Supporting Information

An Unnatural Amino Acid That Mimics Phosphotyrosine

Jingyan Ge[a], Hao Wu[a] and Shao Q. Yao*[a,b,c]

Departments of [a]Chemistry and [b]Biological Sciences, [c]NUS MedChem Program of the Life Sciences Institute, National University of Singapore, 3 Science Drive 3, Singapore 117543, Republic of Singapore

**Corresponding author (chmyaosq@nus.edu.sg)*

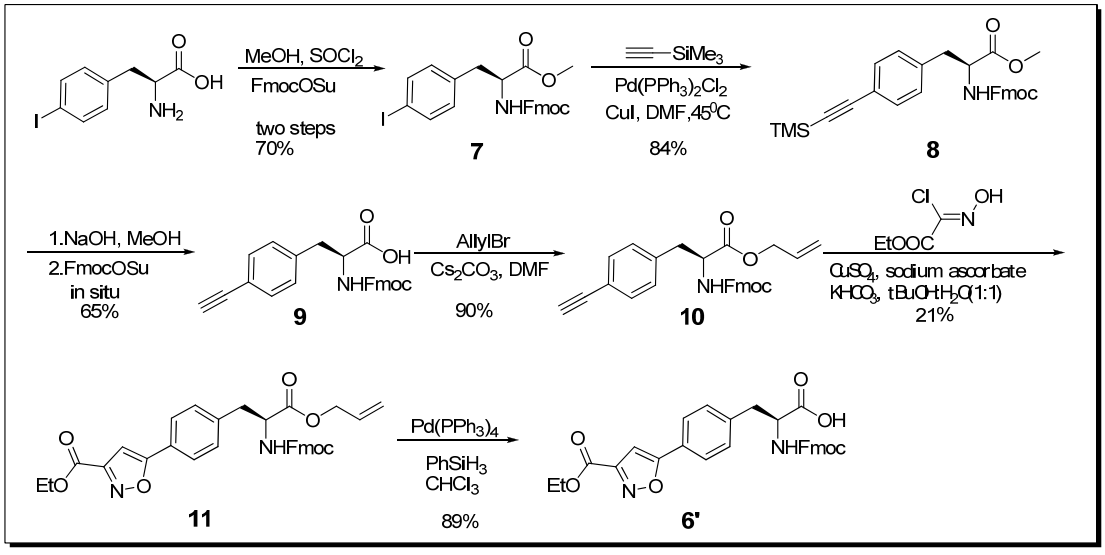
Contents

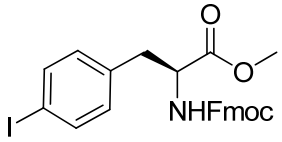
1. **General information**
2. **Synthetic procedures for the unnatural amino acid**
3. **Procedure and LC-MS profiles of inhibitors and SH2-binding peptide**
4. **Procedures and results of cell proliferation**
5. **Procedure and results of fluorescence polarization (FP) experiments**
6. **Procedure and results of docking simulations**
7. **Procedure and results of hydrolytic stability test**
8. **Procedure and results of cell permeability test**
9. **References**
10. **^{1}H and ^{13}C NMR spectra**

1. General Information

All chemicals were purchased from commercial vendors and used without further purifications, unless otherwise noted. ^{1}H NMR spectra and ^{13}C NMR were recorded on Bruker 300 MHz, 500 MHz or DPX-300 NMR spectrometers. Chemical shifts are reported in parts per million referenced with respect to residual solvent ($CDCl_3$ = 7.26 ppm and DMSO-d_6 = 2.5 ppm) or from internal standard tetramethylsilane (TMS = 0.00 ppm). The following abbreviation was used in reporting spectra: s = singlet, d = doublet, t = triplet, q = quarter, m = multiplet, dd = doublet of doublets. High resolution mass spectra (HRMS) were recorded on Finnigan MAT95XL-T using ESI. All solvents used were of HPLC grade, unless otherwise indicated. All reactions requiring anhydrous conditions were conducted under a nitrogen or argon atmosphere in flame-dried glasswares. Dimethylformamide was dried over calcium hydride and distilled under reduced pressure. All LC profiles and mass spectra were recorded on Shimadzu LC-IT-TOF and LC-ESI systems equipped with an autosampler, using reverse-phase Phenomenex Luna 5 C_{18} (2) 100 Å 50 × 3.0 mm column. Proteins were expressed in *E. coli* strain BL21-DE3 and purified by Ni-NTA technology. IC_{50} curves were generated using the Graphpad Prism software *v*. 5 (GraphPad, San Diego, USA).

2. Synthetic procedures for the unnatural amino acid

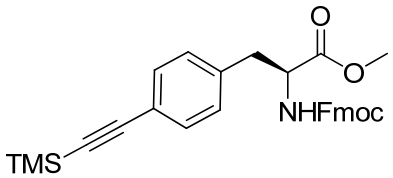




Methyl 2-(((*9H*-fluoren-9-yl)methoxy)carbonylamino)-3-(4-iodophenyl) propanoate (7)

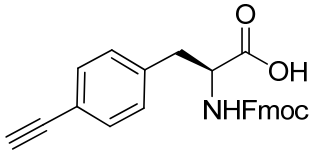
Thionyl Chloride (15 mL) was added dropwise to 4-Iodo-*L*-phenylalanine (12 g, 41.3 mmol, synthesized following reference 1) [1] in a methanol (200 mL) solution at 0 °C.

The yellow solution was refluxed for 2 hours. The solution was evaporated and dissolved in a minimal amount of hot MeOH. Et_2O was added to obtain the white product precipitation. The crude product was dissolved in $NaHCO_3$ (sat.) and dichloromethane (DCM) (20 mL each). To this solution was added Fmoc-OSu (15.2 g, 45.4 mmol), followed by stirring for 5 hours. DCM was evaporated *in vacuo*. The residue was extracted with EtOAc three times. The organic layer was combined, washed with brine and dried over Na_2SO_4. Flash column chromatography (9/1 to 3/1 Hexane/EtOAc gradients) afforded the protected unnatural amino acid **7** as a white solid (15.2 g, 70 % in two steps). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.78 (d, J = 7.38 Hz, 2H), 7.55-7.61 (m, 4H), 7.41 (t, J = 7.40 Hz, 2H), 7.32 (t, J = 7.39 Hz, 2H), 6.81 (d, J = 7.89 Hz, 2H), 5.23 (d, J = 8.22 Hz, NH), 4.63-4.65 (1H, m), 4.34-4.50 (2H, m), 4.20 (t, J = 6.42 Hz, 1H), 3.73 (s, 3H), 2.98-3.08 (m, 2H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 171.6, 155.5, 143.6, 141.3, 137.6, 135.4, 131.2, 127.7, 127.1, 125.0, 120.0, 92.6, 66.8, 54.5, 52.4, 47.14, 37.7. IT-TOF: m/z $[M+H]^+$ calcd: 528.06, found: 528.03.



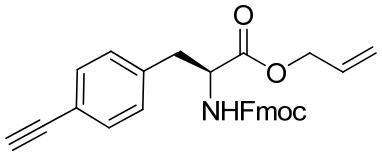
Methyl 2-(((*9H*-fluoren-9-yl)methoxy)carbonylamino)-3-(4-((trimethylsilyl) ethynyl) phenyl)propanoate (8)

DMF (25 mL) was degassed by bubbling with argon in a dry two-neck round bottle flask. Subsequently, compound **7** (1.2 g, 2.28 mmol) was added and the solution was stirred until it became homogenous. $Pd(PPh_3)_2Cl_2$ (16 mg, 0.02 mmol) and CuI (18.9 mg, 0.1 mmol), DIEA (1.2 mL, 6.5 mmol) were subsequently added under an N_2 atmosphere. Trimethylsilyl acetylene (0.63 mL, 4.56 mmol) was then added by a needled syringe. The reaction was heated at 45 ^{o}C for 3 hours before being quenched by addition of an ammonium chloride solution and EtOAc. Then the aqueous phase was extracted with EtOAc. The combined organic phases were washed with brine and dried with Na_2SO_4. After removal of the solvent under reduced pressure, column chromatography on silica gel (1/9 to 1/6 EtOAc/Hexane gradients) afforded **8** as a light yellow solid (84 % yield). ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.75 (d, J = 7.38 Hz, 2H), 7.53-7.55 (broad, 2H), 7.27-7.41 (m, 6H), 7.01 (d, J = 7.71 Hz, 2H), 5.27 (d, J = 8.04 Hz, NH), 4.64 (m, 1H), 4.35-4.46 (m, 2H), 4.19 (t, J = 6.66 Hz, 1H), 3.69 (s, 3H), 3.05-3.10 (m, 2H), 0.25 (s, 9H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 171.6, 155.4, 143.7, 141.3, 136.3, 132.1, 129.1, 127.7, 127.0, 125.0, 121.9, 120.0, 104.7, 94.4, 66.9, 54.6, 52.3, 47.1, 38.1, -0.1. IT-TOF: m/z $[M+Na]^+$ calcd: 520.20, found: 520.17.



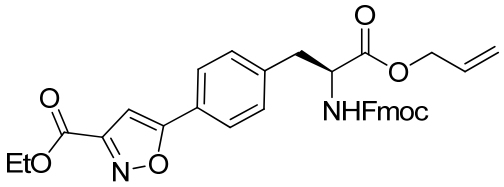
2-(((*9H*-fluoren-9-yl)methoxy)carbonylamino)-3-(4-ethynylphenyl)propanoic acid (9)

The compound **8** (2.0 g, 4 mmol) was dissolved in MeOH (10 mL) and cooled to 0 ^{o}C. NaOH (10 mL, 10 M) was added dropwise. After 5 hours, the *p*H of the solution was adjusted to 8 using HCl and $NaHCO_3$ (sat). Then to the mixture was added Fmoc-OSu (1.6 g, 4.8 mmol) followed by stirring overnight. MeOH was removed *in vacuo* and the aqueous layer was extracted with EtOAc. The combined organic phases were washed with brine and dried with Na_2SO_4. After removal of the solvent under reduced pressure, column chromatography on silica gel (9/1 Hexane/EtOAc to 9/1 EtOAc/MeOH with 0.1 % acetic acid gradients) afforded **9** (1.07 g, 65 %) as a white solid. ^{1}H-NMR (300 MHz, CD_3OD) δ 7.78 (d, *J* = 7.38 Hz, 2H), 7.57 (d, *J* = 7.41 Hz, 2H), 7.38 (t, *J* = 7.56 Hz, 4H), 7.26-7.31 (m, 2H), 7.11 (d, *J* = 8.04 Hz, 2H), 4.13-4.41 (m, 4H), 3.42 (s, 1H), 2.90-3.24 (m, 2H). ^{13}C-NMR (75 MHz, DMSO-d_6) δ 174.8, 158.3, 145.2, 142.5, 139.6, 133.0, 128.7, 128.1, 126.1, 122.1, 120.9, 84.3, 78.5, 67.9, 56.5, 48.3, 38.3, 26.3. IT-TOF: m/z [M+Na] $^{+}$ calcd: 434.15, found: 434.12.



