%; R_f 0.78, ethyl acetate:n-hexane 30:70; white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 12.70 (s, 1H), 7.61 (d, *J* = 8.8 Hz, 1H), 6.41 (dd, *J* = 8.8, 2.5 Hz, 1H), 6.38 (d, *J* = 2.4 Hz, 1H), 4.14 (q, *J* = 7.1 Hz, 2H), 4.03 (t, *J* = 6.2 Hz, 2H), 2.54 (s, 3H), 2.49 (t, *J* = 7.3 Hz, 2H), 2.16-2.07 (m, 2H) 1.25 (t, *J* = 7.1 Hz, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 202.7, 173.1, 165.5, 165.3, 132.4, 114.1, 107.9, 101.5, 67.2, 60.6, 30.8, 26.3, 24.5, 14.3 ppm. MS (ESI) $[M+H]^+$ calcd. $C_{14}H_{19}O_5$ 267.1, found 267.2.

Synthesis of ethyl 4-(4-formyl-3-(((trifluoromethyl)sulfonyl)oxy)phenoxy)butanoate S10



In a 10 ml round bottom flask, ethyl 4-(4-formyl-3-hydroxyphenoxy)butanoate **2g** (555 mg, 2.2 mmol) and triethylamine (920 µl, 6.6 mmol) were dissolved in anhydrous DMF (4 ml). To this mixture, N-phenyl-trifluoromethanesulfonimide (715 mg, 2 mmol) was added portion wise and solution was stirred at room temperature for 4 h. The progress of reaction was monitored by thin layer chromatography. The solvent-solvent extraction of the crude reaction mixture was performed using ethyl acetate:n-hexane (40:60) as organic medium and water. The choice of organic medium ensures removal of non-volatile DMF. The organic layer was dried with anh. Na_2SO_4, concentrated in vacuo. The purification of crude reaction mixture was performed by silica gel flash column chromatography using ethyl acetate:n-hexane (5:95) to isolate **S10** (599 mg, 78% yield; R_f 0.45, ethyl acetate:n-hexane 30:70; brownish liquid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.12 (s, 1H), 7.94 (d, *J* = 8.7 Hz, 1H), 7.01 (dd, *J* = 8.7, 2.1 Hz, 1H), 6.87 (d, *J* = 2.2 Hz, 1H), 4.22-4.07 (m, 4H), 2.53 (t, *J* = 7.1 Hz, 2H), 2.23-2.10 (m, 2H), 1.27 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 185.5, 172.9, 164.7, 151.4, 132.4, 122.0, 118.8 (q, -$SO_2\underline{C}F_3$), 114.7, 108.8, 68.2, 60.8, 30.5, 24.3, 14.3. HRMS (ESI) $[M+H]^+$ calcd. $C_{14}H_{16}F_3O_7S$ 385.0569, found 385.0586.

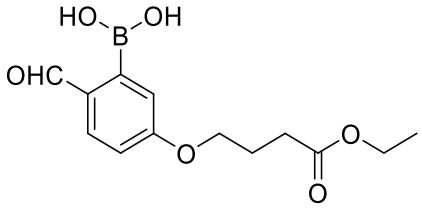
Synthesis of ethyl 4-(4-formyl-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy) butanoate 2e



In a 15 ml Schlenk tube, ethyl 4-(4-formyl-3-(((trifluoromethyl)sulfonyl)oxy)phenoxy) butanoate **S10** (384 mg, 1 mmol), bis(pinacolato)diboron (507 mg, 2 mmol), [1,1'-bis(diphenylphosphino)

ferrocene]palladium (II) chloride (73 mg, 0.1 mmol) and anh. sodium acetate (246 mg, 3 mmol) were dissolved in degassed 1,4-dioxane (5 ml) under inert atmosphere. The reaction mixture was further degassed with nitrogen for 10 minutes and heated at 95 ^{o}C for 6 h. The reaction mixture was concentrated in vacuo and purified by silica gel flash column chromatography using ethyl acetate:n-hexane (5:95) to isolate **2e** (217 mg, 60% yield; R_f 0.63, ethyl acetate:n-hexane 40:60; colorless viscous liquid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.36 (s, 1H), 7.92 (d, *J* = 8.6 Hz, 1H), 7.27 (d, *J* = 2.5 Hz, 1H), 7.00 (dd, *J* = 8.6, 2.5 Hz, 1H), 4.22-4.06 (m, 4H), 2.52 (t, *J* = 7.2 Hz, 2H), 2.20-2.08 (m, 2H), 1.39 (s, 12H), 1.27 (t, *J* = 7.1 Hz, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 193.1, 173.2, 162.8, 134.8, 130.8, 120.7, 116.5, 84.6, 67.1, 60.6, 30.7, 25.0, 24.6, 14.4 ppm. HRMS (ESI) $[M+H]^+$ calcd. $C_{19}H_{28}BO_6$ 363.1979, found 363.1999.

Synthesis of (5-(4-ethoxy-4-oxobutoxy)-2-formylphenyl)boronic acid 2d[8]



In a 5 ml Eppendorf tube, boronic acid (polymer bound, 200 mg, 1-2 mmol/g) and ethyl 4-(4-formyl-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)butanoate **2e** (36 mg, 0.1 mmol) were taken. To this mixture, CH_3CN/1 M HCl (9:1, 2 ml) was added followed by end to end rotation at 25 ^{o}C for 48 h. The purification of crude reaction mixture was performed by reverse phase preparative HPLC to isolate **2d** (10 mg, 36% yield; colorless viscous liquid). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.74 (s, 1H), 7.92 (bs, 2H), 7.84 (d, *J* = 8.5 Hz, 1H), 7.81 (d, *J* = 2.6 Hz, 1H), 7.11 (dd, *J* = 8.5, 2.7 Hz, 1H), 4.21-4.12 (m, 4H), 2.54 (t, *J* = 7.2 Hz, 2H), 2.23-2.10 (m, 2H), 1.27 (t, *J* = 7.1 Hz, 3H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 196.6, 173.1, 163.6, 141.8, 133.1, 125.4, 115.9, 67.3, 60.7, 30.7, 24.5, 14.4 ppm. ^{13}C-HMBC 136.4 ($\underline{C}_{arom}$B). HRMS (ESI) $[M+Na]^+$ calcd. $C_{13}H_{17}BNaO_6$ 303.1016, found 303.0991.

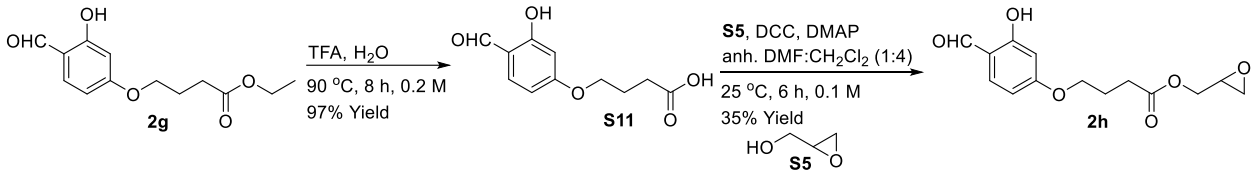
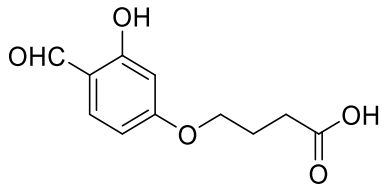


Figure S4. Synthesis of oxiran-2-ylmethyl 4-(4-formyl-3-hydroxyphenoxy)butanoate **2h.**

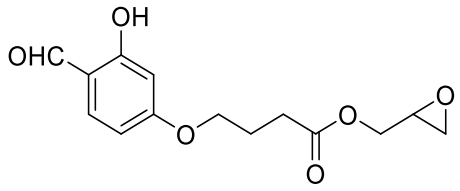
4-(4-formyl-3-hydroxyphenoxy)butanoic acid S11[9]



This compound was synthesized according to the procedure for synthesis of **S4**.

Yield 97%; R_f 0.30, ethyl acetate:n-hexane 30:70; white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.46 (s, 1H), 9.72 (s, 1H), 7.43 (d, *J* = 8.7 Hz, 1H), 6.53 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.42 (d, *J* = 2.2 Hz, 1H), 4.09 (t, *J* = 6.1 Hz, 2H), 2.59 (t, *J* = 7.2 Hz, 2H), 2.21-2.10 (m, 2H) ppm. ^{13}C NMR (101 MHz, DMSO-*d*$_6$) δ 191.1, 174.0, 165.2, 163.1, 132.3, 116.2, 107.7, 101.2, 67.2, 30.0, 24.0 ppm. MS (ESI) $[M+H]^+$ calcd. $C_{11}H_{13}O_5$ 225.1, found 225.1.

Oxiran-2-ylmethyl 4-(4-formyl-3-hydroxyphenoxy)butanoate 2h



In a 25 ml round bottom flask, 4-(4-formyl-3-hydroxyphenoxy)butanoic acid **S11** (224 mg, 1 mmol), DCC (227 mg, 1.1 mmol) and DMAP (37 mg, 0.3 mmol) were dissolved in anh. CH_2Cl_2:DMF (10 ml, 4:1). To this solution, glycidol **S5** (133 µl, 2 mmol) was added and stirred at room temperature for 6 h. The reaction mixture was filtered and the filtrate was concentrated in vacuo. The residue was purified by silica gel flash column chromatography using ethyl acetate:n-hexane (20:80) to give **2h** (98 mg, 35% yield; R_f 0.30, ethyl acetate:n-hexane 30:70; white solid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.45 (s, 1H), 9.71 (s, 1H), 7.43 (d, *J* = 8.7 Hz, 1H), 6.53 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.41 (d, *J* = 2.1 Hz, 1H), 4.45 (dd, *J* = 12.3, 2.9 Hz, 1H), 4.08 (t, *J*

= 6.1, 2H), 3.94 (dd, *J* = 12.3, 6.4 Hz, 1H), 3.27-3.16 (m, 1H), 2.88-2.81 (m, 1H), 2.65 (dd, *J* = 4.8, 2.6 Hz, 1H), 2.58 (t, *J* = 7.2 Hz, 2H), 2.24-2.09 (m, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 194.5, 172.7, 166.1, 164.6, 135.4, 115.4, 108.7, 101.4, 67.3, 65.3, 49.4, 44.8, 30.5, 24.4 ppm. HRMS (ESI) $[M+Na]^+$ calcd. $C_{14}H_{16}NaO_6$ 303.0845, found 303.0843.

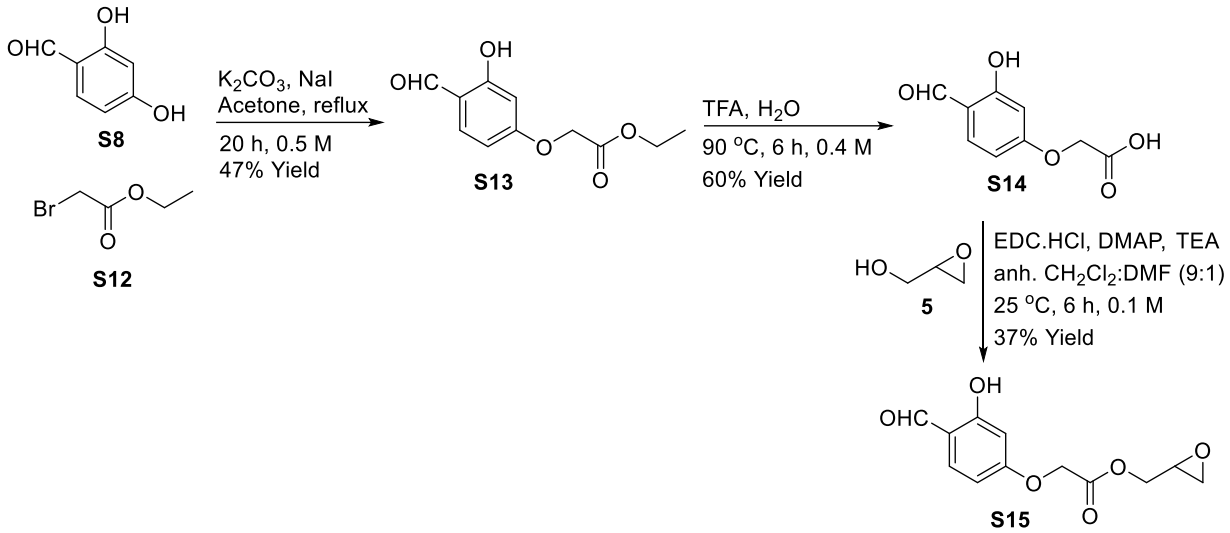
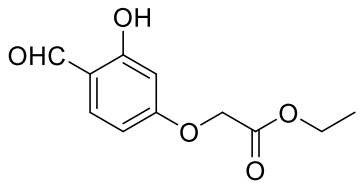


Figure S5. Synthesis of oxiran-2-ylmethyl 2-(4-formyl-3-hydroxyphenoxy)acetate **S15**.

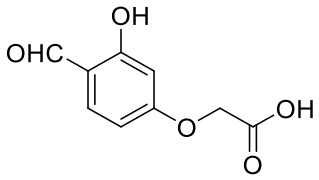
Synthesis of ethyl 2-(4-formyl-3-hydroxyphenoxy)acetate S13[10]



This compound was synthesized according to the procedure for synthesis of **2g.**

Yield 47%; R_f 0.60, ethyl acetate:n-hexane 40:60; colorless liquid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.44 (s, 1H), 9.74 (s, 1H), 7.47 (d, *J* = 8.7 Hz, 1H), 6.59 (dd, *J* = 8.7, 2.4 Hz, 1H), 6.39 (d, *J* = 2.3 Hz, 1H), 4.67 (s, 2H), 4.29 (q, *J* = 7.1 Hz, 2H), 1.31 (t, *J* = 7.1 Hz, 3H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 194.7, 167.9, 164.9, 164.4, 135.6, 116.0, 108.6, 101.6, 65.3, 61.9, 14.3 ppm. MS (ESI) $[M+H]^+$ calcd. $C_{11}H_{13}O_5$ 225.1, found 225.0.

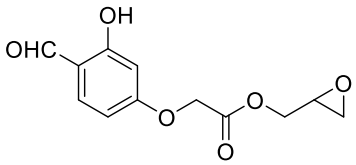
Synthesis of 2-(4-formyl-3-hydroxyphenoxy)acetic acid S14[10]



This compound was synthesized according to the procedure for synthesis of **S4**.

Yield 60%, ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.98 (s, 1H), 10.02 (s, 1H), 7.62 (d, J = 8.7 Hz, 1H), 6.55 (dd, J = 8.7, 2.2 Hz, 1H), 6.44 (d, J = 2.3 Hz, 1H), 4.76 (s, 2H) ppm. ^{13}C NMR (101 MHz, DMSO-d_6) δ 190.9, 169.6, 164.4, 162.9, 132.0, 116.6, 107.6, 101.5, 64.7 ppm. MS (ESI) $[M+H]^+$ calcd. $C_9H_9O_5$ 197.0, found 197.0.

Synthesis of oxiran-2-ylmethyl 2-(4-formyl-3-hydroxyphenoxy)acetate S15



In a 10 ml round bottom flask, 2-(4-formyl-3-hydroxyphenoxy)acetic acid **S14** (98 mg, 0.5 mmol), EDC.HCl (105 mg, 0.55 mmol), DMAP (18 mg, 0.15 mmol), and triethylamine (140 µl, 1 mmol) were dissolved in anh. CH_2Cl_2:DMF (5 ml, 9:1). To this solution, glycidol **5** (66 µl, 1 mmol) was added and stirred at room temperature for 6 h. The reaction mixture was concentrated in vacuo. The residue was purified by silica gel flash column chromatography using ethyl acetate:n-hexane (25:75) to give **S15** (47 mg, 37% yield; R_f 0.83, ethyl acetate:n-hexane 40:60; white solid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.43 (s, 1H), 9.74 (s, 1H), 7.48 (d, J = 8.7 Hz, 1H), 6.59 (dd, J = 8.7, 2.4 Hz, 1H), 6.40 (d, J = 2.3 Hz, 1H), 4.74 (s, 2H), 4.58 (dd, J = 12.2, 2.9 Hz, 1H), 4.04 (dd, J = 12.2, 6.4 Hz, 1H), 3.32-3.18 (m, 1H), 2.91-2.82 (m, 1H), 2.66 (dd, J = 4.8, 2.6 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 194.7, 167.7, 164.7, 164.4, 135.6, 116.0, 108.5, 101.7, 66.1, 65.0, 49.1, 44.7 ppm. HRMS (ESI) $[M+Na]^+$ calcd. $C_{12}H_{12}NaO_6$ 275.0532, found 275.0528.

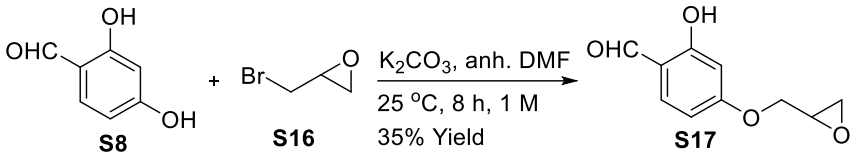
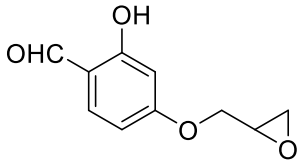


Figure S6. Synthesis of 2-hydroxy-4-(oxiran-2-ylmethoxy)benzaldehyde **S17**.

Synthesis of 2-hydroxy-4-(oxiran-2-ylmethoxy)benzaldehyde S17



In a 5 ml round bottom flask, 2, 4-dihydroxybenzaldehyde **S8** (138 mg, 1 mmol) and K_2CO_3 (138 mg, 1 mmol) were dissolved in anh. DMF (1 ml). To this solution, epibromohydrin **S16** (86 µl, 1 mmol) was added and stirred at room temperature. After 8 h, the solvent-solvent extraction was performed using ethyl acetate:n-hexane (30:70) as organic medium and water work up. The collected organic layers were dried with anh. sodium sulphate, filtered, and concentrated in vacuo. The residue was purified by silica gel flash column chromatography using ethyl acetate:n-hexane (3:97) to give **S17** (68 mg, 35% yield; R_f 0.50, ethyl acetate:n-hexane 40:60; brown semisolid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.45 (s, 1H), 9.73 (s, 1H), 7.44 (d, *J* = 8.7 Hz, 1H), 6.57 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.43 (d, *J* = 2.2 Hz, 1H), 4.31 (dd, *J* = 11.1, 2.9 Hz, 1H), 3.97 (dd, *J* = 11.1, 5.9 Hz, 1H), 3.41-3.32 (m, 1H), 2.97-2.90 (m, 1H), 2.76 (dd, *J* = 4.8, 2.6 Hz, 1H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 194.6, 165.6, 164.5, 135.5, 115.7, 108.8, 101.5, 69.2, 49.8, 44.7 ppm. HRMS (ESI) $[M+H]^+$ calcd. $C_{10}H_{11}O_4$ 195.0657, found 195.0650.

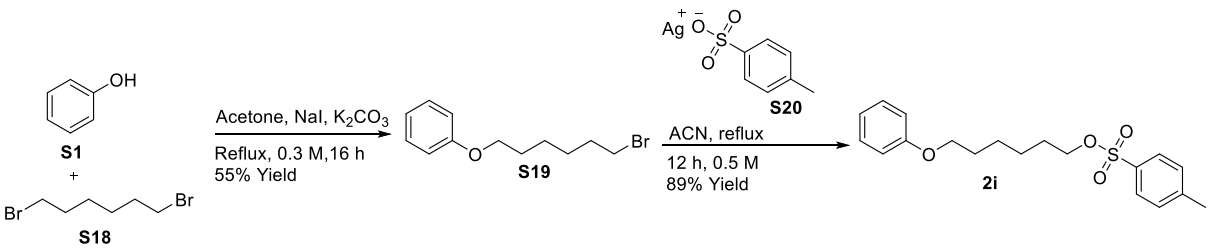
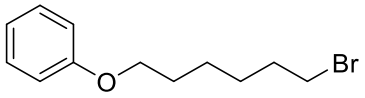


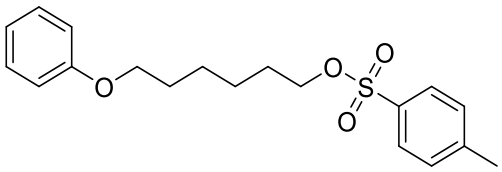
Figure S7. Synthesis of 6-phenoxyhexyl 4-methylbenzenesulfonate **2i**.

Synthesis of ((6-bromohexyl)oxy)benzene S19[11]



In 25 ml round bottom flask, phenol **S1** (188 mg, 2 mmol), potassium carbonate (276 mg, 2 mmol), sodium iodide (298 mg, 2 mmol), and 1,6-dibromohexane **S18** (303 µl, 2 mmol) were dissolved in 10 ml acetone. The reaction mixture was refluxed and the progress of reaction was monitored by TLC. After 16 h, reaction mixture was filtered, concentrated, and purified by silica gel flash column chromatography using n-hexane to isolate **S19** (382 mg, 55% yield; R_f 0.69, ethyl acetate:n-hexane 10:90; colorless liquid). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.31-7.26 (m, 2H), 6.97-6.91 (m, 1H), 6.91-6.85 (m, 2H), 3.96 (t, *J* = 6.4 Hz, 2H), 3.42 (t, *J* = 6.8 Hz, 2H), 1.97-1.85 (m, 2H), 1.85-1.75 (m, 2H), 1.54-1.44 (m, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 159.2, 129.6, 120.7, 114.6, 67.7, 33.9, 32.8, 29.3, 28.1, 25.5. MS (ESI) $[M+H]^+$ calcd. $C_{12}H_{18}BrO$ 257.1, found 257.1.

Synthesis of 6-phenoxyhexyl 4-methylbenzenesulfonate 2i



In 10 ml round bottom flask, ((6-bromohexyl)oxy)benzene **S19** (280 mg, 1.1 mmol), silver tosylate **S20** (337 mg, 1.21 mmol) were taken and dissolved in 5 ml acetone to reflux. The progress of reaction was monitored by TLC. After 12 h, the reaction mixture was filtered, concentrated, and purified by silica gel flash column chromatography (ethyl acetate:n-hexane, 6:94) to isolate **2i** (340 mg, 89% yield; R_f 0.25, ethyl acetate:n-hexane 10:90; White solid). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79 (d, *J* = 8.3 Hz, 2H), 7.33 (d, *J* = 8.0 Hz, 2H), 7.30-7.24 (m, 2H), 6.96-6.90 (m, 1H), 6.90-6.84 (m, 2H), 4.04 (t, *J* = 6.5 Hz, 2H), 3.91 (t, *J* = 6.4 Hz, 2H), 2.44 (s, 3H), 1.78-1.62 (m, 4H), 1.47-1.33 (m, 4H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 159.1, 144.8, 133.3, 129.9, 129.5, 128.0, 120.7, 114.6, 70.6, 67.6, 29.2, 28.9, 25.6, 25.3, 21.8 ppm. HRMS (ESI) $[M+Na]^+$ calcd. $C_{19}H_{24}O_4SNa$ 371.1293, found 371.1301.

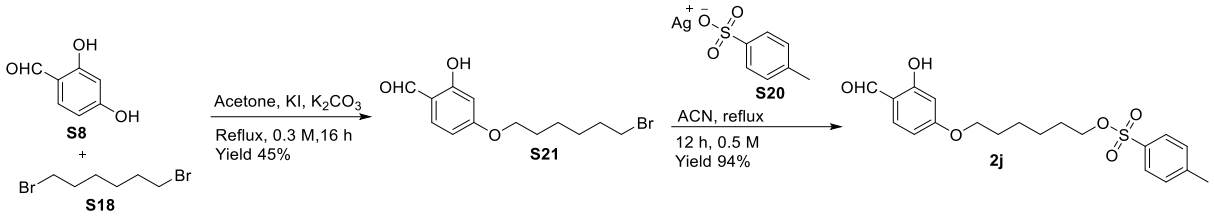
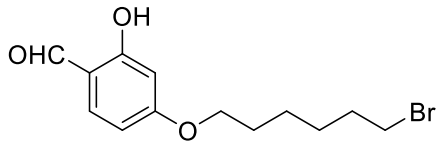


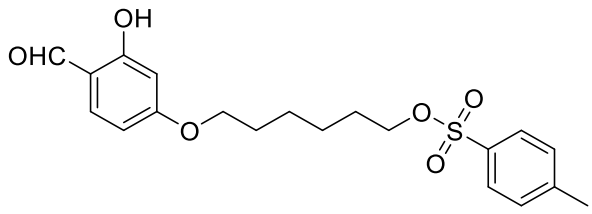
Figure S8. Synthesis of 6-(4-formyl-3-hydroxyphenoxy)hexyl 4-methylbenzenesulfonate **2j**.

4-((6-bromohexyl)oxy)-2-hydroxybenzaldehyde S21



In 25 ml round bottom flask, 2,4-dihyroxybenzaldehyde **S8** (276 mg, 2 mmol), potassium carbonate (276 mg, 2 mmol), sodium iodide (298 mg, 2 mmol), and 1, 6 dibromohexane **S18** (303µl, 2mmol) were dissolved in 10 ml acetone to reflux. The progress of reaction was monitored by TLC. After 16 h, the reaction mixture was filtered, concentrated and purified by silica gel flash column chromatography (ethyl acetate:n-hexane, 2:98) to isolate **S21** (270 mg, 45% yield; R_f 0.53, ethyl acetate:n-hexane 30:70; white solid). ^{1}H NMR (500 MHz, $CDCl_3$) δ 11.48 (s, 1H), 9.70 (s, 1H), 7.42 (d, *J* = 8.7 Hz, 1H), 6.52 (dd, *J* = 8.5, 2.3 Hz, 1H), 6.40 (d, *J* = 2.3 Hz, 1H), 4.01 (t, *J* = 6.4 Hz, 2H), 3.42 (t, *J* = 6.8 Hz, 2H), 1.96-1.86 (m, 2H), 1.85-1.77 (m, 2H), 1.58-1.43 (m, 4H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 194.4, 166.5, 164.6, 135.3, 115.2, 108.8, 101.2, 68.4, 33.8, 32.7, 28.9, 28.0, 25.3 ppm. HRMS (ESI) $[M+H]^+$ calcd. $C_{13}H_{18}BrO_3$ 301.0439, found 301.0453.

6-(4-formyl-3-hydroxyphenoxy)hexyl 4-methylbenzenesulfonate 2j



In a 5 ml round bottom flask of 4-((6-bromohexyl)oxy)-2-hydroxybenzaldehyde **S21** (360 mg, 1.2 mmol) and silver tosylate **S20** (363 mg, 1.3 mmol) were dissolved in 5 ml acetone. The reaction mixture was refluxed and the progress of reaction was monitored by TLC. After 12 h,

the reaction mixture was filtered, concentrated, and purified by silica gel flash column chromatography (ethyl acetate:n-hexane, 15:85) to isolate **2j** (442 mg, 94% yield; R_f 0.57, ethyl acetate:n-hexane 30:70; white solid). ^{1}H NMR (500 MHz, $CDCl_3$) δ 11.47 (s, 1H), 9.71 (s, 1H), 7.78 (d, *J* = 8.3 Hz, 2H), 7.42 (d, *J* = 8.7 Hz, 1H), 7.34 (d, *J* = 8.0 Hz, 2H), 6.50 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.38 (d, *J* = 2.3 Hz, 1H), 4.04 (t, *J* = 6.4 Hz, 2H), 3.96 (t, *J* = 6.4 Hz, 2H), 2.44 (s, 3H), 1.79-1.71 (m, 2H), 1.71-1.64 (m, 2H), 1.49-1.32 (m, 4H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 194.4, 166.3, 164.4, 144.8, 135.3, 133.1, 129.9, 127.8, 115.1, 108.6, 101.1, 70.5, 68.2, 28.7, 28.7, 25.3, 25.1, 21.6 ppm. HRMS (ESI) $[M+H]^+$ calcd. $C_{20}H_{25}O_6S$ 393.1372, found 393.1390.

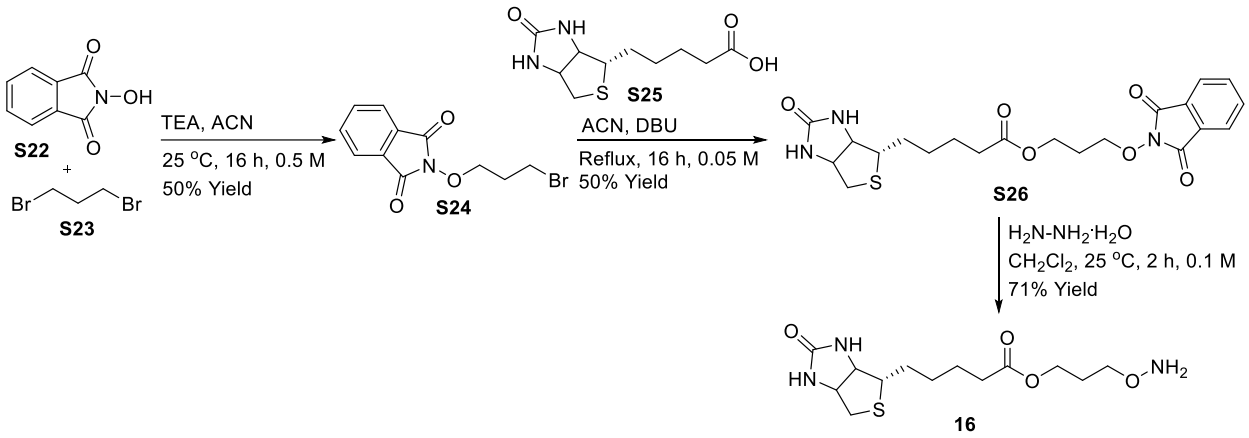
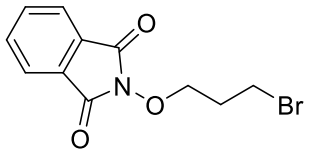


Figure S9. Synthesis of 3-(aminooxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate **16.**

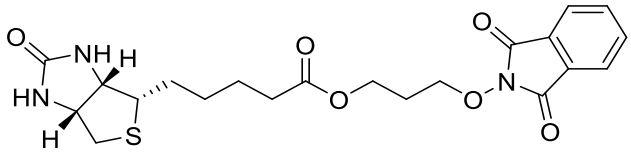
Synthesis of 2-(3-bromopropoxy)isoindoline-1,3-dione S24[12]



In a 250 ml round bottom flask, N-hydroxyphthalimide **S22** (4.894 g, 30 mmol) and triethyl amine (6.09 ml, 60 mmol) were dissolved in acetonitrile (60 ml). To this solution, 1,3-dibromopropane **S23** (8.34 ml, 60 mmol) was added and stirred at 25 °C for 16 h. The reaction mixture was concentrated in vacuo followed by addition of 1 N NaOH solution and ethyl acetate. The organic layer was separated, dried over anh. sodium sulfate, filtered, and concentrated in vacuo. The purification of the crude reaction mixture was performed by silica gel flash column chromatography using ethyl acetate:hexane (3:97) to isolate **S24** (4.259 g, 50% yield; R_f 0.57,

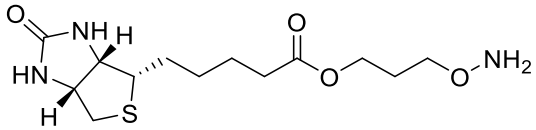
ethyl acetate:n-hexane 30:70; white solid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.89-7.81 (m, 2H), 7.80-7.73 (m, 2H), 4.37 (t, *J* = 5.8 Hz, 2H), 3.71 (t, *J* = 6.5 Hz, 2H), 2.36-2.26 (m, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.7, 134.7, 129.0, 123.7, 76.2, 31.6, 29.4 ppm. MS (ESI) $[M+H]^+$ calcd. $C_{11}H_{11}{}^{79}BrNO_3$ 284.0, found 283.9 and calcd. $C_{11}H_{11}{}^{81}BrNO_3$ 286.0, found 285.9.

Synthesis of 3-((1,3-dioxoisoindolin-2-yl)oxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate S26



In a 5 ml round bottom flask, biotin **S25** (244 mg, 1 mmol), 2-(3-bromopropoxy)isoindoline-1,3-dione **S24** (568 mg, 2 mmol), and DBU (304 µl, 2 mmol) were dissolved in acetonitrile (20 ml). The reaction mixture was refluxed and the progress of the reaction was monitored by TLC. After 16 h, the reaction mixture was concentrated in vacuum. This was followed by solvent-solvent extraction using ethyl acetate and water. The organic fractions were combined, dried on anh. sodium sulfate, filtered, and concentrated on rotary evaporator. The purification of crude reaction mixture was performed by silica gel flash column chromatography (MeOH:DCM, 0.5-5%) to isolate 3-((1,3-dioxoisoindolin-2-yl)oxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate **S26** (224 mg, 50% yield; R_f 0.33, MeOH:DCM 05:95; white solid). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.89-7.82 (m, 2H), 7.80-7.73 (m, 2H), 5.96 (s, 1H), 5.48 (s, 1H), 4.56-4.47 (m, 1H), 4.38-4.26 (m, 5H), 3.23-3.13 (m, 1H), 2.92 (dd, *J* = 12.8, 5.0 Hz, 1H), 2.74 (d, *J* = 12.8 Hz, 1H), 2.36 (t, *J* = 7.4 Hz, 2H), 2.16-2.09 (m, 2H), 1.80-1.60 (m, 4H), 1.54-1.39 (m, 2H) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 173.7, 163.8, 163.7, 134.7, 129.0, 123.7, 75.1, 62.0, 60.7, 60.2, 55.5, 40.7, 34.0, 28.4, 28.3, 27.8, 24.9 ppm. HRMS (ESI) $[M+H]^+$ calcd. $C_{21}H_{26}N_3O_6S$ 448.1542, found 448.1548.

