Supporting Information

Three-Component Protein Modification Using Mercaptobenzaldehyde Derivatives

Yuanwei Dai, Jiaping Weng, Justin George, Huan Chen, Qishan Lin, Jun Wang, Maksim Royzen and Qiang Zhang*

Department of Chemistry, University at Albany, State University of New York

1400 Washington Avenue, Albany, NY 12222 (USA)

E-mail: qzhang5@albany.edu

Table of Contents

1. General Information

^{1}H NMR spectra were recorded at 400 or 600 MHz at ambient temperature with $CDCl_3$ (Cambridge Isotope Laboratories, Inc.) as the solvent unless otherwise stated. ^{13}C NMR spectra were recorded at 101 or 151 MHz at ambient temperature with $CDCl_3$ as the solvent unless otherwise stated. Chemical shifts are reported in parts per million relative to $CDCl_3$ (^{1}H, δ 7.26; ^{13}C, δ 77.0). All ^{13}C NMR spectra were recorded with complete proton decoupling. High-resolution mass spectra were obtained at the University at Albany-SUNY Core Facility Center using an Agilent G6530BA Q-TOF Mass Spectrometer. Analytical thin layer chromatography was performed on SiliCycle silica gel 60 F254 plates and flash column chromatography was performed on SiliCycle silica gel 60 (40–63 mm). Yields refer to chromatographically and spectroscopically pure materials unless otherwise stated. HATU was purchased from Genscript® (Piscataway, New Jersey). Recombinant human insulin was purchased from MP biomedicals (Solon, Ohio). All other reagents were purchased from Sigma-Aldrich, Alfa Aesar, Chemimpex and Oakwood Chemicals. All commercially available materials were used without further purification. All reactions were carried out in oven-dried glassware under an argon atmosphere unless otherwise noted.

HPLC: All separations involved a mobile phase of 0.05% TFA (v/v) in water (solvent A)/acetonitrile (solvent B). Analytical LC-MS analyses were performed using a Waters 2695 Separations Module and an Agilent G6530BA Q-TOF Mass Spectrometer equipped with Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column, and Waters Microsorb 300-5, C4 250 x 2.0 mm columns at a flow rate of 0.3 mL/min. LC-MS analyses were performed using an Agilent G6530BA Q-TOF Mass Spectrometer. Preparative separations were performed using a Dionex Ultimate UHPLC system equipped with a UV detector and Proto reverse phase HPLC column Microsorb 200 C18 (250 × 20 mm) at a flow rate of 10.0 mL/min.

2. General Procedures

A. Automated solid-phase peptide synthesis

Automated peptide synthesis was performed on a Pioneer peptide synthesis system (GEN600611). Peptides were synthesized under standard automated Fmoc protocols using DMF as solvent, deblocking for 5 min in piperidine/DBU/DMF (2:2:96, *v/v/v*), coupling for 25 min ('standard cycle'), or 55 min ('extended cycle') for amino acids such as prolines, valines, threonines, isoleucines and arginines using HATU as coupling reagent.

The following $^{\alpha}N$-Fmoc or $^{\alpha}N$-Boc-protected amino acids from Novabiochem, GL Biochem or CS Bio were employed in SPPS: Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(O^{t}Bu)-OH, Fmoc-Glu(O^{t}Bu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Met-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(^{t}Bu)-OH, Fmoc-

Thr(^{t}Bu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Tyr-OH, Fmoc-Val-OH. (2S,4R)-Fmoc-4-mercaptopyrrolidine-2-carboxylic acid was from PolyPeptide Group.

B. Preparation of regular polypeptides

Upon completion of the automated synthesis on a 0.05 mmol scale, the peptide resin was washed into a peptide synthesis vessel using CH_2Cl_2. Resin was treated with CH_2Cl_2/TFE/AcOH (8:1:1, *v/v/v*) solution 40 min (×3). The resin was then removed by filtration, the combined cleavage solution was concentrated under reduced pressure, and global deprotection was performed under the treatment of TFA/H_2O/TIPS (95:2.5:2.5, *v/v/v*) solution for 20 min to 1 h, and the filtrate was concentrated under a nitrogen atmosphere. The residue was washed with cold diethyl ether to give a white solid, which was then dissolved in a mixture of acetonitrile and water containing 5% of acetic acid. The resulting solution was ready for HPLC purification after filtration.

C. Conjugation with peptides

To a solution of the peptides (0.01 M) in 0.8 mL of THF/phosphate buffer solution (1:9, *v/v*, pH = 6.8) was added 2.2 equivalents of **1** (3.9 mg, 0.0177 mmol) in one portion. The reaction mixture was stirred at room temperature for 8 h and monitored by LC-MS. Upon completion, the reaction was diluted with distilled deionized water (2.0 mL) and further purified using preparative HPLC.

D. One-pot Conjugation and Cu-catalyzed Azide Alkyne Cycloaddition (double-click)

To a solution of glycine methyl ester hydrochloride (2.7 mg, 0.024 mmol, 0.05 M) in 400 μL of phosphate buffer solution (pH = 6.8), was added **1** (8.8 mg, 0.04 mmol, 0.1 M) in 400 μL of *tert*-butanol. After degassing for 30 min, 0.048 mmol of azide compound (benzyl azide, *p*-azidoacetophenone, 4-(azidomethyl)-7-methoxy-2H-chromen-2-one, 1-azido-1-deoxy-β-D-glucopyranoside tetraacetate, 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-glucopyranosyl azide) was added to the mixture, followed by the addition of 5 μL of $CuSO_4$ (0.5 M in deionized water), 5 μL of freshly prepared sodium ascorbate (1 M in deionized water) and 5 μL of TBTA (0.5 M in *tert*-butanol).[1] The reaction mixture was stirred at room temperature for 8 h, monitored by LC-MS. Upon completion, the reaction was diluted by adding 3 mL of H_2O and extracted by ethyl acetate (3 mL×3), combined, dried over Na_2SO_4. The solvents were removed in vacuo, the residue was purified by flash column chromatography to obtain the desired product.

E. One-pot Conjugation and Cu-catalyzed Azide Alkyne Cycloaddition (mono-click)

To a solution of glycine methyl ester hydrochloride (2.7 mg, 0.024 mmol, 0.05 M) in 400 μL of phosphate buffer solution (pH = 6.8), was added **1** (8.8 mg, 0.04 mmol, 0.1 M) in 400 μL of *tert*-butanol. After degassing

for 30 min, 0.024 mmol of azide compound (benzyl azide, *p*-azidoacetophenone, biotin-PEG3-azide, 4-(azidomethyl)-7-methoxy-2H-chromen-2-one, 1-azido-1-deoxy-β-D-glucopyranoside tetraacetate, 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-glucopyranosyl azide) was added to the mixture, followed by the addition of 5 μL of $CuSO_4$ (0.5 M in deionized water), 5 μL of freshly prepared sodium ascorbate (1 M in deionized water) and 5 μL of TBTA (0.5 M in *tert*-butanol). The reaction mixture was stirred at room temperature for 8 h, monitored by LC-MS. Upon completion, the reaction was diluted by adding 3 mL of H_2O and extracted by ethyl acetate (3 mL×3), combined, dried over Na_2SO_4. The solvents were removed in vacuo, the residue was purified by flash column chromatography to obtain the desired product.

F. One-pot Conjugation and CuAAC with peptides

To a solution of peptides (0.025 M) in 200 μL of phosphate buffer solution (pH = 6.8), was added **1** (1.8 mg, 0.04 M) and 0.005 mmol of azide compound (4-(azidomethyl)-7-methoxy-2H-chromen-2-one, 1-Azido-1-deoxy-β-D-glucopyranoside) in 200 μL of *tert*-butanol. After degassing for 30 min, was added to the mixture followed by the addition of 5 μL of $CuSO_4$ (0.1 M in deionized water), 5 μL of freshly prepared sodium ascorbate (0.2 M in deionized water) and 5 μL of TBTA (0.1 M in *tert*-butanol). The reaction mixture was stirred at room temperature for 8 h, monitored by LC-MS. Upon completion, the reaction was diluted with distilled deionized water (2.0 mL) and further purified using preparative HPLC. Mobile phase compositions were: solvent A: 0.1% TFA in H_2O, solvent B: 0.1% TFA in CH_3CN. The sample was eluted using a Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column.

G. High resolution mass spectrometry/mass spectrometry measurements:

*For compounds **11e, 11i-k, 11n-q**:*

High resolution LC/MS/MS analysis was performed on an Agilent G6530BA Q-TOF Mass Spectrometer, operated with electrospray ionization in positive-ion mode. The tandem MS analysis of all the pure compounds was performed on the same Agilent MS instrument in the MS/MS mode using direct injection.

For modified insulin:

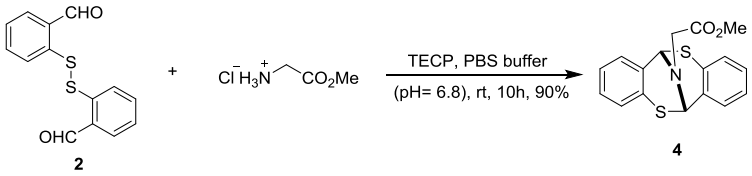
DTT was used to break all disulfide bonds. The procedure was: 5 μL of 1 M DTT in water was added to 50 μL of 1 mg/ml of the intact protein sample in deionized H_2O. The solution was incubated for 1 h at 40 ºC. The sample was diluted 100x in 0.1% TFA H_2O/CH_3CN (80:20, *v/v*), filtered and directly infused into the LC/MS system. The MS/MS experiments were carried out on the same Agilent MS instrument in the MS/MS mode. LC gradient conditions are described as below: The gradient started at 20% B then increasing to 90% B in 15 min at a flow rate of 0.5 mL/min. Mobile phase compositions were: solvent A: 0.1% TFA in H_2O, solvent B: 0.1% TFA in CH_3CN. The sample was eluted using a Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column.

H. Bioassay Method:

The human monocyte cell line THP-1 was purchased from ATCC. Cells were routinely cultured in RPMI-1640 medium in T75 cm2 flask supplemented with 10% (v/v) fetal bovine serum (FBS) and 50×10^{-6} m 2-mercaptoethanol at 37 °C in 5% CO_2 incubator. 0.06 mM AF546 azide (control) or **17** was incubated with 1×10^6 /ml THP-1 cells for 4 h. After washing the cells for 3 times, cells were transferred to a 96 well plate for imaging. The same experiment for both control and **17** was repeated for 3 times. Fluorescence imaging was conducted on an Olympus IX73 inverted fluorescence microscope equipped with a yellow Cy3 filter set. A digital camera (Zyla sCMOS, Andor) mounted on the microscope was used to capture images using Andor SOLIS software.

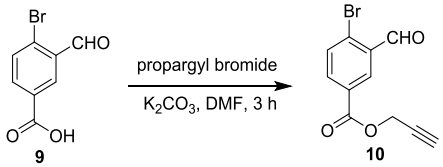
3. Experiment details and data

*Synthesis of **4***



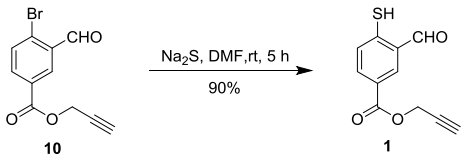
To a solution of glycine methyl ester hydrochloride (30 mg, 0.273 mmol) and tris(2-carboxyethyl)phosphine hydrochloride (210 mg, 0.731 mmol) in 3.6 mL of PBS buffer, was added Na_2HPO_4 to adjust the pH of the solution to 6.8, followed by the addition of 2,2'-dithiodibenzaldehyde **2**[2] (100 mg, 0.364 mmol) in 0.4 mL of tetrahydrofuran. The mixture was stirred at room temperature for 10 h, monitored by TLC. Upon completion, the resulting solution was diluted with 10 mL of water and extracted with ethyl acetate (15 mL×3), combined the organic phase, washed with water and brine. Removing solvents in vacuo, the residue was purified by silica gel column chromatography (hexane: ethyl acetate=5: 1) to give **4** (54.1 mg, 90%) as a white solid. R_f = 0.48 (hexane: EtOAc= 5: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) 7.23 (2 H, s), 7.05 (4 H, s), 6.99 (2 H, s), 5.54 (2 H, s), 3.88 (1 H, d, *J* 17.2), 3.74 (3 H, s), 3.53 (1 H, d, *J* 17.2). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 169.85, 132.66, 129.32, 129.24, 127.84, 127.55, 124.85, 61.87, 54.96, 52.10. **IR** (thin film): cm^{-1} 1740, 1219, 741, 671. **HRMS** (ESI+): calculated for $C_{17}H_{16}NO_2S_2$ $[M+H]^+$ 330.0622 found 330.0633.

*Synthesis of **10***



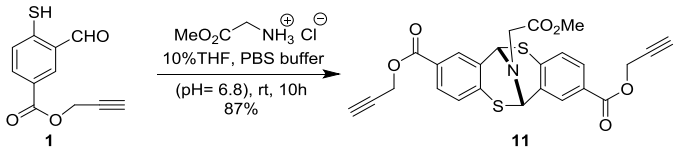
To a solution of commercial available 4-bromo-3-formyl-benzoic acid **9**[3] (184 mg, 0.807 mmol) in 6 mL of dimethyl formamide was added potassium carbonate (233 mg, 1.62 mmol), followed by the addition of propargyl bromide (0.16 mL, 1.06 mmol, 80% in toluene). The mixture was stirred for 3 h at room temperature, TLC showed fully consumption of starting material. The mixture was diluted with 20 mL water, extracted with ethyl acetate (30 mL×3), combined organic phase, washed with water and brine. Remove the solvent in vacuo, the residue was purified by silica gel column chromatography (hexane: ethyl acetate= 2: 1) to give **10** (204 mg, 95%) as a white solid. R_f = 0.64 (hexane: EtOAc= 2: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 10.38 (s, 1H), 8.56 (d, *J* = 2.2 Hz, 1H), 8.12 (dd, *J* = 8.3, 2.2 Hz, 1H), 7.77 (d, *J* = 8.4 Hz, 1H), 4.95 (d, *J* = 2.4 Hz, 2H), 2.55 (t, *J* = 2.5 Hz, 1H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 190.66, 164.17, 135.63, 134.40, 133.70, 132.06, 131.21, 129.57, 77.09, 75.56, 53.03. **IR** (thin film): cm^{-1} 3283, 2870, 2130, 1725, 1695, 1593, 1291,1178, 1098. **HRMS** (ESI+): calculated for $C_{11}H_8BrO_3$ $[M+H]^+$ 266.9657, found: 266.9650.

*Synthesis of **1***



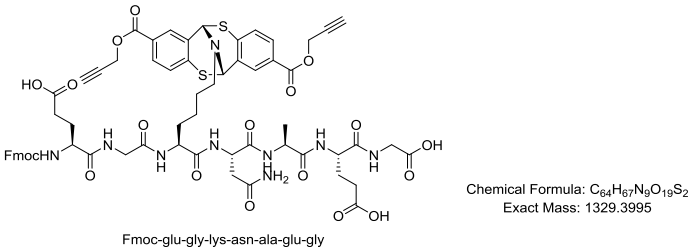
To a solution of compound **10** (650 mg, 2.443 mmol) in 18 mL of dimethyl formamide was added sodium sulfide (381 mg, 4.887 mmol). The mixture was stirred for 5h at room temperature, TLC showed fully consumption of starting material. The mixture was diluted with 20 mL ethyl acetate, washed with 1 M HCl (15 mL×2). The aqueous phase was extracted with ethyl acetate (30 mL×3), combined organic phase, washed with water and brine. Remove the solvent in vacuo, the obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1: 1) to give **1** (484 mg, 90%) as a yellow solid. R_f = 0.27 (hexane: EtOAc= 1: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 10.23 (s, 1H), 8.55 (s, 1H), 8.11 (d, *J* = 8.1 Hz, 1H), 7.81 (d, *J* = 8.1 Hz, 1H), 4.94 (s, 2H), 2.53 (s, 1H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 191.33, 164.07, 145.56, 136.53, 134.58, 133.60, 132.48, 127.74, 126.29, 75.50, 52.93. **IR** (thin film): cm^{-1} 3283, 2930, 2866, 1718, 1697, 1599, 1433, 1371, 1286, 1244, 1180, 1105, 1056, 1024, 758. **HRMS** (ESI+): calculated for $C_{11}H_8O_3S$ $[M+H]^+$ 221.0272, found: 221.0261.

*Synthesis of **11***



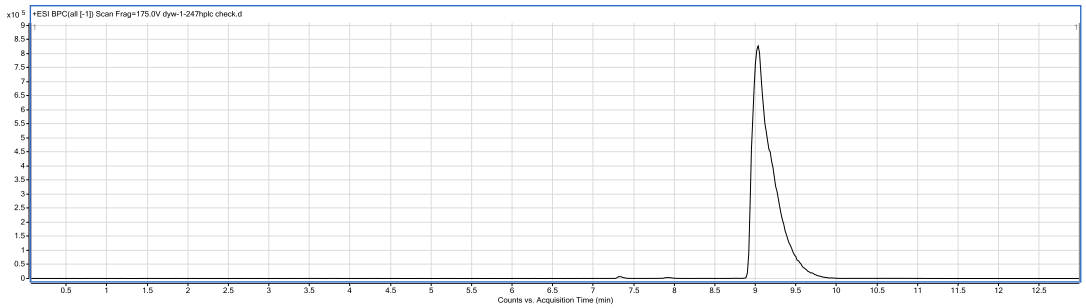
To a solution of the glycine methyl ester hydrochloride (5.1 mg, 0.046 mmol) in 1 mL of phosphate buffer solution (pH = 6.8), was added **1** (20 mg, 0.091 mmol) in 100 μL of tetrahydrofuran. The reaction mixture was stirred at room temperature for 8 h, monitored by TLC. Upon completion, the mixture was diluted by adding 5 mL of H_2O and extracted by ethyl acetate (5 mL×3), combined, dried over Na_2SO_4. The solvents were removed in vacuo, the obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 2: 1) to give **11** (19.7 mg, 88%) as a white solid. R_f = 0.43 (hexane: EtOAc= 2: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.94 (s, 2H), 7.75 (d, *J* = 8.6 Hz, 2H), 7.10 (d, *J* = 8.3 Hz, 2H), 5.68 (s, 2H), 4.99 – 4.78 (m, 4H), 3.86 (d, *J* = 17.2 Hz, 1H), 3.77 (s, 3H), 3.53 (d, *J* = 17.2 Hz, 1H), 2.52 (t, *J* = 2.4 Hz, 1H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 169.40, 164.97, 136.64, 132.50, 130.56, 128.74, 127.78, 126.09, 77.59, 75.14, 61.88, 54.83, 52.53, 52.34. **IR** (thin film): cm^{-1} 3281, 2879, 1749, 1717, 1601, 1260, 1242, 1219, 1109, 759. **HRMS** (ESI+): calculated for $C_{25}H_{20}NO_6S_2$ $[M+H]^+$ 494.0732, found: 494.0722.

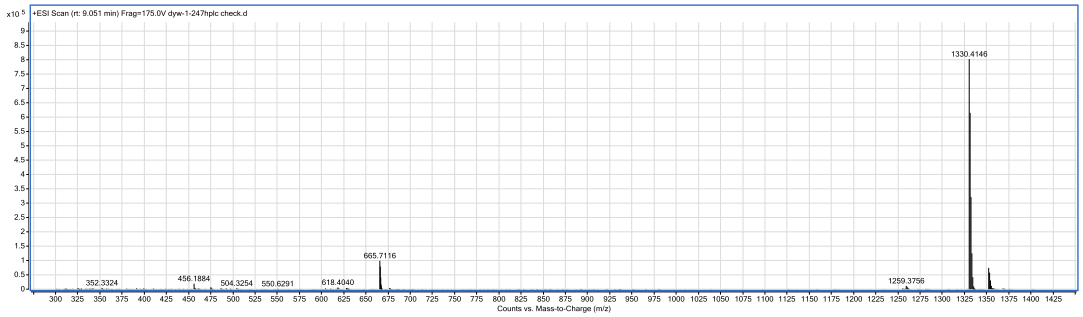
*Synthesis of **11a***



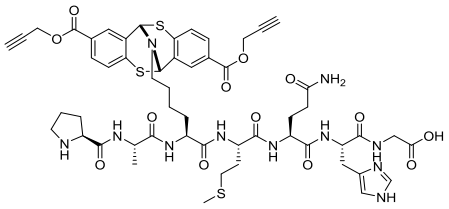
Fmoc-glu-gly-lys-asn-ala-glu-gly

According to General Procedure **A** and **B**, peptide Fmoc-EGKNAEG was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide Fmoc-EGKNAEG was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 95% solvent B over 25 min, Higgins Analytical Proto 200 10 μm 250 × 20 nm C18 column) afforded the desired modified peptide **11a** as a white solid after lyophilization (8.8 mg, 83%). Analytical HPLC: t_R = 9.1 min (50 to 95% solvent B over 13 min, Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{64}H_{68}N_9O_{19}S_2$ $[M+H]^+$ 1330.4073, found: 1330.4146.





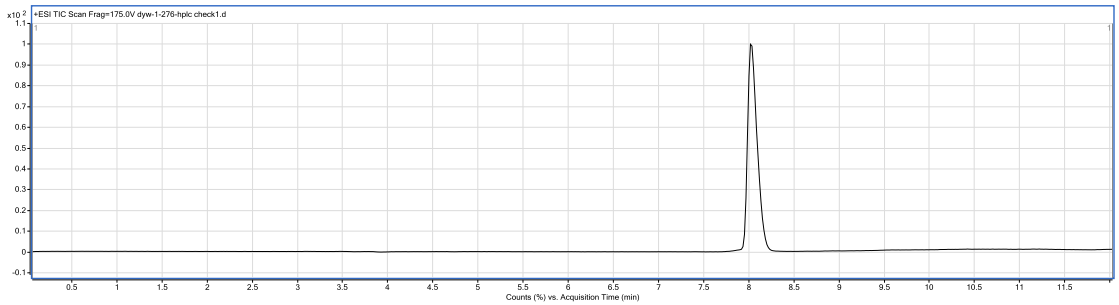
*Synthesis of **11b***

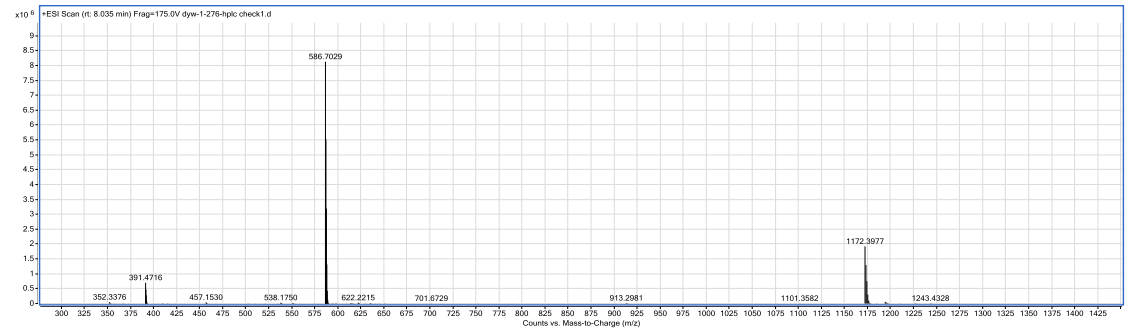


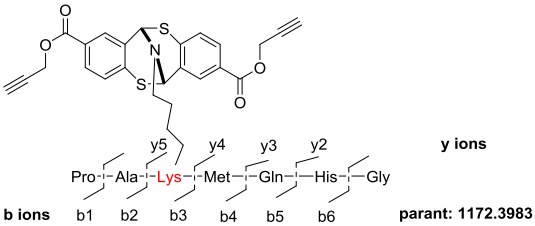
pro-ala-lys-met-gln-his-gly

Chemical Formula: $C_{54}H_{65}N_{11}O_{13}S_3$
Exact Mass: 1171.3925

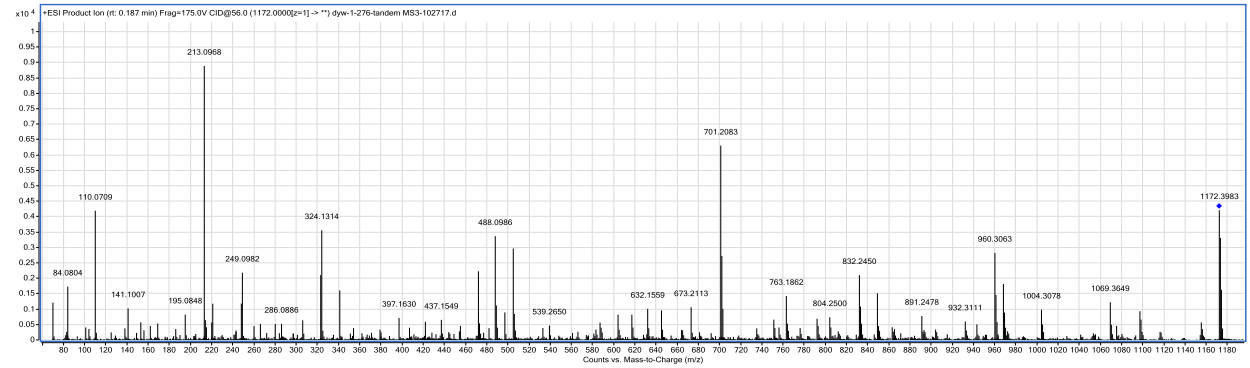
According to General Procedure **A** and **B**, peptide PAKMQHG was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide PAKMQHG was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 95% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **11b** as a white solid after lyophilization (8.2 mg, 88%). Analytical HPLC: t_R = 8.1 min (35 to 95% solvent B over 12 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{54}H_{66}N_{11}O_{13}S_3$ $[M+H]^+$ 1172.4004, found: 1172.3977.





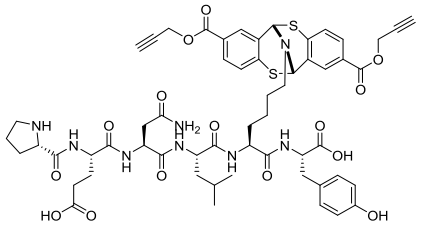


Ion	Cal (Da)	Found (Da)	Ion	Cal (Da)	Found (Da)
b1(-CO)	70.0651	70.0651	b6*	1097.3683	1097.3629
b2	169.0977	169.0968	y2	213.0982	213.0968
b3*	701.2104	701.2083	y3	341.1568	341.1566
b4*	832.2509	832.2450	y4	472.1973	472.1962
b5*	960.3094	960.3063	y5*	1004.3099	1004.3078



MS/MS analysis of modified peptide **11b**. *denotes fragments containing the modification.

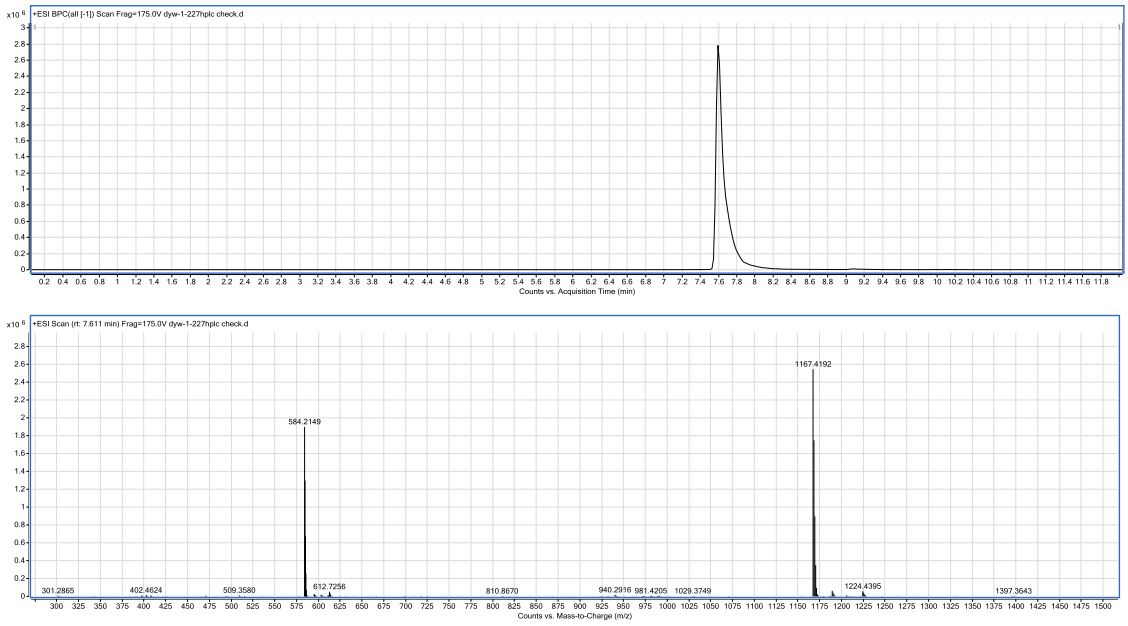
*Synthesis of **11c***



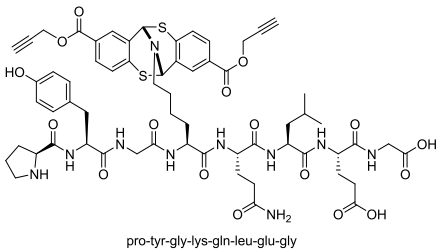
pro-glu-asn-leu-lys-tyr

Chemical Formula: $C_{57}H_{66}N_8O_{15}S_2$
Exact Mass: 1166.4089

According to General Procedure **A** and **B**, peptide PENLKY was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide PENLKY was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 90% solvent B over 25 min, Higgins Analytical Proto 200 10 μm 250 × 20 nm C18 column) afforded the desired modified peptide **11c** as a white solid after lyophilization (8.4 mg, 90%). Analytical HPLC: t_R = 7.6 min (35 to 95% solvent B over 12 min, Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{57}H_{67}N_8O_{15}S_2$ $[M+H]^+$ 1167.4167, found: 1167.4192.



*Synthesis of **11d***

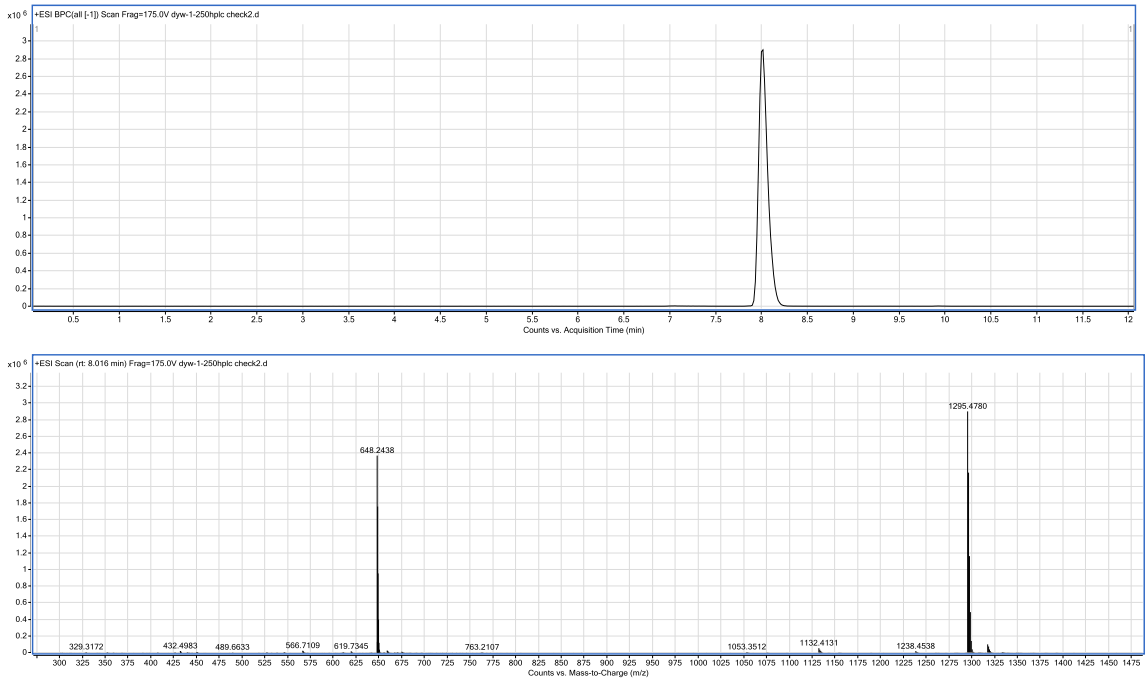


pro-tyr-gly-lys-gln-leu-glu-gly

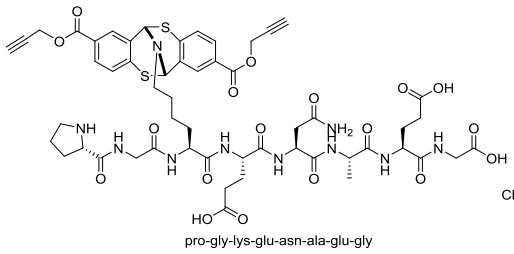
Chemical Formula: $C_{62}H_{74}N_{10}O_{17}S_2$
Exact Mass: 1294.4675

According to General Procedure **A** and **B**, peptide PYGKQLEG was prepared by a three-step protocol on

a 0.05 mmol scale. The conjugation with peptide PYGKQLEG was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (30 to 90% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **11d** as a white solid after lyophilization (8.7 mg, 84%). Analytical HPLC: t_R = 8.1 min (35 to 95% solvent B over 12 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{62}H_{75}N_{10}O_{17}S_2$ $[M+H]^+$ 1295.4753, found: 1295.4780.