Allyl 2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-ethynylphenyl) propanoate (10)

The compound **9** (0.75 g, 1.8 mmol) was dissolved in MeOH (10 mL). Cs_2CO_3 (0.28 g, 0.85 mmol) was added over a period of 10 min. Subsequently, MeOH was removed *in vacuo*. The residue was redissolved in DMF (20 mL). Allyl bromide (0.68 g, 5.6 mmol) was added. After 1 hour, DMF was distilled off under reduced pressure. Water was added and the solution was extracted with EtOAc. The combined organic phase was washed with brine, dried over Na_2SO_4 and concentrated. The crude product was purified by silica gel chromatography (EtOAc/Hexane gradients). The white product **10** was obtained in 90 % yield. ^{1}H-NMR (300 MHz, $CDCl_3$) δ 7.78 (d, *J* = 7.41 Hz, 2H), 7.57 (d, *J* = 7.23 Hz, 2H), 7.39-7.43 (m, 4H), 7.32-7.34 (t, *J* = 7.40 Hz, 2H), 7.06 (d, *J* = 7.86 Hz, 2H), 5.83-5.88 (m, 1H), 5.26-5.35 (m, 3H), 4.61-4.72 (m, 3H), 4.35-4.50 (m, 2H), 4.21 (t, *J* = 6.82 Hz, 1H), 3.08-3.15 (m, 3H). ^{13}C-NMR (75 MHz, $CDCl_3$) δ 170.9, 155.5, 143.8, 141.3, 136.6, 132.3, 131.2, 129.4, 127.7, 127.0, 125.0, 121.0, 120.0, 119.3, 77.3, 83.3, 66.9, 66.2, 54.6, 47.2, 38.1. IT-TOF: m/z [M+Na] $^{+}$ calcd: 474.18, found: 474.15.



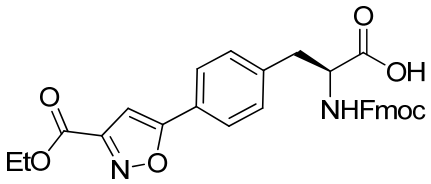
Ethyl 5-(4-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(allyloxy)-3-oxo -propyl)phenyl)isoxazole-3-carboxylate (11) [2]

Method 1: Compound **10** (0.1 g, 0.22 mmol) and ethyl chlorooximidoacetate (33 mg, 0.22 mmol) were dissolved in *t*-BuOH (5 mL). To the resulting solution, H_2O was

added, followed by sodium ascorbate (4.35 mg, 0.02 mmol), copper sulfate (2.75 mg, 0.01 mmol) and finally $KHCO_3$ (90 mg, 0.9 mmol). The reaction mixture was stirred at room temperature for 24 hours. Upon evaporation of *t*-BuOH, the reaction was extract several times with EtOAc. The organic layers were then collected, dried over anhydrous Na_2SO_4 and filtered, and then the solvent was removed under reduced pressure to yield a crude product. The product was purified by flash chromatography to yield **11** (21 %) as a white solid.

Method 2: Compound **10** (0.5 g 1.1 mmol) and ethyl chlorooximidoacetate (0.25 g, 1.6 mmol) were dissolved in THF (20 mL). Triethylamine (0.22 g, 2.2 mmol) was slowly added into solution. The mixture was stirred for 1 day. THF and TEA was evaporated and the residue redissolved in EtOAc. After washing with brine and drying over Na_2SO_4, the organic layer was evaporated. The crude product was purified by flash chromatography to obtain **11** (15%).

^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.60 Hz, 2H), 7.70 (d, *J* = 7.60 Hz, 2H), 7.55-7.57 (m, 2H), 7.40 (t, *J* = 7.40 Hz, 2H), 7.30 (t, *J* = 7.55 Hz, 2H), 7.21 (d, *J* = 8.20 Hz, 2H), 5.85-5.91 (m, 1H), 5.26-5.38 (m, 3H), 4.70-4.74 (m, 1H), 4.63 (d, *J* = 5.7 Hz, 1H), 4.47 (q, *J* = 7.55 Hz, 2H), 4.35-4.39 (m, 1H), 4.19 (t, *J* = 6.95 Hz, 1H), 3.18 (dd, *J* = 13.87, 5.65 Hz, 2H), 1.45 (t, *J* = 6.96 Hz, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$) δ 171.1, 170.8, 159.9, 156.9, 155.5, 143.7, 141.3, 138.9, 131.1, 130.1, 127.7, 127.0, 126.0, 125.4, 125.0, 120.0, 119.4, 99.8, 66.8, 66.2, 62.2, 54.6, 47.1, 38.1, 14.1. IT-TOF: m/z $[M+H]^+$ calcd: 567.21, found: 567.20.



Chemical Formula: $C_{30}H_{26}N_2O_7$

2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-(3-(ethoxycarbonyl)isoxazol-5-yl)phenyl)propanoic acid (6') [3]

Compound **11** (20 mg, 0.03 mmol) was dissolved in DCM under an N_2 atmosphere. Phenylsilane (8 mg, 0.07 mmol) and $Pd(PPh_3)_4$ (1.7 mg, 1.5 μmol) were added. After 30 min, the solution became slightly dark and TLC was showed the reaction was finished. The solvent was removed under reduced pressure and the residue was directly purified by flash chromatography to obtain compound **6'** (89 %). ^{1}H-NMR (300 MHz, DMSO-d_6) δ 13.1 (br, 1H), 7.86 (d, *J* = 7.23 Hz, 4H), 7.61-7.68 (m, 3H), 7.37-7.45 (m, 5H), 7.25-7.31 (m, 2H), 4.39 (q, *J* = 6.9 Hz, 2H), 4.17-4.19 (m, 4H), 2.91-3.19 (m, 2H), 1.35 (t, *J* = 7.05 Hz, 3H). ^{13}C-NMR (75 MHz, DMSO-d_6) δ 173.1, 171.1, 159.4, 156.8, 155.9, 143.7, 141.4, 140.7, 130.1, 127.6, 127.0, 125.7, 125.2, 124.3, 120.1, 100.4, 65.6, 61.2, 55.2, 46.5, 36.3, 14.0. HRMS calcd for $C_{30}H_{25}N_2O_7$: 525.1662, found: 525.1647.

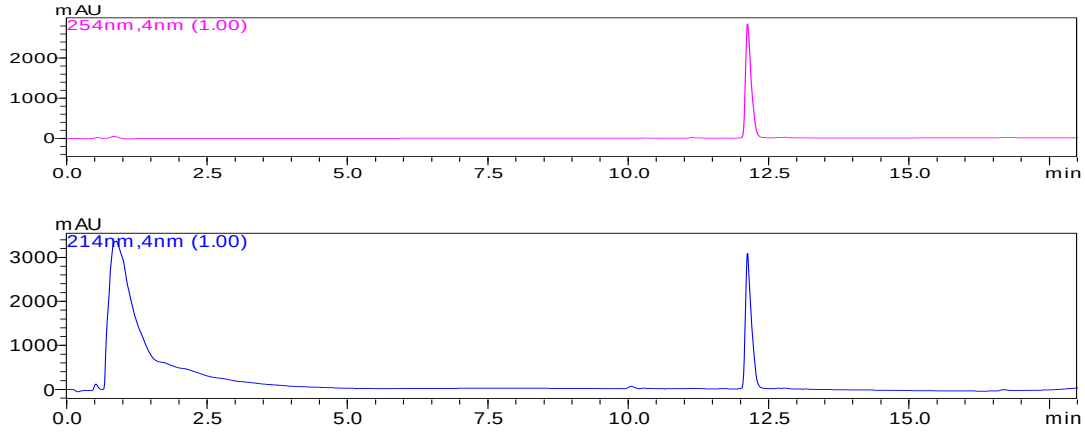
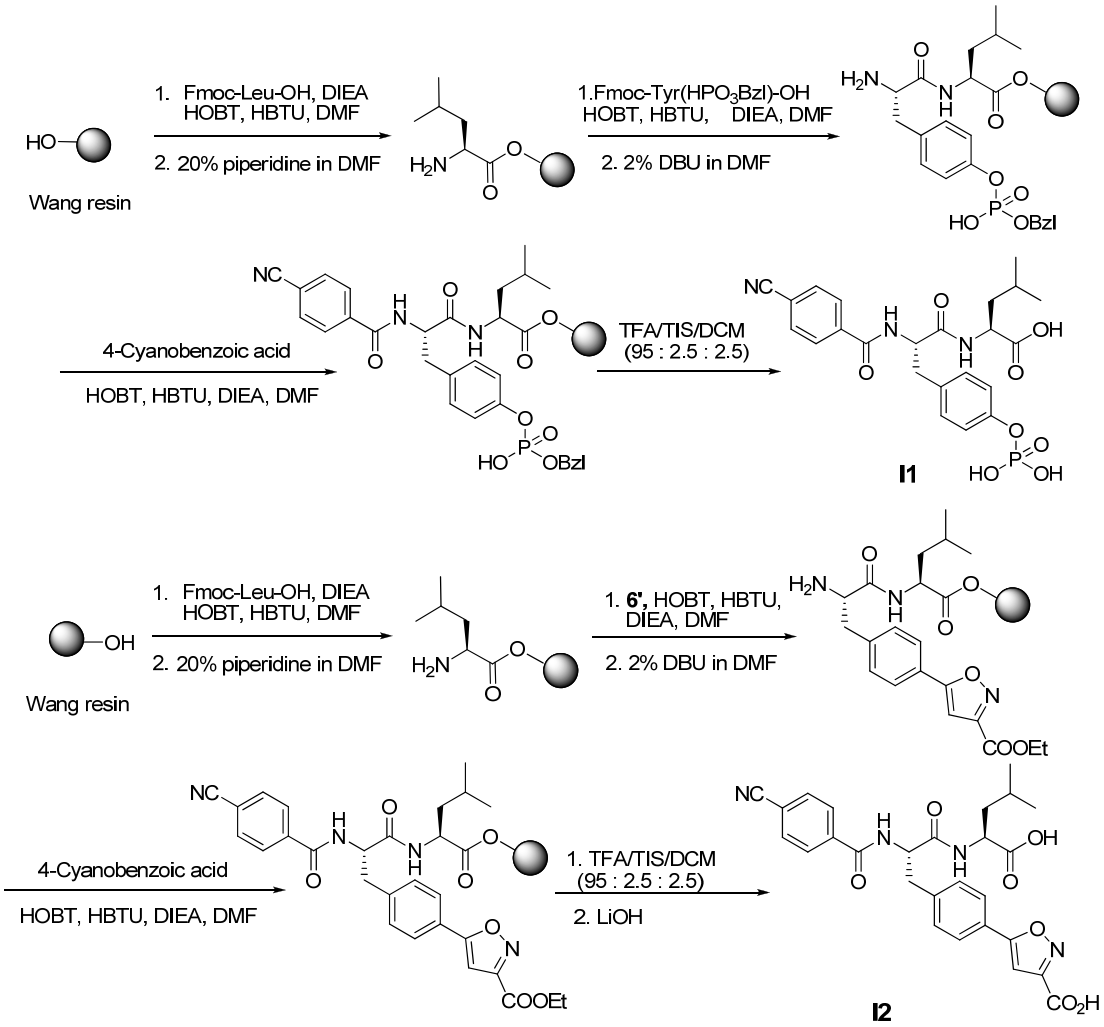


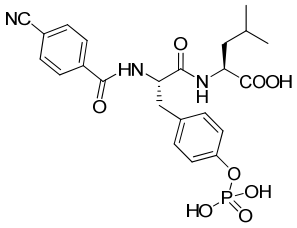
Figure S1. The analytical HPLC results of compound 6'

3. Procedure and LC-MS profiles of inhibitors and SH2-binding peptide

3.1 Synthesis of inhibitors



Scheme S1: Solid phase synthesis of two STAT3 PPI inhibitors.