Synthesis of 3-(aminooxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate 16



In a 5 ml round bottom flask, 3-((1,3-dioxoisoindolin-2-yl)oxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate **S26** (134 mg, 0.3 mmol) in DCM (3 ml) and hydrazine monohydrate (80%, 37 µl, 0.75 mmol) were stirred at room temperature. The progress of the reaction was followed by TLC. After 2 h, the reaction mixture was filtered and concentration of the filtrate in vacuo led to the isolation of 3-(aminooxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate **16** (67 mg, 71% yield; R_f 0.21, MeOH:DCM 05:95; white solid. ^{1}H NMR (500 MHz, D_2O) δ 4.63 (dd, *J* = 7.9, 4.9 Hz, 1H), 4.45 (dd, *J* = 7.9, 4.5 Hz, 1H), 4.21 (t, *J* = 6.3 Hz, 2H), 3.90 (t, *J* = 6.2 Hz, 2H), 3.45-3.27 (m, 1H), 3.02 (dd, *J* = 13.1, 5.0 Hz, 1H), 2.80 (d, *J* = 13.0 Hz, 1H), 2.44 (t, *J* = 7.3 Hz, 2H), 2.08-1.94 (m, 2H), 1.84-1.55 (m, 4H), 1.53-1.37 (m, 2H) ppm. ^{13}C NMR (126 MHz, D_2O) δ 176.9, 165.3, 72.4, 62.1, 62.0, 60.3, 55.3, 39.7, 33.6, 27.9, 27.6, 26.7, 24.1 ppm. HRMS (ESI) $[M+H]^+$ calcd. $C_{13}H_{24}N_3O_4S$ 318.1488, found 318.1467.

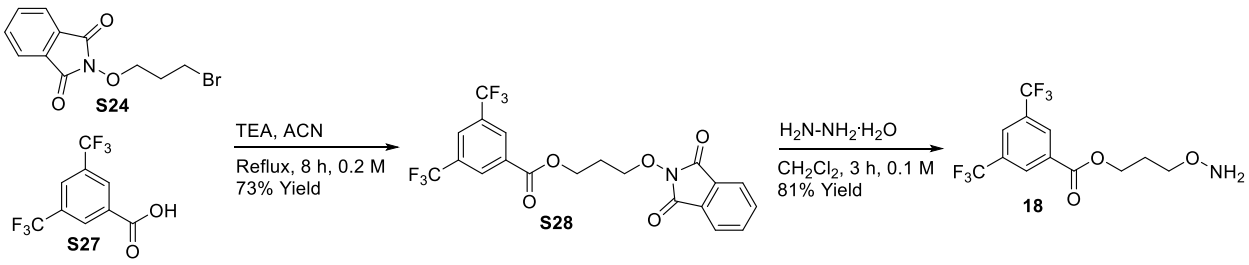
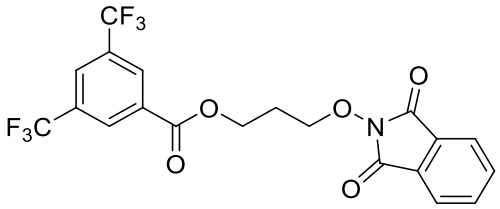


Figure S10. Synthesis of 3-(aminooxy)propyl 3,5-bis(trifluoromethyl) benzoate **18**.

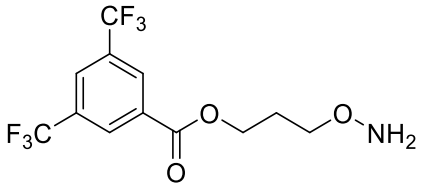
Synthesis of 3-((1,3-dioxoisoindolin-2-yl)oxy)propyl 3,5-bis(trifluoromethyl)benzoate S28



In a 25 ml round bottom flask, 3,5-bis(trifluoromethyl)benzoic acid **S27** (258 mg, 1 mmol), 2-(3-bromopropoxy)isoindoline-1,3-dione **S24** (312 mg, 1.1 mmol) and TEA (418 µl, 3 mmol) were dissolved in acetonitrile (5 ml). The reaction mixture was refluxed and the progress of the reaction was followed by TLC. After 8 h, the reaction mixture was concentrated and purification by silica gel flash column chromatography (ethyl acetate:n-hexane, 2:98) led to the isolation of

3-((1,3-dioxoisoindolin-2-yl)oxy)propyl 3,5-bis(trifluoromethyl)benzoate **S28** (335 mg, 73% yield; R_f 0.67, ethyl acetate:n-hexane 30:70; white solid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.51 (s, 2H), 8.05 (s, 1H), 7.91-7.81 (m, 2H), 7.80-7.72 (m, 2H), 4.72 (t, *J* = 6.3 Hz, 2H), 4.41 (t, *J* = 6.0 Hz, 2H), 2.39-2.21 (m, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.0, 163.7, 134.7, 132.5, 132.3 (q, *J* = 34.1 Hz, 2C), 130.1-129.8 (m, 2C), 129.0, 126.6-126.6 (m, 1C), 123.7, 123.0 (q, *J* = 272.8 Hz, 2C), 74.9, 62.7, 27.8 ppm. ^{19}F NMR (376 MHz, $CDCl_3$) δ -62.94 (TFA was used as an internal standard, -75.70 ppm). HRMS (ESI) $[M+H]^+$ calcd. $C_{20}H_{14}F_6NO_5$ 462.0776, found 462.0775.

Synthesis of 3-(aminooxy)propyl 3,5-bis(trifluoromethyl)benzoate 18



In a 5 ml round bottom flask, 3-((1,3-dioxoisoindolin-2-yl)oxy)propyl 3,5-bis(trifluoromethyl) benzoate **S28** (138 mg, 0.3 mmol) in DCM (3 ml) and hydrazine monohydrate (80%, 37 µl, 0.75 mmol) were stirred at room temperature. The progress of the reaction was followed by TLC. After 3 h, the reaction mixture was filtered and concentration of the filtrate in vacuo led to the isolation of 3-(aminooxy)propyl 3,5-bis(trifluoromethyl)benzoate **18** (80 mg, 81% yield; R_f 0.33, ethyl acetate:n-hexane 30:70; white solid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.48 (s, 2H), 8.07 (s, 1H), 5.43 (bs, 2H), 4.50 (t, *J* = 6.5 Hz, 2H), 3.83 (t, *J* = 6.1 Hz, 2H), 2.19-2.04 (m, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.1, 132.6, 132.4 (q, *J* = 33.9 Hz, 2C), 130.0-129.7 (m, 2C), 126.6-126.3 (m, 1C), 123.02 (q, *J* = 273.0 Hz, 2C), 72.2, 63.6, 27.9 ppm. ^{19}F NMR (376 MHz, $CDCl_3$) δ -62.54 (TFA was used as an internal standard, -75.70 ppm). HRMS (ESI) $[M+H]^+$ calcd. $C_{12}H_{12}F_6NO_3$ 332.0721, found 332.0699.

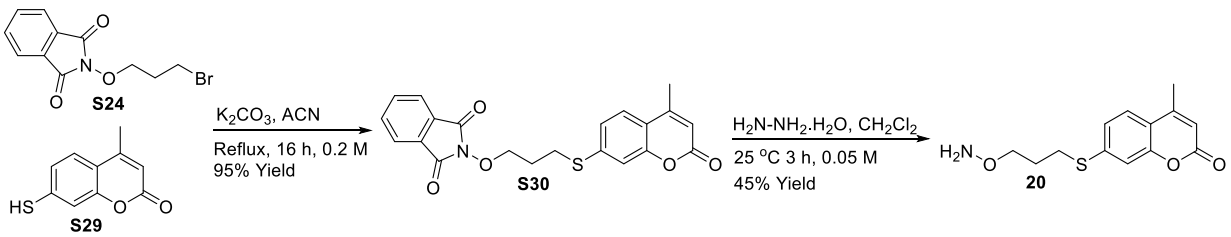
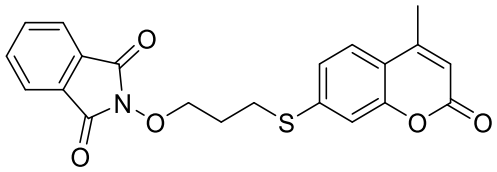


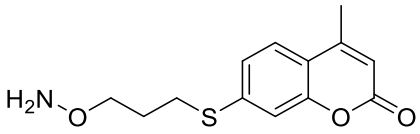
Figure S11. Synthesis of 7-((3-(aminooxy)propyl)thio)-4-methyl-2H-chromen-2-one **20**.

Synthesis of 2-(3-((4-methyl-2-oxo-2H-chromen-7-yl)thio)propoxy)isoindoline-1,3-dione S30



In a 25 ml round bottom flask, 7-mercapto-4-methylcoumarin **S29** (192 mg, 1 mmol), K_2CO_3 (276 mg, 2 mmol), and 2-(3-bromopropoxy)isoindoline-1,3-dione **S24** (568 mg, 2 mmol) were dissolved in degassed acetonitrile (5 ml) and refluxed for 16 h. The reaction mixture was concentrated in vacuo and purified by silica gel flash column chromatography using ethyl acetate:hexane (7:3) to give **S30** (375 mg, 95% yield; R_f 0.37, ethyl acetate:n-hexane 50:50; white solid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90-7.81 (m, 2H), 7.80-7.73 (m, 2H), 7.48 (d, *J* = 8.2 Hz, 1H), 7.26-7.20 (m, 2H), 6.22 (d, *J* = 0.8 Hz, 1H), 4.36 (t, *J* = 5.8 Hz, 2H), 3.35 (t, *J* = 7.1 Hz, 2H), 2.41 (d, *J* = 0.9 Hz, 3H), 2.23-2.08 (m, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.8, 160.7, 154.0, 152.3, 142.6, 134.7, 129.0, 124.9, 123.8, 123.4, 117.5, 114.8, 114.1, 76.6, 28.7, 27.8, 18.7 ppm. HRMS (ESI) $[M+H]^+$ calcd. $C_{21}H_{18}NO_5S$ 396.0906, found 396.0925.

Synthesis of 7-((3-(aminooxy)propyl)thio)-4-methyl-2H-chromen-2-one 20



2-(3-((4-methyl-2-oxo-2H-chromen-7-yl)thio)propoxy)isoindoline-1,3-dione **S30** (237 mg, 0.6 mmol) was dissolved in CH_2Cl_2 (12 ml) in a 50 ml round bottom flask. To this solution, hydrazine monohydrate (80%, 29 µl, 0.6 mmol) was added and stirred at 25 ^{o}C for 3 h. The reaction mixture was filtered and the filtrate was concentrated. The purification of crude reaction mixture was performed by reverse phase preparative HPLC to isolate **20** (76 mg, 45% yield; R_f 0.6, ethyl acetate:n-hexane 50:50; pale green viscous liquid). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.46 (d, *J* = 8.3 Hz, 1H), 7.23-7.13 (m, 2H), 6.18 (d, *J* = 0.9 Hz, 1H), 3.79 (t, *J* = 5.9 Hz, 2H), 3.07 (t, *J* = 7.3 Hz, 2H), 2.40 (d, *J* = 0.8 Hz, 3H), 2.02-1.90 (m, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 160.7, 154.0, 152.3, 143.3, 124.7, 123.1, 117.2, 114.1, 113.9, 73.9, 29.0, 27.8, 18.6 ppm. HRMS (ESI) $[M+H]^+$ calcd. $C_{13}H_{16}NO_3S$ 266.0851, found 266.0841.

4. Experimental data for protein labeling

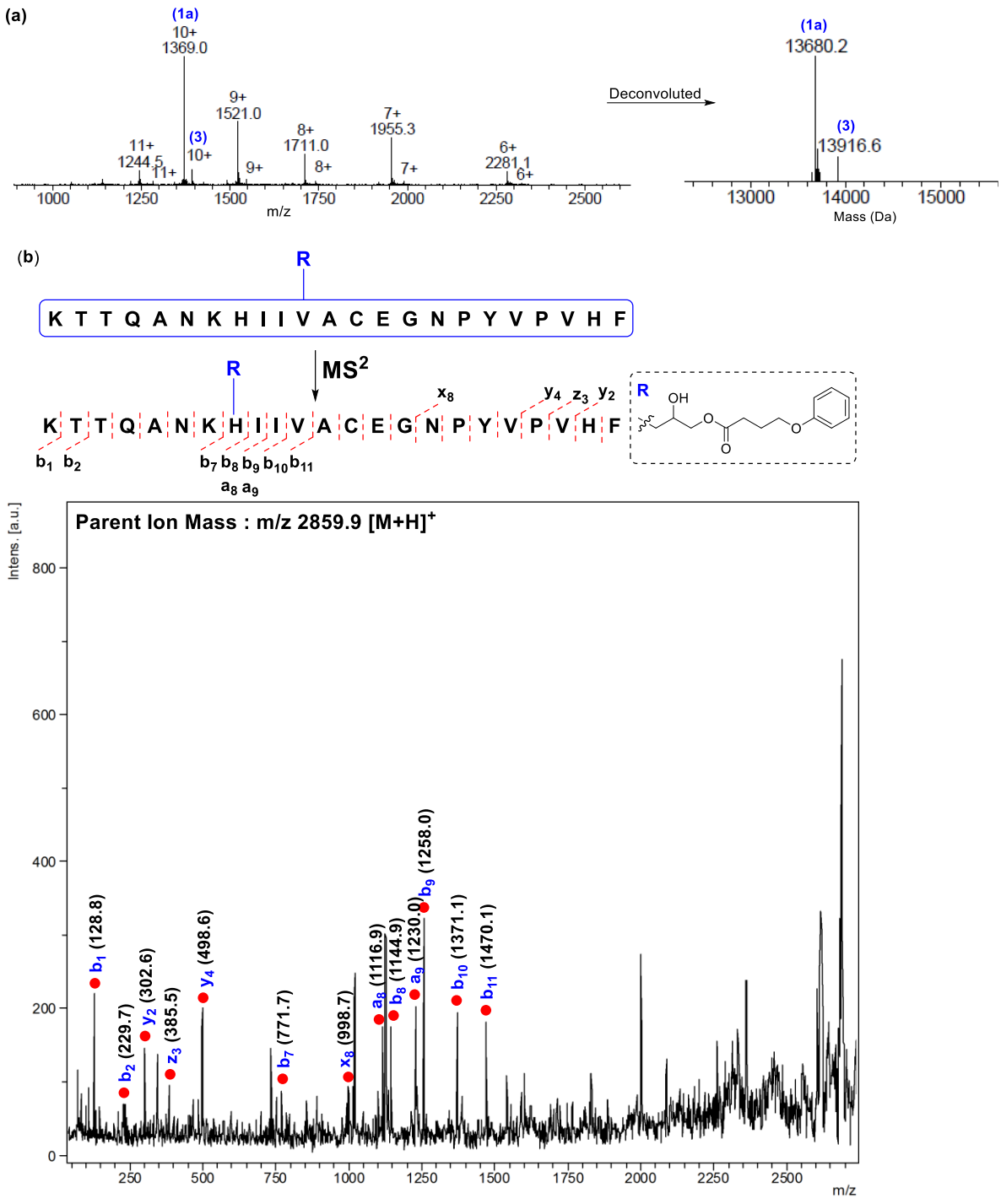


Figure S12. (a) ESI-MS spectrum of labeled RNase A **3**. (b) MS-MS spectrum of labeled KTTQANKHIIVACEGNPYVPVHF (K98-F120, *m/z* 2859.9 $[M+H]^+$) after the digestion of **3** with α-Chymotrypsin. The kinetically labeled site in **3** is H105.

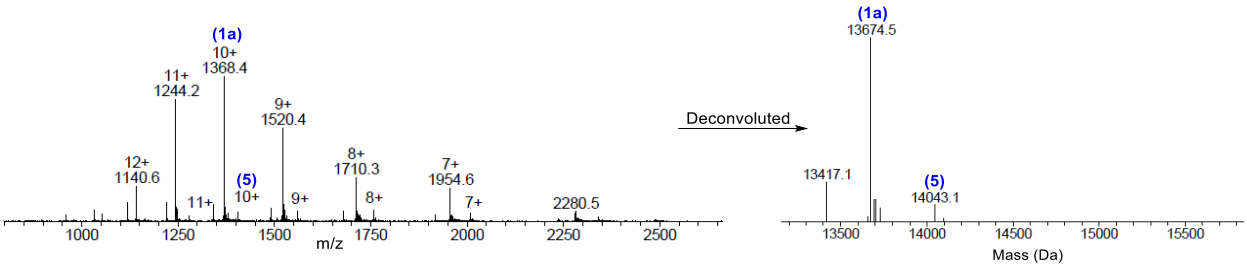


Figure S13. ESI-MS spectrum of the labeled RNase A **5**.

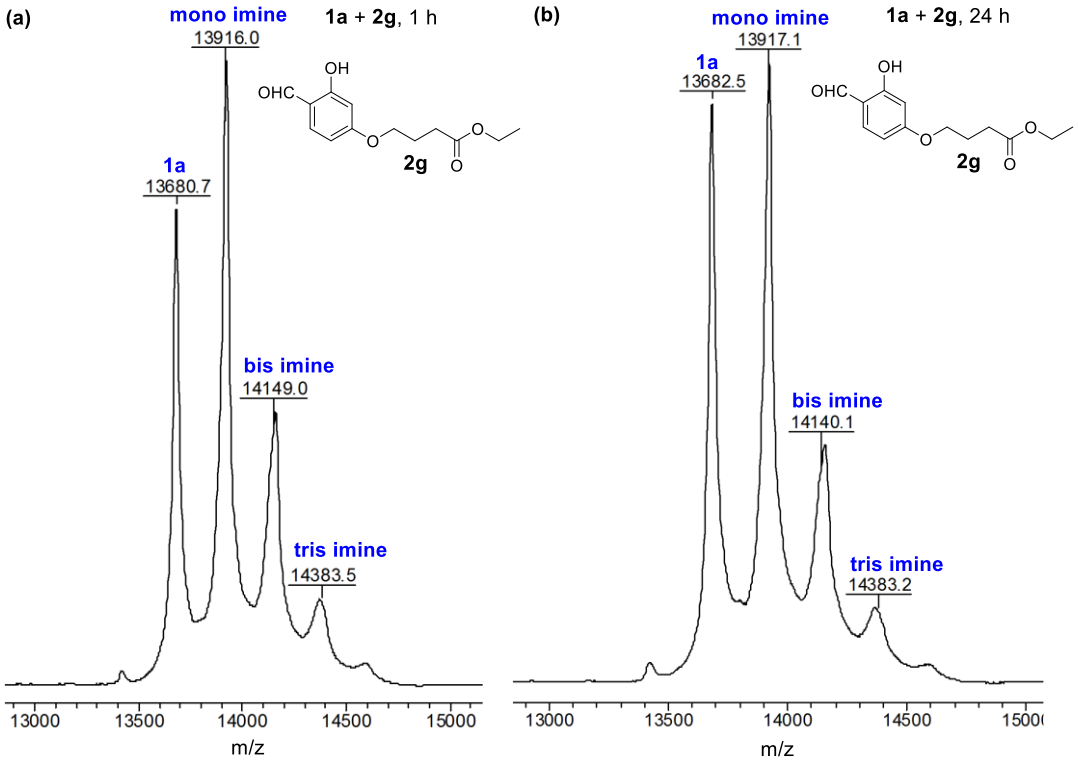


Figure S14. Imine formation with RNase A **1a** (1 equiv.) and reagent **2g** (25 equiv.) MALDI-ToF-MS spectrum (a) after 1 h, (b) after 24 h.

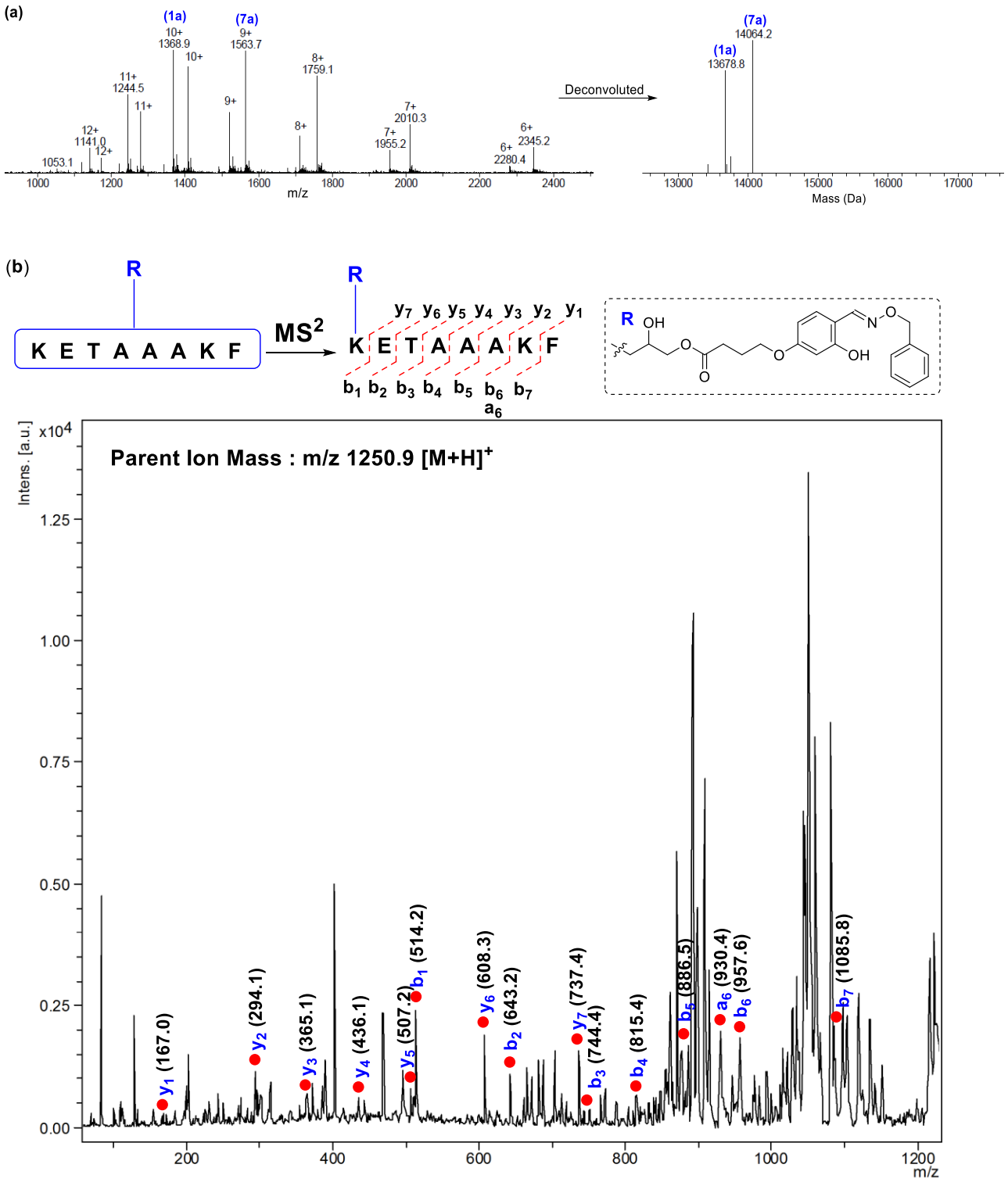


Figure S15. (a) ESI-MS spectrum of the mono-labeled RNase A **7a**. (b) MS-MS spectrum of labeled KETAAAKF (K1-F8, *m/z* 1250.9 $[M+H]^+$) after the digestion of **7a** with α-Chymotrypsin. The site of modification in mono-labeled RNase A **7a** is K1.

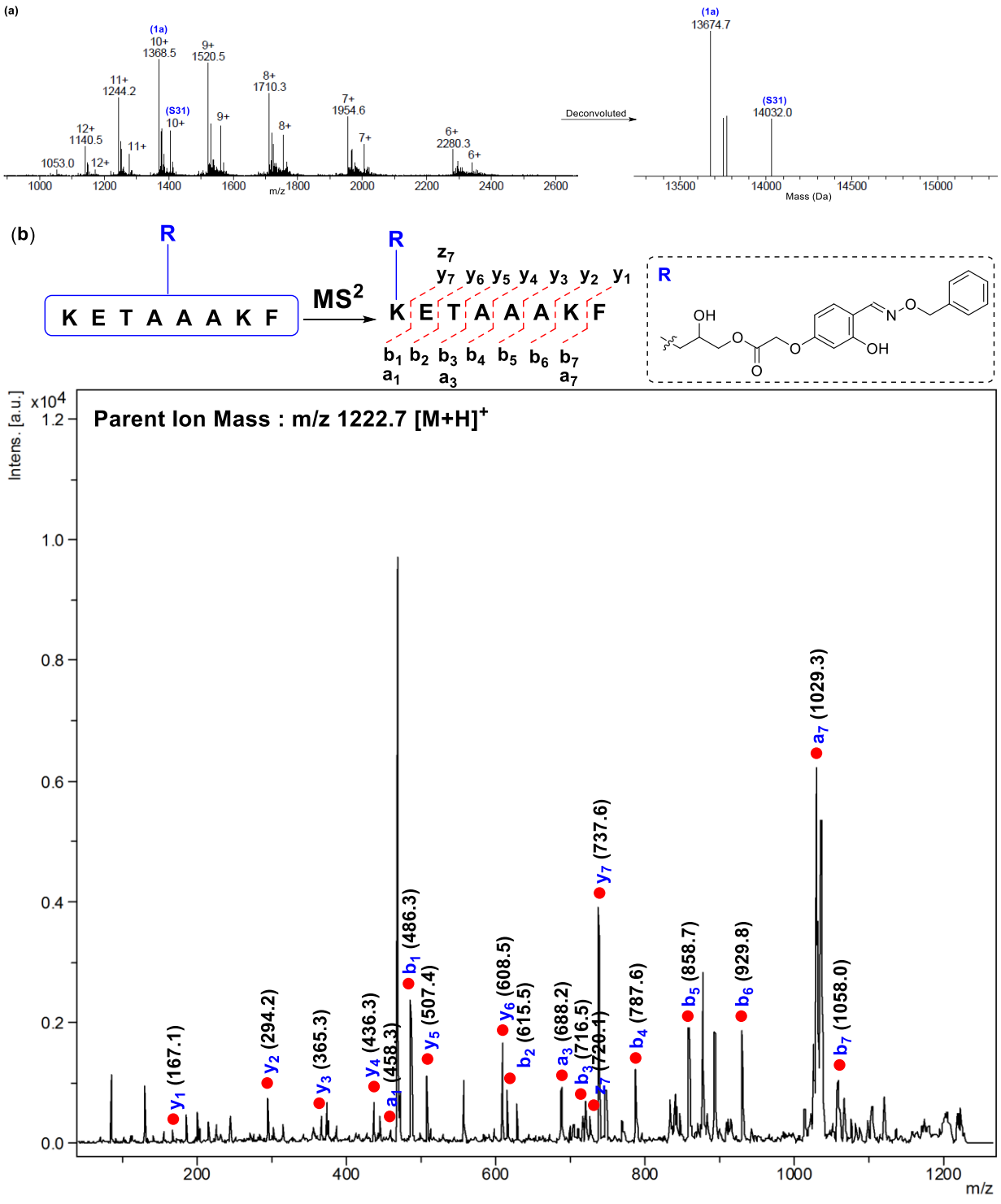


Figure S16. (a) ESI-MS spectrum of the mono-labeled RNase A **S31**. (b) MS-MS spectrum of labeled KETAAAKF (K1-F8, *m/z* 1222.7 $[M+H]^+$) after the digestion of **S31** with α-Chymotrypsin. The site of modification in mono-labeled RNase A **S31** is K1.

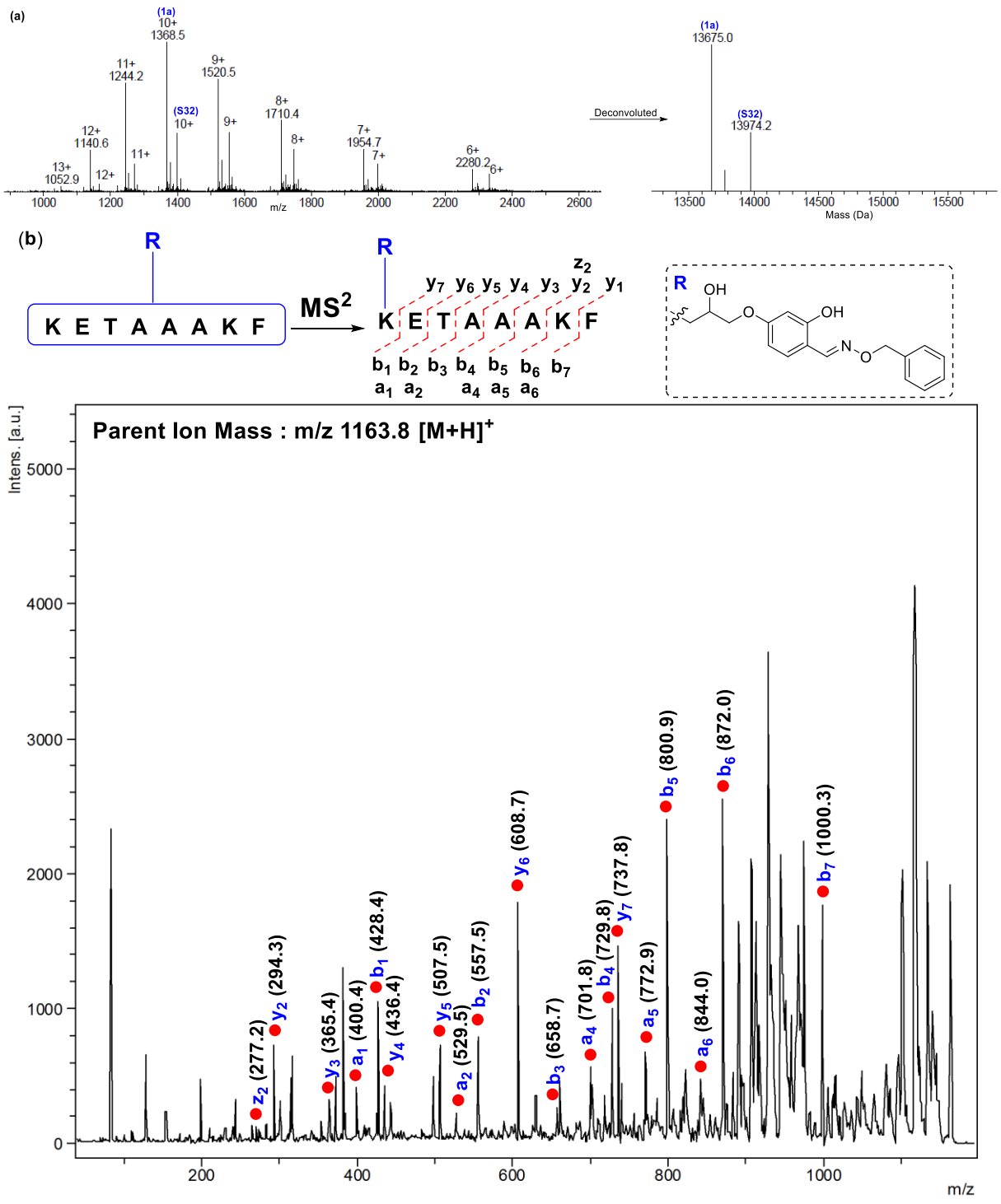


Figure S17. (a) ESI-MS spectrum of the mono-labeled RNase A **S32**. (b) MS-MS spectrum of labeled KETAAAKF (K1-F8, *m/z* 1163.8 $[M+H]^+$) after the digestion of **S32** with α-Chymotrypsin. The site of modification in mono-labeled RNase A **S32** is K1.