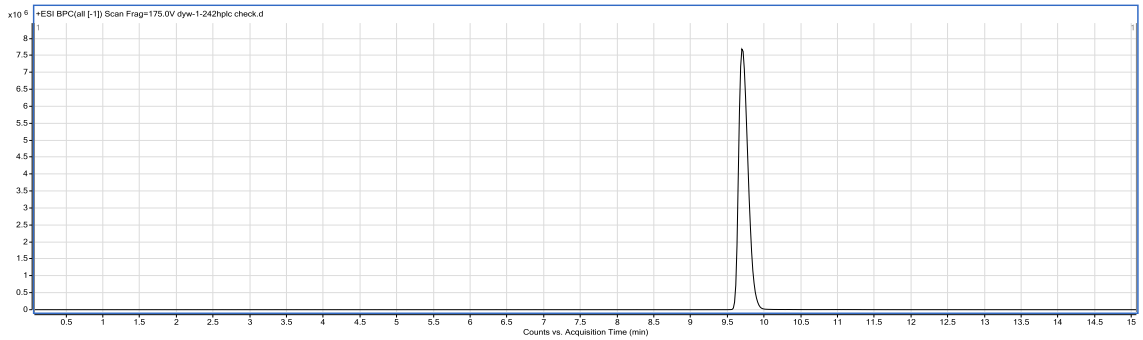
*Synthesis of **11e***

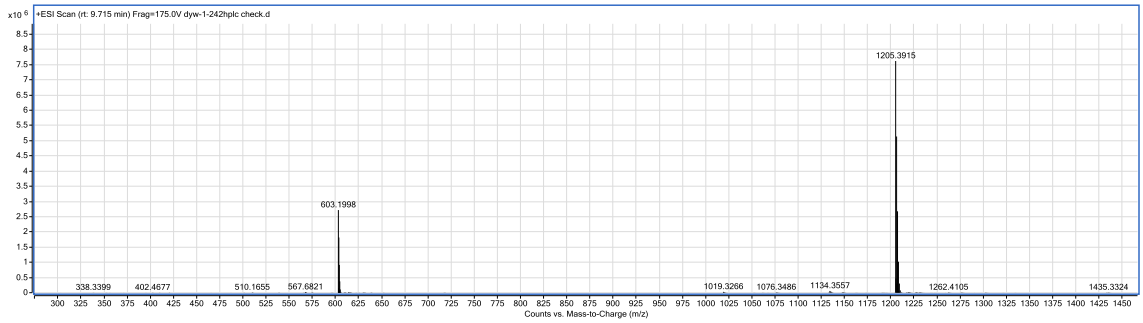


Chemical Formula: $C_{54}H_{64}N_{10}O_{18}S_2$
Exact Mass: 1204.3841

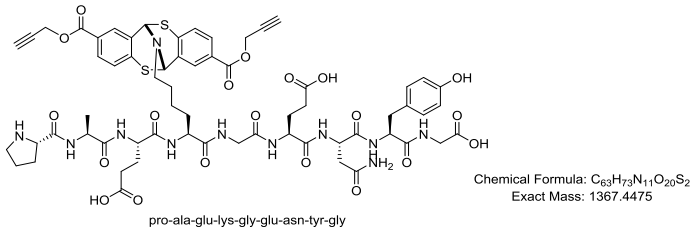
According to General Procedure **A** and **B**, peptide PGKENAEG was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide PGKENAEG was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 90% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **11e** as a white solid after lyophilization (8.7 mg, 90%). Analytical HPLC: t_R = 9.7 min (35 to 95% solvent B over 15 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for

$C_{54}H_{65}N_{10}O_{18}S_2$ $[M+H]^+$ 1205.3920, found: 1205.3915.

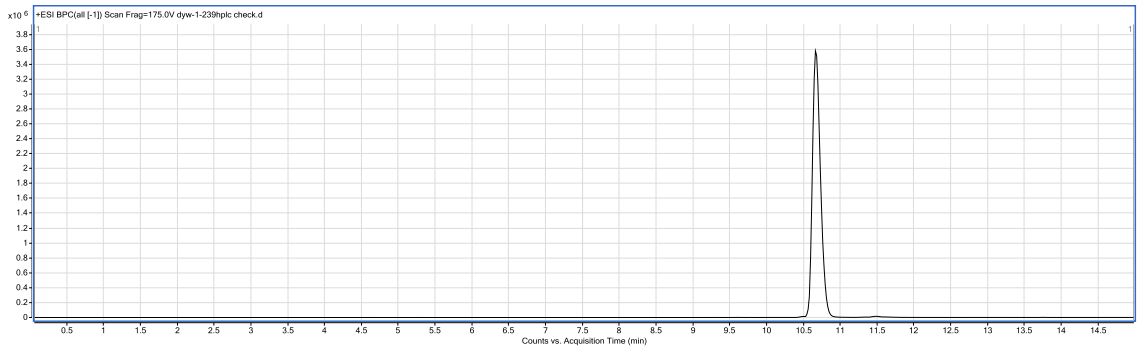


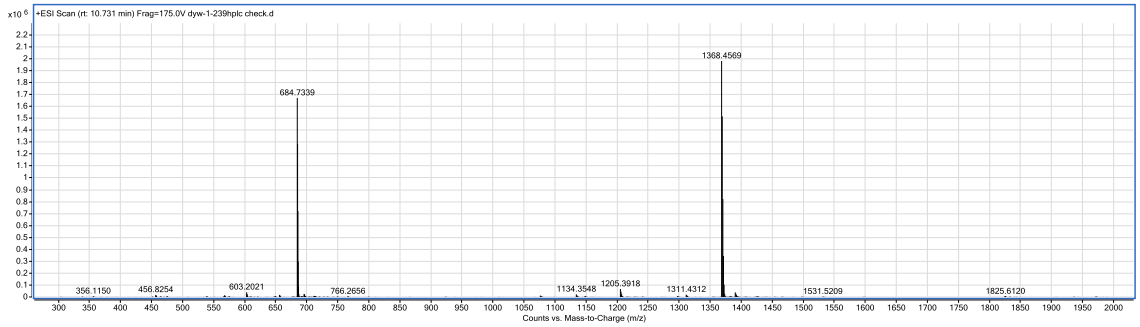


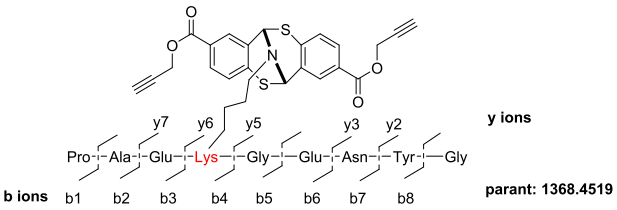
*Synthesis of **11f***



According to General Procedure **A** and **B**, peptide PAEKGENYG was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide PAEKGENYG was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 90% solvent B over 25 min, Higgins Analytical Proto 200 10 μm 250 × 20 nm C18 column) afforded the desired modified peptide **11f** as a white solid after lyophilization (9.4 mg, 86%). Analytical HPLC: t_R = 10.7 min (35 to 95% solvent B over 15 min, Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{63}H_{74}N_{11}O_{20}S_2$ $[M+H]^+$ 1368.4553, found: 1368.4569.





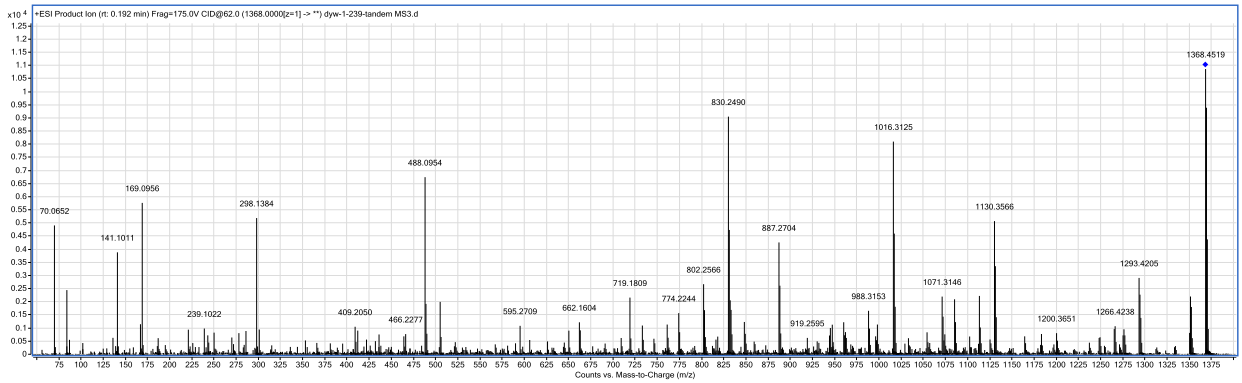


y ions

b ions

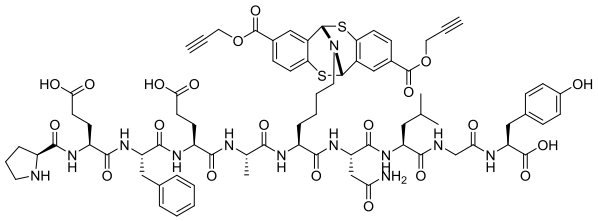
parant: 1368.4519

Ion	Cal (Da)	Found (Da)	Ion	Cal (Da)	Found (Da)
b1(-CO)	70.0651	70.0652	b8*	1293.4233	1293.4205
b2	169.0977	169.0956	y2	239.1026	239.1022
b3	298.1403	298.1384	y3	353.1456	353.1432
b4*	830.2530	830.2490	y5	539.2096	539.2137
b5*	887.2744	887.2704	y6	1071.3223	1071.3146
b6*	1016.3170	1016.3125	y7	1220.3649	1220.3651
b7*	1130.3599	1130.3566			



MS/MS analysis of modified peptide **11f**. *denotes fragments containing the modification.

*Synthesis of **11g***

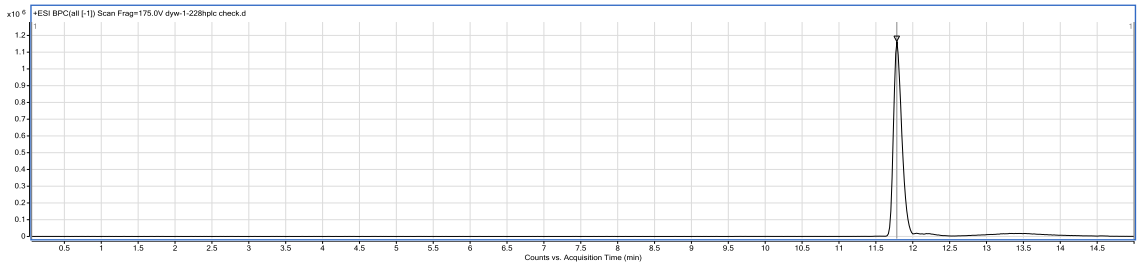


pro-glu-phe-glu-ala-lys-asn-leu-gly-tyr

Chemical Formula: $C_{76}H_{90}N_{12}O_{21}S_2$

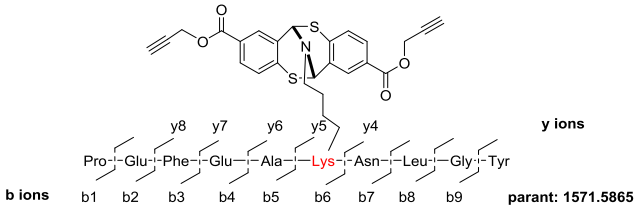
Exact Mass: 1570.5785

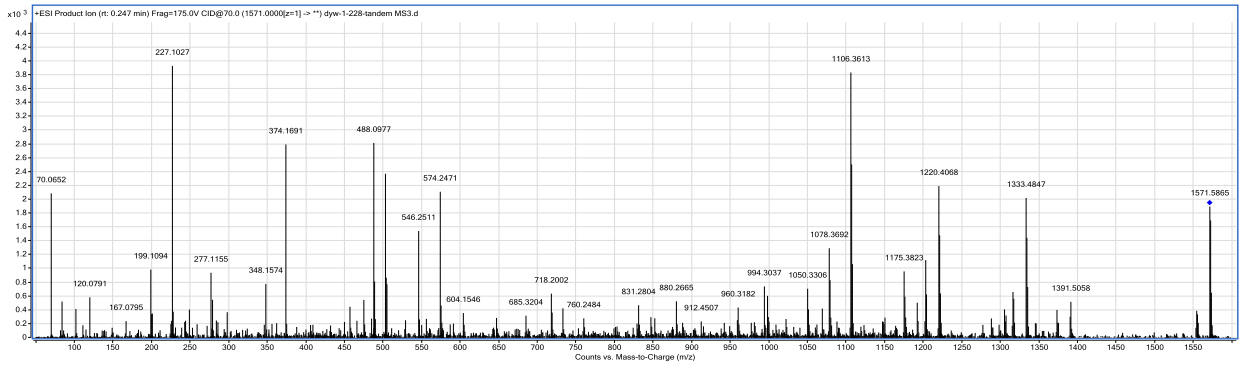
According to General Procedure **A** and **B**, peptide PEFEAKNLGY was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide PEFEAKNLGY was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (40 to 90% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **11g** as a white solid after lyophilization (11.1 mg, 88%). Analytical HPLC: t_R = 11.8 min (35 to 95% solvent B over 15 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{76}H_{91}N_{12}O_{21}S_2$ $[M+H]^+$ 1571.5863, found: 1571.5917.



Ion	Cal (Da)	Found (Da)	Ion	Cal (Da)	Found (Da)
b1(-CO)	70.0651	70.0651	b8*	1333.4904	1333.4867
b2	227.1026	227.1020	b9*	1390.5124	1390.5100
b3	374.1716	374.1687	y4	466.2296	466.2255
b4	503.2142	503.2119	y5	998.3423	998.3450
b5	574.2508	574.2482	y6	1050.3306	1050.3359

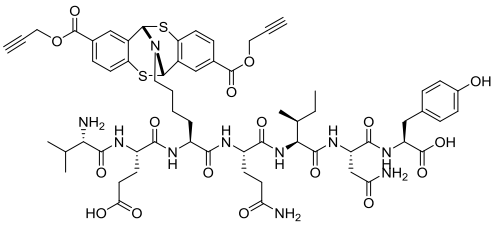
b6*	1106.3634	1106.3601	y7	1198.4220	1198.4277
b7*	1220.4069	1220.4058	y8	1345.4904	1345.4816





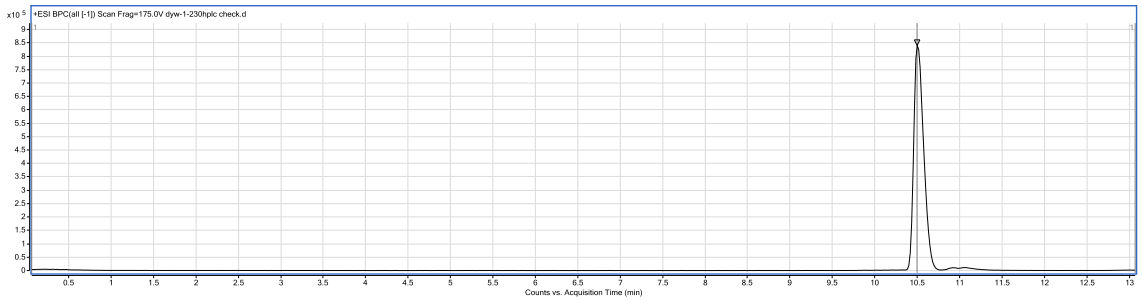
MS/MS analysis of modified peptide **11g**. *denotes fragments containing the modification.

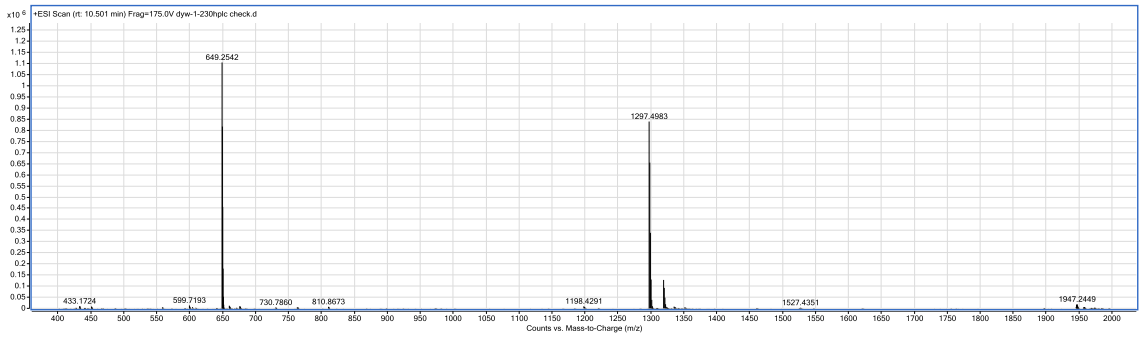
*Synthesis of **11h***

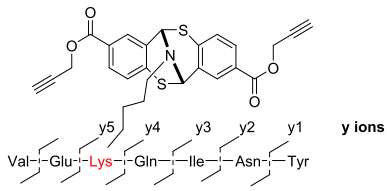


Chemical Formula: $C_{62}H_{76}N_{10}O_{17}S_2$
Exact Mass: 1296.4831

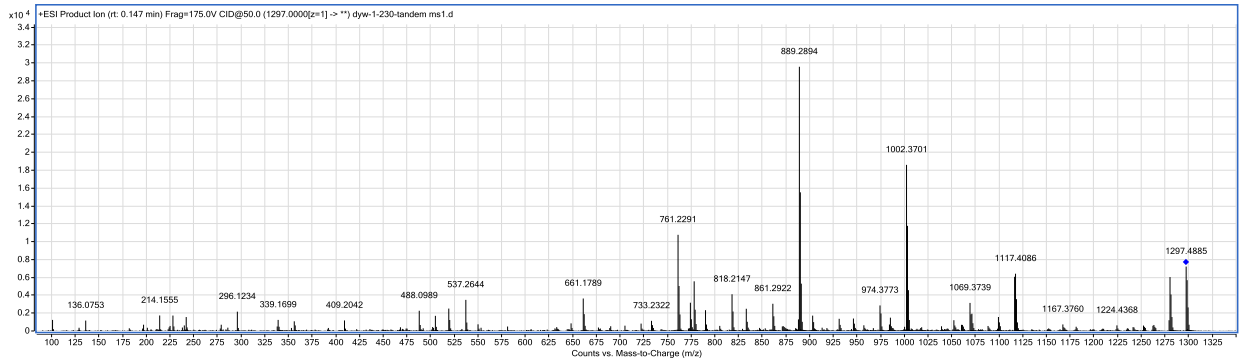
According to General Procedure **A** and **B**, peptide VEKQINY was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide VEKQINY was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 90% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **11h** as a white solid after lyophilization (8.6 mg, 82%). Analytical HPLC: t_R = 10.5 min (35 to 95% solvent B over 13 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{62}H_{77}N_{10}O_{17}S_2$ $[M+H]^+$ 1297.4910, found: 1297.4983.





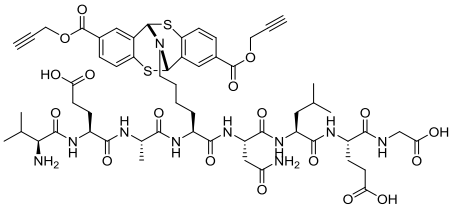


Ion	Cal (Da)	Found (Da)	Ion	Cal (Da)	Found (Da)
b1	101.0841	101.0717	y1	136.0757	136.0753
b2	228.1110	228.1348	y2	296.1241	296.1234
b3*	761.2309	761.2291	y3	409.2082	409.2042
b4*	889.2895	889.2894	y4	537.2667	537.2644
b5*	1002.3736	1002.3701	y5*	1069.3794	1069.3739
b6*	1116.4065	1116.4086			



MS/MS analysis of modified peptide **11h**. *denotes fragments containing the modification.

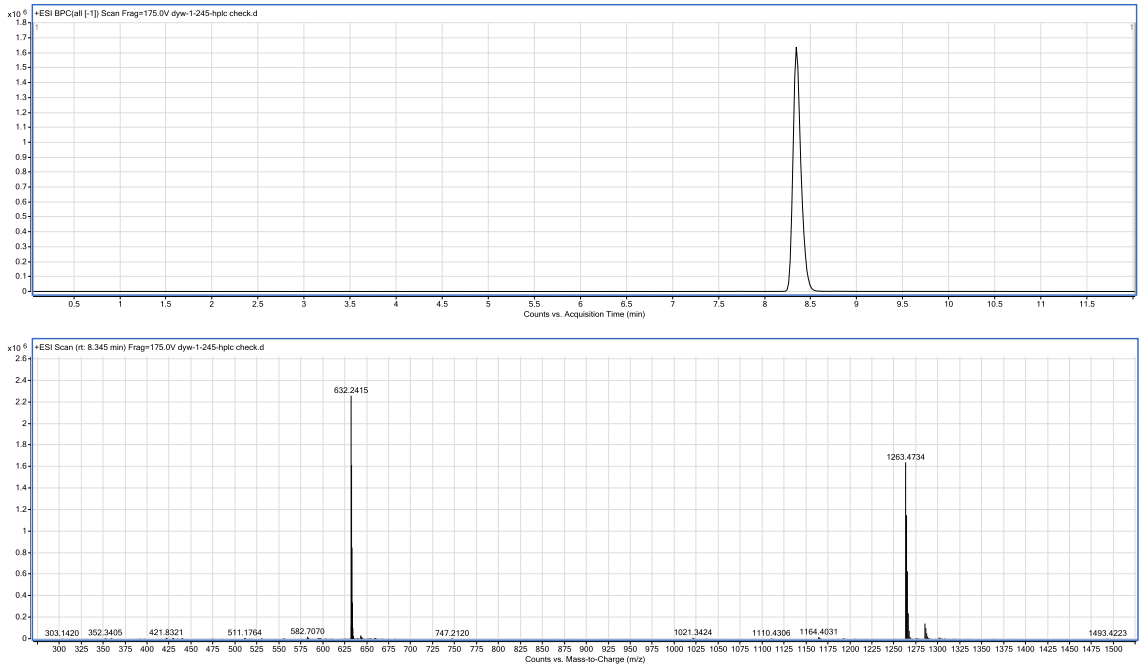
*Synthesis of **11i***

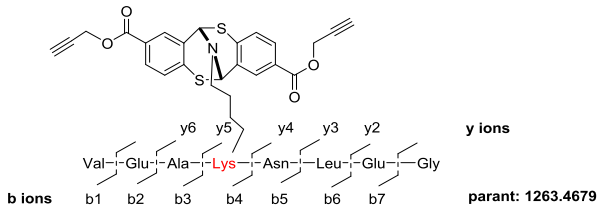


val-glu-ala-lys-asn-leu-glu-gly

Chemical Formula: $C_{58}H_{74}N_{10}O_{18}S_2$
Exact Mass: 1262.4624

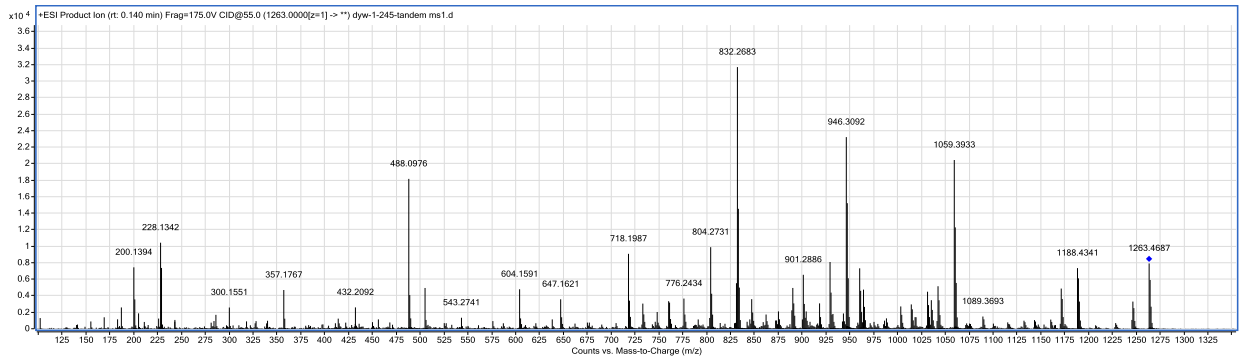
According to General Procedure **A** and **B**, peptide VEAKNLEG was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide VEAKNLEG was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 95% solvent B over 25 min, Higgins Analytical Proto 200 10 μm 250 × 20 nm C18 column) afforded the desired modified peptide **11i** as a white solid after lyophilization (7.9 mg, 78%). Analytical HPLC: t_R = 8.3 min (40 to 95% solvent B over 12 min, Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{58}H_{75}N_{10}O_{18}S_2$ $[M+H]^+$ 1263.4702, found: 1263.4734.





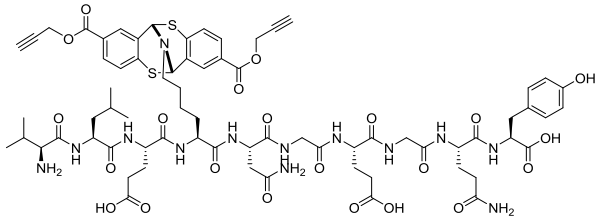
Ion	Cal (Da)	Found (Da)	Ion	Cal (Da)	Found (Da)
b1	101.0841	101.0713	b7*	1188.4376	1188.4341

b2	228.1110	228.1342	y6*	1035.3587	1035.3596
b3	300.1554	300.1551	y5*	964.3216	964.3229
b4*	832.2681	832.2683	y4	432.2089	432.2092
b5*	946.3110	946.3092	y3(-H_2O)	300.1559	300.1551
b6*	1059.3951	1059.3933	y2	188.0553	188.0744

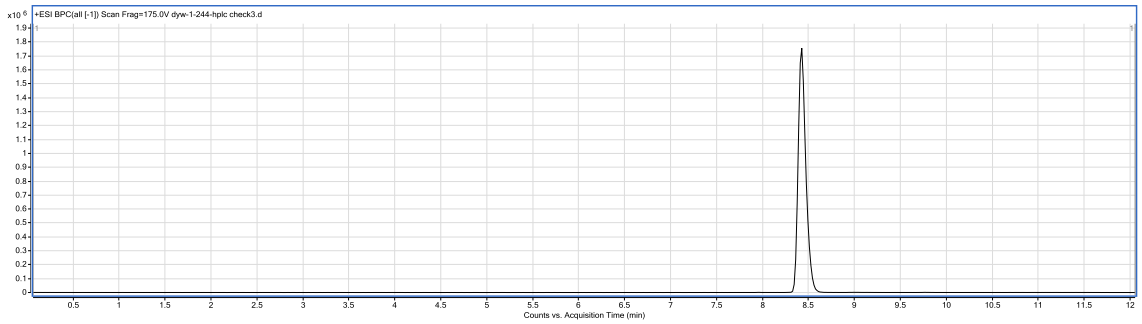


MS/MS analysis of modified peptide **11i**. *denotes fragments containing the modification.

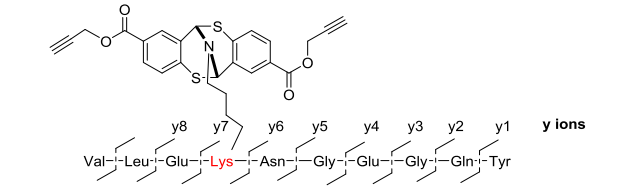
*Synthesis of **11j***

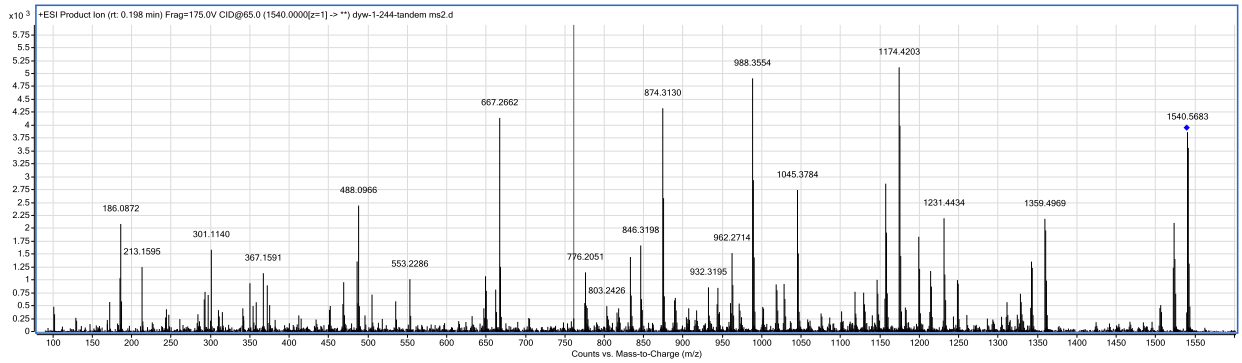


According to General Procedure **A** and **B**, peptide VLEKNGEGQY was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide VLEKNGEGQY was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 95% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **11j** as a white solid after lyophilization (8.8 mg, 71%). Analytical HPLC: t_R = 8.4 min (40 to 95% solvent B over 12 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{71}H_{90}N_{13}O_{22}S_2$ $[M+H]^+$ 1540.5765, found: 1540.5812.



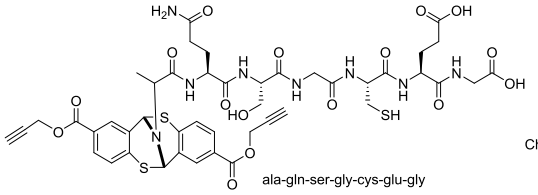
Ion	Cal (Da)	Found (Da)	Ion	Cal (Da)	Found (Da)
b1	101.0841	101.0709	y8*	1328.424	1328.4198
b2	213.1603	213.1595	y7*	1199.3809	1199.3732
b3	342.2029	342.2068	y6	667.2682	667.2662
b4*	874.3156	874.3131	y5	553.2258	553.2286
b5*	988.3585	988.3554	y4	496.2044	496.2050
b6*	1045.3794	1045.3784	y3	367.1618	367.1591
b7*	1174.422	1174.4203	y2	310.1403	310.1408
b8*	1231.4435	1231.4434	y1	136.0757	136.0751
b9*	1359.502	1359.4969			





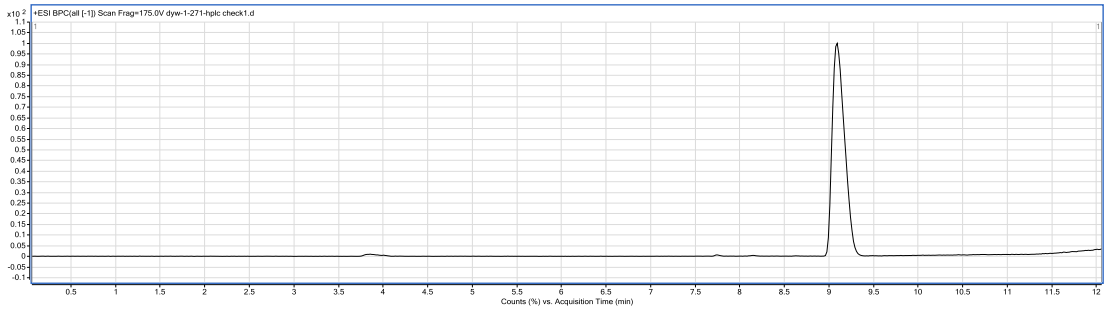
MS/MS analysis of modified peptide **11j**. *denotes fragments containing the modification.

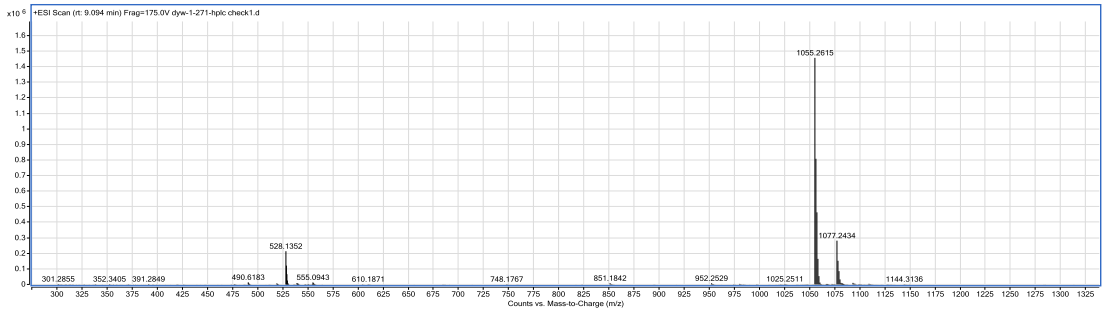
*Synthesis of **11k***

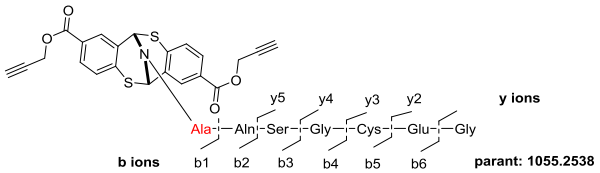


Chemical Formula: $C_{45}H_{50}N_8O_{16}S_3$
Exact Mass: 1054.2507

According to General Procedure **A** and **B**, peptide AQSGCEG was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide AQSGCEG was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 90% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **11k** as a white solid after lyophilization (7.2 mg, 85%). Analytical HPLC: t_R = 9.1 min (35 to 95% solvent B over 12 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{45}H_{51}N_8O_{16}S_3$ $[M+H]^+$ 1055.2585, found: 1055.2615.