Exact Mass: 503.15

Inhibitor **I1**. Wang resin was used to synthesize the inhibitors. First, Wang resin (100 mg, 0.05 mmol, and loading ~0.5 mmol/g) was previously swollen in DMF. Fmoc-Leu-OH (4 equiv) was pre-activated using HOBT (4 equiv), HBTU (4 equiv), DIEA (8 equiv) in DMF. And the resins were added into the solution. The reaction was shaken for 4 hours at room temperature. After that, the resin was washed with DMF (3×), DCM (3×) and DMF (3×). Next, Fmoc-protected amino functionalized resin was treated with 20% piperidine in DMF for 1 hour at room temperature. The rein was washed with DMF (3×), DCM (3×) and DMF (3×). Fmoc-Tyr (HPO_3Bzl)-OH were pre-activated by HOBT (4 equiv), HBTU (4 equiv), DIEA (8 equiv) in DMF. The resin was added into the solution and shaken at room temperature overnight. Subsequently, the resin was washed with DMF (3×), DCM (3×) and DMF (3×). Deprotection of the Fmoc group was carried out with 2% DBU in DMF for 2 hours [4]. And the resin was washed again. 4-Cyanobenzoic acid was coupled with the resin using HOBT (4 equiv), HBTU (4 equiv), and DIEA (8 equiv) in DMF overnight. The resin was then washed with DCM and MeOH before being dried *in vacuo*. The cocktail cleavage solution contains TFA/TIS/DCM (95:2.5:2.5). The resin was treated with the cocktail for 2.5 hours, after which the resin was filtered off and the solvent was removed under reduced pressure. The resulting inhibitor was further purified by preparative HPLC.

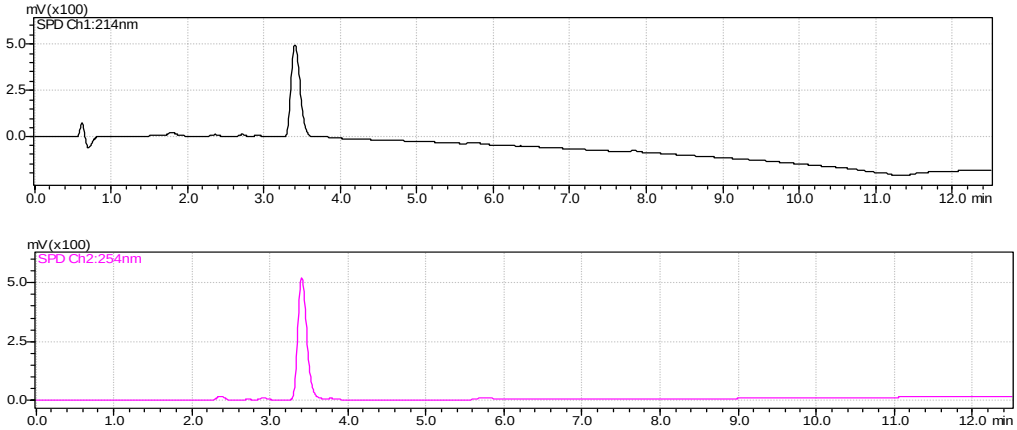
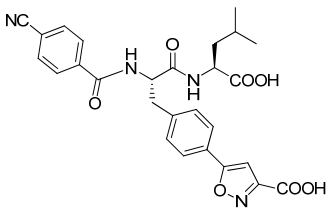


Figure S2. The LCMS results of compound **I1**



Chemical Formula: $C_{27}H_{26}N_4O_7$
Exact Mass: 518.1801

Inhibitor **I2**. The inhibitor **I2** was synthesized similarly as **I1** following procedures as shown above. The unnatural amino acid, compound **6'**, was coupled to the resin as follows. First, **6'** was pre-activated by HOBT (4 equiv), HBTU (4 equiv), DIEA (8 equiv) in DMF. Then resin was added into the solution and the resulting mixture was shaken at room temperature overnight. Subsequently, the resin was washed with DMF (3×), DCM (3×) and DMF (3×). The deprotection of Fmoc group was carried on 2 % DBU in DMF for 2 hours [4]. And the resin was washed again. 4-Cyanobenzoic acid was coupled with the resin using HOBT (4 equiv), HBTU (4 equiv), and DIEA (8 equiv) in DMF overnight. The resin was washed, after coupling, with DMF (3×), DCM (3×) and DMF (3×). At the end of solid phase synthesis, the resin was washed with DCM and MeOH before being dried *in vacuo*. The inhibitor was cleaved off the resin by treatment with the cocktail mixture, TFA/TIS/DCM (95:2.5:2.5), for 2.5 hours. After that, the resin was filtered off and the solvent was removed under reduced pressure to give the protected form of the final product (precursor). The final product was subsequently obtained by treating the precursor with LiOH (2 M) in THF for 2 hours followed by purification with preparative HPLC. ^{1}H-NMR (500 MHz, DMSO-d_6) δ 8.90 (d, J = 8.85 Hz, 1H), 8.42 (d, J = 8.20 Hz, 1H), 7.90-7.92 (m, 4H), 7.84 (d, J = 8.20 Hz, 1H), 7.54 (d, J = 8.85 Hz, 2H), 7.34 (s, 1H), 4.81-4.84 (m, 1H), 4.25-4.29 (m, 1H), 3.01-3.25 (m, 2H), 1.60-1.69 (m, 1H), 1.53-1.59(m, 2H), 0.89 (dd, J = 14.80, 6.30 Hz, 6H) ^{13}C-NMR (125 MHz, DMSO-d_6) δ 173.9, 171.0, 164.9, 160.8, 157.7, 141.4, 139.5, 137.9, 132.3, 130.0, 128.2, 125.6, 124.4, 118.3, 113.7, 100.4, 53.4, 50.4, 36.9, 24.3, 22.8, 21.3. HRMS calcd $C_{27}H_{25}N_4O_7$: m/z $[M+H]^+$ calcd: 517.1724, found: 517.1721.

Based on the ^{1}H NMR result, we found no sign of racemization in the final product **I2**, indicating that during the synthesis of the new amino acid **6'** and the final inhibitor **I2**, the two steps where base treatments were involved (**8** to **9** in Scheme 1 and the LiOH treatment of the cleaved peptide giving **I2** in Scheme 2) did not cause noticeable racemization.

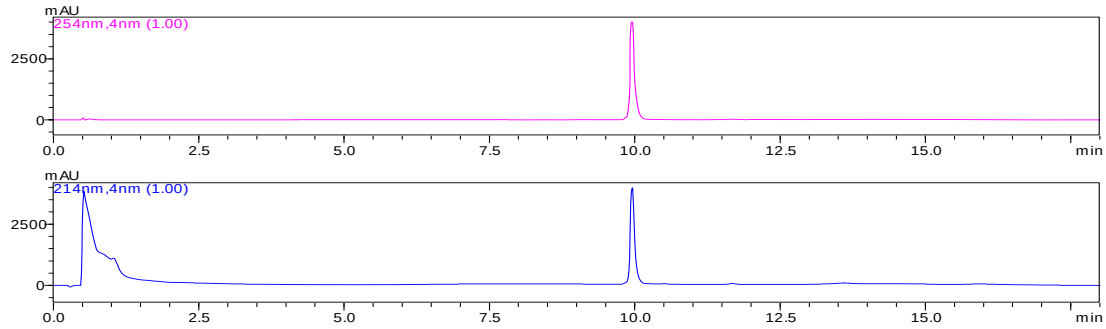


Figure S3. The analytical HPLC results of compound **I2**

3.2 Synthesis of SH2-binding peptide Fluorescein- GY*LPQTV-NH_2

Rink amide resin (100 mg, loading ~ 0.5 mmol/g) was first swollen in DMF and washed with DMF (2×). Before coupling, 20% piperidine in DMF was used to deprotect the Fmoc group for 1 h at RT. Subsequently, the resin was washed with several cylces of DMF and DCM. Fmoc-Val-OH (4 equiv) was pre-activated using HOBT (4 equiv), HBTU (4 equiv), DIEA (8 equiv) in DMF. The resin was put into the solution and shaken for 4 hours. Next, the resin was washed with several cycles of DMF and DCM and treated with 20% piperidine to remove the Fmoc group prior to the next coupling.

The cycle of coupling was repeated with Fmoc-Thr(OtBu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Pro-OH, Fmoc-Leu-OH, Fmoc-Tyr(HPO_3Bzl)-OH, Fmoc-Gly-OH. Final, the resin was coupled with 3', 6'-diacetyl fluorescein (4 equiv). After washing, the acetyl protecting group was removed by treatment with 20% piperidine in DMF. Finally, the resin was washed with DMF (3×), DCM (3×) and MeOH (3×) and dried thoroughly under reduced pressure. A solution of cleavage solution TFA/TIS/DCM (95:2.5:2.5) was added to the resin for 2 hours. The resin was filtered off. The solvent was removed *in vacuo*. The peptide was purified by preparative HPLC.

4. Procedures and results of cell proliferation

T47D cancer cell line was from the National Cancer Institute Developmental Therapeutics Program (NCI60 cell line panel). Cells were cultivated in a T25 flask in RPMI (Invitrogen, Carlsbad, CA) medium with 10% Fetal Bovine Serum (FBS, Gibco Invitrogen) and 100 U/mL Penicillin-Streptomycin (Thermo Scientific, Rockford, IL) and maintained in a humidified 37 ^oC incubator with 5 % CO_2. Subsequently, trypsin with EDTA were used to detach cells detach from the flask. Cells were counted using Hemocytometer and seeded equally into each well of a 96-well plate. Each well contains around 3000~4000 cells in culture medium with 10% FBS. After overnight culture, the medium was aspirated and then treated with different concentrations of **I1** and **I2** in DMSO (500 μM, 250 μM, 125 μM). Staurosporine (STS) was used as a positive control. The same volume of DMSO (final concentration in each well is 0.56%) was used as a negative control. After 2 days, the tetrazolium salt XTT solution (0.2 mg/mL XTT, 25 M phenazine methosulfate (PMS), Roche) was added to the cells. After incubating at 37 ^oC for 20 hours, the absorbance was measured at 450 nm and background absorbance was measured at 650 nm using a Bioteck™ plate reader. Data was presented as the following:

$$viability = \frac{A_{450nm} - A_{650nm}}{a_{450nm} - a_{650nm}}$$

A_{450nm} and A_{650nm} are the absorbance at 450 nm and the background absorbance at 650 nm of the sample wells; a_{450nm} and a_{650nm} are the absorbance at 450nm and the background absorbance at 650nm of the control wells (which only contained DMSO).

5. Procedure and results of fluorescence polarization (FP) experiments

The clone containing the mammalian STAT3 SH2 domain was purchased from Open Biosystems (USA). The construct contains the SH2 domain cloned into a modified pET28 bacterial expression vector. A 6xHis tag is encoded for expression at the N-terminus which gives a final 68 KD fusion protein. The protein was expressed and purified according to protocols provided by the vendor.[5] Figure S1 shows the coomassie gel of fractions containing the purified protein (shown as a band at the expected 68KD position). Only sufficiently pure fractions were collected and used for

the following FP experiments. Fluorescence polarization (FP) experiments were carried out following published procedures.[6] Different concentrations of the inhibitor (1 mM - 7.8 μM) were first incubated with the protein (150 nM) for 30 min. Subsequently, the fluorescently labeled STAT3 peptide (Fluorescein-GY*LPQTV-NH_2) was then added into each well (final conc: 10 nM). The fluorescence readings were then taken with a fluorescence plate reader (Tecan, USA) installed with 2 pairs of polarizers (λ_{ex}: 485 ± 20 nm, λ_{em}: 520 ± 10 nm). From the data, each IC_{50} plot was generated by averaging the duplicates obtained from two independent assays.

