Supplementary Information

Stereo- and regio-defined DNA-encoded chemical libraries enable efficient ligand discovery for conditional CAR-T cell activation and for tumor targeting

Nicholas Favalli[1], Gabriele Bassi[1], Christian Pellegrino[1], Jacopo Millul[2], Roberto De Luca[2], Samuele Cazzamalli[2], Su Yang[3], Anika Trenner[4], Nour Mozzafari[4], Renier Myburgh[5], Mustafa Moroglu[6], Stuart J. Conway[6], Alessandro A. Sartori[4], Markus G. Manz[5], Richard A. Lerner[7], Peter K. Vogt[3], Jörg Scheuermann*[1] and Dario Neri*[1]

AFFILIATIONS

1 Department of Chemistry and Applied Biosciences, Swiss Federal Institute of Technology (ETH Zürich), Zürich, Switzerland. 2 Philochem AG, Otelfingen, Switzerland. 3 Scripps Research Institute, Department of Molecular Medicine, La Jolla, California 92037, USA. 4 Institute of Molecular Cancer Research, University of Zürich, Zürich, Switzerland. 5 Department of Medical Oncology and Hematology, University Hospital Zurich and University of Zurich, Comprehensive Cancer Center Zurich (CCCZ), 8091 Zurich, Switzerland. 6 Department of Chemistry, Chemistry Research Laboratory, University of Oxford, Mansfield Road, Oxford, OX1 3TA, UK. 7 Scripps Research Institute, Department of Chemistry, La Jolla, California 92037, USA.

*) Corresponding authors:

D.N.: phone: +41-44-6337401; e-mail: dario.neri@pharma.ethz.ch

J.S.: phone: +41-44-6337774; e-mail: joerg.scheuermann@pharma.ethz.ch

Table of contents

1. Abbreviations

AC: average counts (AC = total counts / library size);
Boc = tert-Butyloxycarbonyl;
CAIX: Carbonic Anhydrase IX;
CREBBP: cAMP-response element binding protein;
CuAAC: Copper-Catalyzed Azide-Alkyne cycloaddition;
DCM: dichloromethane;
DIPEA: *N*, *N*'-diisopropylethylamine;
DMA: *N*, *N*'-dimethylacetamide;
DMF: *N*, *N*'-dimethylformamide;
DMSO: dimethyl sulfoxide;
EDC: 1-ethyl-3- (3-dimethylaminopropyl) carbodiimide;
EF = Enrichment factor (counts / AC);
EF: enrichment factor;
ELISA: enzyme-linked immunosorbent assay;
FA: formic acid;
FACS buffer: Fluorescein activated cell sorting buffer;
FITC: Fluorescein isothiocyanate isomer I;
Fmoc: 9-fluorenylmethyloxycarbonyl;
FP: fluorescence polarization;
H1047R-PI3K: H1047R-p110 α mutant of p110α/p85α PI3K;
HATU: (1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate;
HOAt: 1-hydroxy-7-azabenzotriazole;
HSA: Human Serum Albumin;
ImN3 = Imidazole-1-sulphonyl-azide;
NHS: *N*-hydroxysuccinimide;
PAGE: polyacrylamide gel electrophoresis;
PB: protein buffer;
PBS: phosphate buffered saline;
PI3K: Phosphoinositide 3-kinases
Pip : Piperidine;
RP: reverse phase
sNHS: *N*-hydroxysulfosuccinimide sodium salt;
TBTA: tris[(1-benzyl-1*H*-1, 2, 3-triazol-4-yl)methyl]amine;
TBTAX3: 4,4',4''-(((nitrilotris(methylene))tris(1*H*-1,2,3-triazole-4,1-diyl))tris(methylene))tribenzoic acid;
TCEP-HCl: tris(2-carboxyethyl)phosphine hydrochloride;
TEA: triethylamine;
TEAA: triethylammonium acetate;
TFA: trifluoroacetic acid;
TMB: 3,3',5,5'-tetramethylbenzidine
TNC: Tenascin-C;
TPPTS: Triphenylphosphine-3,3',3''-trisulfonic acid trisodium salt;
Tris-HCl: tris(hydroxylmethyl) aminomethane hydrochloride;
uPA: urokinase-type plasminogen activator;
wt-PI3K: *wildtype* p110α/p85α PI3K
X-Phos: 2-dicyclohexylphosphino-2',4',6'-triisopropylbiphenyl;

2. Materials and general methods

Reagents. Unless otherwise noted, all reagents and solvents were purchased from commercial sources (ABCR, ACROS, Apollo scientific, Bachem, Enamine, Fluorochem, Sigma-Aldrich, and TCI) and used under the manufacturer's instructions. Oligonucleotides were purchased from DNA Technology (Denmark) and IBA (Germany). Boronic acids and esters stock solutions were purchased from Apollo scientific. Alkynes stock solutions were purchased from Enamine. Carboxylic acids were purchased from several commercial suppliers including ABCR, ChemBridge, Sigma-Aldrich, TCI Europe, Alfa Aesar, Matrix Scientific, Enamine Store and Acros Organics. Water was purified with a Millipore Milli-Q system (Merck). Ligation buffer, DNA-Ligase and high-fidelity Phusion DNA polymerase were purchased from New England Biolabs. PCR purification and gel extraction kits were purchased from Quiagen. All gel images were captured by a Bio-Rad Chemidoc image system. Wang resins were purchased from Bachem. Fmoc-iodo-phenylalanine derivatives were purchased from ABCR. TMB was purchased from Sigma Aldrich.

Software. Databases are managed by InstantJChem (ChemAxon). Selection Fingerprints are evaluated by MATLAB R2019b (mathworks). FP and ELISA data was statistically evaluated using PRISM 8 software. NMR were evaluated using MestreNova 7 software.

Purification methods. Small organic molecules were purified by RP-chromatography (BUCHI) on a C18 40 µM irregular (12 g) column with mQ millipore water 0.1 % FA (buffer A) and Acetonitrile 0,1% FA (buffer B) as mobile phase. Gradient (% of buffer B): 2% for 10 mins., 2% → 100% in 30 mins, 100% for 10 mins. FITC-labelled compounds were purified by semi-preparative HT-RP-HPLC (waters) on Synergi 4µm polar-RP 80Å (150x10 mm) column with mQ millipore water 0.1 % FA (buffer A) and Acetonitrile 0,1% FA (buffer B) as mobile phase. Gradient (% of buffer B): 5% for 2 mins., 5% → 70% in 20 mins, 100% for 2 mins. Oligonucleotide derivatives were purified by semi-preparative HT-RP-HPLC (waters) on Waters XTerra® Shield RP18 (125 Å, 5 µm) column with 0.1 M TEAA pH=7 in mQ millipore water (buffer A) and 0.1 M TEAA pH=7 in mQ millipore water : Acetonitrile = 8 : 1 (buffer B) as mobile phase. Gradient (% of buffer B): 10% for 1 mins., 10% → 20% in 4 mins, 20% → 70% in 10 mins, 100% for 2 mins. Analytical LC traces were registered using Xevo G2-XS QTof

Quadrupole Time of Flight Mass Spectrometer (Waters) LC-MS. ^{1}H and ^{13}C-nuclear magnetic resonance (NMR) spectra were recorded at 298 K on a Bruker 400 MHz, 500 MHz or 600 MHz spectrometer. In case of substantial residual water the water signal was suppressed using presaturation. Chemical shifts are given in parts-per-million (ppm) using residual solvent as the internal standard. Coupling constants (J) are reported in hertz (Hz) and multiplicities are classified by the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quint = quintet, sext = sextet, sept = septet, m = multiplet or unresolved, br. = broad signal.

3. Synthesis of *iodo*-phenyl azido propionic acid scaffolds

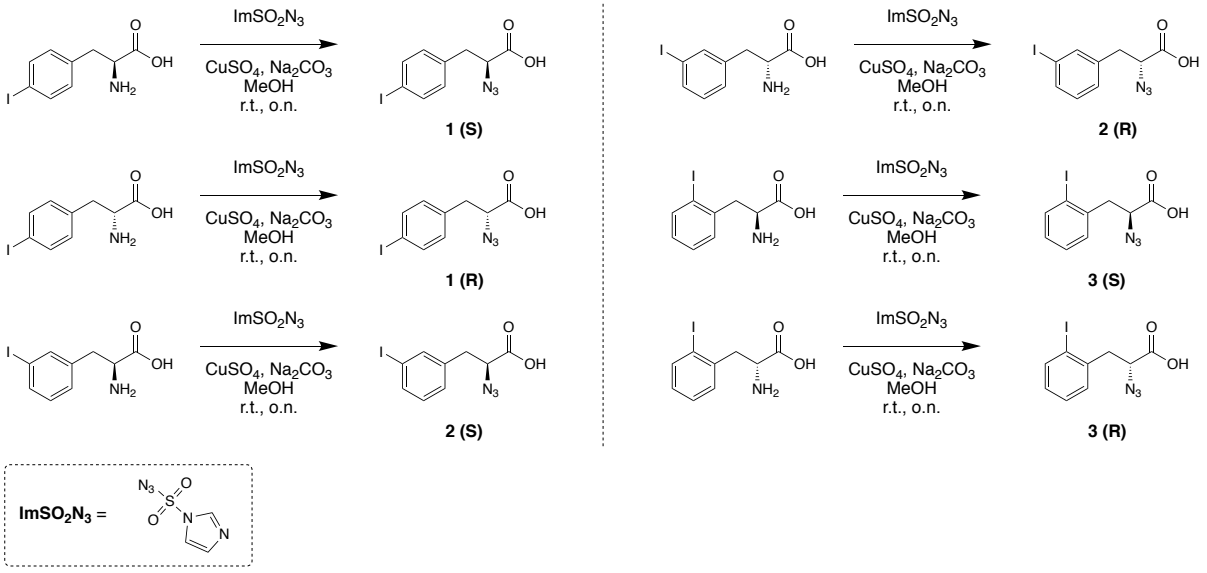


Figure 1: Scheme of synthesis of scaffolds **1 (*S*)**, **1 (*R*)**, **2 (*S*)**, **2 (*R*)**, **3 (*S*)** and **3 (*R*)**.

Synthesis of (*S*)-2-azido-3-(4-iodophenyl)propanoic acid [1 (*S*)], (*R*)-2-azido-3-(4-iodophenyl)propanoic acid [1 (*R*)], (S)-2-azido-3-(3-iodophenyl)propanoic acid [2 (*S*)], (*R*)-2-azido-3-(3-iodophenyl)propanoic acid [2 (*R*)], (*S*)-2-azido-3-(2-iodophenyl)propanoic acid [3 (*S*)] and (*R*)-2-azido-3-(2-iodophenyl)propanoic acid [3 (*R*)].

The commercially available 4-iodo-L-phenylalanine, 4-iodo-D-phenylalanine, 3-iodo-L-phenylalanine, 3-iodo-D-phenylalanine, 2-iodo-L-phenylalanine and 2-iodo-D-phenylalanine (1g each, 3.4 mmol) were dissolved in dry methanol (10 mL). Imidazole-1-sulphonyl-azide hydrochloride (**$ImSO_2N_3$**, 850 mg, 4 mmol), anhydrous potassium carbonate (1.17 g, 8.5 mmol) and anhydrous copper sulphate (25 mg, 0.013 mmol) were added to each reaction and

the resulting mixtures were stirred at room temperature for 24 hours. The reactions were filtered, concentrated under reduced pressure and the products were extracted with ethyl acetate. The pure compounds **1 (*S*)**, **1 (*R*)**, **2 (*S*)**, **2 (*R*)**, **3 (*S*)** and **3 (*R*)** were obtained by RP-chromatography (C18 40 µM irregular, 12 g) . **1 (*S*):** yield = 74% (800 mg, 2.5 mmol). **^{1}H NMR** (400 MHz, DMSO-d6) δ 13.49 (s, 1H), 7.71 – 7.63 (m, 2H), 7.15 – 7.05 (m, 2H), 4.42 (dd, *J* = 8.7, 5.0 Hz, 1H), 3.06 (dd, *J* = 14.2, 5.0 Hz, 1H), 2.89 (dd, *J* = 14.2, 8.7 Hz, 1H). **^{13}C NMR** (101 MHz, DMSO) δ 171.52, 137.51, 137.04, 132.12, 93.18, 62.40, 36.55. **1 (*R*):** yield = 76% (830 mg, 2.6 mmol). **^{1}H NMR** (400 MHz, DMSO-d6) δ 13.47 (s, 1H), 7.71 – 7.63 (m, 2H), 7.13 – 7.04 (m, 2H), 4.42 (dd, *J* = 8.7, 5.0 Hz, 1H), 3.06 (dd, *J* = 14.1, 5.0 Hz, 1H), 2.89 (dd, *J* = 14.2, 8.7 Hz, 1H). **^{13}C NMR** (101 MHz, DMSO) δ 171.51, 137.51, 137.02, 132.12, 93.19, 62.39, 36.55. **2 (*S*):** yield = 63% (678 mg, 2.1 mmol). **^{1}H NMR** (400 MHz, DMSO-d6) δ 7.69 – 7.57 (m, 2H), 7.29 (dt, *J* = 7.6, 1.3 Hz, 1H), 7.12 (t, *J* = 7.7 Hz, 1H), 4.44 (dd, *J* = 8.7, 5.0 Hz, 1H), 3.07 (dd, *J* = 14.2, 5.0 Hz, 1H), 2.88 (dd, *J* = 14.2, 8.7 Hz, 1H), -4.20 (s, 0H). **^{13}C NMR** (101 MHz, DMSO) δ 171.51, 138.21, 135.97, 130.92, 129.21, 95.25, 62.43, 36.46. **2 (*R*):** yield = 85% (915 mg, 2.9 mmol). **^{1}H NMR** (400 MHz, DMSO-d6) δ 7.69 – 7.58 (m, 2H), 7.30 (dt, *J* = 7.7, 1.3 Hz, 1H), 7.13 (t, *J* = 7.7 Hz, 1H), 4.44 (dd, *J* = 8.7, 5.0 Hz, 1H), 3.07 (dd, *J* = 14.2, 5.0 Hz, 1H), 2.89 (dd, *J* = 14.2, 8.7 Hz, 1H). **^{13}C NMR** (101 MHz, DMSO) δ 171.56, 140.03, 138.26, 136.01, 130.97, 129.26, 95.29, 62.47, 36.52. **3 (*S*):** yield = 73% (790 mg, 2.5 mmol). **^{1}H NMR** (400 MHz, DMSO-d6) δ 13.54 (s, 1H), 7.87 (dd, *J* = 7.8, 1.1 Hz, 1H), 7.42 – 7.30 (m, 2H), 7.03 (ddd, *J* = 7.9, 6.7, 2.4 Hz, 1H), 4.35 (dd, *J* = 9.8, 4.8 Hz, 1H), 3.26 (dd, *J* = 14.3, 4.9 Hz, 1H), 3.02 (dd, *J* = 14.3, 9.8 Hz, 1H). **^{13}C NMR** (101 MHz, DMSO) δ 171.41, 139.73, 131.14, 129.50, 128.89, 101.70, 61.77, 41.89. **3 (*R*):** yield = 66% (712 mg, 2.2 mmol). **^{1}H NMR** (400 MHz, DMSO-d6) δ 13.53 (s, 1H), 7.88 (dd, *J* = 7.8, 1.1 Hz, 1H), 7.42 – 7.30 (m, 2H), 7.03 (ddd, *J* = 7.9, 6.7, 2.4 Hz, 1H), 4.34 (dd, *J* = 9.8, 4.8 Hz, 1H), 3.26 (dd, *J* = 14.3, 4.8 Hz, 1H), 3.02 (dd, *J* = 14.3, 9.8 Hz, 1H). **^{13}C NMR** (101 MHz, DMSO) δ 171.40, 139.73, 131.14, 129.50, 128.90, 101.70, 61.77, 41.89.

4. Library synthesis, characterization and purification

4.1 Scaffolds coupling

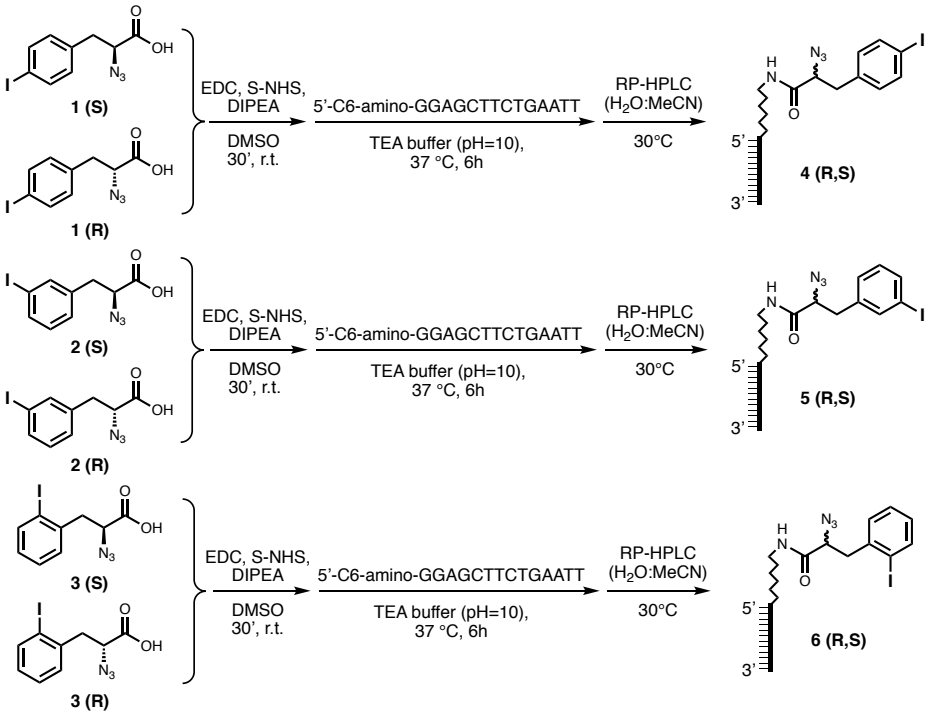


Figure 2: chemical conjugation between *iodo*phenyl azido propionic acid scaffolds and the amino-modified universal 14-mer oligonucleotide. The pair of *S* and *R* enantiomers for each regioisomer (*para, meta* and *ortho* respectively) were mixed in equimolar amount (1:1) and activated for 30 minutes by the EDC/S-NHS/DIPEA method. To the activated scaffolds, a solution of amino-modified oligonucleotide was added. The amide coupling was allowed for 6 hours and the products were precipitated and HPLC purified.

An equimolar mixture for each couple of regio-isomers were prepared: *para*, *meta* and *ortho*-iodo-derivatives [**1(*S*)+1(*R*), 2(*S*)+2(*R*), 3(*S*)+3(*R*)**]. To 100 µmol of each regio-isomer (32 mg, 20 eq. respect to the 5'-aminomodified-DNA) in DMSO (500µL) was added EDC (17 µL, 95 µmol), *sulfo*-NHS (33mg, 150 µmol in 100 µL H_2O) and DIPEA (60µL, 300 µmol). The carboxylic acid activation was put aside for 30 minutes at room temperature. To each activated regio-isomer a solution of **5'-C6-amino-GGAGCTTCTGAATT-3'** (5 µmol in 500 µL of 50 mM TEA/TEA·HCl buffer pH=10) were added and the coupling reactions were heated at 37°C for 6 hours. The reaction was quenched by adding 320 µL of 3M acetate buffer (AcOH/AcONa pH=4.7) and the oligonucleotide-conjugates were precipitate by ethanol (4.3 mL, -20°C for 3 hours). The pellets were re-dissolved in 1 mL of 0.1M TEAA buffer (TEA/AcOH, pH=7.5) and the pure products **4 (*R,S*)**, **5 (*R,S*)** and **6 (*R,S*)** were obtained by RP-HPLC purifications. The three oligonucleotide-conjugates were concentrated and precipitated by ethanol. Compounds **4 (*R,S*)**, **5 (*R,S*)** and **6 (*R,S*)** were isolated with an average yield of 60% (3 µmol).

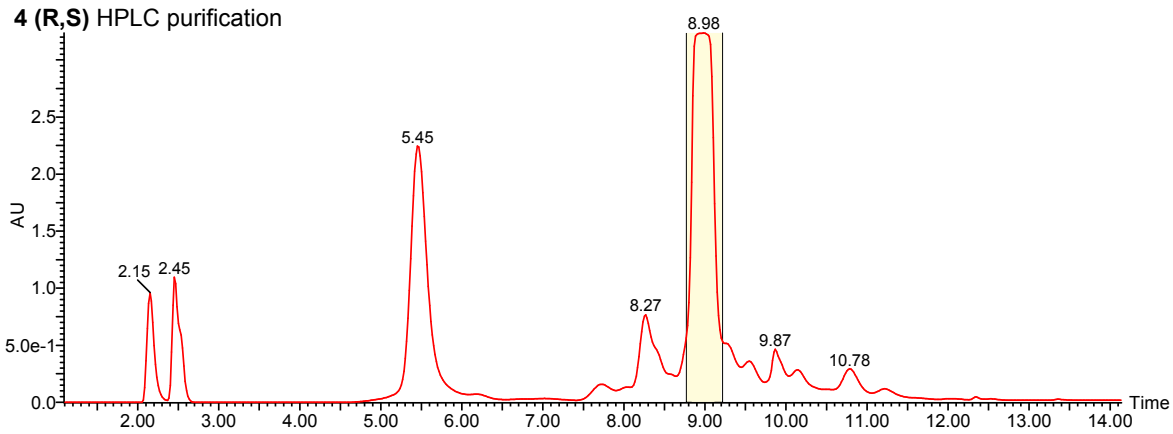


Figure 3: Example of HPLC purification of product **4 (*R,S*)**. HPLC chromatogram registered at λ=260 nm. Peak at 5.45 mins: unreacted starting material (amino-C6-14mer); Peak at 8.98 mins: oligonucleotide-conjugate **4 (*R,S*)**.

4.1.1 LC-MS characterization of 4 (R,S), 5 (R,S) and 6 (R,S)

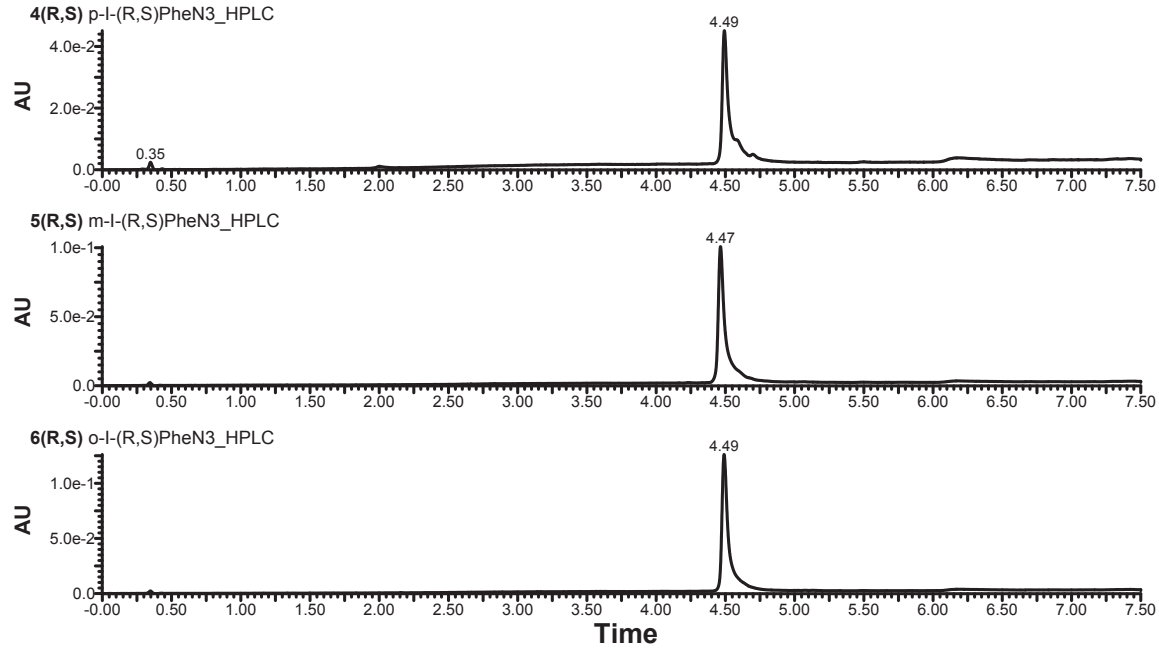


Figure 4: LC-MS chromatogram registered at λ=260 nm of purified compounds **4 (*R,S*)**, **5 (*R,S*)** and **6 (*R,S*)**.

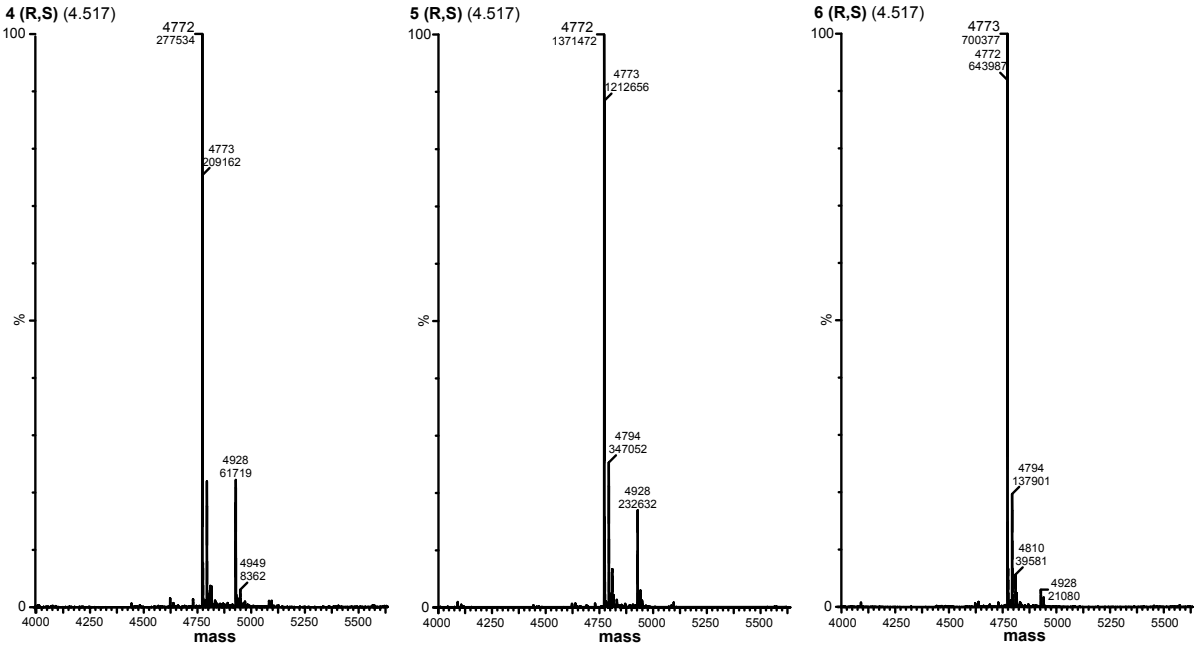


Figure 5: Deconvoluted MS-spectra of products **4 (*R,S*)**, **5 (*R,S*)** and **6 (*R,S*)**.

4.2 Step 1

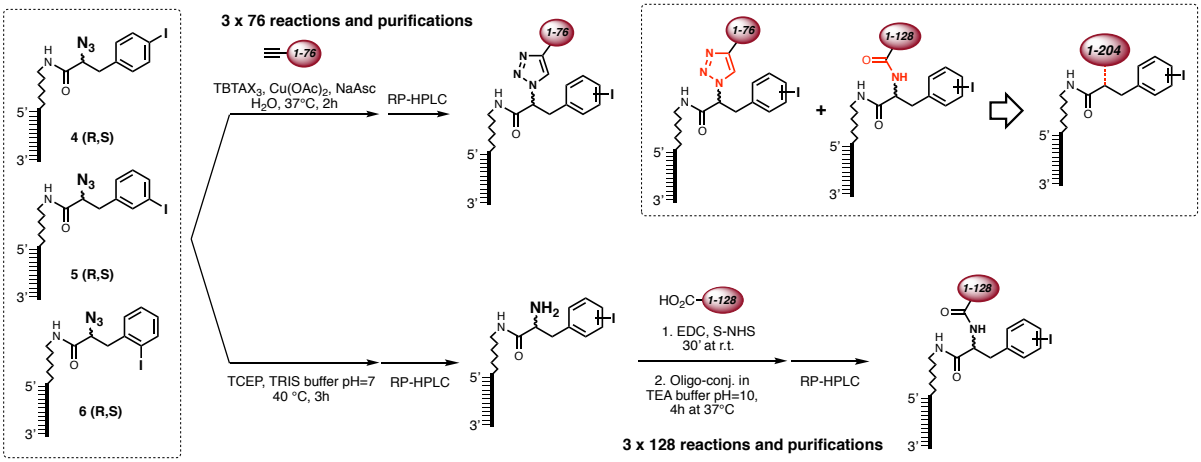


Figure 6: Synthesis of Library step 1. **TBTAX$_3$**: 4,4',4''-(((nitrilotris(methylene))tris(1*H*-1,2,3-triazole-4,1-diyl))tris(methylene))tribenzoic acid; **TCEP**: Tris(2-carboxyethyl)phosphine.

4.2.1 CuAAC reaction (on-DNA)

The reaction conditions were optimized as described in our previous publication[1]. All solvents were degassed in argon atmosphere. 760 nmol of each regio-isomer [**4 (*R,S*)**, **5 (*R,S*)** and **6 (*R,S*)**] were dissolved in mQ millipore water (7.6 mL) and split in 76 (x3) reaction vessels (10 nmol, 100 µL each). The pre-catalyst solution was prepared by mixing 10 mM $Cu(OAc)_2$ (25 µL), 10 mM solution of **TBTAX$_3$** (4,4',4''-(((nitrilotris(methylene))tris(1H-1,2,3-triazole-4,1-diyl))tris(methylene))tribenzoic acid) in 200 mM K_2CO_3 (100 µL) and 2'325 µL of mQ millipore

water, resulting in a 100 µM solution of Cu(II) - $TBTAX_3$ complex[1]. To each reaction vessel the pre-catalyst solution [25 µL, 2.5 nmol of Cu(II)] and 10 mM alkyne solution (**Table 7**) in DMSO (40 µL) were added. The resulting solutions were mixed and the catalyst was activated by adding to each reaction 10 mM sodium-L-ascorbate (40 µL) solution. The reactions were agitated at 35 °C for 3 hours. The reactions were quenched by adding 3M acetate buffer (41µL each reaction) and precipitated with ethanol (740 µL each reaction). The obtained 228 triazole derivatives (76 x 3) were dissolved in TEAA buffer (1 mL each) and individually purified by RP-HPLC.

4.2.2 Staudinger reduction (on-DNA)

1.8 µmol of each regio-isomer [**4 (*R,S*)**, **5 (*R,S*)** and **6 (*R,S*)**] were dissolved in mQ millipore water (0.5 mL) and a 1.0 mL of 200 mM of tris(2-carboxyethyl)phosphine (TCEP) in 500 mM Tris HCl buffer (pH=8.0) was added to each reaction. The reactions were heated at 40°C for 3 hours. The three "free" amino derivatives (*para, meta* and *ortho-iodo*) were purified by RP-HPLC. The purified products were concentrated and precipitated by ethanol (yield = 78%, 1.4 µmol each).

4.2.3 Amide coupling reaction (on-DNA)

Reaction conditions were optimized as described in our previous publication[2]. 50µL of 200 mM carboxylic acid solutions (**Table 7**) were individually activated by adding 500 µL of DMSO, 100 mM EDC in DMSO (90 µL) and 100 mM S-NHS (150 µL) in DMSO:water = 2:1. The activation was left for 30 minutes at room temperature.

1.28 µmol of each regio-isomer [**4 (*R,S*)**, **5 (*R,S*)** and **6 (*R,S*)**] were dissolved in 50 mM TEA buffer pH=10 (12.8 mL) and split in 128 (x3) reaction vessels (10 nmol, 100 µL each). To each vessel the activated carboxylic acid solutions were added (250 µL) and the reactions were kept at 37 °C for 4 hours. The reactions were quenched by adding 3M acetate buffer (70 µL each reaction) and precipitated with ethanol (1.26 mL each reaction). The obtained 384 amides (128 x 3) were dissolved in TEAA buffer (1 mL each) and individually purified by RP-HPLC.

4.2.4 Encoding of step 1

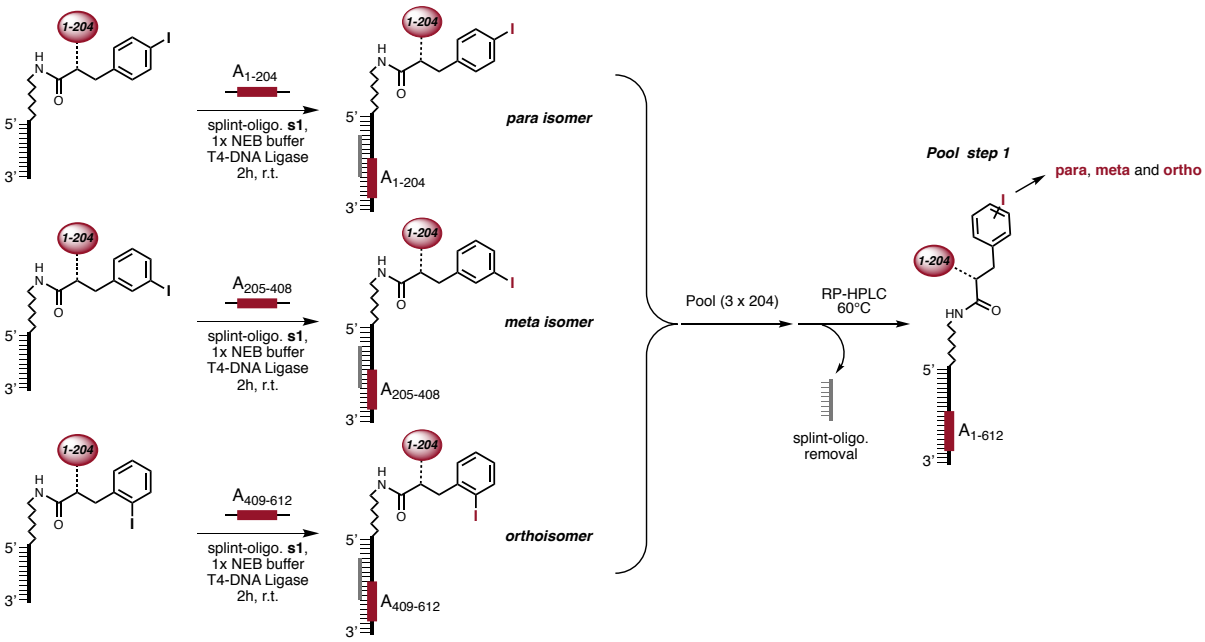


Figure 7: Encoding of Library step 1 by enzymatic splint ligation followed by HPLC purification. **s1**: 3'-CCTCGAAGACTTAAGACACACGAC-5'.

Before the enzymatic ligations, all the purified products were precipitated by ethanol and dried. To the 612 oligonucleotide conjugates (5 nmol each), 500 µM splint-oligonucleotide in mQ millipore water (17 µL each reaction, 8.5 nmol, **s1**: 3'-CCTCGAAGACTTAAGACACACGAC-5'), 200 µM phosphorylated-oligonucleotide codes (37.5 µL, 7.5 nmol, **code A**: 5'-CTGTGTGCTG**XXXXXX**CGAGTCCCATGGCGC-3', x 612 codes) and 6 µL of 10x T4 DNA Ligase Reaction Buffer (500 mM Tris-HCl, 100 mM $MgCl_2$, 100 mM dithiothreitol, 10 mM ATP, pH 7.5) were added. The reactions were heated for 5 minutes at 70 °C, cooled down to room temperature followed by addition of 400 U/mL T4 DNA ligase (1 µL, 0.4 units each reaction). Reactions were kept at room temperature for 2 hours. The enzyme was deactivated by adding 3M acetate buffer pH=4.7 (12 µL) and by heating at 70 °C for 5 minutes. All the crude ligation reactions were analysed by LC-MS (**Figure 8**) before pool. The 612 encoded derivatives were pooled (total volume = 44.7 mL), concentrated, precipitated by ethanol and purified by 60°C RP-HPLC (**Figure 9**).

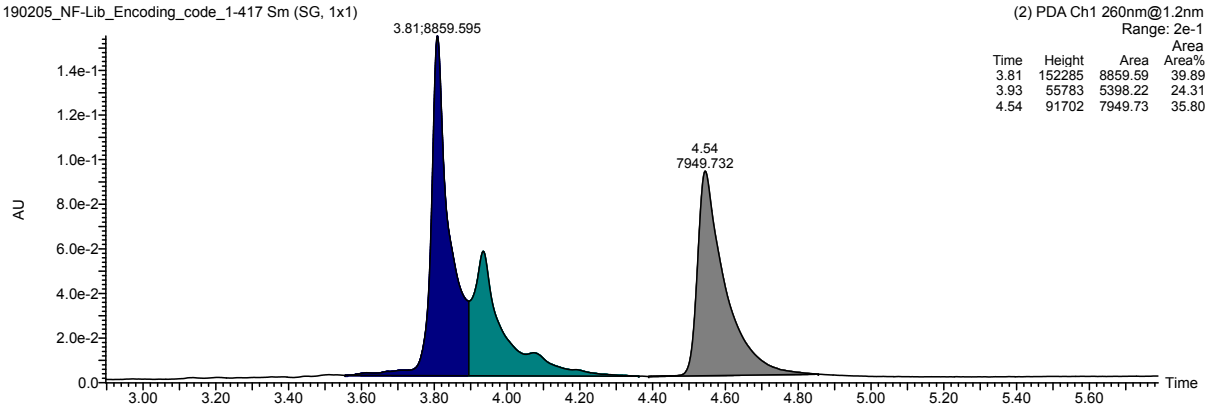


Figure 8: Example of LC-MS analysis of encoding reaction **A417**, chromatogram registered at λ=260 nm. Highlighted in blue, the peak area of the splint oligonucleotide (**s1**), in green the peak area of the unreacted code A417 and in grey the peak area of the encoded product 417. The unreacted 14-mer derivative could not been detected (full conversion).

4.2.5 HPLC purification of step 1

The "**pool step 1**" were re-dissolved in 2mL of 0.1M TEAA buffer (pH=7) and purified by RP-HPLC in the following conditions:

Buffers: Buffer A (0.1M TEAA), Buffer B (MeCN:H_2O=8:1, 0.1M TEAA)

Gradient (% of buffer B)**: 5%** for 1 min., **5%** → **18%** in 14 mins., **18%** → **80%** in 5 mins., **80%** → **100%** in 2 mins., **100%** for 7 mins. Flow = 4.00 mL/min.

Column temperature: 60°C.

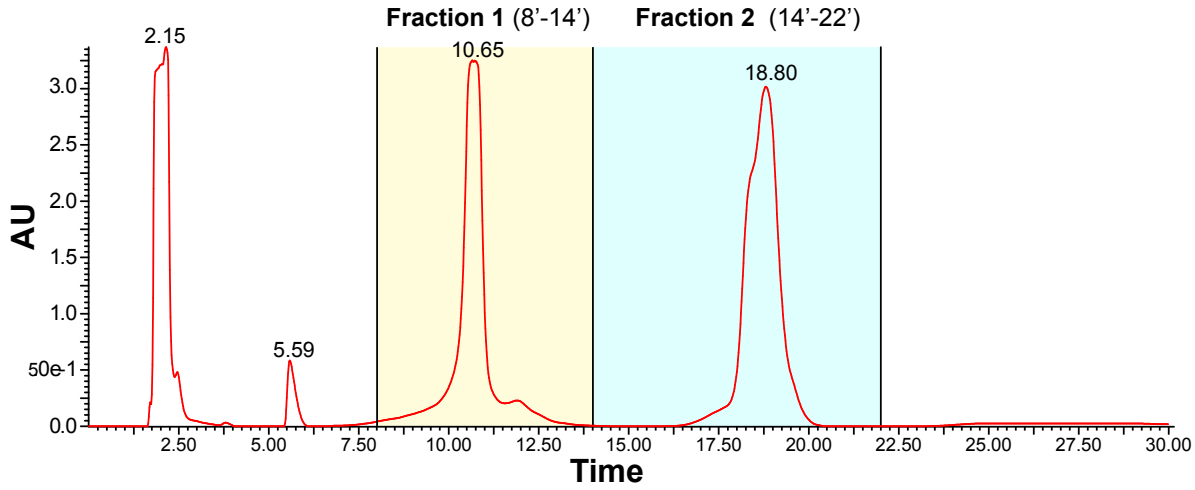


Figure 9: HPLC chromatogram registered at λ=260 nm. Fraction 1 (from 8 to 14 minutes): Splint oligonucleotide (**s1**) and unreacted codes A. Fraction 2 (from 14 to 22 minutes): Encoded step 1 (**pool step 1**).

After HPLC purification all the fractions were analysed by LC-MS. The "**pool step 1**" was isolated with a total yield of 60% (1.8 µmol).

4.2.5 LC-MS characterization of “pool step 1”.

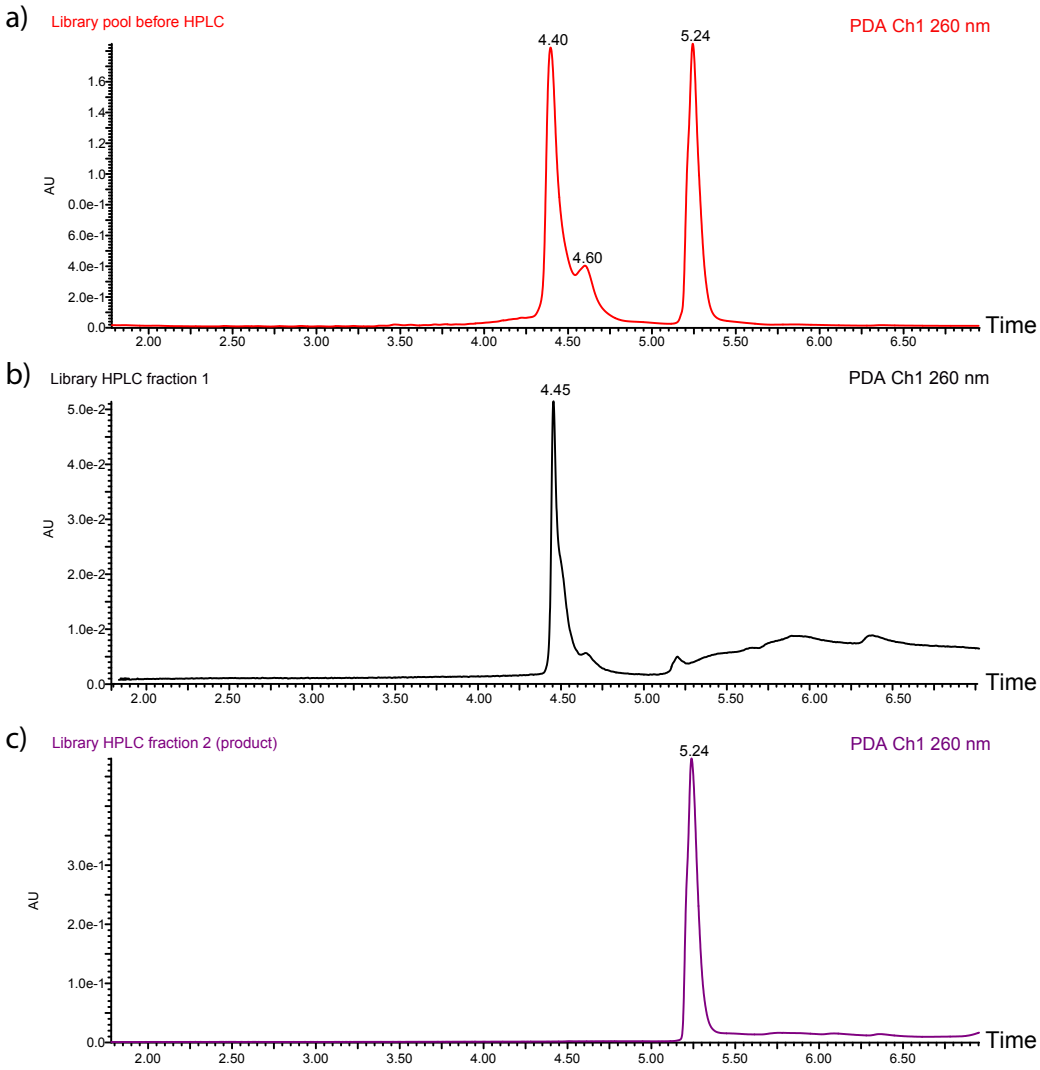


Figure 10: LC-MS chromatograms registered at λ=260 nm of the “**pool step 1**” a) before HPLC purification; b) after HPLC purification (fraction 1); c) after HPLC purification (fraction 2).

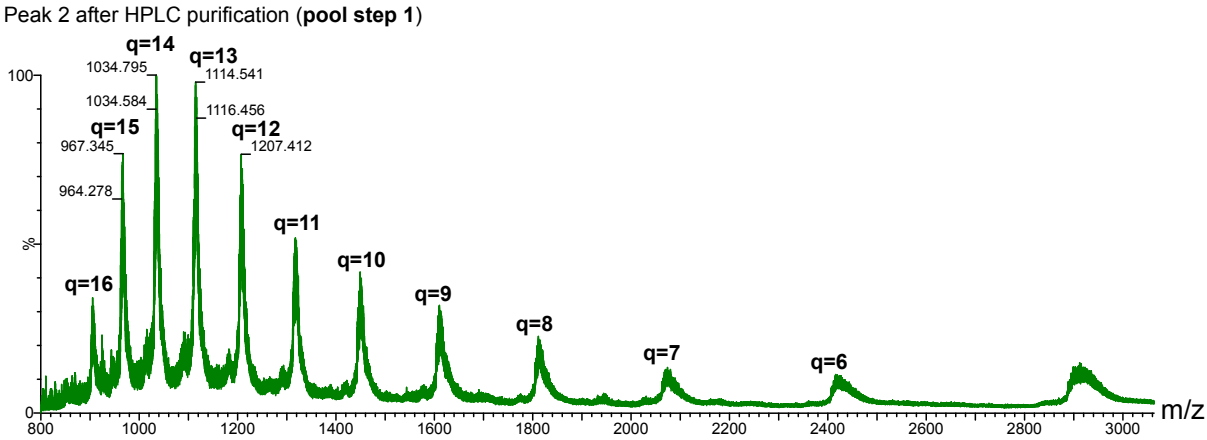


Figure 11: Non-deconvoluted MS-spectra (TOF negative mode) of **pool step 1** after HPLC purification. The average mass of pool is 14.5 KDa, length = 46-mer. **q**: absolute charge.

4.3 Step 2

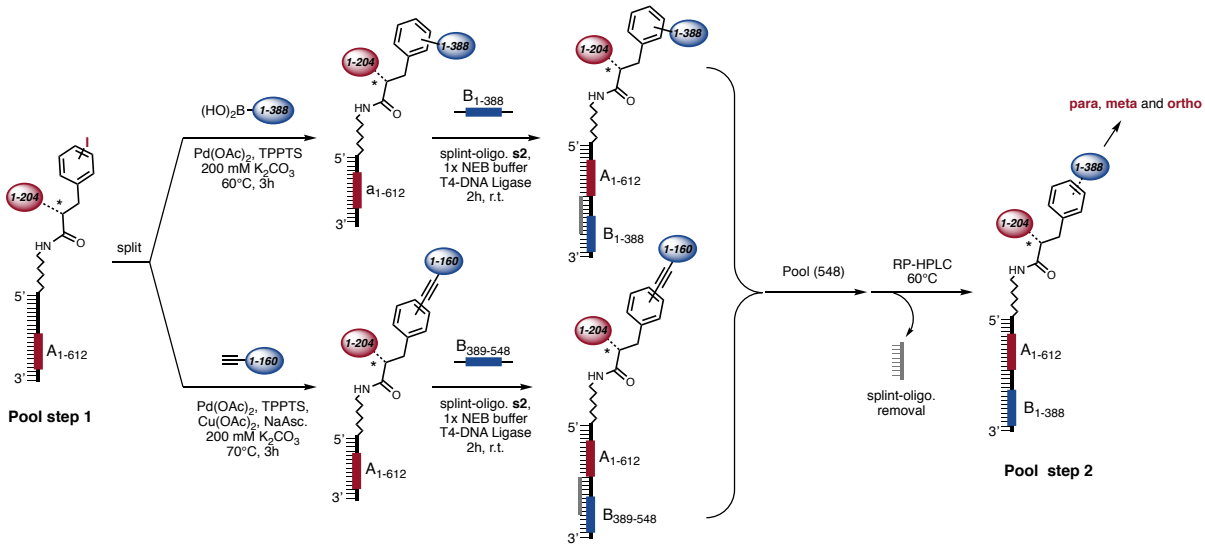


Figure 12: Synthesis and encoding of Library step 2. **s2**: 5'- CGTCGATCCGGCGCCATGG-3'.

1.1 µmol of "**pool step 1**" were dissolved in 200 mM K_2CO_3 (1.1 mL) and split in 548 reaction vessels (2 nmol each, 20 µL) in order to be coupled with 388 boronates (boronic acids and pinacol esters) and with additional 160 alkynes. The second set of building blocks has been chosen after a large screening of boronic acids and alkynes and compounds with conversion higher than 50% and 70% respectively (**Figure 13**) have been included in the library with optimized coupling conditions[1] reported in the next sections.

4.3.1 Protocol of "on-DNA" Suzuki cross coupling

The catalyst solution was prepared by mixing 20 µL of 10 mM palladium (II) acetate in *N,N*-dimethylacetamide (DMA), 100 µL of 100 mM trisodium 3,3',3-phosphine-triyltribenzenesulfonate (TPPTS) in water and 480 µL of water, resulting in a 0.33 mM solution of Pd(0)-TPPTS complex[1]. To each vessel containing 100 µM "**pool step 1**" (2nmol, 20µL) in carbonate were subsequently added the catalyst solution (6 µL of, 2 nmol in Pd) and 200 mM $ArB(OH)_2$ (**Table 8**) in DMA (10 µL). The reactions heated at 60 °C for 3 hours and then quenched by adding 3M acetate buffer (10 µL). The products were precipitated by adding ethanol (140 µL, -20°C).

4.3.2 Protocol of "on-DNA" Sonogashira cross coupling

All solvents were degassed in argon atmosphere. The pre-catalyst solution were prepared by mixing 10mM palladium (II) acetate in DMA (100 µL), 100 mM TPPTS in water (100 µL), 20

mM Copper (II) acetate in water (100 µL) and diluted up to 1 mL with mQ millipore water, resulting in a 1 mM solution of Pd(0)-TPPTS complex and 2 mM solution of Cu(II). To each vessel containing 100 µM "**pool step 1**" (2nmol, 20µL) in 200 mM potassium carbonate, the pre-catalyst solution (4 µL of, 4 nmol in Pd) and 100 mM alkyne (**Table 8**) in DMSO (10 µL) were subsequently added. The copper was reduced by adding a 10 mM solution of sodium L-ascorbate (10 µL) and the resulting solutions were heated at 70 °C for 3 hours. The reactions were quenched by adding 3M acetate buffer (10 µL) and the products were precipitated by adding ethanol (140 µL, -20°C).

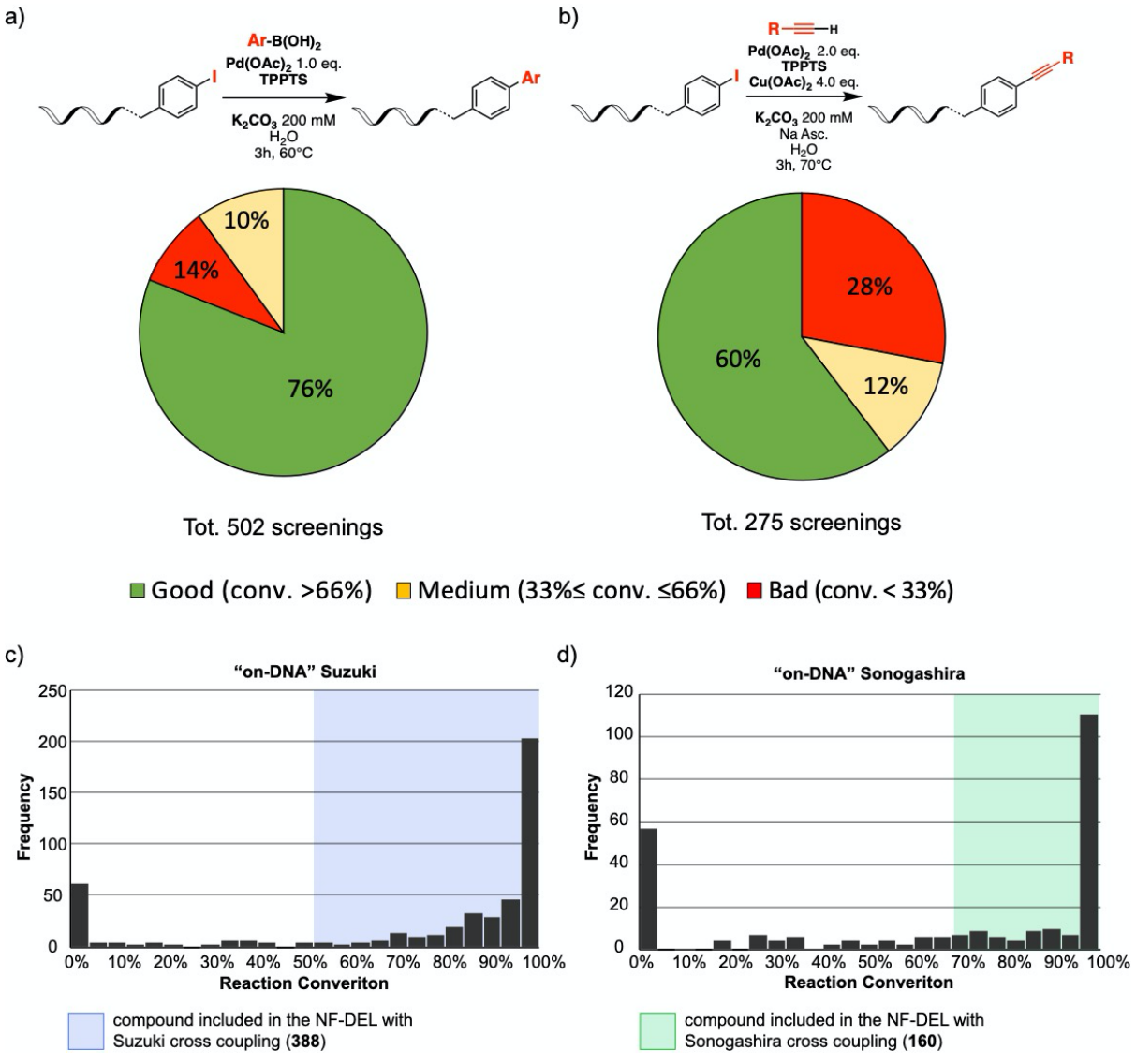


Figure 13: Results of large screening of a) boronates with "on-DNA" Suzuki and b) alkynes with "on-DNA" Sonogashira optimized conditions. Distribution of reaction conversion for c) Sunzuki and d) Sonogashira cross couplings. Only compounds with conversions greater than 50% for Suzuki (tot. 388 boronates) and 70% for Sonogashira (160 alkynes) were included in the library construction.

4.3.3 Encoding of step 2

To the 548 oligonucleotide conjugates (2 nmol each, dry), 500 µM splint-oligonucleotide in mQ millipore water (7 µL each reaction, 3.5 nmol, **s2**: 5'- CGTCGATCCGGCGCCATGG-3'), 200 µM phosphorylated-oligonucleotide codes (15 µL, 3 nmol, **code B**: 5'-GGATCGACG**YYYYYYY**GCGTCAGGCAGC-3'), mQ millipore water (5 µL) and 3 µL of 10x T4 DNA Ligase Reaction Buffer (500 mM Tris-HCl, 100 mM $MgCl_2$, 100 mM dithiothreitol, 10 mM ATP, pH 7.5) were added. The reactions were heated for 5 minutes at 70 °C, cooled down to room temperature followed by addition of 200 U/mL T4 DNA ligase (1 µL, 0.2 units each reaction). The ligation was kept at room temperature for 2 hours. The enzyme was deactivated by adding 3M acetate buffer pH=4.7 (6 µL) and by heating at 70 °C for 5 minutes. All the crude ligation reactions were analysed by LC-MS (**Figure 15**) before pool. The 548 encoded derivatives were pooled (total volume = 20.3 mL), concentrated, precipitated by ethanol and purified by 60°C RP-HPLC (**Figure 14**). The HPLC conditions are described in the section **4.2.5**.

4.3.4 HPLC purification of step 2

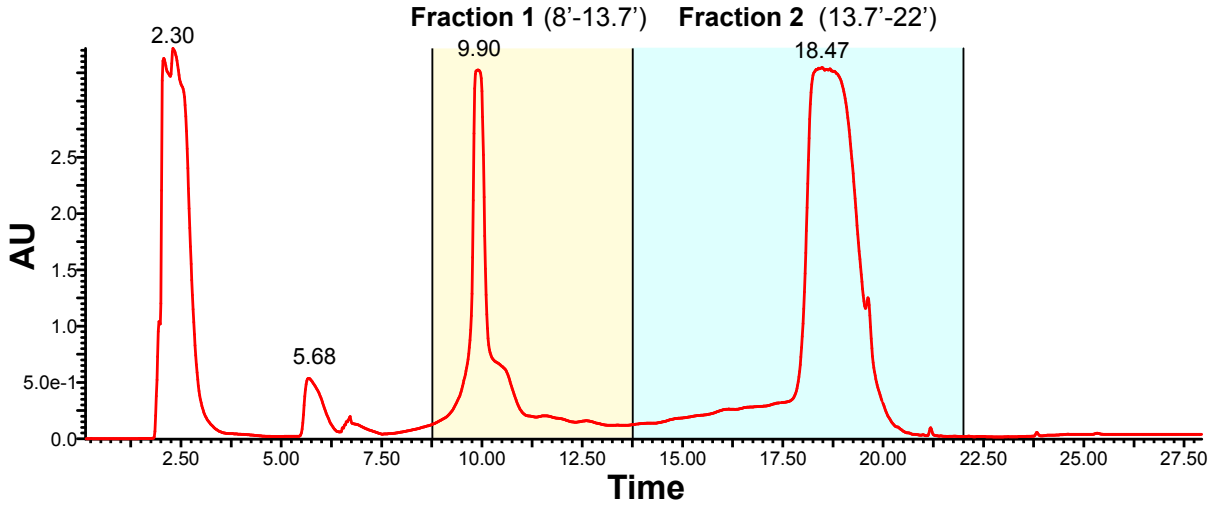


Figure 14: HPLC chromatogram registered at λ=260 nm. Fraction 1 (from 8 to 13.7 minutes): Splint oligonucleotide (**s2**) and unreacted codes B. Fraction 2 (from 13.7 to 22 minutes): Encoded step 2 (**pool step 2**).

After HPLC purification all the fractions were analysed by LC-MS. The "**pool step 2**" was isolated with a total yield of 55% (600 nmol).

4.3.5 LC-MS characterization of “pool step 2”.

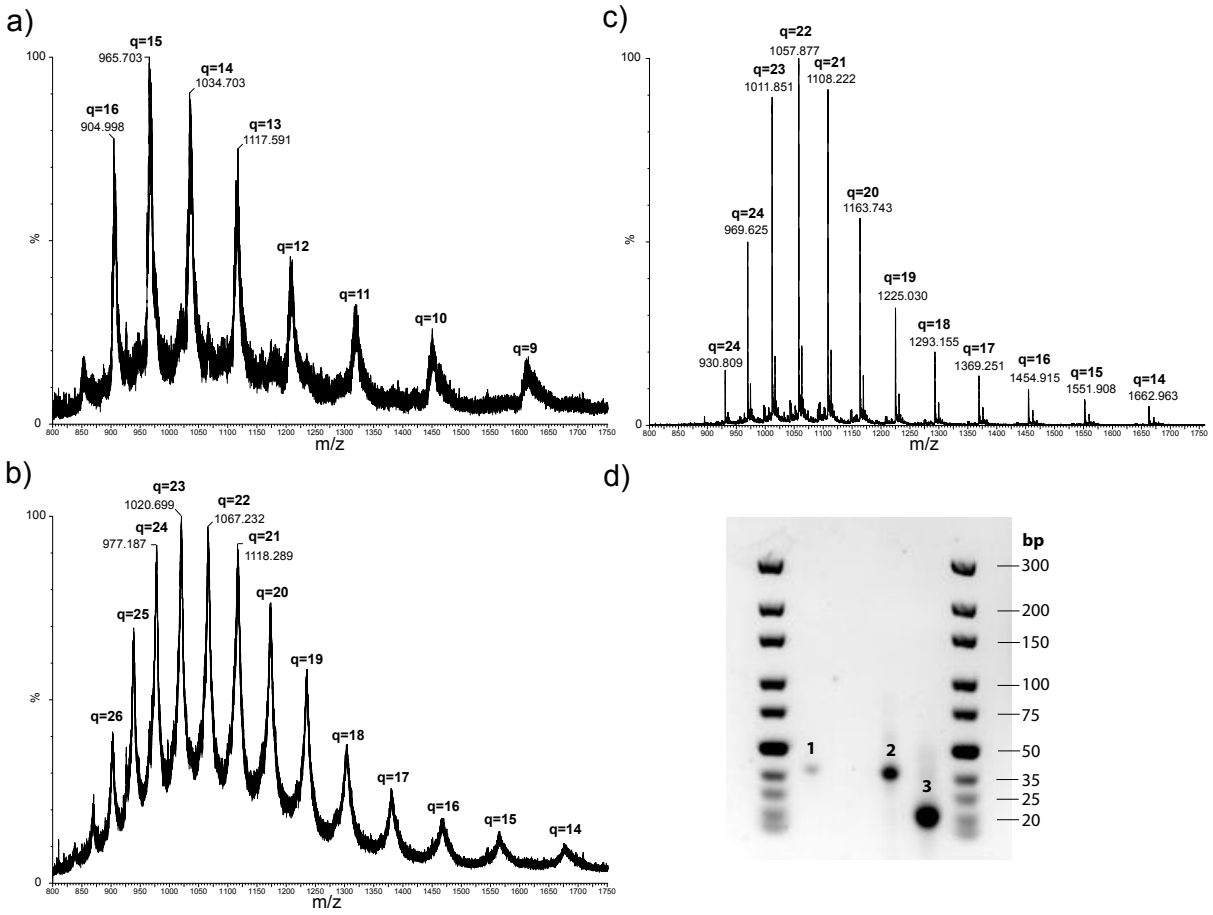


Figure 15: MS analysis of encoding of step 2. Non-deconvoluted spectra (range m/z = 800-1750) of a) unreacted pool step 1 (average MS = 14.5 KDa, charges from -9 to -17); b) encoding reaction **B291** (average MS = 23.5 KDa, charges from -14 to -26); c) ref. oligonucleotide (*o*-I-Phe-CO_2NH- GGAGCTTCTGAATT-A612-B385, MS = 23’297 Da, 74-mer); d) 2% agarose gel where 1 = ref oligonucleotide (37bp), 2 = **pool step 2** (37bp); 3 = pool step 1 (23bp).

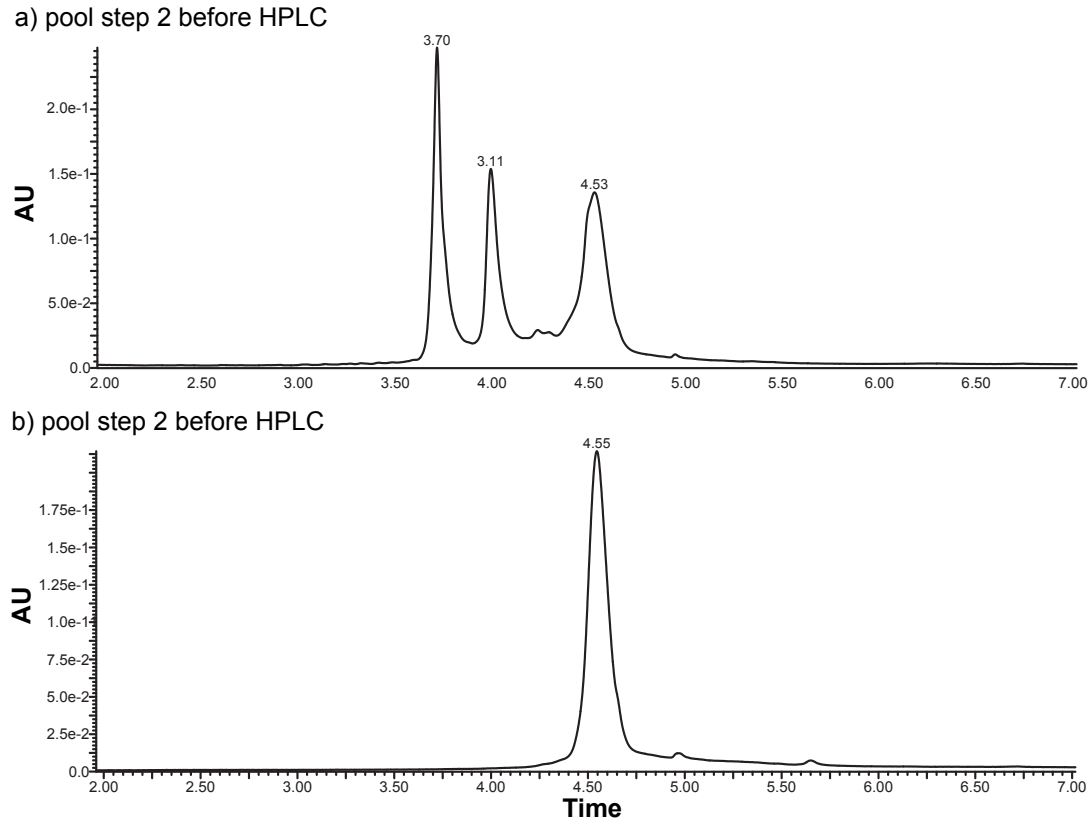


Figure 16: LC-MS chromatograms registered at λ=260 nm of the "**pool step 2**" a) before HPLC purification; b) after HPLC purification (fraction 2, 13.7-22 mins).

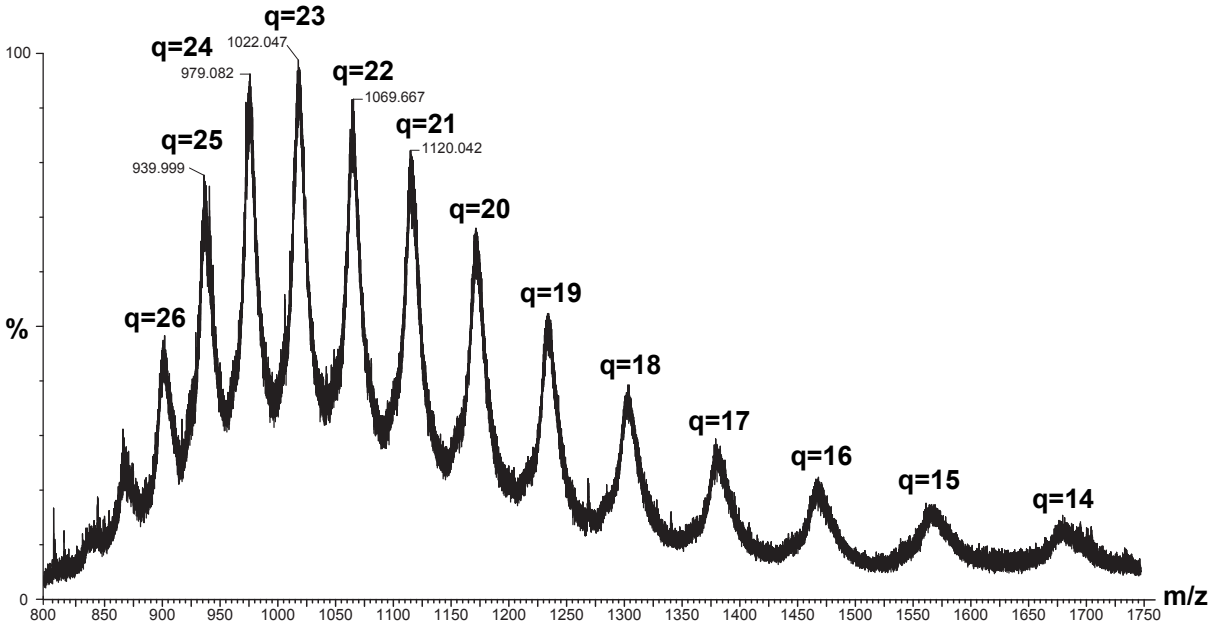


Figure 17: Non-deconvoluted MS-spectra (TOF negative mode) of **pool step 2** after HPLC purification. The average mass of pool is 23.5 KDa, length = 74-mer. **q**: absolute charge.

4.4 Double-strand formation

10 nmol of **(ss)-Library** (10 nmol) and a 12-mer oligonucleotide complementary to the the 3' extremity (5'-GCTGCCTGACGC-3', 20 nmol) were dissolved in 200 µL of 1X NEB buffer 2 (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH = 7.9) and 10 mM deoxynucleotide (dNTP) solution mix was added (20 µL). The annealing was allowed to react at 75°C for 15 minutes. The elongation of the second strand was performed by the addition of 10 units of Klenow polymerase enzyme (5'000 units/ml, 2 µL) at 25°C for 1 hour. The **(ds)-Library** was purified by RP-HPLC at 30 °C.

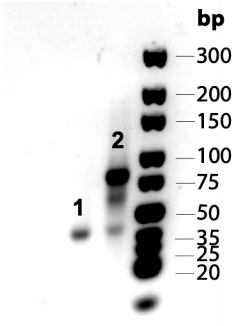


Figure 18: 2% agarose gel where 1 = **ss-Library** (37bp), 2 = **ds-Library** (74bp).

4.5 Library design

Universal oligo.:	5'-GGAGCTTCTGAATT-3'	14-mer
Library step 1:	5'-GGAGCTTCTGAATTCTGTGTGCTG**XXXXXX**CGAGTCCCATGGCGCC-3'	46-mer
Library step 2 (ss):	5'-GGAGCTTCTGAATTCTGTGTGCTG**XXXXXX**CGAGTCCCATGGCGCCGGATCGACG**YYYYYYY**GCGTCAGGCAGC-3'	74-mer
Library step 2 (ds):	5'-GGAGCTTCTGAATTCTGTGTGCTG**XXXXXX**CGAGTCCCATGGCGCCGGATCGACG**YYYYYYY**GCGTCAGGCAGC-3' 3'-CCTCGAAGACTTAAGACACACGAC**XXXXXX**GCTCAGGGTACCGCGGCCTAGCTGC**YYYYYYY**CGCAGTCCGTCG-5'	

Splint s1:	5'- CAGCACACAGAATTCAGAAGCTCC -3'	annealing temp. = 65.2 °C
Splint s2:	5'- CGTCGATCCGGCGCCATGG-3'	annealing temp. = 66.1 °C

5. Affinity Selections

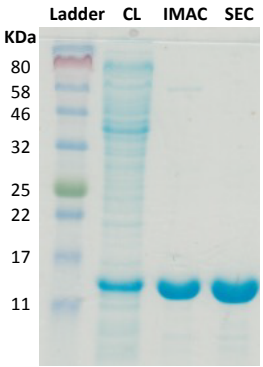
5.1 Proteins for affinity selections

Table 1: List of screened protein targets.

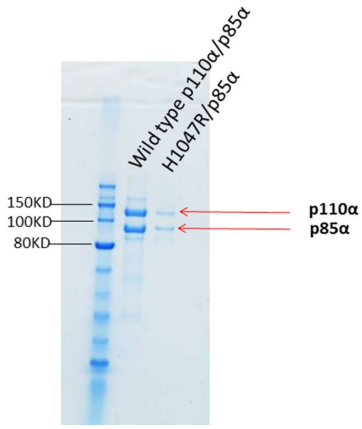
Target	buffer	MW, Da	ε (280 nM), $M^{-1}cm^{-1}$	tag	beads	[protein], µM
CAIX	PBS, pH=7.4	32'732	34'850	biotinylatet	Dynabeads™ MyOne™ Streptavidin C1	2.0
CREBBP	**HEPES1**: 50 mM Hepes, 500 mM NaCl, pH=7.6	26'930	16'673			
wt-PI3K	**HEPES2**: 100 mM NaCl, 20 mM Hepes, 2 mM DTT, pH=7.5.	195K	277'950			
H1047R-PI3K			277'950			
hTNC	PBS, pH=7.4	12'114.52	8'480			
mTNC	PBS, pH=7.4	31'587.18	23'950			
uPA	PBS, pH=7.4	33'000	22'100			
L27E-CtIP	**TRIS**: 20 mM Tris, 150 mM NaCl, 5 mM βmercaptoethanol, pH 8.0.	16757.91 (monomer), 33515.82 (dimer) 66967.8 (tetramer)	22'920			
wt-CtIP						
Albumin	PBS, pH=7.4	66'561	34'445			

5.1.1 CAIX. Recombinant His6-tagged human CAIX was expressed and purified as previously described[3].

5.1.2 CREBBP. The CREBBP bromodomain (Addgene plasmid # 38977) and BRD4(1) bromodomain (Addgene plasmid #38942) constructs were transformed into *E. coli* BL21 (DE3) cells for expression, as described previously[4,5]. The proteins were purified using Immobilized Metal Affinity Chromatography (IMAC) with an HisTrapTM column (GE Healthcare) followed by gel filtration chromatography with Superdex 75 resin (GE Healthcare). The protein purity was assessed by SDS-PAGE.



5.1.3 Wt and H1047R-PI3K Protein Expression, Purification and Biotinylation. We generated pFastBac dual vectors that could co-express His-tagged p110α protein and Avi-tagged p85α protein. The Avi tagged-PI3K vectors were generated from a PI3K pFastBac dual vector received from Peter Shepherd's lab (University of Auckland). The Avi tag was placed at the N terminus of p85α. Then Avi-tagged pFastBac dual vectors were used to generate the baculovirus to infect insect cells. For expression of the PI3K complexes, BTITn-5B1-4 (High Five) cells were infected with baculoviruses encoding the catalytic subunit p110α and regulatory subunit p85α. After 48 h at 28 °C, cells were harvested and washed with ice-cold PBS. The High Five cell pellets were lysed in 20 mM Tris pH 8.0, 100 mM NaCl, 5% glycerol, 10 mM imidazole, and 2 mM β-mercaptoethanol, with one complete EDTA-free protease inhibitor tablet (Roche). Cells were lysed with a 3 minute-probe sonication followed by centrifugation for 1 h at 140,000×g. The supernatant was then passed through a 0.45 µm filter (Advantec). Then a His Ni Resin column (Takara) was used to purify PI3K. The column was washed with up to 50 mM imidazole and then eluted with a buffer containing 20 mM Tris pH 8.0, 100 mM NaCl, 5% glycerol, 300 mM imidazole, and 2 mM β-mercaptoethanol. The His eluate was then loaded onto a 1 mL heparin HP column (GE Healthcare), washed with buffer (20mM Tris pH 8, 100 mM NaCl, 2 mM DTT), and eluted with a 0–100% gradient of buffer (20 mM Tris pH 8, 2 mM DTT, 1 M NaCl). The eluate from the HP column was concentrated to 1 mL using an Amicon 30k centrifugal filter (Millipore) and injected on a Superdex 16/60 200pg gel filtration column (GE Healthcare) pre-equilibrated with buffer (20 mM Hepes pH 7.6, 100 mM NaCl, and 2 mM DTT). Fractions were collected, concentrated, aliquoted and frozen at -80 °C. Biotinylation of PI3K was carried out with the BirA500 kit (AVIDITY). To increase H1047R PI3K protein expression, 150 nM BYL719 was added to the insect cell culture.



5.1.4 TNC. Recombinant His6-tagged human and murine TNC were expressed and purified as previously described[6].

5.1.5 uPA. Recombinant human urokinase (uPA) Protein (His Tag) was purchased (sinobiological, Cat. # 10815-H08H).

5.2.6 Wt and L27E-CtIP. CtIP-NTD (aa 18-145) wt and L27E were cloned and purified as previously described[7]. Briefly, CtIP fragments were PCR-amplified from 3xFLAG-CtIP wt, L27E (pEGFP-C1 backbone)[7,8] and ligated into pET28 MBP-TEV vector (Addgene #69929) upon restriction digest with *BamHI* and *XhoI* (NEB) (primer sequences in Table 1). CtIP-NTD constructs were expressed *E.coli* BL21-CodonPlus-RIL for 20 h at 18°C using 0.5 mM isopropyl-β-D-thiogalactopyranosid (IPTG) and pellets were resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mm NaCl) before snap freezing. After thawing on ice, 1 mM PMSF, protease inhibitor cocktail (Roche) and 0.1 mg/ml lysozyme (Sigma-Aldrich) were added to lysates before stirring for 15 min at 4°C, sonication for 5 min, and ultracentrifugation at 125'000 g for 1 h. Supernatant was loaded onto amylose affinity column (5 ml MBPTrap HP, GE Healthcare) and fusion protein was eluted with 20 mM Tris pH 8.0, 2 mM beta-mercaptoethanol (β-me), 300 mM NaCl, 2 M methyl α-D-glucopyranoside (AMG; Sigma-Aldrich). Subsequently, a buffer exchange with 20 mM Tris pH 8.0, 300 mM NaCl, 5 mM β-me was performed using a HiPrep 26/10 Desalting column (GE Healthcare) and N-terminal His_6-MBP tag was removed by TEV-mediated cleavage at 20°C, overnight, using a five-fold excess of MBP-tagged TEV protease (Gene and Cell Technologies). TEV protease cleavage site-products were captured by amylose affinity chromatography (5 ml MBPTrap HP, GE Healthcare) and preparative size-exclusion chromatography (HiLoad 16/600 Superdex 75, GE Healthcare) in 20 mM Tris pH 8.0, 150 mM NaCl, 5 mM β-me was performed to remove further contaminants.

5.1.7 Human serum albumin. HSA was purchased (Sigma Aldrich, CAS: 70024-90-7).

5.1.8 Protein biotinylation. All the proteins were freshly biotinylated using 3 equivalents of NHS-LC-Biotin (Thermofisher, Cat #: 21336). The reactions were kept at room temperature for 1 hour and the products were purified by PD10 column. The purified proteins were diluted and directly used for affinity selections.

5.2 Selection procedure

Affinity selections were performed with both single (ss) and double strand (ds) library with 10^7 copies of each compound (per selection) as previously described[9]. The selections were performed in duplicate or in triplicate (**Figures 20-24**). The ss-Library and ds-Library were diluted to 110 nM (average conc. of each compound = 0.15 pM) in protein buffer 0.05% tween-20 and 20 µg/mL herring sperm DNA (100 µL). The selections against immobilized protein targets were automated by King Fisher (Thermo Fisher) as previously reported[10].

5.3 PCR amplification and Sequencing

The selection eluates are amplified by two rounds of PCR as previously reported[10] using the following primers:

PCR1-a: 5'-TACACGACGCTCTTCCGATCT XXXXXX GGAGCTTCTGAATTCTGTGTG-3', where X represent a variable region which codify for the selection.
PCR 1-b: 5'-CAGACGTGTGCTCTTCCGATCCGATATGCTGCTGCCTGACGC-3'
PCR2-a: 5'- AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'
PCR2-b: 5'-CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-3'

The PCR products were sequenced by Illumina high-throughput sequencing and the data obtained was processed and analysed as previously reported[9,10].

5.4 Naïve Library

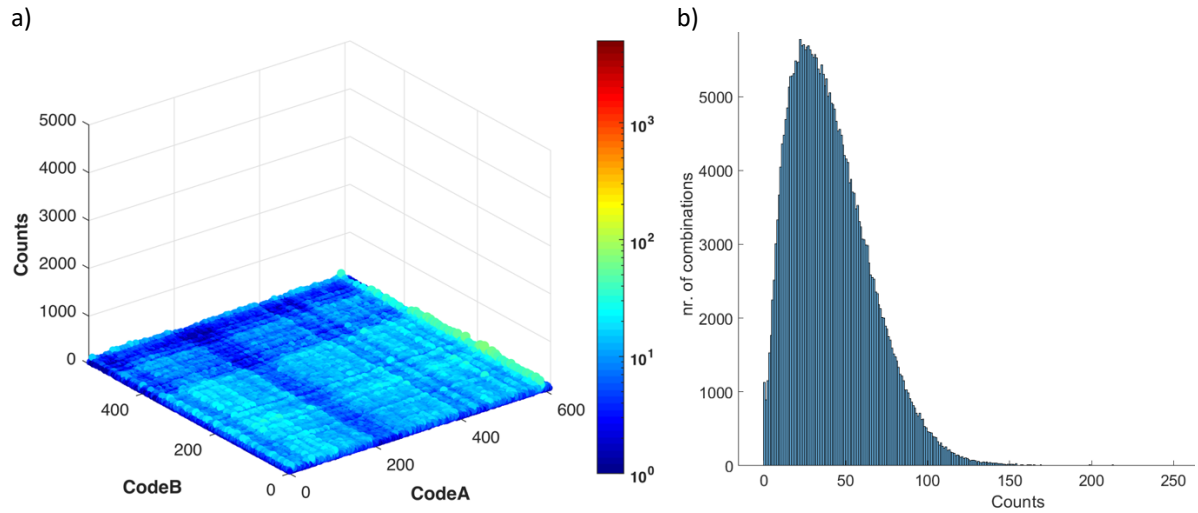


Figure 19: a) Fingerprint of unselected library. The combinations of code A and B are reported in the *xy* plane while the number of counts is visualized on the *z* axis. The average counts (AC) for this selection is 41.07. b) Distribution of counts of Naïve Library (total counts = 13'775'527).

5.4 Fingerprints

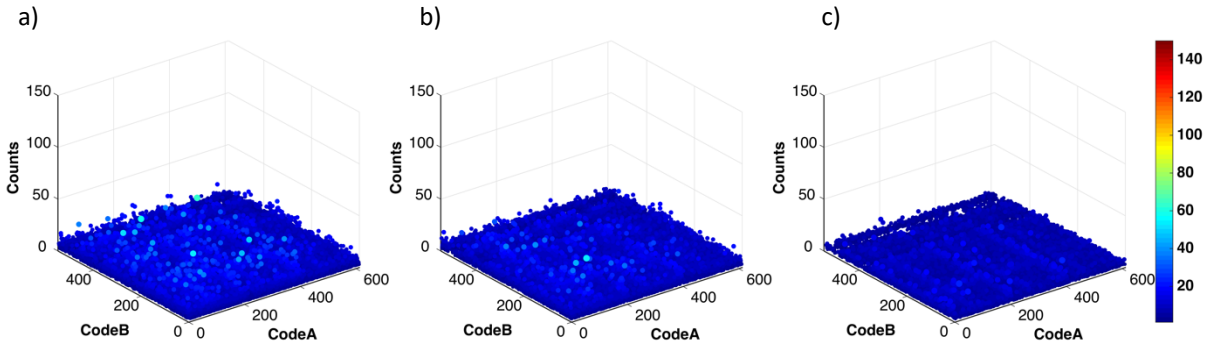


Figure 20: Fingerprint of selections performed against streptavidin beads (no-protein). The average counts (AC) are reported in the **Table 2**.

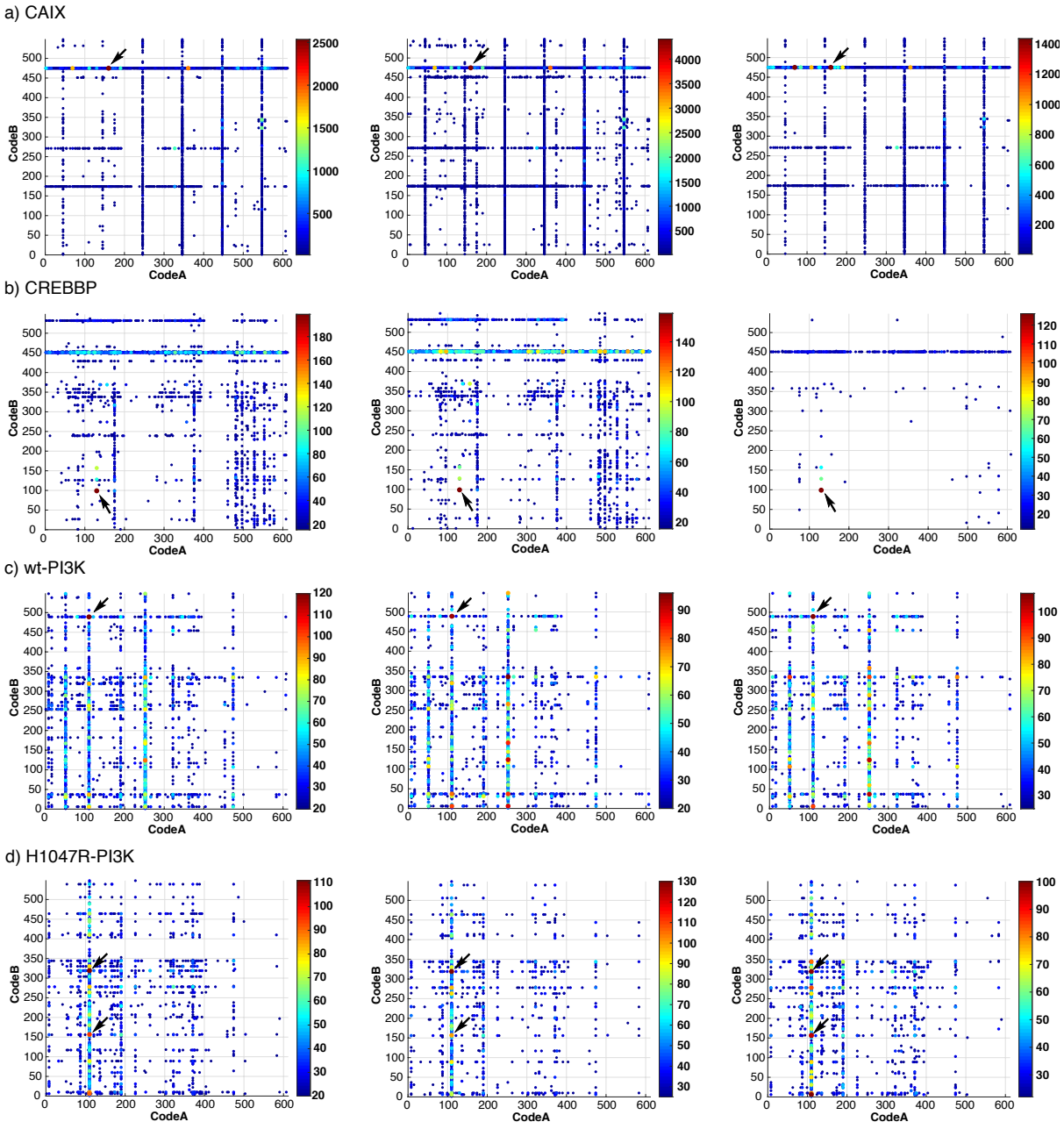


Figure 21: Fingerprints of selections performed in triplicate against a) Carbonic Anhydrase IX, b) CREBBP, c) wt-PI3K and d) H1047R-PI3K. The average counts for each selection are reported in the **Table 2**. The arrows indicate the most enriched combinations which have been resynthesized. Enrichment factors of the most enriched combinations are reported in the **Table 3**.

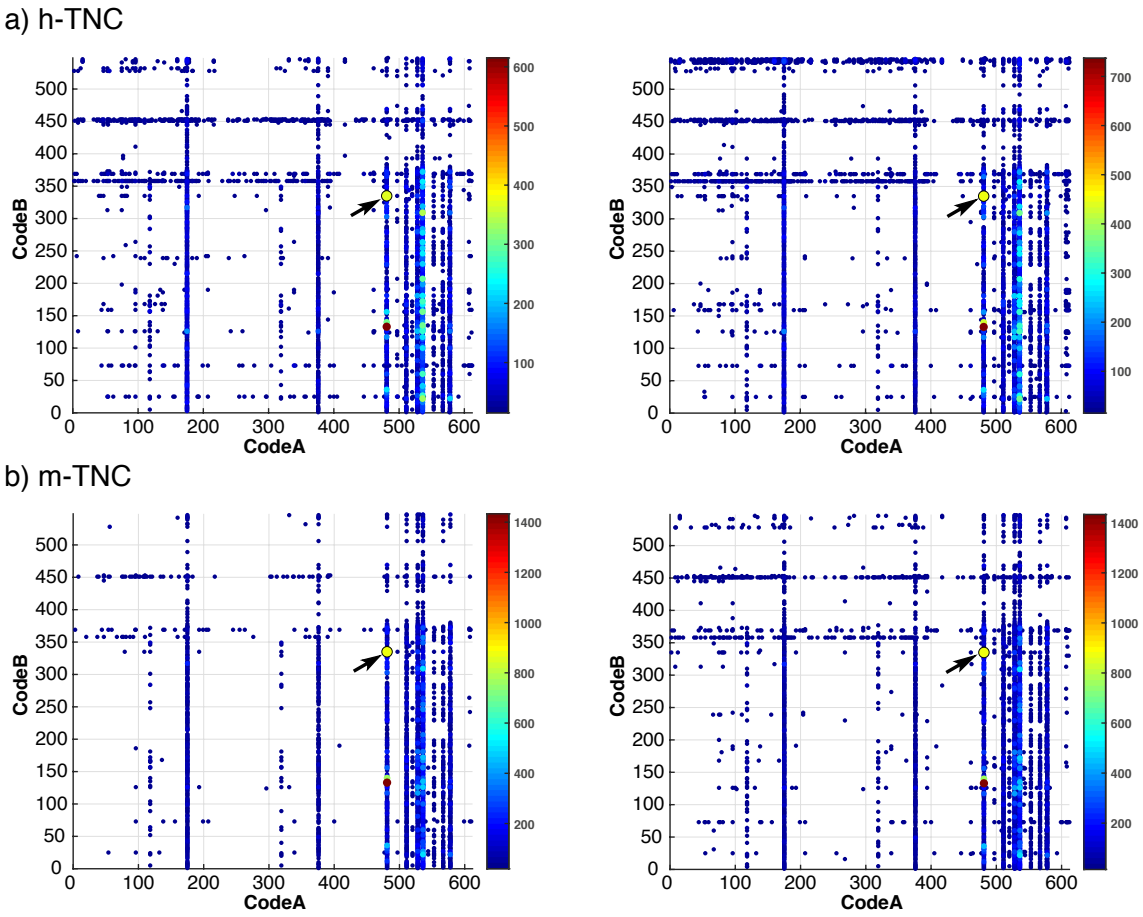


Figure 22: Fingerprints of selections performed in duplicate against a) human-TNC and b) murine-TNC. The average counts for each selection are reported in the **Table 2**. The arrows indicate the most enriched combinations which have been resynthesized. Enrichment factors of the most enriched combinations are reported in the **Table 3**.