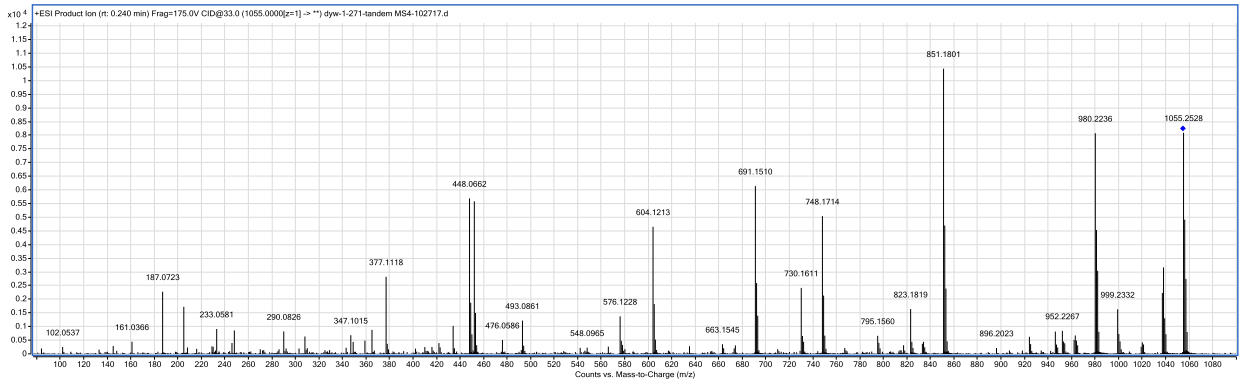






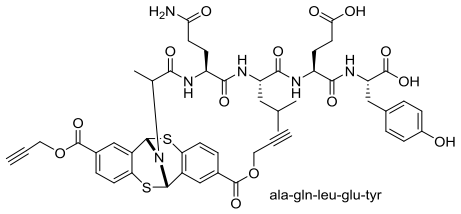
Ion	Cal (Da)	Found (Da)	Ion	Cal (Da)	Found (Da)
b1*	476.0626	476.0586	b6*	980.2265	980.2236

b2*	604.1212	604.1213	y2	205.0819	205.0769
b3*	691.1532	691.1510	y3(-H_2O)	290.0805	290.0826
b4*	748.1747	748.1714	y4(-H_2O)	347.1020	347.1015
b5*	851.1839	851.1801	y5	452.1446	452.1425



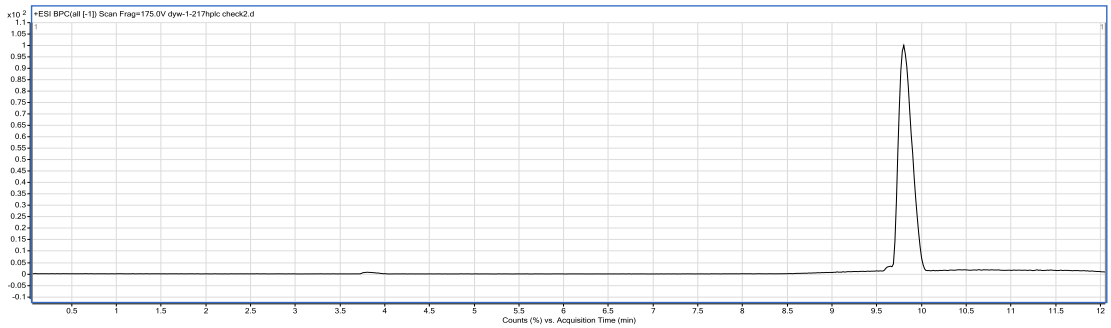
MS/MS analysis of modified peptide **11k**. *denotes fragments containing the modification.

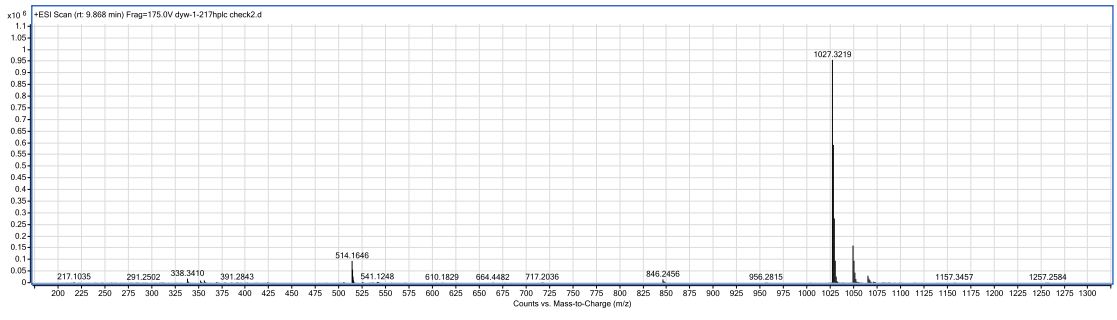
*Synthesis of **11l***



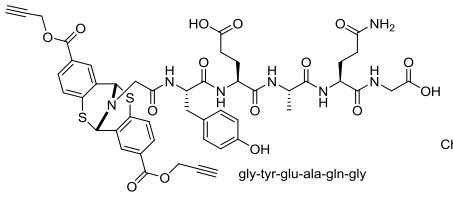
Chemical Formula: $C_{50}H_{54}N_6O_{14}S_2$
Exact Mass: 1026.3139

According to General Procedure **A** and **B**, peptide AQLEY was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide AQLEY was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 90% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **11l** as a white solid after lyophilization (6.6 mg, 81%). Analytical HPLC: t_R = 9.8 min (35 to 95% solvent B over 12 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{50}H_{55}N_6O_{14}S_2$ $[M+H]^+$ 1027.3218, found: 1027.3215.



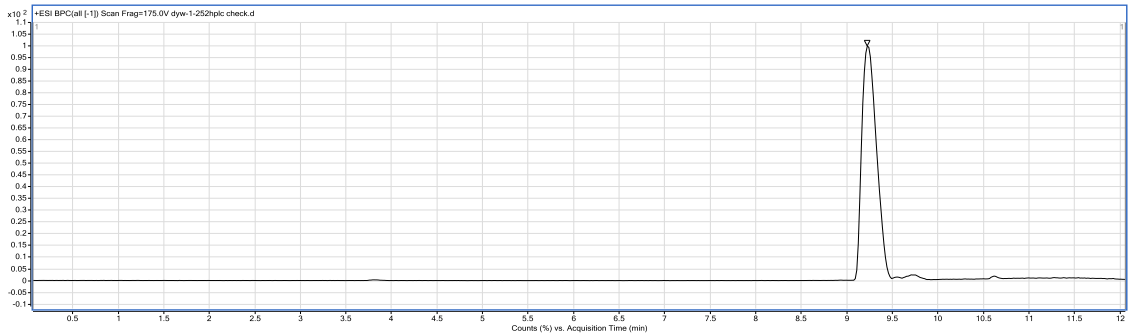


*Synthesis of **11m***

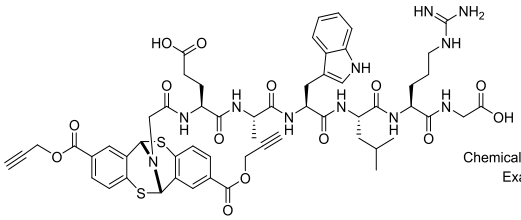


Chemical Formula: $C_{48}H_{49}N_7O_{15}S_2$
Exact Mass: 1027.2728

According to General Procedure **A** and **B**, peptide GYEAQG was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide GYEAQG was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (30 to 90% solvent B over 25 min, Higgins Analytical Proto 200 10 μm 250 × 20 nm C18 column) afforded the desired modified peptide **11m** as a white solid after lyophilization (6.1 mg, 75%). Analytical HPLC: t_R = 9.2 min (35 to 95% solvent B over 12 min, Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{48}H_{50}N_7O_{15}S_2$ $[M+H]^+$ 1028.2806, found: 1028.2805.

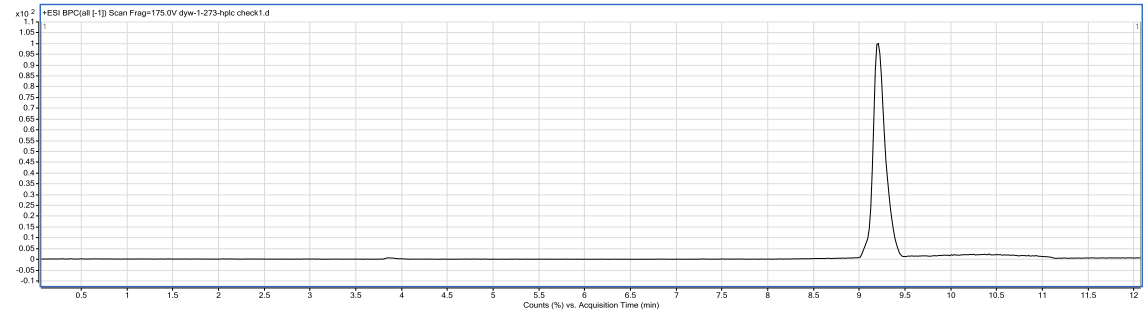


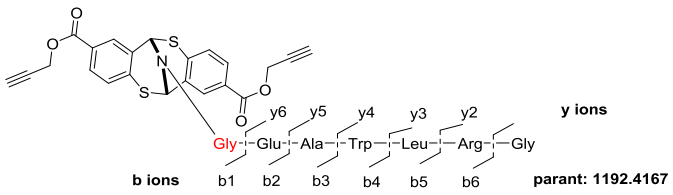
*Synthesis of **11n***



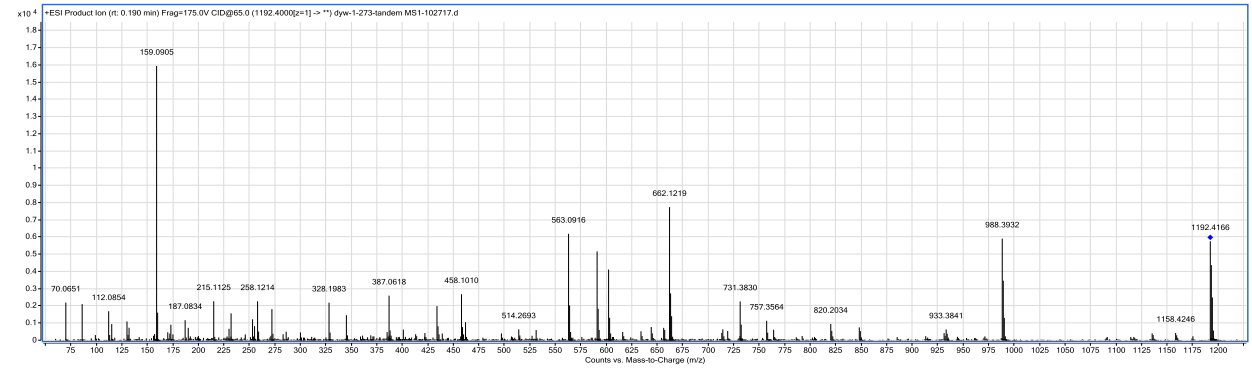
gly-glu-ala-trp-leu-arg-gly

According to General Procedure **A** and **B**, peptide GEAWLRG was prepared by a three-step protocol on a 0.05 mmol scale. The conjugation with peptide GEAWLRG was according to General Procedure **C**. Purification of the crude peptide using preparative HPLC (35 to 95% solvent B over 25 min, Higgins Analytical Proto 200 10 μm 250 × 20 nm C18 column) afforded the desired modified peptide **11n** as a white solid after lyophilization (8.5 mg, 89%). Analytical HPLC: t_R = 9.3 min (35 to 95% solvent B over 12 min, Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{57}H_{66}N_{11}O_{14}S_2$ $[M+H]^+$ 1192.4232, found: 1192.4256.



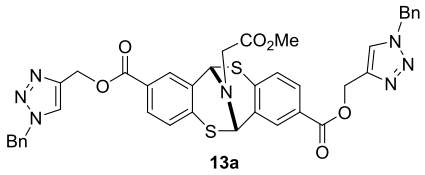


Ion	Cal (Da)	Found (Da)	Ion	Cal (Da)	Found (Da)
b1(-CO)*	434.0515	434.0514	y2(-H_2O)	215.1139	215.1125
b2*	591.0896	591.0881	y3	345.2245	345.2215
b3*	662.1267	662.1219	y4	531.3038	531.2947
b4(-CO)*	820.2105	820.2034	y5	602.3409	602.3394
b5*	933.3946	933.3941	y6	731.3835	731.3830
b6*(+H_2O)	1135.3912	1135.3900			



MS/MS analysis of modified peptide **11n**. *denotes fragments containing the modification.

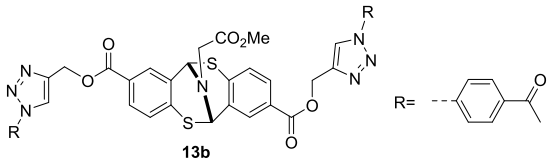
*Synthesis of **13a***



13a

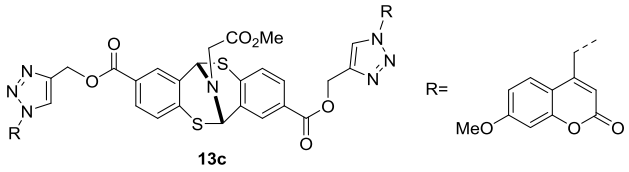
Compound **13a** was prepared according to General Procedure **D**. The obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1: 1) to give **13a** (13.2 mg, 87%) as a white solid. R_f = 0.28 (hexane: EtOAc= 1: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.87 (s, 2H), 7.68 (d, J = 8.4 Hz, 2H), 7.59 (s, 2H), 7.46 – 7.21 (m, 10H), 7.04 (d, J = 8.2 Hz, 2H), 5.62 (s, 2H), 5.52 (s, 4H), 5.40 (q, J = 12.8 Hz, 4H), 3.83 (d, J = 17.1 Hz, 1H), 3.76 (s, 3H), 3.49 (d, J = 17.0 Hz, 1H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 169.37, 165.55, 143.09, 136.36, 134.28, 132.40, 130.44, 129.15, 128.84, 128.62, 128.15, 127.62, 126.37, 123.88, 61.78, 58.01, 54.76, 54.23, 52.29. **IR** (thin film): cm^{-1} 2981, 2889, 2253, 1746, 1715, 1603, 1375, 1239, 1110, 1046, 904, 723, 648. **HRMS** (ESI+): calculated for $C_{39}H_{34}N_7O_6S_2$ $[M+H]^+$ 760.2012, found: 760.2015.

*Synthesis of **13b***



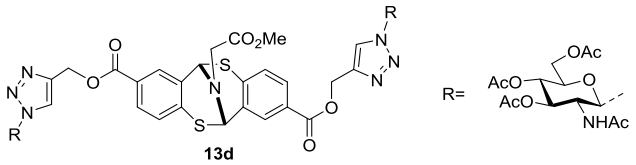
Compound **13b** was prepared according to General Procedure **D**. The obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1: 2) to give **13b** (8.2 mg, 86%) as a white solid. R_f = 0.33 (hexane: EtOAc= 1: 2). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.19 (s, 2H), 8.12 (d, J = 8.4 Hz, 2H), 7.97 – 7.82 (m, 3H), 7.73 (d, J = 8.6 Hz, 2H), 7.07 (d, J = 8.5 Hz, 2H), 5.65 (s, 2H), 5.53 (q, J = 12.6 Hz, 4H), 3.84 (d, J = 17.2 Hz, 1H), 3.75 (s, 3H), 3.50 (d, J = 16.8 Hz, 1H), 2.66 (s, 6H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 196.50, 169.37, 165.66, 143.91, 139.86, 137.04, 136.60, 132.50, 130.53, 130.10, 128.69, 127.75, 126.27, 122.22, 120.18, 61.85, 57.86, 54.78, 29.69, 26.68. **IR** (thin film): cm^{-1} 2952, 2923, 2854, 1717, 1685, 1604, 1442, 1360, 1263, 1243, 1177, 1112, 960, 840. **HRMS** (ESI+): calculated for $C_{41}H_{34}N_7O_8S_2$ $[M+H]^+$ 816.1910, found: 816.1903.

*Synthesis of **13c***



Compound **13c** was prepared according to General Procedure **D**. The obtained residue was purified by silica gel column chromatography (CH_2Cl_2: MeOH= 10: 1) to give **13c** (17.7 mg, 93%) as a white solid. R_f = 0.55 (CH_2Cl_2: MeOH= 10: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.84 (d, *J* = 1.7 Hz, 2H), 7.79 (s, 2H), 7.62 (dd, *J* = 8.3, 1.8 Hz, 2H), 7.49 (d, *J* = 8.8 Hz, 2H), 6.98 (d, *J* = 8.3 Hz, 1H), 6.82 (dd, *J* = 8.8, 2.5 Hz, 2H), 6.78 (d, *J* = 2.5 Hz, 1H), 5.84 (s, 1H), 5.65 (s, 2H), 5.62 (s, 1H), 5.41 (q, *J* = 12.8 Hz, 2H), 3.84 (s, 3H), 3.80 (s, 1H), 3.72 (s, 2H), 3.49 (d, *J* = 17.2 Hz, 1H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 169.41, 165.46, 163.24, 160.19, 155.60, 147.83, 136.49, 132.46, 130.41, 128.55, 127.61, 126.13, 124.76, 124.55, 112.91, 111.89, 110.35, 101.34, 61.71, 57.87, 55.83, 55.82, 54.66, 52.25, 50.18. **IR** (thin film): cm^{-1} 3148, 3093, 2958, 2247, 1710, 1610, 1398, 1285, 1262, 1208, 1148, 1110, 1049, 989, 841, 760, 730. **HRMS** (ESI+): calculated for $C_{47}H_{37}N_7O_{12}S_2Na$ $[M+Na]^+$ 978.1839, found: 978.1823.

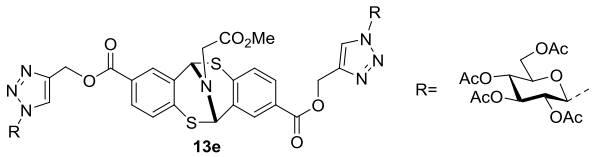
*Synthesis of **13d***



13d

Compound **13d** was prepared according to General Procedure **D**. The obtained residue was purified by silica gel column chromatography (CH_2Cl_2: MeOH= 10: 1) to give **13d** (20.2 mg, 82%) as a white solid. R_f = 0.41 (CH_2Cl_2: MeOH= 10: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.01 (s, 1H), 7.91 (s, 1H), 7.69 (d, J = 8.4 Hz, 2H), 7.02 (d, J = 7.6 Hz, 2H), 6.06 (d, J = 9.3 Hz, 2H), 5.66 (d, J = 8.4 Hz, 2H), 5.53 – 5.38 (m, 3H), 5.24 (t, J = 9.4 Hz, 2H), 4.62 – 4.51 (m, 2H), 4.28 (d, J = 9.7 Hz, 1H), 4.14 (d, J = 12.2 Hz, 1H), 4.02 (d, J = 11.8 Hz, 2H), 3.84 (d, J = 17.1 Hz, 1H), 3.74 (s, 1H), 3.50 (dd, J = 16.8, 6.0 Hz, 1H), 2.06 (s, 9H), 1.73 (d, J = 7.8 Hz, 6H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 170.77, 170.76, 170.58, 170.53, 169.48, 169.46, 169.31, 165.47, 143.17, 143.15, 136.52, 136.50, 132.59, 132.54, 130.54, 130.50, 128.59, 127.63, 127.60, 126.24, 126.22, 123.39, 123.33, 85.86, 74.98, 72.17, 67.96, 61.70, 54.75, 53.49, 52.29, 52.27, 22.84, 22.80, 20.68, 20.61, 20.57. **IR** (thin film): cm^{-1} 2980, 2853, 2251, 1745, 1716, 1602, 1369, 1229, 1106, 1043, 906, 728, 647. **HRMS** (ESI+): calculated for $C_{53}H_{59}N_9O_{22}S_2Na$ $[M+Na]^+$ 1260.3114, found: 1260.3106.

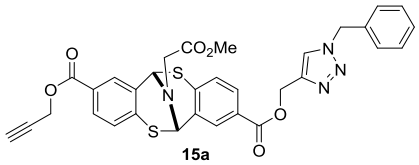
*Synthesis of **13e***



13e

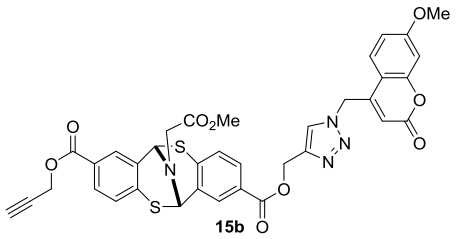
Compound **13e** was prepared according to General Procedure **D**. The obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1: 4) to give **13e** (20.6 mg, 83%) as a white solid. R_f = 0.6 (hexane: EtOAc= 1: 4). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.91 (s, 4H), 7.71 (d, *J* = 8.3 Hz, 2H), 7.06 (d, *J* = 8.3 Hz, 2H), 5.88 (d, *J* = 8.5 Hz, 2H), 5.65 (s, 2H), 5.52 – 5.34 (m, 8H), 5.24 (t, *J* = 9.2 Hz, 2H), 4.30 (dd, *J* = 12.6, 4.9 Hz, 2H), 4.15 (d, *J* = 12.4 Hz, 2H), 4.01 (d, *J* = 10.2 Hz, 2H), 3.84 (d, *J* = 17.1 Hz, 1H), 3.75 (s, 3H), 3.50 (d, *J* = 17.3 Hz, 1H), 2.08 (s, 3H), 2.06 (s, 3H), 2.02 (s, 6H), 1.85 (d, *J* = 9.0 Hz, 6H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 170.48, 169.90, 169.31, 168.87, 168.84, 165.45, 143.52, 136.46, 132.47, 130.51, 128.67, 128.65, 127.68, 126.31, 126.29, 122.46, 85.76, 75.22, 72.56, 70.27, 67.62, 61.80, 61.45, 57.82, 54.80, 52.30, 20.68, 20.51, 20.48, 20.15, 20.12. **IR** (thin film): cm^{-1} 2980, 2889, 2254, 1749, 1717, 1602, 1369, 1216, 1104, 1041, 906, 729. **HRMS** (ESI+): calculated for $C_{53}H_{57}N_7O_{24}S_2Na$ $[M+Na]^+$ 1262.2794, found: 1262.2776.

*Synthesis of **15a***



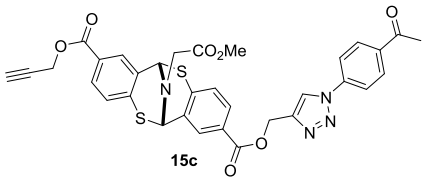
Compound **15a** was prepared according to General Procedure **E**. The obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1.5: 1) to give **15a** (7.2 mg, 58%) as a white solid. R_f = 0.25 (hexane: EtOAc= 2: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.90 (d, *J* = 15.4 Hz, 2H), 7.71 (dd, *J* = 17.3, 8.3 Hz, 2H), 7.58 (s, 1H), 7.33 (dd, *J* = 36.0, 6.7 Hz, 4H), 7.07 (t, *J* = 8.4 Hz, 2H), 5.65 (d, *J* = 10.5 Hz, 2H), 5.52 (s, 2H), 5.41 (q, *J* = 12.7 Hz, 2H), 4.95 – 4.83 (m, 2H), 3.85 (d, *J* = 17.2 Hz, 1H), 3.51 (d, *J* = 17.2 Hz, 1H), 2.52 (s, 1H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 169.39, 165.59, 164.96, 143.13, 136.64, 136.37, 134.29, 132.40, 130.49, 129.18, 128.88, 128.70, 128.19, 127.75, 127.69, 126.44, 126.05, 123.89, 75.14, 61.86, 58.05, 54.82, 54.27, 52.52, 52.33. **IR** (thin film): cm^{-1} 3283, 2982, 2891, 1747, 1717, 1604, 1373, 1241, 1112, 1047, 906, 726. **HRMS** (ESI+): calculated for $C_{32}H_{27}N_4O_6S_2$ $[M+H]^+$ 627.1372, found: 627.1354.

*Synthesis of **15b***



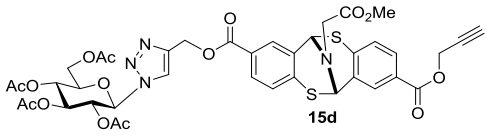
Compound **15b** was prepared according to General Procedure **E**. The obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1: 2) to give **15b** (10.4 mg, 72%) as a white solid. R_f = 0.31 (hexane: EtOAc= 1: 2). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.91 (d, *J* = 12.9 Hz, 2H), 7.77 – 7.66 (m, 1H), 7.51 (d, *J* = 9.7 Hz, 1H), 7.08 (dd, *J* = 8.3, 5.3 Hz, 2H), 6.91 – 6.83 (m, 1H), 5.91 (s, 1H), 5.66 (d, *J* = 6.2 Hz, 2H), 5.44 (q, *J* = 12.7 Hz, 1H), 4.96 – 4.81 (m, 2H), 3.87 (s, 2H), 3.83 (s, 1H), 3.76 (s, 1H), 3.52 (d, *J* = 17.2 Hz, 1H), 2.51 (t, *J* = 2.0 Hz, 1H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 169.40, 165.59, 164.95, 163.37, 160.18, 155.77, 147.61, 143.90, 136.63, 136.59, 132.50, 132.48, 130.54, 128.71, 127.76, 126.23, 126.06, 124.55, 113.05, 112.19, 110.40, 101.47, 75.14, 61.85, 57.89, 55.86, 54.80, 52.52, 52.35, 50.34, 38.13, 31.91, 31.21. **IR** (thin film): cm^{-1} 3278, 2841, 1721, 1598, 1521, 1275, 1258, 1137, 1114, 845, 673. **HRMS** (ESI+): calculated for $C_{36}H_{29}N_4O_9S_2$ $[M+H]^+$ 725.1376, found: 725.1364.

*Synthesis of **15c***



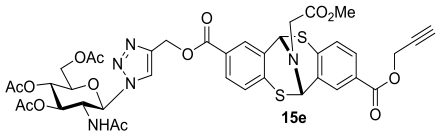
Compound **15c** was prepared according to General Procedure **E**. The obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1: 1) to give **15c** (8.2 mg, 63%) as a white solid. R_f = 0.38 (hexane: EtOAc= 1: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.20 (s, 1H), 8.13 (d, J = 8.7 Hz, 2H), 7.93 (s, 2H), 7.88 (d, J = 8.7 Hz, 2H), 7.74 (d, J = 8.4 Hz, 2H), 7.09 (dd, J = 8.3, 2.7 Hz, 2H), 5.66 (s, 2H), 5.53 (q, J = 12.7 Hz, 2H), 4.97 – 4.81 (m, 2H), 3.85 (d, J = 17.2 Hz, 1H), 3.76 (s, 3H), 3.52 (d, J = 17.2 Hz, 1H), 2.66 (s, 3H), 2.51 (t, J = 2.4 Hz, 1H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 196.52, 169.39, 165.69, 164.94, 143.93, 139.87, 137.03, 136.65, 136.60, 132.52, 132.48, 130.54, 130.10, 128.70, 127.76, 126.25, 126.09, 122.23, 120.19, 77.57, 75.15, 61.87, 57.87, 54.80, 52.53, 52.34, 26.68. **IR** (thin film): cm^{-1} 3281, 2951, 2857, 1718, 1687, 1603, 1442, 1261, 1114, 1039, 961, 731. **HRMS** (ESI+): calculated for $C_{33}H_{27}N_4O_7S_2$ $[M+H]^+$ 655.1321, found: 655.1309.

*Synthesis of **15d***



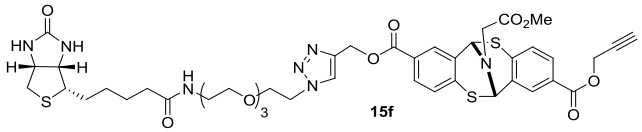
Compound **15d** was prepared according to General Procedure **E**. The obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1: 1.5) to give **15d** (12.9 mg, 74%) as a white solid. R_f = 0.41 (hexane: EtOAc= 1: 1.5). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.92 (s, 3H), 7.72 (t, *J* = 6.6 Hz, 2H), 7.08 (d, *J* = 8.3 Hz, 2H), 5.88 (d, *J* = 8.1 Hz, 1H), 5.66 (d, *J* = 5.2 Hz, 2H), 5.51 – 5.36 (m, 4H), 5.24 (t, *J* = 9.1 Hz, 1H), 4.95 – 4.81 (m, 2H), 4.30 (dd, *J* = 12.8, 4.8 Hz, 1H), 4.15 (d, *J* = 12.9 Hz, 1H), 4.01 (d, *J* = 10.3 Hz, 1H), 3.85 (d, *J* = 17.4 Hz, 1H), 3.52 (d, *J* = 17.3 Hz, 1H), 2.51 (s, 1H), 2.08 (s, 3H), 2.06 (s, 3H), 2.02 (s, 3H), 1.85 (d, *J* = 9.8 Hz, 3H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 170.47, 169.89, 169.37, 169.30, 168.83, 165.44, 164.93, 143.52, 136.62, 136.45, 132.52, 132.43, 130.52, 128.68, 127.71, 126.33, 126.03, 122.40, 85.75, 77.57, 75.21, 75.13, 72.56, 70.27, 67.62, 61.82, 61.46, 54.79, 52.50, 52.32, 20.68, 20.50, 20.48, 20.15, 20.11. **IR** (thin film): cm^{-1} 3285, 2955, 2926, 1713, 1602, 1559, 1370, 1260, 1210, 1106, 839, 760. **HRMS** (ESI+): calculated for $C_{39}H_{38}N_4O_{15}S_2Na$ $[M+Na]^+$ 889.1673, found: 889.1655.

*Synthesis of **15e***



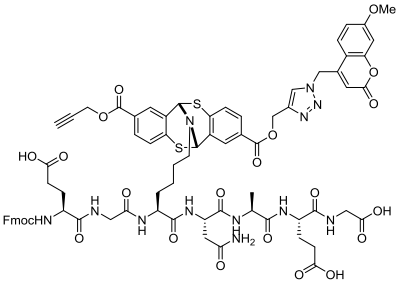
Compound **15e** was prepared according to General Procedure **E**. The obtained residue was purified by silica gel column chromatography (CH_2Cl_2: MeOH= 15: 1) to give **15e** (13.3 mg, 77%) as a white solid. R_f = 0.54 (CH_2Cl_2: MeOH= 15: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.99 (s, 1H), 7.93 (s, 2H), 7.73 (d, *J* = 8.2 Hz, 2H), 7.08 (d, *J* = 8.2 Hz, 2H), 6.03 (d, *J* = 10.2 Hz, 1H), 5.90 (s, 1H), 5.68 (s, 2H), 5.52 – 5.36 (m, 3H), 5.24 (t, *J* = 9.7 Hz, 1H), 4.97 – 4.81 (m, 2H), 4.63 – 4.49 (m, 1H), 4.28 (d, *J* = 12.5 Hz, 1H), 4.14 (d, *J* = 12.1 Hz, 1H), 3.99 (s, 1H), 3.85 (d, *J* = 17.9 Hz, 1H), 3.76 (s, 3H), 3.52 (d, *J* = 17.4 Hz, 1H), 2.51 (s, 1H), 2.07 (s, 9H), 1.75 (d, *J* = 12.9 Hz, 3H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 170.84, 170.57, 169.42, 169.26, 165.43, 164.98, 143.24, 136.69, 136.43, 132.46, 130.56, 128.69, 127.71, 126.01, 85.97, 77.59, 75.14, 72.13, 67.79, 61.85, 61.78, 61.64, 54.81, 53.64, 52.52, 52.33, 22.86, 20.71, 20.59, 20.56. **IR** (thin film): cm^{-1} 3283, 2980, 2926, 2118, 1745, 1717, 1601, 1370, 1230, 1108, 1043, 834, 761. **HRMS** (ESI+): calculated for $C_{39}H_{39}N_5O_{14}S_2Na$ $[M+Na]^+$ 888.1833, found: 888.1830.

*Synthesis of **15f***



Compound **15f** was prepared according to General Procedure **E**. The obtained residue was purified by preparative thin layer chromatography (CH_2Cl_2: MeOH= 10: 1) to give **15f** (11.6 mg, 62%) as a white solid (containing 20% inseparable biotin-PEG3-azide). R_f = 0.35 (CH_2Cl_2: MeOH= 10: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.92 (d, *J* = 4.0 Hz, 2H), 7.89 (s, 1H), 7.72 (t, *J* = 8.8 Hz, 2H), 7.08 (dd, *J* = 8.2, 3.9 Hz, 2H), 6.53 (t, *J* = 5.0 Hz, 1H), 5.83 (s, 1H), 5.68 (d, *J* = 6.5 Hz, 2H), 5.44 (q, *J* = 12.7 Hz, 2H), 5.00 (s, 1H), 4.95 – 4.82 (m, 2H), 4.56 (t, *J* = 4.9 Hz, 2H), 4.50 – 4.45 (m, 1H), 4.30 (s, 1H), 3.90 (d, *J* = 4.9 Hz, 2H), 3.84 (s, 1H), 3.76 (s, 3H), 3.68 – 3.61 (m, 6H), 3.56 – 3.51 (m, 6H), 3.42 (dd, *J* = 9.9, 4.9 Hz, 3H), 3.16 – 3.11 (m, 1H), 2.91 – 2.86 (m, 1H), 2.73 (d, *J* = 8.2 Hz, 1H), 2.52 (s, 1H), 2.19 (t, *J* = 4.0 Hz, 2H), 1.41 (m, 2H), 1.30 (m, 3H). **IR** (thin film): cm^{-1} 3283, 2924, 2866, 2245, 1713, 1697, 1598, 1452, 1440, 1287, 1242, 1179, 1105, 1025, 910, 761, 729. **HRMS** (ESI+): calculated for $C_{43}H_{52}N_7O_{11}S_3$ $[M+H]^+$ 938.2887, found: 938.2908.