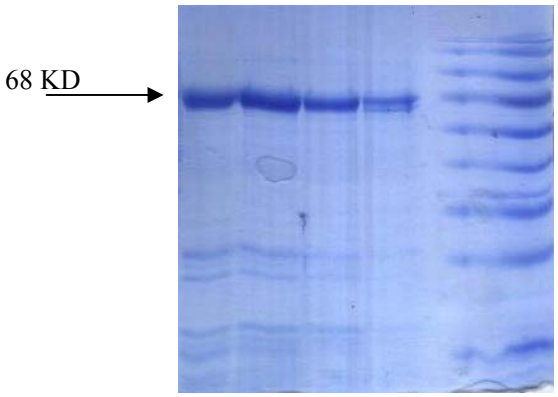


Figure S4. Representative coomassie gel of the purified STAT3 SH2 domain (From left to right, different elution factions 1-4 and MW ladder). Only sufficiently pure fractions were pooled for subsequent FP experiments.

5.1 Procedure and results of inhibition assays of PTP1B

The inhibition of I2 to PTP1B was assayed by measuring the rate of hydrolysis of the fluorogenic substrate, 6,8-difluoromethylumbellifery phosphate (DIFMUP, Invitrogen, USA) in 25 μL reaction volumes in black polypropylene flat-bottom 384-well microtiter plates (Greiner, Germany). Briefly, a two-fold dilution series of **I2**, from approximately 2 mM to 125 μL (final concentrations) was prepared. PTP1B's concentration is the same as STAT3 SH2 domain, 150 nM. The substrate concentration is 25 μM. The results are shown in Figure S5.

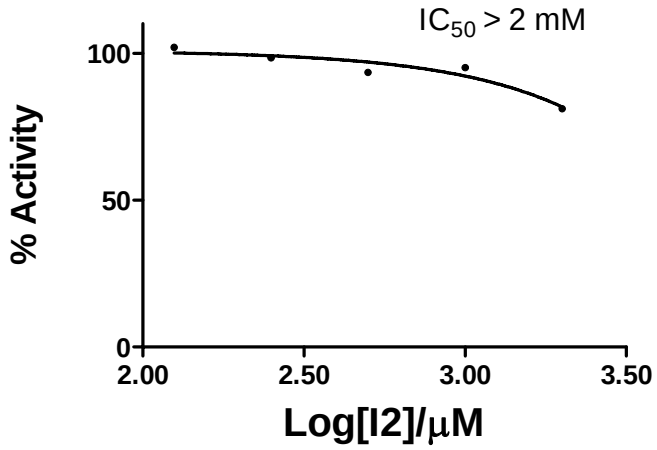


Figure S5. IC_{50} graphs of I2 against PTP1B

6. Procedure and results of docking stimulations

Docking was carried out on an SGI IRIX 6.5 workstation using the SYBYL7.2 suite installed with the FlexX docking software. The mol2 format of the two inhibitors **I1** and **I2** was prepared by Sybyl v7.2 (Tripos, Missouri, USA). STAT3 were obtained from the Protein Data Bank, entry 1BG1 [7]. The docking sphere was set at 10 Å and centered on the Arg 609 residue as binding site. The docking was performed for 100 iteractions and Figure S6 shows the most preferred conformations.

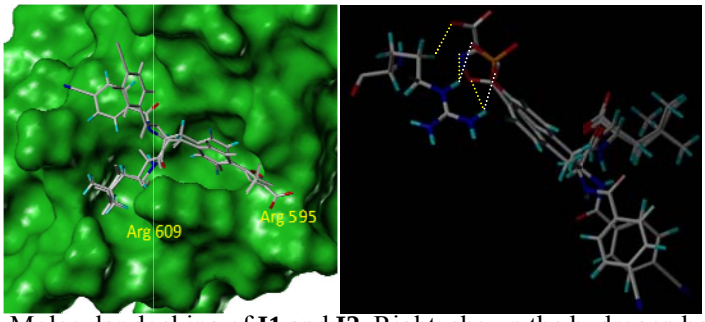
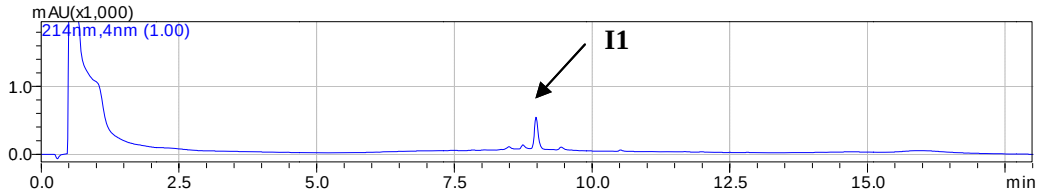


Figure S6. Left: Molecular docking of **I1** and **I2**. Right: shown the hydrogen bond between Arg595 and phosphate (**I1**, white dash bond) , isoxazole carboxylic acid (**I2**, yellow dash) heads.

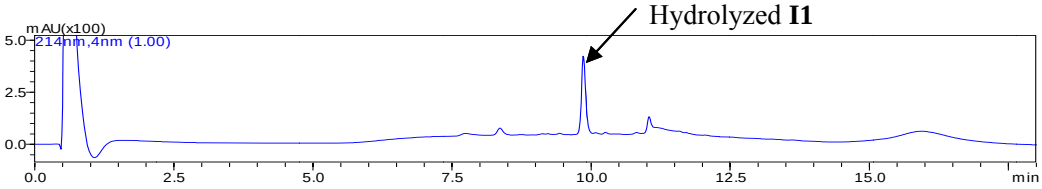
7. Procedure and results of hydro lytic stability test

First, T47D cell lysate was prepared in PBS, *p*H 7.5 by disrupting their cell membranes using needle and syringe. 1 µL of 5 mM **I1** and **I2** in DMSO was added into 100 µL 2 mg/mL cell lysate, respectively. After 15 hours, the lysate was pipetted into Ulracel YM-3 (Millipore) and centrifuge (13000 rpm, 50 min) to get the filtrate and remove most of proteins (larger than 3kD). Then, the composition of filtrate was analyzed by LCMS (Figure S7); **I1** was completely hydrolyzed in 15 h, while most of **I2** remained intake under the same conditions, indicating much more superior hydrolytic stability of **I2**.

a) untreated I1



I1 treated after 15 hrs



b) untreated **I2**

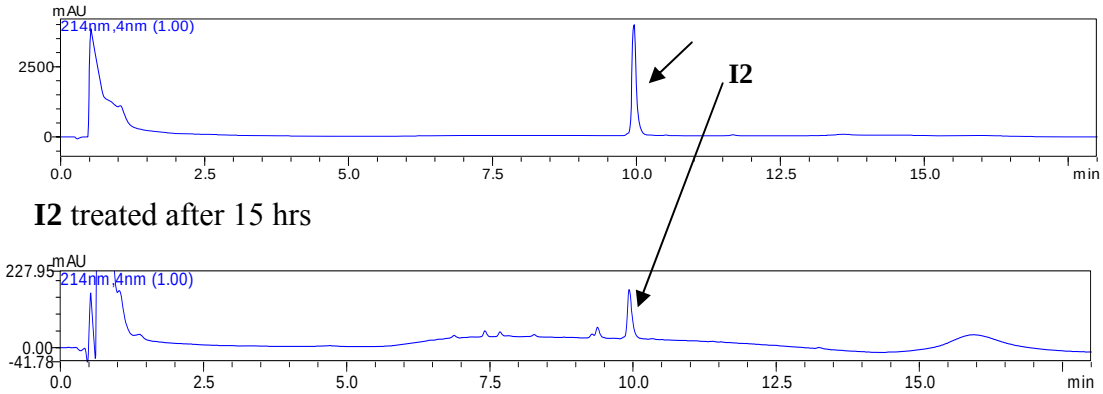


Figure S7. a) LC profiles of untreated (top) and treated I1 (bottom). b) LC profiles of untreated (top) and treated **I2** (bottom).

8. Procedure and results of cell permibility test

MDCK (Madin-Darby canine kidney) cells were used for testing the cell permeability. MDCK was seeded with 600,000~700,000 cells per cm^2 (0.11 cm^2 per well insert, Millipore® #PSHT004R1) and cultured for 3 days before test. When cultures are confluent, remove media and rinse inserts with Hanks' Balanced Salt Solution (HBSS, Gibco® #14025). Transport assay donor solutions consisted of 50 µM **I2** and caffeine in transport medium containing 60 µM Lucifer Yellow (LY) and 1% DMSO. Transport assays were conducted using 75 µL of apical (AP) donor solution and 235 µL of basolateral (BL) acceptor solution (transport medium, pH 7.4, following manufacture descriptions of Millipore® #PSHT004R1). Monolayer was incubated with donor and acceptor solutions for 60 min at 37 °C, 95% humidity. After 60 min, the donor and acceptor solutions of each compound are collected and quantified by HPLC. Lucifer Yellow (LY) was quantified using a fluorescence 96-well plate reader (BioTek® Synergy 4 fluorescence plate reader at Ex = 485 nm and Em = 539 nm). P_{app} (apparent permeability) values were calculated according to the following equation:

$$P_{app} = \left(\frac{dQ}{dt}\right) \times \frac{1}{C_0} \times \frac{1}{A}$$

where dQ/dt is the permeability rate, C_0 is the initial concentration in the donor compartment, and A is the surface area of the filter. Permeability rates were calculated by plotting the percent of initial AP drug mass (peak area) found in the BL compartment versus time and determining the slope of the line. Lucifer yellow (LY) results were used as an internal control for each monolayer to verify tight junction integrity during the entire assay period. Accordingly, LY P_{app} values were quantified from 60 min basolateral samples after background subtraction. As a quality control, results from MDCK monolayer with LY *Papp* > 30 nm/s were not used. The results are shown in Table S1, indicating I2 is indeed reasonable cell-permeable.

Table S1. The results of cell permeability.

	Positive control	Negative control	**I2**
Caffeine / 50 μM	+	-	-
Lucifer Yellow / 60 μM	+	+	+
Compound / 50 μM	-	-	+
DMSO	1%	1%	1%
P_{app} / nm·s^{-1}	1217	n.d.	876

n. d. : not determined.

10. References

[1] Lei, H. Y.; Stoakers, M. S.; Herath, K. P. B.; Lee, J. H.; Schwabacher, A. W.; *J. Org. Chem.*, **1994**, *59*, 4206-4210

[2] Soellner, M. B.; Rawls, K. A.; Grundner, C.; Alber, T.; Ellman, J. A.; *J. Am. Chem. Soc.*, **2007**, *129*, 9613-9615

[3] Vazquez, M. E; Blanco, J. B.; Imperiali, B.; *J. Am. Chem. Soc.*, **2005**, *127*, 1300-1306

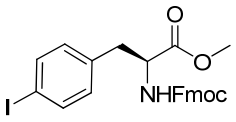
[4] Turkson, J.; Kim, J. S.; Zhang, S.; Yuan, J.; Huang, M.; Glenn, M.; Haura, E.; Sebti, S.; Hamilton, A.; Jove, R.; *Mol. Cancer. Ther.*, **2004**, *3*, 261-269