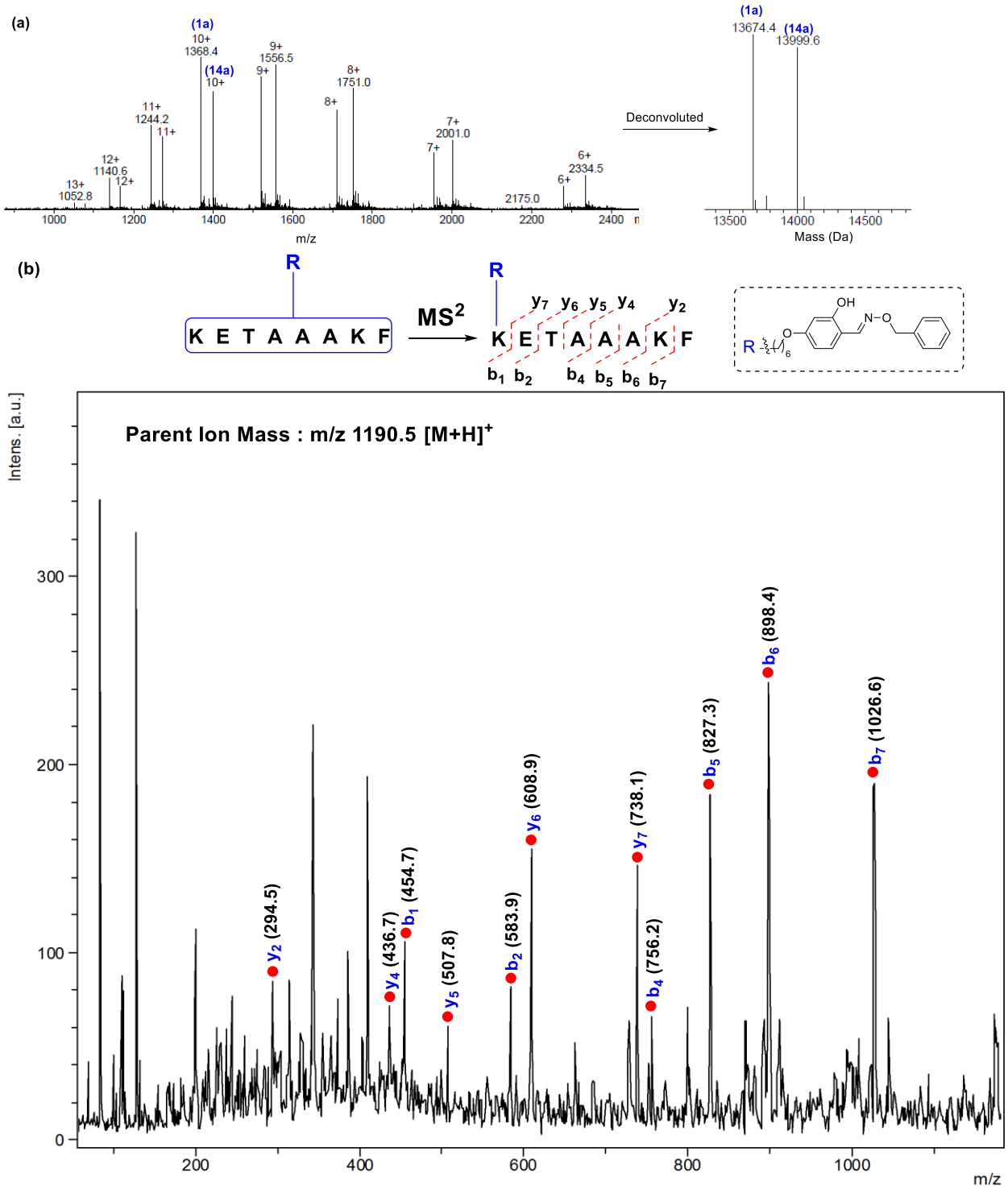


Figure S18. (a) ESI-MS spectrum of the mono-labeled RNase A **14a**. (b) MS-MS spectrum of labeled KETAAAKF (K1-F8, *m/z* 1190.5 $[M+H]^+$) after the digestion of **14a** with α-Chymotrypsin. The site of modification in mono-labeled RNase A **14a** is K1.

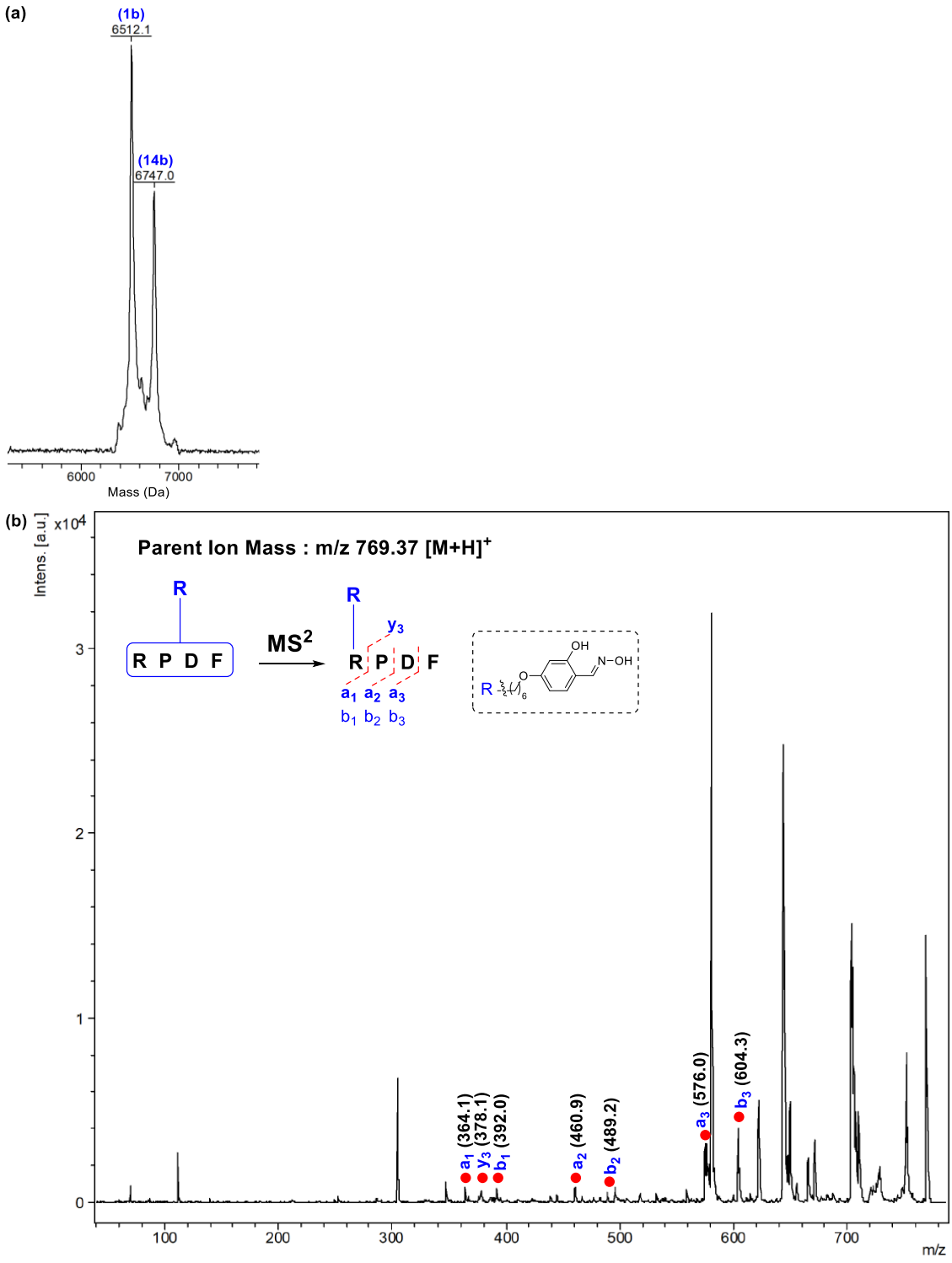


Figure S19. (a) MALDI-ToF-MS spectrum of the labeled aprotinin **14b**. (b) MS-MS spectrum of labeled RPDF (R1-F4, *m/z* 769.37 $[M+H]^+$) after the digestion of **14b** with α-Chymotrypsin. The site of modification in mono-labeled aprotinin **14b** is R1.

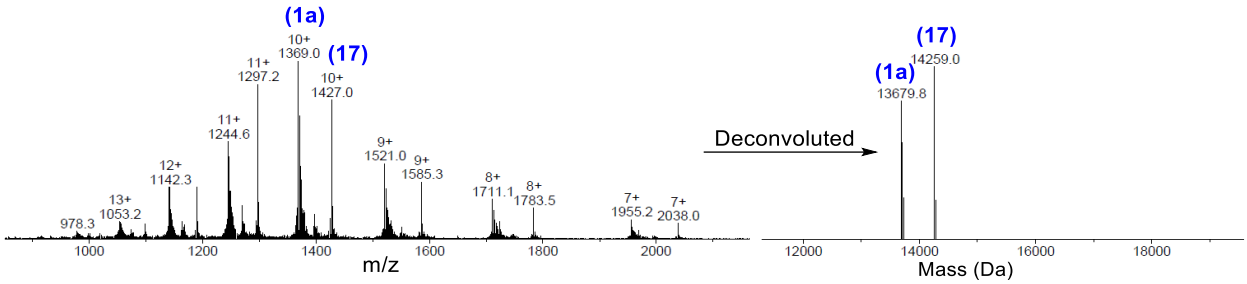


Figure S20. ESI-MS spectrum of biotin tagged RNase A **17**.

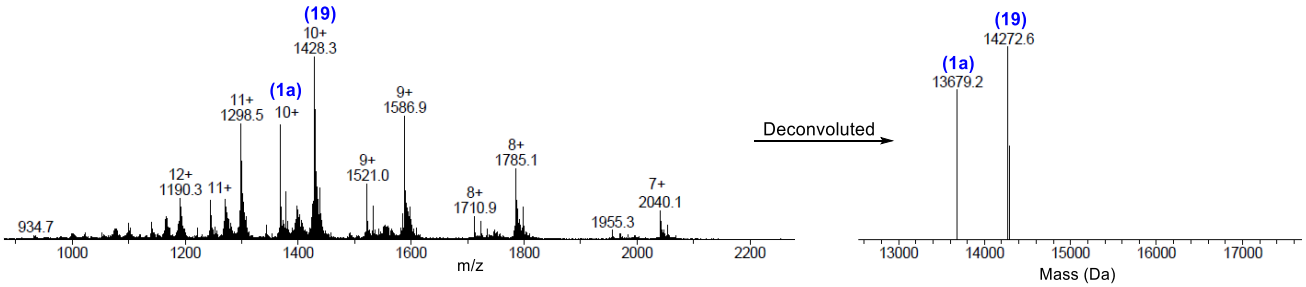


Figure S21. ESI-MS spectrum of ^{19}F-NMR probe tagged RNase A **19**.

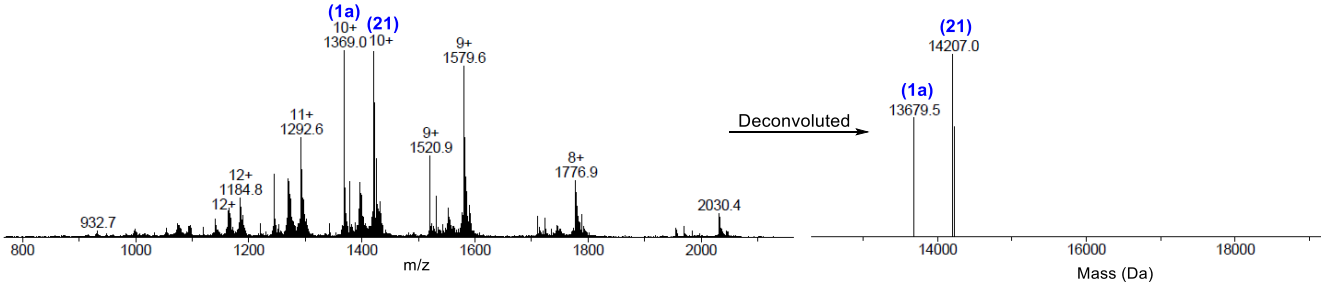


Figure S22. ESI-MS spectrum of coumarin tagged RNase A **21**.

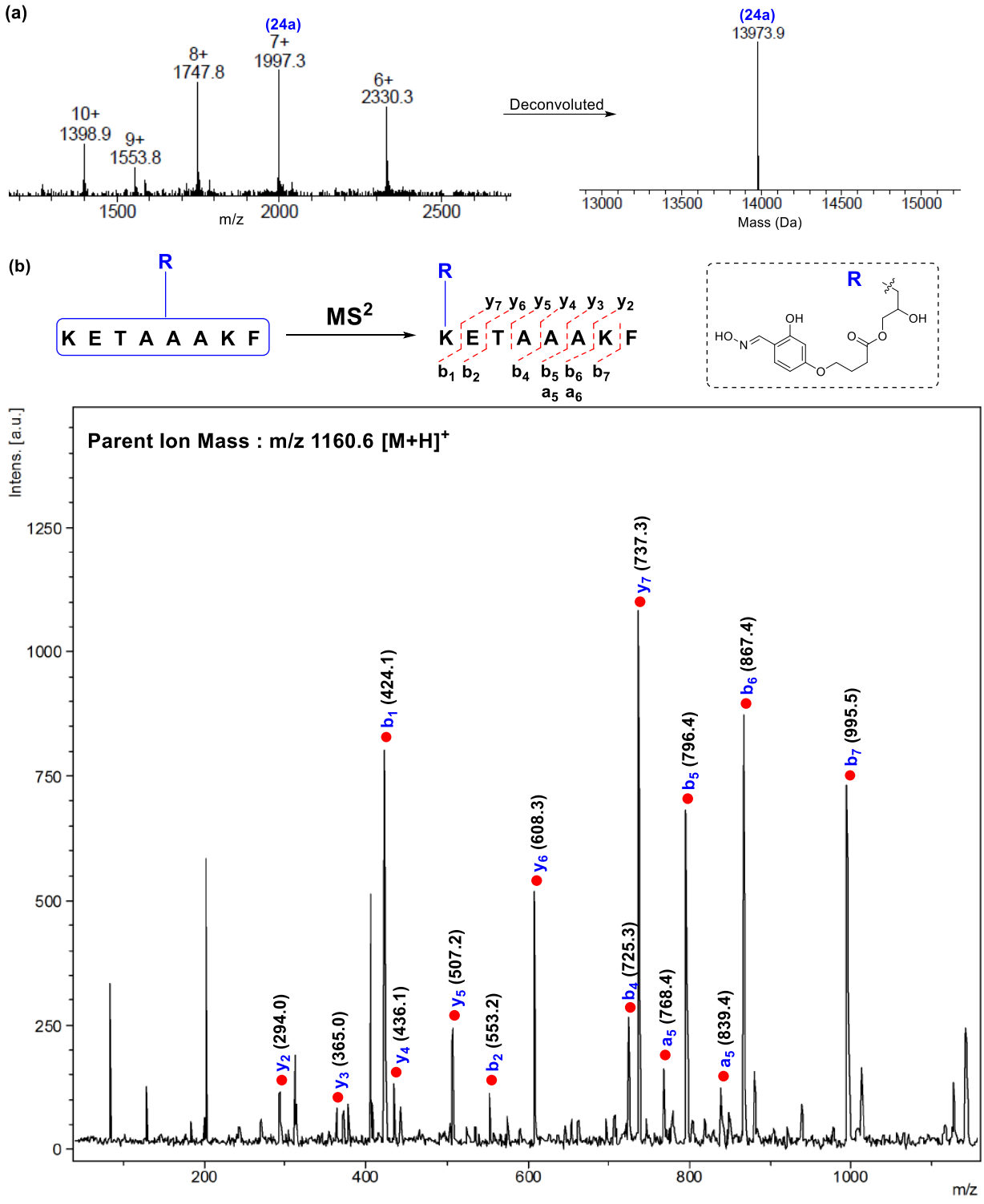


Figure S23. (a) ESI-MS spectrum of the purified mono-labeled RNase A **24a**. (b) MS-MS spectrum of labeled KETAAAKF (K1-F8, *m/z* 1160.6 $[M+H]^+$) after the digestion of **24a** with α-Chymotrypsin.. The site of modification in the mono-labeled RNase A **24a** is K1.

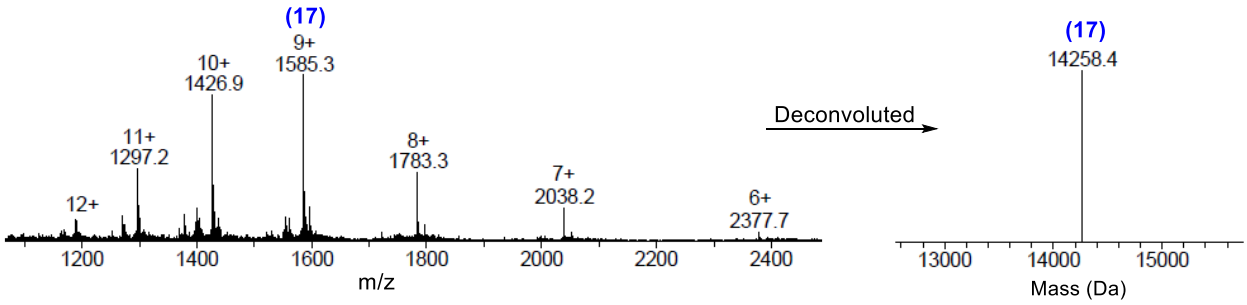


Figure S24. ESI-MS spectrum of the purified biotin tagged RNase A **17**.

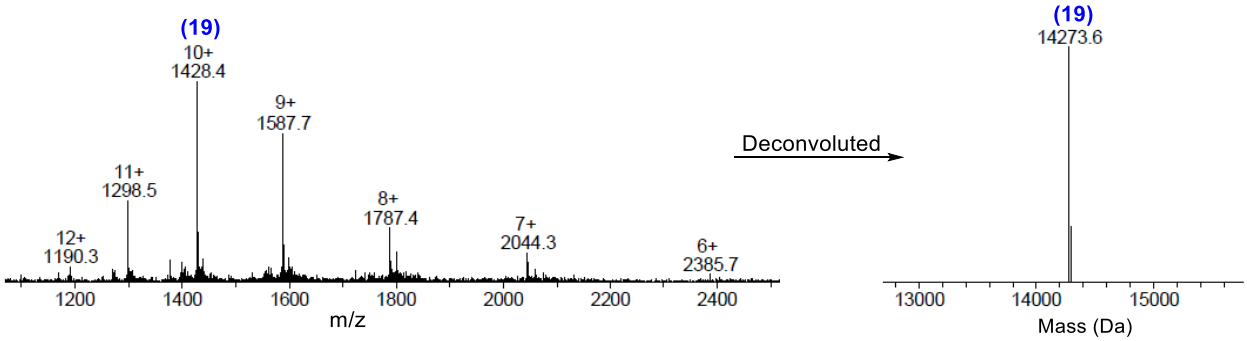


Figure S25. ESI-MS spectrum of the purified ^{19}F NMR probe tagged RNase A **19**.

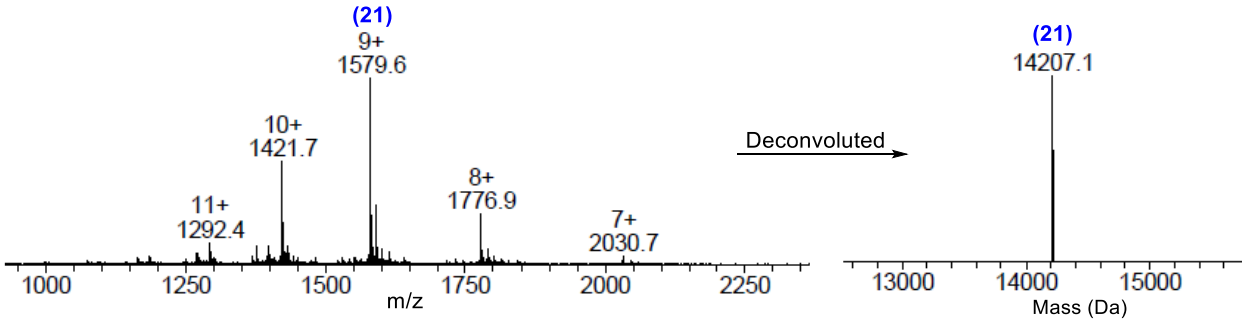


Figure S26. ESI-MS spectrum of the purified coumarin tagged RNase A **21**.

5. Additional data

5a. Conversion with time

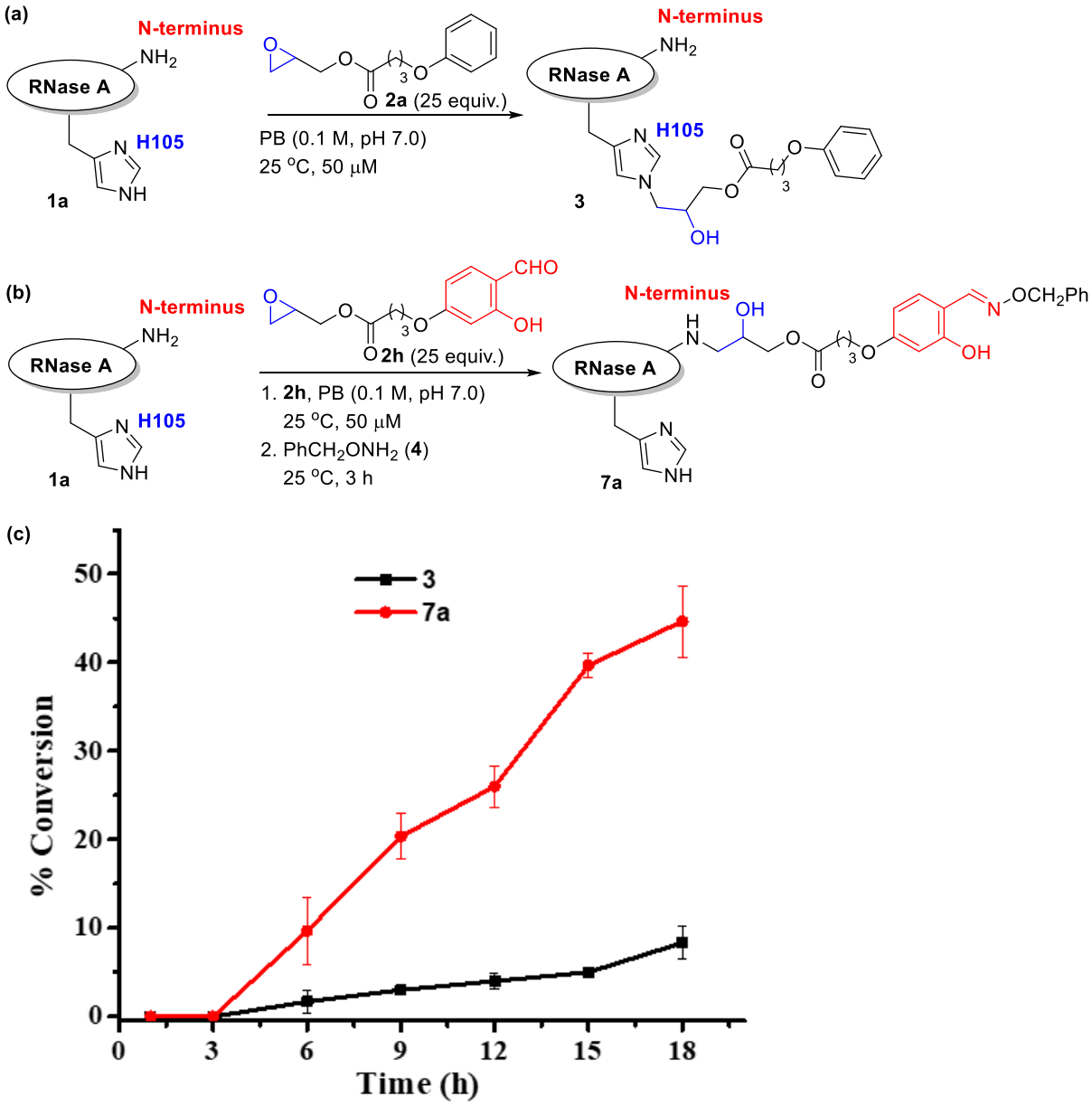


Figure S27. (a) Kinetic labeling of RNase A with reagent **2a**. (b) Chemoselectivity switching in RNase A labeling. (c) Comparison of rate of reaction in product formation **3** versus **7a**.

5b. CD spectra of labeled RNase A

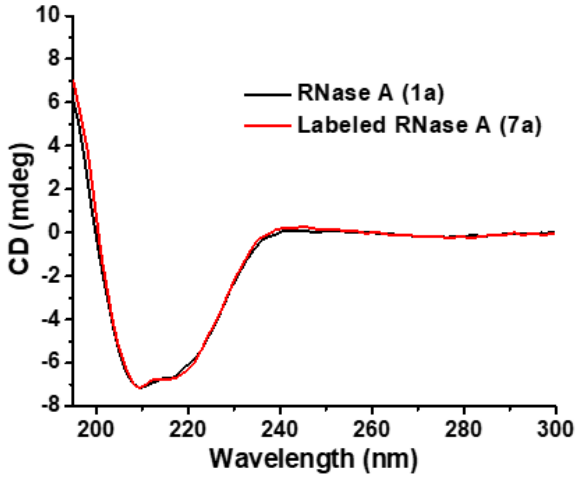


Figure S28. Effect of the N-terminus labeling on structure of RNase A. Circular Dichroism (CD) spectra of RNase A (**1a**) and N-terminus labeled RNase A (**7a**) in phosphate buffer (0.1 M, pH 7.0) at concentration 0.1 mg/ml.

5c. Enzymatic assay of labeled RNase A[13]

Enzymatic activity of RNase A before and after the modification was checked by hydrolysis of Ribonucleic acid (RNA) at 300 nm (A300) using quartz cuvette (path length, 1 cm at 25 ^{o}C). Sodium acetate buffer (pH 5.0, 0.1 M) was prepared with Millipore Grade I water. Freshly prepared solutions were used for the assay. (a) Ribonucleic acid [RNA, 0.1% (w/v), 1 mg/ml] in buffer, (b) RNase A **1a**, (c) labeled RNase A **7a** (10 µg/1 ml, in Millipore Grade I water). Initially, change in the absorption of RNA was monitored at 300 nm using the RNA solution and blank. The absorbance was immediately recorded after a gentle mixing of RNA (500 µl) and RNase A (500 µl). Next, the absorbance was recorded for the mixture of RNA and labeled RNase A **7a**. The activity of the labeled RNase A remains unperturbed.

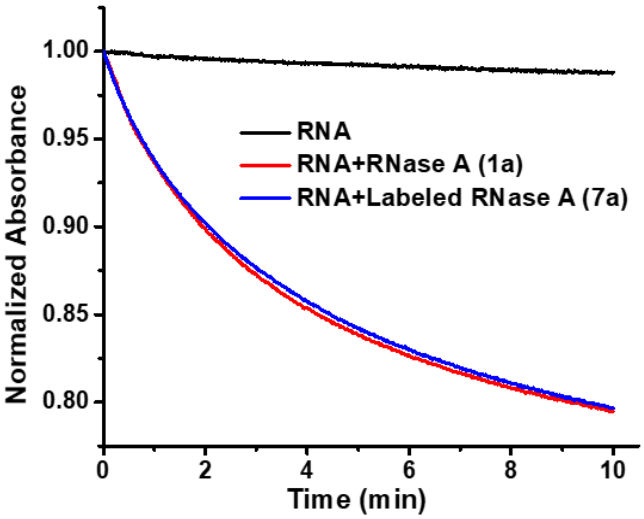


Figure S29. Normalized absorbance spectra of RNA, (RNA+RNase A **1a**), and (RNA+labeled RNase A **7a**) at 300 nm.

5d. Role of the aldehyde

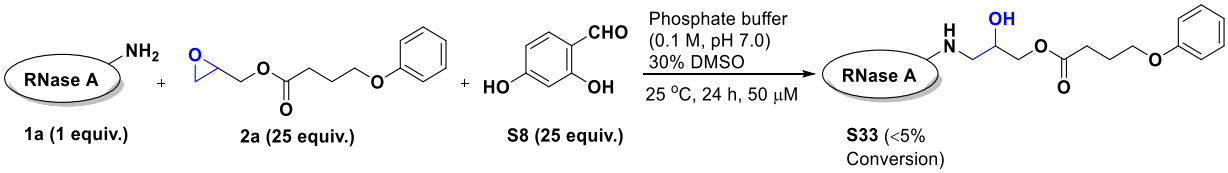


Figure S30. Role of aldehyde in the chemoselectivity switching.

5e. Effect of pH

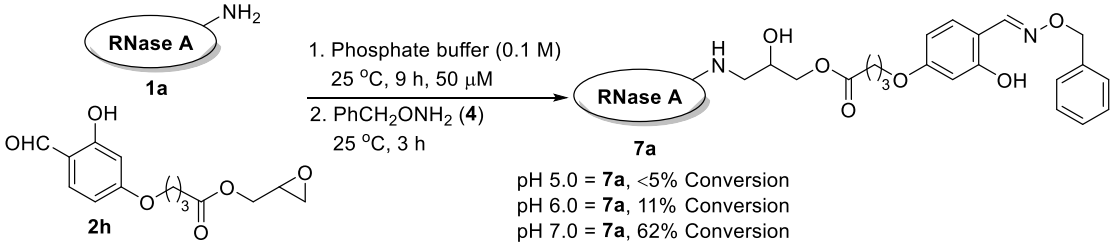


Figure S31. Effect of the pH in the chemoselectivity switching.

5f. Role of the spacer

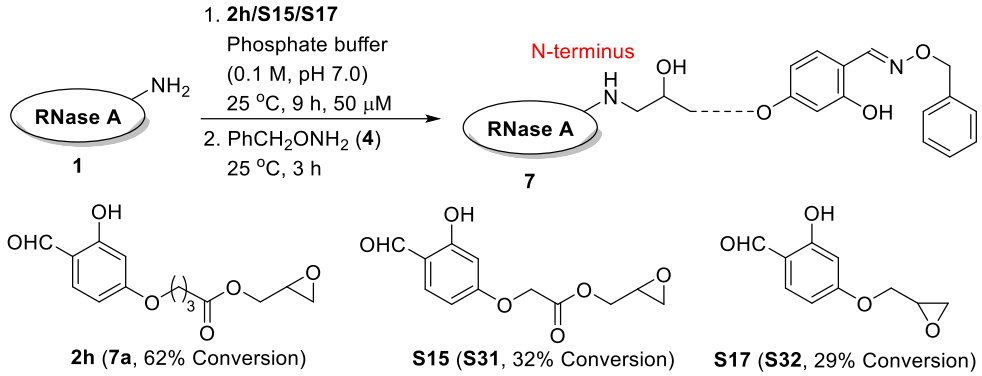


Figure S32. Role of spacer in the chemoselectivity switching. For MS and MS-MS data, see Figures S15-S17.

5g. ^{19}F NMR spectrum

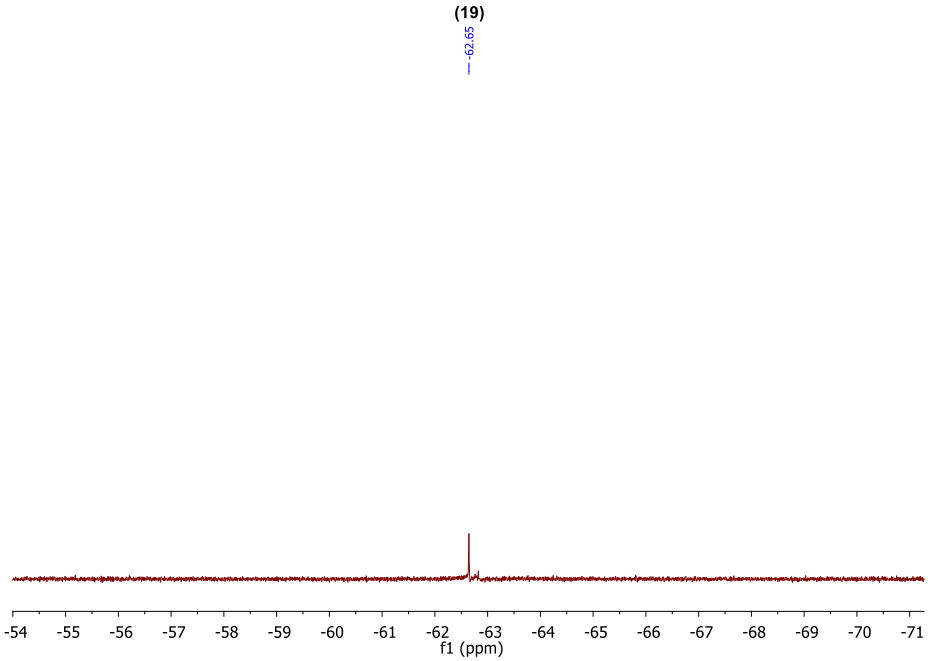


Figure S33. ^{19}F-NMR spectrum of labeled RNase A **19**. (For MS data, see Figure S25)

^{19}F-NMR probe attached RNase A **19** shows a sharp signal at -62.65 ppm by ^{19}F-NMR spectroscopy. Trifluoroacetic acid (0.2 mM) was used as internal standard, -75.45 ppm. The NMR experiment was performed in phosphate buffer (0.1 M, pH 7.0):D_2O (9:1).

