Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.
This journal is © The Royal Society of Chemistry 2018

Electronic Supplementary Information (ESI) for

The discovery of a freezing-induced peptide ligation during the total chemical synthesis of human interferon-ε

Yin–He Yang, Bin Di* and Da–Song Yang*

Jiang Su Key Laboratory of Drug Design and Optimization, Key Laboratory on Protein Chemistry and Structural Biology, China Pharmaceutical University, Nanjing 210009, China.

* Corresponding authors. *E-mail address*: dibin@cpu.edu.cn (B. Di); yangds@cpu.edu.cn (D.-S. Yang)

Contents of Supplementary Information

Detailed experimental procedures

1. General details

Aminomethyl resin (loading: 0.52 mmol/g) was purchased from Tianjin Nankai Hecheng Science & Technology Co., Ltd (Tianjin, China). Boc-*L*-Arg(Tos)-O-CH_2-phi-CH_2-COOH•DCHA and Boc-*L*-Leu-O-CH_2-phi-CH_2-COOH were obtained from PolyPeptide Group (Strasbourg, France). Standard Boc-protected amino acids: Boc-Ala-OH, Boc-Leu-OH•H_2O, Boc-Phe-OH, Boc-Gly-OH, Boc-Met-OH, Boc-Ile-OH, Boc-Val-OH, Boc-Pro-OH, Boc-Gln-OH, Boc-Arg(Mts)-OH, Boc-Glu(OcHex)-OH, Boc-Asp(OcHex)-OH, Boc-Asn(Xan)-OH, Boc-Thz-OH, Boc-Cys(pMeBzl)-OH, Boc-Cys(Acm)-OH, Boc-Lys(2-cholro-Z)-OH, Boc-His(3-Bom)-OH, Boc-Thr(Bzl)-OH, Boc-Tyr(2-Br-Z)-OH, Boc-Ser(Bzl)-OH, 3-tritylmercapto-propionic acid, 1-[bis(dimethylamino)methylene]-1H-benzotriazoliumhexafluorophosphate 3-oxide (HBTU), 1-hydroxybenzotriazole, anhydrous (HOBt), 4-dimethylaminopyridine (DMAP) and *N*,*N*-diisopropylethylamine (DIEA) were purchased from GL Biochem Ltd (Shanghai, China). Boc-Trp(Hoc)-OH was purchased from Peptide Institute, Inc. (Osaka, Japan). Trifluoroacetic acid (TFA), *N*,*N*′-diisopropylcarbodiimide (DIC), sodium hydroxide (NaOH), methoxyamine hydrochloride, disodium hydrogen phosphate dodecahydrate ($Na_2HPO_4 \cdot 12H_2O$), sodium dihydrogen phosphate dihydrate ($NaH_2PO_4 \cdot 2H_2O$) and *p*-cresol were obtained from Aladdin Industrial Corporation (Shanghai, China). 4-Mercaptophenylacetic acid (MPAA), tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and guanidine hydrochloride (Gu•HCl) were purchased from Sigma (St. Louis, MO, USA). Diethyl ether (Et_2O), hydrochloric acid (HCl), ethyl acetate, dichloromethane (DCM), dimethylformamide (DMF) and acetic anhydride were purchased from Nanjing Chemical Reagent Co., Ltd (Nanjing, China). Anhydrous hydrogen fluoride was purchased from Huantai Huiyong Chemical Products Sales Co., Ltd (Zibo, China). Acetonitrile (ACN) was purchased from Tedia Company, Inc. (OH, USA). Liquid nitrogen was purchase from

Air Products and Chemicals, Inc. (Nanjing, China). Deionized water was purified through PL5242 Purelab Classic UV (PALL Co. Ltd., USA) to a resistivity of 18.2 MΩ cm. Sample was weighed on an analytical balance (BSA 124S, AG, Goettingen, Germany). Peptides were synthesized on a CS Bio CS336X automated Peptide Synthesizer (CS Bio Company, CA, USA); analyzed on an Agilent 1100 HPLC system (UV detector, Agilent Technologies, Palo Alto, CA, USA); purified on a Shimadzu SPD-20A preparative HPLC system (UV detector, Shimadzu Corporation, Kyoto, Japan), freeze-dried on a LGJ-18C lyophilizer (Sihuan Corporation, Beijing, China). Mass spectra were acquired on Agilent 6224 Accurate Mass TOF LC/MS mass spectrometer (Agilent Technologies, Palo Alto, CA, USA). Columns for purification including: XBridge^{TM} Prep C_{18}, 5 μm, OBD^{TM} 30 × 150 mm; XBridge^{TM} Prep C_{18}, 10 μm, OBD^{TM} 50 × 250 mm; Symmetry300^{TM} C_4, 5 μm, OBD^{TM} 19 × 250 mm; Xbridge BEH C_{18}, 5 μm, 4.6 × 150 mm, 300Å; and Xbridge BEH C_4, 3.5 μm, 4.6 × 150 mm, 300Å were purchased from Waters Corporation (MA, USA).

2. Peptide synthesis protocol based on Boc chemistry

1) Weigh 0.25 mmol aminomethyl resin and put into a peptide synthesis vessel, swell it with a mixture solvent (DMF: DCM = 1: 1, 15 mL) for 30 min, and then wash with DCM (5 mL) for twice.
2) Weigh the reactants for the synthesis of PAM resin (molar ratio of Boc-AA-O-CH_2-phi-CH_2-COOH: HOBt: DIC: Resin = 2: 1.2: 1.2: 1).
3) Dissolve HOBt and Boc-AA-O-CH_2-phi-CH_2-COOH in mixture solvent (DCM: DMF = 4: 1, the amount of solvents was depending on solubility), dissolve DIC in about 5 mL DCM. Mix the two solutions together and activate in ice-water bath for 30 min. Pour the solution into peptide synthesis vessel containing swollen resin, shake overnight. Drain the solvents and rinse the resin with DMF.
4) Cap unreacted amine by adding 20 mL solution (acetic anhydride: DIEA: DMF = 2: 1: 7) and shake for 30 min. Drain the solvents and wash the resin with DMF thoroughly.

5) Boc deprotection: Add 15–20 mL neat TFA to the reaction vessel, shake for 3 minutes, then drain the TFA and repeat once. Rinse the resin with DMF thoroughly.

6) Coupling [a double coupling strategy is needed for sterically hindered amino acid derivatives, such as, Boc-Arg(Mts)-OH, Boc-Ile-OH, Boc-Glu(OcHex)-OH]: weigh the reactants (molar ratio of amino acid: HBTU: DIEA: resin = 8: 7.5: 20 :1). Dissolve amino acid and HBTU in about 20 mL DMF and then the DIEA was dropwise added into the solution. Shake it for 3 min. Add the HBTU activated amino acid into reaction vessel, shake for 30 min in a water bath at 40℃. Drain the solvents and wash the resin with DMF (For Boc-Gln-OH, use DCM) thoroughly.

7) Repeat steps 5 and 6 to couple on all of the amino acids in order until the sequence is completed. Remove Boc protection of the N-terminal amino acid. Wash the resin thoroughly with DMF, then with DCM/MeOH. Dry the resin under vacuum.

8) HF cleavage: 2 g *p*-cresol, 30 mL HF was added to the resin and then stir in ice water bath for 1 h. After concentrated in vacuo, the ice-cold diethyl ether was added to precipitate the crude peptide products, which were further dissolved with mixture solvent (ACN: H_2O = 1: 1, 0.1% TFA) and then remove resin through filtration.

9) The crude peptides were lyophilized, and purified by preparative HPLC and characterized by ESI-MS.

3. Condensation of peptide fragments based on native chemical ligation

1) Standard NCL ligation buffer: 6 M Gu•HCl, 0.2 M Na_2HPO_4, 50 mM MPAA, 20 mM TCEP.

2) Kinetically controlled ligation (KCL) buffer: 6 M Gu•HCl, 0.2 M Na_2HPO_4, 20 mM TCEP.

3) Unmasking of Thz: 0.2 M $MeONH_2$•HCl, 6 M Gu·HCl, 0.2 M Na_2HPO_4, pH = 4.0, stir for about 4–5 h.

4) Ligation 1: peptide fragments (Fr.1: Fr.2 = 2: 1) were dissolved in NCL ligation buffer, adjust the pH to 8.5 and put in into refrigerator (–20℃). After the ligation was finished, Thz at the N-terminal of sequence was removed.

5) Ligation 2: peptide fragments (Fr.3: product of ligation 1 = 2: 1) were dissolved in NCL ligation buffer, adjust the pH to 7.1 and stir at room temperature. Thz at the N-terminal of product was removed after the finish of ligation.

6) Ligation 3: peptide fragments (Fr.4: Fr.5 = 1: 1) were dissolved in KCL ligation buffer, adjust the pH to 7.1 and stir at room temperature.

7) Ligation 4: peptide fragments (product of ligation 2: product of ligation 3 = 1: 1) were dissolved in NCL ligation buffer, adjust the pH to 7.1 and stir at room temperature.

8) To monitor the ligation reaction, absorb 2 μL of the reaction mixture and quench it by adding 200 μL of solvent (ACN: H_2O = 1: 1, 0.1% TFA). Analyze the sample by analytical HPLC and ESI-MS.

9) After the ligation was finished, stop the reaction by decreasing the pH of reaction buffer (less than 4.0) with 1 M HCl. The crude products were purified by preparative HPLC and characterized by ESI-MS.

10) HPLC analytical system and conditions: Agilent 1100 HPLC, solvent A (water, 0.1% TFA) and solvent B (ACN, 0.1% TFA). The gradient elution was: 0.00–30.00 min, linear from 5 to 60% B. Detector wavelength: 214 nm. The column temperature: 35 ℃, injection volume: 10 μL. The flow rate was 1 mL/min.

4. The ESI-MS characterization results of synthesized peptides

Fr.1, amino acid sequence: **C**IVQVEISR**C(Acm)** LFFVFSLTEK LSKQGRPLND MKQELTTEFR SPR, theoretical mass: 5146.83, multiple charge states: $[M + 4H]^{+4}$ = 1287.72, $[M + 5H]^{+5}$ = 1030.37, $[M + 6H]^{+6}$ = 858.81, $[M + 7H]^{+7}$ = 736.26, $[M + 8H]^{+8}$ = 644.35, which were consistent with measured masses (Figure S2).

Fr.2, amino acid sequence: **C(Thz)**EKLSGTLGS DNLRLQVKMY FRRIHDYLEN QDYST**C(Acm)**AW-**Mpa-L**, theoretical mass: 4839.34, multiple charge states: [M +

3H]$^{+3}$ = 1614.12, [M + 4H]$^{+4}$ = 1210.84, [M + 5H]$^{+5}$ = 968.87, [M + 6H]$^{+6}$ = 807.56, which were consistent with measured masses (Figure S4).

Fr.3, amino acid sequence: **C(Thz)**NISLDGWEE NHTEKFLIQL HQQLEYLEAL MGLE-**Mpa-L**, theoretical mass: 4258.85, multiple charge states: [M + 3H]$^{+3}$ = 1420.62, [M + 4H]$^{+4}$ = 1065.72, [M + 5H]$^{+5}$ = 852.77, which were consistent with measured masses (Figure S6).

Fr.4, amino acid sequence: **C(Thz)**HRKNFLLPQ KSLSPQQYQK GHTLAILHEM LQQIFSLFR-**Mpa-L**, theoretical mass: 4893.81, multiple charge states: [M + 4H]$^{+4}$ = 1224.46, [M + 5H]$^{+5}$ = 979.76, [M + 6H]$^{+6}$ = 816.63, [M + 7H]$^{+7}$ = 700.11, which were consistent with measured masses (Figure S8).

Fr.4 (unmasking of Thz), amino acid sequence: **C**HRKNFLLPQ KSLSPQQYQK GHTLAILHEM LQQIFSLFR-**Mpa-L**, theoretical mass: 4881.79, multiple charge states: [M + 3H]$^{+3}$ = 1628.27, [M + 4H]$^{+4}$ = 1221.46, [M + 5H]$^{+5}$ = 977.36, [M + 6H]$^{+6}$ = 814.63, [M + 7H]$^{+7}$ = 698.40, which were consistent with measured masses (Figure S10).

Fr.5, amino acid sequence: LDLKLIIFQQ RQVNQESLKL LNKLQTLSIQ Q**C(Acm)**L-**Mpa-L**, theoretical mass: 4182.86, multiple charge states: [M + 3H]$^{+3}$ = 1395.29, [M + 4H]$^{+4}$ = 1046.72, [M + 5H]$^{+5}$ = 837.57, which were consistent with measured masses (Figure S12).

Fr.5 (after MPAA exchanging), amino acid sequence: LDLKLIIFQQ RQVNQESLKL LNKLQTLSIQ Q**C(Acm)**L-**MPAA**, theoretical mass: 4131.58, multiple charge states: [M + 3H]$^{+3}$ = 1378.20, [M + 4H]$^{+4}$ = 1033.90, [M + 5H]$^{+5}$ = 827.32, which were consistent with measured masses (Figure S14).

Peptide 1, amino acid sequence: CIVQVEISR**C(Acm)**L, theoretical mass: 1333.45, multiple charge states: [M + H]$^{+}$ = 1334.46, [M + 2H]$^{+2}$ = 667.73, which were consistent with measured masses (Figure S15).

Peptide 2, amino acid sequence: ENQDYST**C(Acm)**AW-**Mpa-L**, theoretical mass: 1488.44, multiple charge states: [M + H]$^{+}$ = 1489.45, [M + 2H]$^{+2}$ = 745.23, which were consistent with measured masses (Figure S16).