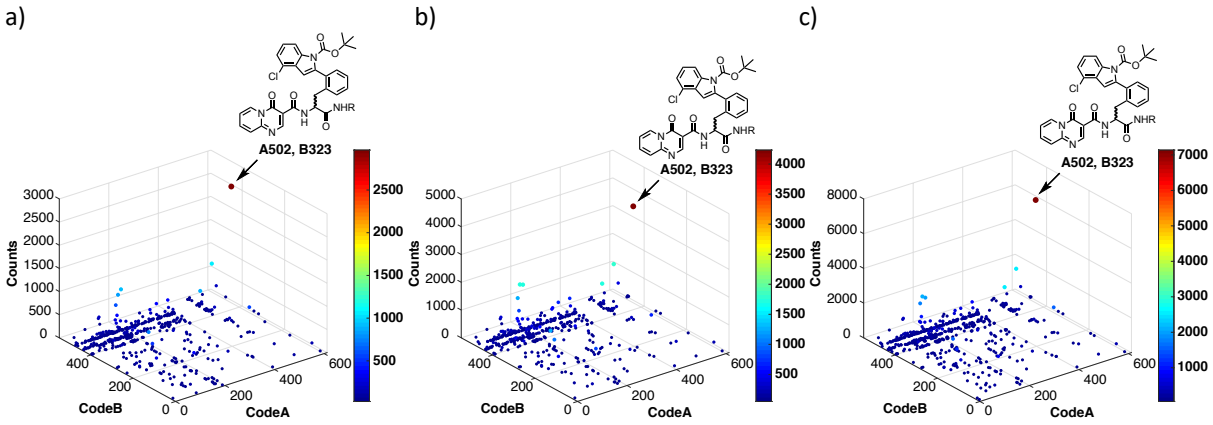


Figure 23: Fingerprints of selections performed in triplicate against human serum albumin. The combination A502/B323 has not been re-synthesised. Average counts and enrichment factors of A502/B323 are reported in **Tables 2** and **3**.

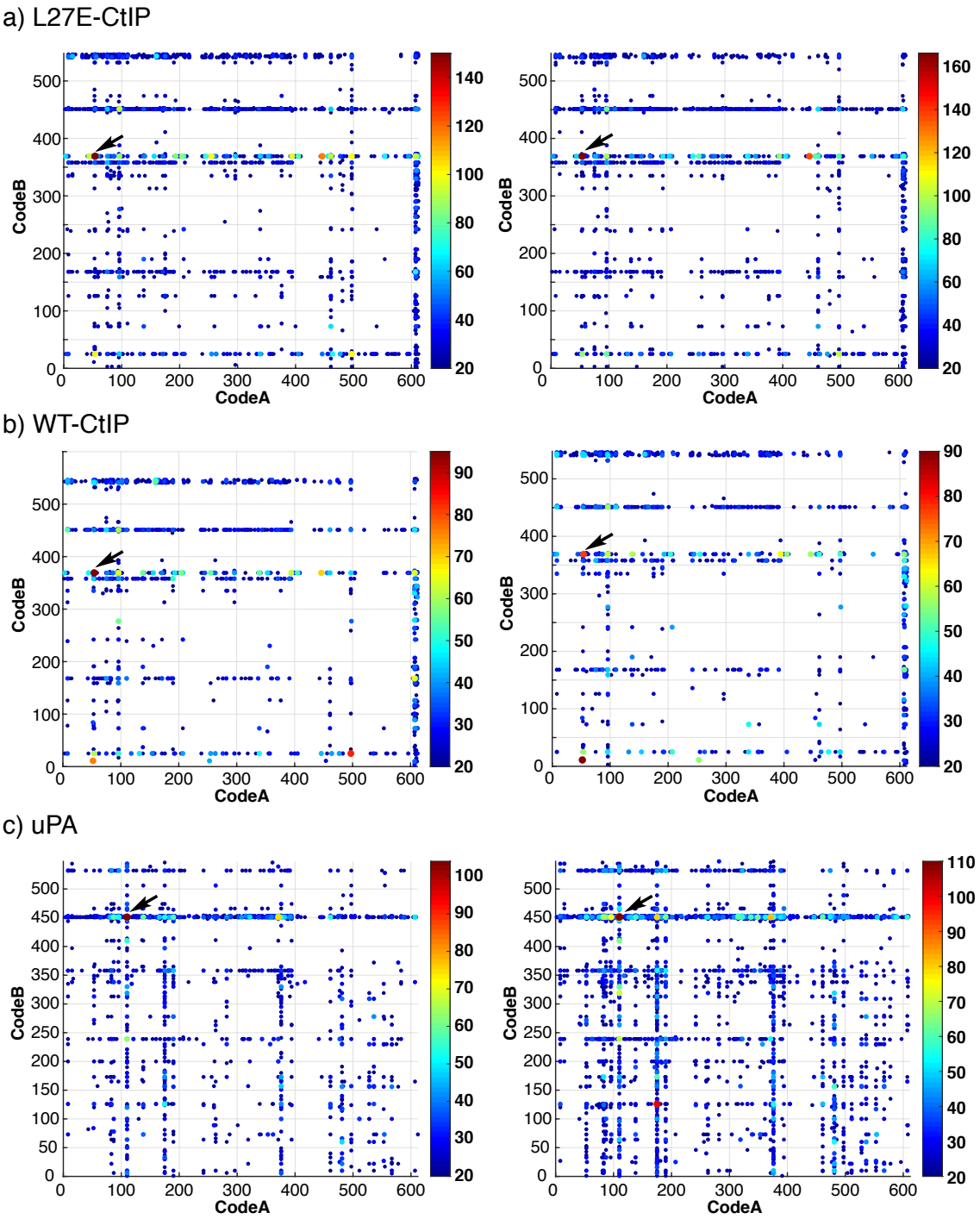


Figure 24: Fingerprints of selections performed in duplicate against a) L27E-CtIP, b) wt-CtIP and c) uPA. The average counts for each selection are reported in the **Table 2**. The arrows indicate the most enriched combinations which have been resynthesized. Enrichment factors of the most enriched combinations are reported in the **Table 3**.

5.6 Selection data analysis

Table 2: detailed analysis of selections. Average counts (AC) are calculated as total counts divided by library size (335'376).

Selection ID	target	Average Count (AC)	total counts	DNA strand	ref. fingerprint
1	Naive	41.07	13'775'527	ss	Figure 19 a
2	No-protein	1.213	406'811	ss	Figure 20 a
3	No-protein	1.11	372'267	ss	Figure 20 b
4	No-protein	1.066	357'511	ss	Figure 20 c
5	HSA	0.9598	321'894	ss	Figure 23 a
6	HSA	1.2	402'451	ss	Figure 23 b
7	HSA	1.393	467'179	ss	Figure 23 c
8	CAIX	1.002	336'047	ss	Figure 21 a (1)
9	CAIX	1.683	564'438	ss	Figure 21 a (2)
10	CAIX	0.8717	292'347	ds	Figure 21 a (3)
11	CAIX	1.536	515'138	ds	-
12	CAIX	0.8141	273'030	ds	-
13	CREBBP	1.201	402'787	ss	Figure 21 b (1)
14	CREBBP	1.171	392'725	ss	Figure 21 b (2)
15	CREBBP	1.213	406'811	ss	Figure 21 b (3)
16	CREBBP	0.9809	328'970	ds	-
17	CREBBP	0.9683	324'745	ds	-
18	CREBBP	0.9911	332'391	ds	-
19	PI3K	1.64	550'017	ss	Figure 21 c (1)
20	PI3K	1.573	527'546	ss	Figure 21 c (2)
21	PI3K	1.707	572'487	ss	Figure 21 c (3)
22	PI3K	0.8	268'301	ds	-
23	PI3K	0.8175	274'170	ds	-
24	PI3K	0.7947	266'523	ds	-
25	H1047R-PI3K	1.544	517'821	ss	Figure 21 d (1)
26	H1047R-PI3K	1.792	600'994	ss	Figure 21 d (2)
27	H1047R-PI3K	1.699	569'804	ss	Figure 21 d (3)
28	H1047R-PI3K	0.9538	319'882	ds	-
29	H1047R-PI3K	0.9639	323'269	ds	-
30	H1047R-PI3K	0.9375	314'415	ds	-
31	wt-CtIP	1.413	473'886	ss	Figure 24 b (1)
32	wt-CtIP	1.417	475'228	ss	Fig 24 b (2)
33	wt-CtIP	1.322	443'367	ds	-
34	wt-CtIP	0.9491	318'305	ds	-
35	L27E-CtIP	1.571	526'876	ss	Figure 24 a (1)
36	L27E-CtIP	1.484	497'698	ss	Figure 24 a (2)
37	L27E-CtIP	0.9663	324'074	ds	-
38	L27E-CtIP	1.044	350'133	ds	-
39	hTNC	1.573	527'546	ss	Figure 22 a (1)
40	hTNC	1.428	478'917	ss	Figure 22 a (2)
41	mTNC	1.542	517'150	ss	Figure 22 b (1)
42	mTNC	1.375	461'142	ss	Figure 22 b (2)
43	uPA	1.561	523'522	ss	Figure 24 c (1)
44	uPA	1.919	643'587	ss	Figure 24 c (2)
45	uPA	1.347	451'751	ds	-
46	uPA	0.9248	310'156	ds	-

Table 3: statistical evaluation of enrichment factors (EF) of the most enriched combinations. EFs are calculated as number of counts for each combination divided by selection's average counts (AC).

ent	Combination (A/B)	Regiochem.	Target	Sel ID	EF #1	EF #2	EF #3	Av. EF	St dev
1	502/323	*ortho*	HSA	5-7	5123	4409	2473	**4002**	1371
2	160/475	*para*	CAIX	8-10	2529	2619	1663	**2270**	528
3	361/475	*meta*	CAIX	8-10	1900	1982	1147	**1677**	460
4	69/475	*para*	CAIX	8-10	1707	1746	1604	**1686**	73
5	130/99	*para*	CREBBP	13-15	132	152	164	**149**	16
6	130/128	*para*	CREBBP	13-15	77	79	73	**76**	3
7	110/489	*para*	wt-PI3K	19-21	73	59	63	**65**	7
8	314/489	*meta*	wt-PI3K	19-21	13	11	16	**13**	3
9	518/489	*ortho*	wt-PI3K	19-21	3	1	0	**1**	2
10	110/157	*para*	wt-PI3K	19-22	37	27	28	**30**	5
11	110/319	*para*	wt-PI3K	19-23	54	45	41	**46**	7
12	110/157	*para*	H1047R-PI3K	25-27	62	61	56	**60**	3
13	110/319	*para*	H1047R-PI3K	25-27	72	73	58	**67**	8
14	110/489	*para*	H1047R-PI3K	25-27	38	30	22	**30**	8
15	54/369	*para*	wt-CtIP	31,32	67	54	16	**46**	27
16	54/369	*para*	L27E-CtIP	37,38	95	112	26	**78**	46
17	481/335	*ortho*	hTNC	39,40	284	257		**270**	19
18	481/335	*ortho*	mTNC	41,42	564	545		**555**	13
19	110/453	*para*	uPA	43,44	44	46		**45**	1

Definitions

(1) $$TC_s = \sum_{i=1}^{612} \sum_{j=1}^{548} SC_s(codeA_i, codeB_j)$$

(2) $$AC_s = \frac{TC_s}{612 \times 548}$$

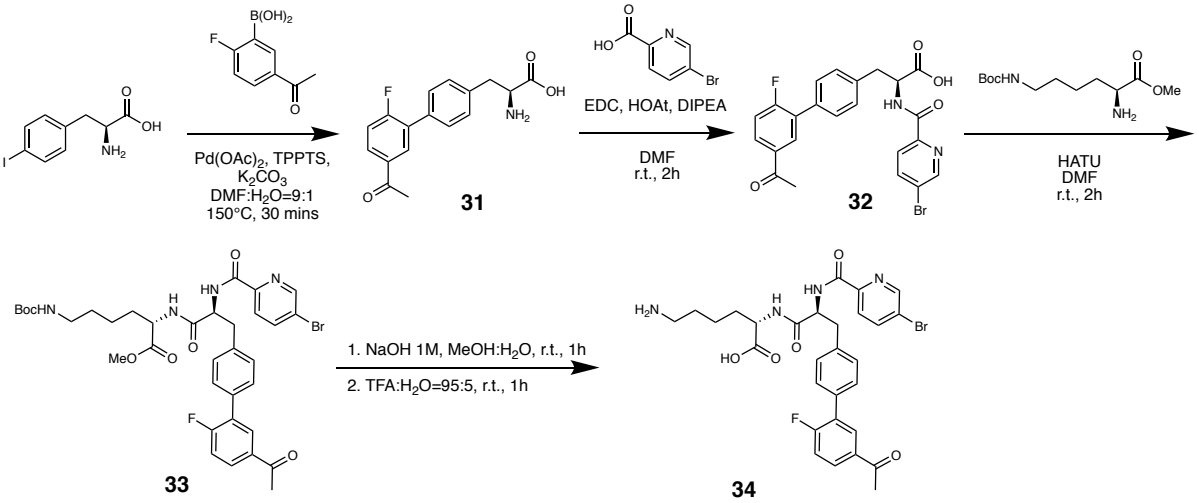
(3) $$EF_{i,j} = \frac{SC_s(codeA_i, CodeB_j)}{AC_s}$$

Equation 1: definition of the total counts (**TC**) for a given selection ***s***, where i and j define the number of diversity elements A and B and SC is the sequence counts. **Equation 2**: definition of the average counts (AC) in a given selection ***s***. **Equation 3**: definition of the enrichment factor (EF) for the *i*-th*, j*-th combination of building blocks A and B.

6. Hit re-synthesis

6.1 Synthesis of CREBBP binders (solution phase)

6.1.2 Synthesis of compound 34



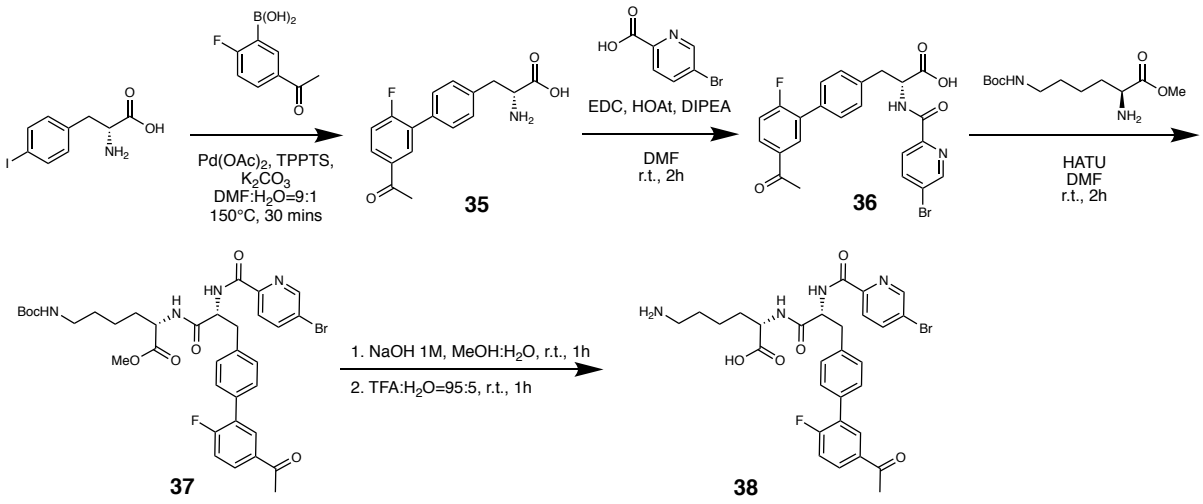
Synthesis of compound 31. Commercially available (S)-2-amino-3-(4-iodophenyl)propanoic acid (200 mg, 0.7 mmol), (5-acetyl-2-fluorophenyl)boronic acid (250 mg, 1.4 mmol), potassium carbonate (390 mg, 2.8 mmol), palladium (II) acetate (31 mg, 0.14 mmol) and triphenylphosphine-3,3',3''-trisulfonic acid trisodium salt (TPPTS, 160 mg, 0.28 mmol) were poured into a round bottom flask and dissolved in DMF:H_2O=2:1. The resulting mixture was heated at 150°C for 30 minutes. The reaction was quenched by 1M HCl and concentrated under reduced pressure. The pure compound **31** was obtained by RP-chromatography (C18 40 µM irregular, 12 g). Yield: 78% (0.55 mmol). **^{1}H NMR** (400 MHz, Methanol-d4) δ 8.15 – 8.10 (m, 1H), 8.09 – 8.03 (m, 1H), 7.62 (d, *J* = 8.3 Hz, 2H), 7.47 (d, *J* = 8.3 Hz, 2H), 7.39 – 7.30 (m, 1H), 4.35 (dd, *J* = 7.6, 5.5 Hz, 1H), 3.46 – 3.35 (m, 1H), 3.27 (dd, *J* = 14.5, 7.5 Hz, 1H), 2.65 (s, 3H). **^{13}C NMR** (101 MHz, MeOD) δ 197.47, 169.74, 163.90, 161.37, 134.42, 134.29, 133.88, 131.00, 129.99, 129.90, 129.48, 129.33, 116.28, 116.04, 53.61, 35.62, 25.34. **m/z** calculated for $C_{17}H_{16}FNO_3$: 301.11, detected (TOF MS ES+): 302.0915.

Synthesis of compound 32. Commercially available 5-bromopicolinic acid (0.8 mmol, 162 mg) was dissolved in dry DMF (5 mL) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, 0.6 mmol, 106 µL), 1-Hydroxy-7-azabenzotriazole (HOAt, 0.8 mmol, 109 mg) and *N,N*-diisopropylethylamine (DIPEA, 3 mmol, 600 µL) were added. The resulting solution was

stirred for 30 minutes at room temperature and **compound 31** was added (120 mg, 0.4mmol). The reaction was kept at room temperature for additional 4 hours. The **pure compound 32 was** obtained by RP-chromatography (C18 40 µM irregular, 12 g). Yield: 60% (0.24 mmol). **^{1}H NMR** (600 MHz, DMSO-d6) δ 8.92 (dd, *J* = 27.5, 8.1 Hz, 1H), 8.56 (d, *J* = 5.2 Hz, 1H), 8.15 (d, *J* = 1.9 Hz, 1H), 8.04 (dd, *J* = 7.7, 2.3 Hz, 1H), 7.98 (ddd, *J* = 8.4, 4.8, 2.3 Hz, 1H), 7.92 (dd, *J* = 5.2, 2.0 Hz, 1H), 7.51 (ddd, *J* = 8.7, 6.8, 2.2 Hz, 2H), 7.49 – 7.41 (m, 1H), 7.37 (dd, *J* = 10.4, 8.2 Hz, 2H), 4.82 – 4.76 (m, 1H), 3.29 (ddd, *J* = 12.1, 6.0, 3.4 Hz, 2H), 2.62 (s, 3H). **^{13}C NMR** (151 MHz, DMSO) δ 196.57, 172.22, 166.14, 162.34, 160.88, 150.50, 149.87, 137.72, 133.67, 132.26, 131.16, 129.41, 128.62, 124.90, 122.08, 116.60, 111.51, 106.36, 53.41, 35.73, 26.71. **m/z** calculated for $C_{23}H_{18}BrFN_2O_4$: 484.04, detected (TOF MS ES+): 484.9627 (^{79}Br), 486.9614 (^{81}Br)

Synthesis of compound 34. Compound **32** (0.21 mmol, 100 mg) was dissolved in dry DMF (2.0 mL) and EDC (0.3 mmol, 53 µL), HOAt (0.3 mmol, 41 mg) and DIPEA (1 mmol, 200 µL) were added. The solution was stirred for 10 minutes at room temperature and methyl N6-(tert-butoxycarbonyl)-L-lysinate hydrochloride (H-Lys(Boc)-OMe, 0.3 mmol, 90 mg) was added. The reaction was kept at room temperature for 4 hours. The reaction was quenched with water and the product was extracted with dichloromethane. The organic phases were combined and dried with anhydrous Na_2CO_3. The solvent was removed under reduce pressure and the crude **compound 33** was dissolved in MeOH and 1M NaOH solution (5 eq., 1.1 mL) was added. The resulting mixture was stirred for 1 hour at room temperature. The reaction was neutralized with 1M HCl and the solvent was removed under reduce pressure. The tert-butoxycarbonyl protective group was finally removed by adding 95:5 TFA:H_2O solution (2 mL) for 1 hour at room temperature. The **pure compound 34** was obtained by RP-chromatography (C18 40 µM irregular, 12 g). Yield: 86% (0.18 mmol). **^{1}H NMR** (600 MHz, DMSO-d6) δ 8.78 – 8.66 (m, 1H), 8.63 (t, *J* = 8.0 Hz, 1H), 8.55 (d, *J* = 5.2 Hz, 1H), 8.13 (d, *J* = 1.8 Hz, 1H), 8.04 (dd, *J* = 7.7, 2.4 Hz, 1H), 7.92 (dd, *J* = 5.2, 2.0 Hz, 1H), 7.78 – 7.63 (m, 3H), 7.50 – 7.45 (m, 2H), 7.45 – 7.41 (m, 1H), 7.41 – 7.36 (m, 2H), 4.88 (dtd, *J* = 19.6, 8.4, 4.6 Hz, 1H), 4.32 – 4.23 (m, 1H), 3.30 – 3.06 (m, 3H), 2.79 (td, *J* = 7.4, 3.7 Hz, 2H), 2.62 (d, *J* = 2.7 Hz, 3H), 1.84 – 1.20 (m, 6H). **^{13}C NMR** (151 MHz, DMSO) δ 196.60, 173.24, 170.32, 166.14, 162.57, 161.99, 160.89, 150.44, 149.87, 139.21, 137.41, 133.68, 132.18, 131.09, 129.65, 128.43, 124.83, 122.09, 116.61, 116.45, 53.50, 51.58, 37.16, 34.22, 30.43, 26.72, 26.42, 22.26. **m/z** calculated for **m/z** calculated for $C_{29}H_{30}BrFN_4O_5$: 612.14, detected (TOF MS ES-): 613.0167 (^{79}Br), 615.3202 (^{81}Br).

6.1.2 Synthesis of compound 38



Synthesis of compound 35. Commercially available (R)-2-amino-3-(4-iodophenyl)propanoic acid (200 mg, 0.7 mmol), (5-acetyl-2-fluorophenyl)boronic acid (250 mg, 1.4 mmol), potassium carbonate (390 mg, 2.8 mmol), palladium (II) acetate (31 mg, 0.14 mmol) and triphenylphosphine-3,3',3''-trisulfonic acid trisodium salt (TPPTS, 160 mg, 0.28 mmol) were poured into a round bottom flask and dissolved in DMF:H_2O=2:1. The resulting mixture was heated at 150°C for 30 minutes. The reaction was quenched by 1M HCl and concentrated under reduced pressure. The pure compound **35** was obtained by RP-chromatography (C18 40 µM irregular, 12 g). Yield: 69% (0.48 mmol), NMR: ^{1}H NMR (400 MHz, Methanol-d4) δ 8.12 (dd, *J* = 7.6, 2.4 Hz, 1H), 8.06 (ddd, *J* = 8.6, 4.8, 2.3 Hz, 1H), 7.61 (dd, *J* = 8.3, 1.7 Hz, 2H), 7.47 (d, *J* = 8.3 Hz, 2H), 7.34 (dd, *J* = 10.3, 8.5 Hz, 1H), 4.35 (dd, *J* = 7.5, 5.5 Hz, 1H), 3.41 (dd, *J* = 14.5, 5.6 Hz, 1H), 3.27 (dd, *J* = 14.5, 7.5 Hz, 1H), 2.65 (s, 3H). ^{13}C NMR (101 MHz, MeOD) δ 197.49, 169.75, 163.89, 161.36, 134.42, 134.27, 133.87, 131.00, 129.98, 129.89, 129.49, 129.33, 116.28, 116.04, 53.62, 35.61, 25.35. **m/z** calculated for $C_{17}H_{16}FNO_3$: 301.11, detected (TOF MS ES+): 302.0911

Synthesis of compound 36. Commercially available 5-bromopicolinic acid (0.8 mmol, 162 mg) was dissolved in dry DMF (5 mL) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, 0.6 mmol, 106 µL), 1-Hydroxy-7-azabenzotriazole (HOAt, 109 mg, 0.8 mmol) and *N,N*-diisopropylethylamine (DIPEA, 3 mmol, 600 µL) were added. The resulting solution was stirred for 30 minutes at room temperature and compound **35** was added (120 mg, 0.4 mmol). The reaction was kept at room temperature for additional 4 hours. The pure compound **36**

was obtained by RP-chromatography (C18 40 µM irregular, 12 g). Yield: 65% (0.26 mmol), **^{1}H NMR** (600 MHz, DMSO-d6) δ 8.87 (d, *J* = 8.0 Hz, 1H), 8.57 – 8.51 (m, 1H), 8.15 (dd, *J* = 2.0, 0.6 Hz, 1H), 8.04 (dd, *J* = 7.8, 2.3 Hz, 1H), 8.00 – 7.95 (m, 1H), 7.91 (dd, *J* = 5.2, 2.0 Hz, 1H), 7.51 – 7.48 (m, 2H), 7.43 (ddd, *J* = 10.5, 8.4, 1.7 Hz, 1H), 7.38 – 7.34 (m, 2H), 4.73 (td, *J* = 7.8, 5.2 Hz, 1H), 3.29 (td, *J* = 14.2, 8.3 Hz, 3H), 2.63 (s, 3H). **^{13}C NMR** (151 MHz, DMSO) δ 196.56, 173.20, 172.21, 162.56, 162.15, 160.88, 150.61, 149.86, 137.87, 133.65, 132.17, 131.17, 129.47, 128.55, 128.05, 124.84, 116.59, 116.43, 53.66, 35.87, 26.71. **m/z** calculated for $C_{23}H_{18}BrFN_2O_4$: 484.04, detected (TOF MS ES-): 484.9540 (^{79}Br), 486.9530 (^{81}Br)

Synthesis of compound 38. Compound **36** (0.21 mmol, 100 mg) was dissolved in dry DMF (2.0 mL) and EDC (0.3 mmol, 53 µL), HOAt (0.3 mmol, 41 mg) and DIPEA (1 mmol, 200 µL) were added. The solution was stirred for 10 minutes at room temperature and methyl N6-(tert-butoxycarbonyl)-L-lysinate hydrochloride (H-Lys(Boc)-OMe, 0.3 mmol, 90 mg) was added. The reaction was kept at room temperature for 4 hours. The reaction was quenched with water and the product was extracted with dichloromethane. The organic phases were combined and dried with anhydrous Na_2CO_3. The solvent was removed under reduce pressure and the crude **compound 37** was dissolved in MeOH and 1M NaOH solution (5 eq., 1.1 mL) was added. The resulting mixture was stirred for 1 hour at room temperature. The reaction was neutralized with 1M HCl and the solvent was removed under reduce pressure. The tert-butoxycarbonyl protective group was finally removed by adding 95:5 TFA:H_2O solution (2 mL) for 1 hour at room temperature. The **pure compound 38** was obtained by RP-chromatography (C18 40 µM irregular, 12 g). Yield: 71% (0.15 mmol). **^{1}H NMR** (600 MHz, DMSO-d6) δ 8.86 – 8.79 (m, 1H), 8.53 (dd, *J* = 5.3, 3.2 Hz, 1H), 8.15 – 8.08 (m, 2H), 8.01 (dt, *J* = 7.7, 2.0 Hz, 1H), 7.96 (ddt, *J* = 6.4, 4.0, 2.1 Hz, 1H), 7.89 (td, *J* = 5.2, 2.0 Hz, 1H), 7.46 (ddd, *J* = 8.0, 6.4, 1.6 Hz, 2H), 7.44 – 7.41 (m, 1H), 7.37 (t, *J* = 7.8 Hz, 2H), 4.90 (dtd, *J* = 39.1, 8.5, 4.7 Hz, 1H), 3.99 (dq, *J* = 13.1, 6.5 Hz, 2H), 3.23 – 3.10 (m, 4H), 2.71 (q, *J* = 7.7 Hz, 2H), 2.60 (s, 3H), 1.75 – 1.57 (m, 2H), 1.51 (dd, *J* = 14.5, 7.5 Hz, 2H), 1.37 – 1.20 (m, 2H). **^{13}C NMR** (151 MHz, DMSO) δ 196.55, 173.78, 169.12, 164.84, 162.54, 161.94, 160.86, 154.90, 150.52, 149.81, 139.23, 137.87, 133.68, 132.03, 131.07, 129.54, 128.38, 127.62, 124.82, 116.57, 114.56, 53.84, 53.53, 38.47, 37.48, 31.42, 26.69, 21.86. **m/z** calculated for $C_{29}H_{30}BrFN_4O_5$: 612.14, detected (TOF MS ES+): 614.1584 (^{79}Br), 616.1569 (^{81}Br).

6.2 Solid phase synthesis

The synthesis of small-molecule ligands of CAIX, wildtype and H1047 PI3K, TNC, CtIP and uPA was performed on solid phase using pre-loaded **Fmoc-L-lys(Boc)-Wang** resin (Bachem, 200-400 mesh, 0.5 mmol/g).

6.2.1 General procedure of Fmoc deprotection

The Fmoc protecting group was removed by incubating three times the resin with piperidine:DMF=1:4 solution (1x30 minutes, 2x10 minutes). After Fmoc deprotection, the resin was washed several times with DMF. The deprotection efficiency was confirmed by TNBS test.

6.2.2 General procedure of amino-acid loading

A solution of Fmoc-protected amino-acid (4 equivalents), O-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HATU, 4 equivalents) and DIPEA (8 equivalents) in dry DMF was added to the "free-amino" peptide resin. The reaction was carried out at room temperature and quenched after four hours by washing the resin several times with DMF. The coupling efficiency was confirmed by TNBS test.

6.2.3 TNBS test

The TNBS (2,4,6-trinitrobenzenesulfonic acid) test can only be used for detecting primary amino groups. The beads turn orange-red in presence of free primary amino group. Few beads were poured into a solution of 2,4,6-trinitrobenzenesulfonic acid in DMF:DIPEA=9:1 and incubated for 5 minutes.

6.2.4 General procedure of amide coupling

The "free-amino" peptide was incubated 4 hours with a solution of activated carboxylic acid (4 equivalents). The carboxylic acids were activated by HATU (4 equivalents) - DIPEA (4 equivalents) procedure. The coupling reaction was quenched by washing the resin several times with DMF. The coupling efficiency was confirmed by TNBS test.

6.2.5 General procedure of azido-transfer.

The resin was swollen in DMSO, and subsequently incubated for 1 hour with 1H-imidazole-1-sulfonyl azide hydrochloride (3 equivalents), DIPEA (9 equivalents) in dry DMSO. The resin was washed several times with DMF and the azido-conversion efficiency was confirmed by TNBS test.

6.2.6 General procedure of CuAAC.

A solution of alkyne (4 equivalents), copper iodide (0.2 equivalents) and tris(benzyltriazolylmethyl)amine (TBTA, 0.25 equivalents) in degassed DMF:TEA = 9:1 was added to the peptide. The reaction was kept overnight at room temperature. The resin was washed with 0.5M EDTA solution pH=8 (2x 5mL), with water (2x 5mL) and several times with DMF. A small portion of the resin was cleaved and the coupling efficiency was confirmed by LC-MS.

6.2.7 General procedure of Suzuki cross-coupling.

Boronic acid (4 equivalents), potassium carbonate (4 equivalents), palladium (II) acetate (0.5 equivalents) and XPhos (0.75 equivalents) were suspended in DMF:water=9:1 and added to the iodo-phenyl peptide derivative. The reaction was kept overnight at room temperature. The resin was washed several times with water and with DMF. A small portion of the resin was cleaved and the coupling efficiency was confirmed by LC-MS.

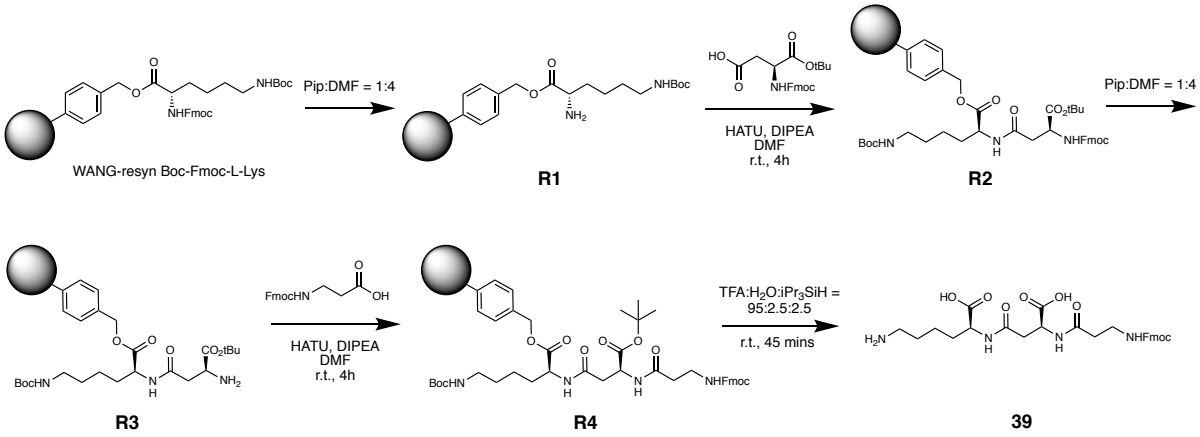
6.2.8 General procedure of Sonogashira cross-coupling.

A solution of alkyne (4 equivalents), copper (I) iodide (0.5 equivalents), palladium (II) acetate (0.5 equivalents) and XPhos (0.75 equivalents) in degassed DMF:TEA=2:1 was added to the iodo-phenyl peptide derivative. The reaction was kept overnight (12-16hrs) at room temperature. The resin was washed with 0.5M EDTA solution pH=8 (2x 5mL), with water (2x 5mL) and several times with DMF. A small fraction of the resin was cleaved and the coupling efficiency was confirmed by LC-MS.

6.2.9 General procedure of resin cleavage and purification

The resin was incubated for 1 hour with a solution of trifluoroacetic acid : water : triisopropylsilane = 95:2.5:2.5 (20 mL / g). The cleavage solution was poured in cold diethyl ether (5 volumes) and the cleaved polypeptide was precipitate for 1 hour at -20°C. The pellet was centrifuged for 20 minutes and the supernatant was discarded. The crude product was dried under reduce pressure and dissolved in mQ Millipore water: acetonitrile = 1:1 mixture (1 mL) and purified by RP-chromatography (C18 40 µM irregular, 12 g) with acetonitrile, 0.1% formic acid (buffer B) : H_2O, 0.1% formic acid (buffer A) as an eluent (2% B for 10 mins; from 2% to 100% B for 45 mins, 100% B for 10 mins).

6.3 Synthesis of tripeptide linker.



The linear tripeptide linker **R4** was assembled on pre-loaded Fmoc-Lys(Boc)-Wang resin (Bachem, 200-400 mesh, 0.5 mmol/g) with the following sequence: H_2N-βAla-Asp-Lys. Fmoc-L-aspartic acid alpha-tert-butyl (ABCR, CAS: 129460-09-9) and Fmoc-beta-alanine (ABCR, CAS: 35737-10-1) were loaded on the resin using the general procedures reported in the section **6.2.1** and **6.2.2**.

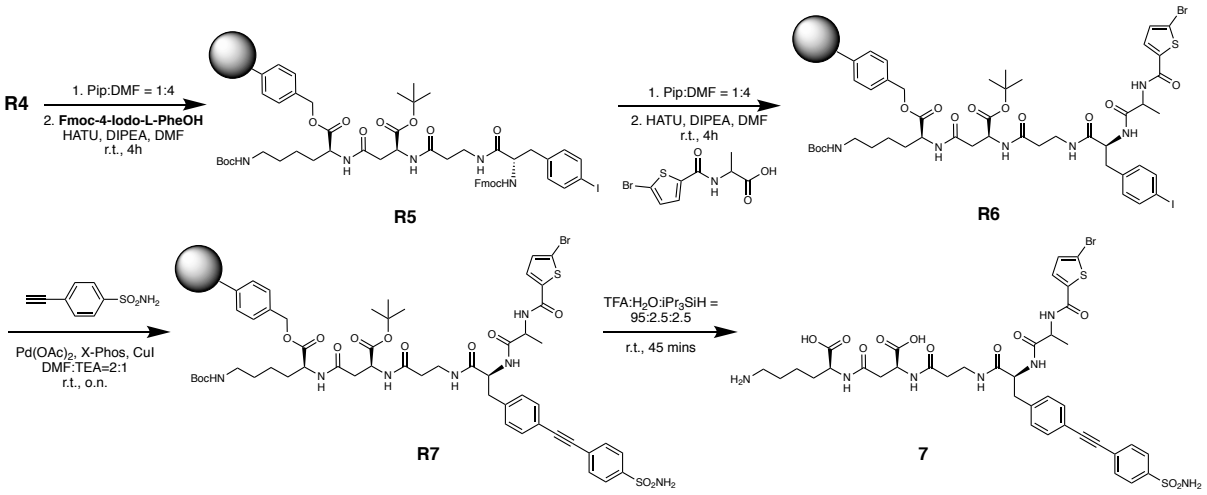
6.3.1 Synthesis of 39

200 mg of resin **R4** (100 µmol) was cleaved as reported in the procedure **6.2.9** and the product **39** was isolated with 36% of yield (36 µmol, 20 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 8.02 (s, 1H), 7.91 – 7.87 (m, 2H), 7.69 (d, *J* = 7.6 Hz, 2H), 7.45 – 7.30 (m, 6H), 4.33 – 4.18 (m, 5H), 4.12 (td, *J* = 8.2, 4.2 Hz, 2H), 3.21 (dt, *J* = 9.6, 7.1 Hz, 3H), 2.75 (d, *J* = 7.5 Hz, 2H), 2.60 (d, *J* = 10.5 Hz, 1H), 2.47 – 2.40 (m, 1H), 2.29 (t, *J* = 7.2 Hz, 2H), 1.77 – 1.62 (m, 1H), 1.50 (dt, *J* = 31.6, 9.6

Hz, 3H), 1.43 – 1.21 (m, 3H). **^{13}C NMR** (151 MHz, DMSO) δ 173.95, 173.41, 169.83, 169.25, 163.89, 155.91, 143.78, 142.43, 140.56, 139.28, 137.29, 128.81, 127.48, 127.18, 126.98, 125.12, 121.27, 119.91, 109.66, 65.33, 51.06, 46.56, 39.80, 38.07, 37.16, 35.83, 30.30, 26.06, 21.58. **m/z** calculated for $C_{28}H_{34}N_4O_8$: 554.24, detected (TOF MS ES+): 555.2159.

6.4 Synthesis of CAIX binders

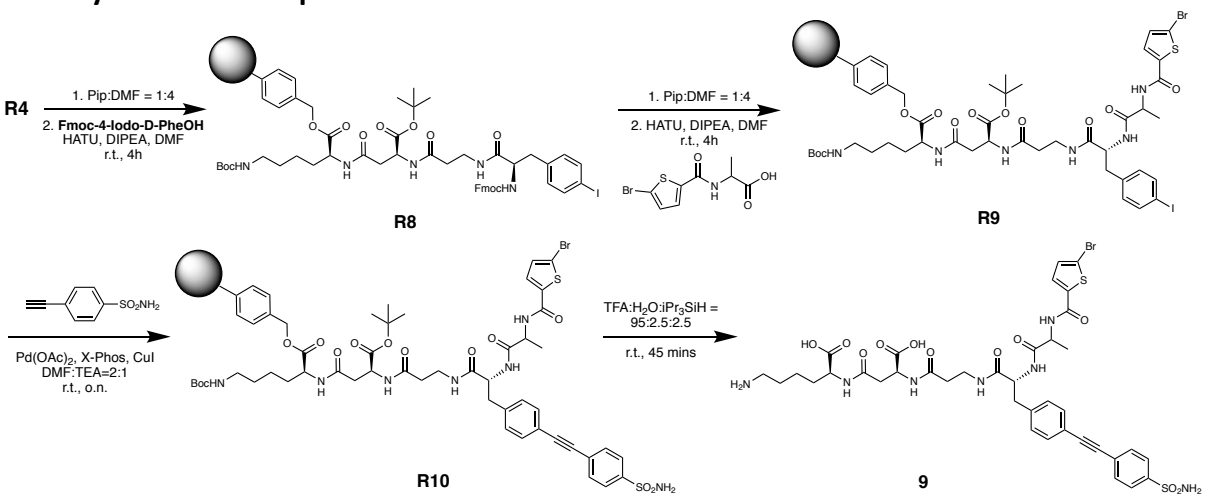
6.4.1 Synthesis of compound 7



Resin **R4** (200 mg, 100 µmol) was deprotected and coupled with Fmoc-4-iodo-L-phenylalanine (CAS: 82565-68-2, 205 mg) via the general procedures (supplementary information **6.2.1** and **6.2.2**) obtaining the resin **R5**. After Fmoc-deprotection, 5-bromothiophene-2-carbonyl (D,L) alanine (CAS: 1396964-88-7, 111 mg) was coupled to the resin using the amide coupling procedure (supplementary information **6.2.4**) obtaining the resin **R6**. 4-ethynylbenzenesulfonamide (CAS: 1788-08-5, 72 mg) was then coupled through Sonogashira cross coupling (supplementary information 6.2.9) yielding resin **R7**. The compound **7** was isolated after cleavage of **R7** and purification with 15% of yield (11 µmol, 10 mg). **m/z** calculated for $C_{38}H_{44}BrN_7O_{11}S_2$: 917.17, detected (TOF MS ES+): 918.1824 (^{79}Br), 920.1838 (^{81}Br). ^{1}H NMR (600 MHz, DMSO-d6) δ 8.32 (s, 1H), 7.89 – 7.80 (m, 3H), 7.78 – 7.65 (m, 4H), 7.37 (d, J = 7.8 Hz, 1H), 7.32 – 7.16 (m, 5H), 4.55 – 4.25 (m, 3H), 4.09 (s, 1H), 3.04 (d, J = 13.2 Hz, 1H), 2.78 (d, J = 42.0 Hz, 3H), 2.29 (d, J = 11.7 Hz, 2H), 1.71 (s, 1H), 1.63 – 1.42 (m, 4H), 1.41 – 1.01 (m, 13H), 0.85 (dd, J = 7.2, 5.8 Hz, 1H). ^{13}C NMR (151 MHz, DMSO) δ 172.00, 171.70, 170.65, 159.80, 159.73, 143.51, 141.48, 141.37, 139.35, 131.63, 131.45, 131.37, 131.07, 130.99, 129.65, 129.41, 125.89, 125.70, 119.29, 116.79, 91.85, 87.70, 69.66, 53.91,

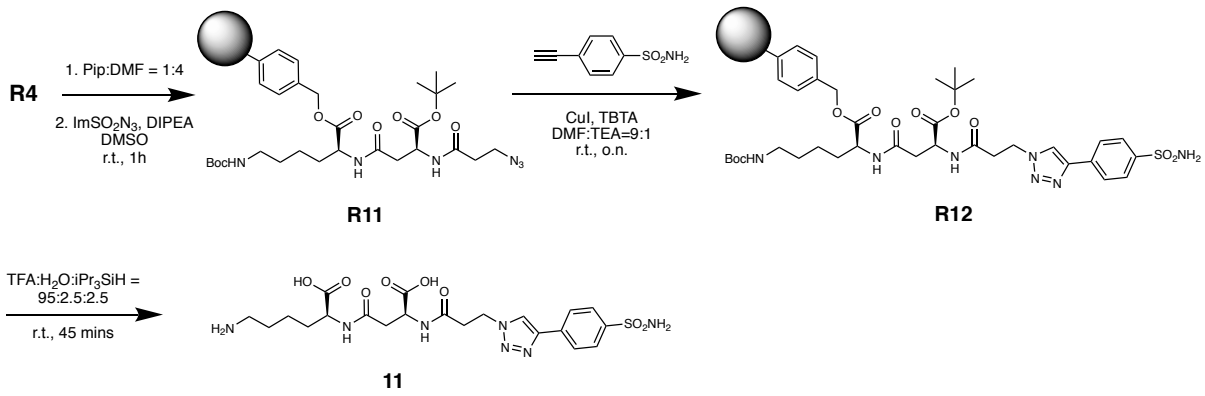
52.72, 51.29, 50.21, 48.92, 38.28, 37.61, 37.44, 35.56, 31.17, 30.79, 28.89, 28.58, 26.36, 21.97, 21.83, 17.51, 17.38, 13.84, 10.69.

6.4.2 Synthesis of compound 9



Resin **R4** (200 mg, 100 µmol) was deprotected and coupled with Fmoc-4-Iodo-D-phenylalanine (CAS: 205526-29-0, 205 mg) using the general procedures (supplementary information **6.2.1** and **6.2.2**) obtaining the resin **R8**. After Fmoc-deprotection, 5-bromothiophene-2-carbonyl (D,L) alanine (CAS: 1396964-88-7, 111 mg) was coupled to the resin using the amide coupling procedure (supplementary information 6.2.4) obtaining the resin **R9**. 4-ethynylbenzenesulfonamide (CAS: 1788-08-5, 72 mg) was than coupled through Sonogashira cross coupling (supplementary information 6.2.9) yielding resin **R10**. The compound **9** was isolated after cleavage of **R10** and purification, with 14% of yield (14 µmol, 13 mg). **m/z** calculated for $C_{38}H_{44}BrN_7O_{11}S_2$: 917.17, detected (TOF MS ES+): 918.1864 (^{79}Br), 929.1859 (^{81}Br). ^{1}H NMR (600 MHz, DMSO-d6) δ 8.47 – 8.14 (m, 3H), 7.89 – 7.81 (m, 2H), 7.75 (s, 1H), 7.74 – 7.69 (m, 1H), 7.46 (s, 1H), 7.37 (d, J = 7.7 Hz, 1H), 7.32 – 7.20 (m, 3H), 4.62 – 4.20 (m, 3H), 4.09 (s, 1H), 3.51 (s, 1H), 3.04 (d, J = 13.2 Hz, 1H), 2.80 (d, J = 66.4 Hz, 3H), 2.28 (s, 2H), 1.96 (d, J = 21.1 Hz, 1H), 1.71 (s, 1H), 1.42 – 1.06 (m, 7H), 0.85 (tt, J = 7.5, 5.2 Hz, 1H). ^{13}C NMR (151 MHz, DMSO) δ 172.01, 171.71, 170.66, 159.73, 143.52, 141.38, 139.36, 131.64, 131.46, 131.38, 131.08, 131.00, 129.66, 125.90, 125.71, 119.30, 117.26, 91.86, 87.71, 69.67, 53.92, 52.73, 51.30, 49.29, 48.93, 38.29, 37.62, 35.58, 31.18, 28.90, 26.37, 21.98, 17.52, 17.39, 13.85.

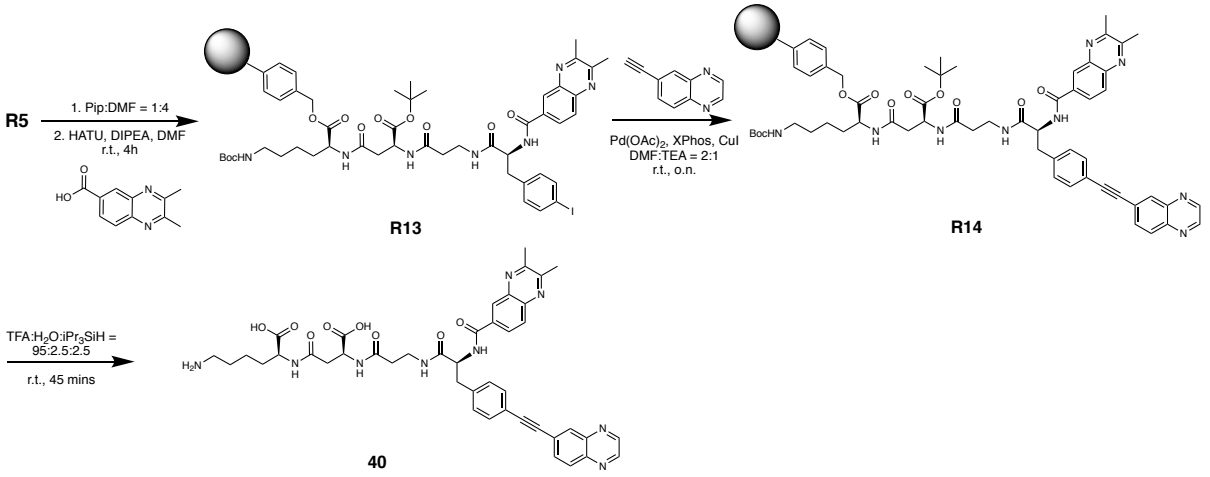
6.4.3 Synthesis of compound 11



200 mg of resin **R4** (100 µmol) were Fmoc-deprotected and incubated with ImN3 (CAS: 952234-37-6, 52 mg) and converted to **R11** using the azido-transfer procedure (Supplementary information **6.2.5**). The resin **R12** was assembled adding 4-ethynylbenzenesulfonamide (CAS: 1788-08-5, 72 mg) using the CuAAC procedure (supplementary information **6.2.6**). The compound **11** was isolated after cleavage of **R12** and purification, with 39% of yield (39 µmol, 21 mg). ^{1}H NMR (600 MHz, DMSO-d6) δ 8.71 (s, 1H), 8.03 (dq, *J* = 8.6, 2.2 Hz, 2H), 7.91 – 7.84 (m, 2H), 7.39 (s, 2H), 4.63 (t, *J* = 6.8 Hz, 2H), 4.30 (dt, *J* = 7.7, 5.6 Hz, 1H), 4.14 (ddd, *J* = 9.4, 7.9, 4.3 Hz, 1H), 2.82 (t, *J* = 6.7 Hz, 2H), 2.75 (t, *J* = 6.8 Hz, 2H), 2.61 – 2.56 (m, 1H), 2.44 (dd, *J* = 14.2, 5.9 Hz, 1H), 1.81 – 1.26 (m, 8H). ^{13}C NMR (151 MHz, DMSO) δ 173.67, 172.96, 169.72, 168.11, 144.84, 142.87, 133.89, 126.26, 125.21, 122.57, 51.50, 50.84, 46.09, 38.07, 37.79, 35.28, 30.05, 25.96, 21.70. **m/z** calculated for $C_{21}H_{29}N_7O_8S$: 539.18, detected (TOF MS ES+): 540.12.59

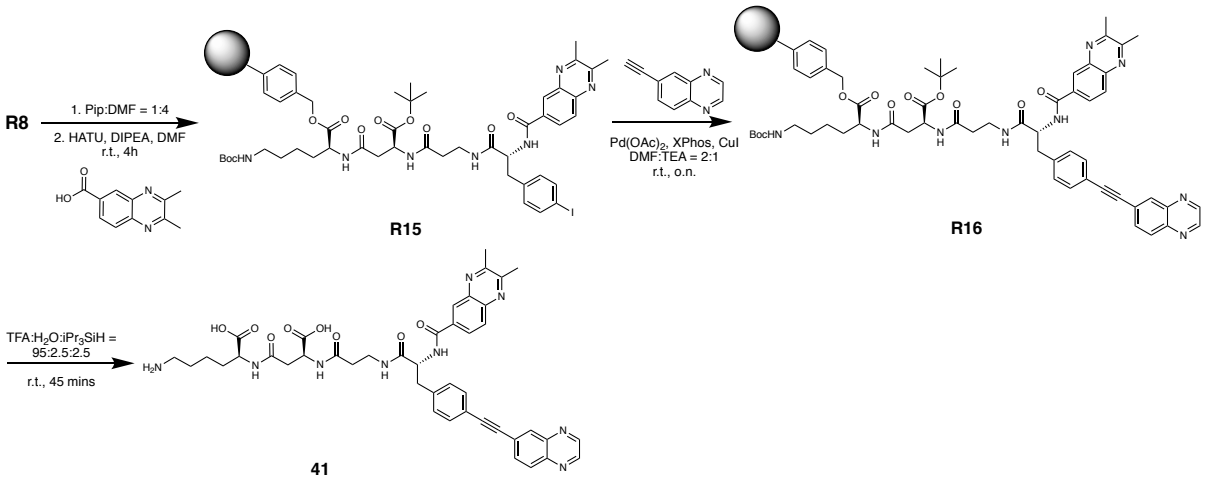
6.5 Synthesis of wt-PI3K binders

6.5.1 Synthesis of compound 40



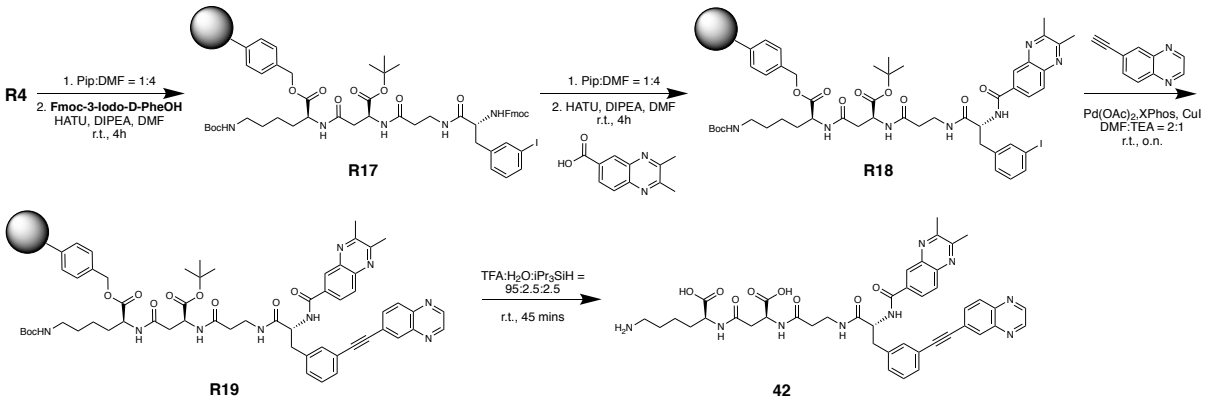
200 mg of resin **R5** (*p*-iodo-L-Phe, 100 µmol) was deprotected and coupled with 2,3-dimethylquinoxaline-6-carboxylic acid (17635-26-6, 81 mg) according with the procedure 6.2.4. The obtained resin **R13** was coupled with 6-ethynylquinoxaline (CAS: 442517-33-1, 62 mg) by Sonogashira cross-coupling (6.2.9) to **R14**. After cleavage and purification, the compound **40** was obtained with 18% of yield (18 µmol, 15 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 9.08 (d, *J* = 8.5 Hz, 1H), 9.01 – 8.93 (m, 2H), 8.48 (d, *J* = 2.0 Hz, 1H), 8.46 – 8.40 (m, 1H), 8.21 (d, *J* = 1.9 Hz, 1H), 8.10 (dddd, *J* = 11.8, 6.6, 4.9, 2.9 Hz, 3H), 8.02 – 7.95 (m, 1H), 7.91 (dd, *J* = 8.7, 1.9 Hz, 1H), 7.87 (d, *J* = 8.0 Hz, 1H), 7.58 – 7.52 (m, 2H), 7.52 – 7.45 (m, 2H), 4.78 (ddt, *J* = 15.1, 8.3, 4.2 Hz, 1H), 4.38 (dt, *J* = 8.1, 5.8 Hz, 1H), 4.19 (td, *J* = 8.4, 4.3 Hz, 1H), 3.42 – 3.30 (m, 4H), 3.22 (td, *J* = 13.9, 6.9 Hz, 2H), 3.10 (dd, *J* = 14.0, 10.9 Hz, 1H), 2.70 (s, 3H), 2.69 (s, 3H), 2.40 – 2.21 (m, 2H), 1.82 – 1.67 (m, 1H), 1.54 (dtd, *J* = 39.3, 14.7, 13.7, 6.9 Hz, 3H), 1.38 (d, *J* = 10.6 Hz, 2H). **^{13}C NMR** (151 MHz, DMSO) δ 173.87, 173.32, 170.91, 169.83, 169.66, 165.43, 155.51, 154.95, 146.55, 146.07, 141.89, 141.79, 141.54, 140.13, 139.50, 133.72, 132.35, 131.56, 131.29, 129.62, 129.12, 128.35, 127.84, 127.40, 127.29, 124.07, 119.41, 92.14, 88.04, 54.93, 51.78, 50.58, 38.27, 37.19, 35.62, 35.35, 30.40, 26.15, 22.81, 22.75, 21.83. **m/z** calculated for $C_{43}H_{45}N_9O_8$: 815.34, detected (TOF MS ES+): 816.1780.

6.5.2 Synthesis of compound 41



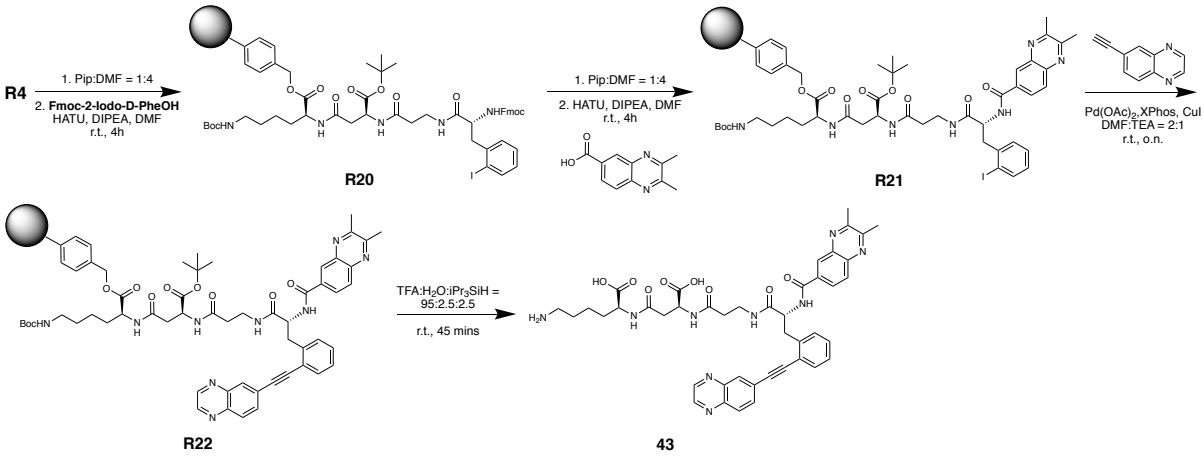
200 mg of resin **R8** (*p*-iodo-D-Phe, 100 µmol) was deprotected and coupled with 2,3-dimethylquinoxaline-6-carboxylic acid (17635-26-6, 81 mg) according to the procedure **6.2.4**. The obtained resin **R15** was coupled with 6-ethynylquinoxaline (CAS: 442517-33-1, 62 mg) by Sonogashira cross-coupling (6.2.9) to **R16**. After cleavage and purification, the compound **41** was obtained with 22% of yield (22 µmol, 18 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 9.18 (d, *J* = 8.5 Hz, 1H), 8.98 (d, *J* = 1.8 Hz, 1H), 8.95 (t, *J* = 1.9 Hz, 1H), 8.51 – 8.46 (m, 2H), 8.22 (d, *J* = 1.9 Hz, 1H), 8.13 – 8.06 (m, 3H), 7.97 (d, *J* = 8.6 Hz, 1H), 7.92 (dd, *J* = 8.6, 1.9 Hz, 1H), 7.74 (d, *J* = 7.7 Hz, 1H), 7.56 – 7.52 (m, 2H), 7.51 – 7.45 (m, 2H), 4.77 (ddd, *J* = 11.0, 8.5, 4.2 Hz, 2H), 4.48 – 4.06 (m, 5H), 3.10 (ddd, *J* = 13.8, 10.9, 3.3 Hz, 3H), 2.76 (p, *J* = 5.8 Hz, 2H), 2.71 (s, 3H), 2.70 (s, 3H), 2.32 (t, *J* = 6.9 Hz, 2H), 1.83 – 1.64 (m, 2H), 1.64 – 1.27 (m, 7H).**^{13}C NMR** (151 MHz, DMSO) δ 173.76, 173.11, 170.91, 169.96, 169.57, 165.38, 155.48, 154.92, 146.56, 146.08, 141.90, 141.80, 141.53, 140.20, 139.51, 133.78, 132.36, 131.57, 131.27, 129.62, 127.79, 127.46, 127.29, 124.08, 119.37, 92.16, 88.00, 55.02, 51.58, 50.98, 38.10, 37.85, 37.16, 35.63, 35.54, 30.13, 25.98, 22.82, 22.77, 21.67. **m/z** calculated for $C_{43}H_{45}N_9O_8$: 815.34, detected (TOF MS ES+): 816.1819.

6.5.3 Synthesis of compound 42



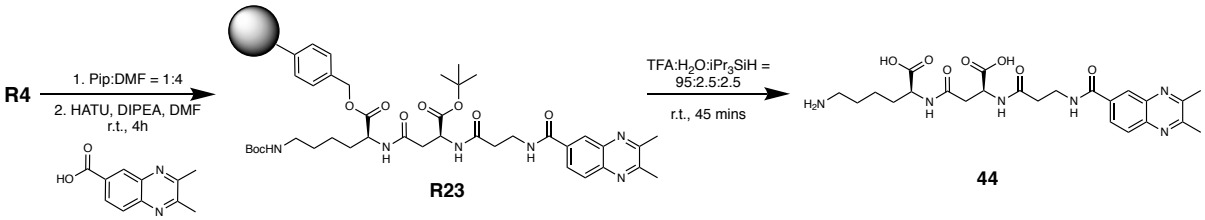
100 mg of resin **R4** (50 µmol) was coupled with Fmoc-3-Iodo-D-phenylalanine (CAS: 478183-67-4, 103 mg) according with the procedure **6.2.2** and subsequently **R17** was deprotected and coupled with 2,3-dimethylquinoxaline-6-carboxylic acid (17635-26-6, 40 mg) following the procedure listed in **6.2.4**. The obtained resin **R18** was coupled with 6-ethynylquinoxaline (CAS: 442517-33-1, 31 mg) by Sonogashira cross-coupling (6.2.9) to **R19**. After cleavage and purification, the compound **42** was obtained with 17% of yield (8.6 µmol, 7 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 9.15 (d, *J* = 8.5 Hz, 1H), 9.00 (d, *J* = 1.8 Hz, 1H), 8.97 (d, *J* = 1.8 Hz, 1H), 8.52 – 8.48 (m, 1H), 8.46 (t, *J* = 5.8 Hz, 1H), 8.16 – 8.13 (m, 1H), 8.11 – 8.07 (m, 2H), 8.07 – 8.03 (m, 1H), 7.98 – 7.93 (m, 1H), 7.90 – 7.85 (m, 1H), 7.73 (d, *J* = 7.7 Hz, 1H), 7.49 (dt, *J* = 7.8, 1.5 Hz, 1H), 7.43 (dt, *J* = 7.7, 1.3 Hz, 1H), 7.35 (t, *J* = 7.7 Hz, 1H), 4.75 (ddd, *J* = 11.0, 8.6, 4.2 Hz, 1H), 4.28 (q, *J* = 6.0 Hz, 1H), 4.18 – 4.12 (m, 1H), 3.20 (dd, *J* = 13.8, 4.1 Hz, 3H), 3.11 – 3.03 (m, 2H), 2.75 (tq, *J* = 12.9, 5.9 Hz, 3H), 2.65 (s, 3H), 2.63 (s, 3H), 2.32 (t, *J* = 6.9 Hz, 2H), 1.72 (d, *J* = 7.2 Hz, 1H), 1.60 – 1.27 (m, 6H). **^{13}C NMR** (151 MHz, DMSO) δ 173.78, 173.16, 170.88, 169.89, 169.53, 165.51, 155.41, 154.86, 146.59, 146.12, 141.87, 141.81, 141.51, 139.52, 139.26, 133.85, 132.42, 132.31, 131.57, 130.28, 129.61, 129.36, 128.48, 127.77, 127.45, 127.32, 123.97, 121.21, 92.08, 87.98, 55.13, 51.67, 50.99, 38.10, 37.95, 36.91, 35.60, 35.51, 30.20, 26.03, 22.77, 22.70, 21.65. **m/z** calculated for $C_{43}H_{45}N_9O_8$: 815.34, detected (TOF MS ES+): 816.1810

6.5.4 compound 43



100 mg of resin **R4** (50 µmol) was coupled with Fmoc-2-Iodo-D-phenylalanine (CAS: 478183-65-2, 103 mg) according with the procedure **6.2.2** and subsequently **R20** was deprotected and coupled with 2,3-dimethylquinoxaline-6-carboxylic acid (17635-26-6, 40 mg) according with the procedure **6.2.4**. The obtained resin **R21** was coupled with 6-ethynylquinoxaline (CAS: 442517-33-1, 31 mg) by Sonogashira cross-coupling (6.2.9) to **R22**. After cleavage and purification, the compound **43** was obtained with 24% of yield (12 µmol, 10 mg). **m/z** calculated for $C_{43}H_{45}N_9O_8$: 815.34, detected (TOF MS ES+): 814.2906. ^{1}H NMR (500 MHz, DMSO-d6) δ 9.08 – 8.85 (m, 2H), 8.53 – 8.40 (m, 1H), 8.40 – 8.34 (m, 1H), 8.28 (s, 1H), 8.16 – 8.01 (m, 4H), 8.00 – 7.90 (m, 1H), 7.57 (dd, J = 45.4, 7.5 Hz, 1H), 7.42 – 7.21 (m, 4H), 7.16 (t, J = 7.4 Hz, 1H), 4.97 (dt, J = 9.4, 4.9 Hz, 1H), 4.69 (ddd, J = 11.4, 8.4, 4.4 Hz, 1H), 4.49 (q, J = 6.9 Hz, 1H), 4.18 (td, J = 9.1, 4.4 Hz, 1H), 3.07 (dt, J = 46.6, 5.7 Hz, 1H), 2.75 (t, J = 5.5 Hz, 2H), 2.73 – 2.66 (m, 6H), 2.33 (p, J = 8.4, 8.0 Hz, 2H), 1.91 (s, 2H), 1.78 – 1.62 (m, 1H), 1.52 (ddq, J = 20.4, 13.3, 7.5, 6.7 Hz, 4H), 1.36 (s, 3H), 1.23 (d, J = 7.0 Hz, 2H), 1.16 (t, J = 7.2 Hz, 4H). ^{13}C NMR (126 MHz, DMSO) δ 171.40, 170.42, 164.58, 158.85, 155.91, 155.34, 147.07, 146.51, 143.00, 142.47, 142.07, 141.25, 140.66, 140.04, 134.37, 133.04, 132.61, 132.45, 130.92, 130.02, 129.62, 129.51, 128.51, 128.26, 128.10, 127.86, 127.26, 124.80, 122.20, 93.09, 91.06, 55.19, 36.84, 36.22, 35.81, 35.76, 28.04, 23.37, 23.31.

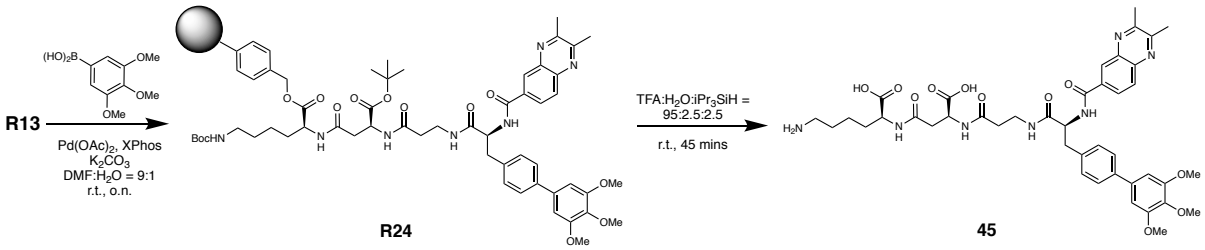
6.5.4 Synthesis of compound 44



100 mg of resin **R4** (50 µmol) was coupled with 2,3-dimethylquinoxaline-6-carboxylic acid (17635-26-6, 40 mg) according with the procedure **6.2.4**. The obtained resin **R23** was cleaved and purified. The compound **44** was obtained with 39% of yield (19 µmol, 10 mg). ^{1}H NMR (600 MHz, DMSO-d6) δ 8.94 (t, *J* = 5.6 Hz, 1H), 8.45 (d, *J* = 1.9 Hz, 1H), 8.14 (dd, *J* = 8.6, 2.0 Hz, 1H), 8.02 (d, *J* = 7.9 Hz, 1H), 7.97 (d, *J* = 8.6 Hz, 1H), 7.87 (d, *J* = 7.7 Hz, 1H), 4.31 (dt, *J* = 7.9, 5.8 Hz, 1H), 4.11 (td, *J* = 8.2, 4.5 Hz, 1H), 3.58 – 3.52 (m, 3H), 2.74 (t, *J* = 7.0 Hz, 2H), 2.68 (s, 6H), 2.60 (dt, *J* = 13.2, 6.8 Hz, 1H), 2.45 (t, *J* = 7.2 Hz, 2H), 1.75 – 1.25 (m, 7H). ^{13}C NMR (151 MHz, DMSO) δ 174.07, 173.57, 169.73, 165.35, 155.36, 154.86, 141.43, 139.55, 134.22, 127.82, 127.12, 52.10, 50.97, 38.19, 37.88, 36.47, 35.48, 30.50, 26.20, 22.73, 21.73.**m/z** calculated for $C_{24}H_{32}N_6O_7$: 516.23, detected (TOF MS ES+): 517.1739.

6.6 Synthesis of H1047R-PI3K binders

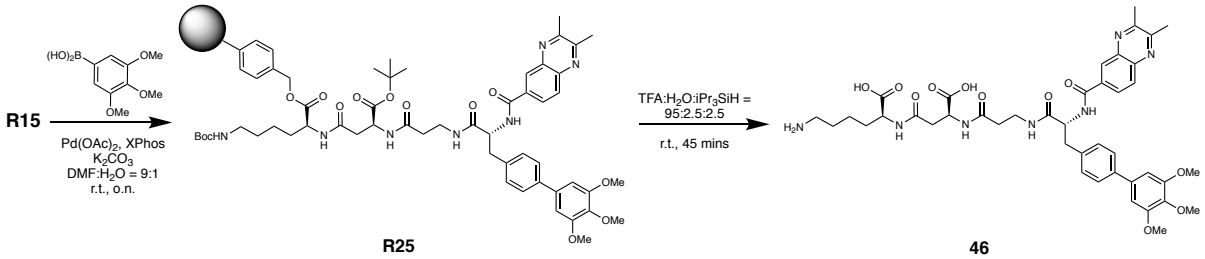
6.6.1 Synthesis of compound 45



200 mg of resin **R13** (L-isomer, 100 µmol) was coupled with (3,4,5-trimethoxyphenyl)boronic acid (CAS: 182163-96-8, 85 mg) by Suzuki cross-coupling (**6.2.8**) to **R24**. After cleavage and purification, the compound **45** was obtained with 19% of yield (19µmol, 16 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 9.09 (d, *J* = 8.5 Hz, 1H), 8.50 (d, *J* = 2.1 Hz, 1H), 8.46 – 8.39 (m, 1H), 8.12 (dd, *J* = 8.7, 2.0 Hz, 1H), 8.09 (dq, *J* = 9.0, 2.1 Hz, 1H), 7.99 – 7.94 (m, 1H), 7.72 (d, *J* = 7.8 Hz, 1H), 7.59 – 7.55 (m, 2H), 7.52 (td, *J* = 9.9, 9.1, 4.3 Hz, 1H), 7.48 – 7.41 (m, 2H), 6.85 (s, 2H), 4.75 (ddd, *J* = 11.0, 8.4, 4.1 Hz, 1H), 4.29 (dt, *J* = 7.8, 5.5 Hz, 1H), 4.18 (td, *J* = 8.6, 4.2 Hz, 1H), 3.81 (s, 6H), 3.65 (s, 3H), 3.18 (dd, *J* = 13.9, 4.1 Hz, 3H), 3.08 (dd, *J* = 13.9, 10.9 Hz, 2H), 2.70

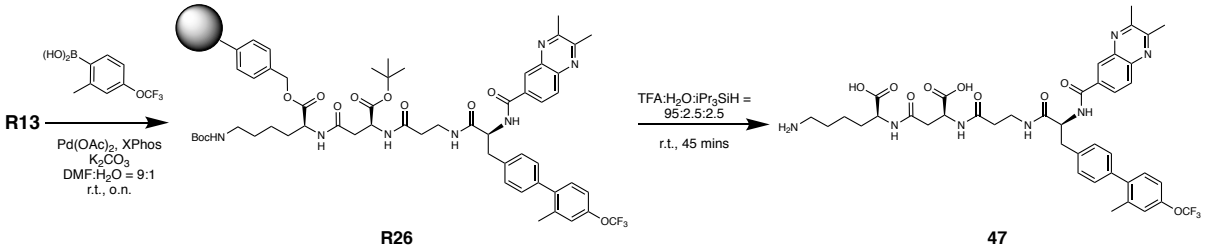
(s, 3H), 2.69 (s, 3H), 2.45 (dd, *J* = 14.0, 5.7 Hz, 1H), 2.35 – 2.28 (m, 2H), 1.73 (dp, *J* = 10.1, 6.4, 4.5 Hz, 1H), 1.60 – 1.31 (m, 6H), 1.27 – 1.20 (m, 1H). **^{13}C NMR** (151 MHz, DMSO) δ 173.69, 173.21, 171.13, 170.02, 169.48, 165.38, 155.51, 154.94, 152.99, 141.53, 139.49, 138.04, 137.61, 136.76, 135.66, 133.76, 131.79, 131.30, 129.47, 128.55, 127.82, 127.43, 127.33, 126.29, 103.81, 59.90, 55.76, 55.30, 51.55, 51.01, 38.10, 36.78, 35.65, 35.51, 30.14, 25.98, 22.81, 22.74, 21.65. **m/z** calculated for $C_{42}H_{51}N_7O_{11}$: 829.36, detected (TOF MS ES+): 830.3386.

6.6.2 Synthesis of compound 46



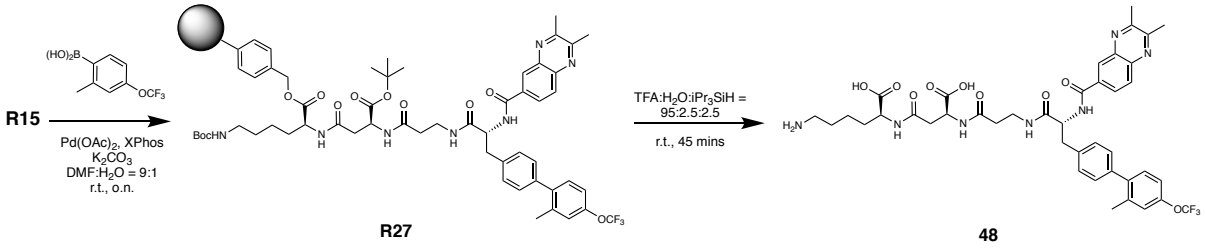
200 mg of resin **R15** (D-isomer, 100 µmol) was coupled with (3,4,5-trimethoxyphenyl)boronic acid (CAS: 182163-96-8, 85 mg) by Suzuki cross-coupling (**6.2.8**) to **R25**. After cleavage and purification, the compound **46** was obtained with 17% of yield (17 µmol, 14 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 9.16 (d, *J* = 8.5 Hz, 1H), 8.55 – 8.49 (m, 1H), 8.47 (t, *J* = 5.8 Hz, 1H), 8.11 (dhept, *J* = 6.6, 2.0 Hz, 2H), 7.96 (t, *J* = 8.3 Hz, 1H), 7.74 (d, *J* = 7.7 Hz, 1H), 7.61 – 7.56 (m, 2H), 7.55 – 7.50 (m, 1H), 7.49 – 7.42 (m, 3H), 6.85 (s, 2H), 4.74 (ddd, *J* = 11.0, 8.3, 4.1 Hz, 1H), 4.29 (dt, *J* = 7.7, 5.6 Hz, 1H), 4.19 (td, *J* = 10.1, 8.6, 4.2 Hz, 1H), 3.81 (d, *J* = 1.7 Hz, 6H), 3.65 (d, *J* = 1.8 Hz, 3H), 3.16 (dd, *J* = 13.9, 4.1 Hz, 2H), 3.09 (dd, *J* = 13.9, 10.9 Hz, 2H), 2.76 (hept, *J* = 6.1, 5.7 Hz, 2H), 2.71 (s, 3H), 2.68 (s, 3H), 2.65 – 2.57 (m, 2H), 2.32 (q, *J* = 7.0, 5.9 Hz, 2H), 1.77 – 1.68 (m, 1H), 1.61 – 1.50 (m, 2H), 1.45 (dq, *J* = 12.4, 6.4, 5.7 Hz, 1H), 1.42 – 1.33 (m, 2H). **^{13}C NMR** (151 MHz, DMSO) δ 173.72, 173.07, 171.16, 169.94, 169.62, 165.38, 155.48, 154.91, 152.99, 141.52, 139.51, 138.03, 137.65, 136.76, 135.66, 133.78, 131.30, 129.48, 128.55, 128.47, 127.78, 127.49, 127.32, 126.28, 103.80, 59.89, 55.75, 55.37, 51.48, 50.87, 38.11, 37.75, 36.76, 35.60, 30.13, 25.97, 22.81, 22.76, 21.68. **m/z** calculated for $C_{42}H_{51}N_7O_{11}$: 829.36, detected (TOF MS ES+): 830.2145.

6.6.3 Synthesis of compound 47



200 mg of resin **R13** (L-isomer, 100 µmol) was coupled with (2-methyl-4-(trifluoromethoxy)phenyl)boronic acid (CAS: 850033-39-5, 88 mg) by Suzuki cross-coupling (**6.2.8**) to **R26**. After cleavage and purification, the compound **47** was obtained with 20% of yield (20 µmol, 17 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 9.18 (s, 1H), 8.46 (qd, *J* = 6.3, 4.8, 2.3 Hz, 2H), 8.14 – 8.06 (m, 2H), 8.00 – 7.91 (m, 2H), 7.72 (dd, *J* = 7.6, 4.1 Hz, 1H), 7.45 (dd, *J* = 7.6, 5.1 Hz, 2H), 7.28 – 7.24 (m, 2H), 7.23 (dd, *J* = 7.7, 5.4 Hz, 2H), 7.18 (dd, *J* = 7.9, 2.7 Hz, 1H), 4.76 (ddd, *J* = 10.9, 8.4, 4.4 Hz, 2H), 4.27 (q, *J* = 6.1 Hz, 1H), 4.16 (td, *J* = 8.6, 4.3 Hz, 1H), 3.13 – 3.06 (m, 3H), 2.70 (s, 3H), 2.68 (s, 3H), 2.31 (dt, *J* = 14.1, 7.1 Hz, 2H), 2.16 (d, *J* = 5.3 Hz, 3H), 1.95 (d, *J* = 20.9 Hz, 1H), 1.73 (q, *J* = 17.8, 9.8 Hz, 2H), 1.61 – 1.41 (m, 4H), 1.41 – 1.29 (m, 3H), 1.22 (d, *J* = 12.9 Hz, 2H). **^{13}C NMR** (151 MHz, DMSO) δ 173.94, 173.30, 171.15, 169.89, 169.63, 165.51, 155.47, 154.91, 147.12, 141.50, 140.34, 139.48, 137.56, 137.47, 134.30, 133.85, 131.04, 130.26, 129.06, 128.53, 127.76, 127.47, 127.33, 122.28, 118.07, 55.20, 52.08, 50.84, 38.15, 38.10, 36.85, 35.62, 35.42, 30.43, 26.14, 22.81, 22.75, 21.84, 20.01. **m/z** calculated for $C_{41}H_{46}F_3N_7O_9$: 837.33, detected (TOF MS ES+): 838.2709

6.6.4 Synthesis of compound 48

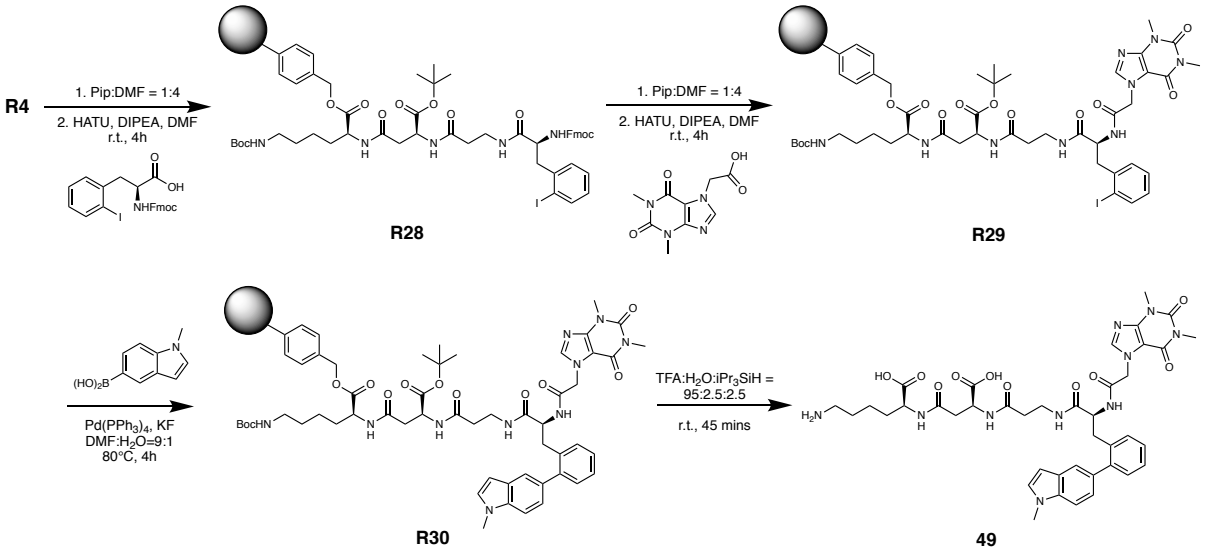


200 mg of resin **R15** (D-isomer, 100 µmol) was coupled with (2-methyl-4-(trifluoromethoxy)phenyl)boronic acid (CAS: 850033-39-5, 88 mg) by Suzuki cross-coupling (**6.2.8**) to **R27**. After cleavage and purification, the compound **48** was obtained with 22% of yield (22 µmol, 18 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 9.17 (d, *J* = 8.3 Hz, 1H), 8.46 (qd, *J* =

6.3, 4.8, 2.3 Hz, 2H), 8.08 (td, *J* = 8.7, 3.0 Hz, 2H), 8.01 – 7.92 (m, 2H), 7.72 (dd, *J* = 7.6, 4.1 Hz, 1H), 7.45 (dd, *J* = 7.6, 5.1 Hz, 2H), 7.29 – 7.21 (m, 5H), 7.18 (dd, *J* = 7.9, 2.7 Hz, 1H), 4.76 (ddd, *J* = 10.9, 8.4, 4.4 Hz, 2H), 4.28 (q, *J* = 6.1 Hz, 1H), 4.16 (td, *J* = 8.6, 4.3 Hz, 1H), 3.10 (dd, *J* = 14.0, 10.7 Hz, 3H), 2.70 (s, 3H), 2.68 (s, 3H), 2.32 (dt, *J* = 14.1, 7.1 Hz, 2H), 2.17 (d, *J* = 5.3 Hz, 3H), 2.02 – 1.87 (m, 1H), 1.80 – 1.16 (m, 11H). **^{13}C NMR** (151 MHz, DMSO) δ 173.80, 173.18, 171.15, 169.95, 169.58, 165.56, 155.46, 154.90, 147.13, 141.50, 140.36, 139.50, 137.59, 137.48, 133.91, 131.05, 129.05, 128.55, 127.75, 127.49, 127.31, 122.29, 120.85, 119.15, 118.08, 69.67, 55.24, 51.68, 51.02, 38.11, 37.89, 36.85, 35.61, 30.20, 26.01, 22.81, 22.76, 21.64, 20.02. **m/z** calculated for $C_{41}H_{46}F_3N_7O_9$: 837.33, detected (TOF MS ES+): 838.2979

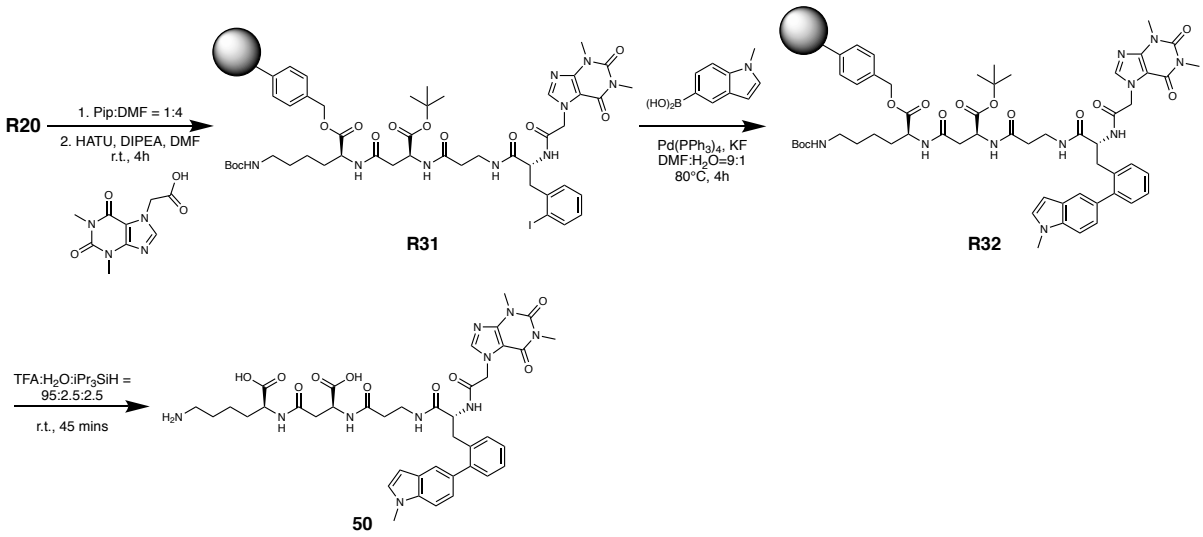
6.7 Synthesis of TNC binders

6.7.1 Synthesis of compound 49



200 mg of resin **R4** (100 µmol) was coupled with Fmoc-2-Iodo-L-phenylalanine (CAS: 210282-32-9, 205 mg) according with the procedure **6.2.2** and subsequently **R28** was deprotected and coupled with theophylline-7-acetic acid (CAS: 652-37-9, 95 mg) according with the procedure 6.2.4. The obtained resin **R29** was coupled with 1-methylindole-5-boronic acid (CAS: 192182-55-1, 70 mg) by Suzuki cross-coupling (6.2.8) to **R30**. After cleavage and purification, the product **49** was obtained with 10% of yield (10 µmol, 8 mg). **m/z** calculated for $C_{40}H_{48}N_{10}O_{10}$: 828.36, detected (TOF MS ES+): 829.3591

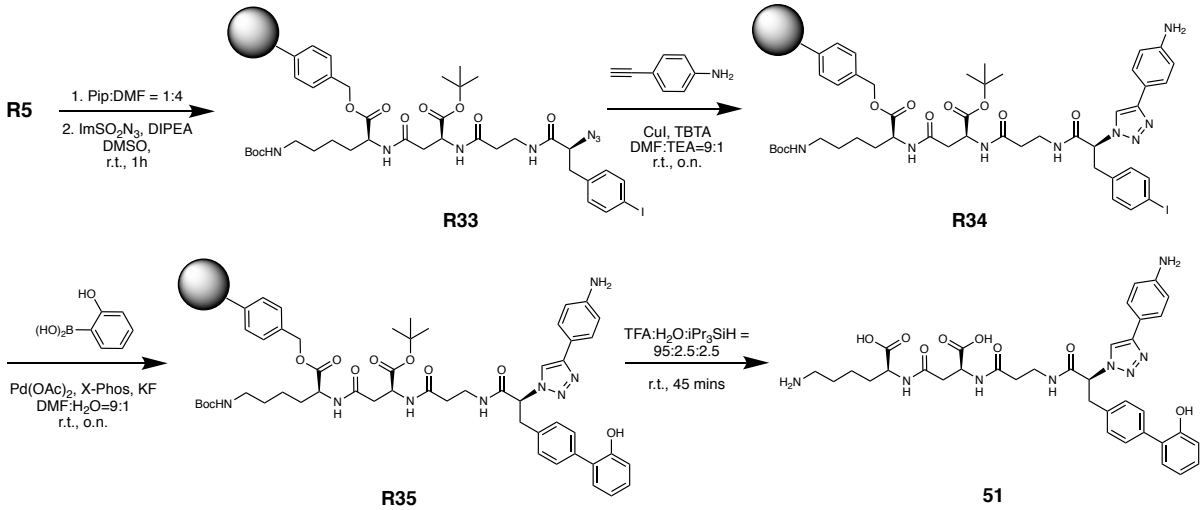
6.7.2 Synthesis of compound 50



200 mg of resin **R20** (*ortho*-Iodo D-isomer, 100 µmol) was coupled with theophylline-7-acetic acid (CAS: 652-37-9, 95 mg) according to the procedure **6.2.4**. The obtained resin **R31** was coupled with 1-methylindole-5-boronic acid (CAS: 192182-55-1, 70 mg) by Suzuki cross-coupling (**6.2.8**) to **R32**. After cleavage and purification, the product **50** was obtained with 12% of yield (12µmol, 10 mg). **m/z** calculated for $C_{40}H_{48}N_{10}O_{10}$: 828.36, detected (TOF MS ES+): 829.8002

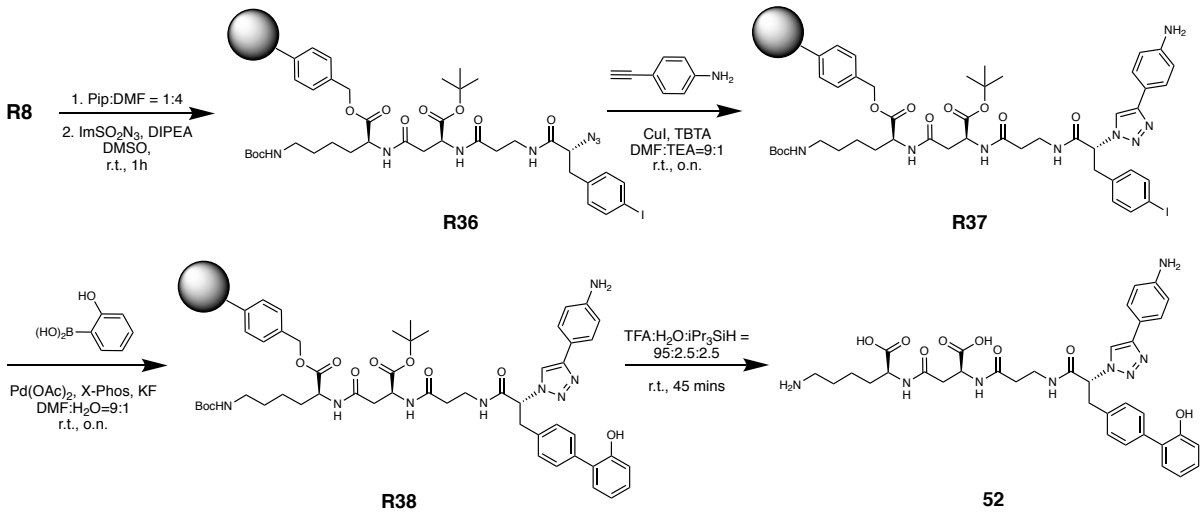
6.8 Synthesis of L27E-CtIP binders

6.8.1 Synthesis of compound 51



200 mg of resin **R5** (L-isomer, 100 µmol) was incubated with ImN3 (CAS: 952234-37-6, 52 mg) and converted to **R33** using the azido-transfer procedure (Supplementary information **6.2.5**). The obtained resin **R33** were coupled with 4-ethynylaniline (CAS: 14235-81-5, 47 mg) by CuAAC procedure **6.2.6**. The obtained resin **R34** was coupled with (2-hydroxyphenyl)boronic acid (CAS: 89466-08-0, 55 mg) by Suzuki cross-coupling (6.2.8) to **R35**. After cleavage and purification, the product **51** was obtained with 42% of yield (42µmol, 30 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 8.82 – 8.68 (m, 1H), 8.53 (d, *J* = 3.8 Hz, 1H), 8.14 (d, *J* = 8.3 Hz, 1H), 7.98 – 7.90 (m, 1H), 7.76 (dd, *J* = 11.6, 7.3 Hz, 1H), 7.65 – 7.59 (m, 1H), 7.49 (ddd, *J* = 17.0, 8.9, 2.9 Hz, 3H), 7.44 – 7.36 (m, 2H), 7.29 – 7.21 (m, 2H), 7.18 (dt, *J* = 7.6, 1.9 Hz, 1H), 7.14 – 7.08 (m, 1H), 6.91 (dd, *J* = 8.1, 1.2 Hz, 1H), 6.81 (tt, *J* = 7.5, 1.9 Hz, 1H), 6.63 – 6.56 (m, 2H), 5.66 – 5.59 (m, 1H), 4.35 (q, *J* = 6.3 Hz, 1H), 4.19 (td, *J* = 8.7, 4.4 Hz, 2H), 3.34 (s, 2H), 3.23 (q, *J* = 6.1, 5.2 Hz, 2H), 2.75 (t, *J* = 6.9 Hz, 3H), 2.61 (t, *J* = 10.3 Hz, 1H), 2.34 – 2.19 (m, 3H), 1.72 (d, *J* = 11.7 Hz, 1H), 1.62 – 1.29 (m, 6H). **^{13}C NMR** (151 MHz, DMSO) δ 173.60, 169.38, 167.42, 167.30, 154.22, 148.41, 146.84, 136.89, 134.39, 132.60, 130.01, 129.12, 128.82, 128.40, 127.22, 126.01, 119.14, 118.60, 118.33, 115.88, 113.80, 63.74, 51.41, 50.41, 45.41, 38.22, 37.02, 35.53, 35.07, 30.19, 25.98, 21.83. **m/z** calculated for $C_{36}H_{42}N_8O_8$: 714.31, detected (TOF MS ES+): 715.1923.