5h. Fluorescence spectra

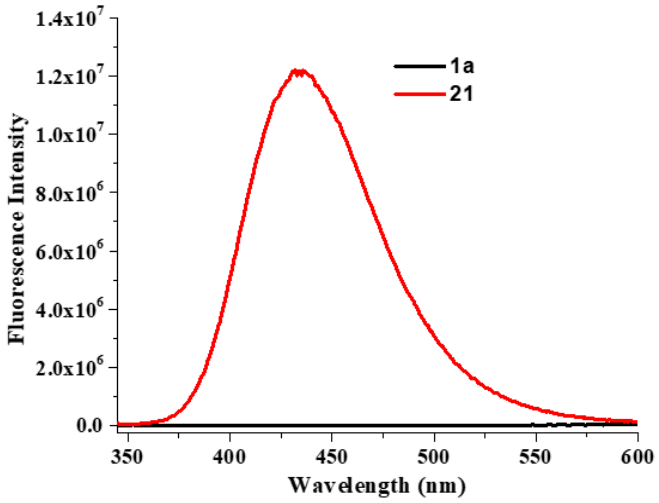


Figure S34. Steady-state fluorescence spectra of RNase A (**1a**) and coumarin tagged RNase A (**21**). In phosphate buffer (0.1 M, pH 7.0), **21** exhibits emission band peaked at 430 nm (excitation at 333 nm). For MS data, see Figure S26.

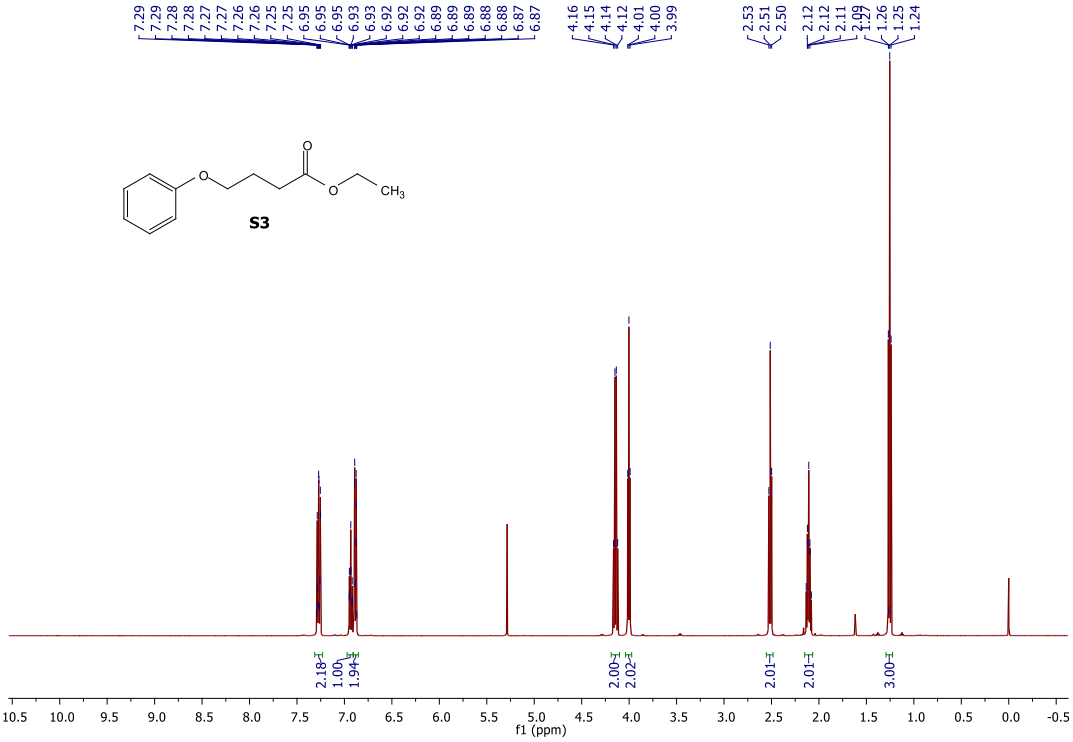
6. Protein sequence

Ribonuclease A from bovine pancreas
PDB ID: 2AAS
amino acid sequence:
KETAAAKFERQHMDSSTSAASSSNYCNQMMKSRNLTKDRCKPVNTFVHESLADVQAV
CSQKNVACKNGQTNCYQSYSTMSITDCRETGSSKYPNCAYKTTQANKHIIVACEGNPY
VPVHFDASV

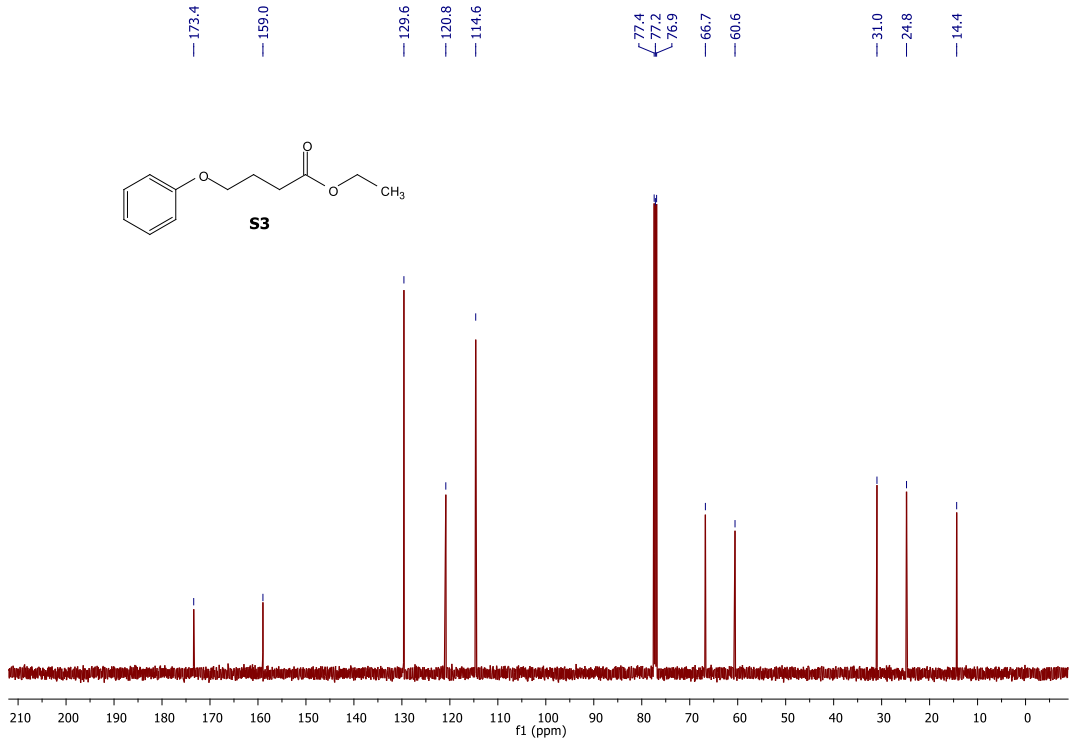
Aprotinin from bovine lung
PDB ID: 3LDI
amino acid sequence:
RPDFCLEPPYTGPCKARIIRYFYNAKAGLCQTFVYGGCRAKRNNFKSAEDCMRTCGGA