Ligation product between peptide 1 and peptide 2: amino acid sequence:

ENQDYST**C(Acm)**AW CIVQVEISR**C(Acm)** L, theoretical mass: 2602.59, multiple charge states: $[M + 2H]^{+2}$ = 1302.30, $[M + 3H]^{+3}$ = 868.54, which were consistent with measured masses (Figure S17).

Product of ligation 1, amino acid sequence: **C(Thz)**EKLSGTLGS DNLRLQVKMY FRRIHDYLEN QDYST**C(Acm)**AWCI VQVEISR**C(Acm)**LF FVFSLTEKLS KQGRPLNDMK QELTTEFRSP R, theoretical mass: 9766.87, multiple charge states: $[M + 7H]^{+7}$ = 1396.27, $[M + 8H]^{+8}$ = 1221.86, $[M + 9H]^{+9}$ = 1086.21, $[M + 10H]^{+10}$ = 977.69, $[M + 11H]^{+11}$ = 888.90, $[M + 12H]^{+12}$ = 814.91, which were consistent with measured masses (Figure S23).

Product of ligation 1 (unmasking of Thz), amino acid sequence: **C**EKLSGTLGS DNLRLQVKMY FRRIHDYLEN QDYST**C(Acm)**AWCI VQVEISR**C(Acm)**LF FVFSLTEKLS KQGRPLNDMK QELTTEFRSP R, theoretical mass: 9754.85, multiple charge states: $[M + 7H]^{+7}$ = 1394.55, $[M + 8H]^{+8}$ = 1220.36, $[M + 9H]^{+9}$ = 1084.87, $[M + 10H]^{+10}$ = 976.48, $[M + 11H]^{+11}$ = 887.80, $[M + 12H]^{+12}$ = 813.90, $[M + 13H]^{+13}$ = 751.37, which were consistent with measured masses (Figure S25).

Product of Ligation 2, amino acid sequence: **C(Thz)**NISLDGWEE NHTEKFLIQL HQQLEYLEAL MGLECEKLSG TLGSDNLRLQ VKMYFRRIHD YLENQDYST**C(Acm)** AWCIVQVEIS R**C(Acm)**LFFVFSLT EKLSKQGRPL NDMKQELTTE FRSPR, theoretical mass: 13794.39, multiple charge states: $[M + 10H]^{+10}$ = 1380.44, $[M + 11H]^{+11}$ = 1255.04, $[M + 12H]^{+12}$ = 1150.53, $[M + 13H]^{+13}$ = 1062.11, $[M + 14H]^{+14}$ = 986.31, $[M + 15H]^{+15}$ = 920.63, $[M + 16H]^{+16}$ = 863.15, $[M + 17H]^{+17}$ = 812.43, which were consistent with measured masses (Figure S27).

Product of Ligation 3, amino acid sequence: LDLKLIIFQQ RQVNQESLKL LNKLQTLSIQ Q**C(Acm)**LCHRKNFL LPQKSLSPQQ YQKGHTLAIL HEMLQQIFSL FR-**Mpa-L**, theoretical mass: 8845.35, multiple charge states: $[M + 6H]^{+6}$ = 1475.22, $[M + 7H]^{+7}$ = 1264.62, $[M + 8H]^{+8}$ = 1106.67, $[M + 9H]^{+9}$ = 983.82, $[M + 10H]^{+10}$ = 885.53, $[M + 11H]^{+11}$ = 805.12, $[M + 12H]^{+12}$ = 738.11, $[M + 13H]^{+13}$ = 681.41, which were consistent with measured masses (Figure S29).

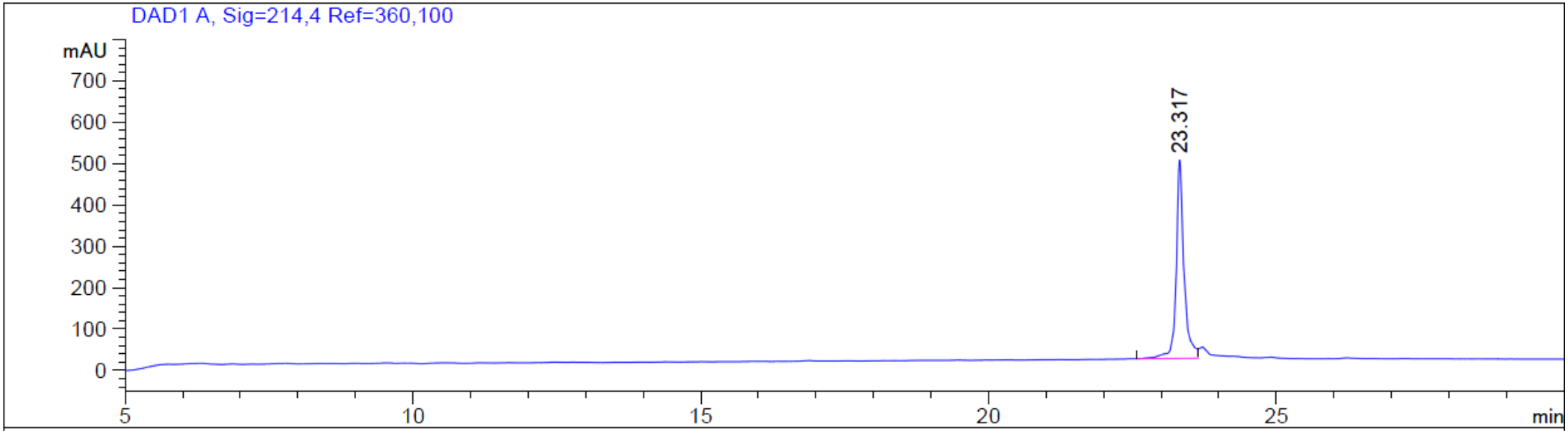

Figure S1. The liquid chromatogram of synthesized peptide Fr.1 of hIFN-ε

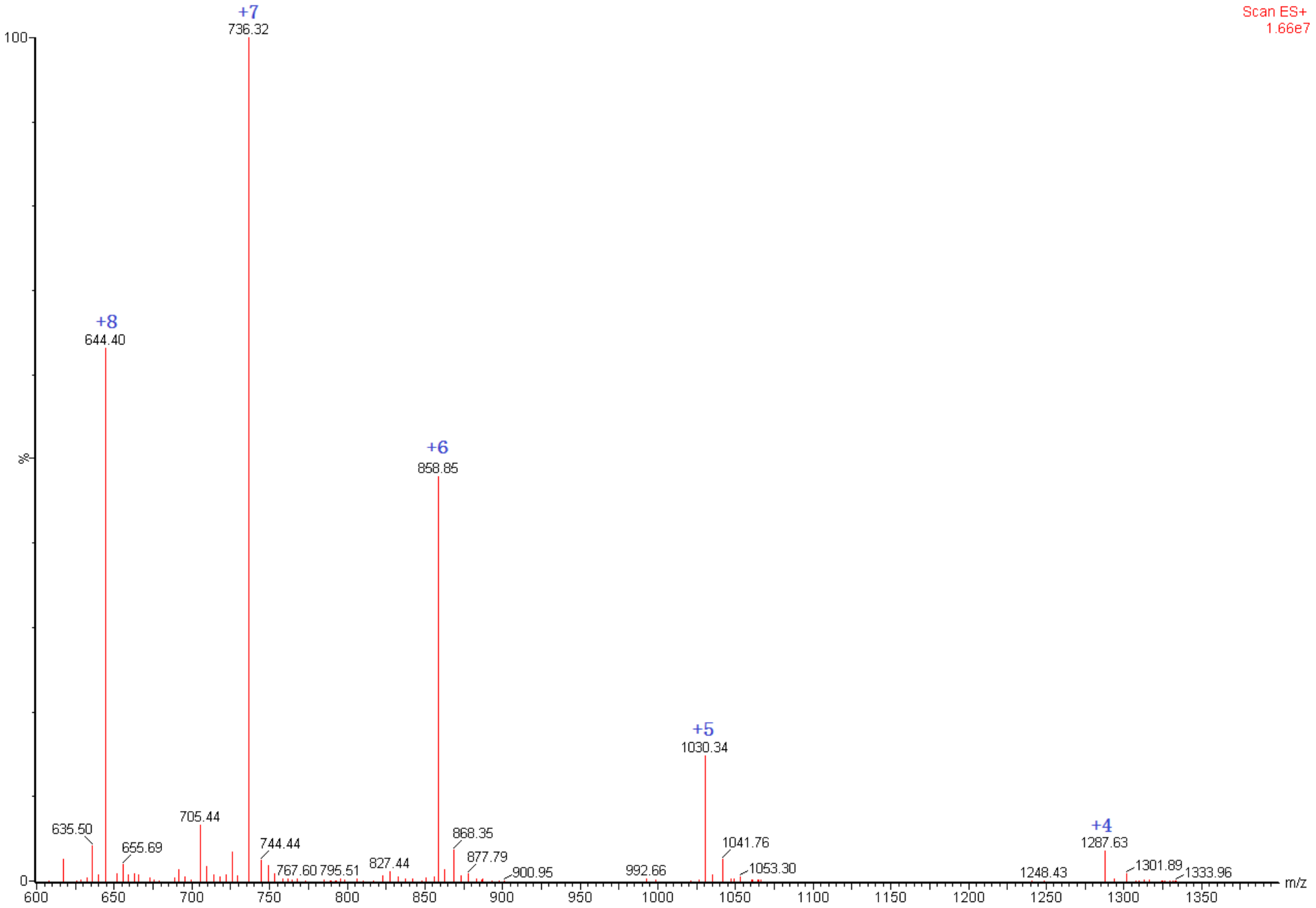

Figure S2. The ESI-MS spectra of synthesized peptide Fr.1 of hIFN-ε

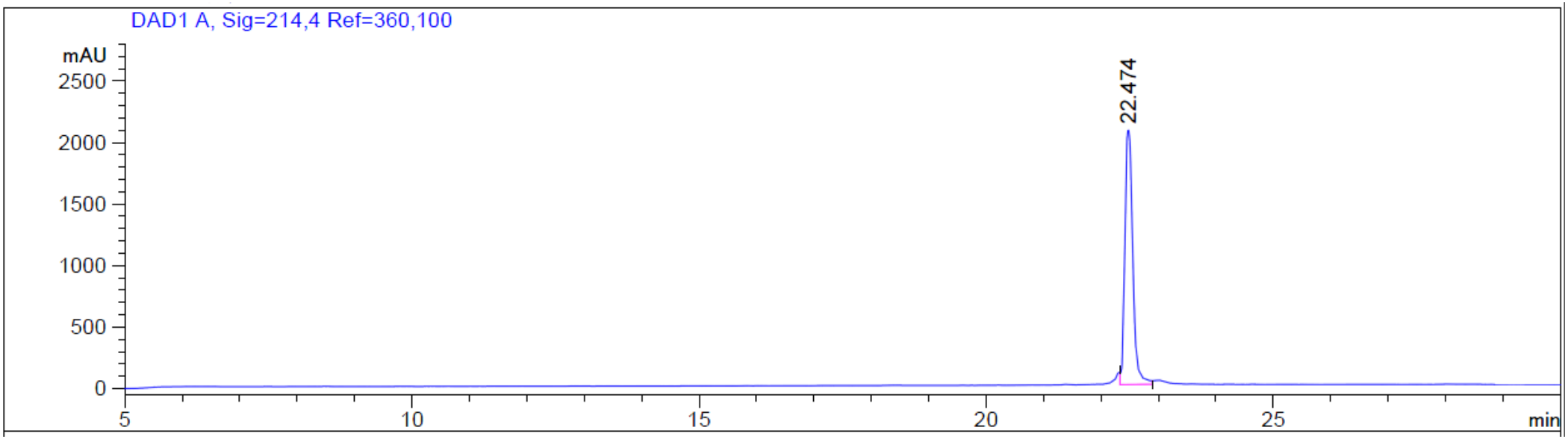

Figure S3. The liquid chromatogram of synthesized peptide Fr.2 of hIFN-ε

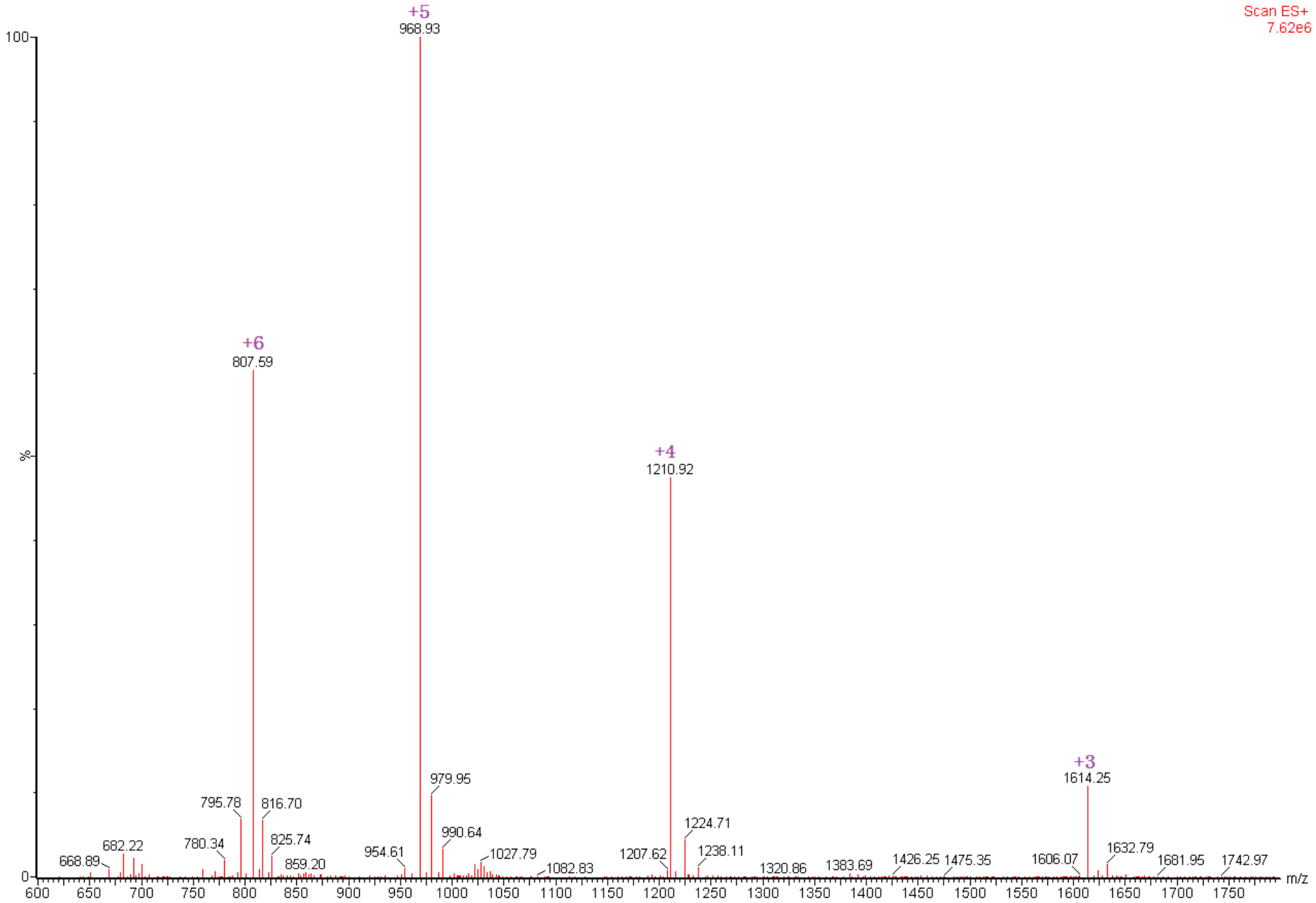

Figure S4. The ESI-MS spectra of synthesized peptide Fr.2 of hIFN-ε

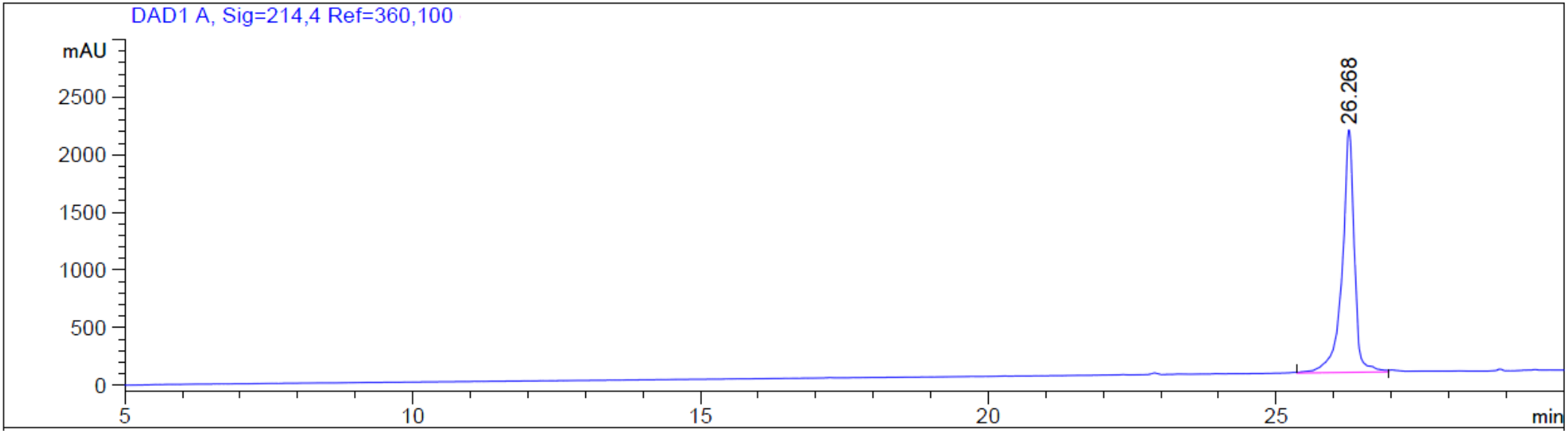

Figure S5. The liquid chromatogram of synthesized peptide Fr.3 of hIFN-ε

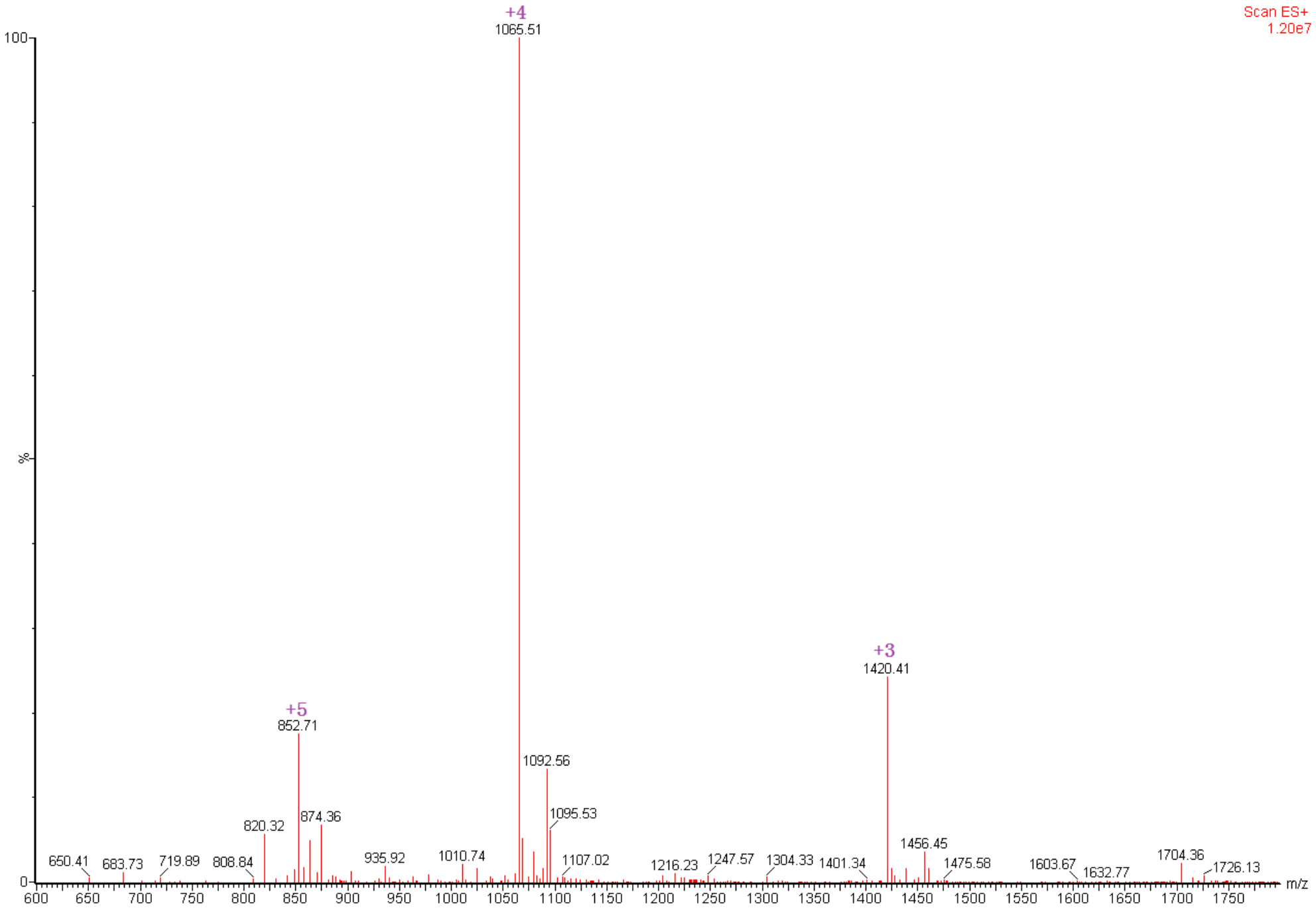

Figure S6. The ESI-MS spectra of synthesized peptide Fr.3 of hIFN-ε

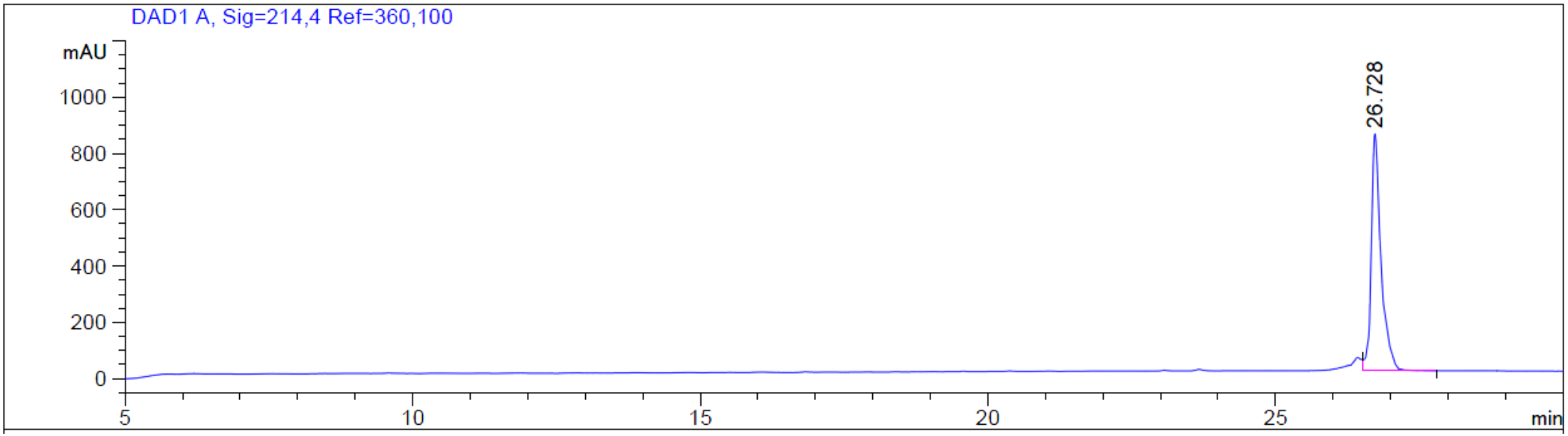

Figure S7. The liquid chromatogram of synthesized peptide Fr.4 of hIFN-ε

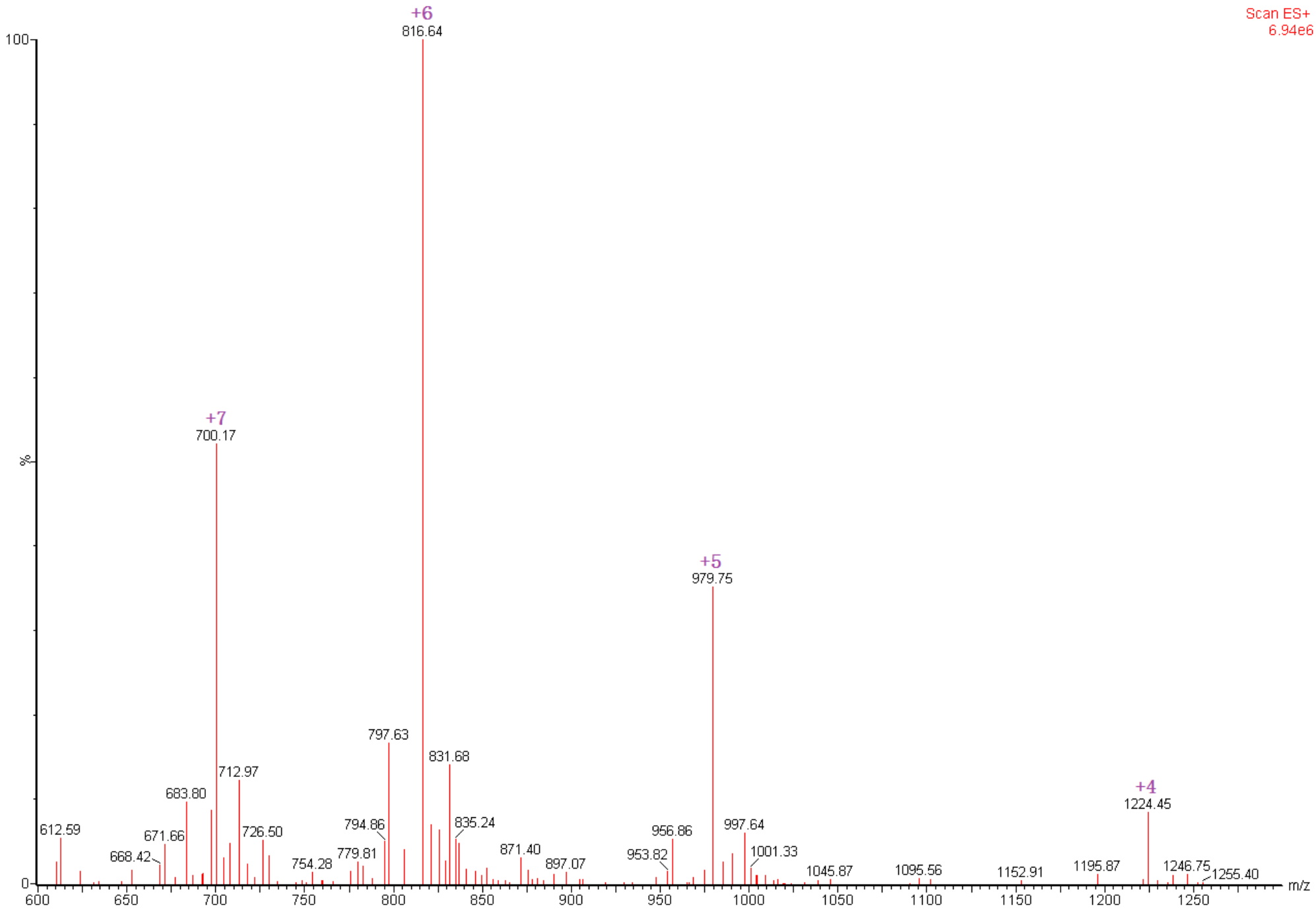

Figure S8. The ESI-MS spectra of synthesized peptide Fr.4 of hIFN-ε

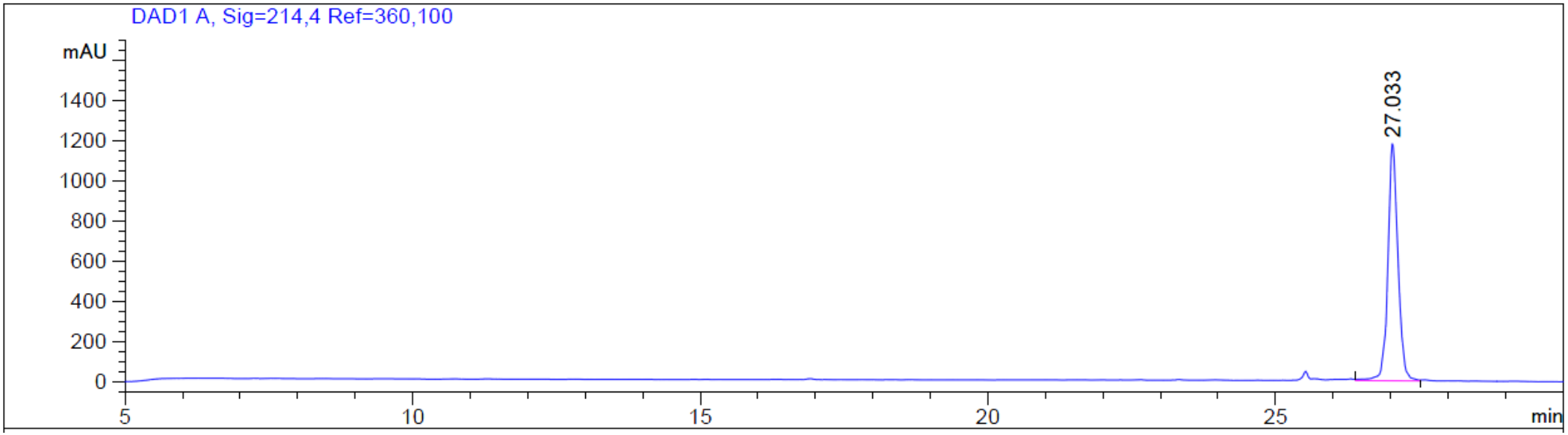

Figure S9. The liquid chromatogram of synthesized peptide Fr.4 (after unmasking of Thz) of hIFN-ε

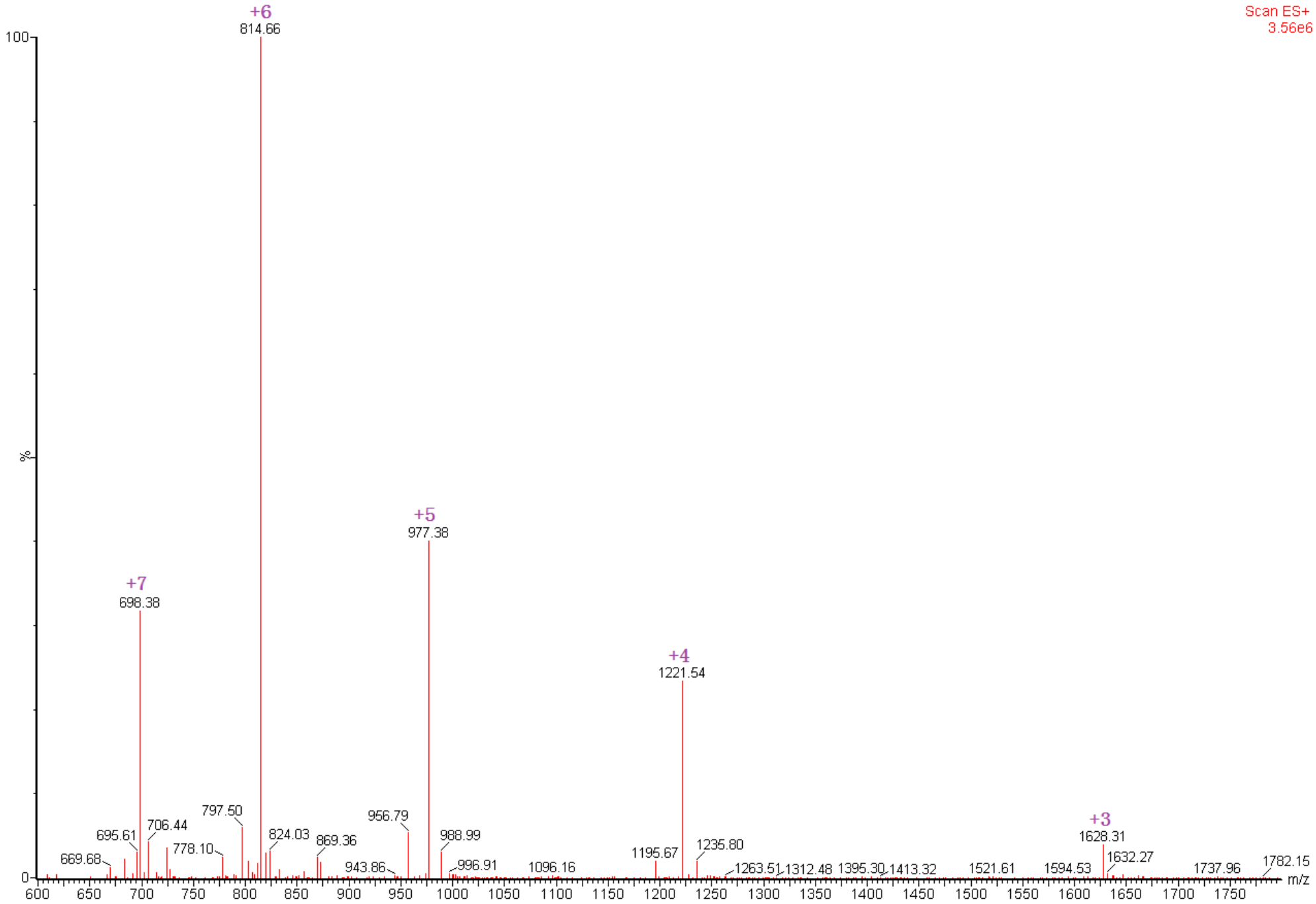

Figure S10. The ESI-MS spectra of synthesized peptide Fr.4 (after unmasking of Thz) of hIFN-ε

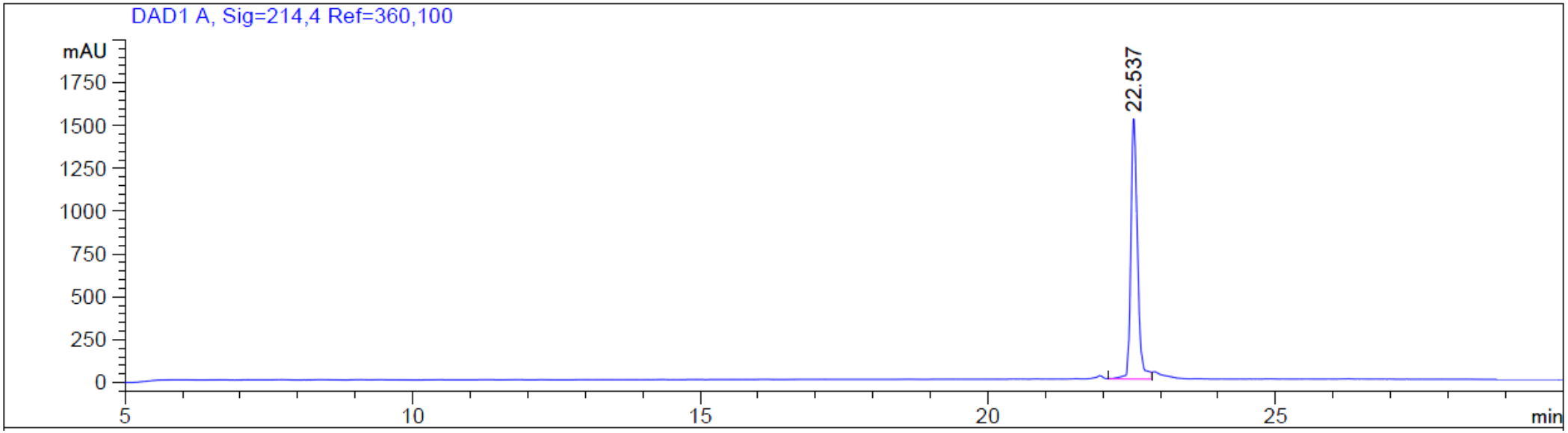

Figure S11. The liquid chromatogram of synthesized peptide Fr.5 of hIFN-ε

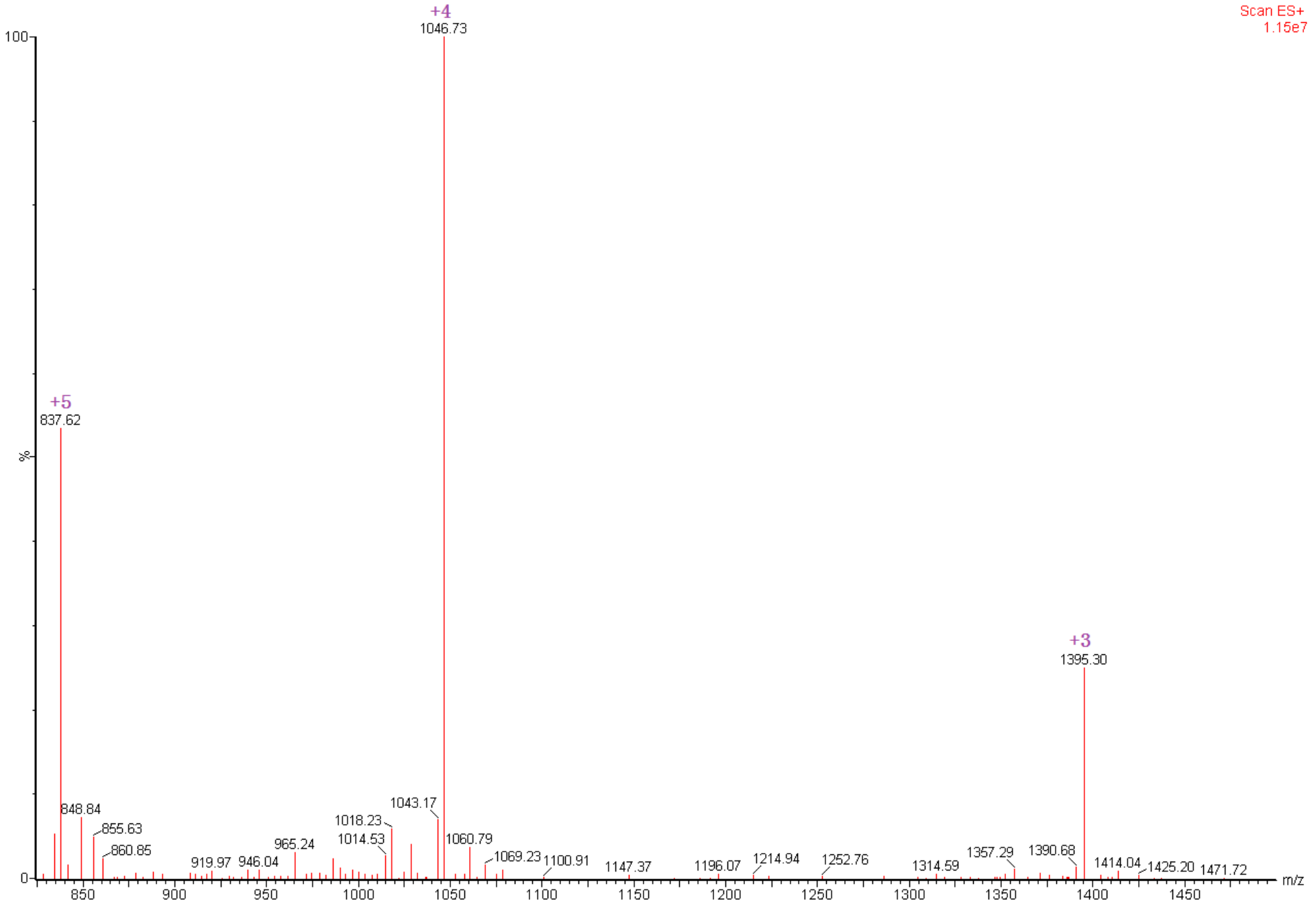

Figure S12. The ESI-MS spectra of synthesized peptide Fr.5 of hIFN-ε

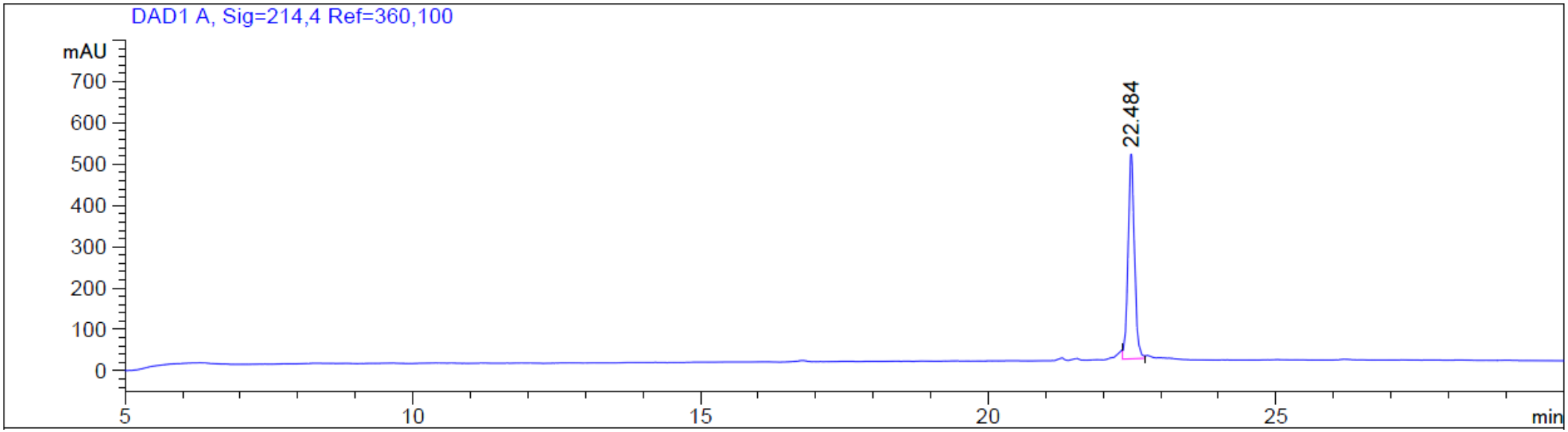

Figure S13. The liquid chromatogram of synthesized peptide Fr.5 (after MPAA exchanging) of hIFN-ε

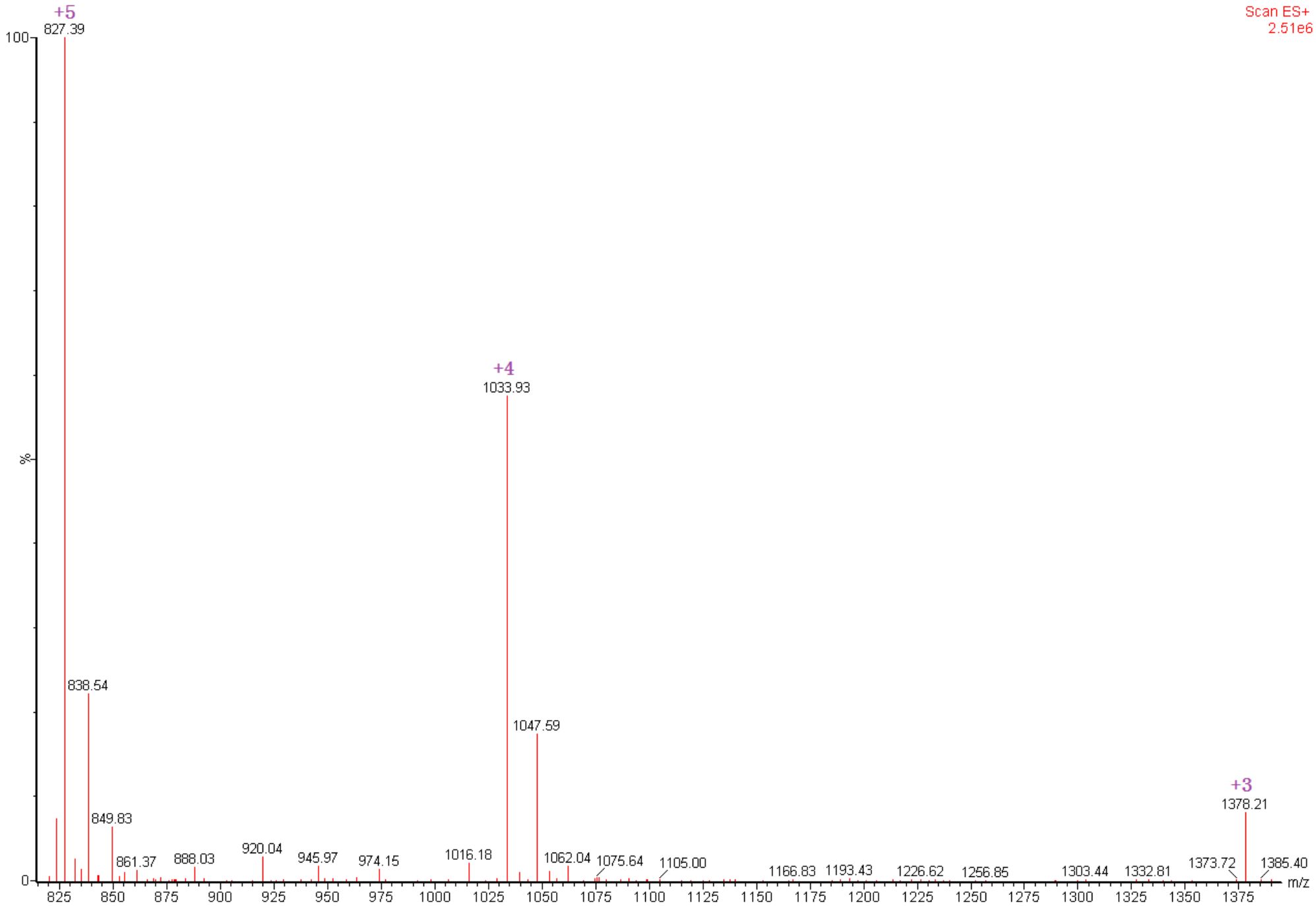

Figure S14. The ESI-MS spectra of synthesized peptide Fr.5 (after MPAA exchanging) of hIFN-ε

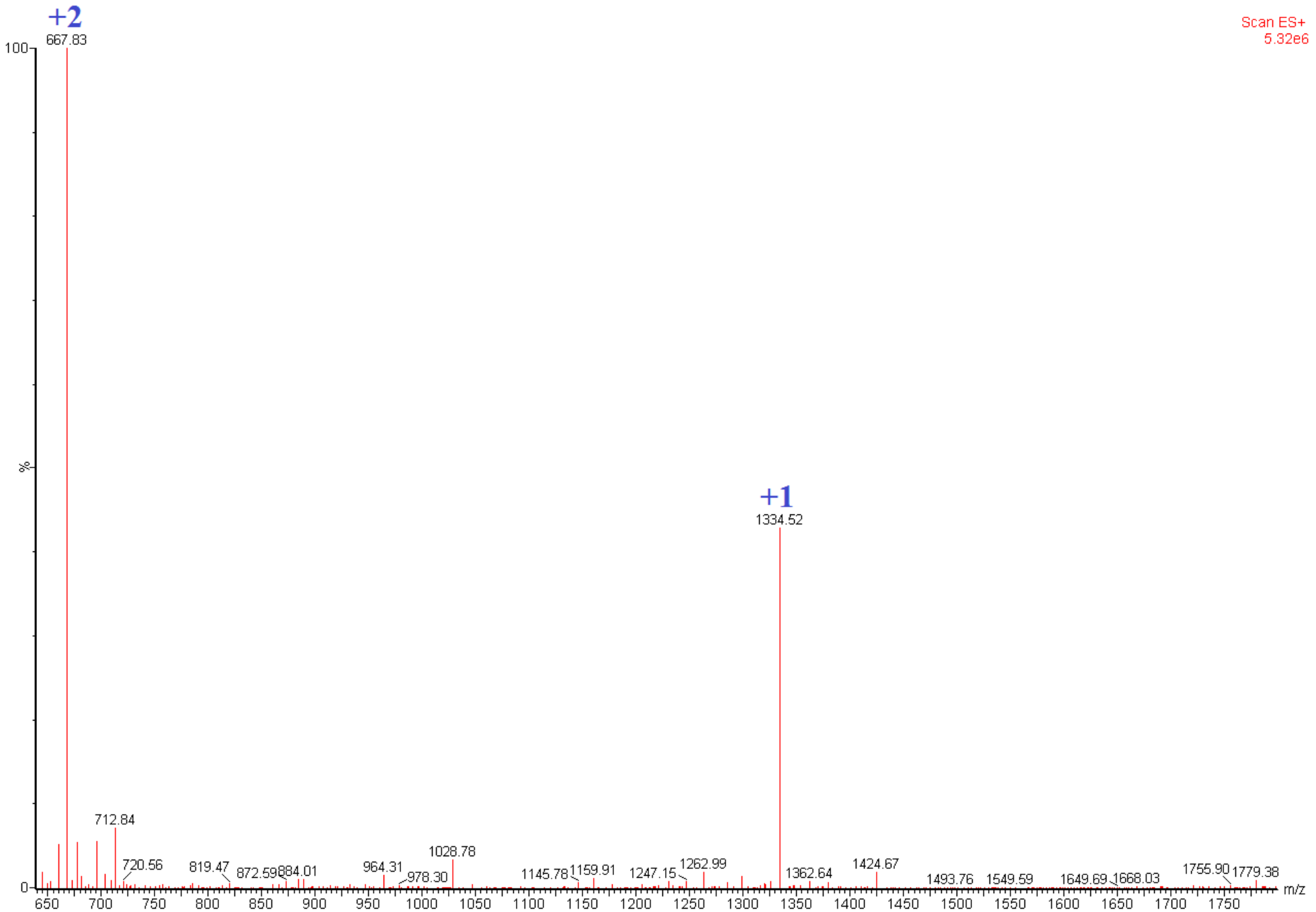

Figure S15. The ESI-MS spectra of peptide 1 [CIVQVEISRC(Acm)L]

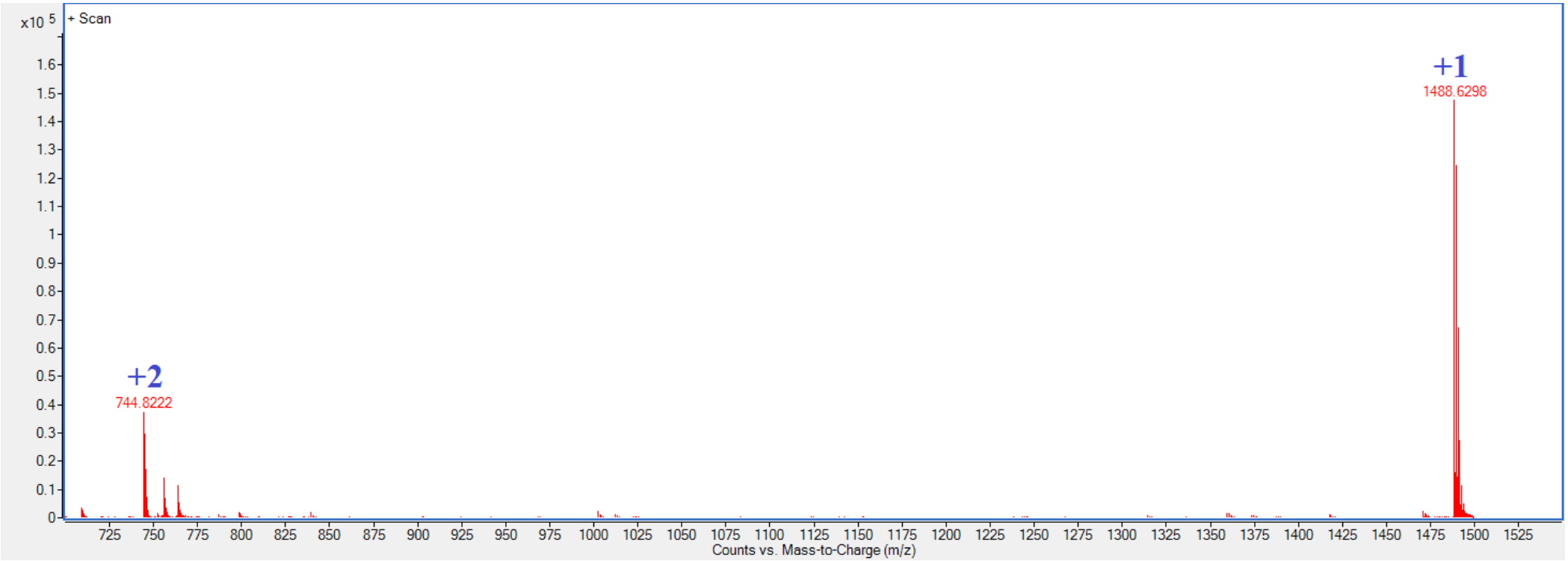

Figure S16. The ESI-MS spectra of peptide 2 [ENQDYSTC(Acm)AW-Mpa-L]

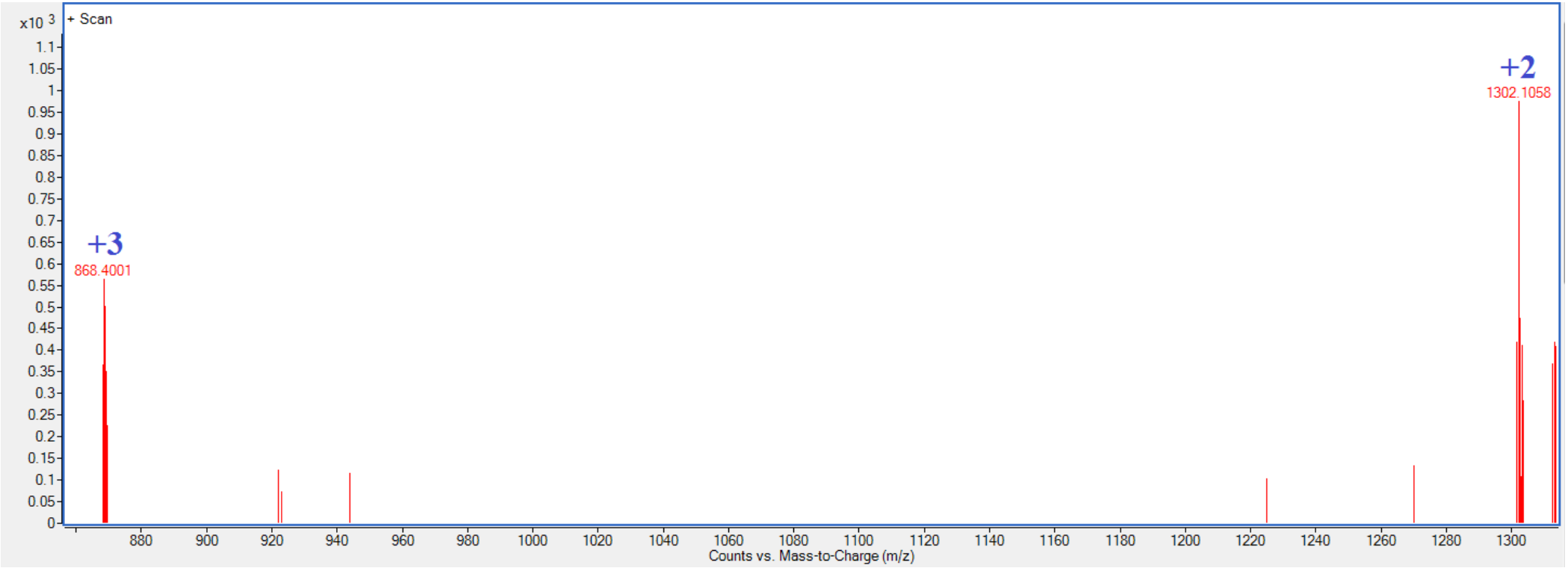

Figure S17. The ESI-MS spectra of ligation product between peptide 1 and peptide 2

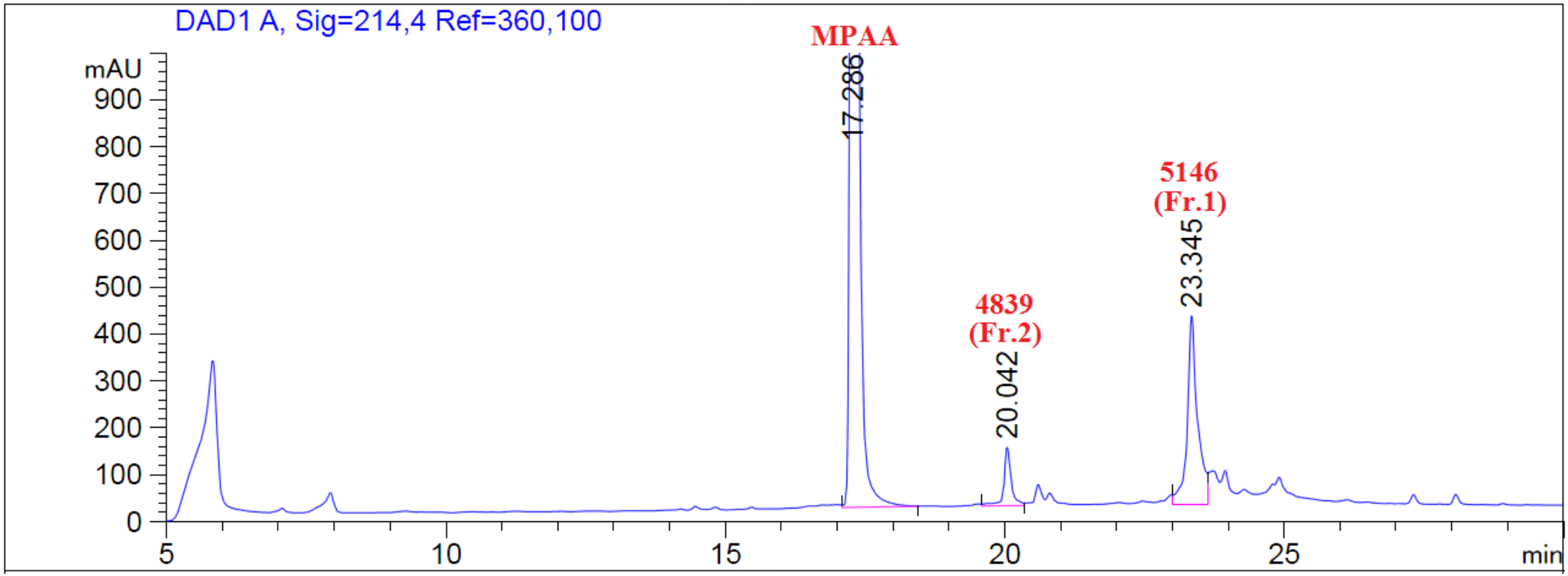

Figure S18. The liquid chromatogram of ligation 1 buffer (0 h)

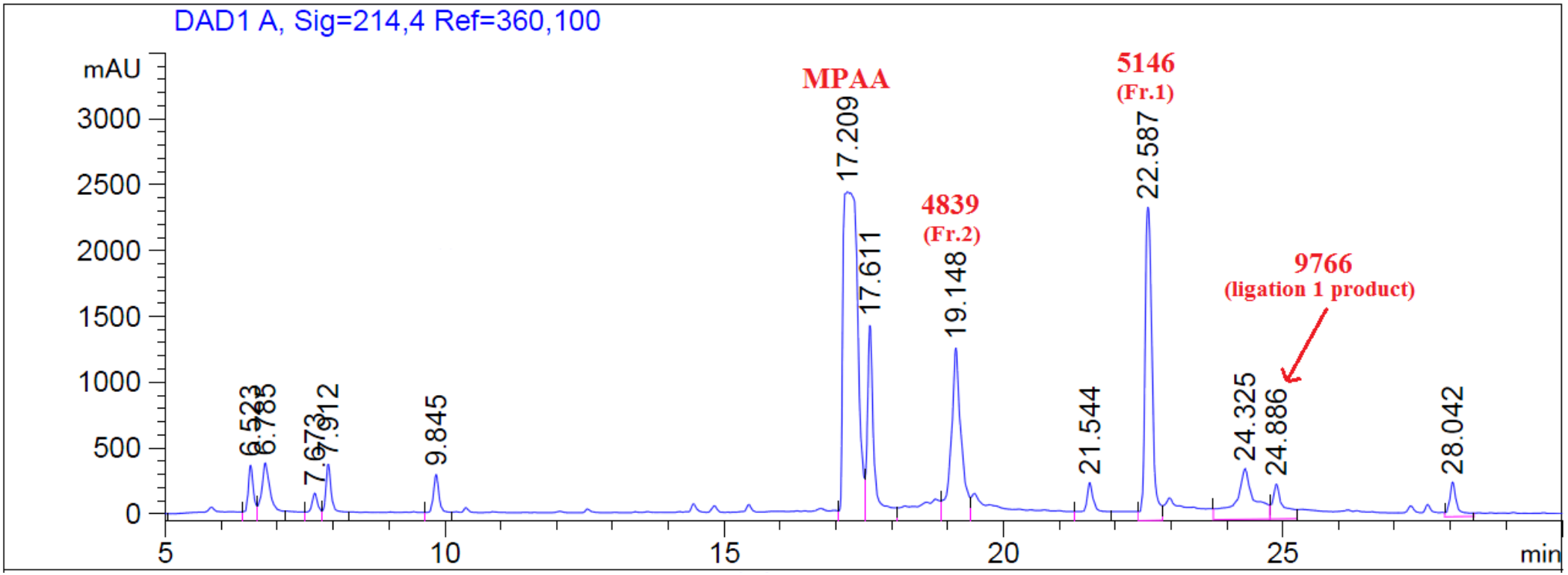

Figure S19. The liquid chromatogram of ligation 1 buffer (pH 8.5, room temperature, 48 h)

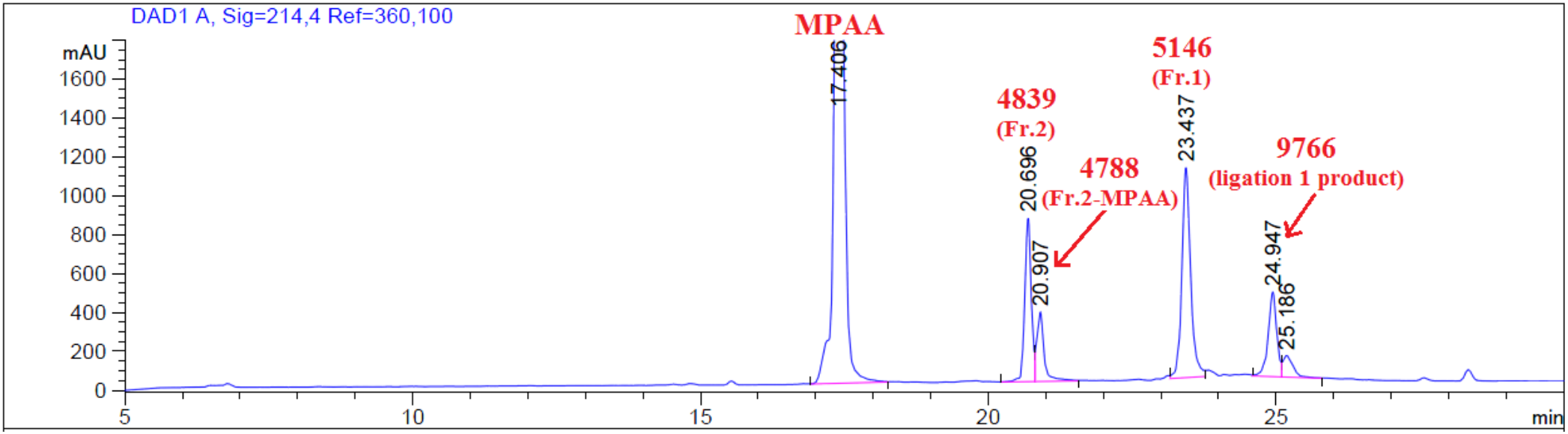

Figure S20. The liquid chromatogram of ligation 1 buffer (pH 8.5, –20℃, 12 h)

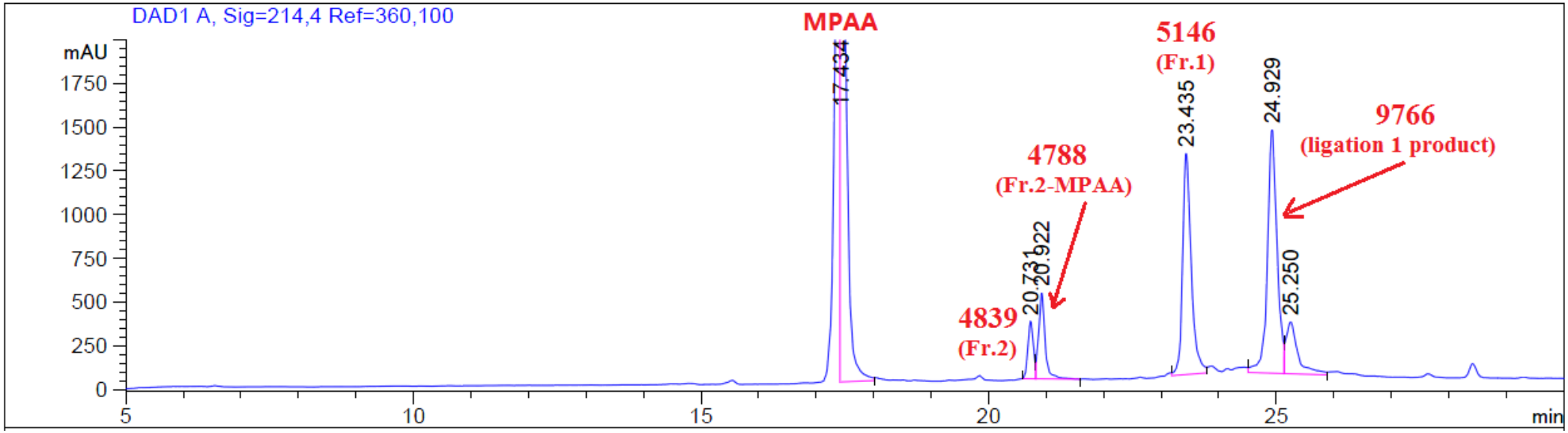

Figure S21. The liquid chromatogram of ligation 1 buffer (pH 8.5, –20℃, 48 h)

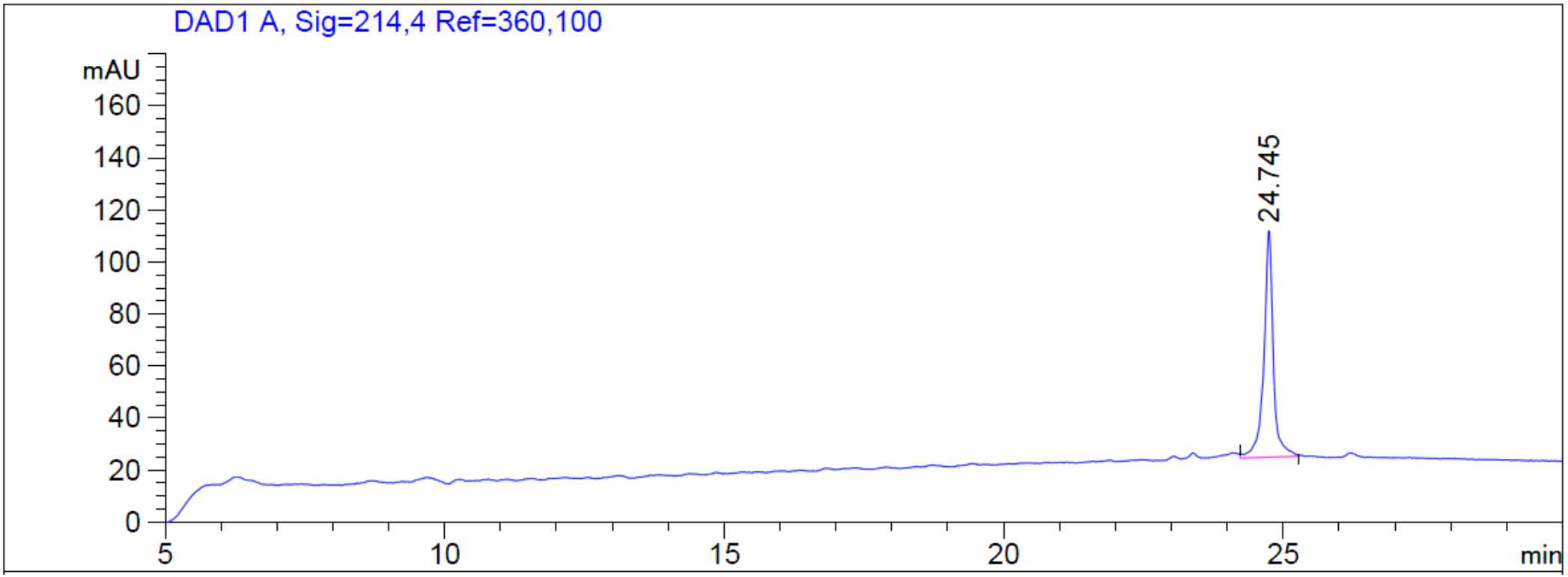

Figure S22. The liquid chromatogram of ligation 1 product

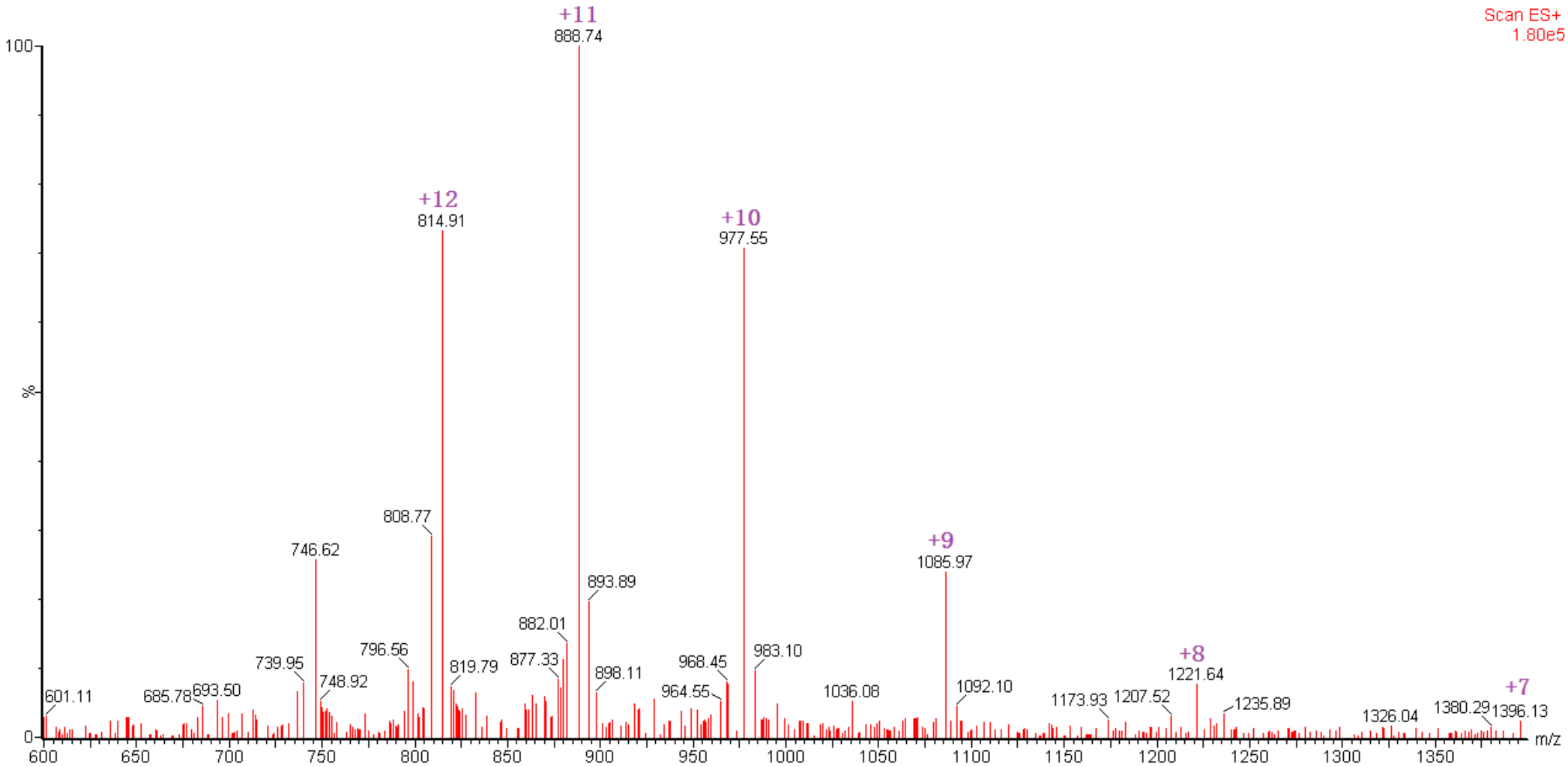

Figure S23. The ESI-MS spectra of ligation 1 product

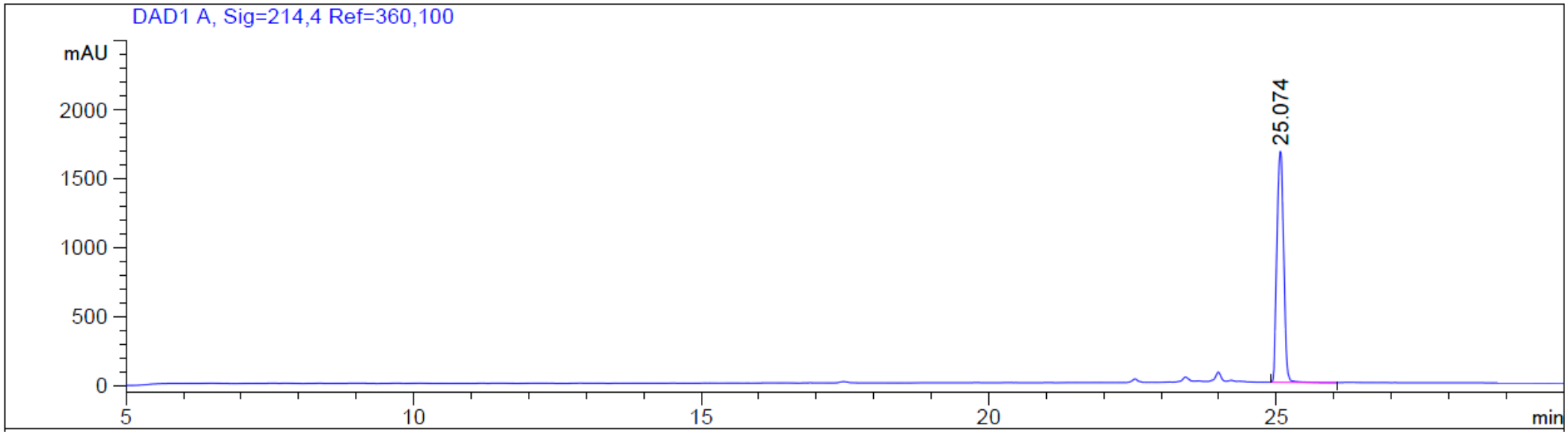

Figure S24. The liquid chromatogram of ligation 1 product (unmasking of Thz)

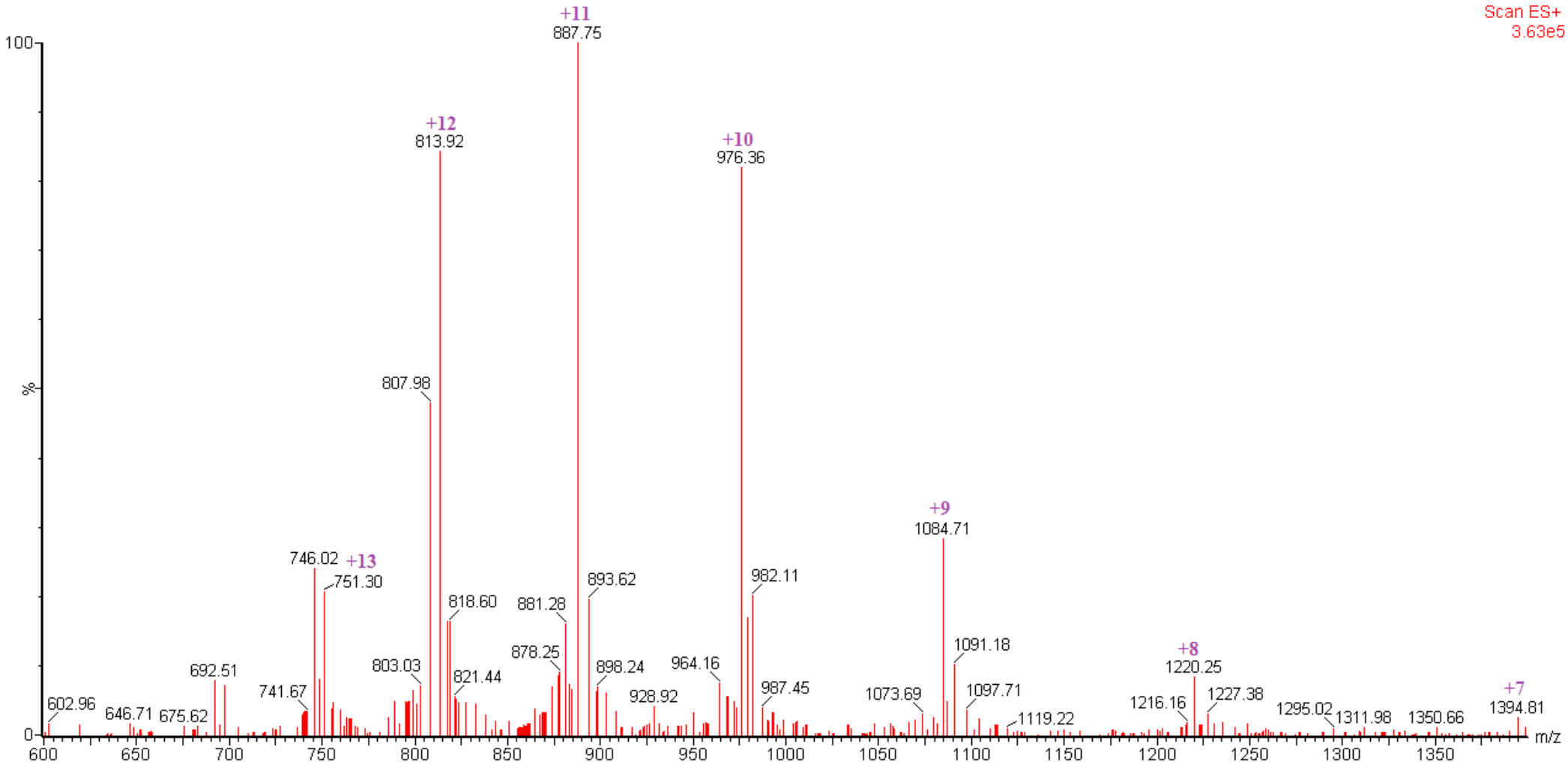

Figure S25. The ESI-MS spectra of ligation 1 product (unmasking of Thz)

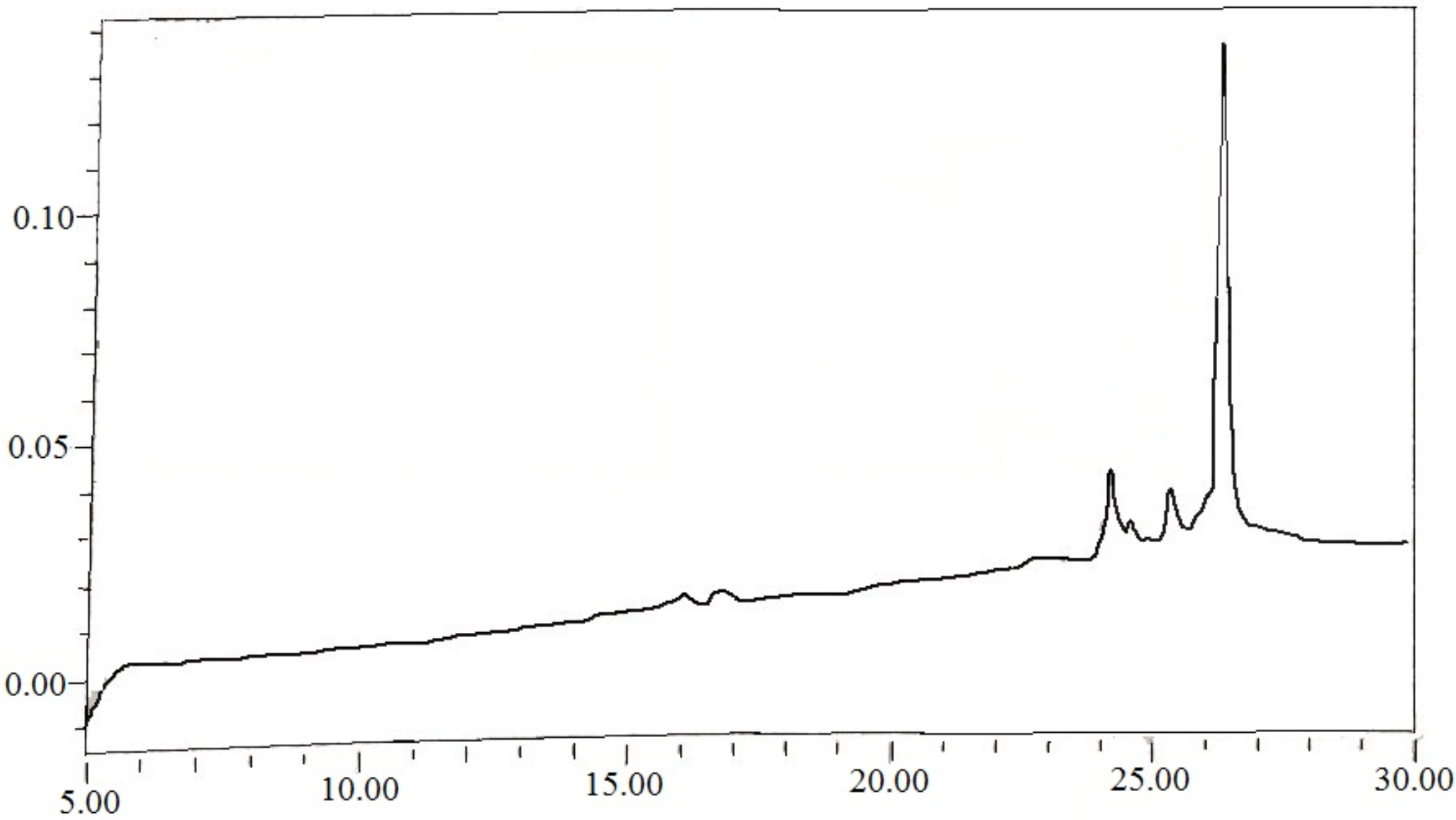

Figure S26. The liquid chromatogram of ligation 2 product

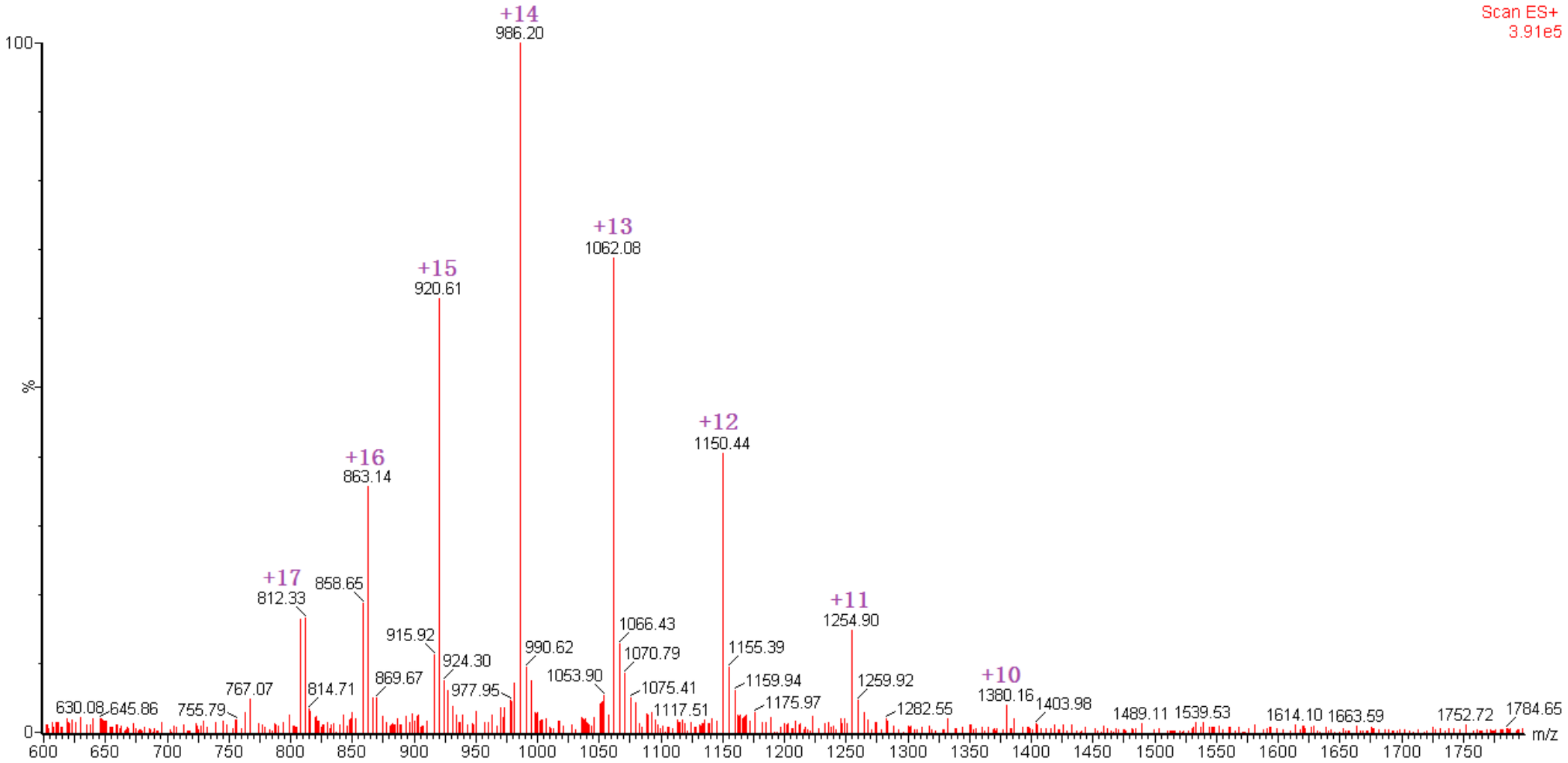

Figure S27. The ESI-MS spectra of ligation 2 product

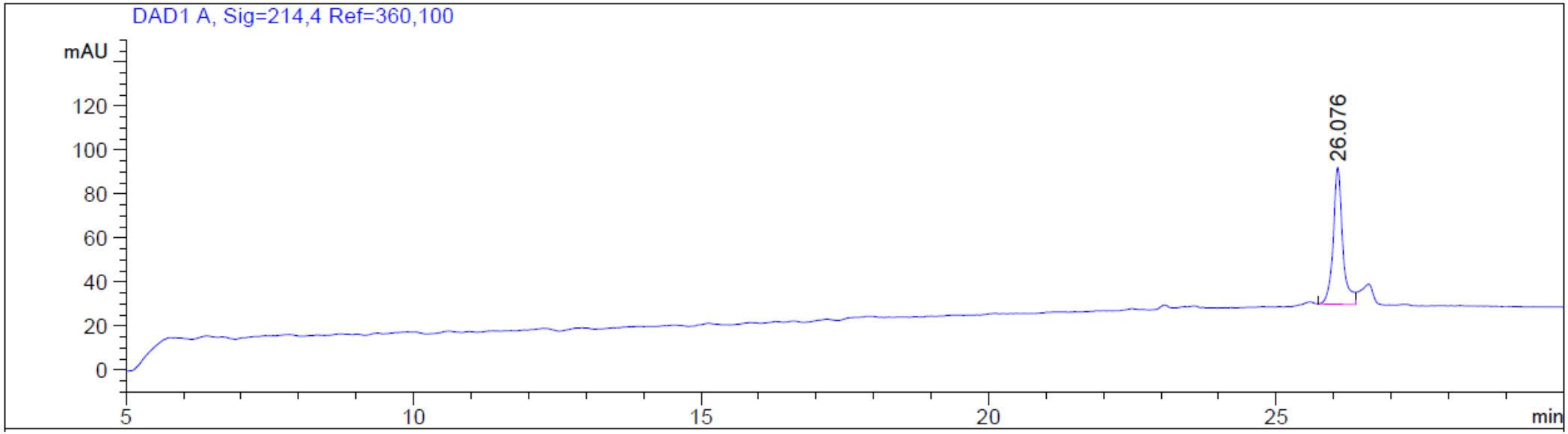

Figure S28. The liquid chromatogram of ligation 3 product

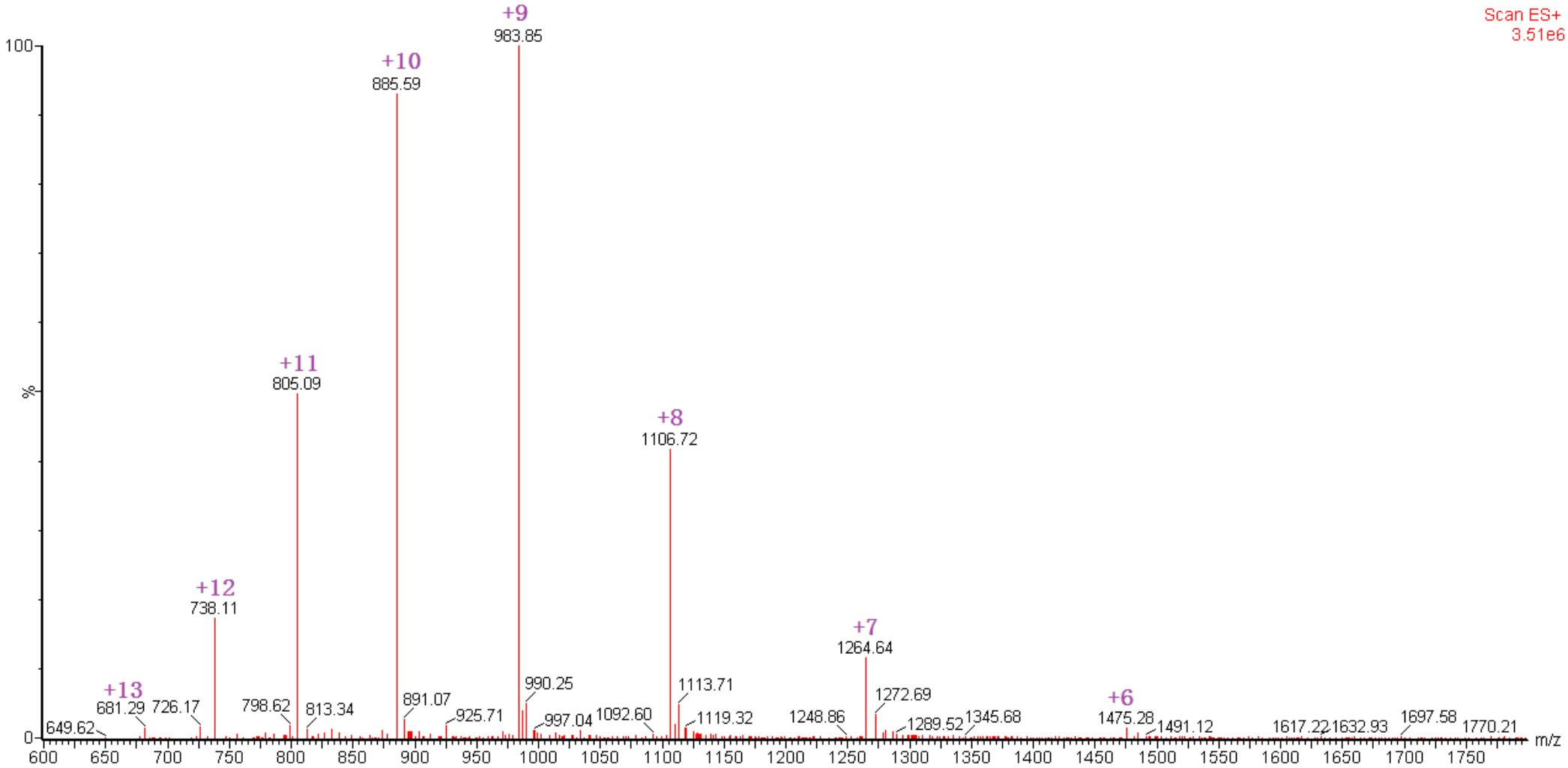

Figure S29. The ESI-MS spectra of ligation 3 product