Supporting Information

Post-translational site-specific protein azidolation with a pyridoxal derivative

Baochuan Wang,[a] Xun Li,[a] Yiwan Wang,[a] Xianxian Mao,[a,b] and Xiaojian Wang*[a]

[a.] Institute of Advanced Synthesis, School of Chemistry and Molecular Engineering, Nanjing Tech University, Nanjing 211816, China
[b.] School of Pharmacy, Jiangsu Health Vocational College, Nanjing 211800, China

Table of Contents

Abbreviations 1
General materials and methods 3
S1. Synthesis and characterization of small compounds 4
S2. Analysis of azidation of glycine derivatives 9
S3. Site-specific azidation of synthesized peptides 14
S4. Site-specific azidation of insulin 38
S5. Imine hydrolysis 42
S6. Site-specific azidation of myoglobin 43
S7. Site-specific azidation of AzoR 45
S8. Examine the activity of N-Gly-AzoR-PLAA 46
S9. General procedure for the SDS-PAGE analysis 47
S10. NMR Spectra 48
S11. UV Spectra 55
References 56

Abbreviations

$CDCl_3$……..……..…..……..…....Deuterated chloroform

CD_3CN……..…….…….....……Acetonitrile-d_3

$CuSO_4$……..…….Copper sulphate

CS_2CO_3.....Cesium carbonate

DCM……..…….Dichloromethane

DMF……..……..……...…….……..N, N-Dimethylformamide

DMSO……..…..…….……Dimethyl sulfoxide

DMSO-d_6……..……..…….…….…….Dimethyl sulfoxide-d_6

EDC·HCl…..…..……1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride

ESI.....Electrospray ionization

Et_3N……..……..……...……..……..…..Triethylamine

EtOAc/EA…….....…….....….....Ethyl acetate

FITC……Fluorescein isothiocyanate isomer I

g……..……..gram

H……..……..…….....Proton

HBTU……...…..…1-Hydroxy Benzotriazole

H_2O…….…….Water

HOBt……..…1-Hydroxybenzotriazole

HPLC……..……..…High performance liquid chromatography

HRMS…….…….....High-resolution mass spectrometry

Hz…….……Hertz

JEOL……..…….Japan Electron Optics Laboratory CO., LTD

M…….……...…….…...……...Molar

MeOH/CH_3OH..……..…Methanol

mg……...…milligram

min……..Minute

mL……..…..Milliliter

mmol…….……..……....Millimoles

MnO_2……..……..……..…Manganese dioxide

N_2……...……...…….....Nitrogen

$NaHCO_3$……...Sodium bicarbonate

NMR……Nuclear magnetic resonance

nmol……..……..……Nanomoles

$Pd(PPh_3)_2Cl_2$……..……Bis(triphenylphosphine)palladium(II) chloride

PE……Petroleu mether

ppm……..……..……...…..parts per million

SDS-PAGE……..…Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

TBTA…(1-(1-Benzyltriazol-4-yl)-N,N-bis[(1-benzyltriazol-4-yl)methyl]methanamine

TFA……..Trifluoroacetic acid

THF…….……..……..……...……Tetrahydrofuran

TLC…...……..…..…….……Thin layer chromatography

Tris……...…Tris(hydroxymethyl)aminoethane

UV……...……..……..……Ultraviolet

µmol……...……..……....…Micr moles

µL……..……..Micro litter

General materials and methods

All chemicals and solvents were purchased from commercial sources and used without further purification unless otherwise indicated. All reactions were performed under anhydrous conditions under an atmosphere of nitrogen. Reactions were monitored by TLC on HSGF254 silica gel plates. Detection was accomplished by examination under UV light (254 nm or 365 nm). Flash chromatography was performed on silica gel (100-200 mesh). ^{1}H NMR spectra were recorded in $CDCl_3$, DMSO-d_6, and CD_3CN on Bruker AVB-400 or JEOL ECZ400S spectrometer at 298K. TMS (δ *(ppm)*$_H$ = 0.00) was used as the internal reference. ^{13}C NMR spectra were recorded in either $CDCl_3$, DMSO-d_6, and CD_3CN at 100 MHz on Bruker AVB-400 or JEOL ECZ400S spectrometer, using the central resonances of $CDCl_3$ (δ *(ppm)*$_C$ = 77.16), DMSO-d_6 (δ *(ppm)*$_C$ = 39.52) or CD_3CN (δ *(ppm)*$_{CD_3}$ = 1.32, δ *(ppm)*$_{CN}$ = 118.26) as the internal references. Chemical shifts were reported in ppm and multiplicities were indicated by s (singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublets), and m (multiplet). Coupling constants, *J*, are reported in hertz (Hz). High-resolution mass spectra (HRMS) were obtained on a Waters ACQUITY UPLC/Xevo G2-XS Qtof system and are reported as m/z (relative intensity). Accurate masses are reported for the molecular ion (M^+) or a suitable fragmention. ESI-MS(/MS) spectra were recorded on a Thermo Liquid chromatography-mass spectrometry (Thermo Fisher Scientific, MSQ PLUS/U3000) equipped with a standard ESI ion source. Data acquisition and analysis were done with the Xcalibur (version 2.0, Thermo quest Finnigan) software package. Both analytical and preparative HPLC were performed on an Elitehplc P230AP system with a D3100 PDA detector.

S1. Synthesis and characterization of small compounds

S1.1 Routes of Synthesis

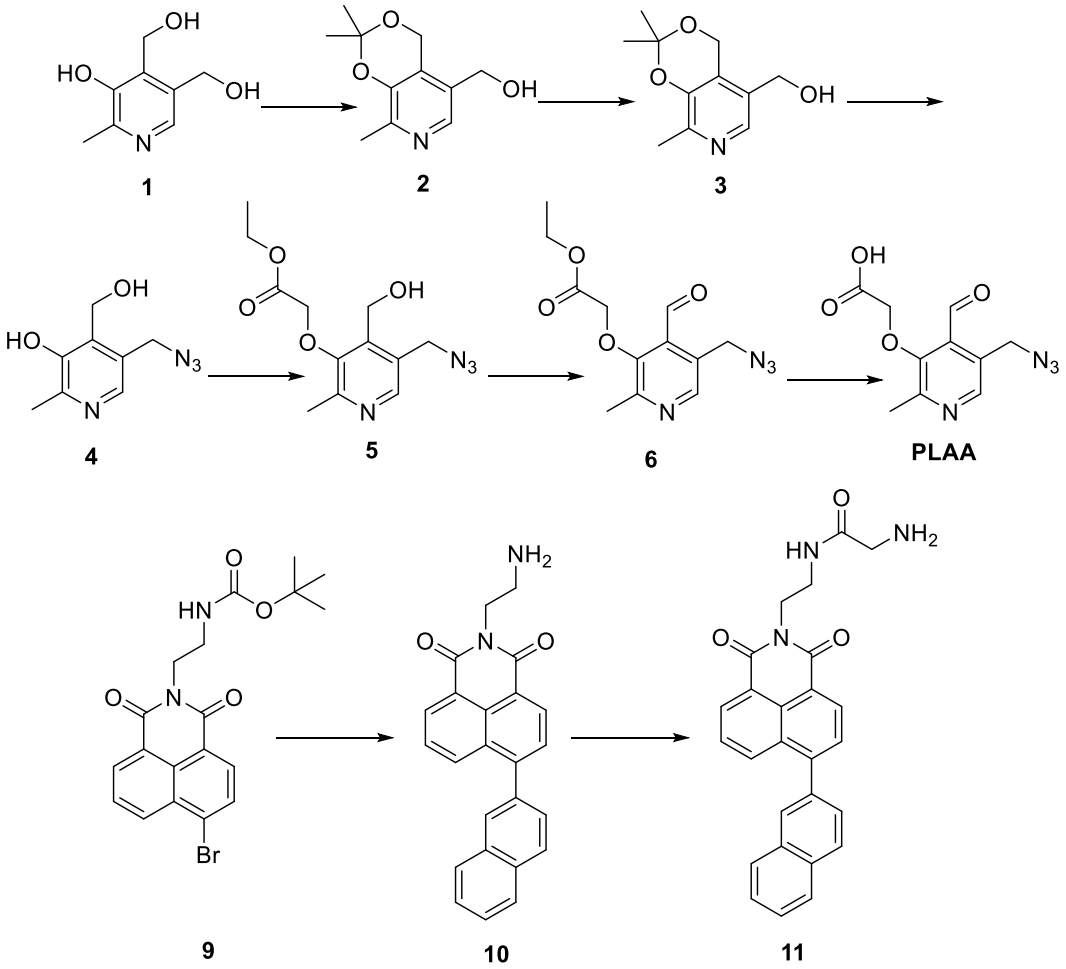


Figure S1. Routes of synthesis.

S1.2 Synthesis of ethyl 2-((5-(azidomethyl)-4-(hydroxymethyl)-2-methylpyridin-3-yl)oxy)acetate (5):

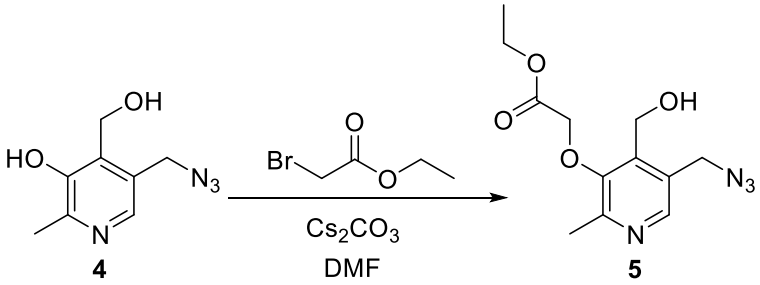


Figure S2. Synthesis of compound **5**.

Compound **1-4** was prepared following the reported method.[1] Compound **5** was prepared as following: in a 50 mL round bottom flask, compound **4** (388.2 mg, 2 mmol),

and cesium carbonate (651.6 mg, 2 mmol) was dissolved in DMF (15 mL). After stirring for 5 minutes, ethyl 2-bromoacetate (317.3 mg, 1.9 mmol) was added. The reaction mixture was allowed to stir at room temperature for 12 hours. The reaction was monitored using thin layer chromatography and upon completion, the reaction mixture was filtered to remove cesium carbonate. The solution was concentrated under vacuum and the product was purified using column chromatography (5:1-1:1, PE/EA) to afford compound **5** (224 mg, 40 %) as a yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$): *δ (ppm)* 8.26 (s, 1H), 4.80 (s, 2H), 4.61 (s, 2H), 4.52 (s, 2H), 4.28 (q, *J* = 7.2 Hz, 2H), 2.56 (s, 3H), 1.32 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): *δ (ppm)* 169.70, 152.84 152.25, 145.73, 140.75, 128.86, 70.36, 61.99, 55.99, 49.86, 20.29, 14.22. HRMS m/z Found: 281.1250, calculated: 281.1250 for $C_{12}H_{17}N_4O_4$ $[M+H]^+$.

S1.3 Synthesis of ethyl 2-((5-(azidomethyl)-4-formyl-2-methylpyridin-3-yl)oxy)acetate (6):

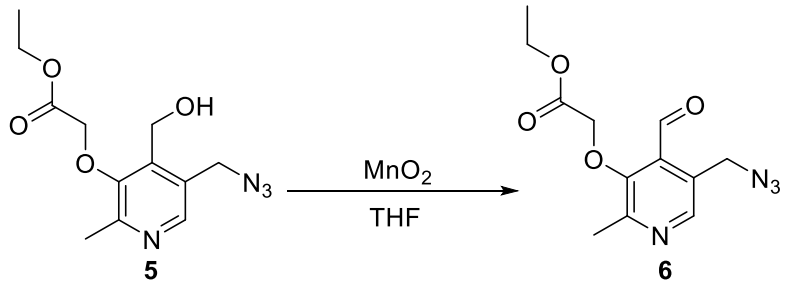


Figure S3. Synthesis of compound **6**.

In a 100 mL round bottom flask, manganese dioxide (MnO_2, 4.68 g, 53.86 mmol) was added to a stirred suspension of compound **5** (303.3 mg, 1.26 mmol) in THF (30 mL) at room temperature. The reaction mixture was stirred for 12 hours and then centrifuged. The supernatant was concentrated under reduced pressure and 30 mL of THF was added to the precipitate after centrifugation with stirring, and then reacted for additional 12 hours. The suspension was centrifuged and the supernatant was concentrated again under reduced pressure. The process was repeated for 2 more times and the crude product was combined and purified via silica gel column chromatography (3:1-2:1, PE/EA) to afford compound **6** (210 mg, 60 %) as a yellow solid. ^{1}H NMR (400 MHz, $CDCl_3$): *δ (ppm)* 10.74 (s, 1H), 8.49 (s, 1H), 4.76 (s, 2H), 4.63 (s, 2H), 4.26 (q, *J*=7.2 Hz, 2H), 2.64 (s, 3H), 1.30 (t, *J* = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, CDCl3): *δ (ppm)* 192.09, 168.11, 154.52, 154.43, 146.28, 132.13, 128.09, 70.76, 61.91, 49.59, 19.91, 14.19. HRMS m/z Found: 279.1093, calculated: 279.1093 for $C_{12}H_{15}N_4O_4$

$[M+H]^+$.

S1.4 Synthesis of 2-((5-(azidomethyl)-4-formyl-2-methylpyridin-3-yl)oxy)acetic acid (PLAA):

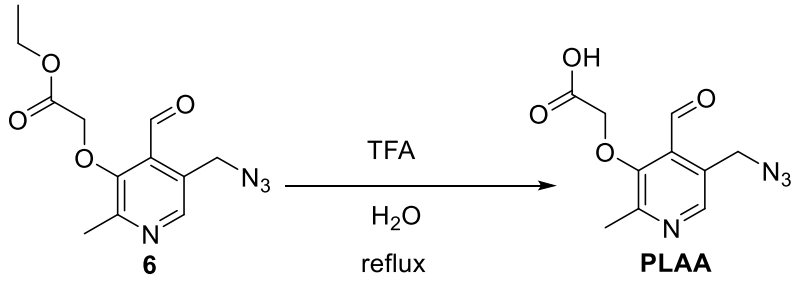


Figure S4. Synthesis of **PLAA**.

In a 50 mL round bottom flask, compound **6** (367.4 mg, 1.2 mmol) was dissolved in H_2O (15 mL). After stirring for 5 minutes, trifluoroacetic acid (0.4 mL, 5.38 mmol) was added. The reaction mixture was allowed to reflux for 12 h. The solution was concentrated under vacuum and the product was purified using column chromatography (1:0-10:1, EA/MeOH) to afford compound **PLAA** (180.5 mg, 60 %). ^{1}H NMR (400 MHz, DMSO-d_6): *δ (ppm)* 10.59 (s, 1H), 8.43 (s, 1H), 4.74 (s, 2H), 4.71 (s, 2H), 2.55 (s, 3H). ^{13}C NMR (100 MHz, DMSO-d_6): *δ (ppm)* 192.70, 170.22, 154.99, 153.94, 145.98, 132.28, 127.25, 70.55, 48.53, 19.65. HRMS m/z Found: 251.0780, calculated: 251.0780 for $C_{10}H_{11}N_4O_4$ $[M+H]^+$.

S1.5 Synthesis of 2-(2-aminoethyl)-6-(naphthalen-2-yl)-1H-benzo[de]isoquinoline-1,3(2H)-dione (10):

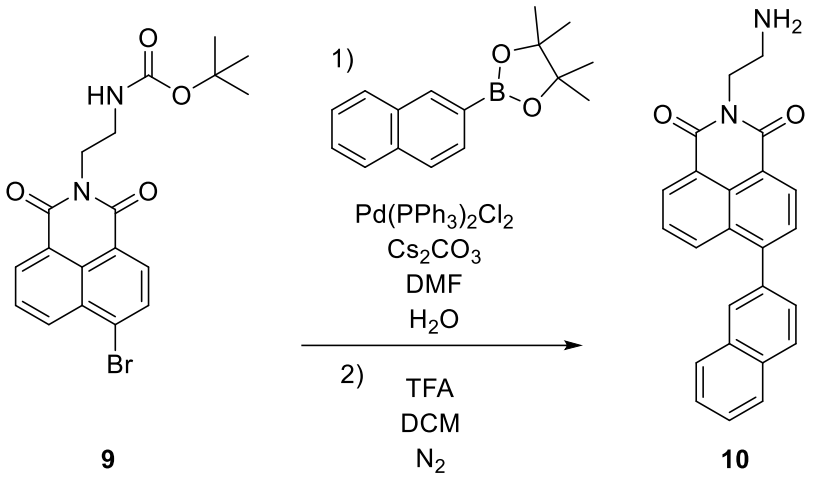


Figure S5. Synthesis of compound **10**.

Compound **9** was prepared following the reported method.[2] Compound **10** was

prepared as following: in a 25 mL round bottom flask, compound **9** (419.3 mg, 1 mmol), and cesium carbonate (1.3 g, 4 mmol) was dissolved in DMF (4 mL) and H_2O (2 mL). After stirring for 5 minutes, pinacol ester (381.2 mg, 1.5 mmol) and palladium bis(triphenylphosphine) dichloride (30 mg, 0.04 mmol) were added. The reaction mixture was allowed to stir at room temperature. The result was monitored by using thin layer chromatography. When the experiment was completed, 200 mL of water and 50 mL of ethyl acetate were used for extraction. The organic phase is concentrated under vacuum to obtain a yellow solid crude product. Then the crude product obtained above was dissolved in DCM (10 mL). TFA (2.5 mL) was added under ice bath condition. The temperature of the reaction solution slowly rises to room temperature. It was monitored by TLC until the end of the reaction. The solution was concentrated under vacuum and the product was purified using column chromatography (1:0-10:1, DCM/MeOH) to afford compound **10** (292.8 mg, 80 %) as a yellow solid. ^{1}H NMR (400 MHz, DMSO-d_6): *δ (ppm)* 8.57 (dd, *J* = 9.9, 7.6 Hz, 2H), 8.32 (d, *J* = 8.5 Hz, 1H), 8.14 (d, *J* = 9.1 Hz, 2H), 8.09 – 8.04 (m, 2H), 7.99-7.91 (m, 3H), 7.86 (t, *J* = 7.9 Hz, 1H), 7.68 (d, *J* = 8.5 Hz, 1H), 7.66 – 7.59 (m, 2H), 4.36 (t, *J* = 5.7 Hz, 2H), 3.20 (t, *J* = 5.7 Hz, 2H). ^{13}C NMR (100 MHz, DMSO-d_6): *δ (ppm)* 164.16, 164.93, 146.13, 135.75, 132.89, 132.62, 132.45, 130.81, 130.38, 129.42, 128.98, 128.39, 128.39, 128.21, 127.76, 127.65, 127.63, 127.01, 126.91, 122.74, 121.64, 37.64, 37.61. HRMS m/z Found: 367.1143, calculated: 367.1147 for $C_{24}H_{19}N_2O_2$ $[M+H]^+$.

S1.6 Synthesis of 2-amino-N-(2-(6-(naphthalen-2-yl)-1,3-dioxo-1H-benzo[de]isoquinolin-2(3H)-yl)ethyl)acetamide (11):

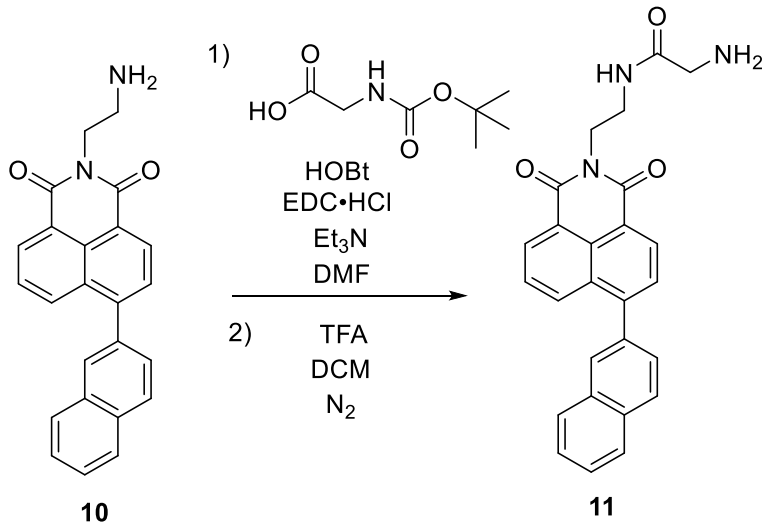


Figure S6. Synthesis of compound **11**.

In a 50 mL round bottom flask, compound **10** (264 mg, 0.72 mmol), and triethylamine (1 mL, 7.2 mmol) was dissolved in DMF (15 mL) under ice bath condition. After stirring for 5 minutes, BOC-Glycine (270 mg, 1.56 mmol) and HOBt (392 mg, 2.9 mmol) were added. After stirring for 10 minutes under ice bath condition, 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (560 mg, 2.9 mmol) was added. The reaction mixture was allowed to stir at room temperature for overnight. The reaction was monitored using thin layer chromatography, upon completion, 200 mL of water and 50 mL of dichloromethane were added. The organic phase was separated and concentrated under vacuum to obtain a yellow solid crude product. This crude product was then dissolved in DCM (5 mL) and TFA (1.3 mL) was added under ice bath condition. The reaction mixture was allowed to slowly warm up to room temperature and monitored by TLC until the end of the reaction. The solution was concentrated under vacuum and the product was purified using column chromatography (1:0-10:1, DCM/MeOH) to afford compound **11** (150 mg, 50 %) as a yellow solid. ^{1}H NMR (400 MHz, DMSO-d_6): δ *(ppm)* 8.55 (dd, J = 11.1, 7.5 Hz, 2H), 8.29 (d, J = 8.5 Hz, 1H), 8.13 (d, J = 8.4 Hz, 2H), 8.08 – 8.03 (m, 2H), 7.90 (d, J = 7.5 Hz, 1H), 7.84 (t, J = 7.4 Hz, 1H), 7.69 (d, J = 8.5 Hz, 1H), 7.65 – 7.60 (m, 2H), 4.19 (t, J = 5.8 Hz, 2H), 3.47 (q, J = 5.9 Hz, 2H), 3.05 (s, 2H). ^{13}C NMR (100 MHz, DMSO-d_6): δ *(ppm)* 163.81, 163.60, 145.94, 135.84, 132.92, 132.62, 132.28, 130.78, 130.36, 129.46, 129.02, 128.42, 128.37, 128.33, 128.17, 127.79, 127.68, 127.00, 126.91, 122.73, 55.03, 44.14, 36.59. HRMS m/z Found: 424.1659, calculated: 424.1661 for $C_{26}H_{22}N_3O_3$ $[M+H]^+$.

S2. Analysis of azidation of glycine derivatives

S2.1 The process and analysis of the azidation of glycine benzyl ester

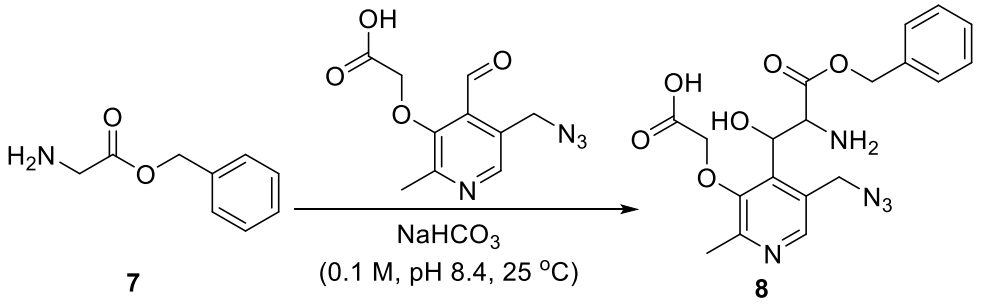


Figure S7. The azidation of **glycine benzyl ester**.

For glycine benzyl ester, in a 2 mL tube, 403 μL of 10 mg/mL benzyl glycinate hydrochloride solution was added in sodium bicarbonate buffer (1.6 mL, 0.1 M, pH 8.4). To this solution, PLAA (20 mg, 0.08 mmol) was added, and the reaction mixture was allowed to stir at 25 ℃. After 24 hours, the reaction mixture was first analyzed by analytical HPLC and LC-MS, and two diastereomers were isolated via preparative HPLC with a combined separated yield of 45%.

Method of analytical HPLC (Column: Elitehplc SinoChrom ODS-BP 5 μm 4.6×150 mm, flow rate: 1 mL/min). MeCN and H_2O were buffered with 0.1% trifluoroacetic acid. Multi-Step gradient: 0-18 min, 2-20% (MeCN); 18-20 min, 20-42.7% (MeCN); 20-27 min, 42.7%-57.3% (MeCN); 27-30 min, 57.3%-98% (MeCN); 30-31min, 98% (MeCN); 31-32 min, 98%-2% (MeCN); 32-35 min, 98% (MeCN). Detection wavelength: 254 nm. Injection volume: 20 μL.

Method of LC-MS (Column: Thermo Acclaim^{TM} 120 C18 3 μm 3×150 mm, flow rate: 0.5 mL/min). MeCN and H_2O were buffered with 0.1% formic acid. Multi-Step gradient: 0-4 min, 5%-95% (MeCN); 4-6 min, 95% (MeCN); 6-6.5 min, 95%-5% (MeCN); 6.5-10 min, 5% (MeCN). Detection wavelength: 254 nm. Injection volume: 5 μL. Parameter setting of Mass: 0-10 min; Mass range, 100-1000; Ionization mode, ESI; Pro temperature, 350 ℃.

Method of preparative HPLC (Column: Elitehplc SinoChrom ODS-BP 5 μm 20×250 mm, flow rate 10 mL/min). MeCN and H2O were buffered with 0.1% trifluoroacetic acid. Multi-Step gradient: 0-18 min, 5-20% (MeCN); 18-24 min, 20-57.3% (MeCN); 24-28 min,

57.3%-59% (MeCN); 28-30 min, 59%-95% (MeCN); 30-32 min, 95% (MeCN); 32-34 min, 95%-5% (MeCN); 34-35 min, 95% (MeCN). Detection wavelength: 254 nm. Injection volume: 5 mL.

For one of the two diastereomers (**8-1**): ^{1}H NMR (400 MHz, CD_3CN): *δ (ppm)* 8.25 (s, 1H), 7.34-7.30(m, 3H), 7.17 – 7.11 (m, 2H), 5.50 (d, *J* = 7.0 Hz, 1H), 5.08 (s, 2H), 4.87 (d, *J* = 7.1 Hz, 1H), 4.72 (d, *J* = 15.4 Hz, 1H), 4.50 (d, *J* = 4.1 Hz, 2H), 4.44 (d, *J* = 15.4 Hz, 1H), 2.49 (s, 3H). ^{13}C NMR (100 MHz, CD_3CN): *δ (ppm)* 171.83, 167.52, 161.48, 161.13, 153.89, 153.62, 145.36, 142.47, 135.31, 131.40, 129.65, 129.22, 71.41, 69.26, 66.29, 57.65, 49.90, 17.57, 16.72. HRMS m/z Found: 416.1571, calculated: 416.1570 for $C_{19}H_{22}N_5O_6$ $[M+H]^+$.

For the other diastereomer (**8-2**): ^{1}H NMR (400 MHz, CD_3CN): *δ (ppm)* 8.17 (s, 1H), 7.34-7.28 (m, 3H), 7.17 – 7.09 (m, 2H), 5.85 (d, *J* = 3.4 Hz, 1H), 5.05 (s, 2H), 4.75 (dd, *J* = 15.1, 6.3 Hz, 2H), 4.62 (d, *J* = 3.4 Hz, 1H), 4.51 (d, *J* = 15.5 Hz, 1H), 4.36 (d, *J* = 14.8 Hz, 1H), 2.55 (s, 3H). ^{13}C NMR (100 MHz, CD_3CN): *δ (ppm)* 172.11, 167.35, 161.73, 161.36, 152.74, 150.24, 146.70, 141.30, 135.23, 133.16, 129.74, 129.48, 129.46, 71.17, 69.32, 69.08, 58.65, 50.11, 17.22. HRMS m/z Found: 416.1569, calculated: 416.1570 for $C_{19}H_{22}N_5O_6$ $[M+H]^+$.

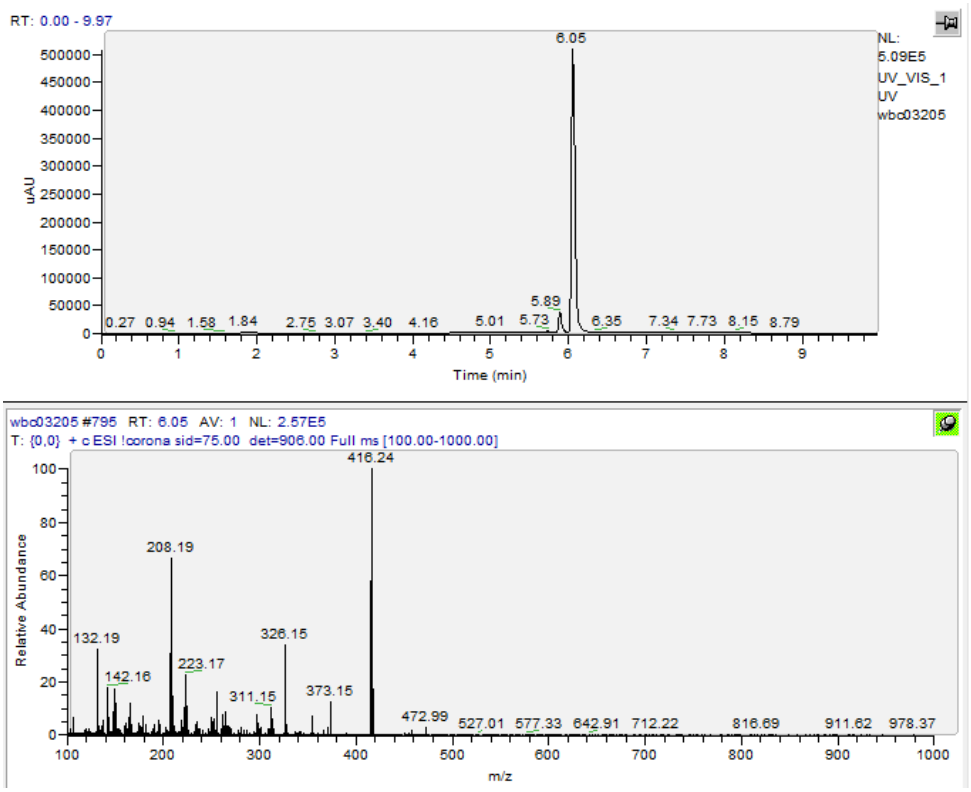


Figure S8. LC-MS analysis of **8-1**.

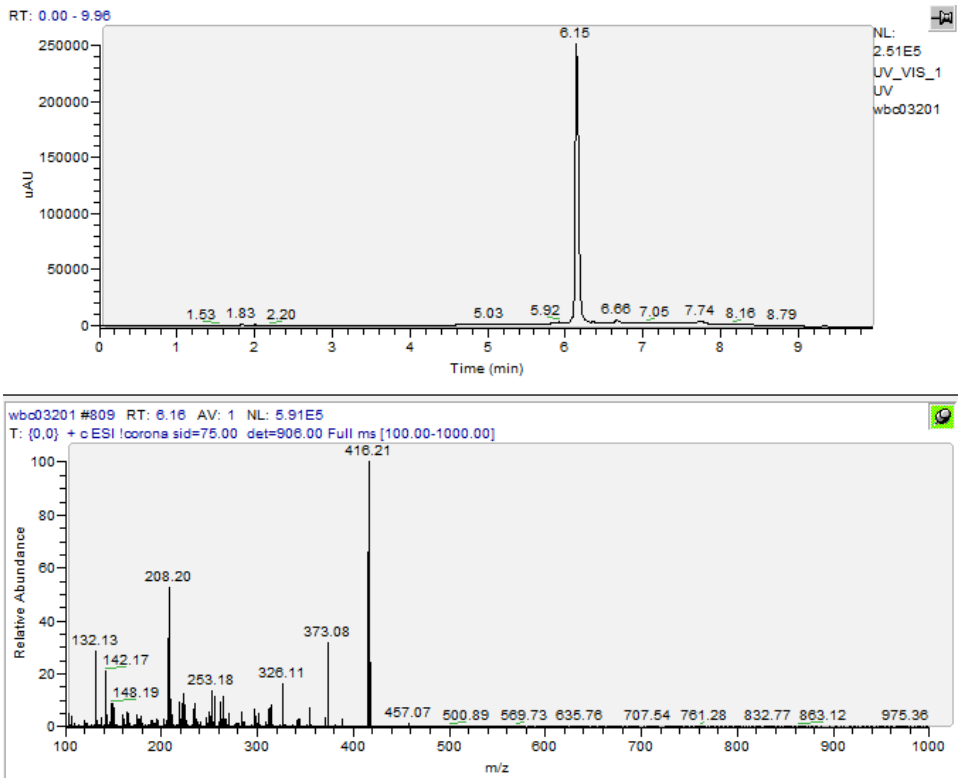
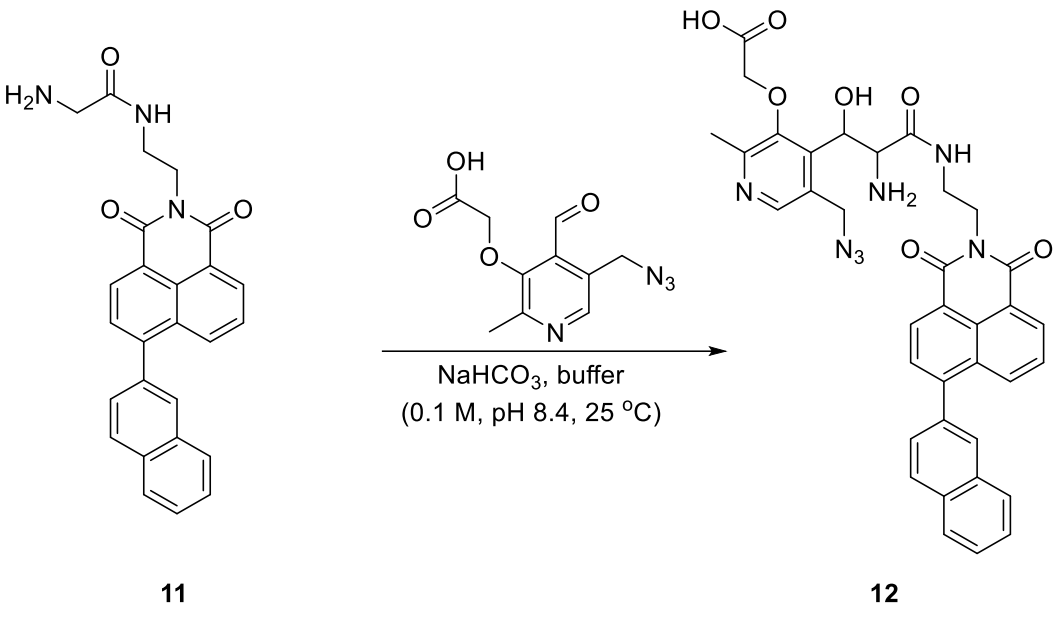


Figure S9. LC-MS analysis of **8-2**.

S2.2 The process and analysis of the azidation of compound 11



11 **12**

Figure S10. The azidation of compound **11** with **PLAA**.

For compound **11**, in a 2 mL tube, 10 μL of 74.28 mg/mL PLAA solution was added in 300 μL of 0.52 mg/mL compound **11** in sodium bicarbonate buffer (0.1 M, pH 8.4)solution (23% DMSO (v/v) in water), and the reaction mixture was allowed to stir at 25 ℃. After rapid mixing of reactants, the initial reaction mixture was analyzed by

analytical HPLC, monitored at 368 nm. The reaction results were monitored every hour for the following 9 hours to monitor the reaction process. The final results of the reaction were analyzed by LC-MS.

Method of analytical HPLC (Column: Elitehplc SinoChrom ODS-BP 5 μm 4.6×200 mm, flow rate: 1 mL/min). MeCN and H_2O were buffered with 0.1% formic acid. Multi-Step gradient: 0-5 min, 20-60% (MeCN); 5-14 min, 60% (MeCN); 14-15 min, 60%-20% (MeCN). Detection wavelength: 368 nm. Injection volume: 20 μL.

Method of LC-MS (Column: Thermo Acclaim™ 120 C18 3 μm 3×150 mm, flow rate: 0.5 mL/min). MeCN and H_2O were buffered with 0.1% formic acid. Multi-Step gradient: 0-7 min, 5%-95% (MeCN); 7-12 min, 95% (MeCN); 12-13.5 min, 95%-5% (MeCN); 13.5-20 min, 5% (MeCN). Detection wavelength: 254 nm. Injection volume: 5 μL. Parameter setting of Mass: 0-20 min; Mass range, 100-1000; Ionization mode, ESI; Pro temperature, 350 ℃.

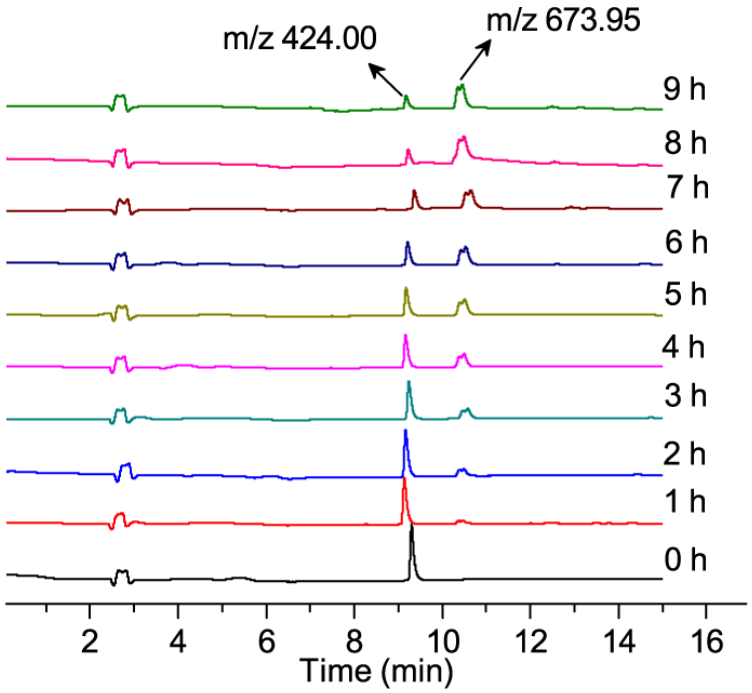


Figure S11. The reaction process of compound **11** with **PLAA**, analyzed by HPLC.

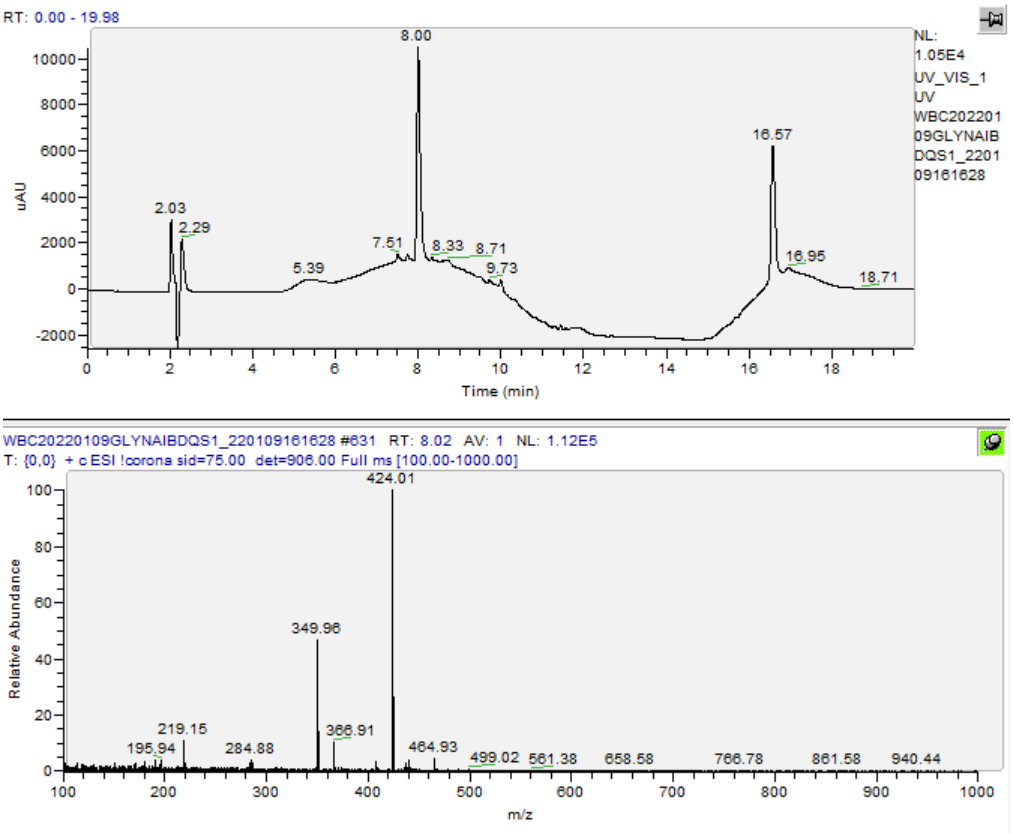


Figure S12. LC-MS analysis of compound **11**.

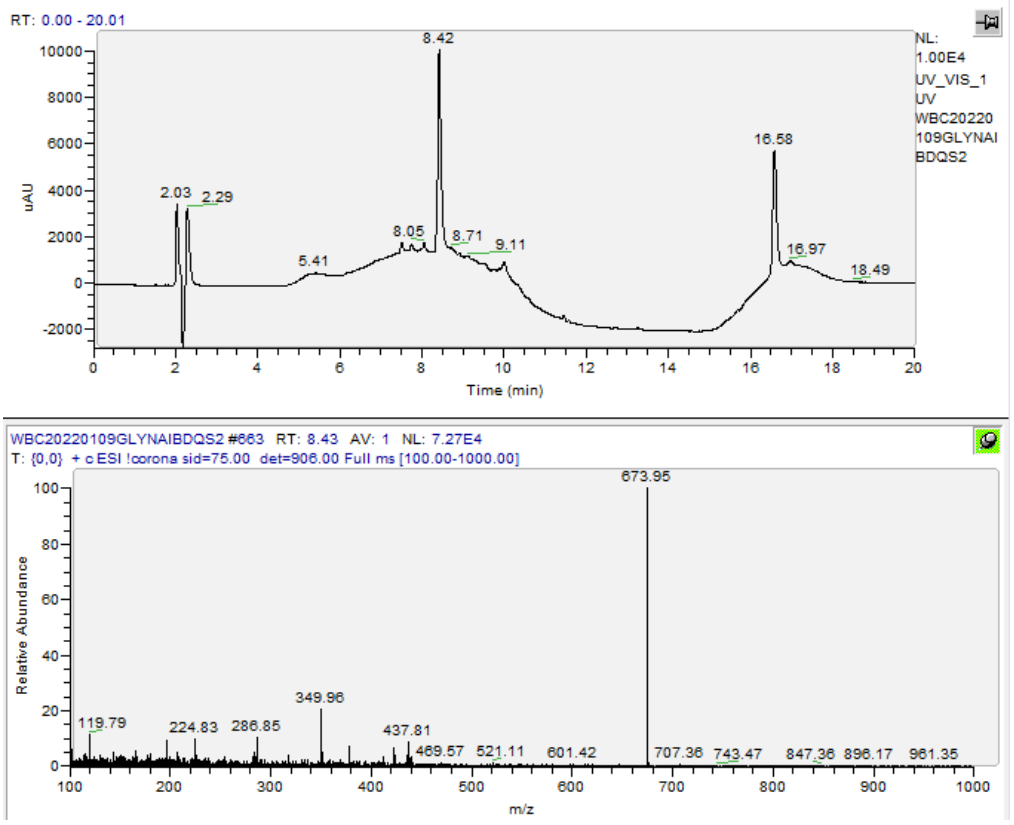


Figure S13. LC-MS analysis of compound **12**.

S3. Site-specific azidation of synthesized peptides

S3.1 Prepration of peptide library

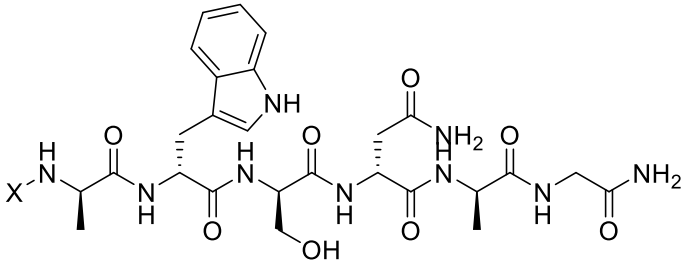


Figure S14. Sequences of the target peptide library.

Model peptides were synthesized through conventional Fmoc solid-phase chemistry on Rink Amide MBHA resin. Resin was swollen before reaction. Coupling reactions were carried out using 5 equivalents of amino acids, 5 equivalents of HOBt, 5 equivalents of HBTU and 10 equivalents of DIPEA in DMF for 2-3 h. Deprotection of Fmoc groups was accomplished by incubation with a 20% v/v piperidine in DMF solution for 20 minutes. Cleavage of peptides from resin and side-chain deprotection were achieved by incubation with a mixture of trifluoroacetic acid (TFA) to H_2O to triisopropylsilane (TIPS) (95:2.5:2.5, v/v/v). The cleavage solution was concentrated under reduced pressure, and anhydrous ether was added to precipitate the crude peptide. Then the purified peptide was used for PLAA modification which separated by preparative HPLC.LC-MS

Method of preparative HPLC (Column: Elitehplc SinoChrom ODS-BP 5 μm 20×250 mm, flow rate: 10 mL/min). MeCN and H_2O were buffered with 0.1% formic acid. Multi-Step gradient: 0-25 min, 5-25% (MeCN); 25-27 min, 25% (MeCN); 27-28 min, 25%-5% (MeCN); 28-30 min, 5% (MeCN). Detection wavelength: 254 nm. Injection volume: 5 mL.

S3.2 Site-specific azidation of peptide

General procedure for the azidation of peptide: in a 1.5 mL tube, 0.57 μmol of the peptide obtained above was dissolved in $NaHCO_3$ buffer (345 μL, 0.1 M, pH 8.4), and pyridoxal azide (PLAA, 5.7 μmol) was added. After reacting at 25 ℃ for 24 hours, the reaction mixture was desalted with cationic affinity resin and analyzed by LC-MS.

Method of LC-MS (Column: Elitehplc SinoChrom ODS-BP 5 μm 4.6×150 mm, flow rate: 1 mL/min). MeCN and H_2O were buffered with 0.1% formic acid. Multi-Step gradient: 0-20 min, 5%-15% (MeCN); 20-30 min, 15%-95% (MeCN); 30-32 min, 95% (MeCN); 33-35 min, 5% (MeCN). Detection wavelength: 254 nm. Injection volume: 5 μL. Parameter setting of Mass: 0-35 min; Mass range, 100-2000; Ionization mode, ESI; Pro temperature, 350 ℃.

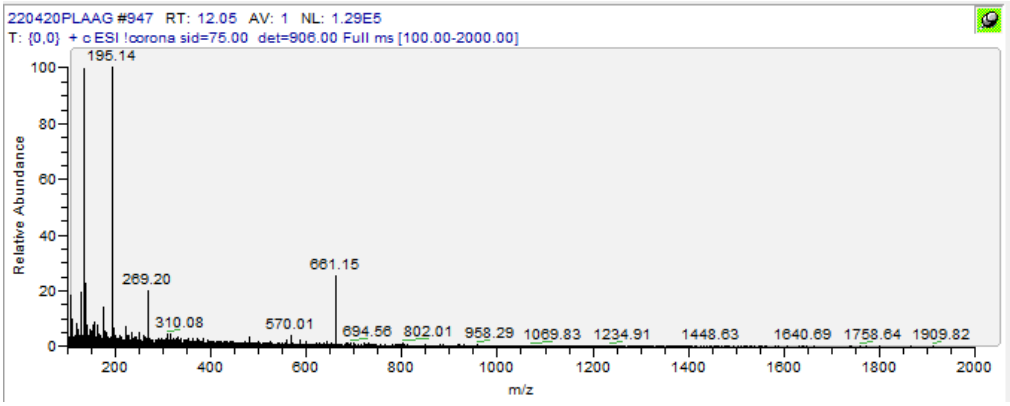


Figure S15. Mass spectrum of N-Gly-peptide.

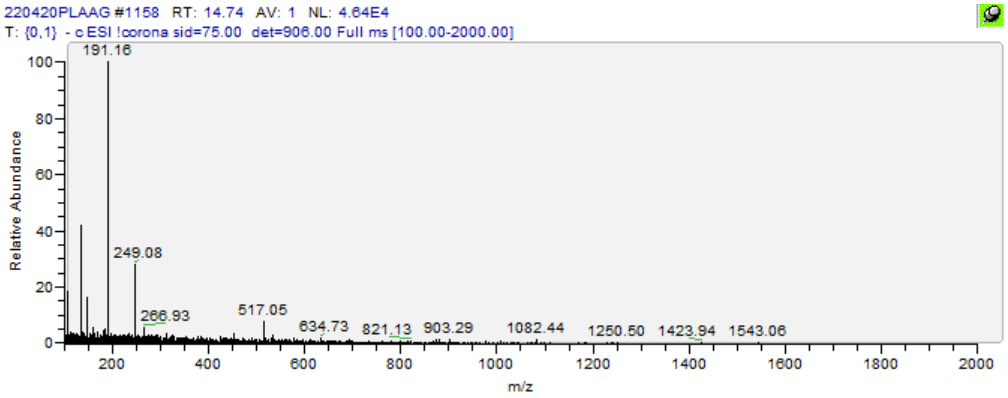
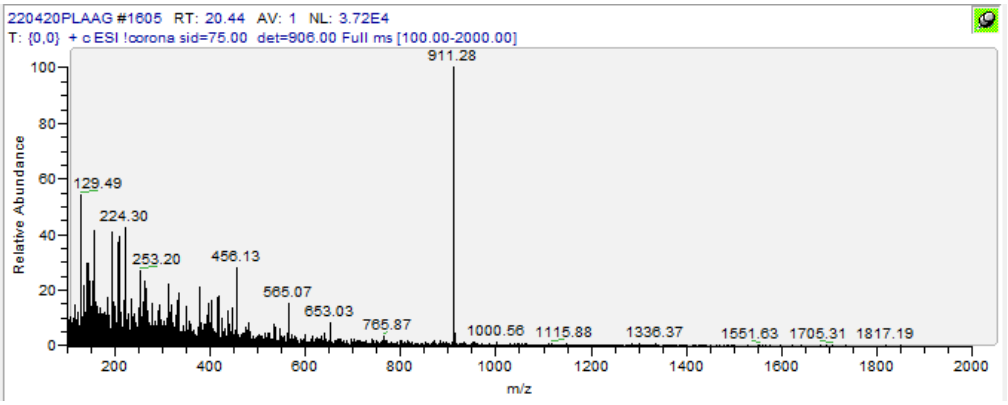


Figure S16. Mass spectrum of **PLAA**.



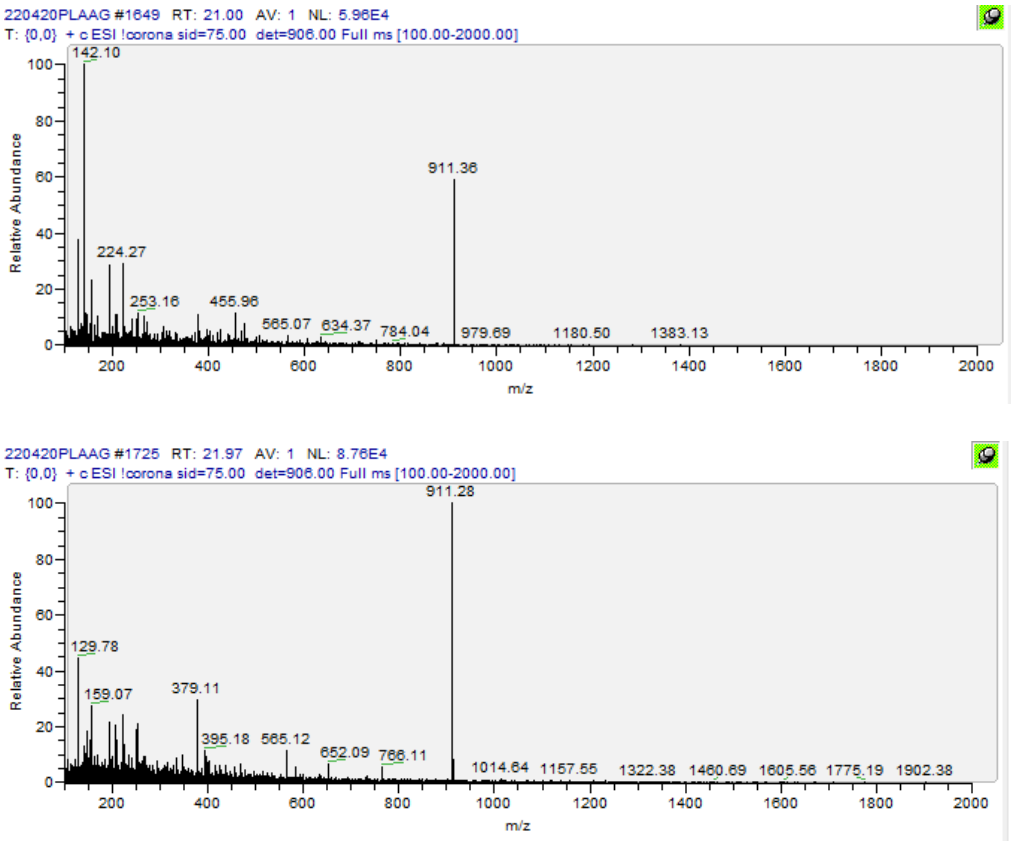


Figure S17. Mass spectrum of N-Gly-peptide-PLAA.

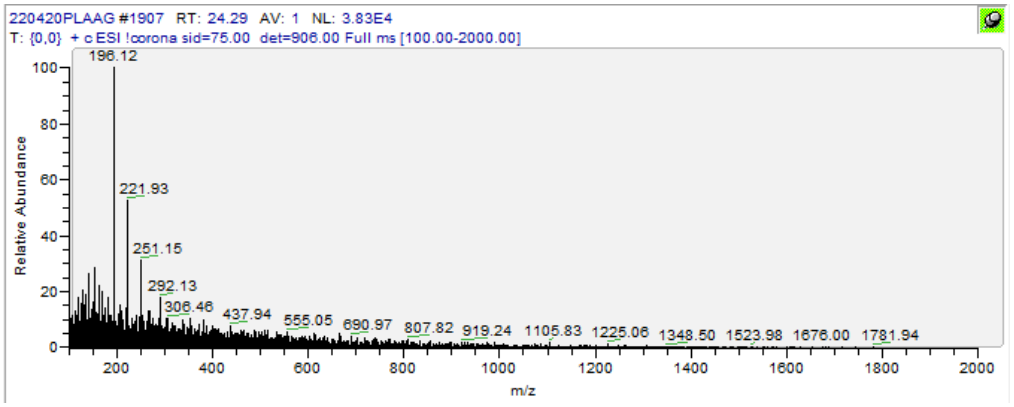


Figure S18. Mass spectrum of **PLAA** possibly from hydrolysis of imine products.

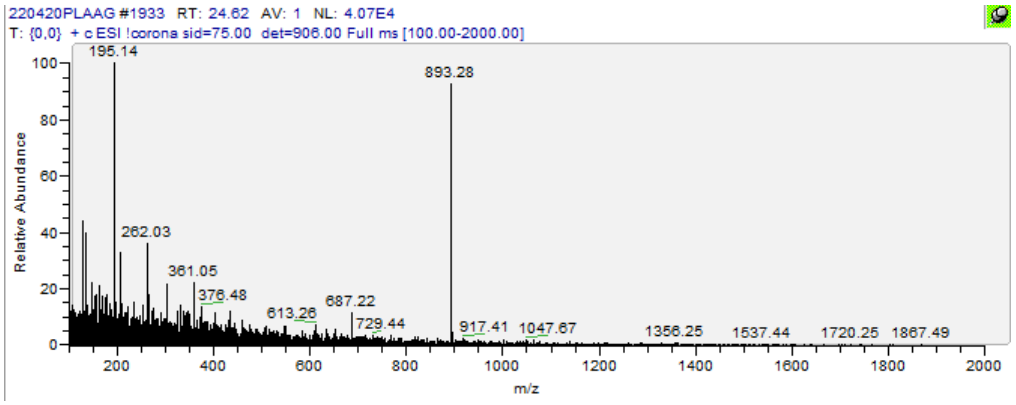


Figure S19. Mass spectrum of imine product of **PLAA** and N-Gly-peptide.

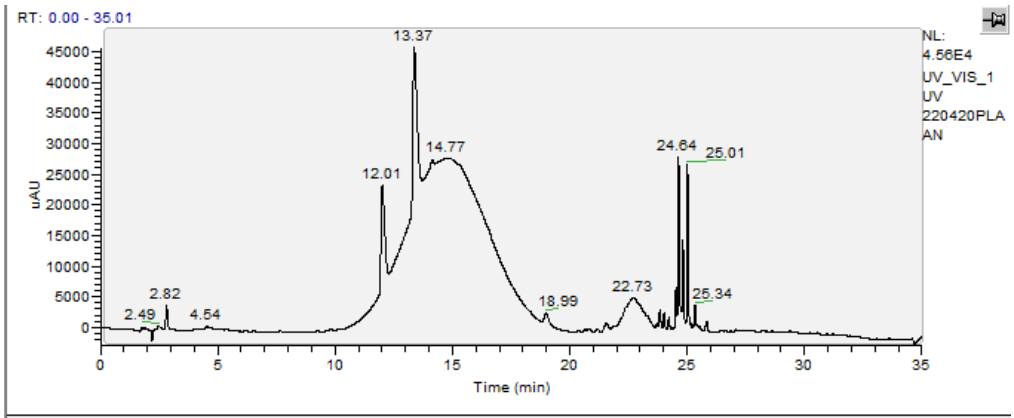
Table 1. The modification of other peptides with PLAA[a]

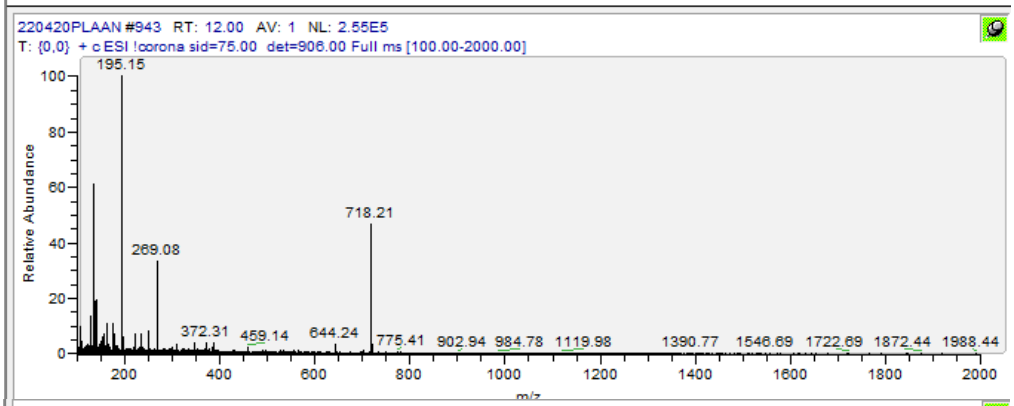
N-terminal of peptide	Peptide	Imine	Transamination	Unidentified byproducts's
Phe	1.00	0.40	0.15	0.07
Ala	1.00	0.72	0.89	0.07
Glu	1.00	0.75	0.13	0.01
His	1.00	1.39	N/D[b]	0.01
Ile	1.00	0.17	N/D	0.04
Lys	1.00	0.47	N/D	0.01
Leu	1.00	0.63	0.29	0.08
Met	1.00	0.90	N/D	0.01
Asn	1.00	0.37	0.70	0.01
Pro	1.00	0.02	N/D	0.01
Gln	1.00	0.55	N/D	0.11
Arg	1.00	0.51	N/D	0.01
Ser	1.00	0.76	N/D	0.01
Val	1.00	0.16	N/D	0.01
Trp	1.00	0.20	N/D	0.06
Tyr	1.00	0.45	0.19	0.01
Asp	1.00	0.83	N/D	0.61
Thr	1.00	0.46	N/D	0.32
Cys[c]	N/D	1.00	N/D	5.66

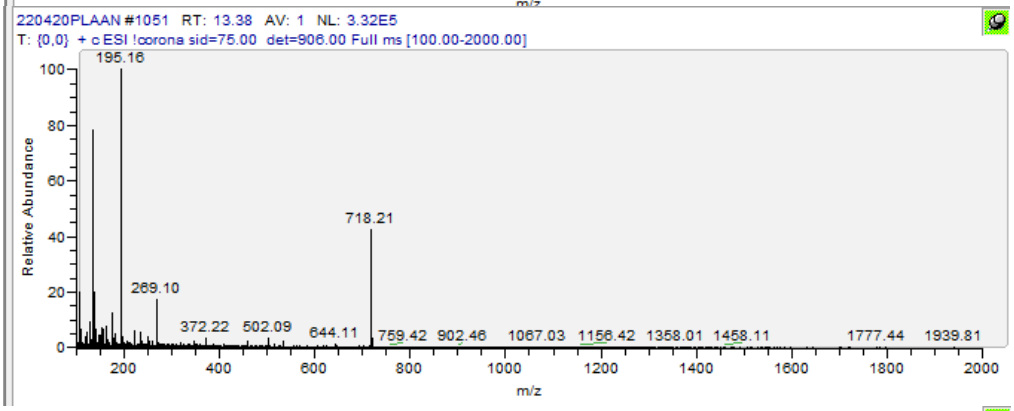
[a] Reactions were monitored by LC-MS, and each peak area was normalized to the peak corresponding the unmodified peptide. For the peptides with peaks overlapped with PLAA peak, the corresponding data were fitted by Origin software to deconvolute the peaks.

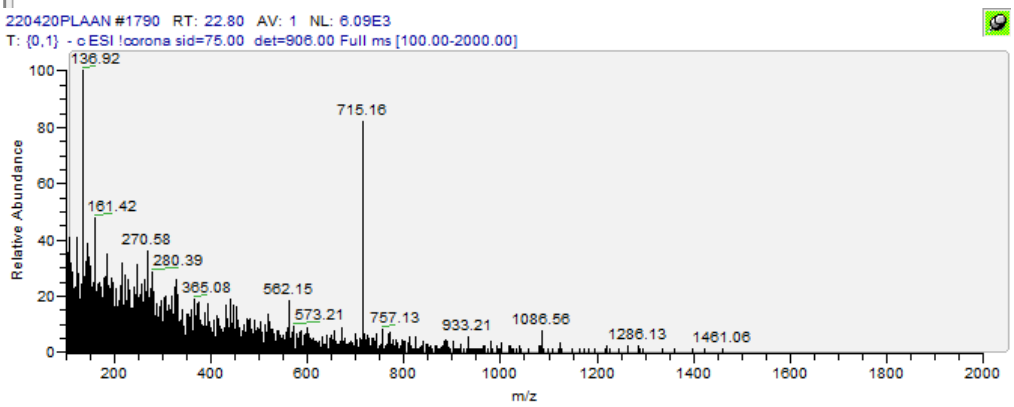
[b] N/D: Not Detected.

[c] No unmodified peptide was observed.


RT: 0.00 - 35.01
NL:
4.56E4
UV_VIS_1
UV
220420PLA
AN
uAU
Time (min)
13.37
14.77
12.01
24.64
25.01
22.73
25.34
18.99
2.82
2.49
4.54


220420PLAAN #943 RT: 12.00 AV: 1 NL: 2.55E5
T: {0,0} + c ESI !corona sid=75.00 det=906.00 Full ms [100.00-2000.00]
Relative Abundance
195.15
718.21
269.08
372.31
459.14
644.24
775.41
902.94
984.78
1119.98
1390.77
1546.69
1722.69
1872.44
1988.44


220420PLAAN #1051 RT: 13.38 AV: 1 NL: 3.32E5
T: {0,0} + c ESI !corona sid=75.00 det=906.00 Full ms [100.00-2000.00]
Relative Abundance
195.16
718.21
269.10
372.22
502.09
644.11
759.42
902.46
1067.03
1156.42
1358.01
1458.11
1777.44
1939.81
m/z


220420PLAAN #1790 RT: 22.80 AV: 1 NL: 6.09E3
T: {0,1} - c ESI !corona sid=75.00 det=906.00 Full ms [100.00-2000.00]
Relative Abundance
136.92
715.16
161.42
270.58
280.39
365.08
562.15
573.21
757.13
933.21
1086.56
1286.13
1461.06
m/z

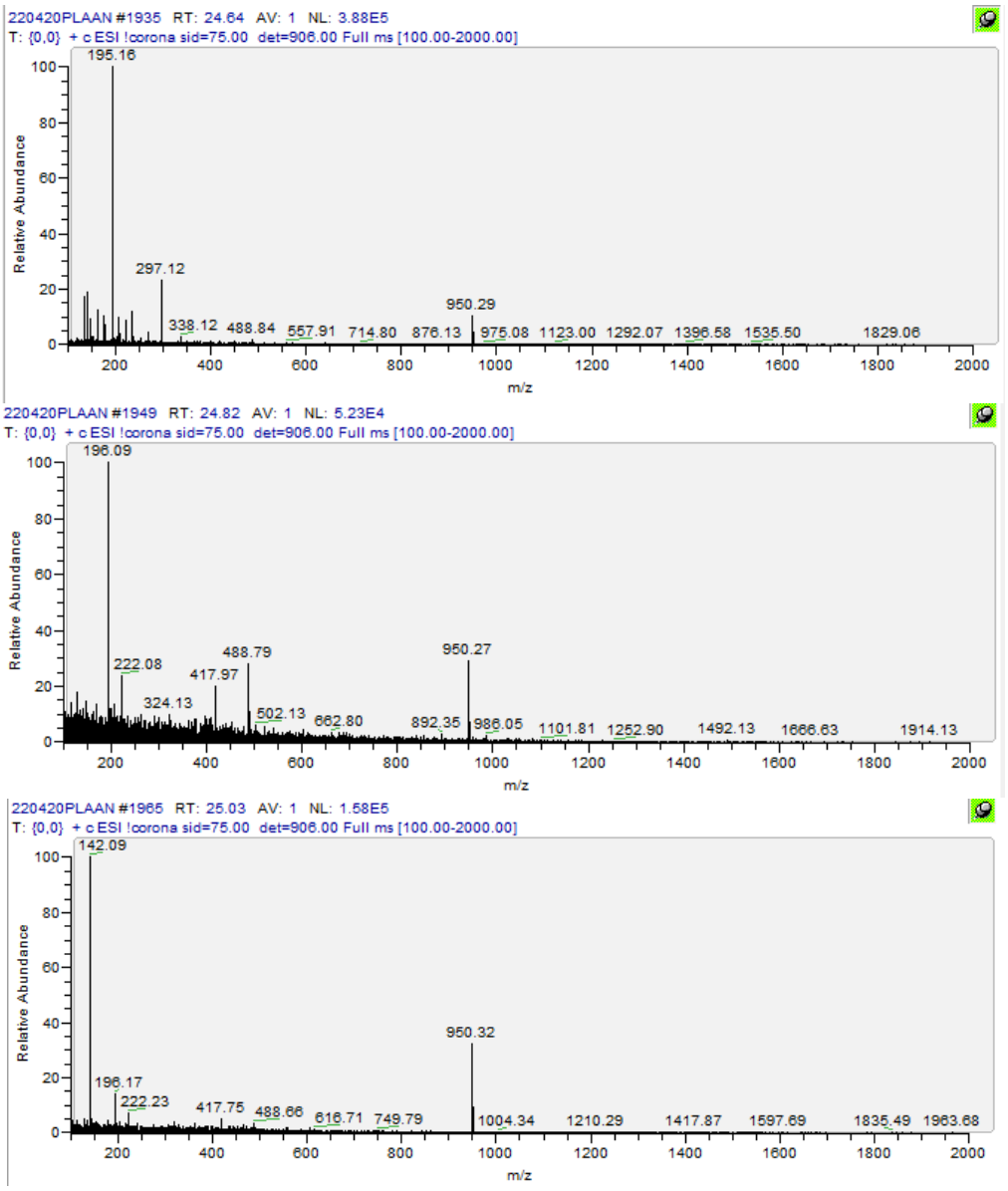


Figure S20. LC-MS analysis of the reaction between N-Asn-peptide and PLAA.

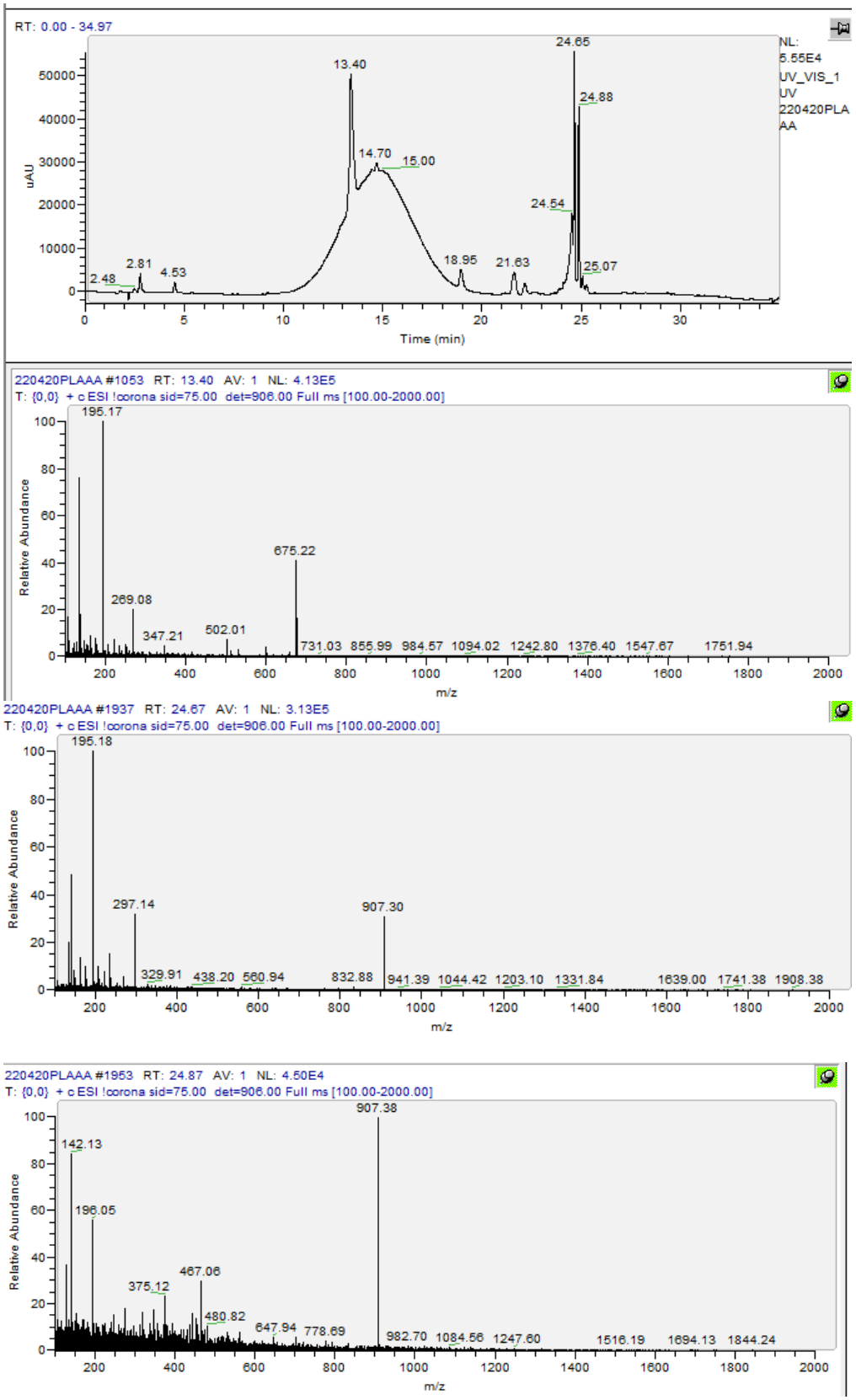


Figure S21. LC-MS analysis of the reaction between N-Ala-peptide and PLAA.

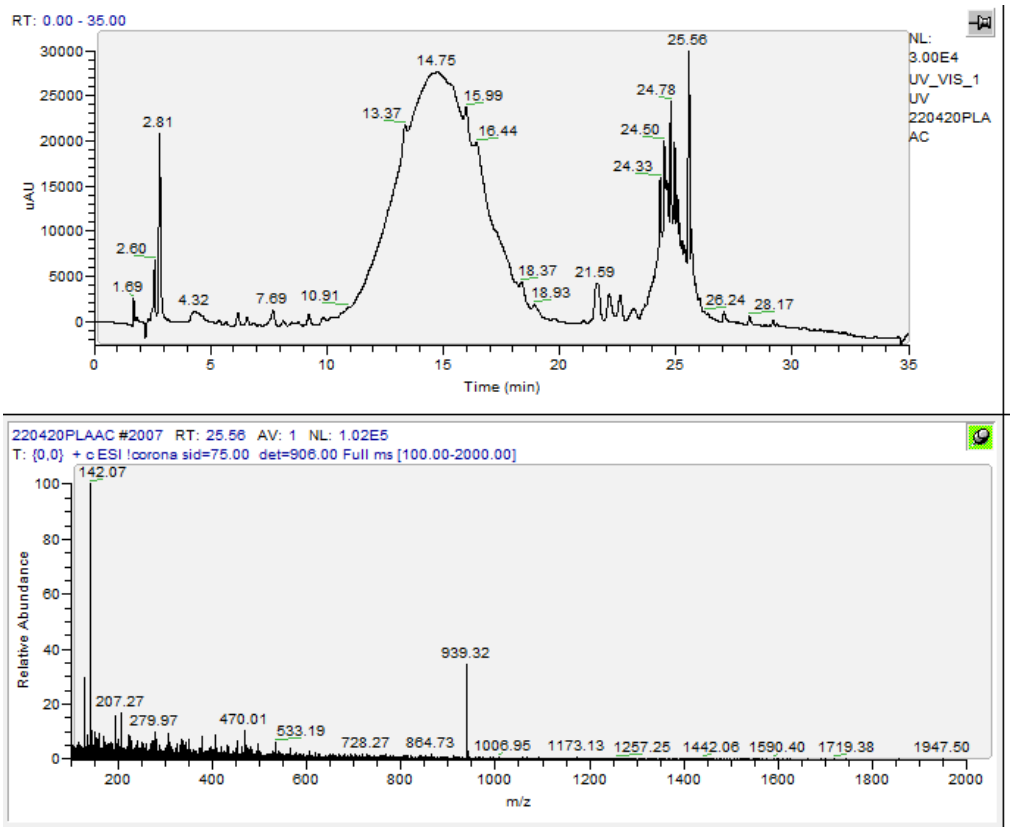
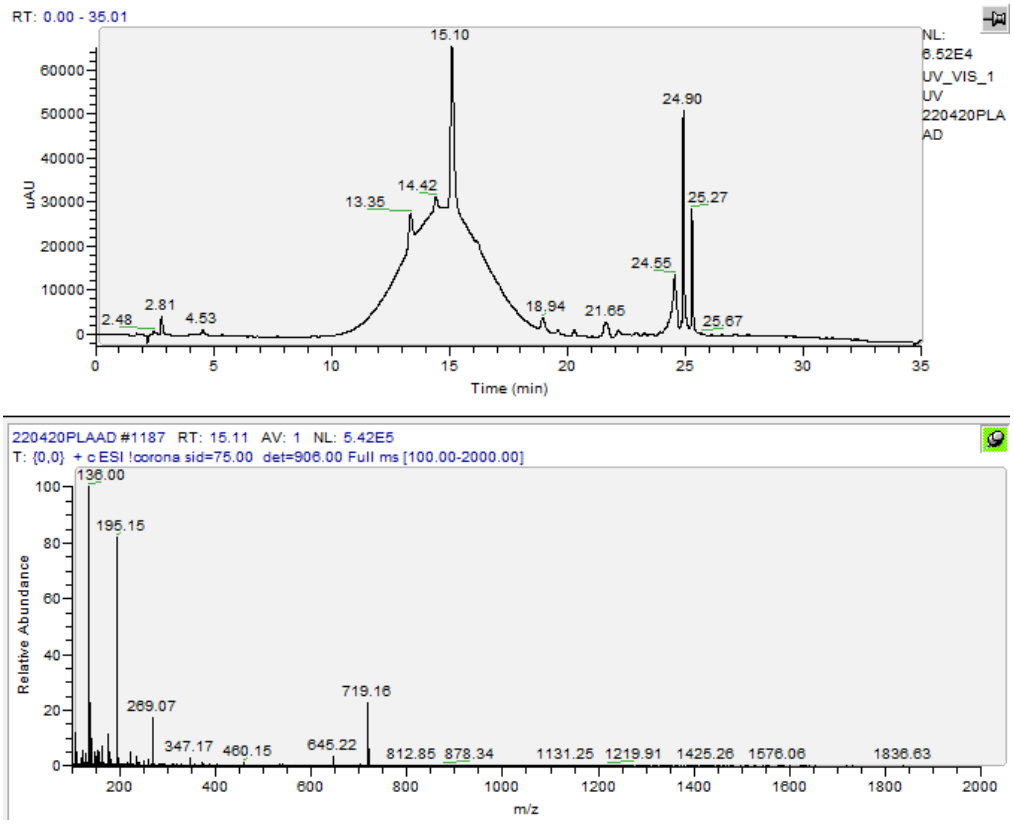


Figure S22. LC-MS analysis of the reaction between N-Cys-peptide and PLAA.



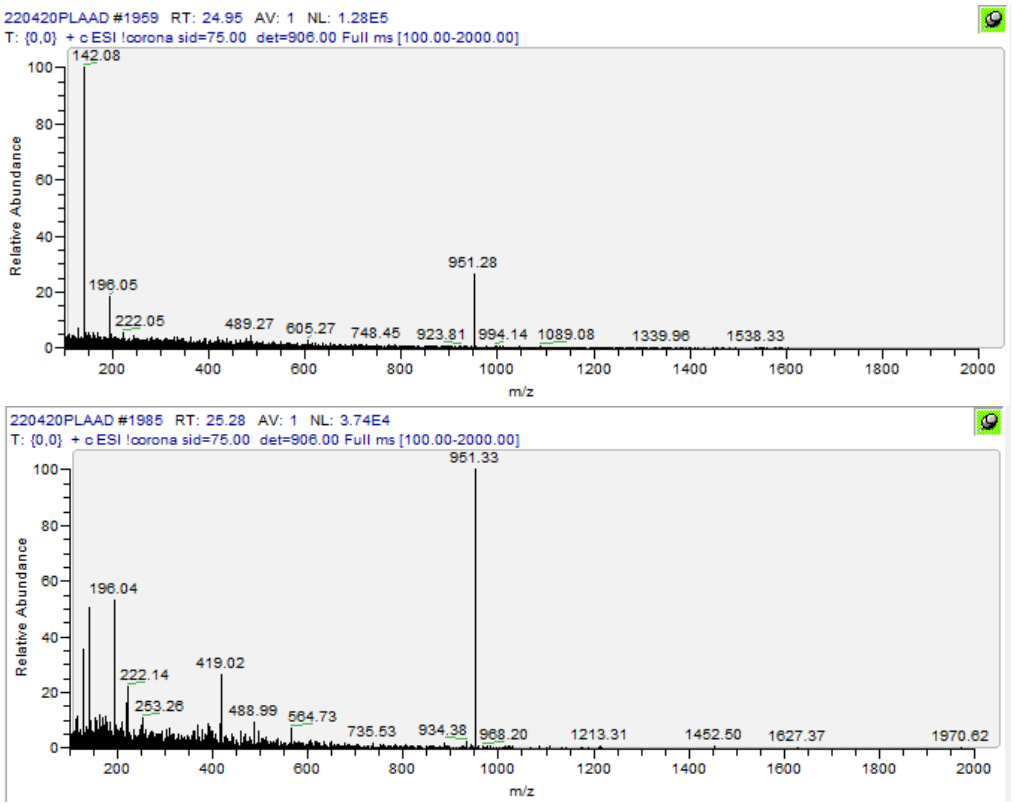
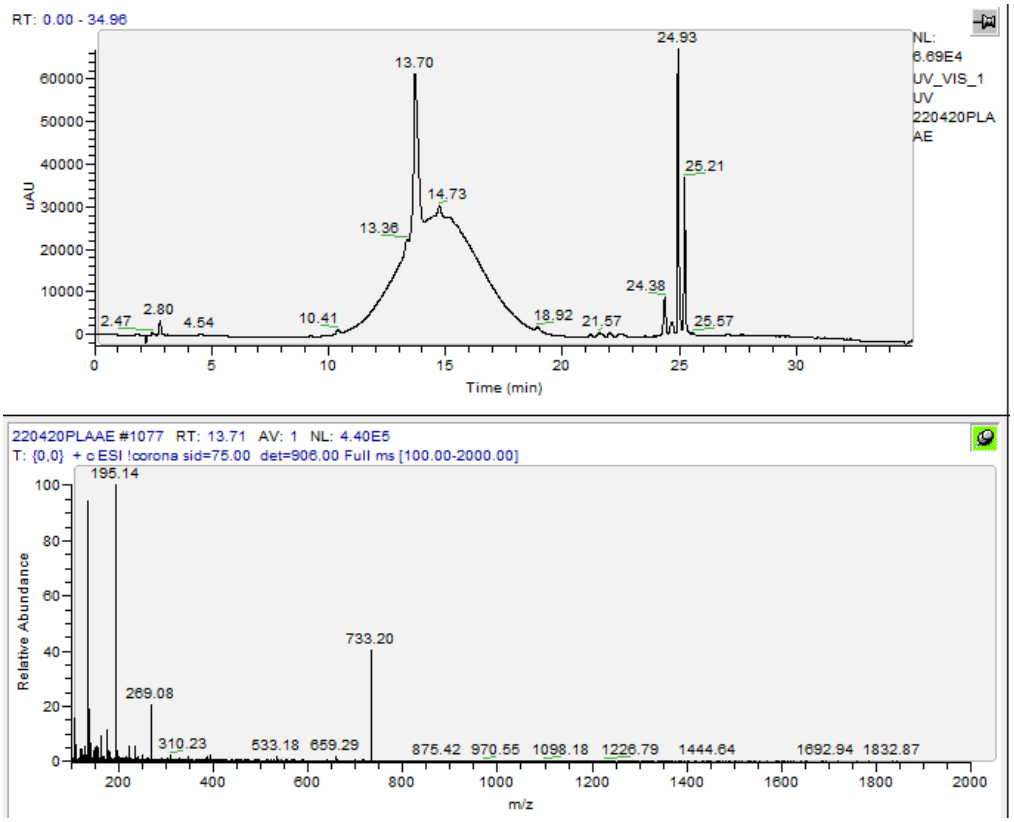


Figure S23. LC-MS analysis of the reaction between N-Asp-peptide and PLAA.



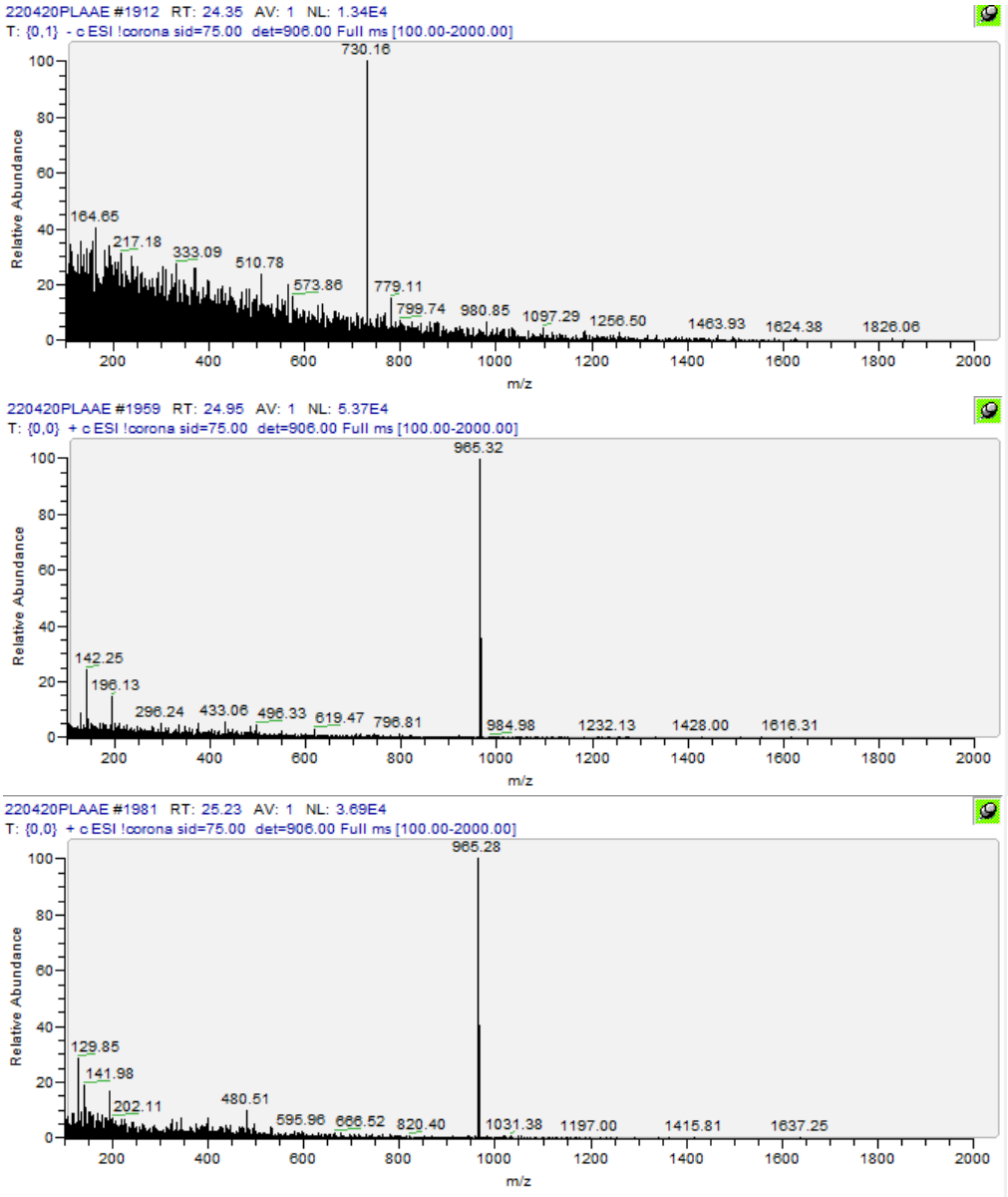


Figure S24. LC-MS analysis of the reaction between N-Glu-peptide and PLAA.

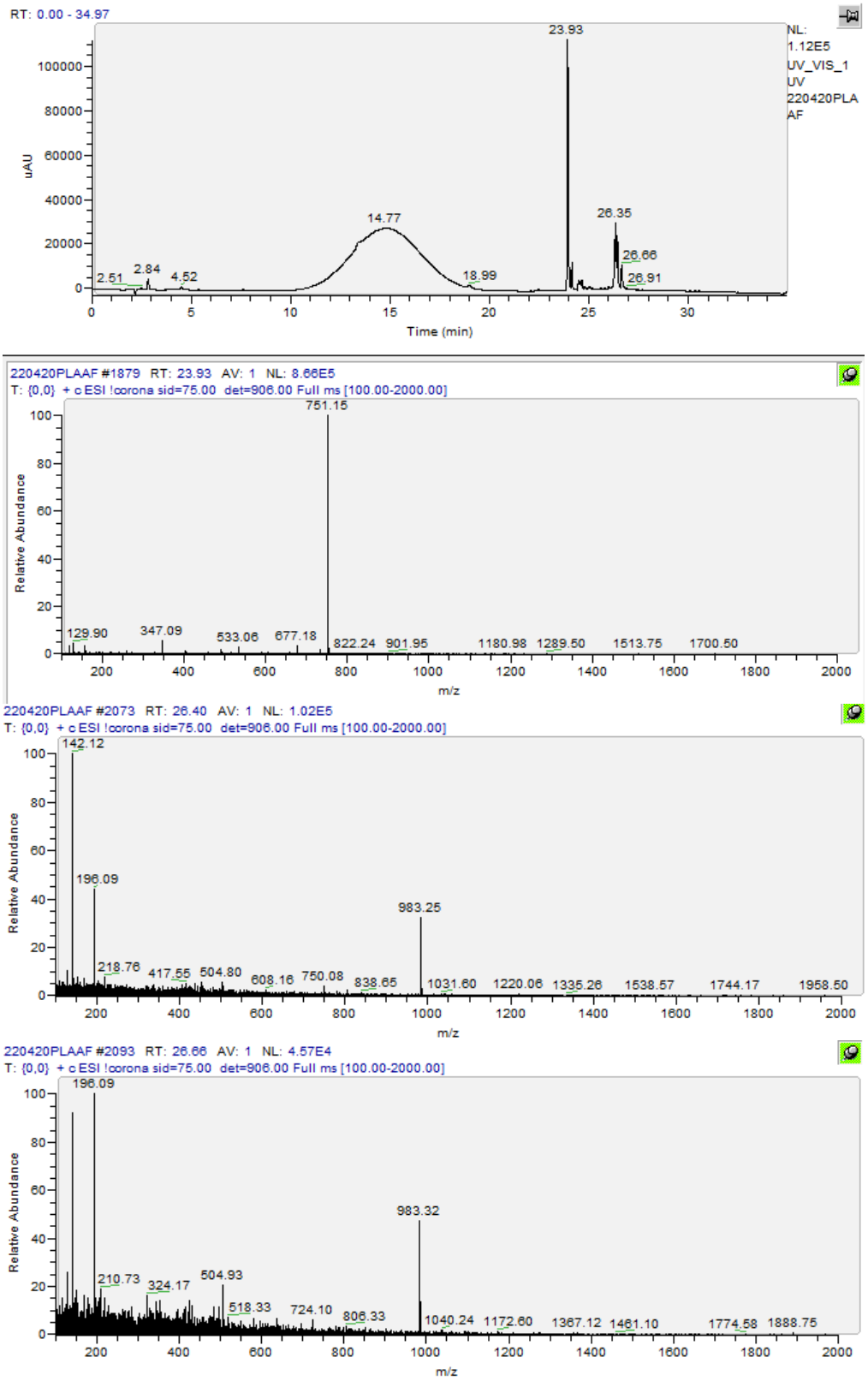
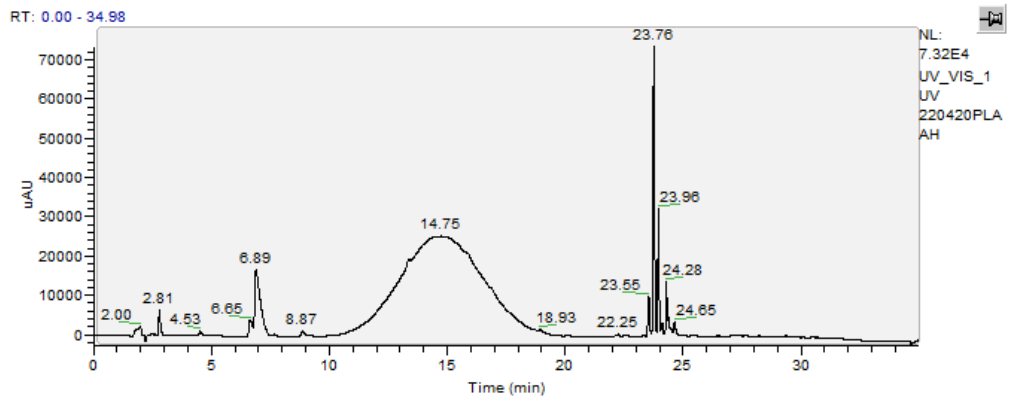
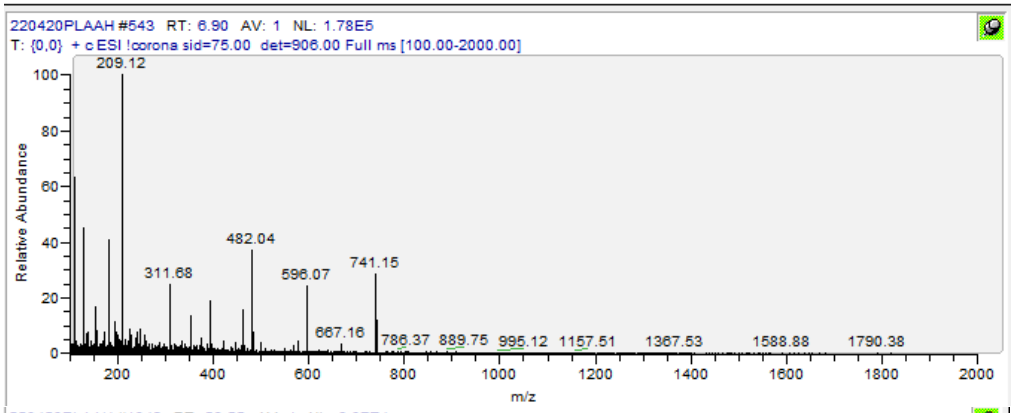
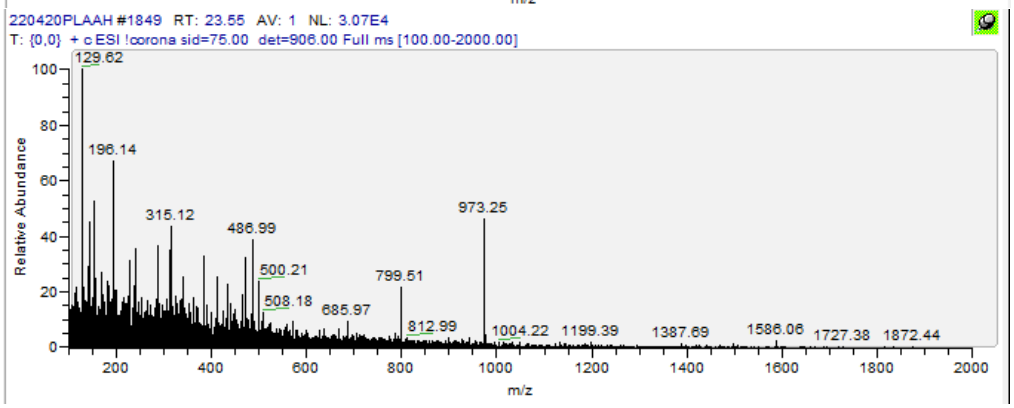
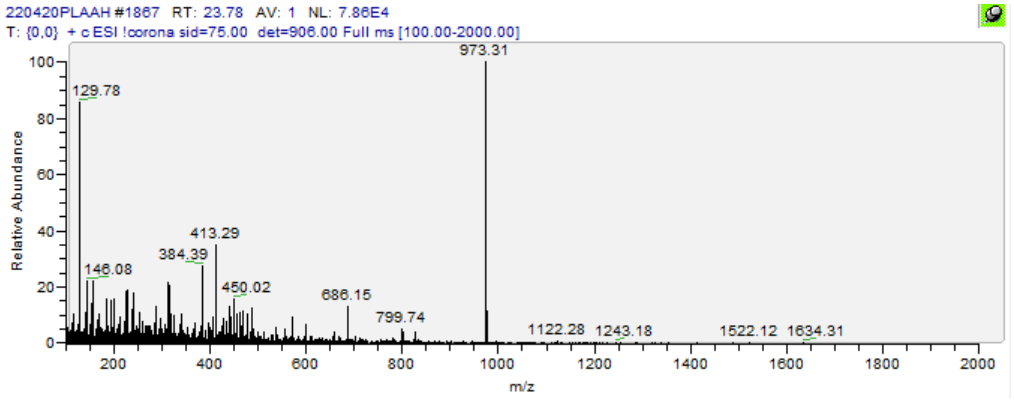


Figure S25. LC-MS analysis of the reaction between N-Phe-peptide and PLAA.


RT: 0.00 - 34.98
NL:
7.32E4
UV_VIS_1
UV
220420PLA
AH
uAU
Time (min)
23.76
14.75
6.89
23.96
24.28
23.55
24.65
2.81
2.00
4.53
6.65
8.87
18.93
22.25


220420PLAAH #543 RT: 6.90 AV: 1 NL: 1.78E5
T: {0,0} + c ESI !corona sid=75.00 det=906.00 Full ms [100.00-2000.00]
Relative Abundance
m/z
209.12
482.04
311.68
596.07
741.15
667.16
786.37
889.75
995.12
1157.51
1367.53
1588.88
1790.38


220420PLAAH #1849 RT: 23.55 AV: 1 NL: 3.07E4
T: {0,0} + c ESI !corona sid=75.00 det=906.00 Full ms [100.00-2000.00]
Relative Abundance
m/z
129.62
196.14
315.12
486.99
973.25
500.21
799.51
508.18
685.97
812.99
1004.22
1199.39
1387.69
1586.06
1727.38
1872.44


220420PLAAH #1867 RT: 23.78 AV: 1 NL: 7.86E4
T: {0,0} + c ESI !corona sid=75.00 det=906.00 Full ms [100.00-2000.00]
Relative Abundance
m/z
973.31
129.78
413.29
384.39
146.08
450.02
686.15
799.74
1122.28
1243.18
1522.12
1634.31

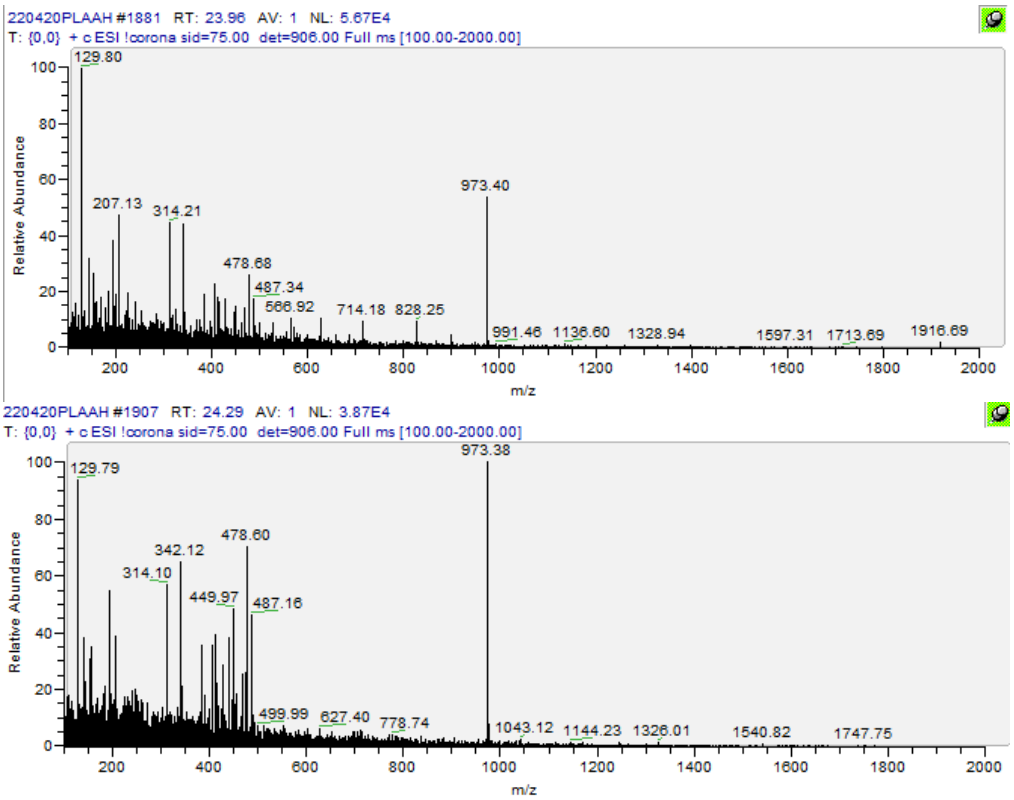
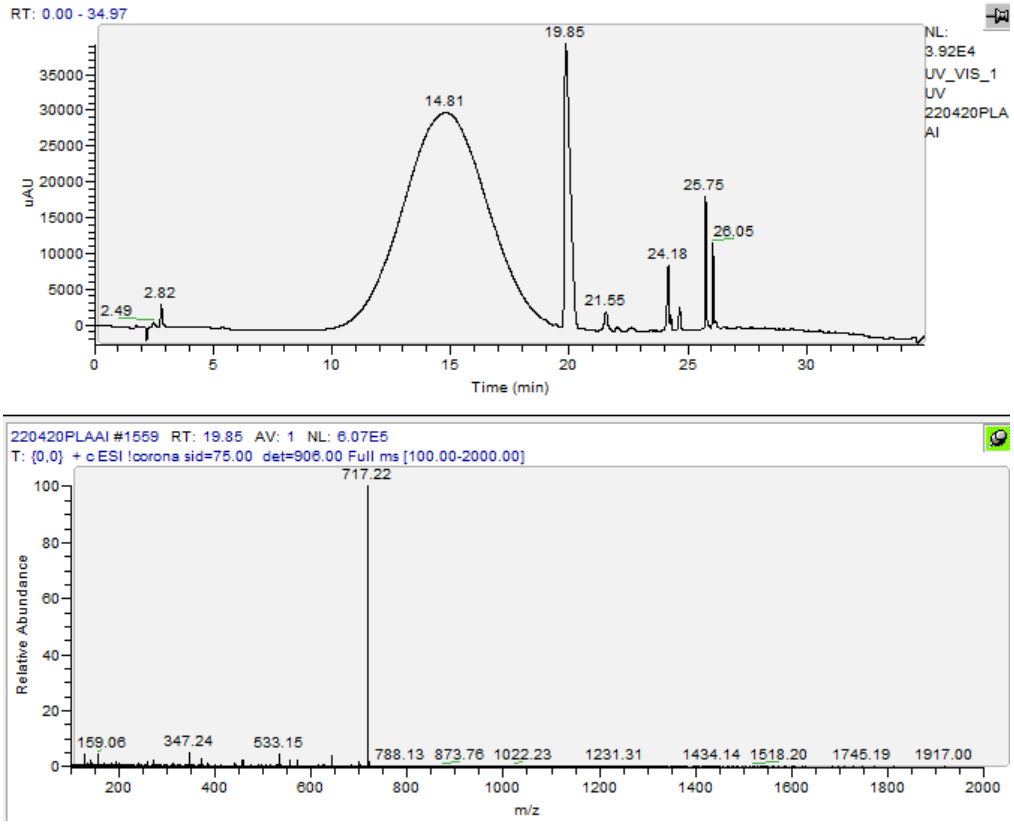


Figure S26. LC-MS analysis of the reaction between N-His-peptide and PLAA.



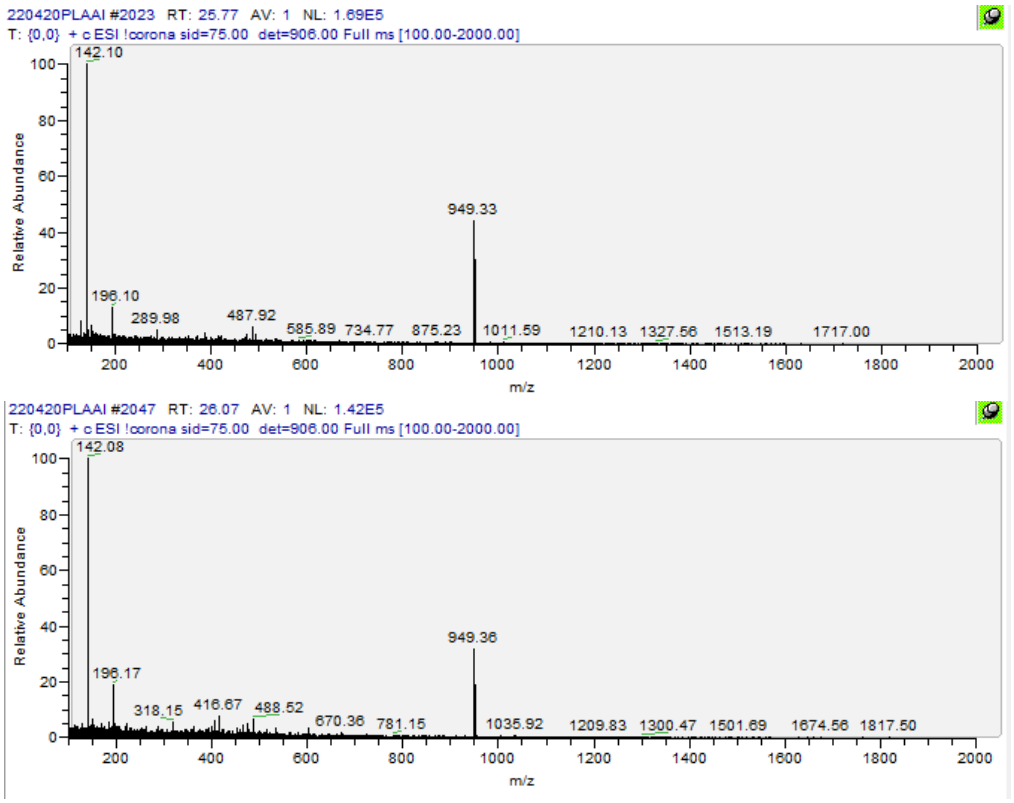
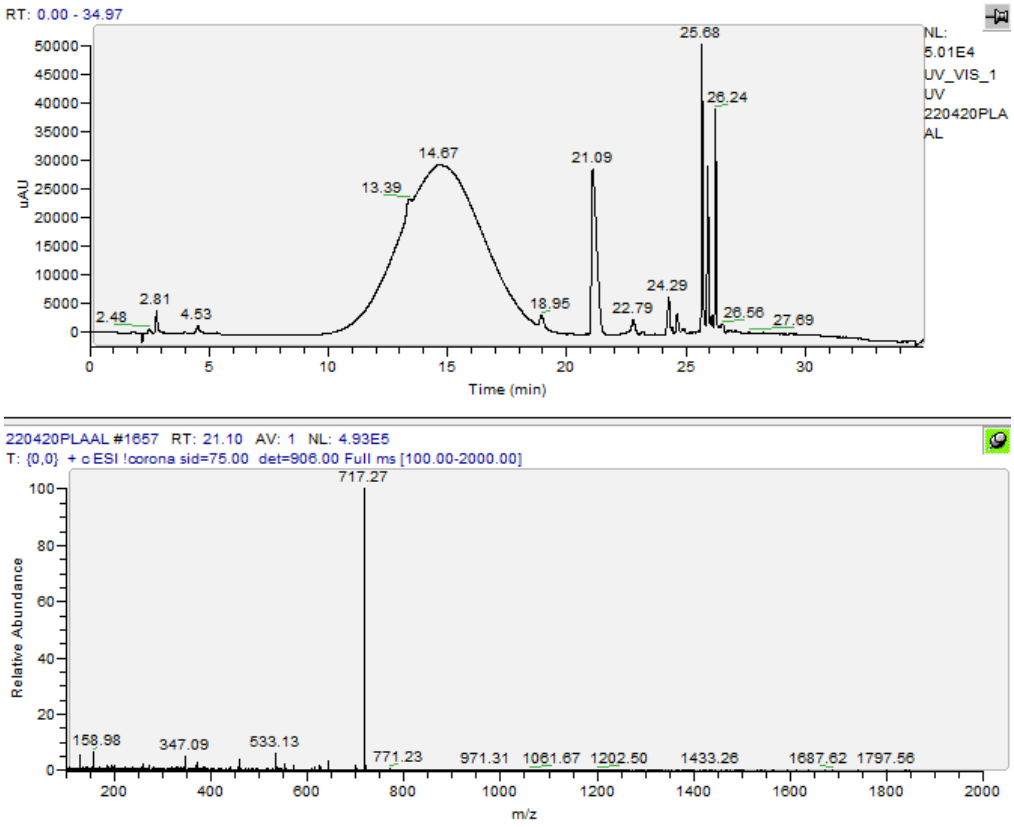


Figure S27. LC-MS analysis of the reaction between N-Ile-peptide and PLAA.



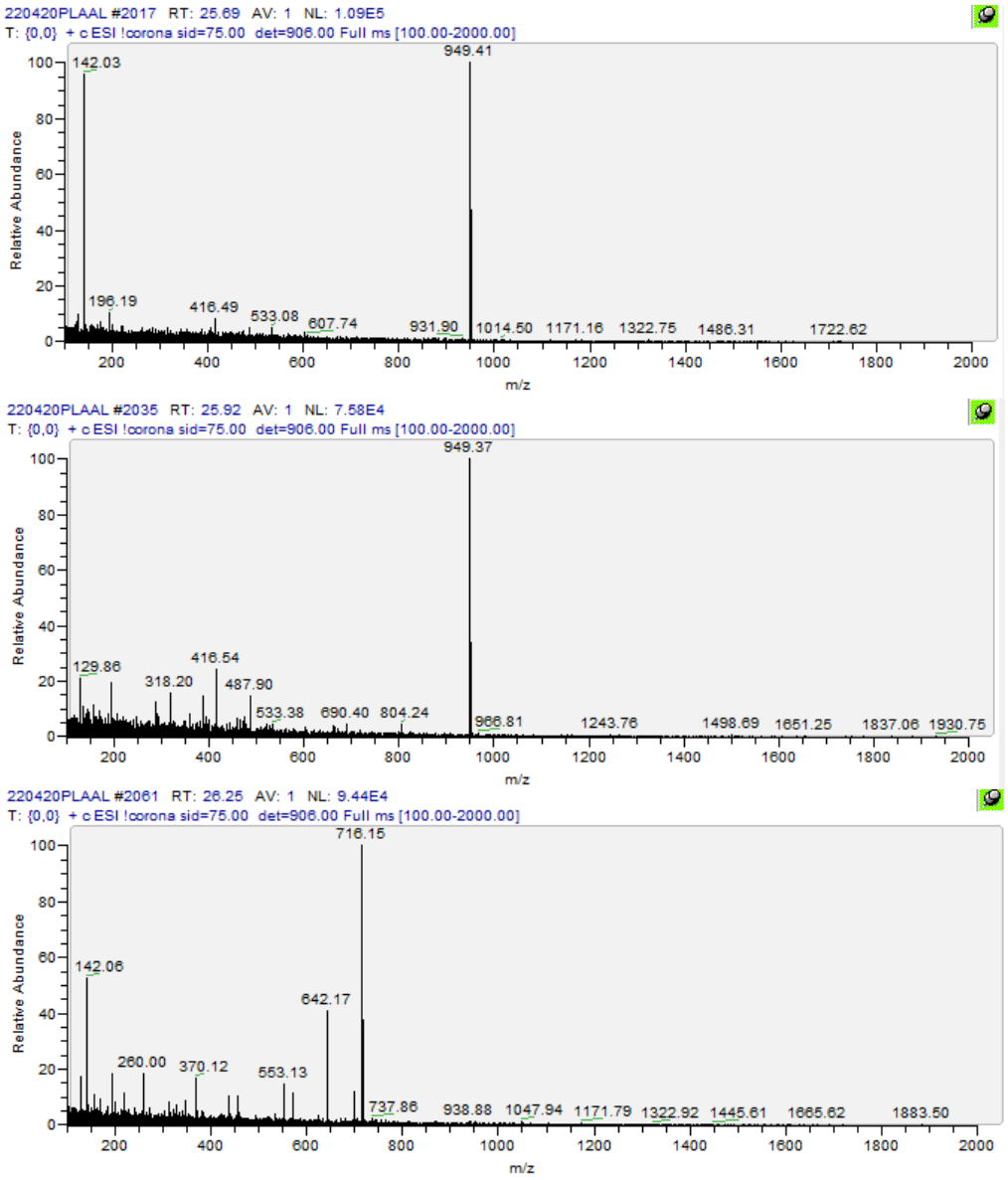


Figure S28. LC-MS analysis of the reaction between N-Leu-peptide and PLAA.

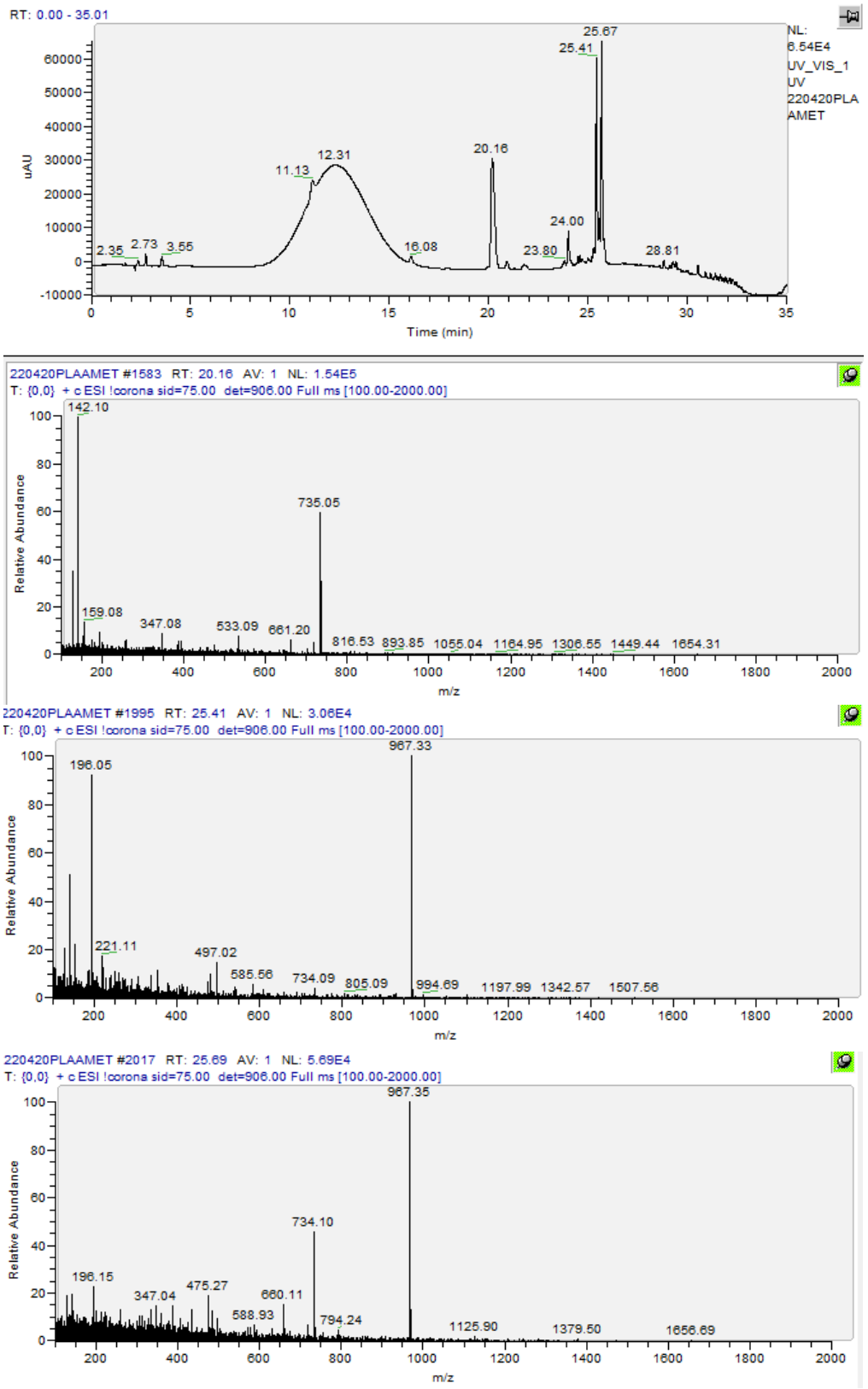


Figure S29. LC-MS analysis of the reaction between N-Met-peptide and PLAA.

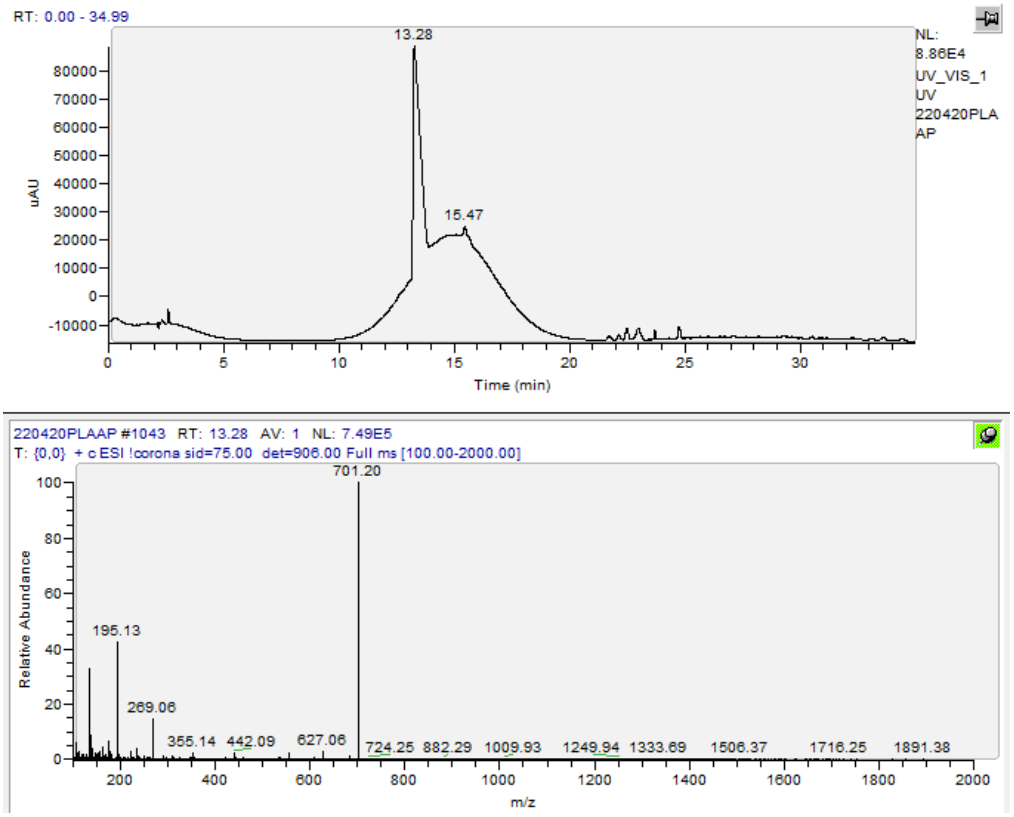
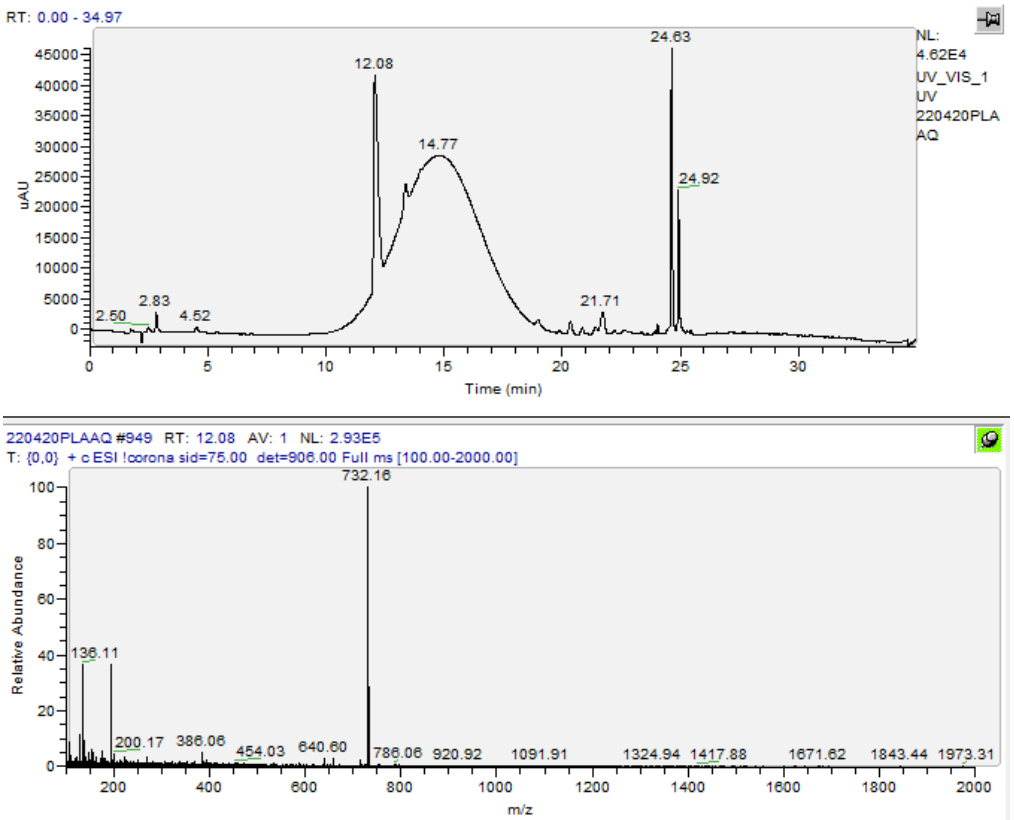


Figure S30. LC-MS analysis of the reaction between N-Pro-peptide and PLAA.



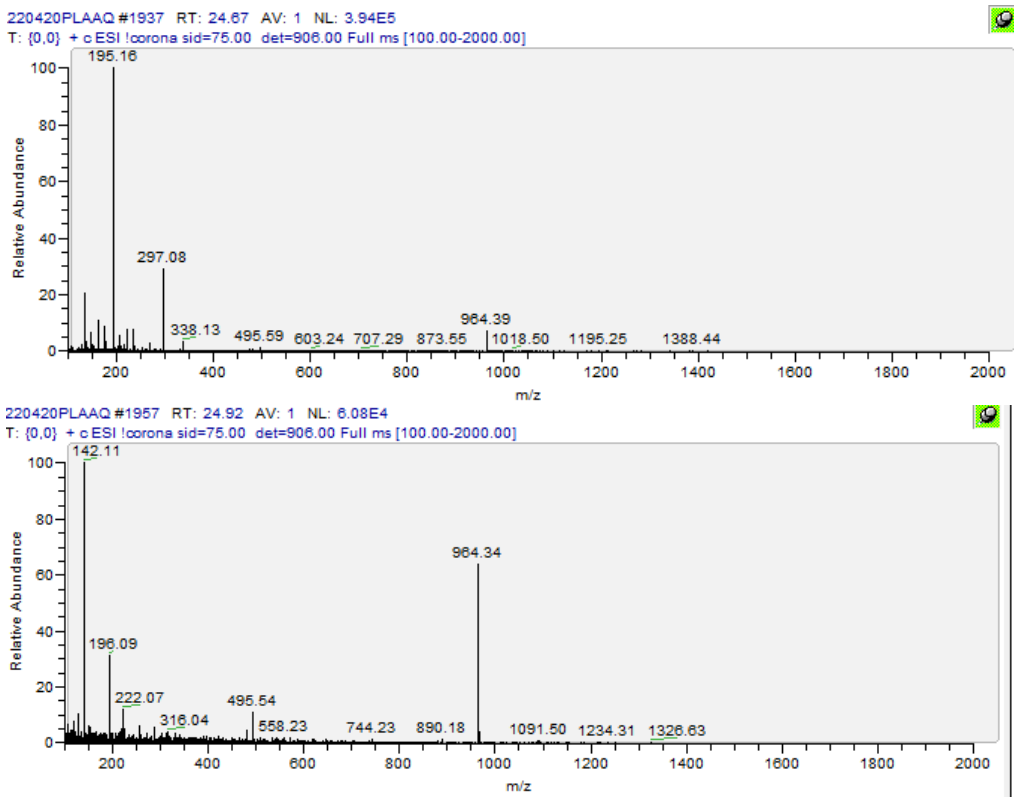
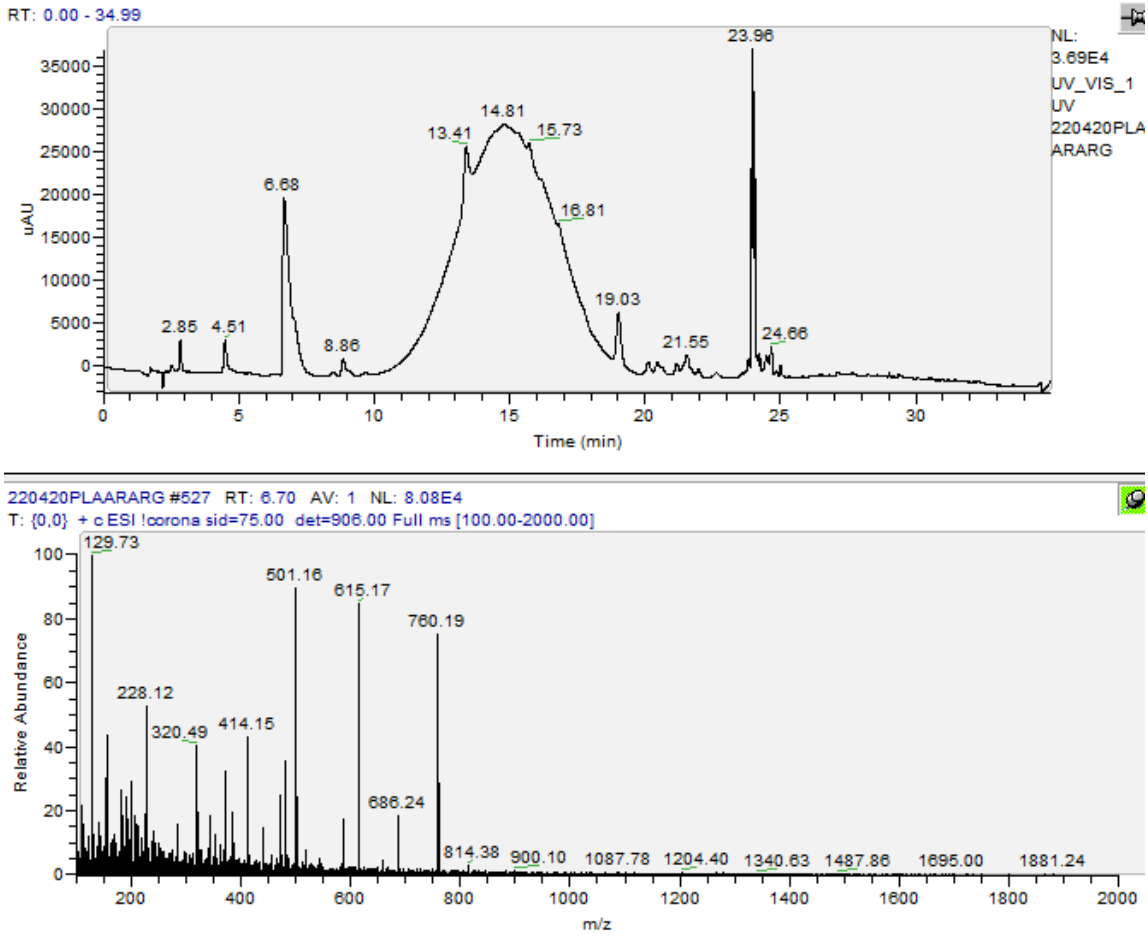


Figure S31. LC-MS analysis of the reaction between N-Gln-peptide and PLAA.



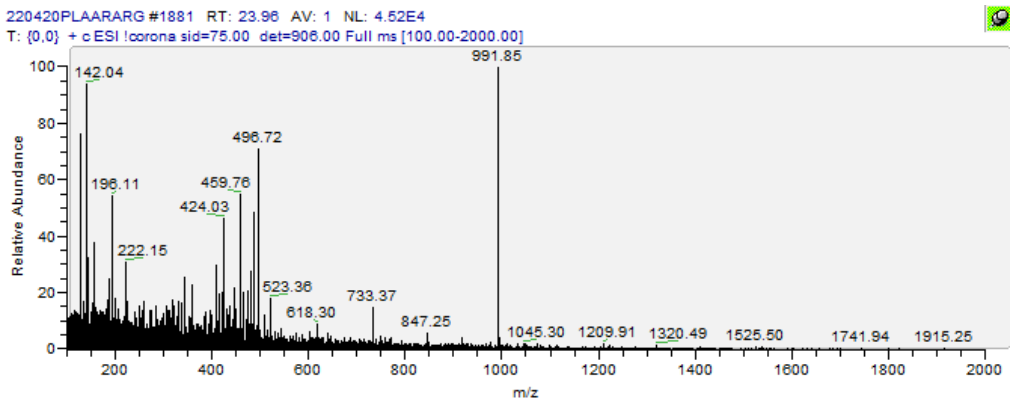
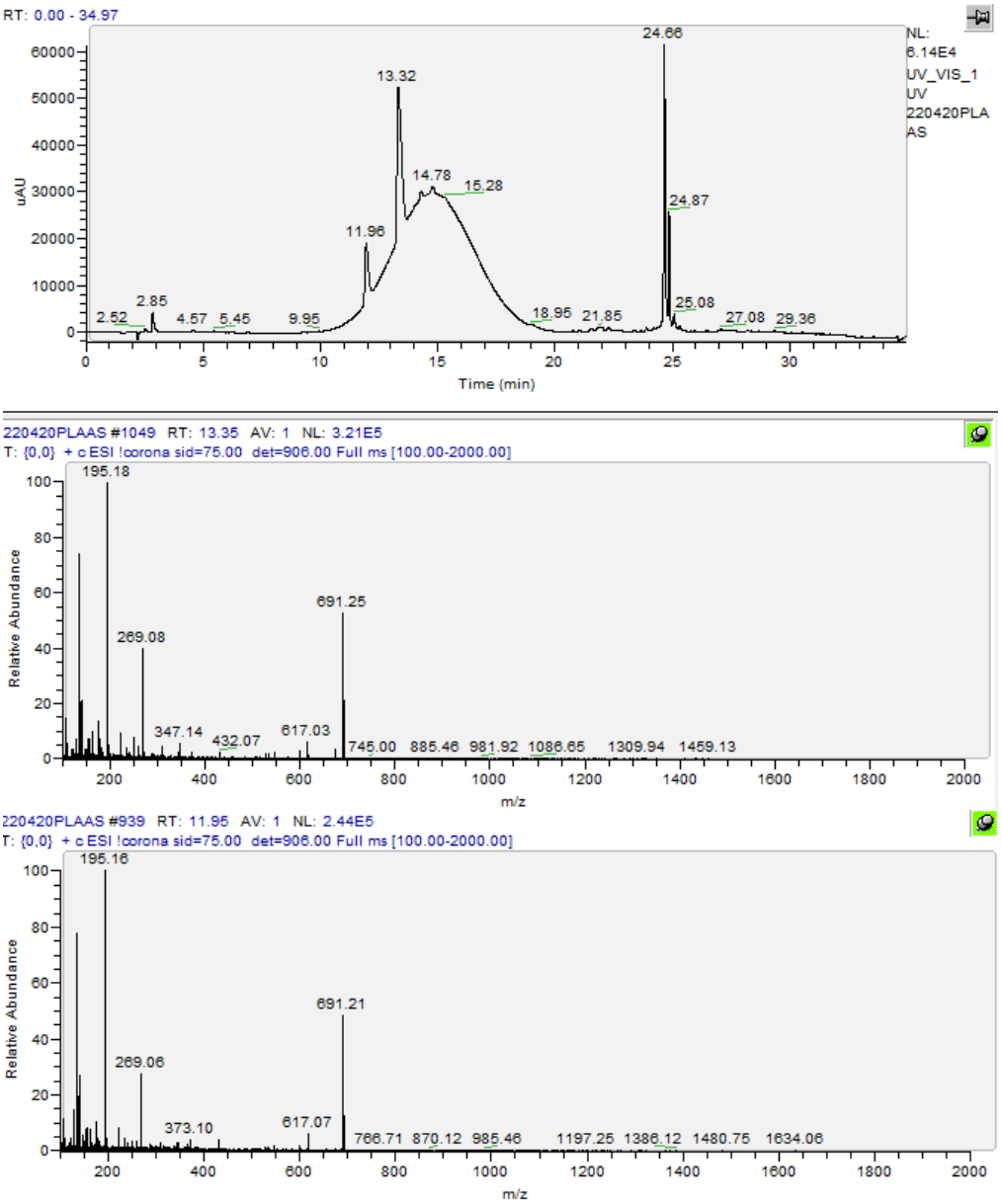


Figure S32. LC-MS analysis of the reaction between N-Arg-peptide and PLAA.



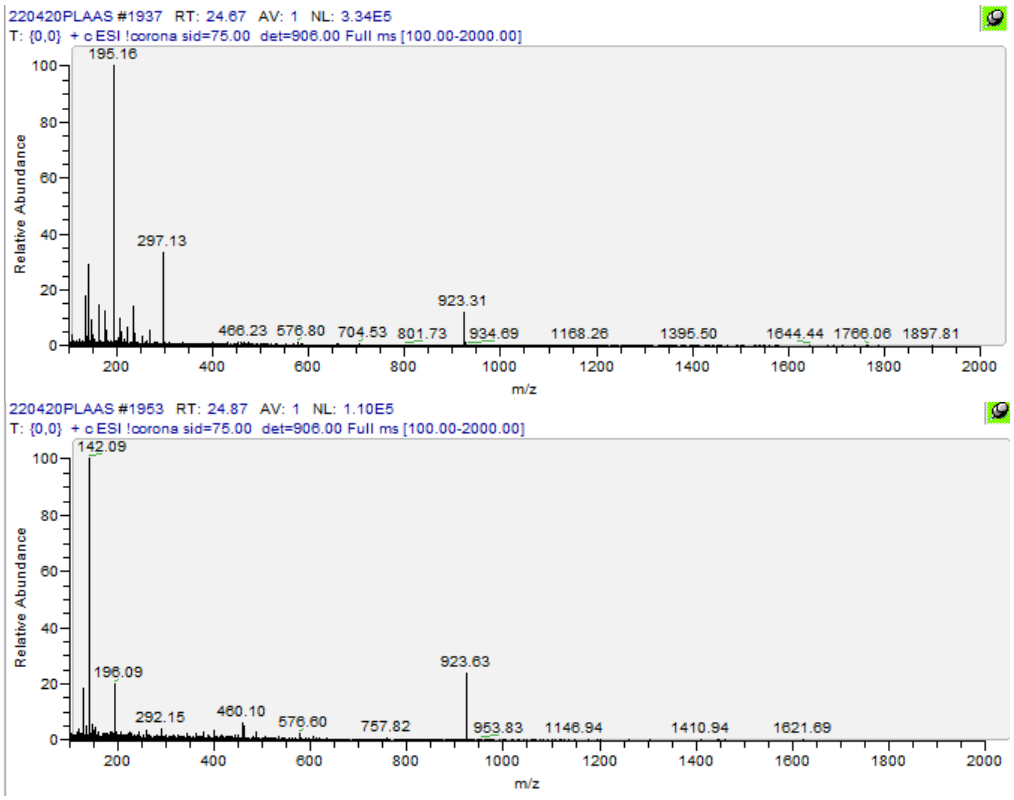
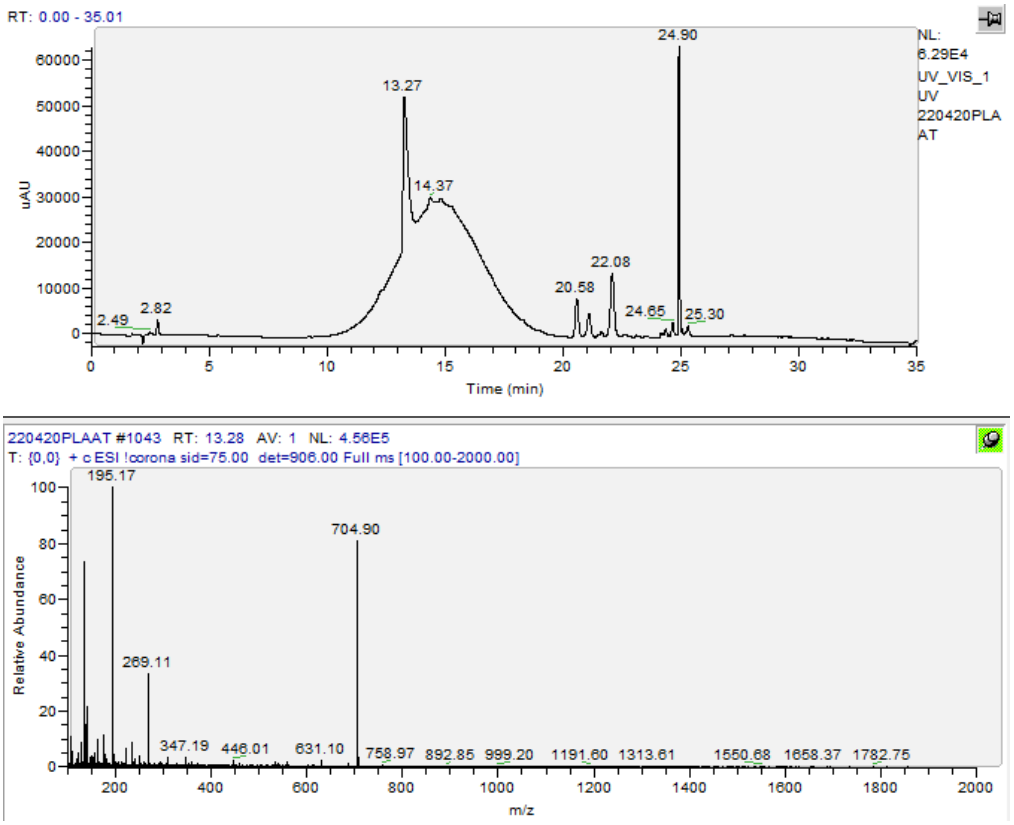


Figure S33. LC-MS analysis of the reaction between N-Ser-peptide and PLAA.



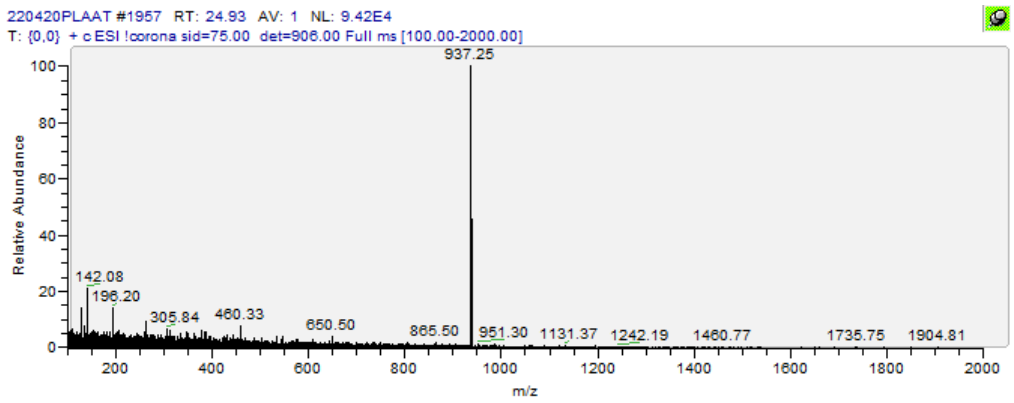
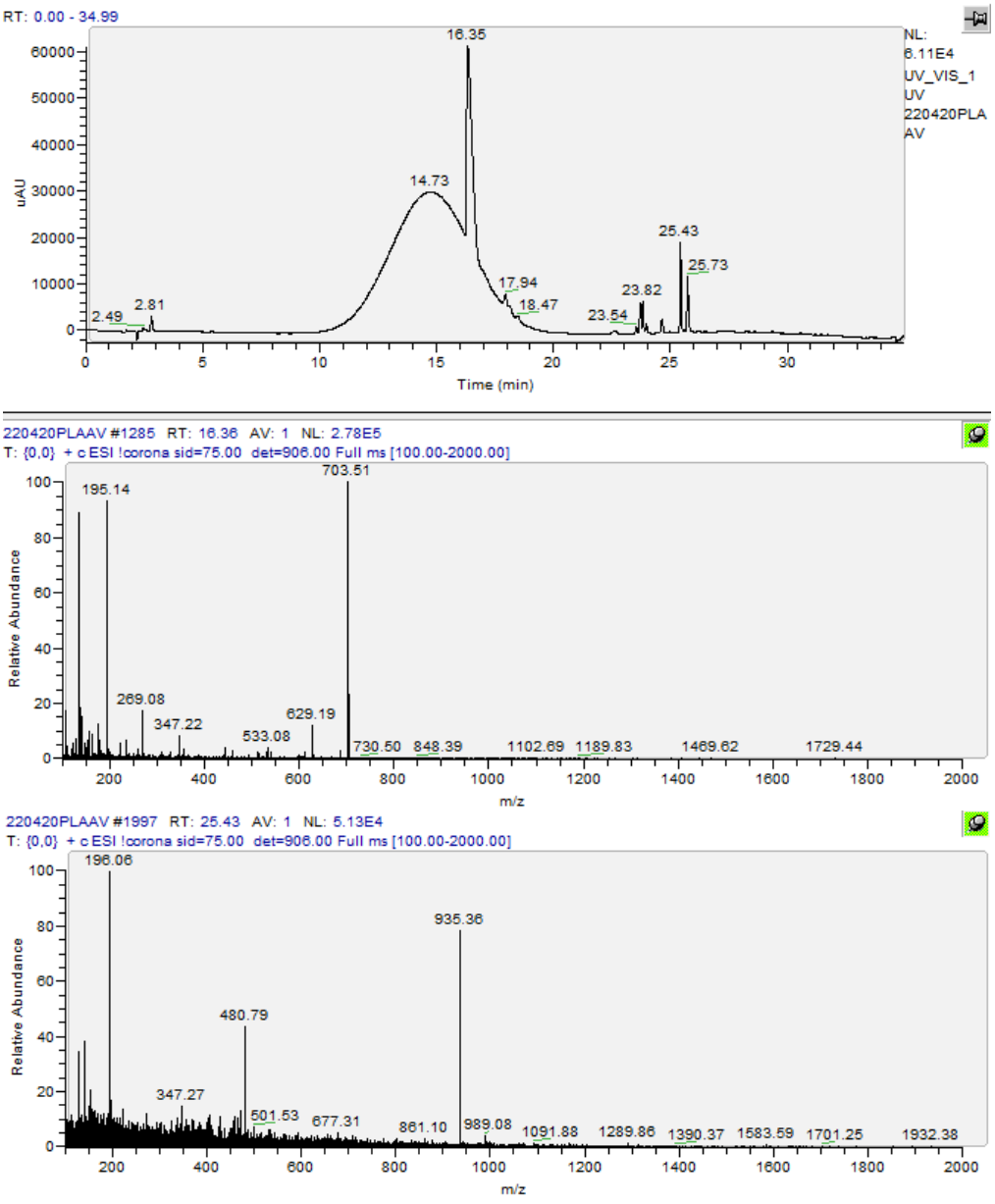


Figure S34. LC-MS analysis of the reaction between N-Thr-peptide and PLAA.



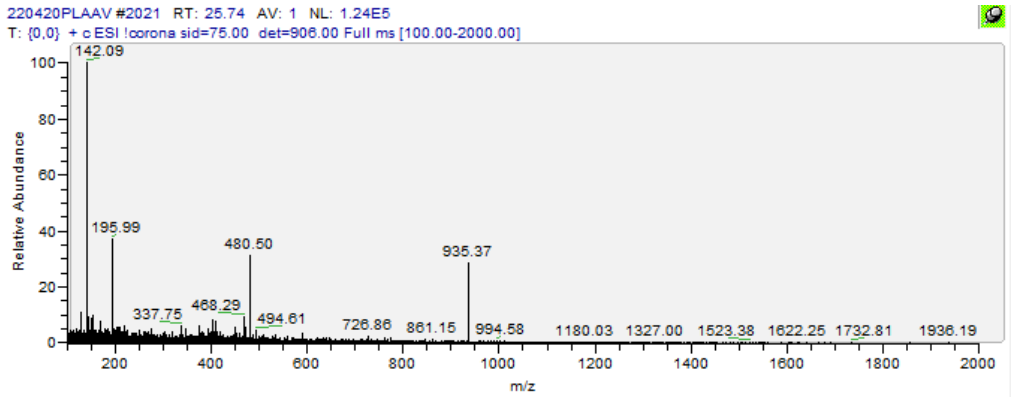


Figure S35. LC-MS analysis of the reaction between N-Val-peptide and PLAA.

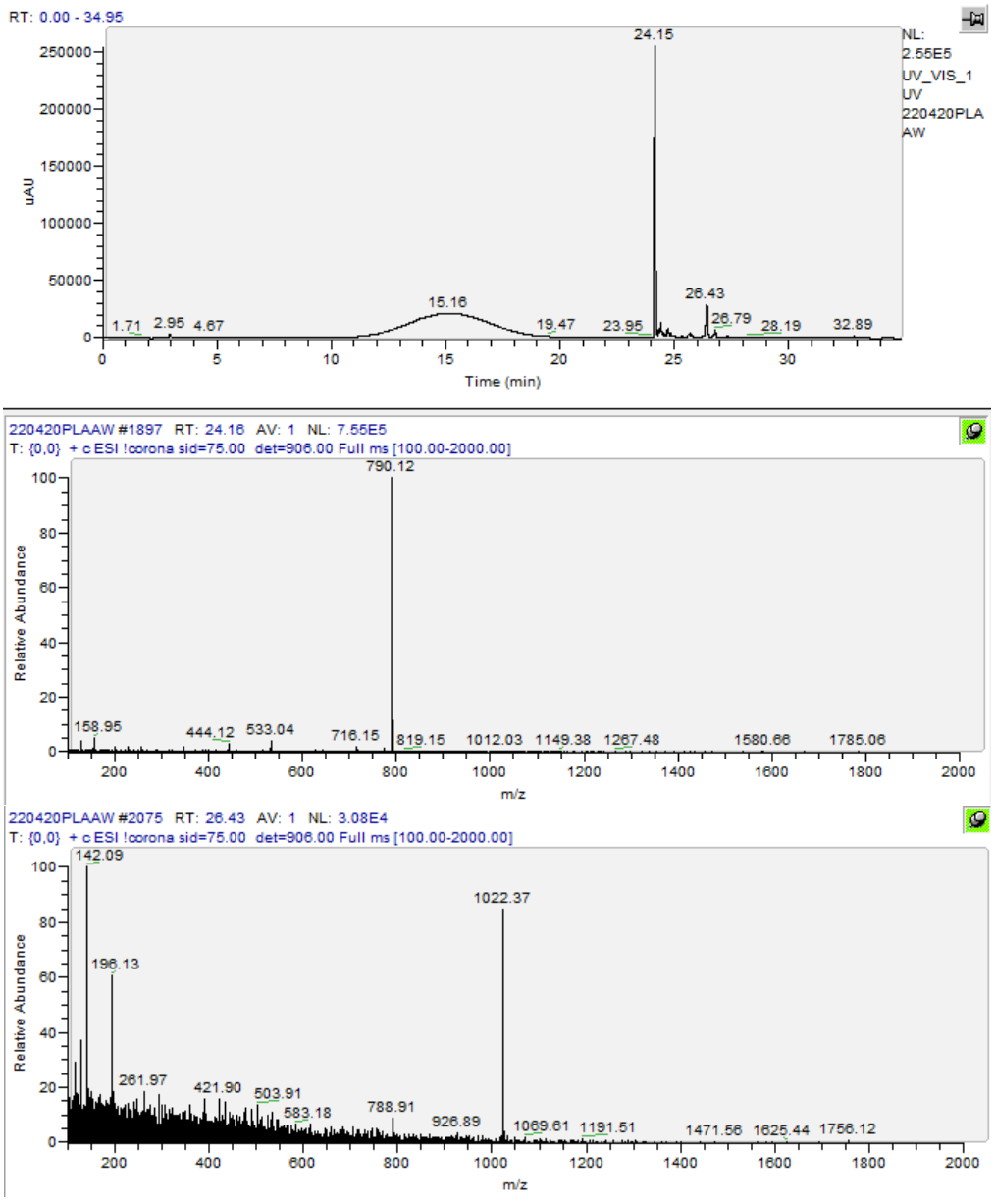


Figure S36. LC-MS analysis of the reaction between N-Trp-peptide and PLAA.

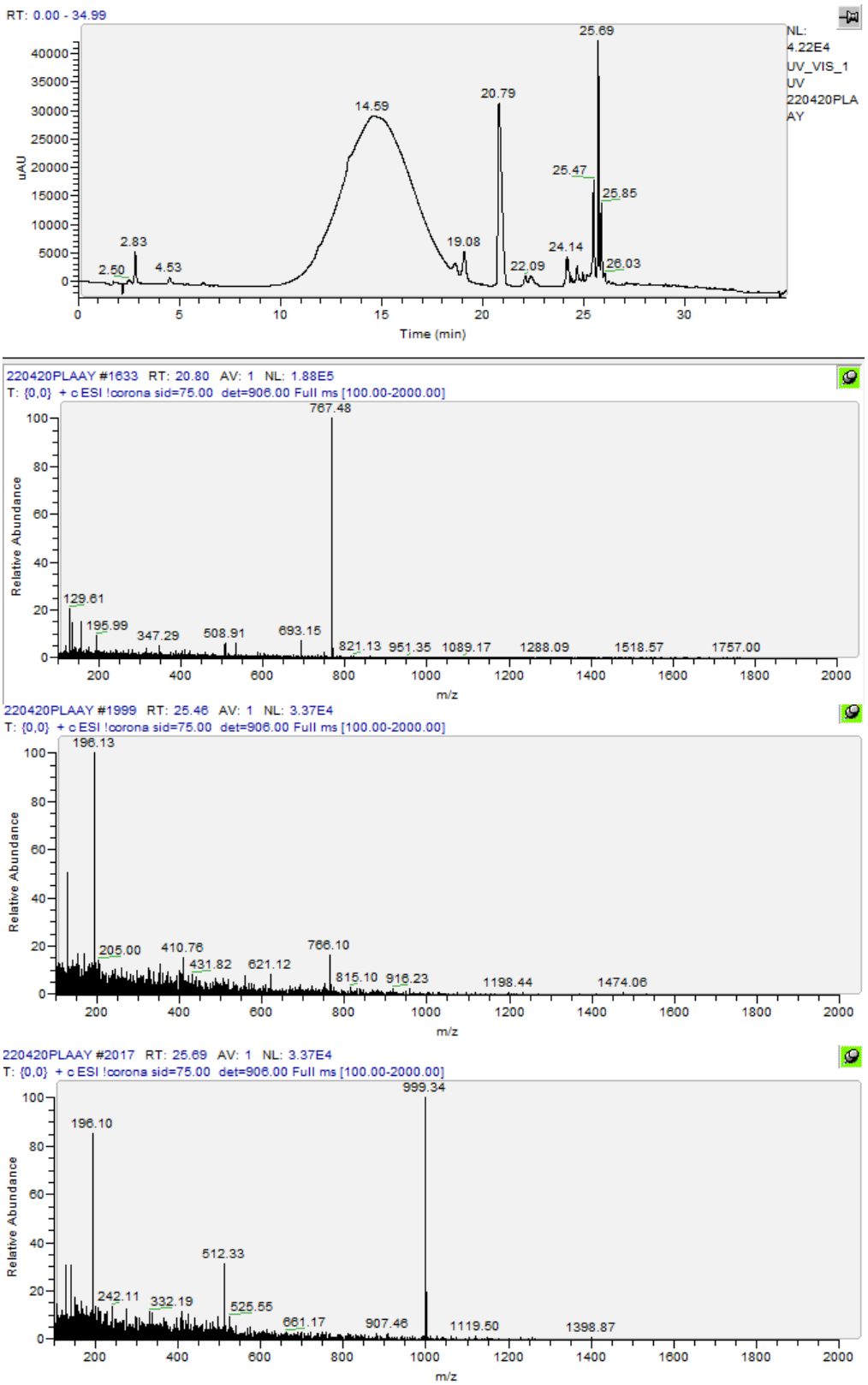


Figure S37. LC-MS analysis of the reaction between N-Tyr-peptide and PLAA.

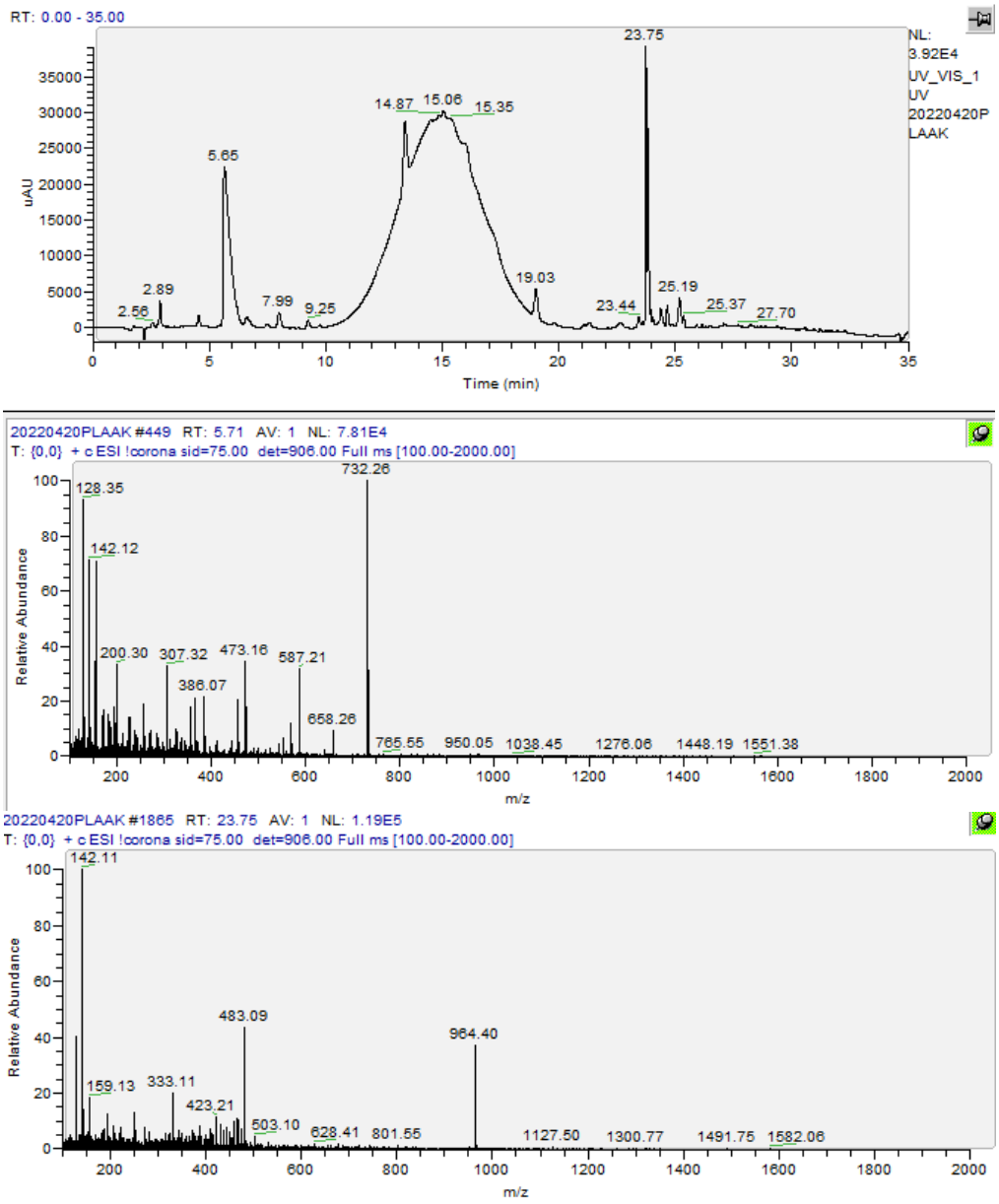
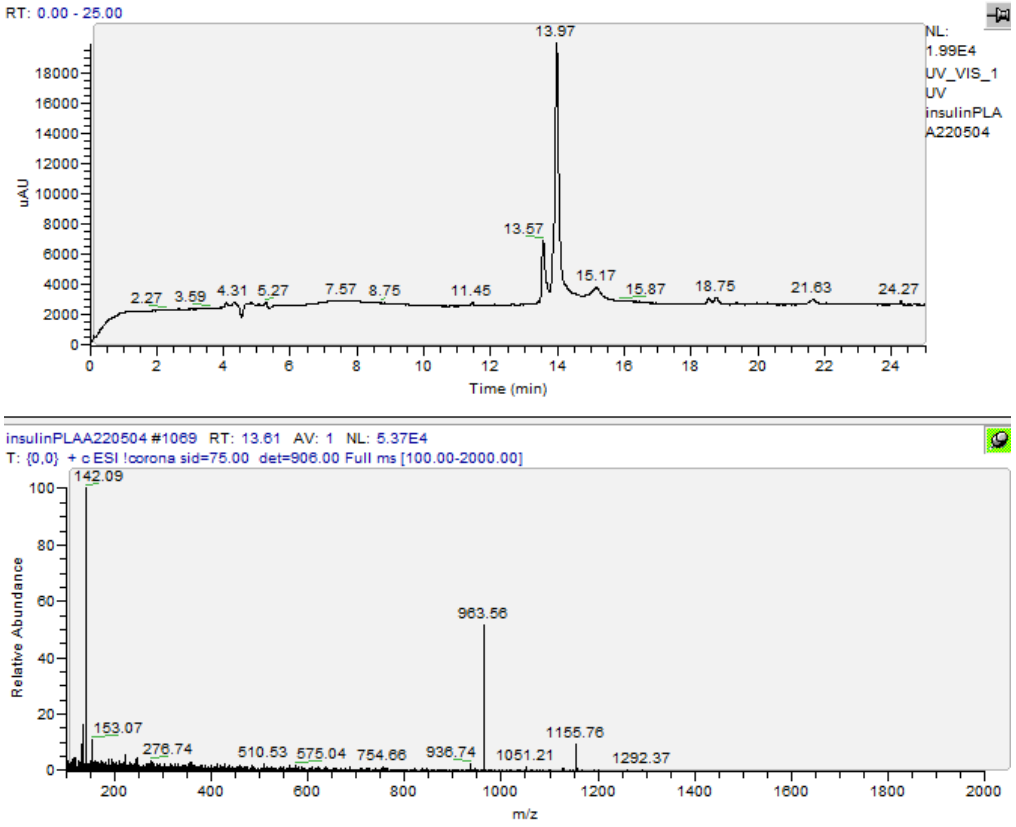


Figure S38. LC-MS analysis of the reaction between N-Lys-peptide and PLAA.

S4. Site-specific azidation of insulin

In a 1.5 mL sample tube, insulin (3.4 mg, 0.59 μmol) was dissolved in $NaHCO_3$ buffer (0.1 M, pH 8.4, 0.5 mL), and PLAA (2.9 mg, 11.80 μmol) was added. After reacting at 25 ℃ for 24 hours, the excess PLAA was removed by dialysis against first 0.1% TFA and then water (MWCO 3 kDa), and modification of insulin was analyzed by LC-MS.

Method of LC-MS (Column: Elitehplc SinoChrom ODS-BP 5 μm 4.6×150 mm, flow rate 1 mL/min). MeCN and H_2O were buffered with 0.1% formic acid. Multi-Step gradient: 0-8 min, 5%-30% (MeCN); 8-20 min, 30%-50% (MeCN); 20-22 min, 50%-95% (MeCN); 22-24 min, 95% (MeCN), 25-25 min, 95%-5% (MeCN). Detection wavelength: 254 nm. Injection volume: 10 μL. Parameter setting of Mass: 0-25 min; Mass range, 100-2000; Ionization mode, ESI; Pro temperature, 350 ℃.



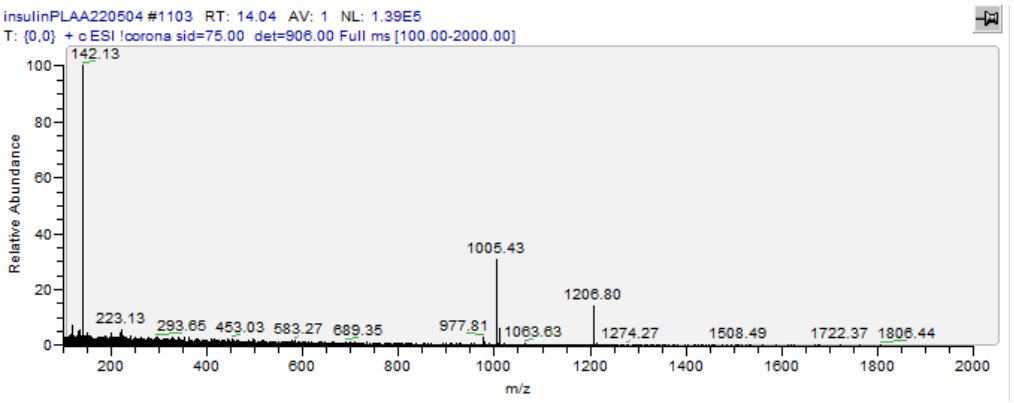


Figure S39. LC-MS analysis of insulin-PLAA.

For further confirmation, the dialyzed insulins (0.136 mg) were added to 100 μL buffer (30 mM Tris·Cl, pH 8.0), and then proteinase K (3 μL, 1.25 mg/mL) was added. The reaction mixture was incubated at 37 ℃ for 24 hours. Then DL-Dithiothreitol (DTT, 3 μL, 6.5 mg/mL) was added in. The reaction mixture was incubated at 37 ℃ for another 4 hours and then analyzed by LC-MS.

Method of LC-MS (Column: Elitehplc SinoChrom ODS-BP 5 μm 4.6×150 mm, flow rate 1 mL/min). MeCN and H_2O were buffered with 0.1% formic acid. Multi-Step gradient: 0-8 min, 5%-30% (MeCN); 8-20 min, 30%-50% (MeCN); 20-30 min, 50%-95% (MeCN); 30-32 min, 95% (MeCN), 32-33 min, 95%-5% (MeCN), 33-35 min, 5% (MeCN). Detection wavelength: 254 nm. Injection volume: 10 μL. Parameter setting of Mass: 0-35 min; Mass range, 100-2000; Ionization mode, ESI; Pro temperature, 350 ℃.

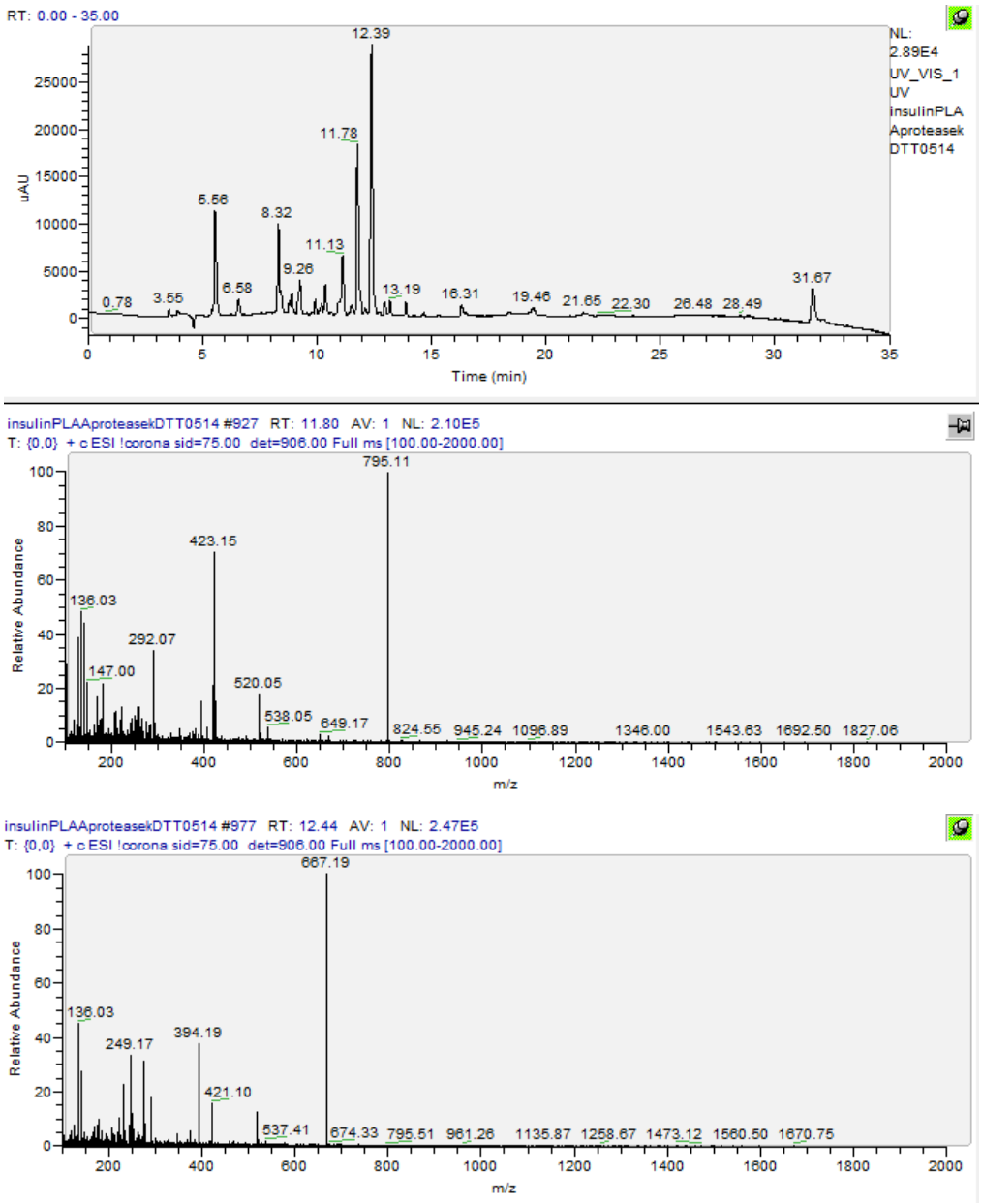


Figure S40. LC-MS analysis of the digestion results of insulin-PLAA.

The correspondence of other digestive products is listed in Table 2.

Table 2 The mass spectrum of other digestive products

Time (min)	Observed	Calculated	Residues
5.56	508.11	508.28	B26-29
8.32	296.12	296.12	A18-19
8.90	368.19	368.23	B10-12
9.26	565.08	565.27	B20-24
9.94	322.01	322.14	A9-11
10.37	416.54	416.25	B27-30
11.13	329.10	329.15	B25-26
11.13	295.14	295.17	A13-14, B15-16, B16-17
11.78	795.11	795.36	PLAA-A1-5
12.39	394.15	394.23	B16-18,
12.39	667.19	667.31	PLAA-A1-4
12.94	538.14	538.25	A16-19
13.19	313.16	313.15	B24-25
13.89	370.14	370.18	B23-25
19.46	378.04	378.19	B1-3

Furthermore, the dialysate of insulins (10 μL) were labeled with 5 equivalent of FITC-yne with copper catalyst (3.91 mM $CuSO_4$, 7.82 mM TBTA, 33 mM sodium ascorbate), and analyzed by SDS–PAGE.

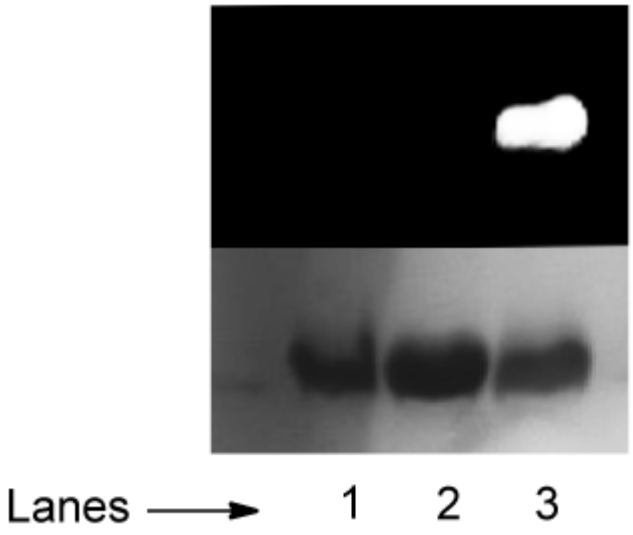


Figure S41. SDS-PAGE of insulins, visualized under UV light (top) and after Coomassie-blue staining (bottom). Wells from left to right: (1) Unmodified insulin reacted with FITC-yne; (2) Insulin; (3) Insulin-PLAA-FITC-yne.

S5. Imine hydrolysis

In a 5 mL sample tube, L-Lysine amide dihydrochloride (20.9 mg, 95.9 μmol) was dissolved in $NaHCO_3$ D2O buffer (0.1 M, pH 8.4, 3 mL), and PLAA (20 mg, 79.9 μmol) was added in. The imine product was determined by NMR after reacting at 25 ℃ for 14 hours. Then the reaction solution was acidified to pD 5 with trifluoroacetic acid and incubated at 25 ℃ for one hour. The the imine was completely hydrolyzed determined by NMR.

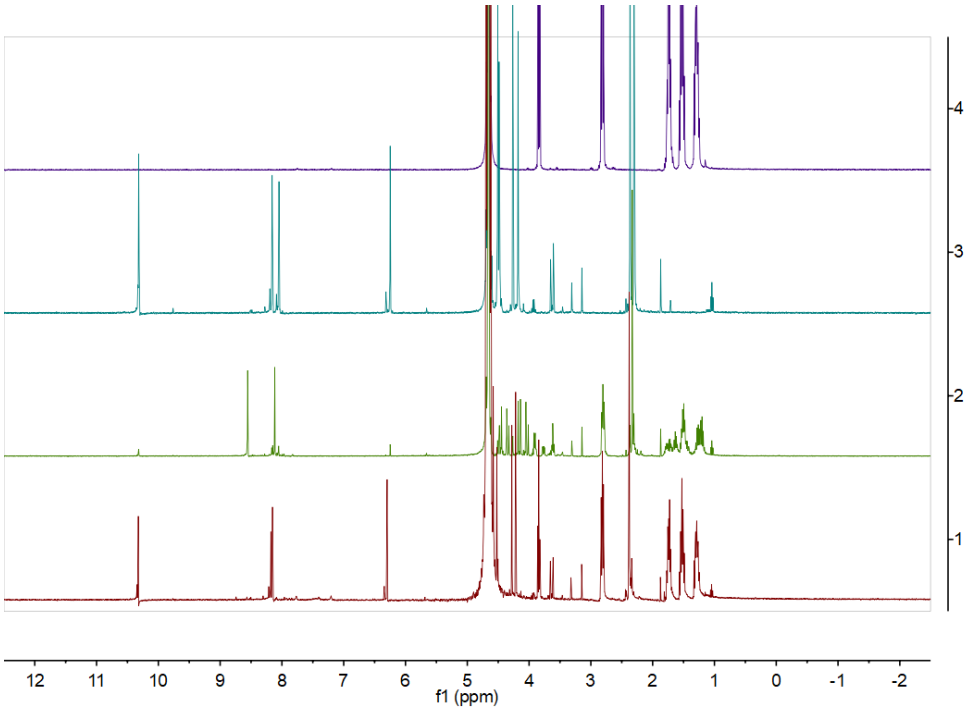


Figure S42. The NMR spectrum of the reaction of L-Lysine amide dihydrochloride and PLAA. From bottom to top: L-Lysine amide dihydrochloride-PLAA (D_2O solution, pD 5); L-Lysine amide dihydrochloride-PLAA ($NaHCO_3$, D_2O solution); PLAA ($NaHCO_3$ D_2O solution); L-Lysine amide dihydrochloride ($NaHCO_3$, D_2O solution).

S6. Site-specific azidation of myoglobin

In a 1.5 mL sample tube, myoglobin (5.7 mg, 0.33 μmol) was dissolved in $NaHCO_3$ buffer (0.1 M, pH 8.4, 0.4 mL), and PLAA (4.2 mg, 16.76 μmol) was added. After reacting at 25 ℃ for 24 hours, the excess PLAA and salts were removed by ultracentrifugation (10 kDa MWCO). The resulting solution of myoglobin (2.3 mg, 100 μL) were added to 39.2 μL buffer (30 mM Tris•Cl, pH 8.0), and then proteinase K (60.8 μL, 1.25 mg/mL) was added. The reaction mixture was incubated at 37 ℃ for 24 hours and then analyzed by LC-MS.

Method of LC-MS (Column: Elitehplc SinoChrom ODS-BP 5 μm 4.6×150 mm, flow rate 1 mL/min). MeCN and H2O were buffered with 0.1% formic acid. Multi-Step gradient: 0-14 min, 5%-95% (MeCN); 14-17 min, 95% (MeCN); 17-18 min, 95%-5% (MeCN); 18-20 min, 5% (MeCN). Detection wavelength: 254 nm. Injection volume: 10 μL. Parameter setting of Mass: 0-20 min; Mass range, 100-2000; Ionization mode, ESI; Pro temperature, 350 ℃.

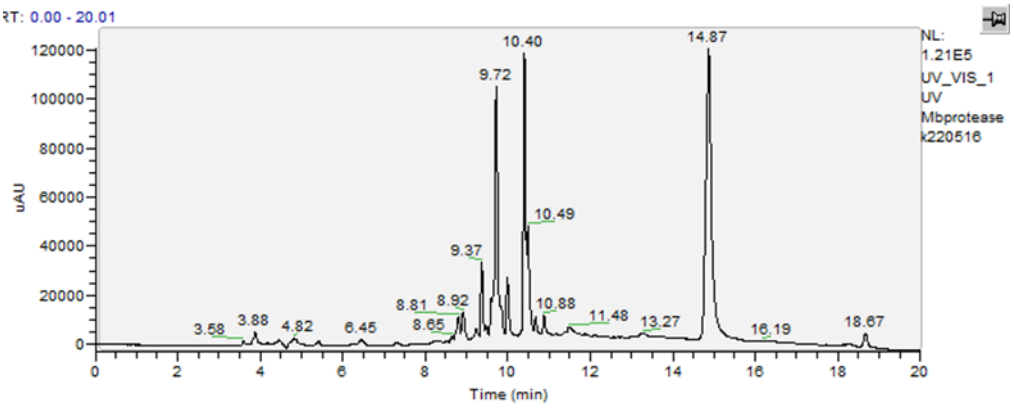


Figure S43. The chromatogram of the digestion of myoglobin.

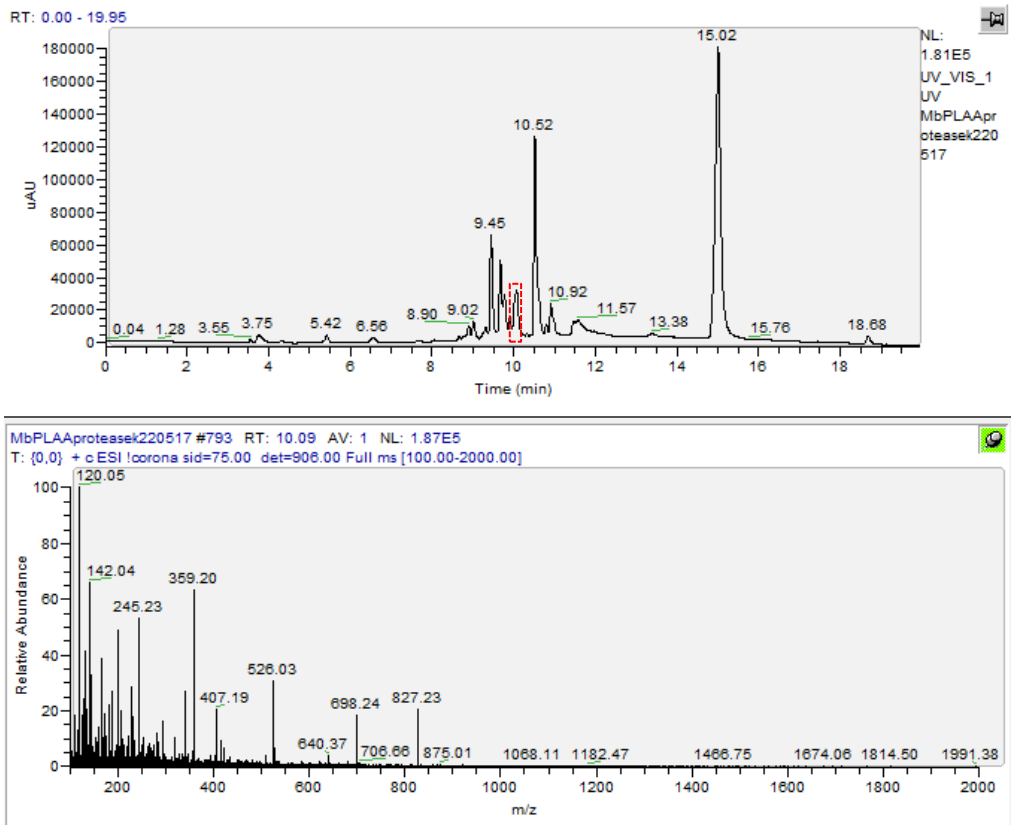


Figure S44. LC-MS analysis of the digestion results of myoglobin-PLAA. The major new peak around 10.1 min had a m/z of 827.26, corresponding to PLAA modified residue 1-6 ([*M*+H]$^+$, calculated 827.23).

S7. Site-specific azidation of AzoR

In a 1.5 mL tube, 10 μL of 1 mg/mL thrombin in thrombin dilution/storage buffer (50 mM sodium citrate, pH 6.5, 200 mM NaCl, 0.1% PEG-8000, 50% glycerol) was added in 130 μL of AzoR (0.78 mg/mL, 3.07 nmol) dissolved in Tris buffer (20 mM Tris, pH 7.3, 300 mM NaCl). After reacting at 20 ℃ for 20 hours, $NaHCO_3$ (1 M, 1.2 μL) and PLAA (10.96 nmol) in DMSO (1 μL) were added in 10 μL of the AzoR-thrombin reaction mixture. After reacting at 25 ℃ for 24 hours, 5 equivalent of FITC-yne and copper catalyst (3.91 mM $CuSO_4$, 7.82 mM TBTA, 33 mM sodium ascorbate) were added and reacted at 25 ℃ for one hour. Then 1.5 μL of 10% formic acid was added and the reaction mixture was further reacted at 25 °C for one hour prior to the SDS–PAGE analysis.

S8. Examine the activity of N-Gly-AzoR-PLAA

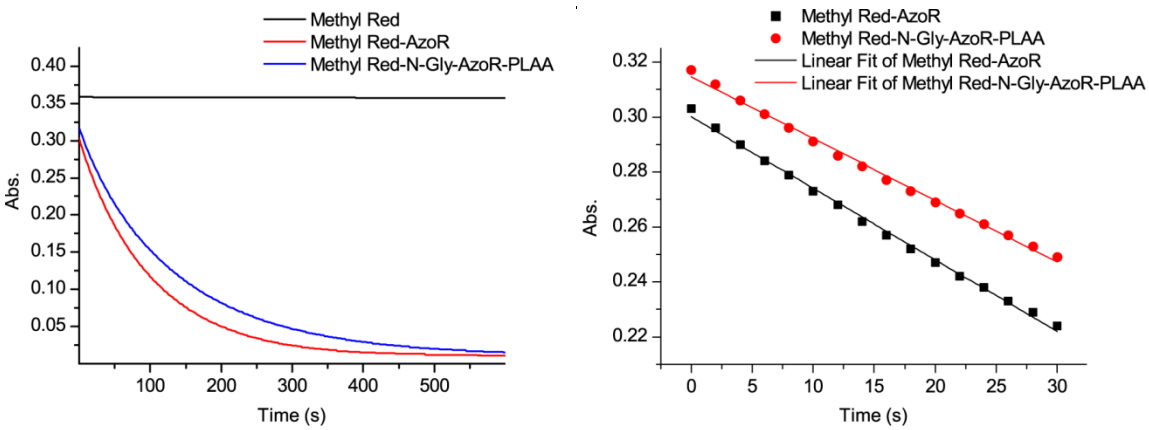


Figure S45. Examine the activity with methyl red degradation assay.

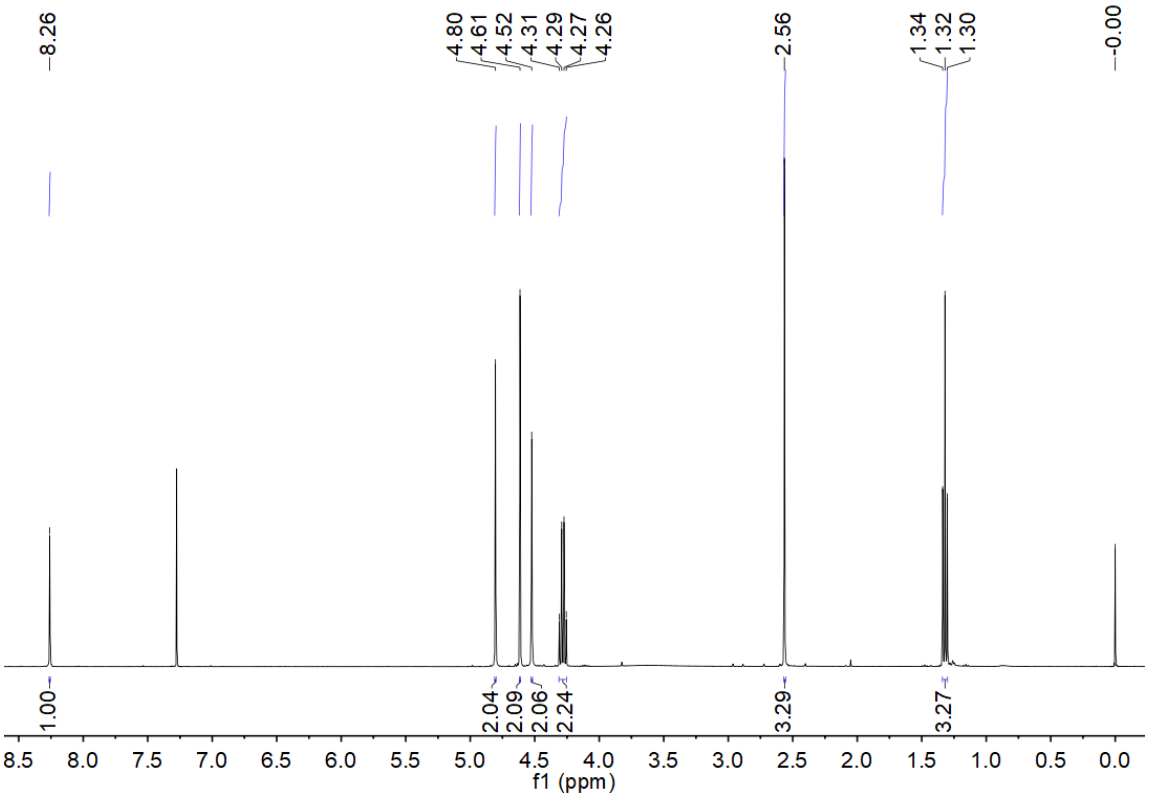
For the methyl red degradation assay, beta-Nicotinamide adenine dinucleotide disodium salt (NADH, 1.06 mg/ml, 100 μL) in water was added in methyl red (5 μmol) in 7.5 mL Tris-NaCl buffer (20 mM Tris, pH 7.3, 300 mM NaCl). To 2 mL of the NADH-methyl red solution, 10 μL of 0.6 mg/mL enzyme solution was added, and the degradation curve of methyl red was measured immediately at 450 nm.

S9. General procedure for the SDS-PAGE analysis

The solution of sample (8.0 μL) was mixed with 5 X loading buffer (2.0 μL) in a 0.2 mL microcentrifuge tube. The samples were loaded onto a gel containing 5% stacking gel and 10% separating gel. The gel was run at 120 V for 15 min, and then at 180 V for 40 min, with the Tris-glycine running buffer (25 mM Tris, 0.192 M glycine, and 0.1% (w/w) SDS, pH 8.3). After SDS-PAGE separation, the gel was washed with DI water and visualized with GenoSens 1850 for fluorescent signal. Then the gel was stained with Coomassie Brilliant Blue R250 and imaged for the protein staining signal.

S10. NMR Spectra

^{1}H NMR of 5



^{13}C NMR of 5

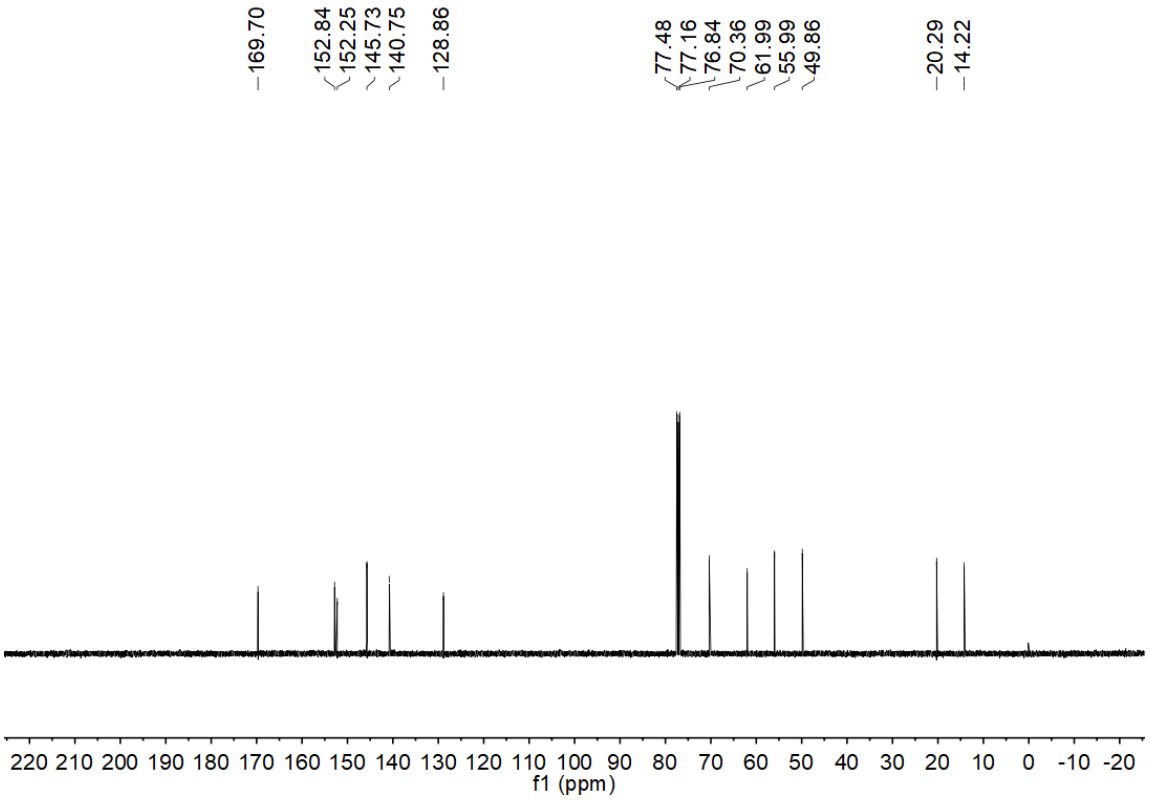
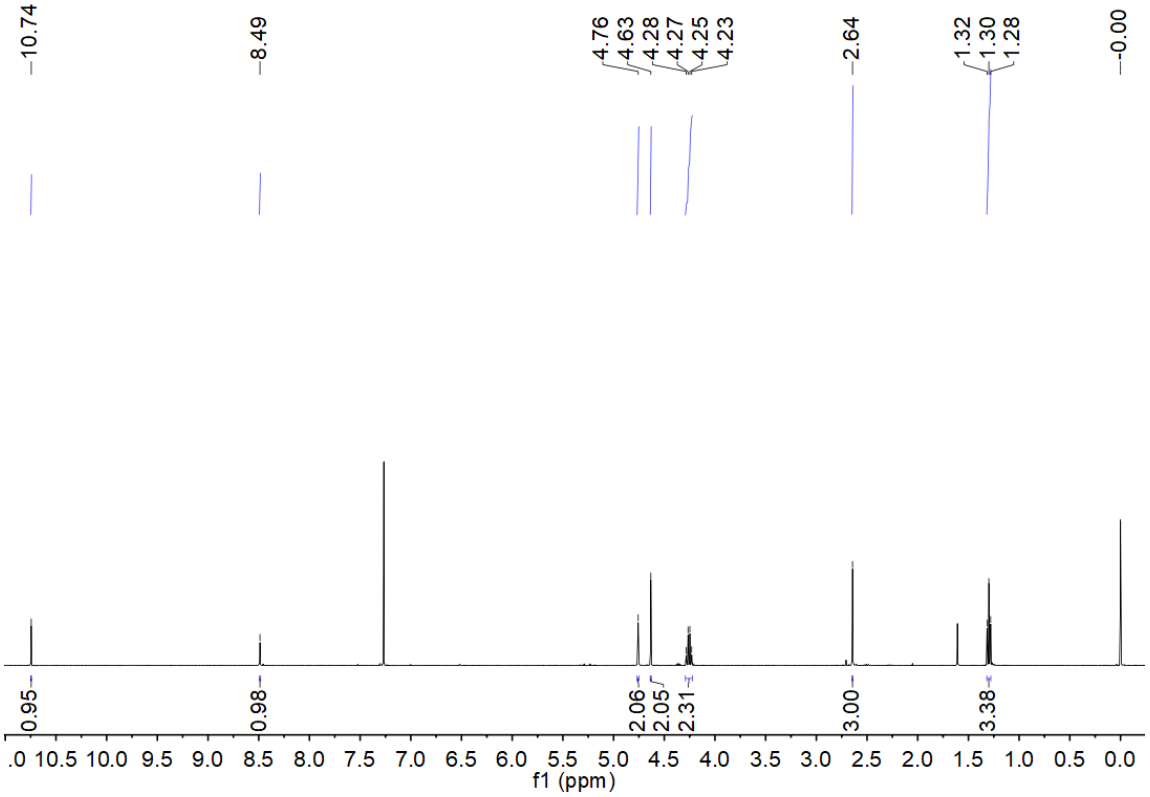


Figure S46. ^{1}H NMR and ^{13}C NMR spectra of **5** in $CDCl_3$.

^{1}H NMR of 6



^{13}C NMR of 6

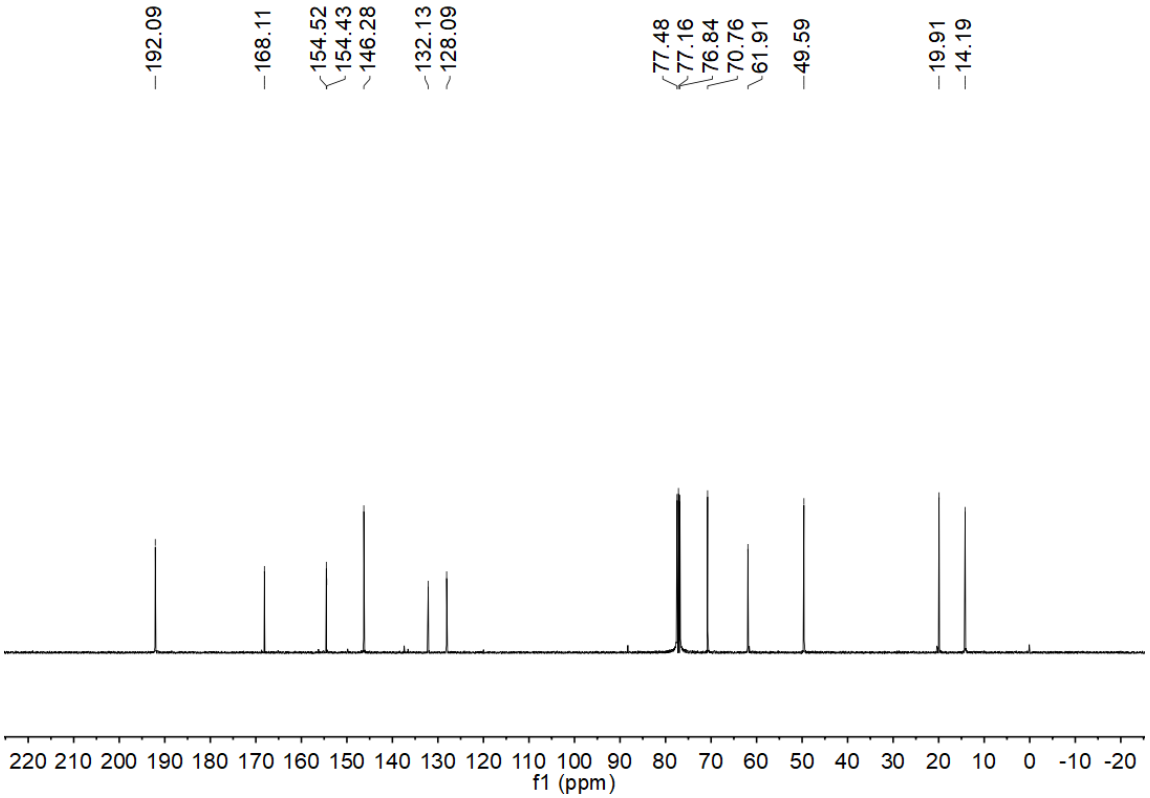
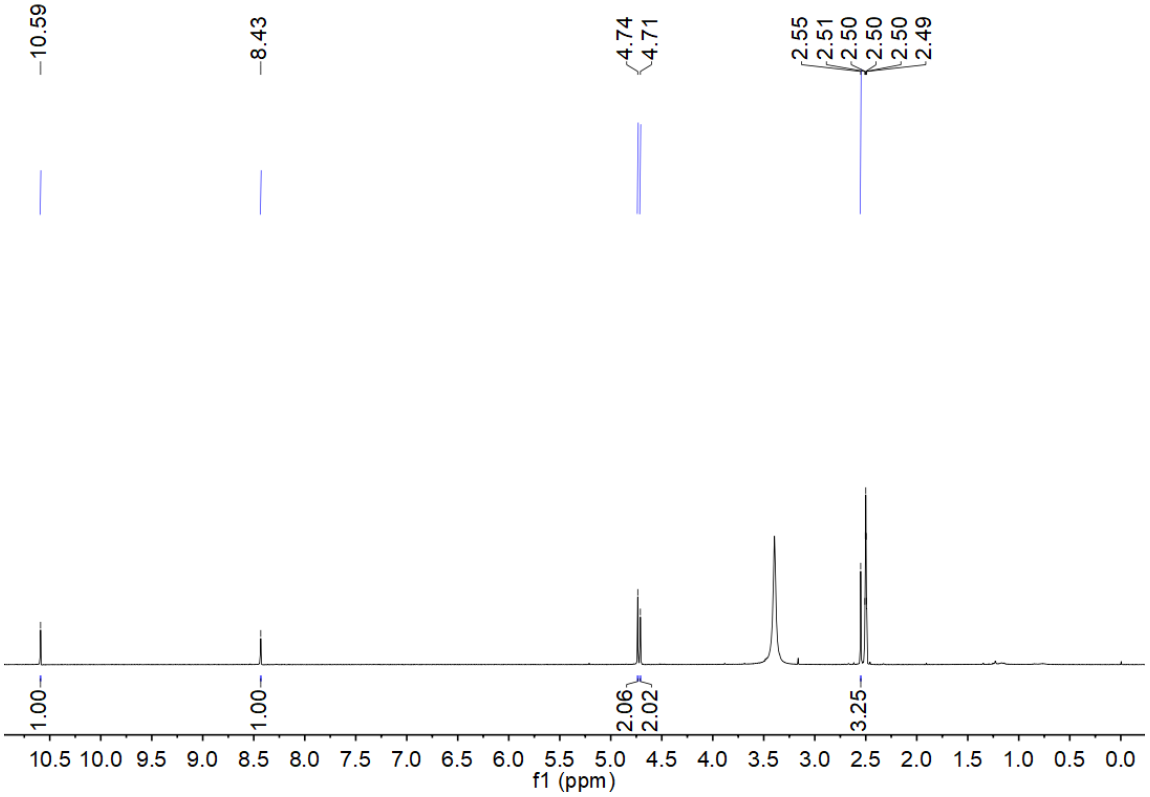


Figure S47. ^{1}H NMR and ^{13}C NMR spectra of **6** in $CDCl_3$.

^{1}H NMR of PLAA



^{13}C NMR of PLAA

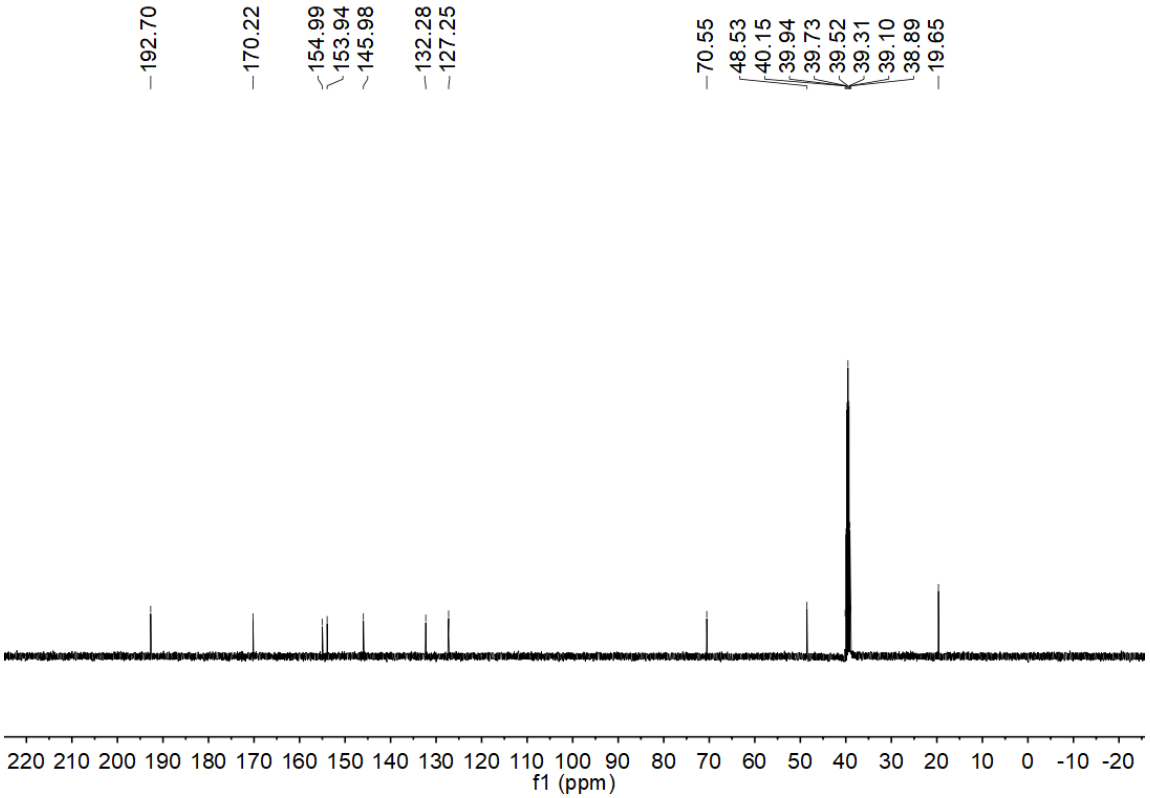
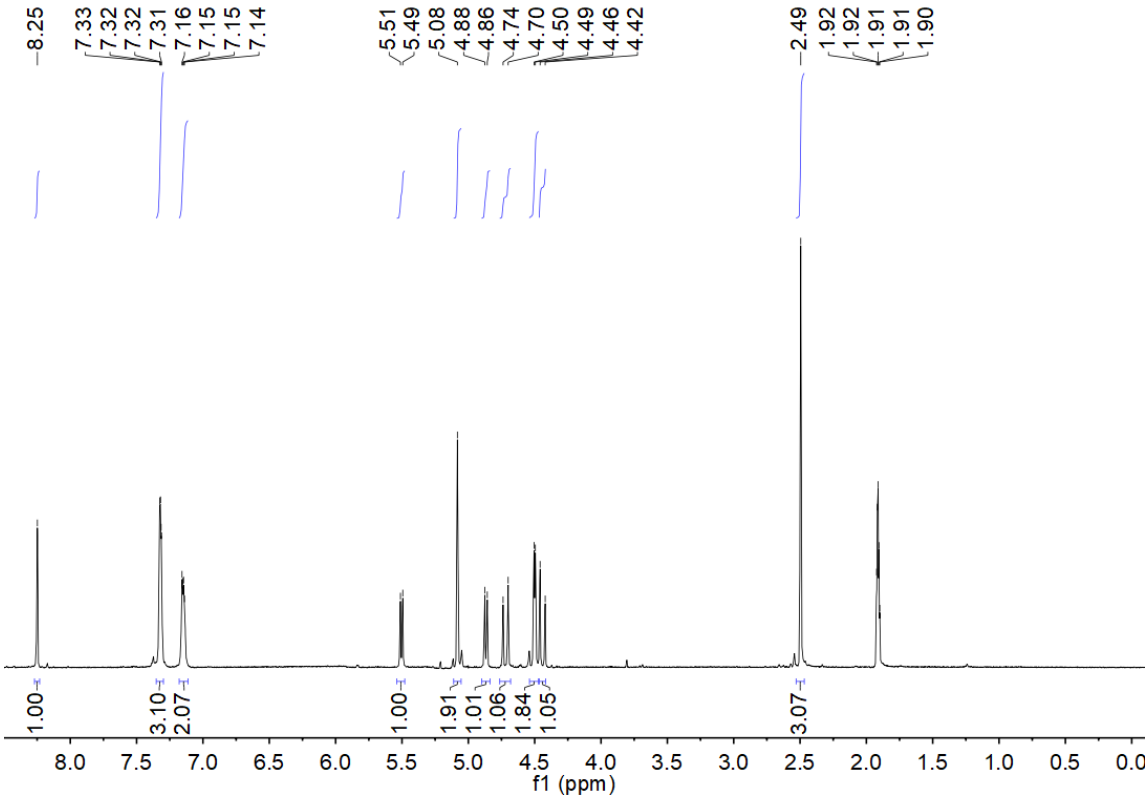


Figure S48. ^{1}H NMR and ^{13}C NMR spectra of **PLAA** in DMSO-d_6.

^{1}H NMR of 8-1



^{13}C NMR of 8-1

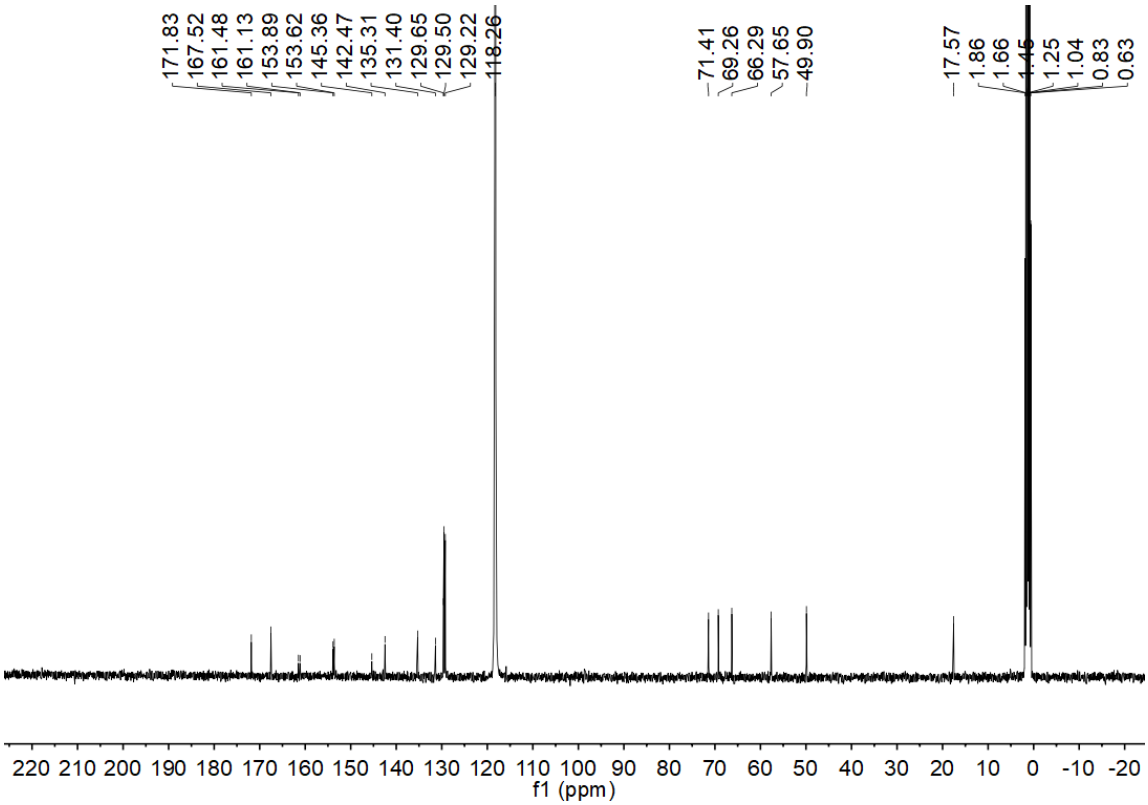
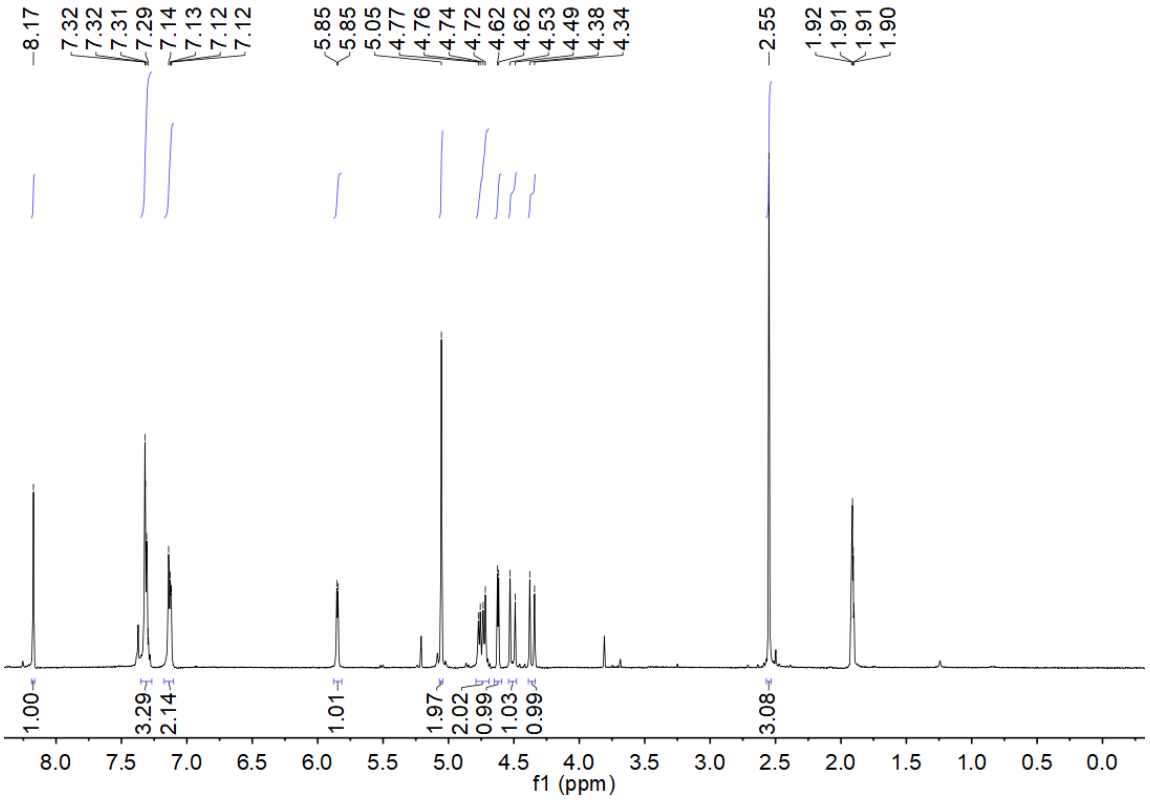


Figure S49. ^{1}H NMR and ^{13}C NMR spectra of **8-1** in CD_3CN.

^{1}H NMR of 8-2



^{13}C NMR of 8-2

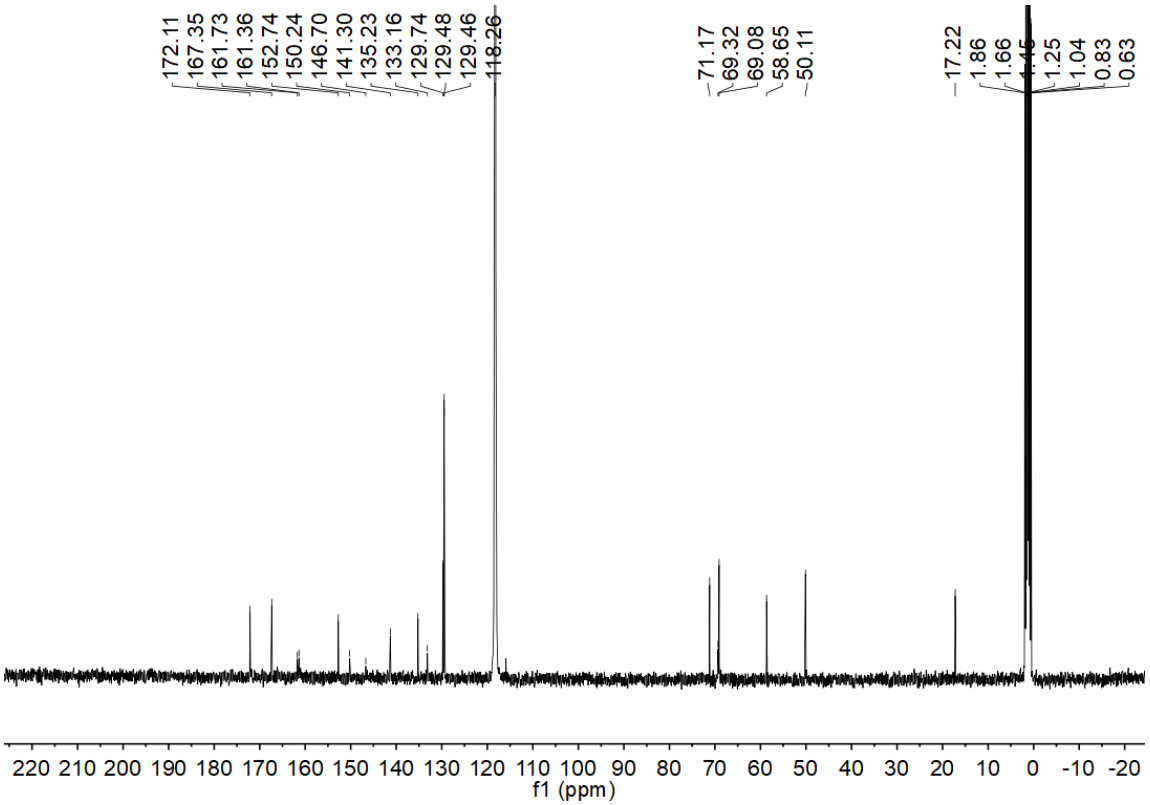
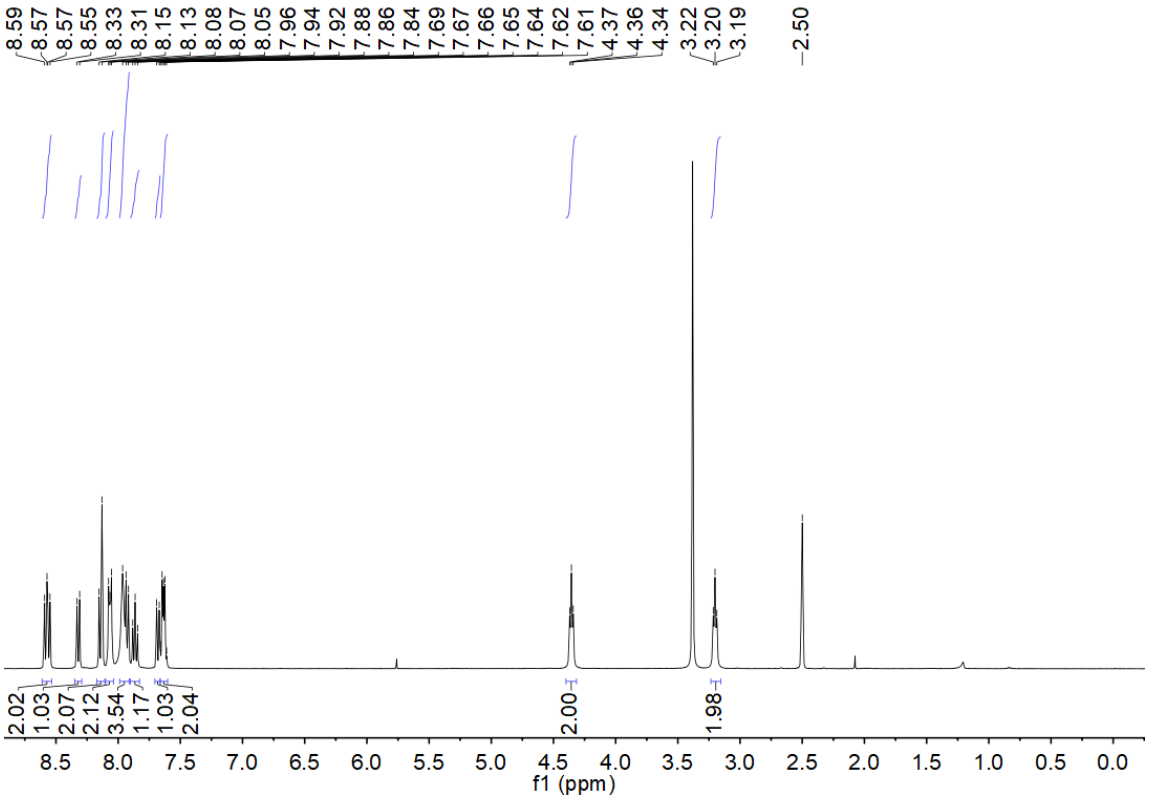


Figure S50. ^{1}H ^{NMR} and ^{13}C NMR spectra of **8-2** in CD_3CN

^{1}H NMR of 10



^{13}C NMR of 10

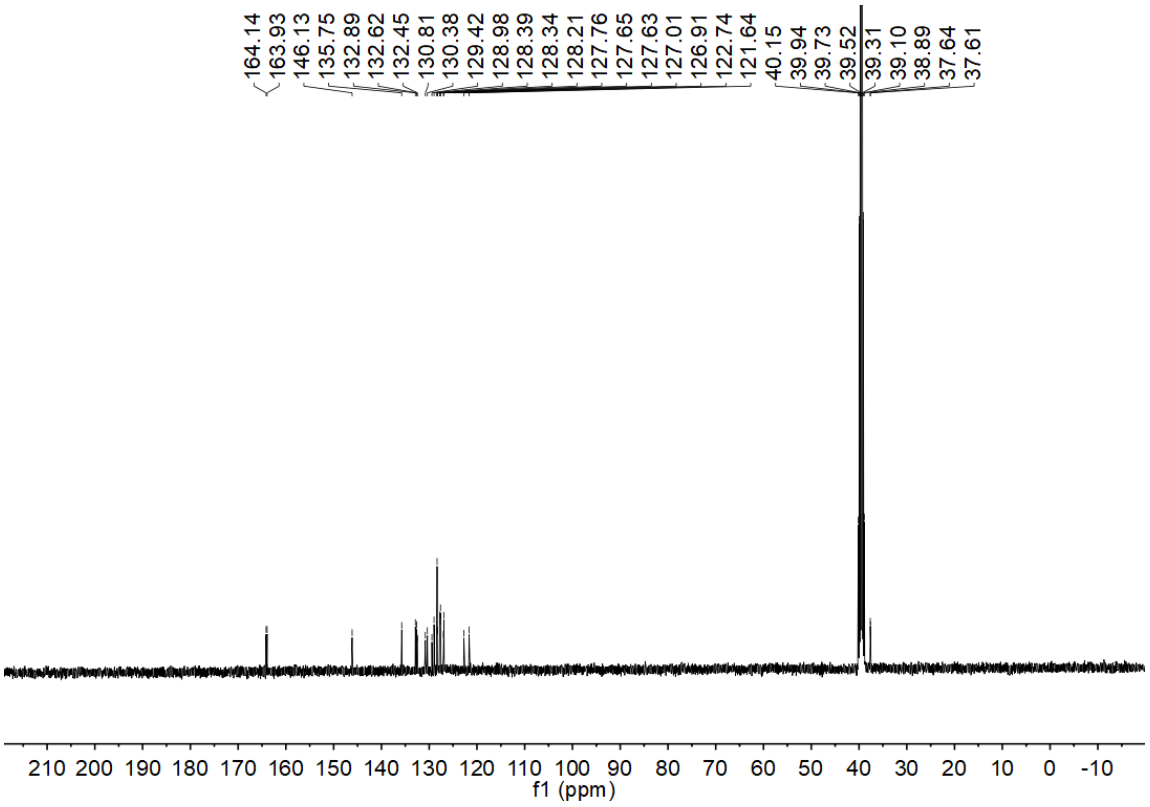
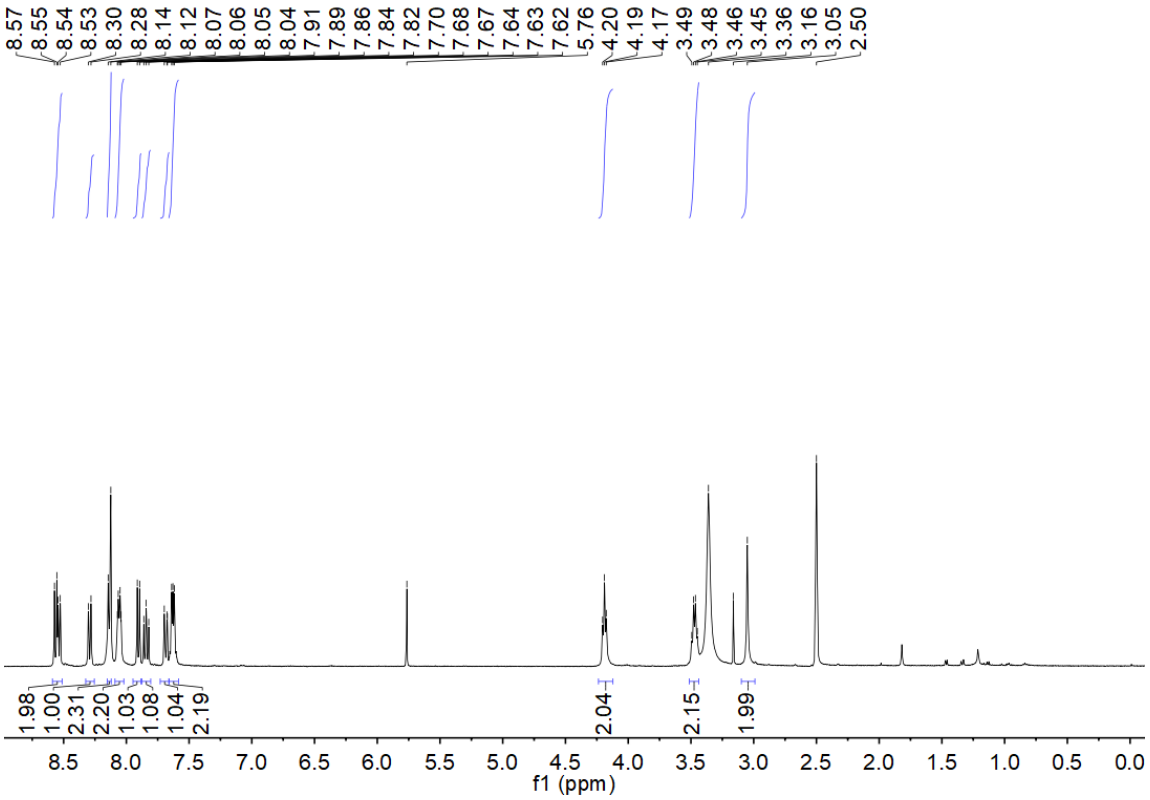


Figure S51. ^{1}H NMR and ^{13}C NMR spectra of **10** in DMSO-d_6.

^{1}H NMR of 11



^{13}C NMR of 11

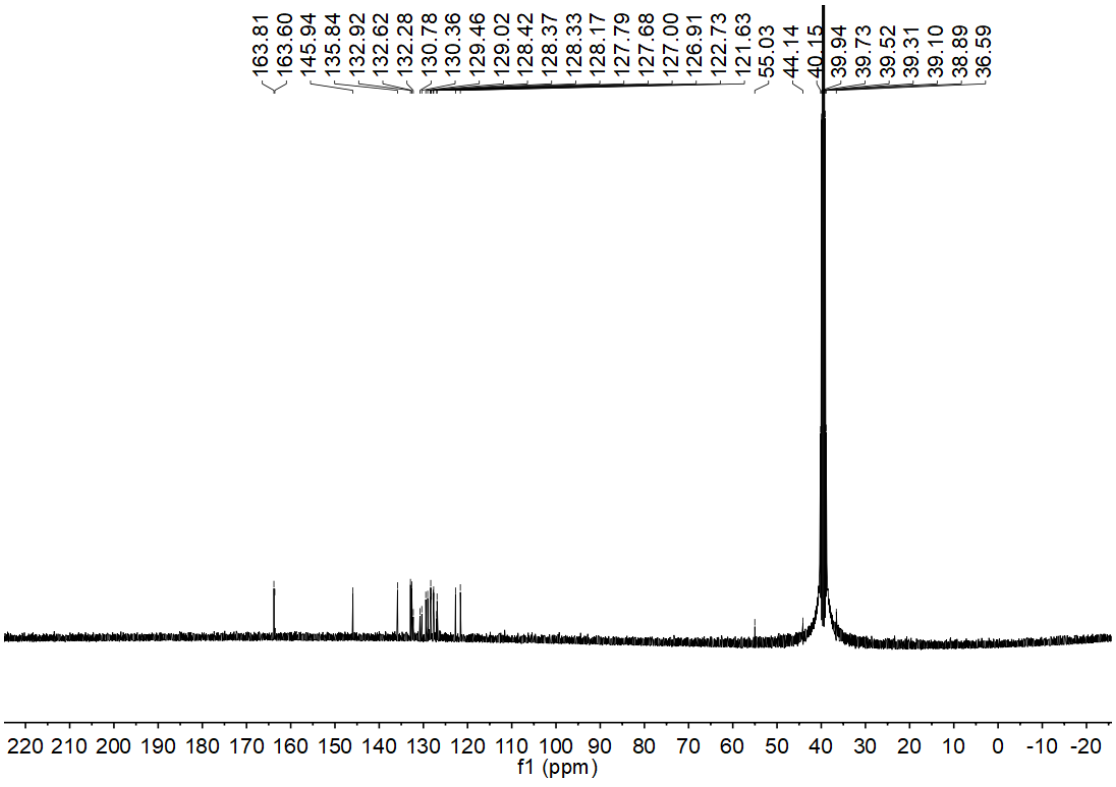


Figure S52. ^{1}H NMR and ^{13}C NMR spectra of **11** in DMSO-d_6.

S11. UV Spectra

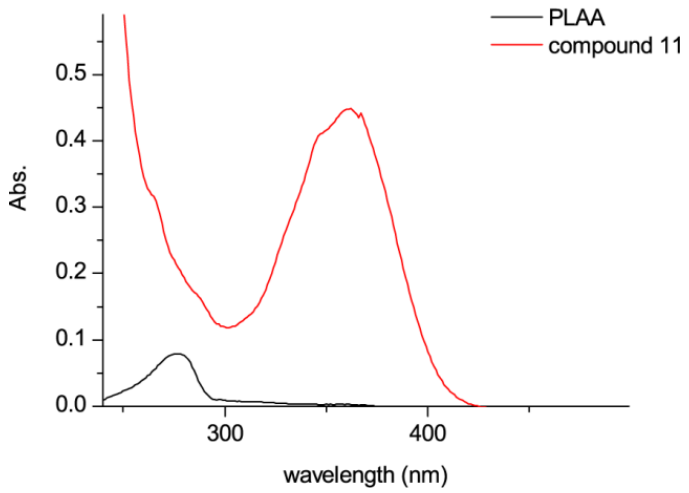


Figure S53. UV-vis absorption spectra of **PLAA** (21 µM, black) and compound **11** (21 µM, red) in MeOH at 25 ℃.

References

1. X. Mao, W. Li, S. Zhu, J. Zou, H. Tian, Y. Duan, Y. Wang, J. Fei and X. Wang, *Chem. Comm.*, 2020, **56**, 7601-7604.
2. X. Zhou, F. Su, H. Lu, P. Senechal-Willis, Y. Tian, R. H. Johnson and D. R. Meldrum, *Biomaterials*, 2012, **33**, 171-180.