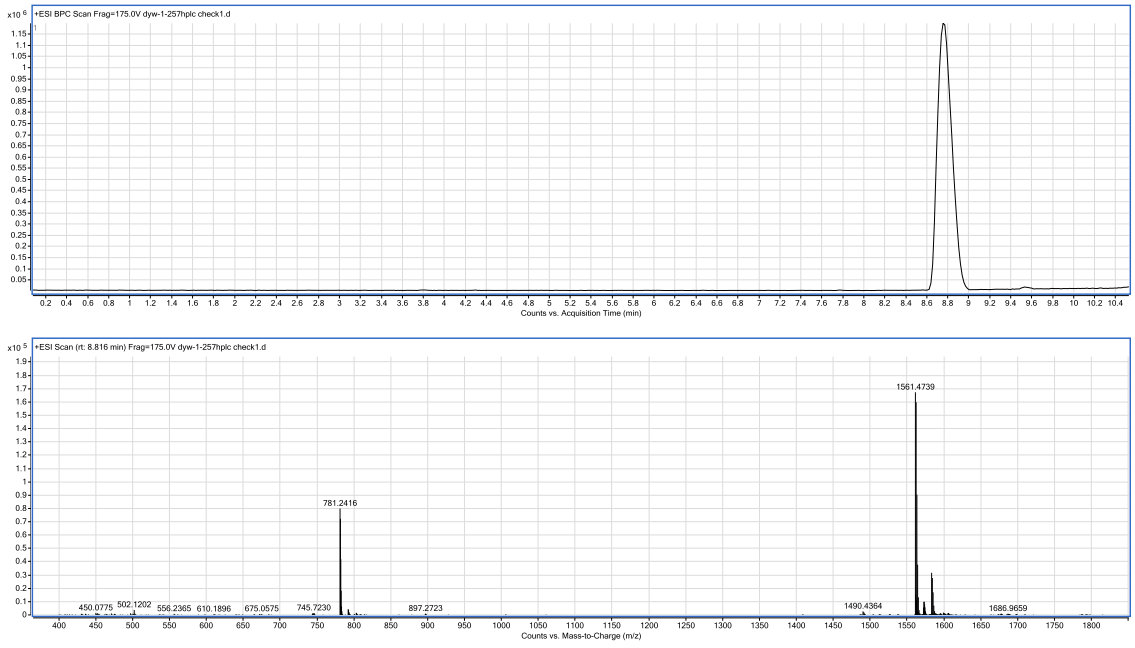
*Synthesis of **15g***



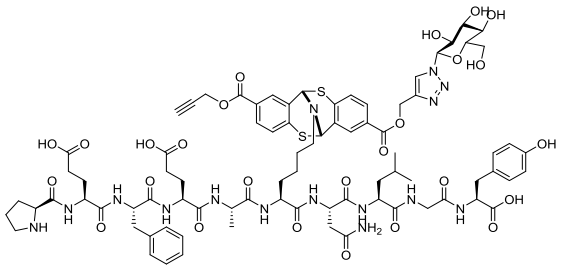
Fmoc-glu-gly-lys-asn-ala-glu-gly

Chemical Formula: $C_{75}H_{76}N_{12}O_{22}S_2$
Exact Mass: 1560.4639

Peptide **15g** was prepared according to General Procedure **F**. Purification of the crude using preparative HPLC (40 to 98% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **15g** as a white solid after lyophilization (3.8 mg, 62%). Analytical HPLC: t_R = 8.8 min (55 to 98% solvent B over 12 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{75}H_{77}N_{12}O_{22}S_2$ $[M+H]^+$ 1561.4717, found: 1561.4739.



*Synthesis of **15h***

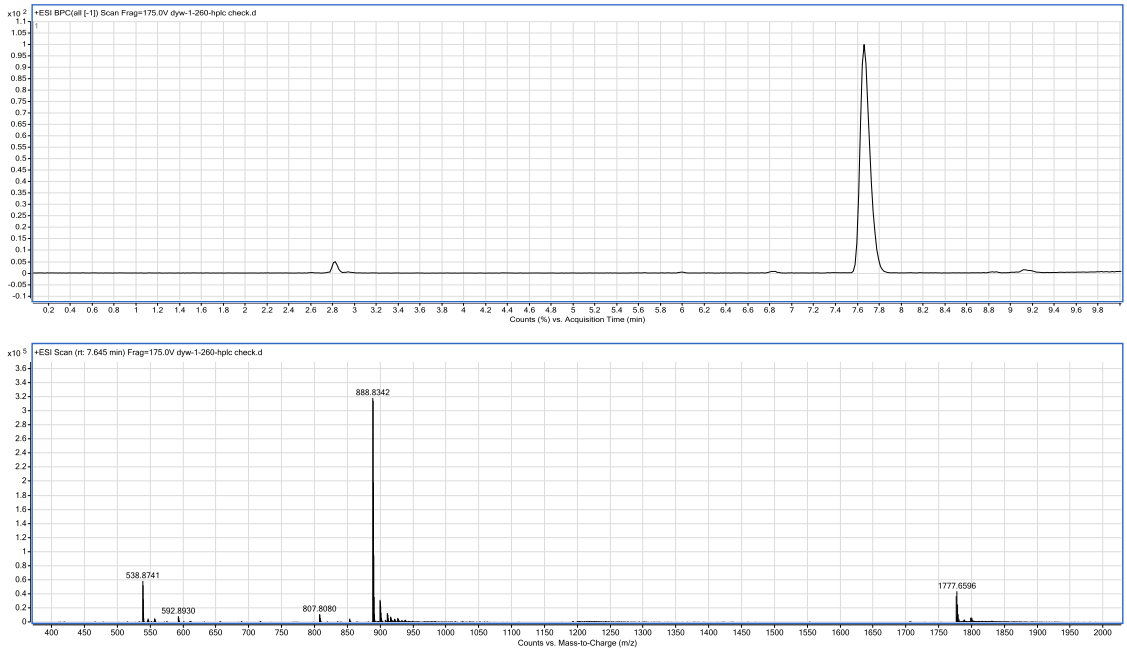


pro-glu-phe-glu-ala-lys-asn-leu-gly-tyr

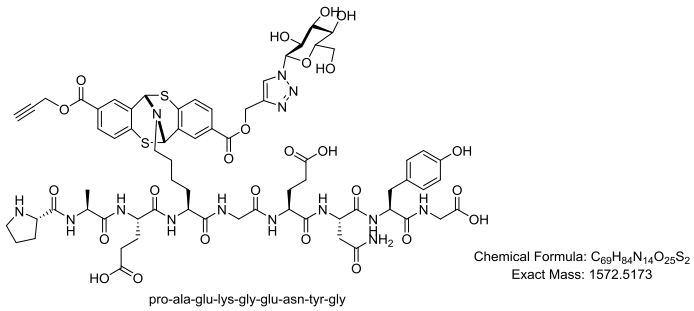
Chemical Formula: $C_{82}H_{101}N_{15}O_{26}S_2$

Exact Mass: 1775.6484

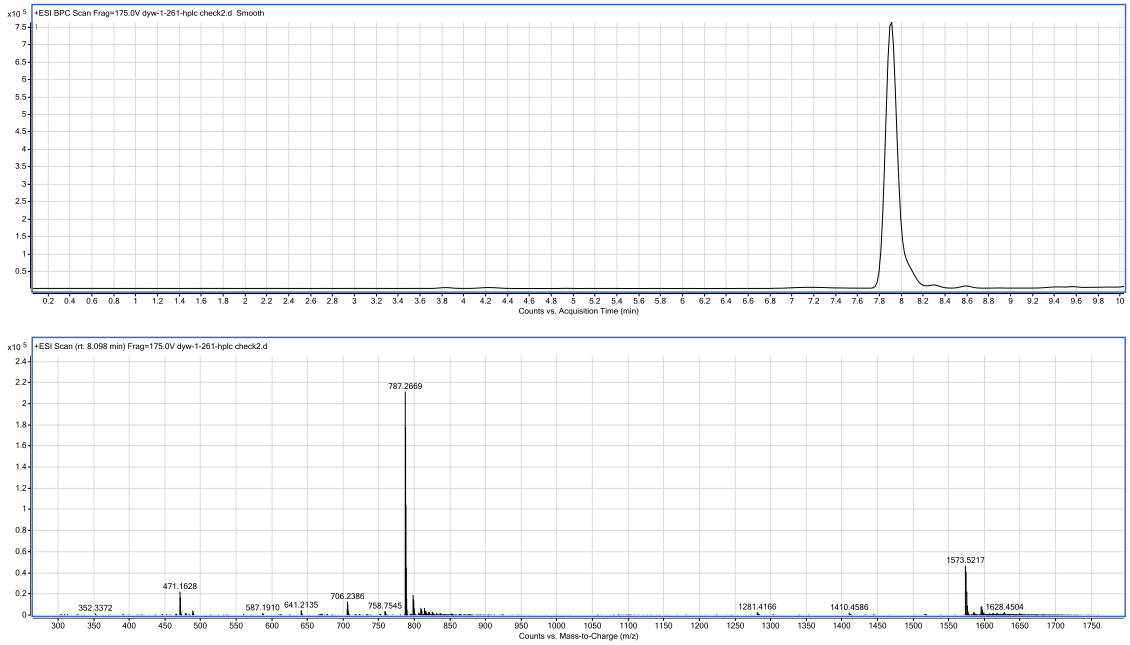
Peptide **15h** was prepared according to General Procedure **F**. Purification of the crude using preparative HPLC (30 to 85% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **15h** as a white solid after lyophilization (4.6 mg, 65%). Analytical HPLC: t_R = 7.7 min (35 to 95% solvent B over 10 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{82}H_{102}N_{15}O_{26}S_2$ $[M+H]^+$ 1776.6562, found: 1776.6596; $C_{82}H_{103}N_{15}O_{26}S_2$ $[M+2H]^{2+}$ 888.8320, found:888.8342.



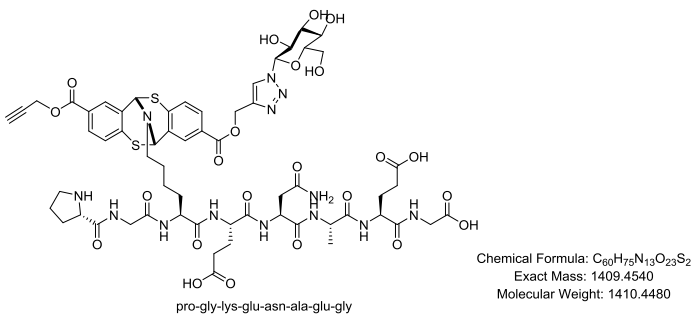
*Synthesis of **15i***



Peptide **15i** was prepared according to General Procedure **F**. Purification of the crude using preparative HPLC (20 to 80% solvent B over 25 min, Higgins Analytical Proto 200 10 µm 250 × 20 nm C18 column) afforded the desired modified peptide **15i** as a white solid after lyophilization (3.9 mg, 61%). Analytical HPLC: t_R = 7.9 min (35 to 95% solvent B over 10 min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{69}H_{85}N_{14}O_{25}S_2$ $[M+H]^+$ 1573.5252, found: 1573.5225.

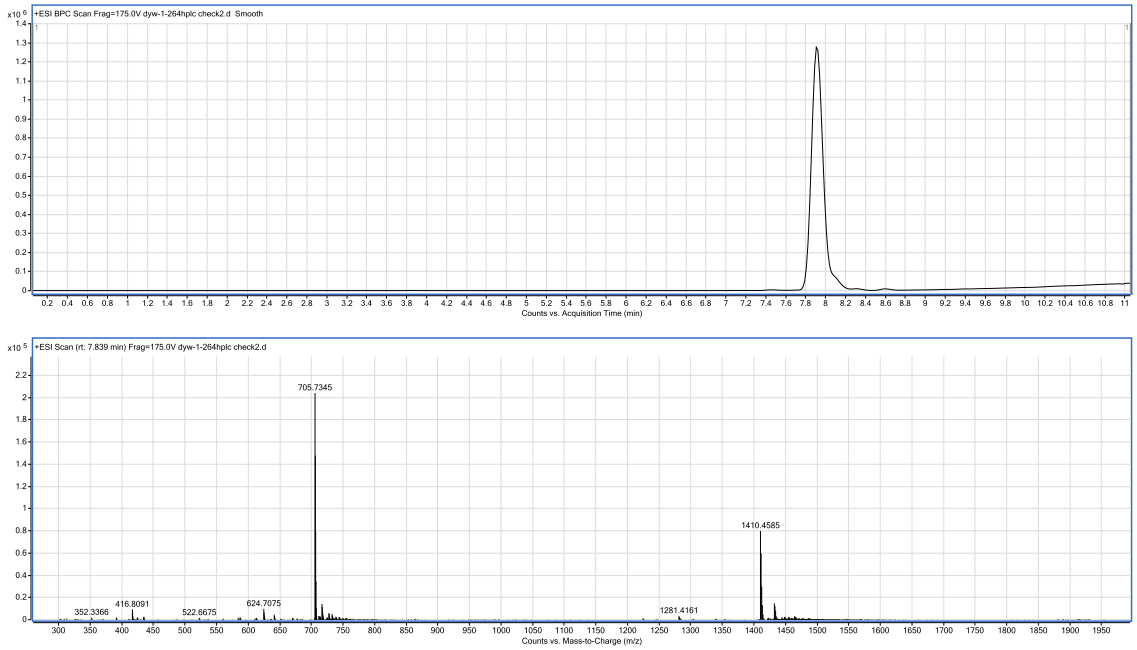


*Synthesis of **15j***

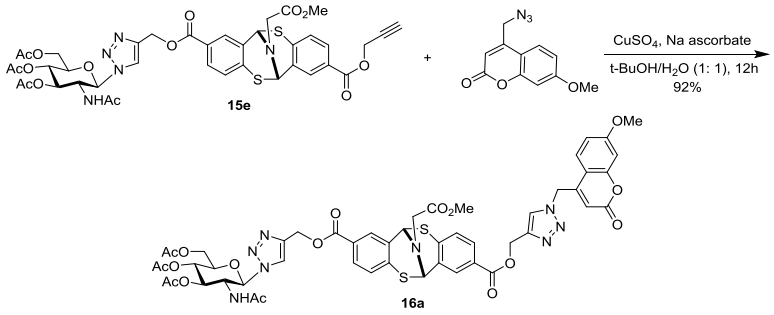


Peptide **15j** was prepared according to General Procedure **F**. Purification of the crude using preparative

HPLC (20 to 80% solvent B over 25 min, Higgins Analytical Proto 200 10 μm 250 × 20 nm C18 column) afforded the desired modified peptide **15j** as a white solid after lyophilization (3.8 mg, 68%). Analytical HPLC: t_R = 7.9 min (35 to 95% solvent B over 11 min, Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column); HRMS (ESI+): calculated for $C_{60}H_{76}N_{13}O_{23}S_2$ $[M+H]^+$ 1410.4618, found: 1410.4585.



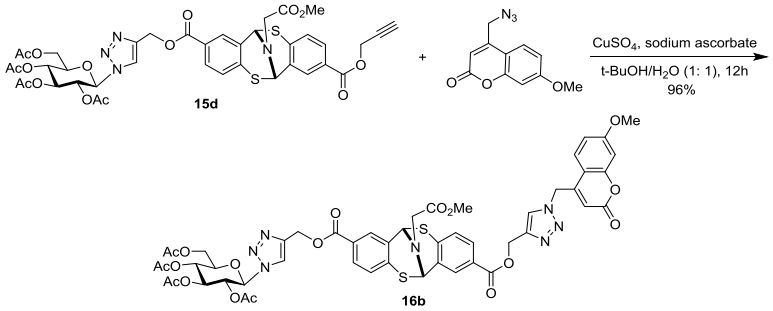
*Synthesis of **16a***



Compound **15e** (5 mg, 0.00578 mmol) and 4-(azidomethyl)-7-methoxy-2H-chromen-2-one (2.7 mg, 0.0115 mmol) was dissolved in 1 mL of tert-butanol/water (1: 1). After degassing for 20 min, 6 μL of $CuSO_4$ (0.1 M in deionized water) was added, followed by the addition of 6 μL of sodium ascorbate (0.2 M in deionized water) and 6 μL of TBTA (0.2 M in *tert*-butanol). The reaction mixture was stirred at room temperature for 12 h. Then the mixture was diluted with ethyl acetate (10 mL), washed with water and brine (6 mL×2), dried over Na_2SO_4. The solvent was removed in vacuo. The obtained residue was purified by silica gel column chromatography (CH_2Cl_2: MeOH= 10: 1) to give **16a** (5.8 mg, 92%) as a white solid. R_f = 0.52 (CH_2Cl_2: MeOH= 10: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.99 (s, 1H), 7.90 (d, J = 14.3 Hz, 2H), 7.74 (s,

1H), 7.69 (t, J = 9.2 Hz, 2H), 7.51 (d, J = 9.5 Hz, 1H), 7.05 (dd, J = 8.3, 3.7 Hz, 2H), 6.90 – 6.83 (m, 2H), 6.02 (d, J = 9.9 Hz, 1H), 5.91 (s, 1H), 5.83 (dd, J = 9.1, 3.3 Hz, 1H), 5.66 (d, J = 5.6 Hz, 4H), 5.24 (t, J = 10.3 Hz, 1H), 4.56 (dd, J = 19.8, 10.1 Hz, 1H), 4.29 (dd, J = 12.7, 4.8 Hz, 1H), 4.14 (d, J = 12.7 Hz, 1H), 3.99 (dd, J = 8.9, 3.7 Hz, 1H), 3.87 (s, 3H), 3.82 (s, 1H), 3.75 (s, 3H), 3.51 (dd, J = 17.2, 2.3 Hz, 1H), 2.09 – 2.04 (m, 10H), 1.75 (d, J = 10.8 Hz, 3H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 170.79, 170.55, 169.42, 169.26, 165.60, 165.39, 163.36, 160.21, 155.75, 147.67, 143.89, 132.46, 130.54, 130.49, 128.66, 128.63, 127.68, 126.16, 124.61, 124.57, 113.04, 112.17, 110.41, 101.46, 85.93, 75.08, 72.13, 67.81, 61.81, 61.64, 57.88, 55.86, 54.78, 53.64, 52.30, 50.32, 22.91, 22.87, 20.71, 20.59, 20.57. **IR** (thin film): cm^{-1} 2927, 2857, 2248, 2118, 1742, 1717, 1611, 1605, 1557, 1373, 1233, 1105, 1045, 910, 761, 732. **HRMS** (ESI+): calculated for $C_{50}H_{49}N_8O_{17}S_2$ $[M+H]^+$ 1097.2657, found: 1097.2635.

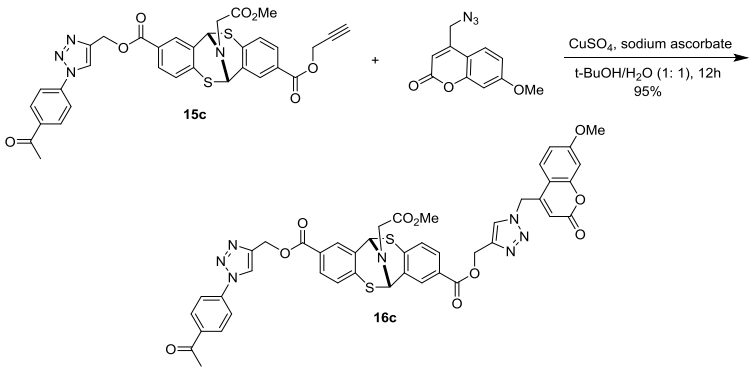
*Synthesis of **16b***



Compound **15d** (10 mg, 0.0115 mmol) and 4-(azidomethyl)-7-methoxy-2H-chromen-2-one (5.9 mg, 0.0254 mmol) was dissolved in 1.5 mL of tert-butanol/water (1: 1). After degassing for 20 min, 10 μL of $CuSO_4$ (0.1 M in deionized water) was added, followed by the addition 10 μL of sodium ascorbate (0.2 M in deionized water) and 10 μL of TBTA (0.1 M in *tert*-butanol). The reaction mixture was stirred at room temperature for 12 h. Then the mixture was diluted with ethyl acetate (15 mL), washed with water and brine (8 mL×2), dried over Na_2SO_4. The solvent was removed in vacuo. The obtained residue was purified by silica gel column chromatography (CH_2Cl_2: MeOH= 15: 1) to give **16b** (12.1 mg, 96%) as a white solid. R_f = 0.46 (CH_2Cl_2: MeOH= 15: 1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.90 (d, *J* = 14.2 Hz, 3H), 7.74 (s, 1H), 7.68 (d, *J* = 8.9 Hz, 2H), 7.50 (d, *J* = 8.3 Hz, 1H), 7.05 (s, 2H), 6.85 (s, 2H), 5.88 (d, *J* = 12.0 Hz, 2H), 5.65 (d, *J* = 7.7 Hz, 4H), 5.50 – 5.38 (m, 6H), 5.24 (t, *J* = 8.9 Hz, 1H), 4.30 (d, *J* = 12.6 Hz, 1H), 4.15 (d, *J* = 14.5 Hz, 1H), 4.01 (d, *J* = 10.7 Hz, 1H), 3.87 (s, 3H), 3.82 (s, 1H), 3.75 (s, 3H), 3.50 (d, *J* = 17.7 Hz, 1H), 2.07 (s, 3H), 2.02 (s, 3H), 1.85 (d, *J* = 9.1 Hz, 3H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 170.49, 169.90, 169.71, 169.39, 169.31, 165.58, 165.44, 163.36, 160.17, 155.76, 147.63, 132.51, 130.51, 128.66, 127.71, 126.30, 126.21, 124.60, 124.55, 113.03, 112.19, 110.40, 101.46, 85.78, 75.23, 72.57, 70.28, 67.63, 61.82, 61.47,

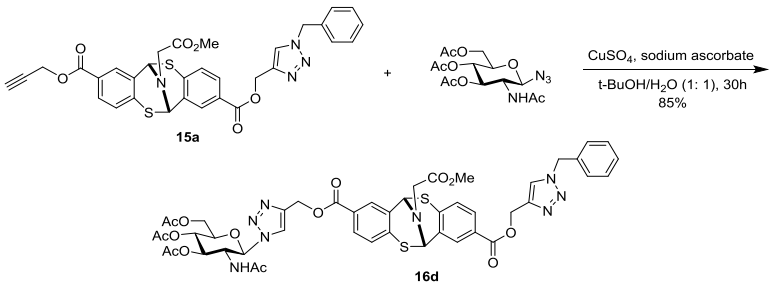
57.89, 55.86, 54.79, 52.31, 50.33, 20.69, 20.51, 20.49, 20.13. **IR** (thin film): cm^{-1} 2955, 2927, 2855, 1748, 1717, 1612, 1375, 1286, 1262, 1224, 1108, 1044, 914, 760, 732. **HRMS** (ESI+): calculated for $C_{50}H_{48}N_7O_{18}S_2$ $[M+H]^+$ 1098.2497, found: 1098.2484.

*Synthesis of **16c***



Compound **15c** (5 mg, 0.00764 mmol) and 4-(azidomethyl)-7-methoxy-2H-chromen-2-one (3.9 mg, 0.0168 mmol) was dissolved in 1 mL of *tert*-butanol/water (1: 1). After degassing for 20 min, 8 μL of $CuSO_4$ (0.1 M in deionized water) was added, followed by the addition 8 μL of sodium ascorbate (0.2 M in deionized water) and 8 μL of TBTA (0.1 M in *tert*-butanol). The reaction mixture was stirred at room temperature for 12h. Then the mixture was diluted with ethyl acetate (10 mL), washed with water and brine (6 mL×2), dried over Na_2SO_4. The solvent was removed in vacuo. The obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1: 4) to give **16c** (6.5 mg, 95%) as a white solid. R_f = 0.58 (hexane: EtOAc= 1: 4). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.20 (s, 1H), 8.12 (d, *J* = 8.4 Hz, 2H), 7.91 (d, *J* = 1.7 Hz, 1H), 7.88 (dd, *J* = 5.2, 3.2 Hz, 3H), 7.77 – 7.64 (m, 3H), 7.51 (d, *J* = 9.3 Hz, 1H), 7.06 (t, *J* = 8.2 Hz, 2H), 6.89 – 6.83 (m, 2H), 5.90 (s, 1H), 5.66 (s, 2H), 5.64 (d, *J* = 6.4 Hz, 2H), 5.53 (q, *J* = 12.8 Hz, 2H), 5.43 (q, *J* = 12.8 Hz, 2H), 3.87 (s, 3H), 3.84 (d, *J* = 17.2 Hz, 1H), 3.75 (s, 3H), 3.50 (d, *J* = 17.2 Hz, 1H), 2.66 (s, 3H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 196.52, 169.39, 165.66, 165.56, 163.36, 160.17, 155.75, 147.64, 143.92, 143.87, 139.86, 137.02, 136.63, 136.54, 132.52, 132.45, 130.51, 130.49, 130.09, 128.67, 127.74, 127.71, 126.24, 126.22, 124.59, 124.54, 122.22, 120.18, 113.04, 112.15, 110.39, 101.46, 61.83, 57.86, 55.86, 54.77, 52.33, 50.32, 26.68. **IR** (thin film): cm^{-1} 3143, 2955, 2927, 2854, 1717, 1688, 1605, 1557, 1286, 1261, 1178, 1112, 1048, 911, 840, 761, 732. **HRMS** (ESI+): calculated for $C_{44}H_{36}N_7O_{10}S_2$ $[M+H]^+$ 886.1965, found: 886.1983.

*Synthesis of **16d***

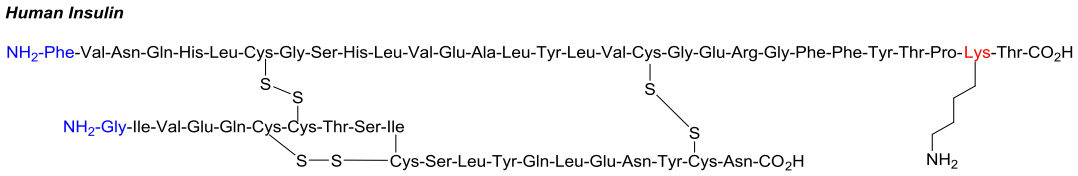


Compound **15a** (4 mg, 0.00639 mmol) and 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-glucopyranosyl azide (4.8 mg, 0.0128 mmol) was dissolved in 1 mL of tert-butanol/water (1: 1). After degassing for 20 min, 6 μL of $CuSO_4$ (0.1 M in deionized water) was added, followed by the addition 6 μL of sodium ascorbate (0.2 M in deionized water) and 6 μL of TBTA (0.1 M in *tert*-butanol). The reaction mixture was stirred at room temperature for 30 h. Then the mixture was diluted with ethyl acetate (10 mL), washed with water and brine (6 mL×2), dried over Na_2SO_4. The solvent was removed in vacuo. The obtained residue was purified by silica gel column chromatography (hexane: ethyl acetate= 1: 3) to give **16d** (5.4 mg, 85%) as a white solid. R_f = 0.27 (hexane: EtOAc= 1: 2). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.98 (d, J = 1.7 Hz, 1H), 7.92 (s, 1H), 7.87 (d, J = 1.4 Hz, 1H), 7.75 – 7.65 (m, 2H), 7.58 (s, 1H), 7.37 (d, J = 7.0 Hz, 2H), 7.32 – 7.26 (m, 2H), 7.09 – 7.01 (m, 2H), 6.04 – 5.96 (m, 1H), 5.74 (dd, J = 9.0, 5.0 Hz, 1H), 5.65 (d, J = 3.7 Hz, 1H), 5.63 (s, 1H), 5.52 (s, 2H), 5.46 – 5.38 (m, 4H), 5.24 (t, J = 9.9 Hz, 1H), 5.10 (dd, J = 12.5, 6.9 Hz, 1H), 4.55 (dd, J = 19.5, 10.1 Hz, 1H), 4.30 – 4.23 (m, 1H), 4.19 – 4.12 (m, 1H), 3.99 (d, J = 10.2 Hz, 1H), 3.84 (d, J = 17.2 Hz, 1H), 3.75 (s, 3H), 3.50 (dd, J = 17.1, 1.9 Hz, 1H), 2.09 – 2.03 (m, 10H), 1.75 (d, J = 12.6 Hz, 3H). **^{13}C NMR** (151 MHz, $CDCl_3$) δ 170.85, 170.55, 170.33, 169.40, 169.24, 165.60, 165.41, 143.23, 143.13, 136.41, 134.29, 132.44, 130.54, 130.47, 129.18, 128.88, 128.65, 128.18, 127.68, 126.37, 123.89, 86.00, 75.15, 74.02, 72.11, 67.73, 61.80, 61.62, 58.03, 54.81, 54.26, 53.68, 52.31, 22.88, 20.71, 20.62, 20.58, 20.56. **IR** (thin film): cm^{-1} 2955, 2927, 2118, 1748, 1716, 1602, 1372, 1260, 1235, 1110, 1045, 915, 761, 732. **HRMS** (ESI+): calculated for $C_{46}H_{47}N_8O_{14}S_2$ $[M+H]^+$ 999.2653, found: 999.2661.

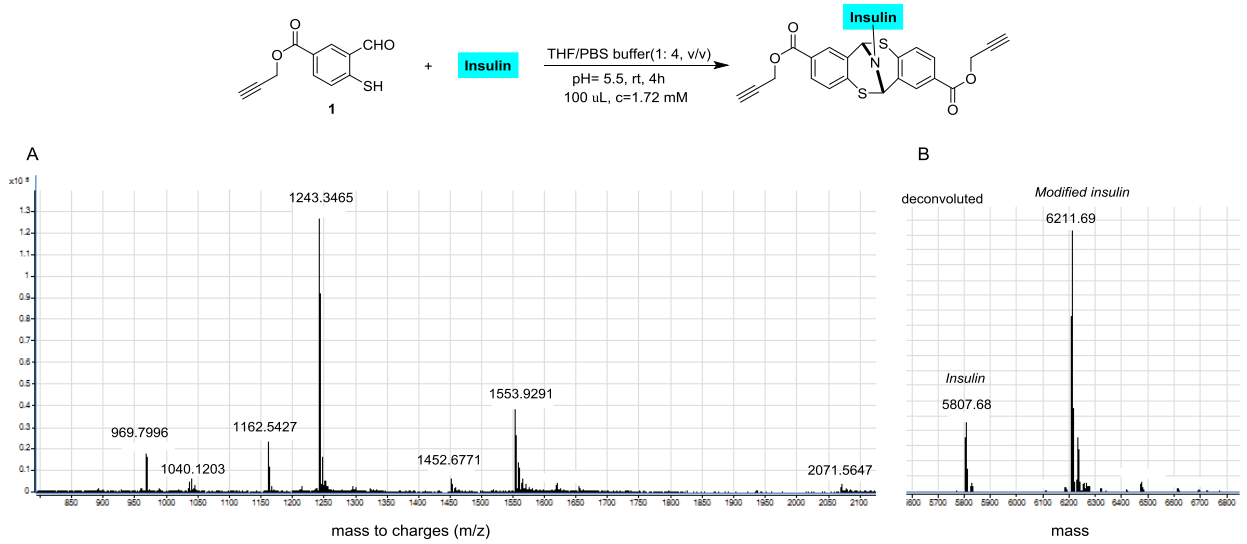
4. Conjugation with recombinant human insulin

4.1 Conjugation with human insulin

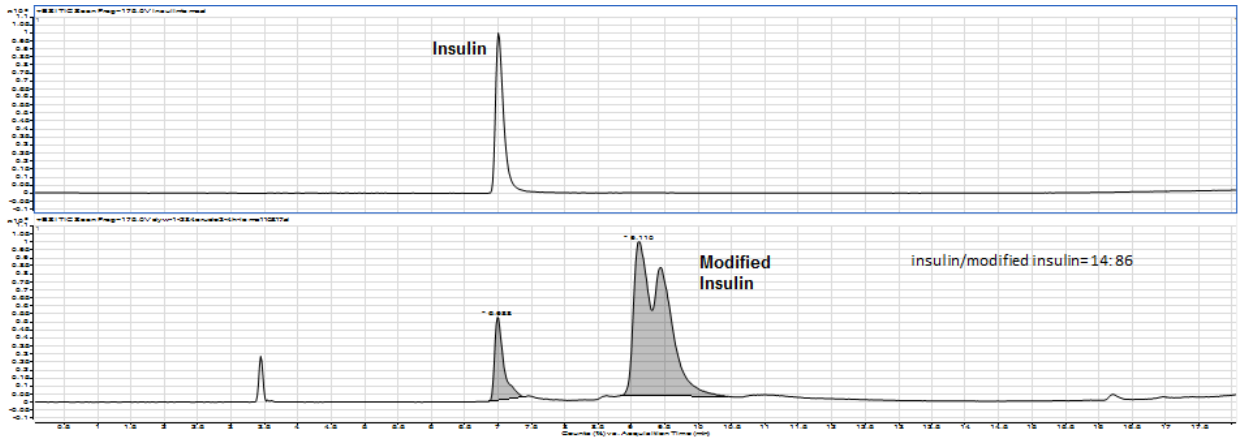
Human insulin (5817.4 Da), mono-modified human insulin (6221.5 Da)



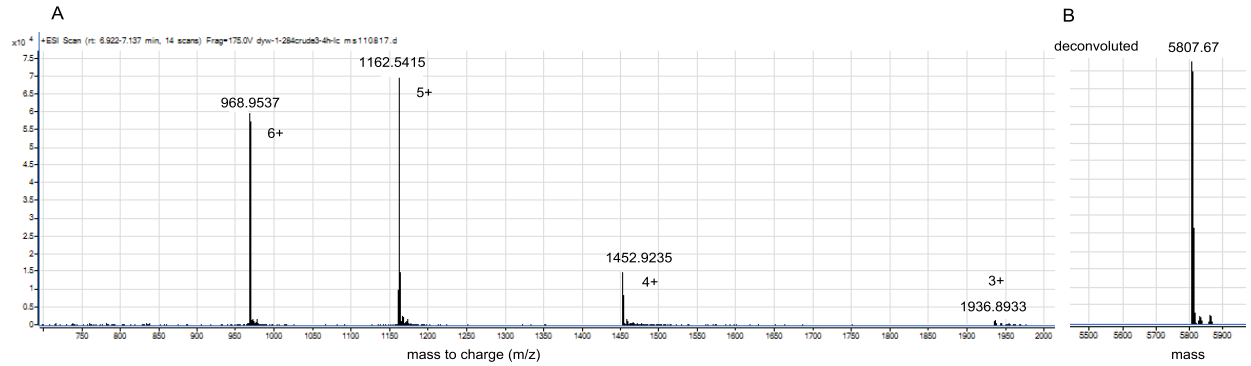
To a solution of human insulin (1.0 mg, 0.000172 mmol) in 80 μL of PBS buffer (pH= 5.5) was added prop-2-yn-1-yl 3-formyl-4-mercaptobenzoate **1** (0.34 mg, 0.0012 mmol) in 20 μL of tetrahydrofuran. The final concentration of human insulin is 1.72 mM, the mixture was stirred at room temperature for 4 h, monitored the reaction by LC-MS. The ratio of insulin/modified insulin is 14: 86 by LC-MS. Two isomeric peaks were detected due to the modification occurred at the *N*-termini of either A chain or B chain. The specific site of the modification was confirmed by LC-MS/MS protocol.