6.8.2 Synthesis of compound 52

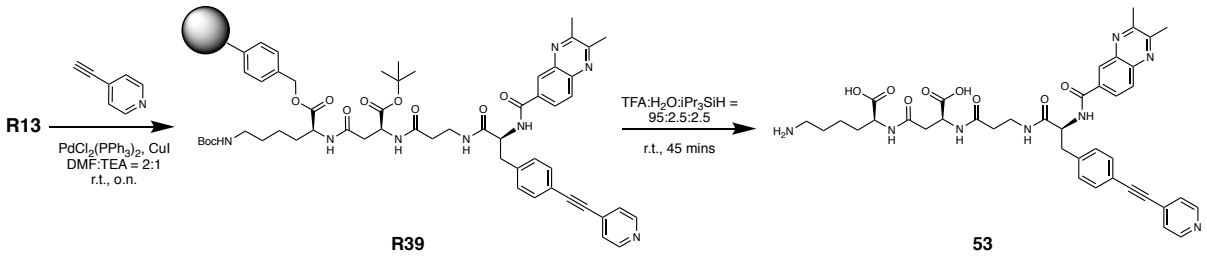


200 mg of resin **R8** (D-isomer, 100 µmol) was incubated with ImN3 (CAS: 952234-37-6, 52 mg) and converted to **R36** using the azido-transfer procedure (Supplementary information 6.2.5). The obtained resin **R36** were coupled with 4-ethynylaniline (CAS: 14235-81-5, 47 mg) by

CuAAC procedure **6.2.6**. The obtained resin **R37** was coupled with (2-hydroxyphenyl)boronic acid (CAS: 89466-08-0, 55 mg) by Suzuki cross-coupling (**6.2.8**) to **R38**. After cleavage and purification, the product **52** was obtained with 38% of yield (38 mol, 27 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 8.70 (dt, *J* = 12.3, 5.7 Hz, 1H), 8.55 – 8.46 (m, 1H), 8.17 (d, *J* = 8.1 Hz, 1H), 7.98 – 7.91 (m, 1H), 7.77 (dddd, *J* = 9.0, 7.5, 6.1, 3.4 Hz, 1H), 7.68 – 7.60 (m, 2H), 7.53 – 7.45 (m, 2H), 7.42 – 7.38 (m, 1H), 7.27 – 7.20 (m, 2H), 7.18 (dd, *J* = 7.6, 1.7 Hz, 1H), 7.11 (ddt, *J* = 9.7, 7.2, 1.7 Hz, 1H), 6.90 (dd, *J* = 8.1, 1.3 Hz, 1H), 6.82 (tt, *J* = 7.4, 1.4 Hz, 1H), 6.60 (dq, *J* = 8.5, 2.2, 1.8 Hz, 2H), 5.66 – 5.55 (m, 1H), 4.42 (q, *J* = 6.3 Hz, 2H), 4.27 – 4.09 (m, 2H), 3.05 (q, *J* = 7.3 Hz, 1H), 2.76 (pd, *J* = 8.0, 5.4, 4.9 Hz, 2H), 2.64 – 2.56 (m, 1H), 2.31 – 2.22 (m, 2H), 1.78 – 1.27 (m, 7H). **^{13}C NMR** (151 MHz, DMSO) δ 173.47, 169.59, 167.49, 154.17, 148.43, 146.85, 136.88, 135.16, 134.30, 132.69, 130.32, 130.04, 129.12, 128.43, 128.12, 127.18, 126.00, 119.18, 118.54, 115.85, 113.78, 63.74, 51.24, 45.55, 38.31, 37.24, 37.03, 35.48, 34.94, 30.25, 26.10, 21.90. **m/z** calculated for $C_{36}H_{42}N_8O_8$: 714.31, detected (TOF MS ES+): 715.1946.

6.9 Synthesis of uPA binders

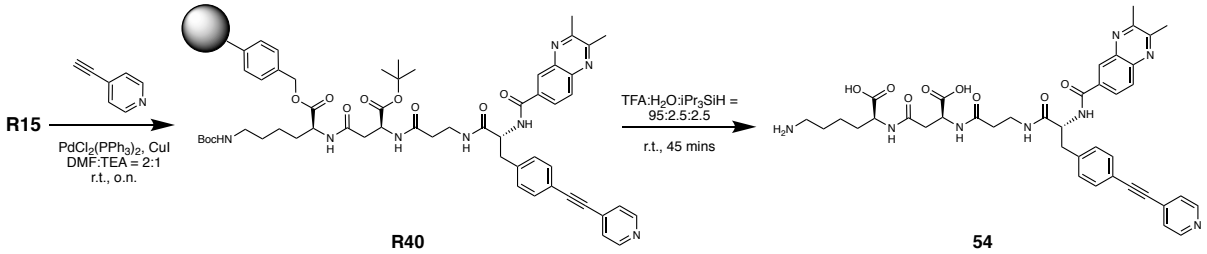
6.9.1 Synthesis of compound 53



200 mg of resin **R13** (L-isomer, 100 µmol) were coupled with 4-ethynylpyridine (CAS: 352530-29-1, 56 mg) by Sonogashira cross-coupling (**6.2.9**) to **R39**. After cleavage and purification, the product **53** was obtained with 27% of yield (27 µmol, 21 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 9.06 (d, *J* = 8.5 Hz, 1H), 8.62 – 8.56 (m, 2H), 8.50 – 8.43 (m, 1H), 8.41 (t, *J* = 5.7 Hz, 1H), 8.13 – 8.04 (m, 2H), 8.02 – 7.90 (m, 2H), 7.74 (d, *J* = 7.9 Hz, 1H), 7.66 – 7.55 (m, 1H), 7.54 – 7.40 (m, 6H), 4.76 (ddd, *J* = 11.0, 8.5, 4.3 Hz, 1H), 4.32 (dt, *J* = 7.9, 5.6 Hz, 1H), 4.19 (td, *J* = 8.7, 4.2 Hz, 1H), 3.21 (dd, *J* = 13.8, 4.2 Hz, 2H), 3.07 (dd, *J* = 13.8, 10.9 Hz, 2H), 2.76 (t, *J* = 6.7 Hz, 2H), 2.71 (s, 3H), 2.70 (s, 3H), 2.65 – 2.60 (m, 1H), 2.31 (t, *J* = 7.2 Hz, 2H), 1.77 – 1.69 (m, 1H), 1.61 – 1.43 (m, 3H), 1.37 (h, *J* = 8.6, 7.9 Hz, 2H). **^{13}C NMR** (151 MHz, DMSO) δ 173.66, 173.18, 170.84, 169.99, 169.51, 167.32, 165.38, 163.47, 155.53, 154.96, 149.77, 141.53, 140.48,

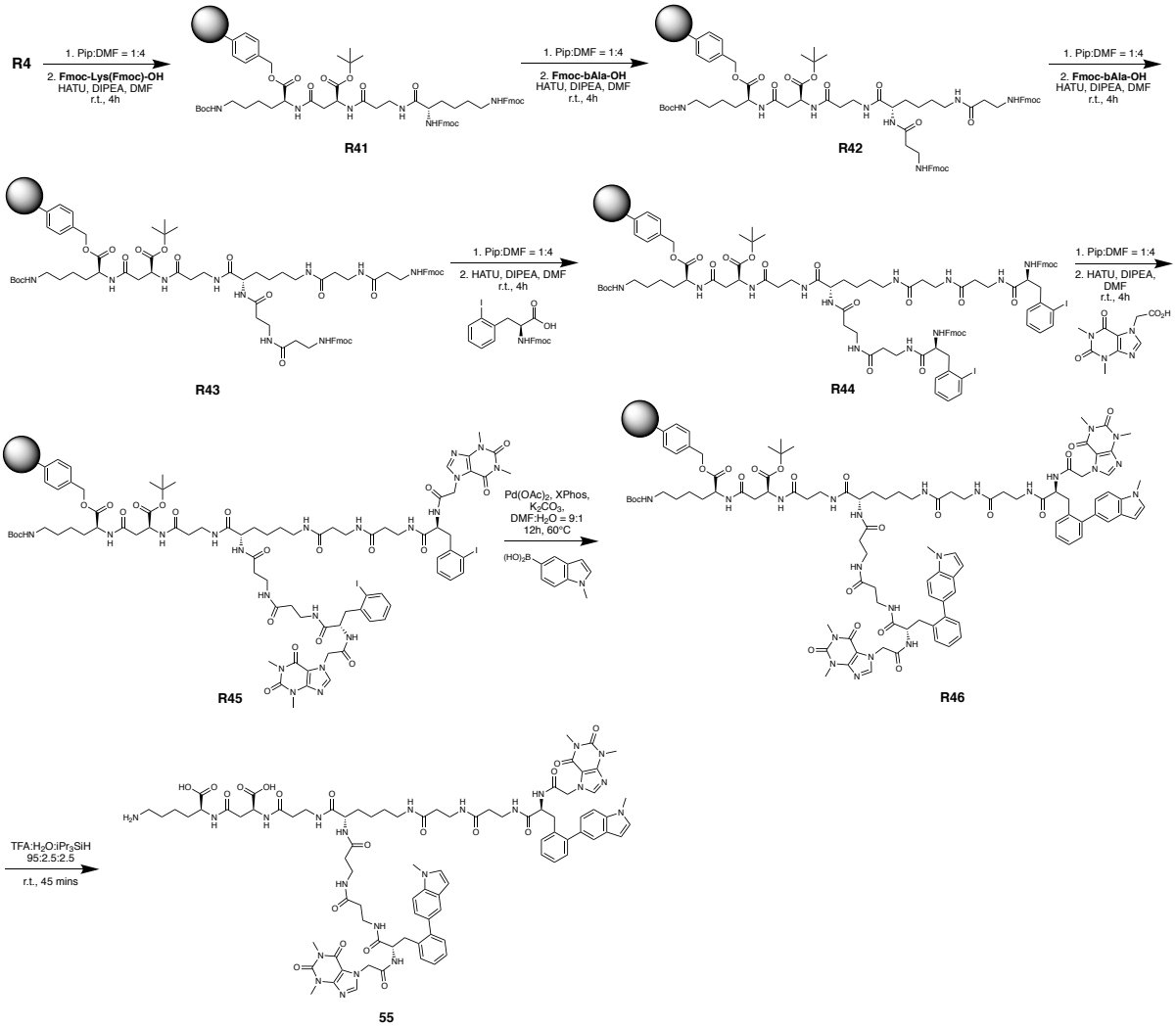
139.48, 133.71, 132.54, 131.35, 130.15, 129.62, 129.11, 128.38, 127.83, 127.37, 125.16, 118.86, 93.65, 86.33, 54.86, 51.47, 50.80, 38.15, 37.17, 35.43, 30.15, 25.97, 22.81, 22.75, 21.71. **m/z** calculated for $C_{40}H_{44}N_8O_8$: 764.33, detected (TOF MS ES+): 765.1988

6.9.2 Synthesis of compound 54



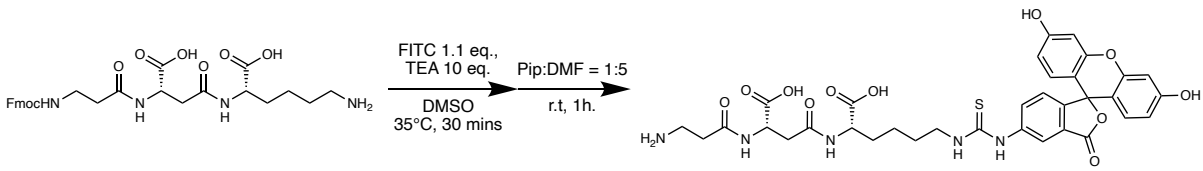
200 mg of resin **R15** (D-isomer, 100 µmol) was coupled with 4-ethynylpyridine (CAS: 352530-29-1, 56 mg) by Sonogashira cross-coupling (**6.2.9**) to **R40**. After cleavage and purification, the product **54** was obtained with 22% of yield (22 µmol, 17 mg). **^{1}H NMR** (600 MHz, DMSO-d6) δ 9.16 (d, *J* = 8.4 Hz, 1H), 8.59 (d, *J* = 4.9 Hz, 2H), 8.53 – 8.43 (m, 2H), 8.13 – 8.05 (m, 2H), 8.01 – 7.92 (m, 2H), 7.73 (d, *J* = 7.7 Hz, 1H), 7.53 – 7.43 (m, 6H), 4.75 (ddd, *J* = 10.9, 8.5, 4.2 Hz, 2H), 4.28 (q, *J* = 6.0 Hz, 1H), 4.17 (td, *J* = 8.6, 4.3 Hz, 1H), 3.19 (dd, *J* = 13.9, 4.3 Hz, 2H), 3.11 – 3.05 (m, 2H), 2.75 (p, *J* = 5.6 Hz, 2H), 2.71 (s, 3H), 2.70 (s, 3H), 2.65 – 2.58 (m, 1H), 2.31 (t, *J* = 6.9 Hz, 2H), 1.78 – 1.68 (m, 1H), 1.60 – 1.31 (m, 6H). **^{13}C NMR** (151 MHz, DMSO) δ 173.74, 173.08, 170.88, 169.95, 169.57, 165.38, 155.49, 154.92, 149.78, 141.53, 140.56, 139.51, 133.76, 132.59, 131.34, 130.14, 129.64, 129.12, 128.39, 127.79, 127.45, 127.28, 125.16, 118.85, 93.66, 86.32, 54.97, 51.54, 50.91, 38.10, 37.15, 35.51, 30.12, 25.97, 22.77, 21.69. **m/z** calculated for $C_{40}H_{44}N_8O_8$: 764.33, detected (TOF MS ES+): 765.2083

6.10 Synthesis of A481/B335 dimer (TNC binder, compound 55)



200 mg of resin **R4** (100 µmol) was deprotected and a dimeric peptide was assembled using the general procedures (supplementary information **6.2.1** and **6.2.2**) with following sequence: Fmoc-L-Lys(Fmoc)-OH (CAS: 78081-87-5, 236 mg), Fmoc-beta-alanine-OH (CAS: 35737-10-1, 8 eq., 250 mg), Fmoc-beta-alanine-OH (8 eq., 250 mg), Fmoc-2-iodo-(L)-Phe-OH (CAS: 210282-32-9, 8 eq., 410 mg). The obtained resin **R44** was subsequently deprotected and coupled with theophiline-7-acetic acid (CAS: 652-37-9, 8 eq., 190 mg) according with the procedure 6.2.3. **R45** was converted to **R46** by Suzuki cross coupling reaction with (1-methyl-1H-indol-5-yl)boronic acid acid (CAS: 192182-55-1, 8 eq., 140 mg) carried out at 60 °C for 12 hours. Finally, the resin **R46** was cleaved and product **55** was purified by RP-HPLC. Yield = 1% (1.1 µmol, 2 mg). **m/z** calculated for $C_{85}H_{104}N_{22}O_{19}$: 1736.78, detected (TOF MS ES+): 1737.7246, 1738.7219.

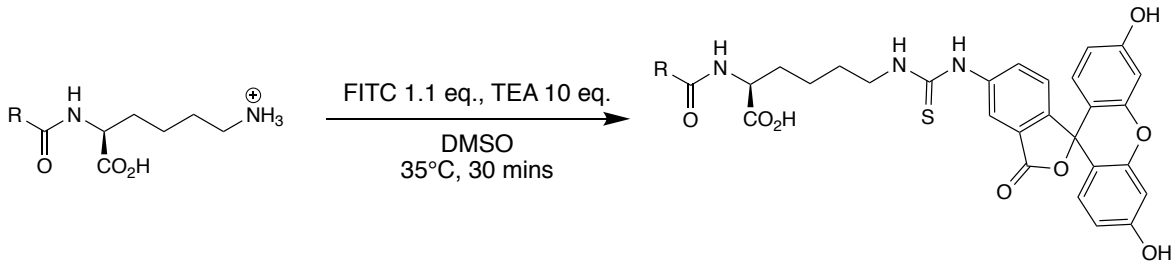
6.11 Synthesis of FITC-labelled negative control (R-NH_2)



39 **56** (H-LFluo)

5.5 mg of compound **39** (10 µmol) was dissolved in 500 µL of dry DMSO and **5-FITC** (CAS: 3326-32-7 4.3 mg, 11 µmol) and TEA (14 µL, 0.1 mmol) were added. The reaction was heated at 35°C for 30 minutes. To the crude reaction piperidine:DMF = 1:5 solution (1 mL) was added and the deprotection was allowed for 1 hour at room temperature. The reaction was concentrated and the pure product **56** was obtained by RP-HPLC (mobile-phase: H_2O:acetonitrile 1% FA from 95:5 to 0:100 in 20 minutes). Compound **56** (H-LFluo) was lyophilized and characterized by LC-MS. Yield = 62 % (6.2 µmol, 4.5 mg).

6.12 Synthesis of FITC-labelled binders



The "Free amino" Lysine derivatives (starting material, **Table 4**) were dissolved in dry DMSO to a final concentration of 10 mM solution. To 100 µL of each binder solution (1 µmol) were added 110 µL of FITC solution (10 mM in DMSO, 3.9 mg/mL) and 1.4 µL of TEA (10 µmol). The reaction was heated at 35°C for 30 minutes. The crude was quenched with formic acid (FA, 1.5 µL, 40 µmol), diluted with H_2O:acetonitrile=1:1 (600 µL) and directly purified by RP-HPLC (mobile-phase: H_2O:acetonitrile 1% FA from 95:5 to 0:100 in 20 minutes). RT(FITC) = 15 mins, RT(FITC-conj) = 10-13 mins. The eluted products were lyophilized and characterized by LC-MS. The obtained products (**Table 4** and **5**) were re-dissolved in the protein buffer (2.00 mL, **Table 4**) and the concentrations were determined by UV-Vis spectrophotometry [the values of ε (λ=498nm) are reported in **Table 4**] after a sample dilution of 1:5.

6.13 Determination of ε-value for FITC-labelled binders

3.6 mg of H-LFluo (**56**) were dissolved in 5 mL of mQ millipore water 2% DMSO (conc = 1 mM) and diluted to 400 µM in PBS, HEPES1, HEPES2 and TRIS buffer. The obtained samples were serially diluted to 200 µM, 133 µM, 80 µM, 44 µM and 15 µM. The absorbance was measured at λ=498 nM by nanodrop (*l*=0.1 cm). The equation of the slope (Abs vs [H-LFluo]) and extinction coefficients (ε) were determined by PRISM (Table 4).

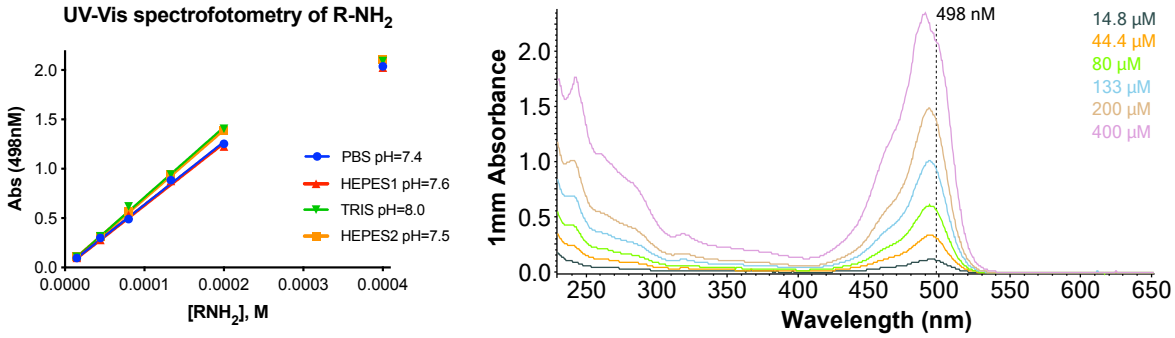
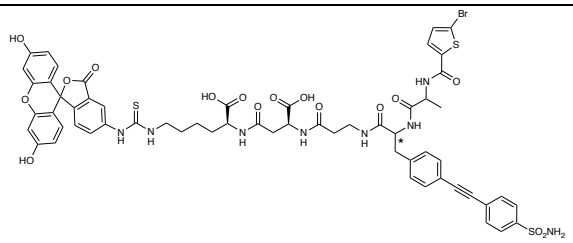
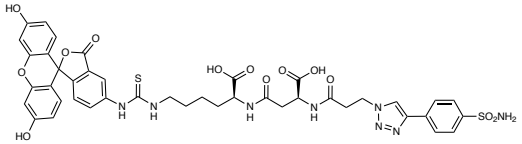
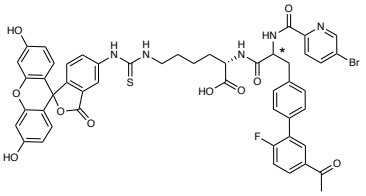
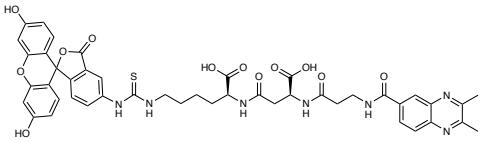
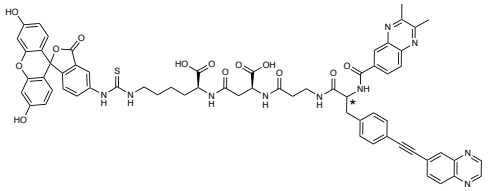
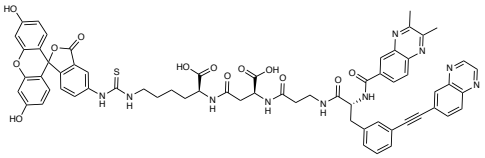
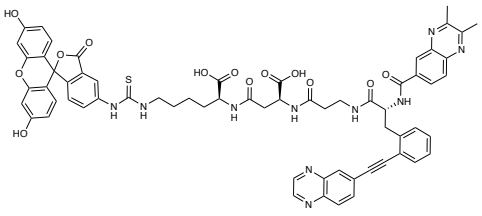
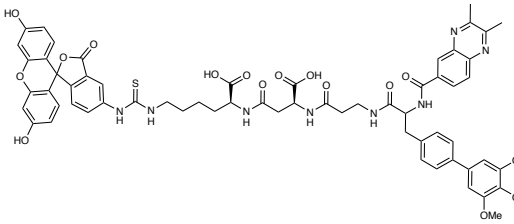
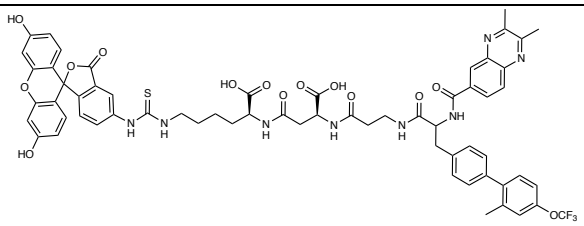
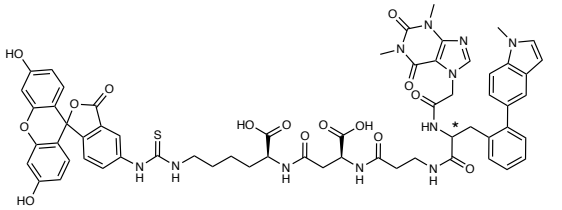
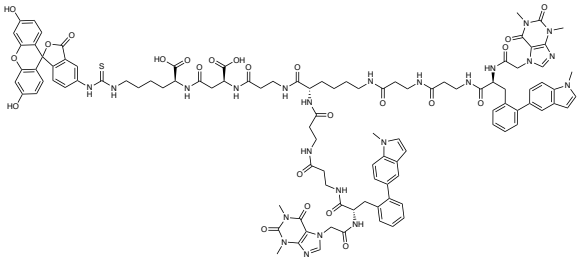
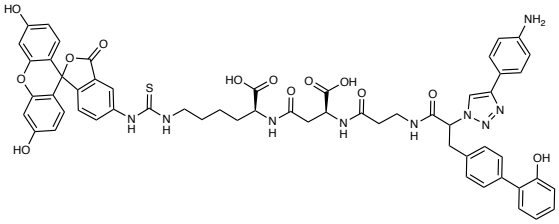
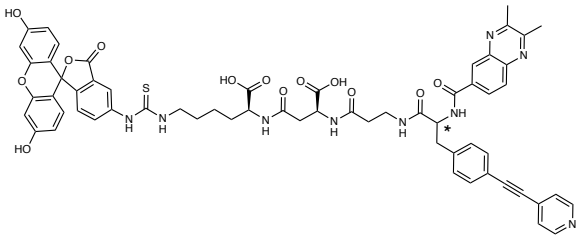
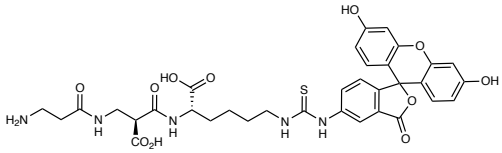


Table 4: Determination of concentration and reaction yields of FITC-labelled derivatives by UV-Vis spectrophotometry. S.M. : starting material.

S.M.	Compound	Target	ε (M^{-1} cm^{-1})	5× Abs (λ=498nm)	[Ligand], µM	Buffer	pH	Yield
7	**8**	CAIX	(6.37±0.09)·10⁴	1.576	247±3	PBS	7.4	49%
9	**10**	CAIX	(6.37±0.09)·10⁴	2.045	321±5	PBS	7.4	64%
11	**12**	CAIX	(6.37±0.09)·10⁴	1.23	193±1	PBS	7.4	39%
34	**13**	CREBBP	(6.3±0.1)·10⁴	1.47	233±4	HEPES1	7.6	47%
38	**14**	CREBBP	(6.3±0.1)·10⁴	2.106	334±5	HEPES1	7.6	67%
40	**16**	PI3K	(6.98±0.04)·10⁴	1.268	182±1	HEPES2	7.5	36%
41	**17**	PI3K	(6.98±0.04)·10⁴	1.234	177±1	HEPES2	7.5	35%
42	**18**	PI3K	(6.98±0.04)·10⁴	1.422	204±1	HEPES2	7.5	41%
43	**19**	PI3K	(6.98±0.04)·10⁴	1.931	277±2	HEPES2	7.5	55%
44	**15**	PI3K	(6.98±0.04)·10⁴	1.764	253±1	HEPES2	7.5	51%
45	**20**	H1047R-PI3K	(6.98±0.04)·10⁴	1.129	162±1	HEPES2	7.5	32%
46	**21**	H1047R-PI3K	(6.98±0.04)·10⁴	2.25	322±2	HEPES2	7.5	64%
47	**22**	H1047R-PI3K	(6.98±0.04)·10⁴	1.638	235±1	HEPES2	7.5	47%
48	**23**	H1047R-PI3K	(6.98±0.04)·10⁴	1.589	228±1	HEPES2	7.5	46%
49	**24**	m/h-TNC	(6.37±0.09)·10⁴	2.05	320±4	PBS	7.4	64%
50	**25**	m/h-TNC	(6.37±0.09)·10⁴	2.113	332±4	PBS	7.4	66%
55	**26**	h-TNC	(6.37±0.09)·10⁴	1.524	239±2	PBS	7.4	47.8
51	**27**	L27E-CtIP	(7.1±0.1)·10⁴	2.015	284±4	TRIS	8.0	57%
52	**28**	L27E-CtIP	(7.1±0.1)·10⁴	1.95	275±4	TRIS	8.0	55%
53	**29**	UPA	(6.37±0.09)·10⁴	1.18	185±3	PBS	7.4	37%
54	**30**	UPA	(6.37±0.09)·10⁴	1.26	198±3	PBS	7.4	40%
39	**56** (H-LFluo)	Neg. contr.			1'000	H_2O, 2%DMSO		

Table 5: MS characterization of FITC labelled derivatives. [a] Stereochemistry of phenylalanine. [b] Exact Mass. [c] Mass (m/z) detected by TOF MS ES- (negative mode).

Compound (*)[a]	Structure	Formula	Calc. MS [b]	Found MS [c]
8 (S)		$C_{59}H_{55}BrN_8O_{16}S_3$	1306.21	1305.2665 1307.3420
10 (R)		$C_{59}H_{55}BrN_8O_{16}S_3$	1306.21	1305.2524 1307.3276
12		$C_{42}H_{40}N_8O_{13}S_2$	928.22	927.1768
13 (S)		$C_{50}H_{41}BrFN_5O_{10}S$	1001.17	1000.2838 1002.2523
14 (R)		$C_{50}H_{41}BrFN_5O_{10}S$	1001.17	1000.1766 1002.1793
15		$C_{45}H_{43}N_7O_{12}S$	905.27	904.2667
16 (S)		$C_{64}H_{56}N_{10}O_{13}S$	1204.37	1203.4991
17 (R)		$C_{64}H_{56}N_{10}O_{13}S$	1204.37	1203.4517
18 (R)		$C_{64}H_{56}N_{10}O_{13}S$	1204.37	1203.3741
19 (R)		$C_{64}H_{56}N_{10}O_{13}S$	1204.37	1201.4917
20 (S)		$C_{63}H_{62}N_8O_{16}S$	1218.40	1217.3973
21 (R)		$C_{63}H_{62}N_8O_{16}S$	1218.40	1217.3934

Compound (*)[a]	Structure	Formula	Calc. MS [b]	Found Ms
22 (S)		$C_{62}H_{57}F_3N_8O_{14}S$	1226.37	1225.5010
23 (R)		$C_{62}H_{57}F_3N_8O_{14}S$	1226.37	1225.3918
24 (S)		$C_{61}H_{59}N_{11}O_{15}S$	1217.39	1216.3901
25 (R)		$C_{61}H_{59}N_{11}O_{15}S$	1217.39	1216.5350
26 (S,S)		$C_{106}H_{115}N_{23}O_{24}S$	2125.82	2125.6755
27 (S)		$C_{57}H_{53}N_9O_{13}S$	1103.35	1102.4734
28 (R)		$C_{57}H_{53}N_9O_{13}S$	1103.35	1202.4735
29 (S)		$C_{61}H_{55}N_9O_{13}S$	1153.36	1152.4971
30 (R)		$C_{61}H_{55}N_9O_{13}S$	1153.36	1152.4985
56 (H-LFluo)		$C_{34}H_{35}N_5O_{11}S$	721.21	720.2874

7. Hit Validation

7.1 Fluorescence polarization

The FITC-labelled compounds were diluted to a final concentration of 50 nM (compounds **8**, **10** and **12** were diluted to 5 nM) and incubated (5µL) for 15 minutes in a black 384-well plate (Greiner small-volume, non-binding) with serial dilutions of protein (5µL each) to a final volume of 10 µL. The fluorescence anisotropy was measured at 535 nm on a Spectra Max Paradigm multimode plate reader (Molecular Devices). The experiments were performed in triplicate and the resulting data was statistically evaluated by Prism 8. The K_d values were obtained by fitting target concentration vs anisotropy using [Inhibitor] vs. response, Variable slope four parameters prism equation. The obtained K_d are reported **in the Table 6**.

7.2 Enzyme-linked immunosorbent assay (ELISA)

The protein (100 µL/well, 200 nM) was incubated overnight at 4°C on a F8 maxisorp (Thermo Scientific) plate. The protein was blocked by adding 4% Milk in PBS (200 µL/well, 30 min at RT) than washed with PB (3x, 200 µL/well). The immobilized protein was incubated for 30 minutes in the dark with serial dilutions of FITC-labelled compound than washed with PB (3x, 200 µL/well). A solution of 200 nM anti-fluorescein antibody (IgG1 human)[11] in 2% Milk-PB was added to each well (100 µL/well) and incubated for additional 30 minutes in the dark. The resulting complex was washed with PB (3x, 200 µL/well) and incubated for additional 30 minutes of protein A-HRP (1µg/mL in 2% Milk-PB, 100µL/well). Each well was washed with PB 0.1% Tween (3x, 200 µL/ well) and with PB (3x, 200 uL/ well). The substrate (**TMB**) was added (100 µL /well) and developed in the dark for 1-5 minutes. The reaction was stopped by adding 50 µL of 1 M sulphuric acid. The absorbance was measured on a Spectra Max Paradigm multimode plate reader (Molecular Devices) at 620-650 nm and 450 nm. The experiments were performed in triplicate and the resulting data was statistically evaluated by Prism 8. The K_d values were obtained by fitting target concentration vs anisotropy using [Inhibitor] vs. response, Variable slope four parameters prism equation. The obtained K_d are reported in the

Table 6.

Table 6: dissociation constants (K_d) values of compounds 8,10,12-30 determined by fluorescence polarization (FP) or ELISA.

Compound ID	Target	Validation assay	K_d
8	CAIX	FP	7.2 ± 0.3 nM
10	CAIX	FP	8.8 ± 0.3 nM
12	CAIX	FP	68 ± 3 nM
13	CREBBP	FP	0.92 ± 0.06 nM
14	CREBBP	FP	6.0 ± 0.7 µM
13	BDR4	FP	>75 µM
14	BDR4	FP	>75 µM
13	BPTF	FP	>100 µM
14	BPTF	FP	>100 µM
15	wt-PI3K	FP	> 3.0 µM
16	wt-PI3K	FP	306 ± 8 nM
17	wt-PI3K	FP	126 ± 2 nM
18	PI3K	FP	> 3.0 µM
19	PI3K	FP	> 3.0 µM
20	PI3K	ELISA	18 ± 3 µM
21	PI3K	ELISA	15.7 ± 0.8 µM
22	PI3K	ELISA	3.5 ± 0.7 µM
23	PI3K	ELISA	2.9 ± 0.4 µM
20	H1047R-PI3K	ELISA	3.6 ± 0.4 µM
21	H1047R-PI3K	ELISA	3.0 ± 0.4 µM
22	H1047R-PI3K	ELISA	0.19 ± 0.02 µM
23	H1047R-PI3K	ELISA	0.37 ± 0.06 µM
24	hTNC	ELISA	40 ± 6 µM
25	hTNC	ELISA	70 ± 10 µM
26	hTNC	ELISA	1.7 ± 0.1 µM
24	mTNC	ELISA	40 ± 10 µM
25	mTNC	ELISA	120 ± 10 µM
27	L27E-CtIP	ELISA	10.9 ± 0.9 µM
28	L27E-CtIP	ELISA	50 ± 10 µM
29	uPA	ELISA	>150 µM
30	uPA	ELISA	>150 µM

7.2.1 CAIX

FITC-labelled compounds **8**, **10**, **AAZ-FITC** conjugate (positive control) and **56** (**L-NHFluo**, negative control) were serially diluted from 1.0 µM to 12.8 pM (dilution 1:5, 8 dilutions) in PBS (100 µL each).

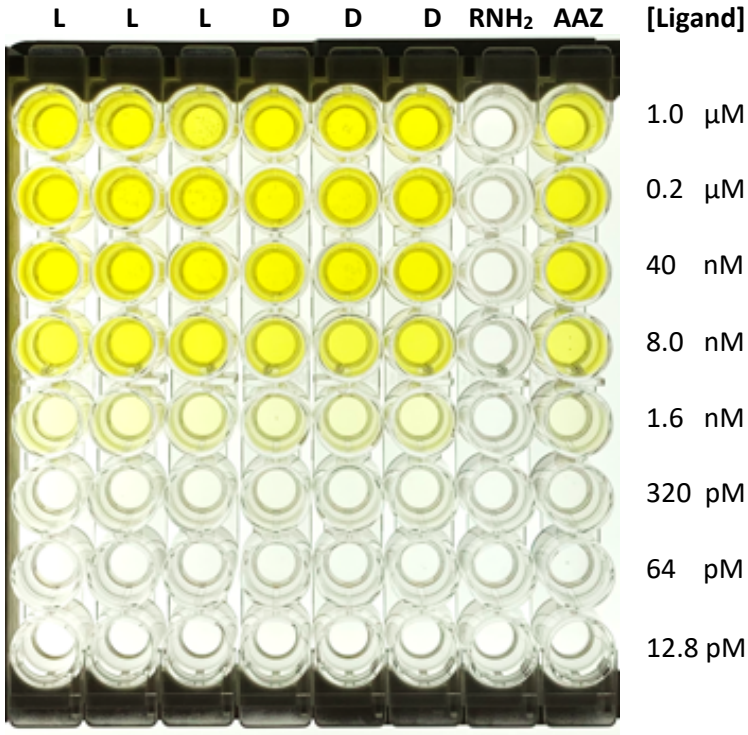
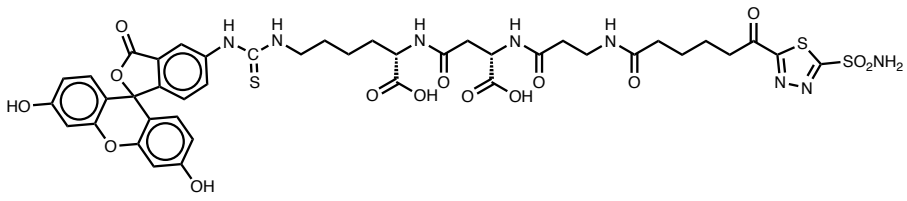


Figure 26: **L**: compound **8**; **D**: compound **10**; **RNH_2**: compound **56**, **AAZ**: positive control (acetazolamide-FITC conjugate).



AAZ-FITC conjugate

7.2.2 CREBBP

FITC-labelled compounds **13**, **14** and **56** (**H-LFluo**, negative control) were serially diluted from 150 µM to 9.16 nM (dilution 1:4, 8 dilutions) in HEPES1 (100 µL each). See **Figure 3c**.

7.2.3 PI3K (wt and H1047R-PI3K)

FITC-labelled compounds **20**, **21**, **22**, **23** and **56** (**H-LFLuo**, negative control) were serially diluted from 150 µM to 1.9 nM (dilution 1:5, 8 dilutions) in HEPES2 (100 µL each) and incubated with immobilized wt-PI3K and H1047R-PI3K. The experiment was repeated in triplicate against immobilized H1047R-PI3K. Compounds **20-23** were serially diluted from 100 µM to 1.3 nM (dilution 1:5, 8 dilutions).

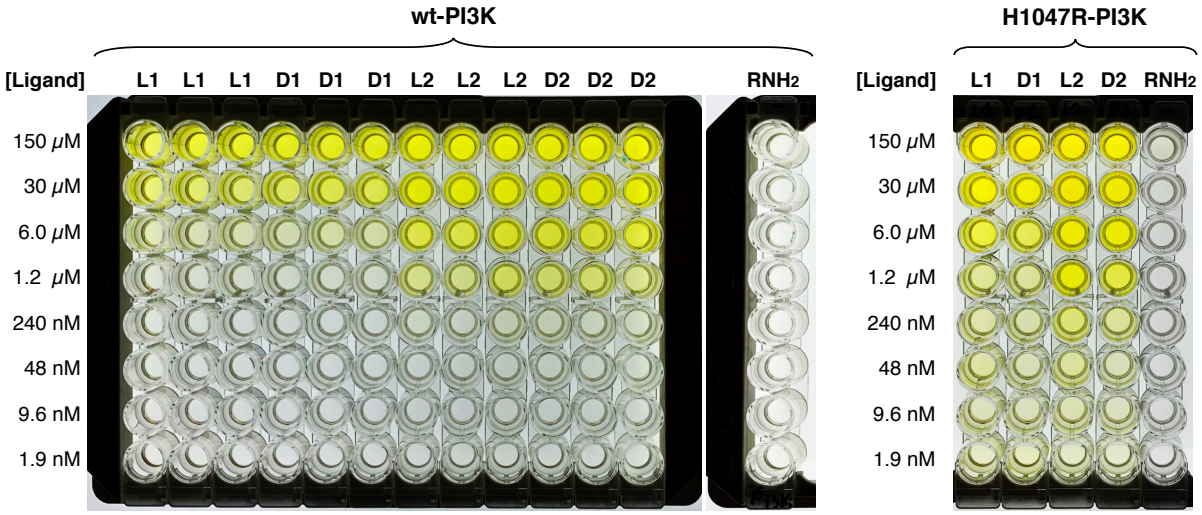


Figure 28: ELISA against immobilized a) wt-PI3K (triplicate) compared with b) H1047R-PI3K. **L1**: compound **20**; **D1**: compound **21**; **L2**: compound **22**; **D1**: compound **23**; **RNH_2**: compound **56**.

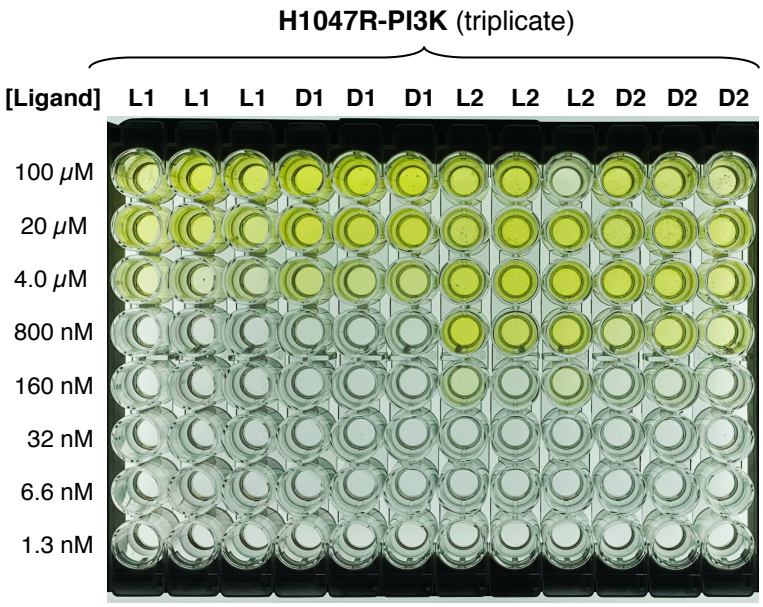


Figure 29: ELISA against immobilized H1047R-PI3K (triplicate). **L1**: compound **20**; **D1**: compound **21**; **L2**: compound **22**; **D1**: compound **23**; **RNH_2**: compound **56**.

7.2.4 Tenascin C

FITC-labelled compounds **24**, **25** and **56** (**H-LFluo**, negative control) were serially diluted from 150 µM to 1.17 µM (dilution 1:2, 8 dilutions) in PBS (100 µL each).

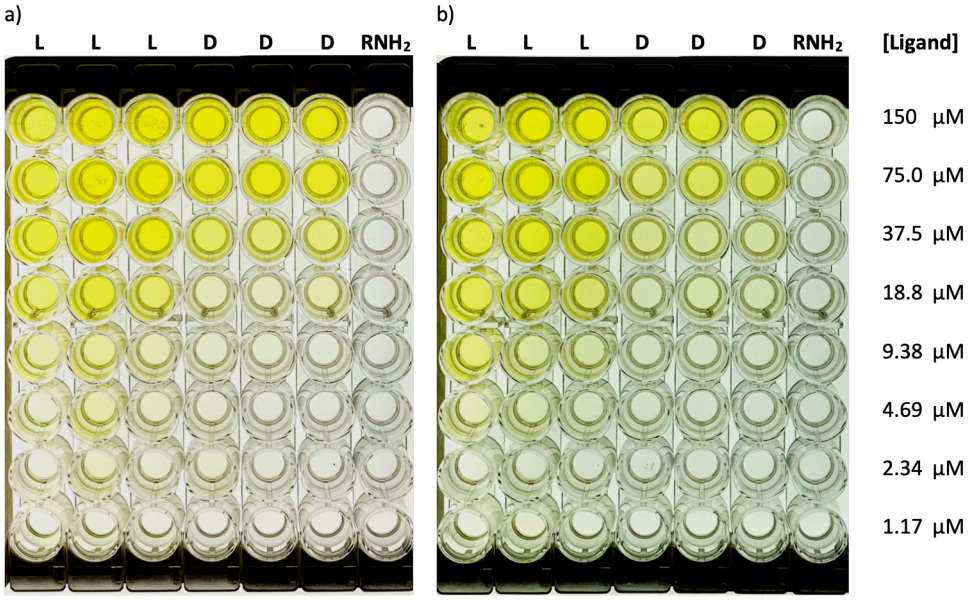


Figure 30: ELISA against immobilized a) human tenascin C (h-TNC) and b) murine tenascin C (m-TNC). **L**: compound **24**; **D**: compound **25**; **RNH_2**: compound **56**.

FITC-labelled compound **26** was serially diluted from 200 µM to 91 nM (dilution 1:3, 8 dilutions) in PBS (100 µL each). Compound **24** was concentrate up to 400 µM and serially diluted to 0.1 µM (dilution 1:3, 8 dilutions) in PBS (100 µL each).

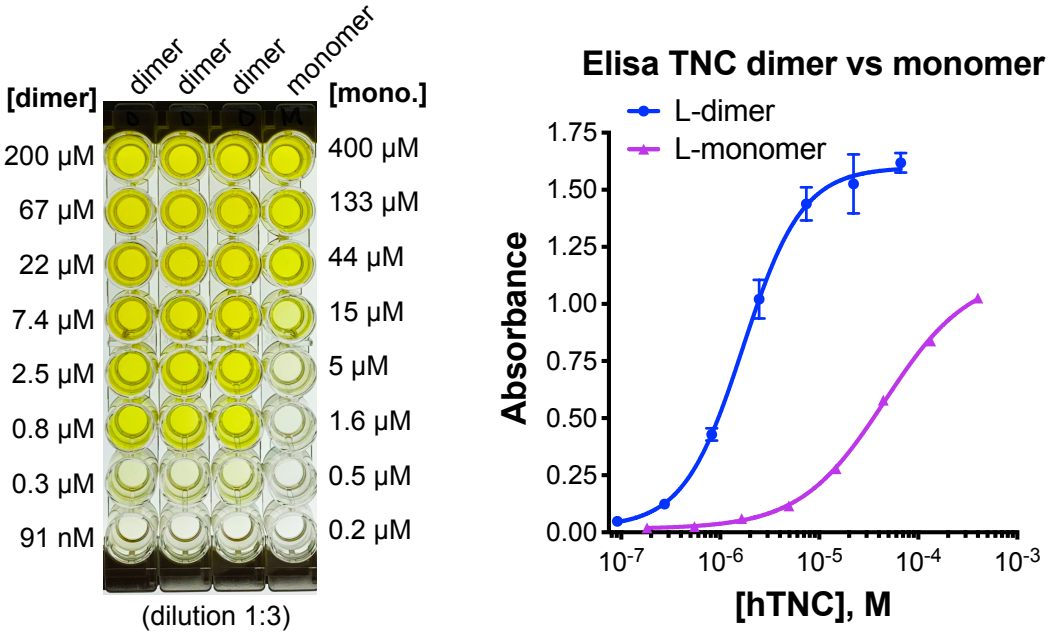


Figure 31: ELISA against immobilized human tenascin C (h-TNC) **dimer**: compound **26**; **monomer**: compound **24** (L-isomer).

7.2.5 L27E-CtIP

FITC-labelled compounds **27**, **28** and **56** (**H-LFluo**, negative control) were serially diluted from 100 µM to 46 nM (dilution 1:3, 8 dilutions) in TRIS buffer (100 µL each).

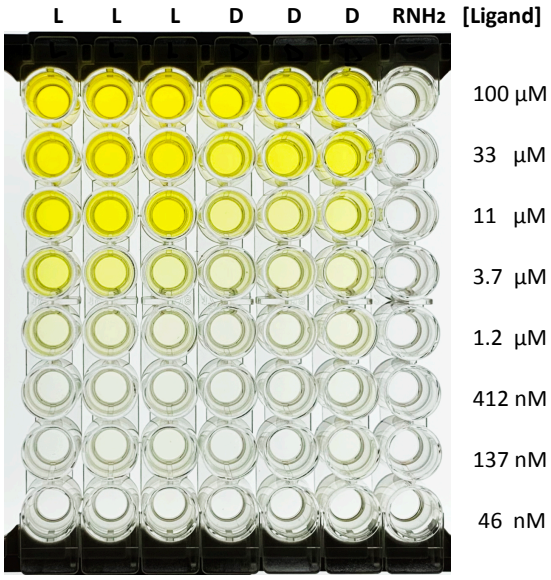


Figure 32: ELISA against immobilized L27E-CtIP. **L**: compound **27**; **D**: compound **28**; **RNH_2**: compound **56**.

7.2.6 uPA

FITC-labelled compounds **29**, **30** and **56** (**H-LFluo**, negative control) were serially diluted from 150 µM to 9.16 nM (dilution 1:4, 8 dilutions) in PBS (100 µL each).

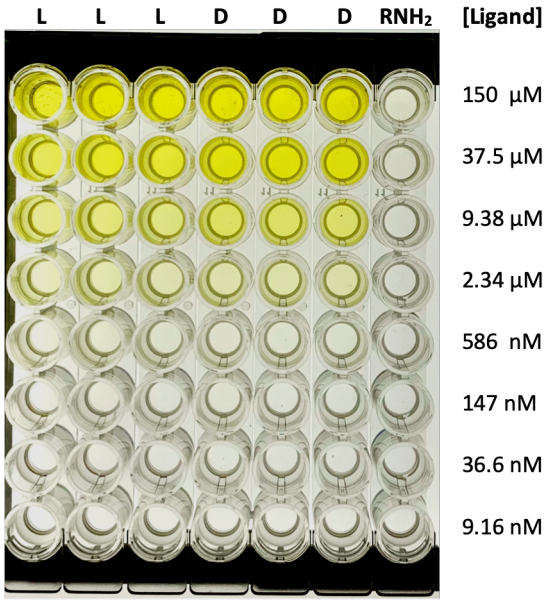


Figure 33: ELISA against immobilized urokinase-plasminogen activator (uPA). **L**: compound **29**; **D**: compound **30**; **RNH_2**: compound **56**.

7.3 Flow cytometry analysis

Cells were detached from culture plates using Accutase cell detachment solution (MERCK, cat. A6964), counted and suspended to a final concentration of 1.5×10^6 cell/100 µL in a solution of FACS buffer (1% bovine serum albumin, 2 mM EDTA, in 500 mL PBS pH 7.4). Aliquots of 3×10^5 cells (200 µL) were spun down and resuspended in solutions of compound **8** and compound **10** (200µL; 50 nM, 10 nM and 2 nM) in FACS buffer and incubated on ice for 1 h. Cells were washed once with 200 µl FACS (1% v/v) / PBS (pH 7.4), spun down, resuspended in a solution of FACS buffer (200 µL) and analyzed via a 2L CytoFLEX Flow Cytometer (Beckman Coulter). FlowJo Version 8.7 (Treestar) was used for the data analysis and visualization.

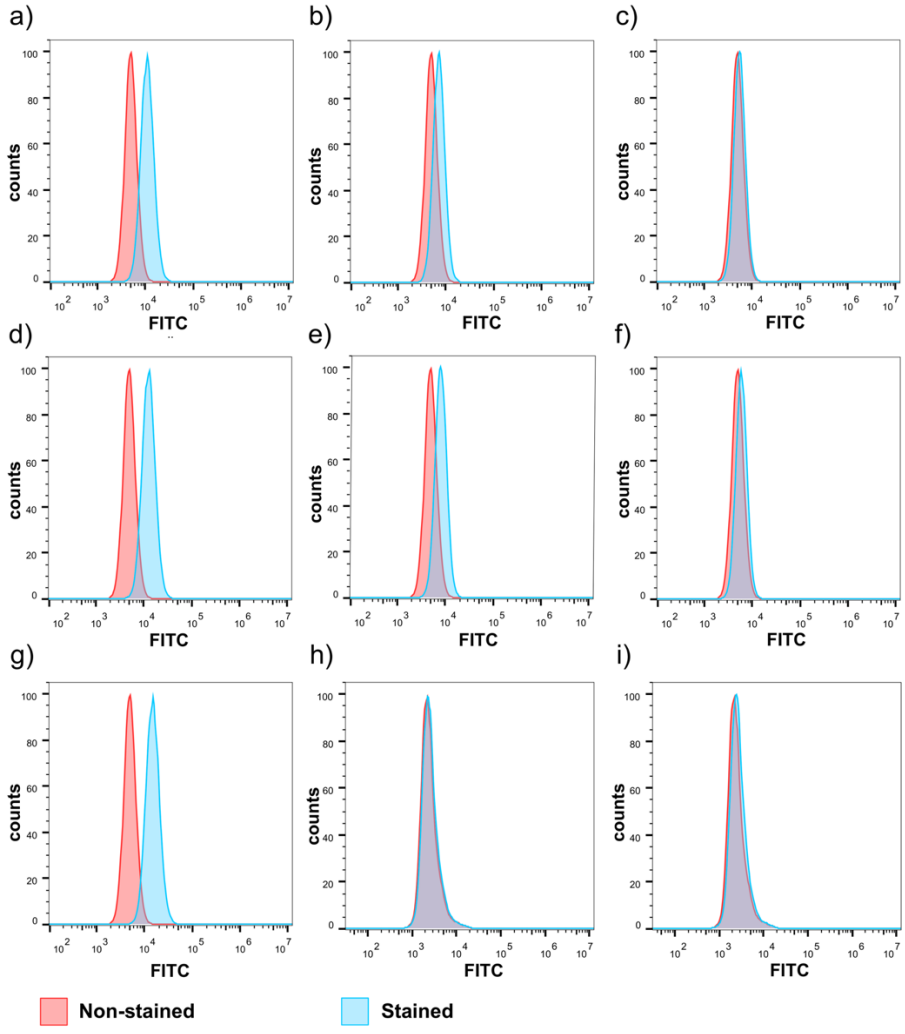


Figure 34: Flow cytometry of a) 50 nM (R)-Phe/A160/B475 on SK-RC-52 cells; b) 10 nM (R)-Phe/A160/B475 on SK-RC-52 cells; c) 2 nM (R)-Phe/A160/B475 on SK-RC-52 cells; d) 50 nM (S)-Phe/A160/B475 on SK-RC-52 cells; e) 10 nM (S)-Phe/A160/B475 on SK-RC-52 cells; f) 2 nM (S)-Phe/A160/B475 on SK-RC-52 cells; g) 100 nM AAZ-Fluorescein on SK-RC-52 cells; h) 50nM (R)-Phe/A160/B475 on HEK293; i) 50nM (S)-Phe/A160/B475 on HEK293. Non-stained: cells (SK-RC-52 or HEK293) without FITC-labelled compound; Stained: cells (SK-RC-52 or HEK293) with FITC-labelled compound.

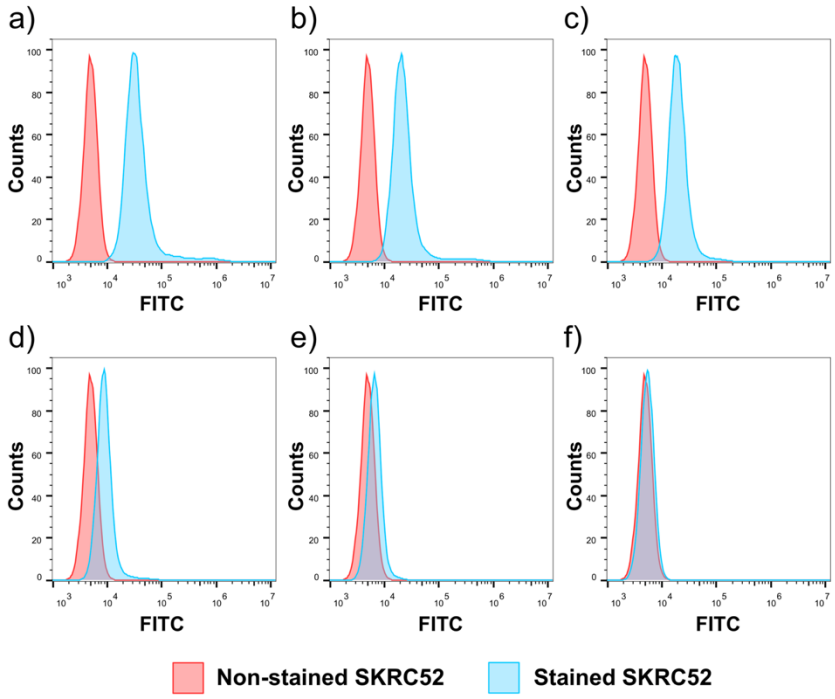


Figure 35: Flow cytometry on SK-RC-52 cells of compound **12** with concentrations of a) 10 µM, b) 5.0 µM, c) 1.0 µM, d) 500 nM, e) 100 nM and f) 10 nM .

7.3.1 Fluorescence enhancement method

Flow cytometry analysis was carried out as previously described in **7.3**. After the incubation of compounds **8** and **10** (200µL, 100 nM), a secondary anti-FITC Rabbit IgG (Ref#4510-7804; BioRad) was added at a concentration of 1:700 from stock solution and incubated for 1 hr on ice. After the washing step, to remove the excess anti-FITC rabbit IgG, the tertiary antibody goat anti-rabbit IgG conjugated to alexa fluo 488 was added to amplify the signal. An additional washing step was implemented, to remove the excess prior to the analysis. The sample was analysed by 2L CytoFLEX Flow Cytometer (Beckman Coulter). FlowJo Version 8.7 (Treestar) was used for the data analysis and visualization (**Figure 2c**).

7.4 UniCAR-T Killing assay

Universal CAR-T cells were produced as previously described[12]. UniCAR T-cells, αCAIX CAR T-cells and non-transduced T cells were thawed and grown in complete (10% Fetal Bovine Serum, 2 mM ultraglutamine and 1% antibiotic-antimycoticum (Gibco, #15240062)) Advanced RPMI (Gibco) without IL-2 (Proleukin, Roche Diagnostics) at a density of 1 x 10^6 cells per mL. On day 0, SK-RC-52 cells were harvested and membrane stained using PKH26 Red

Fluorescent Cell Linker Kit for General Membrane (Sigma-Aldrich) following manufacturer instructions. After the staining procedure was completed, the stained SK-RC-52 cells were seeded at a density of 30'000 cells per well, in a 96 well plate and incubated overnight (37°C, 5% CO_2). The next day, SK-RC-52 cells in extra wells were detached and counted. T-cells or UniCAR-T were resuspended in 100 µL complete Advanced RPMI (Gibco) containing different concentrations of bispecific adapters and added on the SK-RC-52 cells in a 1:1 target to effector cell ratio. Anti-CAIX CAR-Ts and non-transduced T-cells served as a positive and negative control respectively. After addition, the plate was spun down (400 g, 1 min, RT) and incubated for 24 h (37°C, 5% CO_2). The next day, the supernatant was transferred to a round bottom 96 well plate. After washing the wells with 100 µL of PBS, the PBS was also collected. Then, 50 µL Accutase (Millipore) was added to each well and incubated for 5 minutes at 37°C to detach the target cells. The detached cells were added to the corresponding well of the round bottom 96 well plate. The plate was spun down (400 g, 5 min, RT), energetically flicked to remove the supernatant and the pellets were resuspended in 150 µL FACS buffer. After 30 min of incubation with the FACS buffer at 4°C in the dark, the cells were spun down (400 g, 5 min, 4°C). Immediately before measurement, the pellets were resuspended in a 1:10'000 dilution of the live/dead staining TOTO-3 Iodide (ThermoFisher) in FACs buffer, strained (30 µm nylon mesh) and analyzed via flow cytometry (Cytoflex, Beckman Coulter). The flow cytometry data was analyzed using FlowJo software (Treestar).

7.5 Ex-Vivo

All animal experiments were conducted in accordance with Swiss animal welfare laws and regulations under the license number 04/2018 granted by the Veterinäramt des Kantons Zürich. The ex-vivo experiment was performed in athymic BALB/c *nu/nu* mice (8-10 weeks of age, Janvier) bearing subcutaneous SK-RC-52 tumor in the right flank. 50 nmol of Compound **8** were injected in tumor bearing mice and the animals were sacrificed by asphyxiation 1 hour after. Organs and tumors were extracted, flash-frozen and then cut into sections of 10 µm width. A proper staining was performed as described in the manuscript.

8. Appendix

8.1 List of building blocks A

Table 7: List of building blocks A (alkynes and carboxylic acids) and oligonucleotide codes A. (a) Position of iodine on the aromatic ring of phenylalanine.

CdId	Structure (SMILES)	Iodo[a]	Sequence	codons
1	OC(=O)C1=CC(=CN=C1)C#C	p	CTGTGTGCTGGCCTCGCGAGTCCCATGGCGC	GCCTCG
2	BrC1=NC=C(OCC#C)C=C1	p	CTGTGTGCTGTCCGACCGAGTCCCATGGCGC	TCCGAC
3	CNC1=CC=C(OCC#C)C=C1	p	CTGTGTGCTGCAAGTGCGAGTCCCATGGCGC	CAAGTG
4	NC(=O)C1=CC(=CN=C1)C#C	p	CTGTGTGCTGGTCCGCCGAGTCCCATGGCGC	GTCCGC
5	O=C(NCC#C)NC1CC1	p	CTGTGTGCTGGACGACCGAGTCCCATGGCGC	GACGAC
6	Cl.NCCC(O)CCC#C	p	CTGTGTGCTGTTATAGCGAGTCCCATGGCGC	TTATAG
7	OC1=CC(OCC#C)=CC=C1	p	CTGTGTGCTGCCGAAGCGAGTCCCATGGCGC	CCGAAG
8	C#CCN1C=CC2=C1C=CC=C2	p	CTGTGTGCTGGAACCACGAGTCCCATGGCGC	GAACCA
9	FC(F)(F)C1=NC(OCC#C)=CC=C1	p	CTGTGTGCTGAGAGAACGAGTCCCATGGCGC	AGAGAA
10	CC1=NC(=CC=C1)C#C	p	CTGTGTGCTGCATGAGCGAGTCCCATGGCGC	CATGAG
11	Cl.NC(=N)NCC#C	p	CTGTGTGCTGACATTACGAGTCCCATGGCGC	ACATTA
12	Cl.C#CCN1C=CN=C1C1=CC=CS1	p	CTGTGTGCTGGGAATCCGAGTCCCATGGCGC	GGAATC
13	OC(=O)CC(=O)NCC#C	p	CTGTGTGCTGGCCAACCGAGTCCCATGGCGC	GCCAAC
14	BrC1=CC2=C(OCC(=O)N2CC#C)C=C1	p	CTGTGTGCTGGTGGCACGAGTCCCATGGCGC	GTGGCA
15	BrC1=C(OCC#C)C=CC=N1	p	CTGTGTGCTGATAATACGAGTCCCATGGCGC	ATAATA
16	C#CC1=NC2=C(C=CC=C2)N=C1	p	CTGTGTGCTGCGTAAGCGAGTCCCATGGCGC	CGTAAG
17	C#CCNC1CCCC1	p	CTGTGTGCTGTCTTCTCGAGTCCCATGGCGC	TCTTCT
18	C#CC1=CN=CN=C1	p	CTGTGTGCTGATTCGCCGAGTCCCATGGCGC	ATTCGC
19	O=C1NCCN1CC#C	p	CTGTGTGCTGGGAAGGCGAGTCCCATGGCGC	GGAAGG
20	COC1=NC(=CC=C1)C#C	p	CTGTGTGCTGATGGTACGAGTCCCATGGCGC	ATGGTA
21	C#CCN1C=CN=C1	p	CTGTGTGCTGTAATAACGAGTCCCATGGCGC	TAATAA
22	O=C(NCC#C)C1=CNC(=O)C=C1	p	CTGTGTGCTGTTAACCCGAGTCCCATGGCGC	TTAACC
23	C#CC1=CC=NC=C1	p	CTGTGTGCTGAATACGCGAGTCCCATGGCGC	AATACG
24	O=C1COCC(=O)N1CC#C	p	CTGTGTGCTGAAGCGGCGAGTCCCATGGCGC	AAGCGG
25	NC1=C(F)C(F)=C(C#C)C(F)=C1F	p	CTGTGTGCTGCGCTTACGAGTCCCATGGCGC	CGCTTA
26	CC1=NC2=C(C=C1)C=C(C=C2)C#C	p	CTGTGTGCTGTGCTGCCGAGTCCCATGGCGC	TGCTGC
27	CS(=O)(=O)NCCC#C	p	CTGTGTGCTGTGCACACGAGTCCCATGGCGC	TGCACA
28	C#CC1=CN=CS1	p	CTGTGTGCTGCTTGCACGAGTCCCATGGCGC	CTTGCA
29	C#CC1=NC=NC=C1	p	CTGTGTGCTGTTAGACCGAGTCCCATGGCGC	TTAGAC
30	Cl.C#CCNCC1CC1	p	CTGTGTGCTGTGATGACGAGTCCCATGGCGC	TGATGA
31	COCCN(C)CC#C	p	CTGTGTGCTGGATATTCGAGTCCCATGGCGC	GATATT
32	C#CC1=NC2=C(S1)C=CC=C2	p	CTGTGTGCTGACGGTGCGAGTCCCATGGCGC	ACGGTG
33	NC(CO)CC#C	p	CTGTGTGCTGAGGTACCGAGTCCCATGGCGC	AGGTAC
34	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NCC#C	p	CTGTGTGCTGGCCTGCCGAGTCCCATGGCGC	GCCTGC
35	CC1=CC(NC(=O)NCC#C)=NO1	p	CTGTGTGCTGGCTCCGCGAGTCCCATGGCGC	GCTCCG
36	O=S(=O)(NCC#C)C1=CC=CC=C1	p	CTGTGTGCTGGATAACCGAGTCCCATGGCGC	GATAAC
37	COC1=CC(=NC=C1)C#C	p	CTGTGTGCTGGGCGTTCGAGTCCCATGGCGC	GGCGTT
38	NC(C#C)C1CCCCC1	p	CTGTGTGCTGGCGGAACGAGTCCCATGGCGC	GCGGAA
39	FC1=C(C=CN=C1)C#C	p	CTGTGTGCTGGACAATCGAGTCCCATGGCGC	GACAAT
40	Cl.C#CCN1C=NC2=C1C=CC=C2	p	CTGTGTGCTGCTTCACCGAGTCCCATGGCGC	CTTCAC
41	O=CC1=CC=C(OCC#C)C=C1	p	CTGTGTGCTGCGTGGCCGAGTCCCATGGCGC	CGTGGC
42	OC(=O)C1=CC=C(C=C1)C#C	p	CTGTGTGCTGACCTACCGAGTCCCATGGCGC	ACCTAC
43	O=S1(=O)CCN(CC#C)CC1	p	CTGTGTGCTGTTCGAACGAGTCCCATGGCGC	TTCGAA
44	NC1=CC=CC(=C1)C#C	p	CTGTGTGCTGCTCCGACGAGTCCCATGGCGC	CTCCGA
45	NS(=O)(=O)C1=CC=C(C=C1)C#C	p	CTGTGTGCTGAGAGCGCGAGTCCCATGGCGC	AGAGCG
46	C#CCNC1CCC2=C1C=CC=C2	p	CTGTGTGCTGCGTGAACGAGTCCCATGGCGC	CGTGAA
47	C#CC1=NC=CC=C1	p	CTGTGTGCTGTCTCGGCGAGTCCCATGGCGC	TCTCGG
48	NC1(CCCCC1)C#C	p	CTGTGTGCTGTCGGCTCGAGTCCCATGGCGC	TCGGCT
49	CNCC#C	p	CTGTGTGCTGACACTCCGAGTCCCATGGCGC	ACACTC
50	NCC#C	p	CTGTGTGCTGGTGGTGCGAGTCCCATGGCGC	GTGGTG
51	C#CC1=CC=C2N=CC=NC2=C1	p	CTGTGTGCTGGGACTTCGAGTCCCATGGCGC	GGACTT
52	CC1=C(C=O)C2=C(C=CC=C2)N1CC#C	p	CTGTGTGCTGGTACTCCGAGTCCCATGGCGC	GTACTC
53	C#CCN1CCOCC1	p	CTGTGTGCTGAGCTAACGAGTCCCATGGCGC	AGCTAA
54	NC1=CC=C(C=C1)C#C	p	CTGTGTGCTGTAAGCGCGAGTCCCATGGCGC	TAAGCG
55	C#CC1=CC2=C(NC=C2)C=C1	p	CTGTGTGCTGCAGCAGCGAGTCCCATGGCGC	CAGCAG
56	C#CCN1C2=C(C=CC=C2)C2=C1C=CC=C2	p	CTGTGTGCTGCACGAACGAGTCCCATGGCGC	CACGAA
57	OC(CC#C)C(O)=O	p	CTGTGTGCTGTATTATCGAGTCCCATGGCGC	TATTAT
58	CC(NCC#C)C1=CC2=C(OCC(=O)N2)C=C1	p	CTGTGTGCTGATGTCACGAGTCCCATGGCGC	ATGTCA
59	CC(=O)O[C@]1(CCC2C3CCC4=CC(=O)CC[C@@H]4C3CC[C@@]12C)C#C	p	CTGTGTGCTGCTAGGACGAGTCCCATGGCGC	CTAGGA
60	NS(=O)(=O)C1=NN=C(NC(=O)CCC#C)S1	p	CTGTGTGCTGCGGTGCCGAGTCCCATGGCGC	CGGTGC
61	NC(=N)C1=CC=C(CNC(=O)CCC#C)C=C1	p	CTGTGTGCTGGGCAGACGAGTCCCATGGCGC	GGCAGA
62	N	p	CTGTGTGCTGAACCTGCGAGTCCCATGGCGC	AACCTG
63	CCC=CCC(O)=O	p	CTGTGTGCTGTGGTAACGAGTCCCATGGCGC	TGGTAA
64	OCC(O)=O	p	CTGTGTGCTGGAACGGCGAGTCCCATGGCGC	GAACGG

65	CC1=CC=C(C=C1)S(=O)(=O)NCC(O)=O	p	CTGTGTGCTGCAGAACCGAGTCCCATGGCGC	CAGAAC
66	CN(CC(O)=O)C(=O)C1=CC=CC=C1	p	CTGTGTGCTGAAGACCCGAGTCCCATGGCGC	AAGACC
67	CN(C)C1=CC=C(CC(O)=O)C=C1	p	CTGTGTGCTGATATAACGAGTCCCATGGCGC	ATATAA
68	COC1=C(OC)C=C(CCC(O)=O)C=C1	p	CTGTGTGCTGAGAATTCGAGTCCCATGGCGC	AGAATT
69	COC1=CC(CC(O)=O)=CC(Br)=C1O	p	CTGTGTGCTGGCCAGGCGAGTCCCATGGCGC	GCCAGG
70	OC(=O)CC1=CC=C2OCOC2=C1	p	CTGTGTGCTGGCGTAGCGAGTCCCATGGCGC	GCGTAG
71	COC1=CC(CCC(O)=O)=CC(OC)=C1OC	p	CTGTGTGCTGGCGTCTCGAGTCCCATGGCGC	GCGTCT
72	OC(=O)CCC1=CNC2=C1C=CC=C2	p	CTGTGTGCTGACTGAGCGAGTCCCATGGCGC	ACTGAG
73	COC1=CC(OC)=NC(CCC(O)=O)=N1	p	CTGTGTGCTGGCTGTGCGAGTCCCATGGCGC	GCTGTG
74	OC(=O)CCC1=CC=CN=C1	p	CTGTGTGCTGAGTGGACGAGTCCCATGGCGC	AGTGGA
75	OC(=O)CCC1CNC2=C1C=CC=C2	p	CTGTGTGCTGCAGGATCGAGTCCCATGGCGC	CAGGAT
76	OC(=O)CC1NC(=O)NC1=O	p	CTGTGTGCTGCAGGCCCGAGTCCCATGGCGC	CAGGCC
77	OC(=O)CCN1C(=O)OC2=C1C=CC=C2	p	CTGTGTGCTGCTCTAACGAGTCCCATGGCGC	CTCTAA
78	CC1=CN(CC(O)=O)C(=O)NC1=O	p	CTGTGTGCTGAACTTACGAGTCCCATGGCGC	AACTTA
79	CN1C2N=CN(CC(O)=O)C2C(=O)N(C)C1=O	p	CTGTGTGCTGTAGGTTCGAGTCCCATGGCGC	TAGGTT
80	CN(CC(O)=O)S(=O)(=O)C1=CC=CC=C1	p	CTGTGTGCTGTACTTGCGAGTCCCATGGCGC	TACTTG
81	OC(=O)COC1=CC2=C(C=CC(=O)O2)C=C1	p	CTGTGTGCTGACCTCACGAGTCCCATGGCGC	ACCTCA
82	CC1=C(C)C2=C(OC1=O)C=C(OCC(O)=O)C=C2	p	CTGTGTGCTGGTGAACCGAGTCCCATGGCGC	GTGAAC
83	OC(=O)CC1OC2=C(NC1=O)C=CC=C2	p	CTGTGTGCTGAACCGCCGAGTCCCATGGCGC	AACCGC
84	OC(=O)CC1=CC2=C(N1)C=CC=C2	p	CTGTGTGCTGTAGAATCGAGTCCCATGGCGC	TAGAAT
85	COC1=C(CO)C=CC(OCC(O)=O)=C1	p	CTGTGTGCTGCGGAATCGAGTCCCATGGCGC	CGGAAT
86	CC1=CC=C(C=C1)C(=O)CCC(O)=O	p	CTGTGTGCTGTATCAACGAGTCCCATGGCGC	TATCAA
87	OC(=O)CCC1=CC(=O)C2=C(O1)C=CC(Br)=C2	p	CTGTGTGCTGTTGCGGCGAGTCCCATGGCGC	TTGCGG
88	OC(=O)CCC1=NN=C(O1)C1=CC=CC=C1	p	CTGTGTGCTGTATTCACGAGTCCCATGGCGC	TATTCA
89	OC(=O)CNC(=O)C1=CC=CO1	p	CTGTGTGCTGACCTGGCGAGTCCCATGGCGC	ACCTGG
90	OC(=O)CCN1C(=O)COC2=C1C=C(Cl)C=C2	p	CTGTGTGCTGCGACTCCGAGTCCCATGGCGC	CGACTC
91	OC(=O)CCN1C=NC2=C(C=CC=C2)C1=O	p	CTGTGTGCTGGTCGTCCGAGTCCCATGGCGC	GTCGTC
92	OC(=O)CCC1=NC(=NO1)C1=CN=CC=C1	p	CTGTGTGCTGCGCCAACGAGTCCCATGGCGC	CGCCAA
93	OC(=O)CCN1C=CC(=O)NC1=O	p	CTGTGTGCTGTAGTAGCGAGTCCCATGGCGC	TAGTAG
94	CC(=O)C1=C(C)N(CCC(O)=O)N=C1C	p	CTGTGTGCTGTATCGCCGAGTCCCATGGCGC	TATCGC
95	CC1=CC2=C(C=CC=C2)N1CCC(O)=O	p	CTGTGTGCTGACGAGTCGAGTCCCATGGCGC	ACGAGT
96	CC1=CC2=C(C=C1)C(CC(O)=O)C(=O)N2	p	CTGTGTGCTGACTCTGCGAGTCCCATGGCGC	ACTCTG
97	CC1=CC(=O)OC2=C1C=C(OCC(O)=O)C=C2	p	CTGTGTGCTGATCTAGCGAGTCCCATGGCGC	ATCTAG
98	OC(=O)CCC1=NC(=NO1)C1=CC=CO1	p	CTGTGTGCTGCACAGACGAGTCCCATGGCGC	CACAGA
99	OC(=O)C1=C2C=CN=CC2=CC=C1	p	CTGTGTGCTGTGAGTCCGAGTCCCATGGCGC	TGAGTC
100	OC(=O)C1=CN=C2C=CC=CN2C1=O	p	CTGTGTGCTGAACGAGCGAGTCCCATGGCGC	AACGAG
101	OC(=O)C1=CC=C2C=CNC2=C1	p	CTGTGTGCTGAGGCCGCGAGTCCCATGGCGC	AGGCCG
102	COC1=CC=C2NC(=CC2=C1)C(O)=O	p	CTGTGTGCTGCAAGCACGAGTCCCATGGCGC	CAAGCA
103	OC(=O)C1=NNC2=C1C=CC=C2	p	CTGTGTGCTGATACGACGAGTCCCATGGCGC	ATACGA
104	OC(=O)C1=CC=C2NC=NC2=C1	p	CTGTGTGCTGCACTCACGAGTCCCATGGCGC	CACTCA
105	CC1=NC(C)=C(CC(O)=O)C(O)=N1	p	CTGTGTGCTGGAGACTCGAGTCCCATGGCGC	GAGACT
106	OC(=O)C1=CN=C(N=C1)N1CCOCC1	p	CTGTGTGCTGTAGTCCCGAGTCCCATGGCGC	TAGTCC
107	OC(=O)C1=NNC(=C1)C1CC1	p	CTGTGTGCTGAGCAGCCGAGTCCCATGGCGC	AGCAGC
108	COC1=CC2=C(C=C1)C(CC(O)=O)=CO2	p	CTGTGTGCTGCCGGACCGAGTCCCATGGCGC	CCGGAC
109	CC1=C(C=C(O1)S(=O)(=O)N1CCOCC1)C(O)=O	p	CTGTGTGCTGCGCCGCCGAGTCCCATGGCGC	CGCCGC
110	CC1=NC2=CC=C(C=C2N=C1C)C(O)=O	p	CTGTGTGCTGTGTTGTCGAGTCCCATGGCGC	TGTTGT
111	CCCC(=O)C1=CN(CC(O)=O)C2=CC=CC=C12	p	CTGTGTGCTGGTGTGCCGAGTCCCATGGCGC	GTGTGC
112	OCCN1C=NC2=CC(=CC=C12)C(O)=O	p	CTGTGTGCTGTAACGACGAGTCCCATGGCGC	TAACGA
113	OCCC1=CN2N=C(C=C2N=C1)C(O)=O	p	CTGTGTGCTGGCGACCCGAGTCCCATGGCGC	GCGACC
114	OC(=O)C1=CC=C(CN2C=CC=N2)O1	p	CTGTGTGCTGGTTGTTCGAGTCCCATGGCGC	GTTGTT
115	OC(=O)C1CCN(CC2=CC=CO2)CC1	p	CTGTGTGCTGTGGAACCGAGTCCCATGGCGC	TGGAAC
116	NC(=O)CN1CCCC(C1)C(O)=O	p	CTGTGTGCTGCGCTATCGAGTCCCATGGCGC	CGCTAT
117	CC1=CC=CN2C(CC(O)=O)=CN=C12	p	CTGTGTGCTGTGTTCGCGAGTCCCATGGCGC	TGTTCG
118	C[C@@H]1C[C@H](NC(=S)N1)C(O)=O	p	CTGTGTGCTGCTCATGCGAGTCCCATGGCGC	CTCATG
119	OC(=O)[C@@H]1CCC(=O)N1	p	CTGTGTGCTGCTCCTCCGAGTCCCATGGCGC	CTCCTC
120	OC(=O)C1CN(CC2=CN=CC=C2)C(=O)C1	p	CTGTGTGCTGCTGGTCCGAGTCCCATGGCGC	CTGGTC
121	COC1=C2OCC(CC2=CC=C1)C(O)=O	p	CTGTGTGCTGTGCGATCGAGTCCCATGGCGC	TGCGAT
122	CC(=O)C1=C(C)N(CC(O)=O)N=C1C	p	CTGTGTGCTGCTCAGCCGAGTCCCATGGCGC	CTCAGC
123	OC(=O)C1=CC=CC=C1N1CCC(=O)NC1=O	p	CTGTGTGCTGCCGTCACGAGTCCCATGGCGC	CCGTCA
124	COC1=C(Cl)C=C(C=C1)N1CC(CC1=O)C(O)=O	p	CTGTGTGCTGCGTCCGCGAGTCCCATGGCGC	CGTCCG
125	OC(=O)CCCC1=NC(=NO1)C1=CC=NC=C1	p	CTGTGTGCTGTGGTTCCGAGTCCCATGGCGC	TGGTTC
126	CN(C)S(=O)(=O)C1=CC(C(O)=O)=C(C)O1	p	CTGTGTGCTGCAAGGTCGAGTCCCATGGCGC	CAAGGT
127	CC1=CC=CN=C1C(O)=O	p	CTGTGTGCTGGAACACCGAGTCCCATGGCGC	GAACAC
128	CCOC1=C(C=CC=N1)C(O)=O	p	CTGTGTGCTGGGATAGCGAGTCCCATGGCGC	GGATAG
129	OC(=O)C1=CN=C(O)C=C1	p	CTGTGTGCTGACTCCACGAGTCCCATGGCGC	ACTCCA
130	OC(=O)C1=CC=C(Br)C=N1	p	CTGTGTGCTGCTAGTTCGAGTCCCATGGCGC	CTAGTT
131	OC(=O)C1=CC=C(OC2=CC=C3OCOC3=C2)N=C1	p	CTGTGTGCTGAGTGCTCGAGTCCCATGGCGC	AGTGCT
132	OC(=O)C1=NNC(=O)C=C1	p	CTGTGTGCTGAAGGCGCGAGTCCCATGGCGC	AAGGCG
133	CC1=NC2=CC=CC=C2N1CC(O)=O	p	CTGTGTGCTGAGAAGACGAGTCCCATGGCGC	AGAAGA
134	OC(=O)C1=CNN=C1	p	CTGTGTGCTGCGGCGGCGAGTCCCATGGCGC	CGGCGG
135	CC1=CC=C(C(O)=O)C(O)=N1	p	CTGTGTGCTGATAGTCCGAGTCCCATGGCGC	ATAGTC
136	CC1=CC=C(O1)C1=NNC(=C1)C(O)=O	p	CTGTGTGCTGCGTCGTCGAGTCCCATGGCGC	CGTCGT
137	OC(=O)C1=C(N=CC=N1)C(=O)N1CCCCC1	p	CTGTGTGCTGTGCGTACGAGTCCCATGGCGC	TGCGTA
138	OC(=O)C1=CN=C2SC=CN2C1=O	p	CTGTGTGCTGTGGTGGCGAGTCCCATGGCGC	TGGTGG
139	NC(=O)C1(CC1)C(O)=O	p	CTGTGTGCTGGCACAGCGAGTCCCATGGCGC	GCACAG
140	CC1=NN2C(=C1)N=CC(C(O)=O)=C2C	p	CTGTGTGCTGTCAAGGCGAGTCCCATGGCGC	TCAAGG
141	CC1=C(C(O)=O)C(C)=NO1	p	CTGTGTGCTGTTGGATCGAGTCCCATGGCGC	TTGGAT

142	OC(=O)C1=CC=C(OC2=CC=CN=C2)O1	p	CTGTGTGCTGAATGACCGAGTCCCATGGCGC	AATGAC
143	O[C@H](C(O)=O)C1=CC=CC=C1	p	CTGTGTGCTGGGCTCTCGAGTCCCATGGCGC	GGCTCT
144	CNC(=O)C1=CC=C(C=C1)C(O)=O	p	CTGTGTGCTGCAACAACGAGTCCCATGGCGC	CAACAA
145	NS(=O)(=O)C1=CC=C(C=C1)C(O)=O	p	CTGTGTGCTGAATCCTCGAGTCCCATGGCGC	AATCCT
146	OC(=O)C1=C(O)N=CC=C1	p	CTGTGTGCTGTCCACGCGAGTCCCATGGCGC	TCCACG
147	COC1=NN2C(CCC(O)=O)=NN=C2C=C1	p	CTGTGTGCTGGGCCACCGAGTCCCATGGCGC	GGCCAC
148	Cl.CN(C)CC1=CNC2=C1C=CC(=C2)C(O)=O	p	CTGTGTGCTGCATTACCGAGTCCCATGGCGC	CATTAC
149	CC1=C(CCC(O)=O)C(=O)NC(=O)N1	p	CTGTGTGCTGGCTAGCCGAGTCCCATGGCGC	GCTAGC
150	NC(=O)NC(CC(O)=O)C1=CC=CS1	p	CTGTGTGCTGATAAGTCGAGTCCCATGGCGC	ATAAGT
151	OC(=O)C1=CN=C(C=C1)N1C=NC=N1	p	CTGTGTGCTGTGTGAGCGAGTCCCATGGCGC	TGTGAG
152	CC1=NC2=C(C=NN2C(C)=C1)C(O)=O	p	CTGTGTGCTGCGTTGACGAGTCCCATGGCGC	CGTTGA
153	OC(=O)C1=NN(C(=O)C=C1)C1=CC=CC=C1	p	CTGTGTGCTGGACGTGCGAGTCCCATGGCGC	GACGTG
154	OC(=O)C1=CNN=N1	p	CTGTGTGCTGAGATATCGAGTCCCATGGCGC	AGATAT
155	OC(=O)C1CCCN1C(=O)C1CC1	p	CTGTGTGCTGATTACACGAGTCCCATGGCGC	ATTACA
156	CN1NC(=O)C2=C1NC(=O)C(CC(O)=O)=C2C	p	CTGTGTGCTGTTCGCGCGAGTCCCATGGCGC	TTCGCG
157	CC1=C(CC(O)=O)C(=O)NC(N)=N1	p	CTGTGTGCTGCAGTAACGAGTCCCATGGCGC	CAGTAA
158	OC(=O)C1=NNC2=C1CCC2	p	CTGTGTGCTGAGCGACCGAGTCCCATGGCGC	AGCGAC
159	CC1=C(C=NC=N1)C(O)=O	p	CTGTGTGCTGCGTGTGCGAGTCCCATGGCGC	CGTGTG
160	CC(NC(=O)C1=CC=C(Br)S1)C(O)=O	p	CTGTGTGCTGGGACAACGAGTCCCATGGCGC	GGACAA
161	CC1=NNC(C(O)=O)=C1Br	p	CTGTGTGCTGACCAGACGAGTCCCATGGCGC	ACCAGA
162	CN1C2=C(NC(CCC(O)=O)=N2)C(=O)NC1=O	p	CTGTGTGCTGACTCACCGAGTCCCATGGCGC	ACTCAC
163	OC(=O)C1=C(Br)C(=NN1)C1CC1	p	CTGTGTGCTGCCAACCCGAGTCCCATGGCGC	CCAACC
164	OC(=O)C1CC1C(=O)N1CCN(CC1)C1=CC=CC=C1	p	CTGTGTGCTGTGAACGCGAGTCCCATGGCGC	TGAACG
165	NC(=O)C1=CC=C(S1)C(O)=O	p	CTGTGTGCTGTCCGCACGAGTCCCATGGCGC	TCCGCA
166	NC1=NC(Cl)=CC(=C1)C(O)=O	p	CTGTGTGCTGTCCAATCGAGTCCCATGGCGC	TCCAAT
167	OC1CC(N(C1)C(=O)C1=CC=C(F)C=C1)C(O)=O	p	CTGTGTGCTGTAGCGTCGAGTCCCATGGCGC	TAGCGT
168	CCC(NC1=CC=CC=C1)C(O)=O	p	CTGTGTGCTGTACTGACGAGTCCCATGGCGC	TACTGA
169	OC(=O)CCN1C=CNC(=O)C1=O	p	CTGTGTGCTGCACGCGCGAGTCCCATGGCGC	CACGCG
170	OC(=O)CN1C=C2C=CC=CC2=N1	p	CTGTGTGCTGCGGCTTCGAGTCCCATGGCGC	CGGCTT
171	NC1=NNC(C(O)=O)=C1C1=CC=CC=C1	p	CTGTGTGCTGGCGCGGCGAGTCCCATGGCGC	GCGCGG
172	CN1N=C(C(O)=O)C(Br)=C1C	p	CTGTGTGCTGCCACGGCGAGTCCCATGGCGC	CCACGG
173	CC1=NC(=NO1)C1=CC(=CC=C1)C(O)=O	p	CTGTGTGCTGAACTCGCGAGTCCCATGGCGC	AACTCG
174	OC(=O)CN(CC1=CC=CC=C1)CC1=CC=CC=C1	p	CTGTGTGCTGGAGTACCGAGTCCCATGGCGC	GAGTAC
175	OC(=O)C1=C(Br)SC=N1	p	CTGTGTGCTGCTTACCCGAGTCCCATGGCGC	CTTACC
176	OC(=O)C1=CNC(=O)C(Br)=C1	p	CTGTGTGCTGAAGTCTCGAGTCCCATGGCGC	AAGTCT
177	OC(=O)C1=NC2=CC=CC=C2N=C1	p	CTGTGTGCTGTCCAGCCGAGTCCCATGGCGC	TCCAGC
178	NC1=C(N=C(Br)C=N1)C(O)=O	p	CTGTGTGCTGCAACGCCGAGTCCCATGGCGC	CAACGC
179	CC1=CC(=NN1C1=CC(F)=CC=C1)C(O)=O	p	CTGTGTGCTGCTCGTACGAGTCCCATGGCGC	CTCGTA
180	CC1=C(CCC(O)=O)C(=O)N2N=CN=C2N1	p	CTGTGTGCTGACTTCGCGAGTCCCATGGCGC	ACTTCG
181	CCC1=CC2=C(S1)N=CN=C2NCCC(O)=O	p	CTGTGTGCTGTATCTTCGAGTCCCATGGCGC	TATCTT
182	OC(=O)CN1C2=C(CCCC1=O)SC=C2	p	CTGTGTGCTGATCCACCGAGTCCCATGGCGC	ATCCAC
183	CN1C(=O)N(CCC(O)=O)C2=C1C=CC=C2	p	CTGTGTGCTGTGCGCCCGAGTCCCATGGCGC	TGCGCC
184	OC(=O)CN1C2=C(CCCC1=O)C=CC=C2	p	CTGTGTGCTGGTCCTGCGAGTCCCATGGCGC	GTCCTG
185	OC(=O)C1=NC=C(F)C=C1	p	CTGTGTGCTGGGTCCACGAGTCCCATGGCGC	GGTCCA
186	OC(=O)CCN1C=C(Cl)C=N1	p	CTGTGTGCTGGAGCAACGAGTCCCATGGCGC	GAGCAA
187	OC(=O)C1CSC2(CCC(=O)N12)C1=CC=CC=C1	p	CTGTGTGCTGCGGTAGCGAGTCCCATGGCGC	CGGTAG
188	CC1=NN(C(C)=C1CC(O)=O)C1=CC=CC=C1	p	CTGTGTGCTGGAAGTTCGAGTCCCATGGCGC	GAAGTT
189	OC(=O)C1=NC=C(CC2=CC=CC=C2)C=C1	p	CTGTGTGCTGTTAGCTCGAGTCCCATGGCGC	TTAGCT
190	CCOC(=O)C1=C(NCC#C)N=C(Cl)C=C1	p	CTGTGTGCTGTTCTATCGAGTCCCATGGCGC	TTCTAT
191	Cl.C#CCNCC1=CC=CO1	p	CTGTGTGCTGAATCTCCGAGTCCCATGGCGC	AATCTC
192	Cl.NC(CC#C)CC(F)(F)F	p	CTGTGTGCTGAGTCTACGAGTCCCATGGCGC	AGTCTA
193	COC1=C(N)C=C(C=C1)C#C	p	CTGTGTGCTGACCAAGCGAGTCCCATGGCGC	ACCAAG
194	Cl.NC(C#C)C1CCOCC1	p	CTGTGTGCTGTGTAGCCGAGTCCCATGGCGC	TGTAGC
195	Cl.NC1(CCC1)C#C	p	CTGTGTGCTGATACCTCGAGTCCCATGGCGC	ATACCT
196	OC(=O)C1=CC2=C(C=C1)N(CC2)C(=O)C#C	p	CTGTGTGCTGCACCTACGAGTCCCATGGCGC	CACCTA
197	NNC(=O)CCC#C	p	CTGTGTGCTGACTTATCGAGTCCCATGGCGC	ACTTAT
198	CN1CCN(CCC#C)CC1	p	CTGTGTGCTGTTGATCCGAGTCCCATGGCGC	TTGATC
199	Cl.C#CC1CNC1	p	CTGTGTGCTGCGATGTCGAGTCCCATGGCGC	CGATGT
200	CN(C)CCC#C	p	CTGTGTGCTGATGCTCCGAGTCCCATGGCGC	ATGCTC
201	C#CCCNC1COC1	p	CTGTGTGCTGGCGCCACGAGTCCCATGGCGC	GCGCCA
202	OC1=CC=C(C(O)=C1)C2=CSC(NC(CCCC#C)=O)=N2	p	CTGTGTGCTGCCTAACCGAGTCCCATGGCGC	CCTAAC
203	OC1=C(O)C=CC(C(CN2C=C(C(NCC#C)=O)C=N2)=O)=C1	p	CTGTGTGCTGGCGGCGCGAGTCCCATGGCGC	GCGGCG
204	OC1=C(O)C=CC(C(CN2C(C=C(C(NCC#C)=O)C=C3)=C3N=C2)=O)=C1	p	CTGTGTGCTGGTGCAGCGAGTCCCATGGCGC	GTGCAG
205	OC(=O)C1=CC(=CN=C1)C#C	m	CTGTGTGCTGCTAACTCGAGTCCCATGGCGC	CTAACT
206	BrC1=NC=C(OCC#C)C=C1	m	CTGTGTGCTGAATCAGCGAGTCCCATGGCGC	AATCAG
207	CNC1=CC=C(OCC#C)C=C1	m	CTGTGTGCTGGGATCACGAGTCCCATGGCGC	GGATCA
208	NC(=O)C1=CC(=CN=C1)C#C	m	CTGTGTGCTGGCAGCACGAGTCCCATGGCGC	GCAGCA
209	O=C(NCC#C)NC1CC1	m	CTGTGTGCTGTATAGACGAGTCCCATGGCGC	TATAGA
210	Cl.NCCC(O)CCC#C	m	CTGTGTGCTGAGGCGCCGAGTCCCATGGCGC	AGGCGC
211	OC1=CC(OCC#C)=CC=C1	m	CTGTGTGCTGCACTGTCGAGTCCCATGGCGC	CACTGT
212	C#CCN1C=CC2=C1C=CC=C2	m	CTGTGTGCTGTCACATCGAGTCCCATGGCGC	TCACAT
213	FC(F)(F)C1=NC(OCC#C)=CC=C1	m	CTGTGTGCTGTCTGCGCGAGTCCCATGGCGC	TCTGCG
214	CC1=NC(=CC=C1)C#C	m	CTGTGTGCTGCGAGATCGAGTCCCATGGCGC	CGAGAT
215	Cl.NC(=N)NCC#C	m	CTGTGTGCTGTGGCGACGAGTCCCATGGCGC	TGGCGA
216	Cl.C#CCN1C=CN=C1C1=CC=CS1	m	CTGTGTGCTGGCGATGCGAGTCCCATGGCGC	GCGATG
217	OC(=O)CC(=O)NCC#C	m	CTGTGTGCTGCAATCCCGAGTCCCATGGCGC	CAATCC
218	BrC1=CC2=C(OCC(=O)N2CC#C)C=C1	m	CTGTGTGCTGCATCATCGAGTCCCATGGCGC	CATCAT

219	BrC1=C(OCC#C)C=CC=N1	m	CTGTGTGCTGCATCCACGAGTCCCATGGCGC	CATCCA
220	C#CC1=NC2=C(C=CC=C2)N=C1	m	CTGTGTGCTGGTCGATCGAGTCCCATGGCGC	GTCGAT
221	C#CCNC1CCCC1	m	CTGTGTGCTGGTACGTCGAGTCCCATGGCGC	GTACGT
222	C#CC1=CN=CN=C1	m	CTGTGTGCTGTGCATGCGAGTCCCATGGCGC	TGCATG
223	O=C1NCCN1CC#C	m	CTGTGTGCTGGCACCTCGAGTCCCATGGCGC	GCACCT
224	COC1=NC(=CC=C1)C#C	m	CTGTGTGCTGAAGTTGCGAGTCCCATGGCGC	AAGTTG
225	C#CCN1C=CN=C1	m	CTGTGTGCTGAATATACGAGTCCCATGGCGC	AATATA
226	O=C(NCC#C)C1=CNC(=O)C=C1	m	CTGTGTGCTGGAGCGCCGAGTCCCATGGCGC	GAGCGC
227	C#CC1=CC=NC=C1	m	CTGTGTGCTGTCTGTTCGAGTCCCATGGCGC	TCTGTT
228	O=C1COCC(=O)N1CC#C	m	CTGTGTGCTGGTTGAACGAGTCCCATGGCGC	GTTGAA
229	NC1=C(F)C(F)=C(C#C)C(F)=C1F	m	CTGTGTGCTGTGTCCTCGAGTCCCATGGCGC	TGTCCT
230	CC1=NC2=C(C=C1)C=C(C=C2)C#C	m	CTGTGTGCTGGGTCTCCGAGTCCCATGGCGC	GGTCTC
231	CS(=O)(=O)NCCC#C	m	CTGTGTGCTGGCGCACCGAGTCCCATGGCGC	GCGCAC
232	C#CC1=CN=CS1	m	CTGTGTGCTGAGTTGCCGAGTCCCATGGCGC	AGTTGC
233	C#CC1=NC=NC=C1	m	CTGTGTGCTGAGCGGTCGAGTCCCATGGCGC	AGCGGT
234	Cl.C#CCNCC1CC1	m	CTGTGTGCTGACTGTACGAGTCCCATGGCGC	ACTGTA
235	COCCN(C)CC#C	m	CTGTGTGCTGTCATACCGAGTCCCATGGCGC	TCATAC
236	C#CC1=NC2=C(S1)C=CC=C2	m	CTGTGTGCTGGGAGTGCGAGTCCCATGGCGC	GGAGTG
237	NC(CO)CC#C	m	CTGTGTGCTGAACACACGAGTCCCATGGCGC	AACACA
238	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NCC#C	m	CTGTGTGCTGTCCTCCCGAGTCCCATGGCGC	TCCTCC
239	CC1=CC(NC(=O)NCC#C)=NO1	m	CTGTGTGCTGAGTCATCGAGTCCCATGGCGC	AGTCAT
240	O=S(=O)(NCC#C)C1=CC=CC=C1	m	CTGTGTGCTGCTACTACGAGTCCCATGGCGC	CTACTA
241	COC1=CC(=NC=C1)C#C	m	CTGTGTGCTGTCCATACGAGTCCCATGGCGC	TCCATA
242	NC(C#C)C1CCCCC1	m	CTGTGTGCTGTTGCCACGAGTCCCATGGCGC	TTGCCA
243	FC1=C(C=CN=C1)C#C	m	CTGTGTGCTGGACATACGAGTCCCATGGCGC	GACATA
244	Cl.C#CCN1C=NC2=C1C=CC=C2	m	CTGTGTGCTGTTGTCTCGAGTCCCATGGCGC	TTGTCT
245	O=CC1=CC=C(OCC#C)C=C1	m	CTGTGTGCTGTACTACCGAGTCCCATGGCGC	TACTAC
246	OC(=O)C1=CC=C(C=C1)C#C	m	CTGTGTGCTGGCAAGTCGAGTCCCATGGCGC	GCAAGT
247	O=S1(=O)CCN(CC#C)CC1	m	CTGTGTGCTGATGGACCGAGTCCCATGGCGC	ATGGAC
248	NC1=CC=CC(=C1)C#C	m	CTGTGTGCTGGATTGACGAGTCCCATGGCGC	GATTGA
249	NS(=O)(=O)C1=CC=C(C=C1)C#C	m	CTGTGTGCTGTACATCCGAGTCCCATGGCGC	TACATC
250	C#CCNC1CCC2=C1C=CC=C2	m	CTGTGTGCTGCAGCCTCGAGTCCCATGGCGC	CAGCCT
251	C#CC1=NC=CC=C1	m	CTGTGTGCTGCGGTCTCGAGTCCCATGGCGC	CGGTCT
252	NC1(CCCCC1)C#C	m	CTGTGTGCTGGATCGTCGAGTCCCATGGCGC	GATCGT
253	CNCC#C	m	CTGTGTGCTGTATACCCGAGTCCCATGGCGC	TATACC
254	NCC#C	m	CTGTGTGCTGGTAACACGAGTCCCATGGCGC	GTAACA
255	C#CC1=CC=C2N=CC=NC2=C1	m	CTGTGTGCTGACGGCACGAGTCCCATGGCGC	ACGGCA
256	CC1=C(C=O)C2=C(C=CC=C2)N1CC#C	m	CTGTGTGCTGGATGCGCGAGTCCCATGGCGC	GATGCG
257	C#CCN1CCOCC1	m	CTGTGTGCTGCGCAGGCGAGTCCCATGGCGC	CGCAGG
258	NC1=CC=C(C=C1)C#C	m	CTGTGTGCTGCTAATCCGAGTCCCATGGCGC	CTAATC
259	C#CC1=CC2=C(NC=C2)C=C1	m	CTGTGTGCTGTGATCCCGAGTCCCATGGCGC	TGATCC
260	C#CCN1C2=C(C=CC=C2)C2=C1C=CC=C2	m	CTGTGTGCTGATAGATCGAGTCCCATGGCGC	ATAGAT
261	OC(CC#C)C(O)=O	m	CTGTGTGCTGCCTCCTCGAGTCCCATGGCGC	CCTCCT
262	CC(NCC#C)C1=CC2=C(OCC(=O)N2)C=C1	m	CTGTGTGCTGCGGAGACGAGTCCCATGGCGC	CGGAGA
263	CC(=O)O[C@]1(CCC2C3CCC4=CC(=O)CC[C@@H]4C3CC[C@@]12C)C#C	m	CTGTGTGCTGGGTGTACGAGTCCCATGGCGC	GGTGTA
264	NS(=O)(=O)C1=NN=C(NC(=O)CCC#C)S1	m	CTGTGTGCTGTGGACTCGAGTCCCATGGCGC	TGGACT
265	NC(=N)C1=CC=C(CNC(=O)CCC#C)C=C1	m	CTGTGTGCTGGCCGGACGAGTCCCATGGCGC	GCCGGA
266	N	m	CTGTGTGCTGTGCCGTCGAGTCCCATGGCGC	TGCCGT
267	CCC=CCC(O)=O	m	CTGTGTGCTGAACGGACGAGTCCCATGGCGC	AACGGA
268	OCC(O)=O	m	CTGTGTGCTGATATTGCGAGTCCCATGGCGC	ATATTG
269	CC1=CC=C(C=C1)S(=O)(=O)NCC(O)=O	m	CTGTGTGCTGGACCTCCGAGTCCCATGGCGC	GACCTC
270	CN(CC(O)=O)C(=O)C1=CC=CC=C1	m	CTGTGTGCTGCATGGACGAGTCCCATGGCGC	CATGGA
271	CN(C)C1=CC=C(CC(O)=O)C=C1	m	CTGTGTGCTGCGAATGCGAGTCCCATGGCGC	CGAATG
272	COC1=C(OC)C=C(CCC(O)=O)C=C1	m	CTGTGTGCTGACATAGCGAGTCCCATGGCGC	ACATAG
273	COC1=CC(CC(O)=O)=CC(Br)=C1O	m	CTGTGTGCTGCCAGCGCGAGTCCCATGGCGC	CCAGCG
274	OC(=O)CC1=CC=C2OCOC2=C1	m	CTGTGTGCTGCGCACTCGAGTCCCATGGCGC	CGCACT
275	COC1=CC(CCC(O)=O)=CC(OC)=C1OC	m	CTGTGTGCTGCCGCAACGAGTCCCATGGCGC	CCGCAA
276	OC(=O)CCC1=CNC2=C1C=CC=C2	m	CTGTGTGCTGGGCGAACGAGTCCCATGGCGC	GGCGAA
277	COC1=CC(OC)=NC(CCC(O)=O)=N1	m	CTGTGTGCTGTTATTACGAGTCCCATGGCGC	TTATTA
278	OC(=O)CCC1=CC=CN=C1	m	CTGTGTGCTGATCACTCGAGTCCCATGGCGC	ATCACT
279	OC(=O)CCC1CNC2=C1C=CC=C2	m	CTGTGTGCTGGACGGTCGAGTCCCATGGCGC	GACGGT
280	OC(=O)CC1NC(=O)NC1=O	m	CTGTGTGCTGCGGACCCGAGTCCCATGGCGC	CGGACC
281	OC(=O)CCN1C(=O)OC2=C1C=CC=C2	m	CTGTGTGCTGGCTCAACGAGTCCCATGGCGC	GCTCAA
282	CC1=CN(CC(O)=O)C(=O)NC1=O	m	CTGTGTGCTGGCCGAGCGAGTCCCATGGCGC	GCCGAG
283	CN1C2N=CN(CC(O)=O)C2C(=O)N(C)C1=O	m	CTGTGTGCTGAGTAGTCGAGTCCCATGGCGC	AGTAGT
284	CN(CC(O)=O)S(=O)(=O)C1=CC=CC=C1	m	CTGTGTGCTGAAGGTCCGAGTCCCATGGCGC	AAGGTC
285	OC(=O)COC1=CC2=C(C=CC(=O)O2)C=C1	m	CTGTGTGCTGAACAACCGAGTCCCATGGCGC	AACAAC
286	CC1=C(C)C2=C(OC1=O)C=C(OCC(O)=O)C=C2	m	CTGTGTGCTGAGTTAGCGAGTCCCATGGCGC	AGTTAG
287	OC(=O)CC1OC2=C(NC1=O)C=CC=C2	m	CTGTGTGCTGCCTCTCCGAGTCCCATGGCGC	CCTCTC
288	OC(=O)CC1=CC2=C(N1)C=CC=C2	m	CTGTGTGCTGAGCGCACGAGTCCCATGGCGC	AGCGCA
289	COC1=C(CO)C=CC(OCC(O)=O)=C1	m	CTGTGTGCTGACCGTTCGAGTCCCATGGCGC	ACCGTT
290	CC1=CC=C(C=C1)C(=O)CCC(O)=O	m	CTGTGTGCTGAAGCCACGAGTCCCATGGCGC	AAGCCA
291	OC(=O)CCC1=CC(=O)C2=C(O1)C=CC(Br)=C2	m	CTGTGTGCTGCGAGTACGAGTCCCATGGCGC	CGAGTA
292	OC(=O)CCC1=NN=C(O1)C1=CC=CC=C1	m	CTGTGTGCTGTCAACTCGAGTCCCATGGCGC	TCAACT
293	OC(=O)CNC(=O)C1=CC=CO1	m	CTGTGTGCTGAGCATACGAGTCCCATGGCGC	AGCATA
294	OC(=O)CCN1C(=O)COC2=C1C=C(Cl)C=C2	m	CTGTGTGCTGGAATGCCGAGTCCCATGGCGC	GAATGC
295	OC(=O)CCN1C=NC2=C(C=CC=C2)C1=O	m	CTGTGTGCTGCCACACCGAGTCCCATGGCGC	CCACAC

296	OC(=O)CCC1=NC(=NO1)C1=CN=CC=C1	m	CTGTGTGCTGCAGACACGAGTCCCATGGCGC	CAGACA
297	OC(=O)CCN1C=CC(=O)NC1=O	m	CTGTGTGCTGTCGTATCGAGTCCCATGGCGC	TCGTAT
298	CC(=O)C1=C(C)N(CCC(O)=O)N=C1C	m	CTGTGTGCTGGGTACGCGAGTCCCATGGCGC	GGTACG
299	CC1=CC2=C(C=CC=C2)N1CCC(O)=O	m	CTGTGTGCTGACCGCGCGAGTCCCATGGCGC	ACCGCG
300	CC1=CC2=C(C=C1)C(CC(O)=O)C(=O)N2	m	CTGTGTGCTGACGCAGCGAGTCCCATGGCGC	ACGCAG
301	CC1=CC(=O)OC2=C1C=C(OCC(O)=O)C=C2	m	CTGTGTGCTGGCTAAGCGAGTCCCATGGCGC	GCTAAG
302	OC(=O)CCC1=NC(=NO1)C1=CC=CO1	m	CTGTGTGCTGAGCTCCCGAGTCCCATGGCGC	AGCTCC
303	OC(=O)C1=C2C=CN=CC2=CC=C1	m	CTGTGTGCTGGATACACGAGTCCCATGGCGC	GATACA
304	OC(=O)C1=CN=C2C=CC=CN2C1=O	m	CTGTGTGCTGCCTAGTCGAGTCCCATGGCGC	CCTAGT
305	OC(=O)C1=CC=C2C=CNC2=C1	m	CTGTGTGCTGCAATGACGAGTCCCATGGCGC	CAATGA
306	COC1=CC=C2NC(=CC2=C1)C(O)=O	m	CTGTGTGCTGCTTGGTCGAGTCCCATGGCGC	CTTGGT
307	OC(=O)C1=NNC2=C1C=CC=C2	m	CTGTGTGCTGTAGATACGAGTCCCATGGCGC	TAGATA
308	OC(=O)C1=CC=C2NC=NC2=C1	m	CTGTGTGCTGGTATGACGAGTCCCATGGCGC	GTATGA
309	CC1=NC(C)=C(CC(O)=O)C(O)=N1	m	CTGTGTGCTGATGGCTCGAGTCCCATGGCGC	ATGGCT
310	OC(=O)C1=CN=C(N=C1)N1CCOCC1	m	CTGTGTGCTGTTCAGGCGAGTCCCATGGCGC	TTCAGG
311	OC(=O)C1=NNC(=C1)C1CC1	m	CTGTGTGCTGCTTCTTCGAGTCCCATGGCGC	CTTCTT
312	COC1=CC2=C(C=C1)C(CC(O)=O)=CO2	m	CTGTGTGCTGTGAATACGAGTCCCATGGCGC	TGAATA
313	CC1=C(C=C(O1)S(=O)(=O)N1CCOCC1)C(O)=O	m	CTGTGTGCTGATTATTCGAGTCCCATGGCGC	ATTATT
314	CC1=NC2=CC=C(C=C2N=C1C)C(O)=O	m	CTGTGTGCTGACGATACGAGTCCCATGGCGC	ACGATA
315	CCCC(=O)C1=CN(CC(O)=O)C2=CC=CC=C12	m	CTGTGTGCTGTAGGCACGAGTCCCATGGCGC	TAGGCA
316	OCCN1C=NC2=CC(=CC=C12)C(O)=O	m	CTGTGTGCTGGGTTGGCGAGTCCCATGGCGC	GGTTGG
317	OCCC1=CN2N=C(C=C2N=C1)C(O)=O	m	CTGTGTGCTGCCAATTCGAGTCCCATGGCGC	CCAATT
318	OC(=O)C1=CC=C(CN2C=CC=N2)O1	m	CTGTGTGCTGACACGTCGAGTCCCATGGCGC	ACACGT
319	OC(=O)C1CCN(CC2=CC=CO2)CC1	m	CTGTGTGCTGCACACCCGAGTCCCATGGCGC	CACACC
320	NC(=O)CN1CCCC(C1)C(O)=O	m	CTGTGTGCTGGGCACCCGAGTCCCATGGCGC	GGCACC
321	CC1=CC=CN2C(CC(O)=O)=CN=C12	m	CTGTGTGCTGGCCTATCGAGTCCCATGGCGC	GCCTAT
322	C[C@@H]1C[C@H](NC(=S)N1)C(O)=O	m	CTGTGTGCTGGTGCCTCGAGTCCCATGGCGC	GTGCCT
323	OC(=O)[C@@H]1CCC(=O)N1	m	CTGTGTGCTGAGCCAGCGAGTCCCATGGCGC	AGCCAG
324	OC(=O)C1CN(CC2=CN=CC=C2)C(=O)C1	m	CTGTGTGCTGGACGCACGAGTCCCATGGCGC	GACGCA
325	COC1=C2OCC(CC2=CC=C1)C(O)=O	m	CTGTGTGCTGTAACCTCGAGTCCCATGGCGC	TAACCT
326	CC(=O)C1=C(C)N(CC(O)=O)N=C1C	m	CTGTGTGCTGAGGACACGAGTCCCATGGCGC	AGGACA
327	OC(=O)C1=CC=CC=C1N1CCC(=O)NC1=O	m	CTGTGTGCTGCTGGAACGAGTCCCATGGCGC	CTGGAA
328	COC1=C(Cl)C=C(C=C1)N1CC(CC1=O)C(O)=O	m	CTGTGTGCTGACTACTCGAGTCCCATGGCGC	ACTACT
329	OC(=O)CCCC1=NC(=NO1)C1=CC=NC=C1	m	CTGTGTGCTGTCCTGTCGAGTCCCATGGCGC	TCCTGT
330	CN(C)S(=O)(=O)C1=CC(C(O)=O)=C(C)O1	m	CTGTGTGCTGCCTATGCGAGTCCCATGGCGC	CCTATG
331	CC1=CC=CN=C1C(O)=O	m	CTGTGTGCTGACGAACCGAGTCCCATGGCGC	ACGAAC
332	CCOC1=C(C=CC=N1)C(O)=O	m	CTGTGTGCTGAGGCAACGAGTCCCATGGCGC	AGGCAA
333	OC(=O)C1=CN=C(O)C=C1	m	CTGTGTGCTGTGACTGCGAGTCCCATGGCGC	TGACTG
334	OC(=O)C1=CC=C(Br)C=N1	m	CTGTGTGCTGAGTAACCGAGTCCCATGGCGC	AGTAAC
335	OC(=O)C1=CC=C(OC2=CC=C3OCOC3=C2)N=C1	m	CTGTGTGCTGCCAGAACGAGTCCCATGGCGC	CCAGAA
336	OC(=O)C1=NNC(=O)C=C1	m	CTGTGTGCTGAACAGGCGAGTCCCATGGCGC	AACAGG
337	CC1=NC2=CC=CC=C2N1CC(O)=O	m	CTGTGTGCTGAAGAGACGAGTCCCATGGCGC	AAGAGA
338	OC(=O)C1=CNN=C1	m	CTGTGTGCTGTTAAGACGAGTCCCATGGCGC	TTAAGA
339	CC1=CC=C(C(O)=O)C(O)=N1	m	CTGTGTGCTGCTCGCTCGAGTCCCATGGCGC	CTCGCT
340	CC1=CC=C(O1)C1=NNC(=C1)C(O)=O	m	CTGTGTGCTGCTGCGCCGAGTCCCATGGCGC	CTGCGC
341	OC(=O)C1=C(N=CC=N1)C(=O)N1CCCCC1	m	CTGTGTGCTGTCACTACGAGTCCCATGGCGC	TCACTA
342	OC(=O)C1=CN=C2SC=CN2C1=O	m	CTGTGTGCTGAATTAACGAGTCCCATGGCGC	AATTAA
343	NC(=O)C1(CC1)C(O)=O	m	CTGTGTGCTGACATCCCGAGTCCCATGGCGC	ACATCC
344	CC1=NN2C(=C1)N=CC(C(O)=O)=C2C	m	CTGTGTGCTGAAGCACCGAGTCCCATGGCGC	AAGCAC
345	CC1=C(C(O)=O)C(C)=NO1	m	CTGTGTGCTGTTGGCCCGAGTCCCATGGCGC	TTGGCC
346	OC(=O)C1=CC=C(OC2=CC=CN=C2)O1	m	CTGTGTGCTGTGAAGTCGAGTCCCATGGCGC	TGAAGT
347	O[C@H](C(O)=O)C1=CC=CC=C1	m	CTGTGTGCTGAGTGTCCGAGTCCCATGGCGC	AGTGTC
348	CNC(=O)C1=CC=C(C=C1)C(O)=O	m	CTGTGTGCTGCGAAGCCGAGTCCCATGGCGC	CGAAGC
349	NS(=O)(=O)C1=CC=C(C=C1)C(O)=O	m	CTGTGTGCTGTTGTTGCGAGTCCCATGGCGC	TTGTTG
350	OC(=O)C1=C(O)N=CC=C1	m	CTGTGTGCTGGTCCAACGAGTCCCATGGCGC	GTCCAA
351	COC1=NN2C(CCC(O)=O)=NN=C2C=C1	m	CTGTGTGCTGGTGTAACGAGTCCCATGGCGC	GTGTAA
352	Cl.CN(C)CC1=CNC2=C1C=CC(=C2)C(O)=O	m	CTGTGTGCTGTCTTAGCGAGTCCCATGGCGC	TCTTAG
353	CC1=C(CCC(O)=O)C(=O)NC(=O)N1	m	CTGTGTGCTGTAGCTGCGAGTCCCATGGCGC	TAGCTG
354	NC(=O)NC(CC(O)=O)C1=CC=CS1	m	CTGTGTGCTGCACGGCCGAGTCCCATGGCGC	CACGGC
355	OC(=O)C1=CN=C(C=C1)N1C=NC=N1	m	CTGTGTGCTGCTACATCGAGTCCCATGGCGC	CTACAT
356	CC1=NC2=C(C=NN2C(C)=C1)C(O)=O	m	CTGTGTGCTGGTCAGTCGAGTCCCATGGCGC	GTCAGT
357	OC(=O)C1=NN(C(=O)C=C1)C1=CC=CC=C1	m	CTGTGTGCTGACGTTCCGAGTCCCATGGCGC	ACGTTC
358	OC(=O)C1=CNN=N1	m	CTGTGTGCTGATTCCGCGAGTCCCATGGCGC	ATTCCG
359	OC(=O)C1CCCN1C(=O)C1CC1	m	CTGTGTGCTGCACCGGCGAGTCCCATGGCGC	CACCGG
360	CN1NC(=O)C2=C1NC(=O)C(CC(O)=O)=C2C	m	CTGTGTGCTGTCAGCCCGAGTCCCATGGCGC	TCAGCC
361	CC1=C(CC(O)=O)C(=O)NC(N)=N1	m	CTGTGTGCTGGGTGGTCGAGTCCCATGGCGC	GGTGGT
362	OC(=O)C1=NNC2=C1CCC2	m	CTGTGTGCTGCAACTTCGAGTCCCATGGCGC	CAACTT
363	CC1=C(C=NC=N1)C(O)=O	m	CTGTGTGCTGATATGCCGAGTCCCATGGCGC	ATATGC
364	CC(NC(=O)C1=CC=C(Br)S1)C(O)=O	m	CTGTGTGCTGGCATGGCGAGTCCCATGGCGC	GCATGG
365	CC1=NNC(C(O)=O)=C1Br	m	CTGTGTGCTGGGAACTCGAGTCCCATGGCGC	GGAACT
366	CN1C2=C(NC(CCC(O)=O)=N2)C(=O)NC1=O	m	CTGTGTGCTGATCATCCGAGTCCCATGGCGC	ATCATC
367	OC(=O)C1=C(Br)C(=NN1)C1CC1	m	CTGTGTGCTGAGTATGCGAGTCCCATGGCGC	AGTATG
368	OC(=O)C1CC1C(=O)N1CCN(CC1)C1=CC=CC=C1	m	CTGTGTGCTGGCGAGACGAGTCCCATGGCGC	GCGAGA
369	NC(=O)C1=CC=C(S1)C(O)=O	m	CTGTGTGCTGTATGCTCGAGTCCCATGGCGC	TATGCT
370	NC1=NC(Cl)=CC(=C1)C(O)=O	m	CTGTGTGCTGCTTCGGCGAGTCCCATGGCGC	CTTCGG
371	OC1CC(N(C1)C(=O)C1=CC=C(F)C=C1)C(O)=O	m	CTGTGTGCTGGGAGACCGAGTCCCATGGCGC	GGAGAC
372	CCC(NC1=CC=CC=C1)C(O)=O	m	CTGTGTGCTGGGCGCGCGAGTCCCATGGCGC	GGCGCG

373	OC(=O)CCN1C=CNC(=O)C1=O	m	CTGTGTGCTGCGGCCACGAGTCCCATGGCGC	CGGCCA
374	OC(=O)CN1C=C2C=CC=CC2=N1	m	CTGTGTGCTGTTATGTCGAGTCCCATGGCGC	TTATGT
375	NC1=NNC(C(O)=O)=C1C1=CC=CC=C1	m	CTGTGTGCTGTCTGAACGAGTCCCATGGCGC	TCTGAA
376	CN1N=C(C(O)=O)C(Br)=C1C	m	CTGTGTGCTGCTGTGACGAGTCCCATGGCGC	CTGTGA
377	CC1=NC(=NO1)C1=CC(=CC=C1)C(O)=O	m	CTGTGTGCTGACCGGCCGAGTCCCATGGCGC	ACCGGC
378	OC(=O)CN(CC1=CC=CC=C1)CC1=CC=CC=C1	m	CTGTGTGCTGCCTTCCCGAGTCCCATGGCGC	CCTTCC
379	OC(=O)C1=C(Br)SC=N1	m	CTGTGTGCTGGCCGCCCGAGTCCCATGGCGC	GCCGCC
380	OC(=O)C1=CNC(=O)C(Br)=C1	m	CTGTGTGCTGGCTGCTCGAGTCCCATGGCGC	GCTGCT
381	OC(=O)C1=NC2=CC=CC=C2N=C1	m	CTGTGTGCTGTTGCACCGAGTCCCATGGCGC	TTGCAC
382	NC1=C(N=C(Br)C=N1)C(O)=O	m	CTGTGTGCTGGAATTACGAGTCCCATGGCGC	GAATTA
383	CC1=CC(=NN1C1=CC(F)=CC=C1)C(O)=O	m	CTGTGTGCTGCCATTCCGAGTCCCATGGCGC	CCATTC
384	CC1=C(CCC(O)=O)C(=O)N2N=CN=C2N1	m	CTGTGTGCTGGCACGACGAGTCCCATGGCGC	GCACGA
385	CCC1=CC2=C(S1)N=CN=C2NCCC(O)=O	m	CTGTGTGCTGTTACCGCGAGTCCCATGGCGC	TTACCG
386	OC(=O)CN1C2=C(CCCC1=O)SC=C2	m	CTGTGTGCTGACGTAACGAGTCCCATGGCGC	ACGTAA
387	CN1C(=O)N(CCC(O)=O)C2=C1C=CC=C2	m	CTGTGTGCTGCCAGGCCGAGTCCCATGGCGC	CCAGGC
388	OC(=O)CN1C2=C(CCCC1=O)C=CC=C2	m	CTGTGTGCTGTCAGAGCGAGTCCCATGGCGC	TCAGAG
389	OC(=O)C1=NC=C(F)C=C1	m	CTGTGTGCTGAGACACCGAGTCCCATGGCGC	AGACAC
390	OC(=O)CCN1C=C(Cl)C=N1	m	CTGTGTGCTGGGTCAGCGAGTCCCATGGCGC	GGTCAG
391	OC(=O)C1CSC2(CCC(=O)N12)C1=CC=CC=C1	m	CTGTGTGCTGAGAACCCGAGTCCCATGGCGC	AGAACC
392	CC1=NN(C(C)=C1CC(O)=O)C1=CC=CC=C1	m	CTGTGTGCTGATGAGCCGAGTCCCATGGCGC	ATGAGC
393	OC(=O)C1=NC=C(CC2=CC=CC=C2)C=C1	m	CTGTGTGCTGACTTGACGAGTCCCATGGCGC	ACTTGA
394	CCOC(=O)C1=C(NCC#C)N=C(Cl)C=C1	m	CTGTGTGCTGAGGAGGCGAGTCCCATGGCGC	AGGAGG
395	Cl.C#CCNCC1=CC=CO1	m	CTGTGTGCTGTCGTCGCGAGTCCCATGGCGC	TCGTCG
396	Cl.NC(CC#C)CC(F)(F)F	m	CTGTGTGCTGGCTTCACGAGTCCCATGGCGC	GCTTCA
397	COC1=C(N)C=C(C=C1)C#C	m	CTGTGTGCTGTCAGGTCGAGTCCCATGGCGC	TCAGGT
398	Cl.NC(C#C)C1CCOCC1	m	CTGTGTGCTGCGAACACGAGTCCCATGGCGC	CGAACA
399	Cl.NC1(CCC1)C#C	m	CTGTGTGCTGCGCGAGCGAGTCCCATGGCGC	CGCGAG
400	OC(=O)C1=CC2=C(C=C1)N(CC2)C(=O)C#C	m	CTGTGTGCTGCCTCAGCGAGTCCCATGGCGC	CCTCAG
401	NNC(=O)CCC#C	m	CTGTGTGCTGGGCCTACGAGTCCCATGGCGC	GGCCTA
402	CN1CCN(CCC#C)CC1	m	CTGTGTGCTGTCGACACGAGTCCCATGGCGC	TCGACA
403	Cl.C#CC1CNC1	m	CTGTGTGCTGCATATCCGAGTCCCATGGCGC	CATATC
404	CN(C)CCC#C	m	CTGTGTGCTGAGGATCCGAGTCCCATGGCGC	AGGATC
405	C#CCCNC1COC1	m	CTGTGTGCTGAAGGAACGAGTCCCATGGCGC	AAGGAA
406	OC1=CC=C(C(O)=C1)C2=CSC(NC(CCCC#C)=O)=N2	m	CTGTGTGCTGTCTGGCCGAGTCCCATGGCGC	TCTGGC
407	OC1=C(O)C=CC(C(CN2C=C(C(NCC#C)=O)C=N2)=O)=C1	m	CTGTGTGCTGCGATAACGAGTCCCATGGCGC	CGATAA
408	OC1=C(O)C=CC(C(CN2C(C=C(C(NCC#C)=O)C=C3)=C3N=C2)=O)=C1	m	CTGTGTGCTGCCAAGACGAGTCCCATGGCGC	CCAAGA
409	OC(=O)C1=CC(=CN=C1)C#C	o	CTGTGTGCTGCTCGACCGAGTCCCATGGCGC	CTCGAC
410	BrC1=NC=C(OCC#C)C=C1	o	CTGTGTGCTGGTTACTCGAGTCCCATGGCGC	GTTACT
411	CNC1=CC=C(OCC#C)C=C1	o	CTGTGTGCTGTGCTAGCGAGTCCCATGGCGC	TGCTAG
412	NC(=O)C1=CC(=CN=C1)C#C	o	CTGTGTGCTGTCGGTACGAGTCCCATGGCGC	TCGGTA
413	O=C(NCC#C)NC1CC1	o	CTGTGTGCTGGCTTGTCGAGTCCCATGGCGC	GCTTGT
414	Cl.NCCC(O)CCC#C	o	CTGTGTGCTGCCGCGTCGAGTCCCATGGCGC	CCGCGT
415	OC1=CC(OCC#C)=CC=C1	o	CTGTGTGCTGCGATCGCGAGTCCCATGGCGC	CGATCG
416	C#CCN1C=CC2=C1C=CC=C2	o	CTGTGTGCTGATGAATCGAGTCCCATGGCGC	ATGAAT
417	FC(F)(F)C1=NC(OCC#C)=CC=C1	o	CTGTGTGCTGCATACTCGAGTCCCATGGCGC	CATACT
418	CC1=NC(=CC=C1)C#C	o	CTGTGTGCTGCGACAGCGAGTCCCATGGCGC	CGACAG
419	Cl.NC(=N)NCC#C	o	CTGTGTGCTGGAAGAACGAGTCCCATGGCGC	GAAGAA
420	Cl.C#CCN1C=CN=C1C1=CC=CS1	o	CTGTGTGCTGGGTTAACGAGTCCCATGGCGC	GGTTAA
421	OC(=O)CC(=O)NCC#C	o	CTGTGTGCTGACGCCTCGAGTCCCATGGCGC	ACGCCT
422	BrC1=CC2=C(OCC(=O)N2CC#C)C=C1	o	CTGTGTGCTGAATGTGCGAGTCCCATGGCGC	AATGTG
423	BrC1=C(OCC#C)C=CC=N1	o	CTGTGTGCTGACAAGCCGAGTCCCATGGCGC	ACAAGC
424	C#CC1=NC2=C(C=CC=C2)N=C1	o	CTGTGTGCTGGACTGGCGAGTCCCATGGCGC	GACTGG
425	C#CCNC1CCCC1	o	CTGTGTGCTGACGGATCGAGTCCCATGGCGC	ACGGAT
426	C#CC1=CN=CN=C1	o	CTGTGTGCTGATCTGACGAGTCCCATGGCGC	ATCTGA
427	O=C1NCCN1CC#C	o	CTGTGTGCTGCTCTCCCGAGTCCCATGGCGC	CTCTCC
428	COC1=NC(=CC=C1)C#C	o	CTGTGTGCTGAACGCCCGAGTCCCATGGCGC	AACGCC
429	C#CCN1C=CN=C1	o	CTGTGTGCTGCAGGTACGAGTCCCATGGCGC	CAGGTA
430	O=C(NCC#C)C1=CNC(=O)C=C1	o	CTGTGTGCTGGATCTACGAGTCCCATGGCGC	GATCTA
431	C#CC1=CC=NC=C1	o	CTGTGTGCTGGCATAACGAGTCCCATGGCGC	GCATAA
432	O=C1COCC(=O)N1CC#C	o	CTGTGTGCTGAGCACGCGAGTCCCATGGCGC	AGCACG
433	NC1=C(F)C(F)=C(C#C)C(F)=C1F	o	CTGTGTGCTGACCGAACGAGTCCCATGGCGC	ACCGAA
434	CC1=NC2=C(C=C1)C=C(C=C2)C#C	o	CTGTGTGCTGAACCAACGAGTCCCATGGCGC	AACCAA
435	CS(=O)(=O)NCCC#C	o	CTGTGTGCTGGCTGACCGAGTCCCATGGCGC	GCTGAC
436	C#CC1=CN=CS1	o	CTGTGTGCTGCAGTCGCGAGTCCCATGGCGC	CAGTCG
437	C#CC1=NC=NC=C1	o	CTGTGTGCTGCACGTTCGAGTCCCATGGCGC	CACGTT
438	Cl.C#CCNCC1CC1	o	CTGTGTGCTGTCACGCCGAGTCCCATGGCGC	TCACGC
439	COCCN(C)CC#C	o	CTGTGTGCTGAAGCTTCGAGTCCCATGGCGC	AAGCTT
440	C#CC1=NC2=C(S1)C=CC=C2	o	CTGTGTGCTGCTCACACGAGTCCCATGGCGC	CTCACA
441	NC(CO)CC#C	o	CTGTGTGCTGGCCTTACGAGTCCCATGGCGC	GCCTTA
442	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NCC#C	o	CTGTGTGCTGCGCGTCCGAGTCCCATGGCGC	CGCGTC
443	CC1=CC(NC(=O)NCC#C)=NO1	o	CTGTGTGCTGGCCATTCGAGTCCCATGGCGC	GCCATT
444	O=S(=O)(NCC#C)C1=CC=CC=C1	o	CTGTGTGCTGCCGTGGCGAGTCCCATGGCGC	CCGTGG
445	COC1=CC(=NC=C1)C#C	o	CTGTGTGCTGACTATCCGAGTCCCATGGCGC	ACTATC
446	NC(C#C)C1CCCCC1	o	CTGTGTGCTGACAACACGAGTCCCATGGCGC	ACAACA
447	FC1=C(C=CN=C1)C#C	o	CTGTGTGCTGTGAGCACGAGTCCCATGGCGC	TGAGCA
448	Cl.C#CCN1C=NC2=C1C=CC=C2	o	CTGTGTGCTGCTCAATCGAGTCCCATGGCGC	CTCAAT
449	O=CC1=CC=C(OCC#C)C=C1	o	CTGTGTGCTGCGCCTGCGAGTCCCATGGCGC	CGCCTG

450	OC(=O)C1=CC=C(C=C1)C#C	o	CTGTGTGCTGGTAGTACGAGTCCCATGGCGC	GTAGTA
451	O=S1(=O)CCN(CC#C)CC1	o	CTGTGTGCTGGACACGCGAGTCCCATGGCGC	GACACG
452	NC1=CC=CC(=C1)C#C	o	CTGTGTGCTGTCATTGCGAGTCCCATGGCGC	TCATTG
453	NS(=O)(=O)C1=CC=C(C=C1)C#C	o	CTGTGTGCTGGACCAGCGAGTCCCATGGCGC	GACCAG
454	C#CCNC1CCC2=C1C=CC=C2	o	CTGTGTGCTGATAACGCGAGTCCCATGGCGC	ATAACG
455	C#CC1=NC=CC=C1	o	CTGTGTGCTGATAGCACGAGTCCCATGGCGC	ATAGCA
456	NC1(CCCCC1)C#C	o	CTGTGTGCTGCTCTGGCGAGTCCCATGGCGC	CTCTGG
457	CNCC#C	o	CTGTGTGCTGCCTGATCGAGTCCCATGGCGC	CCTGAT
458	NCC#C	o	CTGTGTGCTGACAGACCGAGTCCCATGGCGC	ACAGAC
459	C#CC1=CC=C2N=CC=NC2=C1	o	CTGTGTGCTGCCTCGACGAGTCCCATGGCGC	CCTCGA
460	CC1=C(C=O)C2=C(C=CC=C2)N1CC#C	o	CTGTGTGCTGGCAGTCCGAGTCCCATGGCGC	GCAGTC
461	C#CCN1CCOCC1	o	CTGTGTGCTGGGCGGCCGAGTCCCATGGCGC	GGCGGC
462	NC1=CC=C(C=C1)C#C	o	CTGTGTGCTGAATCGACGAGTCCCATGGCGC	AATCGA
463	C#CC1=CC2=C(NC=C2)C=C1	o	CTGTGTGCTGATACAGCGAGTCCCATGGCGC	ATACAG
464	C#CCN1C2=C(C=CC=C2)C2=C1C=CC=C2	o	CTGTGTGCTGCTGAGTCGAGTCCCATGGCGC	CTGAGT
465	OC(CC#C)C(O)=O	o	CTGTGTGCTGCGCGGACGAGTCCCATGGCGC	CGCGGA
466	CC(NCC#C)C1=CC2=C(OCC(=O)N2)C=C1	o	CTGTGTGCTGCGCAACCGAGTCCCATGGCGC	CGCAAC
467	CC(=O)O[C@]1(CCC2C3CCC4=CC(=O)CC[C@@H]4C3CC[C@@]12C)C#C	o	CTGTGTGCTGCATCTGCGAGTCCCATGGCGC	CATCTG
468	NS(=O)(=O)C1=NN=C(NC(=O)CCC#C)S1	o	CTGTGTGCTGTATGTCCGAGTCCCATGGCGC	TATGTC
469	NC(=N)C1=CC=C(CNC(=O)CCC#C)C=C1	o	CTGTGTGCTGCAGCGACGAGTCCCATGGCGC	CAGCGA
470	N	o	CTGTGTGCTGTAAGTACGAGTCCCATGGCGC	TAAGTA
471	CCC=CCC(O)=O	o	CTGTGTGCTGGGACGCCGAGTCCCATGGCGC	GGACGC
472	OCC(O)=O	o	CTGTGTGCTGACTGCCCGAGTCCCATGGCGC	ACTGCC
473	CC1=CC=C(C=C1)S(=O)(=O)NCC(O)=O	o	CTGTGTGCTGGACAGCCGAGTCCCATGGCGC	GACAGC
474	CN(CC(O)=O)C(=O)C1=CC=CC=C1	o	CTGTGTGCTGGTTATACGAGTCCCATGGCGC	GTTATA
475	CN(C)C1=CC=C(CC(O)=O)C=C1	o	CTGTGTGCTGGAGGAGCGAGTCCCATGGCGC	GAGGAG
476	COC1=C(OC)C=C(CCC(O)=O)C=C1	o	CTGTGTGCTGAATTGGCGAGTCCCATGGCGC	AATTGG
477	COC1=CC(CC(O)=O)=CC(Br)=C1O	o	CTGTGTGCTGCTCCAGCGAGTCCCATGGCGC	CTCCAG
478	OC(=O)CC1=CC=C2OCOC2=C1	o	CTGTGTGCTGGCAACGCGAGTCCCATGGCGC	GCAACG
479	COC1=CC(CCC(O)=O)=CC(OC)=C1OC	o	CTGTGTGCTGGATAGGCGAGTCCCATGGCGC	GATAGG
480	OC(=O)CCC1=CNC2=C1C=CC=C2	o	CTGTGTGCTGTACCATCGAGTCCCATGGCGC	TACCAT
481	COC1=CC(OC)=NC(CCC(O)=O)=N1	o	CTGTGTGCTGGTCTCACGAGTCCCATGGCGC	GTCTCA
482	OC(=O)CCC1=CC=CN=C1	o	CTGTGTGCTGCCGAGCCGAGTCCCATGGCGC	CCGAGC
483	OC(=O)CCC1CNC2=C1C=CC=C2	o	CTGTGTGCTGGAGATCCGAGTCCCATGGCGC	GAGATC
484	OC(=O)CC1NC(=O)NC1=O	o	CTGTGTGCTGCCGGTTCGAGTCCCATGGCGC	CCGGTT
485	OC(=O)CCN1C(=O)OC2=C1C=CC=C2	o	CTGTGTGCTGGGAGGACGAGTCCCATGGCGC	GGAGGA
486	CC1=CN(CC(O)=O)C(=O)NC1=O	o	CTGTGTGCTGTATCCGCGAGTCCCATGGCGC	TATCCG
487	CN1C2N=CN(CC(O)=O)C2C(=O)N(C)C1=O	o	CTGTGTGCTGATGCGTCGAGTCCCATGGCGC	ATGCGT
488	CN(CC(O)=O)S(=O)(=O)C1=CC=CC=C1	o	CTGTGTGCTGCAATAGCGAGTCCCATGGCGC	CAATAG
489	OC(=O)COC1=CC2=C(C=CC(=O)O2)C=C1	o	CTGTGTGCTGTAGGACCGAGTCCCATGGCGC	TAGGAC
490	CC1=C(C)C2=C(OC1=O)C=C(OCC(O)=O)C=C2	o	CTGTGTGCTGCTGTATCGAGTCCCATGGCGC	CTGTAT
491	OC(=O)CC1OC2=C(NC1=O)C=CC=C2	o	CTGTGTGCTGTCGTGACGAGTCCCATGGCGC	TCGTGA
492	OC(=O)CC1=CC2=C(N1)C=CC=C2	o	CTGTGTGCTGGCGCTTCGAGTCCCATGGCGC	GCGCTT
493	COC1=C(CO)C=CC(OCC(O)=O)=C1	o	CTGTGTGCTGGGCCGGCGAGTCCCATGGCGC	GGCCGG
494	CC1=CC=C(C=C1)C(=O)CCC(O)=O	o	CTGTGTGCTGGATTCTCGAGTCCCATGGCGC	GATTCT
495	OC(=O)CCC1=CC(=O)C2=C(O1)C=CC(Br)=C2	o	CTGTGTGCTGCACTTCCGAGTCCCATGGCGC	CACTTC
496	OC(=O)CCC1=NN=C(O1)C1=CC=CC=C1	o	CTGTGTGCTGTAACTCCGAGTCCCATGGCGC	TAACTC
497	OC(=O)CNC(=O)C1=CC=CO1	o	CTGTGTGCTGACAGCTCGAGTCCCATGGCGC	ACAGCT
498	OC(=O)CCN1C(=O)COC2=C1C=C(Cl)C=C2	o	CTGTGTGCTGTAATGGCGAGTCCCATGGCGC	TAATGG
499	OC(=O)CCN1C=NC2=C(C=CC=C2)C1=O	o	CTGTGTGCTGAATGGTCGAGTCCCATGGCGC	AATGGT
500	OC(=O)CCC1=NC(=NO1)C1=CN=CC=C1	o	CTGTGTGCTGCCGCCGCGAGTCCCATGGCGC	CCGCCG
501	OC(=O)CCN1C=CC(=O)NC1=O	o	CTGTGTGCTGTAAGGCCGAGTCCCATGGCGC	TAAGGC
502	CC(=O)C1=C(C)N(CCC(O)=O)N=C1C	o	CTGTGTGCTGAAGTGCCGAGTCCCATGGCGC	AAGTGC
503	CC1=CC2=C(C=CC=C2)N1CCC(O)=O	o	CTGTGTGCTGTCAATCCGAGTCCCATGGCGC	TCAATC
504	CC1=CC2=C(C=C1)C(CC(O)=O)C(=O)N2	o	CTGTGTGCTGCACCACCGAGTCCCATGGCGC	CACCAC
505	CC1=CC(=O)OC2=C1C=C(OCC(O)=O)C=C2	o	CTGTGTGCTGCGTATACGAGTCCCATGGCGC	CGTATA
506	OC(=O)CCC1=NC(=NO1)C1=CC=CO1	o	CTGTGTGCTGATCCTACGAGTCCCATGGCGC	ATCCTA
507	OC(=O)C1=C2C=CN=CC2=CC=C1	o	CTGTGTGCTGGTGTCGCGAGTCCCATGGCGC	GTGTCG
508	OC(=O)C1=CN=C2C=CC=CN2C1=O	o	CTGTGTGCTGTTGAAGCGAGTCCCATGGCGC	TTGAAG
509	OC(=O)C1=CC=C2C=CNC2=C1	o	CTGTGTGCTGTCCGTGCGAGTCCCATGGCGC	TCCGTG
510	COC1=CC=C2NC(=CC2=C1)C(O)=O	o	CTGTGTGCTGGGTGCCCGAGTCCCATGGCGC	GGTGCC
511	OC(=O)C1=NNC2=C1C=CC=C2	o	CTGTGTGCTGATGTGGCGAGTCCCATGGCGC	ATGTGG
512	OC(=O)C1=CC=C2NC=NC2=C1	o	CTGTGTGCTGGCAATACGAGTCCCATGGCGC	GCAATA
513	CC1=NC(C)=C(CC(O)=O)C(O)=N1	o	CTGTGTGCTGTACACTCGAGTCCCATGGCGC	TACACT
514	OC(=O)C1=CN=C(N=C1)N1CCOCC1	o	CTGTGTGCTGACAGGACGAGTCCCATGGCGC	ACAGGA
515	OC(=O)C1=NNC(=C1)C1CC1	o	CTGTGTGCTGACCACCCGAGTCCCATGGCGC	ACCACC
516	COC1=CC2=C(C=C1)C(CC(O)=O)=CO2	o	CTGTGTGCTGGATTAGCGAGTCCCATGGCGC	GATTAG
517	CC1=C(C=C(O1)S(=O)(=O)N1CCOCC1)C(O)=O	o	CTGTGTGCTGCCTACACGAGTCCCATGGCGC	CCTACA
518	CC1=NC2=CC=C(C=C2N=C1C)C(O)=O	o	CTGTGTGCTGATGATGCGAGTCCCATGGCGC	ATGATG
519	CCCC(=O)C1=CN(CC(O)=O)C2=CC=CC=C12	o	CTGTGTGCTGACAATGCGAGTCCCATGGCGC	ACAATG
520	OCCN1C=NC2=CC(=CC=C12)C(O)=O	o	CTGTGTGCTGAGCCGACGAGTCCCATGGCGC	AGCCGA
521	OCCC1=CN2N=C(C=C2N=C1)C(O)=O	o	CTGTGTGCTGTTAGTGCGAGTCCCATGGCGC	TTAGTG
522	OC(=O)C1=CC=C(CN2C=CC=N2)O1	o	CTGTGTGCTGGTAATGCGAGTCCCATGGCGC	GTAATG
523	OC(=O)C1CCN(CC2=CC=CO2)CC1	o	CTGTGTGCTGTTCTTCCGAGTCCCATGGCGC	TTCTTC
524	NC(=O)CN1CCCC(C1)C(O)=O	o	CTGTGTGCTGAAGAAGCGAGTCCCATGGCGC	AAGAAG
525	CC1=CC=CN2C(CC(O)=O)=CN=C12	o	CTGTGTGCTGCTATACCGAGTCCCATGGCGC	CTATAC
526	C[C@@H]1C[C@H](NC(=S)N1)C(O)=O	o	CTGTGTGCTGCAAGACCGAGTCCCATGGCGC	CAAGAC

527	OC(=O)[C@@H]1CCC(=O)N1	o	CTGTGTGCTGACGACGCGAGTCCCATGGCGC	ACGACG
528	OC(=O)C1CN(CC2=CN=CC=C2)C(=O)C1	o	CTGTGTGCTGTCGATTCGAGTCCCATGGCGC	TCGATT
529	COC1=C2OCC(CC2=CC=C1)C(O)=O	o	CTGTGTGCTGCCACCACGAGTCCCATGGCGC	CCACCA
530	CC(=O)C1=C(C)N(CC(O)=O)N=C1C	o	CTGTGTGCTGTTCAACCGAGTCCCATGGCGC	TTCAAC
531	OC(=O)C1=CC=CC=C1N1CCC(=O)NC1=O	o	CTGTGTGCTGCCATATCGAGTCCCATGGCGC	CCATAT
532	COC1=C(Cl)C=C(C=C1)N1CC(CC1=O)C(O)=O	o	CTGTGTGCTGCCGACTCGAGTCCCATGGCGC	CCGACT
533	OC(=O)CCCC1=NC(=NO1)C1=CC=NC=C1	o	CTGTGTGCTGTAAGATCGAGTCCCATGGCGC	TAAGAT
534	CN(C)S(=O)(=O)C1=CC(C(O)=O)=C(C)O1	o	CTGTGTGCTGGTGATTCGAGTCCCATGGCGC	GTGATT
535	CC1=CC=CN=C1C(O)=O	o	CTGTGTGCTGAATGCACGAGTCCCATGGCGC	AATGCA
536	CCOC1=C(C=CC=N1)C(O)=O	o	CTGTGTGCTGATTCAACGAGTCCCATGGCGC	ATTCAA
537	OC(=O)C1=CN=C(O)C=C1	o	CTGTGTGCTGAGTTCACGAGTCCCATGGCGC	AGTTCA
538	OC(=O)C1=CC=C(Br)C=N1	o	CTGTGTGCTGAATAGCCGAGTCCCATGGCGC	AATAGC
539	OC(=O)C1=CC=C(OC2=CC=C3OCOC3=C2)N=C1	o	CTGTGTGCTGGTAGCGCGAGTCCCATGGCGC	GTAGCG
540	OC(=O)C1=NNC(=O)C=C1	o	CTGTGTGCTGAGACGGCGAGTCCCATGGCGC	AGACGG
541	CC1=NC2=CC=CC=C2N1CC(O)=O	o	CTGTGTGCTGTTAATTCGAGTCCCATGGCGC	TTAATT
542	OC(=O)C1=CNN=C1	o	CTGTGTGCTGAACATTCGAGTCCCATGGCGC	AACATT
543	CC1=CC=C(C(O)=O)C(O)=N1	o	CTGTGTGCTGAGCAATCGAGTCCCATGGCGC	AGCAAT
544	CC1=CC=C(O1)C1=NNC(=C1)C(O)=O	o	CTGTGTGCTGGTCTACCGAGTCCCATGGCGC	GTCTAC
545	OC(=O)C1=C(N=CC=N1)C(=O)N1CCCCC1	o	CTGTGTGCTGTAACAGCGAGTCCCATGGCGC	TAACAG
546	OC(=O)C1=CN=C2SC=CN2C1=O	o	CTGTGTGCTGAGCGTGCGAGTCCCATGGCGC	AGCGTG
547	NC(=O)C1(CC1)C(O)=O	o	CTGTGTGCTGCGAGCCCGAGTCCCATGGCGC	CGAGCC
548	CC1=NN2C(=C1)N=CC(C(O)=O)=C2C	o	CTGTGTGCTGCTGATACGAGTCCCATGGCGC	CTGATA
549	CC1=C(C(O)=O)C(C)=NO1	o	CTGTGTGCTGGTGAGGCGAGTCCCATGGCGC	GTGAGG
550	OC(=O)C1=CC=C(OC2=CC=CN=C2)O1	o	CTGTGTGCTGGAGCCGCGAGTCCCATGGCGC	GAGCCG
551	O[C@H](C(O)=O)C1=CC=CC=C1	o	CTGTGTGCTGGGTAATCGAGTCCCATGGCGC	GGTAAT
552	CNC(=O)C1=CC=C(C=C1)C(O)=O	o	CTGTGTGCTGAATAATCGAGTCCCATGGCGC	AATAAT
553	NS(=O)(=O)C1=CC=C(C=C1)C(O)=O	o	CTGTGTGCTGTAGACGCGAGTCCCATGGCGC	TAGACG
554	OC(=O)C1=C(O)N=CC=C1	o	CTGTGTGCTGCCTTAACGAGTCCCATGGCGC	CCTTAA
555	COC1=NN2C(CCC(O)=O)=NN=C2C=C1	o	CTGTGTGCTGTATAAGCGAGTCCCATGGCGC	TATAAG
556	Cl.CN(C)CC1=CNC2=C1C=CC(=C2)C(O)=O	o	CTGTGTGCTGCGACCTCGAGTCCCATGGCGC	CGACCT
557	CC1=C(CCC(O)=O)C(=O)NC(=O)N1	o	CTGTGTGCTGAGCCTCCGAGTCCCATGGCGC	AGCCTC
558	NC(=O)NC(CC(O)=O)C1=CC=CS1	o	CTGTGTGCTGCTTAGACGAGTCCCATGGCGC	CTTAGA
559	OC(=O)C1=CN=C(C=C1)N1C=NC=N1	o	CTGTGTGCTGGCCACACGAGTCCCATGGCGC	GCCACA
560	CC1=NC2=C(C=NN2C(C)=C1)C(O)=O	o	CTGTGTGCTGTGTATTCGAGTCCCATGGCGC	TGTATT
561	OC(=O)C1=NN(C(=O)C=C1)C1=CC=CC=C1	o	CTGTGTGCTGATCCGGCGAGTCCCATGGCGC	ATCCGG
562	OC(=O)C1=CNN=N1	o	CTGTGTGCTGAGGTGACGAGTCCCATGGCGC	AGGTGA
563	OC(=O)C1CCCN1C(=O)C1CC1	o	CTGTGTGCTGGTAAGCCGAGTCCCATGGCGC	GTAAGC
564	CN1NC(=O)C2=C1NC(=O)C(CC(O)=O)=C2C	o	CTGTGTGCTGACACCGCGAGTCCCATGGCGC	ACACCG
565	CC1=C(CC(O)=O)C(=O)NC(N)=N1	o	CTGTGTGCTGAGACCACGAGTCCCATGGCGC	AGACCA
566	OC(=O)C1=NNC2=C1CCC2	o	CTGTGTGCTGTTACAACGAGTCCCATGGCGC	TTACAA
567	CC1=C(C=NC=N1)C(O)=O	o	CTGTGTGCTGGACTCCCGAGTCCCATGGCGC	GACTCC
568	CC(NC(=O)C1=CC=C(Br)S1)C(O)=O	o	CTGTGTGCTGTGGCAGCGAGTCCCATGGCGC	TGGCAG
569	CC1=NNC(C(O)=O)=C1Br	o	CTGTGTGCTGCTAGAGCGAGTCCCATGGCGC	CTAGAG
570	CN1C2=C(NC(CCC(O)=O)=N2)C(=O)NC1=O	o	CTGTGTGCTGACTAGGCGAGTCCCATGGCGC	ACTAGG
571	OC(=O)C1=C(Br)C(=NN1)C1CC1	o	CTGTGTGCTGTAGAGCCGAGTCCCATGGCGC	TAGAGC
572	OC(=O)C1CC1C(=O)N1CCN(CC1)C1=CC=CC=C1	o	CTGTGTGCTGTTCGGCCGAGTCCCATGGCGC	TTCGGC
573	NC(=O)C1=CC=C(S1)C(O)=O	o	CTGTGTGCTGGTATCCCGAGTCCCATGGCGC	GTATCC
574	NC1=NC(Cl)=CC(=C1)C(O)=O	o	CTGTGTGCTGGACTAACGAGTCCCATGGCGC	GACTAA
575	OC1CC(N(C1)C(=O)C1=CC=C(F)C=C1)C(O)=O	o	CTGTGTGCTGCTAAGGCGAGTCCCATGGCGC	CTAAGG
576	CCC(NC1=CC=CC=C1)C(O)=O	o	CTGTGTGCTGGTGCGACGAGTCCCATGGCGC	GTGCGA
577	OC(=O)CCN1C=CNC(=O)C1=O	o	CTGTGTGCTGGTTCATCGAGTCCCATGGCGC	GTTCAT
578	OC(=O)CN1C=C2C=CC=CC2=N1	o	CTGTGTGCTGACGCGACGAGTCCCATGGCGC	ACGCGA
579	NC1=NNC(C(O)=O)=C1C1=CC=CC=C1	o	CTGTGTGCTGAGAGGCCGAGTCCCATGGCGC	AGAGGC
580	CN1N=C(C(O)=O)C(Br)=C1C	o	CTGTGTGCTGGATGATCGAGTCCCATGGCGC	GATGAT
581	CC1=NC(=NO1)C1=CC(=CC=C1)C(O)=O	o	CTGTGTGCTGGACCGACGAGTCCCATGGCGC	GACCGA
582	OC(=O)CN(CC1=CC=CC=C1)CC1=CC=CC=C1	o	CTGTGTGCTGGGCAAGCGAGTCCCATGGCGC	GGCAAG
583	OC(=O)C1=C(Br)SC=N1	o	CTGTGTGCTGGAATATCGAGTCCCATGGCGC	GAATAT
584	OC(=O)C1=CNC(=O)C(Br)=C1	o	CTGTGTGCTGGAAGCCCGAGTCCCATGGCGC	GAAGCC
585	OC(=O)C1=NC2=CC=CC=C2N=C1	o	CTGTGTGCTGACACAACGAGTCCCATGGCGC	ACACAA
586	NC1=C(N=C(Br)C=N1)C(O)=O	o	CTGTGTGCTGCGGCACCGAGTCCCATGGCGC	CGGCAC
587	CC1=CC(=NN1C1=CC(F)=CC=C1)C(O)=O	o	CTGTGTGCTGCAGATTCGAGTCCCATGGCGC	CAGATT
588	CC1=C(CCC(O)=O)C(=O)N2N=CN=C2N1	o	CTGTGTGCTGTCATCACGAGTCCCATGGCGC	TCATCA
589	CCC1=CC2=C(S1)N=CN=C2NCCC(O)=O	o	CTGTGTGCTGTTCCTTCGAGTCCCATGGCGC	TTCCTT
590	OC(=O)CN1C2=C(CCCC1=O)SC=C2	o	CTGTGTGCTGCAGAGGCGAGTCCCATGGCGC	CAGAGG
591	CN1C(=O)N(CCC(O)=O)C2=C1C=CC=C2	o	CTGTGTGCTGCAACCGCGAGTCCCATGGCGC	CAACCG
592	OC(=O)CN1C2=C(CCCC1=O)C=CC=C2	o	CTGTGTGCTGCAGCTCCGAGTCCCATGGCGC	CAGCTC
593	OC(=O)C1=NC=C(F)C=C1	o	CTGTGTGCTGCACAAGCGAGTCCCATGGCGC	CACAAG
594	OC(=O)CCN1C=C(Cl)C=N1	o	CTGTGTGCTGAACTATCGAGTCCCATGGCGC	AACTAT
595	OC(=O)C1CSC2(CCC(=O)N12)C1=CC=CC=C1	o	CTGTGTGCTGCTGCTGCGAGTCCCATGGCGC	CTGCTG
596	CC1=NN(C(C)=C1CC(O)=O)C1=CC=CC=C1	o	CTGTGTGCTGTCCTAACGAGTCCCATGGCGC	TCCTAA
597	OC(=O)C1=NC=C(CC2=CC=CC=C2)C=C1	o	CTGTGTGCTGGAGTCACGAGTCCCATGGCGC	GAGTCA
598	CCOC(=O)C1=C(NCC#C)N=C(Cl)C=C1	o	CTGTGTGCTGCTGACGCGAGTCCCATGGCGC	CTGACG
599	Cl.C#CCNCC1=CC=CO1	o	CTGTGTGCTGTCGCTCCGAGTCCCATGGCGC	TCGCTC
600	Cl.NC(CC#C)CC(F)(F)F	o	CTGTGTGCTGAGATTCCGAGTCCCATGGCGC	AGATTC
601	COC1=C(N)C=C(C=C1)C#C	o	CTGTGTGCTGGATGGCCGAGTCCCATGGCGC	GATGGC
602	Cl.NC(C#C)C1CCOCC1	o	CTGTGTGCTGAATTCCCGAGTCCCATGGCGC	AATTCC
603	Cl.NC1(CCC1)C#C	o	CTGTGTGCTGGGCTTCCGAGTCCCATGGCGC	GGCTTC

604	OC(=O)C1=CC2=C(C=C1)N(CC2)C(=O)C#C	o	CTGTGTGCTGGAATCGCGAGTCCCATGGCGC	GAATCG
605	NNC(=O)CCC#C	o	CTGTGTGCTGCGACGACGAGTCCCATGGCGC	CGACGA
606	CN1CCN(CCC#C)CC1	o	CTGTGTGCTGGGACCGCGAGTCCCATGGCGC	GGACCG
607	Cl.C#CC1CNC1	o	CTGTGTGCTGGCGAATCGAGTCCCATGGCGC	GCGAAT
608	CN(C)CCC#C	o	CTGTGTGCTGATTAAGCGAGTCCCATGGCGC	ATTAAG
609	C#CCCNC1COC1	o	CTGTGTGCTGGAGTGTCGAGTCCCATGGCGC	GAGTGT
610	OC1=C(O)C=CC(C(CN2C(C=C(C(NCC#C)=O)C=C3)=C3N=C2)=O)=C1	o	CTGTGTGCTGCTATCACGAGTCCCATGGCGC	CTATCA
611	OC1=C(C=CC(C(CN2C3=C(N=C2)C=C(C(NCC#C)=O)C=C3)=O)=C1)O	o	CTGTGTGCTGGCAGATCGAGTCCCATGGCGC	GCAGAT
612	OC1=C(O)C=CC(C(CN2C(C=CC=C3C(NCC#C)=O)=C3N=C2)=O)=C1	o	CTGTGTGCTGTGTCACCGAGTCCCATGGCGC	TGTCAC

8.2 List of building blocks B

Table 8: List of building blocks B (boronates and alkynes) and oligonucleotide codes B.

CdId	BB2_ID	smiles	code	codon
1	boronate_1	OB(O)c1cc(ccc1Cl)C#N	CGGATCGACGGTCTCACGCGTCAGGCAGC	GTCTCAC
2	boronate_3	COc1ccc(B(O)O)c(F)c1	CGGATCGACGGTCGTACGCGTCAGGCAGC	GTCGTAC
3	boronate_4	O.Nc1cccc(c1)B(O)O	CGGATCGACGCTCATTGGCGTCAGGCAGC	CTCATTG
4	boronate_5	Cc1cc(ccc1F)B(O)O	CGGATCGACGGTAGAGAGCGTCAGGCAGC	GTAGAGA
5	boronate_6	COc1ccc(cc1)B(O)O	CGGATCGACGGTTACCTGCGTCAGGCAGC	GTTACCT
6	boronate_7	Cc1cc(F)cc(c1)B(O)O	CGGATCGACGAGTAATTGCGTCAGGCAGC	AGTAATT
7	boronate_8	OCc1ccc(cc1)B(O)O	CGGATCGACGAGTGAGCGCGTCAGGCAGC	AGTGAGC
8	boronate_9	OB(O)c1ccc(OC(F)F)cc1	CGGATCGACGACTGATAGCGTCAGGCAGC	ACTGATA
9	boronate_10	COc1ccc(cc1C)B(O)O	CGGATCGACGACGTATAGCGTCAGGCAGC	ACGTATA
10	boronate_11	COc1ccc(Cl)cc1B(O)O	CGGATCGACGTTCTCCTGCGTCAGGCAGC	TTCTCCT
11	boronate_12	Cc1cccc(C)c1B(O)O	CGGATCGACGAACTGCTGCGTCAGGCAGC	AACTGCT
12	boronate_15	COc1cc(ccc1Cl)B(O)O	CGGATCGACGCACACACGCGTCAGGCAGC	CACACAC
13	boronate_16	Cl.NCc1ccc(cc1)B(O)O	CGGATCGACGCACGTGTGCGTCAGGCAGC	CACGTGT
14	boronate_17	OB(O)c1cccc(C#N)c1F	CGGATCGACGCGAGGTGGCGTCAGGCAGC	CGAGGTG
15	boronate_18	OB(O)c1cccc(c1)C(F)F	CGGATCGACGAAGCGAGGCGTCAGGCAGC	AAGCGAG
16	boronate_19	Cc1cccc(B(O)O)c1F	CGGATCGACGGCGCATGGCGTCAGGCAGC	GCGCATG
17	boronate_20	OB(O)c1c(F)cc(O)cc1F	CGGATCGACGGTTGGTCGCGTCAGGCAGC	GTTGGTC
18	boronate_21	Cc1ccc(F)c(c1)B(O)O	CGGATCGACGCTTCTCTGCGTCAGGCAGC	CTTCTCT
19	boronate_22	OB(O)c1cccc(C=O)c1F	CGGATCGACGTTGCACGGCGTCAGGCAGC	TTGCACG
20	boronate_23	COCc1ccc(cc1)B(O)O	CGGATCGACGGAGTAGAGCGTCAGGCAGC	GAGTAGA
21	boronate_24	COc1ccccc1B(O)O	CGGATCGACGATGTGAGGCGTCAGGCAGC	ATGTGAG
22	boronate_25	COCc1ccccc1B(O)O	CGGATCGACGAACGTATGCGTCAGGCAGC	AACGTAT
23	boronate_26	COc1cccc(c1)B(O)O	CGGATCGACGTCCGGCTGCGTCAGGCAGC	TCCGGCT
24	boronate_27	CC(C)Oc1ccc(cc1)B(O)O	CGGATCGACGTGATGATGCGTCAGGCAGC	TGATGAT
25	boronate_28	OB(O)c1ccc(O)c(Cl)c1	CGGATCGACGTGTGGACGCGTCAGGCAGC	TGTGGAC
26	boronate_29	OB(O)c1cccc(c1)[N+](=O)[O-]	CGGATCGACGGTAGTGCGCGTCAGGCAGC	GTAGTGC
27	boronate_30	OB(O)c1cc(F)ccc1C=C	CGGATCGACGGCAACACGCGTCAGGCAGC	GCAACAC
28	boronate_31	OB(O)c1cc(F)ccc1C=O	CGGATCGACGAAGACCGGCGTCAGGCAGC	AAGACCG
29	boronate_33	OB(O)c1cccc(c1)C(=O)O	CGGATCGACGAGAGAGAGCGTCAGGCAGC	AGAGAGA
30	boronate_34	OCCNS(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGTCGAGATGCGTCAGGCAGC	TCGAGAT
31	boronate_35	Cc1ccc(Cl)c(c1)B(O)O	CGGATCGACGCCGACTTGCGTCAGGCAGC	CCGACTT
32	boronate_36	CCOc1ccc(C)cc1B(O)O	CGGATCGACGTGAGATAGCGTCAGGCAGC	TGAGATA
33	boronate_37	OB(O)c1cc(F)cc(C=O)c1	CGGATCGACGTTGGCGTGCGTCAGGCAGC	TTGGCGT
34	boronate_38	COc1cccc(B(O)O)c1F	CGGATCGACGAATCCTCGCGTCAGGCAGC	AATCCTC
35	boronate_39	OB(O)c1cncc(F)c1	CGGATCGACGCACGTACGCGTCAGGCAGC	CACGTAC
36	boronate_40	OB(O)c1cccc2ncccc12	CGGATCGACGCACACGAGCGTCAGGCAGC	CACACGA
37	boronate_42	OB(O)c1ccc(F)c(C=O)c1	CGGATCGACGCGTAACAGCGTCAGGCAGC	CGTAACA
38	boronate_43	OCc1cc(F)cc(c1)B(O)O	CGGATCGACGAATTCCGGCGTCAGGCAGC	AATTCCG
39	boronate_44	CS(=O)(=O)c1ccccc1B(O)O	CGGATCGACGGCGTTACGCGTCAGGCAGC	GCGTTAC
40	boronate_45	OB(O)c1cc(O)ccc1Cl	CGGATCGACGCTCCATTGCGTCAGGCAGC	CTCCATT
41	boronate_47	OB(O)c1ccc(CC#N)cc1	CGGATCGACGCGCCGGTGCGTCAGGCAGC	CGCCGGT
42	boronate_48	Cc1cc(Cl)ccc1B(O)O	CGGATCGACGGTAAGACGCGTCAGGCAGC	GTAAGAC
43	boronate_49	OB(O)c1cccc(CC(=O)O)c1	CGGATCGACGGCTGAATGCGTCAGGCAGC	GCTGAAT
44	boronate_50	CCNC(=O)c1ccc(F)c(c1)B(O)O	CGGATCGACGATAAGGTGCGTCAGGCAGC	ATAAGGT
45	boronate_51	OB(O)c1ccc(C(=O)O)c(F)c1	CGGATCGACGATCATTCGCGTCAGGCAGC	ATCATTC
46	boronate_52	OB(O)c1ccc(F)cc1C=O	CGGATCGACGAGCGAGTGCGTCAGGCAGC	AGCGAGT
47	boronate_54	OB(O)c1ccc(cc1)C(F)F	CGGATCGACGCCAGACTGCGTCAGGCAGC	CCAGACT
48	boronate_55	Cc1ccc(cc1Cl)B(O)O	CGGATCGACGTGACCAGGCGTCAGGCAGC	TGACCAG
49	boronate_56	OB(O)c1ccc(cc1)[N+](=O)[O-]	CGGATCGACGGCCTACAGCGTCAGGCAGC	GCCTACA
50	boronate_57	COc1ccc(Cl)c(c1)B(O)O	CGGATCGACGGCCTCGTGCGTCAGGCAGC	GCCTCGT
51	boronate_59	COC(=O)c1cncc(c1)B(O)O	CGGATCGACGTGTCGTTGCGTCAGGCAGC	TGTCGTT
52	boronate_60	COCc1cccc(c1)B(O)O	CGGATCGACGGTCTGAAGCGTCAGGCAGC	GTCTGAA
53	boronate_61	OB(O)c1ccc(cc1)C(=O)O	CGGATCGACGCCGTACTGCGTCAGGCAGC	CCGTACT
54	boronate_62	OB(O)c1ccc2ncccc2c1	CGGATCGACGAGGTGTCGCGTCAGGCAGC	AGGTGTC
55	boronate_63	OB(O)c1ccc(Cl)cc1	CGGATCGACGTGCCTGGGCGTCAGGCAGC	TGCCTGG
56	boronate_65	COc1ccncc1B(O)O	CGGATCGACGAGGATGCGCGTCAGGCAGC	AGGATGC
57	boronate_66	OB(O)c1ccccc1C(=O)O	CGGATCGACGGTTATGCGCGTCAGGCAGC	GTTATGC
58	boronate_67	COc1cc(Cl)ccc1B(O)O	CGGATCGACGGTAGGAAGCGTCAGGCAGC	GTAGGAA
59	boronate_68	Cc1ccc(B(O)O)c(F)c1	CGGATCGACGGTGTCGTGCGTCAGGCAGC	GTGTCGT

60	boronate_69	Cc1ccc(Cl)cc1B(O)O	CGGATCGACGGCTCCTTGCGTCAGGCAGC	GCTCCTT
61	boronate_70	OB(O)c1cc(C=O)ccc1F	CGGATCGACGTTCTGAGGCGTCAGGCAGC	TTCTGAG
62	boronate_71	OCc1ccccc1B(O)O	CGGATCGACGTCATGGAGCGTCAGGCAGC	TCATGGA
63	boronate_72	COc1cc(ccc1F)B(O)O	CGGATCGACGATCGTAAGCGTCAGGCAGC	ATCGTAA
64	boronate_73	OB(O)C1=CCCCC1	CGGATCGACGGACTTATGCGTCAGGCAGC	GACTTAT
65	boronate_74	Cc1ccc(cc1F)B(O)O	CGGATCGACGCAACGTTGCGTCAGGCAGC	CAACGTT
66	boronate_75	Cc1cc(F)ccc1B(O)O	CGGATCGACGCGATACTGCGTCAGGCAGC	CGATACT
67	boronate_76	COc1cc(C)c(cc1C)B(O)O	CGGATCGACGTACGATGGCGTCAGGCAGC	TACGATG
68	boronate_77	OB(O)c1ccc(CC#N)cc1F	CGGATCGACGCCAGTGTGCGTCAGGCAGC	CCAGTGT
69	boronate_79	CCOc1ccc(B(O)O)c(C)c1	CGGATCGACGCCTGGTGGCGTCAGGCAGC	CCTGGTG
70	boronate_80	Cc1c(Cl)cccc1B(O)O	CGGATCGACGCCAGTTGGCGTCAGGCAGC	CCAGTTG
71	boronate_83	OB(O)c1ccc(Cl)c(c1)C#N	CGGATCGACGTCATCGTGCGTCAGGCAGC	TCATCGT
72	boronate_84	OB(O)c1ccc2ccncc2c1	CGGATCGACGATATATCGCGTCAGGCAGC	ATATATC
73	boronate_86	COc1ccc(B(O)O)c(OC)c1	CGGATCGACGGTGCCGAGCGTCAGGCAGC	GTGCCGA
74	boronate_87	NC(=O)c1ccc(cc1)B(O)O	CGGATCGACGCAGACCAGCGTCAGGCAGC	CAGACCA
75	boronate_89	OB(O)c1c(F)cccc1C=O	CGGATCGACGCGATTGCGCGTCAGGCAGC	CGATTGC
76	boronate_90	COc1ccc(OC)c(c1)B(O)O	CGGATCGACGTACCTACGCGTCAGGCAGC	TACCTAC
77	boronate_91	OB(O)c1cncc(Cl)c1	CGGATCGACGGATGAGCGCGTCAGGCAGC	GATGAGC
78	boronate_92	OB(O)c1cccc(F)c1Cl	CGGATCGACGCAGGTTCGCGTCAGGCAGC	CAGGTTC
79	boronate_93	COc1ccc(F)c(c1)B(O)O	CGGATCGACGATAACTAGCGTCAGGCAGC	ATAACTA
80	boronate_95	OB(O)c1ccc(C=O)cc1F	CGGATCGACGACGTCCGGCGTCAGGCAGC	ACGTCCG
81	boronate_96	OB(O)c1cccc(CCl)c1	CGGATCGACGGTGCATAGCGTCAGGCAGC	GTGCATA
82	boronate_97	OB(O)c1ccc(Cl)cc1C=O	CGGATCGACGGAATCAAGCGTCAGGCAGC	GAATCAA
83	boronate_98	OCc1ccc(B(O)O)c(F)c1	CGGATCGACGACTTGCGGCGTCAGGCAGC	ACTTGCG
84	boronate_99	COc1cccc(Cl)c1B(O)O	CGGATCGACGTGTTCGTGCGTCAGGCAGC	TGTTCGT
85	boronate_100	CC(C)c1ccccc1B(O)O	CGGATCGACGCCTCCGCGCGTCAGGCAGC	CCTCCGC
86	boronate_101	COc1cc(Cl)cc(c1)B(O)O	CGGATCGACGCTCATATGCGTCAGGCAGC	CTCATAT
87	boronate_102	CCOc1ccc(cc1F)B(O)O	CGGATCGACGCTGAAGGGCGTCAGGCAGC	CTGAAGG
88	boronate_103	OB(O)c1cc(F)cc(F)c1	CGGATCGACGTCTAGCTGCGTCAGGCAGC	TCTAGCT
89	boronate_104	OB(O)c1ccc(cc1)C2CC2	CGGATCGACGTTCCTGTGCGTCAGGCAGC	TTCCTGT
90	boronate_105	Cc1ccc(cc1N)B(O)O	CGGATCGACGGACTGGAGCGTCAGGCAGC	GACTGGA
91	boronate_106	COc1c(Cl)cccc1B(O)O	CGGATCGACGTTAACCGGCGTCAGGCAGC	TTAACCG
92	boronate_107	OB(O)c1ccc(C=O)cc1	CGGATCGACGAGACTGAGCGTCAGGCAGC	AGACTGA
93	boronate_108	Cc1ccccc1B(O)O	CGGATCGACGATCTTGCGCGTCAGGCAGC	ATCTTGC
94	boronate_109	Nc1cc(ccc1F)B(O)O	CGGATCGACGCTAAGGCGCGTCAGGCAGC	CTAAGGC
95	boronate_110	OB(O)c1ccc(F)c(Cl)c1	CGGATCGACGATCGCATGCGTCAGGCAGC	ATCGCAT
96	boronate_111	OB(O)c1cc(C=O)ccc1Cl	CGGATCGACGAAGTCCAGCGTCAGGCAGC	AAGTCCA
97	boronate_112	OB(O)c1ccc(F)cc1Cl	CGGATCGACGCATTACGGCGTCAGGCAGC	CATTACG
98	boronate_113	OB(O)c1cccc(O)c1	CGGATCGACGATAGCCTGCGTCAGGCAGC	ATAGCCT
99	boronate_115	CC(=O)c1ccc(F)c(c1)B(O)O	CGGATCGACGCCAGGTAGCGTCAGGCAGC	CCAGGTA
100	boronate_116	COc1ccc(B(O)O)c(C=O)c1	CGGATCGACGAGTAGTAGCGTCAGGCAGC	AGTAGTA
101	boronate_117	OB(O)c1ccc(F)c(O)c1	CGGATCGACGTATGGAGGCGTCAGGCAGC	TATGGAG
102	boronate_118	Cc1ccc(cc1)B(O)O	CGGATCGACGAGCACGAGCGTCAGGCAGC	AGCACGA
103	boronate_119	OB(O)c1cccc(F)c1	CGGATCGACGAATTGCAGCGTCAGGCAGC	AATTGCA
104	boronate_120	OB(O)c1ccccc1	CGGATCGACGCAGATTGGCGTCAGGCAGC	CAGATTG
105	boronate_122	OB(O)c1cc(ccc1F)C#N	CGGATCGACGGTCCAAGGCGTCAGGCAGC	GTCCAAG
106	boronate_123	OB(O)c1ccc(nc1)C(F)(F)F	CGGATCGACGATGCGCTGCGTCAGGCAGC	ATGCGCT
107	boronate_124	OB(O)c1ccccc1F	CGGATCGACGACATAGTGCGTCAGGCAGC	ACATAGT
108	boronate_125	CNC(=O)c1ccc(cc1Cl)B(O)O	CGGATCGACGATAGAGCGCGTCAGGCAGC	ATAGAGC
109	boronate_127	CC(C)c1cccc(c1)B(O)O	CGGATCGACGTATGTCGGCGTCAGGCAGC	TATGTCG
110	boronate_128	Cc1ccc(cn1)B(O)O	CGGATCGACGTATCATTGCGTCAGGCAGC	TATCATT
111	boronate_130	OB(O)c1cc(F)cc(Cl)c1	CGGATCGACGCCTTCCGGCGTCAGGCAGC	CCTTCCG
112	boronate_131	Cc1cccc(B(O)O)c1C	CGGATCGACGTGCGTCGGCGTCAGGCAGC	TGCGTCG
113	boronate_132	OB(O)c1cc(Cl)ccc1F	CGGATCGACGAGAAGTGGCGTCAGGCAGC	AGAAGTG
114	boronate_134	CCCc1ccc(cc1)B(O)O	CGGATCGACGCGTAGGAGCGTCAGGCAGC	CGTAGGA
115	boronate_135	COc1ccc(F)cc1B(O)O	CGGATCGACGCTGTTAGGCGTCAGGCAGC	CTGTTAG
116	boronate_136	OB(O)c1ccc2cc[nH]c2c1	CGGATCGACGACGATCAGCGTCAGGCAGC	ACGATCA
117	boronate_137	Cc1ccc(B(O)O)c(C)c1	CGGATCGACGTCGTACAGCGTCAGGCAGC	TCGTACA
118	boronate_139	OB(O)c1ccc(F)c2ccccc12	CGGATCGACGCTATTATGCGTCAGGCAGC	CTATTAT
119	boronate_141	CCc1ccc(cc1)B(O)O	CGGATCGACGCGCAGGCGCGTCAGGCAGC	CGCAGGC
120	boronate_142	OB(O)c1cc(F)cc(c1)C#N	CGGATCGACGTAGCTTCGCGTCAGGCAGC	TAGCTTC
121	boronate_143	OB(O)c1cc(F)c(F)c(F)c1	CGGATCGACGCCTTCTCGCGTCAGGCAGC	CCTTCTC
122	boronate_144	OB(O)c1cc(F)ccc1F	CGGATCGACGCTGGCCGGCGTCAGGCAGC	CTGGCCG
123	boronate_145	OB(O)c1cc(O)cc(F)c1	CGGATCGACGGTCGAGCGCGTCAGGCAGC	GTCGAGC
124	boronate_146	Nc1cccc(c1)B(O)O	CGGATCGACGTAGATTAGCGTCAGGCAGC	TAGATTA
125	boronate_147	CNC(=O)c1ccc(cc1)B(O)O	CGGATCGACGTTGATACGCGTCAGGCAGC	TTGATAC
126	boronate_148	OB(O)c1cccc(C=O)c1	CGGATCGACGGCACAATGCGTCAGGCAGC	GCACAAT
127	boronate_149	OB(O)c1cccc(F)c1F	CGGATCGACGGCTTGAGGCGTCAGGCAGC	GCTTGAG
128	boronate_150	CC(=O)c1cccc(c1)B(O)O	CGGATCGACGTACTTGGGCGTCAGGCAGC	TACTTGG
129	boronate_151	OB(O)c1cc(Cl)ccc1Cl	CGGATCGACGCCACATAGCGTCAGGCAGC	CCACATA
130	boronate_152	CCc1ccccc1B(O)O	CGGATCGACGGACAGTCGCGTCAGGCAGC	GACAGTC
131	boronate_154	OB(O)c1ccc(F)cc1F	CGGATCGACGCGCGTTAGCGTCAGGCAGC	CGCGTTA
132	boronate_155	COc1ccc(cc1OC)B(O)O	CGGATCGACGATCTCCGGCGTCAGGCAGC	ATCTCCG
133	boronate_156	OB(O)c1ccccc1Cl	CGGATCGACGCTTGCACGCGTCAGGCAGC	CTTGCAC
134	boronate_157	CCS(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGTGTCACTGCGTCAGGCAGC	TGTCACT
135	boronate_158	COc1cccc(B(O)O)c1OC	CGGATCGACGGTGCGTGGCGTCAGGCAGC	GTGCGTG
136	boronate_159	CC(=O)c1cccc1B(O)O	CGGATCGACGACGCATCGCGTCAGGCAGC	ACGCATC

137	boronate_160	COc1cc(cc(F)c1F)B(O)O	CGGATCGACGGACGCGTGCGTCAGGCAGC	GACGCGT
138	boronate_161	OB(O)c1ccc(F)cc1	CGGATCGACGAGCGACGGCGTCAGGCAGC	AGCGACG
139	boronate_162	CCOC(=O)c1cccc(B(O)O)c1F	CGGATCGACGGCCGTAGGCGTCAGGCAGC	GCCGTAG
140	boronate_163	OCc1cc(ccc1F)B(O)O	CGGATCGACGCTCAGCAGCGTCAGGCAGC	CTCAGCA
141	boronate_164	Cc1cccc(c1)B(O)O	CGGATCGACGCTTACCAGCGTCAGGCAGC	CTTACCA
142	boronate_165	Cl.Nc1cccc(c1)B(O)O	CGGATCGACGGCAGGTGGCGTCAGGCAGC	GCAGGTG
143	boronate_166	CS(=O)c1ccccc1B(O)O	CGGATCGACGCCGGCTGGCGTCAGGCAGC	CCGGCTG
144	boronate_167	OB(O)c1cccc(c1)C#N	CGGATCGACGCAACAACGCGTCAGGCAGC	CAACAAC
145	boronate_168	CS(=O)(=O)Nc1ccccc1B(O)O	CGGATCGACGCGTTCAGGCGTCAGGCAGC	CGTTCAG
146	boronate_169	COc1c(Cl)cc(cc1Cl)B(O)O	CGGATCGACGCCACGAAGCGTCAGGCAGC	CCACGAA
147	boronate_170	OB(O)c1ccc(F)nc1	CGGATCGACGCGGAGAGGCGTCAGGCAGC	CGGAGAG
148	boronate_171	NC(=O)c1ccc(cc1F)B(O)O	CGGATCGACGGTCATGAGCGTCAGGCAGC	GTCATGA
149	boronate_173	OB(O)c1cc(Cl)cc(Cl)c1	CGGATCGACGACTGACGGCGTCAGGCAGC	ACTGACG
150	boronate_176	NC(=O)c1cc(cc(c1)[N+](=O)[O-])B(O)O	CGGATCGACGTGACGGAGCGTCAGGCAGC	TGACGGA
151	boronate_177	COc1ccc(OC)c(B(O)O)c1OC	CGGATCGACGTCTTATTGCGTCAGGCAGC	TCTTATT
152	boronate_178	OB(O)c1cccc(OC(F)F)c1	CGGATCGACGATACTACGCGTCAGGCAGC	ATACTAC
153	boronate_180	COC(=O)c1ccc(cc1OC)B(O)O	CGGATCGACGTAGCCGTGCGTCAGGCAGC	TAGCCGT
154	boronate_181	OB(O)c1ccc(F)c(F)c1	CGGATCGACGTCCACGGGCGTCAGGCAGC	TCCACGG
155	boronate_182	OCc1ccc(Cl)c(c1)B(O)O	CGGATCGACGCGGCTTCGCGTCAGGCAGC	CGGCTTC
156	boronate_183	Cc1c(F)cccc1B(O)O	CGGATCGACGTGTGCTTGCGTCAGGCAGC	TGTGCTT
157	boronate_184	COc1cc(cc(OC)c1OC)B(O)O	CGGATCGACGTTGTCTTGCGTCAGGCAGC	TTGTCTT
158	boronate_185	OB(O)c1ccc(CBr)cc1	CGGATCGACGATAGTCAGCGTCAGGCAGC	ATAGTCA
159	boronate_186	OB(O)c1cccc(F)c1O	CGGATCGACGTGGAGTAGCGTCAGGCAGC	TGGAGTA
160	boronate_187	COc1cc(OC)cc(c1)B(O)O	CGGATCGACGCGGATGGGCGTCAGGCAGC	CGGATGG
161	boronate_188	CC(=O)c1ccc(cc1)B(O)O	CGGATCGACGGCTACCAGCGTCAGGCAGC	GCTACCA
162	boronate_189	OB(O)c1ccc(F)c(c1)C#N	CGGATCGACGACAACGAGCGTCAGGCAGC	ACAACGA
163	boronate_190	OB(O)c1ccc(cc1[N+](=O)[O-])C(=O)O	CGGATCGACGTTGATGAGCGTCAGGCAGC	TTGATGA
164	boronate_191	OB(O)c1ccccc1C2CC2	CGGATCGACGGAGCCGCGCGTCAGGCAGC	GAGCCGC
165	boronate_192	OB(O)c1ccc2[nH]ccc2c1	CGGATCGACGGATTAATGCGTCAGGCAGC	GATTAAT
166	boronate_193	OB(O)c1ccc(OCC#N)cc1	CGGATCGACGGATGCGGGCGTCAGGCAGC	GATGCGG
167	boronate_194	CC(=O)NCc1ccc(cc1)B(O)O	CGGATCGACGTAACGTAGCGTCAGGCAGC	TAACGTA
168	boronate_195	CC(C)(C)OC(=O)n1cccc1B(O)O	CGGATCGACGACGCGTTGCGTCAGGCAGC	ACGCGTT
169	boronate_197	OB(O)c1ccc(Cl)c(Cl)c1	CGGATCGACGCGGCCGTGCGTCAGGCAGC	CGGCCGT
170	boronate_198	CC(C)(C)NS(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGATGGTTAGCGTCAGGCAGC	ATGGTTA
171	boronate_199	OB(O)c1cccc(c1)C(F)(F)F	CGGATCGACGGTCGCCGGCGTCAGGCAGC	GTCGCCG
172	boronate_200	Cc1nn(C)c2cc(ccc12)B(O)O	CGGATCGACGTACGTCAGCGTCAGGCAGC	TACGTCA
173	boronate_201	OB(O)c1cc(F)cc(CC#N)c1	CGGATCGACGAGCTCTTGCGTCAGGCAGC	AGCTCTT
174	boronate_203	Cc1cc(ccc1B(O)O)S(=O)(=O)N	CGGATCGACGCATTGTTGCGTCAGGCAGC	CATTGTT
175	boronate_204	OB(O)c1ccc(C=O)s1	CGGATCGACGACAGGAAGCGTCAGGCAGC	ACAGGAA
176	boronate_205	OB(O)c1ccccc1C#N	CGGATCGACGTCTATGCGCGTCAGGCAGC	TCTATGC
177	boronate_206	OB(O)c1oc2ccccc2c1	CGGATCGACGTAATACAGCGTCAGGCAGC	TAATACA
178	boronate_207	COc1ccc(C=O)cc1B(O)O	CGGATCGACGTACTCGAGCGTCAGGCAGC	TACTCGA
179	boronate_208	CS(=O)(=O)c1cccc(c1)B(O)O	CGGATCGACGTGTCTAGGCGTCAGGCAGC	TGTCTAG
180	boronate_209	OB(O)c1cccc(Cl)c1F	CGGATCGACGTTCGCTCGCGTCAGGCAGC	TTCGCTC
181	boronate_210	CC(=O)Nc1cccc(c1)B(O)O	CGGATCGACGGCTCAGTGCGTCAGGCAGC	GCTCAGT
182	boronate_211	OB(O)c1cccc(OC(F)(F)F)c1F	CGGATCGACGTACAGGAGCGTCAGGCAGC	TACAGGA
183	boronate_212	Cc1ccc(cc1B(O)O)S(=O)(=O)N2CCCC2	CGGATCGACGCCGGAAGGCGTCAGGCAGC	CCGGAAG
184	boronate_213	CN(C)S(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGCCACGCTGCGTCAGGCAGC	CCACGCT
185	boronate_214	CS(=O)(=O)Nc1cccc(c1)B(O)O	CGGATCGACGAGATAAGGCGTCAGGCAGC	AGATAAG
186	boronate_215	OB(O)c1ccc(OC2CCCCO2)cc1	CGGATCGACGATCAGACGCGTCAGGCAGC	ATCAGAC
187	boronate_216	CC(C)(C)OC(=O)c1ccc(cc1)B(O)O	CGGATCGACGATCTCAAGCGTCAGGCAGC	ATCTCAA
188	boronate_217	CS(=O)(=O)c1ccc(B(O)O)c(F)c1	CGGATCGACGGTAACTCGCGTCAGGCAGC	GTAACTC
189	boronate_218	OB(O)c1ccc2cn[nH]c2c1	CGGATCGACGACAGTAGGCGTCAGGCAGC	ACAGTAG
190	boronate_219	OB(O)c1ccc(cc1)N2CCOCC2	CGGATCGACGCGCTGTAGCGTCAGGCAGC	CGCTGTA
191	boronate_220	OB(O)c1ccccc1C=O	CGGATCGACGGAGCCATGCGTCAGGCAGC	GAGCCAT
192	boronate_221	OB(O)c1cc(F)ccc1Cl	CGGATCGACGACTCTTGGCGTCAGGCAGC	ACTCTTG
193	boronate_223	OB(O)c1ccc(Cl)cc1F	CGGATCGACGTGTAGAGGCGTCAGGCAGC	TGTAGAG
194	boronate_224	OB(O)c1ccccc1OCC#N	CGGATCGACGAGAACAAGCGTCAGGCAGC	AGAACAA
195	boronate_226	OB(O)c1ccccc1C(F)(F)F	CGGATCGACGATCTACTGCGTCAGGCAGC	ATCTACT
196	boronate_227	Cc1c(cc(cc1[N+](=O)[O-])[N+](=O)[O-])B(O)O	CGGATCGACGGATAACGGCGTCAGGCAGC	GATAACG
197	boronate_228	COc1ncc(cn1)B(O)O	CGGATCGACGTTACGCTGCGTCAGGCAGC	TTACGCT
198	boronate_229	CC(C)(C#N)c1ccc(cc1)B(O)O	CGGATCGACGTAGACGAGCGTCAGGCAGC	TAGACGA
199	boronate_230	CNC(=O)c1cccc(c1)B(O)O	CGGATCGACGAAGCATGGCGTCAGGCAGC	AAGCATG
200	boronate_232	OB(O)c1ccc(cc1)C(=O)N2CCC(=O)CC2	CGGATCGACGGACTTCAGCGTCAGGCAGC	GACTTCA
201	boronate_233	CC(C)(C)OC(=O)NCc1ccc(F)cc1B(O)O	CGGATCGACGATGTGTTGCGTCAGGCAGC	ATGTGTT
202	boronate_234	OB(O)c1ccc(Cl)cc1Cl	CGGATCGACGGTATGGAGCGTCAGGCAGC	GTATGGA
203	boronate_235	OB(O)c1cccc(CC#N)c1F	CGGATCGACGCGGTTAAGCGTCAGGCAGC	CGGTTAA
204	boronate_237	OB(O)c1ccc(Cl)c(F)c1	CGGATCGACGTTACGAGGCGTCAGGCAGC	TTACGAG
205	boronate_238	OB(O)c1cccnc1Cl	CGGATCGACGATTCTGGGCGTCAGGCAGC	ATTCTGG
206	boronate_239	OB(O)c1cn[nH]c1	CGGATCGACGCTATGCGGCGTCAGGCAGC	CTATGCG
207	boronate_240	COc1ccc(B(O)O)c(c1)C(F)(F)F	CGGATCGACGCCGTTATGCGTCAGGCAGC	CCGTTAT
208	boronate_241	COc1cncc(c1)B(O)O	CGGATCGACGACGCTGTGCGTCAGGCAGC	ACGCTGT
209	boronate_242	COc1cc2ccc(cc2cn1)B(O)O	CGGATCGACGCTAGTGAGCGTCAGGCAGC	CTAGTGA
210	boronate_243	OB(O)c1cc(ccc1Cl)C(F)(F)F	CGGATCGACGTTACAGGGCGTCAGGCAGC	TTACAGG
211	boronate_244	O.CN(C)c1ccc(cn1)B(O)O	CGGATCGACGGTAGCCAGCGTCAGGCAGC	GTAGCCA
212	boronate_245	CS(=O)(=O)NCc1ccc(cc1)B(O)O	CGGATCGACGCCTGTCGGCGTCAGGCAGC	CCTGTCG
213	boronate_246	OB(O)c1cc(Cl)cc(c1)C(F)(F)F	CGGATCGACGGCACGTCGCGTCAGGCAGC	GCACGTC

214	boronate_248	CC(C)(C)NS(=O)(=O)c1cccc(c1)B(O)O	CGGATCGACGTTGGCTAGCGTCAGGCAGC	TTGGCTA
215	boronate_249	OB(O)c1cccc(CC(=O)O)c1F	CGGATCGACGTTAAGTGGCGTCAGGCAGC	TTAAGTG
216	boronate_250	OB(O)c1ccc(CC(=O)O)cc1F	CGGATCGACGTAGGAACGCGTCAGGCAGC	TAGGAAC
217	boronate_251	CC(C)OC(=O)c1cccc(c1)B(O)O	CGGATCGACGGTGATGGGCGTCAGGCAGC	GTGATGG
218	boronate_253	OB(O)c1cc(OC(F)(F)F)ccc1F	CGGATCGACGACAAGAGGCGTCAGGCAGC	ACAAGAG
219	boronate_256	OB(O)c1ccc2OC(F)(F)C(F)(F)Oc2c1	CGGATCGACGACATGATGCGTCAGGCAGC	ACATGAT
220	boronate_257	OB(O)c1ccc(cc1)C(F)(F)F	CGGATCGACGTTGCCTGGCGTCAGGCAGC	TTGCCTG
221	boronate_258	OB(O)c1ccc(Cl)cc1C(F)(F)F	CGGATCGACGGAGTCTTGCGTCAGGCAGC	GAGTCTT
222	boronate_259	COc1cc(C)c(cn1)B(O)O	CGGATCGACGCGCGGAGGCGTCAGGCAGC	CGCGGAG
223	boronate_260	CS(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGATCTGTCGCGTCAGGCAGC	ATCTGTC
224	boronate_261	OB(O)c1ccc(Cl)c(c1)C(=O)O	CGGATCGACGCGTAGTGGCGTCAGGCAGC	CGTAGTG
225	boronate_262	OB(O)c1ccc(F)nc1F	CGGATCGACGGAGCAGGGCGTCAGGCAGC	GAGCAGG
226	boronate_263	CC(C)(C)OC(=O)c1cccc(c1)B(O)O	CGGATCGACGAGCTGAAGCGTCAGGCAGC	AGCTGAA
227	boronate_264	OCCNC(=O)c1cccc(c1)B(O)O	CGGATCGACGTCGCCGCGCGTCAGGCAGC	TCGCCGC
228	boronate_265	OB(O)c1ccc(OC(F)(F)F)cc1	CGGATCGACGCGCCTGCGCGTCAGGCAGC	CGCCTGC
229	boronate_266	COc1ccc(cc1B(O)O)C(F)(F)F	CGGATCGACGAACGACTGCGTCAGGCAGC	AACGACT
230	boronate_267	OB(O)c1cccc2cc(Cl)ncc12	CGGATCGACGGCGTATTGCGTCAGGCAGC	GCGTATT
231	boronate_268	OB(O)c1cccc2cnc(Cl)cc12	CGGATCGACGCACGGTCGCGTCAGGCAGC	CACGGTC
232	boronate_269	OB(O)c1cc(O)cc(OC(F)(F)F)c1	CGGATCGACGGAGCAACGCGTCAGGCAGC	GAGCAAC
233	boronate_270	OB(O)c1ccc(cc1Cl)C(F)(F)F	CGGATCGACGATTAAGTGCGTCAGGCAGC	ATTAAGT
234	boronate_271	COc1cc2c(cccc2cn1)B(O)O	CGGATCGACGGAGTATGGCGTCAGGCAGC	GAGTATG
235	boronate_274	OB(O)c1cc2ccccc2nc1Cl	CGGATCGACGCCTTGGCGCGTCAGGCAGC	CCTTGGC
236	boronate_275	Cn1ncc2cc(F)c(cc12)B(O)O	CGGATCGACGCTTCCTCGCGTCAGGCAGC	CTTCCTC
237	boronate_276	COc1ncc(B(O)O)c(OC)n1	CGGATCGACGTTGGTCGGCGTCAGGCAGC	TTGGTCG
238	boronate_279	OB(O)c1ccc(c(Cl)c1)C(F)(F)F	CGGATCGACGGTGGCTTGCGTCAGGCAGC	GTGGCTT
239	boronate_280	OB(O)c1cccc2c(Cl)nccc12	CGGATCGACGTCAACTCGCGTCAGGCAGC	TCAACTC
240	boronate_281	CC(C)(C)Oc1ncccc1B(O)O	CGGATCGACGGTGAGTCGCGTCAGGCAGC	GTGAGTC
241	boronate_282	Cn1ncc2cccc(B(O)O)c12	CGGATCGACGAATGATGGCGTCAGGCAGC	AATGATG
242	boronate_284	OB(O)c1cocc1	CGGATCGACGCCAGCTCGCGTCAGGCAGC	CCAGCTC
243	boronate_285	OB(O)c1ccc(cc1)S(=O)(=O)n2cccc2	CGGATCGACGAACAAGGGCGTCAGGCAGC	AACAAGG
244	boronate_286	OB(O)c1ccc2OCCOc2c1	CGGATCGACGGCTGCCGGCGTCAGGCAGC	GCTGCCG
245	boronate_287	Cl.COc1ccncc1B(O)O	CGGATCGACGTTGTCCGGCGTCAGGCAGC	TTGTCCG
246	boronate_288	CC(C)(C)OCc1cccc(c1)B(O)O	CGGATCGACGGCGTAGGGCGTCAGGCAGC	GCGTAGG
247	boronate_289	OB(O)c1cccc2[nH]ncc12	CGGATCGACGGAACCTGGCGTCAGGCAGC	GAACCTG
248	boronate_290	CNC(=O)c1cc(ccc1F)B(O)O	CGGATCGACGAGGTAGGGCGTCAGGCAGC	AGGTAGG
249	boronate_291	OB(O)c1cccc2cnccc12	CGGATCGACGCAATATTGCGTCAGGCAGC	CAATATT
250	boronate_292	COC(=O)Nc1ccc(cc1)B(O)O	CGGATCGACGATCACTGGCGTCAGGCAGC	ATCACTG
251	boronate_293	COc1ccc(cn1)B(O)O	CGGATCGACGCAACGAGGCGTCAGGCAGC	CAACGAG
252	boronate_294	OB(O)c1ccc(cc1)C2CC(=O)NN2	CGGATCGACGCGAACGCGCGTCAGGCAGC	CGAACGC
253	boronate_296	OB(O)c1ccccc1S(=O)(=O)N2CCOCC2	CGGATCGACGGCTCGCTGCGTCAGGCAGC	GCTCGCT
254	boronate_297	CC1(C)OB(OC1(C)C)c2ccc(N)cc2	CGGATCGACGTAACTAGGCGTCAGGCAGC	TAACTAG
255	boronate_298	CCN(CC)S(=O)(=O)c1ccc(C)c(c1)B(O)O	CGGATCGACGCTTCTACGCGTCAGGCAGC	CTTCTAC
256	boronate_299	COc1ncccc1B(O)O	CGGATCGACGATATGCTGCGTCAGGCAGC	ATATGCT
257	boronate_301	Cn1ncc2cc(ccc12)B(O)O	CGGATCGACGGCAAGCTGCGTCAGGCAGC	GCAAGCT
258	boronate_302	CCOc1ncccc1B(O)O	CGGATCGACGCGTGATTGCGTCAGGCAGC	CGTGATT
259	boronate_303	CC1(C)OB(OC1(C)C)c2ccc(cc2F)C#N	CGGATCGACGCCGAATCGCGTCAGGCAGC	CCGAATC
260	boronate_304	Cl.COC(=O)c1cc(N)cc(c1)B2OC(C)(C)C(C)(C)O2	CGGATCGACGTCAGGCGGCGTCAGGCAGC	TCAGGCG
261	boronate_305	OB(O)c1cccc(CBr)c1F	CGGATCGACGCATGCGTGCGTCAGGCAGC	CATGCGT
262	boronate_306	OB(O)c1cc(F)cc(c1)C(F)(F)F	CGGATCGACGTGGAAGCGCGTCAGGCAGC	TGGAAGC
263	boronate_307	Cc1cc(ccc1B(O)O)C(F)(F)F	CGGATCGACGCTTAACTGCGTCAGGCAGC	CTTAACT
264	boronate_308	CCOc1ccc(Cl)cc1B(O)O	CGGATCGACGAAGGCGTGCGTCAGGCAGC	AAGGCGT
265	boronate_309	O.OB(O)c1ccc(F)nc1F	CGGATCGACGAGGTTCTGCGTCAGGCAGC	AGGTTCT
266	boronate_310	OB(O)c1ccc(C(=O)O)c(Cl)c1	CGGATCGACGTTACAATGCGTCAGGCAGC	TTACAAT
267	boronate_311	CC1CCC(=CC1)B(O)O	CGGATCGACGCGACGACGCGTCAGGCAGC	CGACGAC
268	boronate_312	COc1ccc(C=O)c(c1)B(O)O	CGGATCGACGCGTGAAGGCGTCAGGCAGC	CGTGAAG
269	boronate_314	CC(C)(C)OC(=O)Nc1cccc(c1)B(O)O	CGGATCGACGACACCGGGCGTCAGGCAGC	ACACCGG
270	boronate_315	OB(O)c1ccc(Cl)c(c1)[N+](=O)[O-]	CGGATCGACGAACCTTAGCGTCAGGCAGC	AACCTTA
271	boronate_316	NS(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGAGTTCGGGCGTCAGGCAGC	AGTTCGG
272	boronate_317	COc1ccc(cc1B(O)O)C(=O)O	CGGATCGACGTGATTCTGCGTCAGGCAGC	TGATTCT
273	boronate_320	OB(O)c1ccc2[nH]ncc2c1	CGGATCGACGCTCGAGTGCGTCAGGCAGC	CTCGAGT
274	boronate_321	CC1(C)OB(OC1(C)C)c2ccccc2OC(F)F	CGGATCGACGTACACTCGCGTCAGGCAGC	TACACTC
275	boronate_322	OB(O)c1cc(OC(F)(F)F)ccc1Cl	CGGATCGACGCGTGTACGCGTCAGGCAGC	CGTGTAC
276	boronate_323	Cc1cc(Cl)ncc1B(O)O	CGGATCGACGCCGCTGAGCGTCAGGCAGC	CCGCTGA
277	boronate_324	OB(O)c1ccc(s1)C(=O)O	CGGATCGACGACGCGCAGCGTCAGGCAGC	ACGCGCA
278	boronate_325	OB(O)c1cn(c2ccccc12)S(=O)(=O)c3ccccc3	CGGATCGACGGCCGCTTGCGTCAGGCAGC	GCCGCTT
279	boronate_326	OB(O)c1ccsc1	CGGATCGACGCACCTCTGCGTCAGGCAGC	CACCTCT
280	boronate_327	CC(C)(C)OC(=O)Nc1ccc(cc1)B(O)O	CGGATCGACGGCTTCACGCGTCAGGCAGC	GCTTCAC
281	boronate_329	OB(O)c1ccc2OCCCOc2c1	CGGATCGACGGACCAGCGCGTCAGGCAGC	GACCAGC
282	boronate_330	CNC(=O)c1ccc(cc1F)B(O)O	CGGATCGACGTCTTCAGGCGTCAGGCAGC	TCTTCAG
283	boronate_331	CC1(C)OB(OC1(C)C)c2ccc(F)c(F)c2F	CGGATCGACGTAGTCGGGCGTCAGGCAGC	TAGTCGG
284	boronate_332	CS(=O)(=O)c1cncc(c1)B(O)O	CGGATCGACGGTGTCAAGCGTCAGGCAGC	GTGTCAA
285	boronate_333	COc1ncc(F)cc1B(O)O	CGGATCGACGTAGGTCTGCGTCAGGCAGC	TAGGTCT
286	boronate_334	OB(O)c1ccc2cccnc2c1	CGGATCGACGGAACTACGCGTCAGGCAGC	GAACTAC
287	boronate_335	OB(O)c1cnc2ccccc2c1	CGGATCGACGCTGTAGTGCGTCAGGCAGC	CTGTAGT
288	boronate_336	OB(O)c1cccc2ccncc12	CGGATCGACGCAATTGGGCGTCAGGCAGC	CAATTGG
289	boronate_337	Cc1ccncc1B(O)O	CGGATCGACGATACAGTGCGTCAGGCAGC	ATACAGT
290	boronate_339	COc1ccc(B(O)O)c(OC)n1	CGGATCGACGAAGCAATGCGTCAGGCAGC	AAGCAAT

291	boronate_340	OB(O)c1ccc(cc1)C(=O)NCc2occc2	CGGATCGACGGCGTGCGGCGTCAGGCAGC	GCGTGCG
292	boronate_343	CN(C)C(=O)c1ccc(cc1)B(O)O	CGGATCGACGCGGTATGGCGTCAGGCAGC	CGGTATG
293	boronate_344	OB(O)c1ccc(Cl)c(c1)C(F)(F)F	CGGATCGACGTATCGGAGCGTCAGGCAGC	TATCGGA
294	boronate_346	CN(C)C(=O)c1ccc(cc1Cl)B(O)O	CGGATCGACGTGGTAACGCGTCAGGCAGC	TGGTAAC
295	boronate_347	OB(O)c1ccc(cc1)C(=O)NCCN2CCOCC2	CGGATCGACGGCGGAGAGCGTCAGGCAGC	GCGGAGA
296	boronate_348	OB(O)c1ccc(Cl)nc1Cl	CGGATCGACGAGCTAACGCGTCAGGCAGC	AGCTAAC
297	boronate_349	OB(O)c1occc1	CGGATCGACGTGCCTTCGCGTCAGGCAGC	TGCCTTC
298	boronate_350	Cc1onc(C)c1B(O)O	CGGATCGACGTGCCGGCGCGTCAGGCAGC	TGCCGGC
299	boronate_351	OB(O)c1cccnc1C(F)(F)F	CGGATCGACGACTGCTCGCGTCAGGCAGC	ACTGCTC
300	boronate_352	CN(C)C(=O)c1ccccc1B(O)O	CGGATCGACGTCCTATAGCGTCAGGCAGC	TCCTATA
301	boronate_354	Cn1ncc2ccc(cc12)B(O)O	CGGATCGACGAACGGTTGCGTCAGGCAGC	AACGGTT
302	boronate_355	CCNS(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGACTCCGAGCGTCAGGCAGC	ACTCCGA
303	boronate_357	OB(O)c1ccc2occc2c1	CGGATCGACGCACGCAAGCGTCAGGCAGC	CACGCAA
304	boronate_358	OB(O)c1ccc(OC(F)(F)F)c(F)c1	CGGATCGACGGCTCTATGCGTCAGGCAGC	GCTCTAT
305	boronate_359	OB(O)c1cc(Cl)cc(c1)C(=O)N2CCOCC2	CGGATCGACGAGAACCGGCGTCAGGCAGC	AGAACCG
306	boronate_360	CC(C)S(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGATGATTGGCGTCAGGCAGC	ATGATTG
307	boronate_361	Cl.Cc1ccc(cn1)B(O)O	CGGATCGACGGTGTACTGCGTCAGGCAGC	GTGTACT
308	boronate_362	CC(C)CC(=O)NS(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGGCTGCGCGCGTCAGGCAGC	GCTGCGC
309	boronate_363	Cc1ccc(B(O)O)c(C)n1	CGGATCGACGGCTTGTAGCGTCAGGCAGC	GCTTGTA
310	boronate_366	CCOc1ccc(OC)cc1B(O)O	CGGATCGACGGCTACTGGCGTCAGGCAGC	GCTACTG
311	boronate_368	OCCNC(=O)c1ccc(cc1F)B(O)O	CGGATCGACGAGCATCTGCGTCAGGCAGC	AGCATCT
312	boronate_369	OB(O)c1ccccc1OC(F)(F)F	CGGATCGACGTATGCACGCGTCAGGCAGC	TATGCAC
313	boronate_370	Cl.Nc1cc(cc(c1)B(O)O)C#N	CGGATCGACGTTGACATGCGTCAGGCAGC	TTGACAT
314	boronate_371	CN(C)C(=O)c1cccc(c1)B(O)O	CGGATCGACGCCGTTCGGCGTCAGGCAGC	CCGTTCG
315	boronate_372	OB(O)c1cccc(OC(F)(F)F)c1	CGGATCGACGTATGATCGCGTCAGGCAGC	TATGATC
316	boronate_374	COc1ccc(cc1C(=O)O)B(O)O	CGGATCGACGCGTCAACGCGTCAGGCAGC	CGTCAAC
317	boronate_376	CNS(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGGCTCACAGCGTCAGGCAGC	GCTCACA
318	boronate_377	CC(C)OC(=O)c1ccc(cc1)B(O)O	CGGATCGACGACGAGAAGCGTCAGGCAGC	ACGAGAA
319	boronate_378	Cc1cc(OC(F)(F)F)ccc1B(O)O	CGGATCGACGTATATGAGCGTCAGGCAGC	TATATGA
320	boronate_379	OB(O)c1cccnc1F	CGGATCGACGTAGTAGTGCGTCAGGCAGC	TAGTAGT
321	boronate_380	CC(C)(C)OCc1ccc(cc1)B(O)O	CGGATCGACGCACAACGGCGTCAGGCAGC	CACAACG
322	boronate_381	COc1cc(OC)c(B(O)O)c(OC)c1	CGGATCGACGCAGTACAGCGTCAGGCAGC	CAGTACA
323	boronate_382	CC(C)(C)OC(=O)n1c(cc2c(Cl)cccc12)B(O)O	CGGATCGACGACCGGACGCGTCAGGCAGC	ACCGGAC
324	boronate_383	Cc1cc2[nH]ncc2cc1B(O)O	CGGATCGACGGCAGATTGCGTCAGGCAGC	GCAGATT
325	boronate_384	OB(O)c1ccc2cnc(Cl)cc2c1	CGGATCGACGCCACCACGCGTCAGGCAGC	CCACCAC
326	boronate_385	OB(O)c1cccc(c1)C(=O)N2CCOCC2	CGGATCGACGAGAAGCAGCGTCAGGCAGC	AGAAGCA
327	boronate_386	Cc1c(ccc2c1cnn2C)B(O)O	CGGATCGACGTGCGATCGCGTCAGGCAGC	TGCGATC
328	boronate_387	OB(O)c1cccc2ccnc(Cl)c12	CGGATCGACGAGGCAACGCGTCAGGCAGC	AGGCAAC
329	boronate_388	Cn1ncc2c(cccc12)B(O)O	CGGATCGACGCATCAGCGCGTCAGGCAGC	CATCAGC
330	boronate_389	COc1ccc2c(c1)cc(B(O)O)n2C(=O)OC(C)(C)C	CGGATCGACGGTTAGATGCGTCAGGCAGC	GTTAGAT
331	boronate_390	CC(C)NC(=O)c1cc(ccc1F)B(O)O	CGGATCGACGCTGCAGCGCGTCAGGCAGC	CTGCAGC
332	boronate_392	Cn1cc(cn1)B(O)O	CGGATCGACGTCTTCCTGCGTCAGGCAGC	TCTTCCT
333	boronate_393	Cc1c(cnn1C)B(O)O	CGGATCGACGATCTGGTGCGTCAGGCAGC	ATCTGGT
334	boronate_394	OB(O)c1ccc2nccn2c1	CGGATCGACGTGAACGGGCGTCAGGCAGC	TGAACGG
335	boronate_395	Cn1ccc2cc(ccc12)B(O)O	CGGATCGACGCGCCACAGCGTCAGGCAGC	CGCCACA
336	boronate_396	Cc1ccc2c(c1)cc(B(O)O)n2C(=O)OC(C)(C)C	CGGATCGACGTCATAGCGCGTCAGGCAGC	TCATAGC
337	boronate_398	Cl.Nc1cc(ccc1B(O)O)C#N	CGGATCGACGACCGATCGCGTCAGGCAGC	ACCGATC
338	boronate_400	CCCC\C=C\B(O)O	CGGATCGACGACCGTCAGCGTCAGGCAGC	ACCGTCA
339	boronate_401	CC(C)(C)OC(=O)n1c(cc2cc(ccc12)C#N)B(O)O	CGGATCGACGTACTAAGGCGTCAGGCAGC	TACTAAG
340	boronate_402	Cn1c(cc2ccc(Cl)cc12)B(O)O	CGGATCGACGGCACTAAGCGTCAGGCAGC	GCACTAA
341	boronate_405	OB(O)c1oc(C=O)cc1	CGGATCGACGTAGGTGGGCGTCAGGCAGC	TAGGTGG
342	boronate_406	OB(O)\C=C\c1ccccc1	CGGATCGACGAGATATTGCGTCAGGCAGC	AGATATT
343	boronate_408	OB(O)c1cccc(Cl)c1Cl	CGGATCGACGATTGTAGGCGTCAGGCAGC	ATTGTAG
344	boronate_409	CC(C)(C)OC(=O)n1c(cc2ccccc12)B(O)O	CGGATCGACGTCACTCTGCGTCAGGCAGC	TCACTCT
345	boronate_411	CC(C)(C)OC(=O)n1c(cc2ccc(cc12)C#N)B(O)O	CGGATCGACGAACGCTAGCGTCAGGCAGC	AACGCTA
346	boronate_413	CC1(C)CCC(=CC1)B(O)O	CGGATCGACGTATGGCTGCGTCAGGCAGC	TATGGCT
347	boronate_415	Cc1cncc(c1)B(O)O	CGGATCGACGCCTTGAAGCGTCAGGCAGC	CCTTGAA
348	boronate_451	CC(=O)NC1=CC=C(C=C1)B(O)O	CGGATCGACGCTCGCAGGCGTCAGGCAGC	CTCGCAG
349	boronate_452	CC(=C)B1OC(C)(C)C(C)(C)O1	CGGATCGACGAGAGAATGCGTCAGGCAGC	AGAGAAT
350	boronate_453	OB(O)C1=CC=C(C=C1)C(=O)N1CCCCC1	CGGATCGACGTCCTACGGCGTCAGGCAGC	TCCTACG
351	boronate_454	CC1(C)OB(OC1(C)C)C1=CN=CC2=C1C=CC=C2	CGGATCGACGAGCTCGCGCGTCAGGCAGC	AGCTCGC
352	boronate_458	CC1(C)OB(OC1(C)C)C1=CN=C(Cl)C=C1	CGGATCGACGCAAGCCTGCGTCAGGCAGC	CAAGCCT
353	boronate_459	OB(O)C1=CC2=C(OCO2)C=C1	CGGATCGACGATGTAGCGCGTCAGGCAGC	ATGTAGC
354	boronate_460	CC1(C)OB(OC1(C)C)C1=CN=C(C=C1)N1CCOCC1	CGGATCGACGGAAGGCTGCGTCAGGCAGC	GAAGGCT
355	boronate_461	CC(C)(C)OC(=O)N1CCN(CC1)C1=CC=C(C=C1)B1OC(C)(C)C(C)(C)O1	CGGATCGACGTCATGACGCGTCAGGCAGC	TCATGAC
356	boronate_462	OB(O)C1=CC=C(CNC(=O)C(F)(F)F)C=C1	CGGATCGACGTTCTAGCGCGTCAGGCAGC	TTCTAGC
357	boronate_463	OB(O)C1=C(SC=C1)C=O	CGGATCGACGAACGTGAGCGTCAGGCAGC	AACGTGA
358	boronate_464	CC(C)(C)OC(=O)N1C=CC=C1B(O)O	CGGATCGACGACAATTAGCGTCAGGCAGC	ACAATTA
359	boronate_466	COC1=CC=C(B(O)O)C(OC)=C1	CGGATCGACGGAATGTCGCGTCAGGCAGC	GAATGTC
360	boronate_2	OB(O)c1ccc(C=O)c(F)c1	CGGATCGACGTCGCCATGCGTCAGGCAGC	TCGCCAT
361	boronate_13	CC(=O)c1ccc(cc1F)B(O)O	CGGATCGACGGAATGGTGCGTCAGGCAGC	GAATGGT
362	boronate_32	OB(O)c1cccc(F)c1C=O	CGGATCGACGATGCCGGGCGTCAGGCAGC	ATGCCGG
363	boronate_82	Cc1oc(cc1)B(O)O	CGGATCGACGCTTGATAGCGTCAGGCAGC	CTTGATA
364	boronate_88	OB(O)c1cccnc1	CGGATCGACGGATCGGCGCGTCAGGCAGC	GATCGGC
365	boronate_110	COc1c(C=O)cccc1B(O)O	CGGATCGACGGATAAGTGCGTCAGGCAGC	GATAAGT
366	boronate_114	OB(O)c1cc(Cl)ccc1O	CGGATCGACGACCACTCGCGTCAGGCAGC	ACCACTC
367	boronate_129	OB(O)c1ccc(cc1)C#N	CGGATCGACGGCCAAGAGCGTCAGGCAGC	GCCAAGA

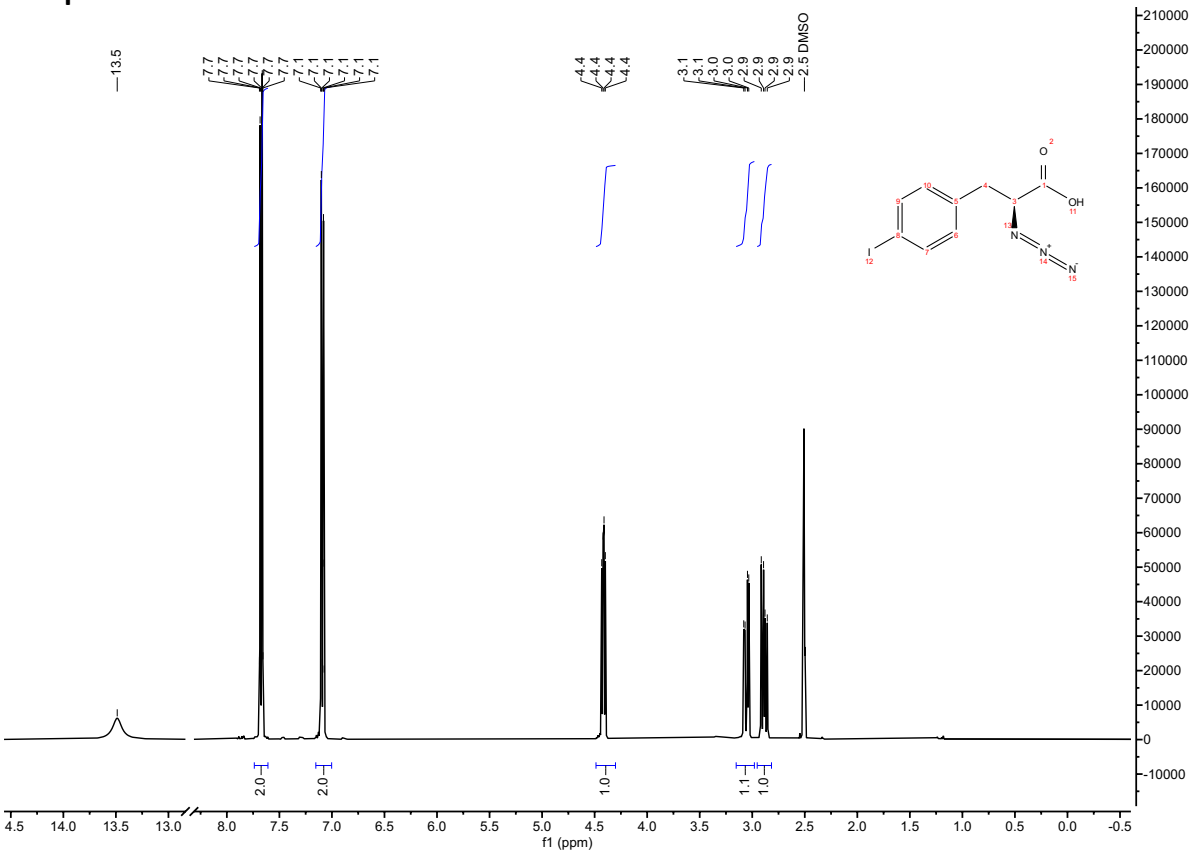
368	boronate_133	OB(O)c1cc(cc(c1)[N+](=O)[O-])C(=O)O	CGGATCGACGCGACTAAGCGTCAGGCAGC	CGACTAA
369	boronate_153	OB(O)c1ccccc1O	CGGATCGACGCAACAGGGCGTCAGGCAGC	CAACAGG
370	boronate_175	OB(O)c1ccc(C#N)c(F)c1	CGGATCGACGCGCTCACGCGTCAGGCAGC	CGCTCAC
371	boronate_179	CS(=O)(=O)Nc1ccc(cc1)B(O)O	CGGATCGACGTGGATAGGCGTCAGGCAGC	TGGATAG
372	boronate_231	CS(=O)(=O)c1ccc(cc1F)B(O)O	CGGATCGACGACGGCATGCGTCAGGCAGC	ACGGCAT
373	boronate_236	NNC(=O)c1ccc(cc1)B(O)O	CGGATCGACGTGCAAGTGCGTCAGGCAGC	TGCAAGT
374	boronate_247	OB(O)c1cncnc1	CGGATCGACGAATAATAGCGTCAGGCAGC	AATAATA
375	boronate_273	CSc1ncc(cn1)B(O)O	CGGATCGACGTGTAGGCGCGTCAGGCAGC	TGTAGGC
376	boronate_295	OB(O)c1ccc2cnccc2c1	CGGATCGACGACTAACAGCGTCAGGCAGC	ACTAACA
377	boronate_365	OB(O)C1=CCCC1	CGGATCGACGCATATACGCGTCAGGCAGC	CATATAC
378	boronate_375	CC(C)NS(=O)(=O)c1ccc(cc1)B(O)O	CGGATCGACGCCTCGGTGCGTCAGGCAGC	CCTCGGT
379	boronate_414	Cc1csc(c1)B(O)O	CGGATCGACGGACTCCGGCGTCAGGCAGC	GACTCCG
380	boronate_456	CCOC(=O)C1=CC=C(C=C1)B1OC(C)(C)C(C)(C)O1	CGGATCGACGCGCTATCGCGTCAGGCAGC	CGCTATC
381	boronate_457	CC1(C)OB(OC1(C)C)C1=CC=C2NC=NC2=C1	CGGATCGACGATCGCGAGCGTCAGGCAGC	ATCGCGA
382	boronate_434	CC1(C)OB(C2=CC(NC3=C4C=CC=C3)=C4C=C2)OC1(C)C	CGGATCGACGCGAGAGCGCGTCAGGCAGC	CGAGAGC
383	boronate_435	NC1=NC=C(B2OC(C)(C)C(C)(C)O2)C=N1	CGGATCGACGAAGAGGAGCGTCAGGCAGC	AAGAGGA
384	boronate_436	CCC/C=C/B(O)O	CGGATCGACGCAGAGATGCGTCAGGCAGC	CAGAGAT
385	boronate_437	CC1(C)OB(C2=CC=C(N3CCNCC3)N=C2)OC1(C)C	CGGATCGACGCGCAATTGCGTCAGGCAGC	CGCAATT
386	boronate_439	O=C(N1CCN(C2=CC=C(B3OC(C)(C)C(C)(C)O3)C=N2)CC1)OC(C)(C)C	CGGATCGACGTGGTACGGCGTCAGGCAGC	TGGTACG
387	boronate_441	OB(C1=CN=CC2=C1C=CC=C2)O	CGGATCGACGACGTCGAGCGTCAGGCAGC	ACGTCGA
388	boronate_443	CC1(C)C(C)(C)OB(C2=CC=CC3=C2C=CN3)O1	CGGATCGACGCCAAGGTGCGTCAGGCAGC	CCAAGGT
389	alkyne_1	C#CCN1C=CC2=C1C=CC=C2	CGGATCGACGGAACCGTGCGTCAGGCAGC	GAACCGT
390	alkyne_2	ClC1=C(NC(=O)NCC#C)C=CC=C1	CGGATCGACGAATCGTTGCGTCAGGCAGC	AATCGTT
391	alkyne_4	Cl.NCCC(O)CCC#C	CGGATCGACGTAGGATTGCGTCAGGCAGC	TAGGATT
392	alkyne_5	FC1=C(NC(=O)NCC#C)C=CC=C1	CGGATCGACGAGTACATGCGTCAGGCAGC	AGTACAT
393	alkyne_6	CCOC(=O)C1=C(NCC#C)N=C(Cl)C=C1	CGGATCGACGCGGAACTGCGTCAGGCAGC	CGGAACT
394	alkyne_7	Cl.NC1=C(C=CC=C1)C(=O)NCC#C	CGGATCGACGCCGACGCGCGTCAGGCAGC	CCGACGC
395	alkyne_8	O=C(NCC#C)NC1=CC=CC=C1	CGGATCGACGTGCTTAGGCGTCAGGCAGC	TGCTTAG
396	alkyne_10	BrC1=CC2=C(C=C1)N(CC#C)C(=O)C2=O	CGGATCGACGTCCGCAGGCGTCAGGCAGC	TCCGCAG
397	alkyne_12	OC(=O)C1=C(NCC#C)N=C(Cl)C=C1	CGGATCGACGGTACATTGCGTCAGGCAGC	GTACATT
398	alkyne_13	O=C1N(CC#C)C2=C(C=CC=C2)C1=O	CGGATCGACGCTTAAGCGCGTCAGGCAGC	CTTAAGC
399	alkyne_14	Cl.C#CCNCC1=CC=CO1	CGGATCGACGTGCTGCAGCGTCAGGCAGC	TGCTGCA
400	alkyne_16	CC(O)CC#C	CGGATCGACGGACTATTGCGTCAGGCAGC	GACTATT
401	alkyne_17	OC(=O)C1=CC(=CN=C1)C#C	CGGATCGACGAATACTTGCGTCAGGCAGC	AATACTT
402	alkyne_18	COC1=C(Br)C=C(C=C1)C#C	CGGATCGACGCCGAGGAGCGTCAGGCAGC	CCGAGGA
403	alkyne_19	BrC1=CC2=C(OCC(=O)N2CC#C)C=C1	CGGATCGACGCAGTTCTGCGTCAGGCAGC	CAGTTCT
404	alkyne_20	ClC1=CC=CC(Cl)=C1C#C	CGGATCGACGGCCTAGCGCGTCAGGCAGC	GCCTAGC
405	alkyne_21	OC1=CC(OCC#C)=CC=C1	CGGATCGACGGCCATCTGCGTCAGGCAGC	GCCATCT
406	alkyne_22	Cl.NC(CC#C)CC(F)(F)F	CGGATCGACGGTTACGAGCGTCAGGCAGC	GTTACGA
407	alkyne_23	OC(=O)[C@@H]1CCCN1CC#C	CGGATCGACGTCGGTGCGCGTCAGGCAGC	TCGGTGC
408	alkyne_24	O=C1NCCN1CC#C	CGGATCGACGAATATGTGCGTCAGGCAGC	AATATGT
409	alkyne_25	NC(=O)C1=CC(=CN=C1)C#C	CGGATCGACGTACGTTCGCGTCAGGCAGC	TACGTTC
410	alkyne_26	C#CCN1C2=C(C=CC=C2)C2=C1C=CC=C2	CGGATCGACGAGTCCAAGCGTCAGGCAGC	AGTCCAA
411	alkyne_27	COC1=C(N)C=C(C=C1)C#C	CGGATCGACGTCAAGTTGCGTCAGGCAGC	TCAAGTT
412	alkyne_29	COC1=C(C=C(Cl)C=C1)C#C	CGGATCGACGCATCGTGGCGTCAGGCAGC	CATCGTG
413	alkyne_30	NC(=O)NC1=CC(=CC=C1)C#C	CGGATCGACGCCGGAGCGCGTCAGGCAGC	CCGGAGC
414	alkyne_32	COC1=C(F)C=C(C=C1)C#C	CGGATCGACGGCGAGTTGCGTCAGGCAGC	GCGAGTT
415	alkyne_33	FC(F)OC1=C(C=CC=C1)C#C	CGGATCGACGGAGGACAGCGTCAGGCAGC	GAGGACA
416	alkyne_34	CCN1C=C(C=N1)C#C	CGGATCGACGTGTGTGGGCGTCAGGCAGC	TGTGTGG
417	alkyne_37	O=C(NCC#C)NC1CC1	CGGATCGACGGCATCCAGCGTCAGGCAGC	GCATCCA
418	alkyne_38	ClC1=CC=C(NC(=O)NCC#C)C=C1	CGGATCGACGCGGATTAGCGTCAGGCAGC	CGGATTA
419	alkyne_39	OC(=O)C1=CC=C(C=C1)S(=O)(=O)NCC#C	CGGATCGACGGTGACTGGCGTCAGGCAGC	GTGACTG
420	alkyne_41	O=C(NCC#C)C1=CNC(=O)C=C1	CGGATCGACGGTGTTGAGCGTCAGGCAGC	GTGTTGA
421	alkyne_42	COC1=CC(C#C)=C(Cl)C=C1	CGGATCGACGTGAGCGTGCGTCAGGCAGC	TGAGCGT
422	alkyne_43	FC1=CC(=CC(F)=C1)C#C	CGGATCGACGTCGGCGGGCGTCAGGCAGC	TCGGCGG
423	alkyne_45	CC1=CC(NC(=O)CNCC#C)=CC=C1	CGGATCGACGGATCTCTGCGTCAGGCAGC	GATCTCT
424	alkyne_46	OC(=O)CC(=O)NCC#C	CGGATCGACGTGAGTATGCGTCAGGCAGC	TGAGTAT
425	alkyne_47	CC1=NC(=CS1)C#C	CGGATCGACGTTCCACAGCGTCAGGCAGC	TTCCACA
426	alkyne_48	Cl.NC(C#C)C1CCOCC1	CGGATCGACGTAATGCTGCGTCAGGCAGC	TAATGCT
427	alkyne_49	Cl.NC1(CCC1)C#C	CGGATCGACGTCTAATCGCGTCAGGCAGC	TCTAATC
428	alkyne_50	CCC(C)(O)C#C	CGGATCGACGACACACTGCGTCAGGCAGC	ACACACT
429	alkyne_53	CC1=NC(=CC=C1)C#C	CGGATCGACGGAACAGAGCGTCAGGCAGC	GAACAGA
430	alkyne_55	CC1=CC(NC(=O)NCC#C)=CC=C1	CGGATCGACGTCGGAGTGCGTCAGGCAGC	TCGGAGT
431	alkyne_56	C#CCNC1CCCC1	CGGATCGACGGCCACAGGCGTCAGGCAGC	GCCACAG
432	alkyne_57	C#CC1=CN=CN=C1	CGGATCGACGTGGTGAAGCGTCAGGCAGC	TGGTGAA
433	alkyne_60	ClC1=C(C=CC=C1)C#C	CGGATCGACGAATCGAAGCGTCAGGCAGC	AATCGAA
434	alkyne_61	ClC1=CC2=C(C=C1)N(CC(=O)NCC#C)C(=O)C2=O	CGGATCGACGTCGATAAGCGTCAGGCAGC	TCGATAA
435	alkyne_62	O=C1CCCCN1CC#C	CGGATCGACGGAGAGCTGCGTCAGGCAGC	GAGAGCT
436	alkyne_63	C#CC1=CC=NN1	CGGATCGACGTTACATAGCGTCAGGCAGC	TTACATA
437	alkyne_66	O=C1NC(CC#C)C(=O)N1	CGGATCGACGTTGCTAAGCGTCAGGCAGC	TTGCTAA
438	alkyne_67	O=S(=O)(NCC#C)C1=CC=CC=C1	CGGATCGACGACTCGAGGCGTCAGGCAGC	ACTCGAG
439	alkyne_68	FC1=C(F)C=C(NC(=O)NCC#C)C=C1	CGGATCGACGCCTAACGGCGTCAGGCAGC	CCTAACG
440	alkyne_69	Cl.C#CCN1C=NC2=C1C=CC=C2	CGGATCGACGACTTCTGGCGTCAGGCAGC	ACTTCTG
441	alkyne_70	COC1=NC(=CC=C1)C#C	CGGATCGACGAATCTAGGCGTCAGGCAGC	AATCTAG
442	alkyne_71	OC(=O)C1(CCC2=C1C=CC=C2)NC(=O)NCC#C	CGGATCGACGAAGCTGAGCGTCAGGCAGC	AAGCTGA
443	alkyne_72	CC1=C(OCC#C)C=CC=C1	CGGATCGACGAGTATTCGCGTCAGGCAGC	AGTATTC
444	alkyne_74	BrC1=CC2=C(OCCC2NCC#C)C=C1	CGGATCGACGAGGCATTGCGTCAGGCAGC	AGGCATT

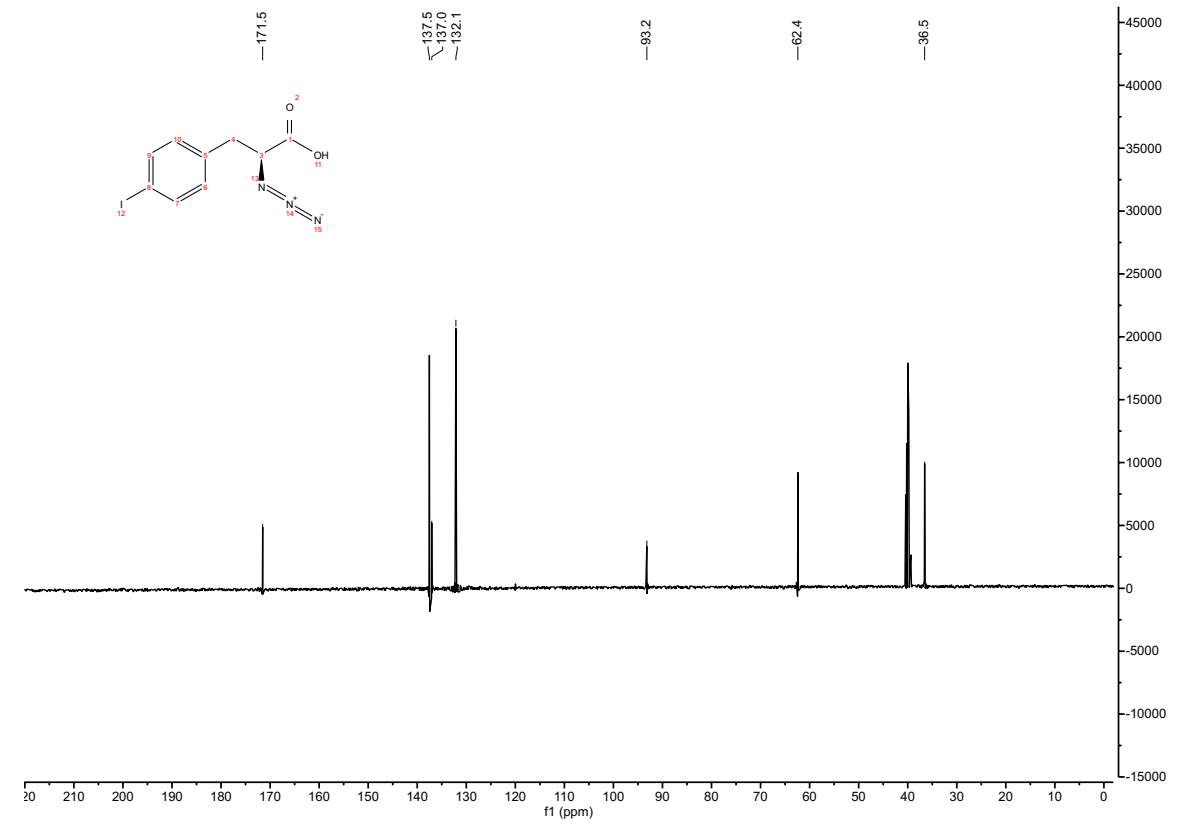
445	alkyne_77	NNC(=O)CCC#C	CGGATCGACGACGGACAGCGTCAGGCAGC	ACGGACA
446	alkyne_78	CC(C)(NS(C)(=O)=O)C#C	CGGATCGACGGTACGCAGCGTCAGGCAGC	GTACGCA
447	alkyne_81	Cl.NC(CC#C)C(N)=O	CGGATCGACGATCCGAAGCGTCAGGCAGC	ATCCGAA
448	alkyne_82	COC1=CC(=NC=C1)C#C	CGGATCGACGGTGAGAAGCGTCAGGCAGC	GTGAGAA
449	alkyne_84	CC1=NC2=C(C=C1)C=C(C=C2)C#C	CGGATCGACGGTCGGTTGCGTCAGGCAGC	GTCGGTT
450	alkyne_88	CC(O)(C#C)C1=C(Br)C=CC=C1	CGGATCGACGGCGAGACGCGTCAGGCAGC	GCGAGAC
451	alkyne_90	NC1=C(C=C(O)C=C1)C#C	CGGATCGACGAGCTAGAGCGTCAGGCAGC	AGCTAGA
452	alkyne_91	Cl.C#CCNCC1CC1	CGGATCGACGAGAGTACGCGTCAGGCAGC	AGAGTAC
453	alkyne_92	C#CC1=CC=NC=C1	CGGATCGACGATATTGAGCGTCAGGCAGC	ATATTGA
454	alkyne_93	C#CC1=CN=CS1	CGGATCGACGAGTCATAGCGTCAGGCAGC	AGTCATA
455	alkyne_94	CN1CCN(CCC#C)CC1	CGGATCGACGCTAGTCGGCGTCAGGCAGC	CTAGTCG
456	alkyne_95	OC(CC#C)C1CC1	CGGATCGACGTTAGAAGGCGTCAGGCAGC	TTAGAAG
457	alkyne_96	Cl.C#CC1CNC1	CGGATCGACGATCTAAGGCGTCAGGCAGC	ATCTAAG
458	alkyne_97	OC(CC1=CC=CC=C1)C#C	CGGATCGACGCATTAGTGCGTCAGGCAGC	CATTAGT
459	alkyne_98	COCCN(C)CC#C	CGGATCGACGGCAGTTAGCGTCAGGCAGC	GCAGTTA
460	alkyne_99	C#CCN1CCCC1	CGGATCGACGTCTGTCAGCGTCAGGCAGC	TCTGTCA
461	alkyne_101	CN1C=C(C=N1)C#C	CGGATCGACGCTGTATCGCGTCAGGCAGC	CTGTATC
462	alkyne_102	COCC(N)C#C	CGGATCGACGCGGTCCAGCGTCAGGCAGC	CGGTCCA
463	alkyne_103	CN(C)CCC#C	CGGATCGACGATGTATGGCGTCAGGCAGC	ATGTATG
464	alkyne_104	ClC1=C(Cl)C(=CC=C1)C#C	CGGATCGACGTAACGATGCGTCAGGCAGC	TAACGAT
465	alkyne_105	C#CCCNC1COC1	CGGATCGACGGTTCTCAGCGTCAGGCAGC	GTTCTCA
466	alkyne_106	FC1=C(C=CN=C1)C#C	CGGATCGACGCGCGTCTGCGTCAGGCAGC	CGCGTCT
467	alkyne_107	CS(=O)(=O)NCCC#C	CGGATCGACGTGACCTCGCGTCAGGCAGC	TGACCTC
468	alkyne_108	NC(CO)CC#C	CGGATCGACGTTCGATAGCGTCAGGCAGC	TTCGATA
469	alkyne_109	CC1(C)OB(OC1(C)C)C#C	CGGATCGACGTATTAGCGCGTCAGGCAGC	TATTAGC
470	alkyne_110	NC1=C(F)C(F)=C(C#C)C(F)=C1F	CGGATCGACGACCAGGCGCGTCAGGCAGC	ACCAGGC
471	alkyne_114	OC(=O)C1=CC(=CC=C1)S(=O)(=O)NCC#C	CGGATCGACGGTTCAGCGCGTCAGGCAGC	GTTCAGC
472	alkyne_115	CC1=CC=C(NC(=O)NCC#C)C=C1	CGGATCGACGTGCTTGTGCGTCAGGCAGC	TGCTTGT
473	alkyne_123	O=S1(=O)CCN(CC#C)CC1	CGGATCGACGTCATCTAGCGTCAGGCAGC	TCATCTA
474	alkyne_124	NC1=CC=CC(=C1)C#C	CGGATCGACGTCTGTATGCGTCAGGCAGC	TCTGTAT
475	alkyne_125	NS(=O)(=O)C1=CC=C(C=C1)C#C	CGGATCGACGCATCCGAGCGTCAGGCAGC	CATCCGA
476	alkyne_126	C#CCNC1CCC2=C1C=CC=C2	CGGATCGACGTCGACCTGCGTCAGGCAGC	TCGACCT
477	alkyne_127	C#CCN1CCOCC1	CGGATCGACGCAATCGAGCGTCAGGCAGC	CAATCGA
478	alkyne_128	COC1=CC(=CC=C1)C#C	CGGATCGACGAATACACGCGTCAGGCAGC	AATACAC
479	alkyne_129	C#CC1=NC=CC=C1	CGGATCGACGACTATAGGCGTCAGGCAGC	ACTATAG
480	alkyne_130	NC1(CCCCC1)C#C	CGGATCGACGACGTAATGCGTCAGGCAGC	ACGTAAT
481	alkyne_131	CN(CC#C)CC1=CC=CC=C1	CGGATCGACGGCTGTTGGCGTCAGGCAGC	GCTGTTG
482	alkyne_132	OC(=O)CNCC#C	CGGATCGACGCCTCAATGCGTCAGGCAGC	CCTCAAT
483	alkyne_133	CNCC#C	CGGATCGACGTCACTGCGCGTCAGGCAGC	TCACTGC
484	alkyne_134	NCC#C	CGGATCGACGTGTATGTGCGTCAGGCAGC	TGTATGT
485	alkyne_135	NC1=CC=C(C=C1)C#C	CGGATCGACGTAGCCAAGCGTCAGGCAGC	TAGCCAA
486	alkyne_136	CCCC#C	CGGATCGACGCGCTTATGCGTCAGGCAGC	CGCTTAT
487	alkyne_137	OC(=O)CCC#C	CGGATCGACGCTCGTCAGCGTCAGGCAGC	CTCGTCA
488	alkyne_138	OCCC#C	CGGATCGACGCGAACTGGCGTCAGGCAGC	CGAACTG
489	alkyne_140	C#CC1=CC=C2N=CC=NC2=C1	CGGATCGACGTCTCTAAGCGTCAGGCAGC	TCTCTAA
490	alkyne_141	OC(C#C)C1=CC=CC=C1	CGGATCGACGATGGATCGCGTCAGGCAGC	ATGGATC
491	alkyne_142	CC(C)(N)C#C	CGGATCGACGAGCCTATGCGTCAGGCAGC	AGCCTAT
492	alkyne_143	C#CC1=CC2=C(NC=C2)C=C1	CGGATCGACGTACGCATGCGTCAGGCAGC	TACGCAT
493	alkyne_146	C#CC1=CC=CC=C1	CGGATCGACGAACCGATGCGTCAGGCAGC	AACCGAT
494	alkyne_147	C#CC1=CC=CN=C1	CGGATCGACGAGGTACAGCGTCAGGCAGC	AGGTACA
495	alkyne_148	CN(C)C1=CC=C(C=C1)C#C	CGGATCGACGGAACTTAGCGTCAGGCAGC	GAACTTA
496	alkyne_149	NC1=CC=CC=C1C#C	CGGATCGACGCCAGCAGGCGTCAGGCAGC	CCAGCAG
497	alkyne_150	C#CC1=CSC=C1	CGGATCGACGTGTGCCAGCGTCAGGCAGC	TGTGCCA
498	alkyne_151	FC(F)(F)C1=CC=CC=C1C#C	CGGATCGACGGCCAACGGCGTCAGGCAGC	GCCAACG
499	alkyne_152	OC(C#C)(C1=CC=CC=C1)C(F)(F)F	CGGATCGACGGCCGGAAGCGTCAGGCAGC	GCCGGAA
500	alkyne_155	OC(=O)C#C	CGGATCGACGGTAATTAGCGTCAGGCAGC	GTAATTA
501	alkyne_156	OCCCC#C	CGGATCGACGCTCGTGGGCGTCAGGCAGC	CTCGTGG
502	alkyne_157	COC1=CC=C(C#C)C(C)=C1	CGGATCGACGTCGCTACGCGTCAGGCAGC	TCGCTAC
503	alkyne_159	OC1(CCCCC1)C#C	CGGATCGACGTCAGCTGGCGTCAGGCAGC	TCAGCTG
504	alkyne_160	OC(C#C)(C1=CC=CC=C1)C1=CC=CC=C1	CGGATCGACGTGAGGTCGCGTCAGGCAGC	TGAGGTC
505	alkyne_161	C#CC1=CSC=C1	CGGATCGACGTCTTAGAGCGTCAGGCAGC	TCTTAGA
506	alkyne_162	BrC1=CC=C(C=C1)C#C	CGGATCGACGTTCAATCGCGTCAGGCAGC	TTCAATC
507	alkyne_163	CCCCCC(O)C#C	CGGATCGACGTCCGTGTGCGTCAGGCAGC	TCCGTGT
508	alkyne_164	CN1C=NC=C1C#C	CGGATCGACGTTCATGGGCGTCAGGCAGC	TTCATGG
509	alkyne_165	C#CC1CCCCC1	CGGATCGACGTGGTAGAGCGTCAGGCAGC	TGGTAGA
510	alkyne_166	CC(O)(C=C)C#C	CGGATCGACGTTGCATCGCGTCAGGCAGC	TTGCATC
511	alkyne_171	NC(=N)C1=CC=C(CNC(=O)CCC#C)C=C1	CGGATCGACGCTACATGGCGTCAGGCAGC	CTACATG
512	alkyne_9	CC1=CC(NC(=O)NCC#C)=NO1	CGGATCGACGCGCATCGGCGTCAGGCAGC	CGCATCG
513	alkyne_31	C#CCN1C=CN=C1	CGGATCGACGCGCATAAGCGTCAGGCAGC	CGCATAA
514	alkyne_51	C#CC1=C2CCCCC2=CC=C1	CGGATCGACGCGTTAATGCGTCAGGCAGC	CGTTAAT
515	alkyne_52	FC(F)(F)C1=NC(OCC#C)=CC=C1	CGGATCGACGTAAGAGCGCGTCAGGCAGC	TAAGAGC
516	alkyne_54	Cl.NC(=N)NCC#C	CGGATCGACGACATTAAGCGTCAGGCAGC	ACATTAA
517	alkyne_65	C#CC1=CC(=CC=C1)C1=CC=CC=C1	CGGATCGACGGAAGGACGCGTCAGGCAGC	GAAGGAC
518	alkyne_73	OC(=O)C1=CC2=C(C=C1)N(CC2)C(=O)C#C	CGGATCGACGGTGAACGGCGTCAGGCAGC	GTGAACG
519	alkyne_76	NC(C#C)C1CCCCC1	CGGATCGACGCTGGATGGCGTCAGGCAGC	CTGGATG
520	alkyne_85	CNC(=O)C#C	CGGATCGACGCCTCGACGCGTCAGGCAGC	CCTCGAC
521	alkyne_100	ClC1=CC(=CC(Cl)=C1)C#C	CGGATCGACGCACACTTGCGTCAGGCAGC	CACACTT

522	alkyne_112	CC(=O)C(CC#C)C(C)=O	CGGATCGACGTATTCATGCGTCAGGCAGC	TATTCAT
523	alkyne_116	CC[C@]12CC[C@H]3[C@@H](CCC4=CC(=O)CC[C@H]34)[C@@H]1CC[C@@]2(O)C#C	CGGATCGACGAGCTGTGGCGTCAGGCAGC	AGCTGTG
524	alkyne_117	OC(CC#C)C(O)=O	CGGATCGACGTATTGACGCGTCAGGCAGC	TATTGAC
525	alkyne_120	CC(=O)O[C@]1(CCC2C3CCC4=CC(=O)CC[C@@H]4C3CC[C@@]12C)C#C	CGGATCGACGTCACGGTGCGTCAGGCAGC	TCACGGT
526	alkyne_139	C#CC1CC1	CGGATCGACGCGGTGGAGCGTCAGGCAGC	CGGTGGA
527	alkyne_144	CC1=C(C=O)C2=C(C=CC=C2)N1CC#C	CGGATCGACGGTTGACAGCGTCAGGCAGC	GTTGACA
528	alkyne_169	NS(=O)(=O)C1=NN=C(NC(=O)CCC#C)S1	CGGATCGACGCGTCCATGCGTCAGGCAGC	CGTCCAT
529	alkyne_3	NC(=S)NC1=CC(=CC=C1)C#C	CGGATCGACGCTAAGAGGCGTCAGGCAGC	CTAAGAG
530	alkyne_11	CN(CC#C)C1=NC=C(N)N=C1	CGGATCGACGGTATTCGGCGTCAGGCAGC	GTATTCG
531	alkyne_35	C#CCN1C=CN=C1C1=CC=CS1	CGGATCGACGGCGCTGGGCGTCAGGCAGC	GCGCTGG
532	alkyne_36	C#CC1=NC2=C(C=CC=C2)N=C1	CGGATCGACGATGCCACGCGTCAGGCAGC	ATGCCAC
533	alkyne_40	C#CCNCC1=NC2=C(C=CC=C2)C=C1	CGGATCGACGTAGCAGCGCGTCAGGCAGC	TAGCAGC
534	alkyne_58	C#CCN1C=NC=N1	CGGATCGACGGAAGCTCGCGTCAGGCAGC	GAAGCTC
535	alkyne_64	CNC1=CC=C(OCC#C)C=C1	CGGATCGACGGTCCATCGCGTCAGGCAGC	GTCCATC
536	alkyne_75	CC(NCC#C)C1=CC2=C(OCC(=O)N2)C=C1	CGGATCGACGCCAGCCAGCGTCAGGCAGC	CCAGCCA
537	alkyne_80	C#CC1=NC=CS1	CGGATCGACGTCACATGGCGTCAGGCAGC	TCACATG
538	alkyne_145	O=CC1=CNC2=C1C=C(C=C2)C#C	CGGATCGACGCTACACTGCGTCAGGCAGC	CTACACT
539	alkyne_158	C#CC1=CC=C(C=C1)C1=CC=CC=C1	CGGATCGACGATGTTGTGCGTCAGGCAGC	ATGTTGT
540	alkyne_167	FC(F)(F)C1=CC(=CC(=C1)C#C)C(F)(F)F	CGGATCGACGCATAGAGGCGTCAGGCAGC	CATAGAG
541	alkyne_M7_1	OC1=C(O)C=CC(C(CN2C(C=C(C(NCC#C)=O)C=C3)=C3N=C2)=O)=C1	CGGATCGACGTTGTGACGCGTCAGGCAGC	TTGTGAC
542	alkyne_M8_1	OC1=C(C=CC(C(CN2C3=C(N=C2)C=C(C(NCC#C)=O)C=C3)=O)=C1)O	CGGATCGACGGCGGTGTGCGTCAGGCAGC	GCGGTGT
543	alkyne_M9_1	OC1=C(O)C=CC(C(CN2C(C=CC=C3C(NCC#C)=O)=C3N=C2)=O)=C1	CGGATCGACGGTACTGGGCGTCAGGCAGC	GTACTGG
544	alkyne_M7_2	OC1=C(O)C=CC(C(CN2C(C=C(C(NCC#C)=O)C=C3)=C3N=C2)=O)=C1	CGGATCGACGTCGTCTCGCGTCAGGCAGC	TCGTCTC
545	alkyne_M8_2	OC1=C(C=CC(C(CN2C3=C(N=C2)C=C(C(NCC#C)=O)C=C3)=O)=C1)O	CGGATCGACGTACCACTGCGTCAGGCAGC	TACCACT
546	alkyne_M9_2	OC1=C(O)C=CC(C(CN2C(C=CC=C3C(NCC#C)=O)=C3N=C2)=O)=C1	CGGATCGACGAGTCTCAGCGTCAGGCAGC	AGTCTCA
547	no_BB	[H]	CGGATCGACGTGTTGCTGCGTCAGGCAGC	TGTTGCT
548	iodo_SM	[I]	CGGATCGACGTGGCAATGCGTCAGGCAGC	TGGCAAT

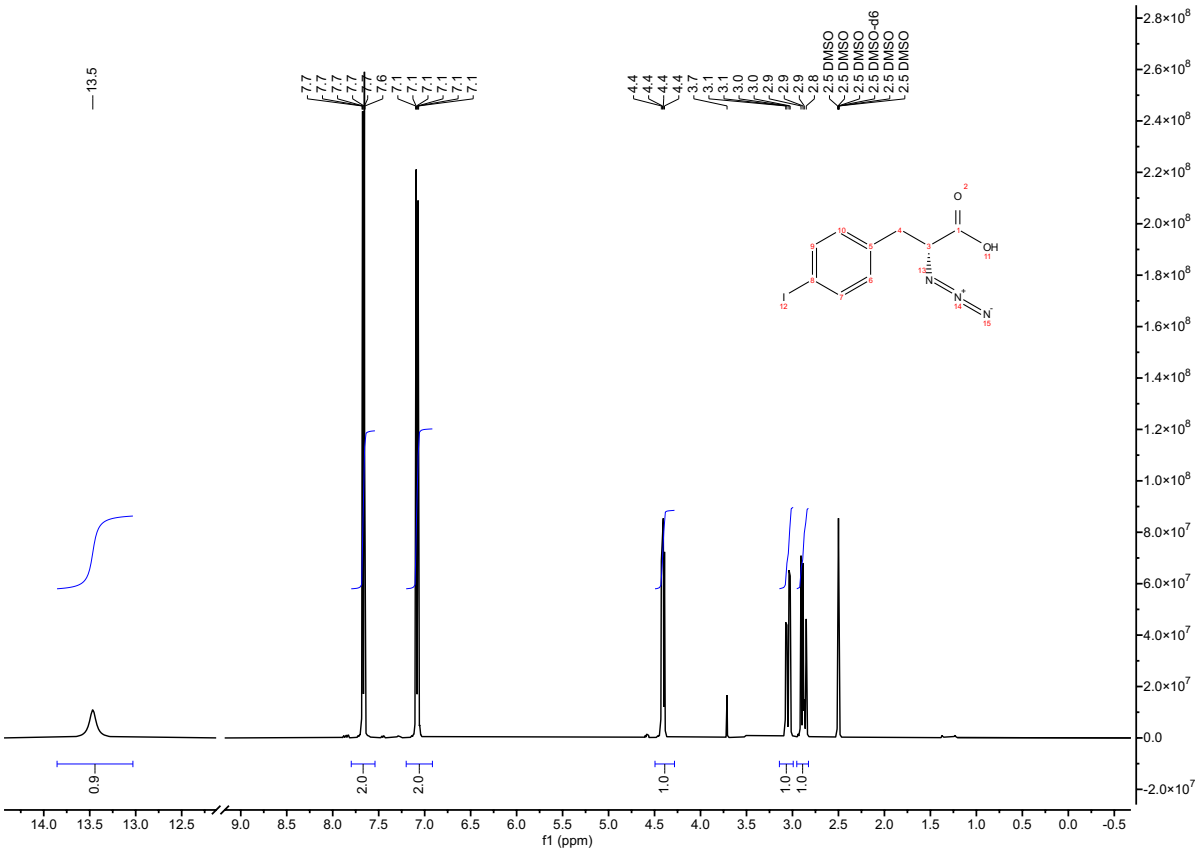
8.3 ^{1}H-NMR and ^{13}C-NMR

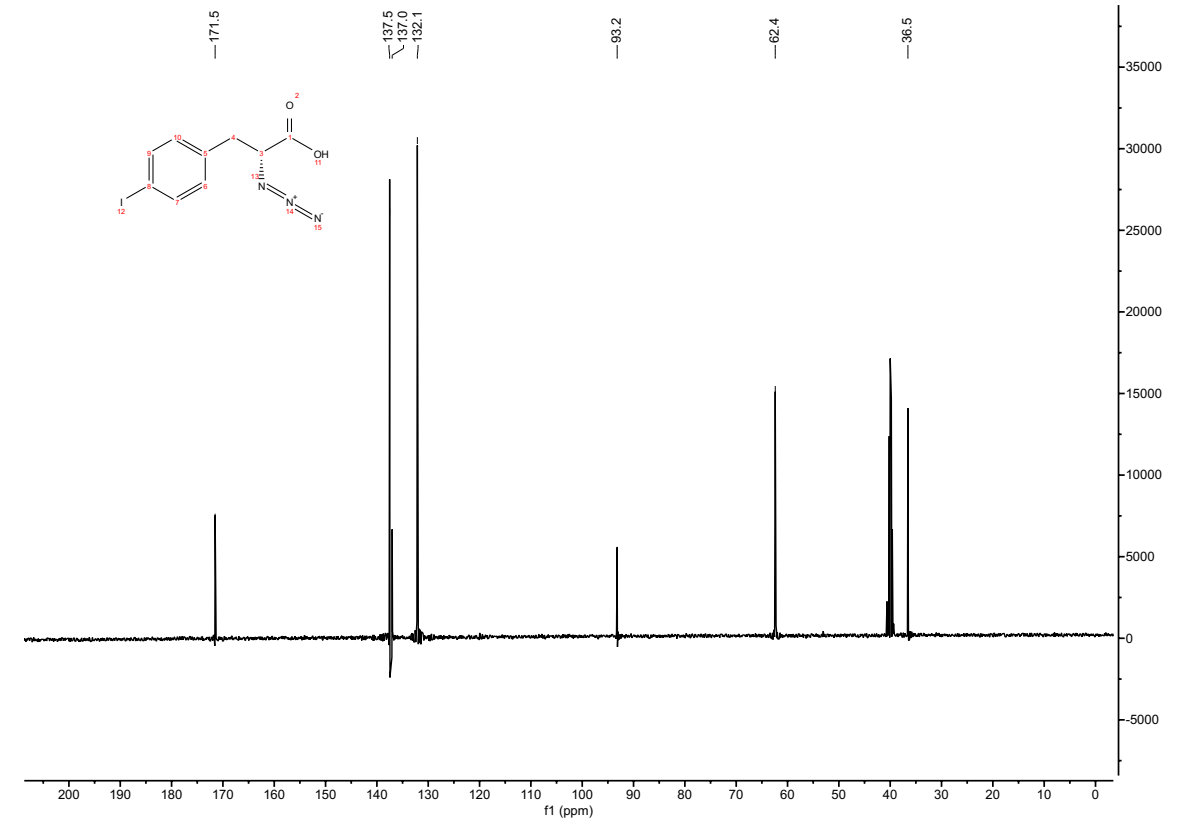
Compound 1S



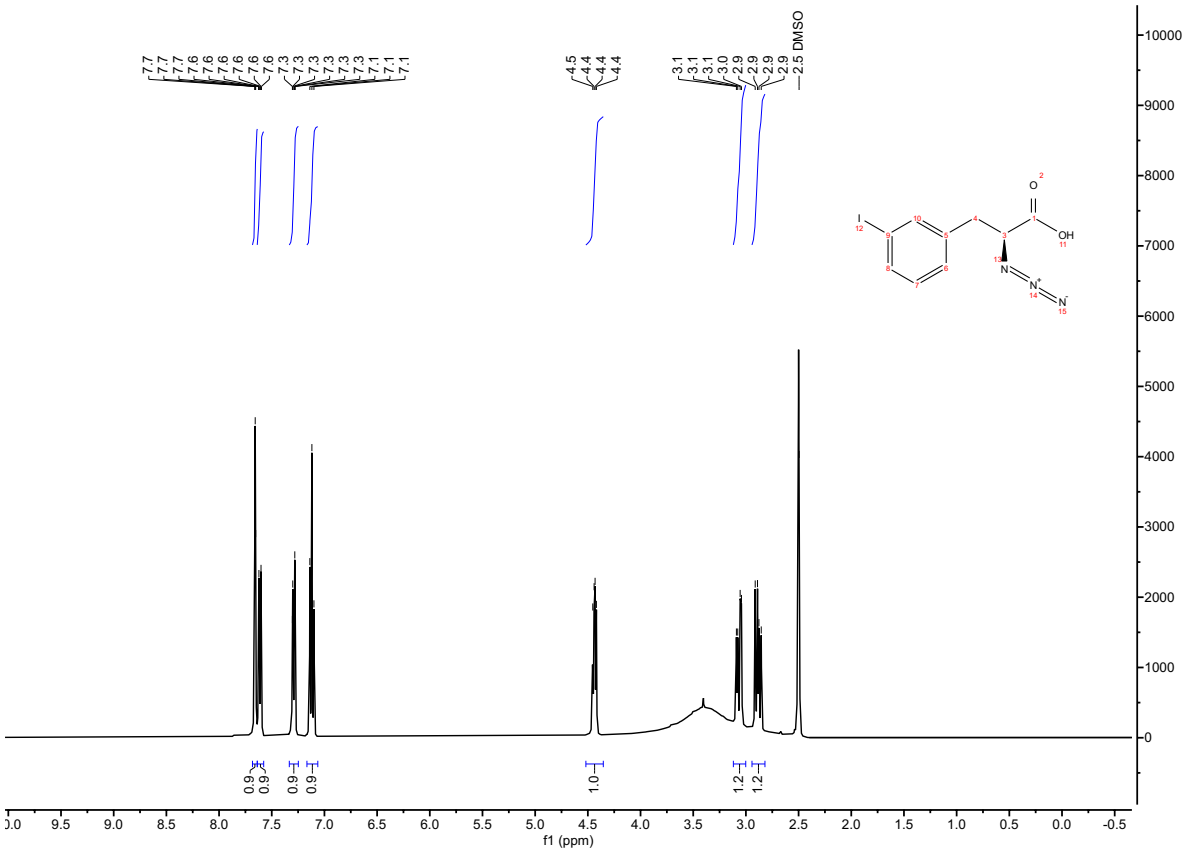


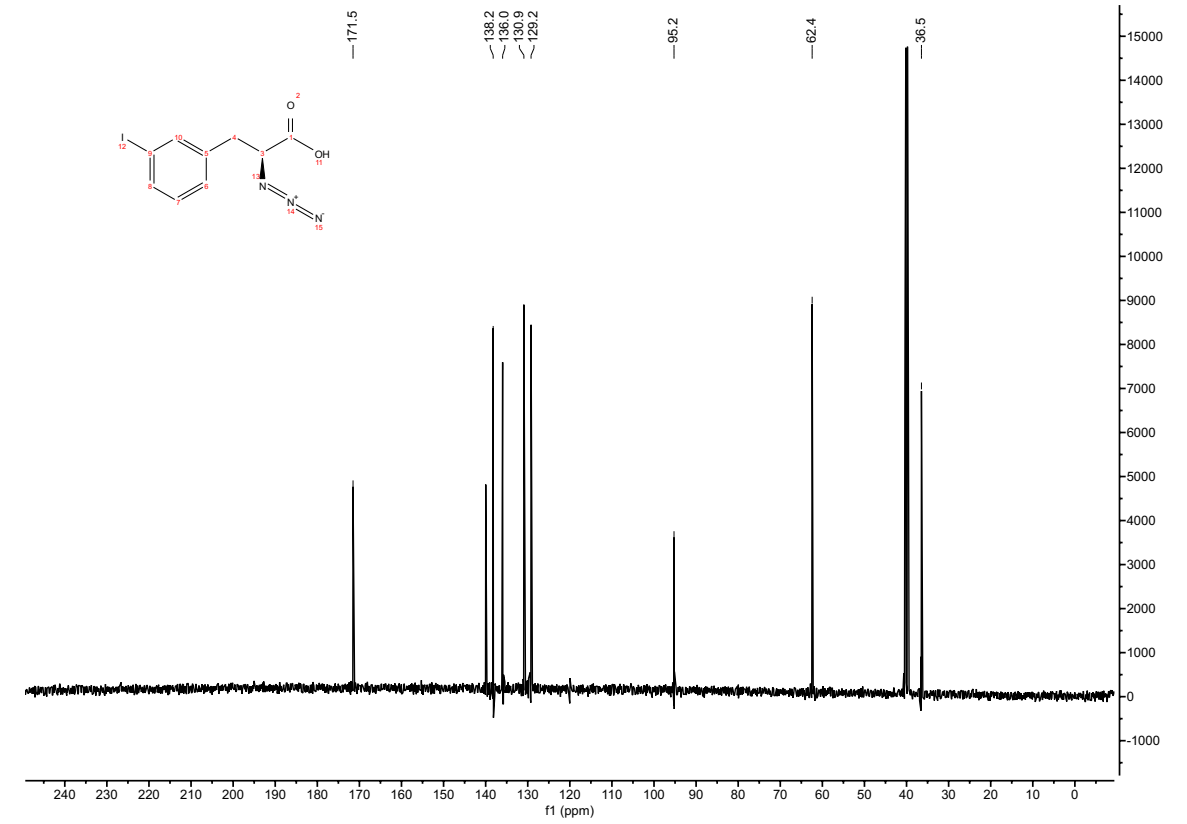
Compound 1R


13.5
7.7
7.7
7.7
7.7
7.7
7.6
7.1
7.1
7.1
7.1
7.1
7.1
4.4
4.4
4.4
4.4
3.7
3.1
3.1
3.0
3.0
2.9
2.9
2.9
2.8
2.5 DMSO
2.5 DMSO
2.5 DMSO
2.5 DMSO-d6
2.5 DMSO
2.5 DMSO
2.5 DMSO
0.9
2.0
2.0
1.0
1.0
1.0
f1 (ppm)

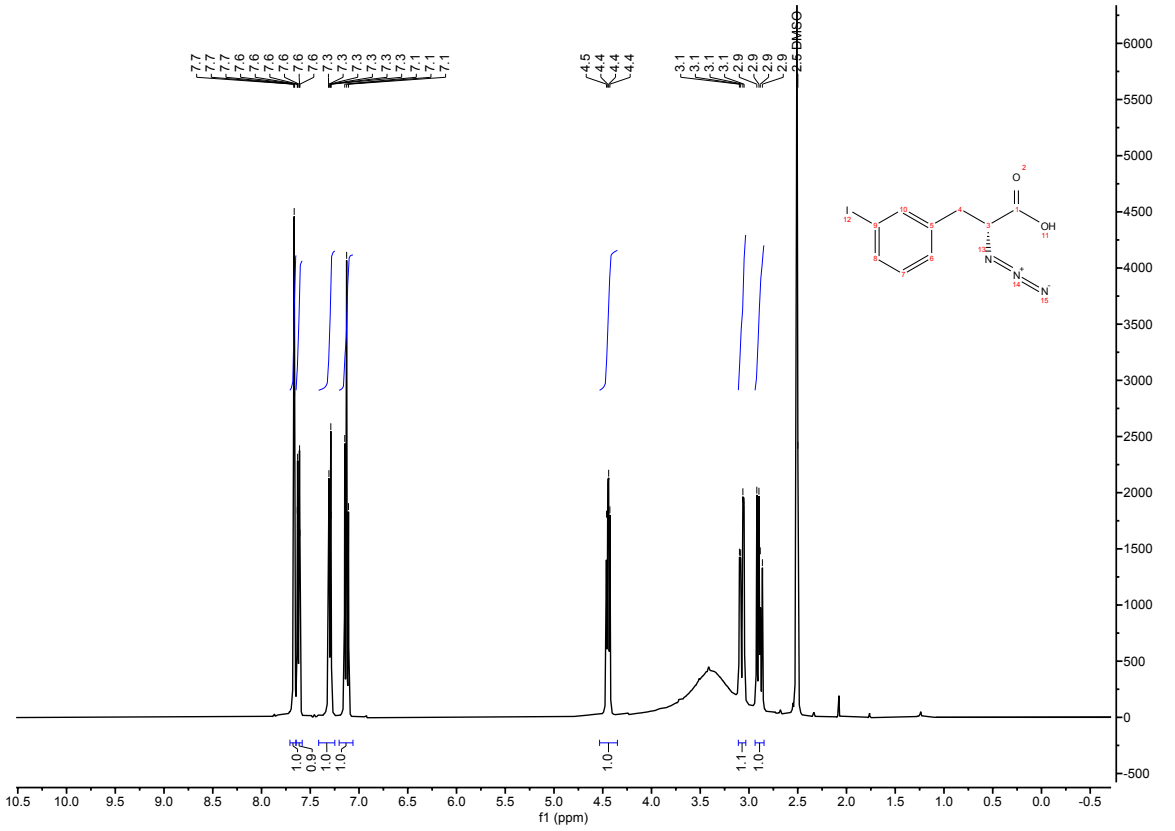

171.5
137.5
137.0
132.1
93.2
62.4
36.5
f1 (ppm)