7. Spectral data

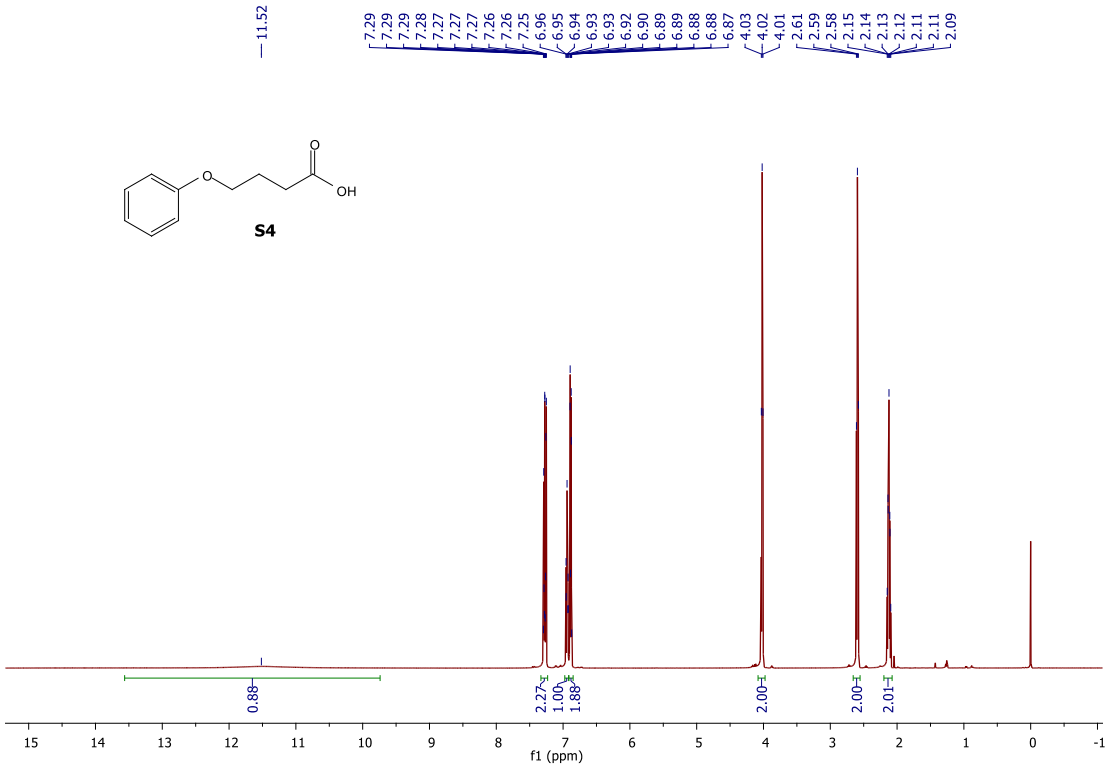
^{1}H NMR



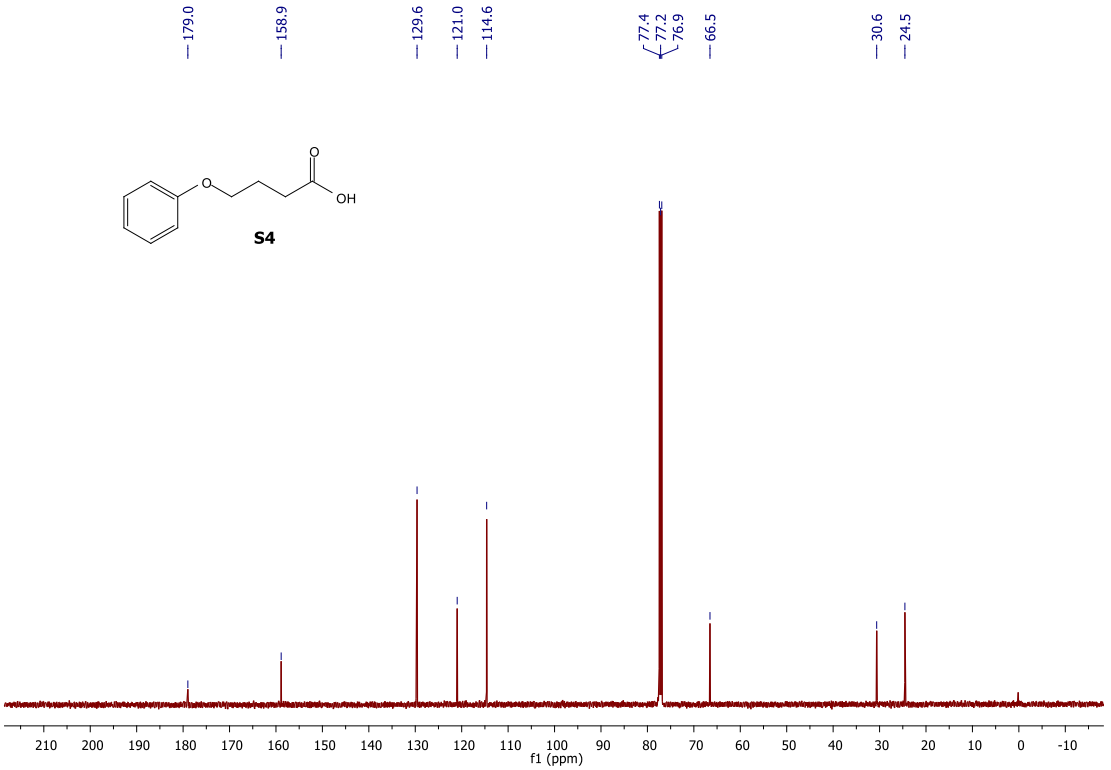
^{13}C NMR



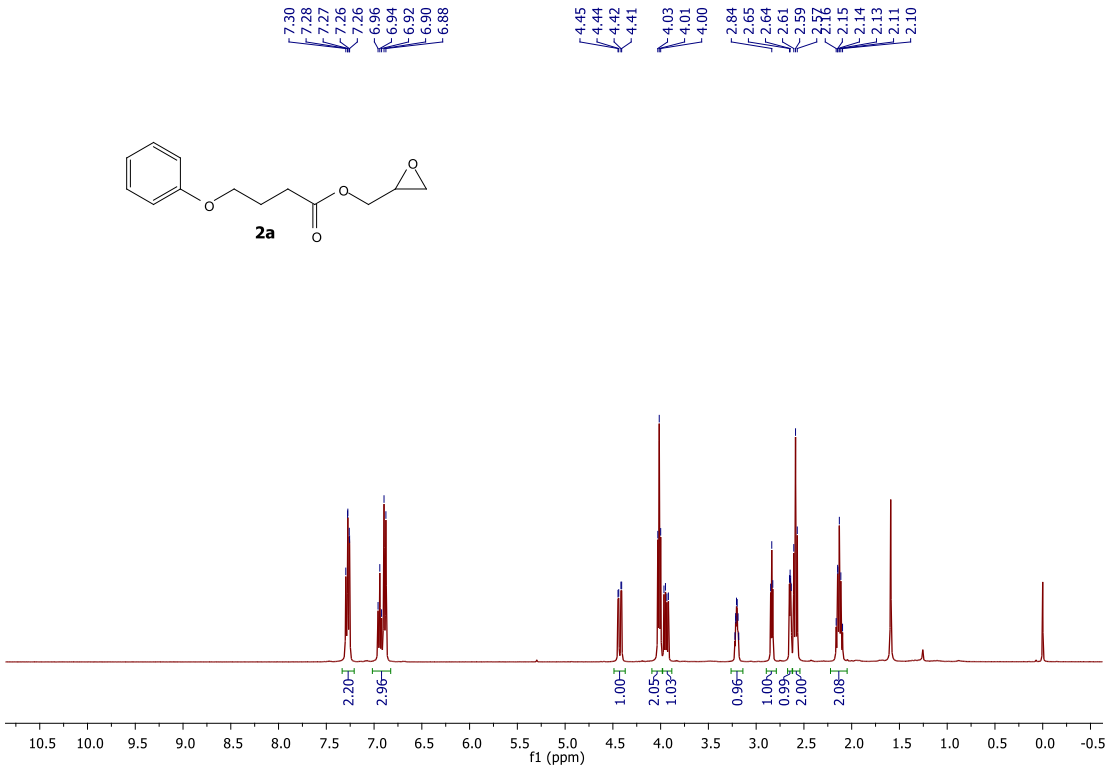
^{1}H NMR



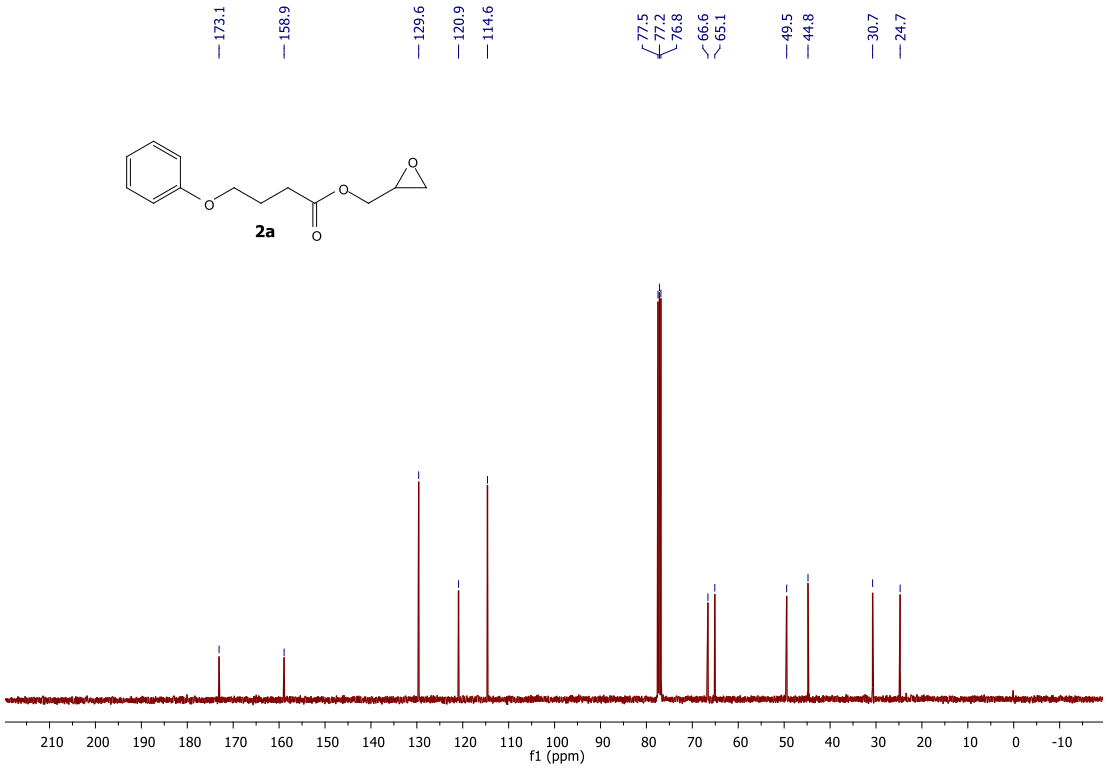
^{13}C NMR



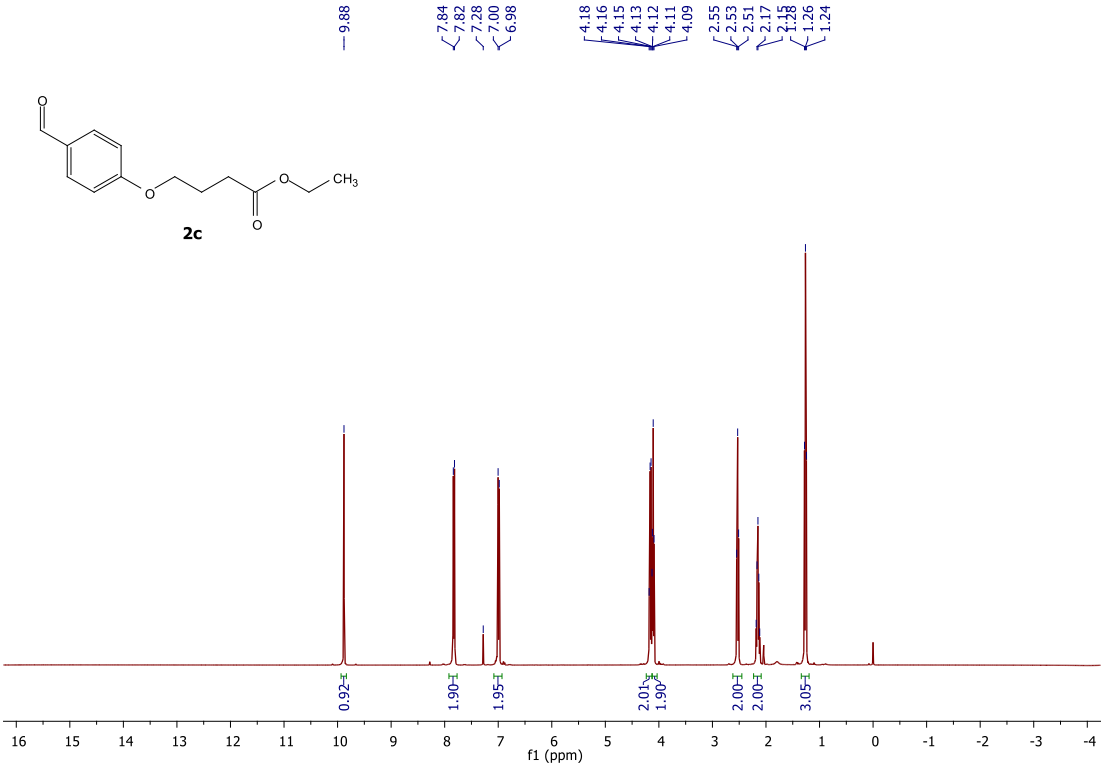
^{1}H NMR

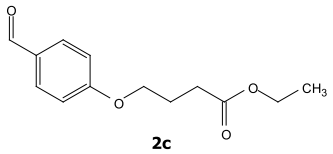


^{13}C NMR



^{1}H NMR



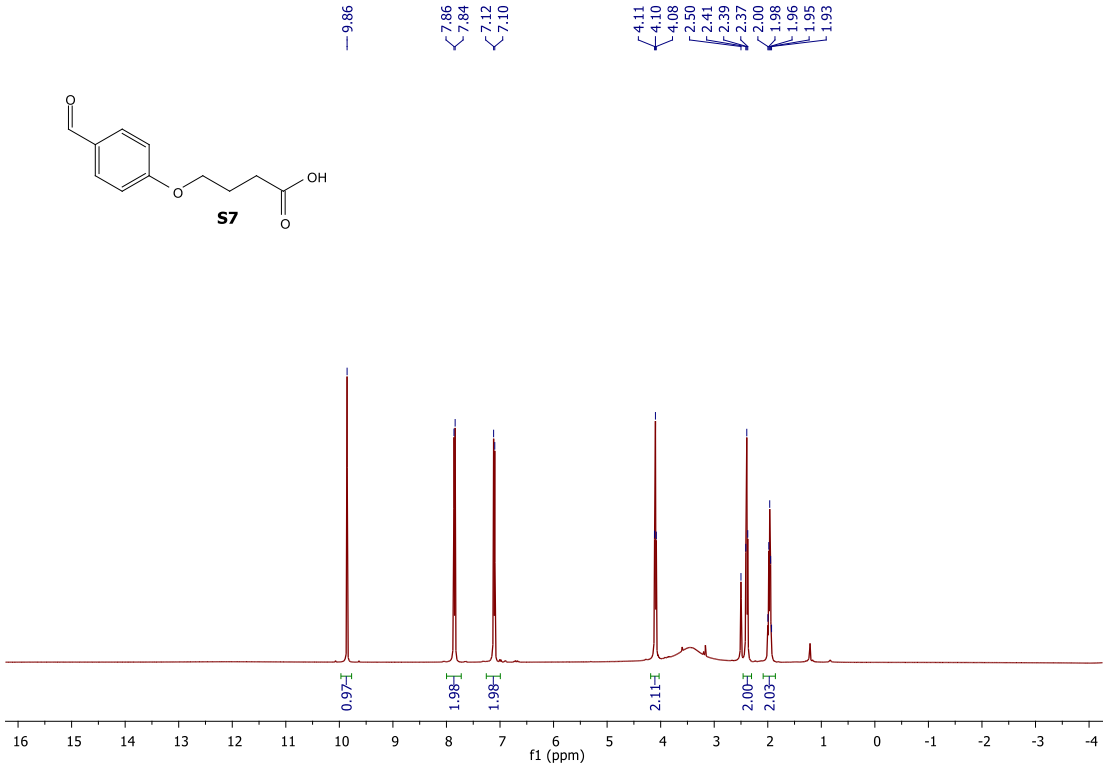


2c

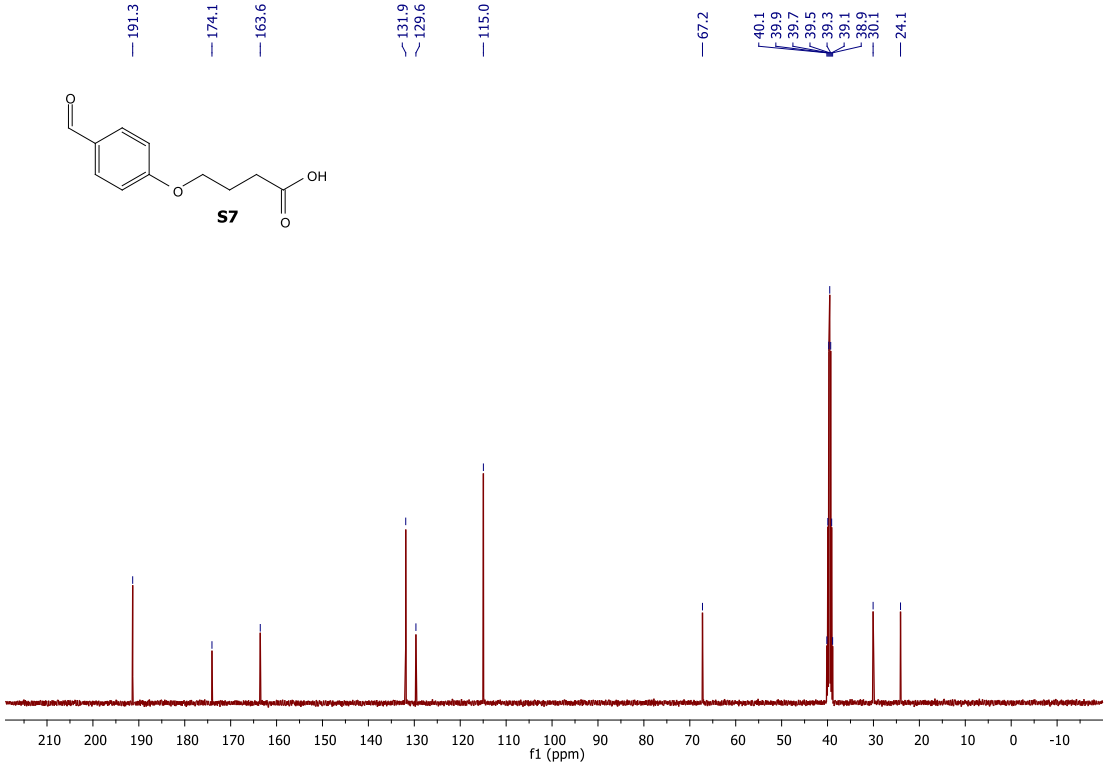
^{13}C NMR

2c

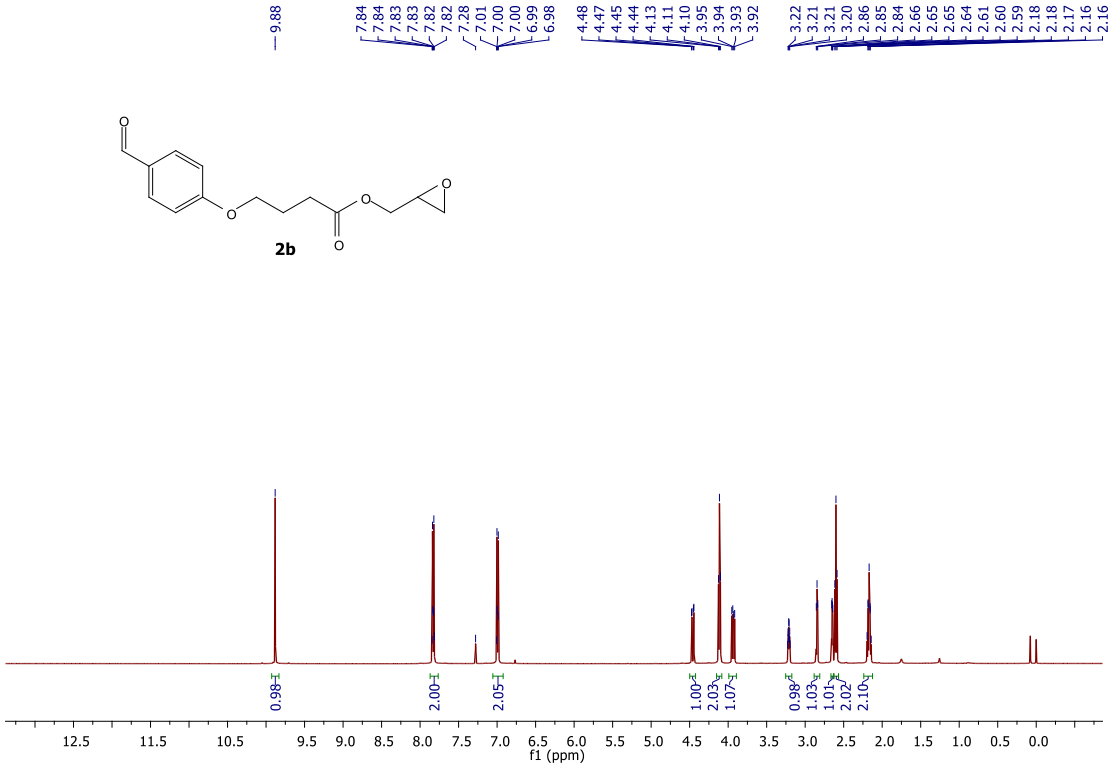
^{13}C NMR



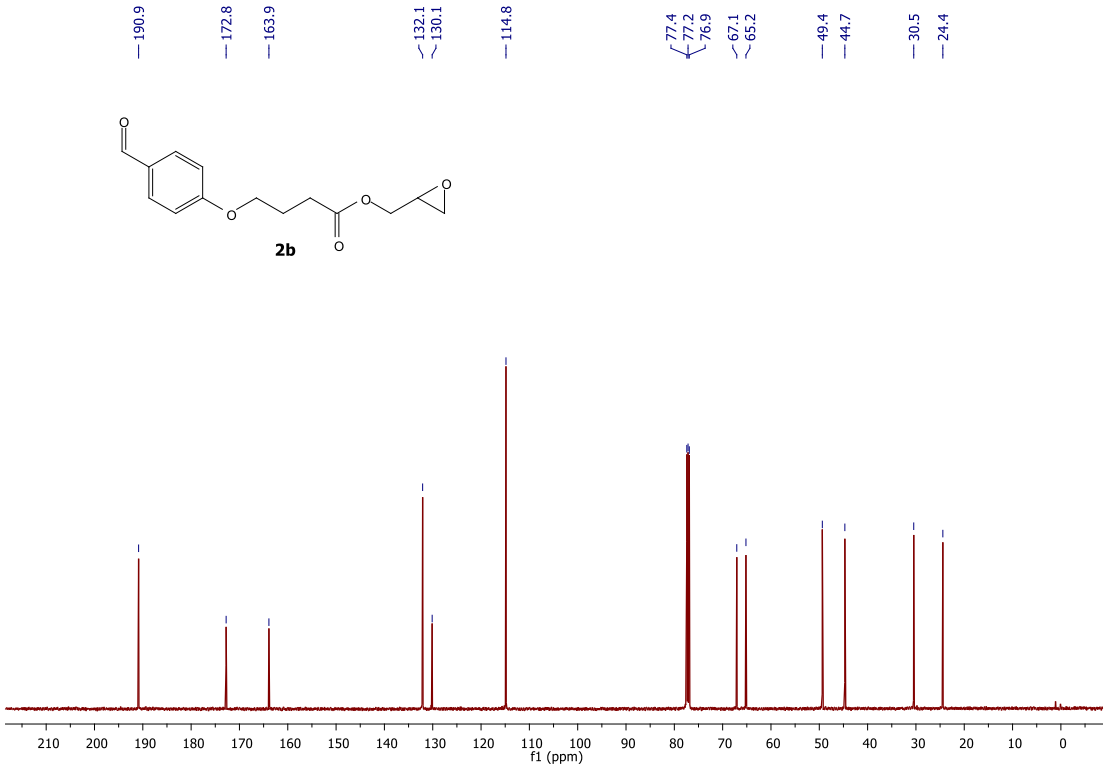
^{13}C NMR



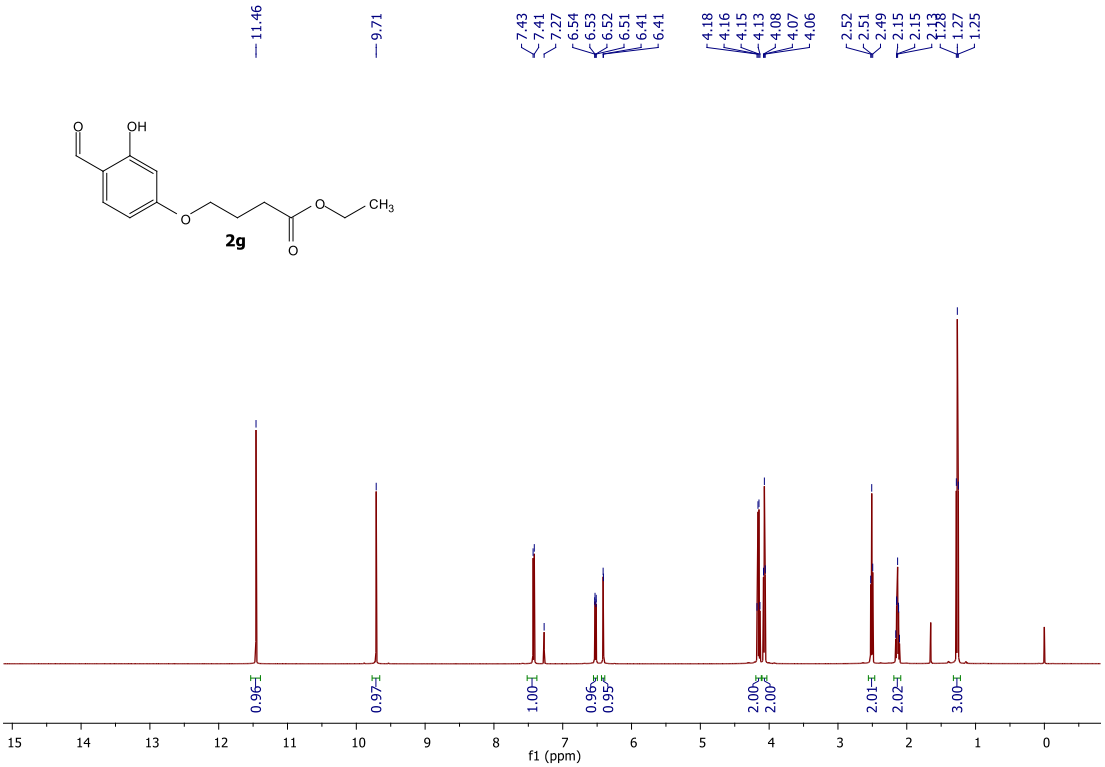
^{1}H NMR



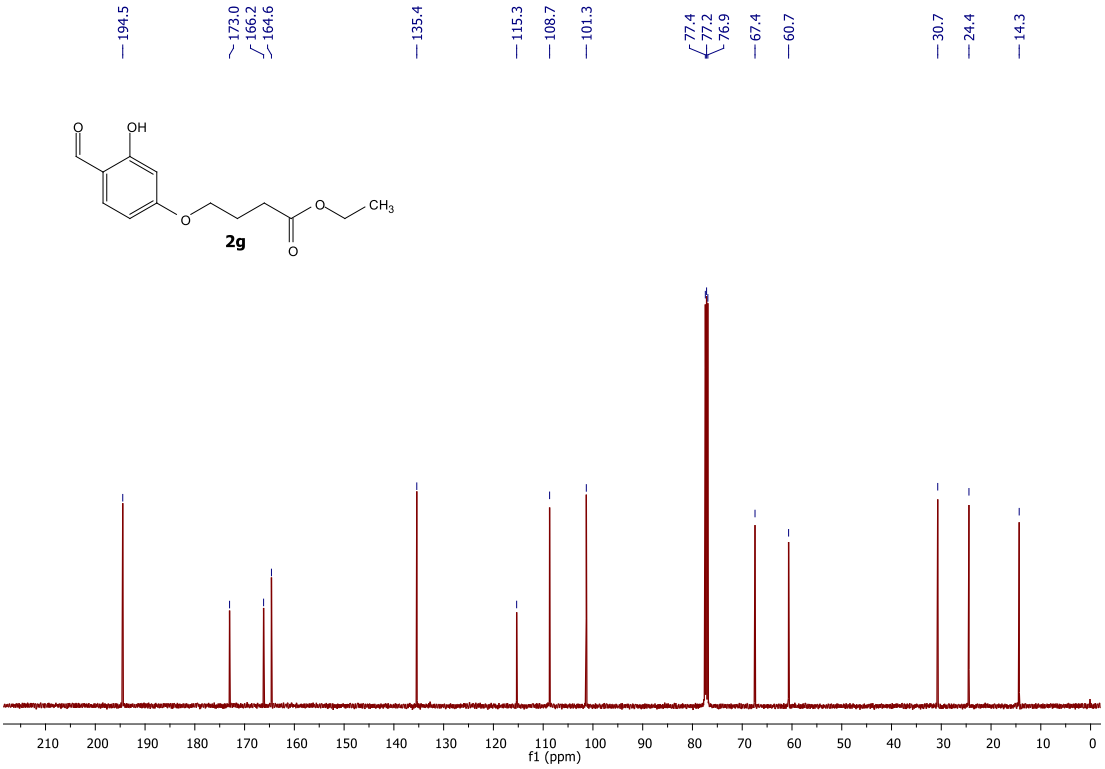
^{13}C NMR



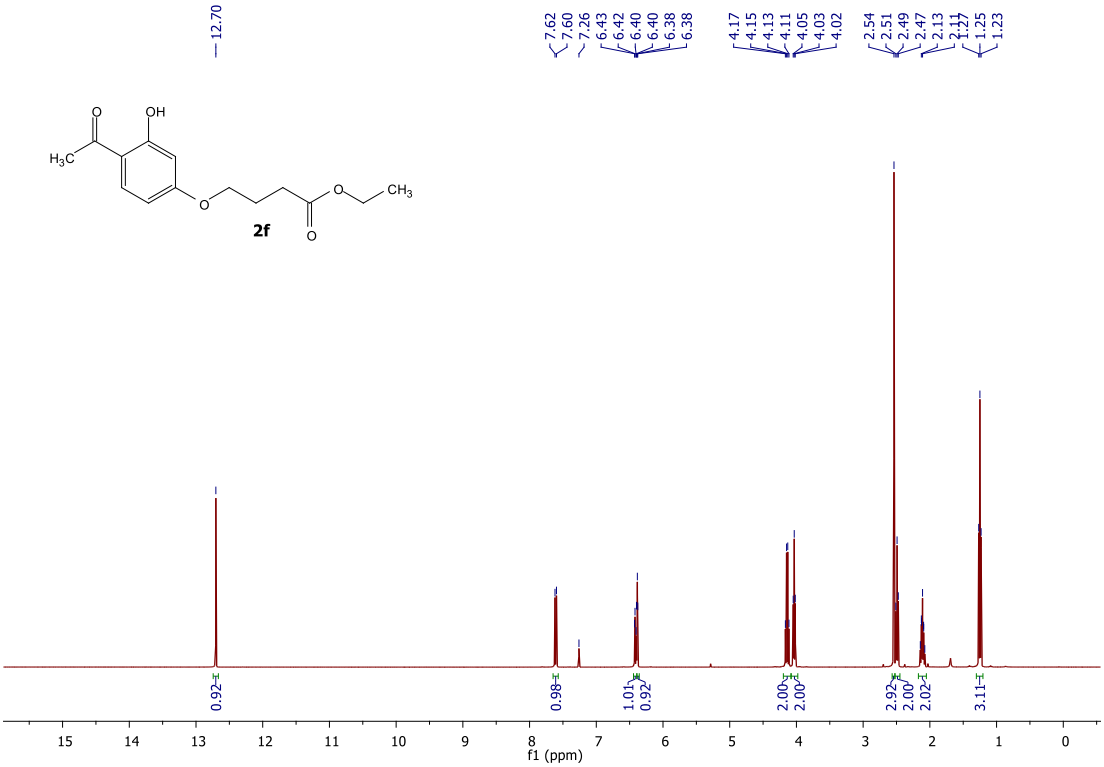
^{1}H NMR



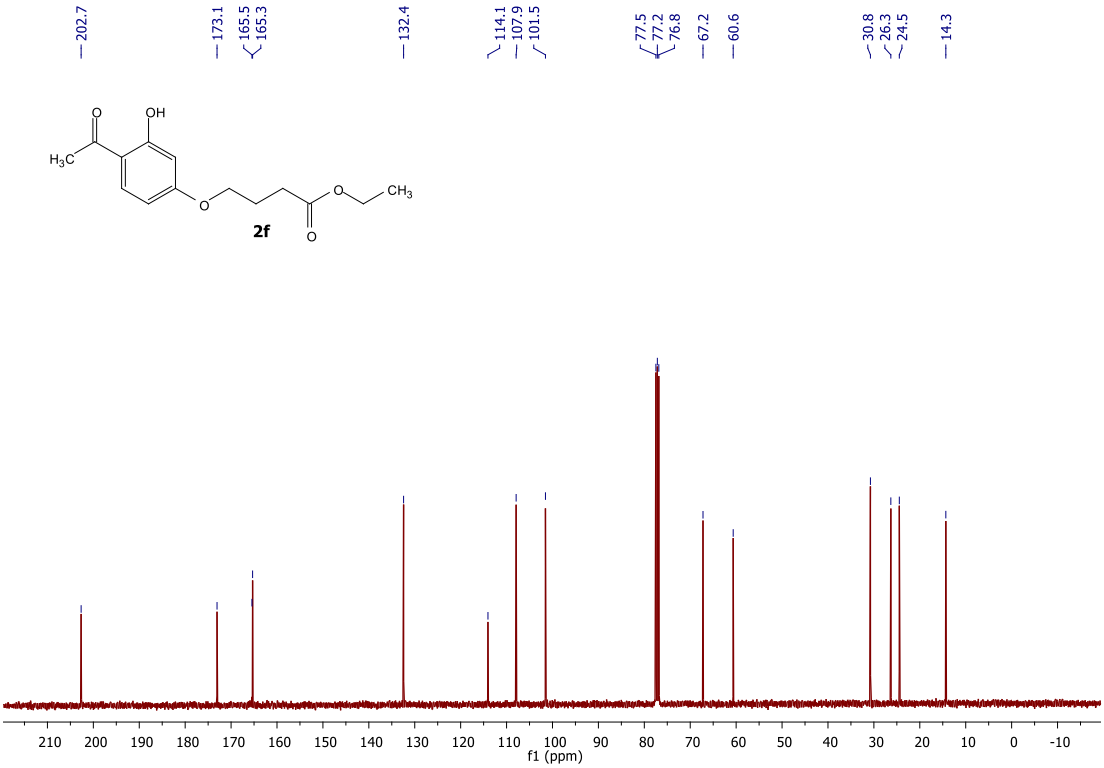
^{13}C NMR



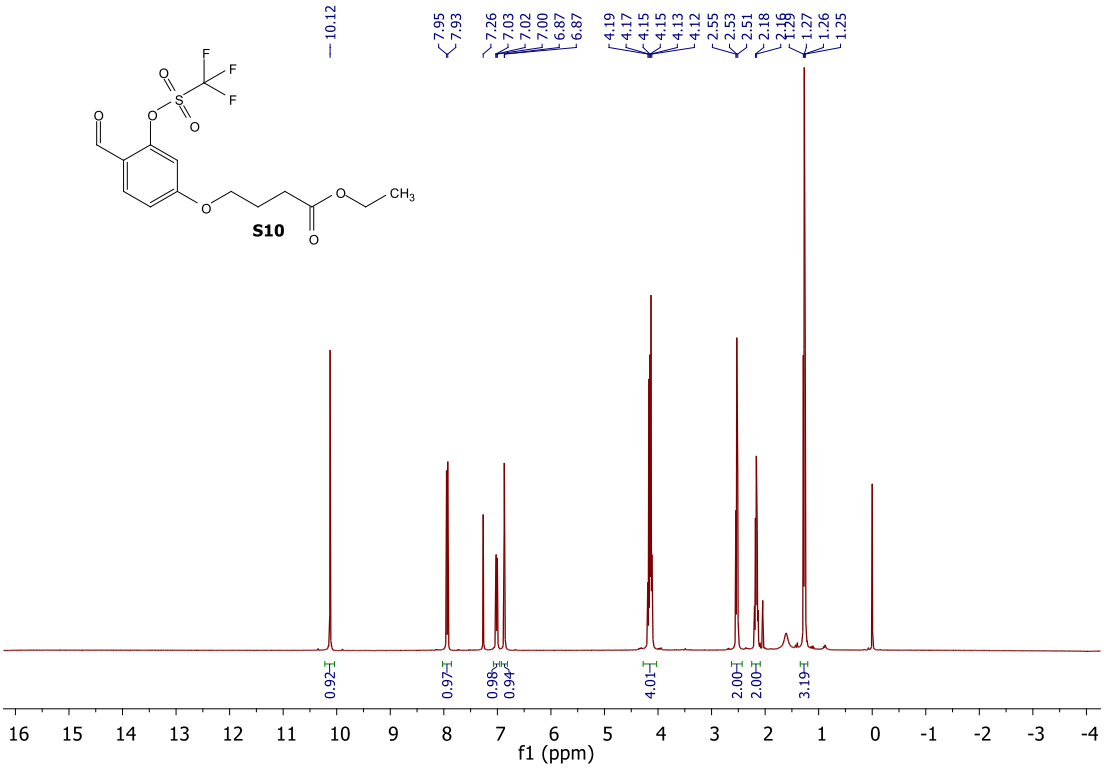
^{1}H NMR



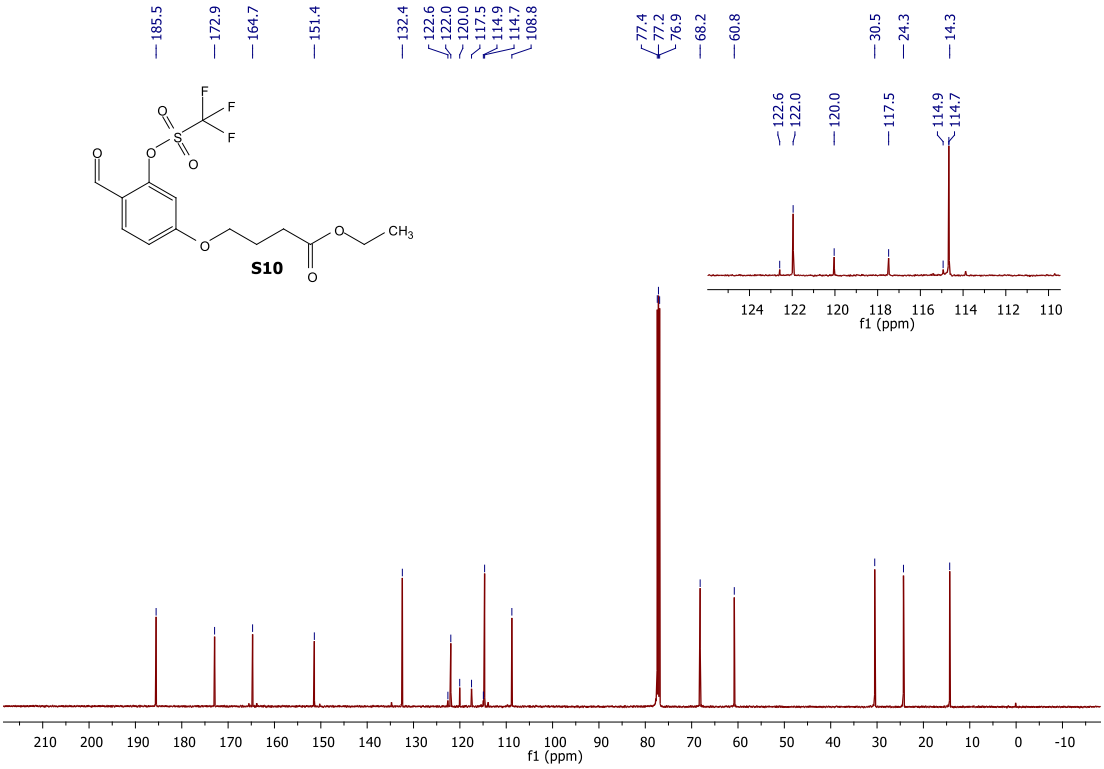
^{13}C NMR



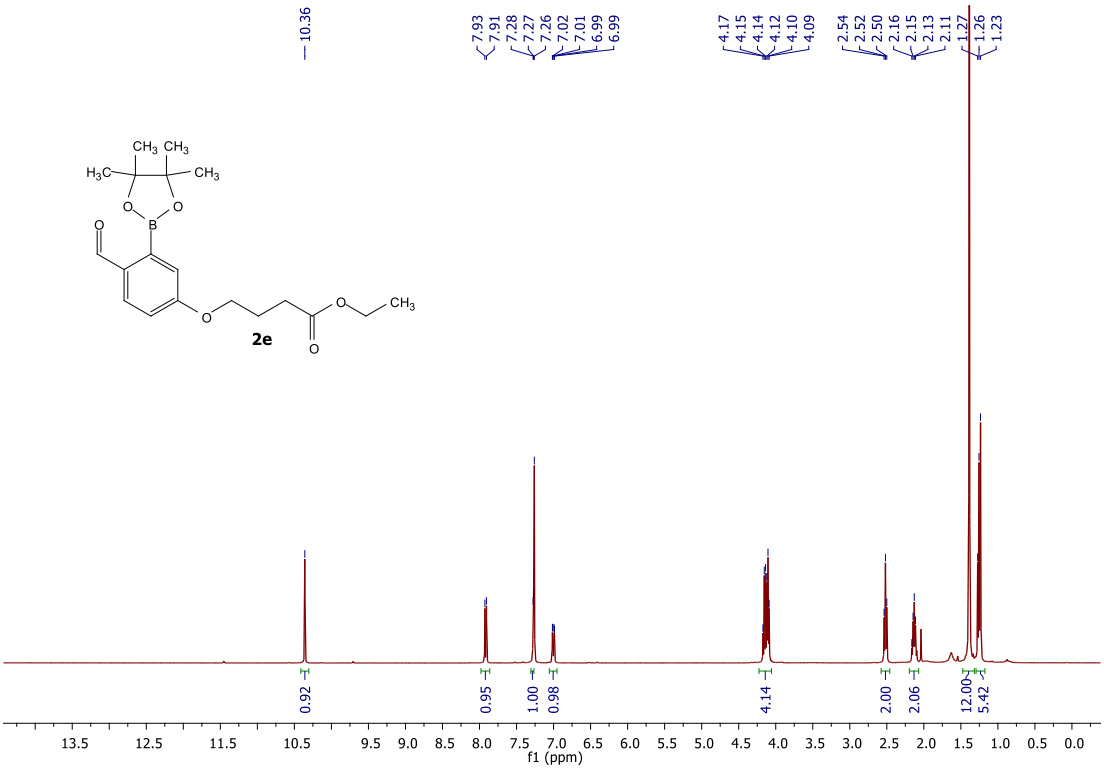
^{1}H NMR



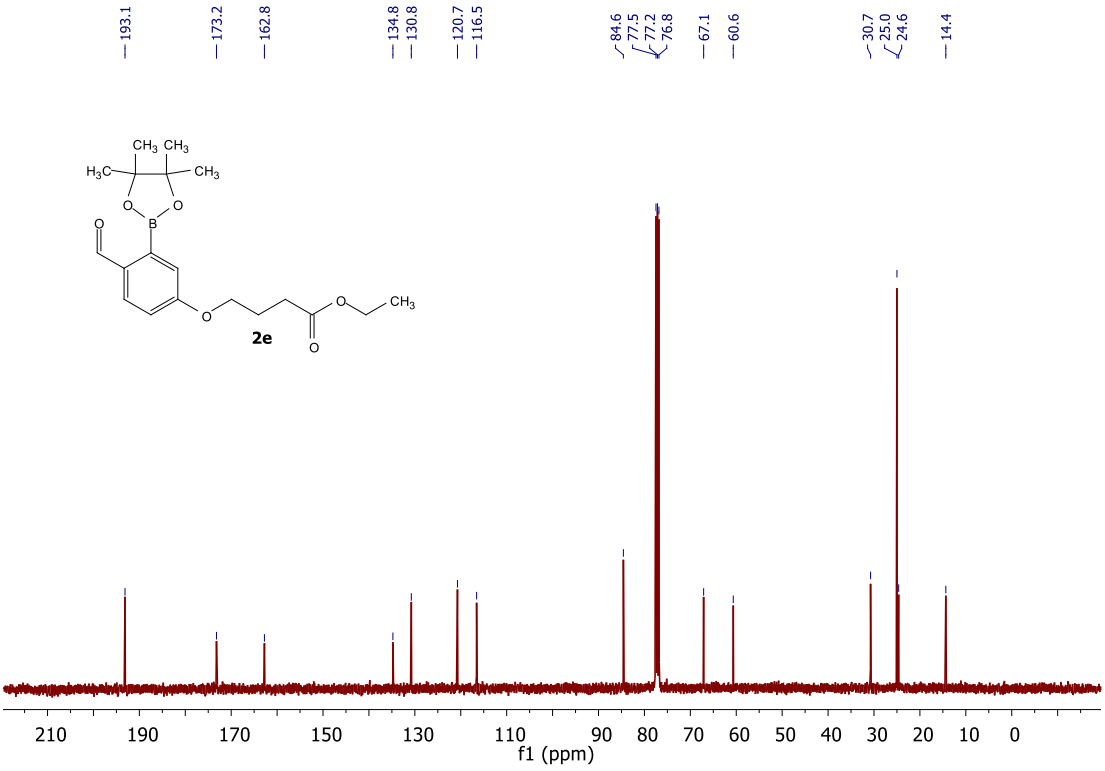
^{13}C NMR



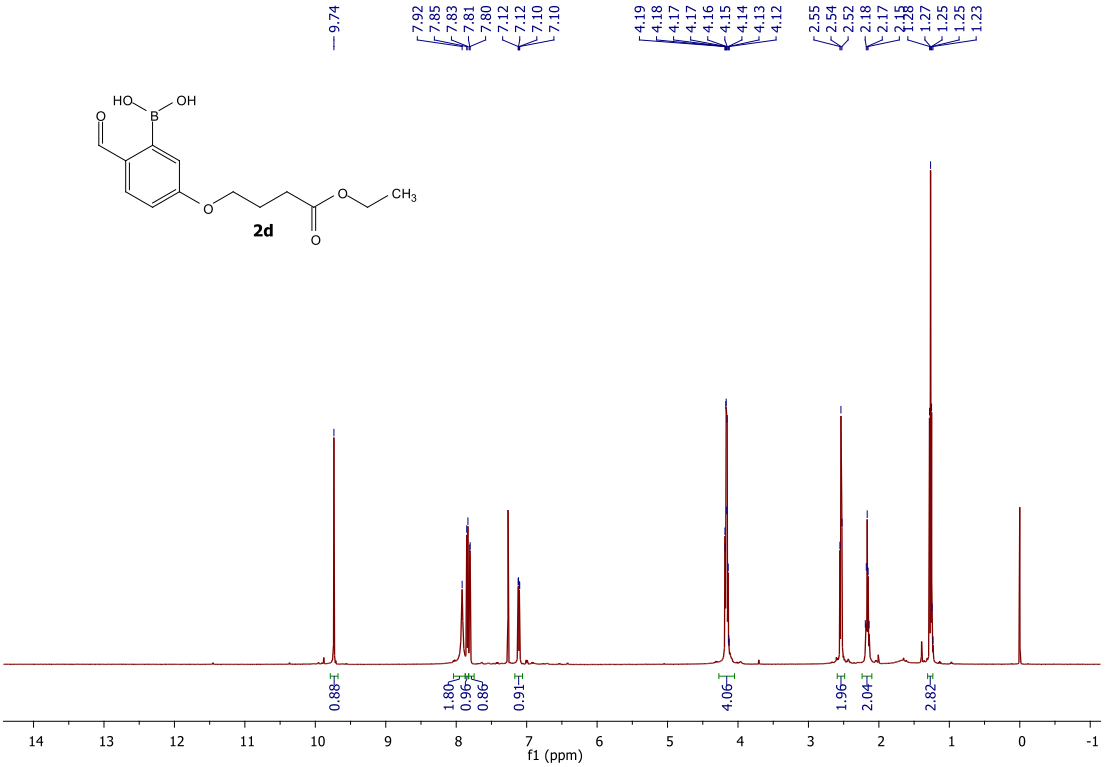
^{1}H NMR



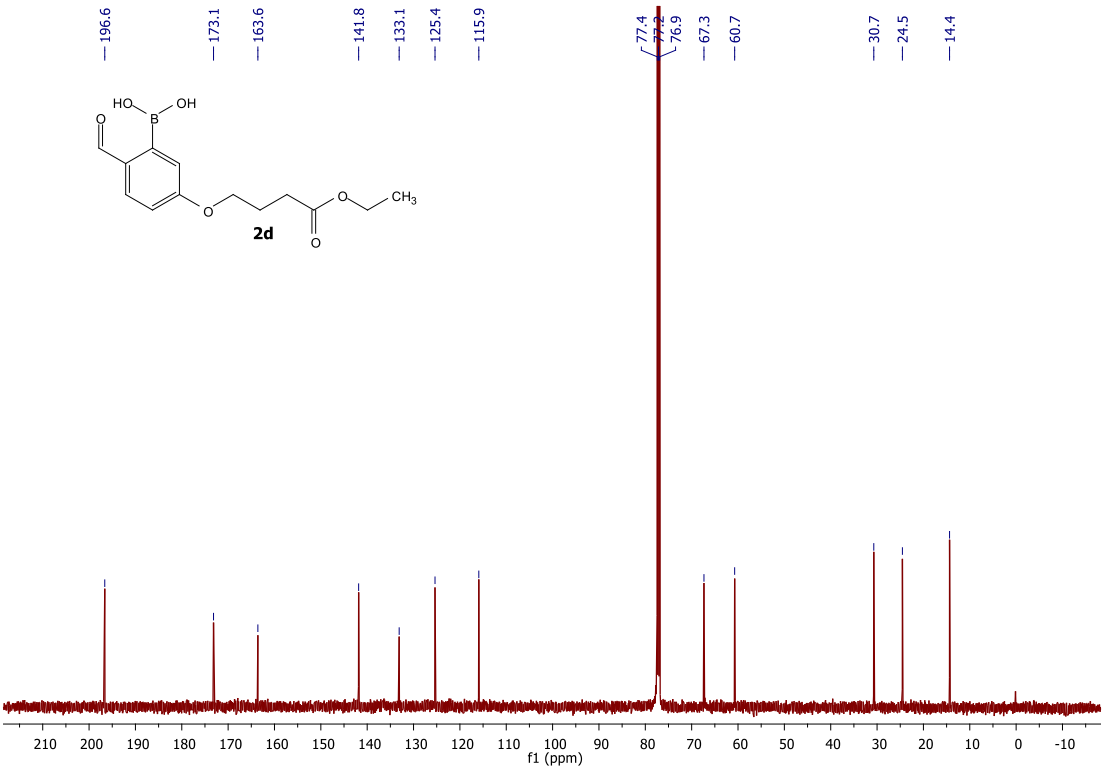
^{13}C NMR



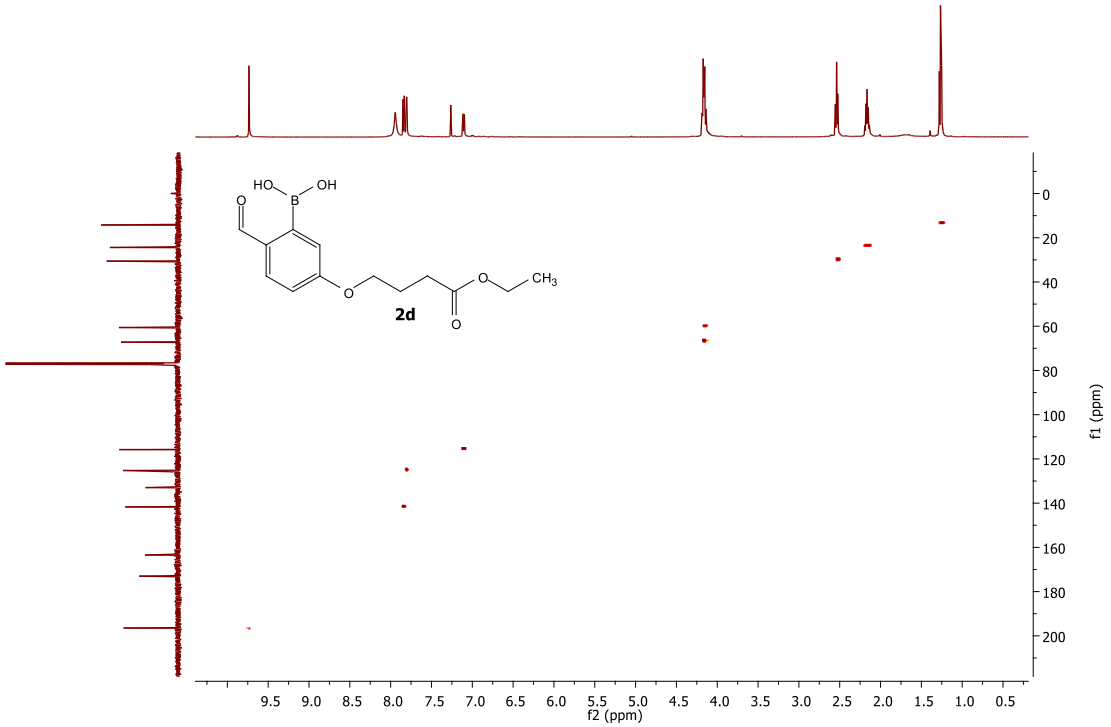
^{1}H NMR



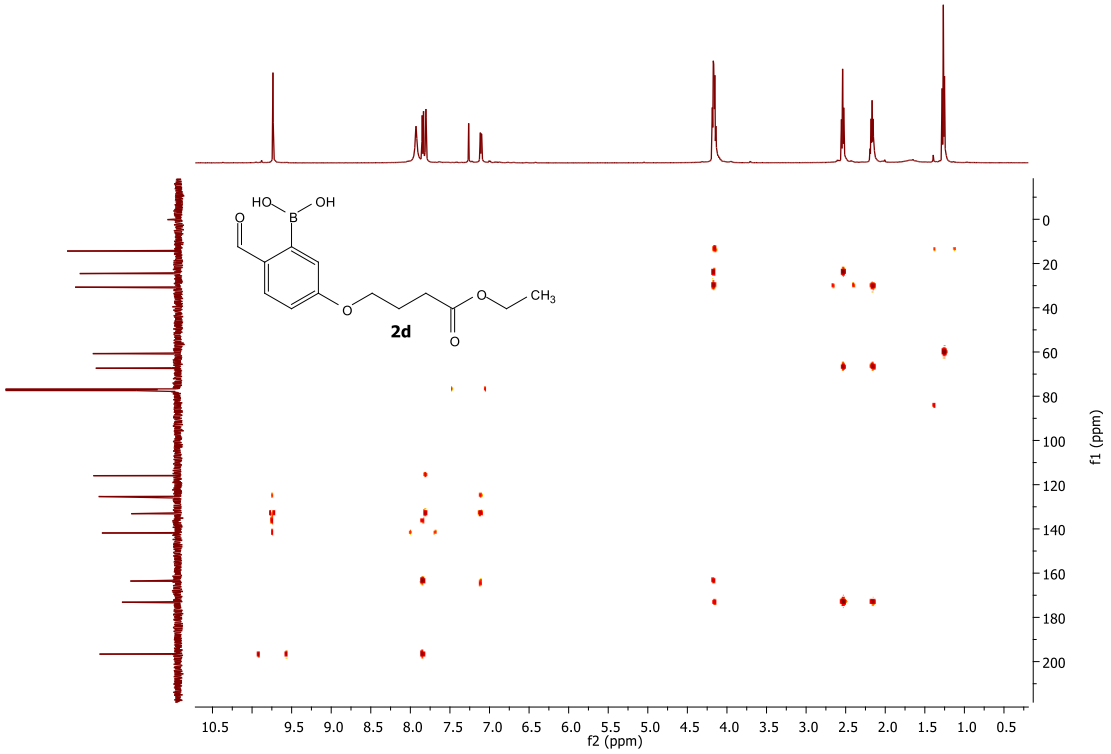
^{13}C NMR



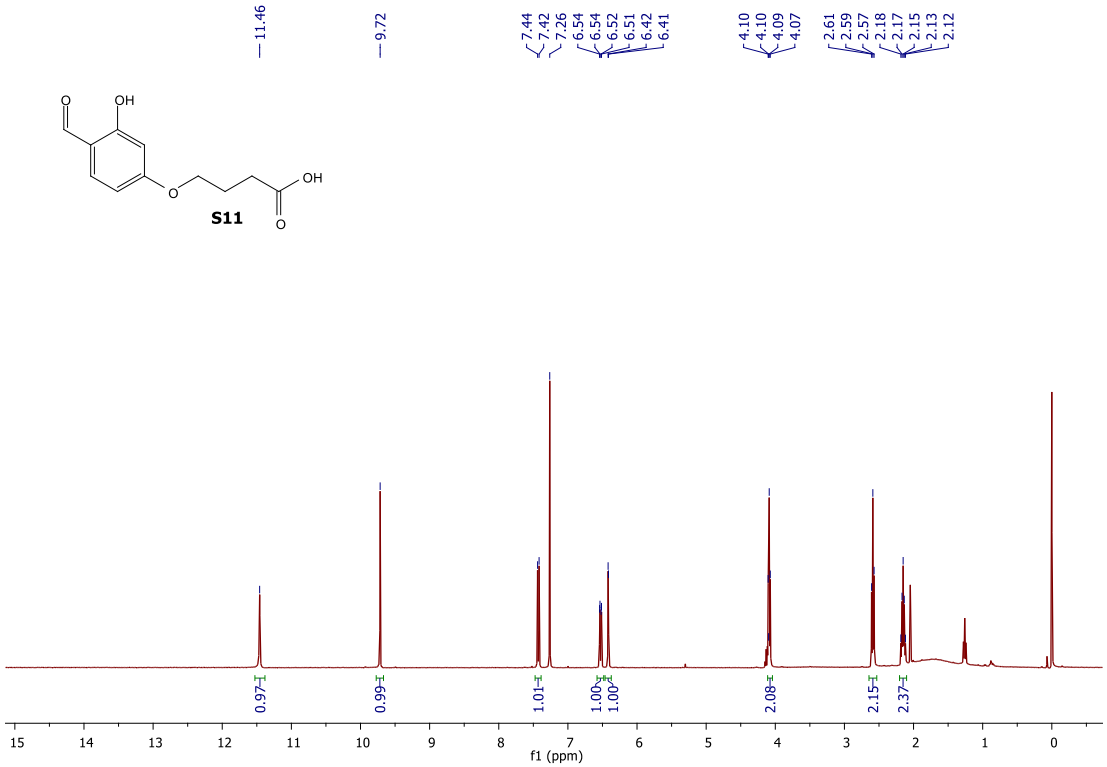
^{1}H-^{13}C correlation (HSQC)



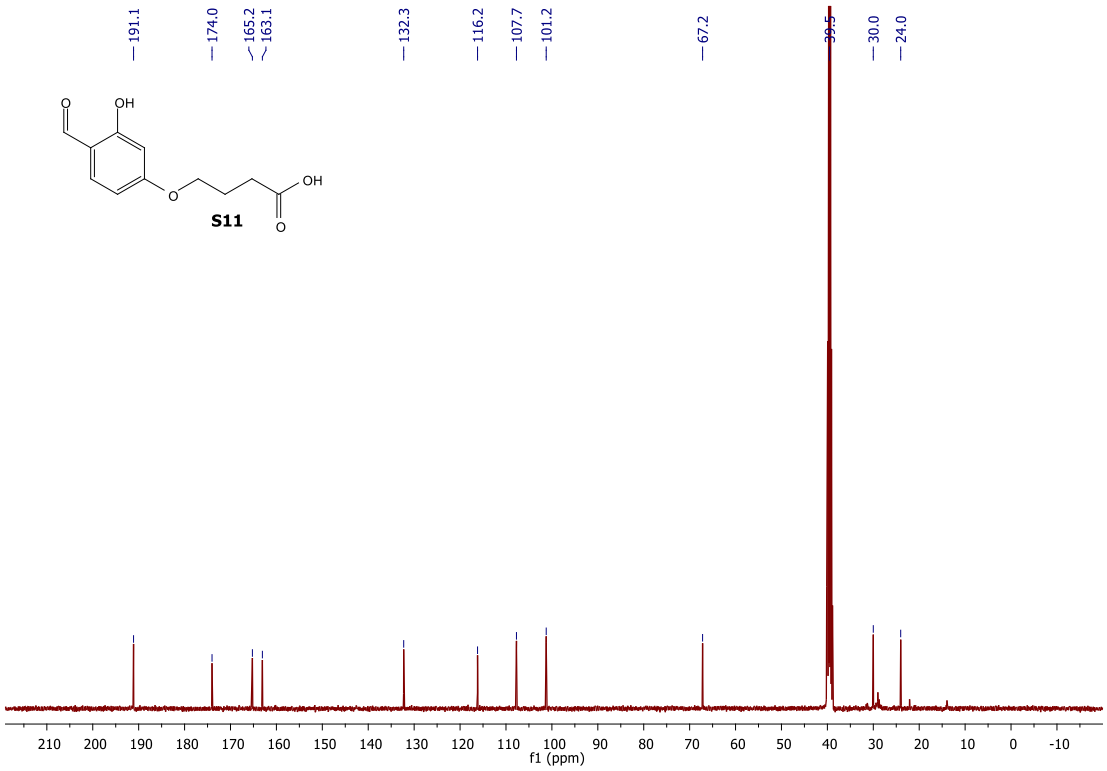
^{1}H-^{13}C correlation (HMBC)



