

Supplementary Table 1.
Data collection and refinement statistics for PDB entry 4OGA

| Human-insulin-bound μ R + Fab 83-7 (PDB 4OGA) | |
|---|------------------------|
| Data collection | |
| Space group | <i>P</i> 23 |
| Cell dimensions | 169.04, 169.04, 169.04 |
| <i>a, b, c</i> (Å) | 90, 90, 90 |
| α, β, γ (°) | 19.78-3.50 (3.60-3.50) |
| Resolution (Å) | 0.214 (14.8) |
| <i>R</i> _{merge} | 12.1 (0.17) |
| <i>I</i> / σ (<i>I</i>) | 0.999 (0.123) |
| <i>CC</i> _{1/2} | not reported |
| Completeness (%) | 38.7 (16.7) |
| Redundancy | 1 |
| Refinement | |
| Resolution (Å) | 19.78-3.50 |
| Resolution (Å) | 19092 |
| No. reflections | - |
| <i>R</i> _{work} / <i>R</i> _{free} | 0.264 / 0.284 |
| No. atoms | |
| Protein | 4552 |
| Carbohydrate / ion | 141 |
| <i>B</i> factors | |
| Protein | 237 |
| Carbohydrate / ion | 270 |
| R.m.s. deviations | |
| Bond lengths (Å) | 0.010 |
| Bond angles (°) | 1.3 |

^aValues in parentheses are for highest-resolution shell.

Diffraction intensities were collected from three single crystals and then merged. For further detail, see Menting, JG *et al.*, *Proc Natl Acad Sci U S A* **111**, E3395-E3404 (2014).

Supplementary Table 2

R.m.s.d.'s between and within structures of Con-Ins-G1-bound μ IR and human-insulin-bound μ IR¹

(a) R.m.s.d. (\AA) for C α atoms of Con-Ins G1 (or hIns) residues (A1-A20, B7-B19) after rigid overlay of hIR domain L1 residues 1-150:

| | | Con-Ins-G1-bound μ IR + Fv 83-7 | | Human-insulin-bound μ IR + Fv 83-7 | | | | Human-insulin-bound zippered ectodomain + Fv 83-7 ¹ |
|---------------------------|---------|--|---------|---|---------|---------|---------|--|
| | | Mono. 1 | Mono. 2 | Mono. 1 | Mono. 2 | Mono. 3 | Mono. 4 | |
| ConInsG1. μ IR.Fv83-7 | Mono. 1 | - | 1.4 | 1.3 | 2.2 | 1.4 | 2.1 | 1.7 |
| | Mono. 2 | 1.4 | - | 1.5 | 1.2 | 1.4 | 1.1 | 1.7 |
| hIns. μ IR.Fv83-7 | Mono. 1 | 1.3 | 1.5 | - | 1.6 | 0.3 | 1.4 | 0.8 |
| | Mono. 2 | 2.2 | 1.2 | 1.6 | - | 1.4 | 0.3 | 1.7 |
| | Mono. 3 | 1.4 | 1.4 | 0.3 | 1.4 | - | 1.2 | 0.9 |
| | Mono. 4 | 2.1 | 1.1 | 1.4 | 0.3 | 1.2 | - | 1.7 |

(b) R.m.s.d. (\AA) for C α atoms of Con-Ins G1 (or hIns) residues (A1-A20, B7-B19) after rigid overlay of same residues:

| | | Con-Ins-G1-bound μ IR + Fv 83-7 | | Human-insulin-bound μ IR + Fv 83-7 | | | | Human-insulin-bound zippered ectodomain + Fv 83-7 ¹ |
|---------------------------|---------|--|---------|---|---------|---------|---------|--|
| | | Mono. 1 | Mono. 2 | Mono. 1 | Mono. 2 | Mono. 3 | Mono. 4 | |
| ConInsG1. μ IR.Fv83-7 | Mono. 1 | - | 0.2 | 0.8 | 0.9 | 0.8 | 0.9 | 0.9 |
| | Mono. 2 | 0.2 | - | 0.8 | 0.9 | 0.8 | 0.8 | 0.9 |
| hIns. μ IR.Fv83-7 | Mono. 1 | 0.8 | 0.8 | - | 0.3 | 0.3 | 0.3 | 0.5 |
| | Mono. 2 | 0.9 | 0.9 | 0.3 | - | 0.2 | 0.2 | 0.6 |
| | Mono. 3 | 0.8 | 0.8 | 0.3 | 0.2 | - | 0.2 | 0.6 |
| | Mono. 4 | 0.9 | 0.8 | 0.3 | 0.2 | 0.2 | - | 0.6 |

¹ PDB entry 6HN5 (Weis, F *et al.*, *Nat Commun* **9**, 4420 (2018)).

² Overlay and r.m.s.d. calculations performed using LSQMAN (Kleywegt, GJ & Jones, TA, *CCP4/ESF-EACBM News Protein Crystallogr* **31**, 9-14 (1994)).

³ See Supplementary Table 3 for definition of monomer chains.

Supplementary Table 3:
Modelled residues within the crystal structures of μ IR complexes

| | Con-Ins-G1-bound μ IR + Fv 83-7 | | Human-insulin-bound μ IR + Fv 83-7 | | | |
|----------------------|--|------------------------------|--|-------------------------------------|------------------------------|-------------------------------------|
| | Copy 1 | Copy 2 | Copy 1 | Copy 2 | Copy 3 | Copy 4 |
| A chain | A1-20 | G1-20 | A1-21 | G1-21 | M1-21 | S1-21 |
| B chain ¹ | B4-19 | H4-19 | B1-27 | H3-27 | N2-27 | T2-27 |
| IR310.T | E5-159, 168-265, 276-309 | K5-159 168-265 276-309 | E4-161, 168-266, 274-309 | K4-162, 168-266, 274-309 | Q5-310 | W4-161, 167-268, 274-309, |
| IR-A α CT | 705-719 | 705-715 | F704-719 | L704-719 | R704-719 | X704-719 |
| Asn16 ² | | | NAG \times 2 FUC | NAG \times 2 | NAG | NAG \times 2 FUC BMA |
| Asn25 | | | NAG \times 2 FUC BMA | NAG \times 2 FUC | NAG \times 2 BMA | NAG \times 2 FUC BMA |
| Asn111 | | | NAG | NAG | NAG | NAG |
| Asn215 | | | NAG | NAG | NAG | NAG \times 2 |
| Asn255 | | | NAG \times 2 FUC BMA | NAG \times 2 FUC BMA MAN | NAG \times 2 FUC BMA | NAG \times 2 FUC BMA MAN |
| Asn295 | | | - | - | - | - |
| 83-7 VH | 1-118 | 1-117 | 1-121 | 1-117 | 1-122 | 1-117 |
| 83-7 VL | 1-114 | 1-111 | -1-114 | 0-112 | 0-113 | -2-111 |
| Solvent | 4 \times SO ₄ ²⁻ | | 18 \times H ₂ O | | | |

¹For compatibility with the hIns B-chain sequence, residues of the B chain of Con-Ins G1 are numbered -1, 0, 1, ..., 21.

²NAG: N-acetylglucosamine; FUC: fucose; BMA: β -D-mannose; MAN: α -D-mannose.

Supplementary Table 4**Variation of site 1 insulin-binding residues across fish and human insulin receptors¹**

| <i>H. sapiens</i> | <i>P. olivaceus</i> | <i>T. nigroviridis</i> | <i>D. rerio</i> |
|-------------------|---------------------|------------------------|-----------------|
| Gln342 | Lys | Lys | Lys |
| Lys40 | Arg | Arg | Thr |
| Asp707 | Asn | Asn | Asn |
| Val712 | Glu | Glu | Glu |
| Val715 | Glu | Glu | Glu |
| Pro716 | Ile | Leu | Leu |
| Arg717 | Lys | Lys | Arg |

¹ The remaining residues of the primary binding site of the receptor (*viz.*, Asp12, Arg14, Asn15, Leu36, Leu37, Phe39, Leu62, Phe64, Arg65, Phe88, Phe89, Try91, Val94, Phe96, Glu97, Arg118, Glu120, Lys121, Thr704, Phe705, Glu706, Tyr708, His710, Asn711, Val713 and Phe714) are conserved across the four species shown.

² Human insulin receptor isoform A numbering.

Detailed Methods for chemical synthesis and characterization of insulin.

Chemical and Protein Materials

Insulin was purchased from Life Technologies unless otherwise specified. Peptides were synthesized via Fmoc solid phase peptide synthesis. N,N-Diisopropylethylamine (DIEA), triisopropylsilane, L-ascorbic acid, acetic acid (AcOH), iodine, tryptophan, piperidine, methanol (MeOH), urea and dichloromethane (DCM) were purchased from Sigma-Aldrich and used directly. Fmoc-protected amino acids and 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU) were purchased from Chem-Impex Int'l. Inc. Boc-Ser[Fmoc-Thr(tBu)] was purchased from AAPPTec. 2-Chlorotriyl chloride resin was purchased from ChemPep. ChemMatrix® Rink amide resin was purchased from Biotage. Dimethylformamide (DMF), trifluoroacetic acid (TFA), acetonitrile (ACN) and ethyl ether were purchased from Fisher Scientific and used as supplied.

HPLC and LC/MS

All crude peptides were purified with a water/acetonitrile gradient in 0.1% TFA on an Agilent 1260 HPLC system. Fractions collected from HPLC were analyzed by LC/MS on a XBridge C18 5- μm (50 \times 2.1 mm) column at 0.4 mL.min⁻¹ with a water/acetonitrile gradient in 0.1% formic acid on an Agilent 6120 Quadrupole LC/MS system. Fractions containing targeted product (based on LC/MS) were collected and lyophilized.

General RP-HPLC conditions

Method A: Individual chains were purified by a Preparative C18 (2) Column (Luna®, 5 μm , 250 x 21.2 mm) with a linear gradient from 20% aqueous ACN (0.1% TFA) to 50% aqueous ACN (0.1% TFA) over 40 min at a flow rate of 5 mL.min⁻¹ for A chains and from 30% aqueous ACN (0.1% TFA) to 60% aqueous ACN (0.1% TFA) over 40 min at a flow rate of 5 mL.min⁻¹ for B chains.