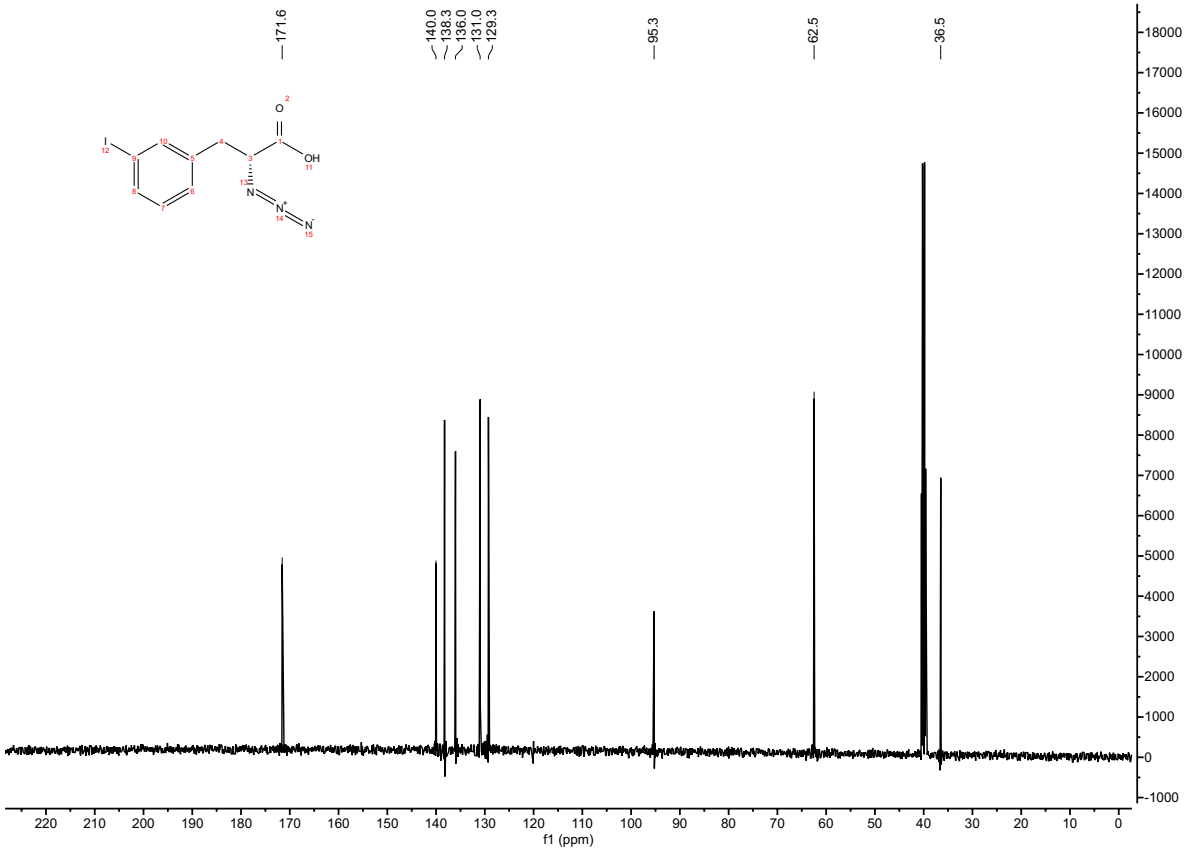
Compound 2S


7.7
7.7
7.7
7.6
7.6
7.6
7.6
7.6
7.6
7.3
7.3
7.3
7.3
7.3
7.3
7.1
7.1
7.1
4.5
4.4
4.4
4.4
3.1
3.1
3.1
3.0
3.0
2.9
2.9
2.9
2.9
2.5 DMSO
0.9
0.9
0.9
0.9
1.0
1.2
1.2
f1 (ppm)

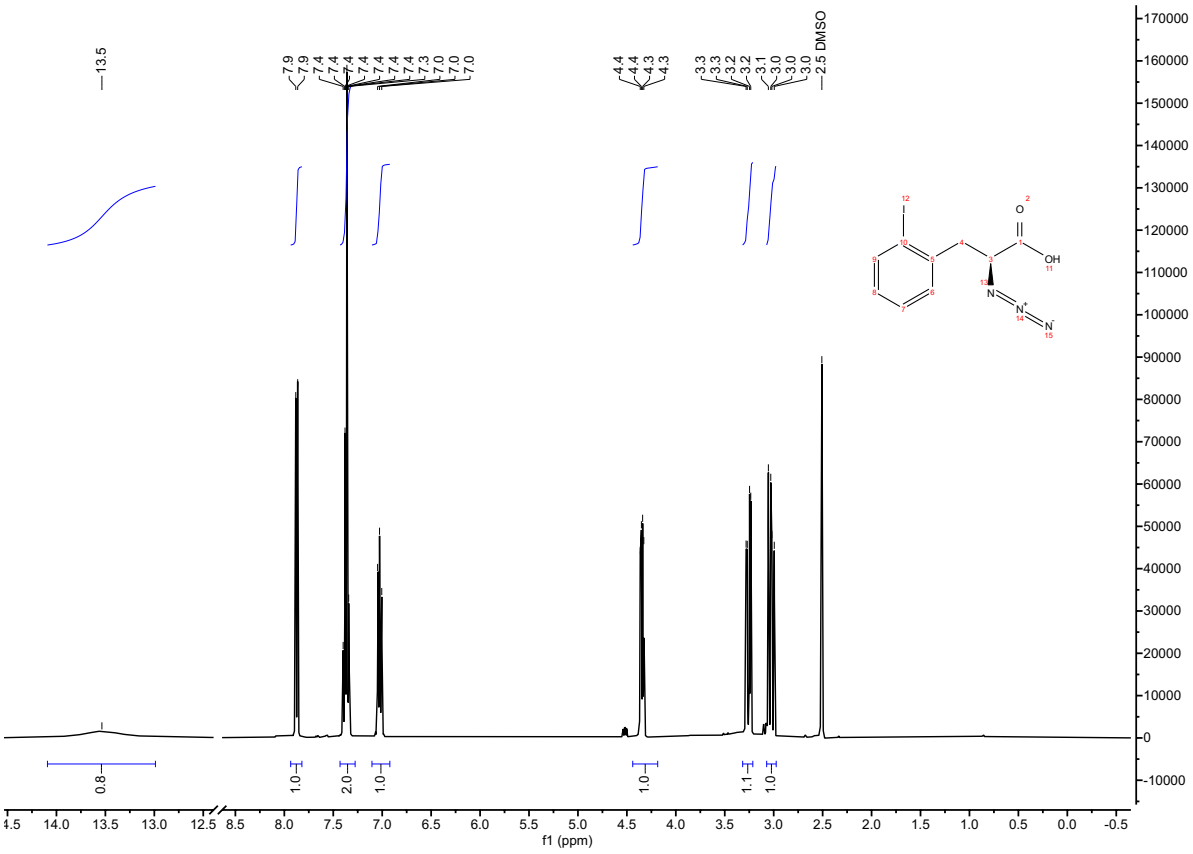

171.5
138.2
136.0
130.9
129.2
95.2
62.4
36.5
f1 (ppm)

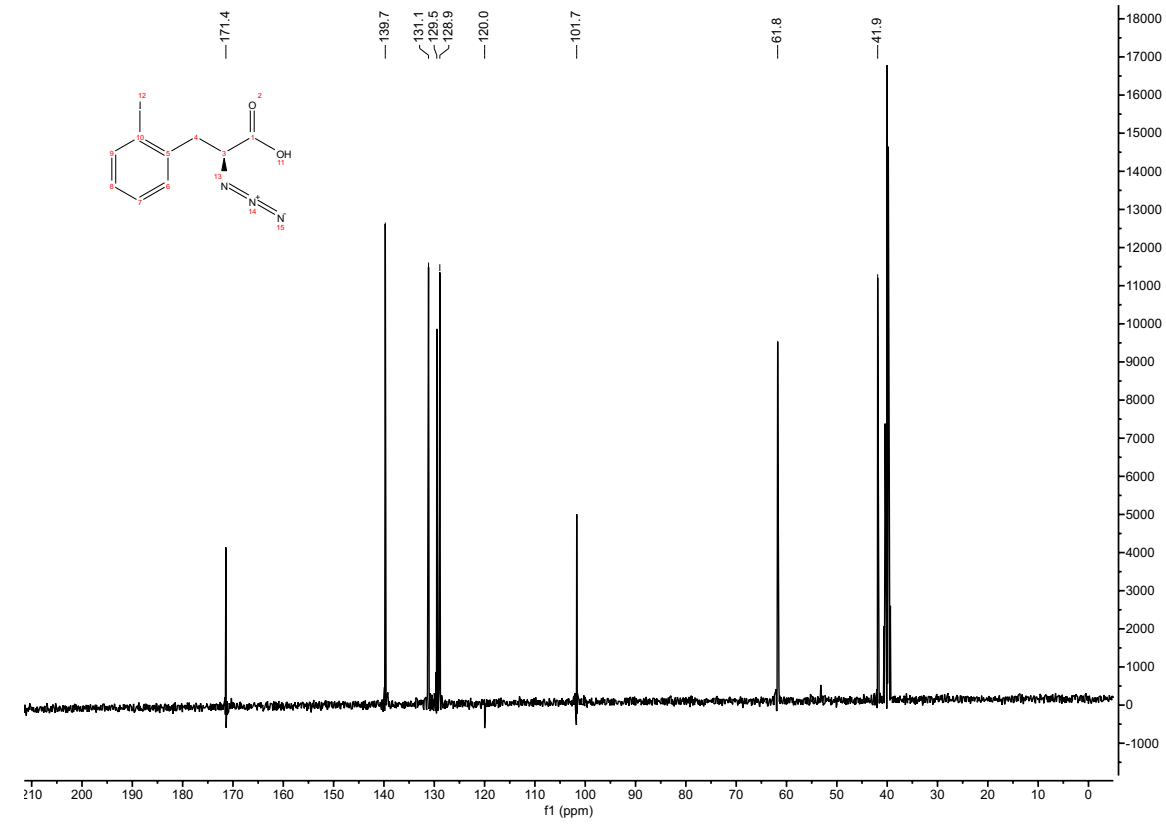
Compound 2R


7.7
7.7
7.7
7.6
7.6
7.6
7.6
7.6
7.6
7.3
7.3
7.3
7.3
7.3
7.3
7.1
7.1
7.1
4.5
4.4
4.4
4.4
3.1
3.1
3.1
3.1
2.9
2.9
2.9
2.9
2.5 DMSO
1.0
0.9
1.0
1.0
1.0
1.1
1.0
f1 (ppm)

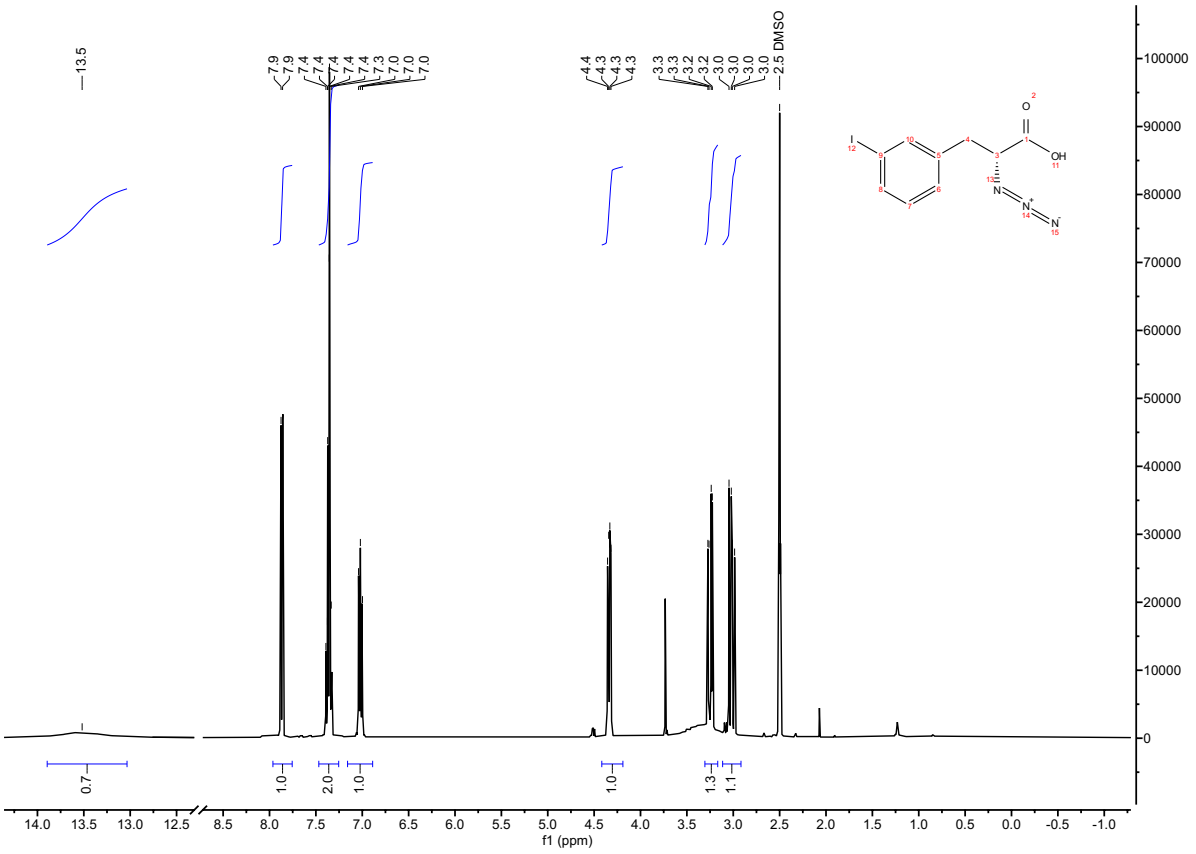

171.6
140.0
138.3
136.0
131.0
129.3
95.3
62.5
36.5
f1 (ppm)

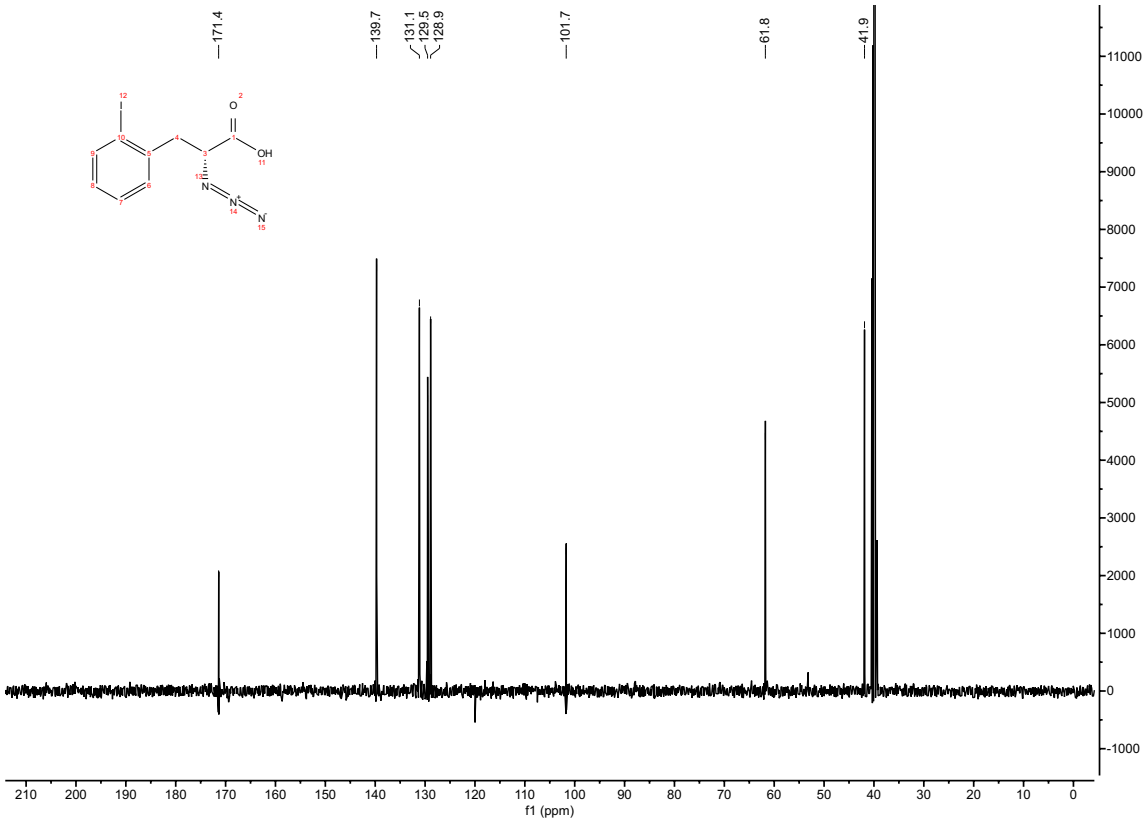
Compound 3S


13.5
7.9
7.9
7.4
7.4
7.4
7.4
7.4
7.4
7.4
7.3
7.0
7.0
7.0
4.4
4.4
4.3
4.3
3.3
3.3
3.2
3.2
3.1
3.0
3.0
3.0
2.5 DMSO
0.8
1.0
2.0
1.0
1.0
1.1
1.0
f1 (ppm)

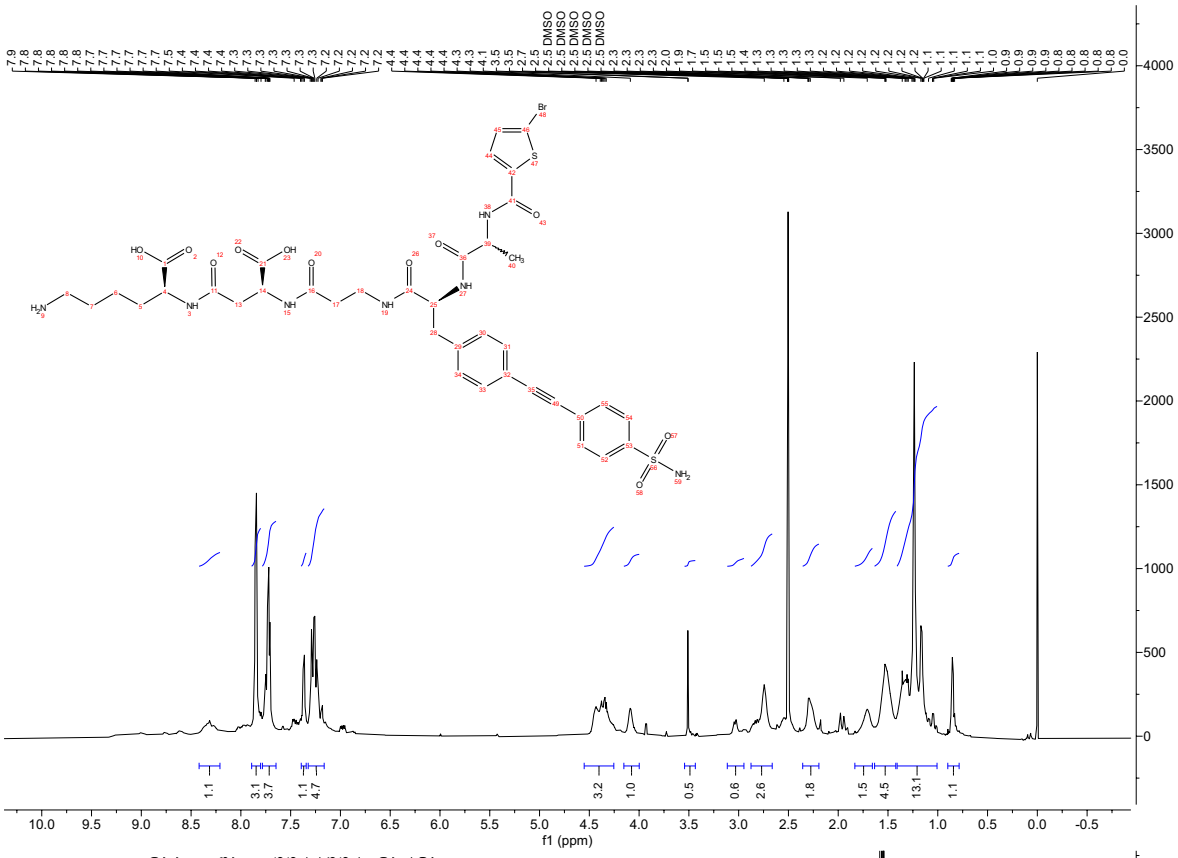

171.4
139.7
131.1
129.5
128.9
120.0
101.7
61.8
41.9
f1 (ppm)

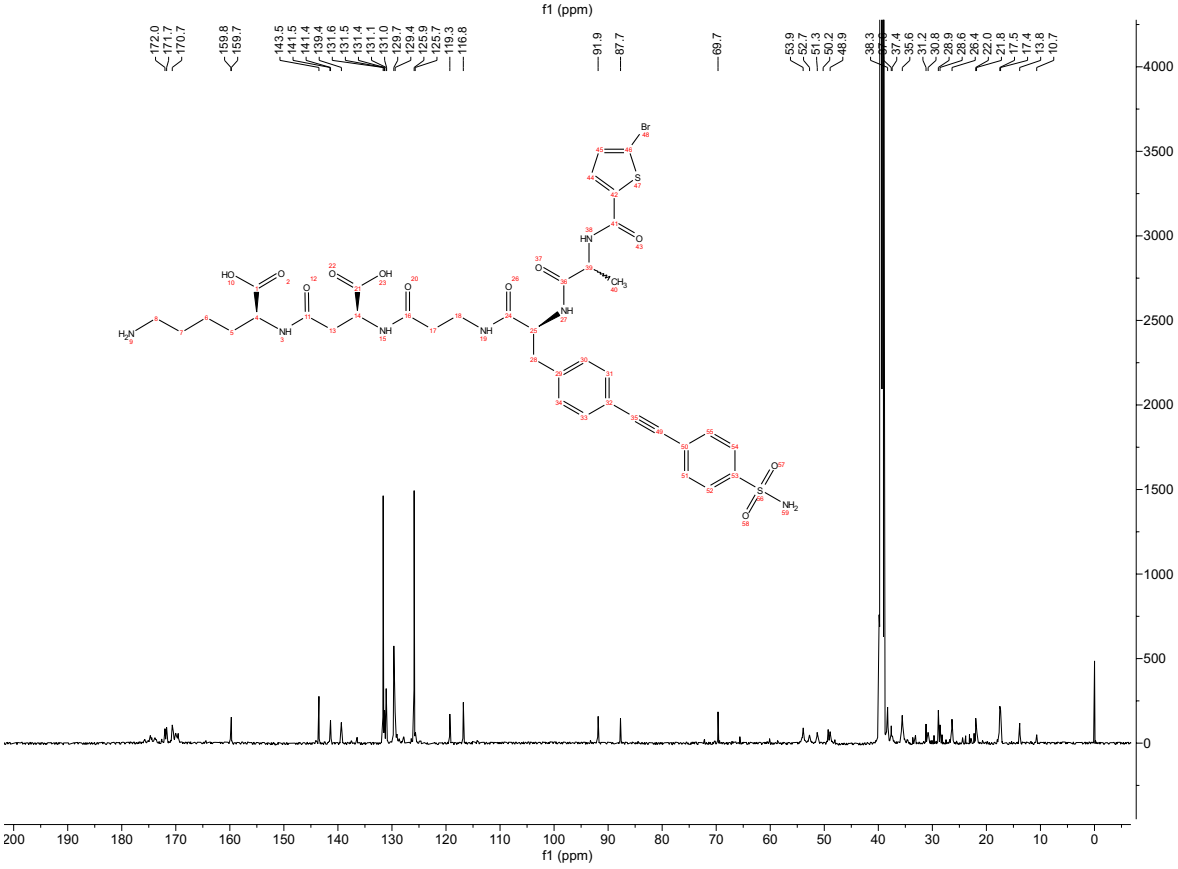
Compound 3R


13.5
7.9
7.9
7.4
7.4
7.4
7.4
7.3
7.0
7.0
7.0
4.4
4.3
4.3
4.3
3.3
3.3
3.2
3.2
3.0
3.0
3.0
3.0
2.5 DMSO
0.7
1.0
2.0
1.0
1.0
1.3
1.1
f1 (ppm)

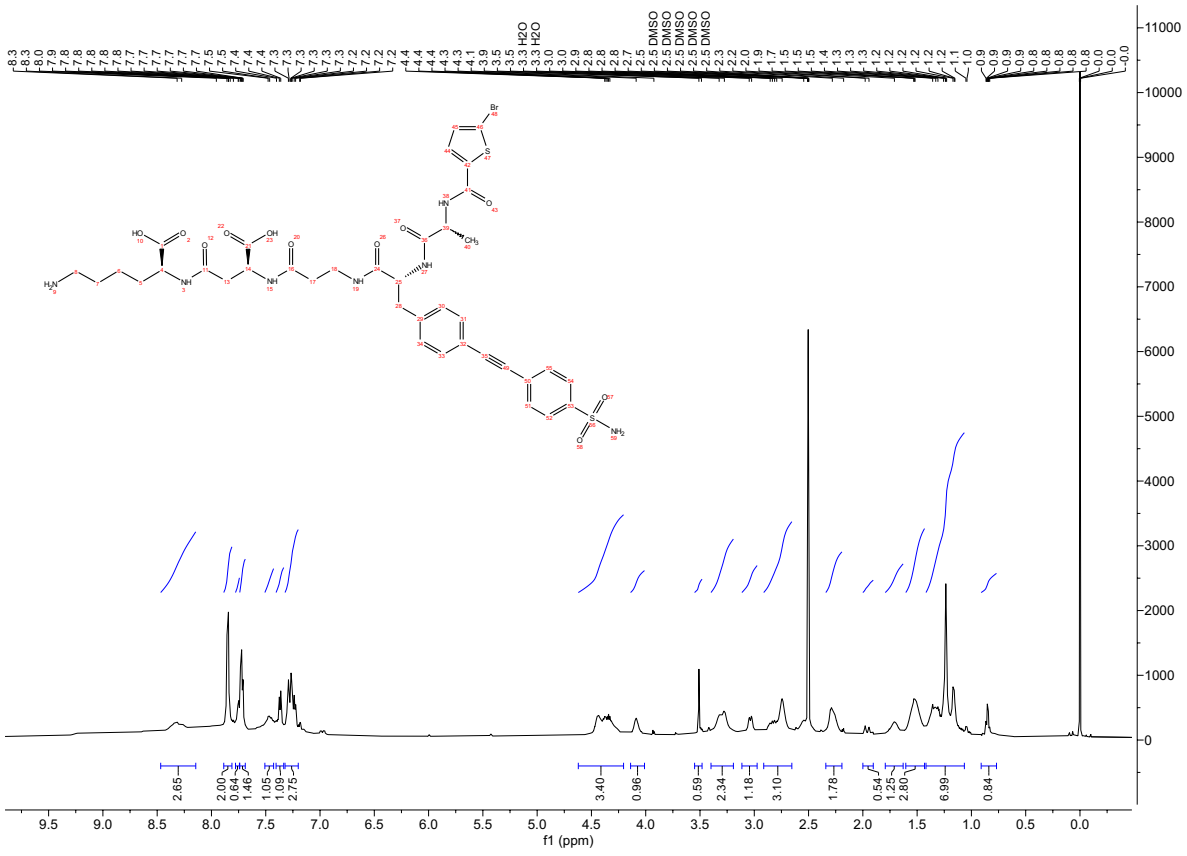

171.4
139.7
131.1
129.5
128.9
101.7
61.8
41.9
f1 (ppm)

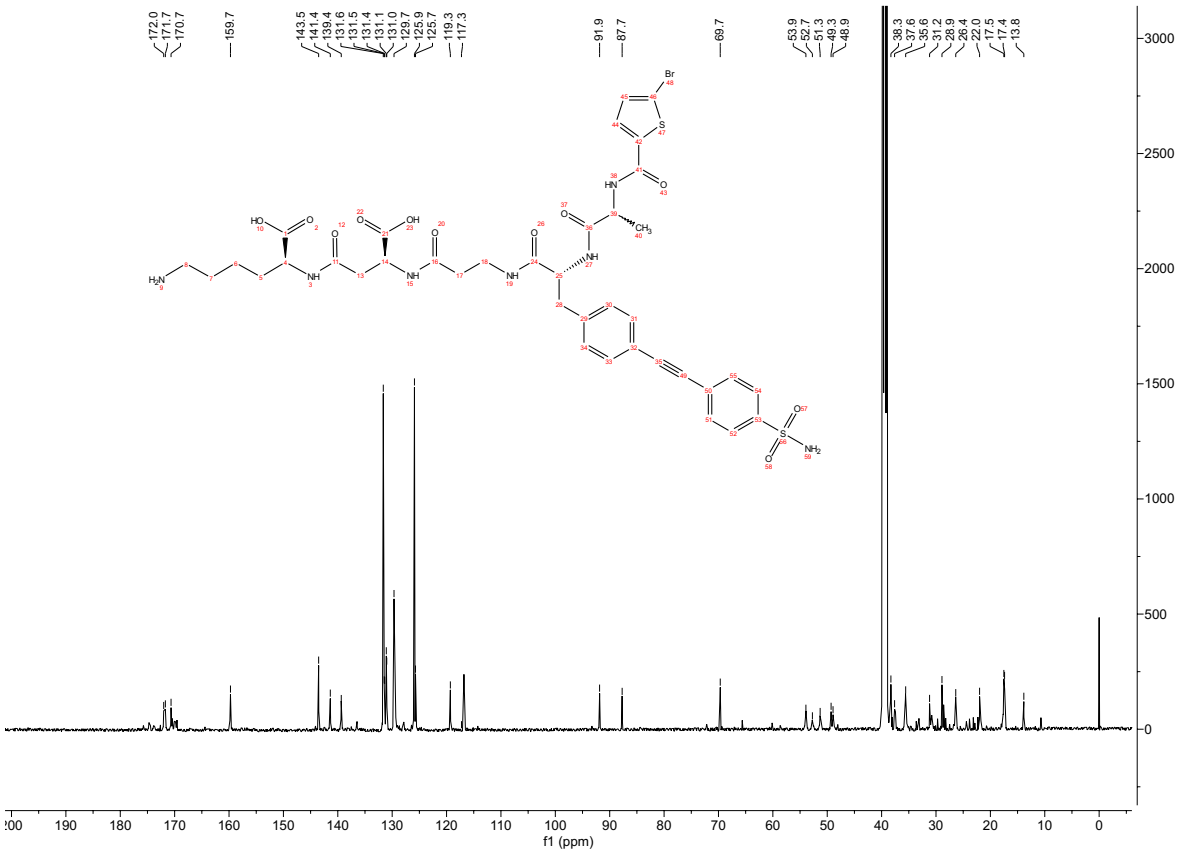
Compound 7



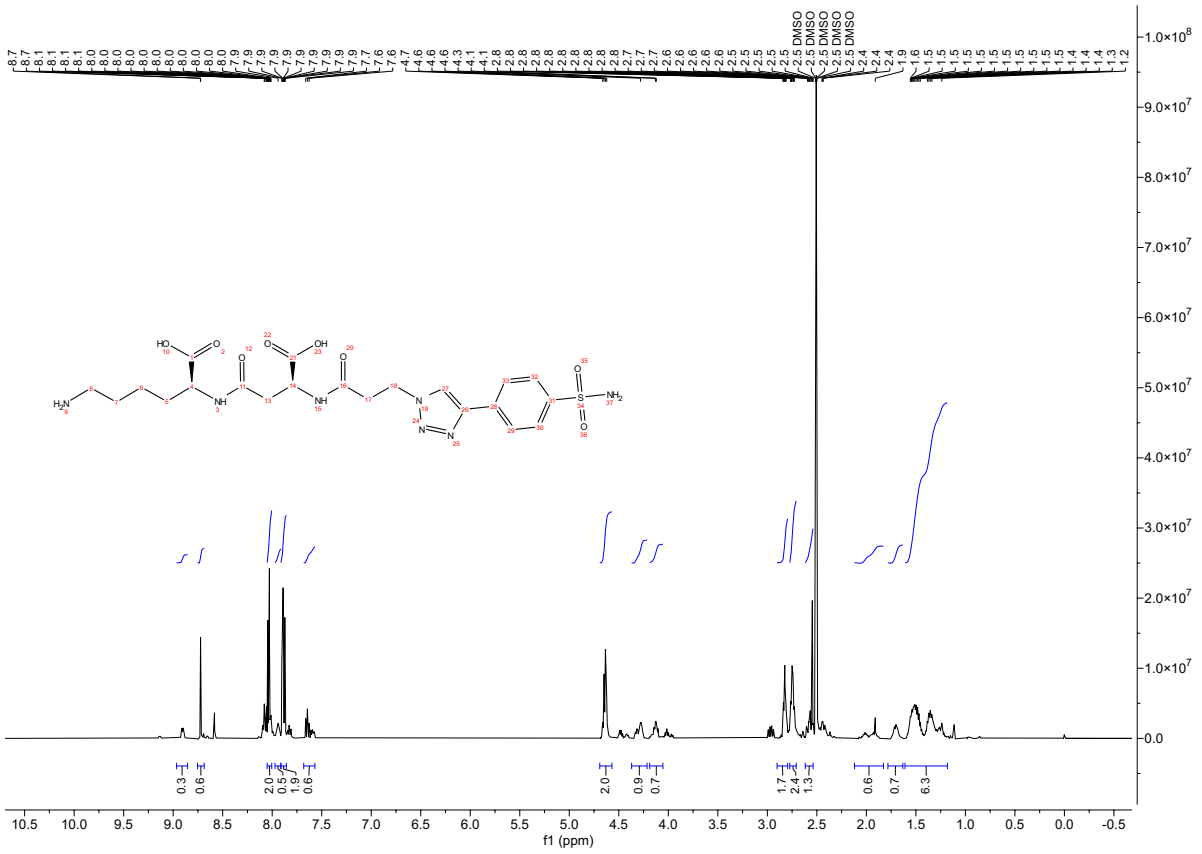


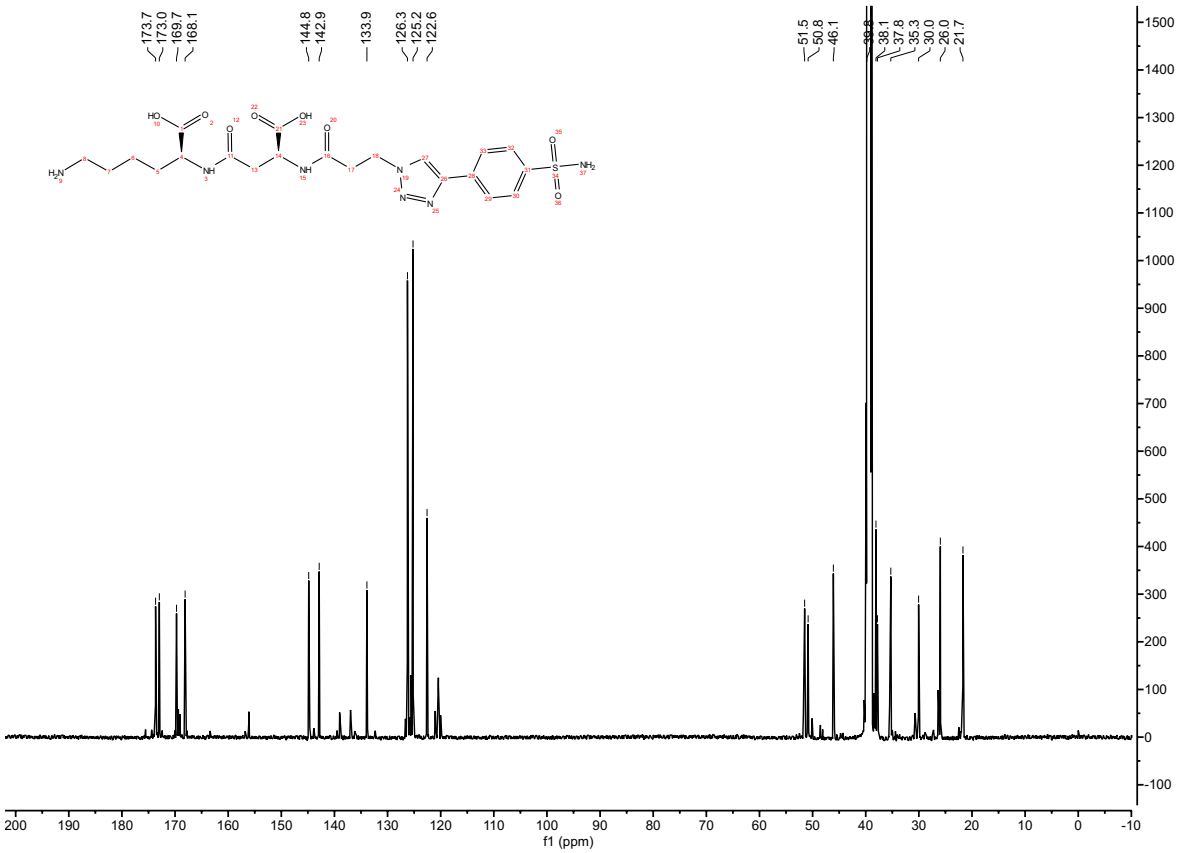
Compound 9



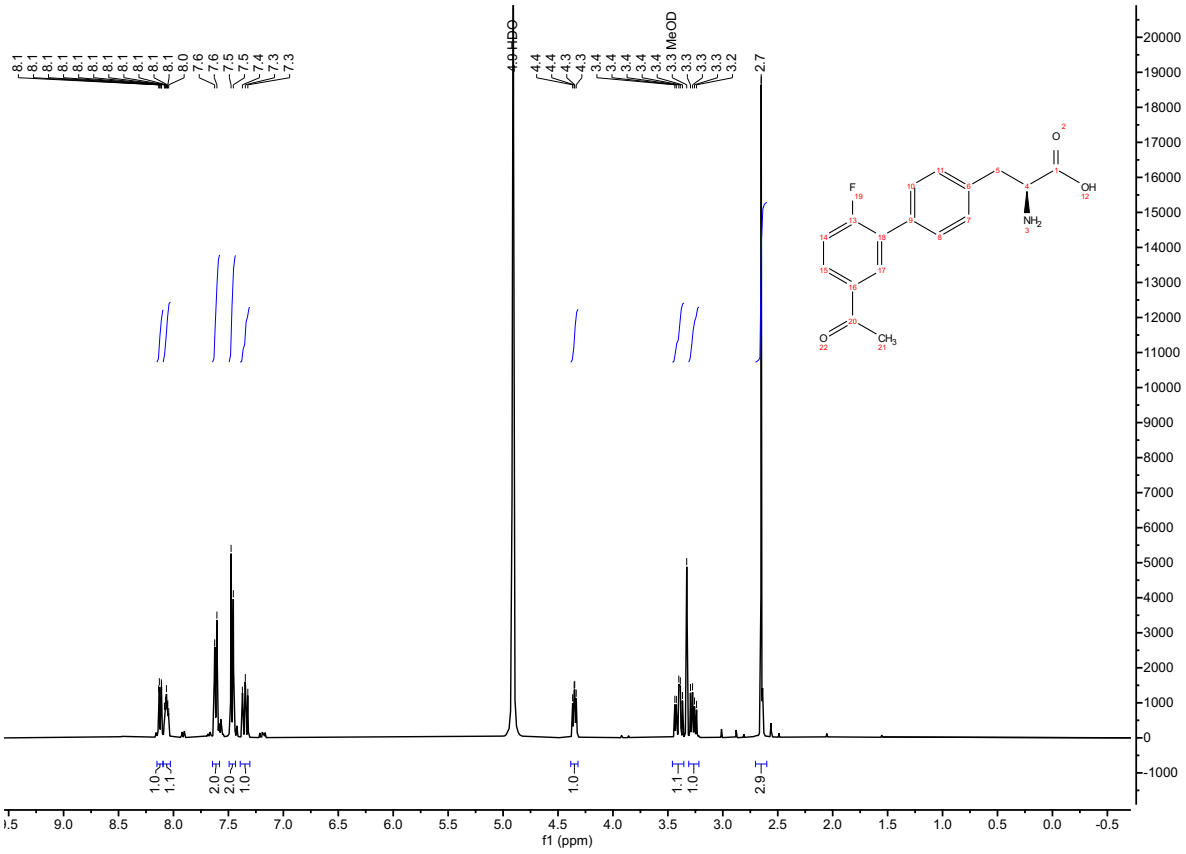


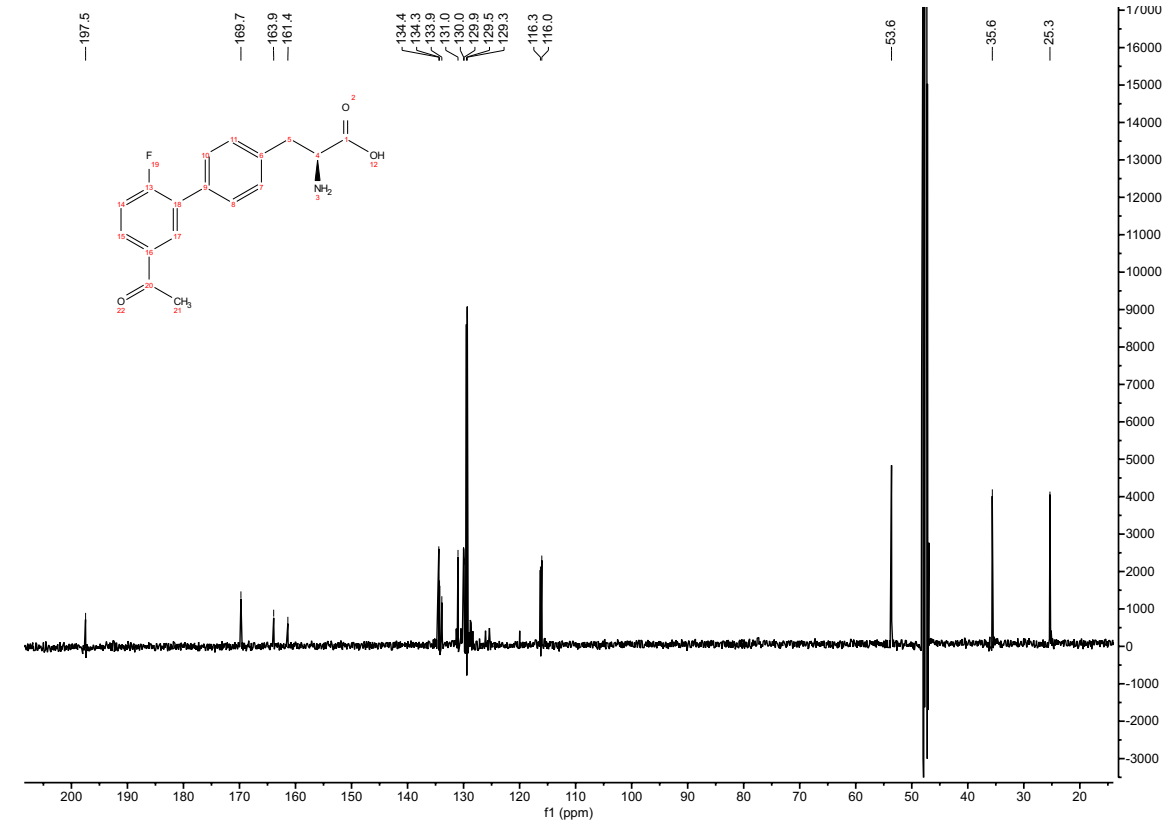
Compound 11


DMSO
f1 (ppm)

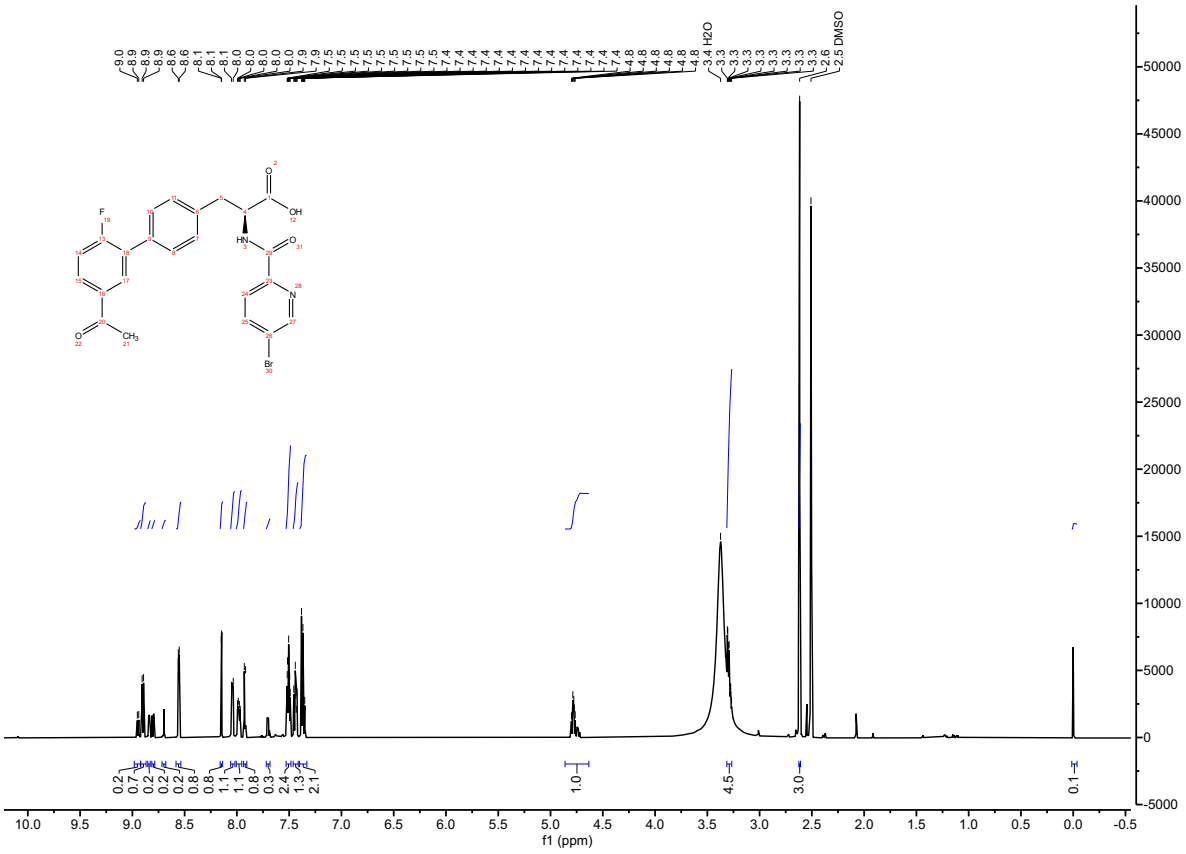

173.7
173.0
169.7
168.1
144.8
142.9
133.9
126.3
125.2
122.6
51.5
50.8
46.1
38.1
37.8
35.3
30.0
26.0
21.7
f1 (ppm)

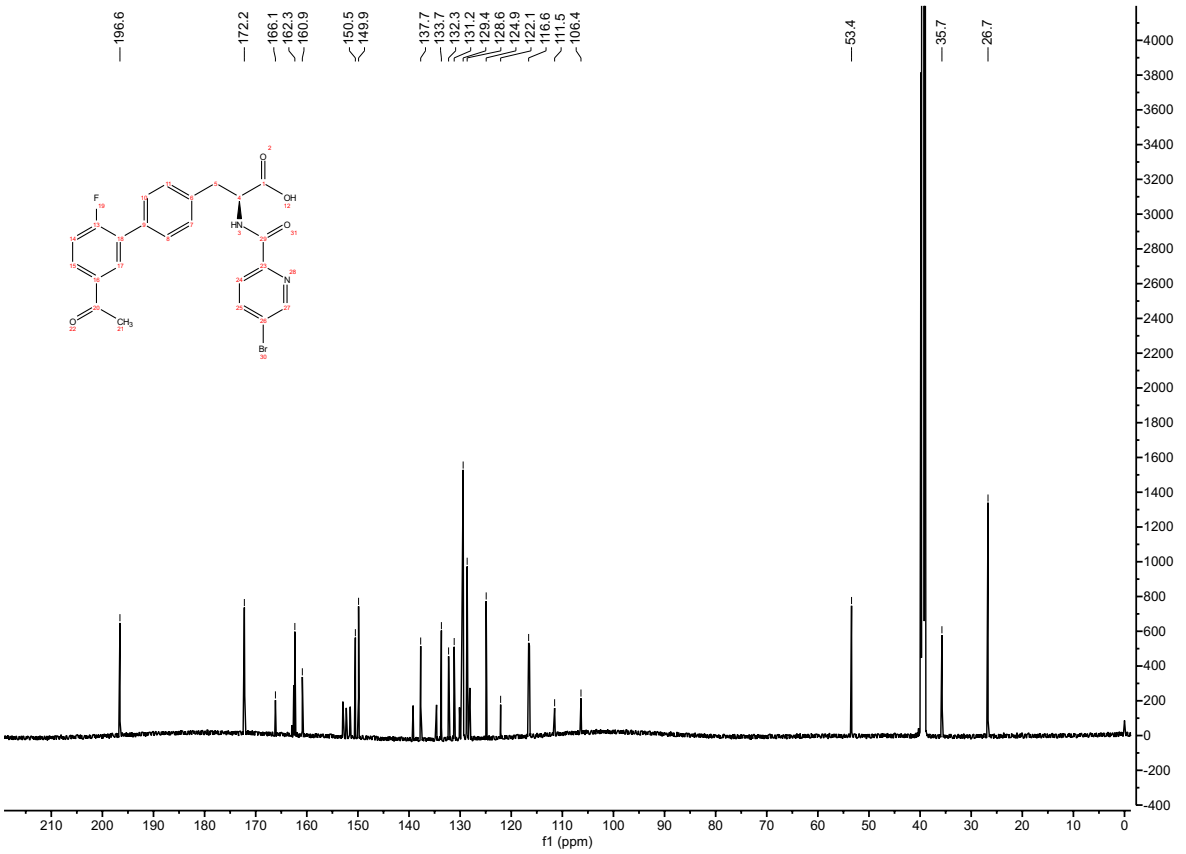
Compound 31


4.9 HDO
3.3 MeOD
2.7
f1 (ppm)

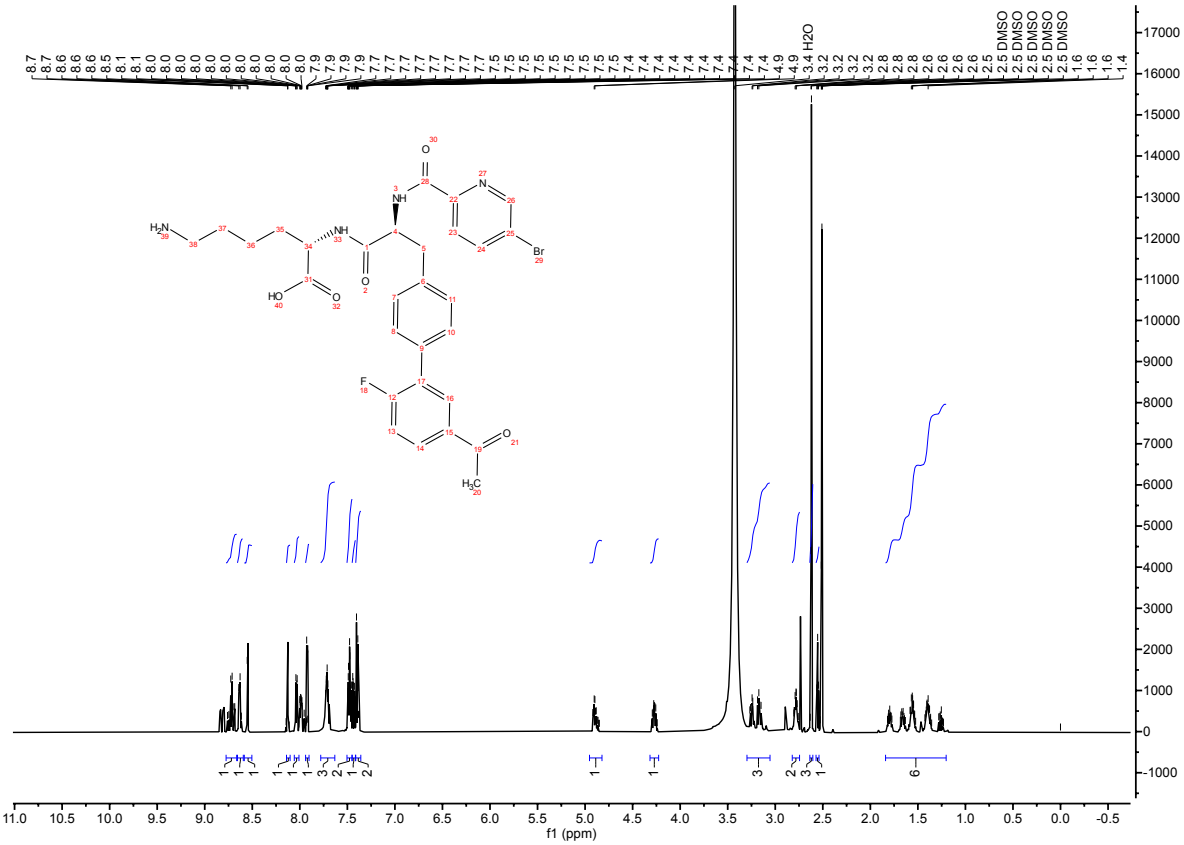

197.5
169.7
163.9
161.4
134.4
134.3
133.9
131.0
130.0
129.9
129.5
129.3
116.3
116.0
53.6
35.6
25.3
f1 (ppm)

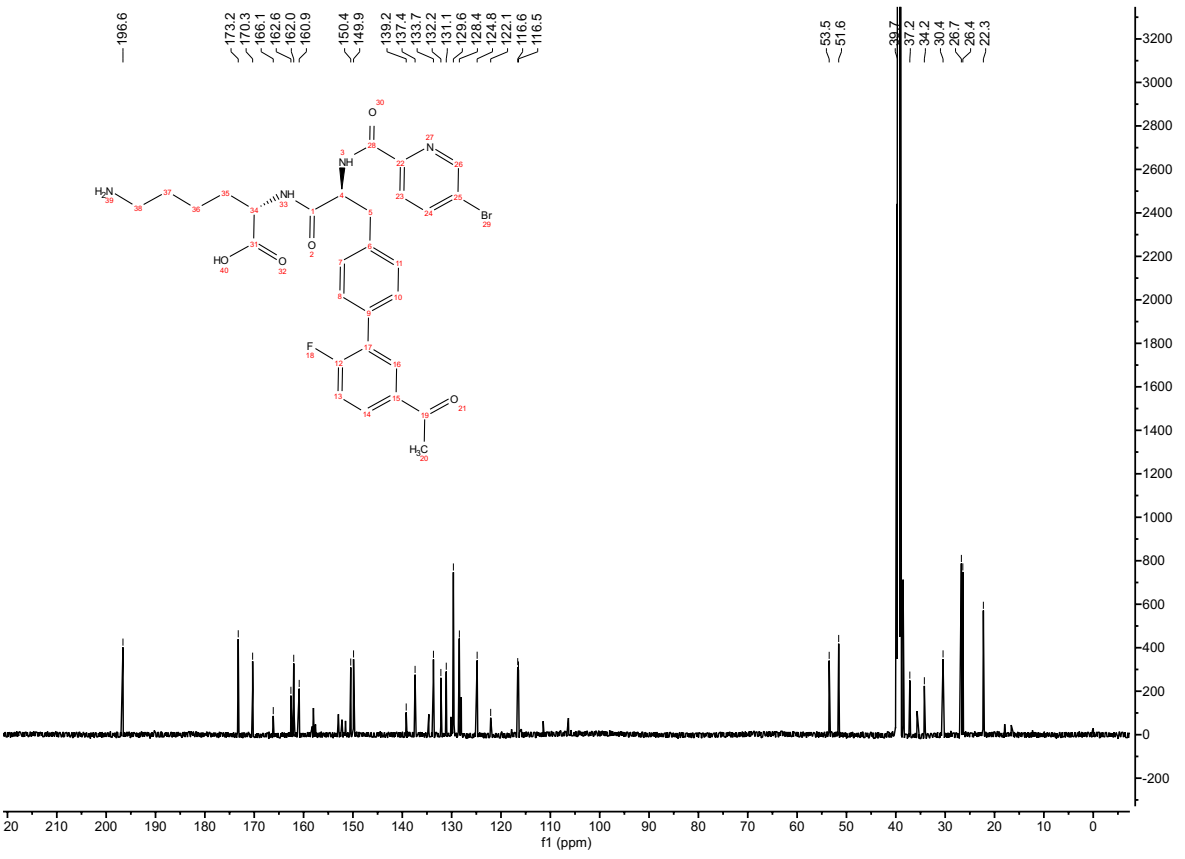
Compound 32


H2O
DMSO
f1 (ppm)

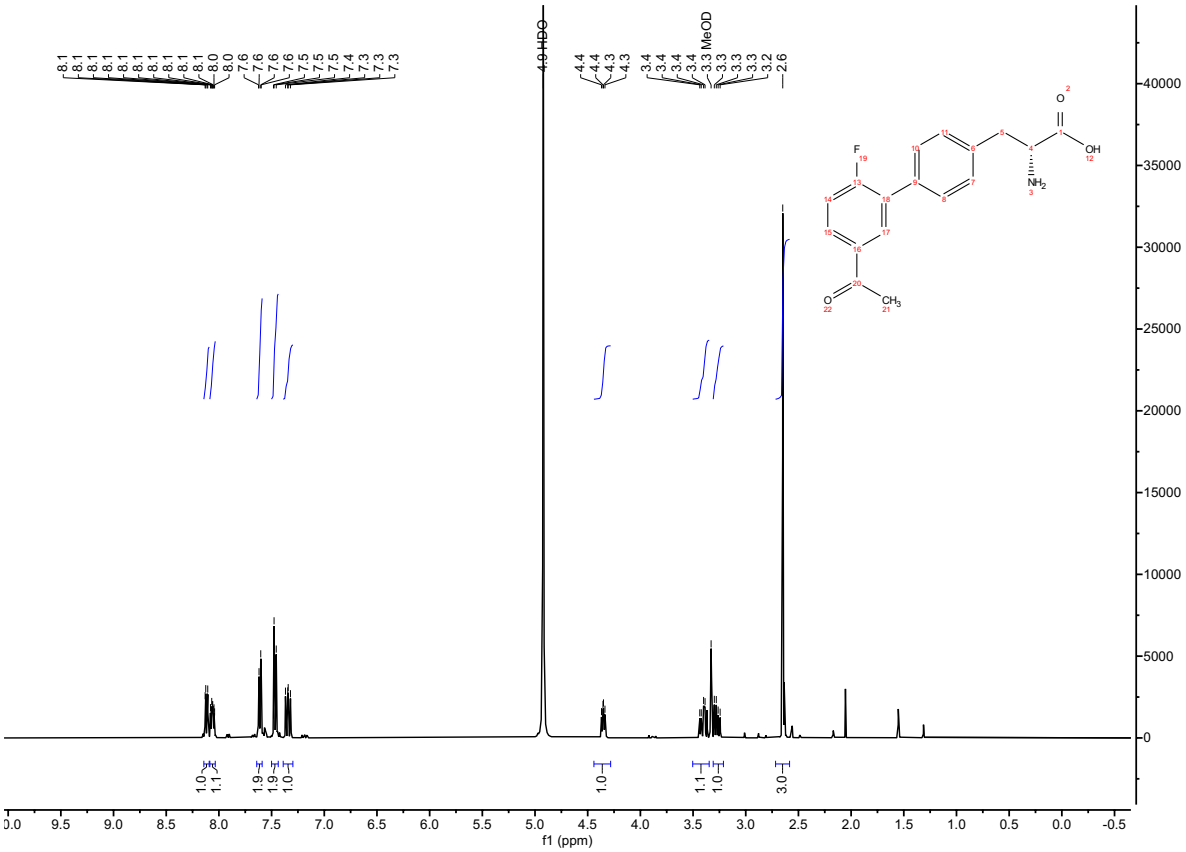

196.6
172.2
166.1
162.3
160.9
150.5
149.9
137.7
133.7
132.3
131.2
129.4
128.6
124.9
122.1
116.6
111.5
106.4
53.4
35.7
26.7
f1 (ppm)

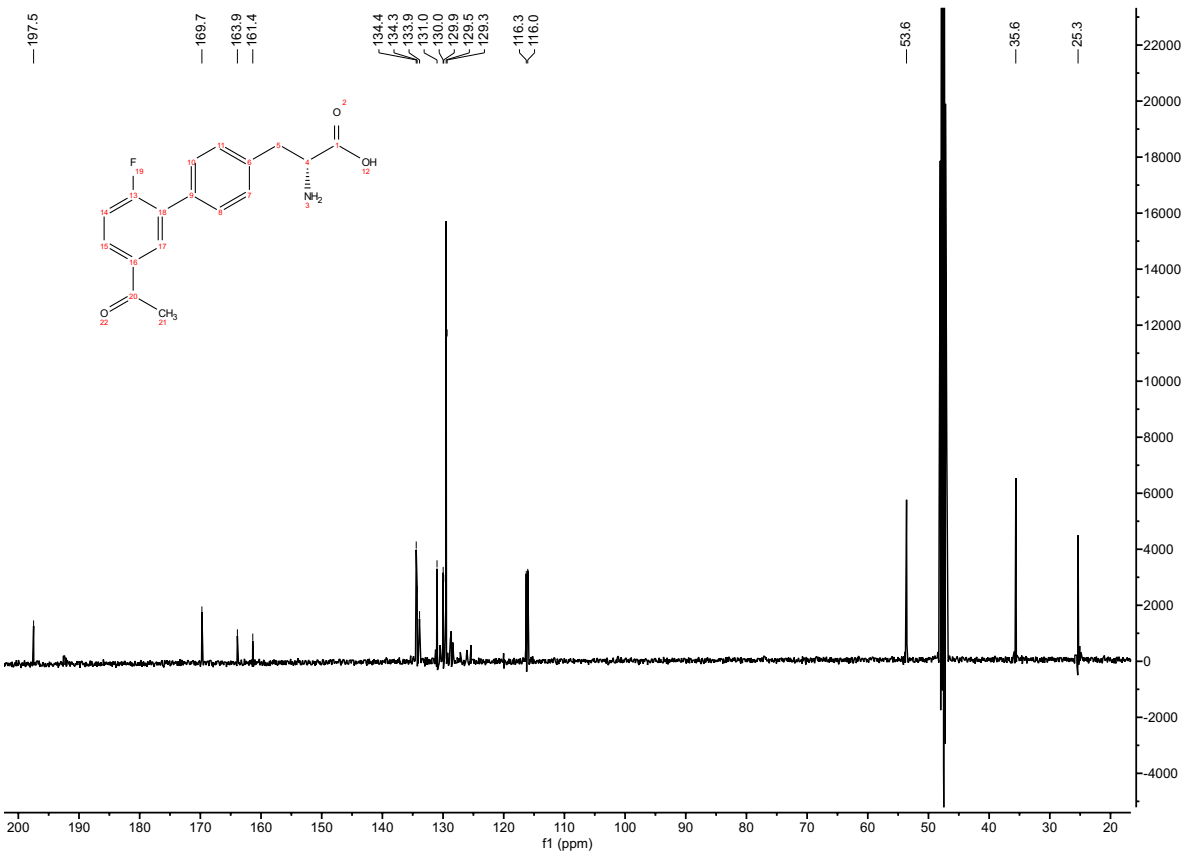
Compound 34


3.4 H2O
2.5 DMSO
f1 (ppm)

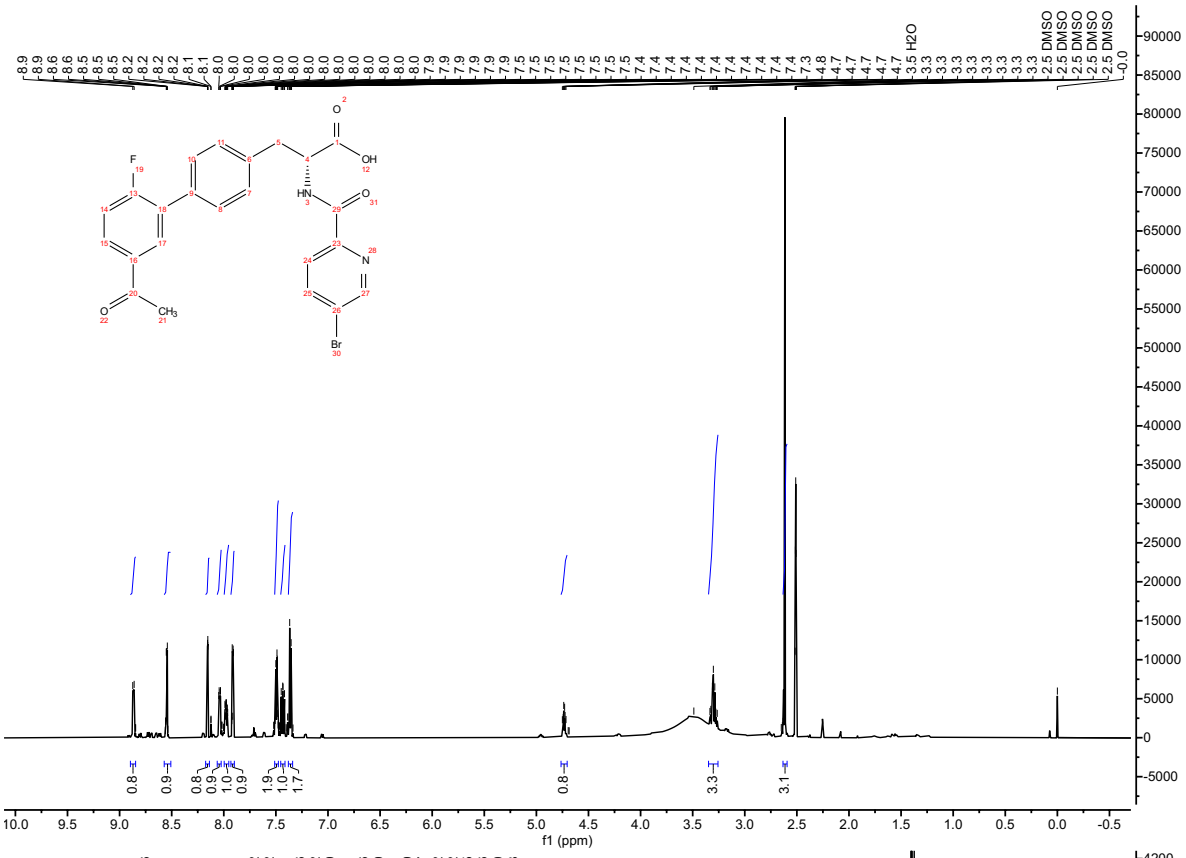

196.6
173.2
170.3
166.1
162.6
162.0
160.9
150.4
149.9
139.2
137.4
133.7
132.2
131.1
129.6
128.4
124.8
122.1
116.6
116.5
53.5
51.6
39.7
37.2
34.2
30.4
26.7
26.4
22.3
f1 (ppm)

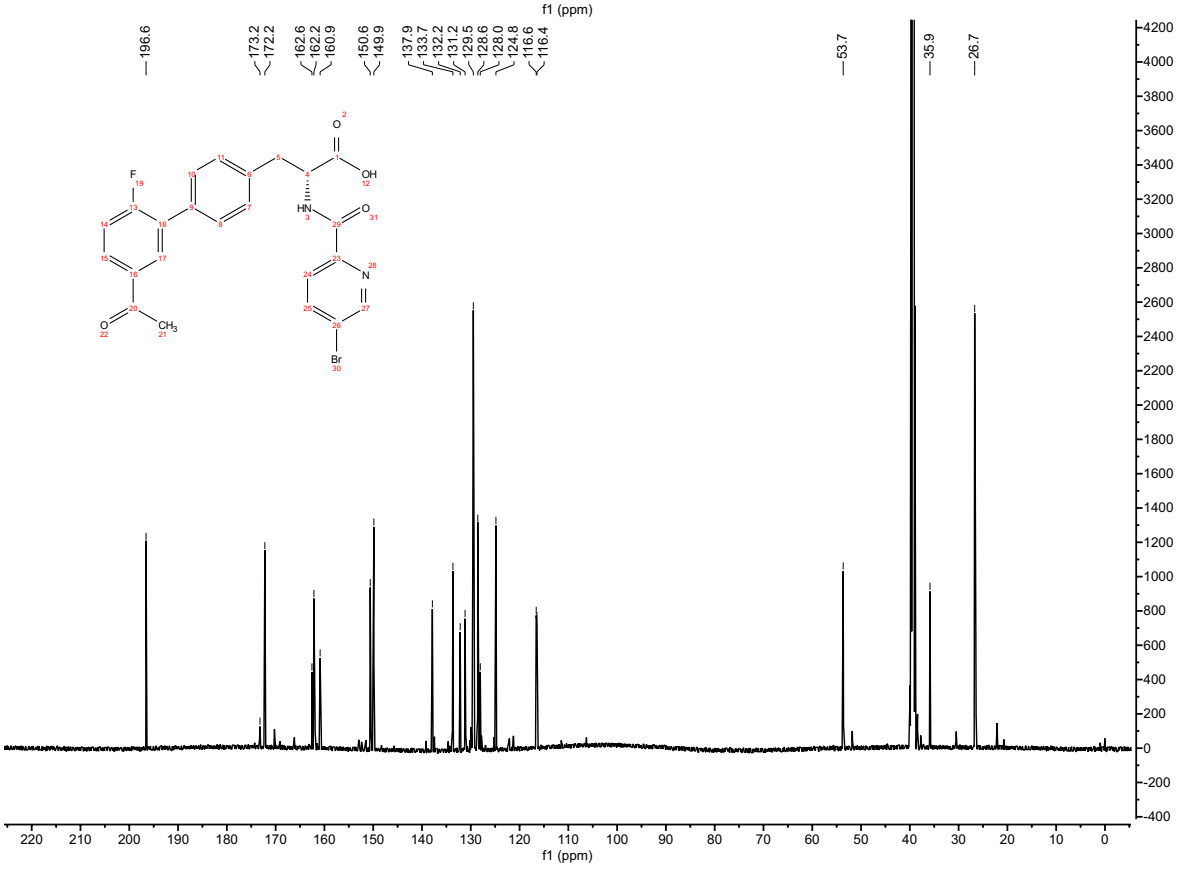
Compound 35


4.9 HDO
3.3 MeOD
f1 (ppm)

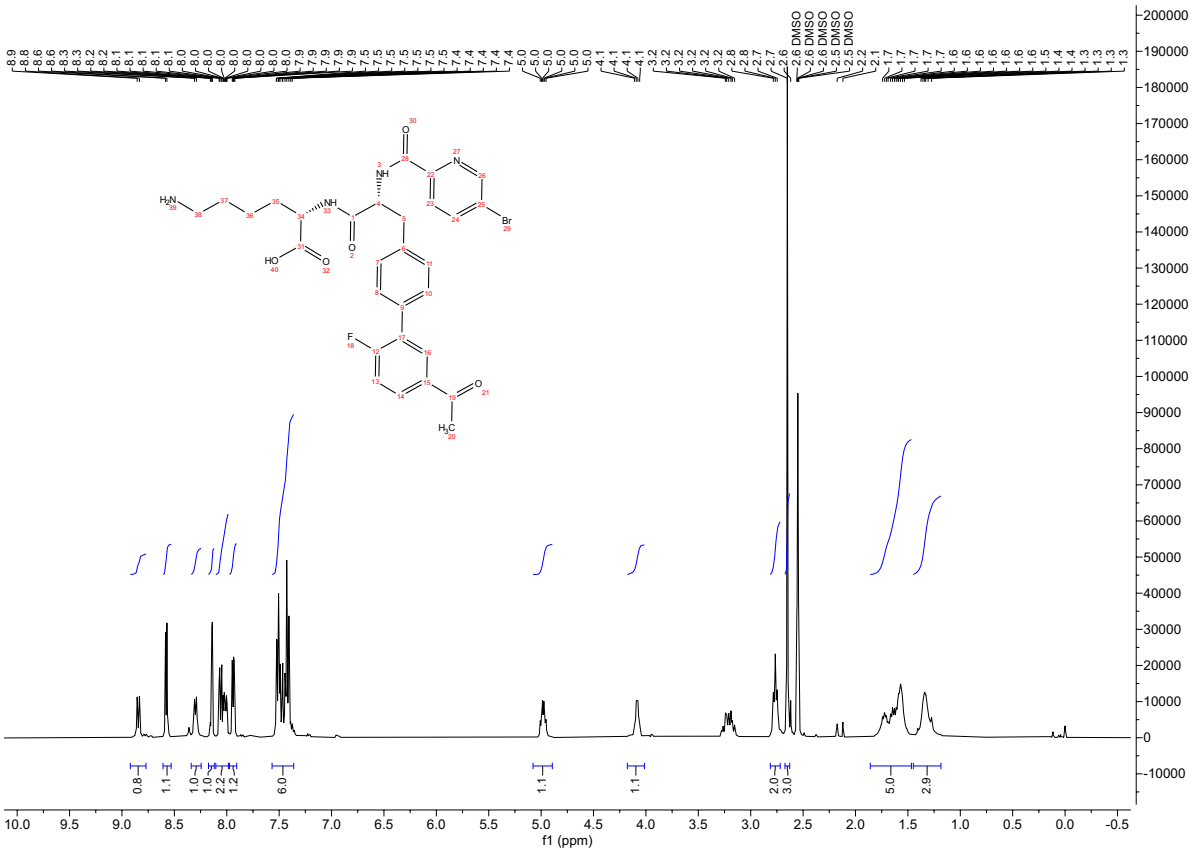

197.5
169.7
163.9
161.4
134.4
134.3
133.9
131.0
130.0
129.9
129.5
129.3
116.3
116.0
53.6
35.6
25.3
f1 (ppm)

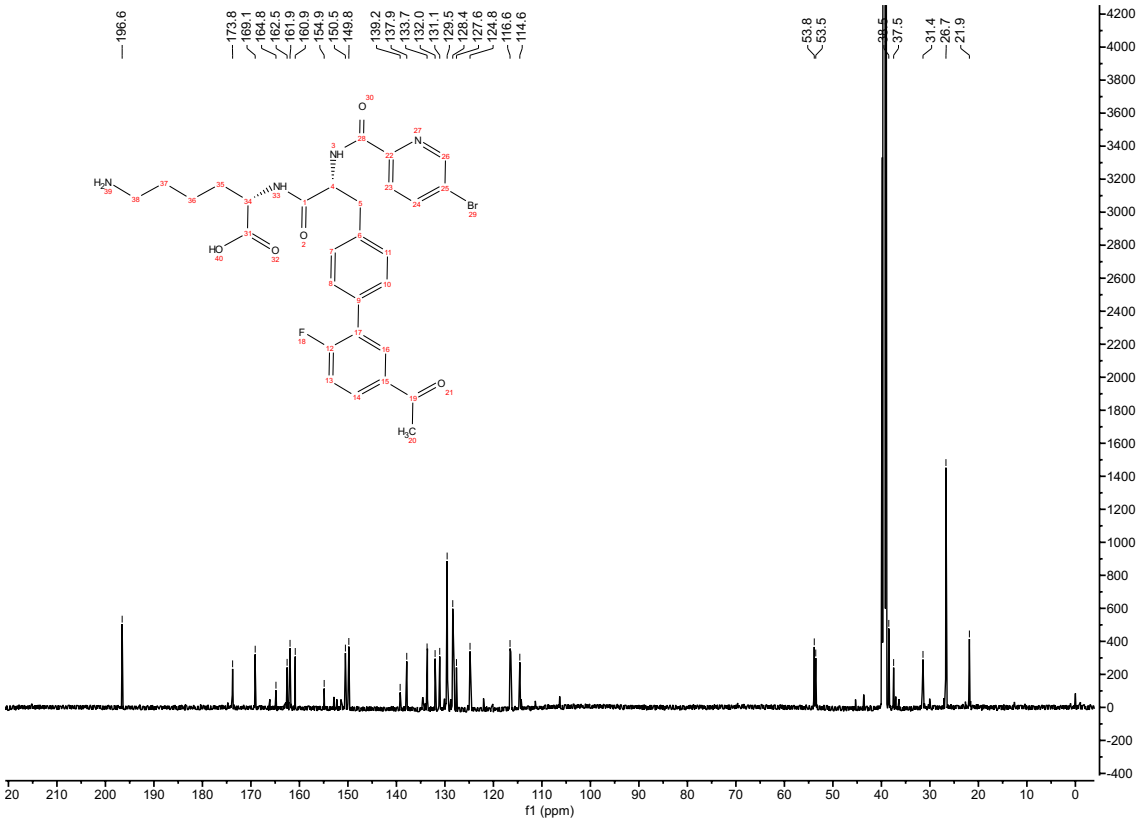
Compound 36


2.5 DMSO
3.5 H2O
0.8
0.9
0.8
0.9
1.0
0.9
1.9
1.0
1.7
0.8
3.3
3.1
f1 (ppm)

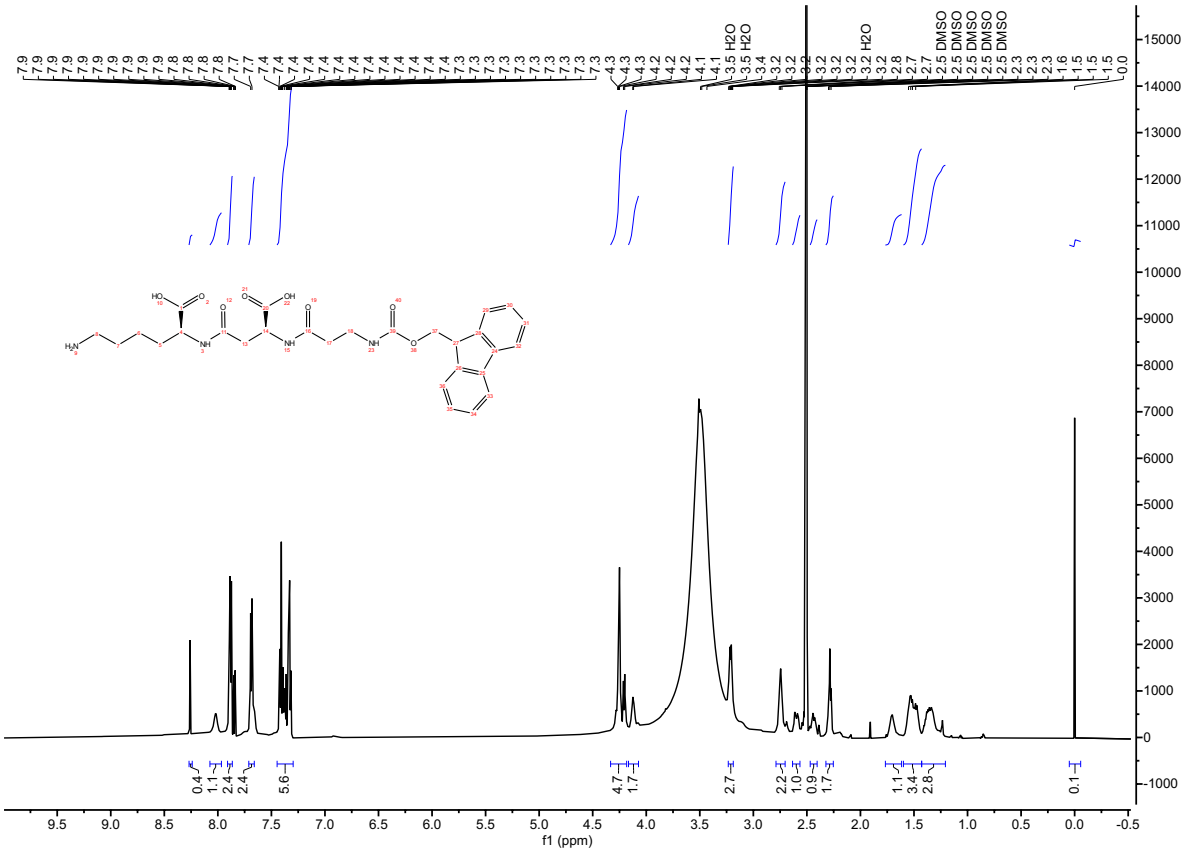

196.6
173.2
172.2
162.6
162.2
160.9
150.6
149.9
137.9
133.7
132.2
131.2
129.5
128.6
128.0
124.8
116.6
116.4
53.7
35.9
26.7
f1 (ppm)

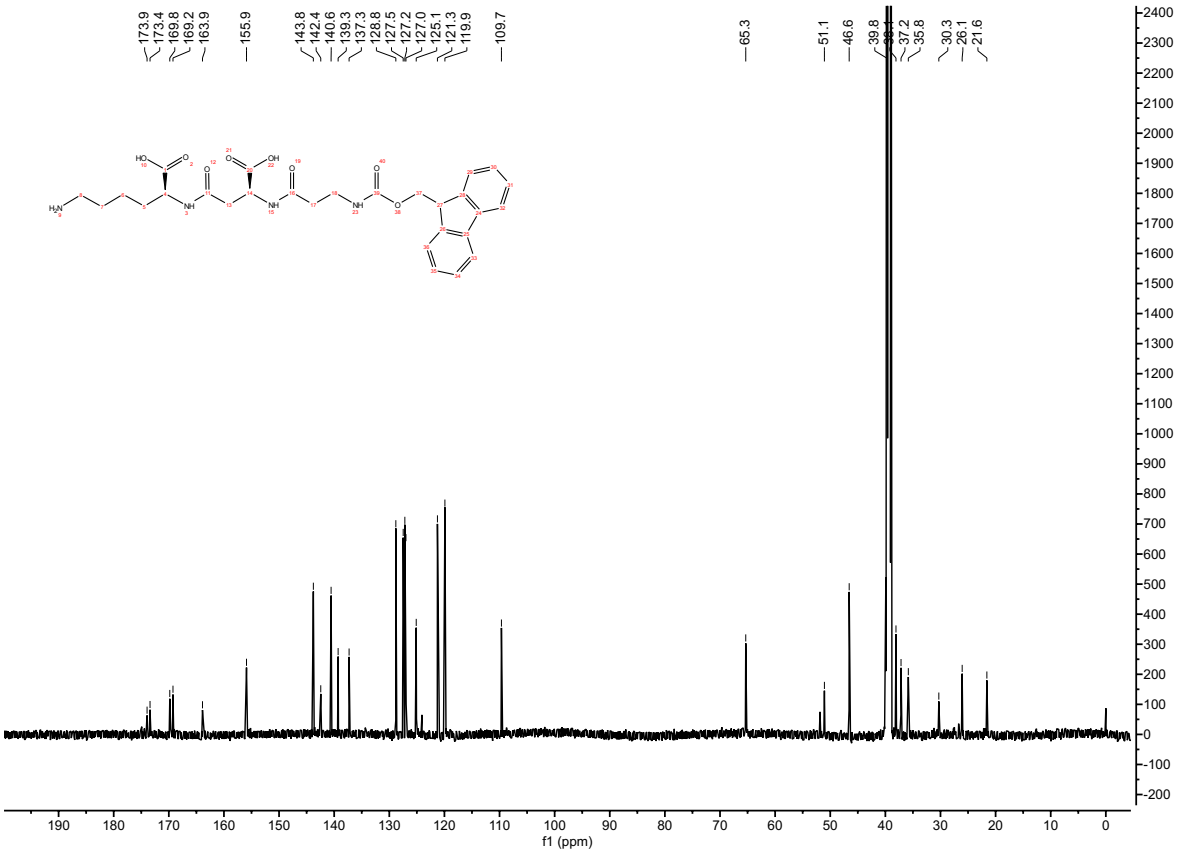
Compound 38



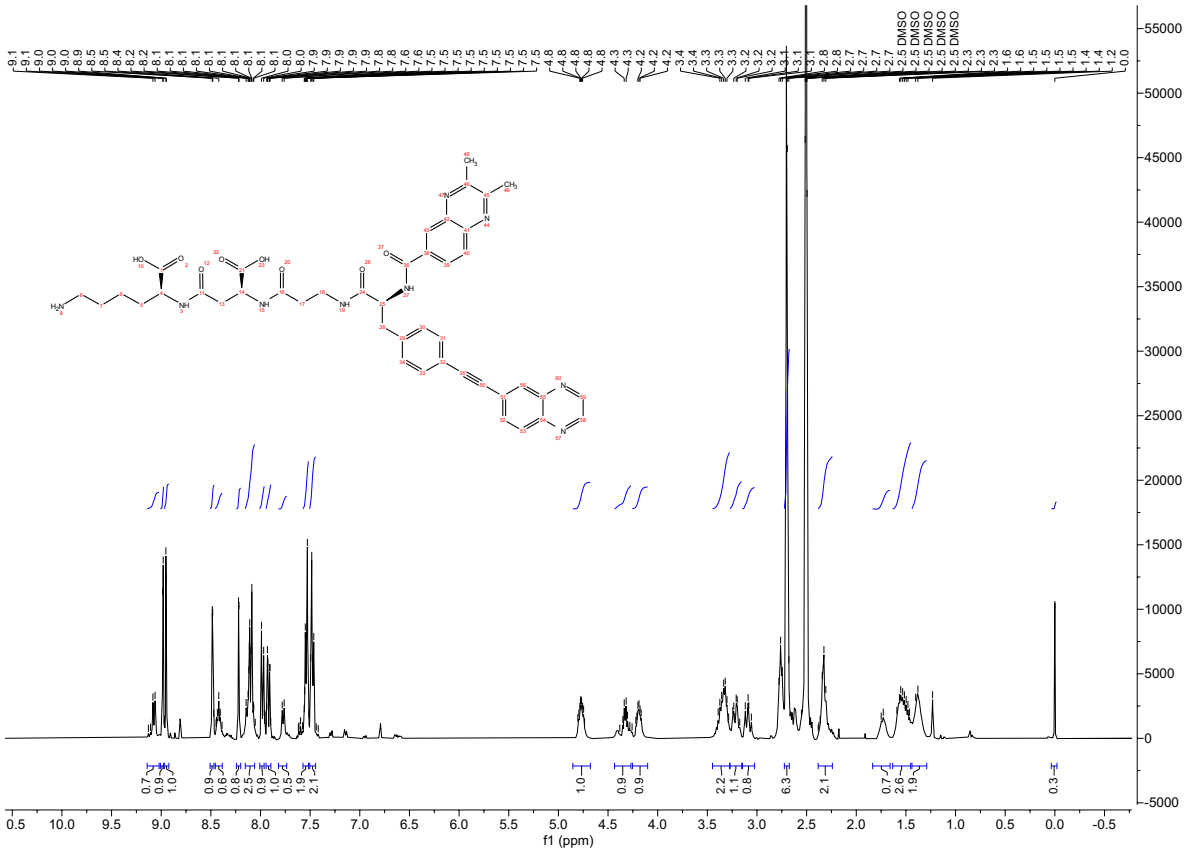


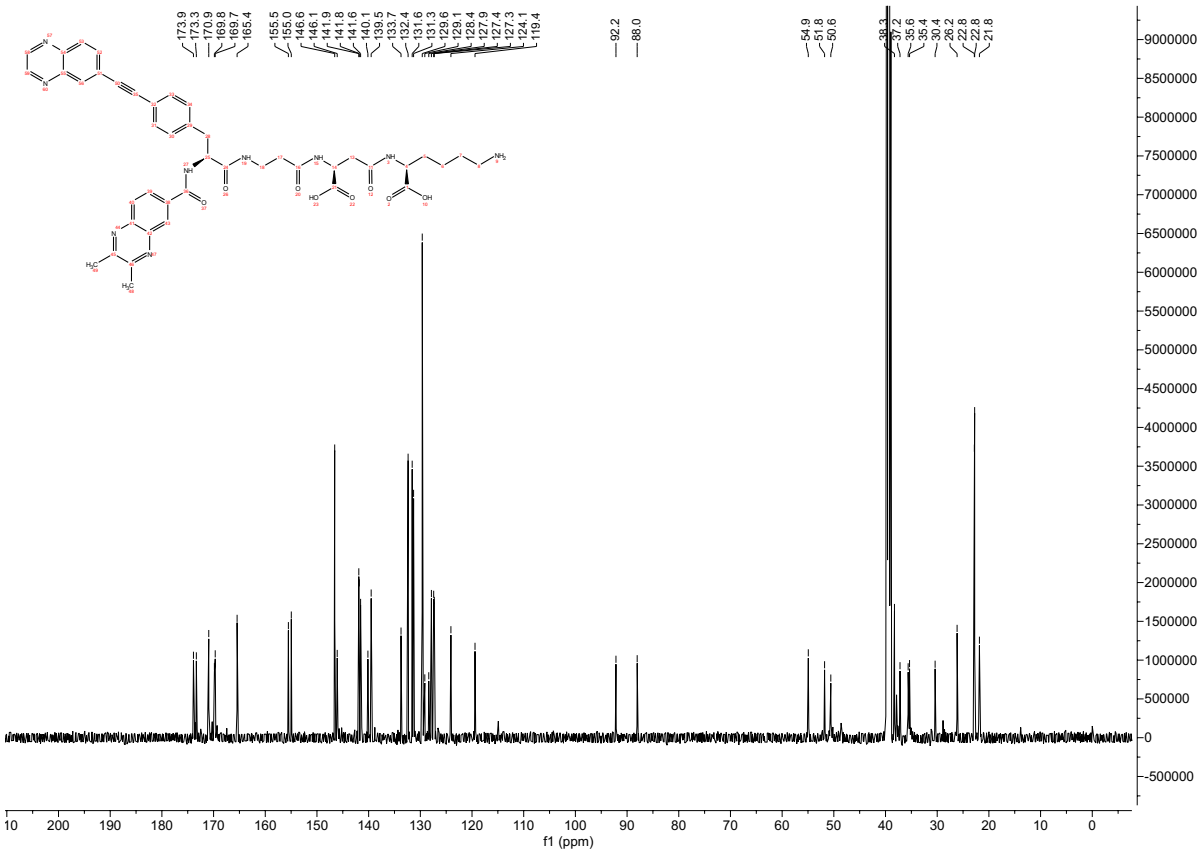
Compound 39



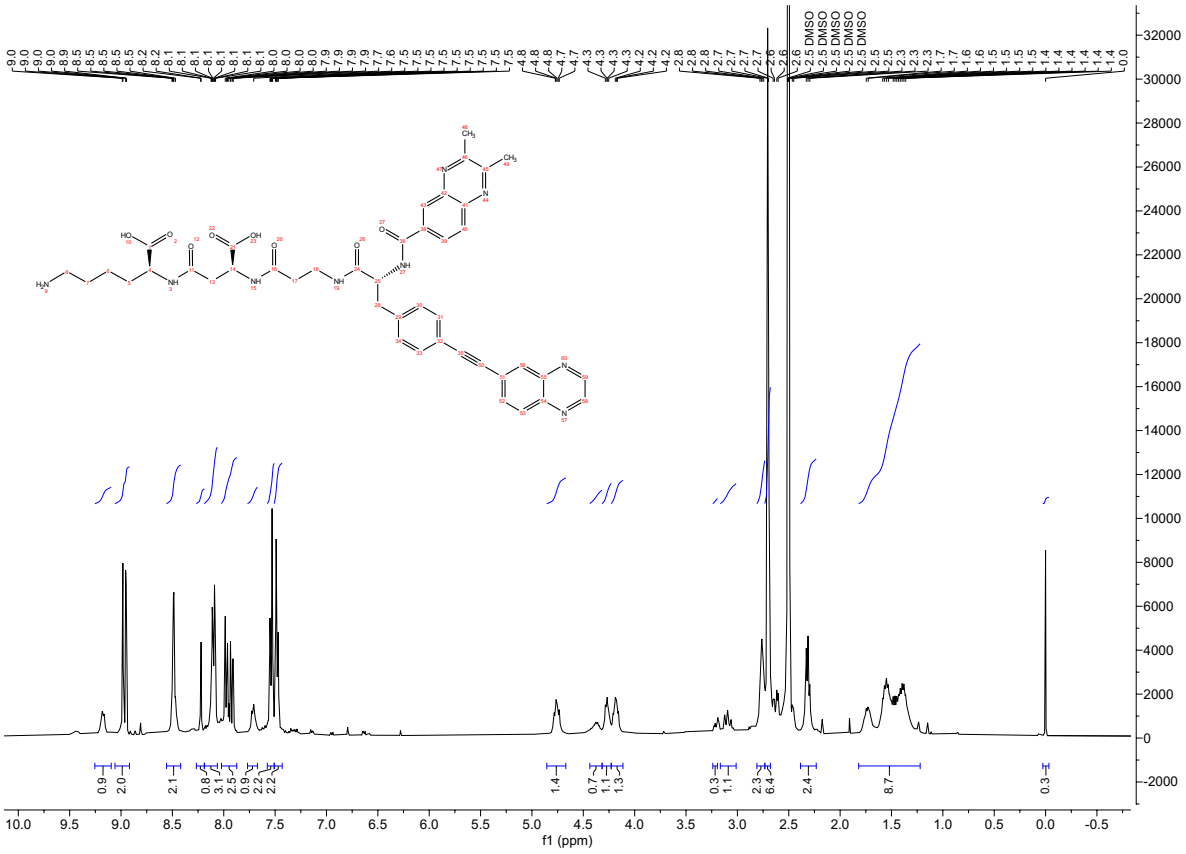


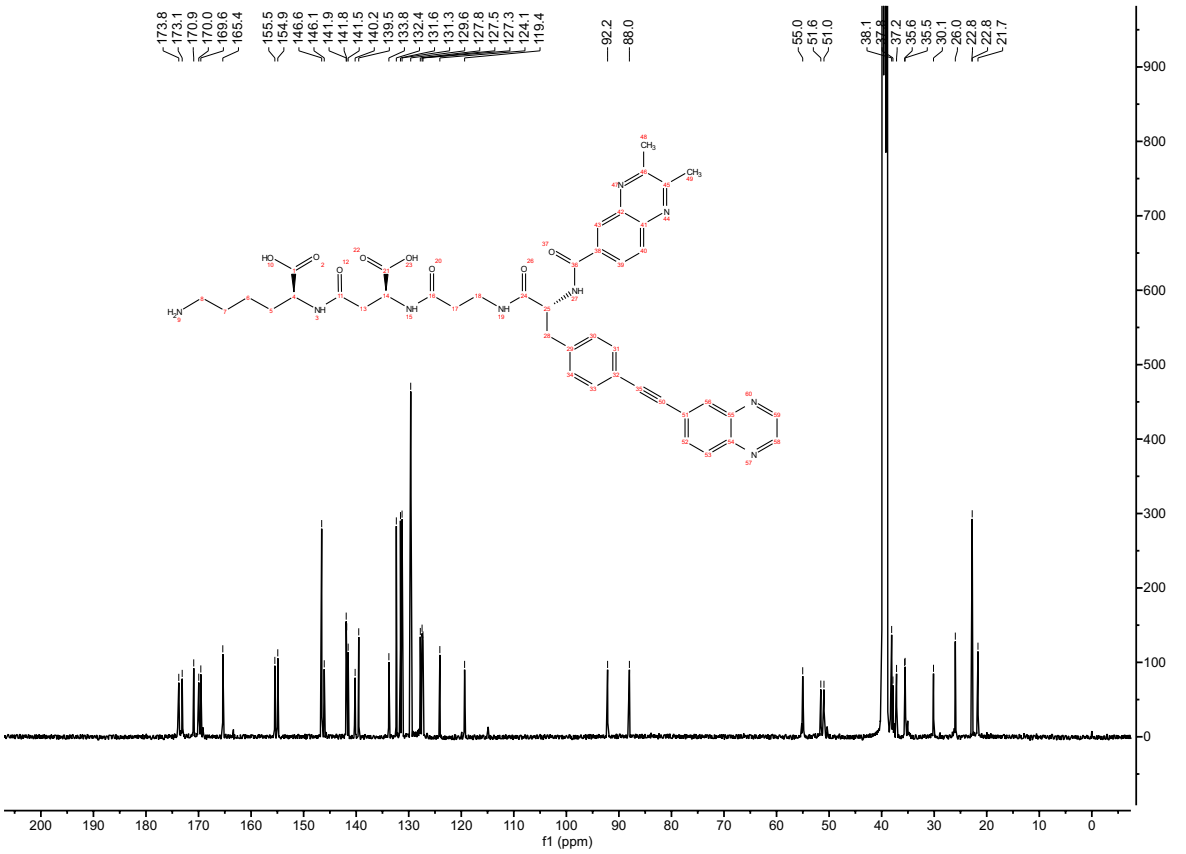
Compound 40



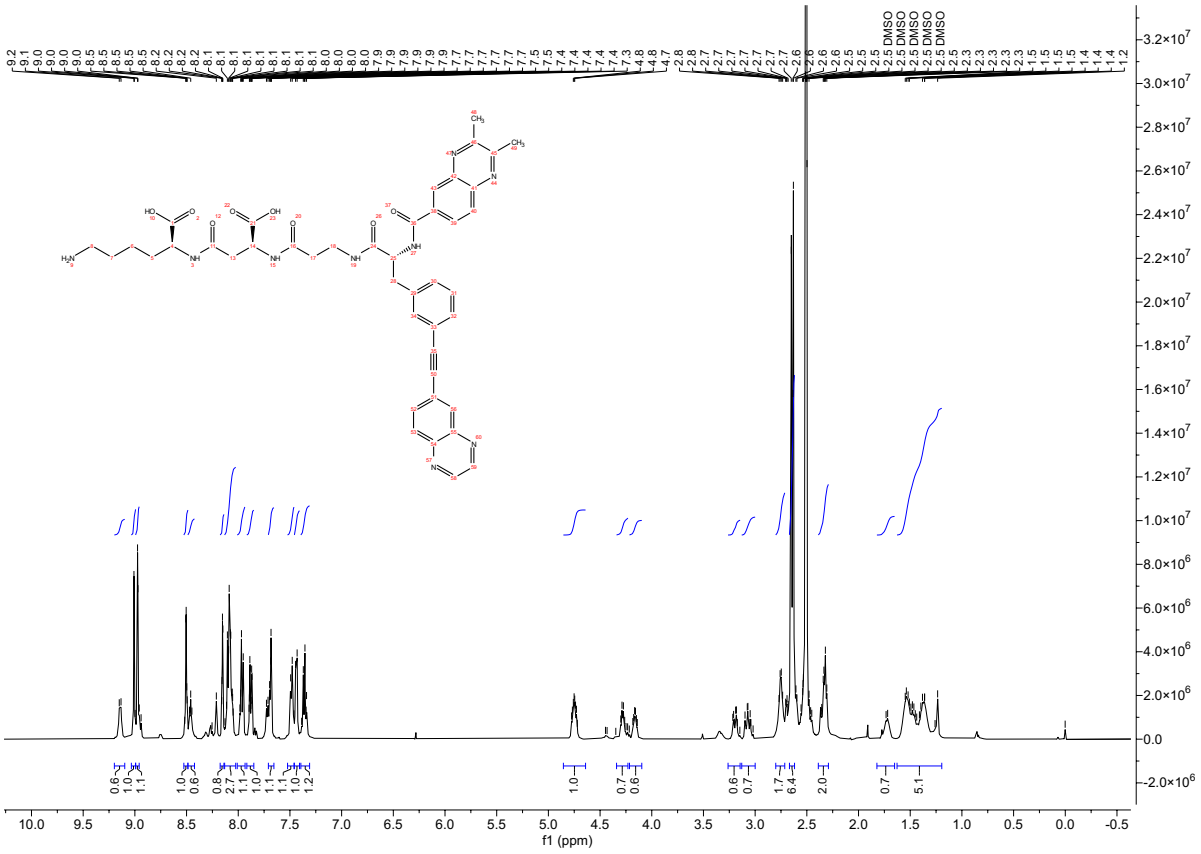


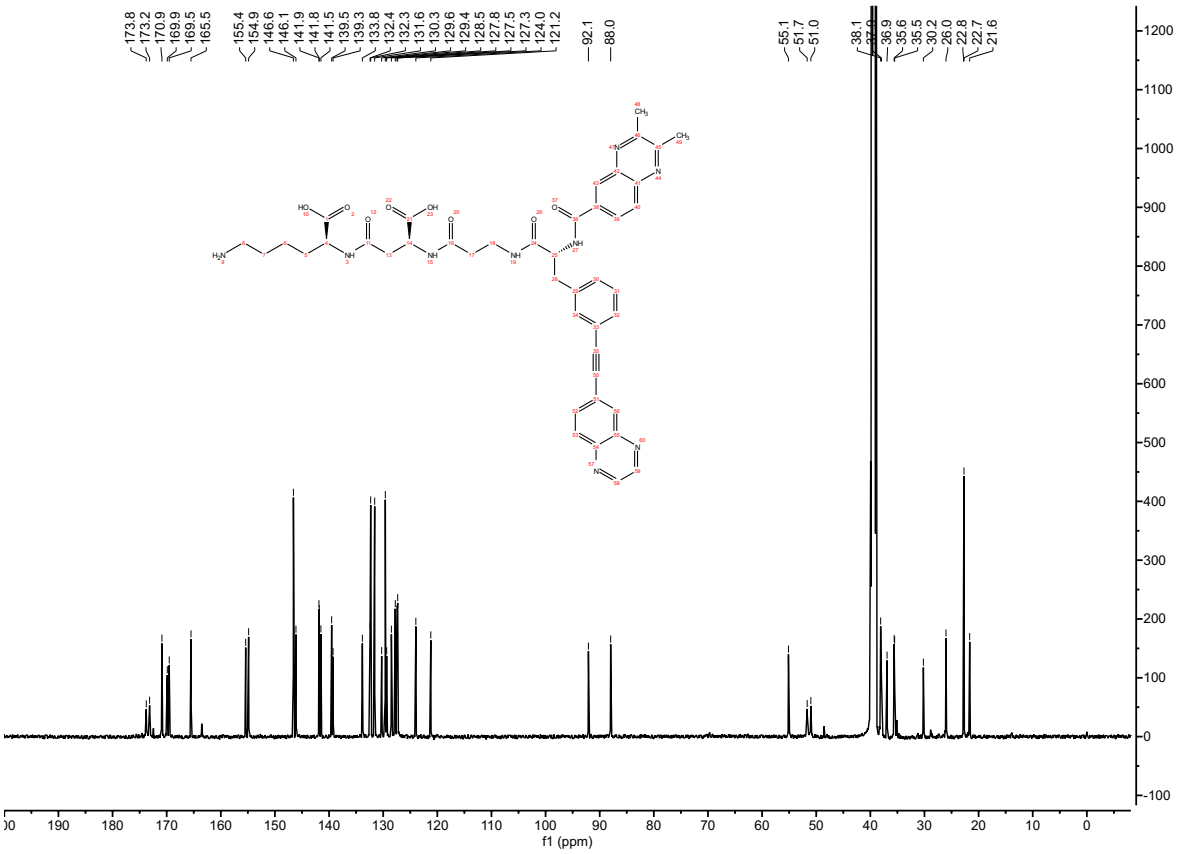
Compound 41



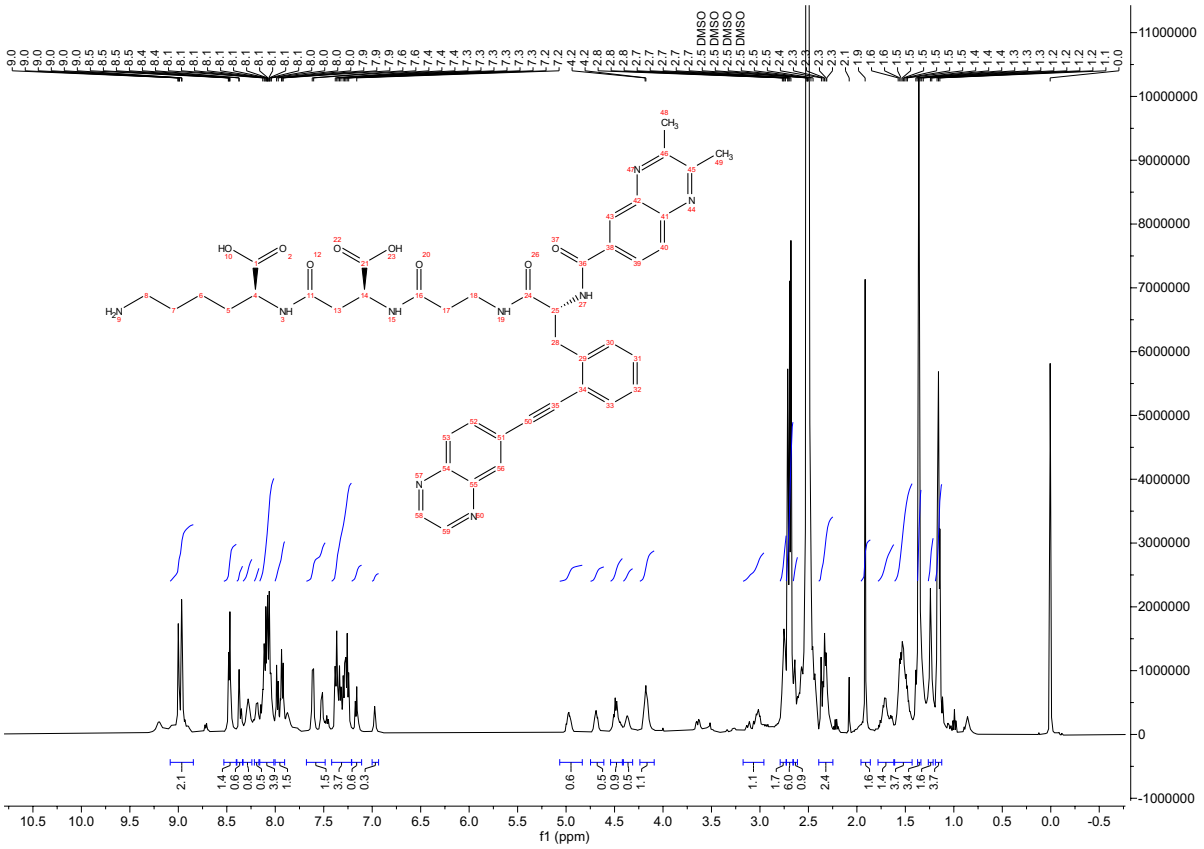


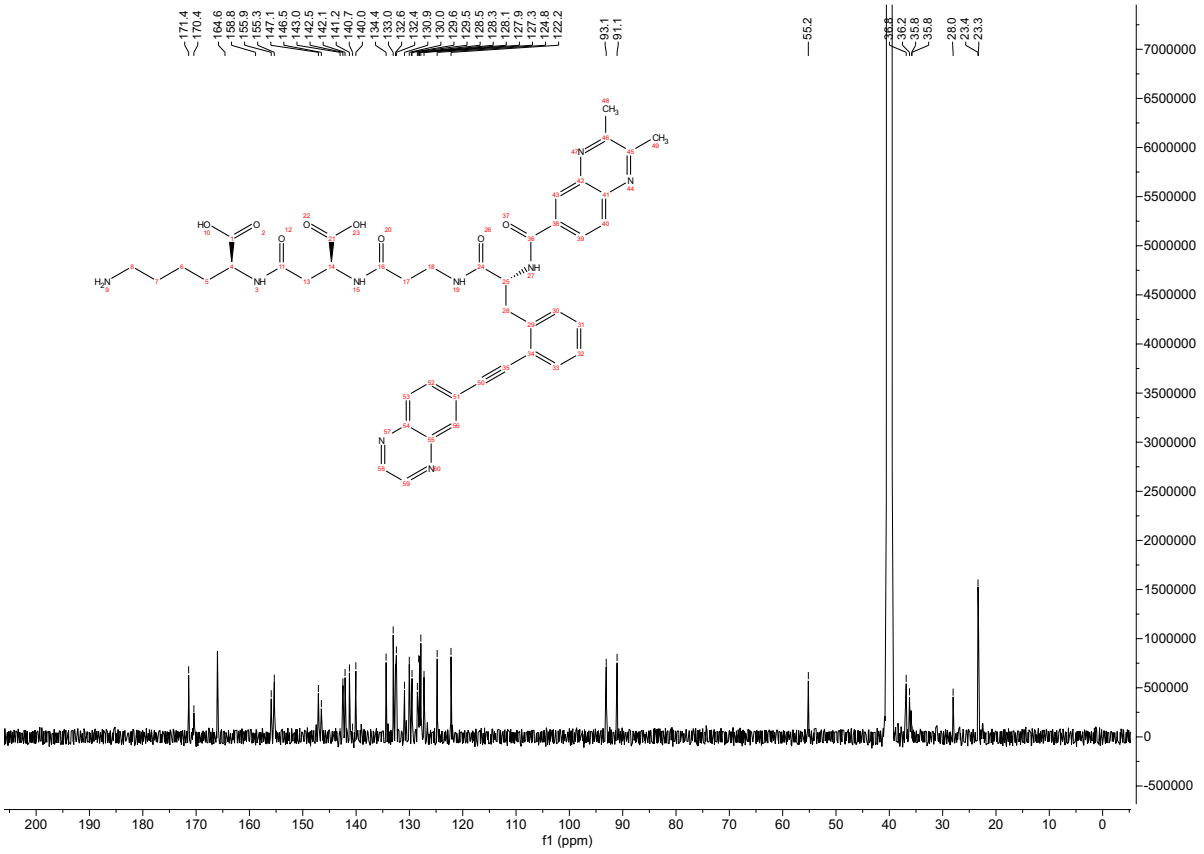
Compound 42



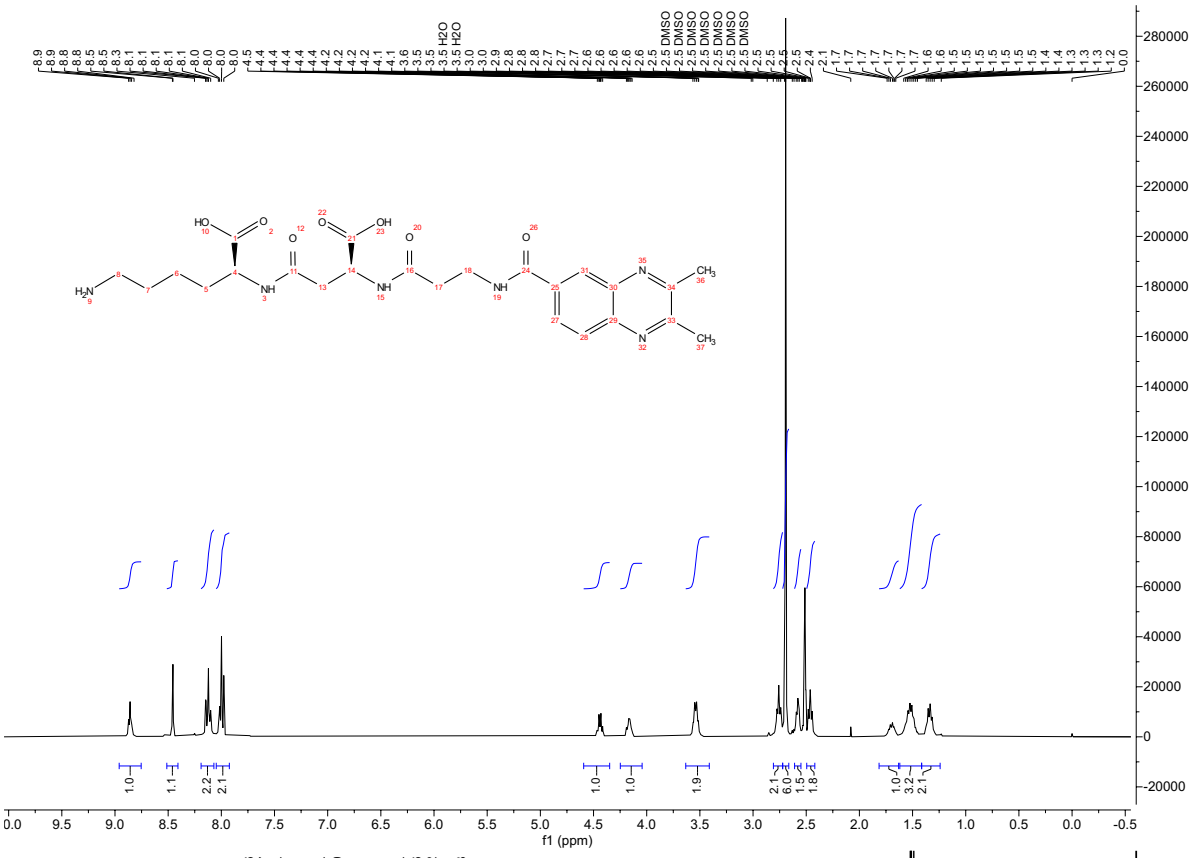


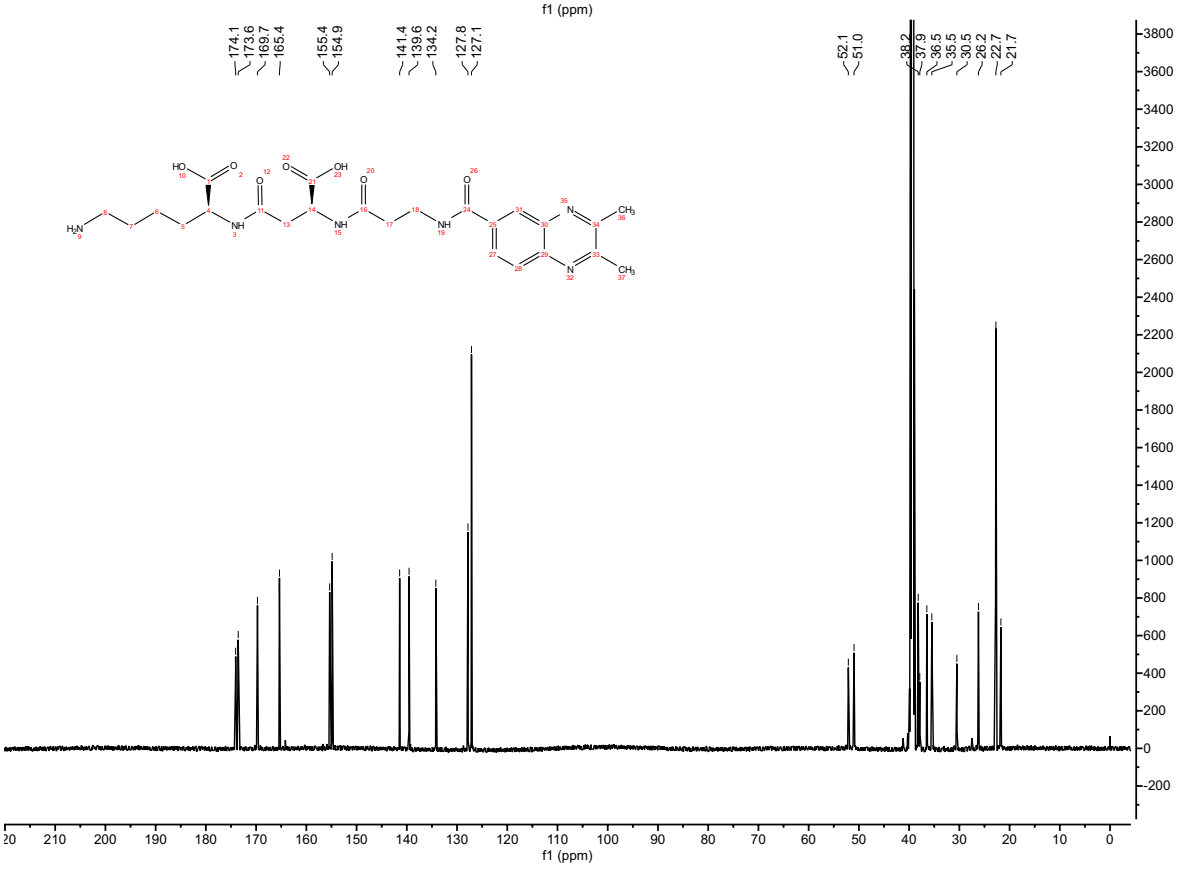
Compound 43



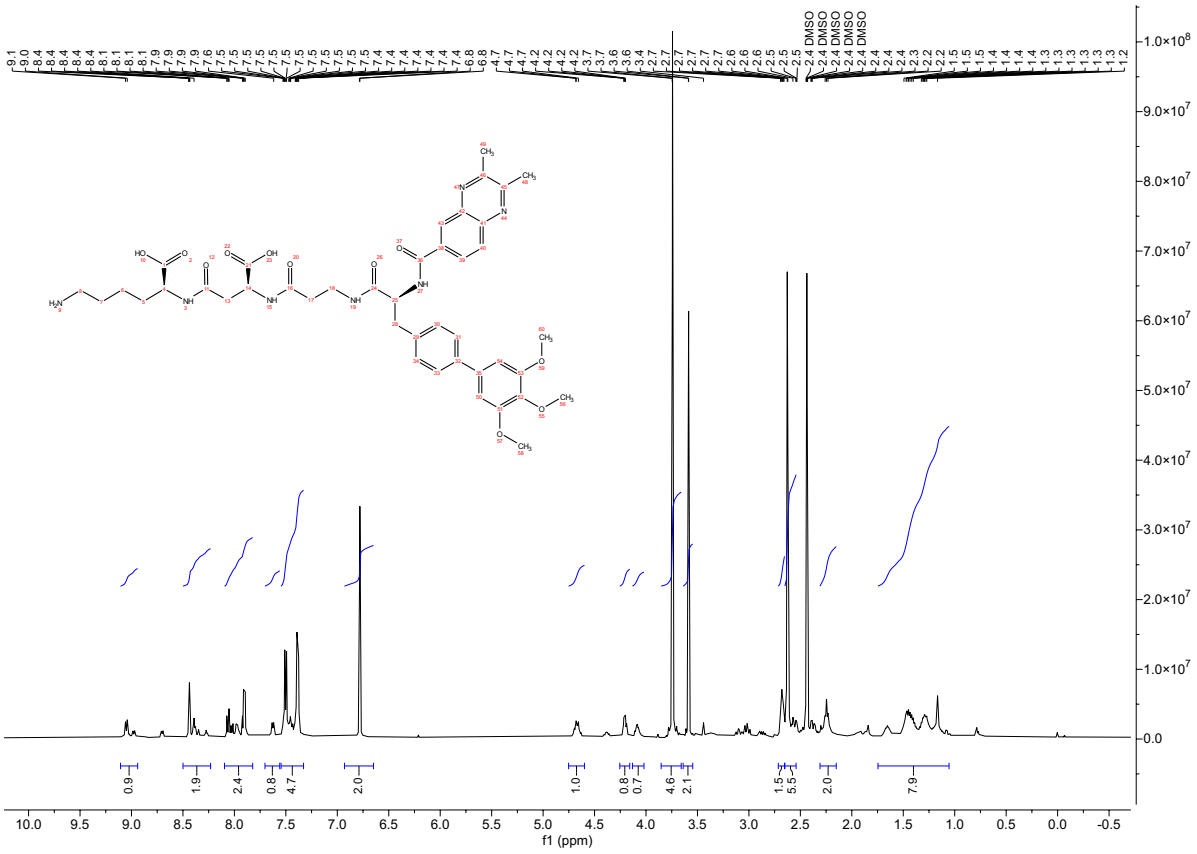


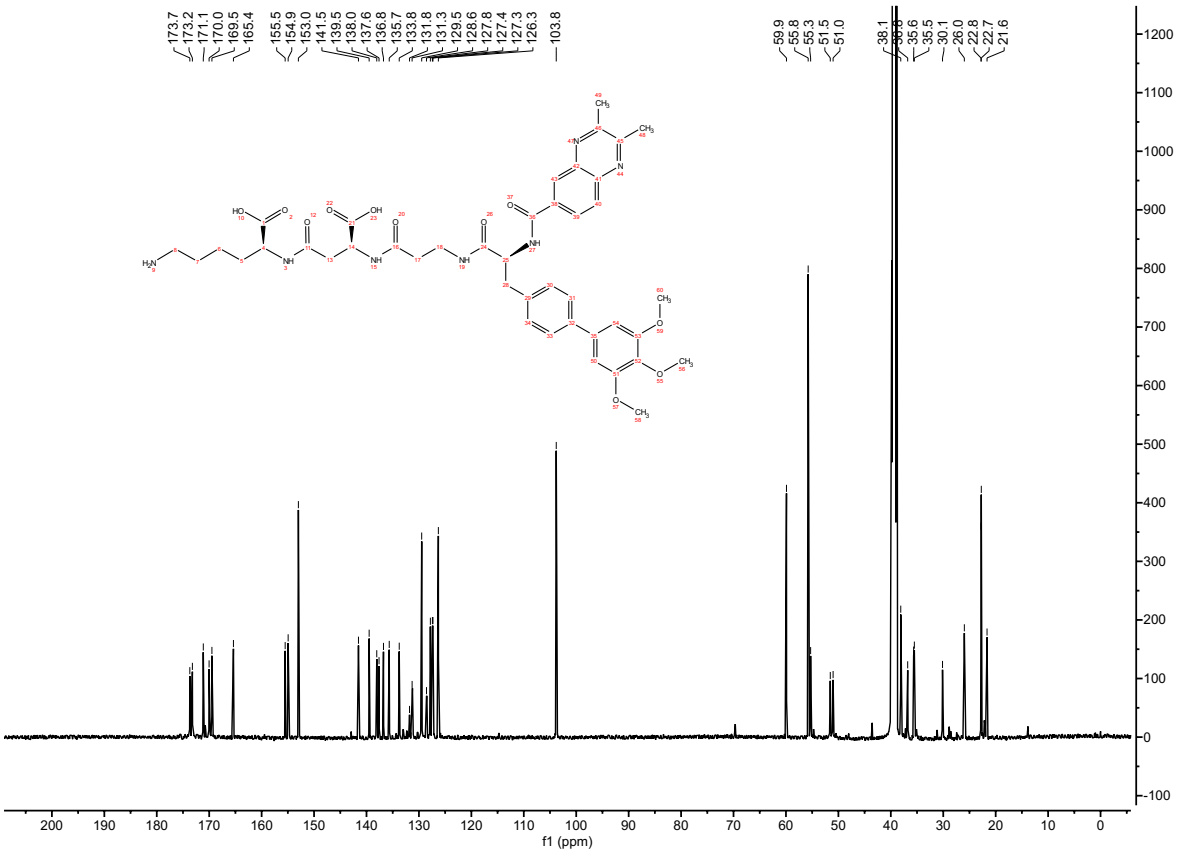
Compound 44


f1 (ppm)

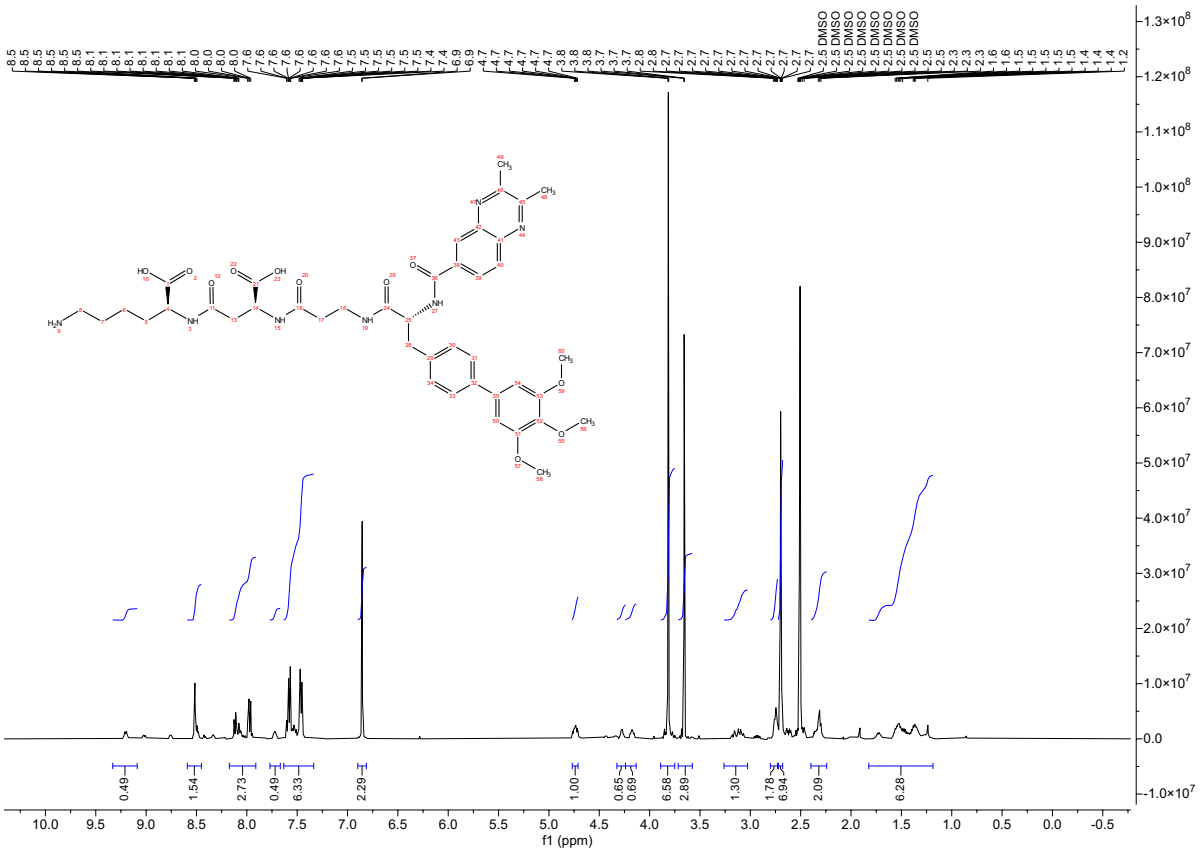

174.1
173.6
169.7
165.4
155.4
154.9
141.4
139.6
134.2
127.8
127.1
52.1
51.0
38.2
37.9
36.5
35.5
30.5
26.2
22.7
21.7
f1 (ppm)

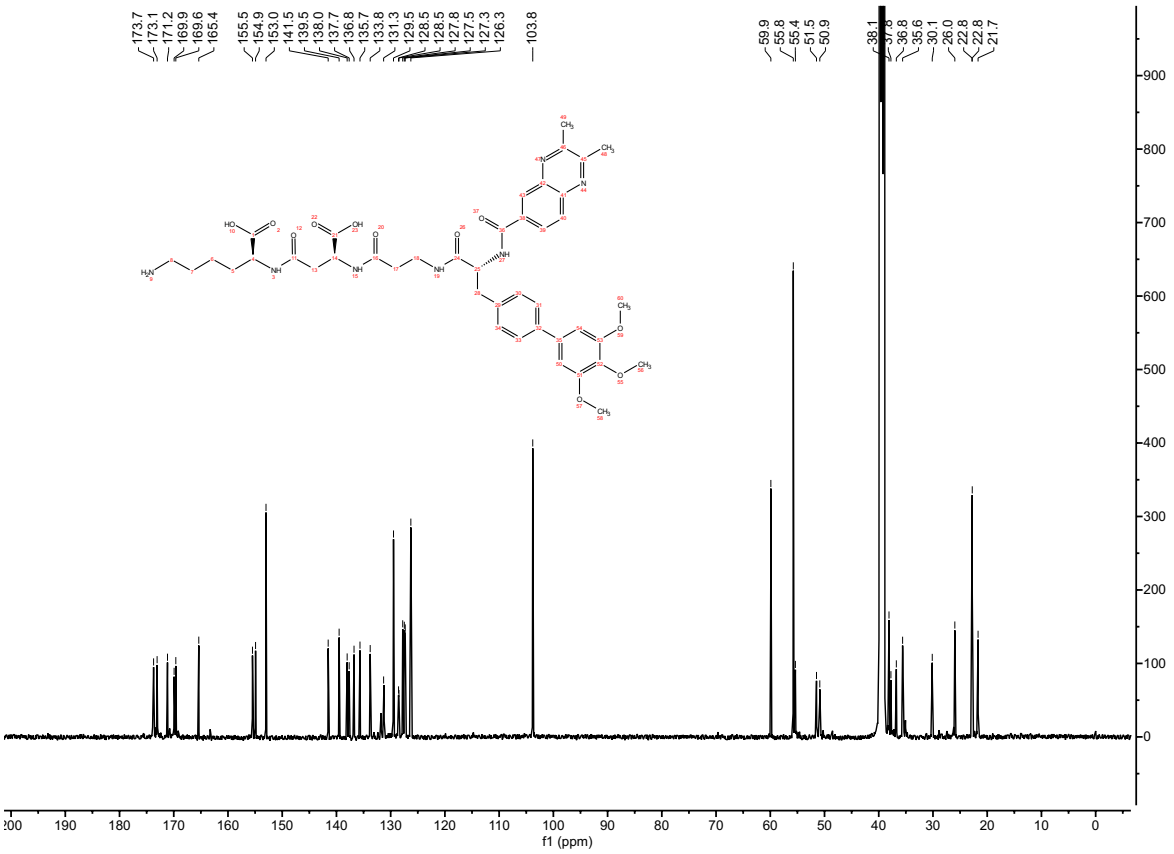
Compound 45



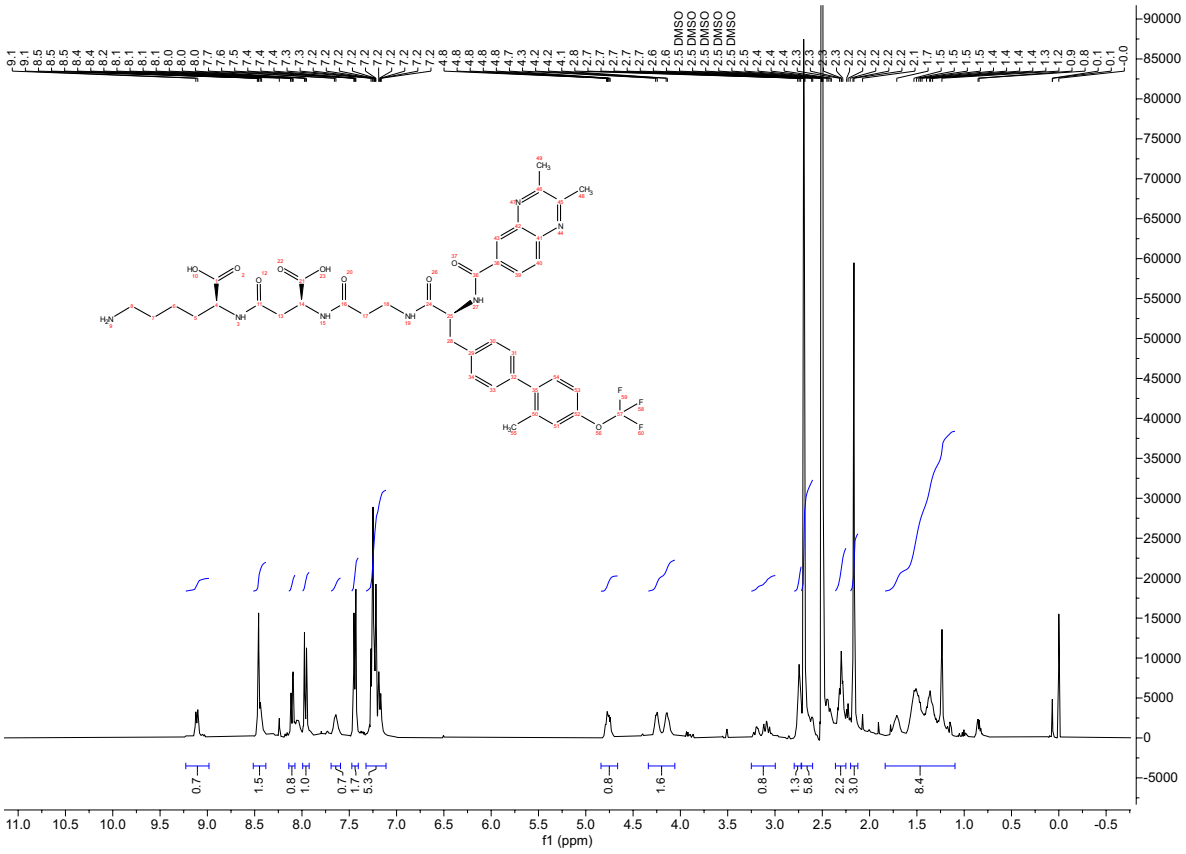


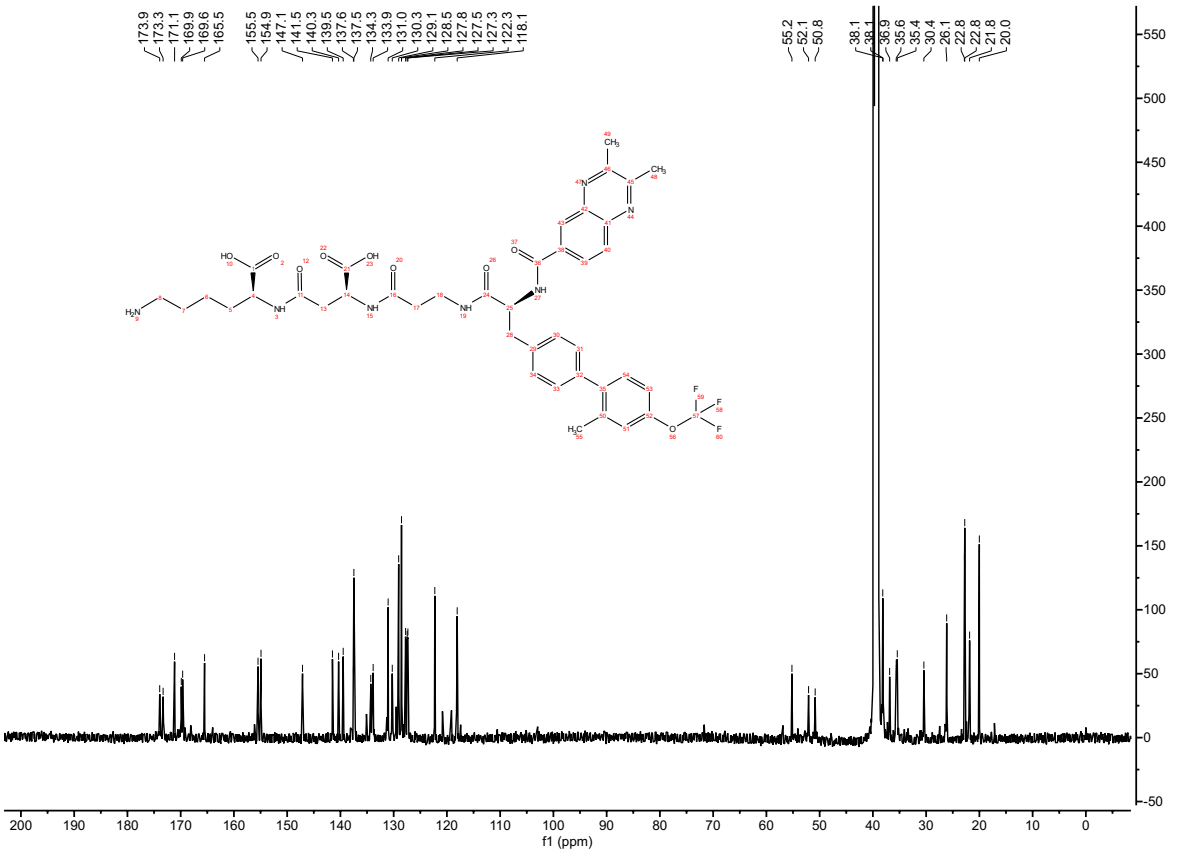
Compound 46



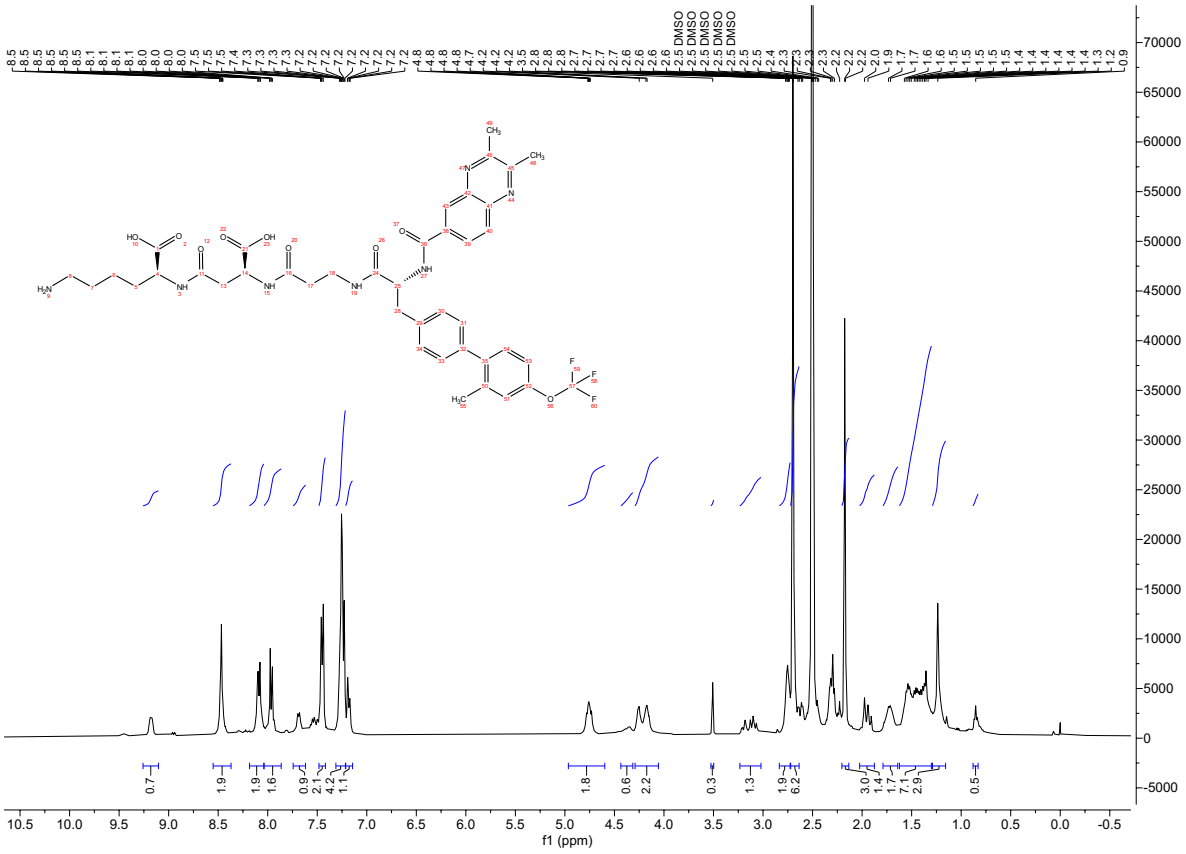


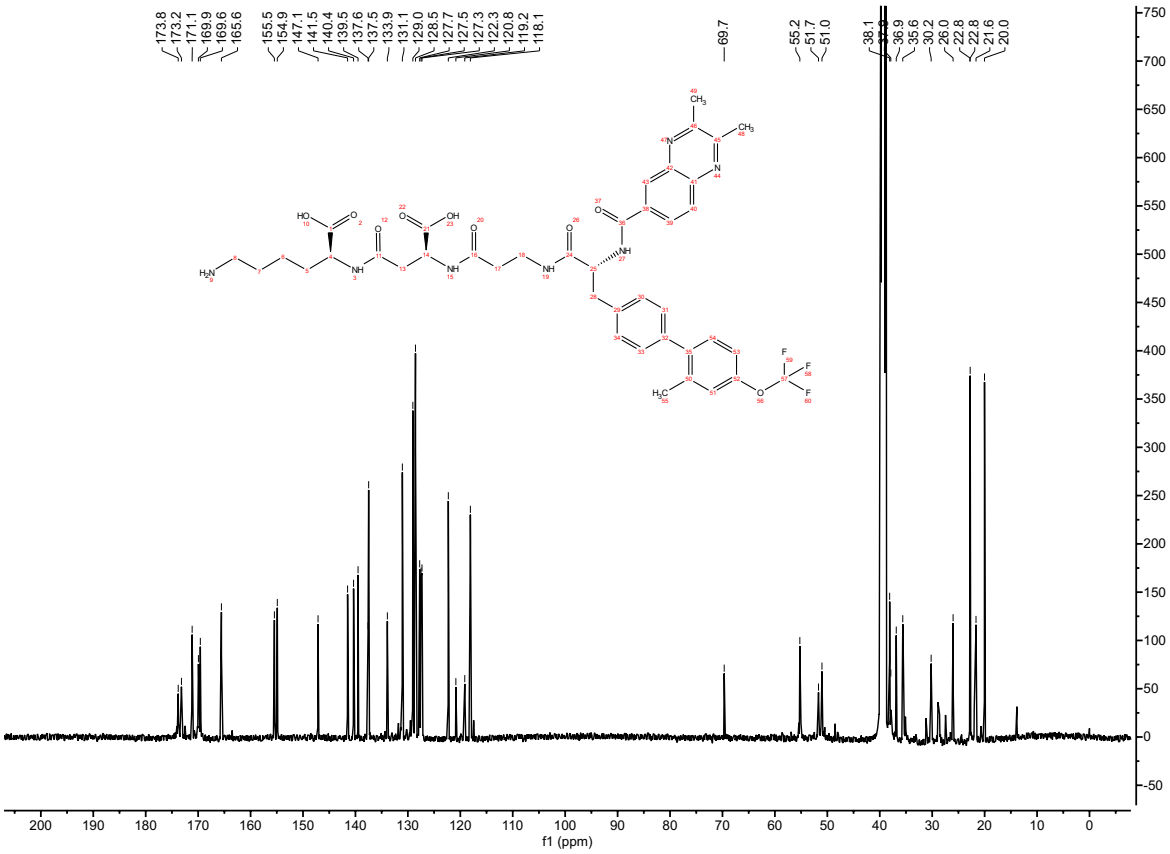
Compound 47



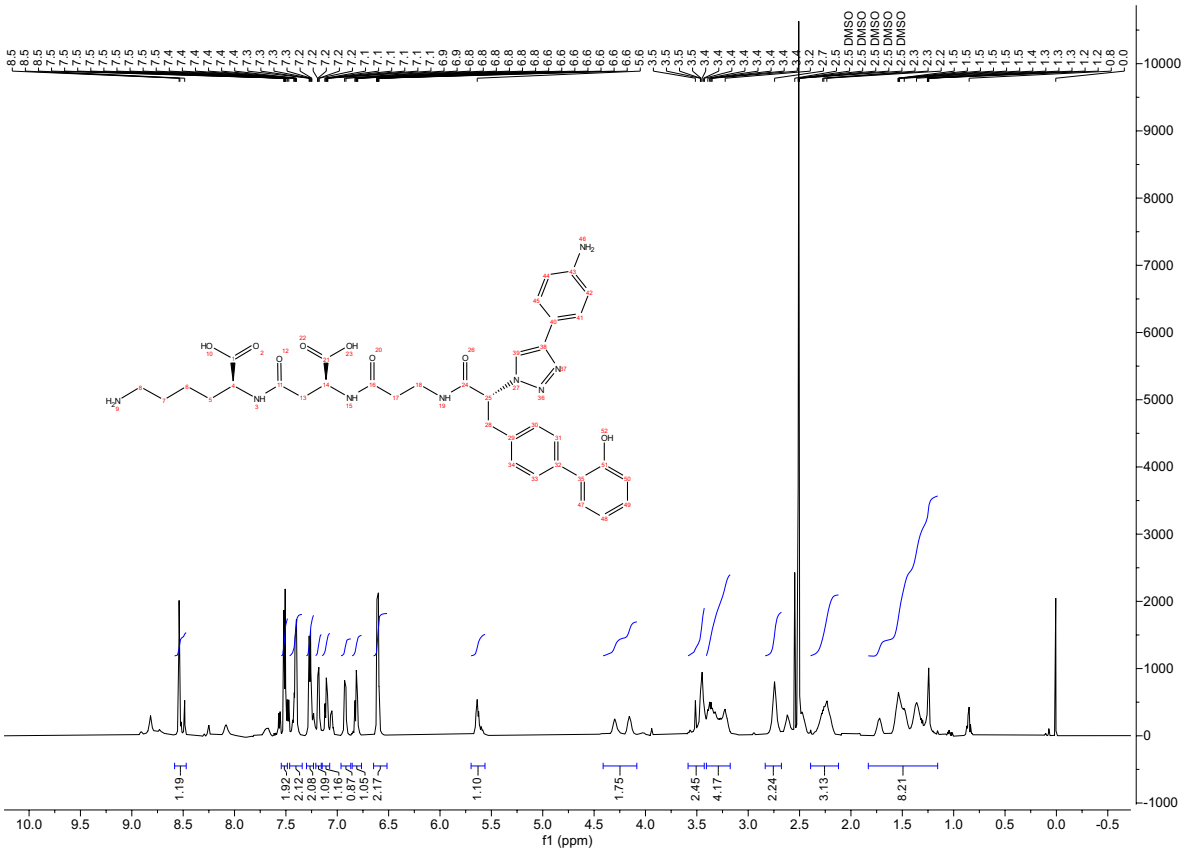


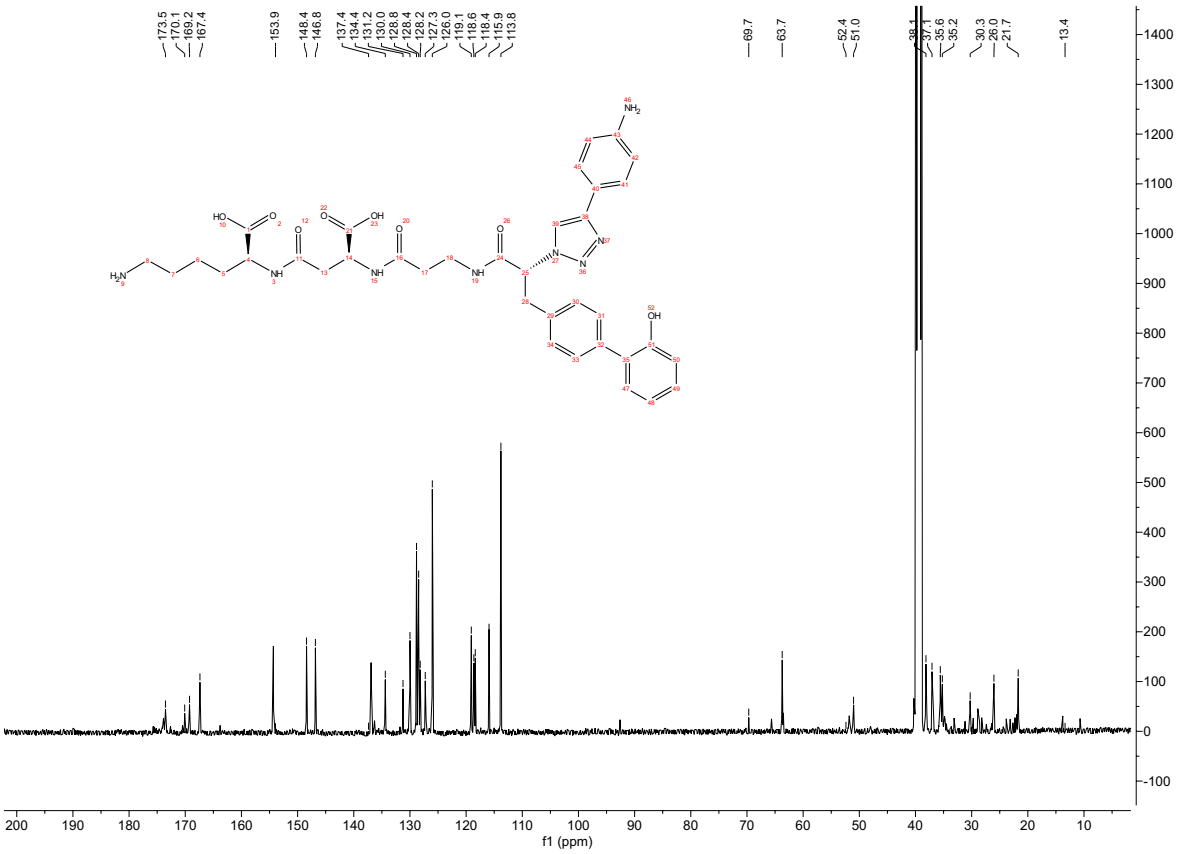
Compound 48



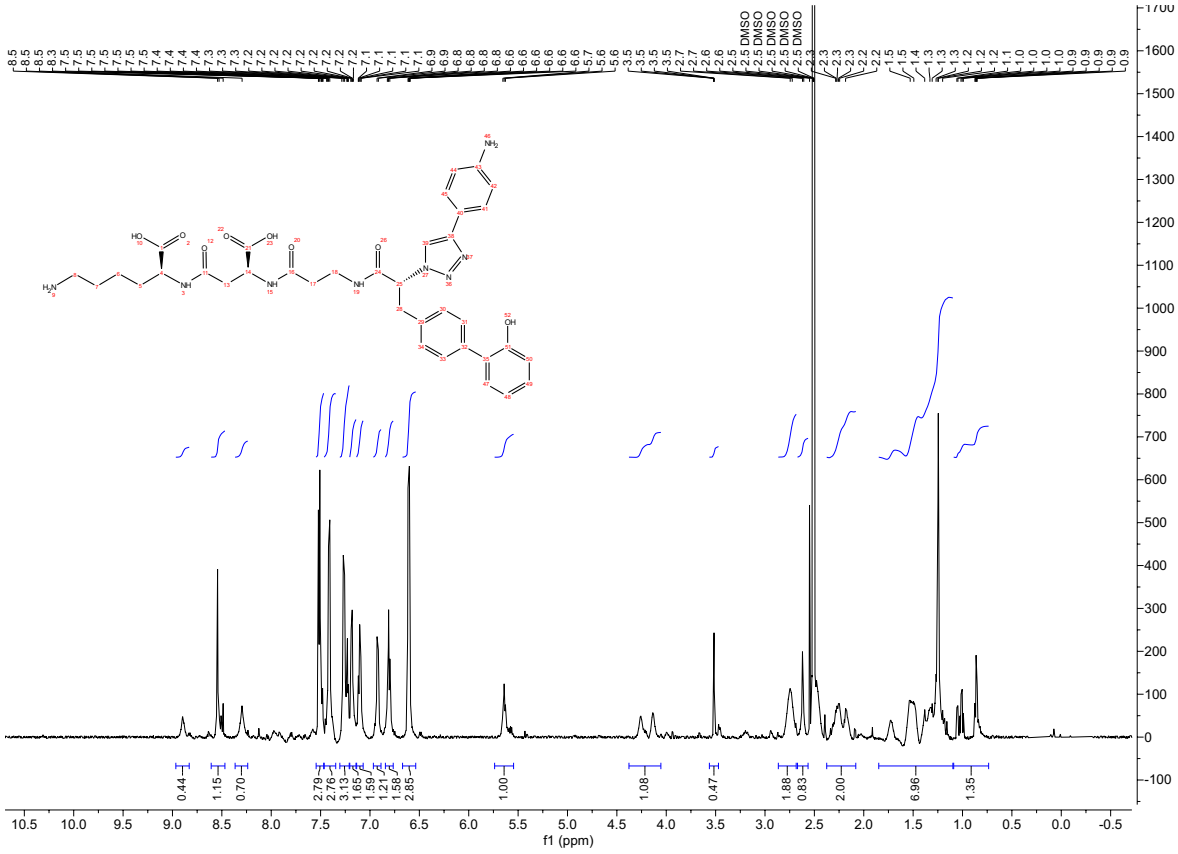


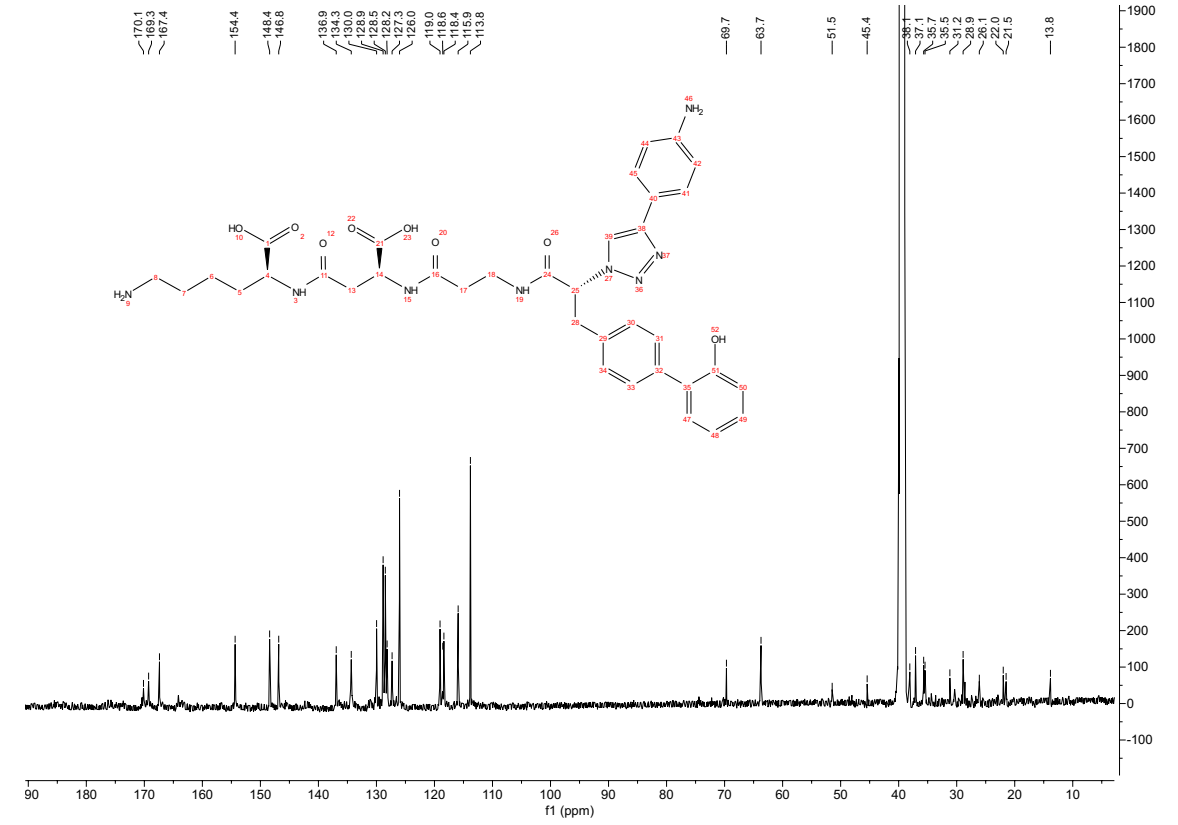
Compound 51



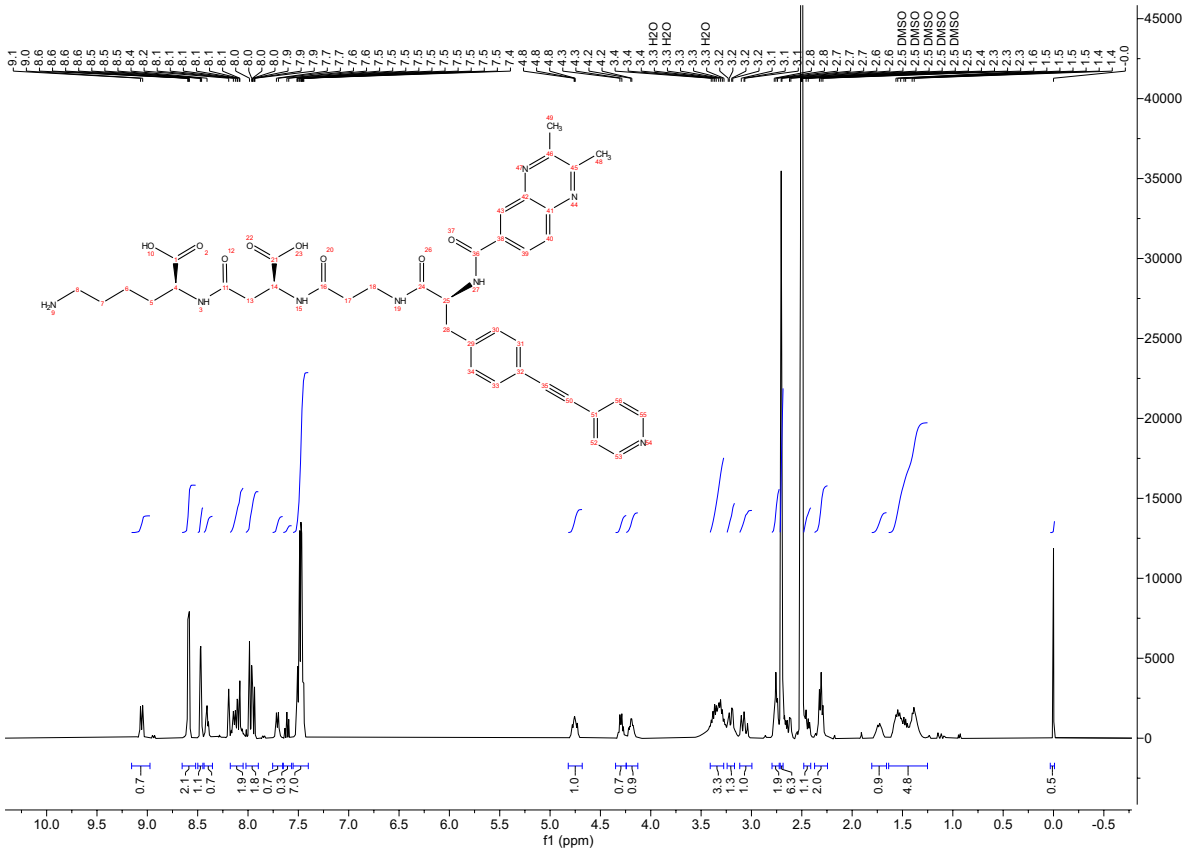


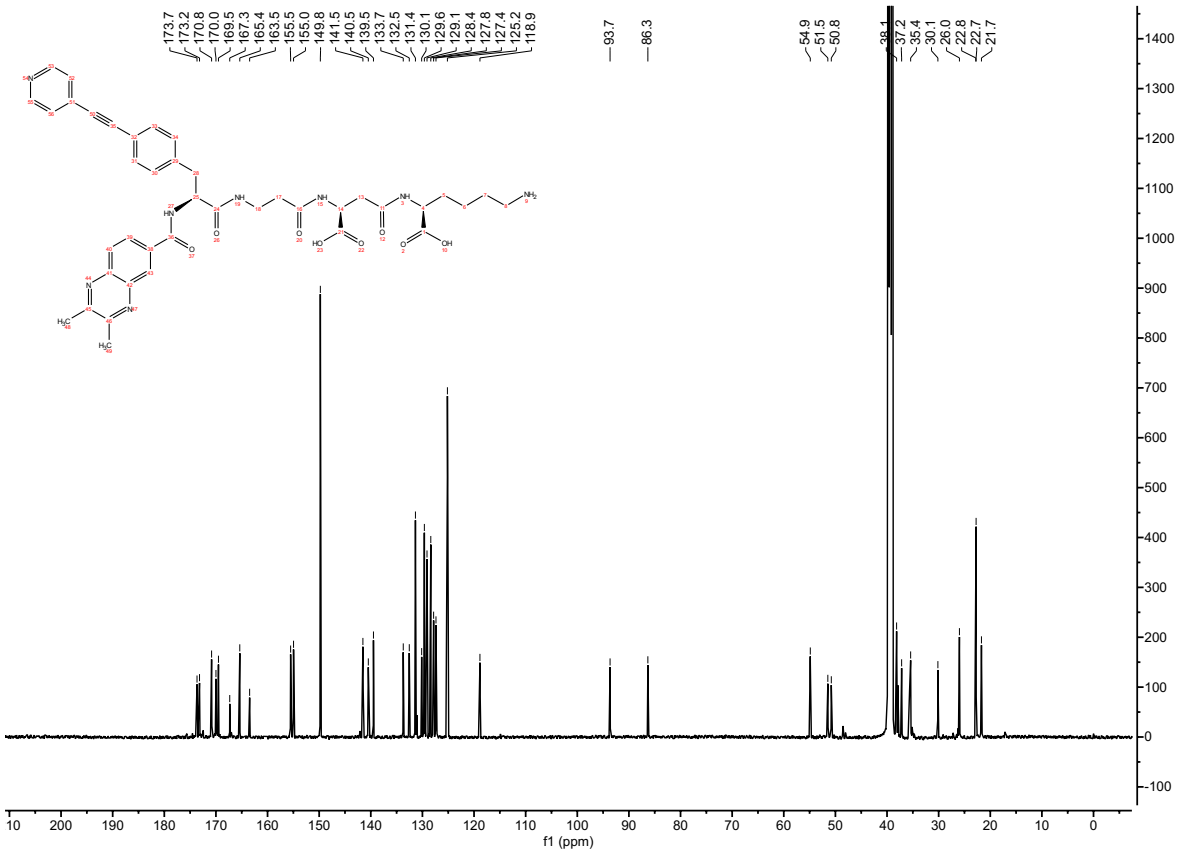
Compound 52


2.5 DMSO
0.44
1.15
0.70
2.79
2.76
3.13
1.65
1.59
1.21
1.58
2.85
1.00
1.08
0.47
1.88
0.83
2.00
6.96
1.35
f1 (ppm)

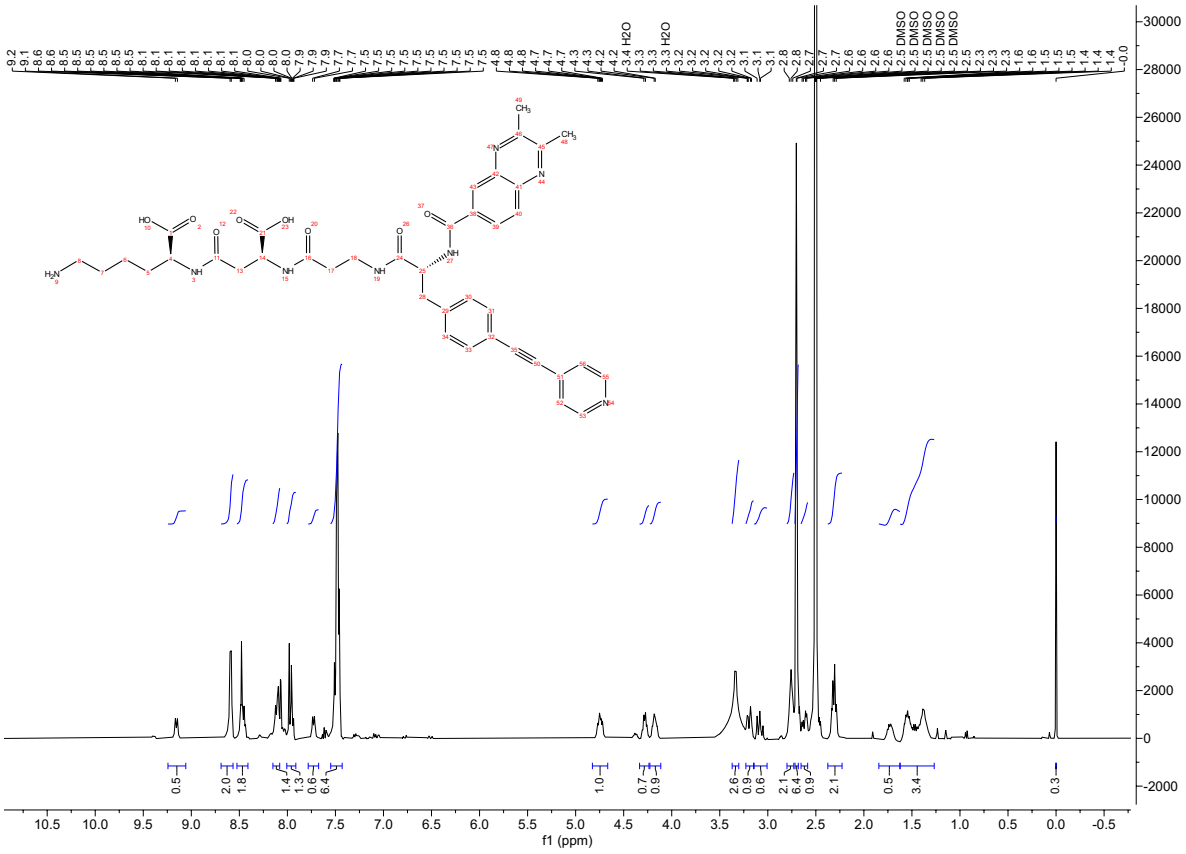

170.1
169.3
167.4
154.4
148.4
146.8
136.9
134.3
130.0
128.9
128.5
128.2
127.3
126.0
119.0
118.6
118.4
115.9
113.8
69.7
63.7
51.5
45.4
39.1
37.1
35.7
35.5
31.2
28.9
26.1
22.0
21.5
13.8
f1 (ppm)

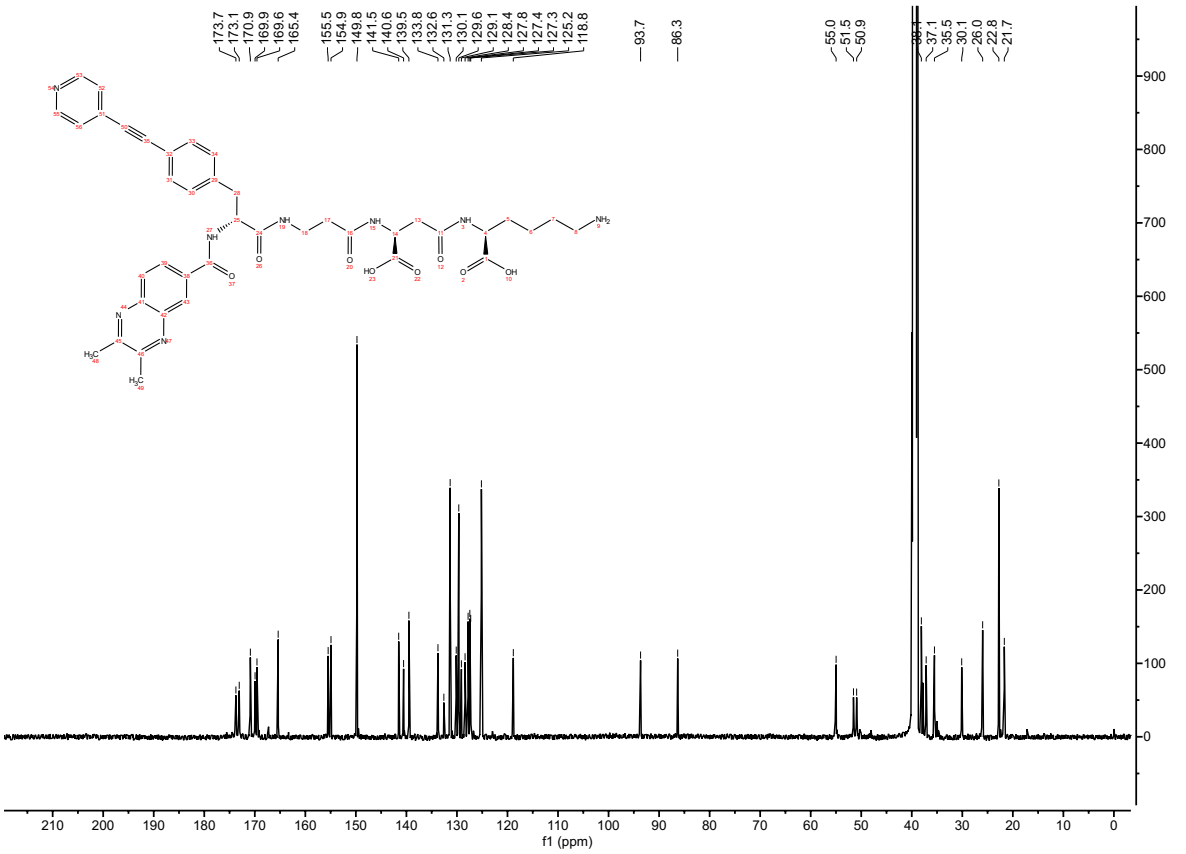
Compound 53





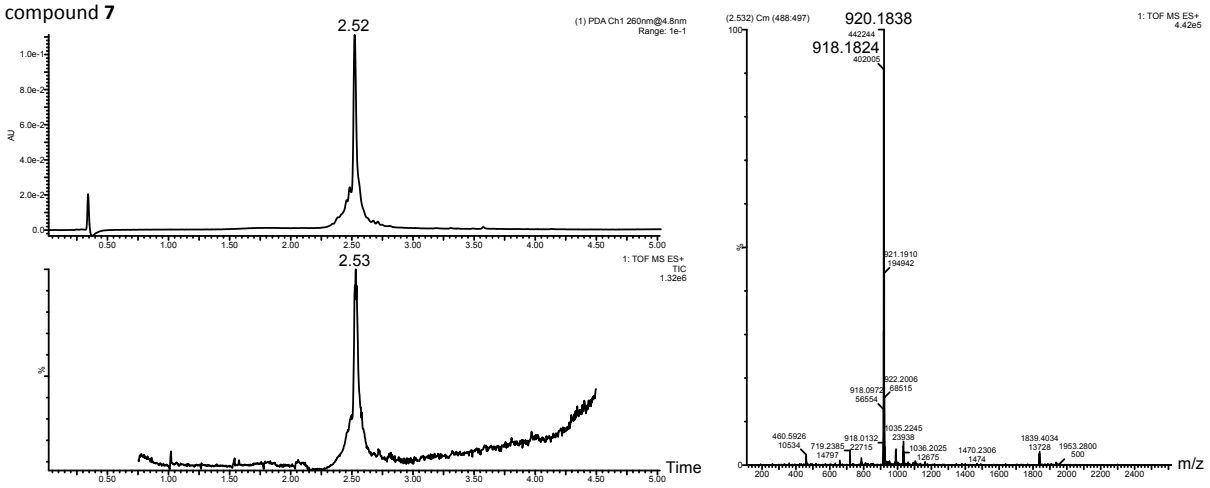
Compound 54



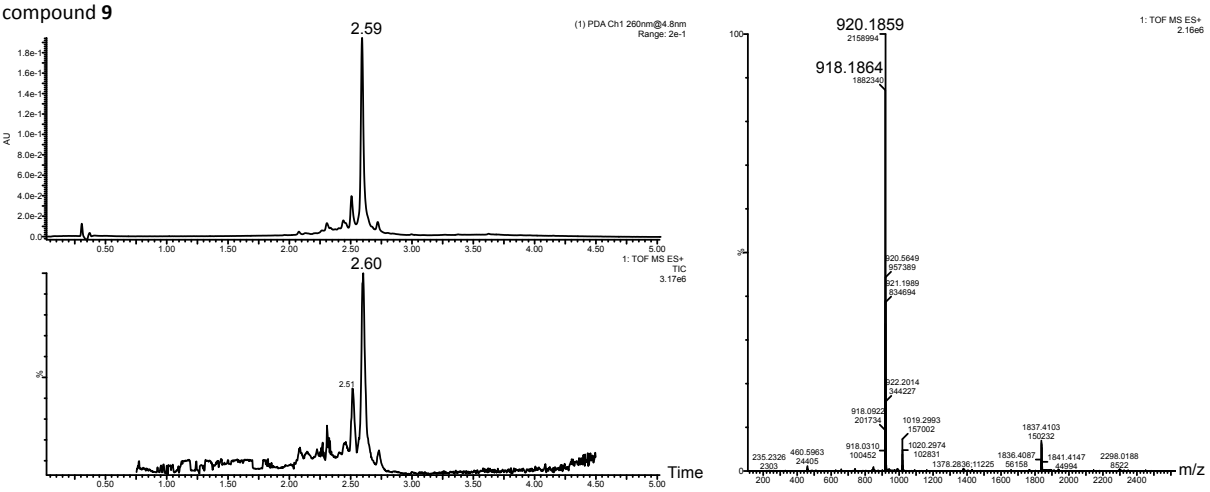


8.4 LC-MS

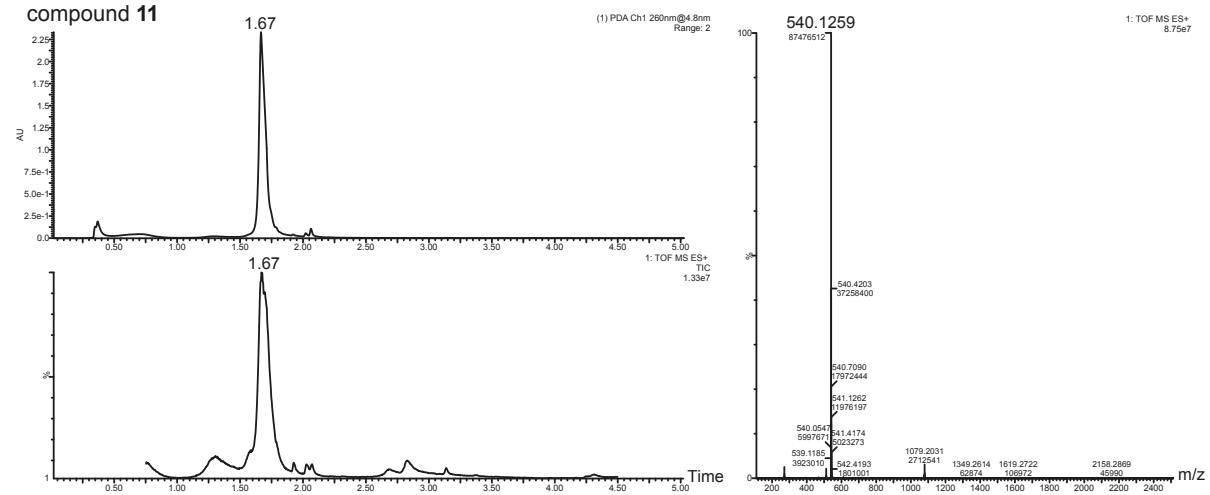
compound **7**



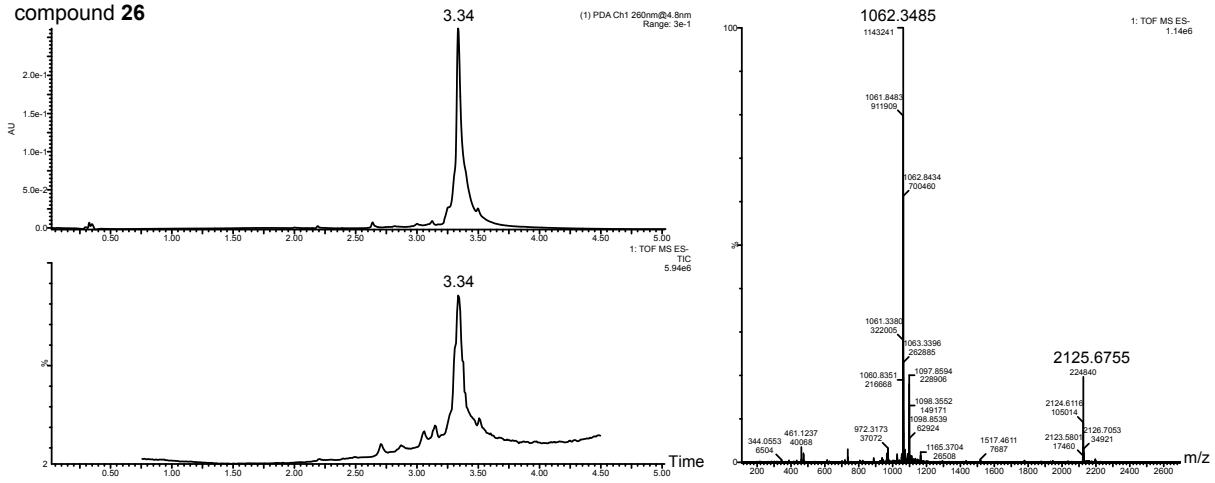
compound **9**



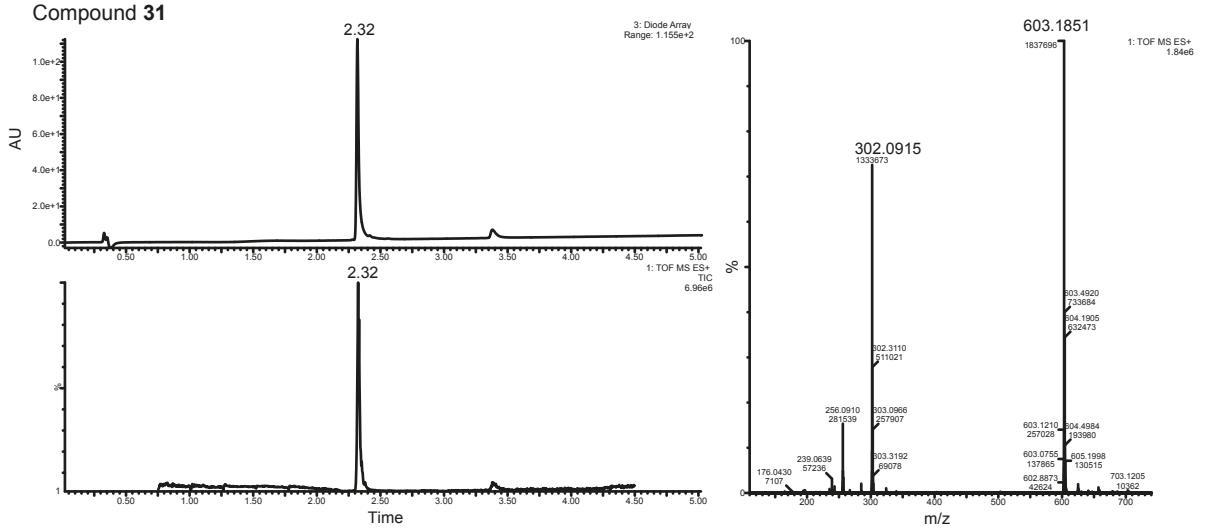
compound **11**



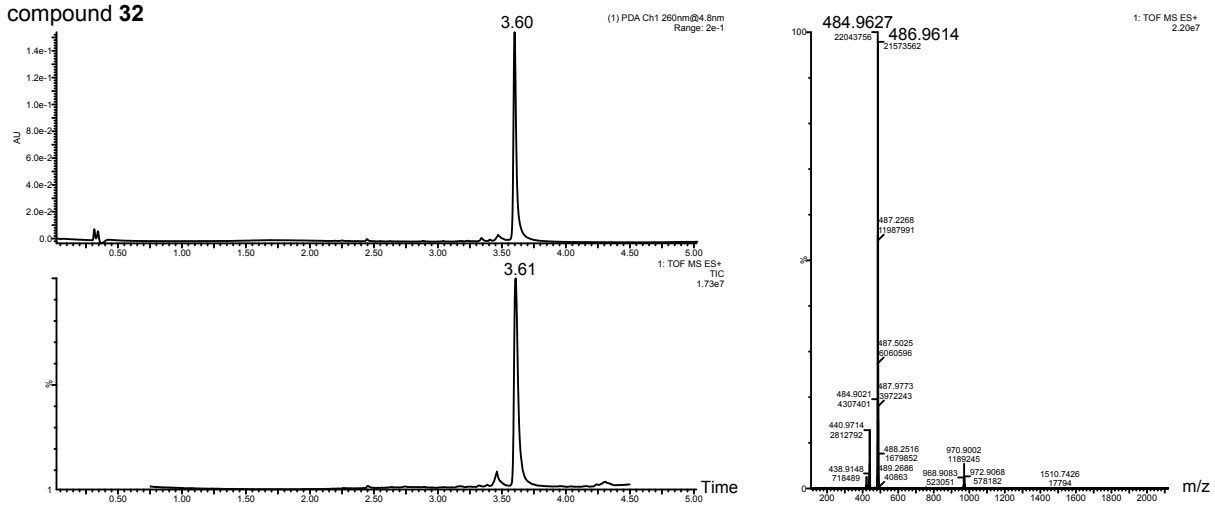
compound **26**



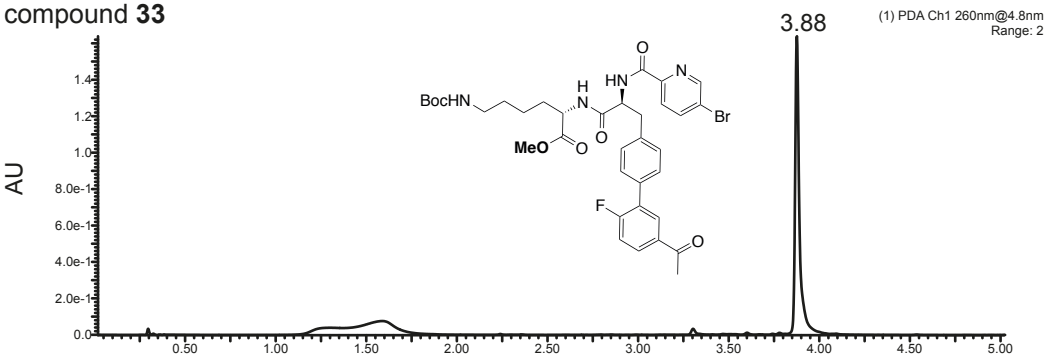
Compound **31**



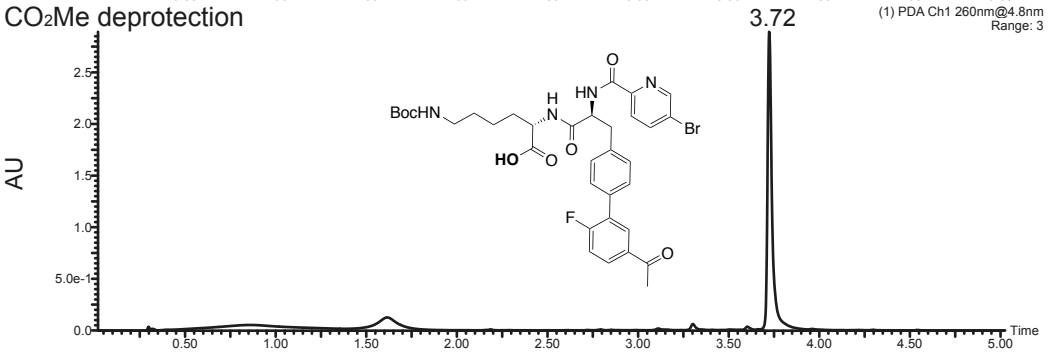
compound **32**



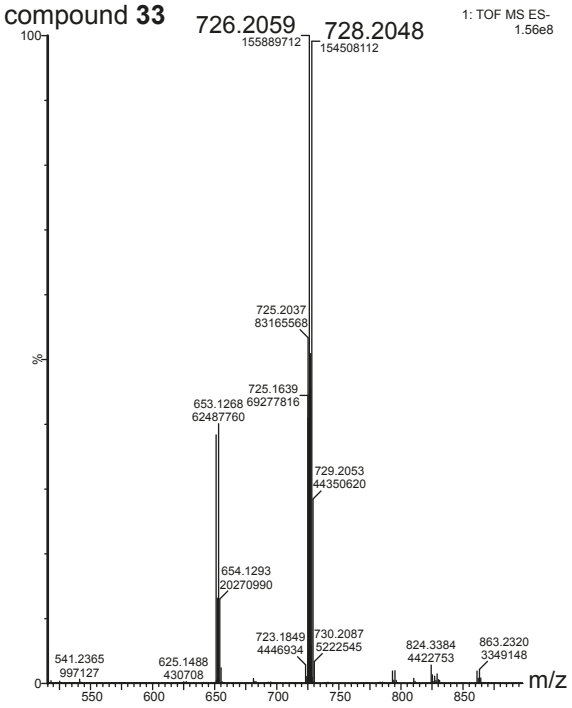
compound **33**



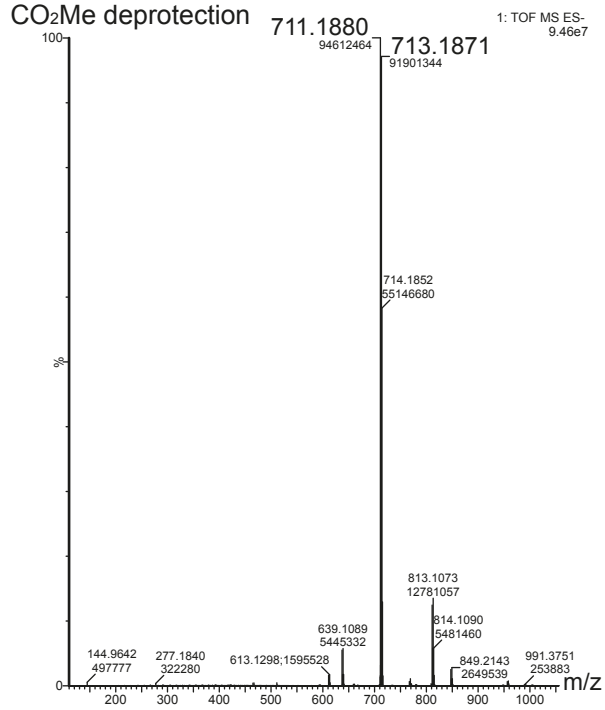
CO_2Me deprotection



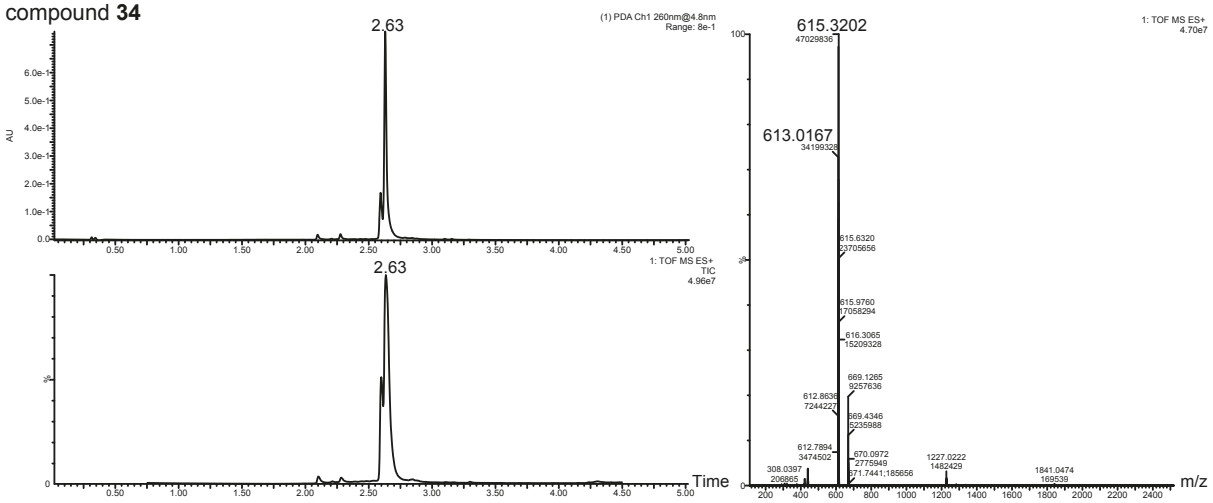
compound **33**



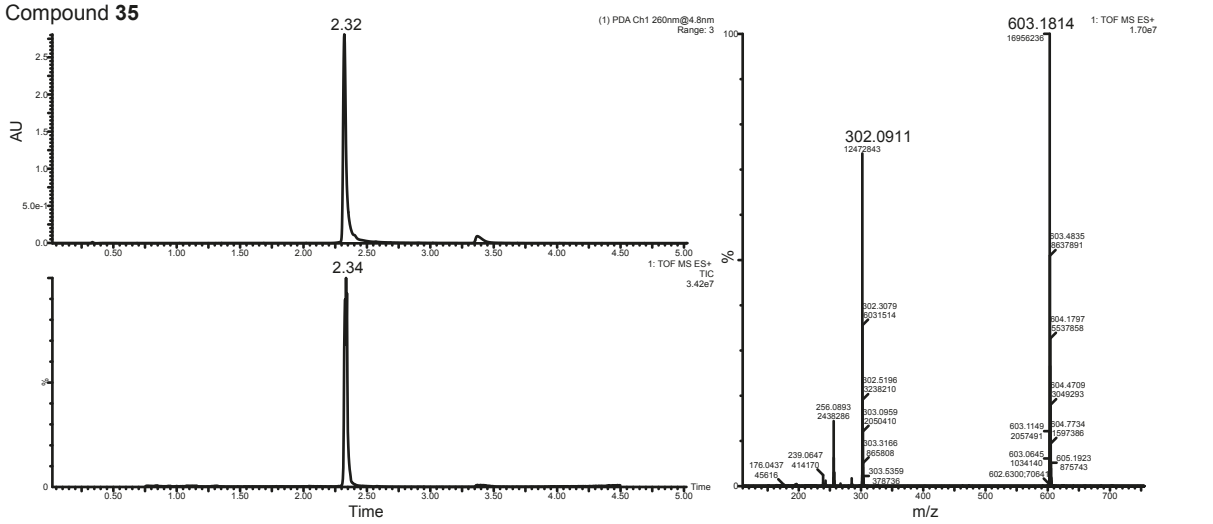
CO_2Me deprotection



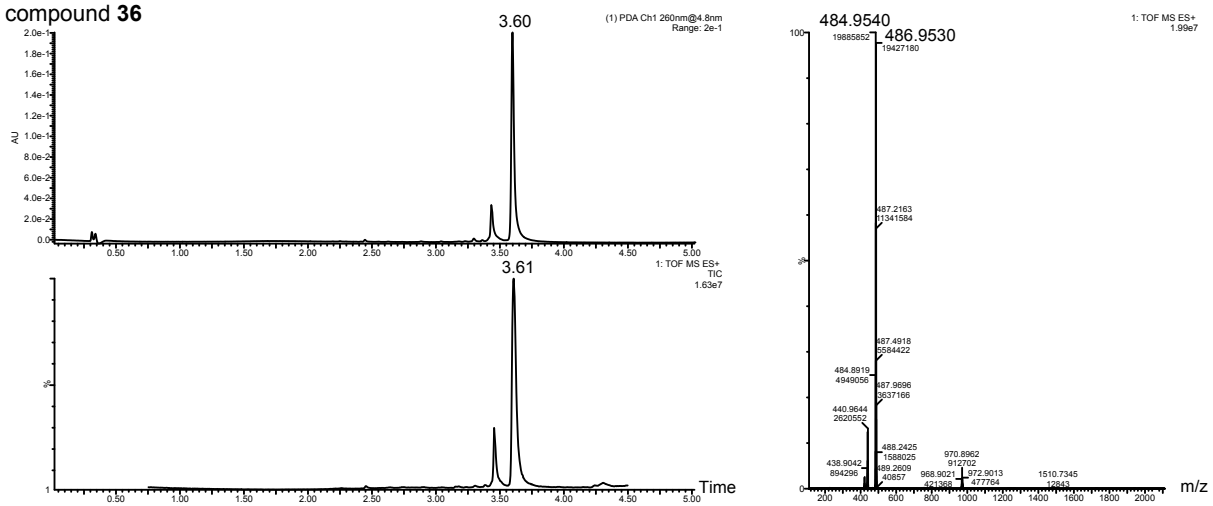
compound **34**



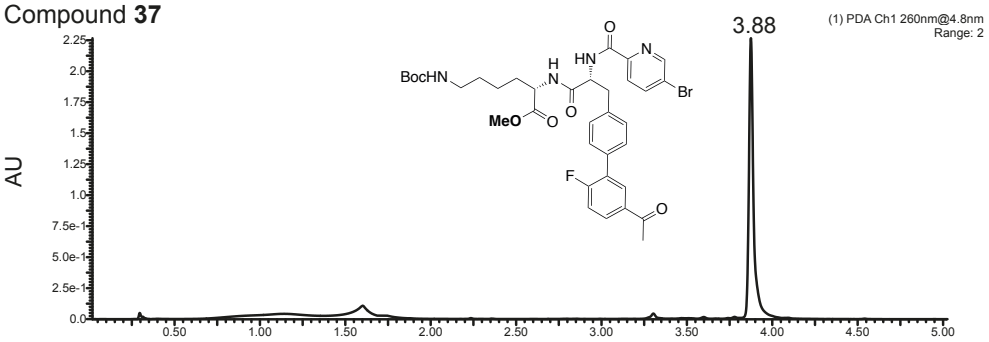
Compound **35**



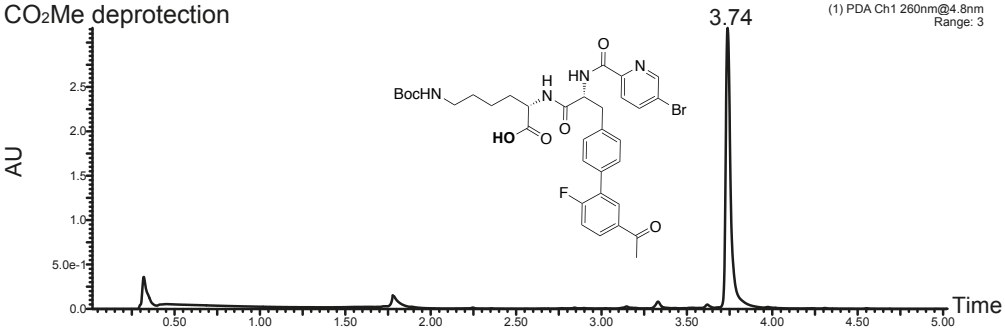
compound **36**



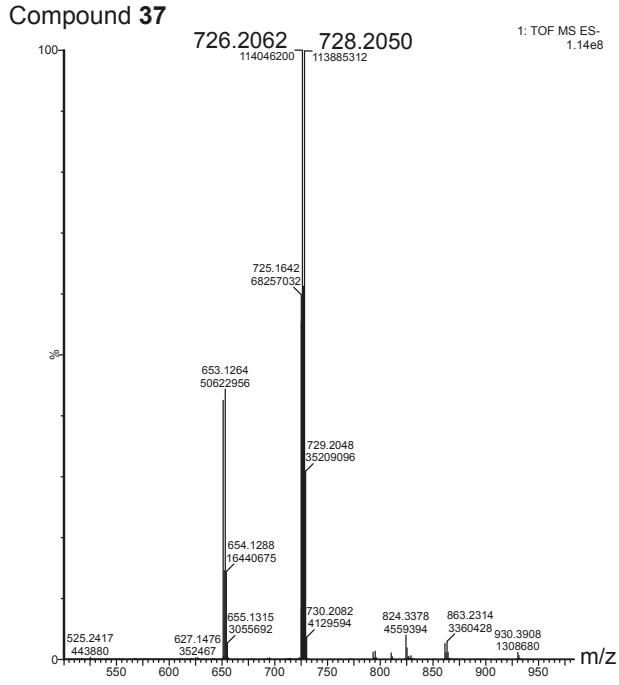
Compound **37**



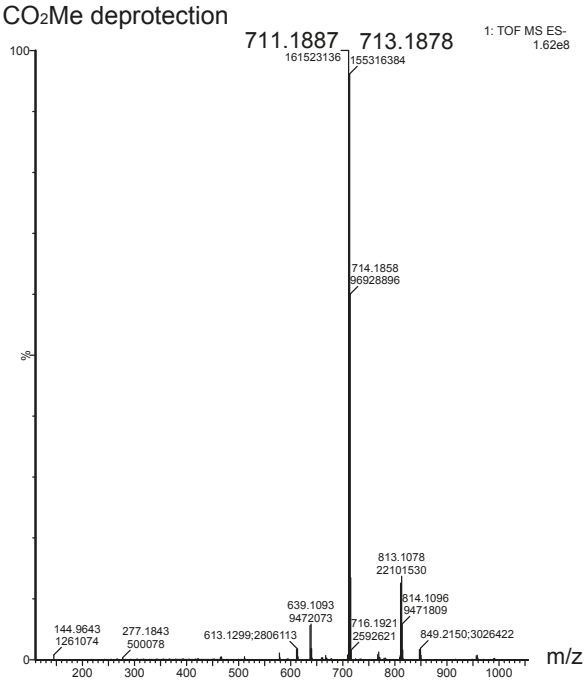
CO_2Me deprotection



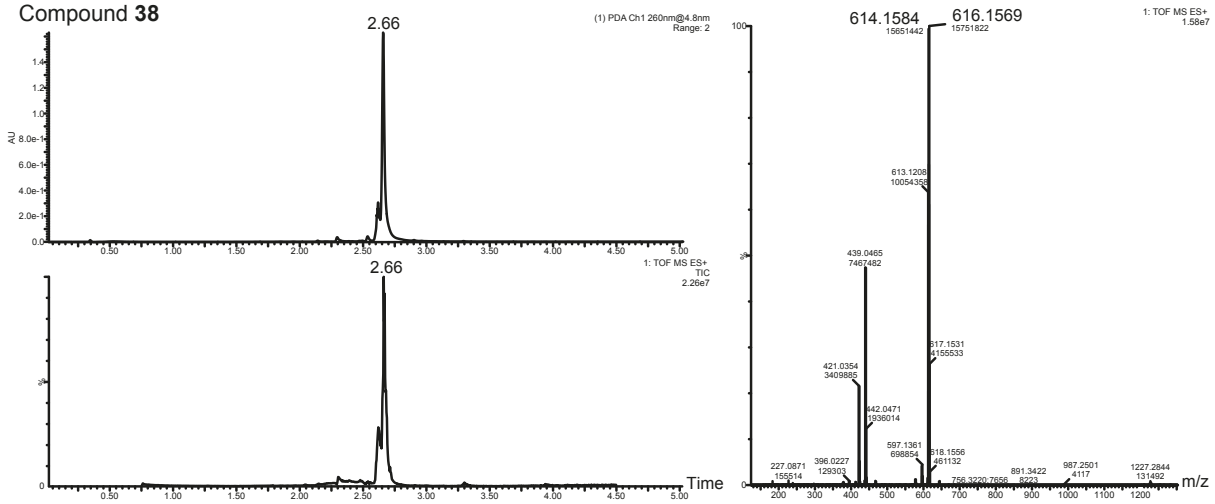
Compound **37**



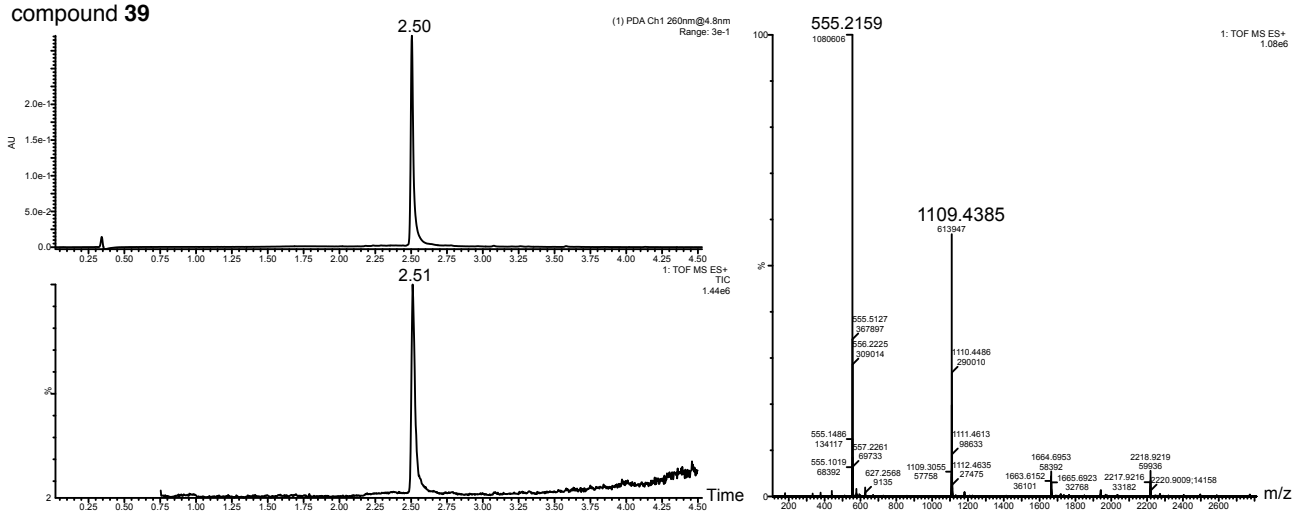
CO_2Me deprotection



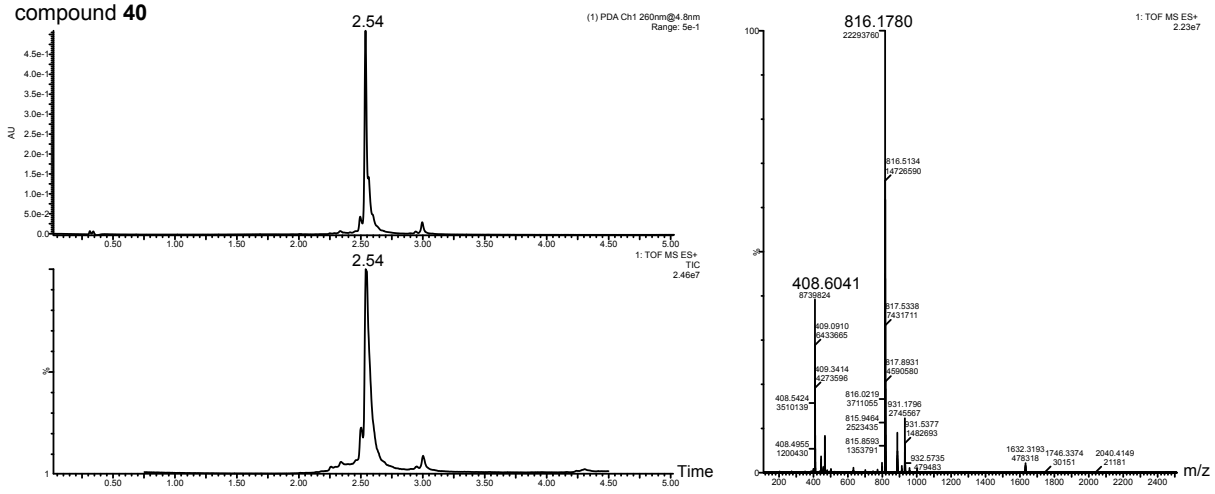
Compound **38**



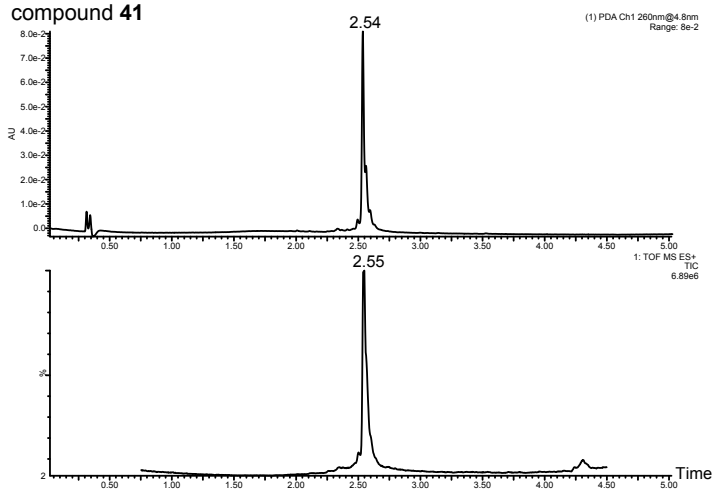
compound **39**

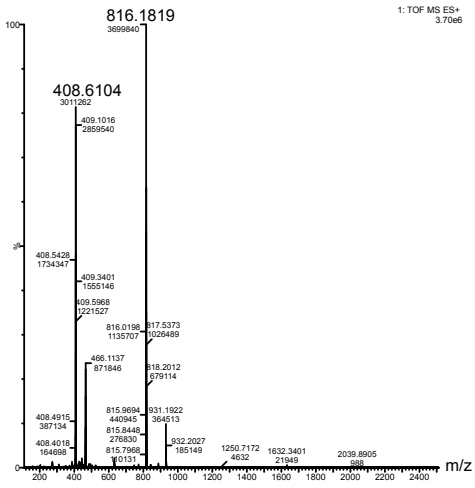


compound **40**

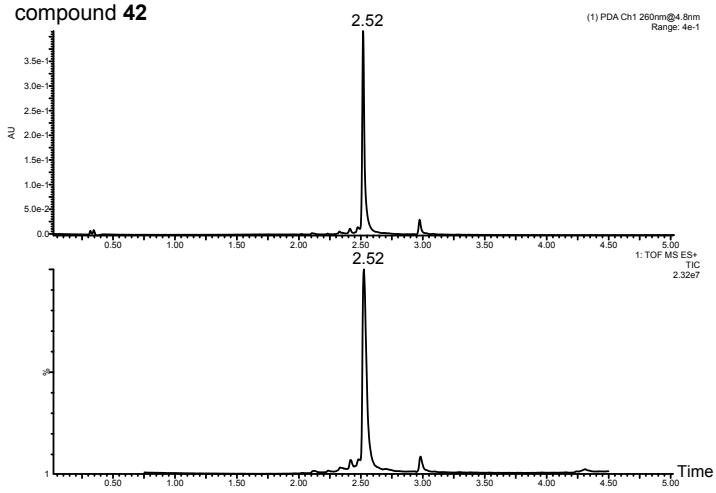


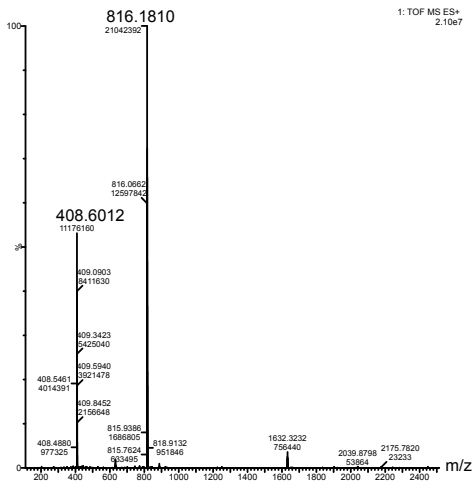
compound **41**



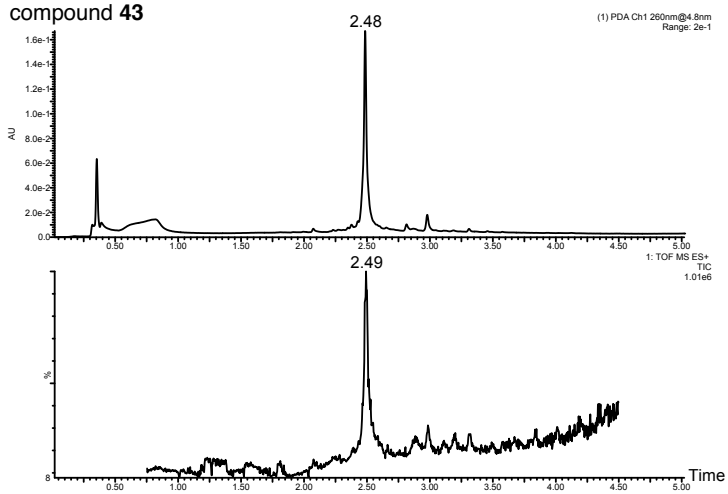


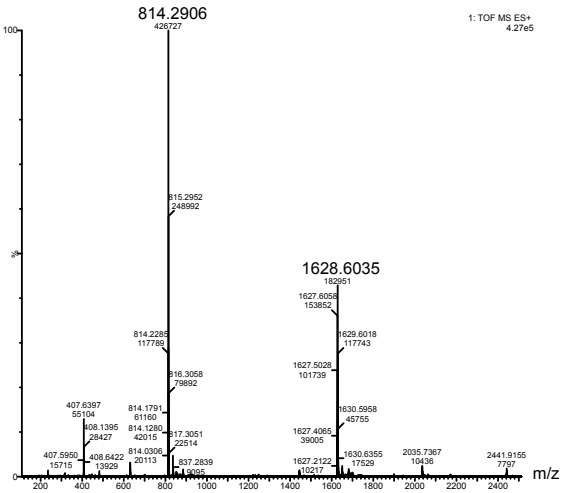
compound **42**



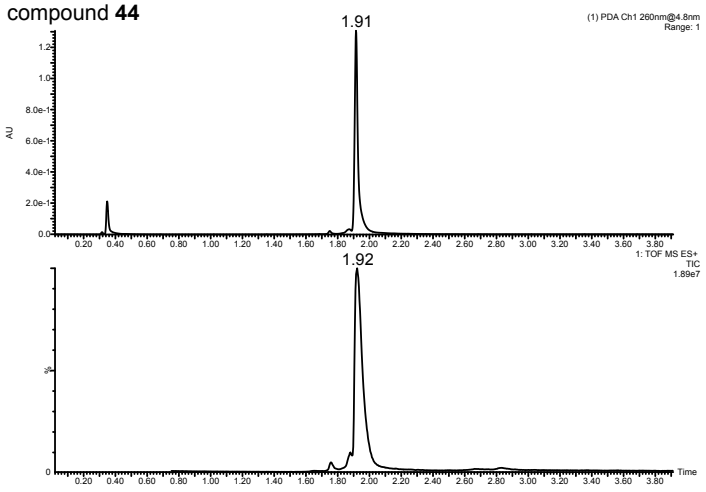


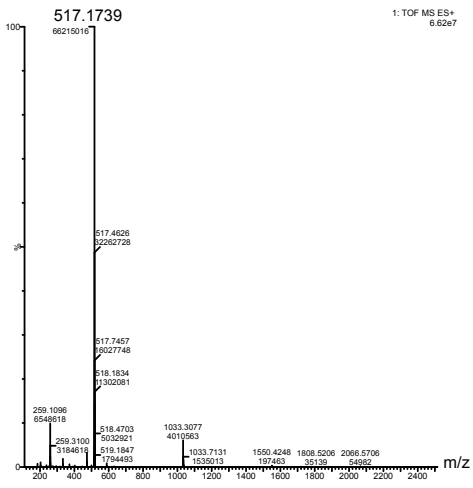
compound **43**



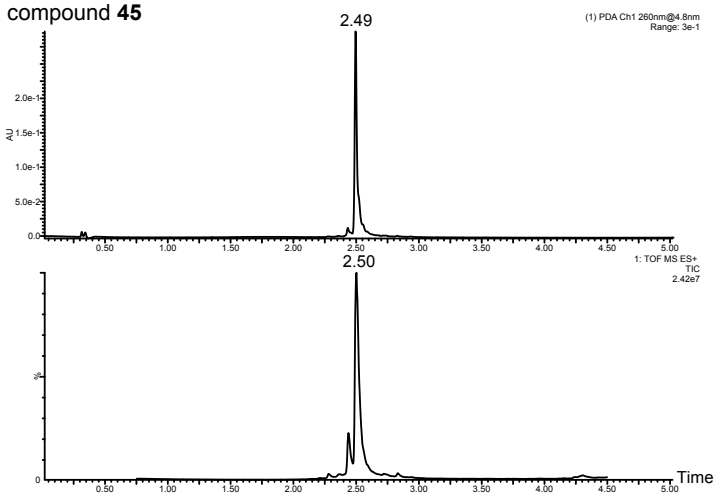


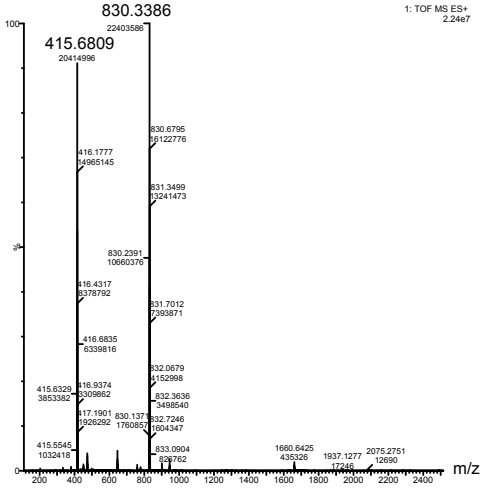
compound **44**



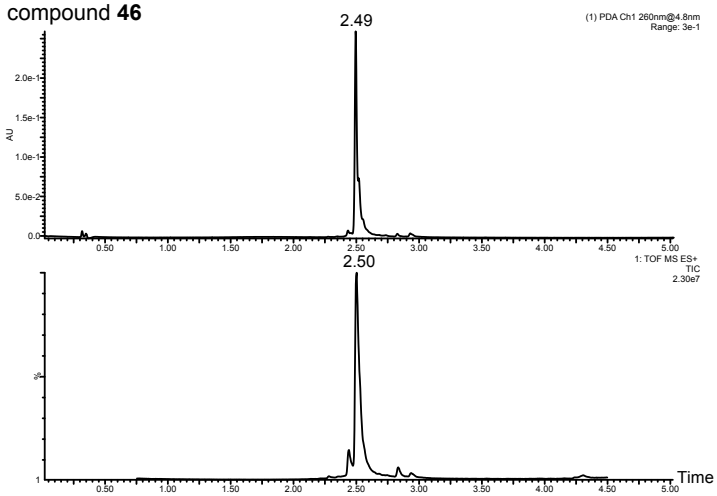


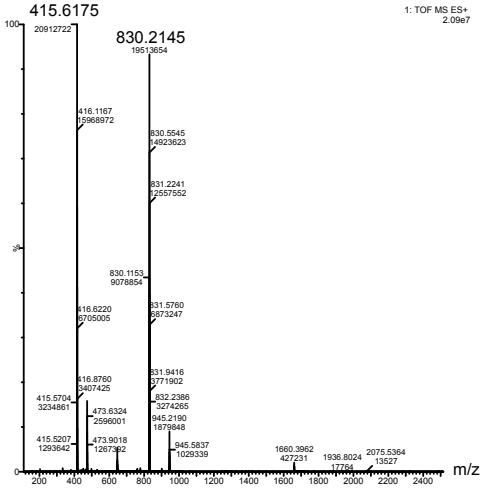
compound **45**



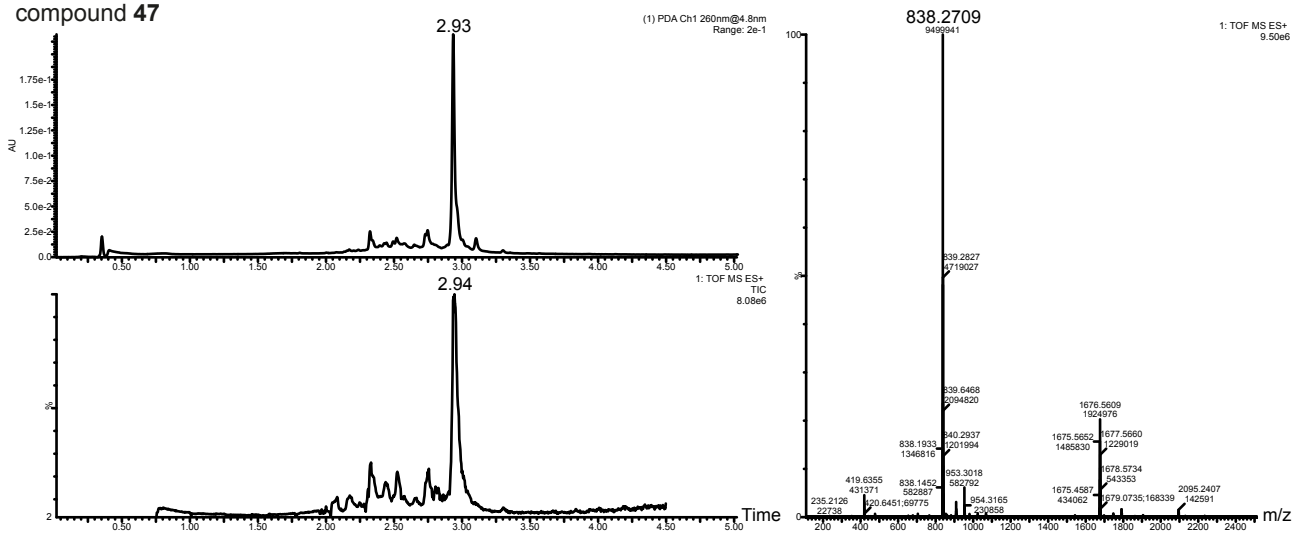


compound **46**

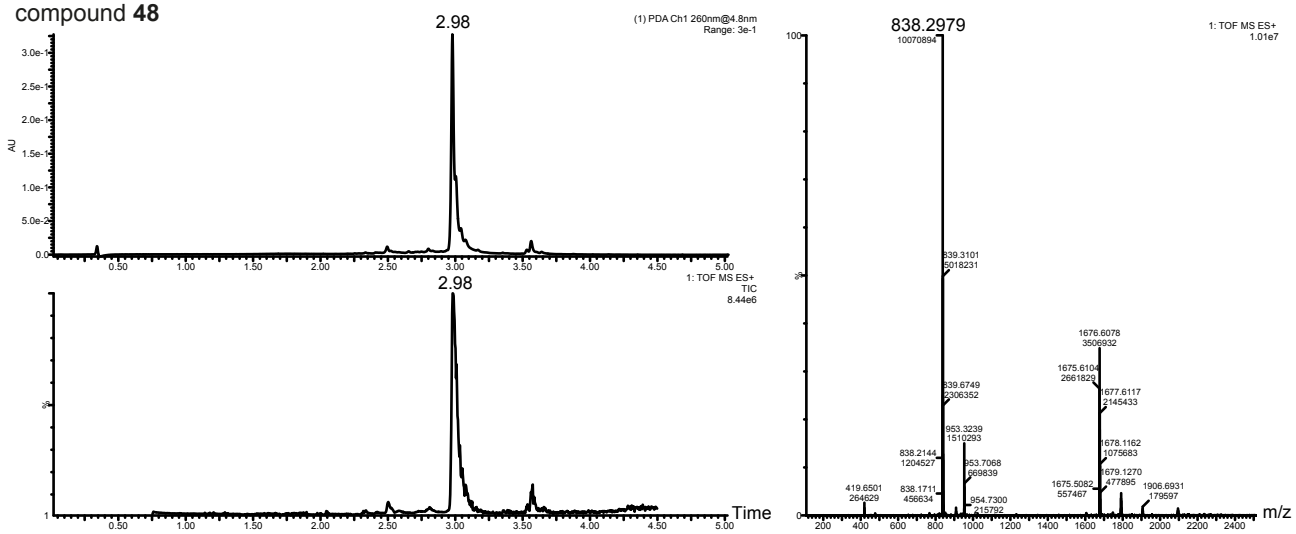




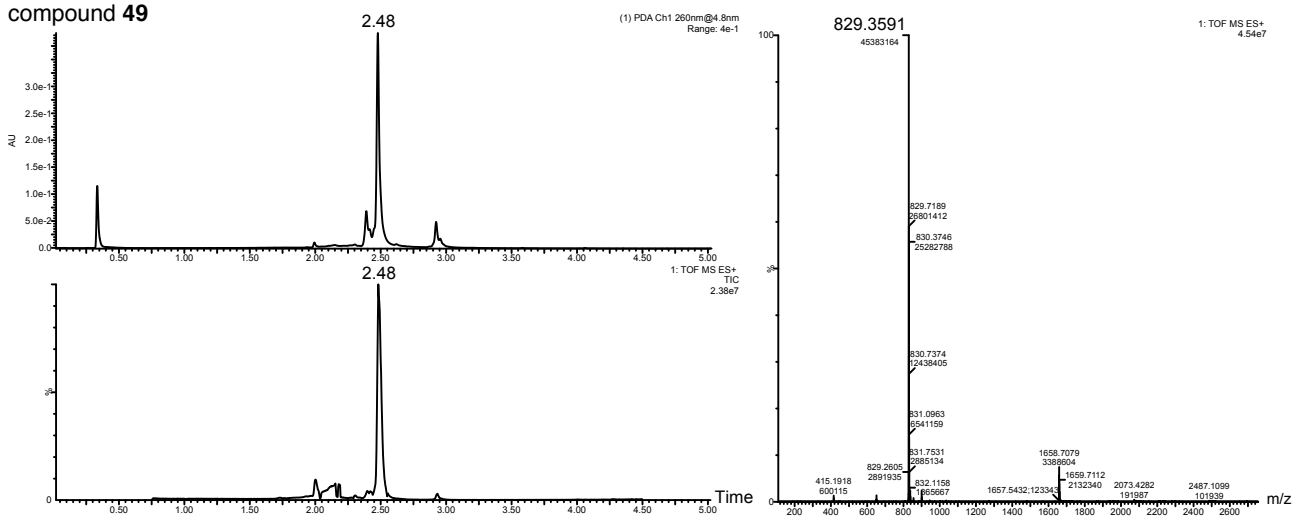
compound **47**



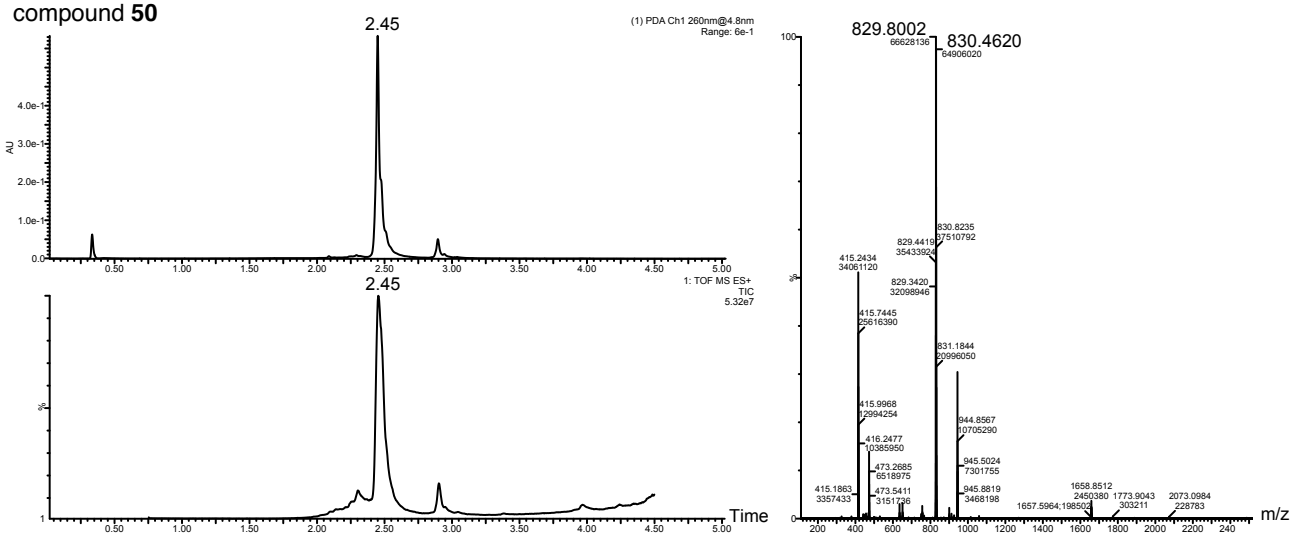
compound **48**



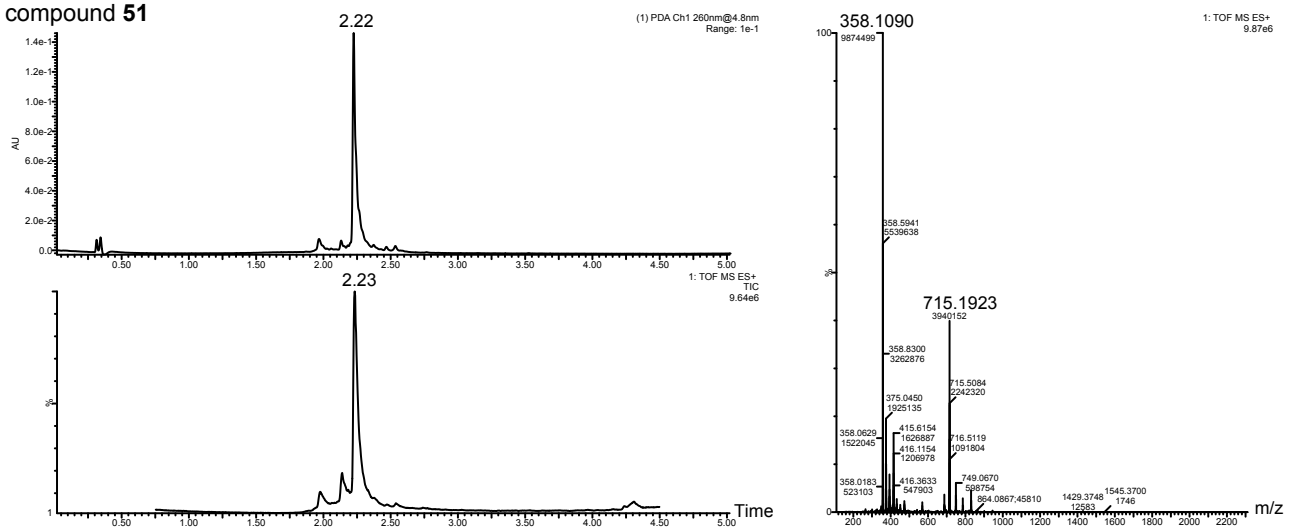
compound **49**



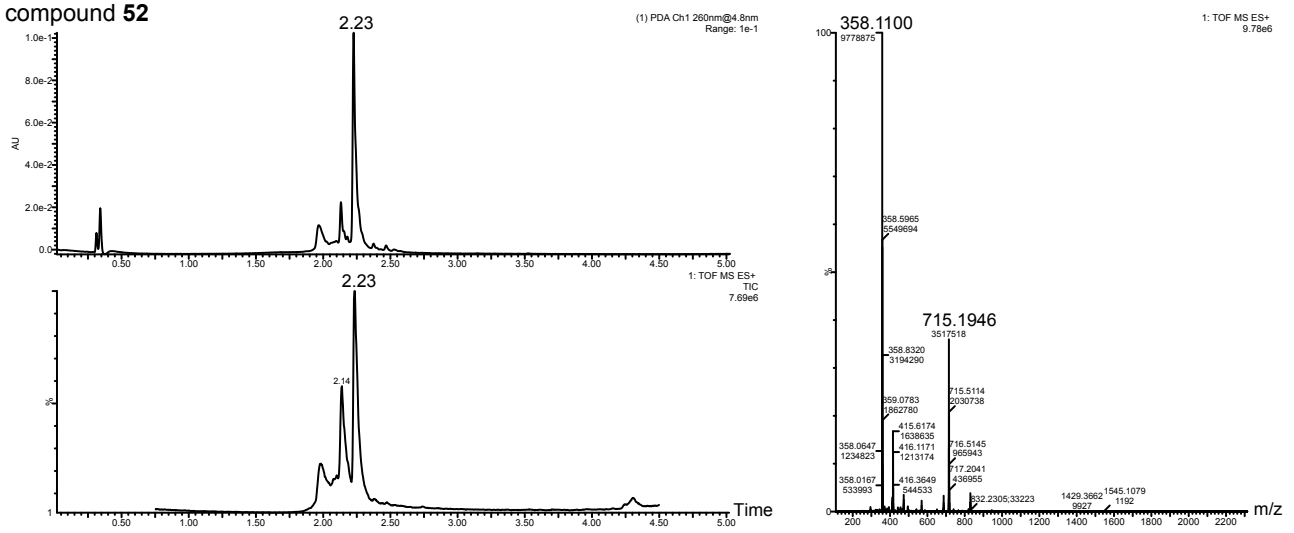
compound **50**



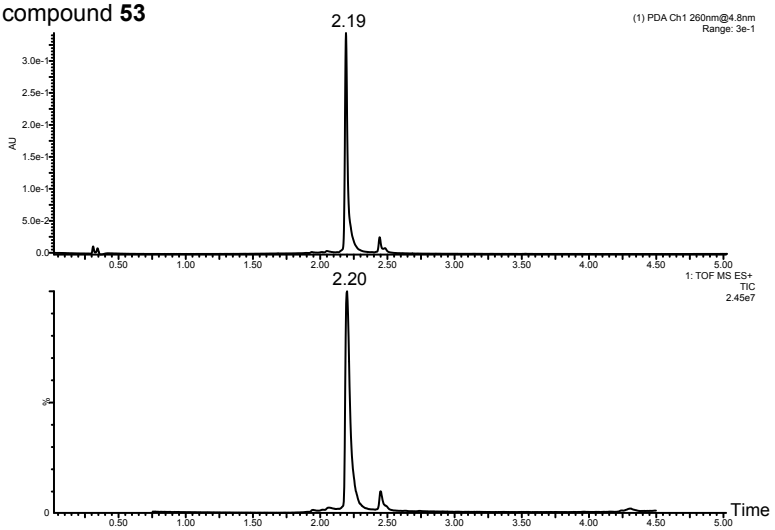
compound **51**

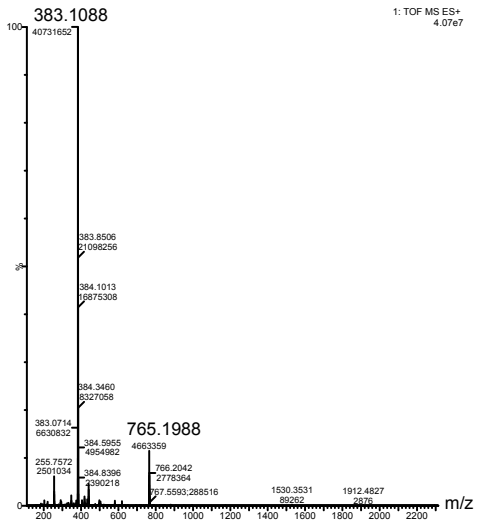


compound **52**

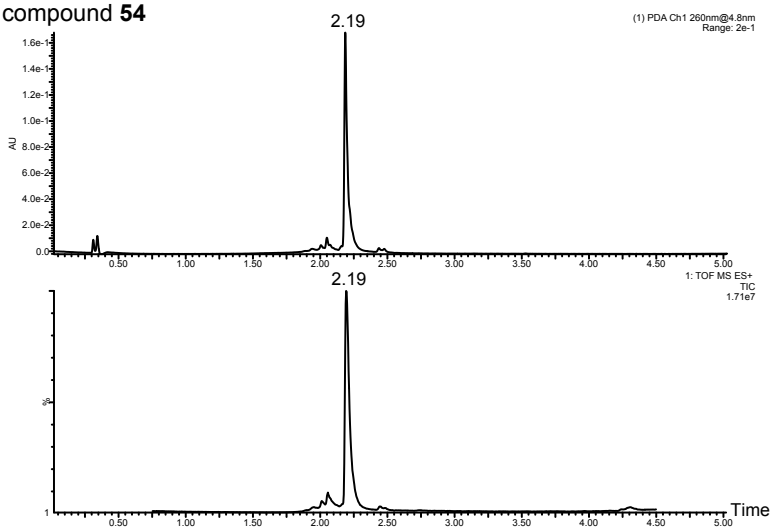


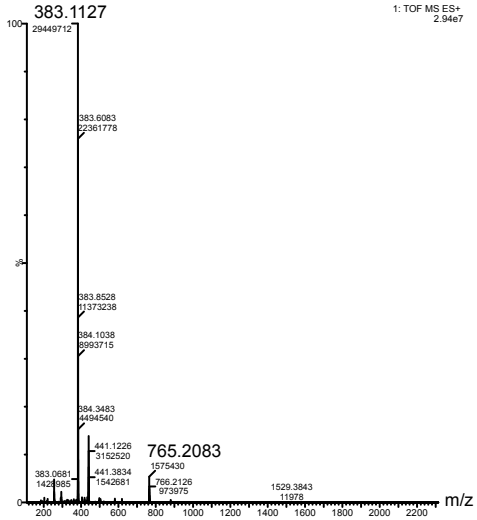
compound **53**



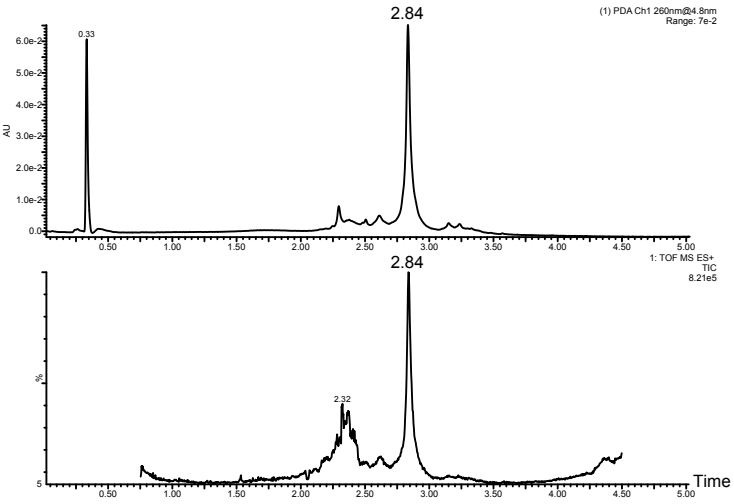


compound **54**

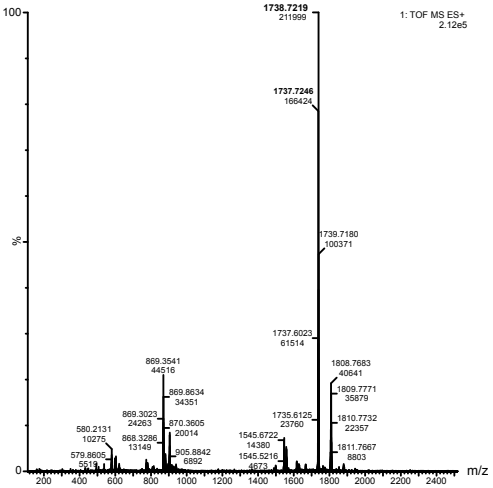




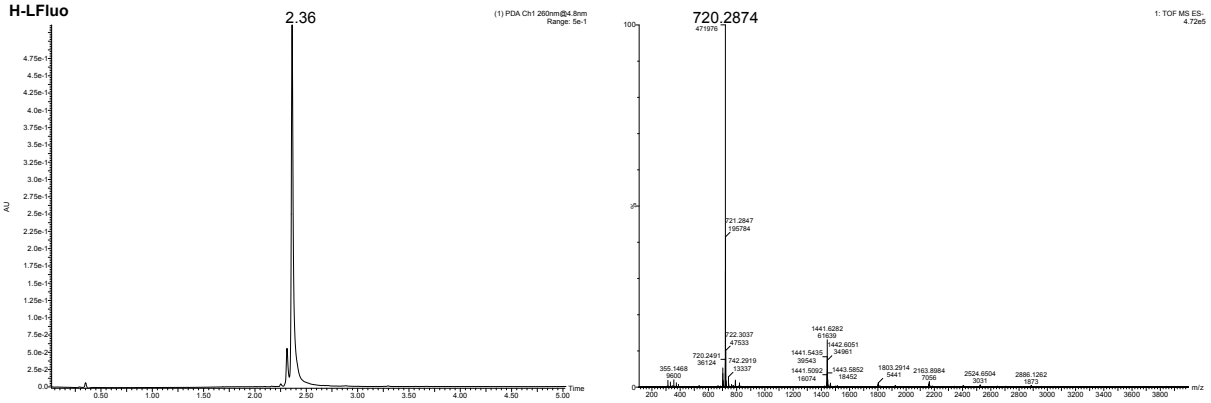
A481/B335 (dimer compound **55**)



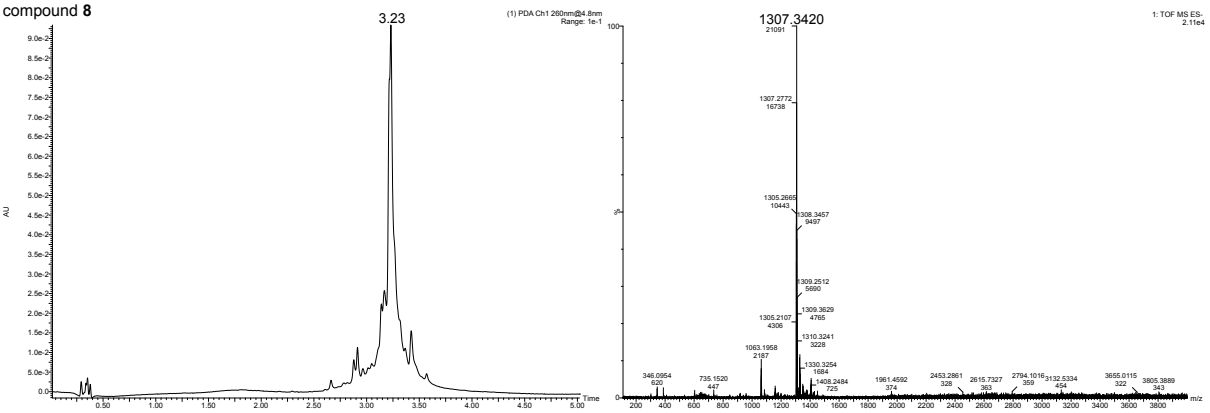
A481/B335 dimer (compound **55**)
570 (2.838) Cm (564:580)



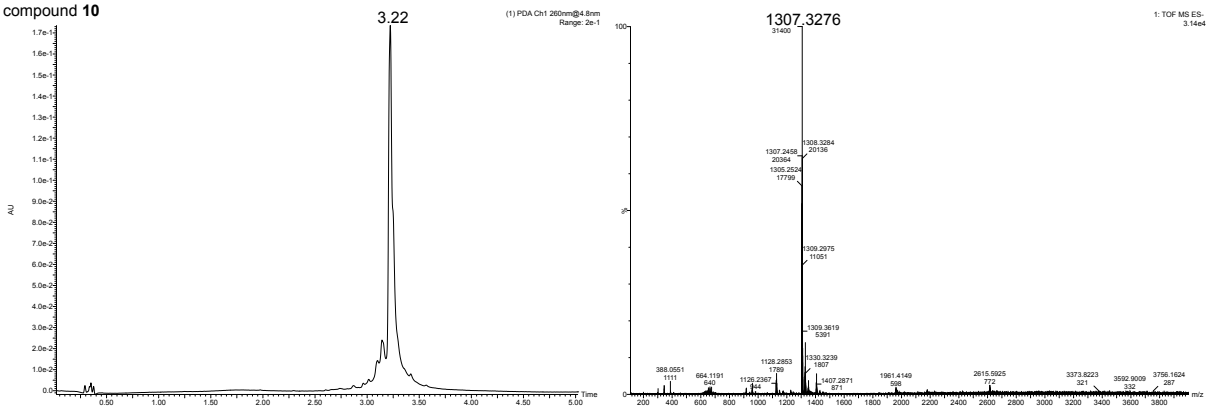
8.3.1 LC-MS of FITC labelled compounds



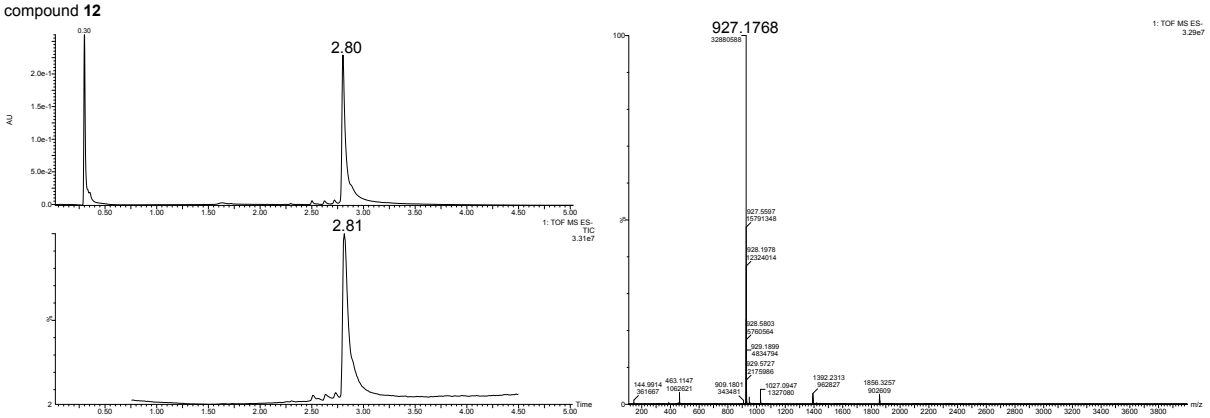
compound **8**



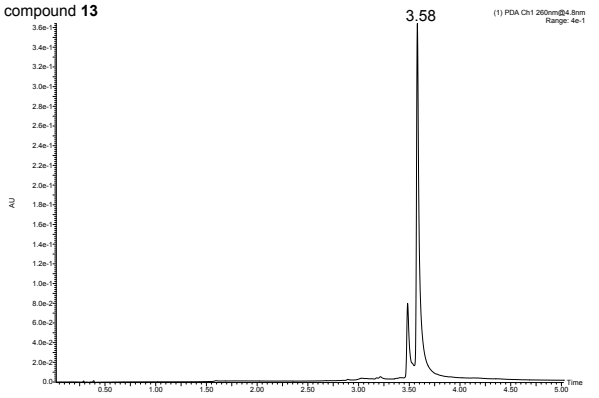
compound **10**

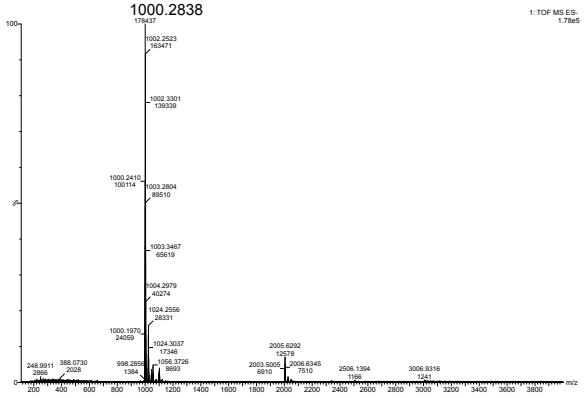


compound **12**

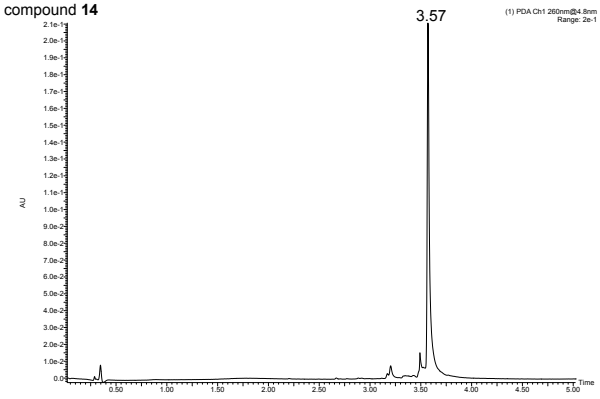


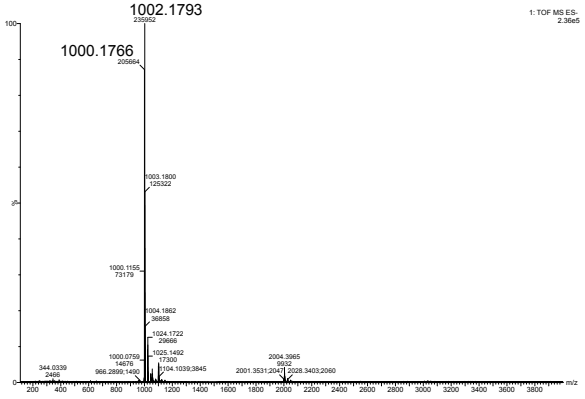
compound **13**



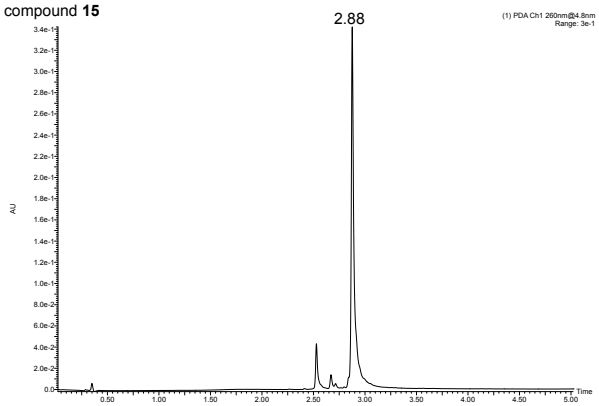


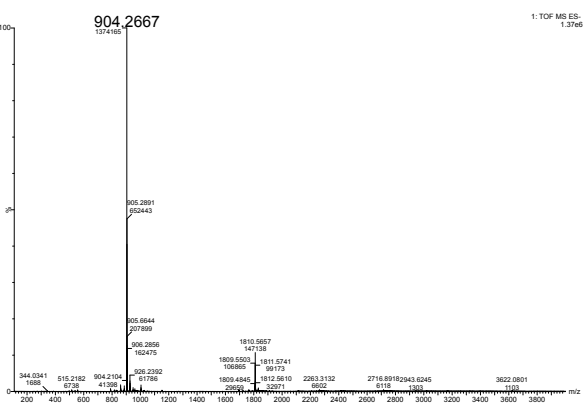
compound **14**



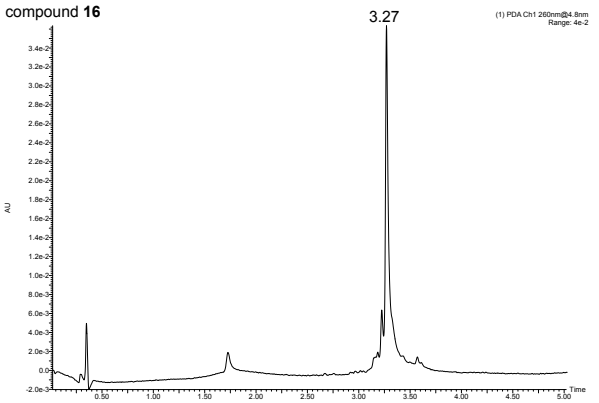


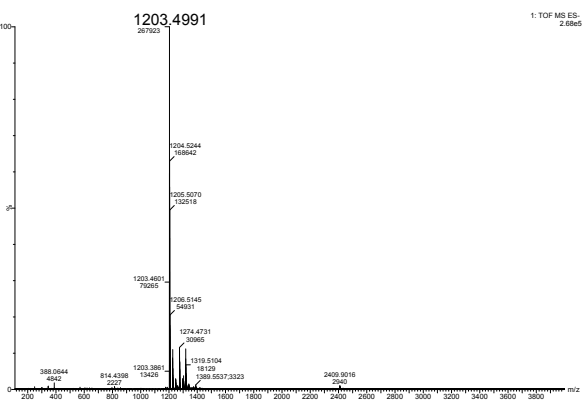
compound **15**



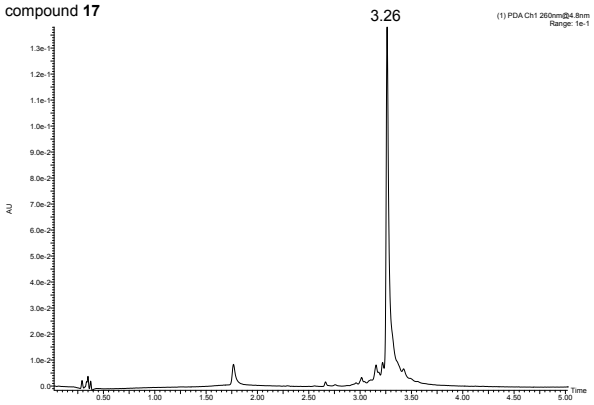


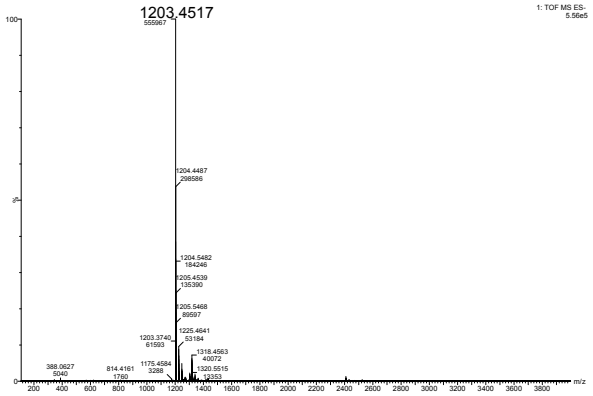
compound **16**



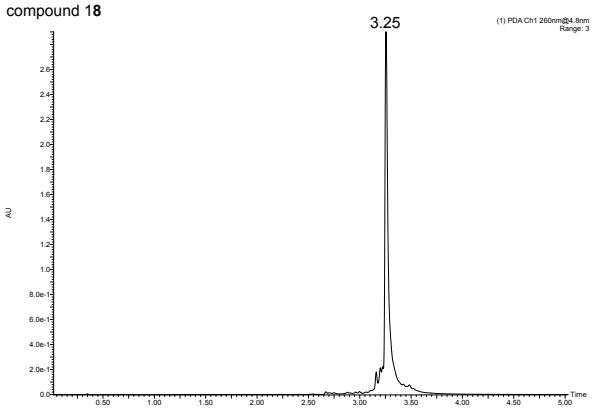


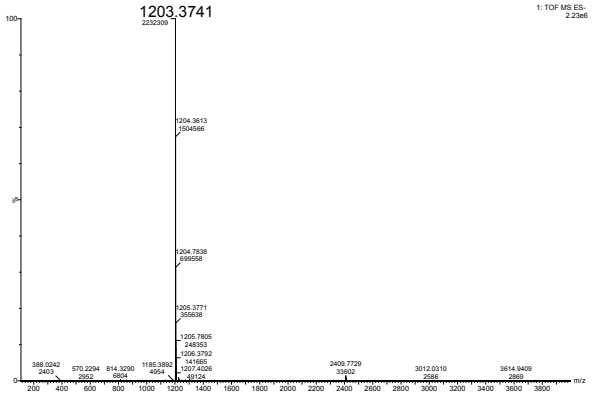
compound **17**



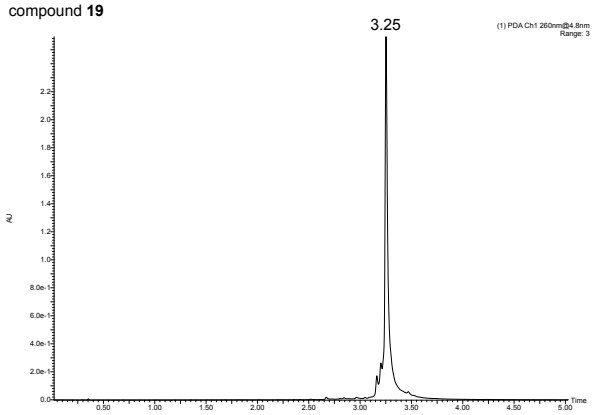


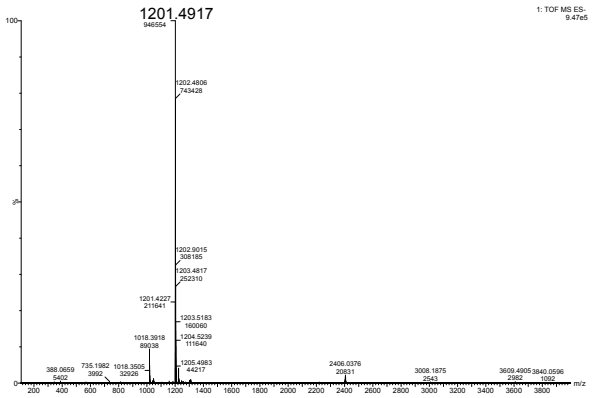
compound **18**



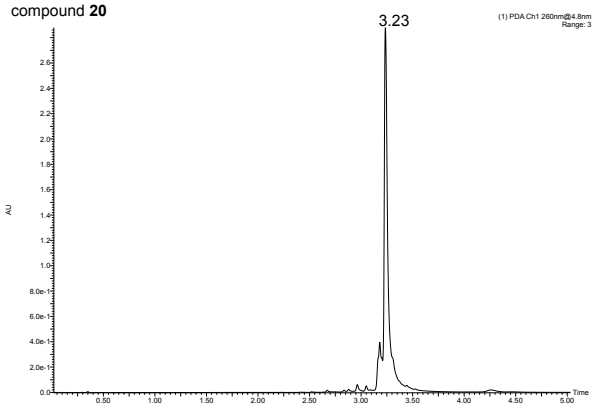


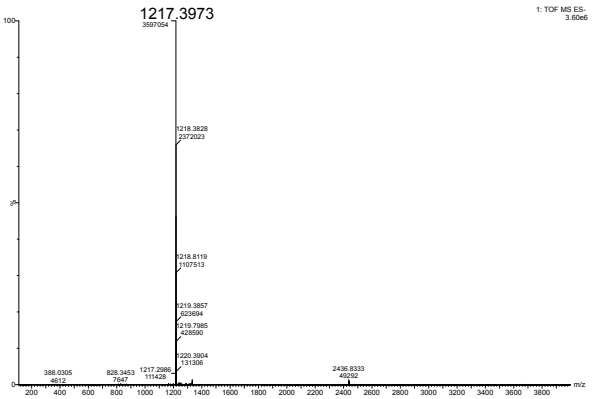
compound **19**



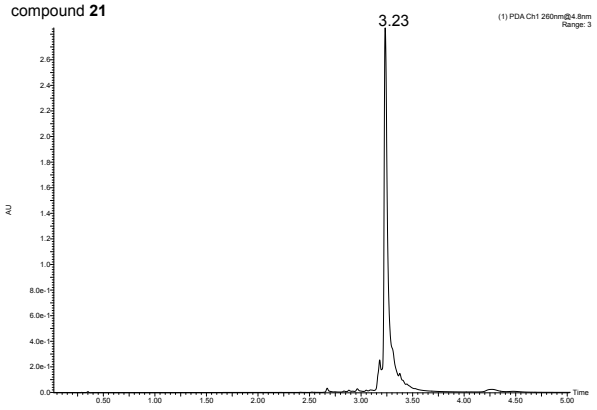


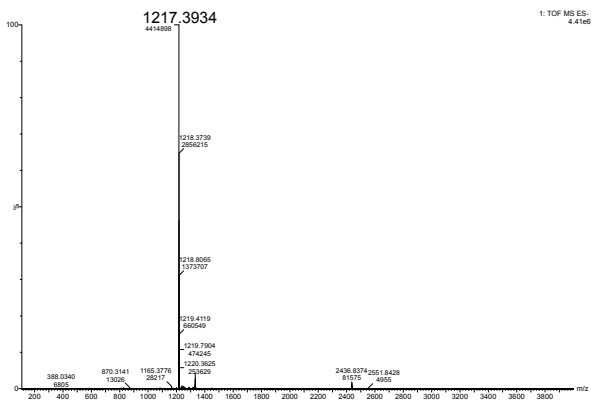
compound **20**



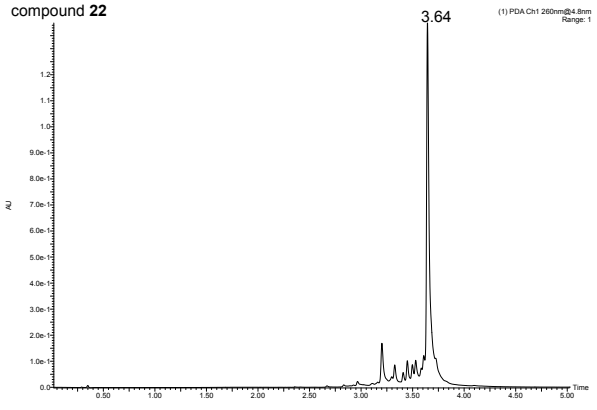


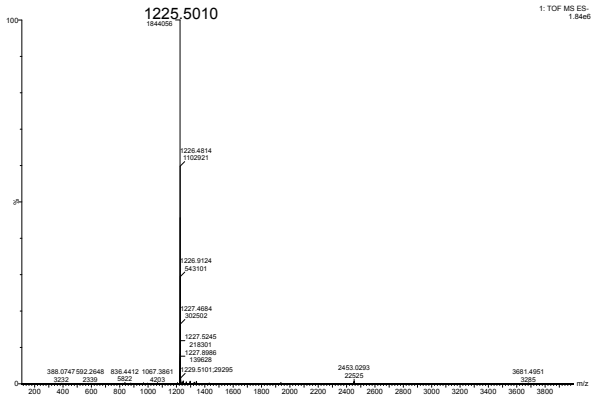
compound **21**



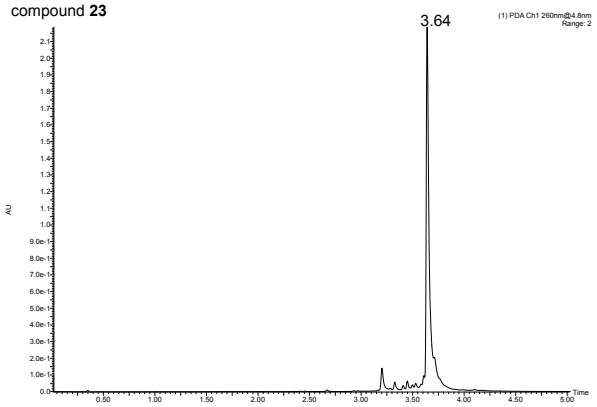


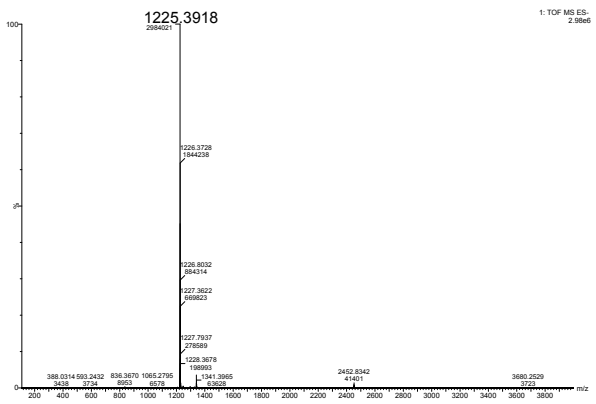
compound **22**



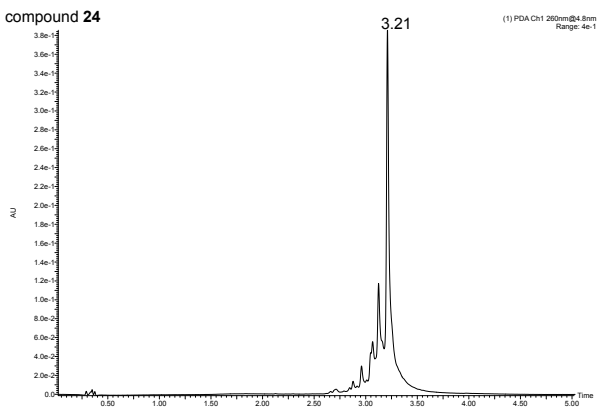


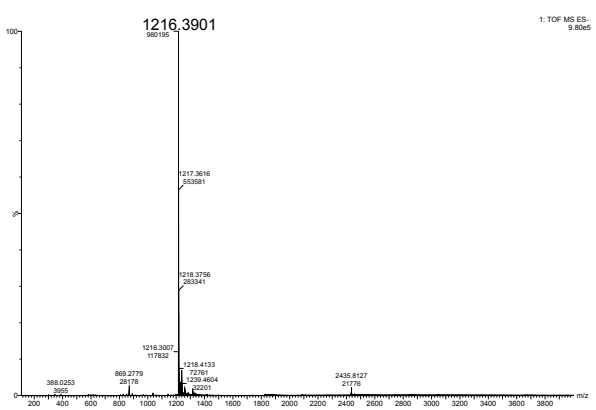
compound **23**



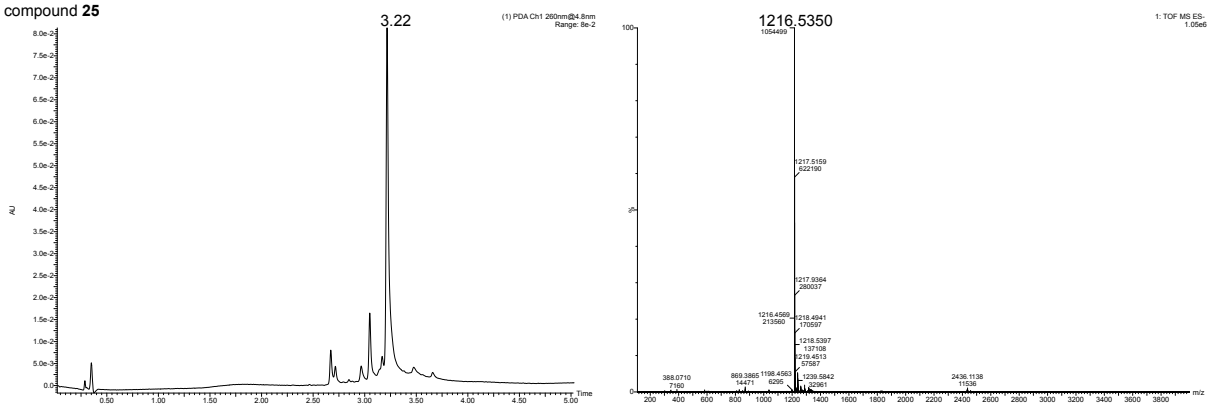


compound **24**

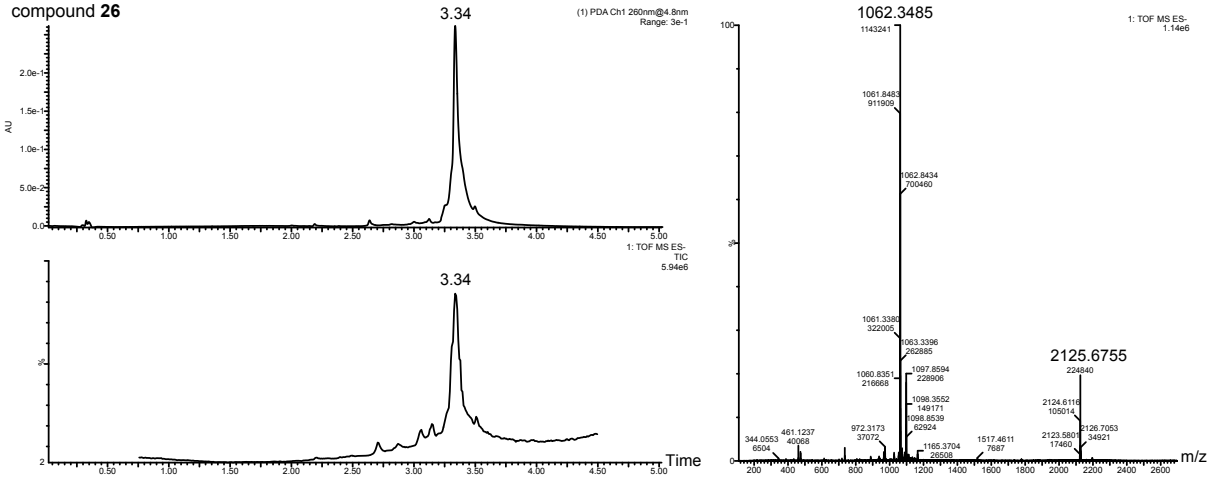




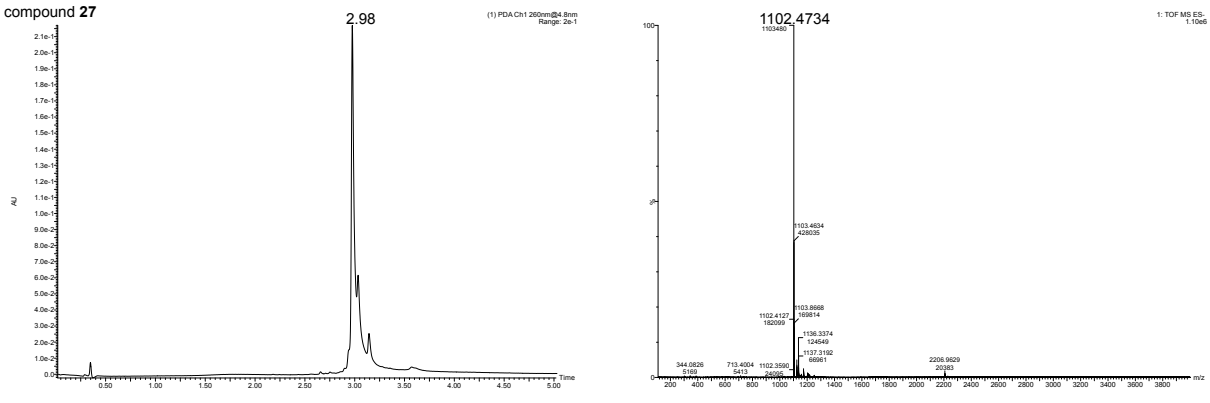
compound **25**



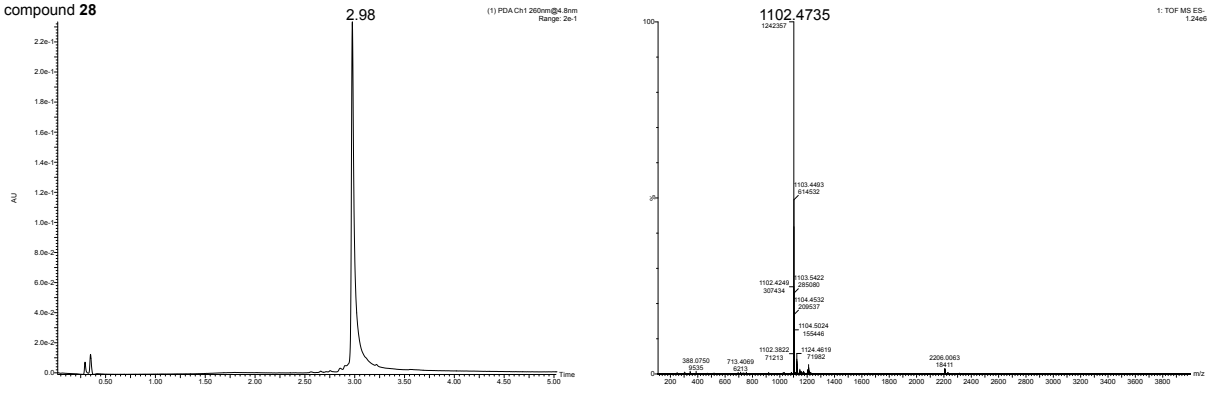
compound **26**



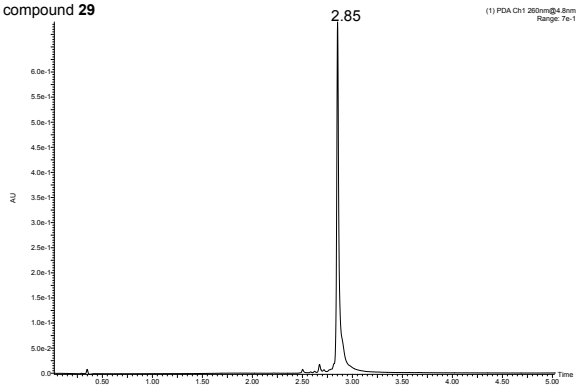
compound **27**

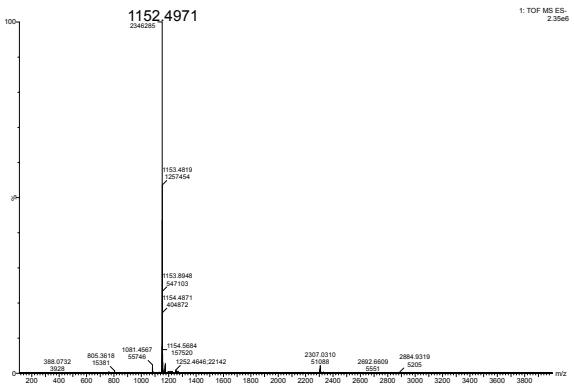


compound **28**

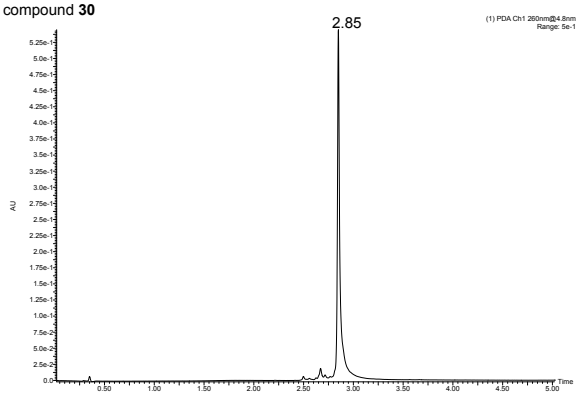


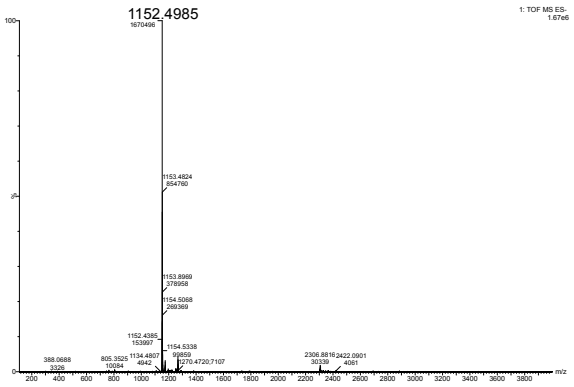
compound **29**





compound **30**





9. References

1 Favalli, N., Bassi, G., Zanetti, T., Scheuermann, J. & Neri, D. Screening of Three Transition Metal-Mediated Reactions Compatible with DNA-Encoded Chemical Libraries. *Helvetica Chimica Acta* **102**, doi:10.1002/hlca.201900033 (2019).

2 Li, Y. *et al.* Optimized Reaction Conditions for Amide Bond Formation in DNA-Encoded Combinatorial Libraries. *ACS Comb Sci* **18**, 438-443, doi:10.1021/acscombsci.6b00058 (2016).

3 Ahlskog, J. K. *et al.* Human monoclonal antibodies targeting carbonic anhydrase IX for the molecular imaging of hypoxic regions in solid tumours. *Br J Cancer* **101**, 645-657, doi:10.1038/sj.bjc.6605200 (2009).

4 Rooney, T. P. *et al.* A series of potent CREBBP bromodomain ligands reveals an induced-fit pocket stabilized by a cation-pi interaction. *Angew Chem Int Ed Engl* **53**, 6126-6130, doi:10.1002/anie.201402750 (2014).

5 Hewings, D. S. *et al.* 3,5-dimethylisoxazoles act as acetyl-lysine-mimetic bromodomain ligands. *J Med Chem* **54**, 6761-6770, doi:10.1021/jm200640v (2011).

6 Brack, S. S., Silacci, M., Birchler, M. & Neri, D. Tumor-targeting properties of novel antibodies specific to the large isoform of tenascin-C. *Clin Cancer Res* **12**, 3200-3208, doi:10.1158/1078-0432.CCR-05-2804 (2006).

7 Godau, J. *et al.* Identification of a miniature Sae2/Ctp1/CtIP ortholog from Paramecium tetraurelia required for sexual reproduction and DNA double-strand break repair. *DNA Repair (Amst)* **77**, 96-108, doi:10.1016/j.dnarep.2019.03.011 (2019).

8 Davies, O. R. *et al.* CtIP tetramer assembly is required for DNA-end resection and repair. *Nat Struct Mol Biol* **22**, 150-157, doi:10.1038/nsmb.2937 (2015).

9 Bassi, G. *et al.* Comparative evaluation of DNA-encoded chemical selections performed using DNA in single-stranded or double-stranded format. *Biochem Biophys Res Commun* **533**, 223-229, doi:10.1016/j.bbrc.2020.04.035 (2020).

10 Decurtins, W. *et al.* Automated screening for small organic ligands using DNA-encoded chemical libraries. *Nat Protoc* **11**, 764-780, doi:10.1038/nprot.2016.039 (2016).

11 Midelfort, K. S. *et al.* Substantial energetic improvement with minimal structural perturbation in a high affinity mutant antibody. *J Mol Biol* **343**, 685-701, doi:10.1016/j.jmb.2004.08.019 (2004).

12 Myburgh, R. *et al.* Anti-human CD117 CAR T-cells efficiently eliminate CD117-expressing healthy and malignant hematopoietic cells. *Leukemia* https://doi.org/10.1038/s41375-020-0818-9 (2020).