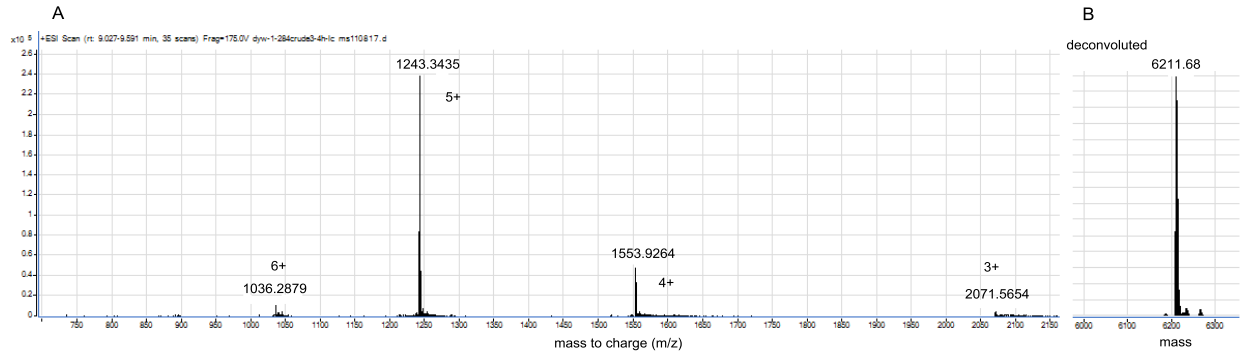
ESI-MS spectrum of the crude modification reaction of human insulin with **1**. (A) Raw spectrum. (B) Spectrum after deconvolution. Insulin: ($C_{257}H_{383}N_{65}O_{77}S_6$) 5807.63 calculated, 5807. 68 found; modified insulin: 6211.65(5807.63+404.02) calculated, 6211.69 found.



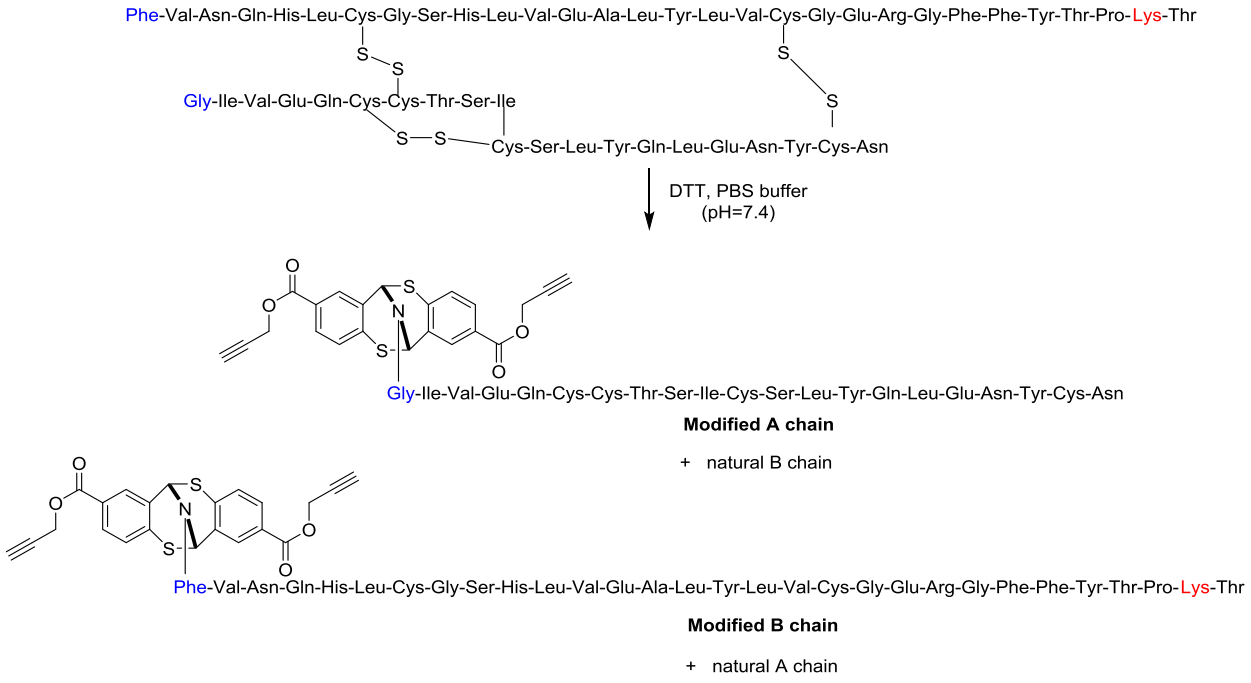
HPLC trace of pure human insulin sample(top) and the crude reaction mixture of human insulin with **1** (down). HPLC condition: 20 to 70% solvent B over 17 min, flow rate 0.5 mL/min, Higgins Analytical Proto 200 10 μm 250 × 4.6 nm C18 column.



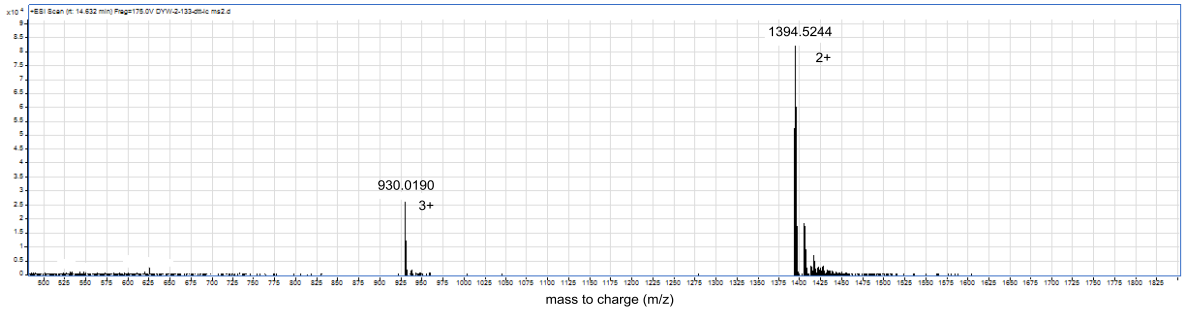
ESI-MS spectra of insulin. (A) Raw spectrum. (B) Spectrum after deconvolution. Insulin: ($C_{257}H_{383}N_{65}O_{77}S_6$) 5807.63 calculated, 5807. 67 found.



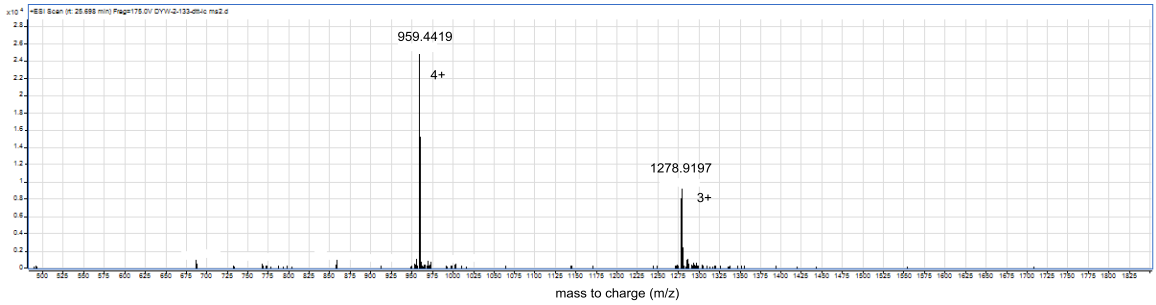
ESI-MS spectrum of modified insulin (+404 Da adduct). (A) Raw spectrum. (B) Spectrum after deconvolution. modified insulin: 6211.65(5807.63+404.02) calculated, 6211.69 found.



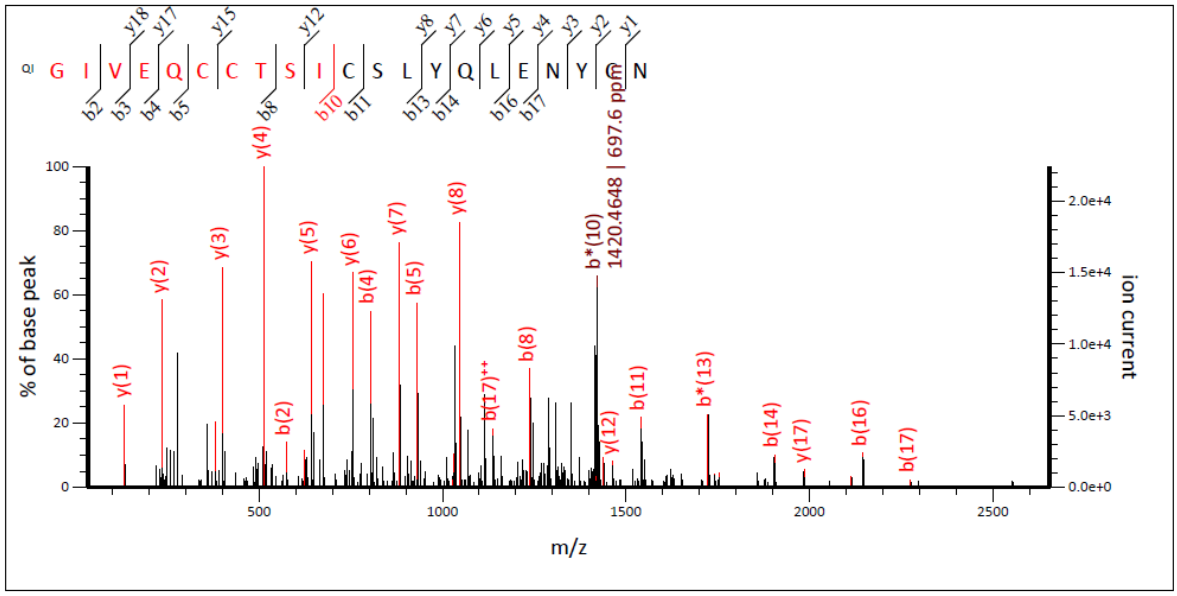
After adding DTT to cleavage disulfide bonds, modified A chain (Mr(calc)= 2786.0177):



After adding DTT to cleavage disulfide bonds, modified B chain (Mr(calc)= 3831.7023):



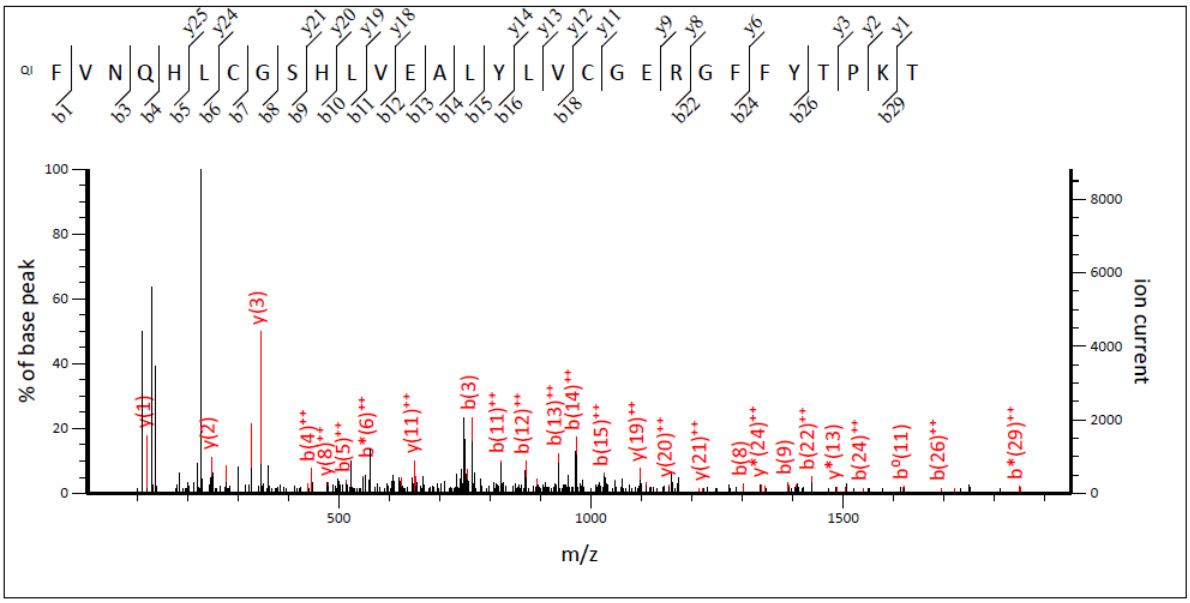
MS/MS spectrum of modified A chain:



Monoisotopic mass of neutral peptide Mr(calc): 2786.0177
Variable modifications:
N-term: QiangM (N-term)
Ions Score: 117 **Expect:** 1.9e-012
Matches: 37/218 fragment ions using 46 most intense peaks

#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	462.0464	231.5269					G							21
2	*575.1305*	288.0689					I	2325.9858	1163.4966	2308.9593	1154.9833	2307.9753	1154.4913	20
3	*674.1989*	337.6031					V	2212.9018	1106.9545	2195.8752	1098.4413	2194.8912	1097.9492	19
4	*803.2415*	402.1244			785.2309	393.1191	E	*2113.8334*	1057.4203	2096.8068	1048.9070	2095.8228	1048.4150	18
5	*931.3001*	466.1537	914.2735	457.6404	913.2895	457.1484	Q	*1984.7908*	992.8990	1967.7642	984.3857	1966.7802	983.8937	17
6	1034.3093	517.6583	1017.2827	509.1450	1016.2987	508.6530	C	1856.7322	928.8697	1839.7056	920.3565	1838.7216	919.8645	16
7	*1137.3185*	569.1629	1120.2919	560.6496	1119.3079	560.1576	C	*1753.7230*	877.3651	1736.6965	868.8519	1735.7124	868.3599	15
8	*1238.3661*	619.6867	1221.3396	611.1734	1220.3556	610.6814	T	1650.7138	825.8605	1633.6873	817.3473	1632.7033	816.8553	14
9	1325.3982	663.2027	1308.3716	654.6894	1307.3876	654.1974	S	1549.6661	775.3367	1532.6396	766.8234	1531.6556	766.3314	13
10	*1438.4822*	719.7448	1421.4557	711.2315	1420.4717	710.7395	I	*1462.6341*	731.8207	1445.6076	723.3074	1444.6235	722.8154	12
11	*1541.4914*	771.2493	1524.4649	762.7361	1523.4808	762.2441	C	1349.5500	675.2787	1332.5235	666.7654	1331.5395	666.2734	11
12	1628.5234	814.7654	1611.4969	806.2521	1610.5129	805.7601	S	1246.5409	623.7741	1229.5143	615.2608	1228.5303	614.7688	10
13	*1741.6075*	871.3074	1724.5810	862.7941	1723.5969	862.3021	L	1159.5088	580.2581	1142.4823	571.7448	1141.4983	571.2528	9
14	*1904.6708*	952.8391	1887.6443	944.3258	1886.6603	943.8338	Y	*1046.4248*	523.7160	1029.3982	515.2027	1028.4142	514.7107	8
15	2032.7294	1016.8683	2015.7029	1008.3551	2014.7188	1007.8631	Q	*883.3614*	442.1844	866.3349	433.6711	865.3509	433.1791	7
16	*2145.8135*	1073.4104	2128.7869	1064.8971	2127.8029	1064.4051	L	*755.3029*	378.1551	738.2763	369.6418	737.2923	369.1498	6
17	*2274.8561*	1137.9317	2257.8295	1129.4184	2256.8455	1128.9264	E	*642.2188*	321.6130	625.1923	313.0998	624.2082	312.6078	5
18	2388.8990	1194.9531	2371.8724	1186.4399	2370.8884	1185.9479	N	*513.1762*	257.0917	496.1497	248.5785			4
19	2551.9623	1276.4848	2534.9358	1267.9715	2533.9518	1267.4795	Y	*399.1333*	200.0703	382.1067	191.5570			3
20	2654.9715	1327.9894	2637.9450	1319.4761	2636.9609	1318.9841	C	*236.0700*	118.5386	219.0434	110.0253			2
21							N	*133.0608*	67.0340	116.0342	58.5207			1

MS/MS spectrum of modified B chain:

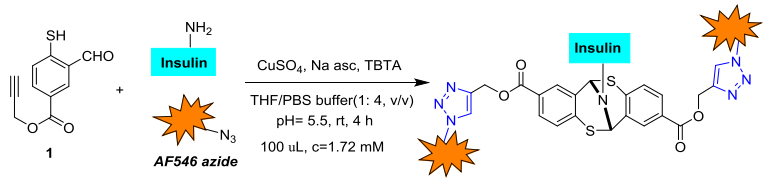


Monoisotopic mass of neutral peptide Mr(calc): 3831.7023
Variable modifications:
N-term: QiangM (N-term)
Ions Score: 55 **Expect:** 2.8e-006
Matches: 63/326 fragment ions using 160 most intense peaks

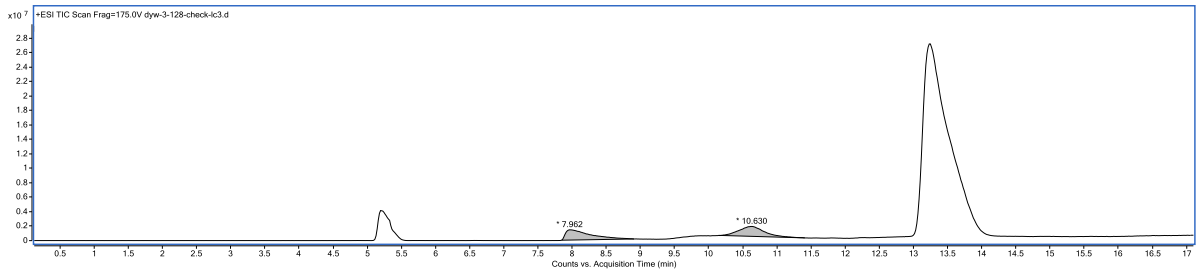
#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	552.0934	*276.5503*					F							30
2	**651.1618**	326.0845					V	3281.6234	1641.3154	3264.5969	1632.8021	3263.6129	1632.3101	29
3	**765.2047**	383.1060	748.1782	374.5927			N	3182.5550	1591.7812	3165.5285	1583.2679	3164.5445	1582.7759	28
4	**893.2633**	*447.1353*	876.2368	438.6220			Q	3068.5121	1534.7597	3051.4855	1526.2464	3050.5015	1525.7544	27
5	1030.3222	*515.6647*	1013.2957	507.1515			H	2940.4535	1470.7304	2923.4270	1462.2171	2922.4430	1461.7251	26
6	1143.4063	572.2068	1126.3797	563.6935			L	2803.3946	1402.2009	2786.3681	1393.6877	2785.3840	1393.1957	25
7	**1246.4155**	*623.7114*	1229.3889	615.1981			C	2690.3105	**1345.6589**	2673.2840	1337.1456	2672.3000	1336.6536	24
8	**1303.4369**	*652.2221*	1286.4104	643.7088			G	2587.3014	1294.1543	2570.2748	1285.6410	2569.2908	1285.1490	23
9	**1390.4690**	695.7381	1373.4424	687.2248	1372.4584	686.7328	S	2530.2799	1265.6436	2513.2533	1257.1303	2512.2693	1256.6383	22
10	1527.5279	*764.2676*	1510.5013	755.7543	1509.5173	755.2623	H	2443.2479	**1222.1276**	2426.2213	1213.6143	2425.2373	1213.1223	21
11	1640.6119	*820.8096*	1623.5854	812.2963	1622.6014	811.8043	L	2306.1890	**1153.5981**	2289.1624	1145.0848	2288.1784	1144.5928	20
12	1739.6804	*870.3438*	1722.6538	861.8305	1721.6698	861.3385	V	2193.1049	**1097.0561**	2176.0783	1088.5428	2175.0943	1088.0508	19
13	1868.7229	*934.8651*	1851.6964	926.3518	1850.7124	925.8598	E	2094.0365	**1047.5219**	2077.0099	1039.0086	2076.0259	1038.5166	18
14	1939.7601	*970.3837*	1922.7335	961.8704	1921.7495	961.3784	A	1964.9939	983.0006	1947.9673	974.4873	1946.9833	973.9953	17
15	2052.8441	*1026.9257*	2035.8176	1018.4124	2034.8336	1017.9204	L	1893.9568	947.4820	1876.9302	938.9687	1875.9462	938.4767	16
16	2215.9075	*1108.4574*	2198.8809	1099.9441	2197.8969	1099.4521	Y	1780.8727	890.9400	1763.8462	882.4267	1762.8621	881.9347	15
17	2328.9915	1164.9994	2311.9650	1156.4861	2310.9810	1155.9941	L	*1617.8094*	809.4083	1600.7828	800.8951	1599.7988	800.4030	14
18	2428.0599	*1214.5336*	2411.0334	1206.0203	2410.0494	1205.5283	V	*1504.7253*	752.8663	1487.6988	744.3530	1486.7147	743.8610	13
19	2531.0691	1266.0382	2514.0426	1257.5249	2513.0585	1257.0329	C	*1405.6569*	703.3321	1388.6304	694.8188	1387.6463	694.3268	12
20	2588.0906	1294.5489	2571.0640	1286.0357	2570.0800	1285.5436	G	*1302.6477*	**651.8275**	1285.6212	643.3142	1284.6371	642.8222	11
21	2717.1332	1359.0702	2700.1066	1350.5569	2699.1226	1350.0649	E	*1245.6263*	**623.3168**	1228.5997	614.8035	1227.6157	614.3115	10
22	2873.2343	*1437.1208*	2856.2077	1428.6075	2855.2237	1428.1155	R	*1116.5837*	558.7955	1099.5571	550.2822	1098.5731	549.7902	9
23	2930.2557	1465.6315	2913.2292	1457.1182	2912.2452	1456.6262	G	960.4825	**480.7449**	943.4560	472.2316	942.4720	471.7396	8
24	3077.3242	*1539.1657*	3060.2976	1530.6524	3059.3136	1530.1604	F	903.4611	452.2342	886.4345	443.7209	885.4505	443.2289	7
25	3224.3926	1612.6999	3207.3660	1604.1867	3206.3820	1603.6946	F	*756.3927*	378.7000	739.3661	370.1867	738.3821	369.6947	6
26	3387.4559	*1694.2316*	3370.4294	1685.7183	3369.4453	1685.2263	Y	609.3243	305.1658	592.2977	296.6525	591.3137	296.1605	5
27	3488.5036	1744.7554	3471.4770	1736.2422	3470.4930	1735.7501	T	*446.2609*	223.6341	429.2344	215.1208	428.2504	214.6288	4
28	3585.5563	1793.2818	3568.5298	1784.7685	3567.5458	1784.2765	P	*345.2132*	173.1103	328.1867	164.5970	327.2027	164.1050	3
29	3713.6513	1857.3293	3696.6248	1848.8160	3695.6407	1848.3240	K	*248.1605*	124.5839	231.1339	116.0706	230.1499	115.5786	2
30							T	*120.0655*	60.5364			102.0550	51.5311	1

MS/MS data supported the modification of human insulin occurred at *N*-termini of the insulin instead of the lysine residue.

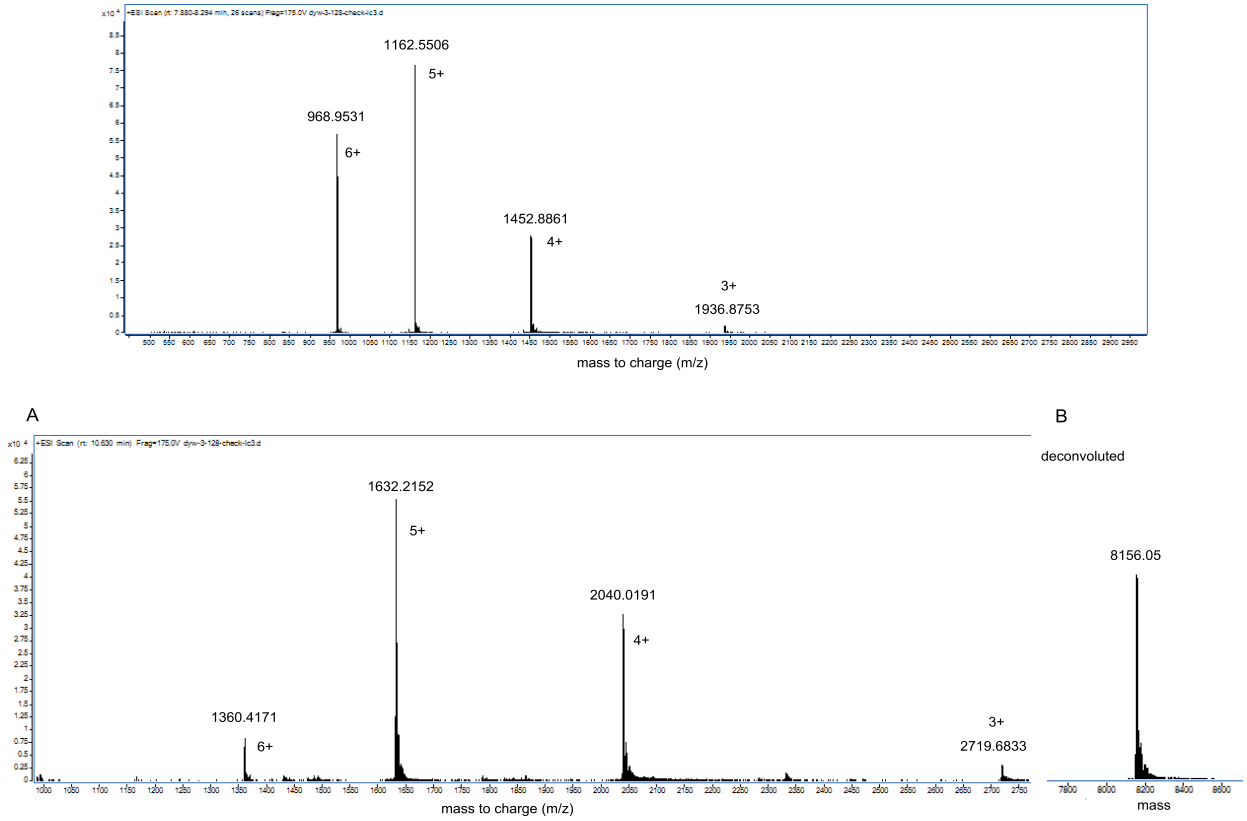
4.2 Three-component one-pot conjugation and CuAAC



To a solution of human insulin (0.5 mg, 0.000086 mmol) in 40 μL of PBS buffer (pH= 5.5) was added prop-2-yn-1-yl 3-formyl-4-mercaptobenzoate **1** (0.24 mg, 0.0008 mmol) in 10 μL of tetrahydrofuran. The final concentration of human insulin is 1.72 mM, the mixture was stirred at room temperature for 1 h. Then, 1.5 μL of 0.01 M $CuSO_4$, 3 μL of 0.01 M sodium ascorbate and 3 μL of 0.01 M TBTA was added to the mixture sequentially, monitored the reaction by LC-MS. The conversion of the one-pot conjugation is about 50% by LC-MS.



HPLC trace of the crude reaction mixture of one-pot conjugation, the ratio of insulin/product is 51:49. HPLC condition: 25 to 70% solvent B over 10 min, 14 min 95% solvent B. flow rate 0.5 mL/min, Higgins Analytical Proto 200 10 µm 250 × 4.6 nm C18 column. 5.2 min, salt peak; 8 min insulin; 10.5 min one-pot reaction product, 13.5 min AF546 azide.



ESI-MS spectrum of one-pot product (+2348.6 Da adduct). (A) Raw spectrum. (B) Spectrum after deconvolution. Product: 8156.2 calculated, 8156.05 found.

5. X-Ray Crystal Structure of compound 4

Compound 4 (CCDC 1825270)

Data collection was performed on a Bruker D8 VENTURE X-ray diffractometer with PHOTON 100 CMOS shutterless mode detector equipped with a Mo-target X-ray tube (λ = 0.71073 Å) at T = 100(2) K. Data reduction and integration were performed with the Bruker software package SAINT (version 8.38A). [4] Data were corrected for absorption effects using the empirical methods as implemented in SADABS (version 2016/2). [5] The structure was solved by SHELXT [6] and

refined by full-matrix least-squares procedures using the Bruker SHELXTL (version 2017/1) [7] software package. All non-hydrogen atoms were refined anisotropically. H-atoms were included at calculated positions and refined as riders, with U_{iso}(H) = 1.2 U_{eq}(C) and U_{iso}(H) = 1.5 U_{eq}(C) for methyl groups. Further crystal and data collection details are listed in Table 1. The structure of the compound is shown in Figure 1. The compound crystallizes in the $P2_1/n$ (No. 14).

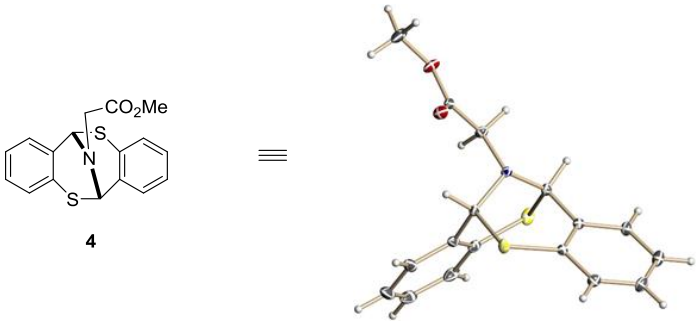


Figure 1. ORTEP drawing of **str2527** drawn with thermal ellipsoids at the 50% probability level. Color scheme used: C: grey; O: red; H: white; N: blue; S: yellow.

Table 1. Experimental details

Crystal data	
Chemical formula	$C_{17}H_{15}NO_2S_2$
M_r	329.42
Crystal system, space group	Monoclinic, $P2_1/n$
Temperature (K)	100
a, *b*, *c* (Å)	15.2896 (5), 7.1208 (2), 15.8248 (5)
β (°)	118.319 (1)
V (Å^3)	1516.71 (8)
Z	4
Radiation type	Mo *K*α
μ (mm^{-1})	0.36

Crystal size (mm)	0.50 × 0.17 × 0.03
Data collection	
Diffractometer	Bruker D8 Venture PHOTON 100 CMOS diffractometer
Absorption correction	Multi-scan *SADABS2016*/2 (Bruker,2016/2) was used for absorption correction. Krause, L., Herbst-Irmer, R., Sheldrick G.M. & Stalke D., J. Appl. Cryst. 48 (2015) 3-10
T_{min}, T_{max}	0.935, 1
No. of measured, independent and observed [$I > 2\sigma(I)$] reflections	51025, 6047, 4936
R_{int}	0.041
$(\sin \theta/\lambda)_{max}$ (Å^{-1})	0.781
Refinement	
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.042, 0.091, 1.08
No. of reflections	6047
No. of parameters	200
No. of restraints	0
H-atom treatment	H-atom parameters constrained
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.52, −0.31

Computer programs: *APEX3* v.2017.3-0 (Bruker AXS Inc., 2017), *SAINT* V.8.38A (Bruker AXS Inc., 2017), SHELXT 2014/5 (Sheldrick, 2014), *SHELXL2017*/1 (Sheldrick, 2017), Xshell v.6.3.1 (Bruker AXS Inc., 2016).

6. References:

1. V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem., Int. Ed*., **2002**, *41*, 2596–2599.

2. C. LingáTung, T. Clarence, X. Li, *Org. Biomol. Chem.*, **2015**,*13*, 6922-6926.

3. A. Burrows, C. Frost, M. Mahon, C. Richardson, *Angew. Chem. Int. Ed*. **2008**, *47*, 8482 –8486.

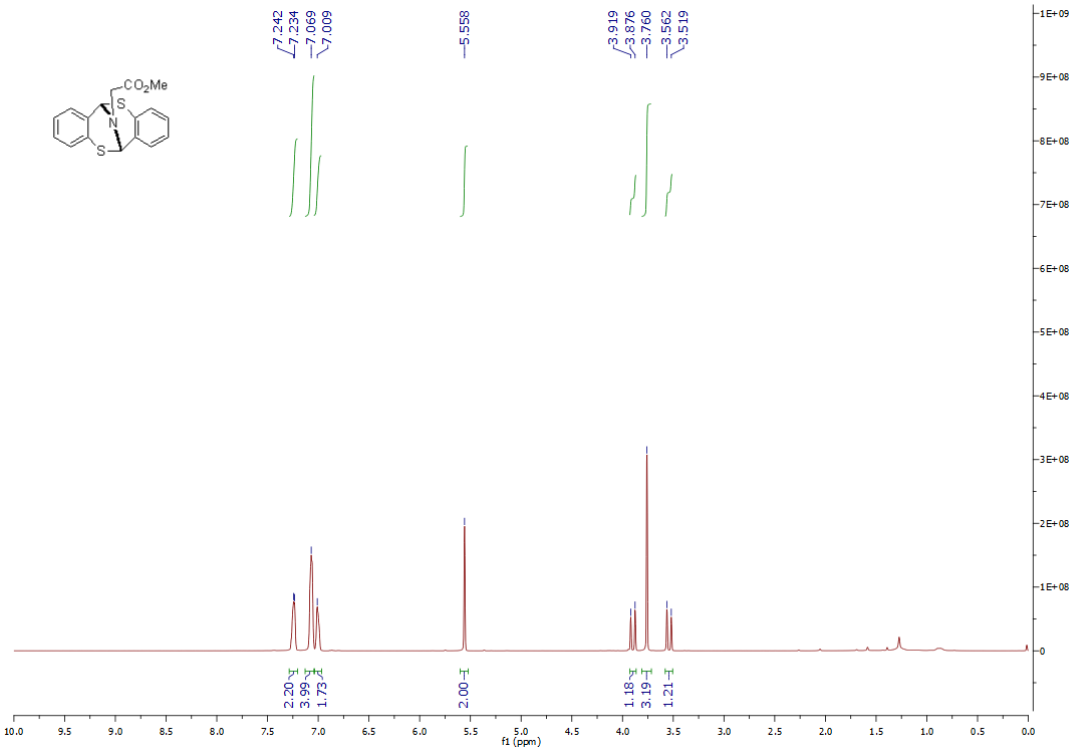
4. SAINT; part of Bruker APEX3 software package (version 2017.3-0): Bruker AXS, 2017.