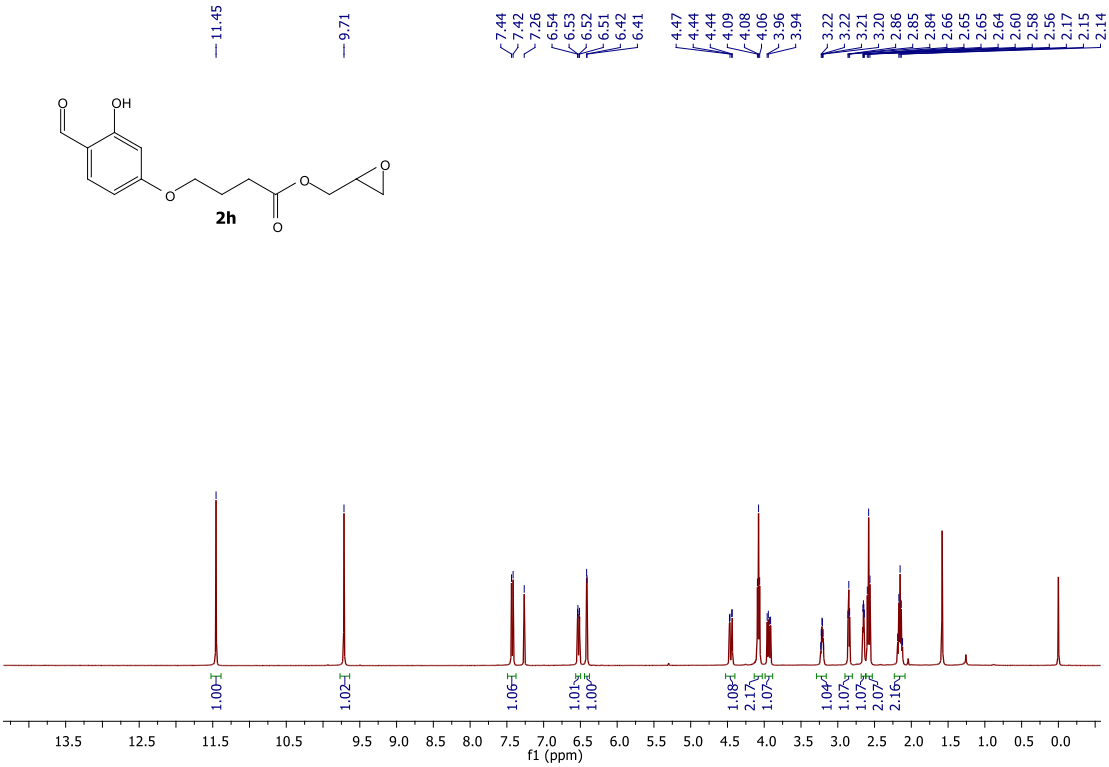
^{1}H NMR



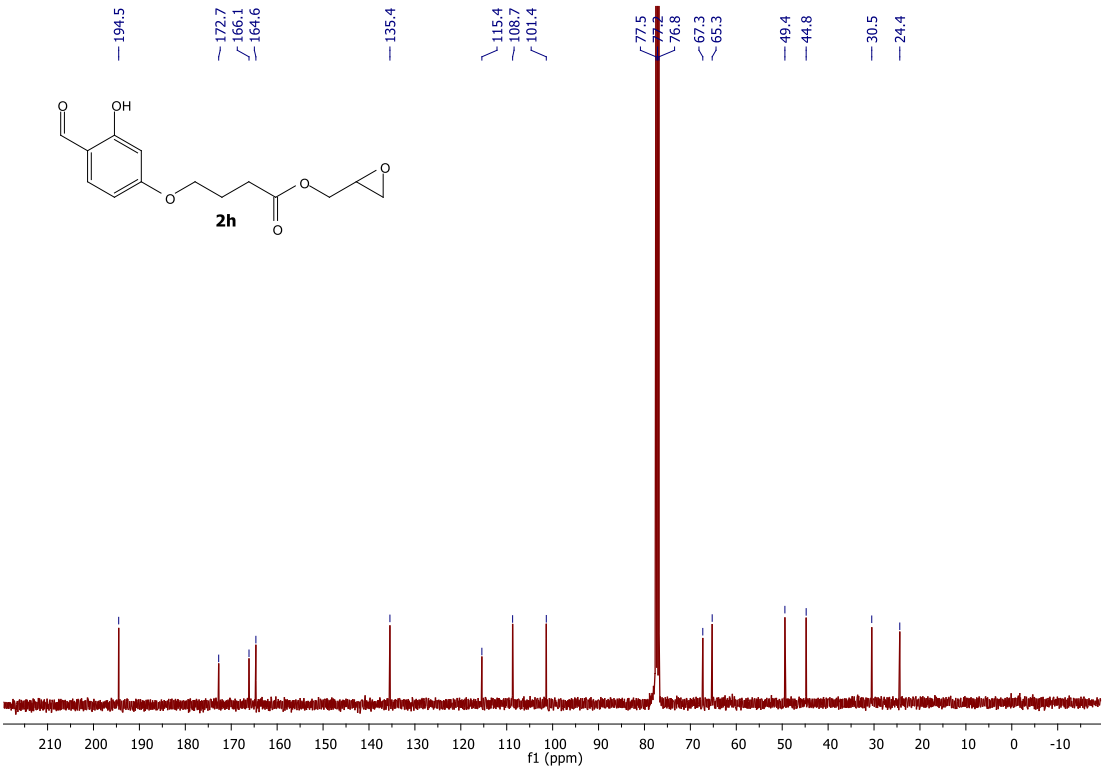
^{13}C NMR



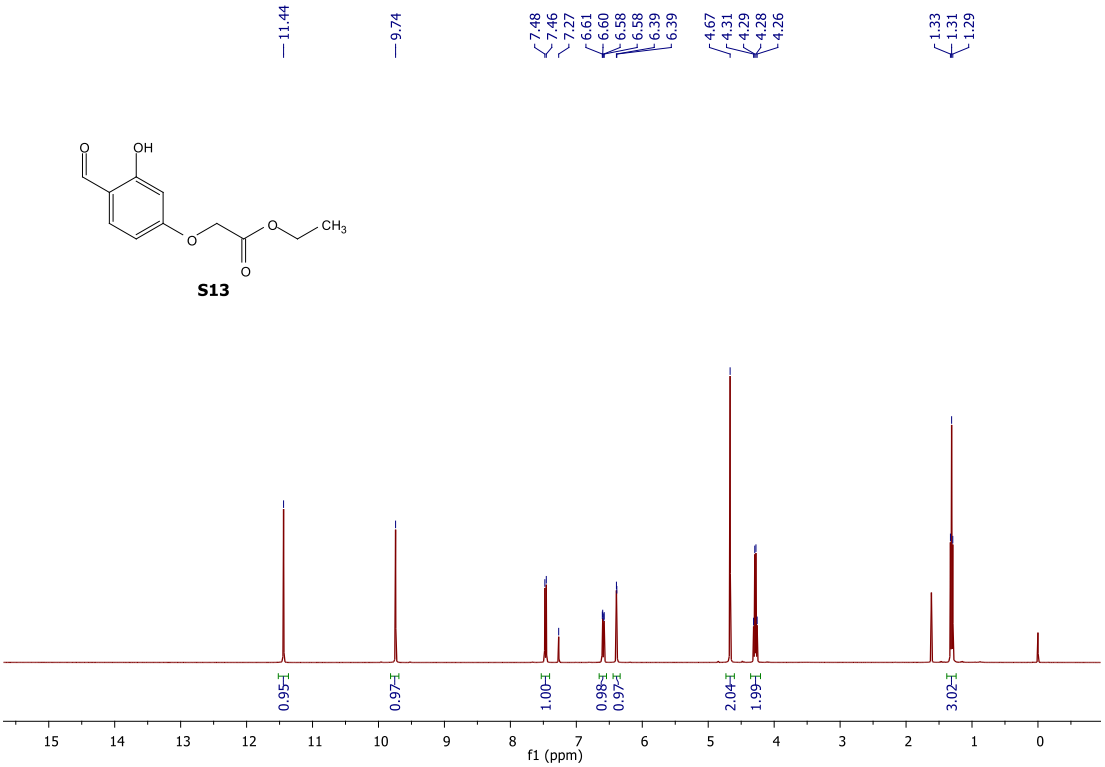
^{1}H NMR



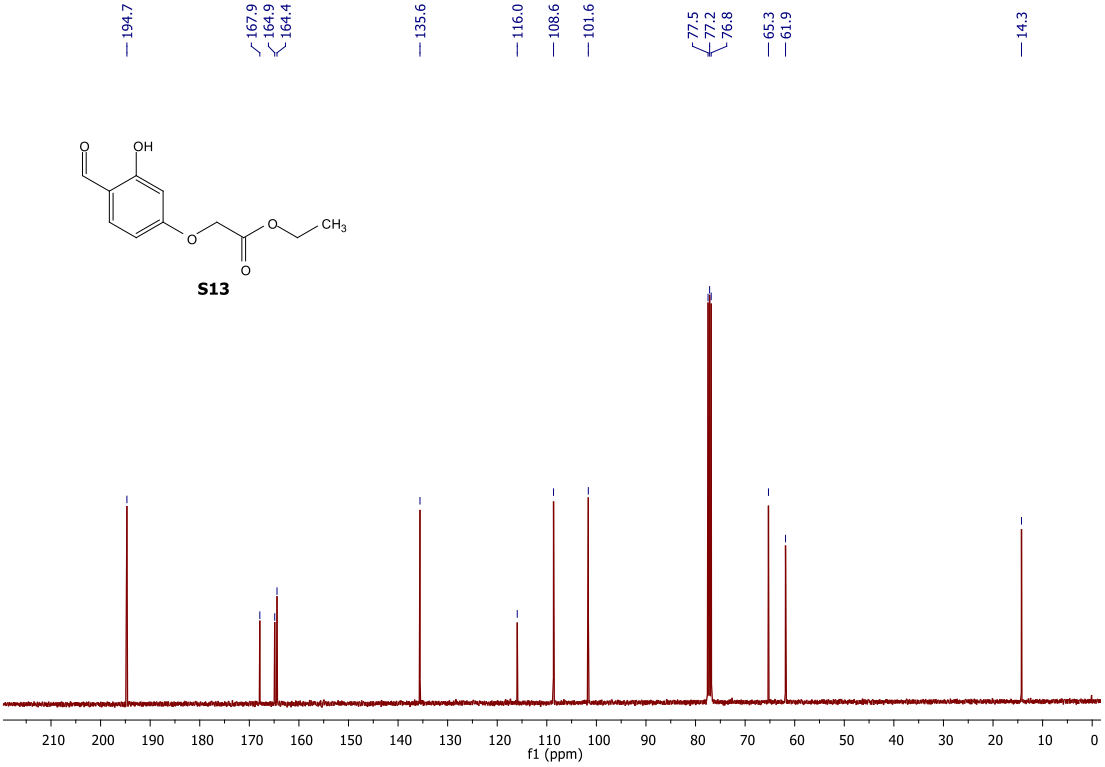
^{13}C NMR



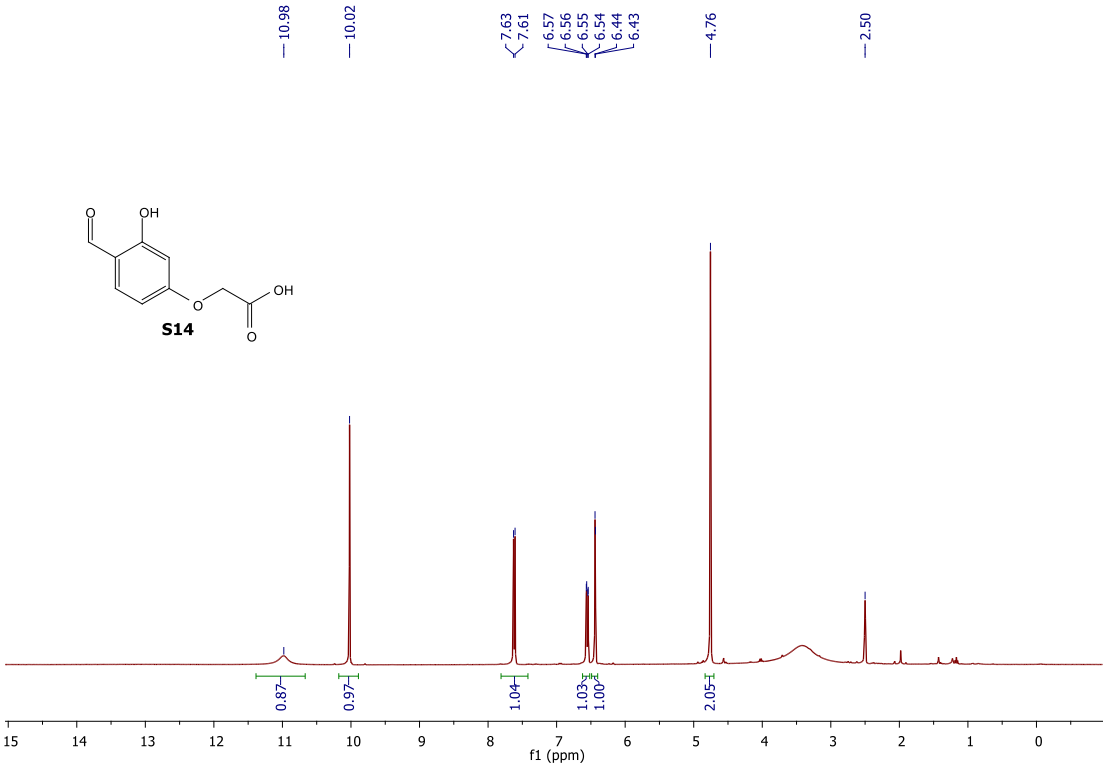
^{1}H NMR



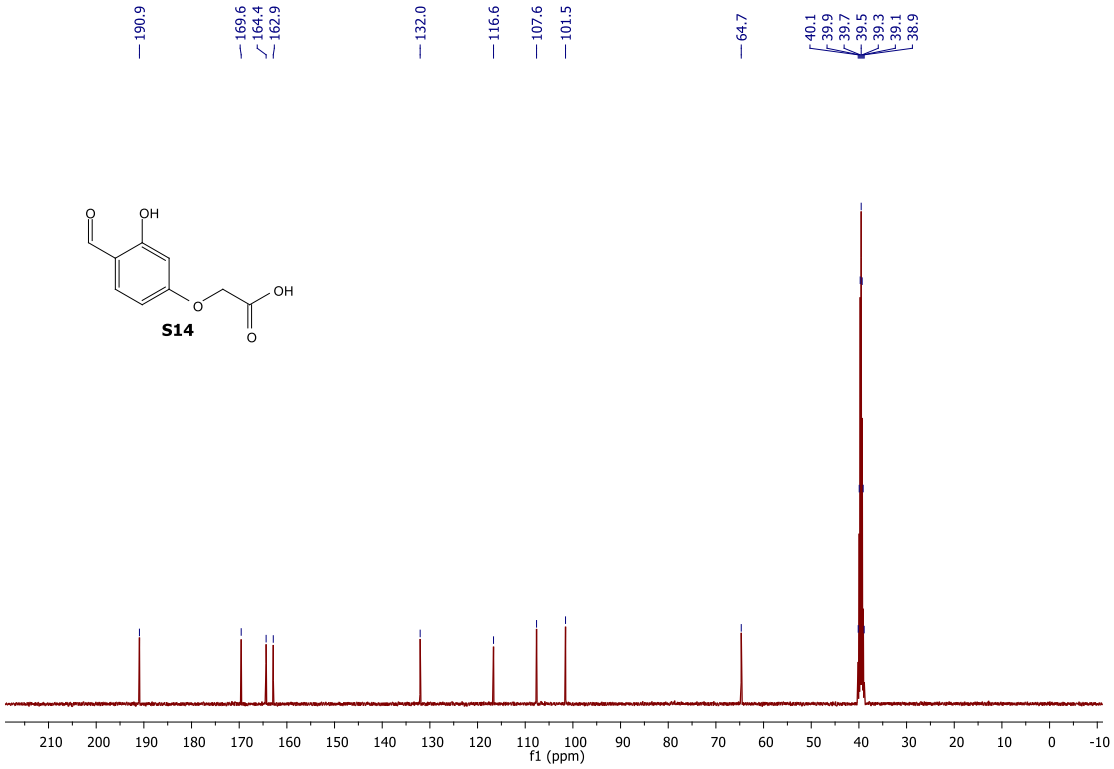
^{13}C NMR



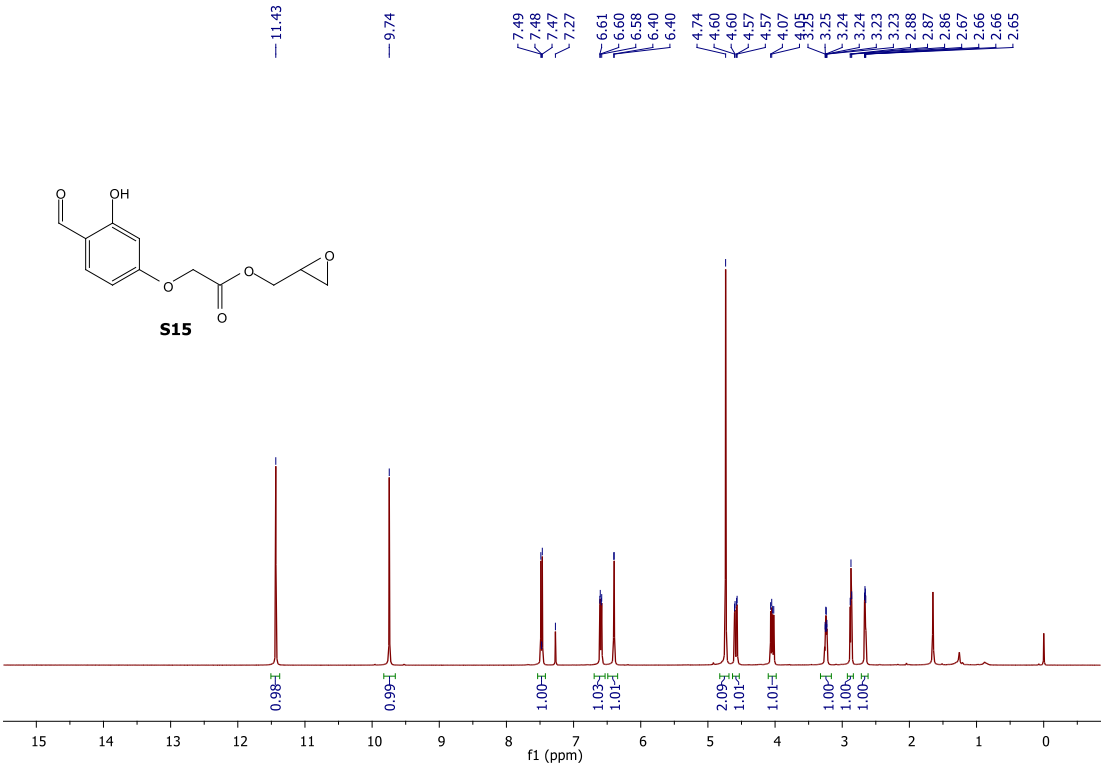
^{1}H NMR



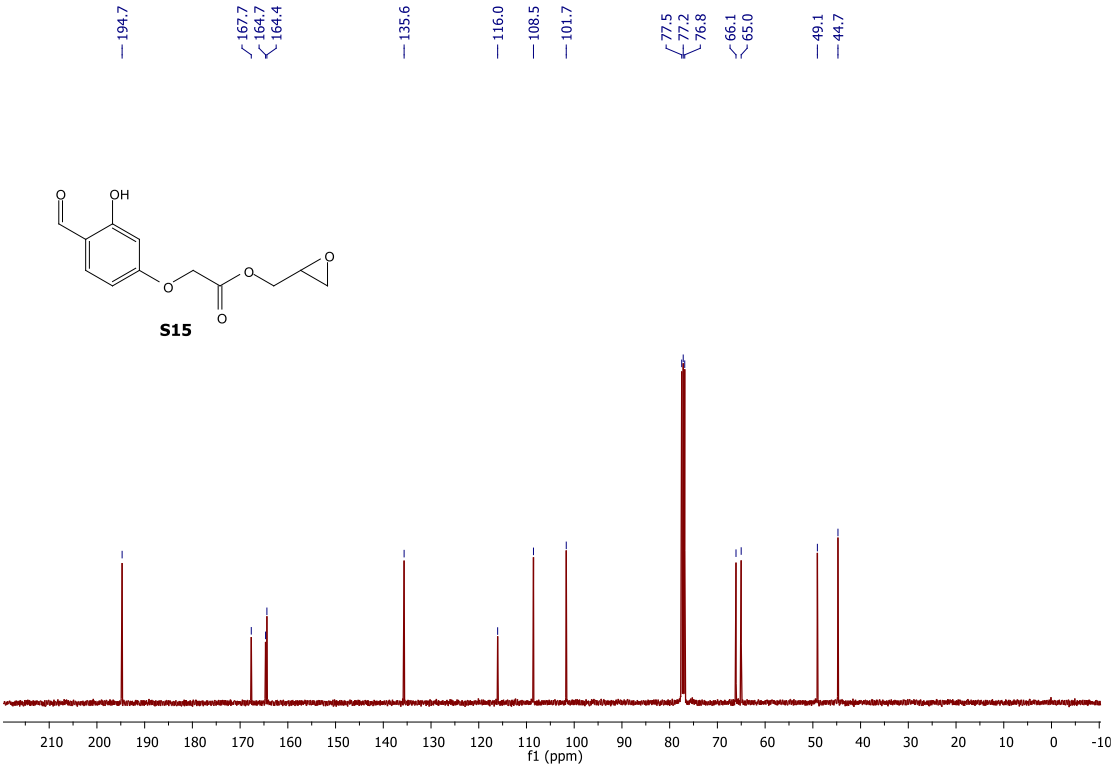
^{13}C NMR



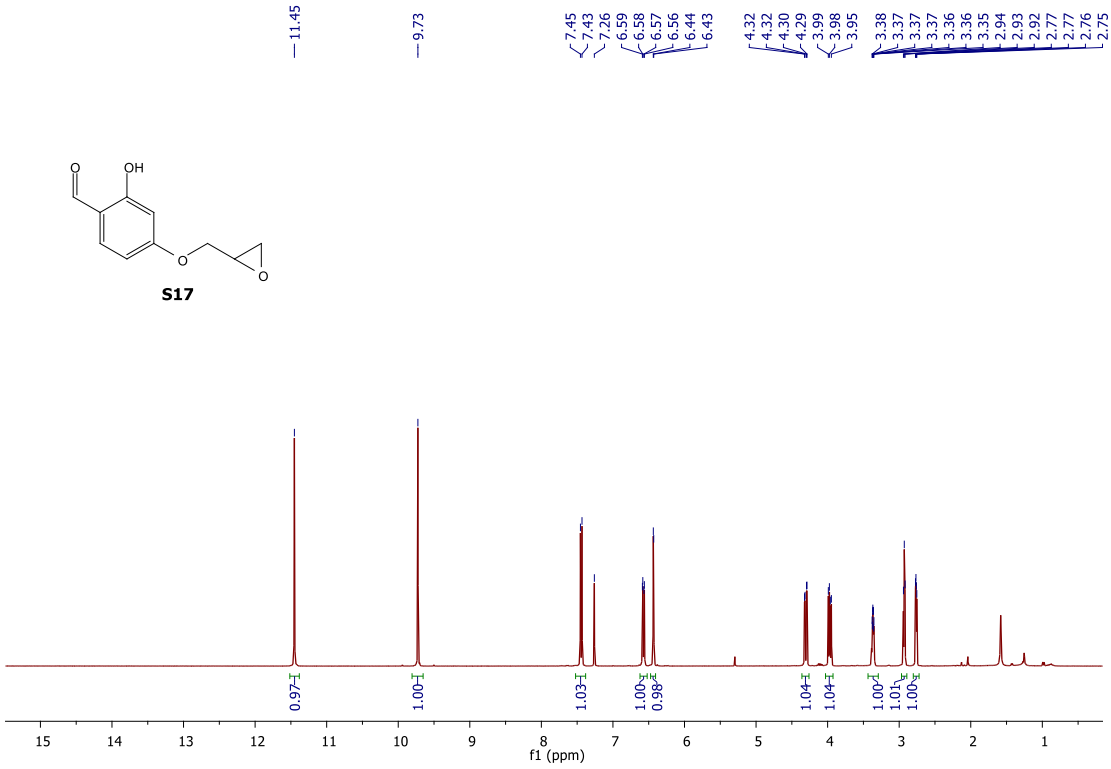
^{1}H NMR



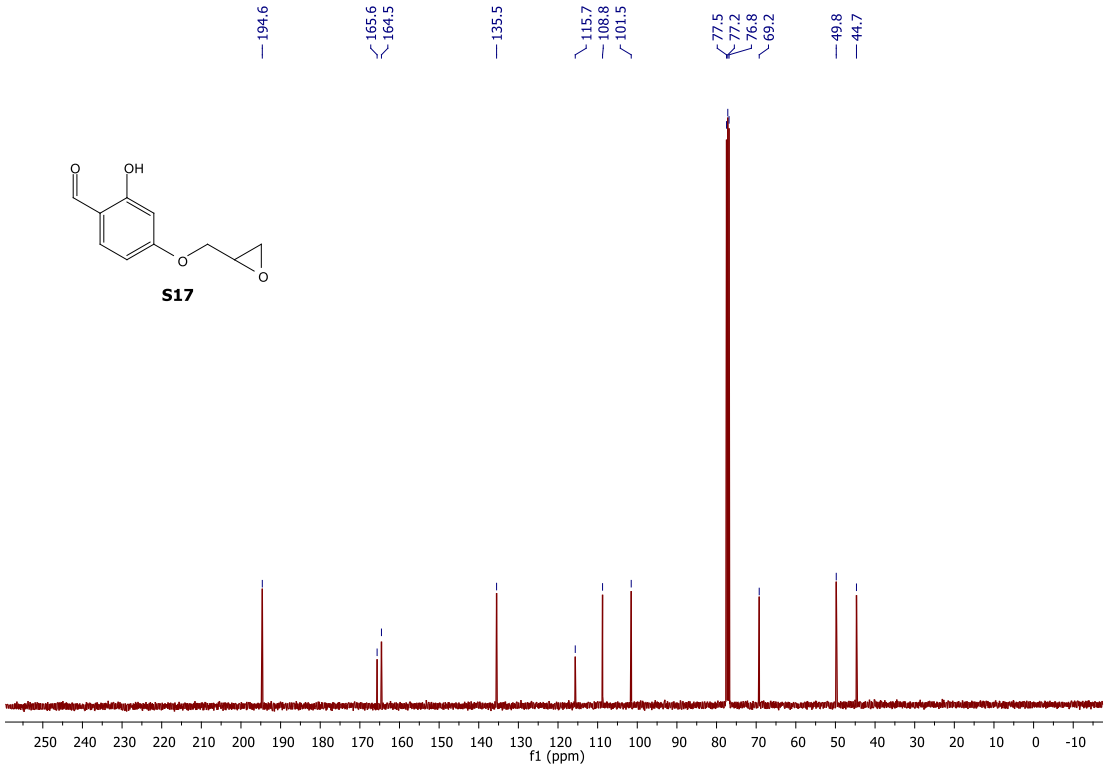
^{13}C NMR



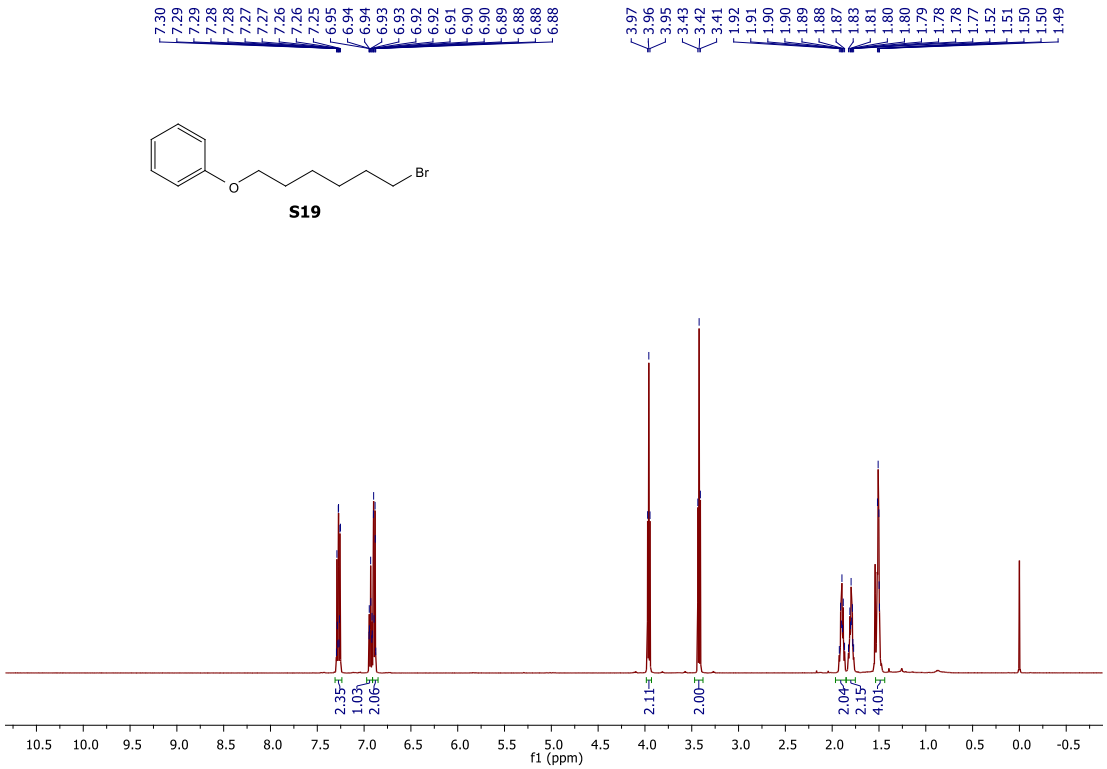
^{1}H NMR



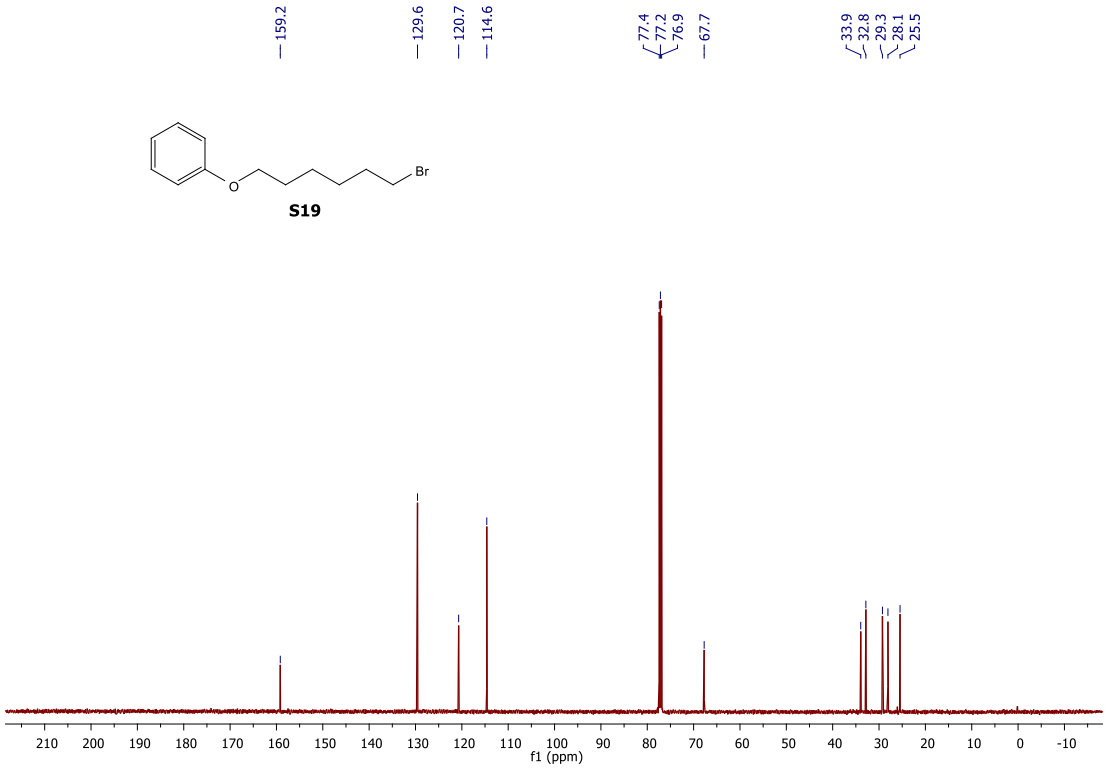
^{13}C NMR



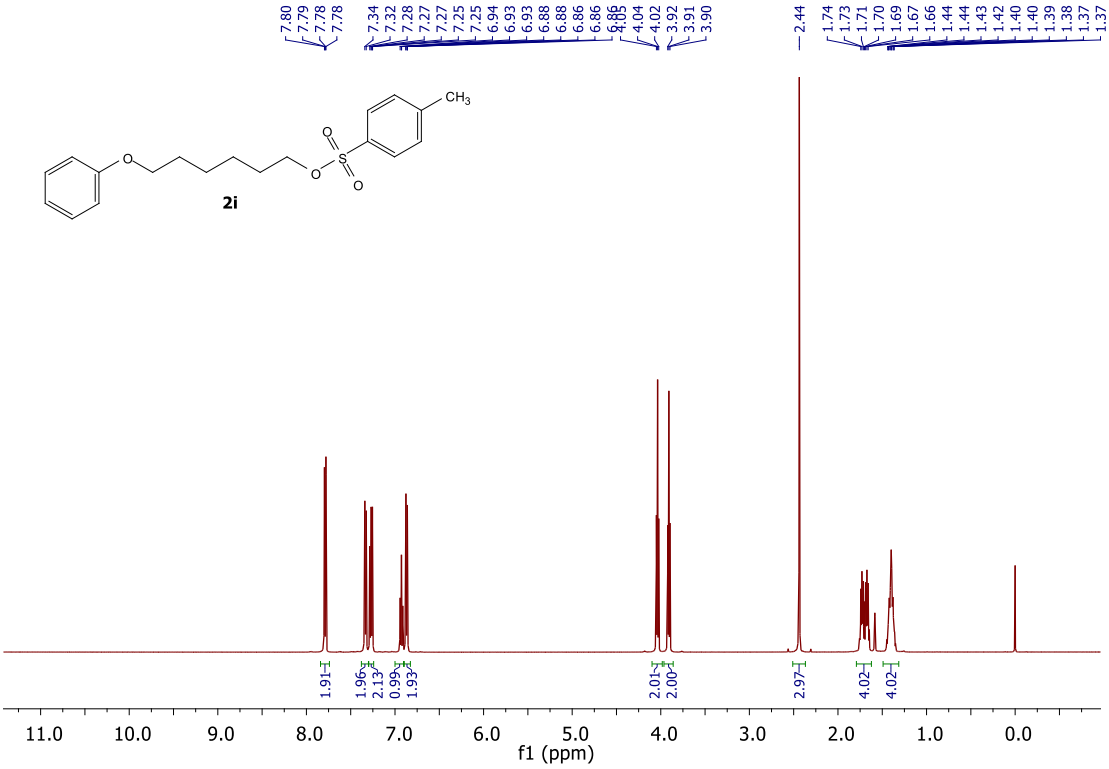
^{1}H NMR



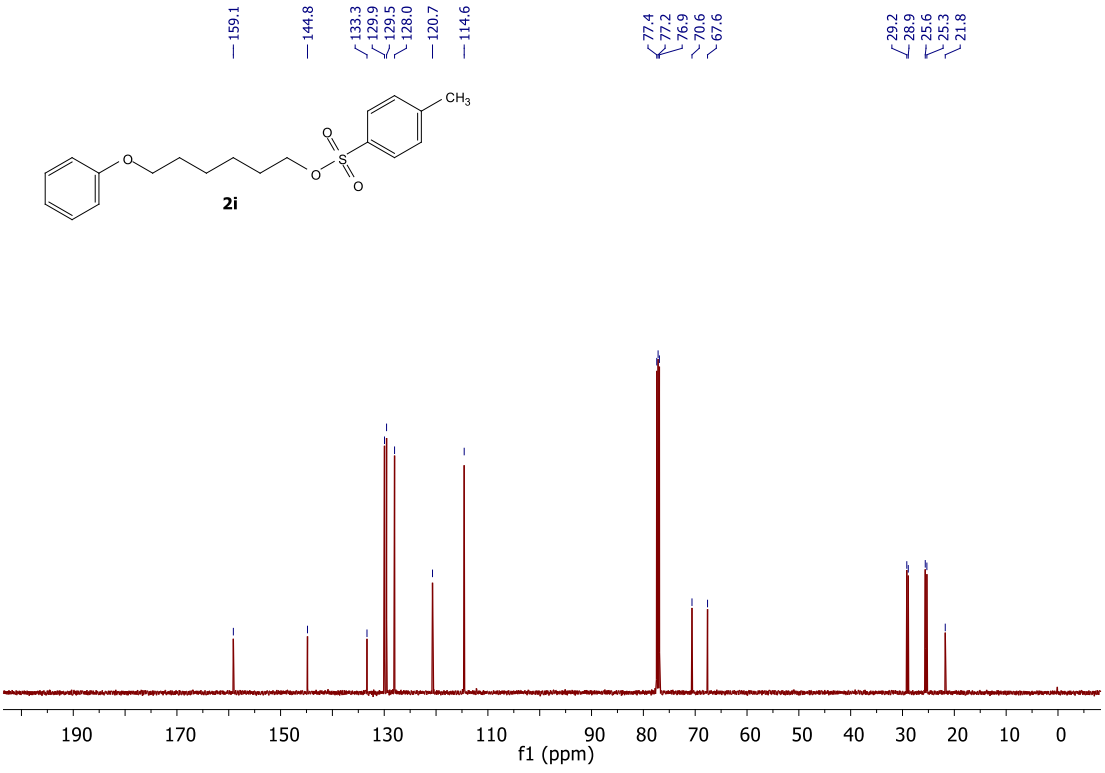
^{13}C NMR



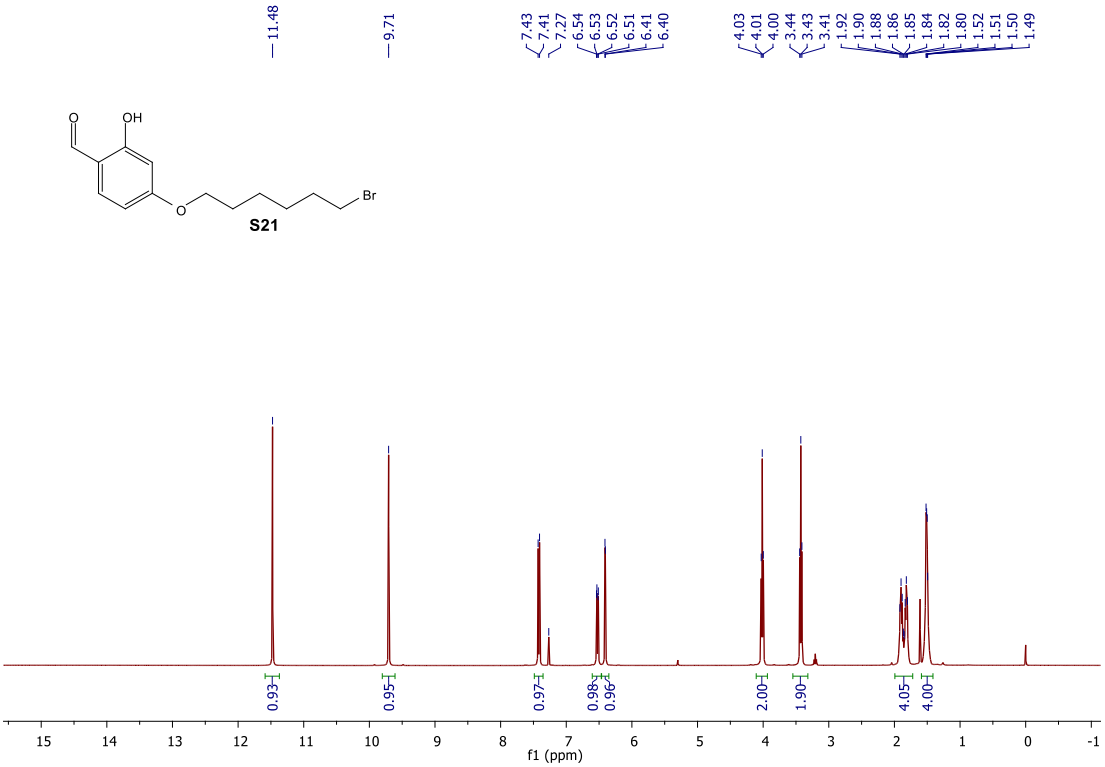
^{1}H NMR



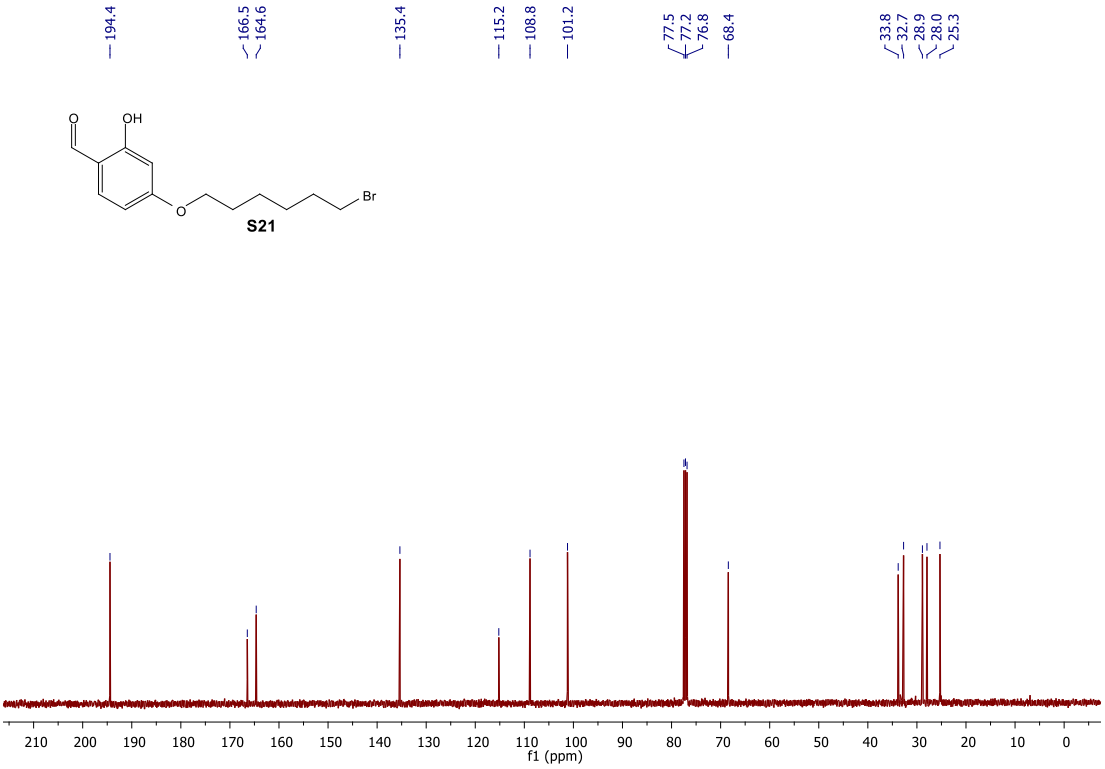
^{13}C NMR



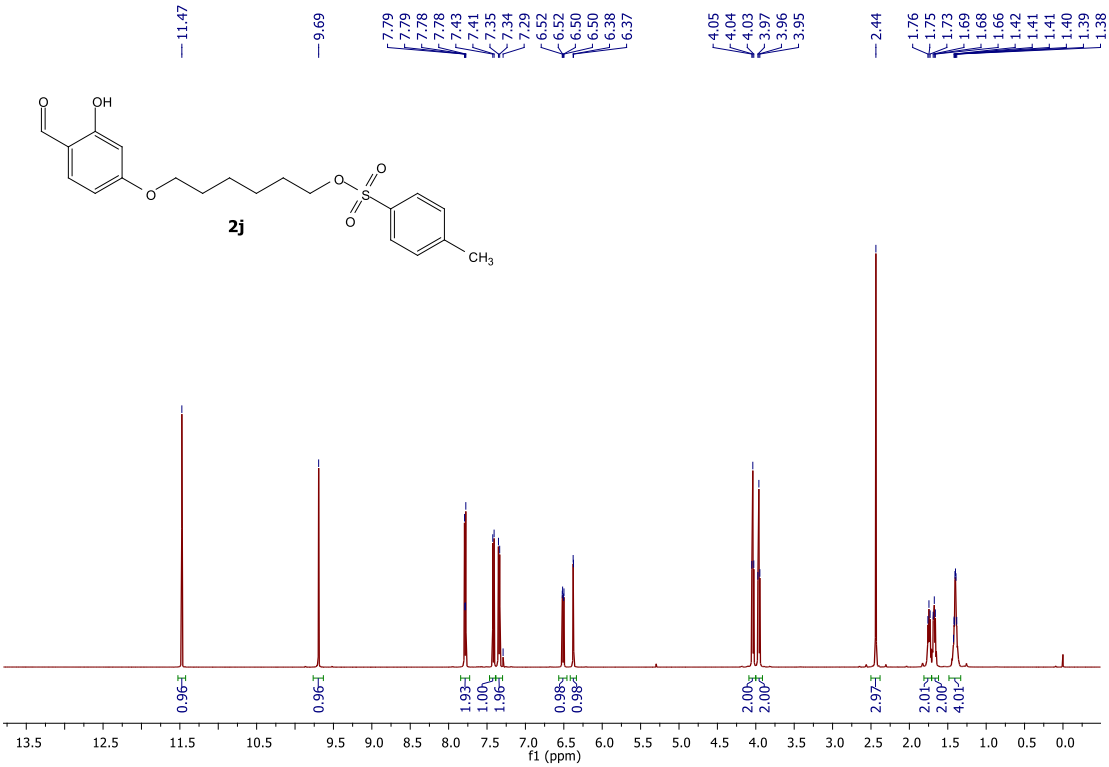
^{1}H NMR



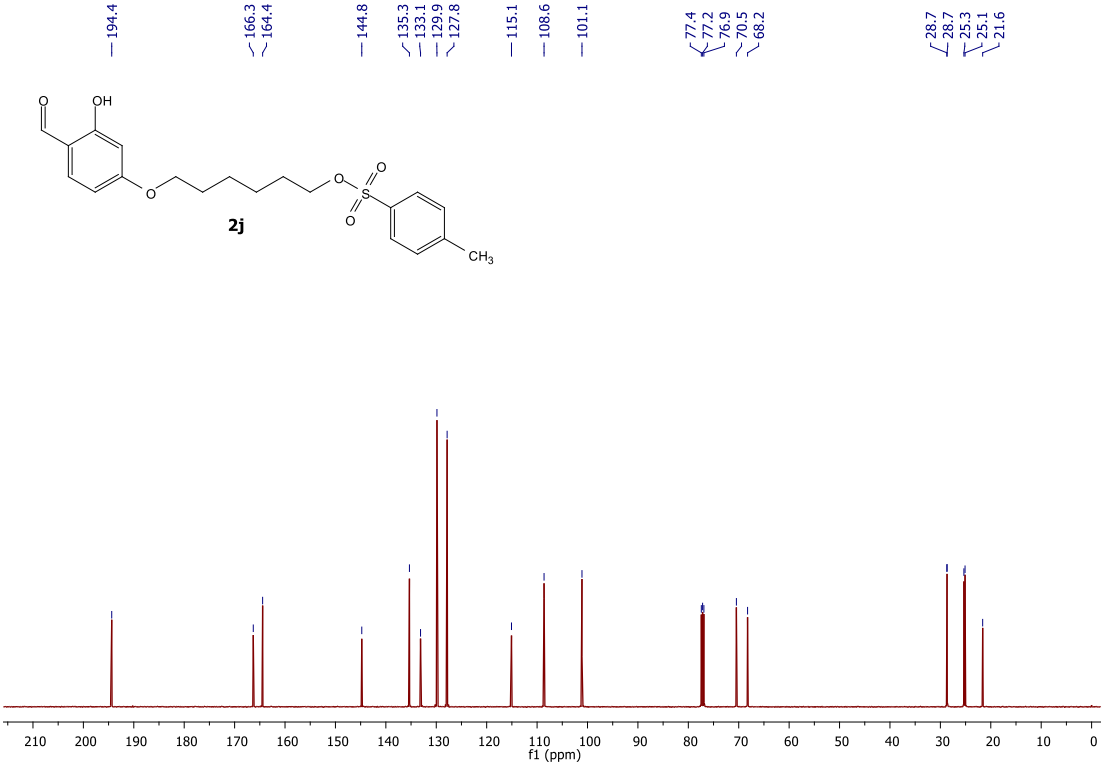
^{13}C NMR



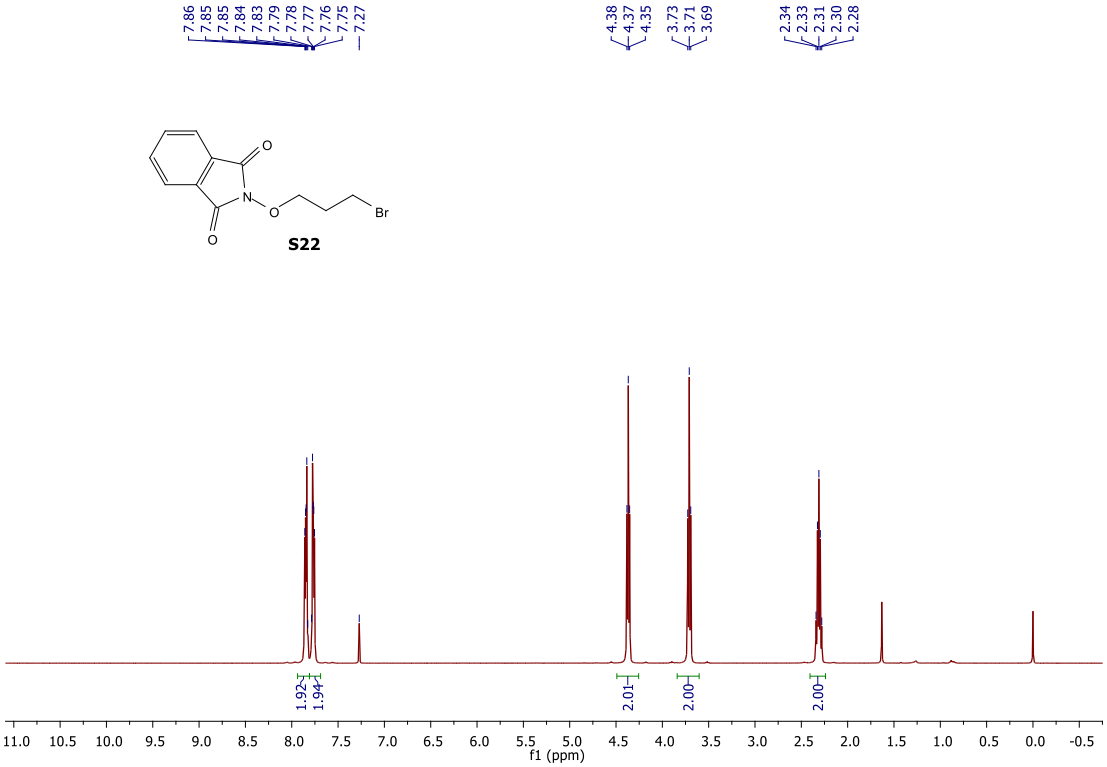
^{1}H NMR



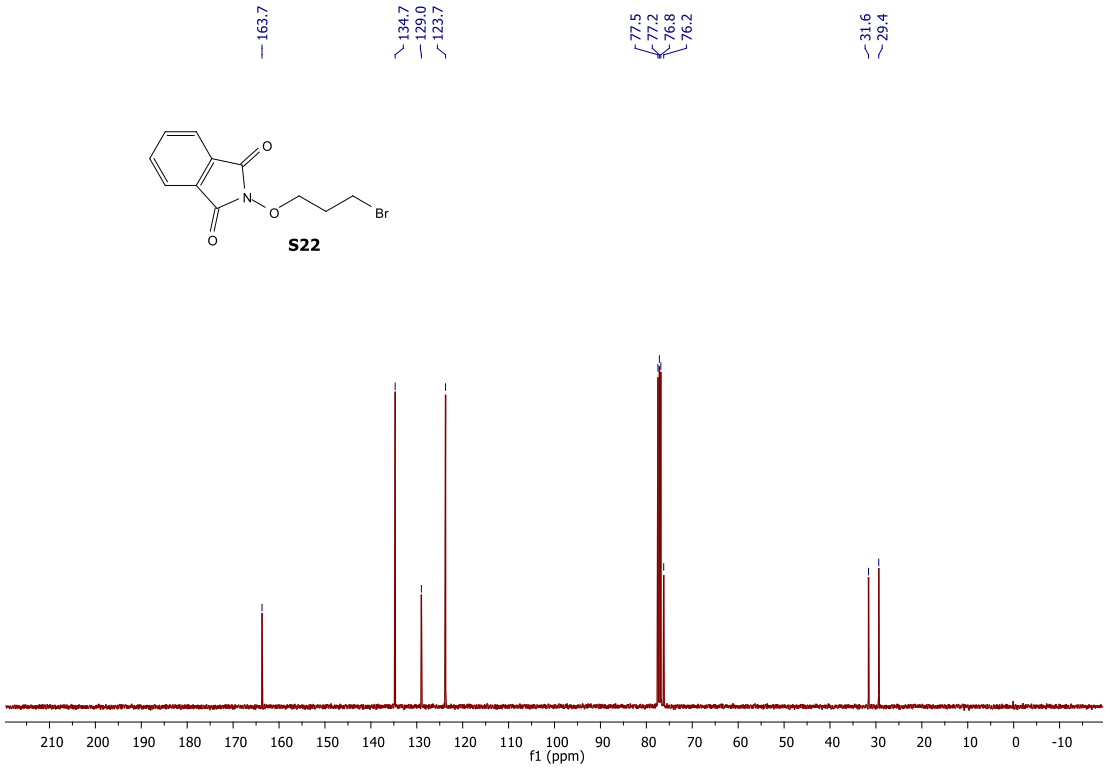
^{13}C NMR



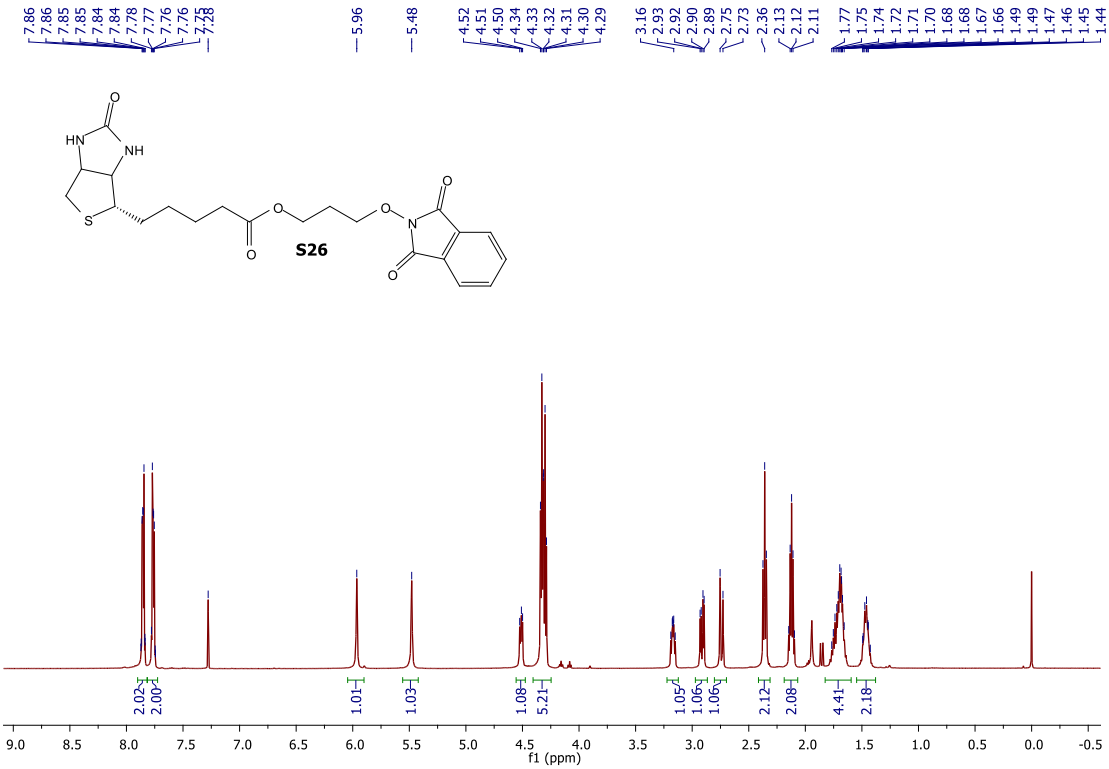
^{1}H NMR



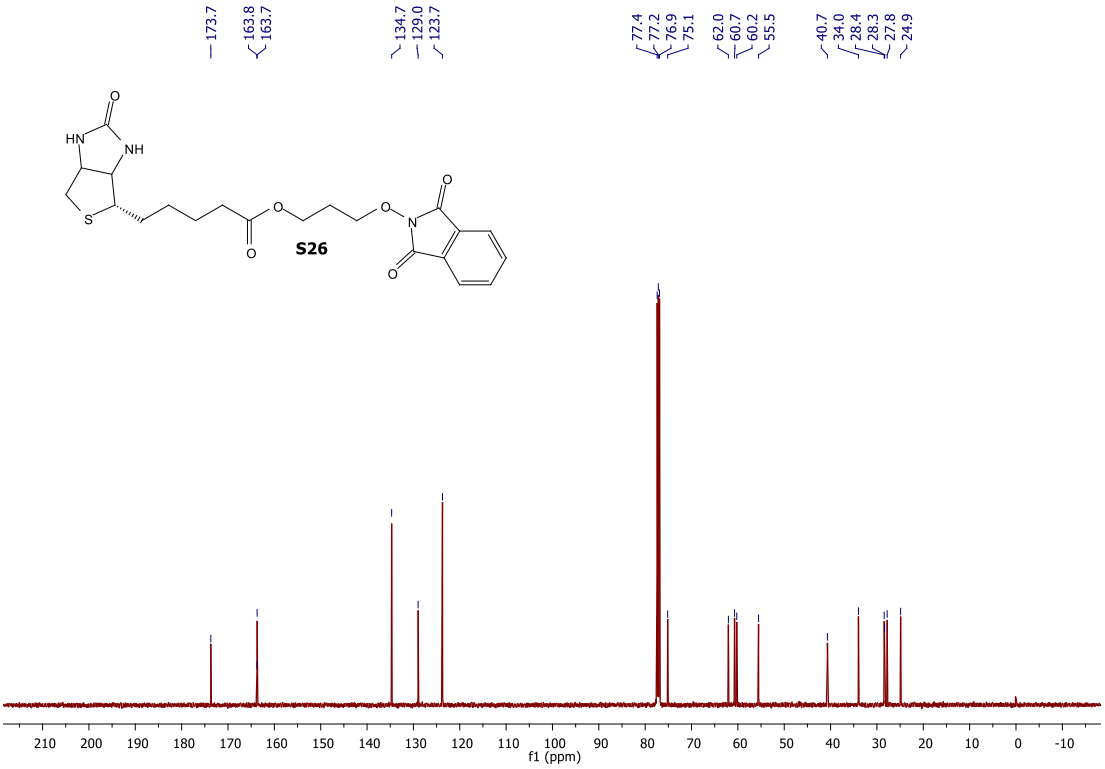
^{13}C NMR



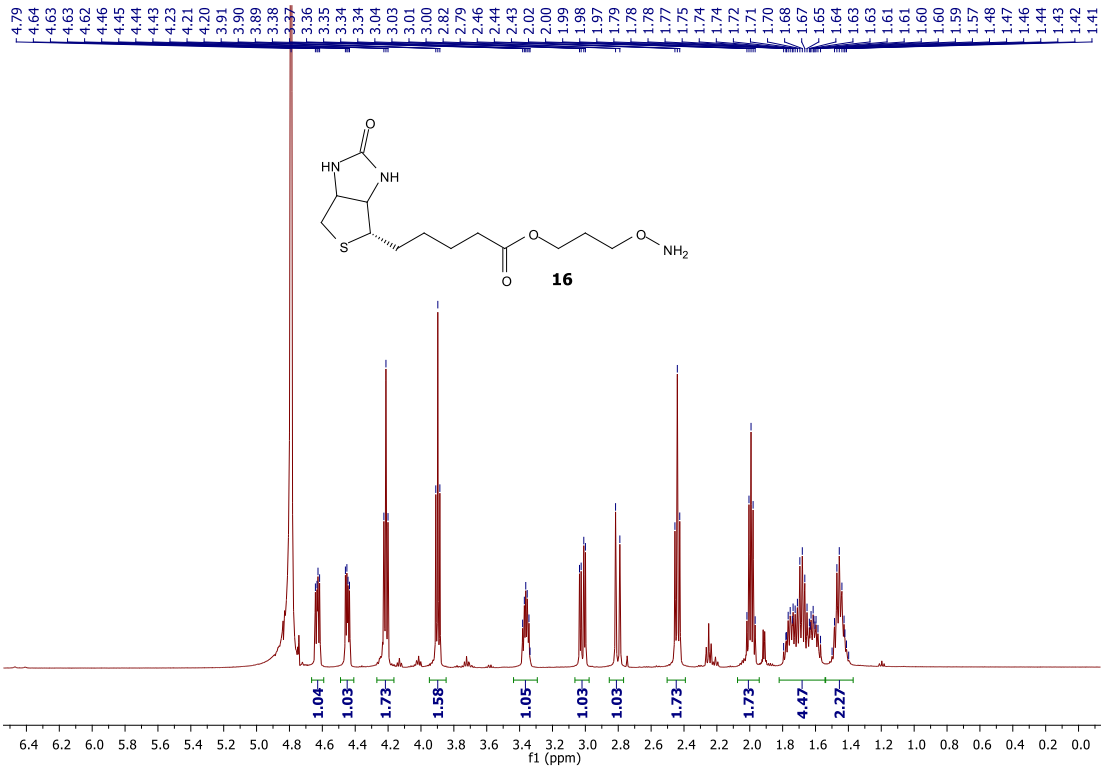
^{1}H NMR



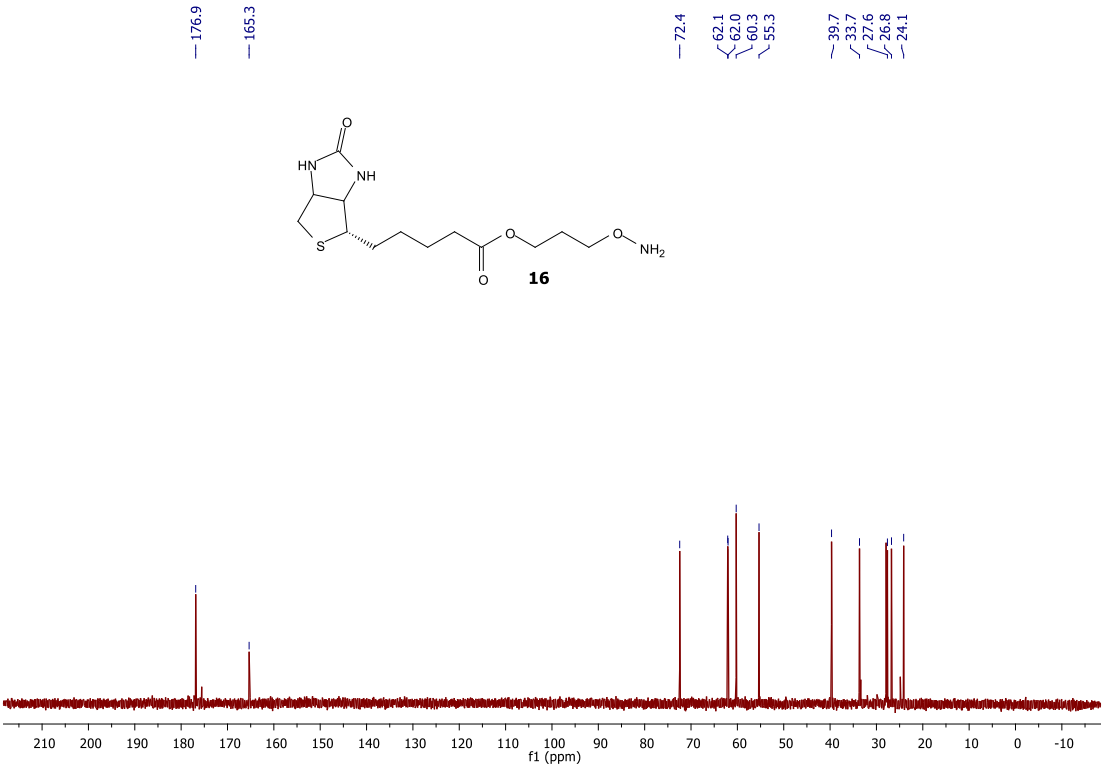
^{13}C NMR



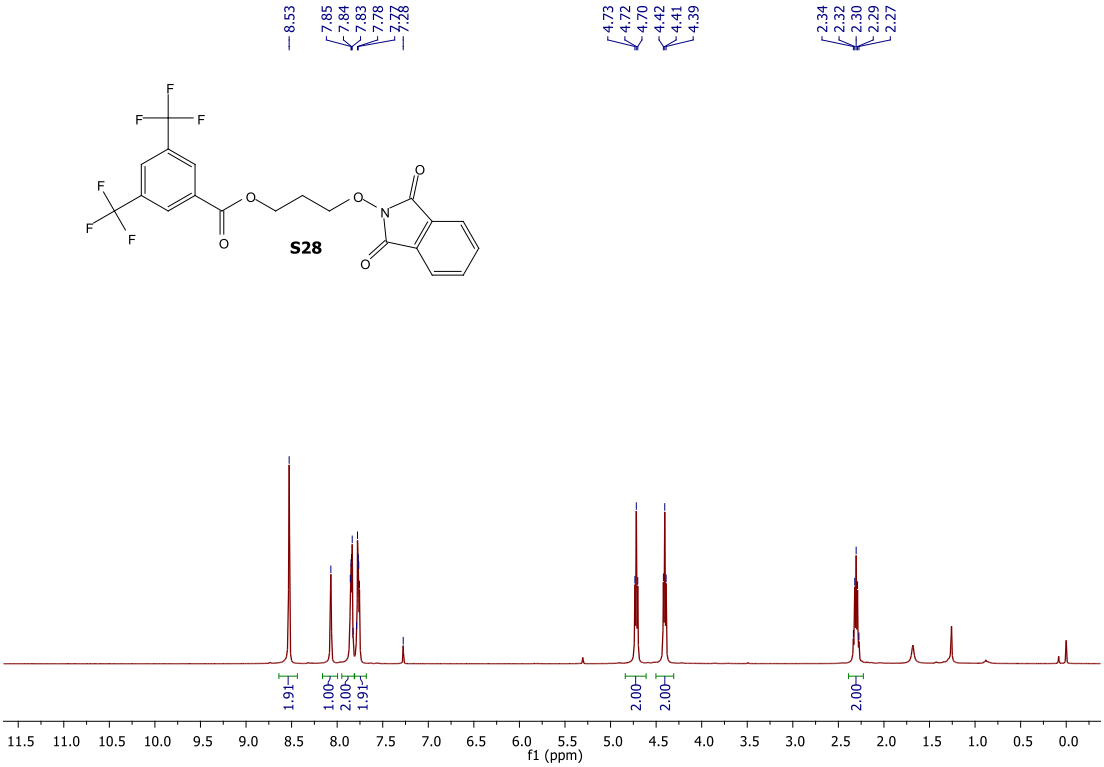
^{1}H NMR



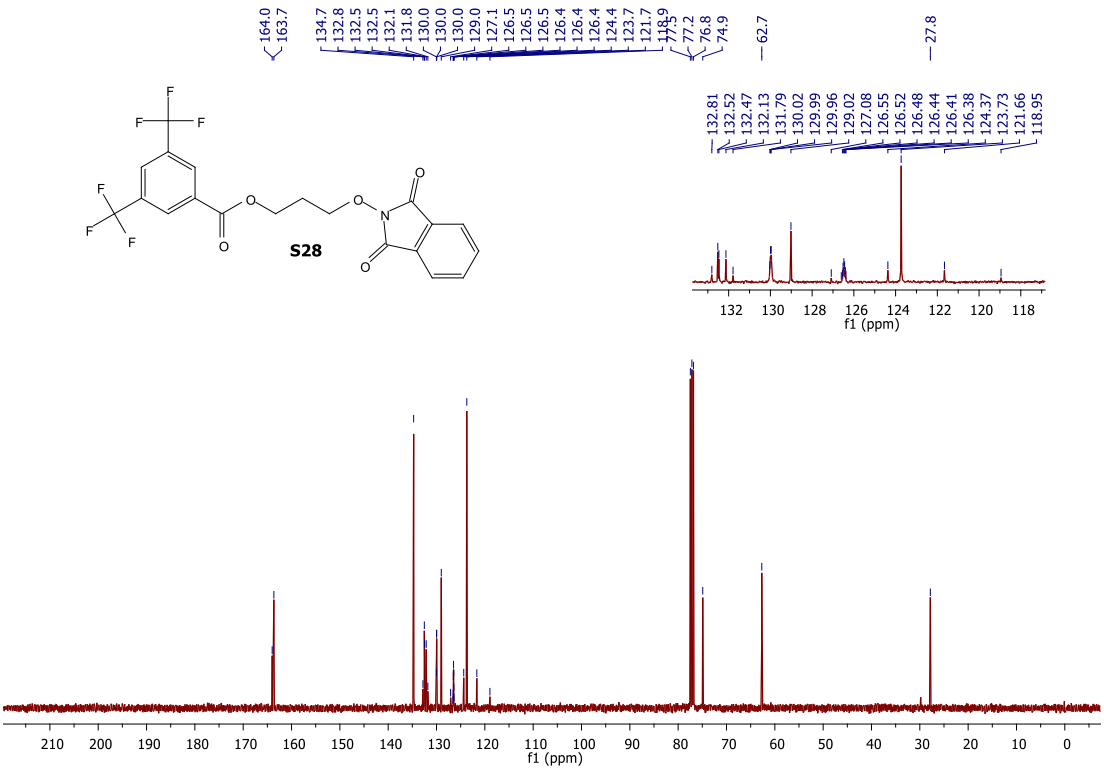
^{13}C NMR



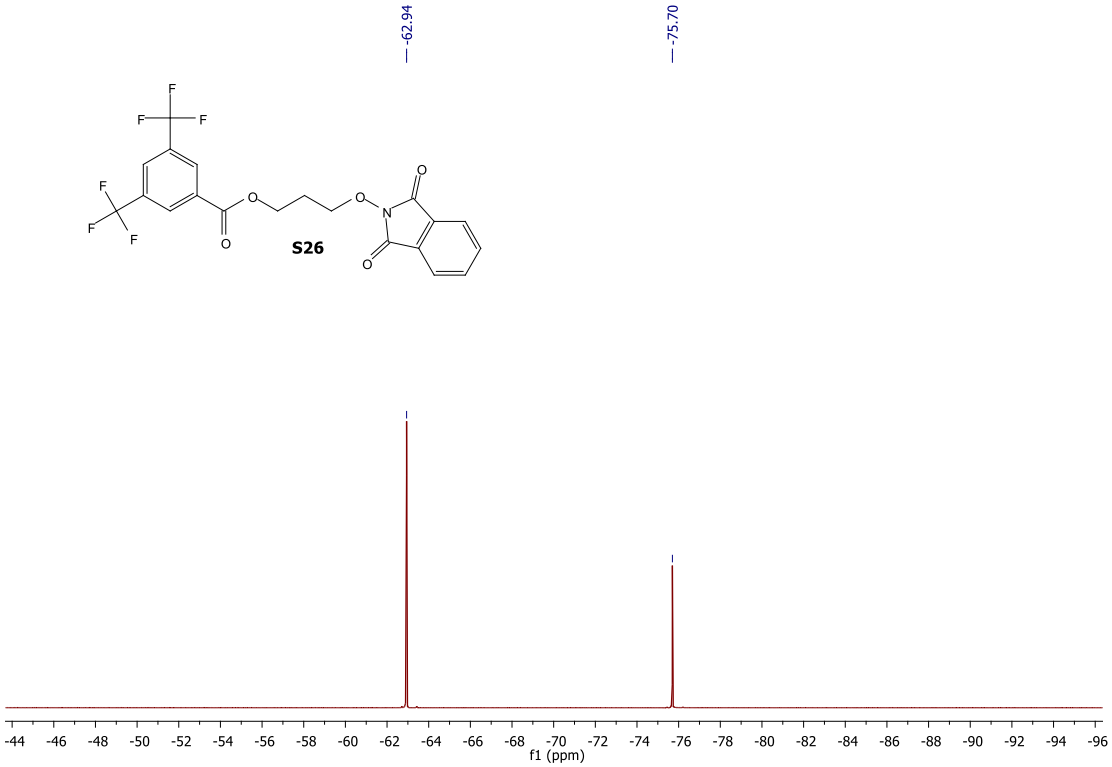
^{1}H NMR



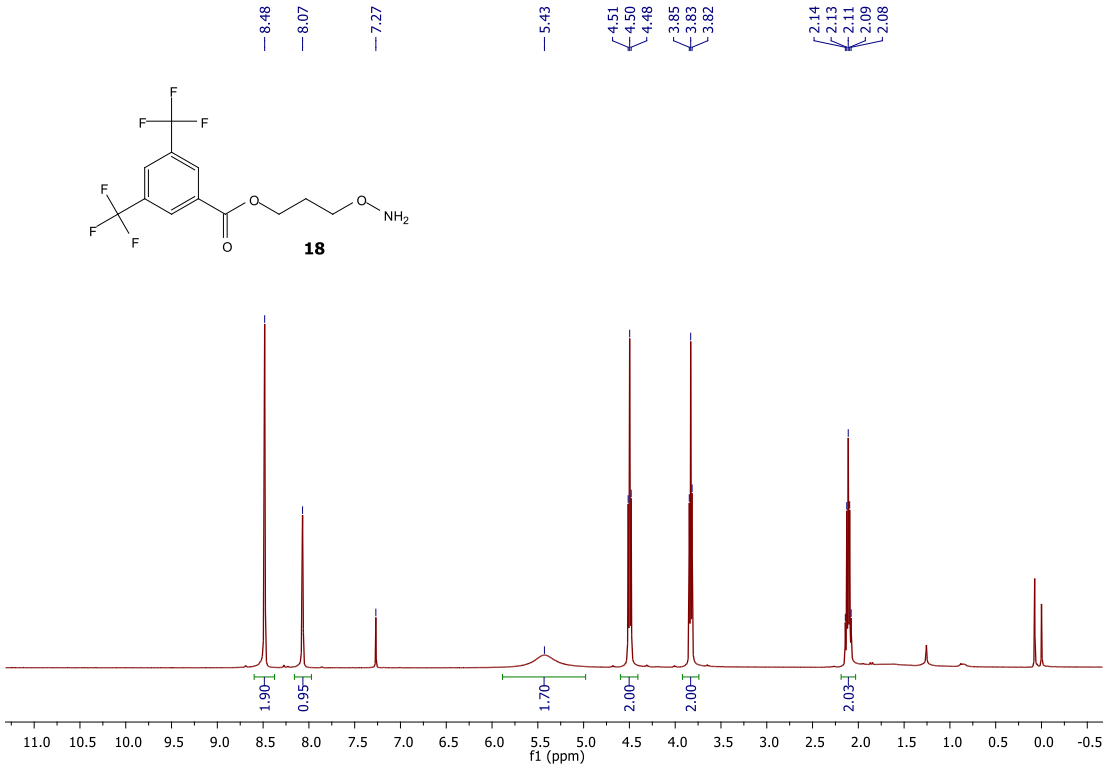
^{13}C NMR



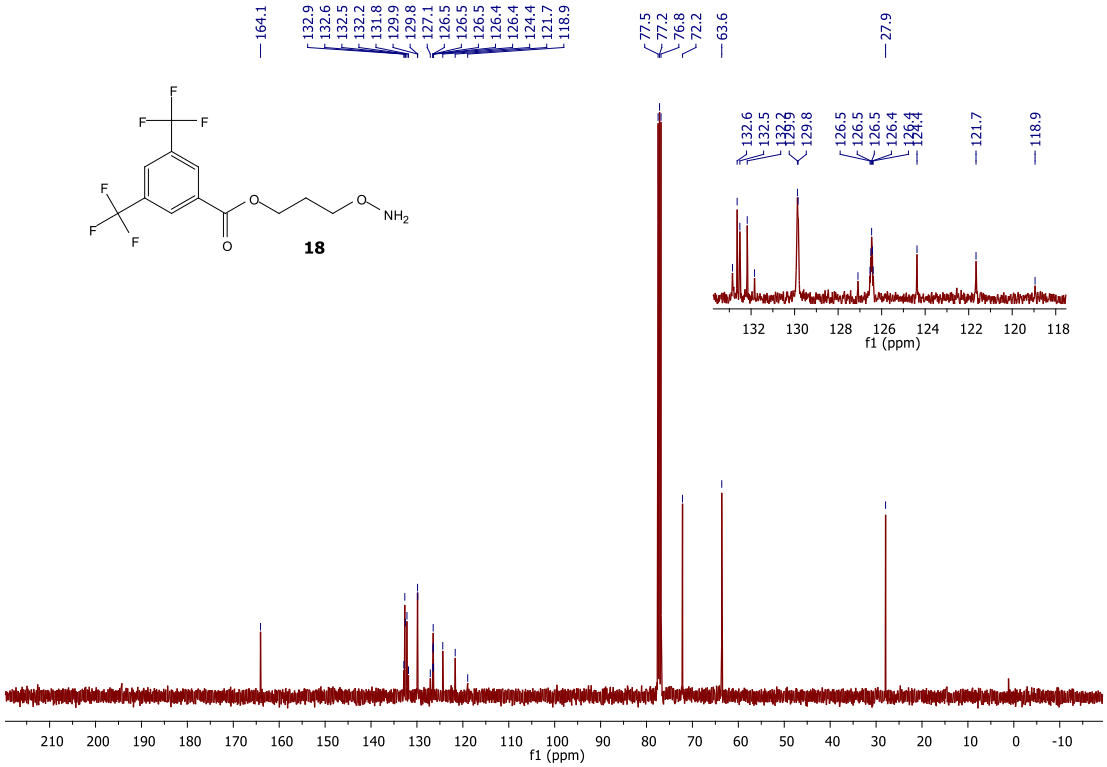
^{19}F NMR



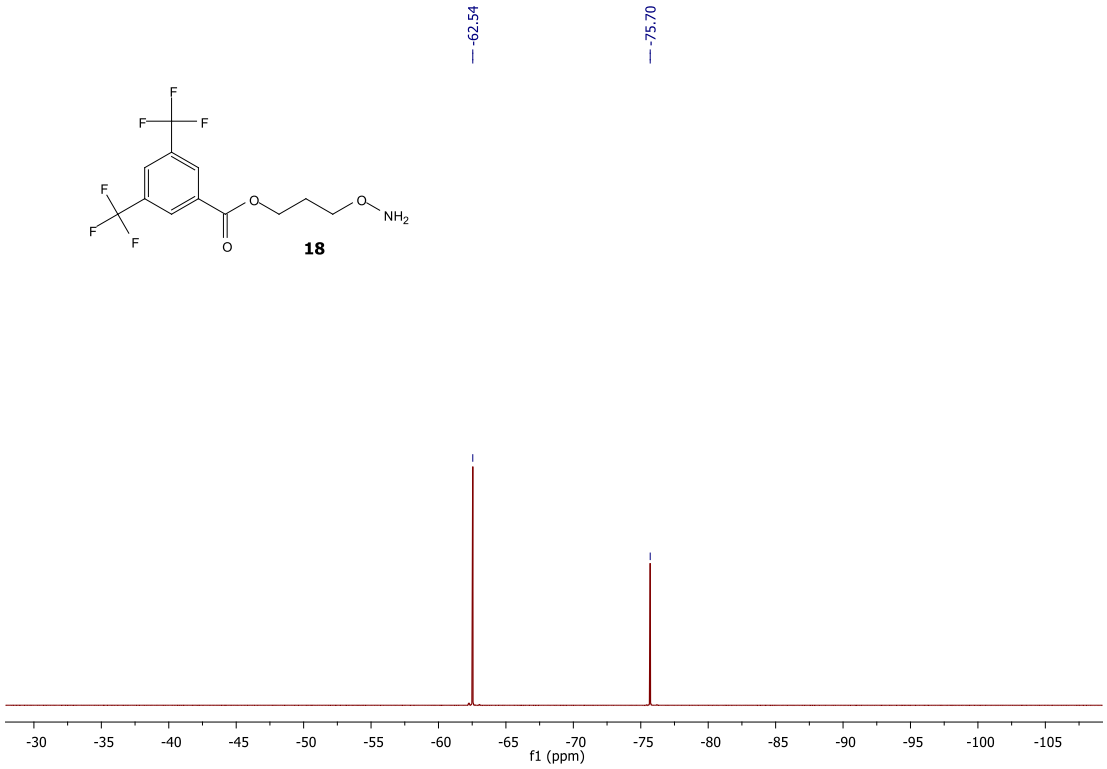
^{1}H NMR



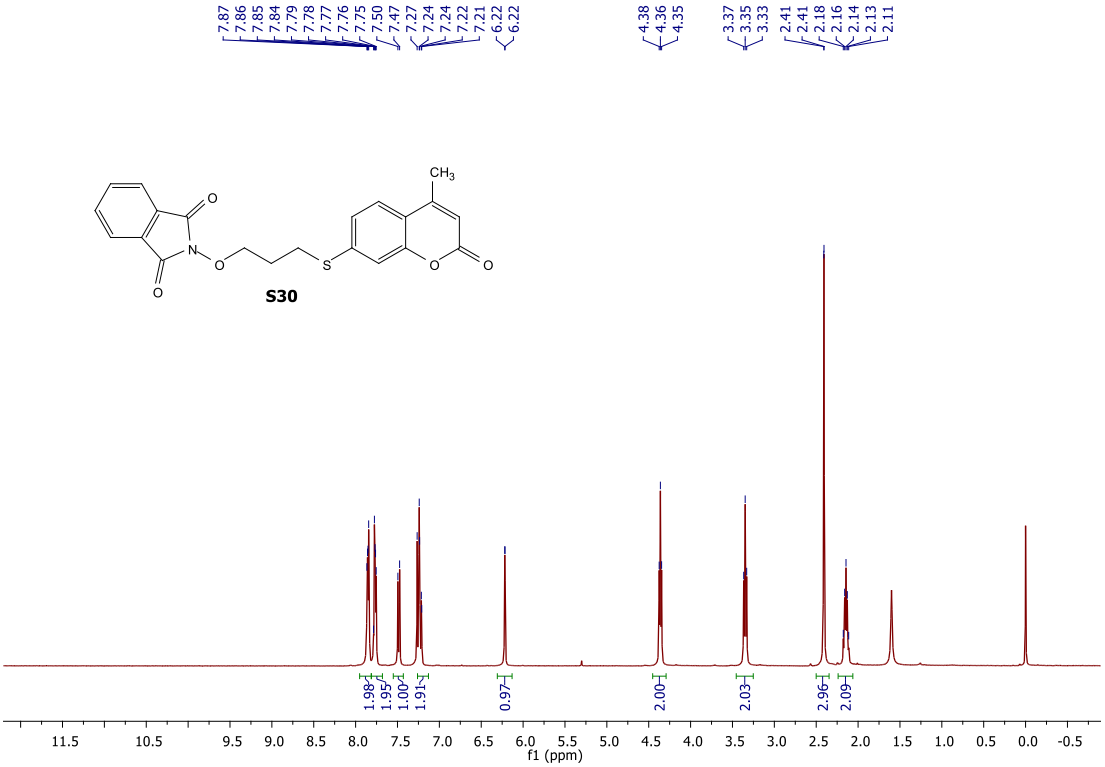
^{13}C NMR



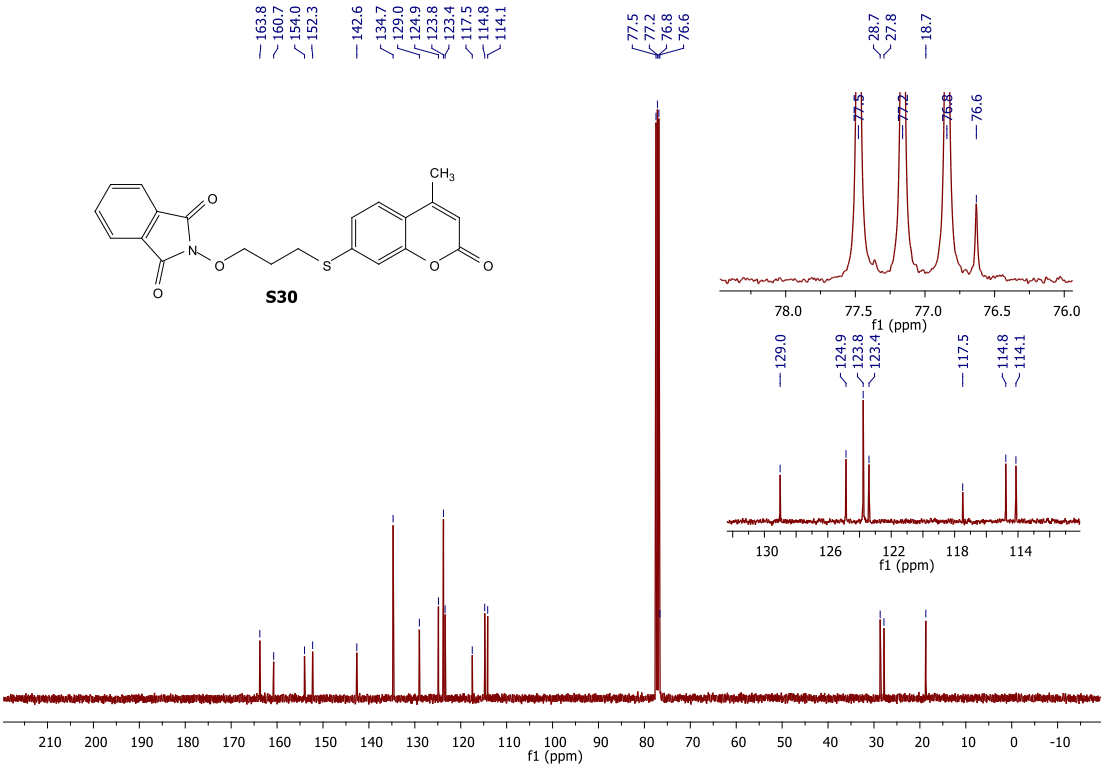
^{19}F NMR



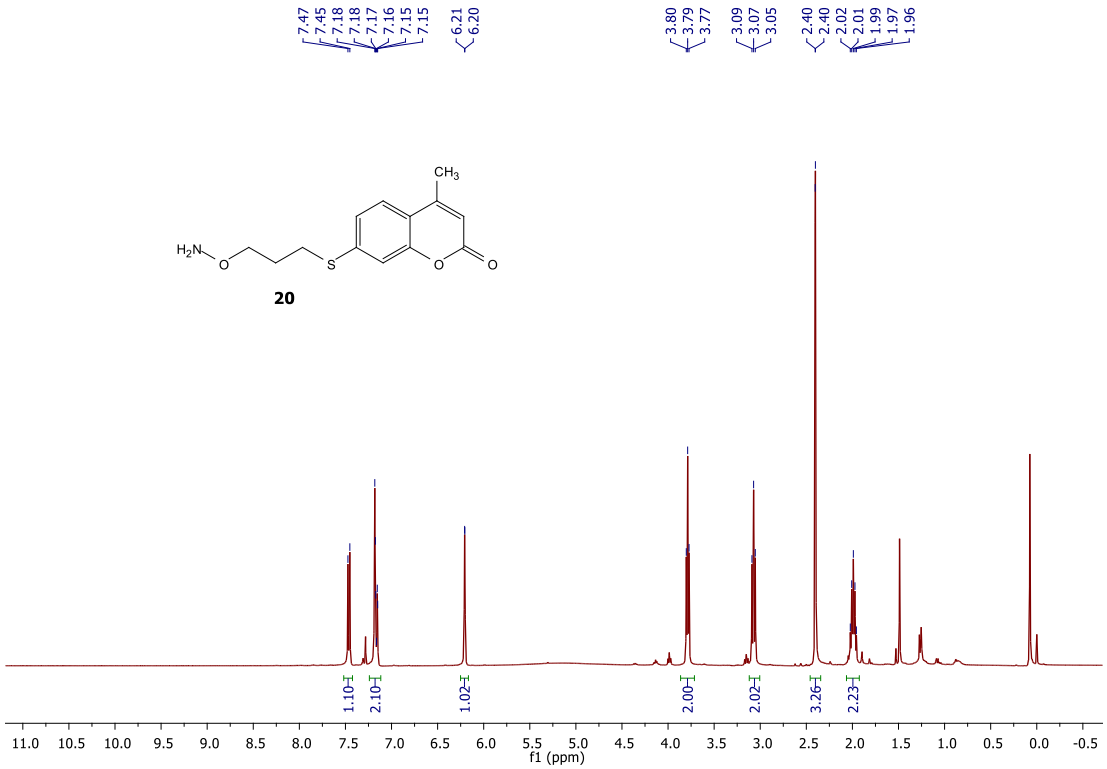
^{1}H NMR



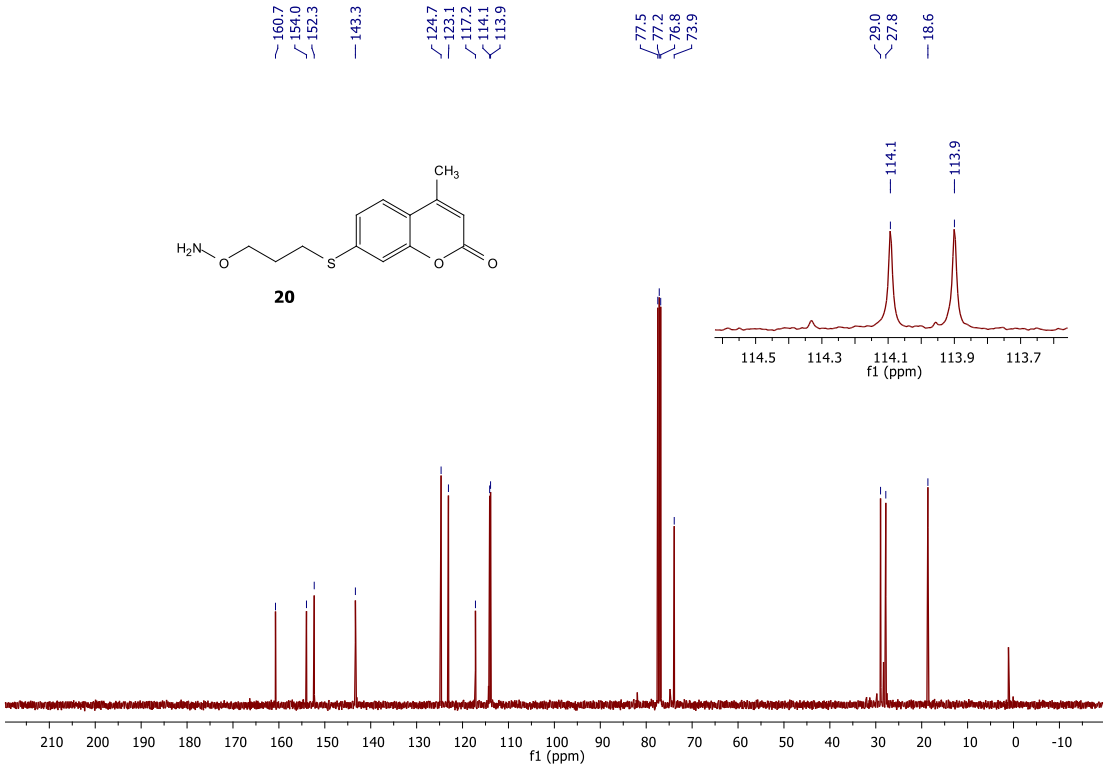
^{13}C NMR



^{1}H NMR



^{13}C NMR



8. References

[1] M. Kinter and N. E. Sherman, Protein sequencing and identification using tandem mass spectrometry. *Wiley Interscience*, **2000**.
[2] A. Bhattacharjya, P. Klumphu and B. H. Lipshutz, *Org. Lett.*, 2015, **17**, 1122-1125.
[3] L. Candish, E. A. Standley, A. Gomez-Suarez, S. Mukherjee and F. Glorius, *Chem.–Eur. J.*, 2016, **22**, 9971-9974.
[4] I. D. Kostas, A. G. Coutsolelos, G. Charalambidis and A. Skondra, *Tetrahedron Lett.*, 2007, **48**, 6688-6691.
[5] X. L. Liu, H. J. Chen, Y. Y. Yang, Y. Wu and J. You, *Eur. J. Org. Chem.*, 2014, 3451-3459.
[6] P. M. S. D. Cal, R. F. M. Frade, V. Chudasama, C. Cordeiro, S. Caddick and P. M. P. Gois, *Chem. Commun.*, 2014, **50**, 5261-5263.
[7] T. Okayama, A. Burritt and V. J. Hruby, *Org. Lett.*, 2000, **2**, 1787-1790.
[8] T. E. Pennington, C. Kardiman and C. A. Huttona, *Tetrahedron Lett.*, 2004, **45**, 6657-6660.
[9] M. G. Thomas, C Lawson, N. M. Allanson, B. W. Leslie, J. R. Bottomley, A. Mcbride and O. A. Olusanya, *Bioorg. Med. Chem. Lett*., 2003, **13**, 423-426.
[10] D. Sougata, N. Dey and S. Bhattacharya, *Chem. Commun.*, 2017, **53**, 2371-2374.
[11] L. Huang, A. Shi, F. He and X. Li, *Bioorg. Med. Chem.*, 2010, **18**, 1244-1251.
[12] F. Ito, S. Ando, M. Iuchi, T. Ukari, M. Takasaki and K. Yamaguchi, *Tetrahedron* 2011, **67**, 8009-8013.
[13] Link for the RNase A enzymatic assay by using the Ribonucleic acid. http://www.sigmaaldrich.com/technical-documents/protocols/biology/enzymatic-assay-of ribonuclease-a.html