Method B: All folded peptides and final products were purified by a Phenomenex semi-preparative C18 Column (5 μm , 250 x 10 mm) with a linear gradient from 20% aqueous ACN (0.1% TFA) to 50% aqueous ACN (0.1% TFA) over 35 min at a flow rate of 3 mL.min⁻¹.

Peptide Synthesis

Peptides were synthesized via Fmoc solid phase peptide synthesis on peptide synthesizer (Altra; Biotage, Inc). Automated peptide synthesis was carried out in a 10 mL reactor vial with the following protocols (for 0.1 mmol scale). For Fmoc deprotection: (i) 4.5 mL of 20% piperidine in DMF; (ii) mix 2 \times 3 min (new solvent delivered for each mixing cycle). For amino acid coupling: (i) 1.25 mL of 0.4 M Fmoc-protected amino acid in DMF; (ii) 1.225 mL of 0.4 M HATU; (iii) 1.0 mL of 1.0 M DIPEA in DMF; and (iv) mix for 10 min at 70 °C (for cysteine and histidine coupling: mix for 10 min at 50 °C; for arginine coupling: mix for 10 min at 50 °C and coupling twice). For DMF washing (performed between deprotection and coupling steps): (i) 4.5 mL of DMF; (ii) mix

45 s. Upon completion of the peptide chain, resins were washed with DCM and dried (using vacuum) for 30 min. Peptide was then cleaved from the resin by exposure to cleavage cocktail for 2 h, which was prepared with 12.5 mL TFA, 330 μ L water, 330 μ L TIS. The peptide was precipitated with ethyl ether at 4 °C, followed by HPLC purification and lyophilization.

General procedure for A-chains

The synthesis of A chains were conducted on Rink amide chemMatrix resin using peptide synthesizer through a standard Fmoc_HATU_DIEA method. The resulting resin-bound A chains were treated with 25% β -mercaptoethanol in DMF (vol/vol, 5 mL) for 2.0h at rt. This step was repeated once. The resulting resins were washed with DMF (5 mL x 3) and DCM (5 mL x 3), and treated with DTNP (10 equiv.) in DCM (5 mL) for 1.0 h at rt. After washing by DMF (5 mL x 3) and DCM (5 mL x 3), the resins were treated with 1% TFA, 5% TIS containing DCM (5 mL) for 2 min with 5 repeats. The resins were washed DMF (5 mL x 3) and DCM (5 mL x 3), and gently agitated in DCM (5 mL) for another 1.0 h before cleavage and purification. DOI A-chains were obtained after purification and lyophilization.

General procedure for B chain

B-chain syntheses were conducted on 2-chlorotrityl chloride resin using standard Fmoc_HATU_DIEA method. The first amino acids were synthesized manually. Cleavage was conducted by treating these resins with 5 mL TFA solution that contained 2.5% TIS, 2.5% H₂O and 10 equiv of DTDP at rt, with shaking for 2 h. The resins were filtered off; the filtrate was precipitated with cold ether (35 mL). The precipitates were collected by centrifugation, and then washed with cold ether (35 mL x 3). Crude B chains were dissolved 0.05% TFA containing aqueous acetonitrile (ACN/H₂O: 50/50 vol/vol, 40 mL) and purified by preparative C18 column.

General procedure for preparation of analogs without tryptophan by one-pot method

A chain (0.005 mmol) and B chain (0.006 mmol) were mixed in 6 M urea, 0.2 M NH₄HCO₃ buffer (pH 8, 2 mL). The mixture was left at rt for 10 min before being treated with a freshly prepared solution of iodine (12.7 mg, 0.1 mmol, 20 equiv based on A chain) in AcOH (3.0 mL). The resulting solution was gently agitated at rt for 5 min before the addition of 1 M ascorbic acid (0.1 mL) until the iodine color (purple) disappeared. The solution was diluted by H₂O (10 mL) and purified. After lyophilization of the fractions, DOI analogs were obtained.

Procedure for preparation of DOI analog with tryptophan by two-step method

A chain (0.005 mmol) and B chain (0.006 mmol) were combined as described in the previous section.

Disulfide (A6-A11, A20-B19) A-B dimer (0.001 mmol) was dissolved in a mixture solvent of AcOH (1.0 mL) and H₂O (1.0 mL) at rt, followed by the addition of 100 mm aqueous hydrochloric acid (1.0 mL). Subsequently, 100 μ L (5 equiv) of tryptophan solution (50 mm in H₂O) and 1.0 mL (20 equiv) of iodine solution (20 mm in methanol) were added sequentially. The resulting solution was stirred for 30 min at room temperature and added ascorbic acid (50 μ L) until the solution turned colorless. HPLC purification and Freeze-drying afforded DOI analog with tryptophan.

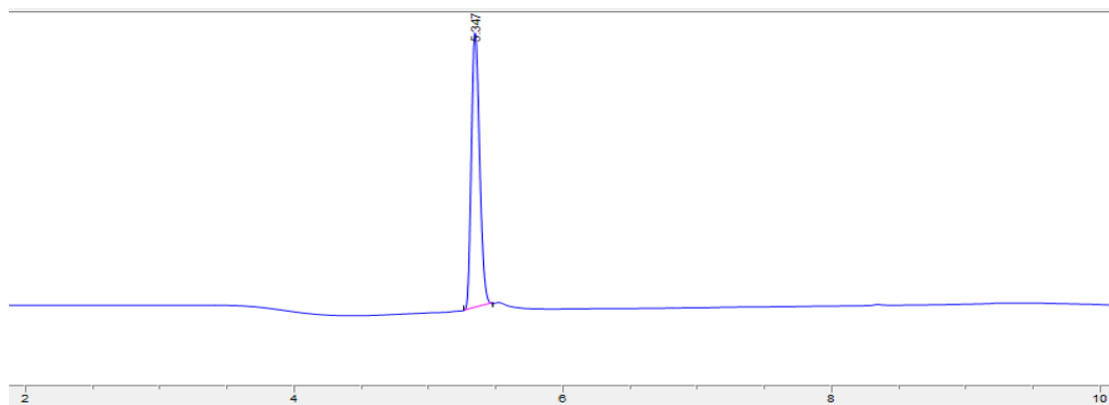
Copy of LC chromatogram, MS spectrum, and trypsin digested segments MS spectrum.



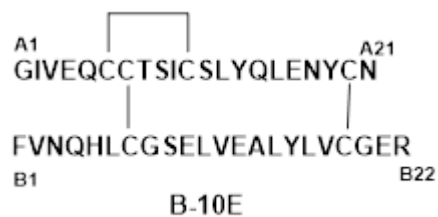
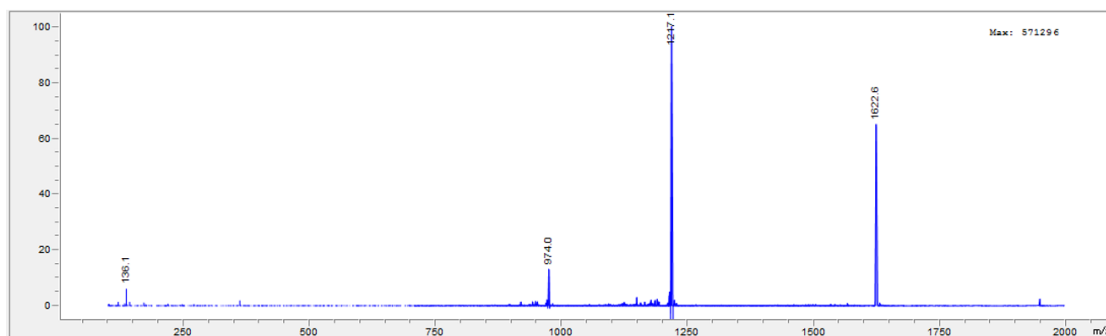
Expected MS: 4865.5

Observed MS: 4864.6

LC for DOI:



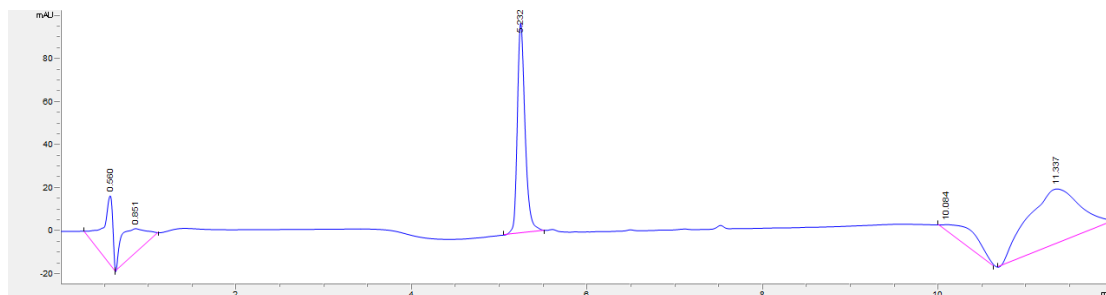
MS Spectrum for DOI:



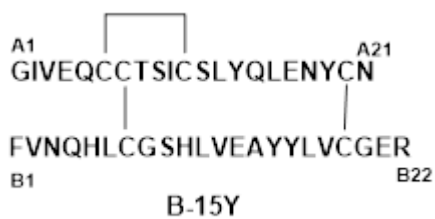
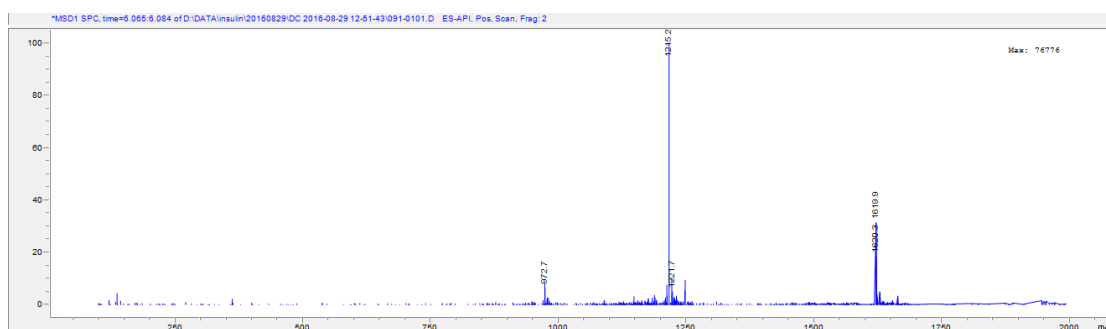
Expected MS: 4857.5