5. SADABS; part of Bruker APEX3 software package (version 2017.3-0): Bruker AXS, 2017.

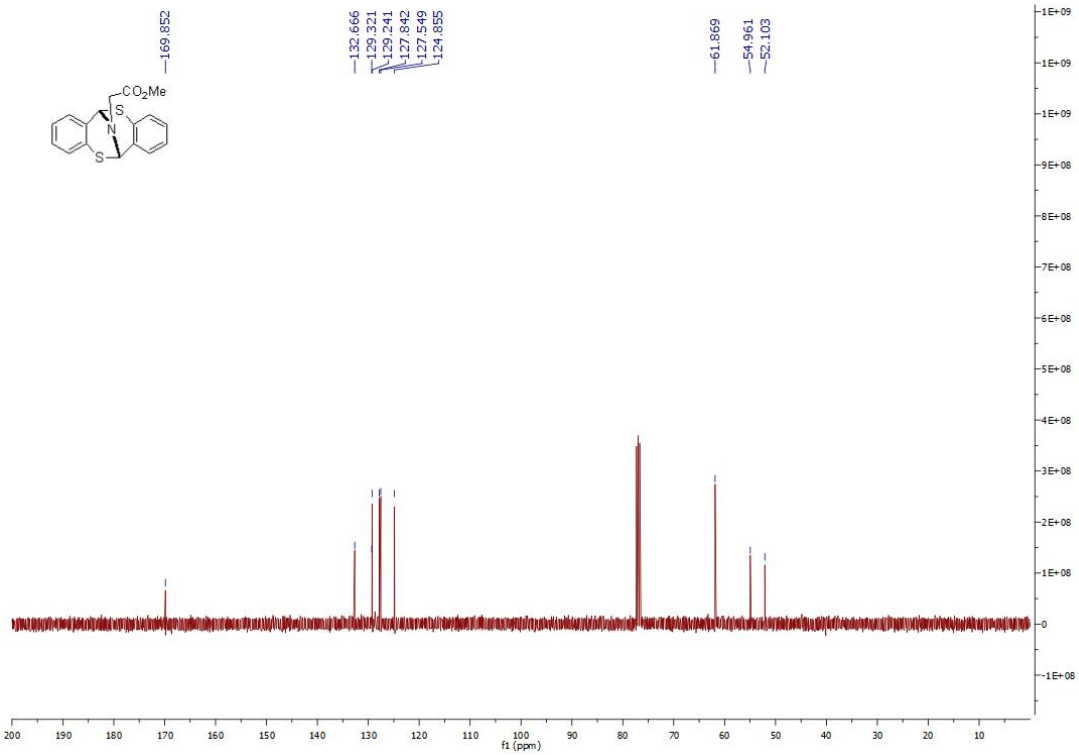
6. SHELXT; Version 2014/5: G. M. Sheldrick, *Acta Cryst.* **2015**, *A71*, 3-8.

7. XL refinement program version 2016/6: G. M. Sheldrick, *Acta Cryst.* **2015**, *C71*, 3-8.

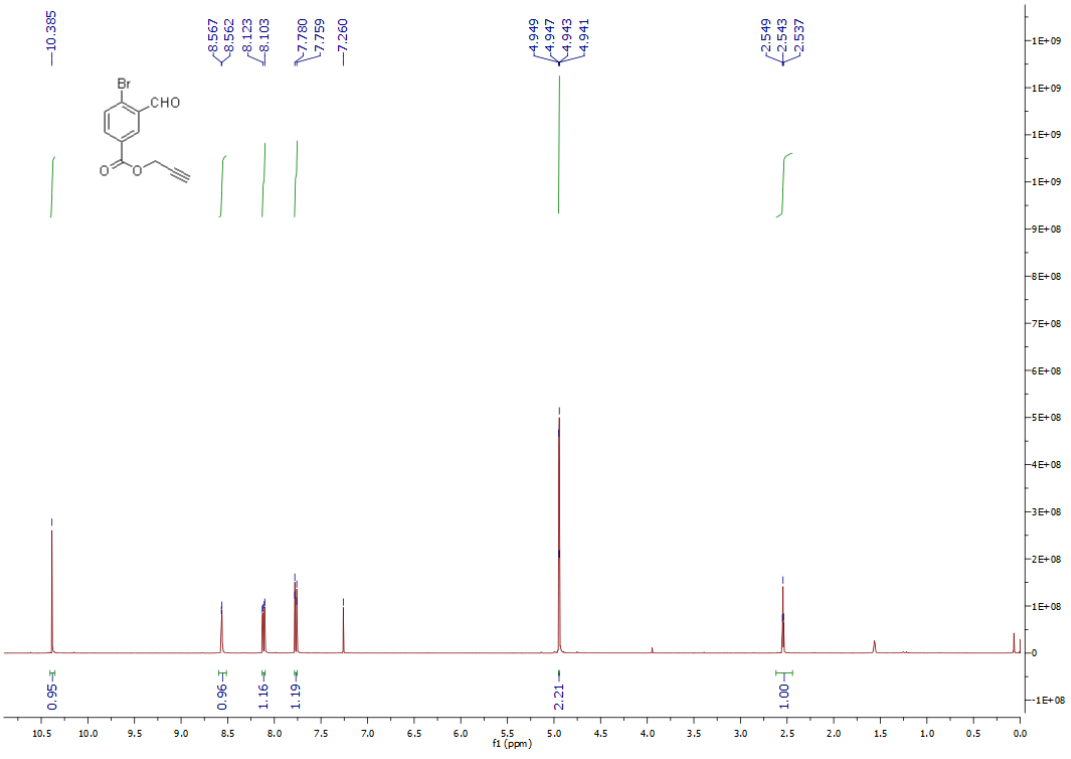
7. NMR spectra of new compounds



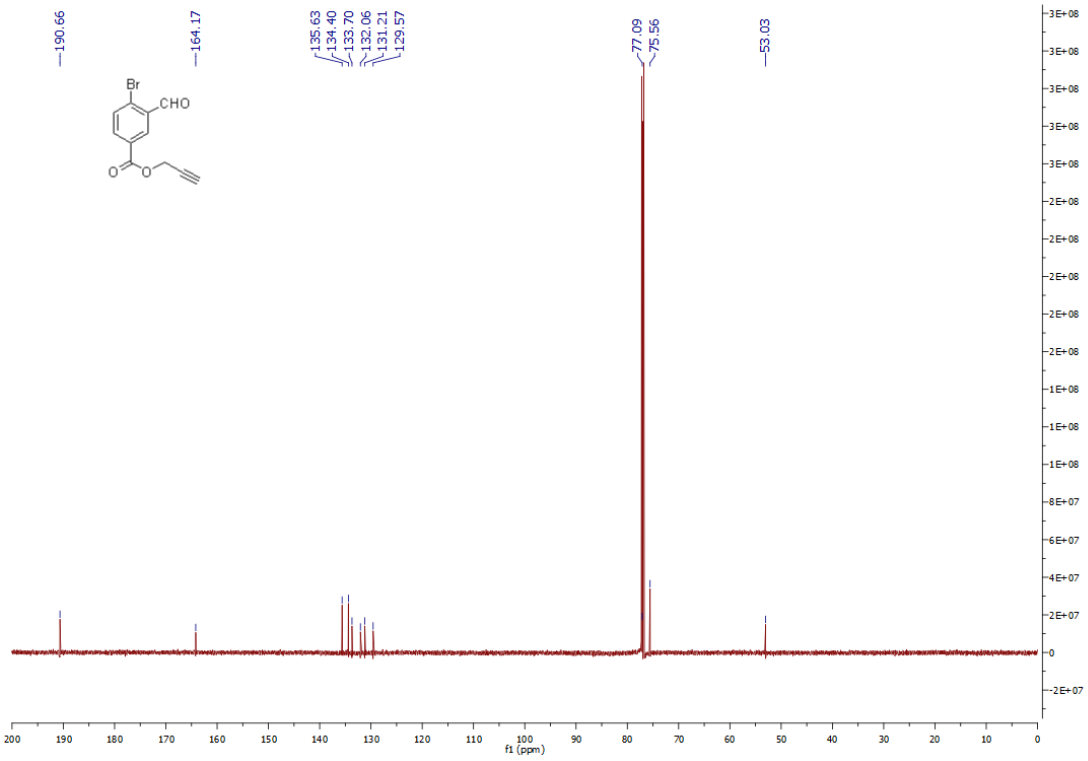
^{1}H NMR spectrum of **4** in $CDCl_3$ (400 MHz).



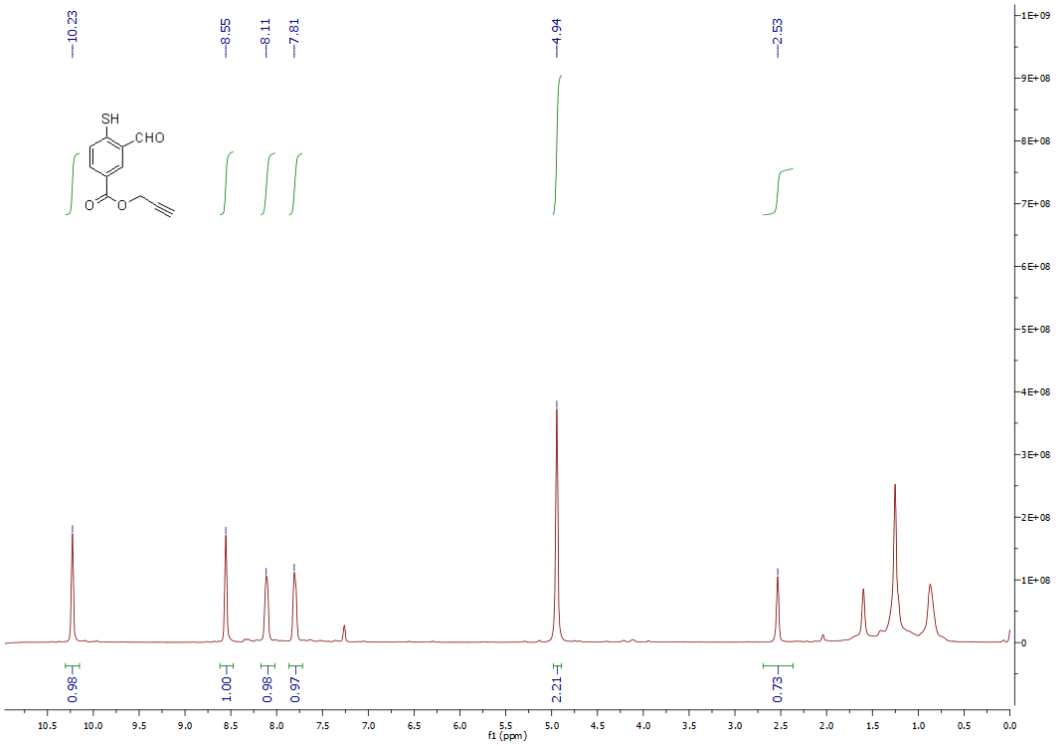
^{13}C NMR spectrum of **4** in $CDCl_3$ (101 MHz).



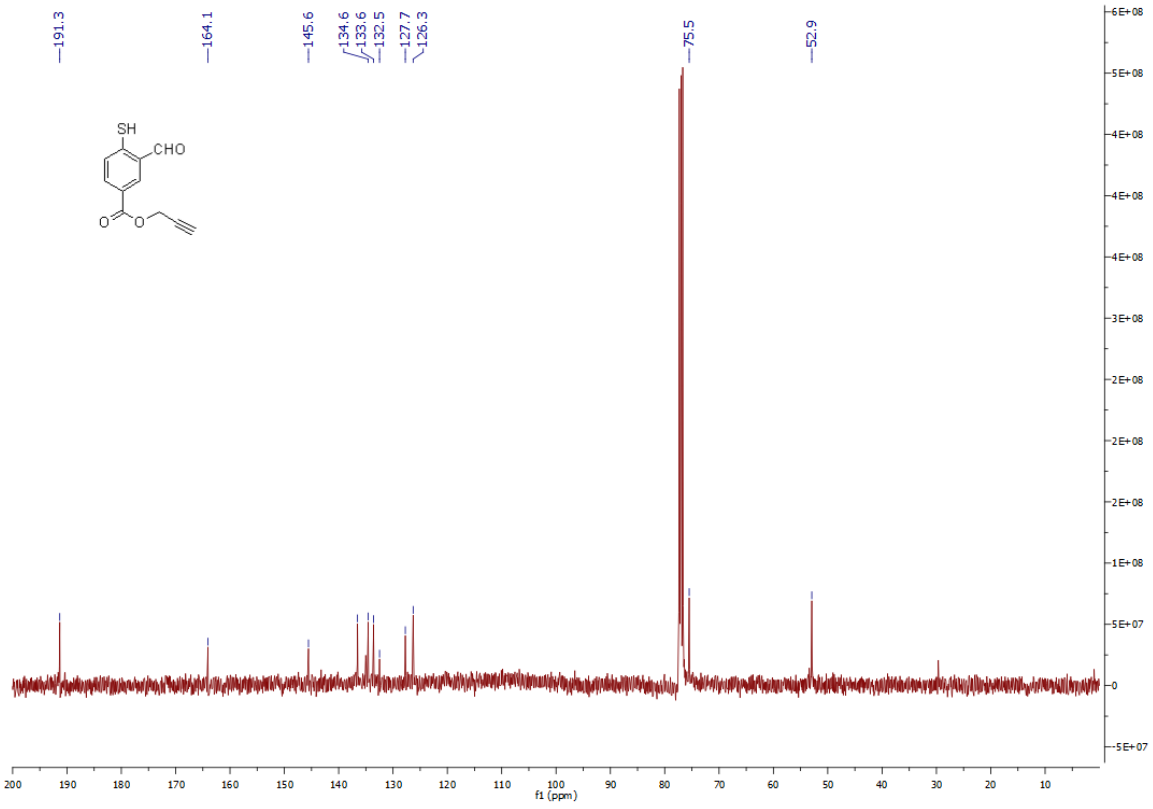
^{1}H NMR spectrum of **10** in $CDCl_3$ (500 MHz).



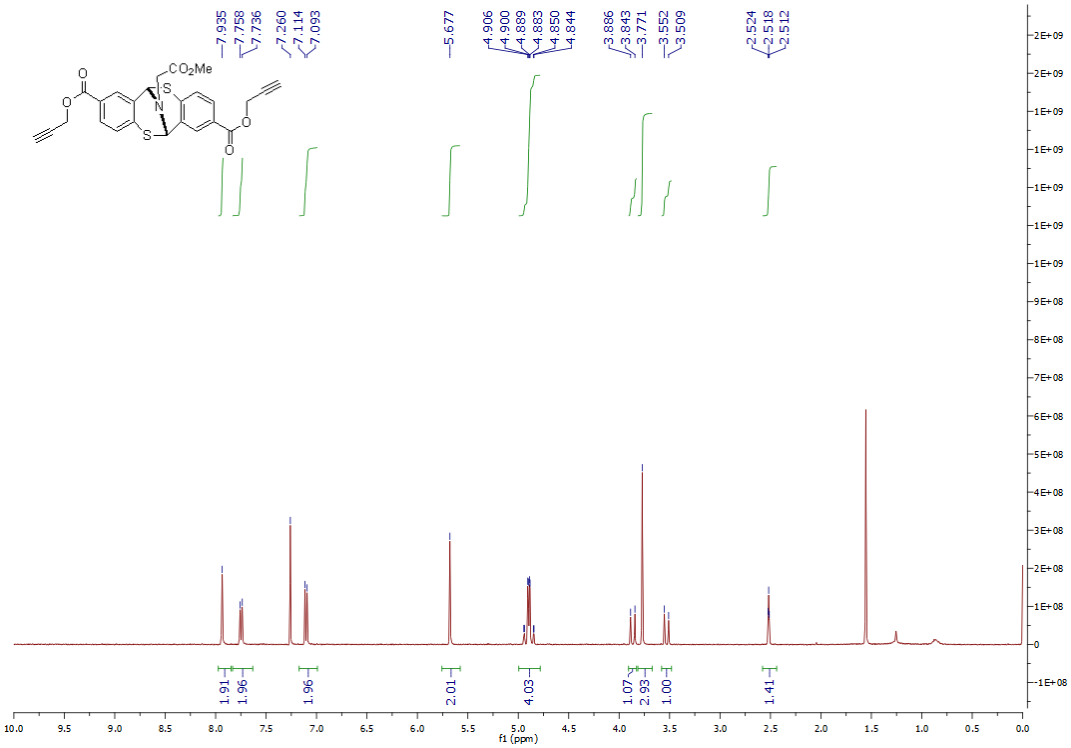
^{13}C NMR spectrum of **10** in $CDCl_3$ (126 MHz).



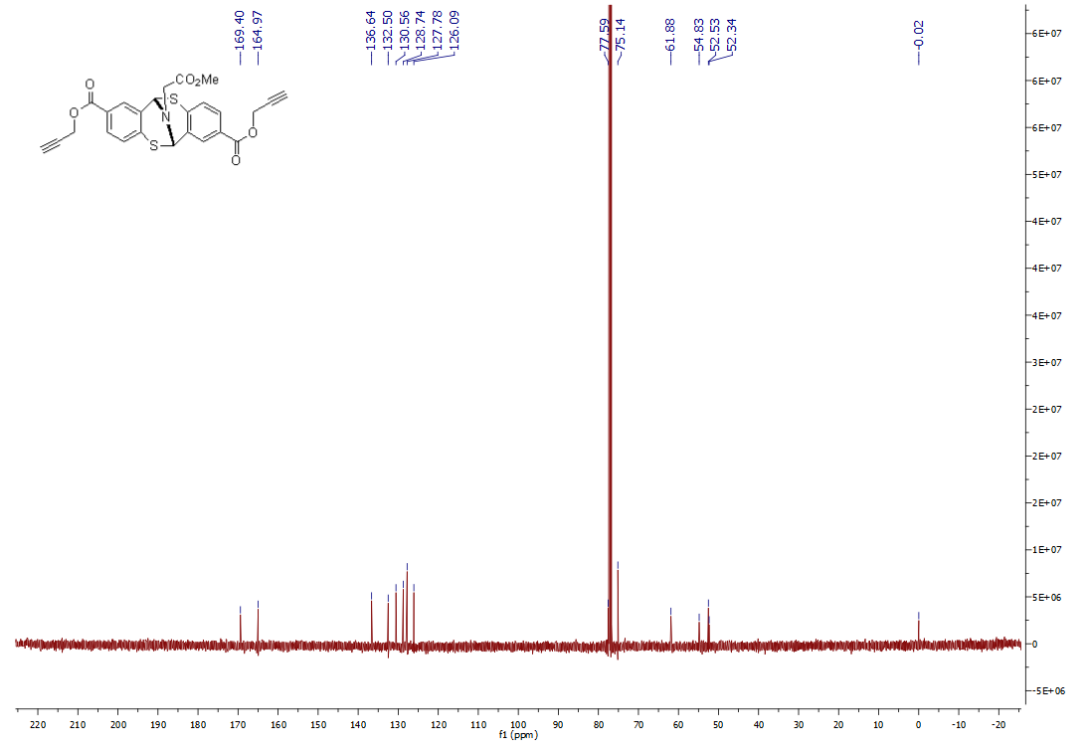
^{1}H NMR spectrum of **1** in $CDCl_3$ (400 MHz).



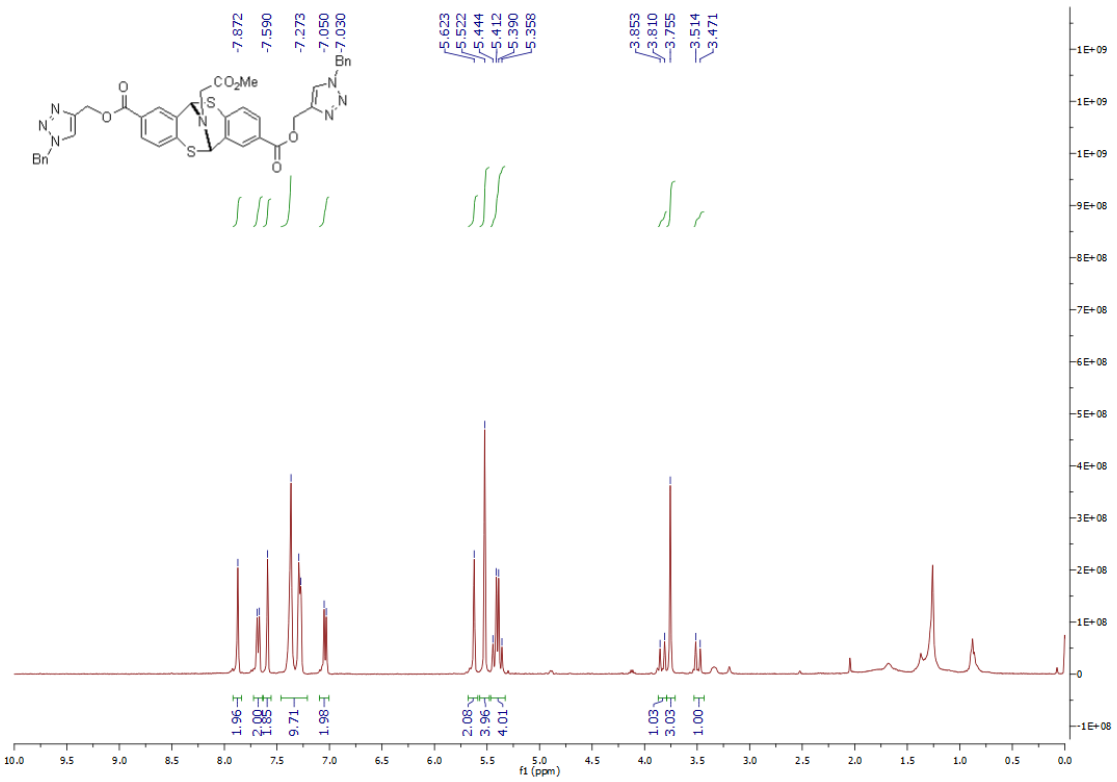
^{13}C NMR spectrum of **1** in $CDCl_3$ (101 MHz).



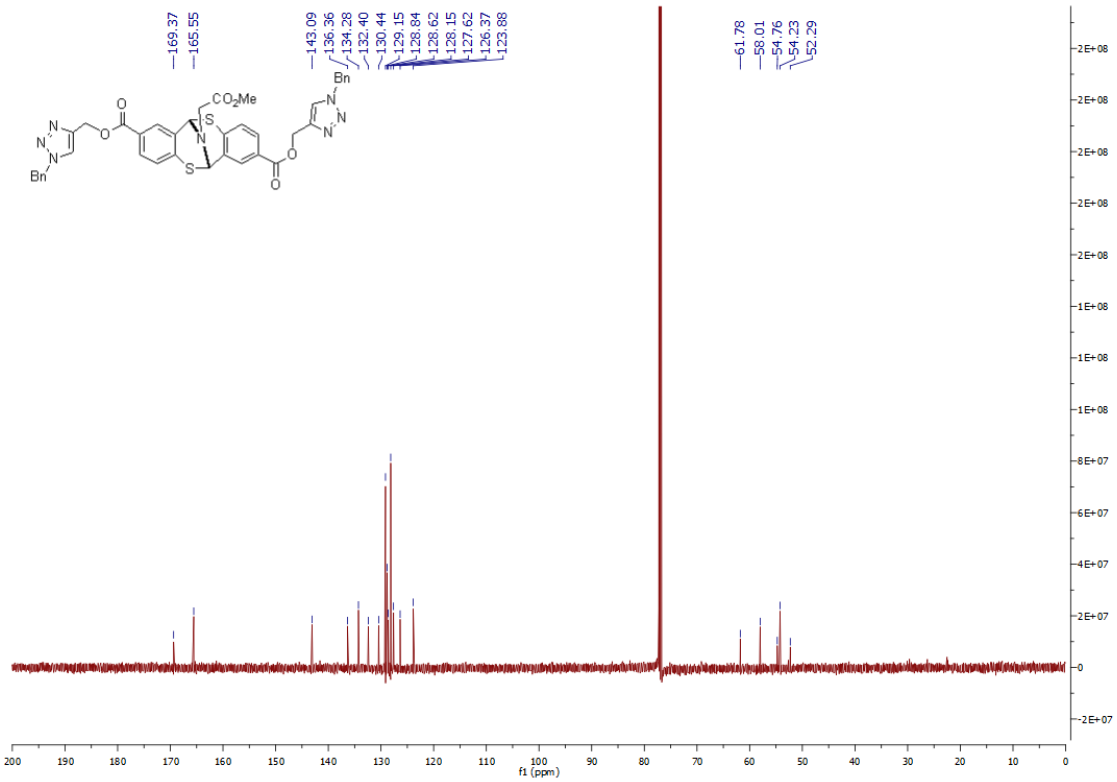
^{1}H NMR spectrum of **11** in $CDCl_3$ (400 MHz).



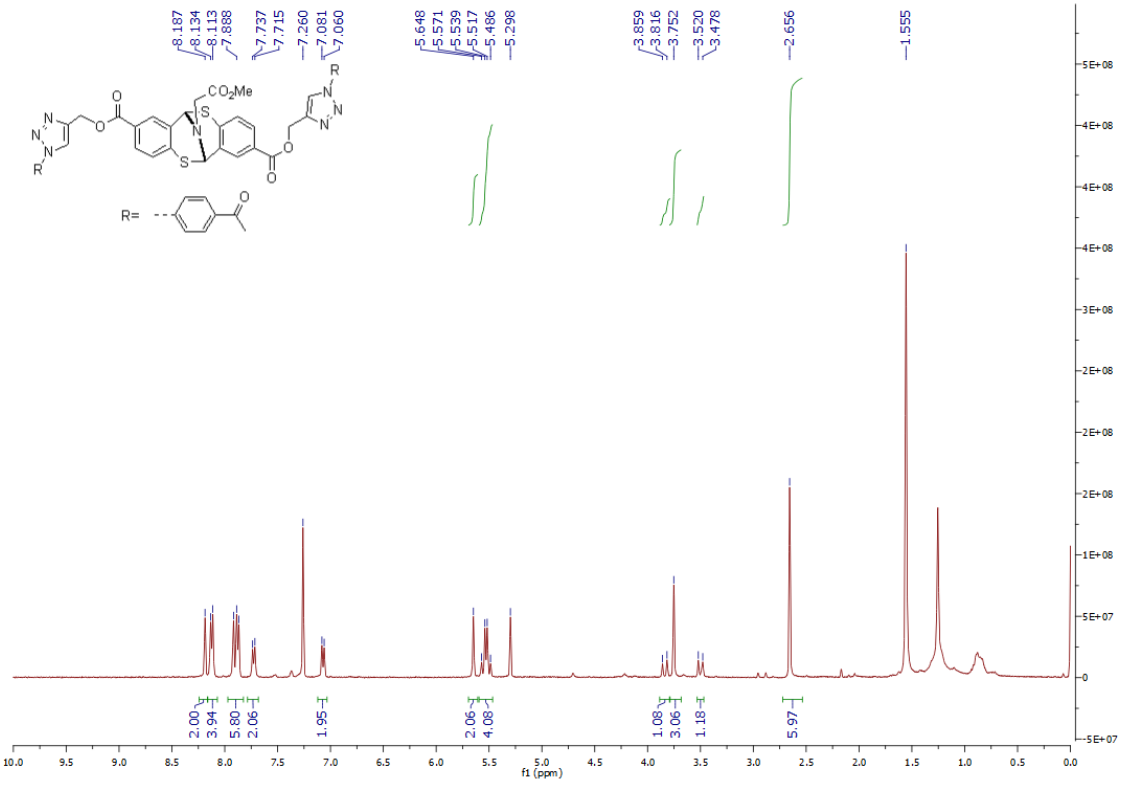
^{13}C NMR spectrum of **11** in $CDCl_3$ (151 MHz).



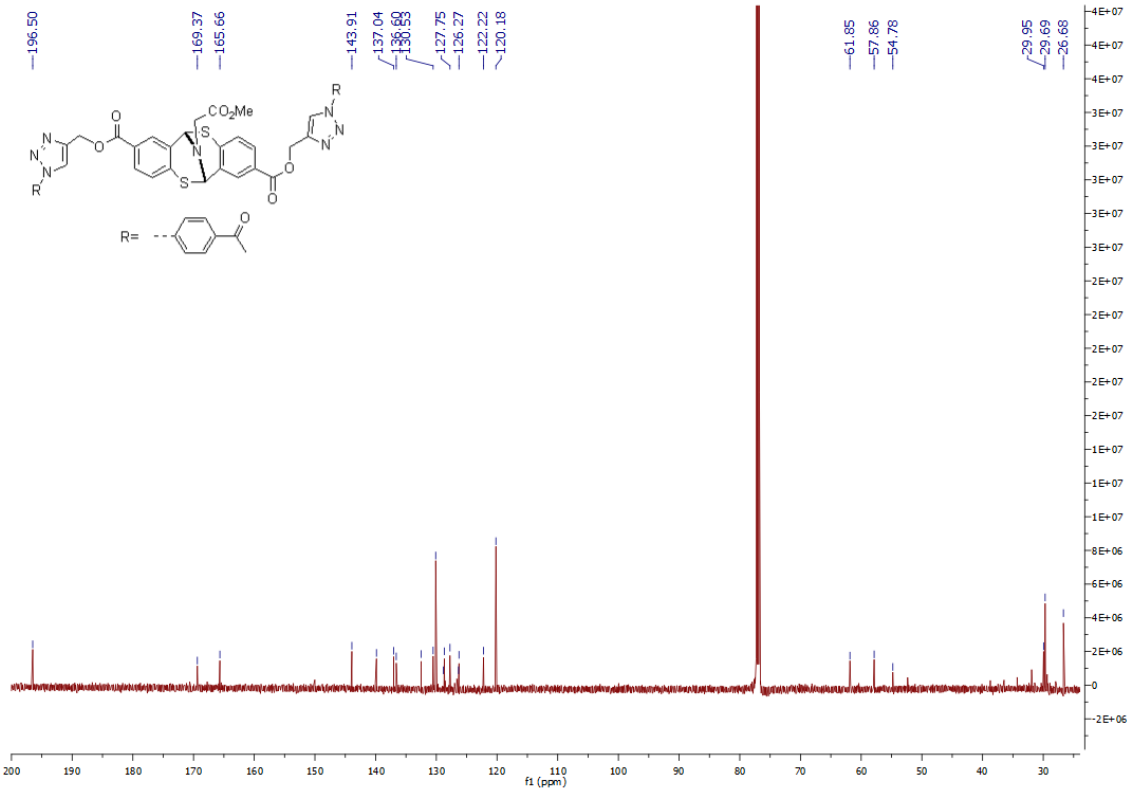
^{1}H NMR spectrum of **13a** in $CDCl_3$ (400 MHz).



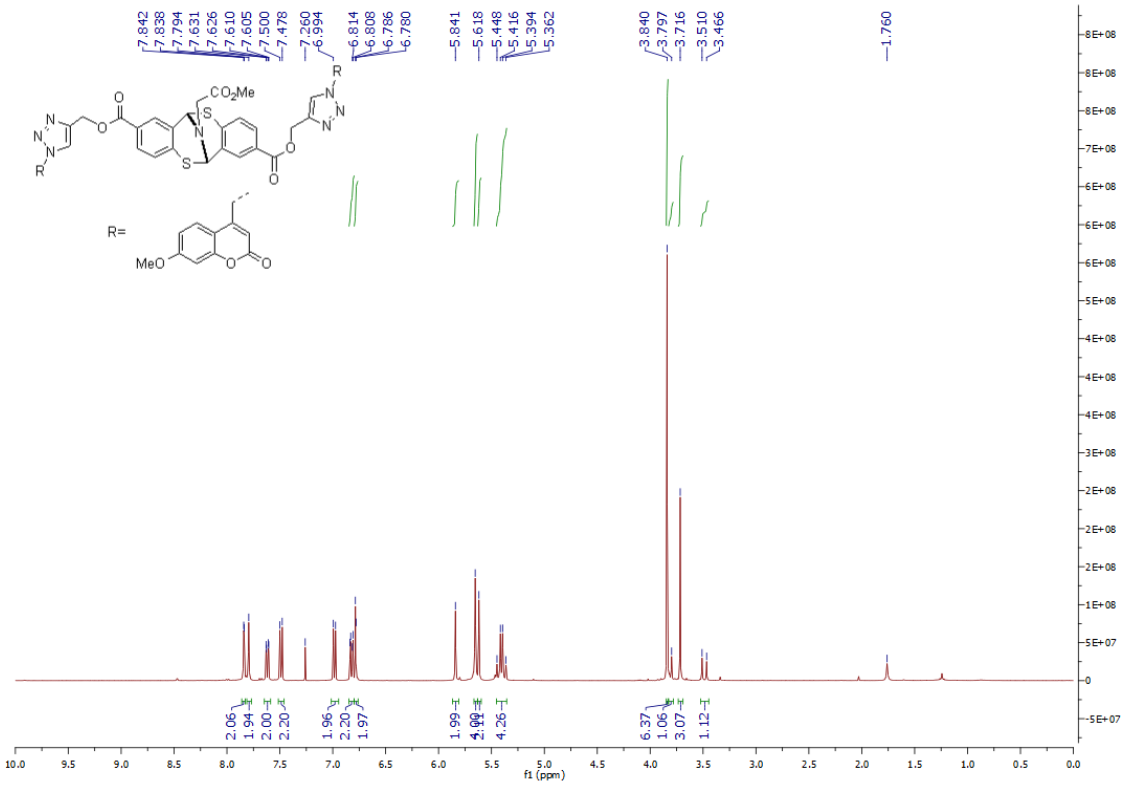
^{13}C NMR spectrum of **13a** in $CDCl_3$ (151 MHz).



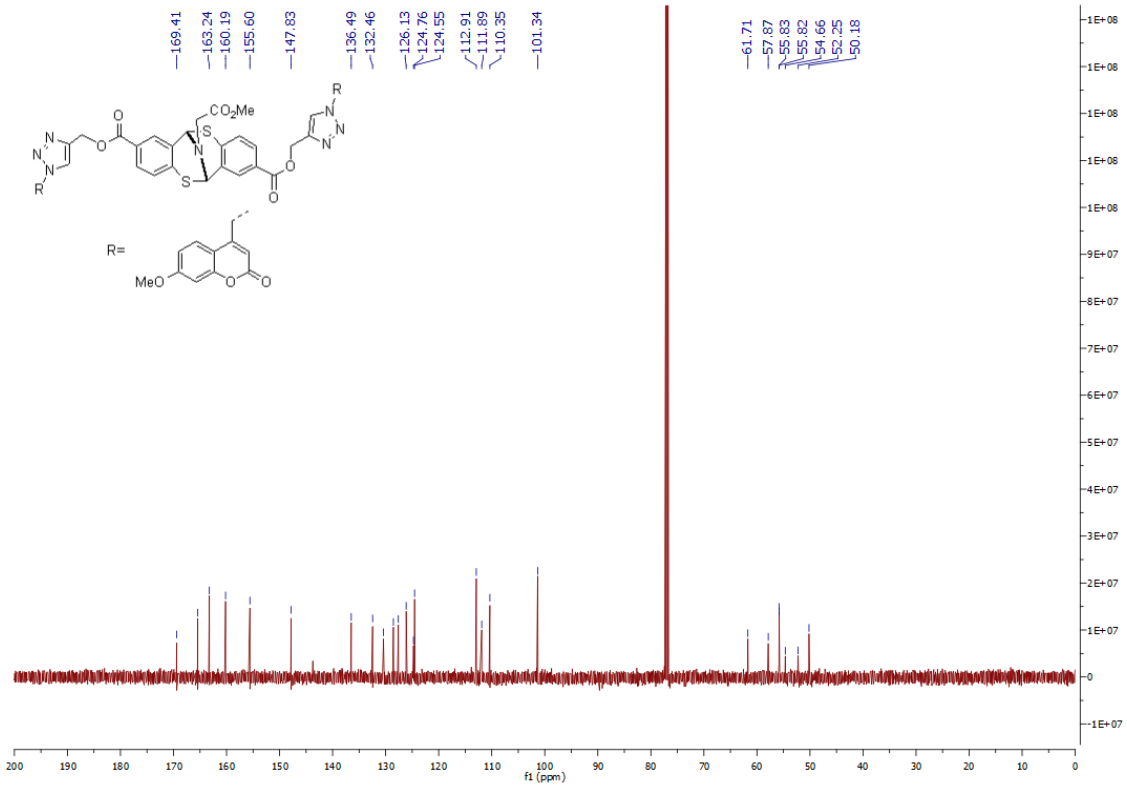
^{1}H NMR spectrum of **13b** in $CDCl_3$ (400 MHz).



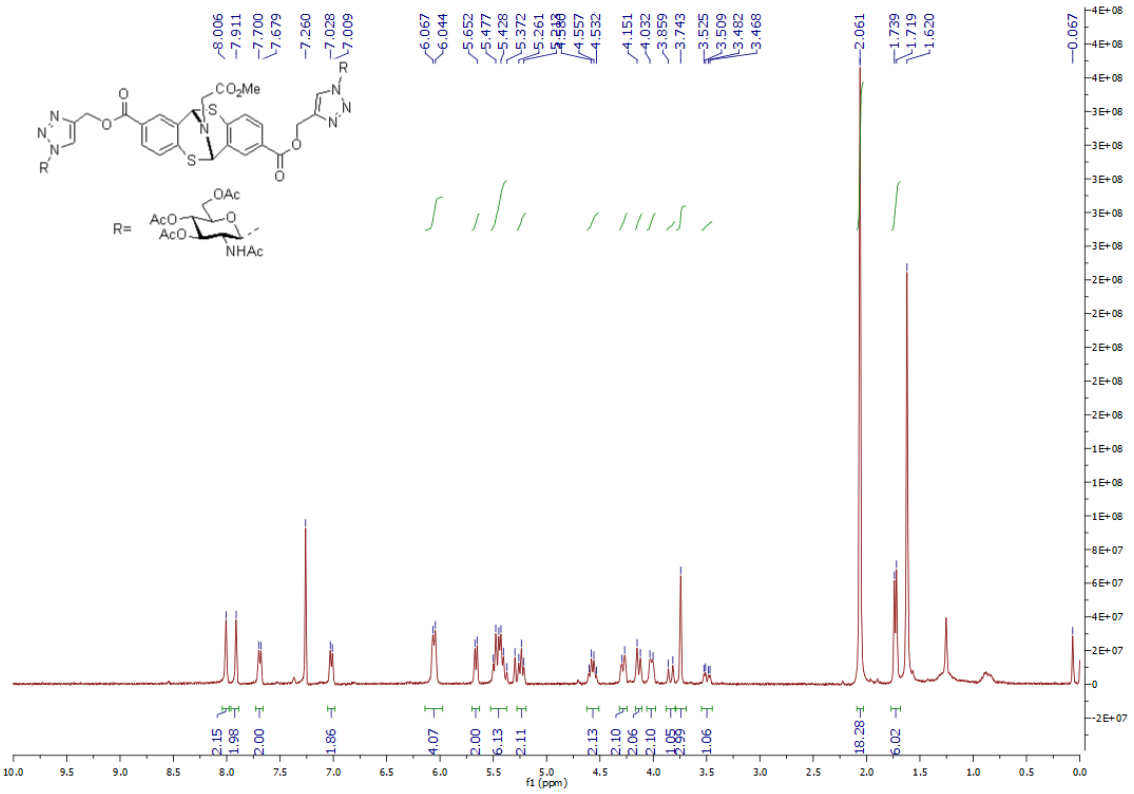
^{13}C NMR spectrum of **13b** in $CDCl_3$ (151 MHz).



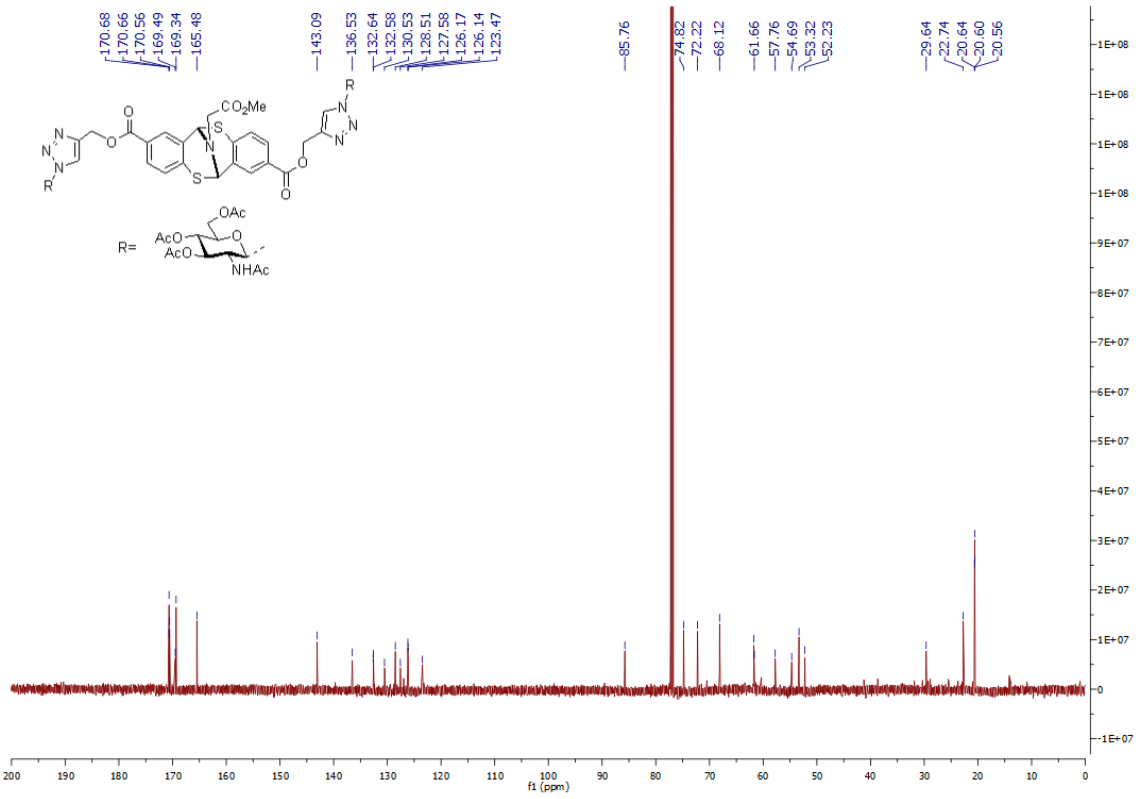
^{1}H NMR spectrum of **13c** in $CDCl_3$ (400 MHz).



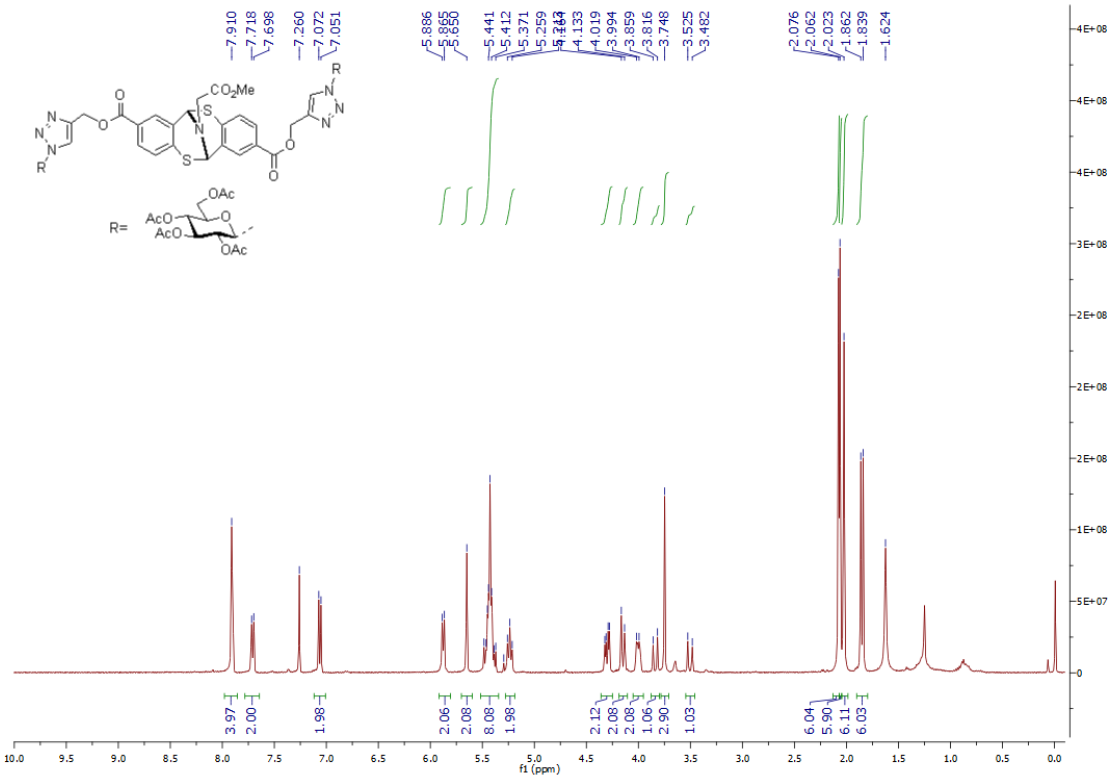
^{13}C NMR spectrum of **13c** in $CDCl_3$ (151 MHz).



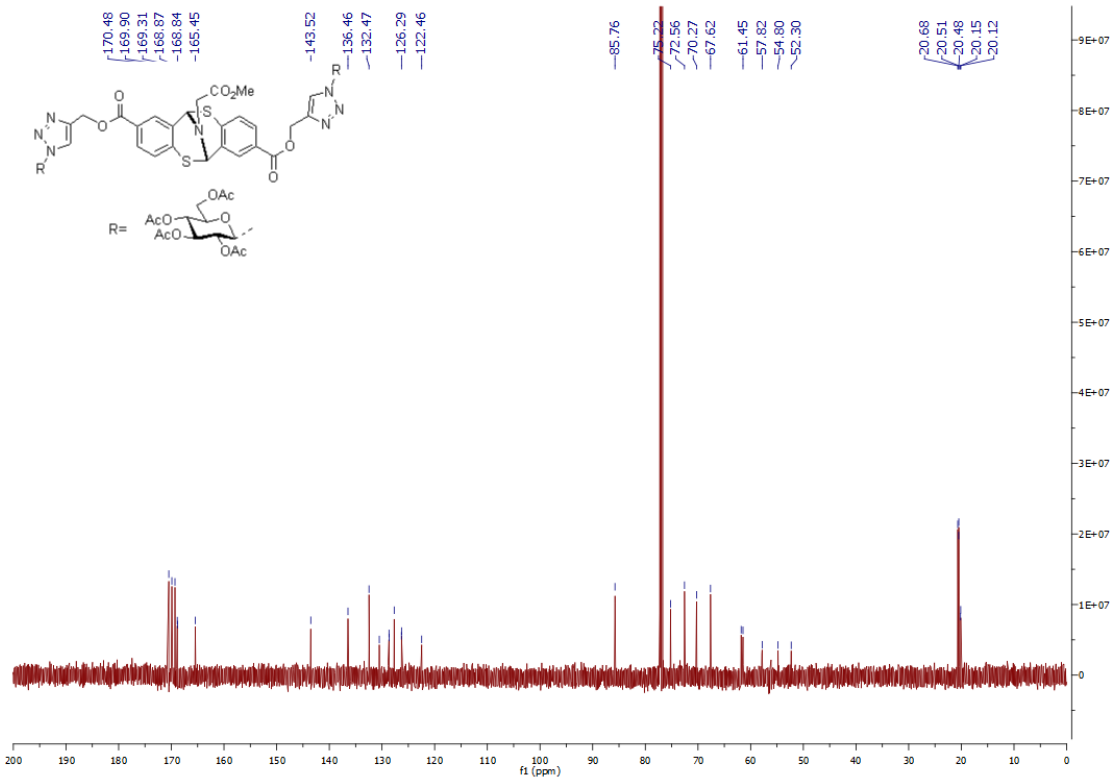
^{1}H NMR spectrum of **13d** in $CDCl_3$ (400 MHz).



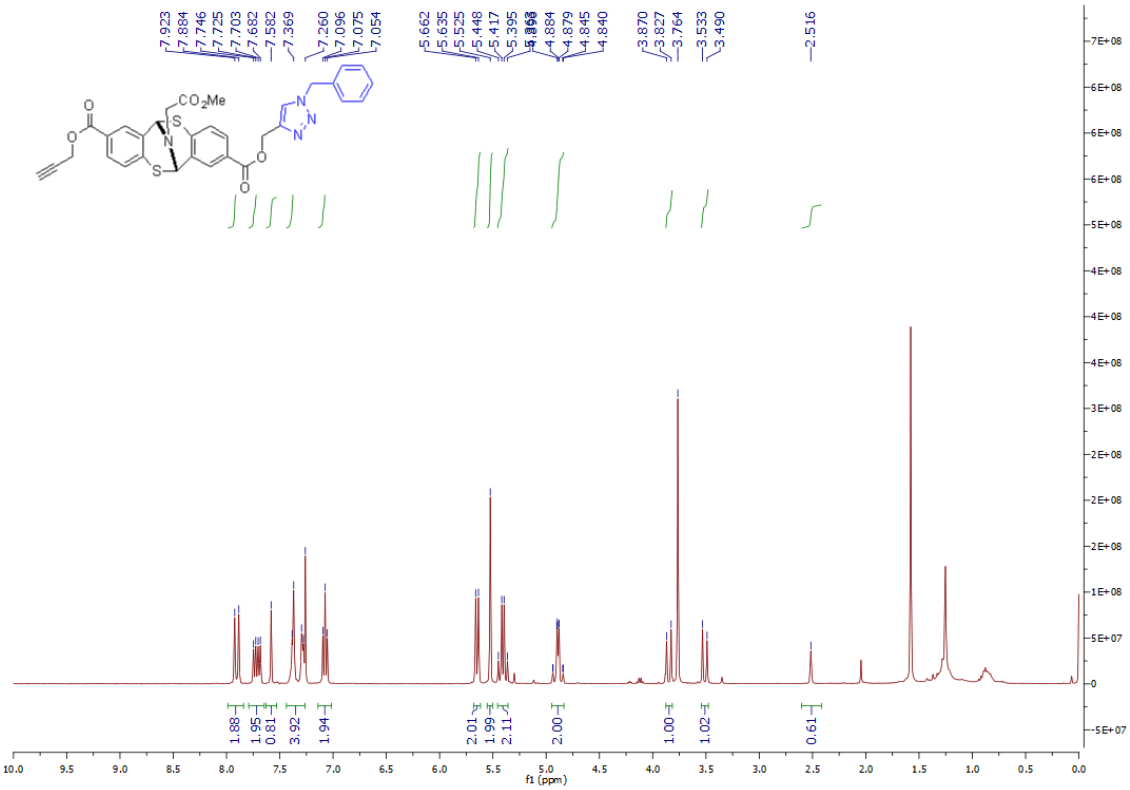
^{13}C NMR spectrum of **13d** in $CDCl_3$ (151 MHz).



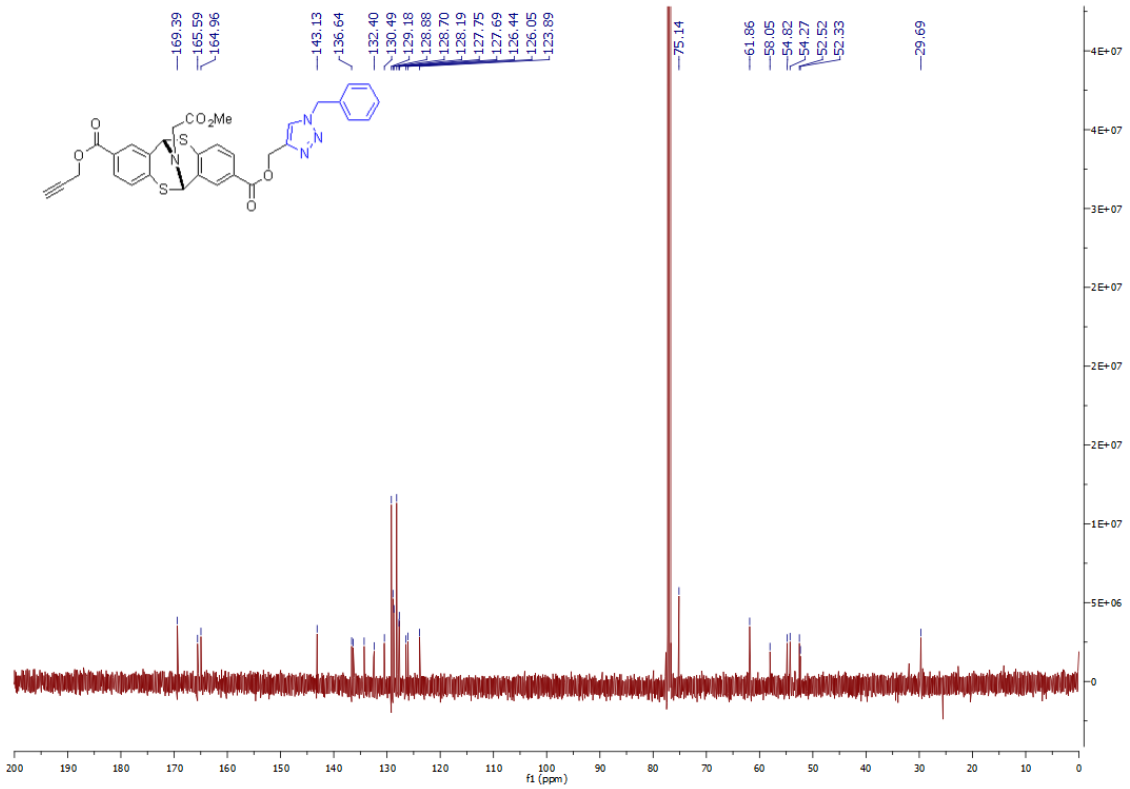
^{1}H NMR spectrum of **13e** in $CDCl_3$ (400 MHz).



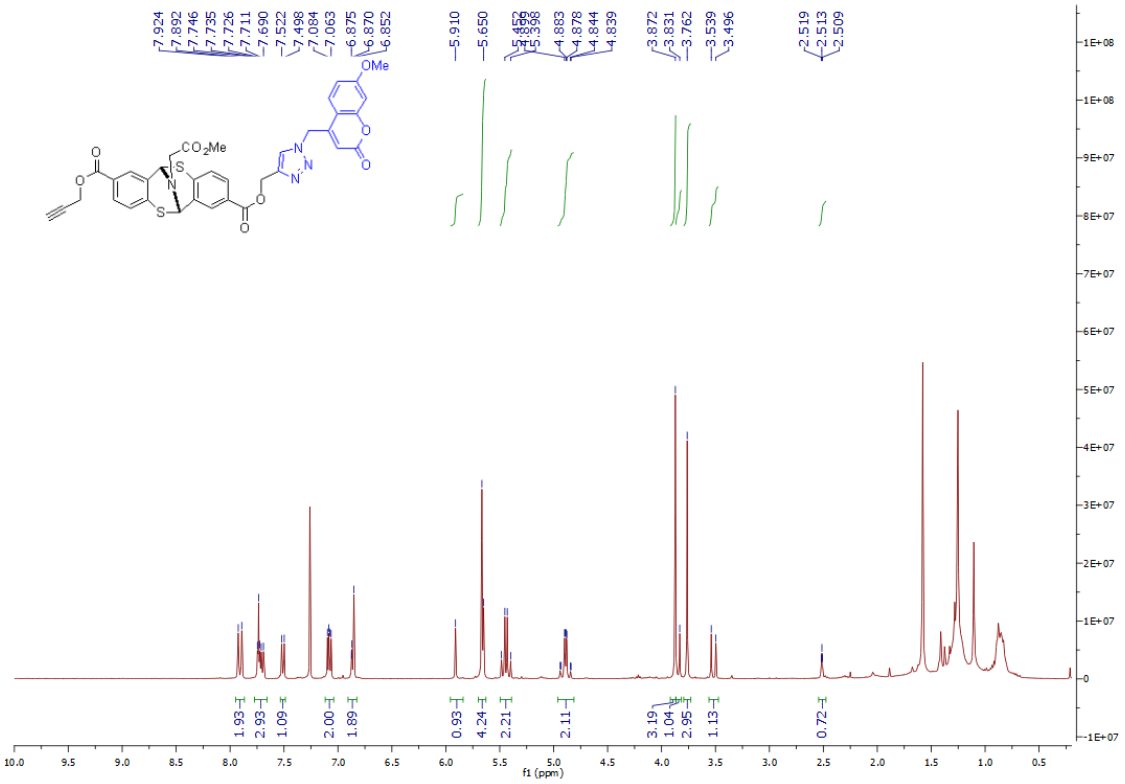
^{13}C NMR spectrum of **13e** in $CDCl_3$ (151 MHz).



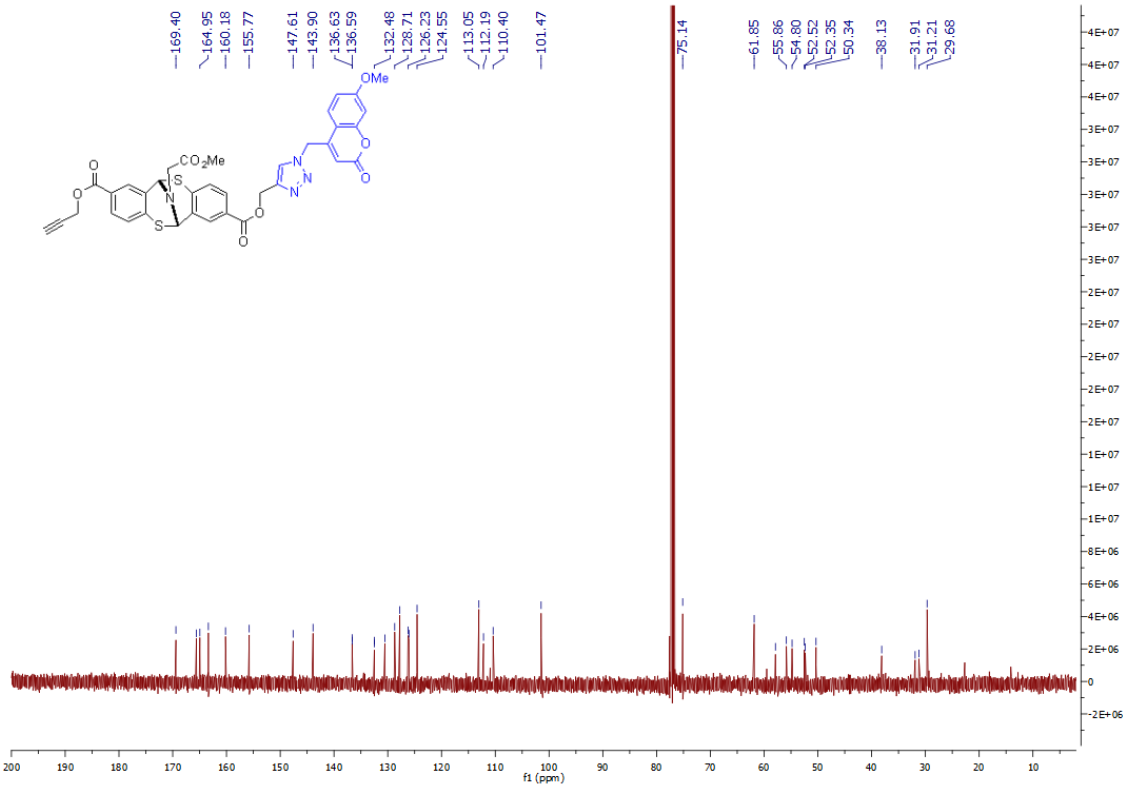
^{1}H NMR spectrum of **15a** in $CDCl_3$ (400 MHz).



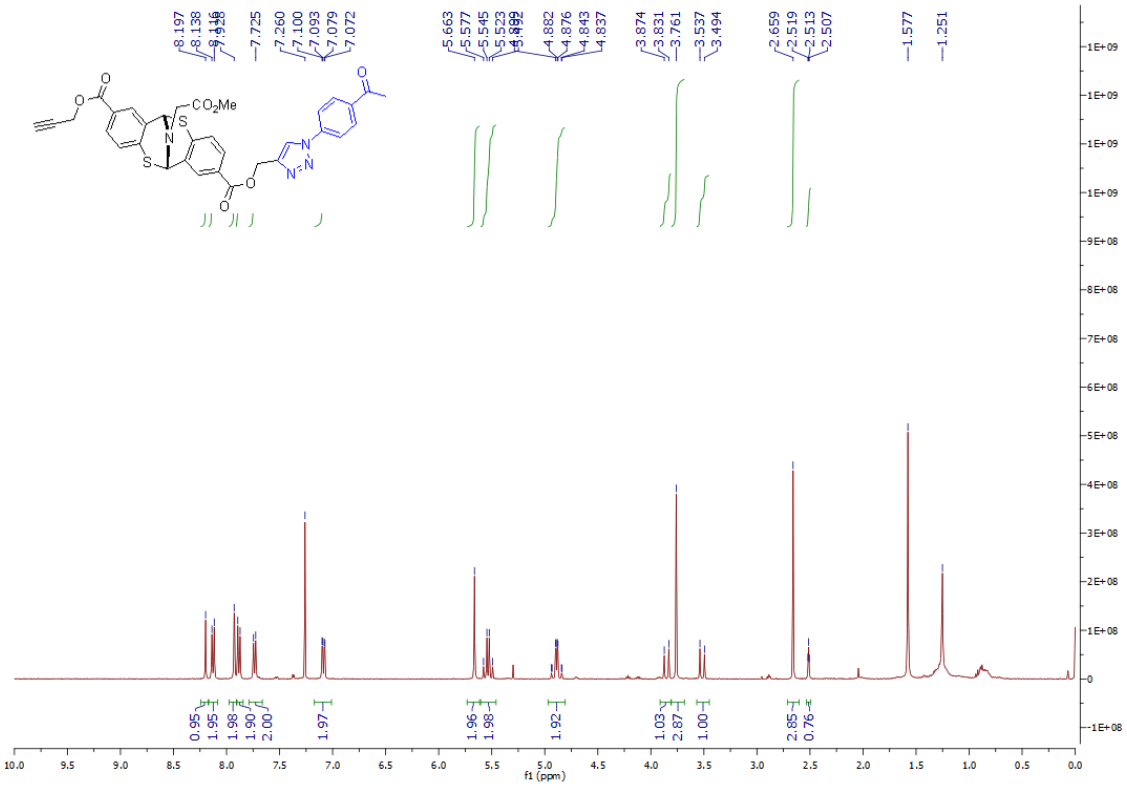
^{13}C NMR spectrum of **15a** in $CDCl_3$ (151 MHz).



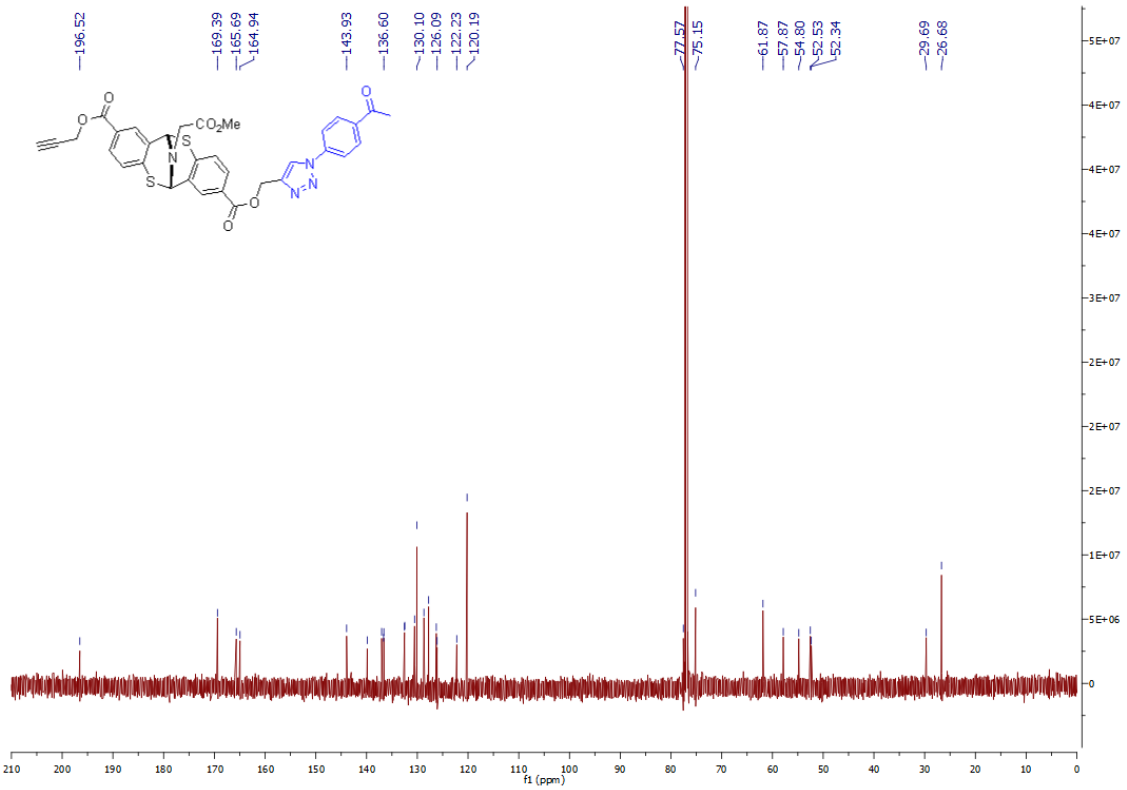
^{1}H NMR spectrum of **15b** in $CDCl_3$ (400 MHz).



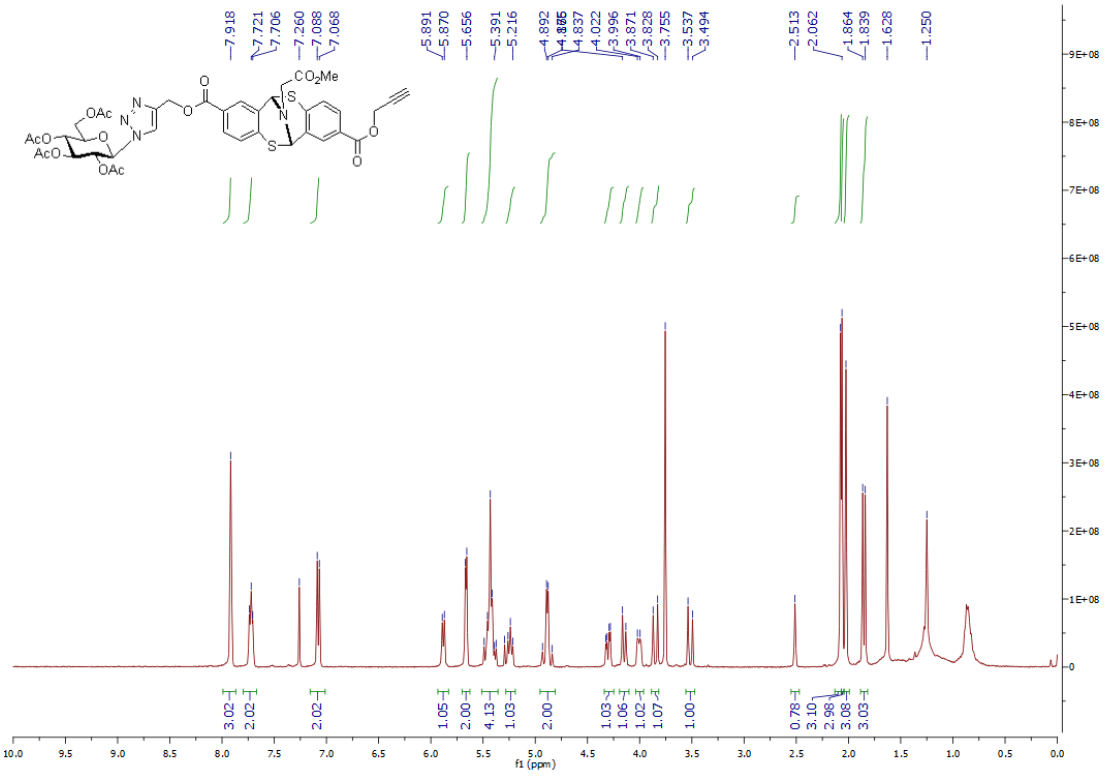
^{13}C NMR spectrum of **15b** in $CDCl_3$ (151 MHz).