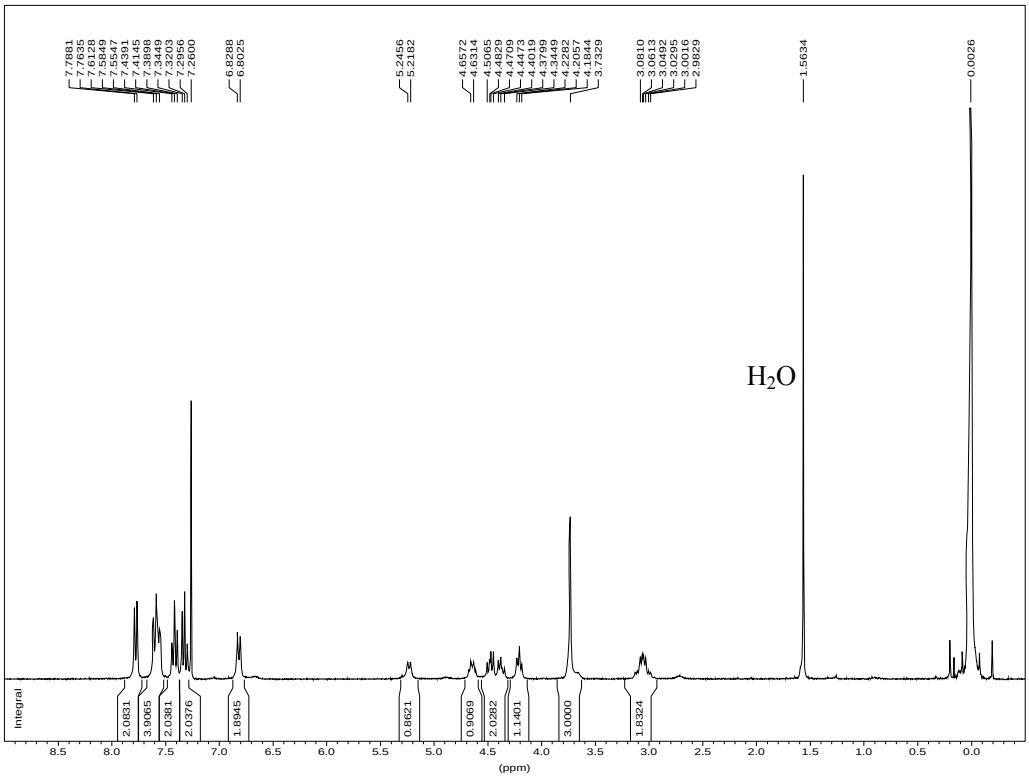
[5] (a) Lu, C. H. S.; Sun, H.; Bakar, F. B. A.; Uttamchandani, M.; Zhou, W.; Liou, Y.-C.; Yao, S. Q.; *Angew. Chem. Intl. Ed.*, **2008**, *47*, 7438-7441. (b) http://www.openbiosystems.com

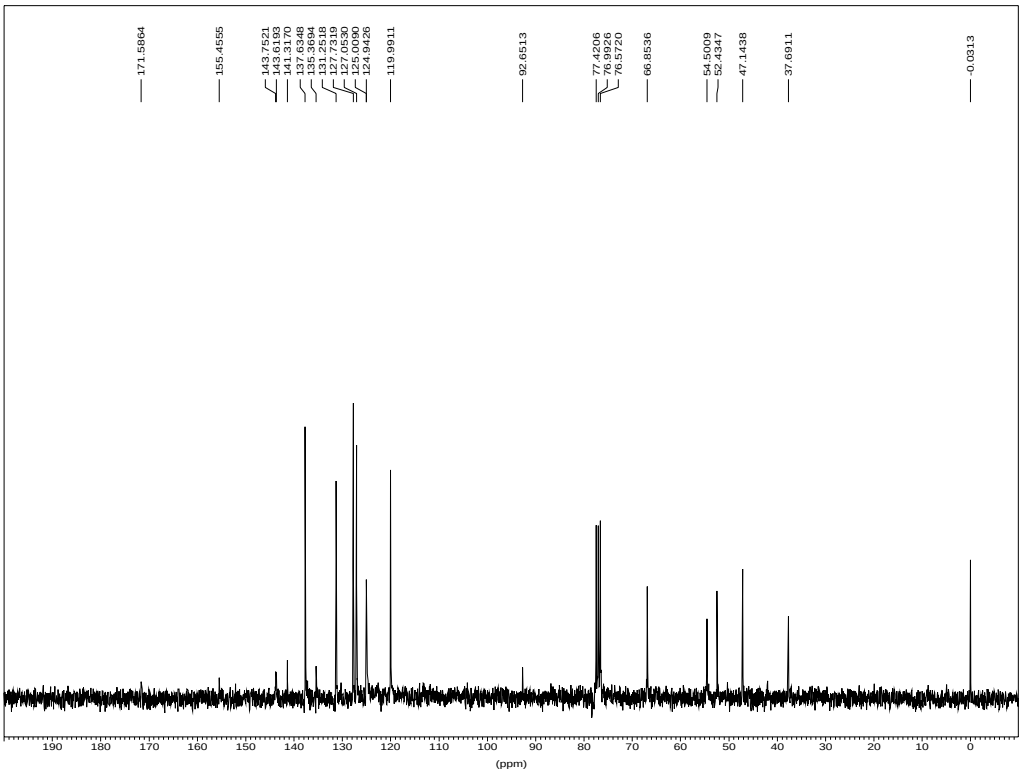
[6] Schust,J.; Berg, T.; *Anal. Biochem.*, **2004**, *330*, 114-118

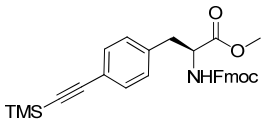
[7] (a) Becker, S.; Groner, B.; Muller, C. W. *Nature*, **1998**, *394*, 145-151. (b) Coleman. IV, D.R.; Ren, Z.; Mandal, P. K.; Cameron, A. G.; Dyer, G. A.; Muranjan, S.; Campell, M.; Chen, X. McMurray. J. S. *J. Med. Chem.,* **2005**, *48*, 6661-6670

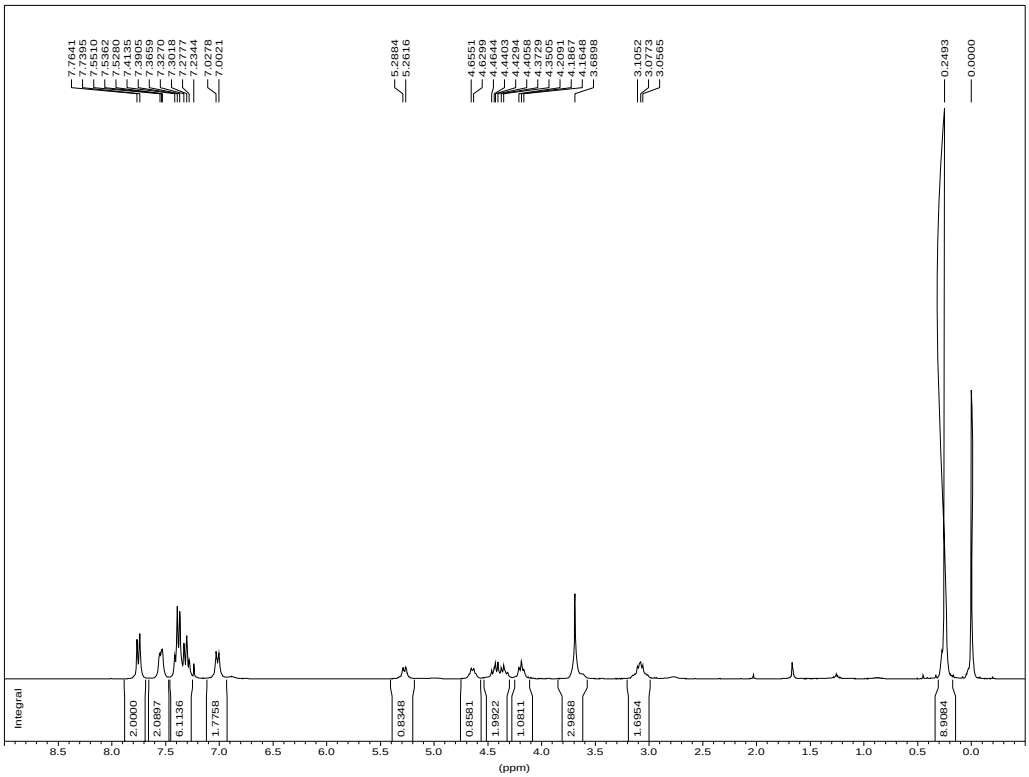
[8] Irvine,J. D.; Takahashi, L.; Lockhart, K.; Cheong, J.; Tolan, J. W.; Selick, H. E.; Grove, J. R.. *J. Pharm. Sci.,* **1999**, *88*, 28-33

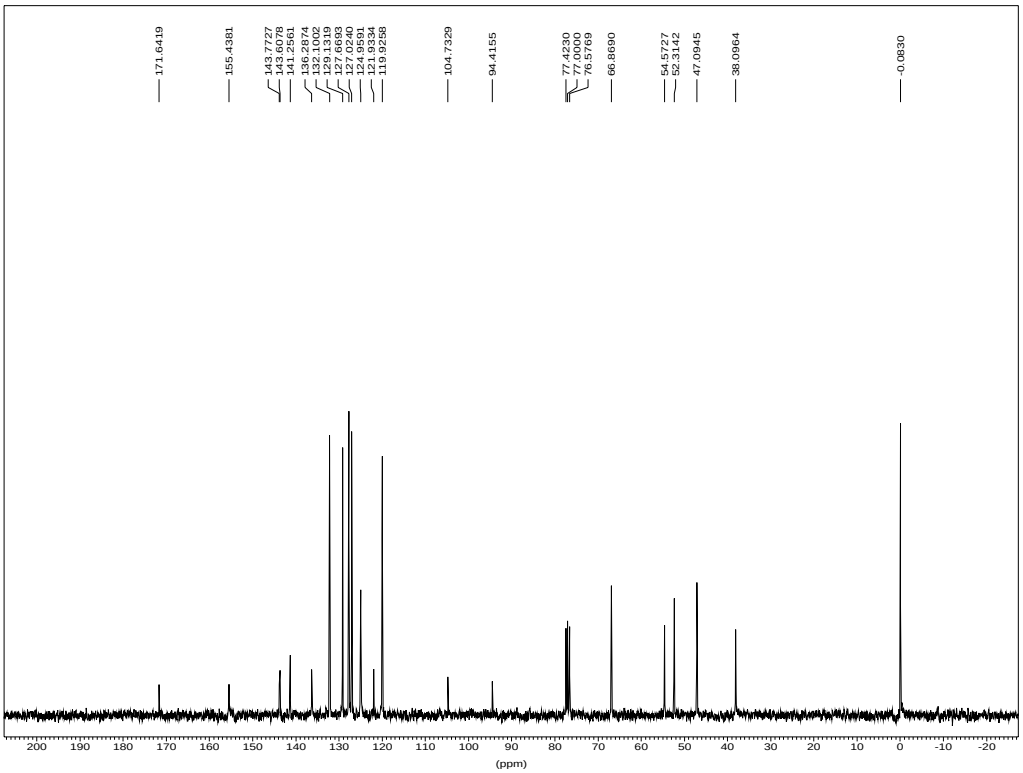

O
O
NHFmoc
I

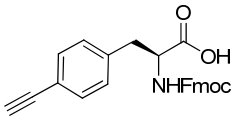

7.7881
7.7635
7.6128
7.5849
7.5547
7.4401
7.4145
7.3898
7.3449
7.3203
7.2966
7.2600
6.8288
6.8025
5.2456
5.2182
4.6572
4.6314
4.5065
4.4829
4.4709
4.4473
4.4219
4.3999
4.3449
4.2282
4.2057
4.1814
3.7329
3.0810
3.0613
3.0492
3.0295
3.0056
2.9829
1.5634
0.0026
H2O
Integral
2.0831
3.9065
2.0381
2.0376
1.8945
0.8621
0.9069
2.0282
1.1401
3.0000
1.8324
(ppm)

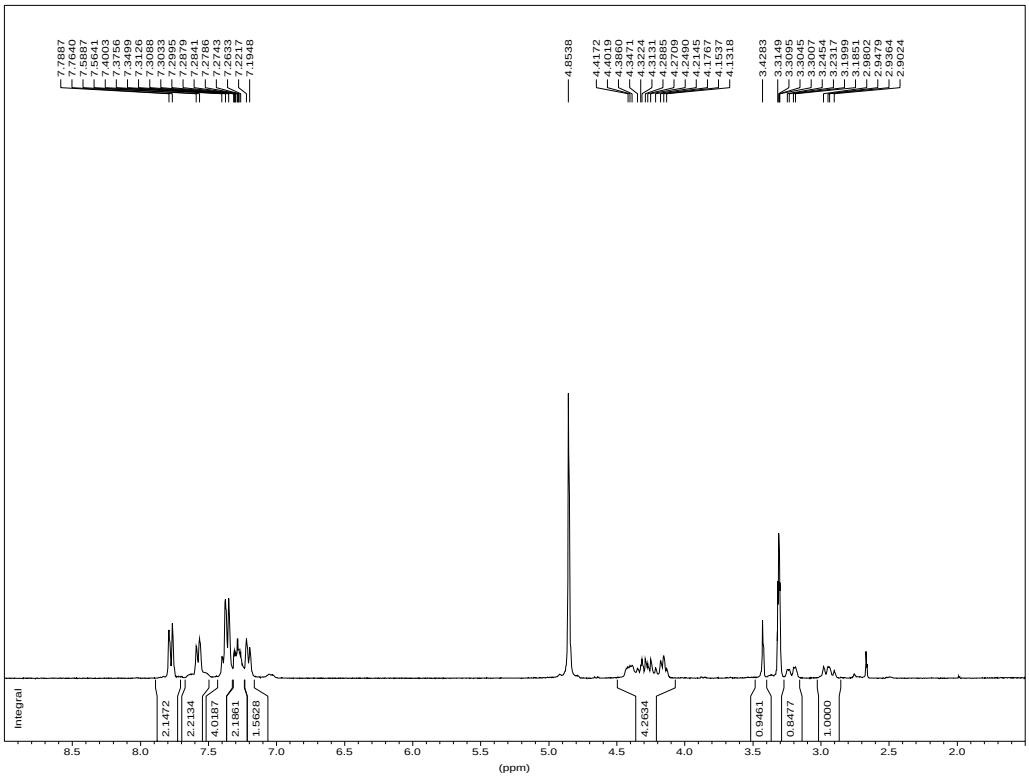

171.5864
155.4555
143.7521
143.6193
141.3170
137.6348
135.3694
131.2518
127.7210
127.0530
125.0090
124.9426
119.9911
92.6513
77.4206
76.9926
76.5720
66.8536
54.5009
52.4347
47.1438
37.6911
-0.0313
(ppm)

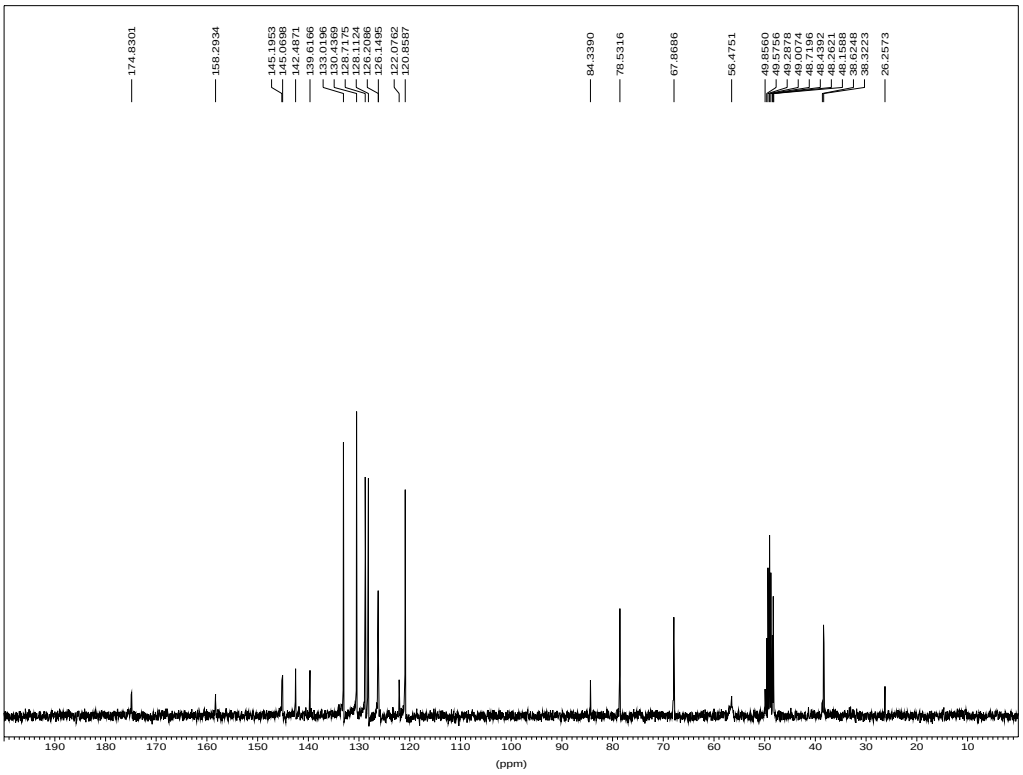

O
O
NHFmoc
TMS

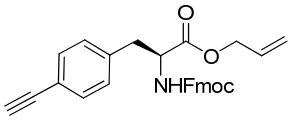

7.7641
7.7295
7.5510
7.5362
7.5280
7.4135
7.3985
7.3659
7.3270
7.3018
7.2777
7.2344
7.0278
7.0021
5.2884
5.2616
4.6551
4.6299
4.4644
4.4403
4.4204
4.4058
4.3729
4.3505
4.2091
4.1897
4.1648
3.6898
3.1052
3.0773
3.0565
0.2493
0.0000
Integral
2.0000
2.0897
6.1136
1.7758
0.8348
0.8581
1.9922
1.0811
2.9868
1.6954
8.9084
(ppm)

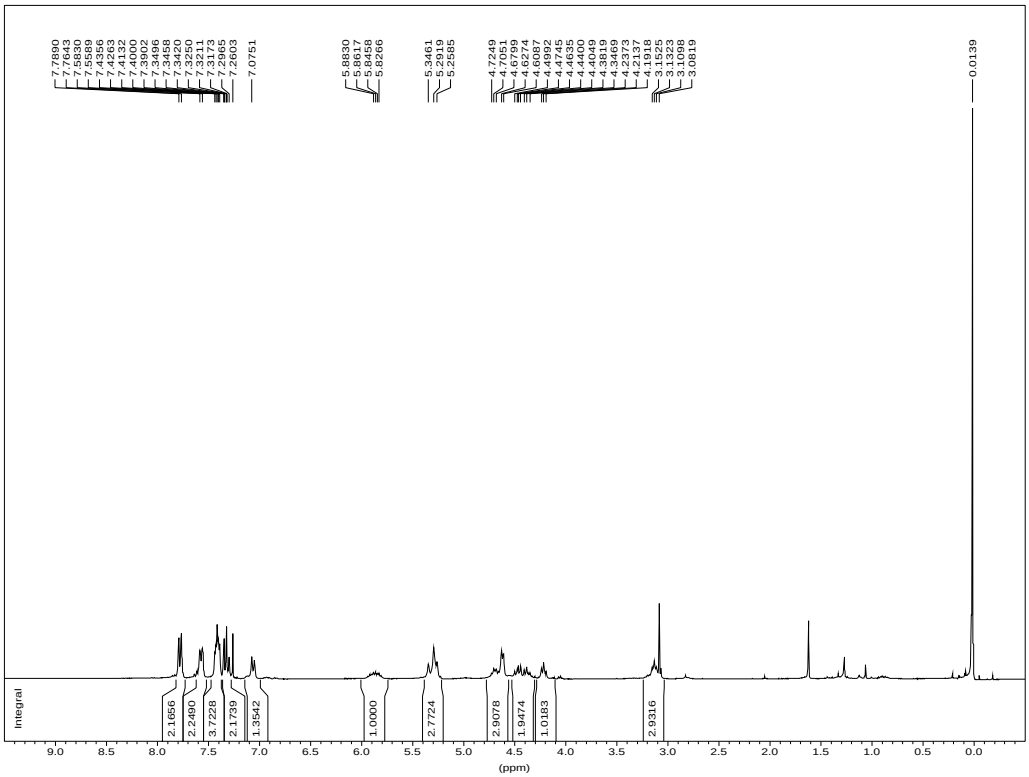

171.6419
155.4381
143.7727
143.6078
141.2561
136.2874
132.1002
129.1319
127.6693
127.0240
124.9591
121.9334
119.9258
104.7329
94.4155
77.4230
77.0000
76.5769
66.8690
54.5727
52.3142
47.0945
38.0964
-0.0830
(ppm)

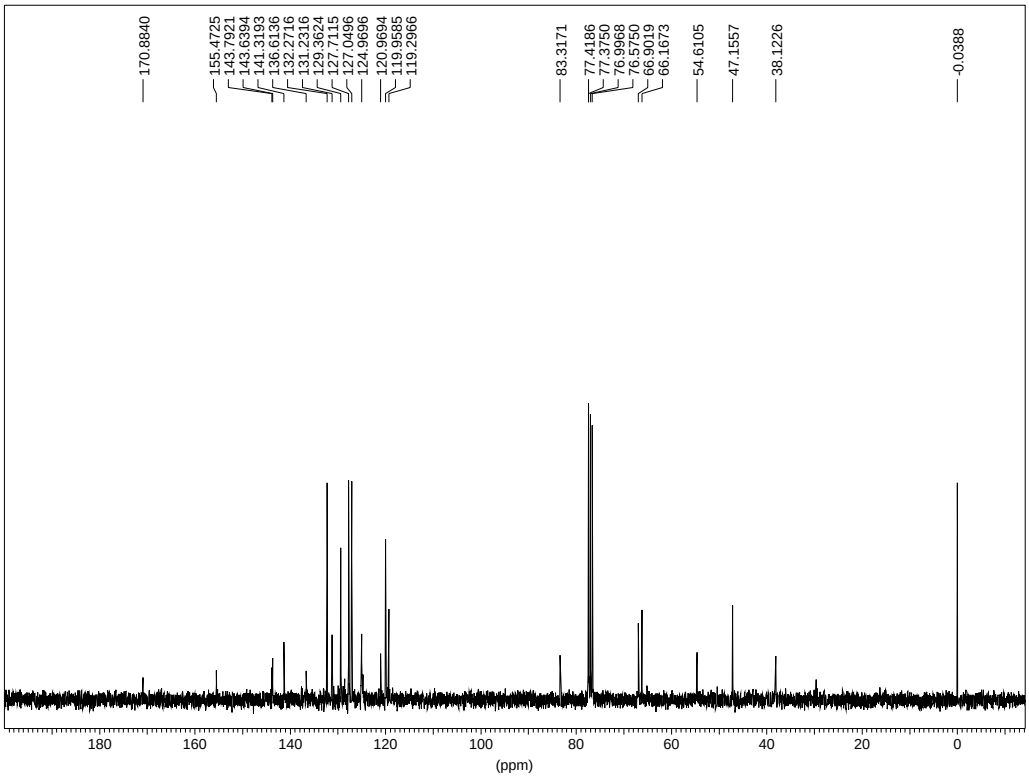

O
OH
NHFmoc

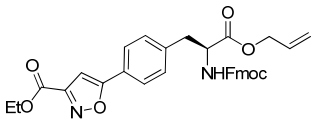

7.7887
7.7640
7.5687
7.5641
7.4003
7.3756
7.3499
7.3126
7.3088
7.3033
7.2995
7.2879
7.2841
7.2786
7.2743
7.2633
7.2217
7.1948
4.8538
4.4172
4.4019
4.3860
4.3471
4.3274
4.3131
4.2885
4.2709
4.2500
4.2145
4.1767
4.1537
4.1318
3.4283
3.3149
3.3095
3.3045
3.3007
3.2454
3.2317
3.1999
3.1851
2.9922
2.9479
2.9364
2.9024
Integral
2.1472
2.2134
4.0187
2.1861
1.5628
4.2634
0.9461
0.8477
1.0000
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
(ppm)