Observed MS: 4855.9

LC for **B 10E**:



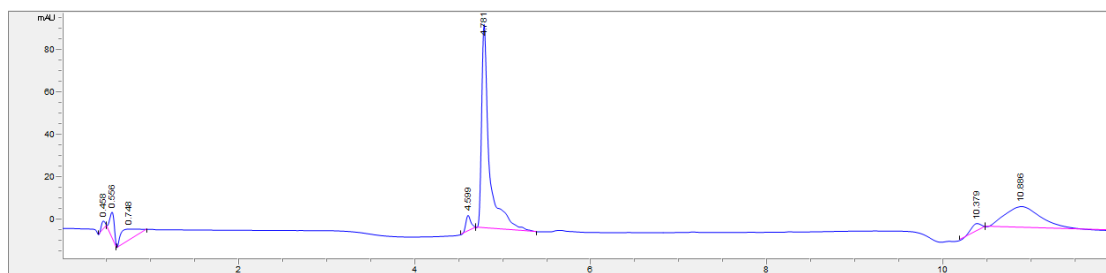
MS Spectrum for **B 10E**:



Expected MS: 4915.6

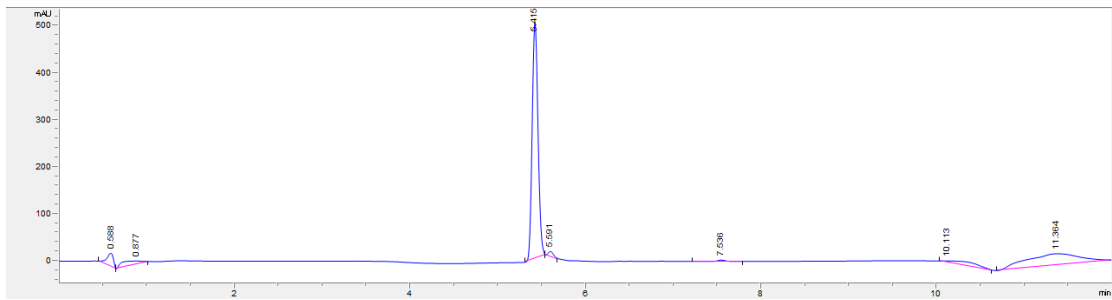
Observed MS: 4915.3

LC for **B 15Y**:

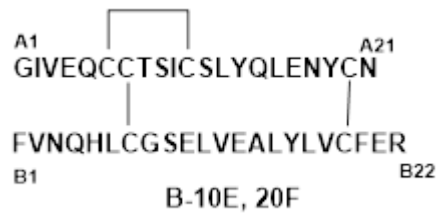
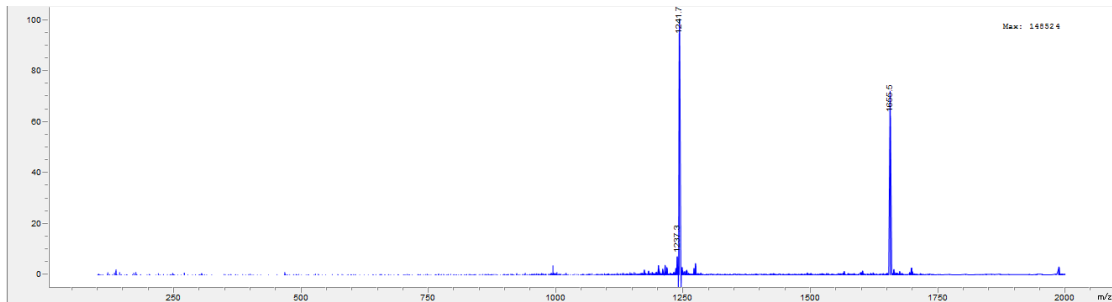


MS Spectrum for **B 15Y**:

LC for **B 10E 20Y DOI:**



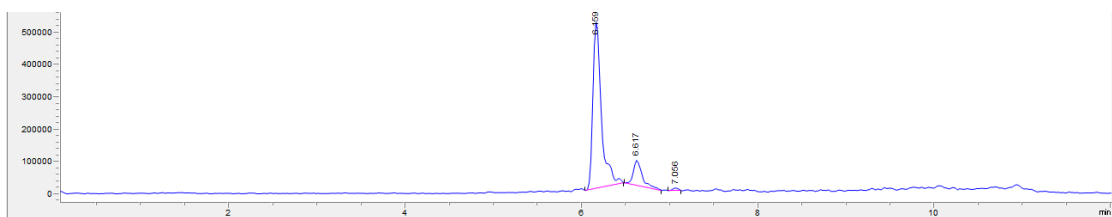
MS Spectrum for **B 10E 20Y DOI:**



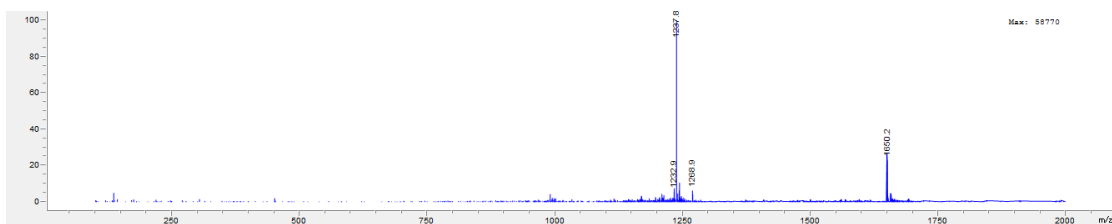
Expected MS: 4947.6

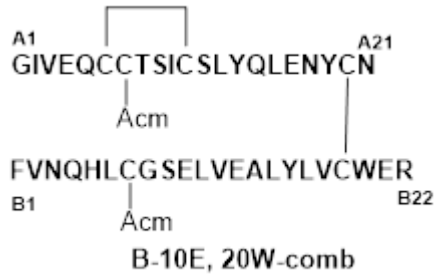
Observed MS: 4947.4

LC for **B 10E 20F DOI:**



MS for **B 10E 20F DOI:**

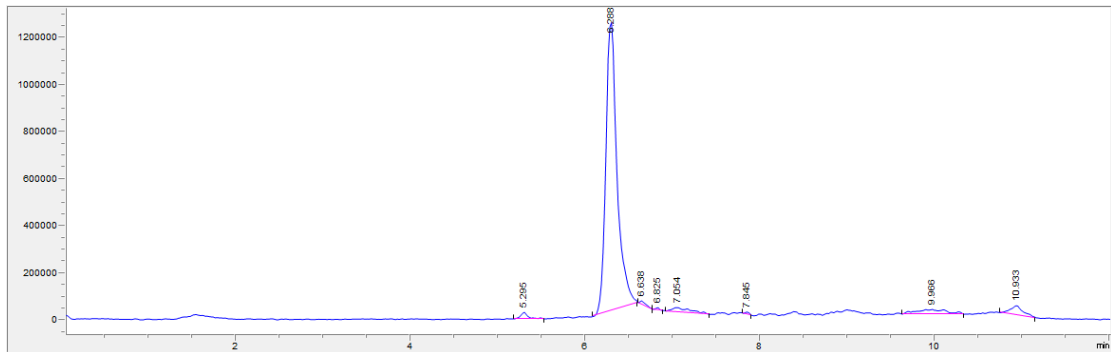




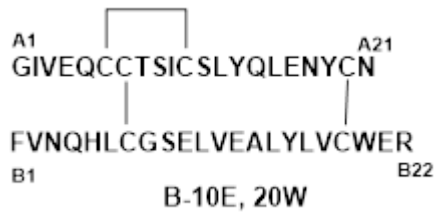
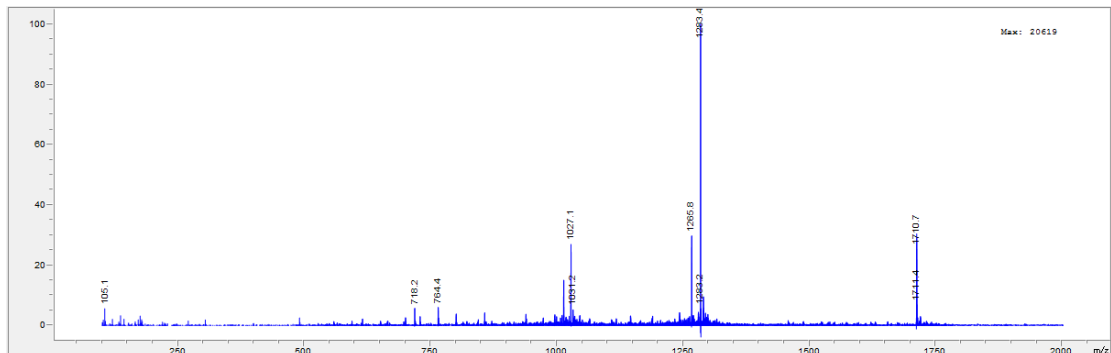
Expected MS: 5130.7

Observed MS: 5129.8

LC for B 10E 20W DOI comb:



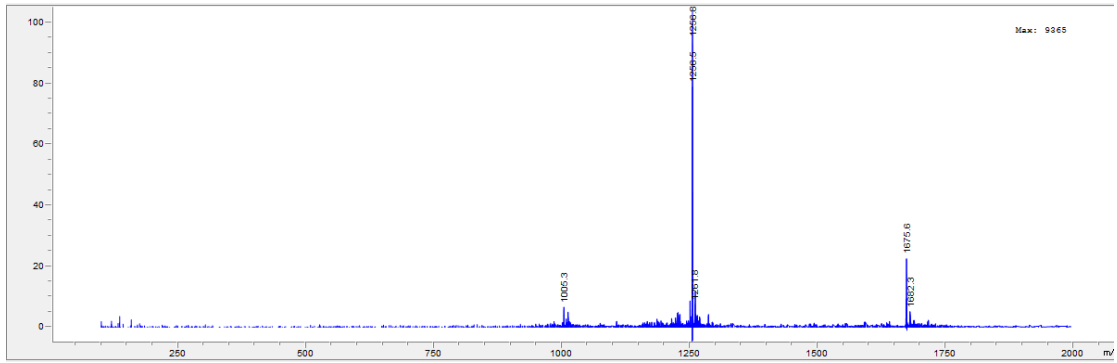
MS for B 10E 20W DOI comb:



Expected MS: 4986.7

Observed MS: 4985.5

LC for B 10E 20W DOI:



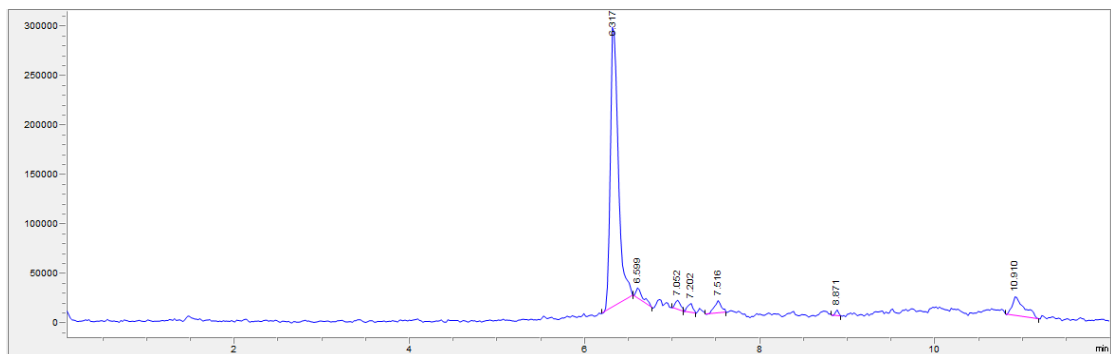
X= L-Phe(4-tBu)-OH

10E, 20-(4-tBu)-F

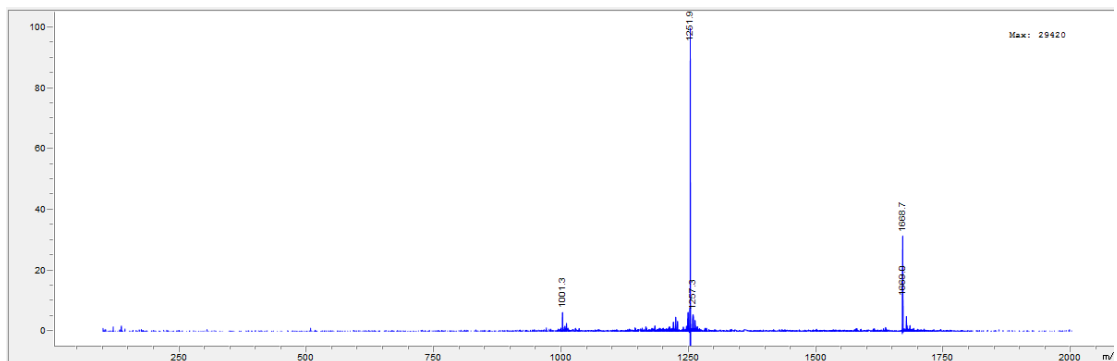
Expected MS: 5003.6

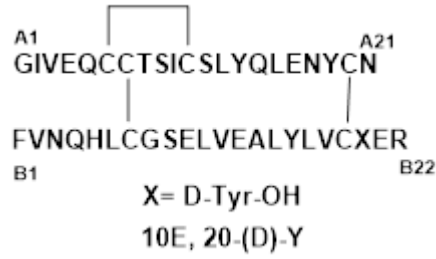
Observed MS: 5002.7

LC for **B 10E 20 (4-t-Bu)-F DOI:**



MS for **B 10E 20 (4-t-Bu)-F DOI:**

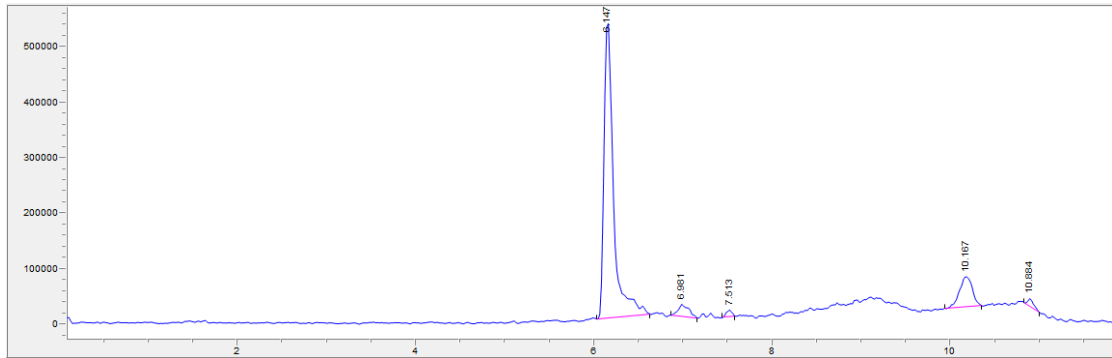




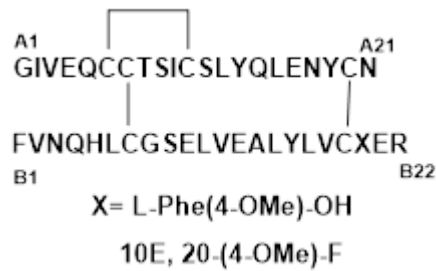
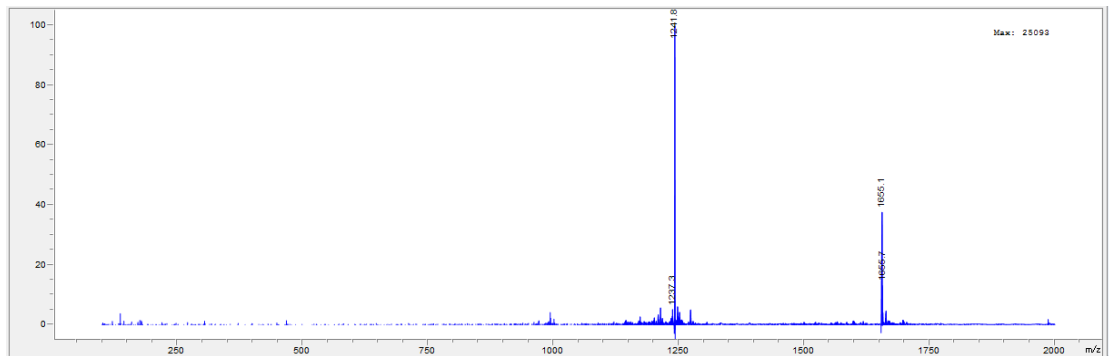
Expected MS: 4963.6

Observed MS: 4962.8

LC for B 10E 20(D)Y DOI:



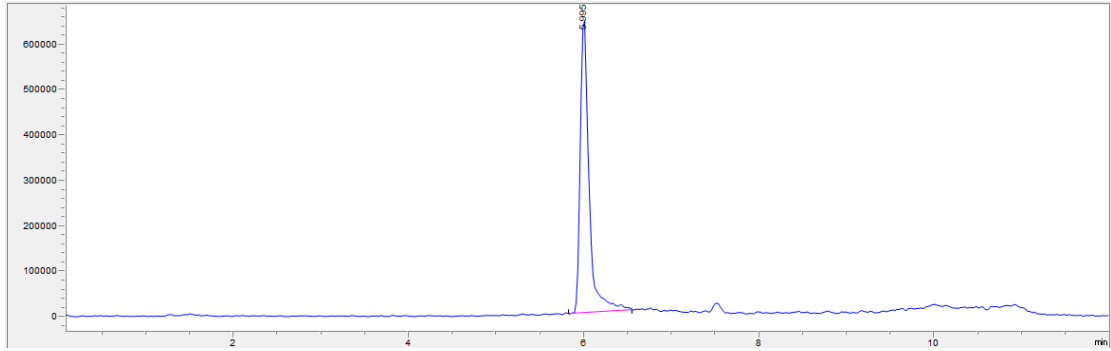
MS for B 10E 20(D)Y DOI:



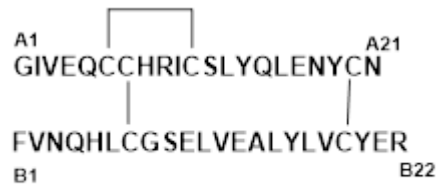
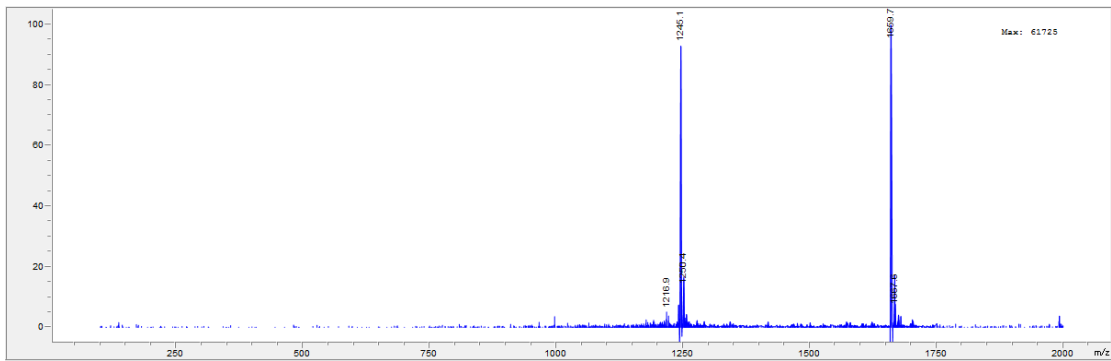
Expected MS: 4976.6

Observed MS: 4976.2

LC for B 10E 20(OMe)F DOI:



MS for **B 10E 20(OMe)F DOI:**

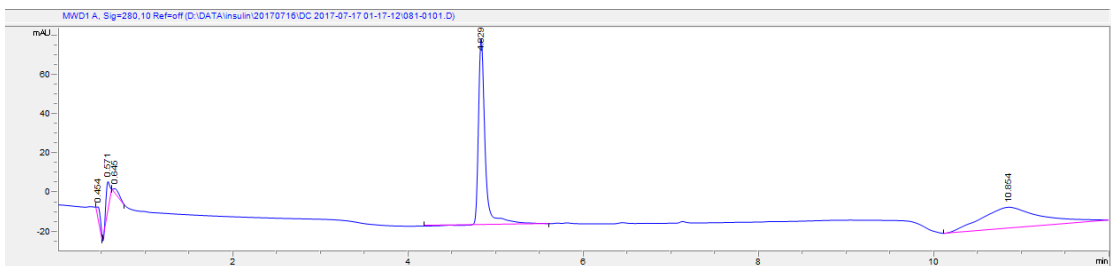


HR, B10E, 20Y

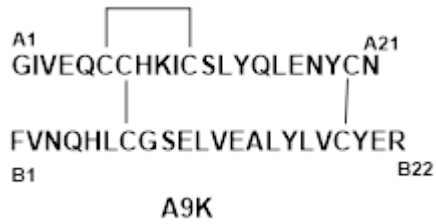
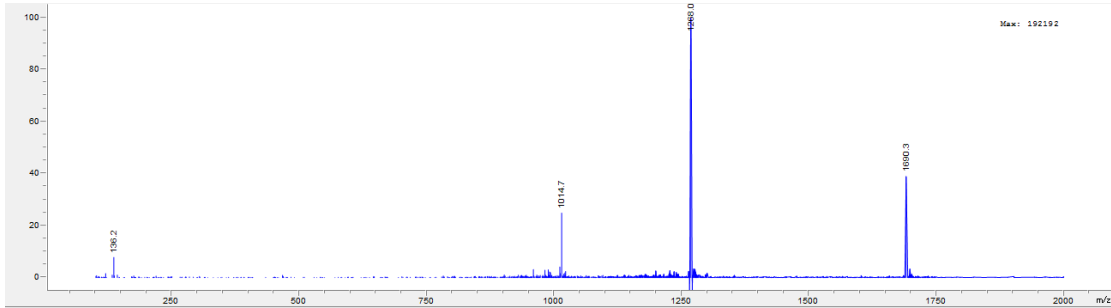
Expected MS: 5068.9

Observed MS: 5068.1,

LC for **A HR B 10E 20YDOI:**



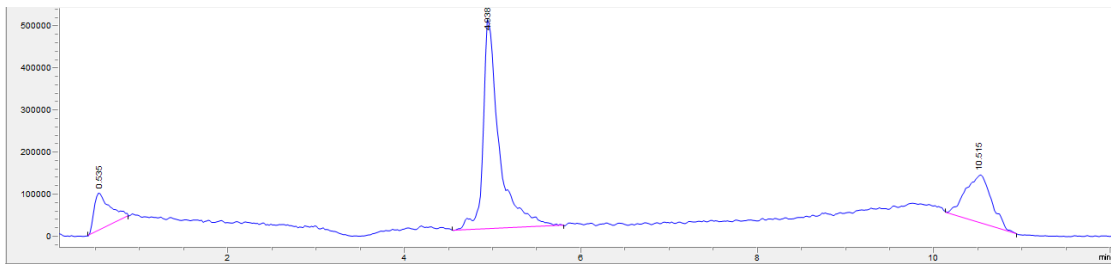
MS for **A HR B 10E 20YDOI:**



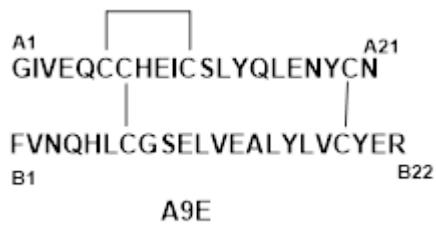
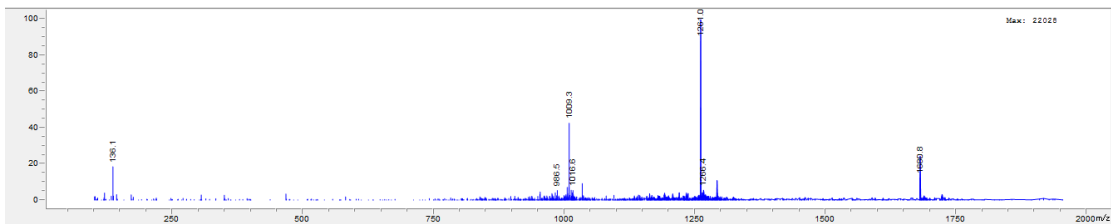
Expected MS: 5040.8

Observed MS: 5040.3

LC for **A9K DOI:**



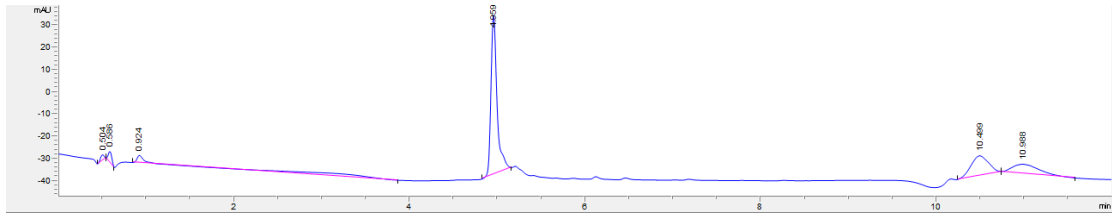
MS for **A9K DOI:**



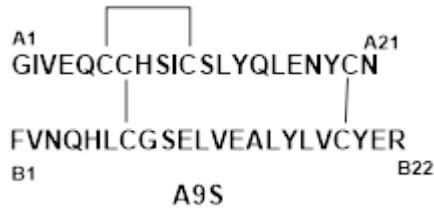
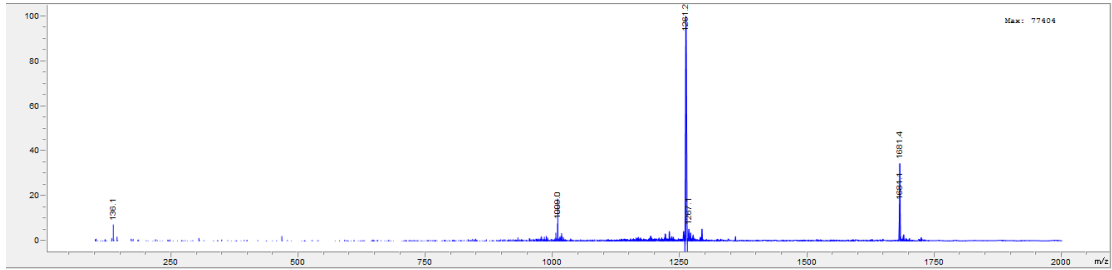
Expected MS: 5041.7

Observed MS: 5041.1

LC for **A9E DOI:**



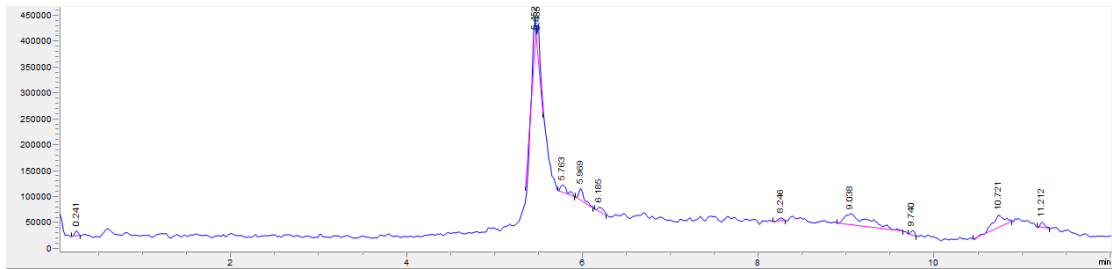
MS for A9E DOI:



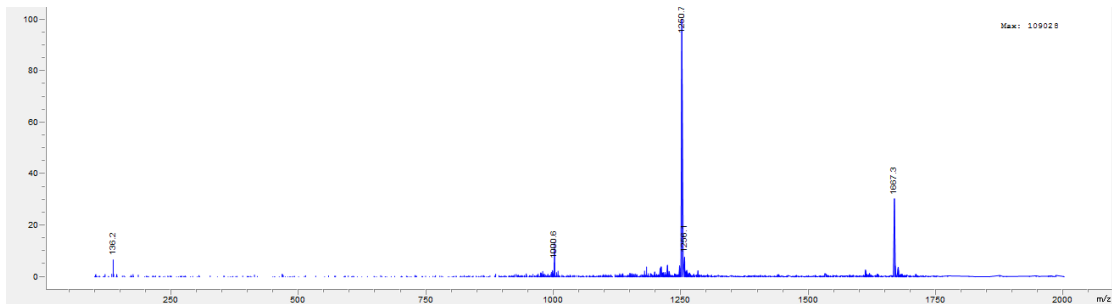
Expected MS: 4999.7

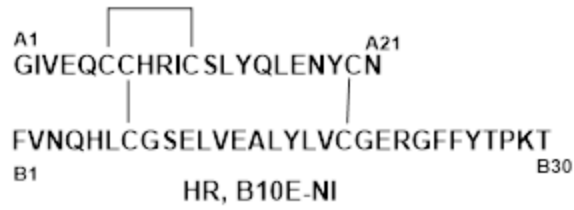
Observed MS: 4998.6

LC for A9S DOI:



MS for A9S DOI:

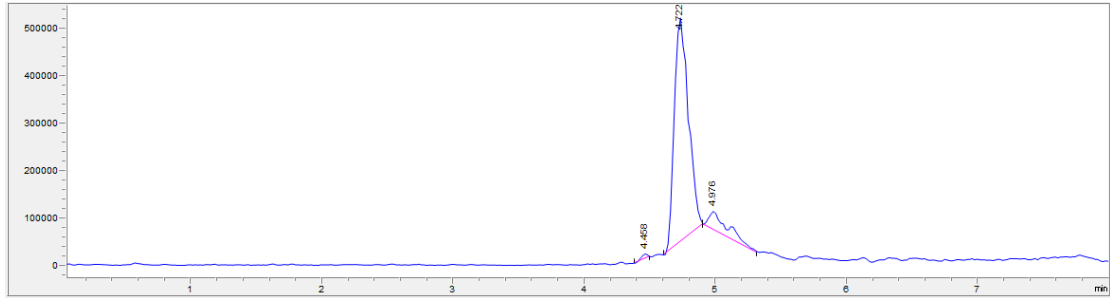




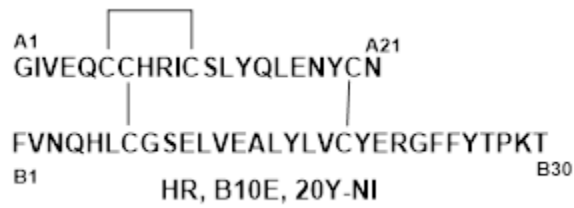
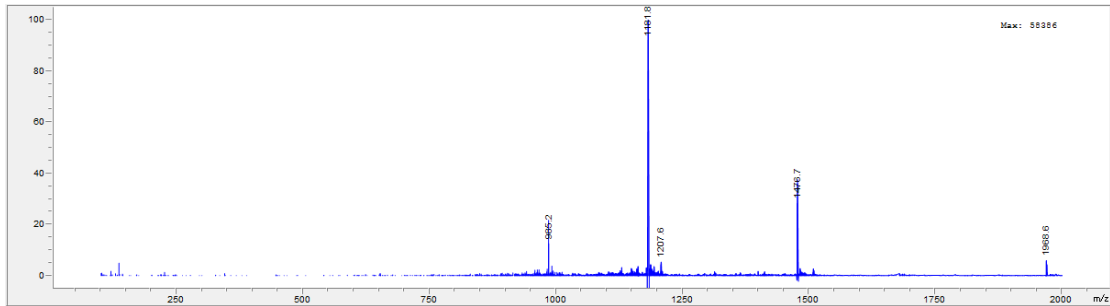
Expected MS: 5904.7