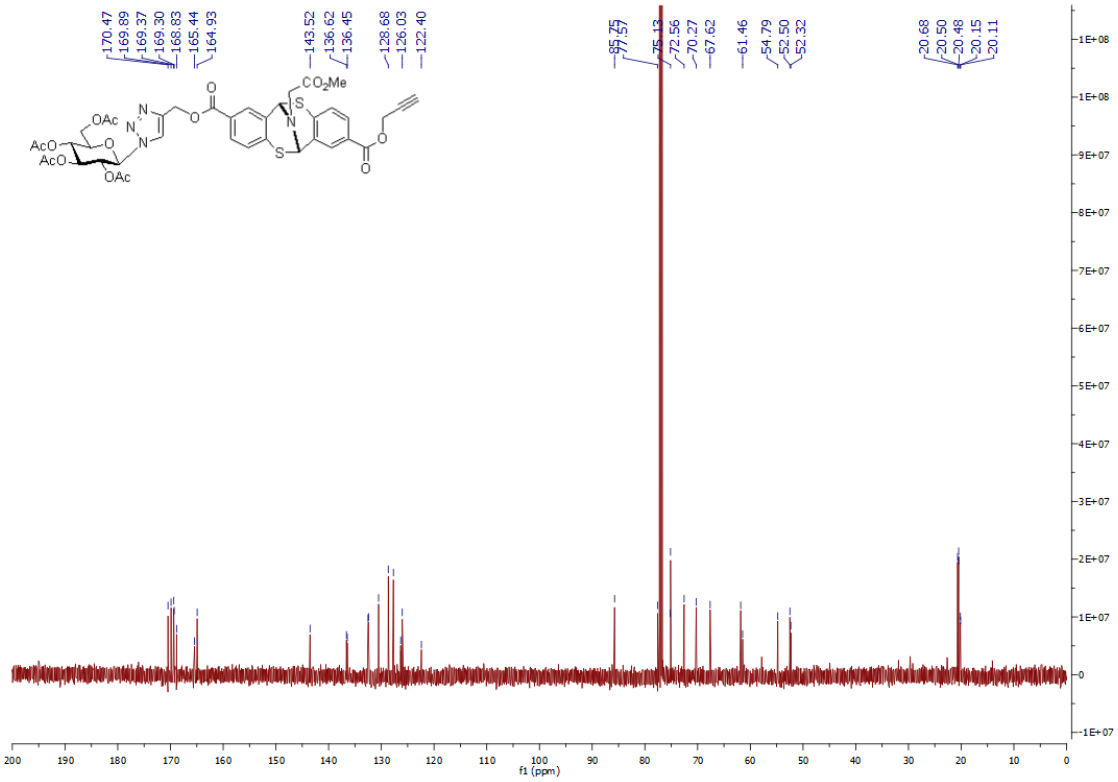
^{1}H NMR spectrum of **15c** in $CDCl_3$ (400 MHz).



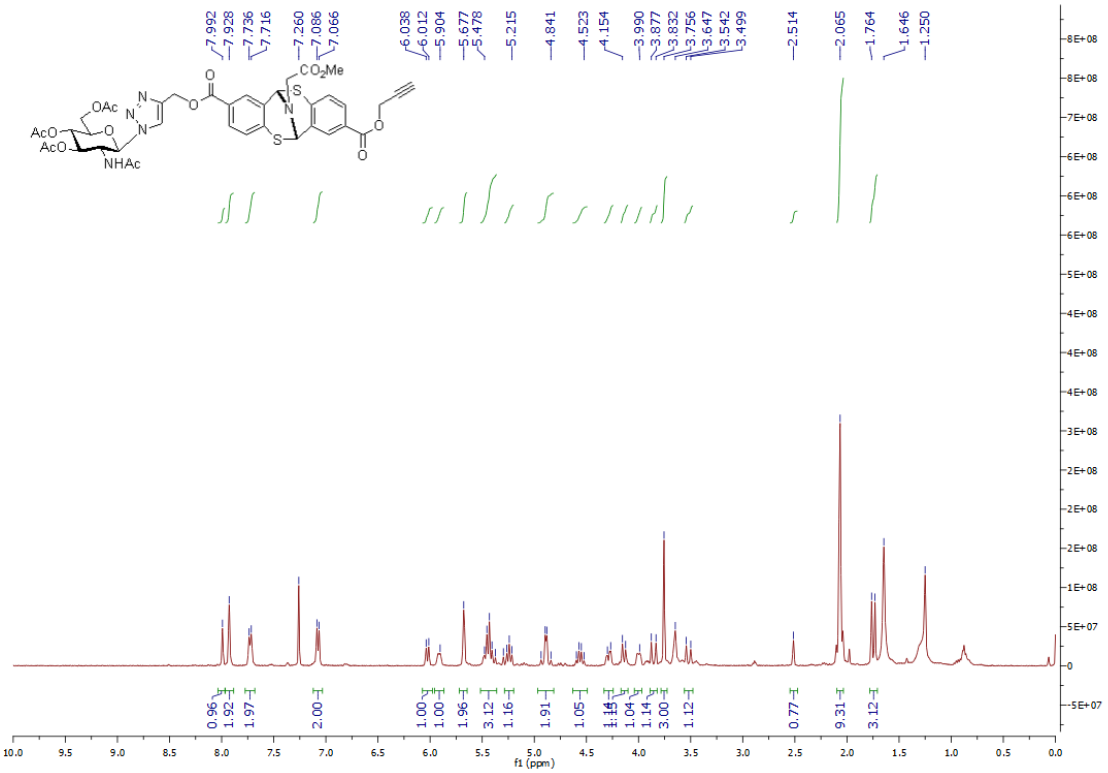
^{13}C NMR spectrum of **15c** in $CDCl_3$ (151 MHz).



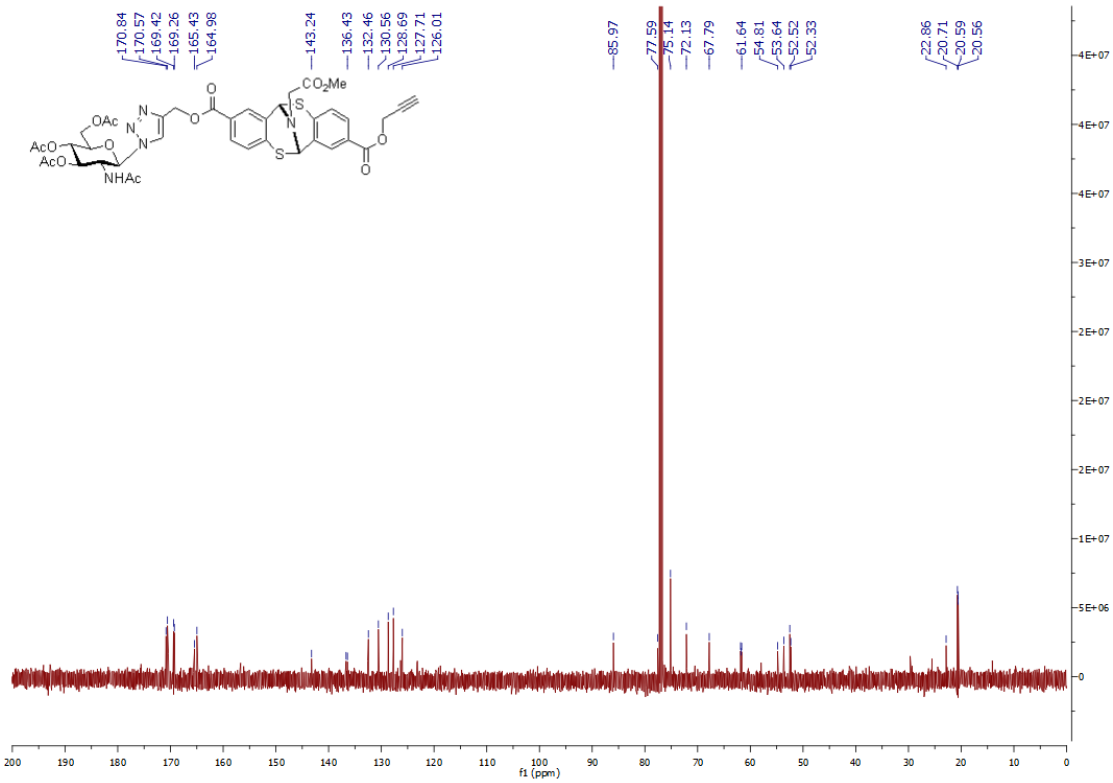
^{1}H NMR spectrum of **15d** in $CDCl_3$ (400 MHz).



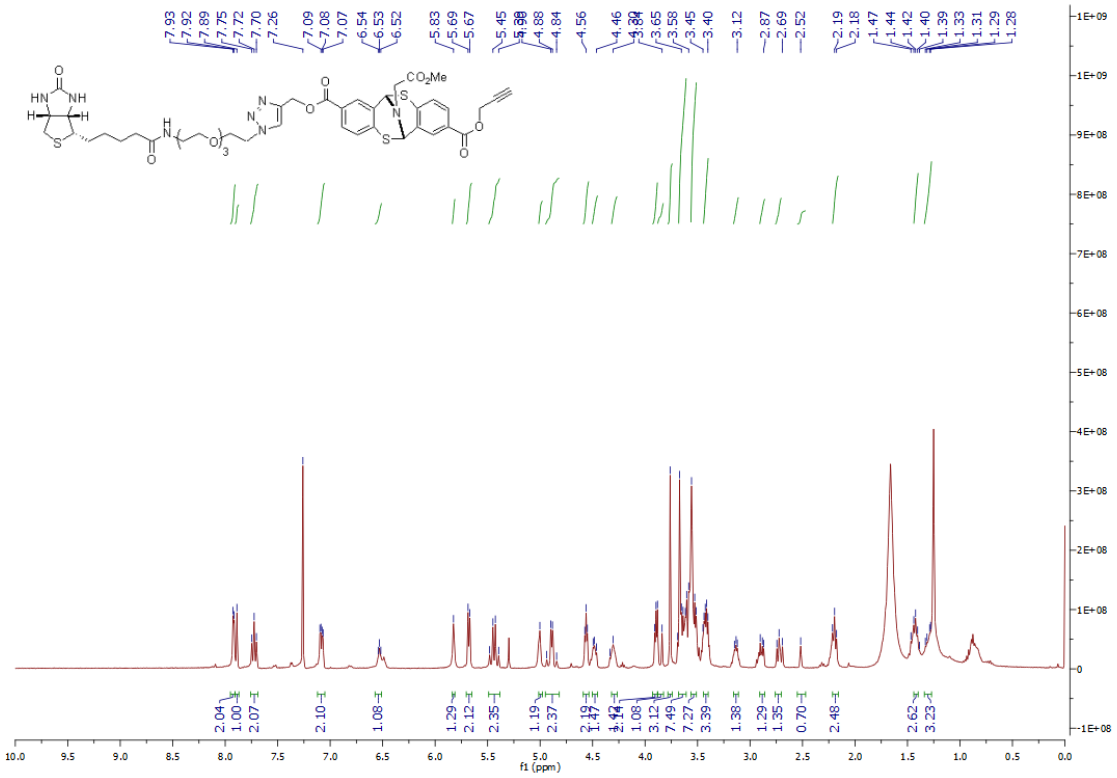
^{13}C NMR spectrum of **15d** in $CDCl_3$ (151 MHz).



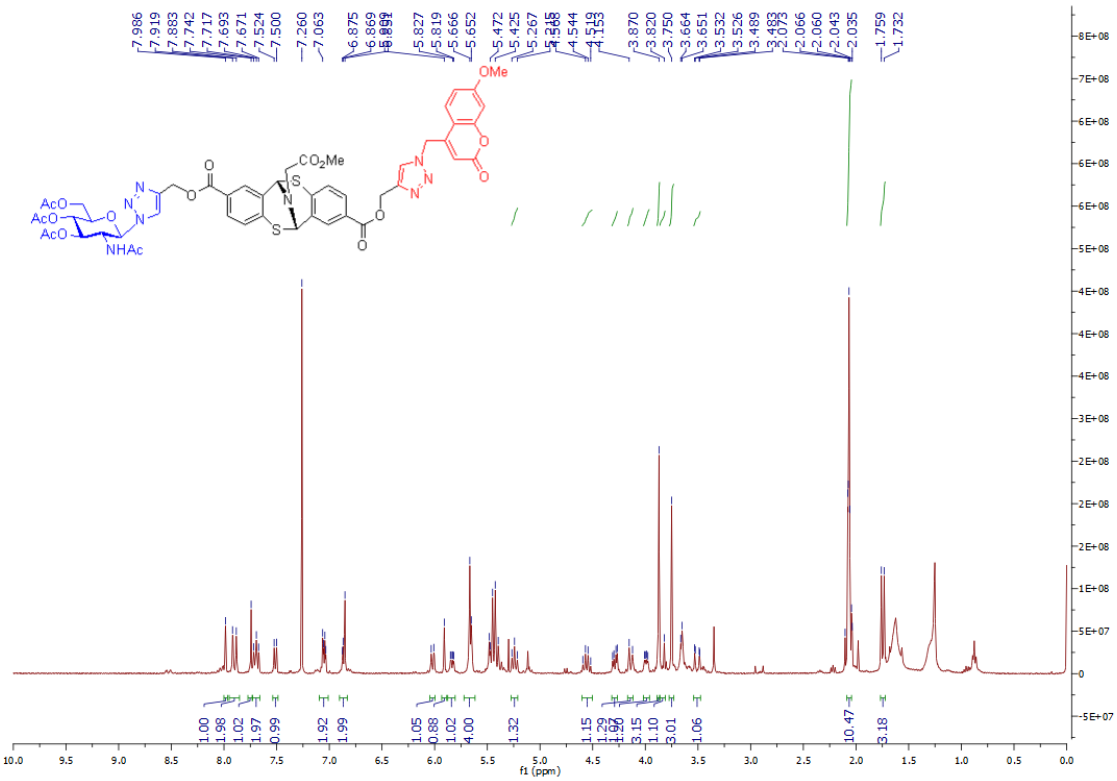
^{1}H NMR spectrum of **15e** in $CDCl_3$ (400 MHz).



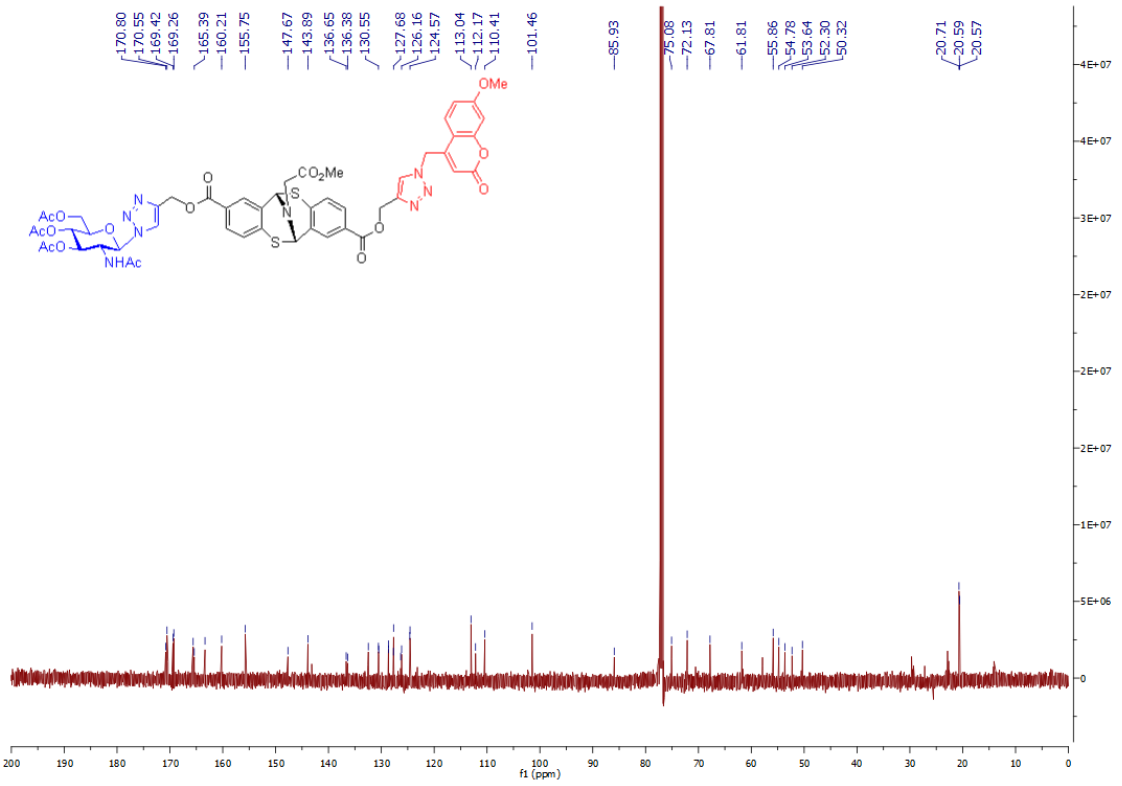
^{13}C NMR spectrum of **15e** in $CDCl_3$ (151 MHz).



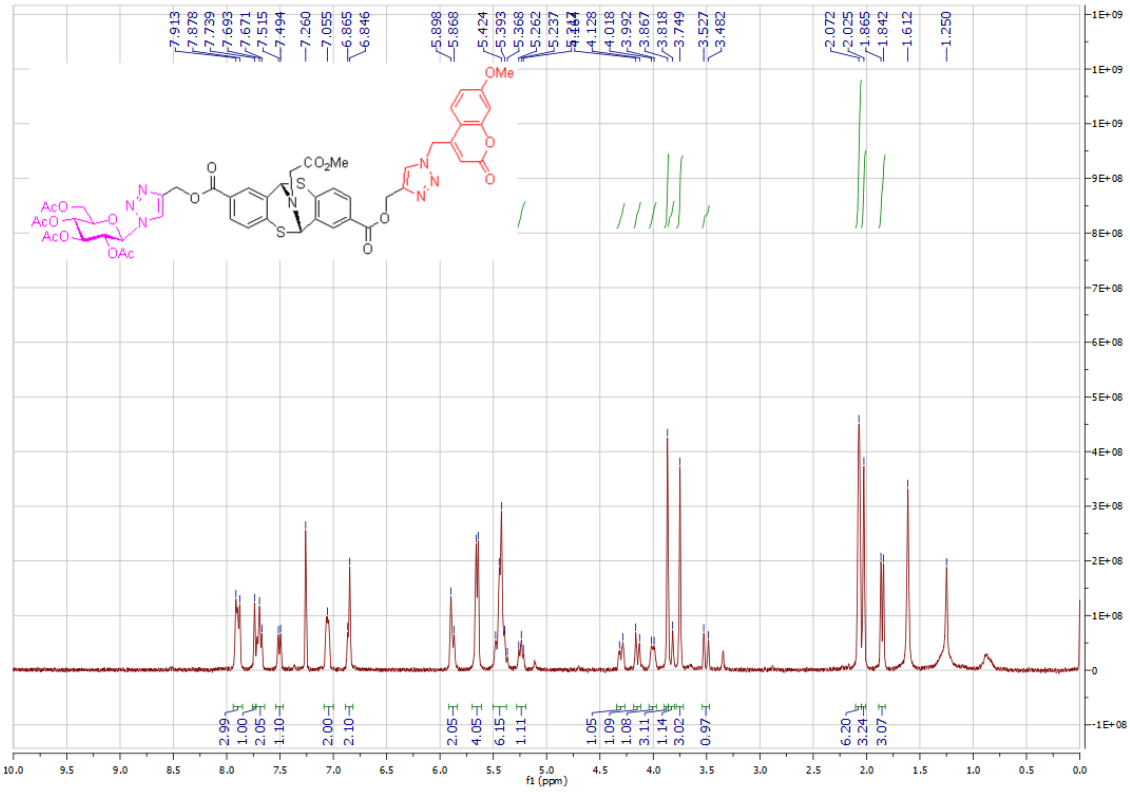
^{1}H NMR spectrum of **15f** in $CDCl_3$ (400 MHz).



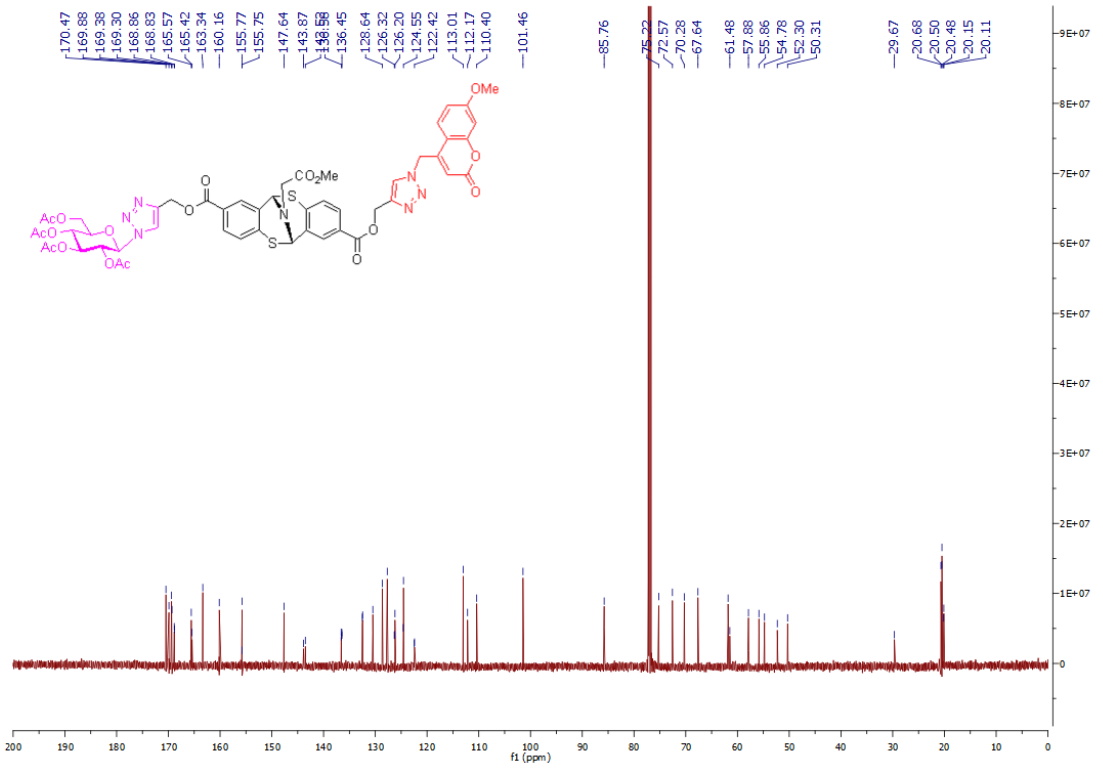
^{1}H NMR spectrum of **16a** in $CDCl_3$ (400 MHz).



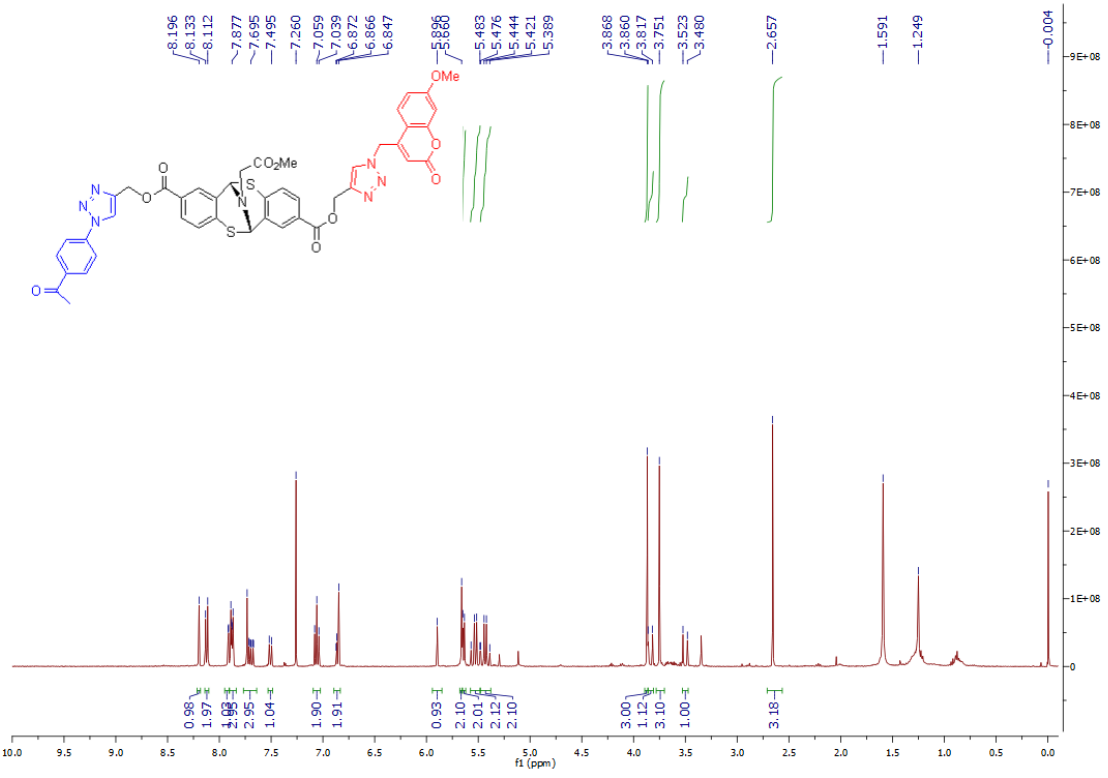
^{13}C NMR spectrum of **16a** in $CDCl_3$ (151 MHz).



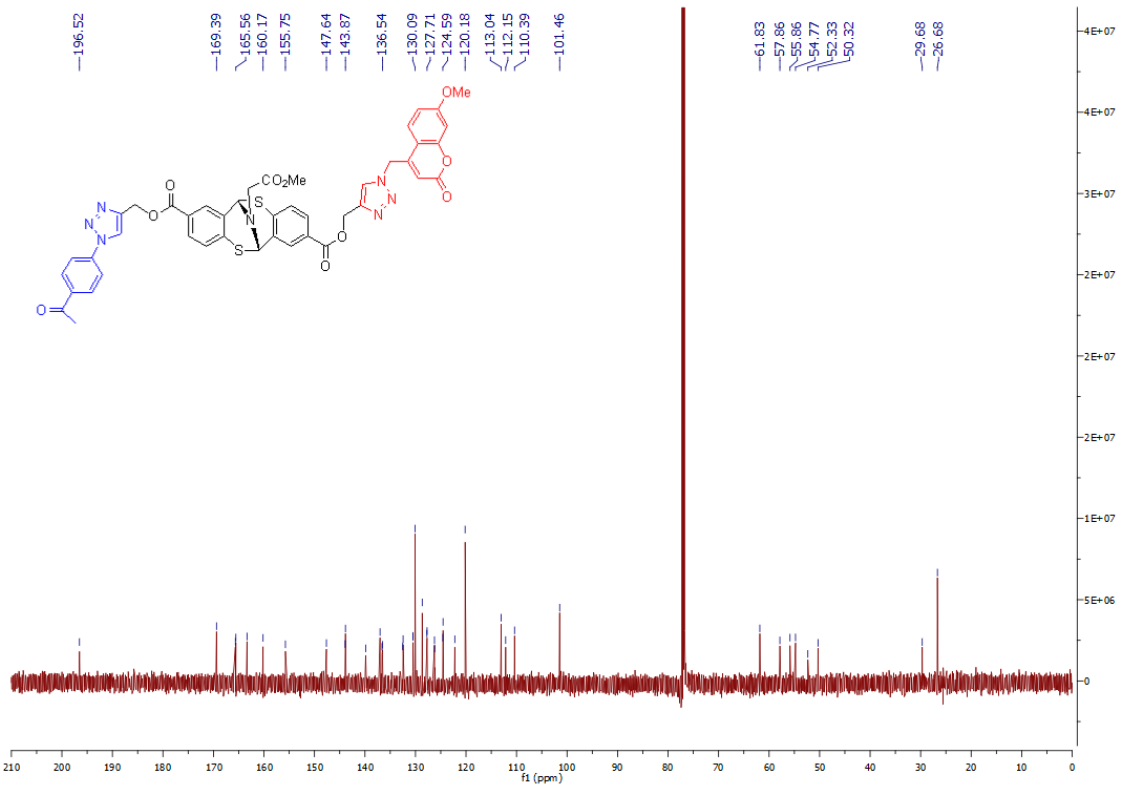
^{1}H NMR spectrum of **16b** in $CDCl_3$ (400 MHz).



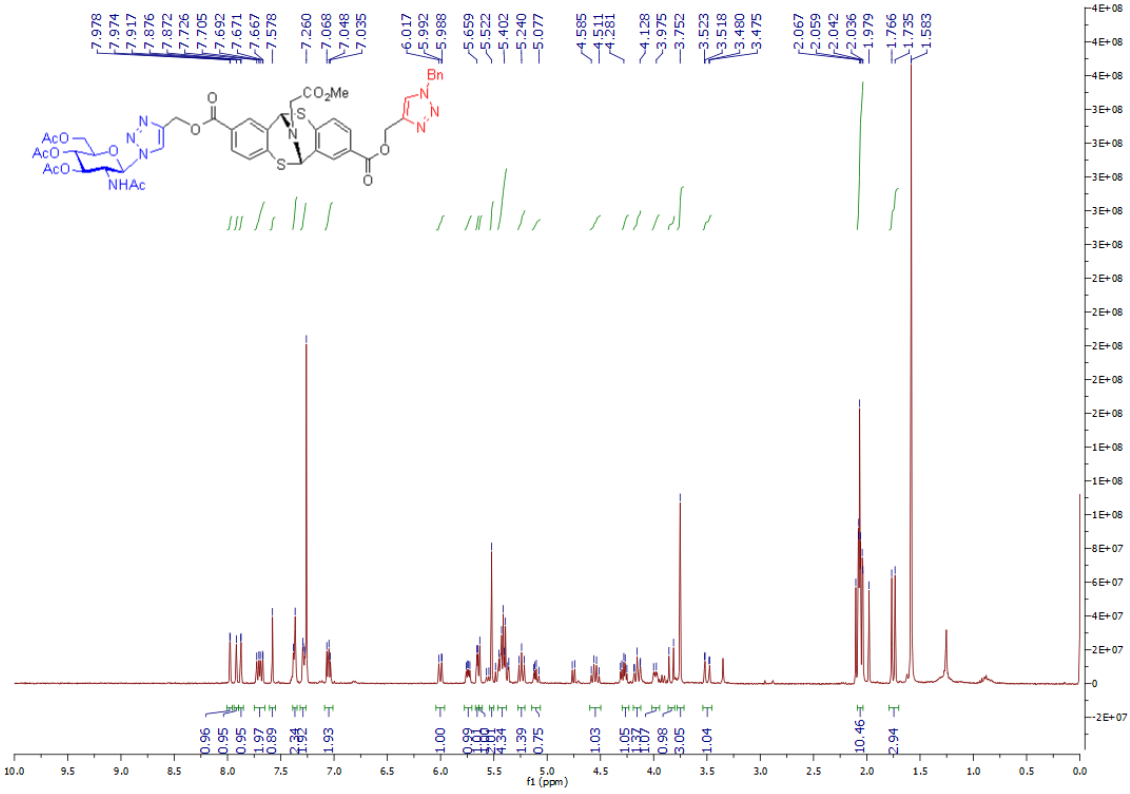
^{13}C NMR spectrum of **16b** in $CDCl_3$ (151 MHz).



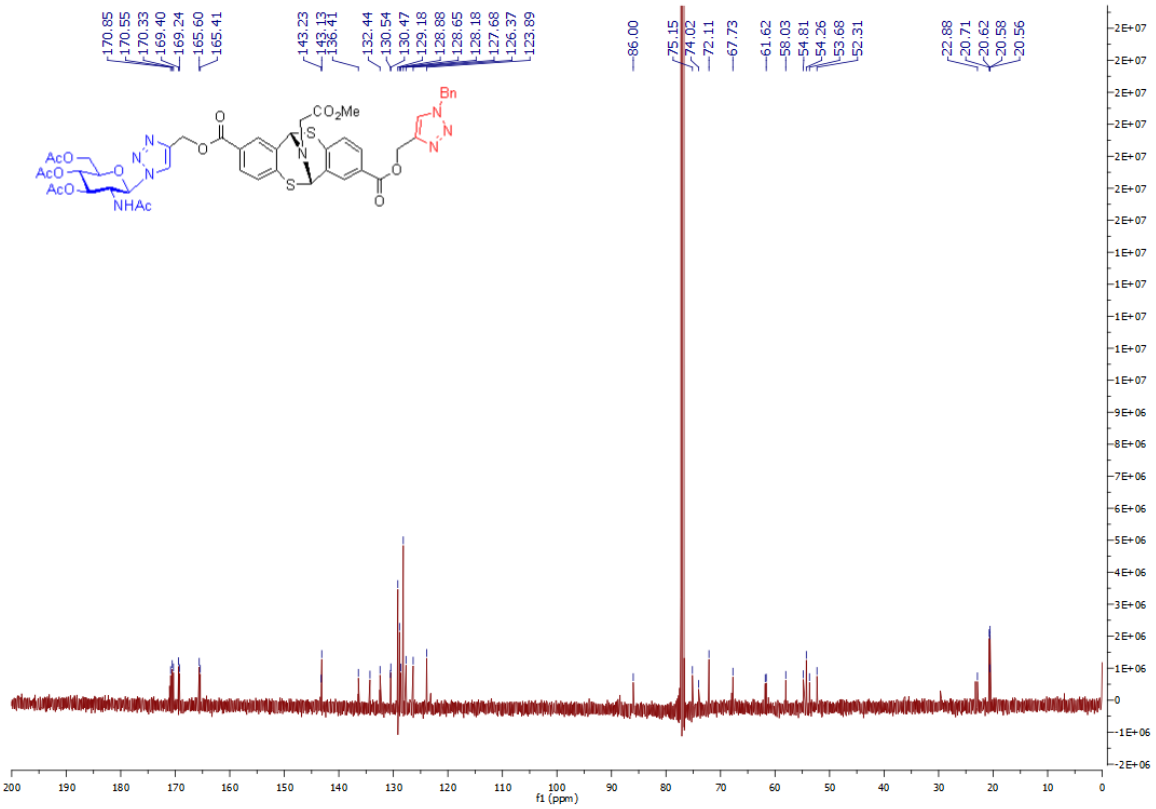
^{1}H NMR spectrum of **16c** in $CDCl_3$ (400 MHz).



^{13}C NMR spectrum of **16c** in $CDCl_3$ (151 MHz).



^{1}H NMR spectrum of **16d** in $CDCl_3$ (400 MHz).



^{13}C NMR spectrum of **16d** in $CDCl_3$ (151 MHz).