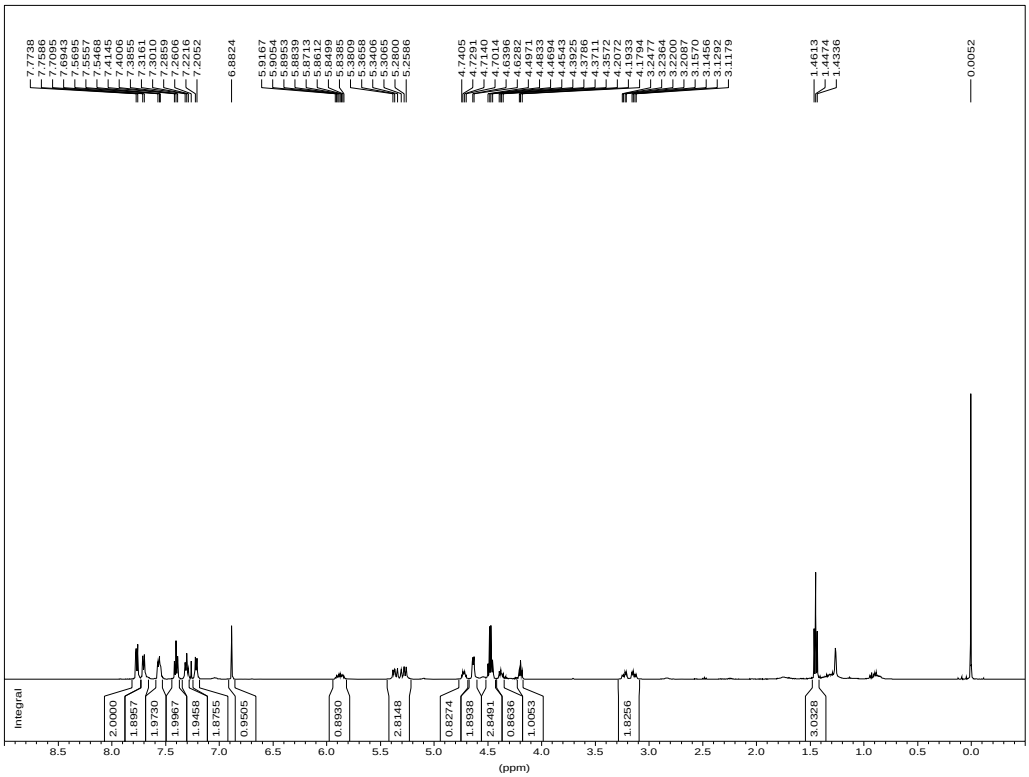

174.8301
158.2934
145.1953
145.0698
142.4871
139.6166
133.0196
130.4698
128.7175
128.1124
126.2086
126.1495
122.0762
120.8587
84.3390
78.5316
67.8686
56.4751
49.8560
49.5756
49.2878
49.0074
48.7196
48.4392
48.2621
48.1588
38.9254
38.3223
26.2573
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
(ppm)

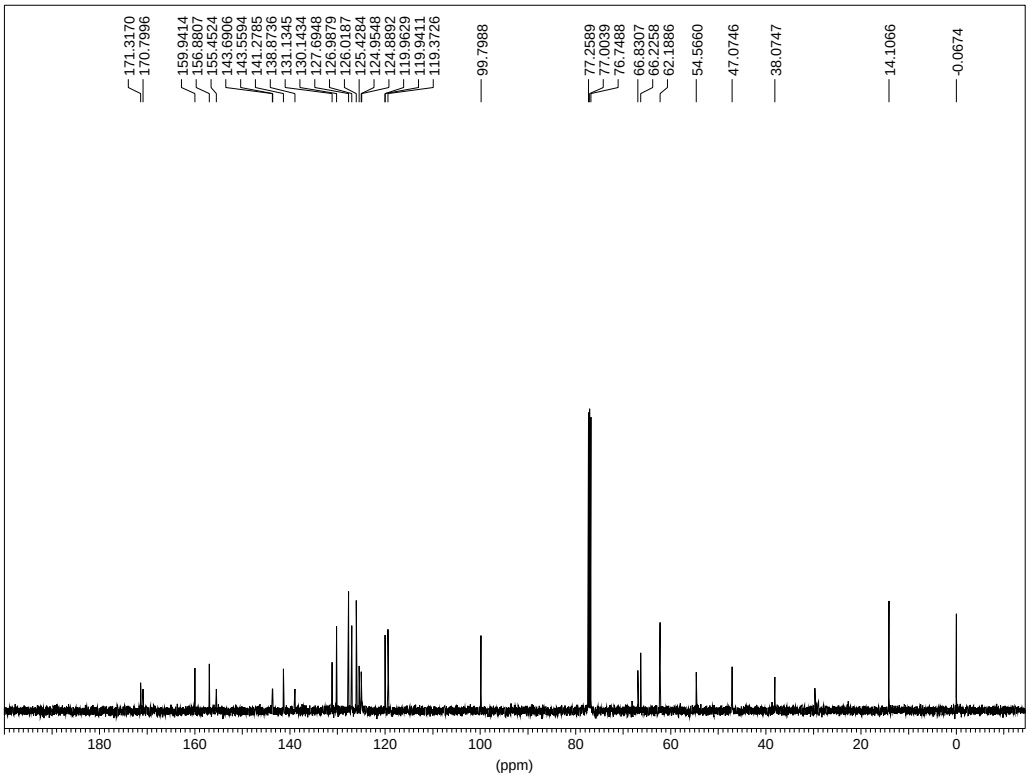

O
O
NHFmoc

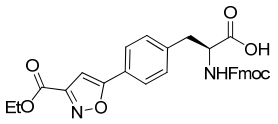

7.7890
7.7643
7.5830
7.5589
7.4356
7.4063
7.4132
7.4000
7.3902
7.3496
7.3466
7.3420
7.3250
7.3211
7.3173
7.2965
7.2603
7.0751
5.8830
5.8617
5.8458
5.8266
5.3461
5.2919
5.2585
4.7249
4.7051
4.6799
4.6274
4.6087
4.4992
4.4745
4.4635
4.4400
4.4049
4.3819
4.3466
4.2373
4.2137
4.1918
3.1525
3.1323
3.1098
3.0819
0.0139
Integral
2.1656
2.2490
3.7228
2.1739
1.3542
1.0000
2.7724
2.9078
1.9474
1.0183
2.9316
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
(ppm)

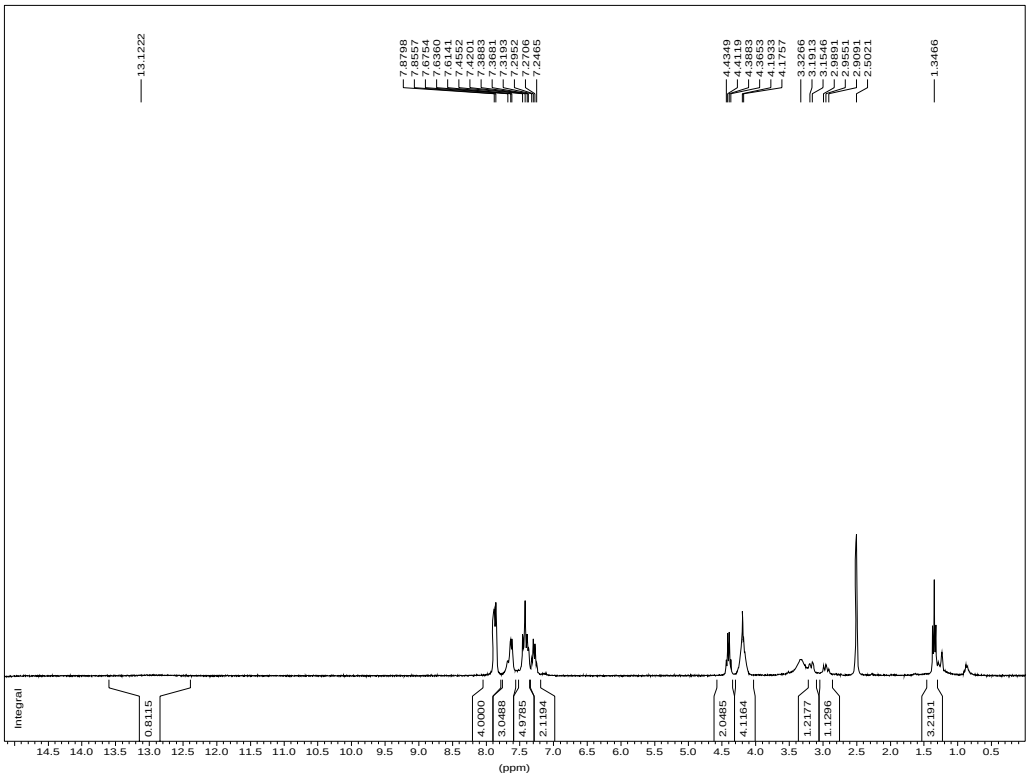

170.8840
155.4725
143.7921
143.6394
141.3193
136.6136
132.2716
131.2316
129.3624
127.7115
127.0496
124.9696
120.9694
119.9585
119.2966
83.3171
77.4186
77.3750
76.9968
76.5750
66.9019
66.1673
54.6105
47.1557
38.1226
-0.0388
180
160
140
120
100
80
60
40
20
0
(ppm)

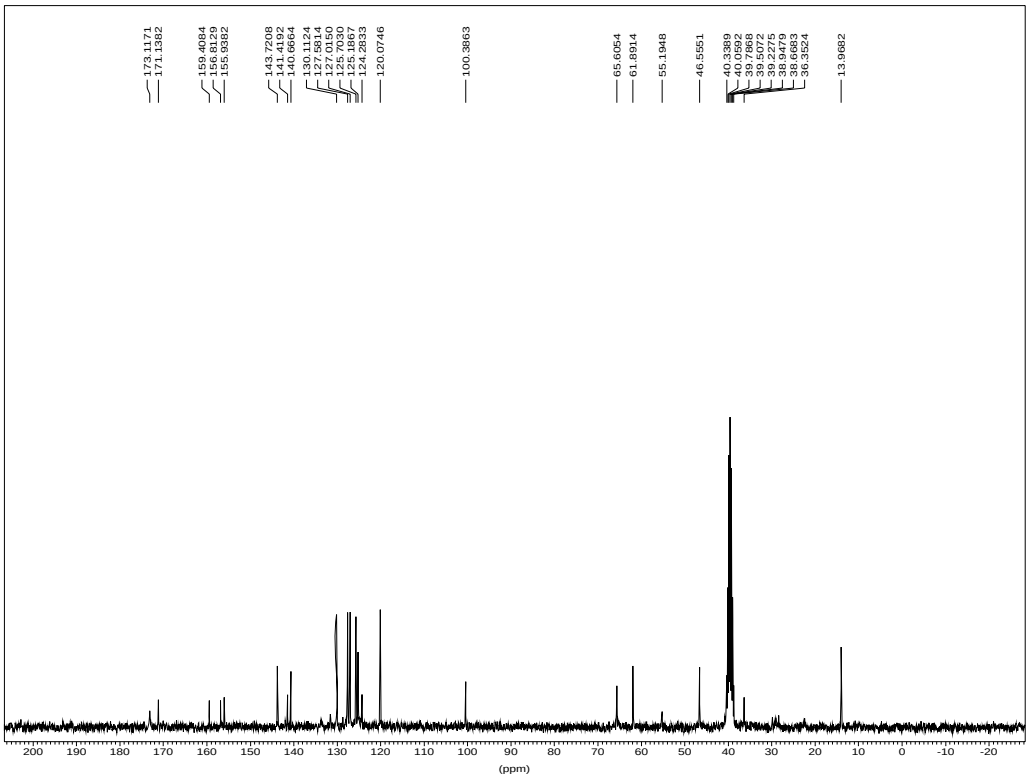

O
O
O
NHFmoc
EtO
N
O

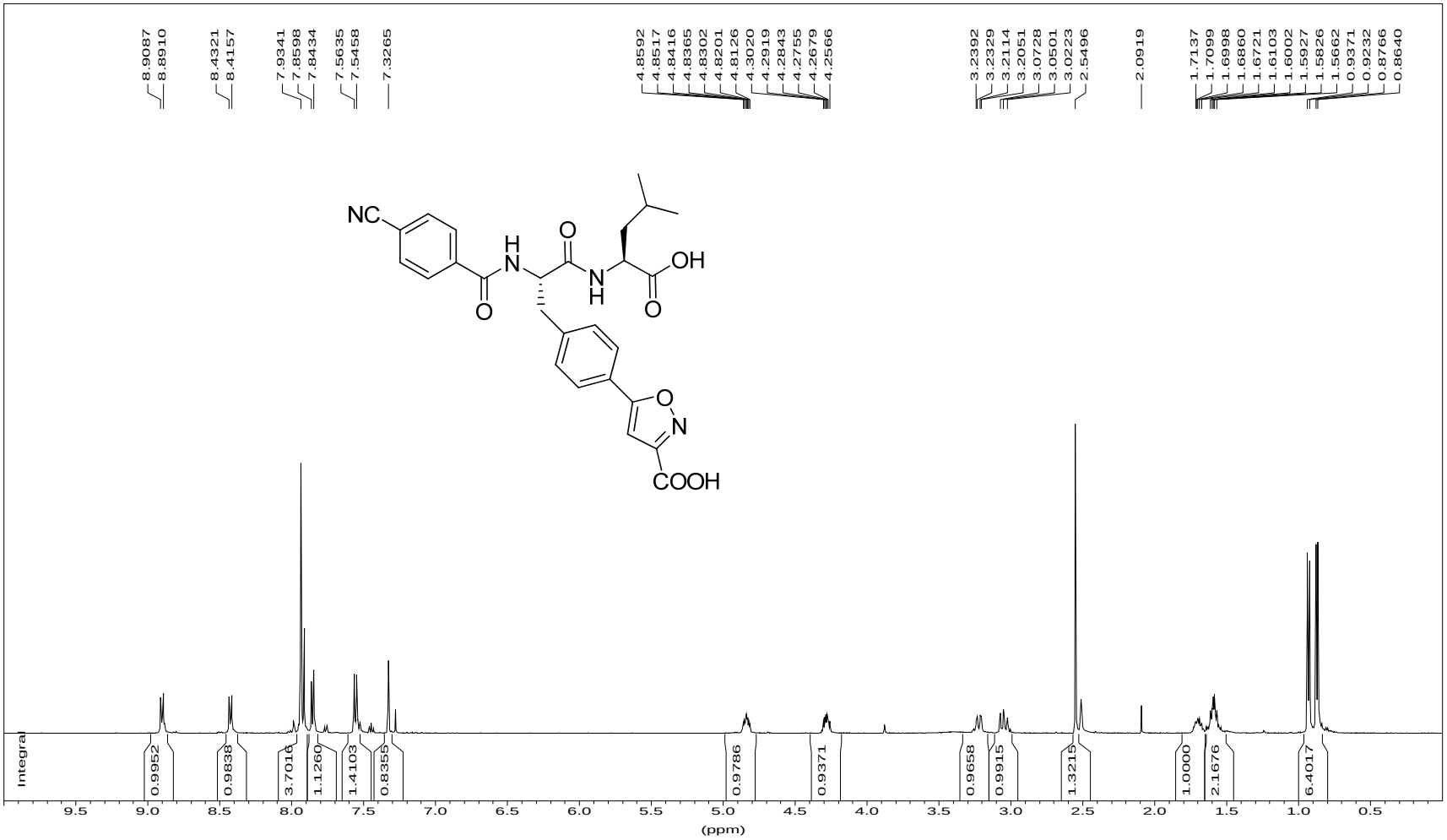

7.7738
7.7586
7.7095
7.6943
7.5695
7.5567
7.5468
7.4145
7.4006
7.3855
7.3161
7.2910
7.2859
7.2606
7.2216
7.2052
6.8824
5.9167
5.9054
5.8953
5.8839
5.8713
5.8612
5.8499
5.8385
5.3869
5.3768
5.3406
5.3065
5.2800
5.2586
4.7405
4.7291
4.7140
4.7014
4.6396
4.6292
4.4971
4.4833
4.4694
4.4543
4.3935
4.3766
4.3711
4.3572
4.2072
4.1933
4.1794
3.2477
3.2364
3.2200
3.2057
3.1570
3.1456
3.1292
3.1179
1.4613
1.4474
1.4336
0.0052
Integral
2.0000
1.8957
1.9730
1.9967
1.9458
1.8755
0.9505
0.8930
2.8148
0.8274
1.8938
2.8491
0.8636
1.0053
1.8256
3.0328
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
(ppm)

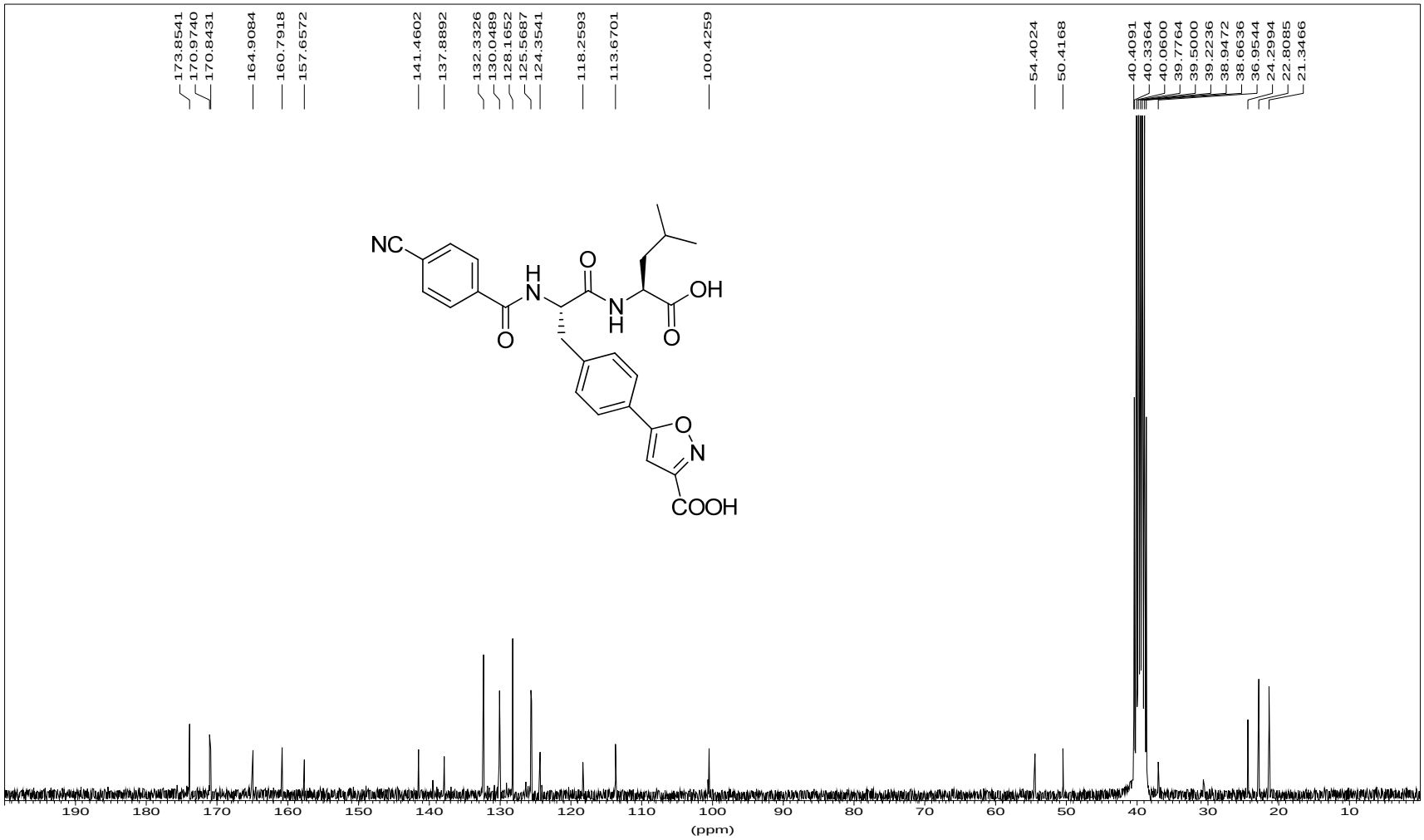

171.3170
170.7996
159.9414
156.8807
155.4524
143.6906
143.5594
141.2785
138.8736
131.1345
130.1434
127.6948
126.9879
126.0187
125.4284
124.9548
124.8892
119.9629
119.9411
119.3726
99.7988
77.2589
77.0039
76.7488
66.8307
66.2258
62.1886
54.5660
47.0746
38.0747
14.1066
-0.0674
180
160
140
120
100
80
60
40
20
0
(ppm)


O
OH
NHFmoc
O
EtO
N
O


13.1222
7.8798
7.8557
7.6754
7.6360
7.6141
7.4452
7.4201
7.3883
7.3681
7.3193
7.2952
7.2706
7.2465
4.4349
4.4119
4.3883
4.3653
4.1933
4.1757
3.3266
3.1943
3.1546
2.9891
2.9551
2.9091
2.5021
1.3466
Integral
0.8115
4.0000
3.0488
4.9785
2.1194
2.0485
4.1164
1.2177
1.1296
3.2191
(ppm)


173.1171
171.1382
159.4084
156.8129
155.9382
143.7208
141.4492
140.6864
130.1124
127.5814
127.0150
125.7030
125.1867
124.2853
120.0746
100.3863
65.6054
61.8914
55.1948
46.5551
40.3389
40.0592
39.7868
39.5072
39.2275
38.9479
38.6683
36.3524
13.9682
(ppm)


8.9087
8.8910
8.4321
8.4157
7.9341
7.8598
7.8434
7.5635
7.5458
7.3265
4.8592
4.8517
4.8416
4.8365
4.8302
4.8201
4.8126
4.3020
4.2919
4.2843
4.2755
4.2679
4.2566
3.2392
3.2329
3.2114
3.2051
3.0728
3.0501
3.0223
2.5496
2.0919
1.7137
1.7099
1.6998
1.6860
1.6721
1.6103
1.6002
1.5927
1.5826
1.5662
0.9371
0.9232
0.8766
0.8640
NC
H
N
O
O
N
H
O
OH
O
O
N
COOH
Integral
0.9952
0.9838
3.7016
1.1260
1.4103
0.8355
0.9786
0.9371
0.9658
0.9915
1.3215
1.0000
2.1676
6.4017
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
(ppm)


173.8541
170.9740
170.8431
164.9084
160.7918
157.6572
141.4602
137.8892
132.3326
130.0489
128.1652
125.5687
124.3541
118.2593
113.6701
100.4259
54.4024
50.4168
40.4091
40.3364
40.0600
39.7764
39.5000
39.2236
38.9472
38.6636
36.9544
24.2994
22.8085
21.3466
NC
H
N
O
N
H
OH
O
O
O
N
COOH
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
(ppm)