Observed MS: 5904.1

LC for **HR B10E-NI**:



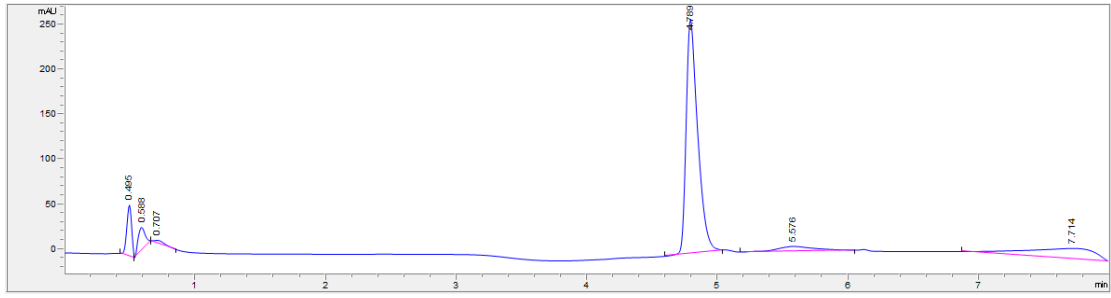
MS for **HR B10E-NI**:



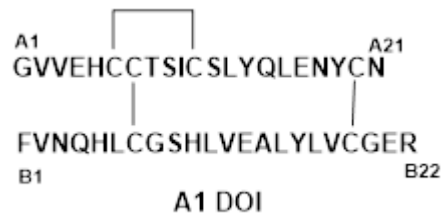
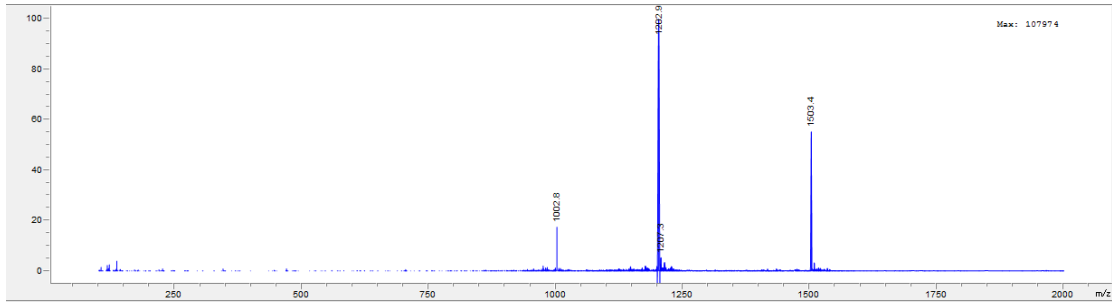
Expected MS: 6010.8

Observed MS: 6009.9

LC for **HR B10E, 20Y-NI**:



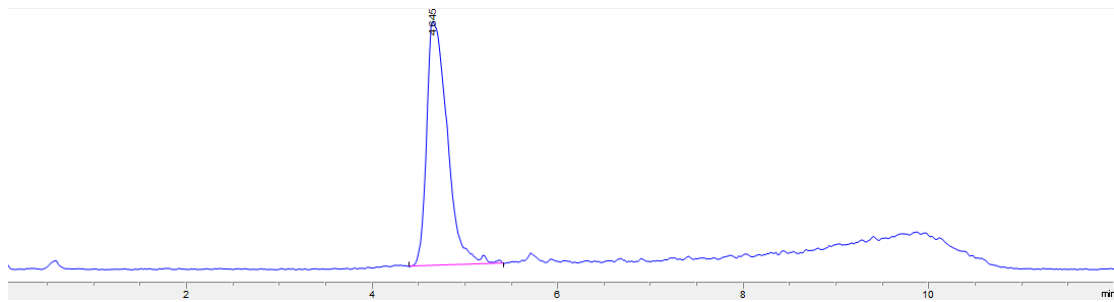
MS for HR B10E, 20Y-NI:



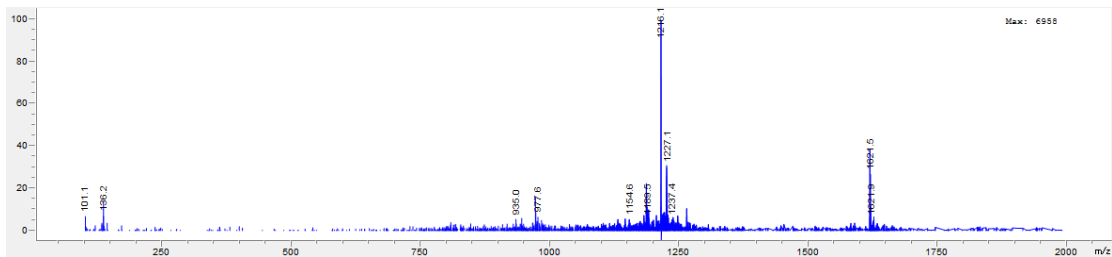
Expected MS: 4860.5

Observed MS: 4860.9,

LC for A1 DOI:



MS for A1 DOI:

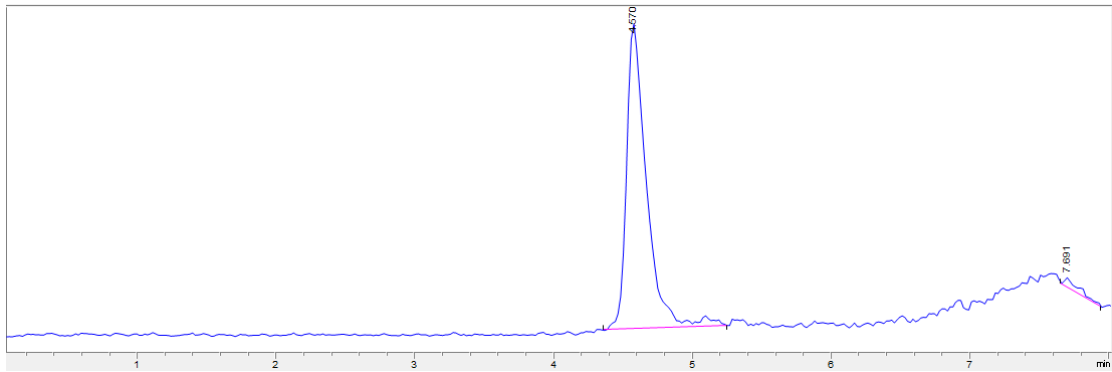




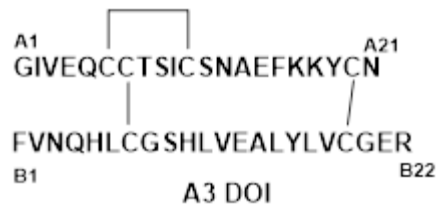
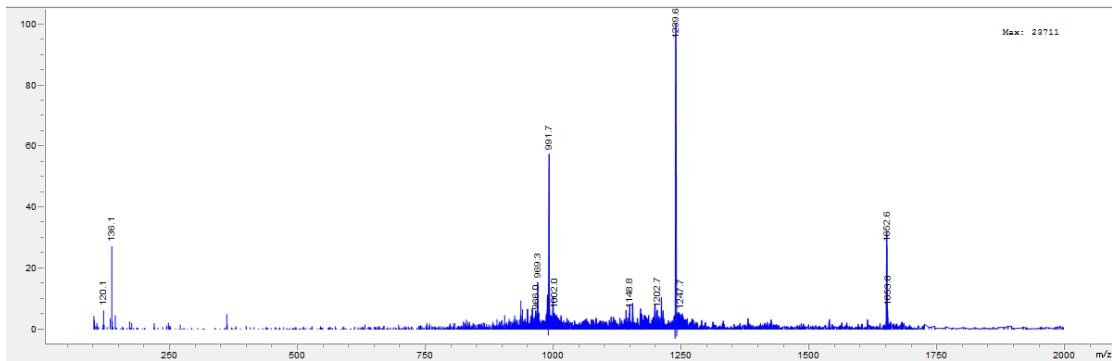
Expected MS: 4954.6

Observed MS: 4954.2

LC for A2 DOI:



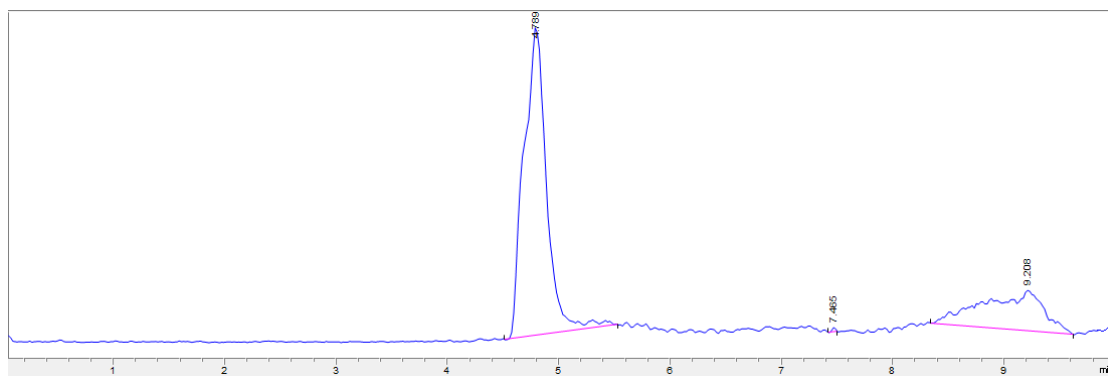
MS for A2 DOI:



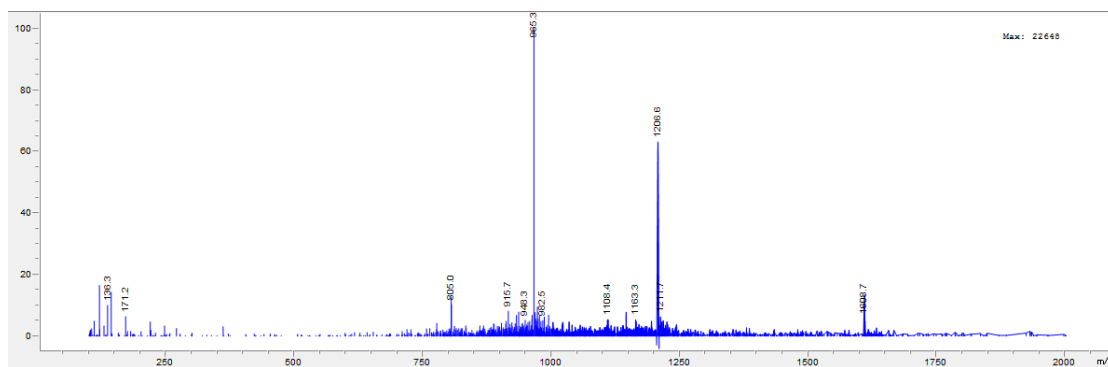
Expected MS: 4822.5

Observed MS: 4822.3

LC for A3 DOI:



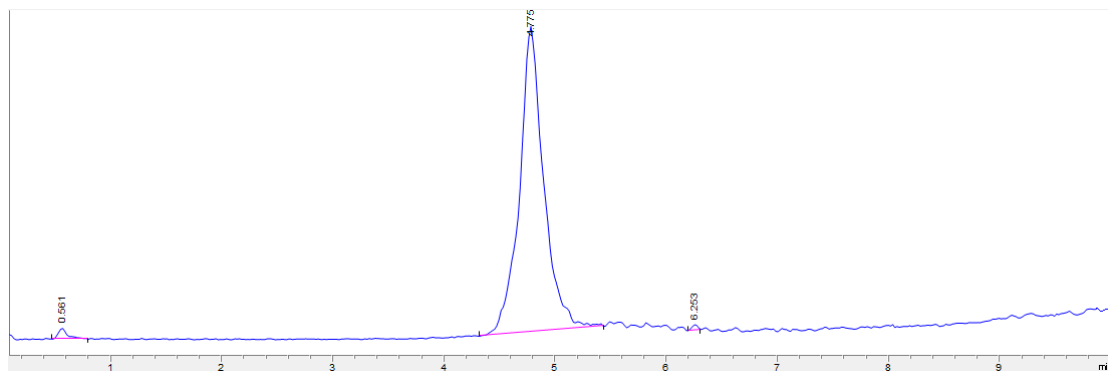
MS for **A3 DOI**:



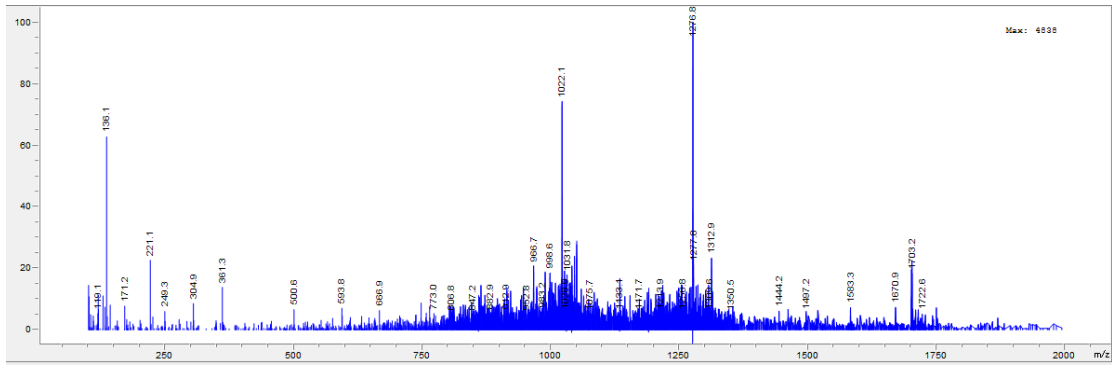
Expected MS: 5109.8

Observed MS: 5105.4

LC for **B1 DOI**:



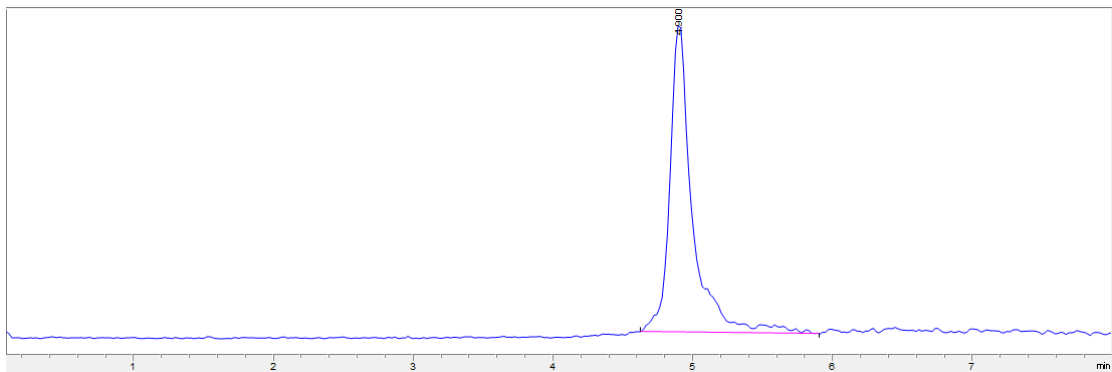
MS for **B1 DOI**:



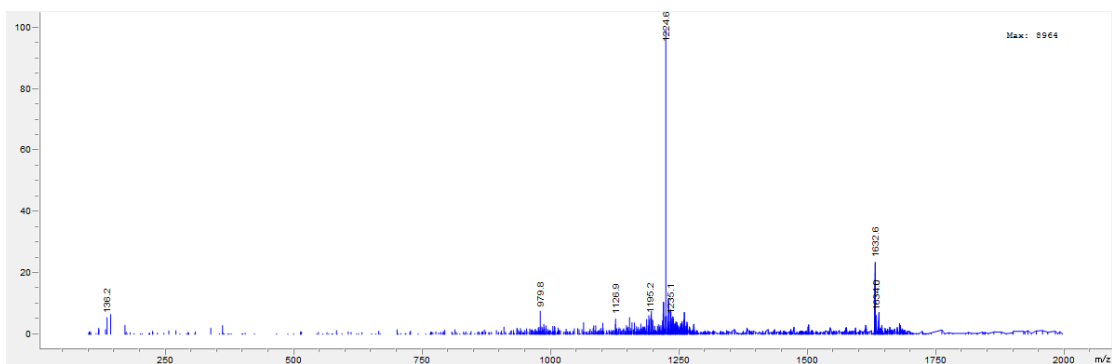
Expected MS: 4894.5

Observed MS: 4894.6

LC for **B2 DOI**:



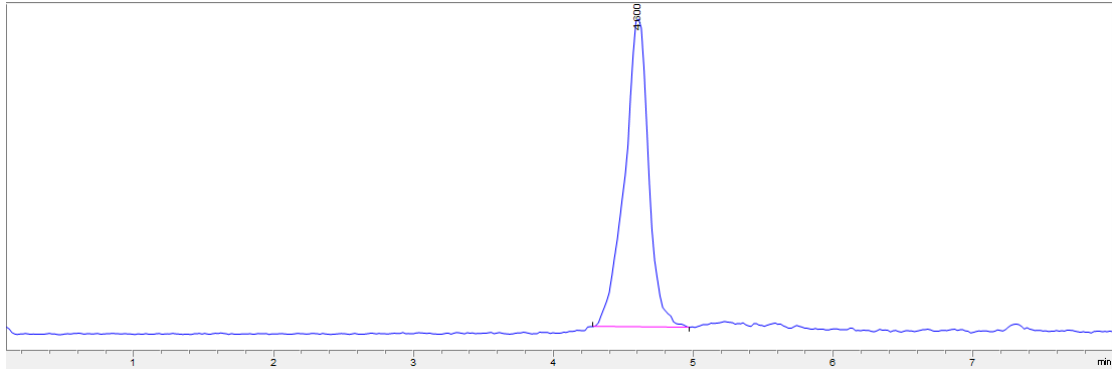
MS for **B2 DOI**:



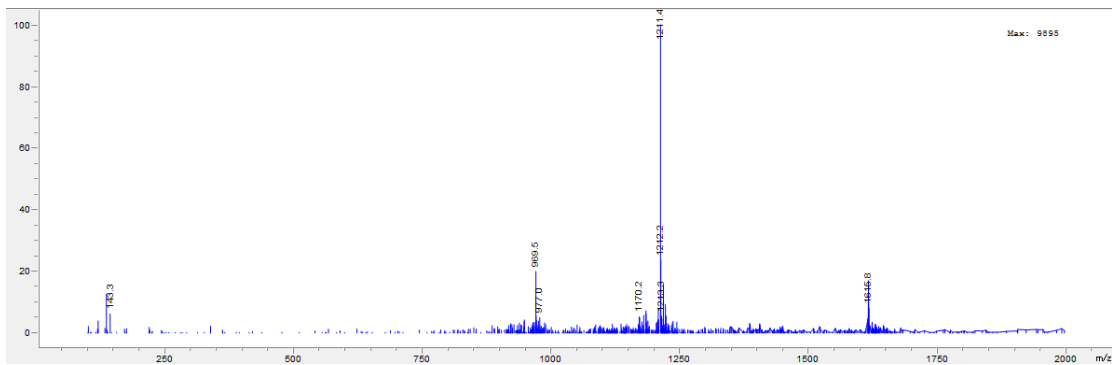
Expected MS: 4842.5

Observed MS: 4842.8

LC for **B3 DOI:**



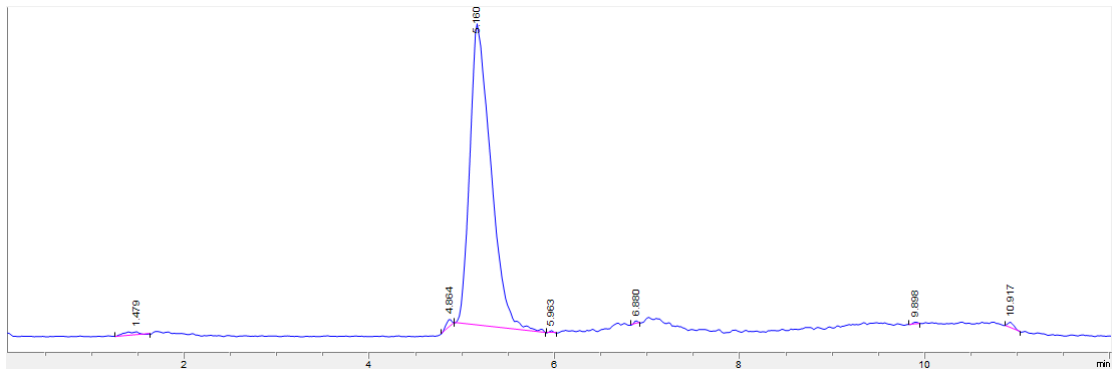
MS for **B3 DOI:**



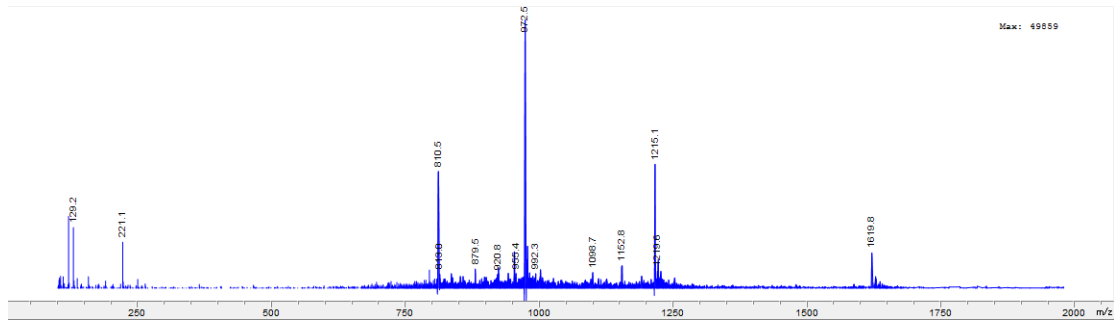
Expected MS: 4858.5

Observed MS: 4857.1

LC for **15 A 20A:**



MS for **15A 20A:**

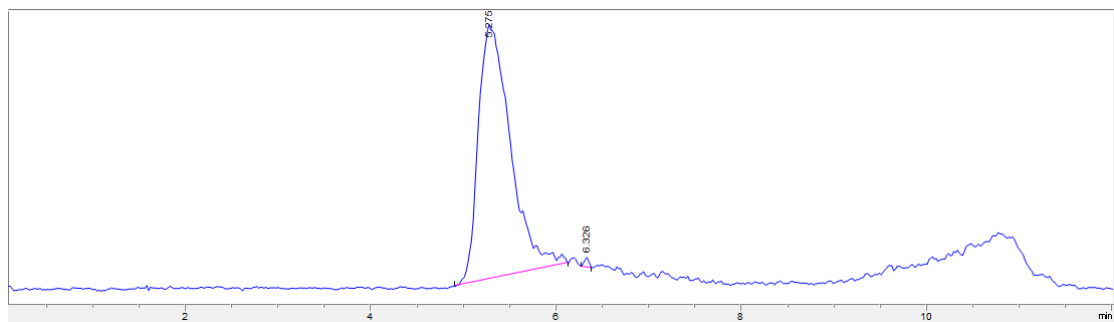


15A, 20Y

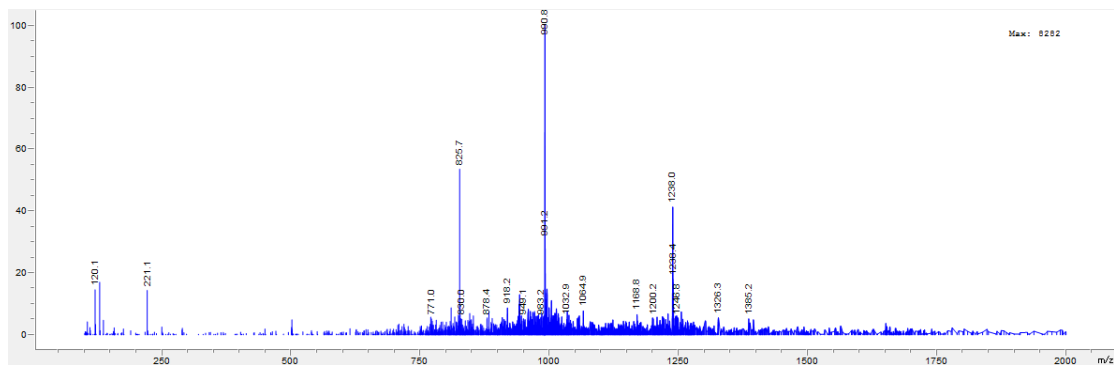
Expected MS: 4950.6

Observed MS: 4948.4

LC for 15A 20Y:



MS for 15A 20Y:

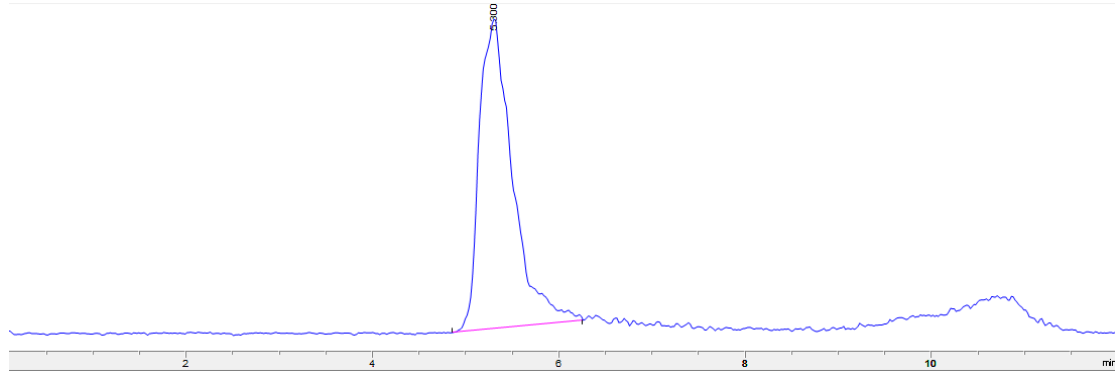


15Y, 20A

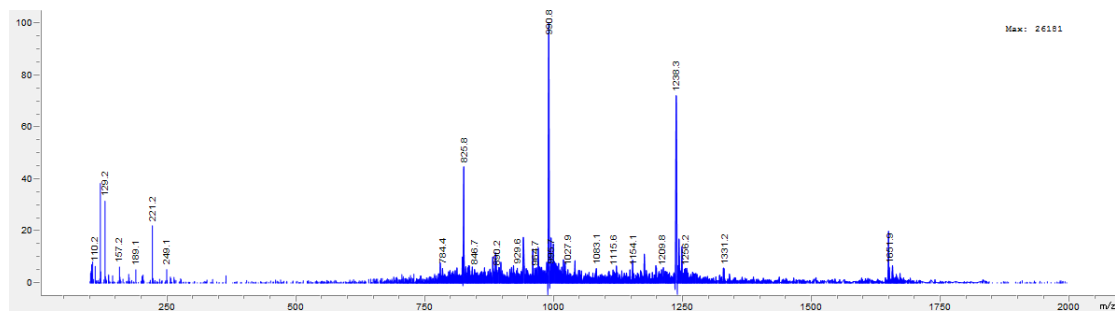
Expected MS: 4950.6

Observed MS: 4948.8

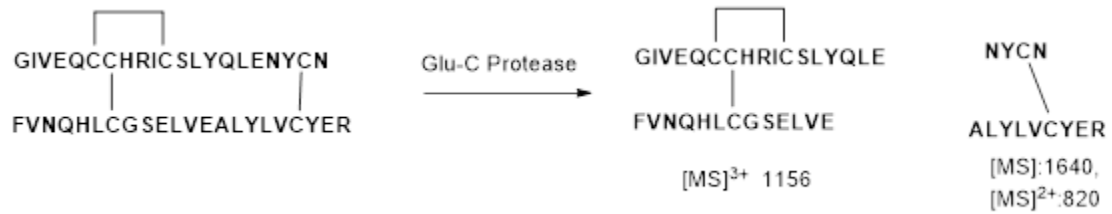
LC for 15 Y 20A:



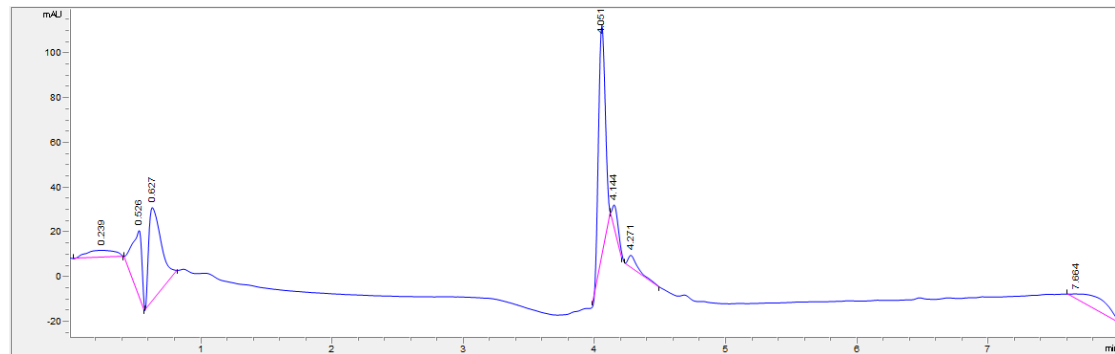
MS for 15Y 20A:



Glu-C digestion of HR, B10E, 20Y DOI:



LC for Glu-C digestion of B10E, 20Y DOI:



MS for Glu-C digestion of B10E, 20Y DOI:

