Supporting Information

Cyclic Peptide Production Using a Macrocyclase with Enhanced Substrate Promiscuity and Relaxed Recognition Determinants

Cristina N Alexandru-Crivac[1,2], Christian Umeobika[1], Niina Leikoski[3], Jouni Jokela[3], Kirstie A. Rickaby[1], André M. Grilo[1,2], Peter Sjö[4], Alleyn T Plowright[5], Mohannad Idress[1,2], Eike Siebs[1], Ada Nneoyi-Egbe[1], Matti Wahlsten[3], Kaarina Sivonen[3], Marcel Jaspars[1], Laurent Trembleau[1]*, David P Fewer[3]* & Wael E Houssen[1,2,6]*

[1]Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Meston Walk, Aberdeen AB24 3UE (UK).

[2]Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD (UK).

[3]Microbiology and Biotechnology Division, Department of Food and Environmental Sciences, P.O.Box 56, Viikki Biocenter, Viikinkaari 9, 00014, University of Helsinki (Finland)

[4]Medicinal Chemistry, Respiratory, Inflammation & Autoimmunity, Innovative Medicines and Early Development Biotech Unit, AstraZeneca, Pepparedsleden 1, Mölndal, 43183 (Sweden).

[5]Medicinal Chemistry, Cardiovascular and Metabolic Diseases Innovative Medicines and Early Development Biotech Unit, AstraZeneca, Pepparedsleden 1, Mölndal, 43183 (Sweden).

[6]Pharmacognosy Department, Faculty of Pharmacy, Mansoura University, Mansoura 35516 (Egypt).

*Correspondence should be addressed to

Wael Houssen (w.houssen@abdn.ac.uk), David P. Fewer (david.fewer@helsinki.fi) or Laurent Trembleau (l.trembleau@abdn.ac.uk)

Experimental Data

Peptide Synthesis

Fmoc amino acid derivatives, 2-(1H-benzotriazol-1-yl)-1, 1, 3, 3 - tetramethyluronium hexafluorophosphate (HBTU), 2-chlorotrityl chloride resin, Fmoc-Phe-Ser(ψ^{Me,Me}Pro)-OH, and Fmoc-thioproline were purchased from Novabiochem, Merck Biosciences, UK. Trifluoroacetic acid (TFA), N,N-diisopropylethylamine (DIEA), N,N-dimethylformamide (DMF) and piperidine were obtained from Alfa Aesar, UK, and used without further purification. Fmoc-Ser(ψPro)-OH was prepared from L-serine, formaldehyde, Fmoc chloride, and *N,N*-diisopropylethylamine (all from Sigma-Aldrich) using an adapted literature procedure for the synthesis of Boc-Ser(ψPro)-OH [Vishnumaya, Monika Raj; Singh, Vinod K. *J. Org. Chem.*, **2009**, *74*, 4289 - 4297].

The linear peptide substrates (**1**-**18**) were synthesized manually in-house using the standard Fmoc-based solid phase peptide synthesis (SPPS) strategy. The attachment of the first amino acid to the 2-chlorotrityl resin was conducted using a molar ratio of amino acid: DIEA: resin of 1.2: 4.8: 1 in DMF. Subsequent amino acids were sequentially coupled following removal of the Fmoc protecting group at each cycle. Fmoc deprotection steps were carried out with 20% piperidine in DMF (v/v) for 6 min (3x: 2min each); coupling reactions were performed in DMF using a molar ratio of amino acid: HBTU: DIEA: resin of 5: 5: 10: 1. Reactions were monitored using the Kaiser test. Pseudoproline residues were introduced using the corresponding Fmoc-Phe-Ser(ψ^{Me,Me}Pro)-OH dipeptide acid, Fmoc-Ser(ψPro)-OH or Fmoc-thioproline.

The peptides were cleaved from the support and side-chain protecting groups removed by treatment with a mixture consisting of 95% TFA, 2.5% triisopropylsilane (TIPS), 2.5% H_2O (20 mL of mixture per g of peptide resin, 5 hours at room temperature). The resin was then filtered and washed with TFA. The combined filtrates were concentrated under reduced pressure. The peptide was precipitated with cold diethyl ether and recovered by centrifugation. The peptide sequence was verified by MS-MS analysis.

As expected, peptide **17** containing the Ser(ψ^{Me,Me}Pro) residue did not survive the traditionally used cleavage cocktail (95% TFA, 2.5% TIPS, 2.5% H_2O), but it was successfully prepared on 2-chlorotrityl resin using side-chain protecting groups sensitive to mildly acidic conditions for the tyrosine and aspartic acid residues [Fmoc-Tyr(2-ClTrt)-OH and Fmoc-Asp(O-2-PhiPr)-OH, Novabiochem, Merck Biosciences, UK]. The peptide was cleaved from the resin under mild acidic conditions (1% TFA / 5% TIS in DCM) and precipitated in cold diethyl ether. The ψ^{Me,Me}Pro residue was not affected under these mildly acidic conditions (see Figure S47).

Substrate **19** is prepared from a precursor linear peptide by proteolytic cleavage with Tobacco Etch Virus (TEV) protease and subsequent purification using Nickel affinity chromatography. The percusor peptide is produced recombinantly in *Escherichia coli* BL21(DE3) and contains the substrate sequence fused with an N-terminal TEV-cleavable Small Ubiquitin-like Modifier (SUMO) tag.

Protein Purification

Full length codon optimised $OscG_{mac}$ (*Oscillatoria* sp. PCC 6506) with an N-terminal TEV protease-cleavable His6-tag cloned in pJexpress411 vector was purchased from Atum (Figure S4). PatG mutants were generated using Q5® Site-Directed Mutagenesis Kit (New England Biolabs) and following the manufacturer protocol. Proteins were expressed in *E. coli* BL21 (DE3) grown on auto-induction medium (Formedium Terrific broth base containing trace elements) for 48 h at 20 °C, with shaking at 200 rpm. Cells were harvested by centrifugation at 4,000 g, 4 °C, for 15 min, and re-suspended in lysis buffer (500 mM NaCl, 20 mM Tris, pH 8.0, 20 mM imidazole pH 8.0 and 3 mM β-mercaptoethanol (BME) with the addition of complete EDTA-free protease inhibitor tablets (Roche) and DNase (Sigma) at 0.4 mg g−1 wet cells). Cells were lysed by passage through a cell disruptor at 30 kPsi and each lysate was cleared by centrifugation at 40,000 x g, 4 °C for 45 min. followed by filtration through a 0.45 µm membrane filter. Each cleared lysate was loaded onto a Ni-sepharose 6 FF column (GE Healthcare) prewashed with lysis buffer. The column was washed with lysis buffer and the protein was eluted with elution buffer (500 mM NaCl, 20 mM Tris-HCl pH 8.0, 250 mM imidazole pH 8.0, 3 mM BME). Each protein was then passed through a desalting column (Desalt 26/10, GE Healthcare), into buffer containing 350 mM NaCl and 20 mM Tris, pH 8. It was further purified via size-exclusion chromatography, using a HiLoad 16/600 Superdex 75 column (GE Healthcare) and the same buffer. The proteins were concentrated using Vivaspin 20 MWCO 30 kDa concentrators (GE Healthcare) and stored at -70°C for enzymatic assays. The purity of each protein was confirmed by SDS-PAGE analysis (Figures S5 and S56) and its identity was confirmed by mass spectrometry (MS).

Macrocyclization reactions

Macrocyclization reactions were prepared for 150 µM peptide substrate and 35 µM enzyme in buffer containing 500 mM NaCl, 10 mM bicine pH 7.5 and 5% DMSO. Reaction mixtures were incubated for 120 h at 30°C. Samples were analysed by LC-ESIMS. Control samples were prepared by incubation of the peptide substrates in the aforementioned buffer. No cyclic peptides were detected in control samples, suggesting there was no spontaneous cyclization occurring independent of the enzyme.

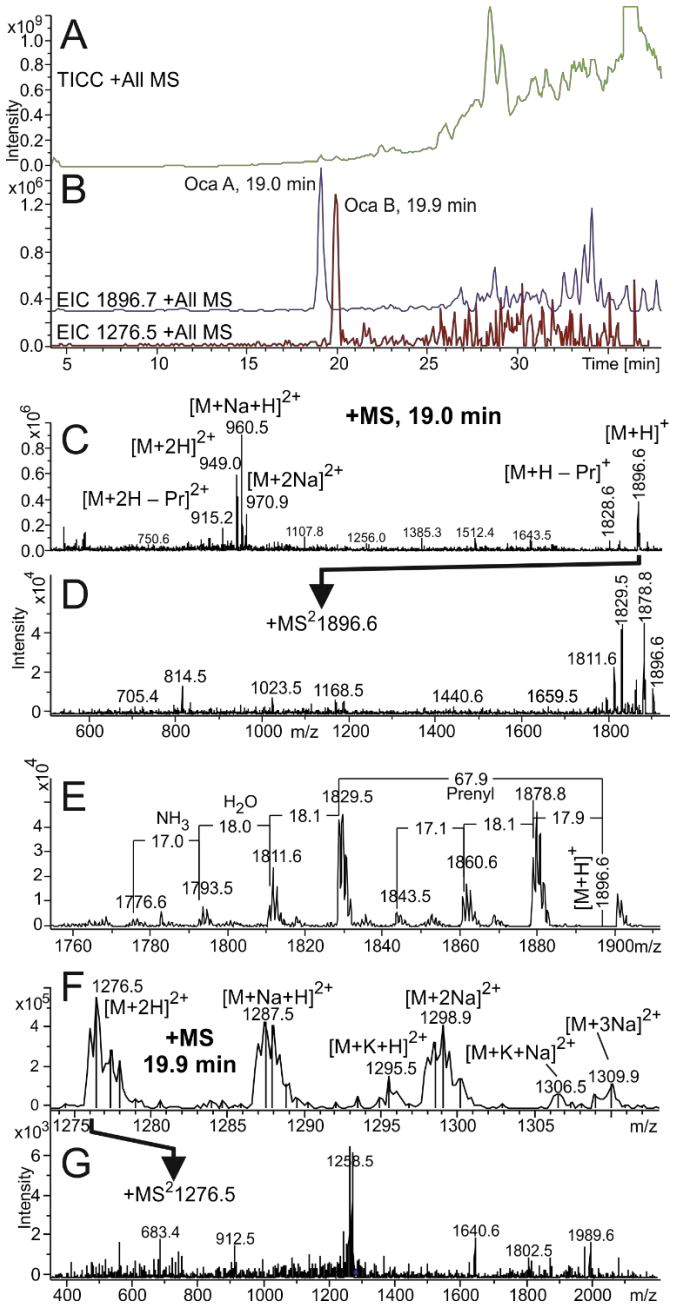
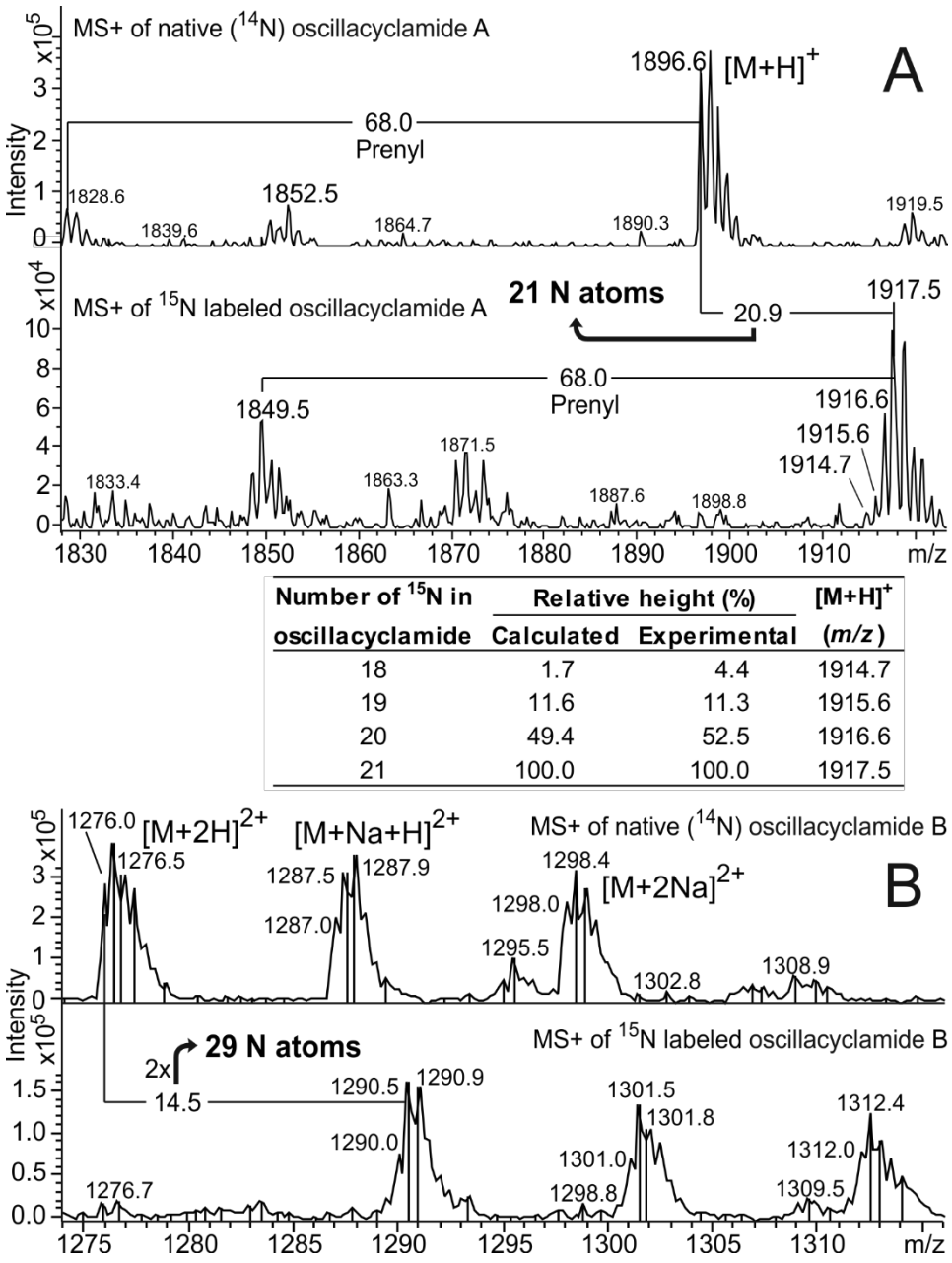


Figure S1. LC-ITMS analysis of methanolic extract from the *Oscillatoria* sp. PCC 6506. A: Total ion current chromatogram (TICC). B: Extracted ion *m/z* 1896.7 and *m/z* 1276.5 chromatogram (EIC) showing the elution of oscillacyclamide A at 19.0 min and oscillacyclamide B at 19.9 min. C: Mass spectrum at 19.0 min showing single and double charged oscillacyclamide A ions with and without a prenyl (Pr) unit. D: Product ion spectrum from the protonated oscillacyclamide *m/z* 1896.6. D: Enlargement of the product ion spectrum showing more clearly the loss of water and prenyl from the protonated oscillacyclamide A. F: The mass spectrum at 19.9 min showing double charged oscillacyclamide B ions. G: The product ion spectrum from the double protonated oscillacyclamide *m/z* 1276.5.



Number of ^{15}N in oscillacyclamide	Relative height (%) Calculated	Experimental	$[M+H]^+$ (*m/z*)
18	1.7	4.4	1914.7
19	11.6	11.3	1915.6
20	49.4	52.5	1916.6
21	100.0	100.0	1917.5

Figure S2. LC-ITMS analysis of native (containing ^{14}N) and labelled (with 98+ atom % ^{15}N urea (ISOTEC) as the only nitrogen source) oscillacyclamides A (A) and B (B). Comparison of the mass spectra shows that the ^{15}N labeling increased the mass of protonated oscillacyclamide A by 20.9 mass units and oscillacyclamide B by 29 mass units which corresponds to 21 and 29 N atoms in oscillacyclamide A and B, respectively. The table analysis shows that fully (*m/z* 1917.5) and partially ^{15}N labeled oscillacyclamide A$[M+H]^+$ peak heights match excellently to the heights of the peaks calculated by assuming 97.7 % labeling degree.

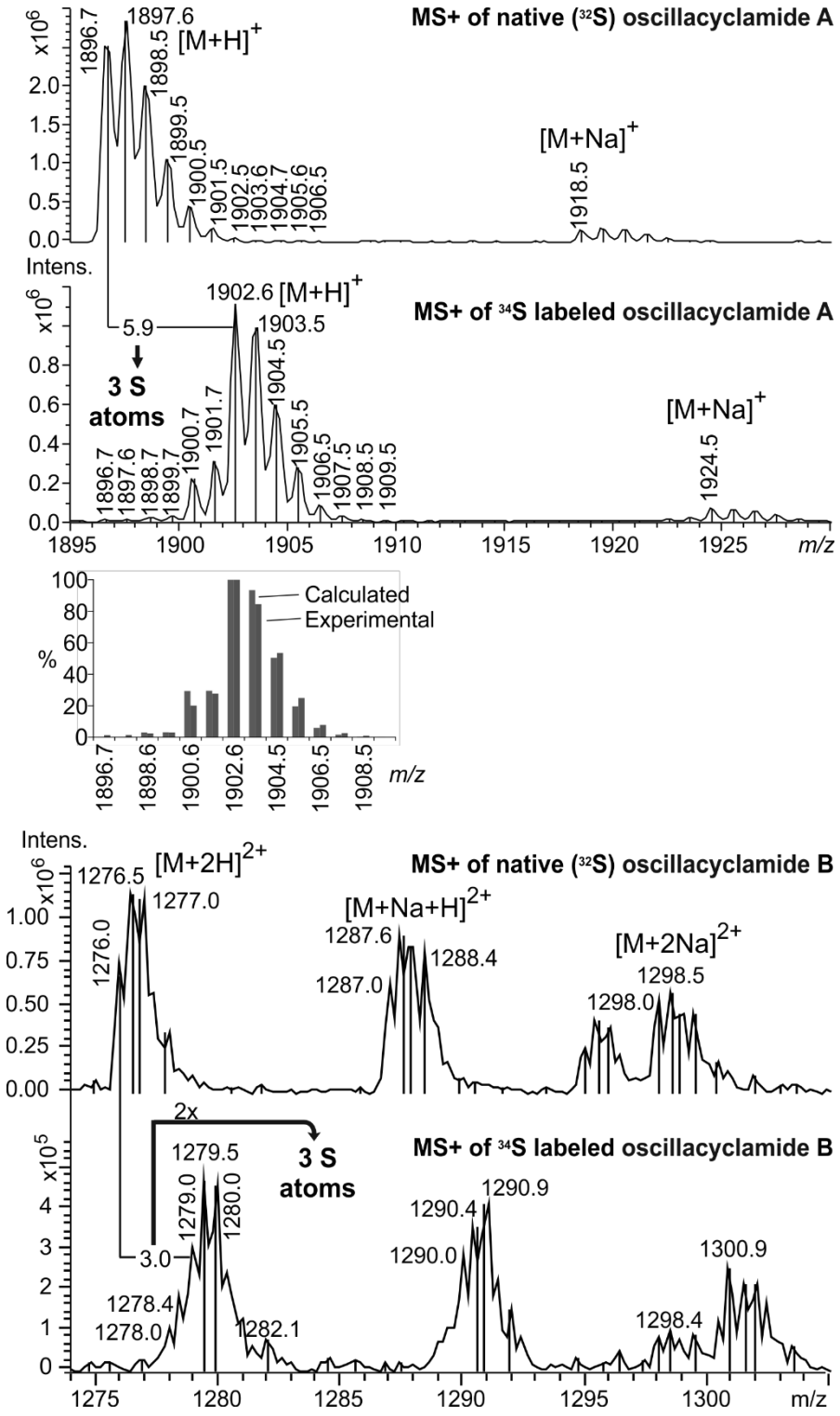


Figure S3. LC-ITMS analysis of native (containing ^{32}S) and labelled (with 90 atom % ^{34}S $MgSO_4$ (ICON) as the only sulphur source) oscillacyclamide A and B. Comparison of the mass spectra shows that the ^{34}S labeling increased the mass of protonated oscillacyclamides by 5.9 – 6.0 mass units which corresponds to 3 S atoms in both oscillacyclamides. The column chart shows that fully (*m/z* 1902.6) and partially ^{15}N labeled oscillacyclamide A $[M+H]^+$ and their isotope peaks heights match the heights of the peaks calculated by assuming 90 % labeling degree.

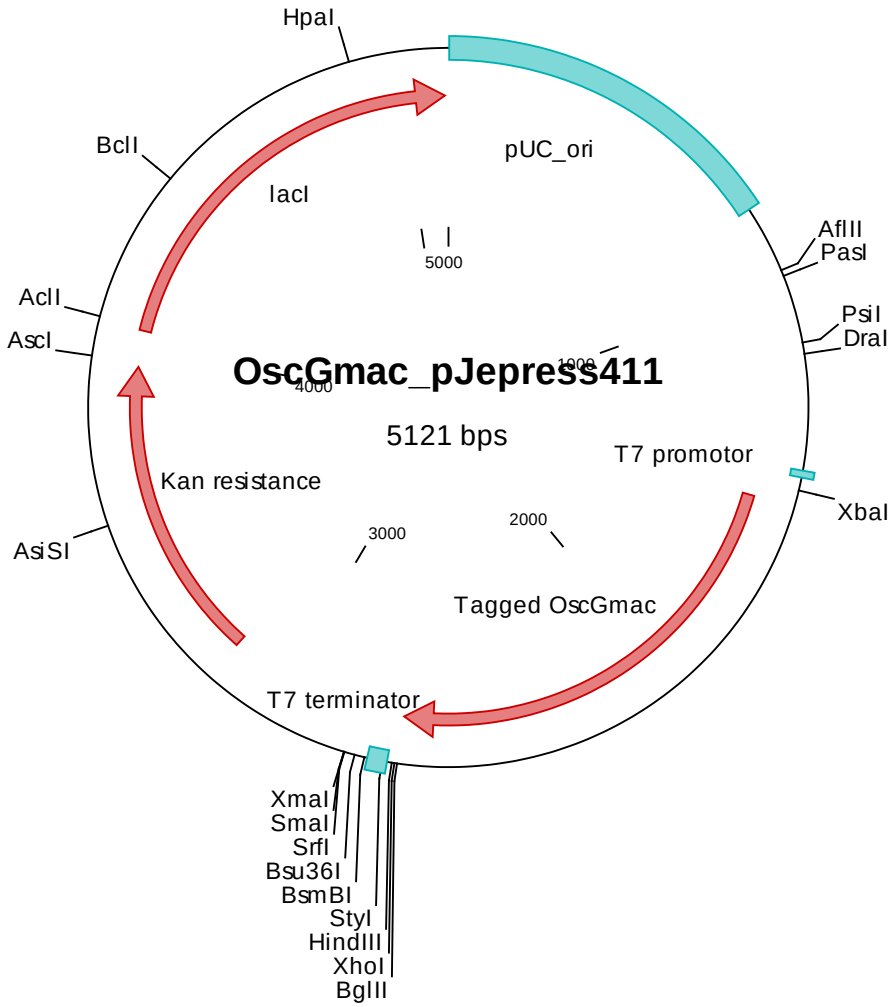


Figure S4: Map of $OscG_{mac}$ construct in pJexpress 411 (Atum, USA).

Codon Optimised DNA sequence of $OscG_{mac}$

ATGCACCATCACCACCATCATGAGAATCTGTATTTCCAGGGTTCCGGCGAGCAAGAGTC
ACGTGAGGCTGGCGAGCAAGGCTCTATTGAGAGCAAGGGTAGCACCAGCTTCACCAG
CAGCAACTTAGTGGTCCCGAATCTGACGAATTTCATTACGCCGAGCGCGGCACCGATG
AACAACGCAGACAAACTGCCTGGTCTGTATGATCTGTGGGCGCATACCAAAGGCGACC
ACGAAATCACGATTGTTATCCTGGATGGTAACGCCGACCTGGAGCGTAGCTGCTTTCA
GGGTGCGAATATCAGCAAGATTTTTCCGTACTGGCATGAAACCCCGGAACCGATCGCG
TTGGAGTACTATGAAGCATTTCTGGAAATCGAGAAATCCGGTGAGAAAGGTGAGGCGA
AAGCGAAAAAGTTGCAAGCGGCCGTGCCGGAAGCGATTCTGAATCGCCTGAAGGGTAA
CTTTCACGCCACCCACATCATCAGCACCATTATTGGTCAGCATGGTTCTCCGGTTCCGG
GCATCGCCCCACGTTGTCGTGCGATCAATATCCCGATTAACACTACTGGTGATAACGGC

GAGTTCATTAGCCCTATCAATCTGACCCGTGCGTTCGAGCTGGCCATGAAACTGGGCG
CGAACATCATTCACTGCGCTGCTTGTTGTGCGACCCAGACGGGCATTGCACACGATTT
GCTGGCACGCGCGGTTAAGAACTGCCAAGATAACAATATCCTGATCGTGGCGCCGACC
GGTAACGACAAAGGTGAGTGCTGGTGCATTCCGGCCATTCTGCCGGGTGTCCTGGGT
GCCGGCATGATGAAAGACAACGGTAAACCGGCGAATTACAGCAATTGGGGTGGCAATT
ACCAGCACGACGGCATTCTGGCACCGGGCGAGAATATTCTGGGTGCGCAGCCAACGA
CCGAAGAAACCAAGCTGAGCCAAGGTACGTCGTGCGCAGCACCGATCGTCACGGGTG
TTAGCGCATTGTTCCTGTCCCTGCAACTGCAGCGTGGCGAGAAGCCAAATGCTGAAGC
TGTGCGCCAGGCAATCCTGAACAGCGCGATCCCGTGTGATCCGGAAGAAATTGAAGAA
CCGGAGCGTTGTTTGCGCGGCAAGCTGAACATTCCGGGTGCGTACCAACTGCTCACCG
GTAAGTAA

His-Tagged OscG$_{mac}$

MHHHHHHENLYFQGSGEQESREAGEQGSIESKGSTSFTSSNLVVPNLTNFITPSAAPMNN
ADKLPGLYDLWAHTKGDHEITIVILDGNADLERSCFQGANISKIFPYWHETPEPIALEYYEAFL
EIEKSGEKGEAKAKKLQAAVPEAIL*NRLKGNFHATHIISTIIGQHGSPVPGIAPRCRAINIPINT
TGDNGEFISPINLTRAFELAMKLGANIIHCAACCATQTGIAHDLLARAVKNCQDNNILIVAPTG
NDKGECWCIPAILPGVLGAGMMKDNGKPANYSNWGGNYQHDGILAPGENILGAQPTTEET
KLSQGTSCAAPIVTGVSALFLSLQLQRGEKPNAEAVRQAILNSAIPCDPEEIEEPERCLRGKL
NIPGAYQLLTGK

*Underlined sequence in red is the double helix insertion

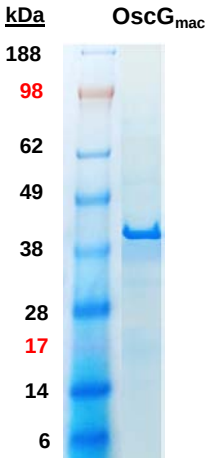


Figure S5: SDS PAGE gel showing the purified OscG$_{mac}$ protein against SeeBlue® Plus2 Pre-Stained protein standard (left lane)

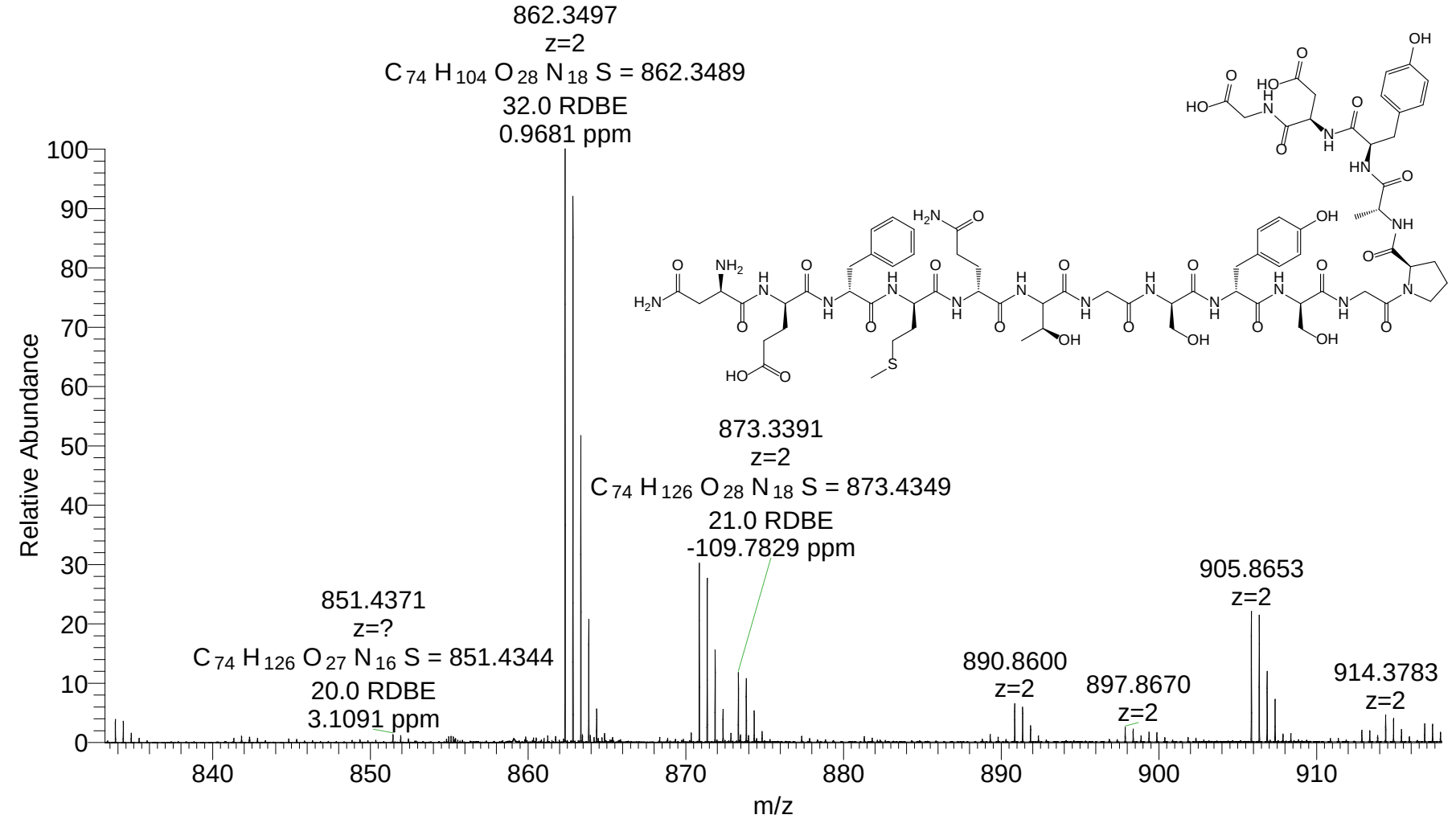


Figure S6: HR LCMS of the peptide substrate H-Asn-Glu-Phe-Met-Gln-Thr-Gly-Ser-Tyr-Ser-Gly-Pro-Ala-Tyr-Asp-Gly-OH (1).

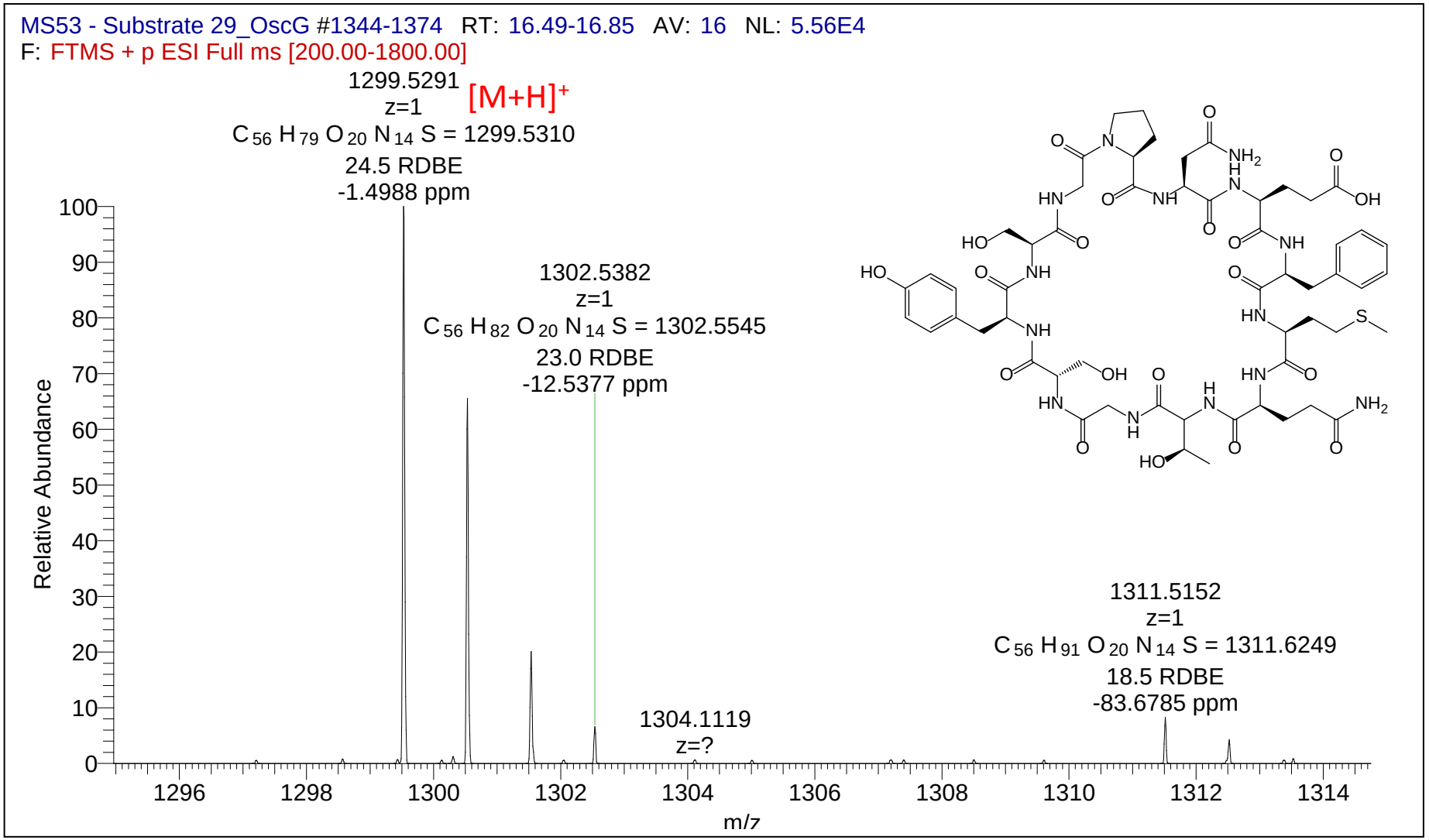


Figure S7: HR LCMS of the cyclic peptide product cyclo-[Asn-Glu-Phe-Met-Gln-Thr-Gly-Ser-Tyr-Ser-Gly-Pro-] (1) generated with $OscG_{mac}$.

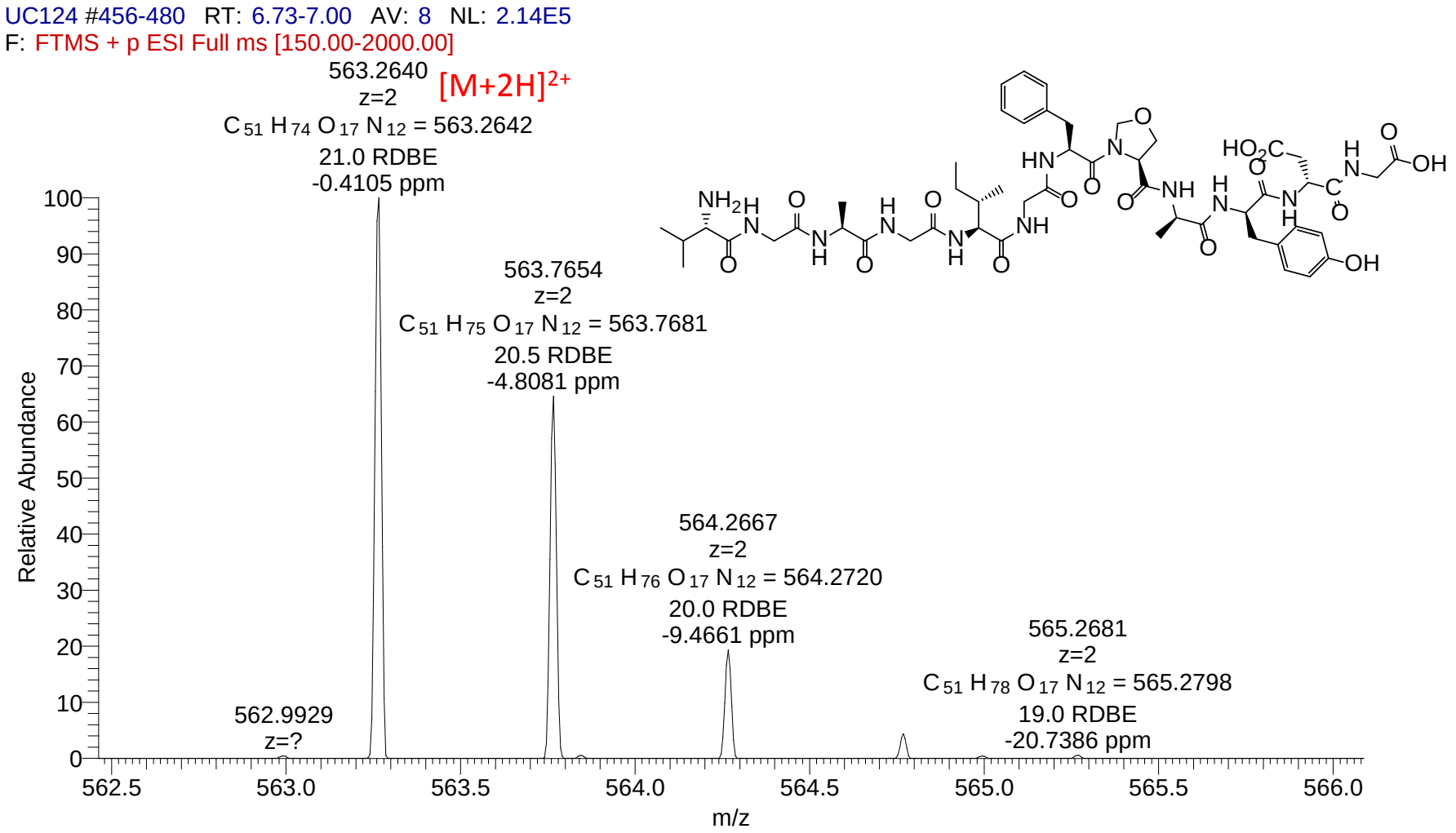


Figure S8: HR LCMS of the peptide substrate H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Ser(ψPro)-Ala-Tyr-Asp-Gly-OH (2).

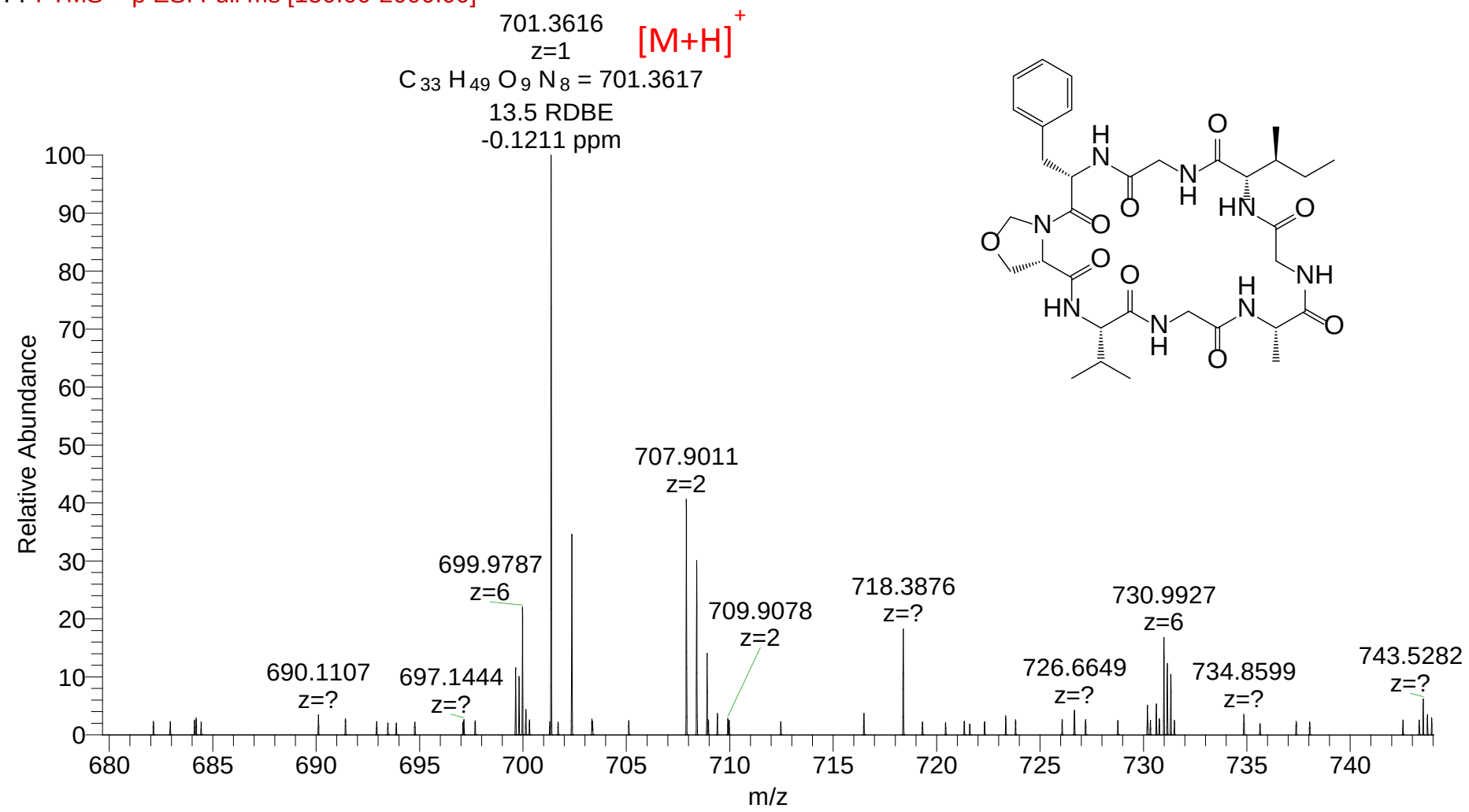


Figure S9: HR LCMS of the cyclic peptide product cyclo-[Val-Gly-Ala-Gly-Ile-Gly-Phe-Ser(ψPro)-] (2) generated with $OscG_{mac}$.

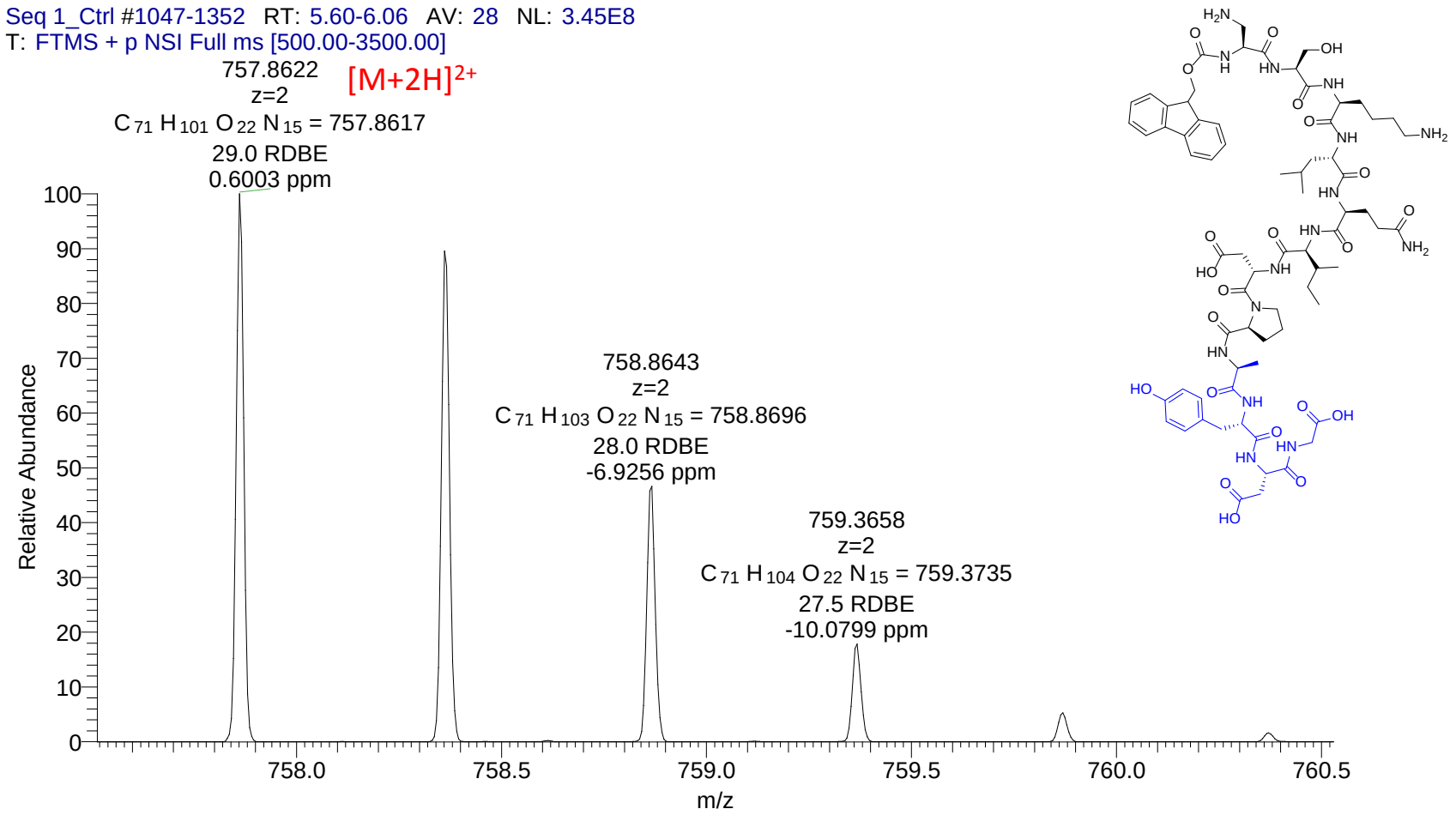


Figure S10: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH (3).

Seq 1_Ctrl #1245-5667 RT: 5.89-26.25 AV: 17 NL: 4.36E7
T: Average spectrum MS2 757.86 (1245-5667)

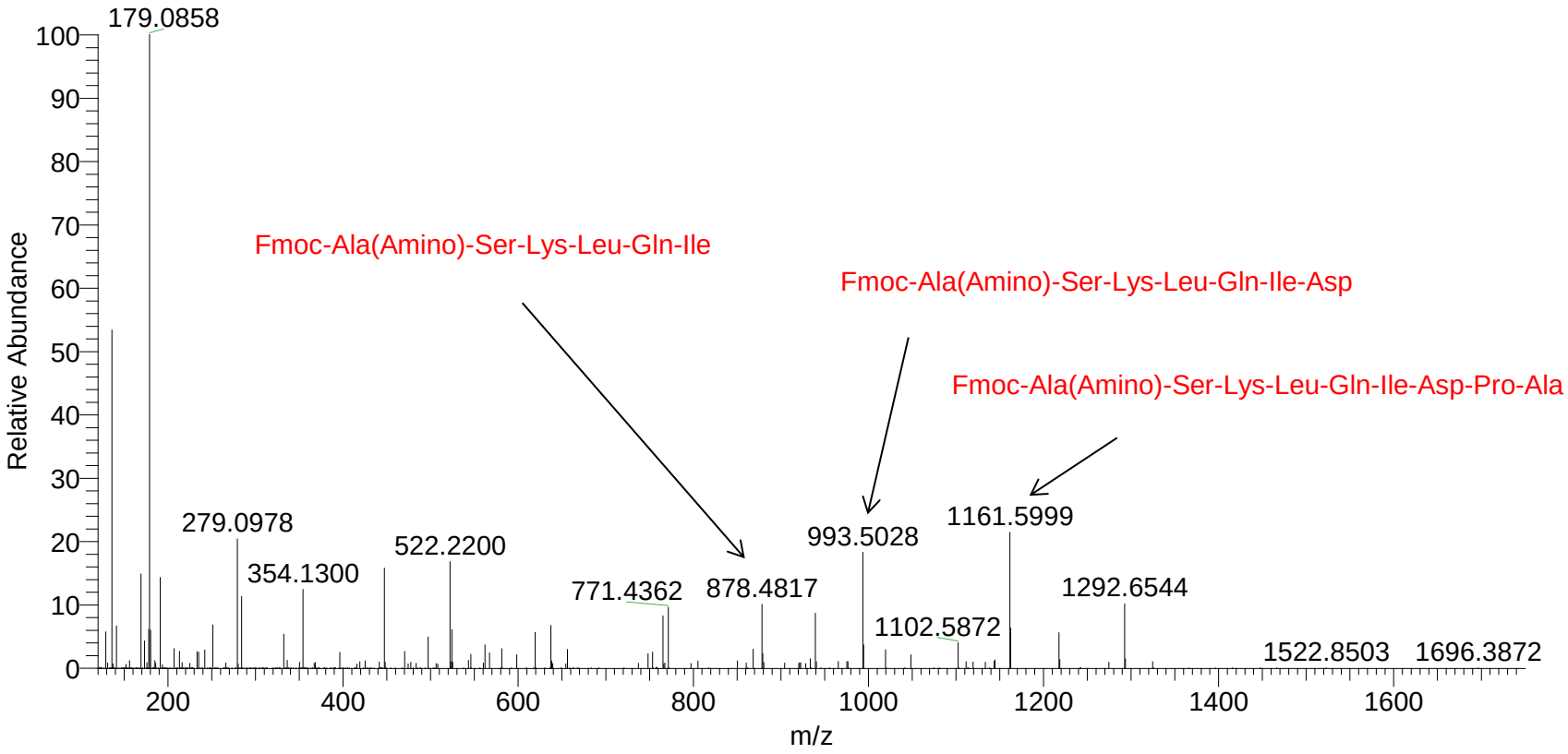


Figure S11: MS MS of the peptide substrate Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH (3).

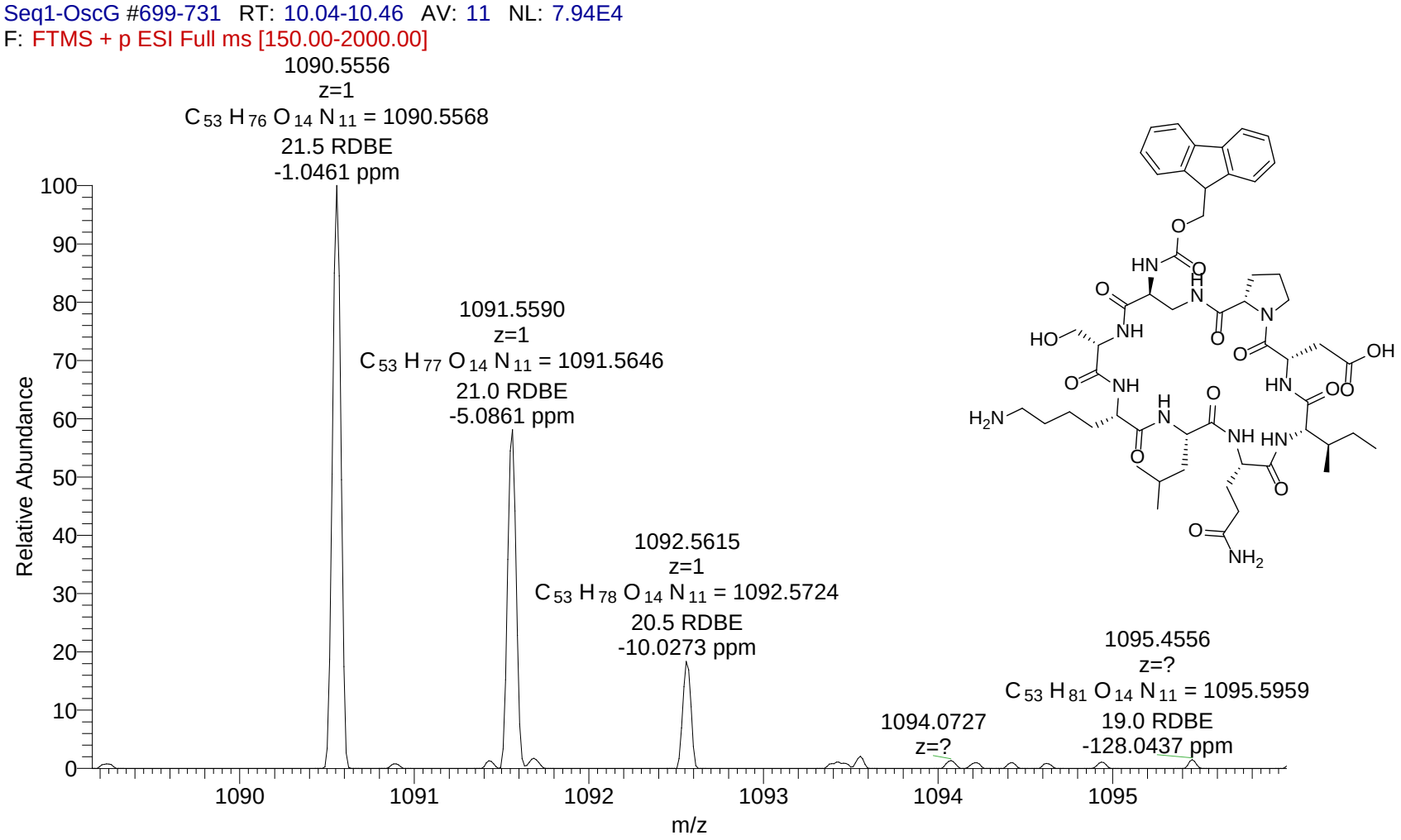


Figure S12: HR LCMS of the cyclic peptide product Cyclo-[Dpr(Fmoc)-Ser-Lys-Leu-Gln-Ile-Asp-Pro-] (3) generated with $OscG_{mac}$.

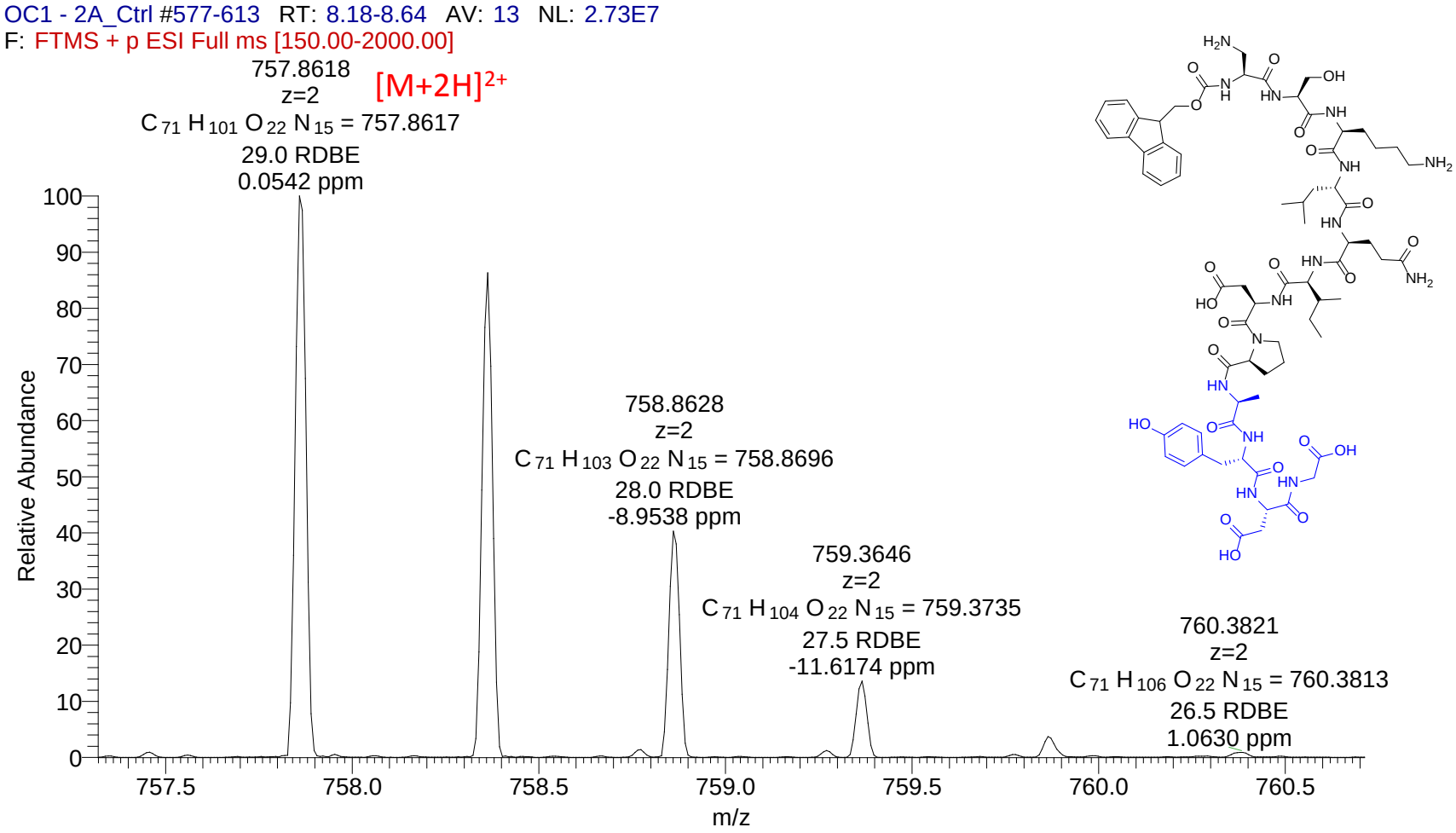


Figure S13: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH (4).

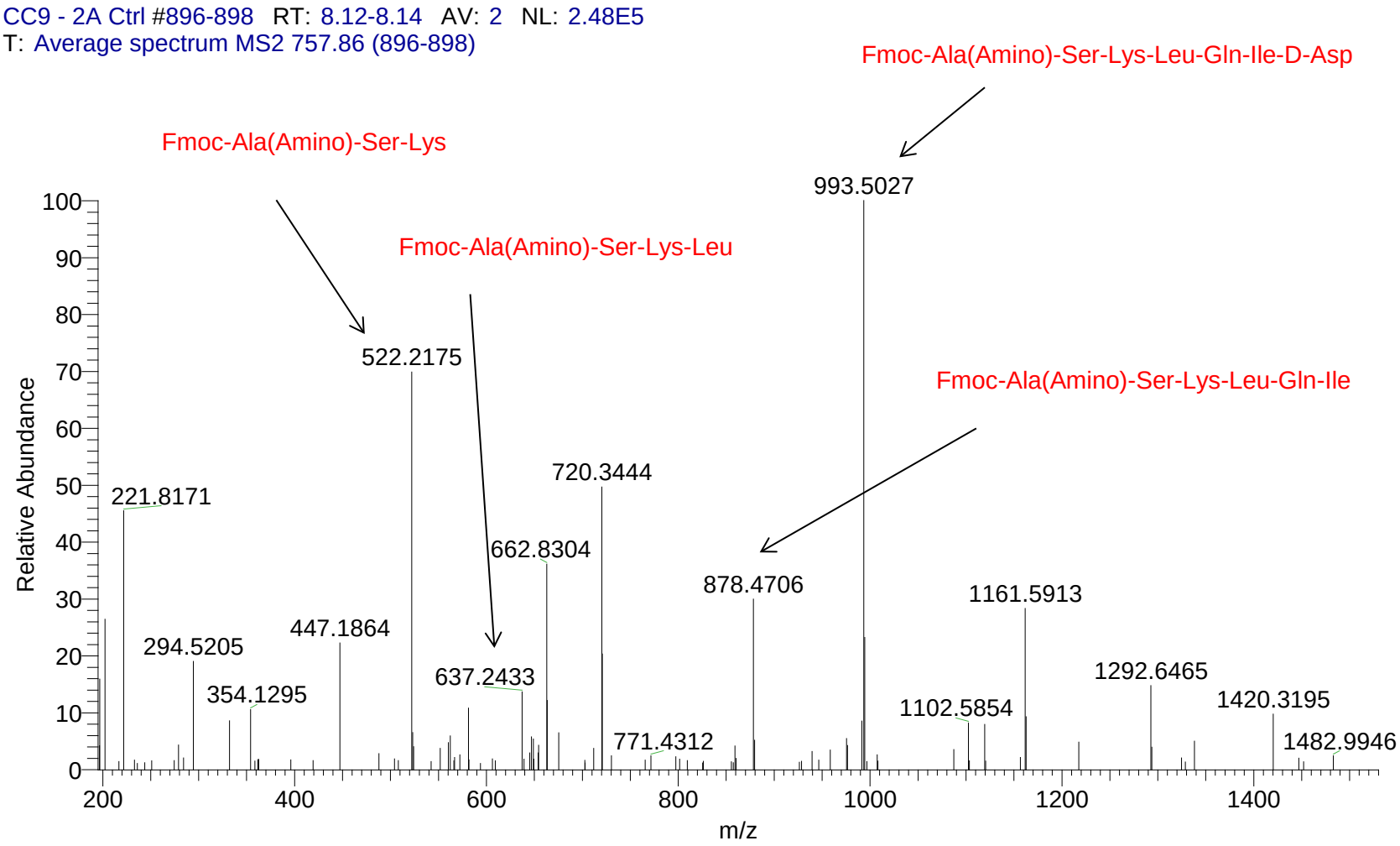


Figure S14: MS MS of the peptide substrate Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH (4).

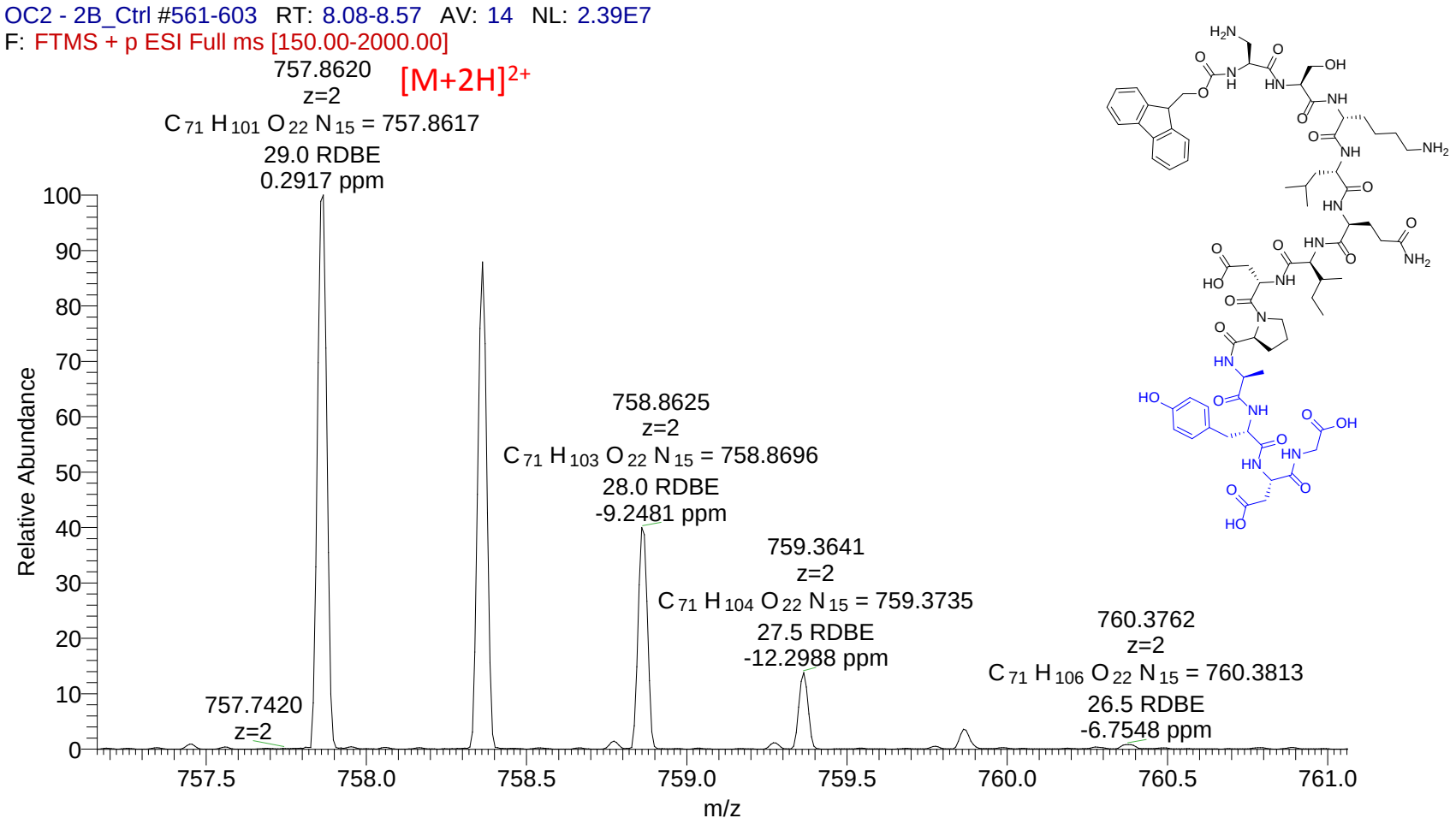


Figure S15: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH (5).

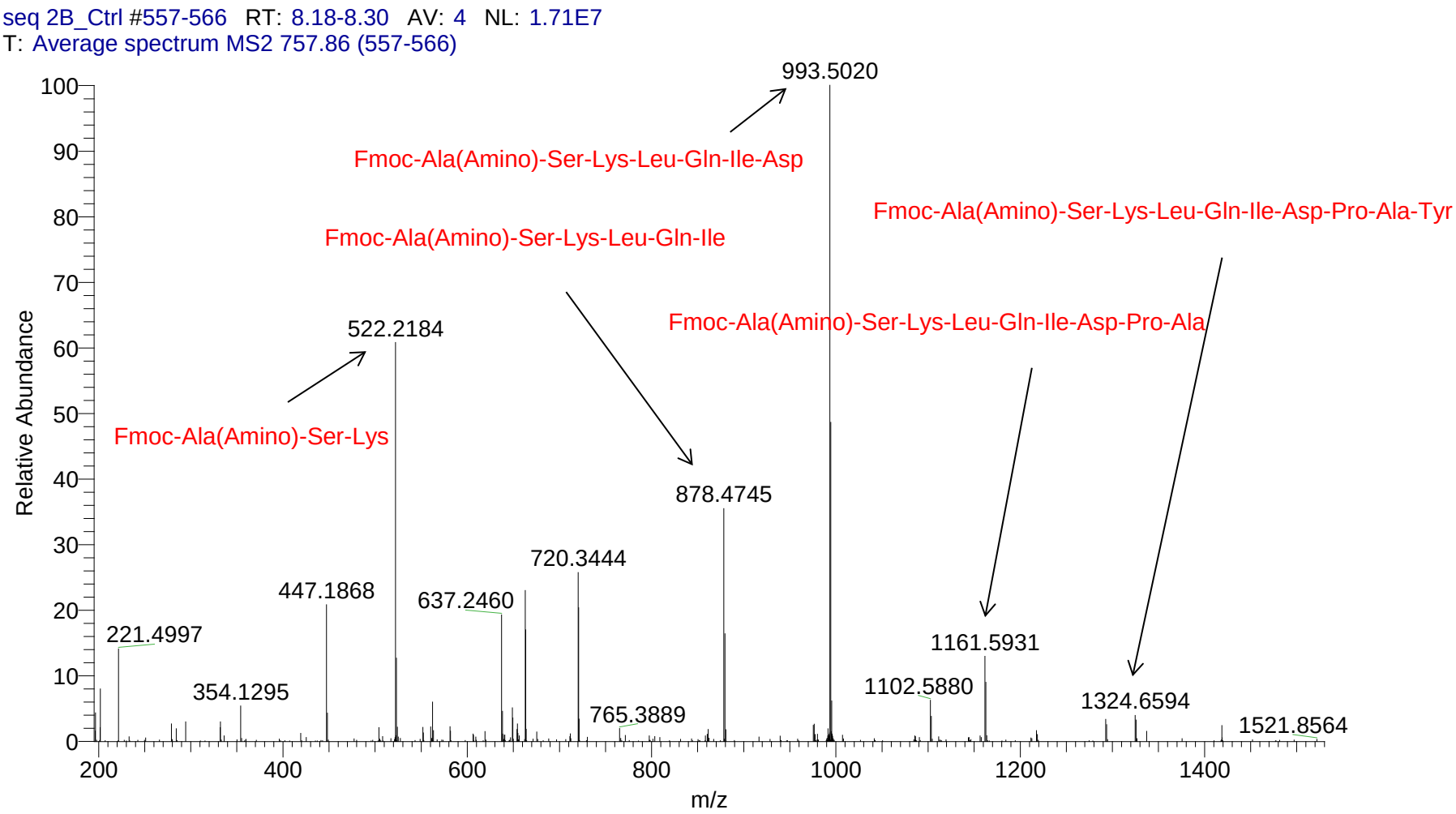


Figure S16: MS MS of the peptide substrate Fmoc-Ala(Amino)-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH (5).

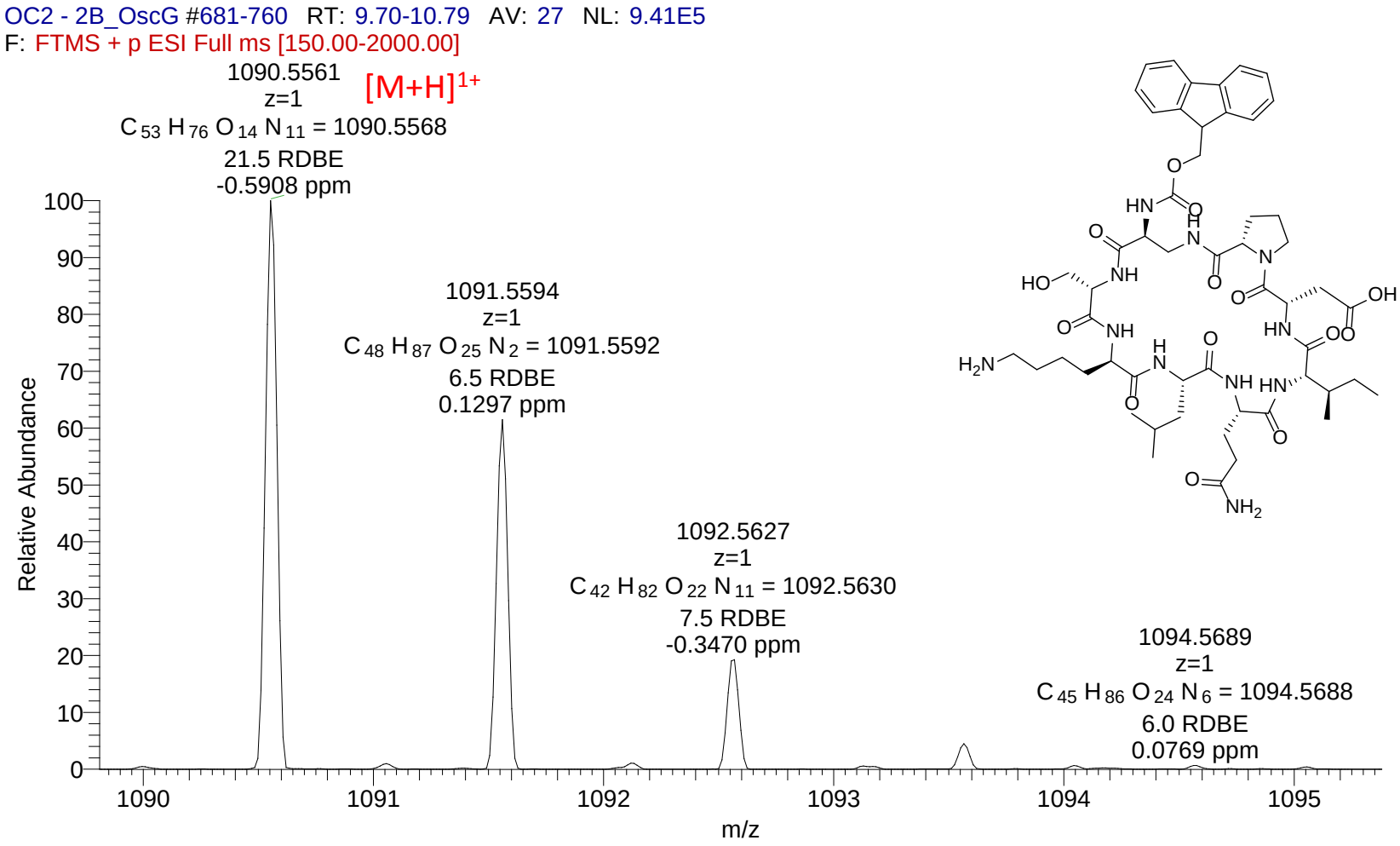


Figure S17: HR LCMS of the cyclic peptide product cyclo-[Dpr(Fmoc)-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-] (5) generated with $OscG_{mac}$.

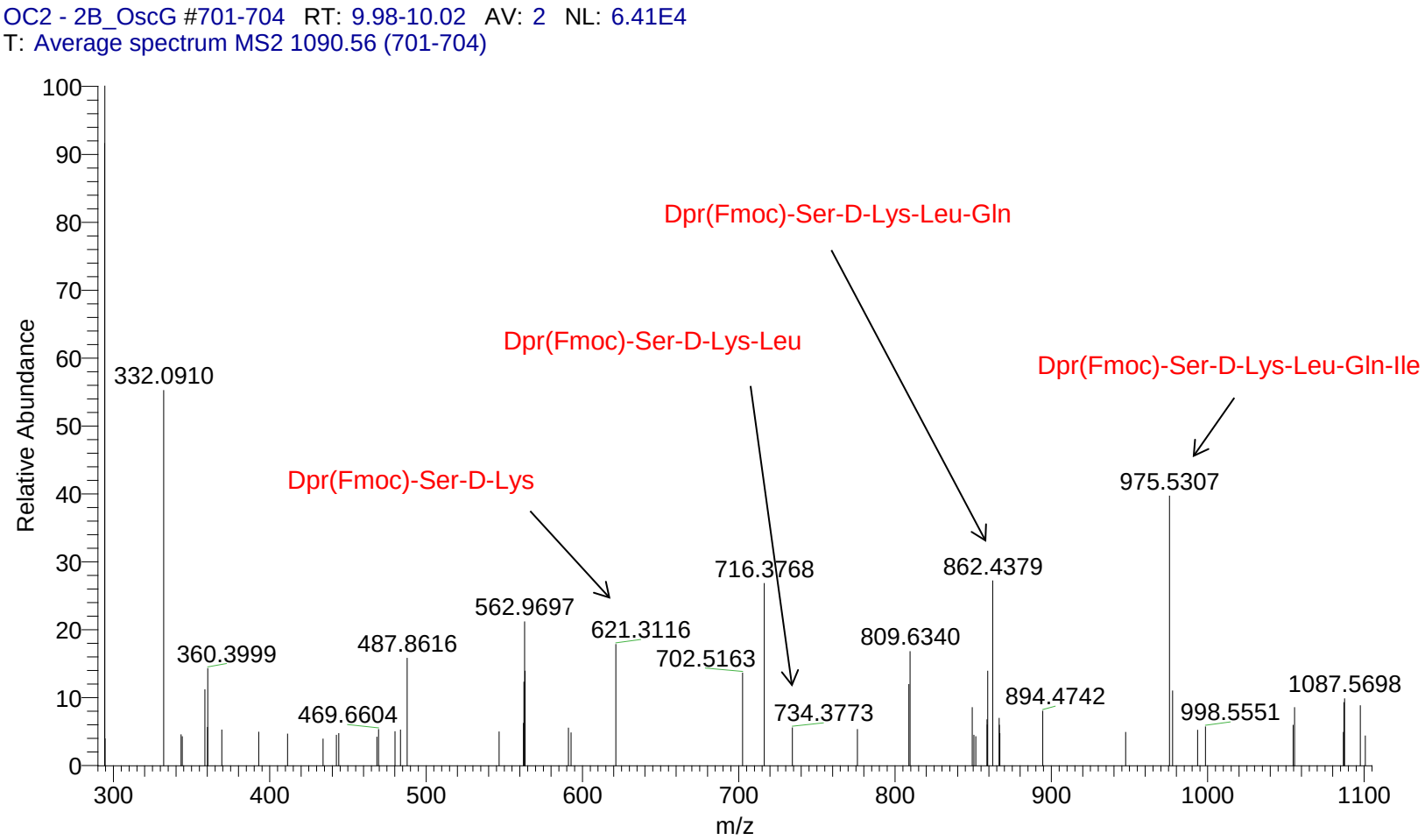


Figure S18: MS MS of the cyclic peptide product cyclo-[Dpr(Fmoc)-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-] (5).

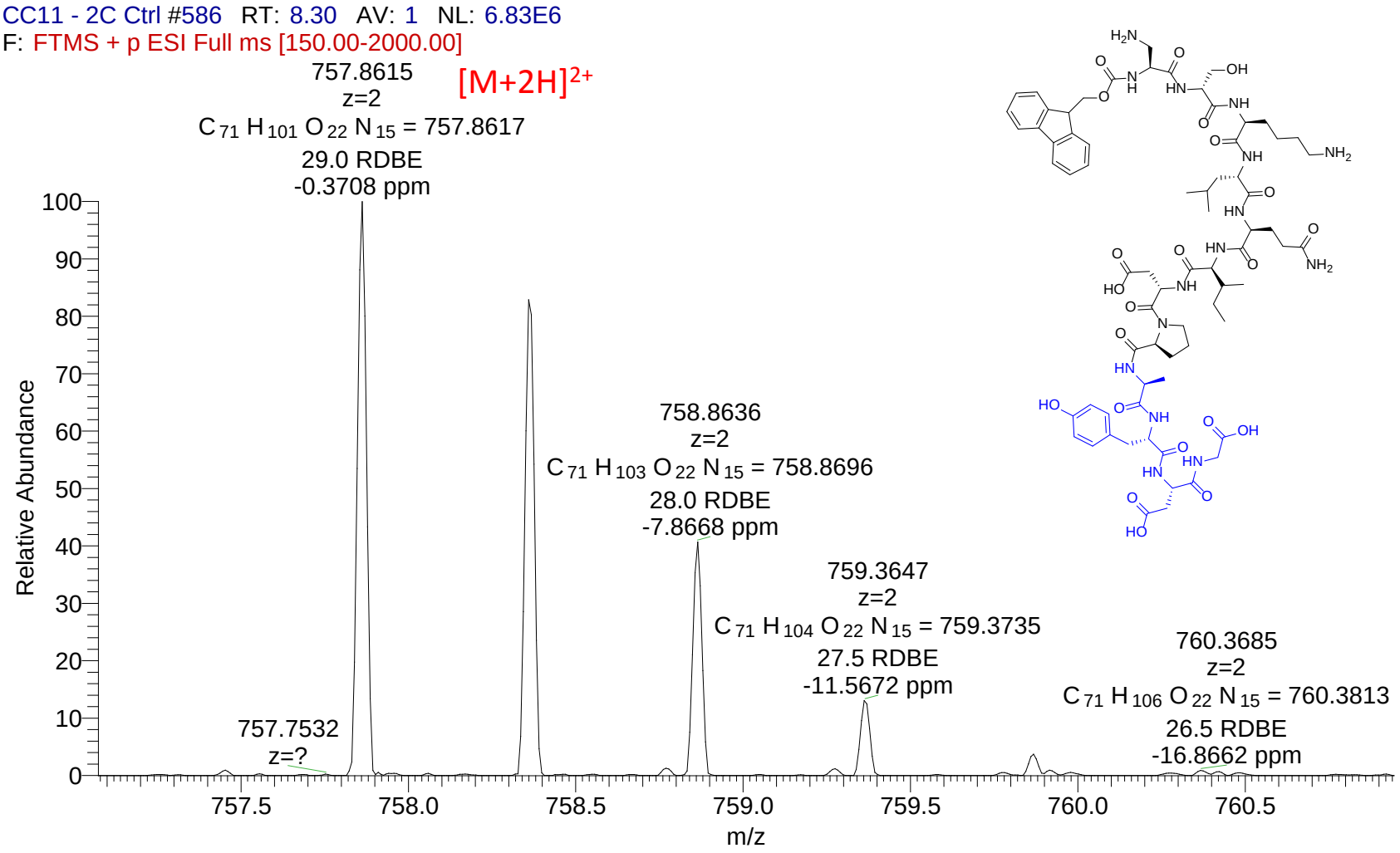


Figure S19: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-D-Ser-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH (6).

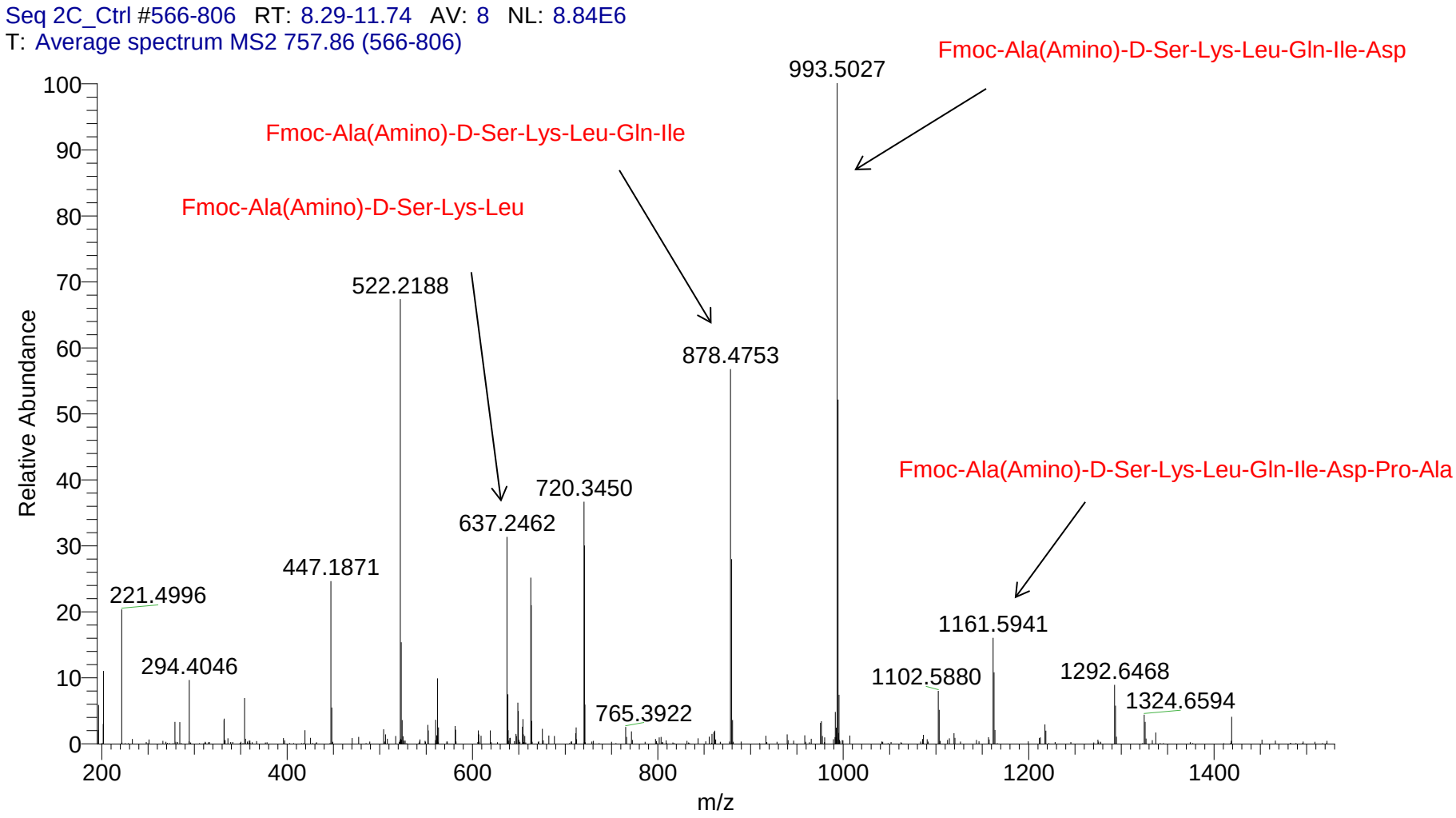


Figure S20: MS MS of the peptide substrate Fmoc-Ala(Amino)-D-Ser-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH (6).

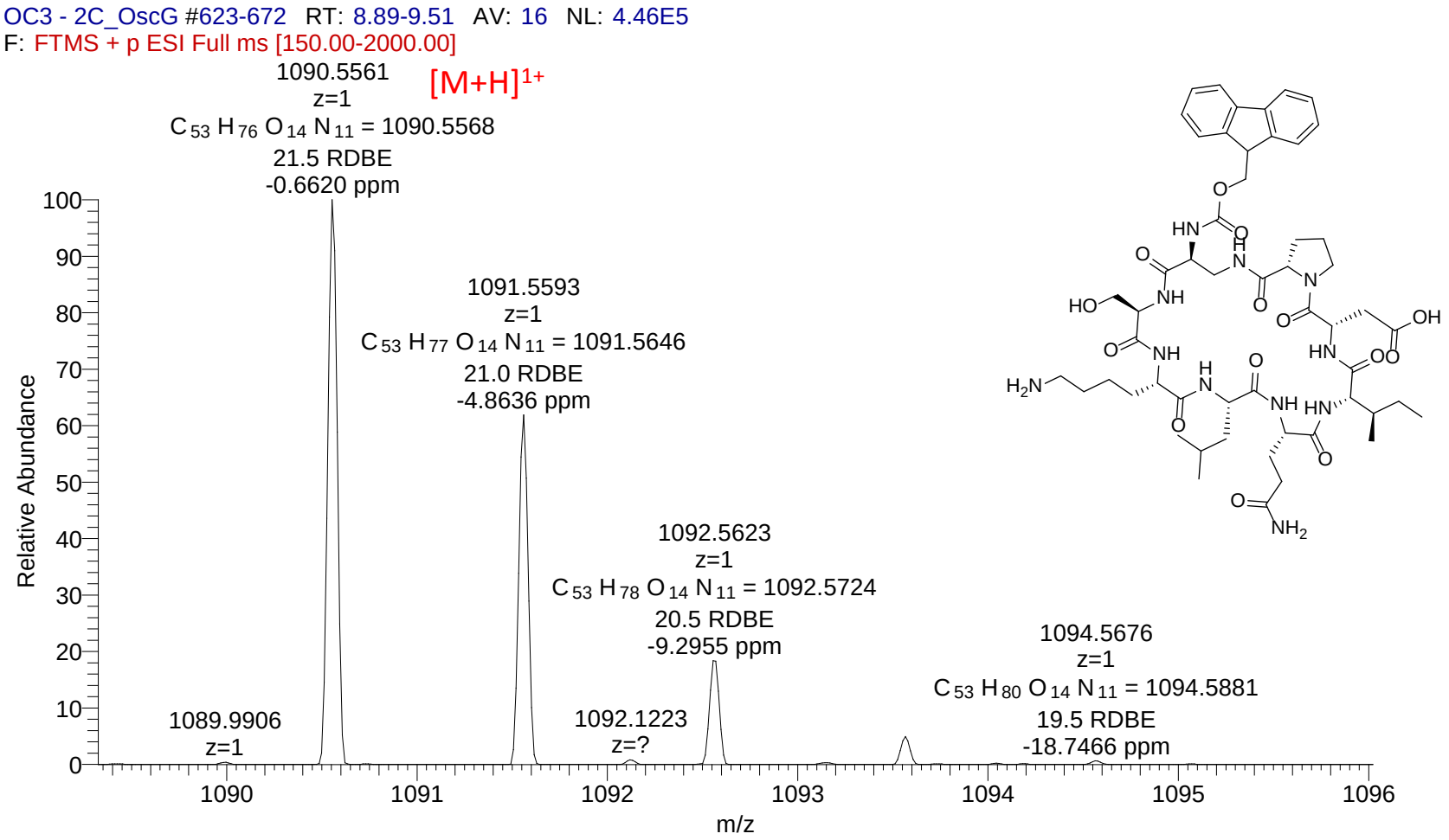


Figure S21: HR LCMS of the cyclic peptide product cyclo-[Dpr(Fmoc)-D-Ser-Lys-Leu-Gln-Ile-Asp-Pro-] (6) generated with $OscG_{mac}$.

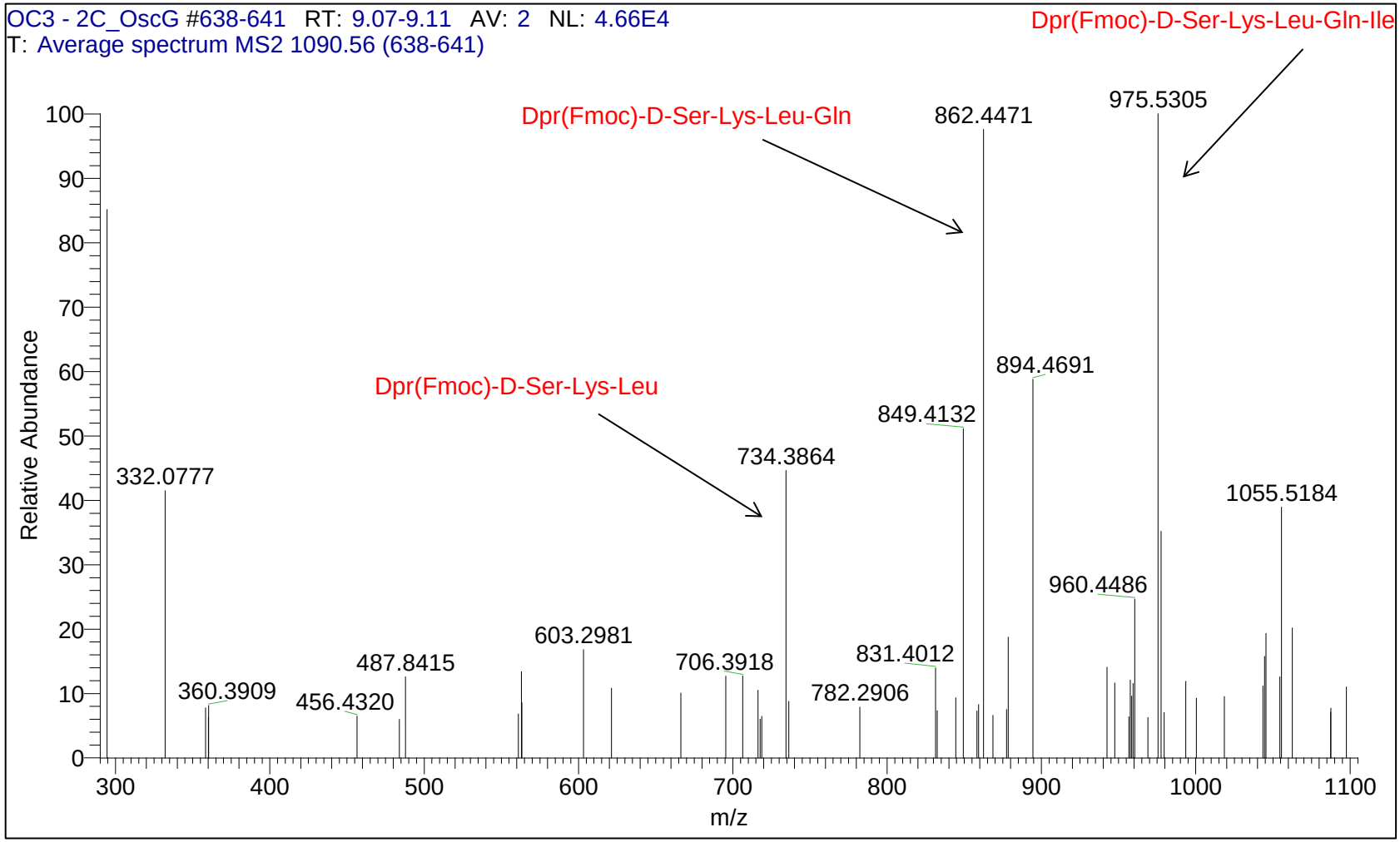


Figure S22: MS MS of the cyclic peptide product cyclo-[Dpr(Fmoc)-D-Ser-Lys-Leu-Gln-Ile-Asp-Pro-] (6).

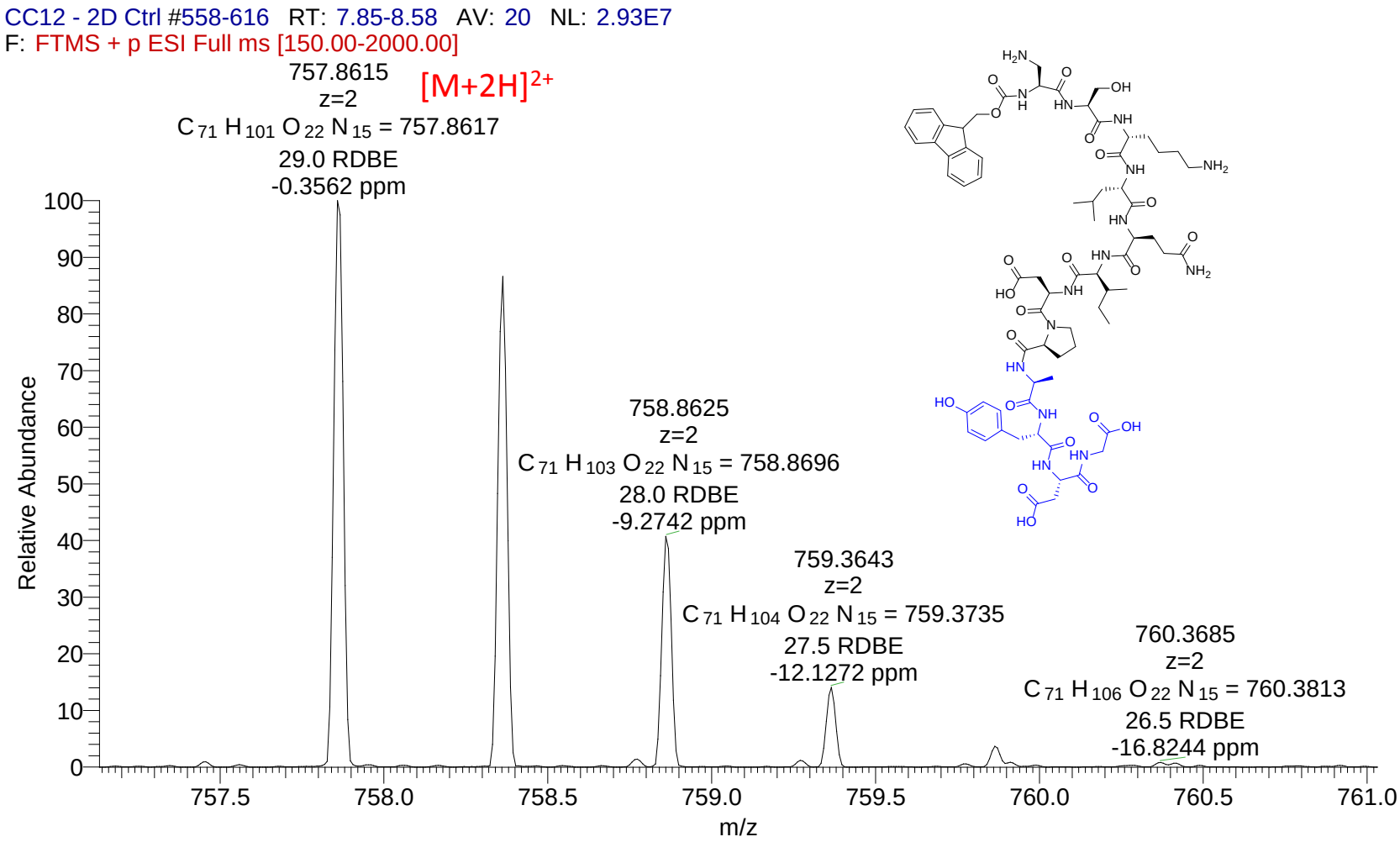


Figure S23: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-Ser-D-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH (7).

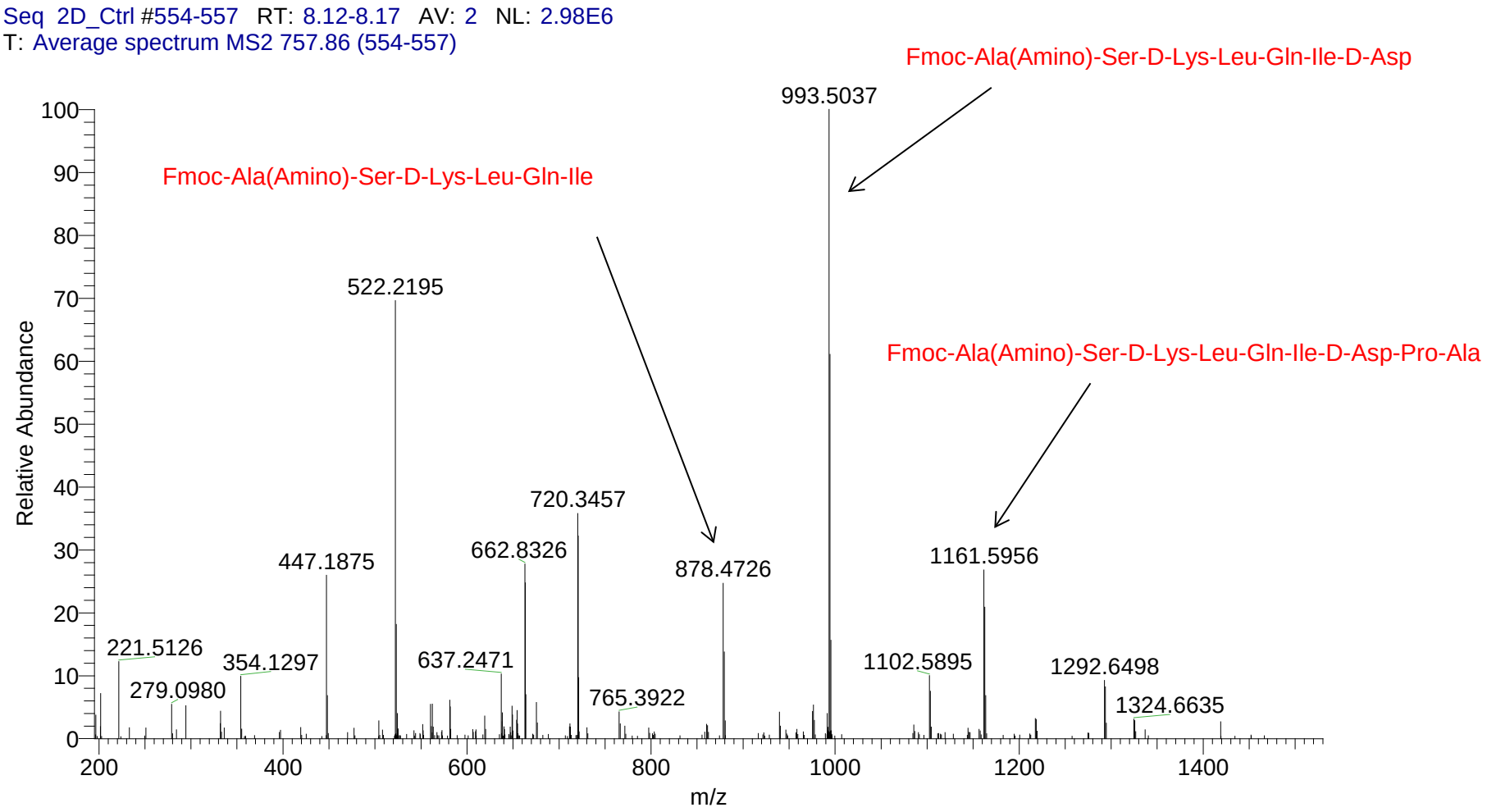


Figure S24: MS MS of the peptide substrate Fmoc-Ala(Amino)-Ser-D-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH (7).

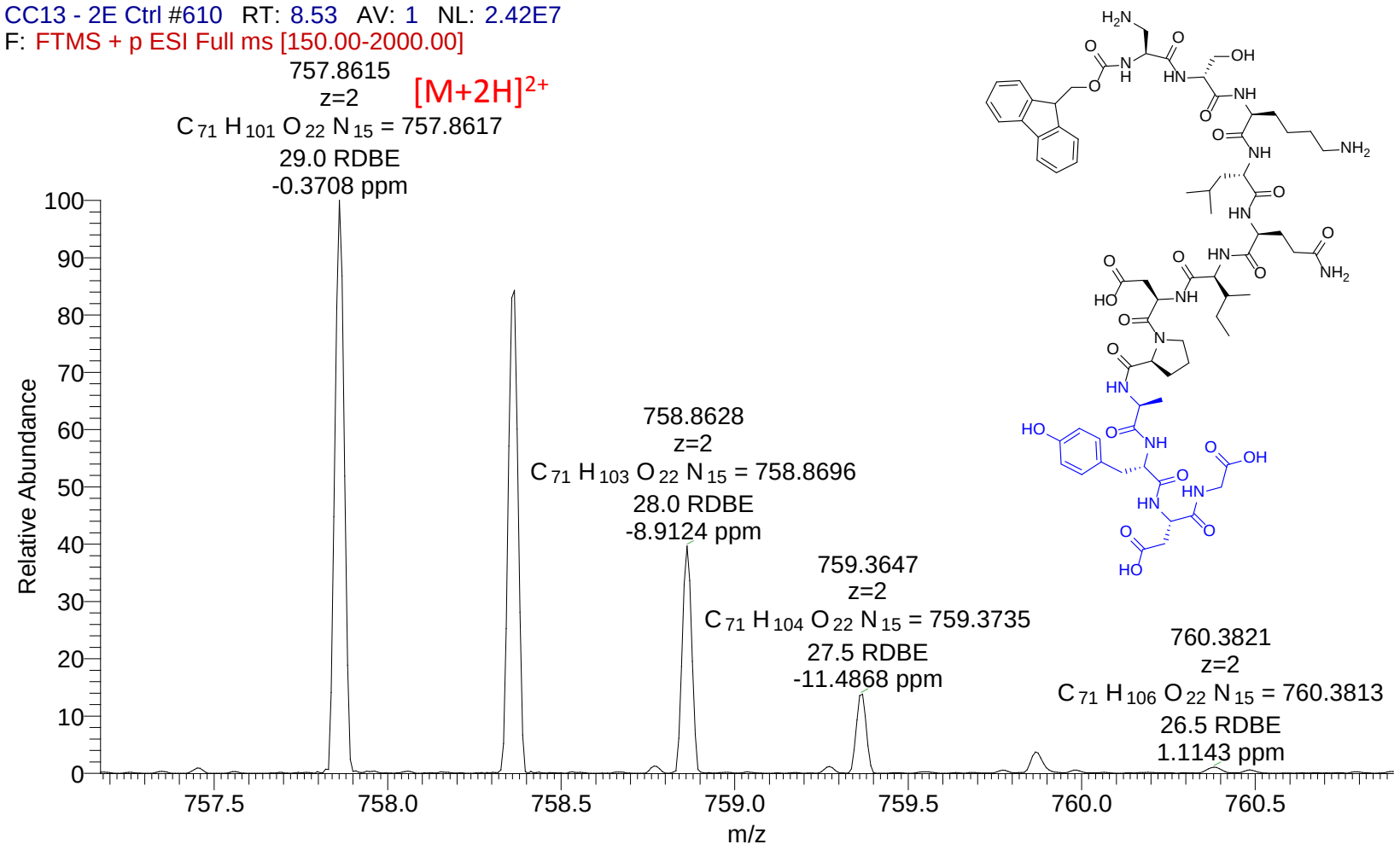


Figure S25: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-D-Ser-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH (8).

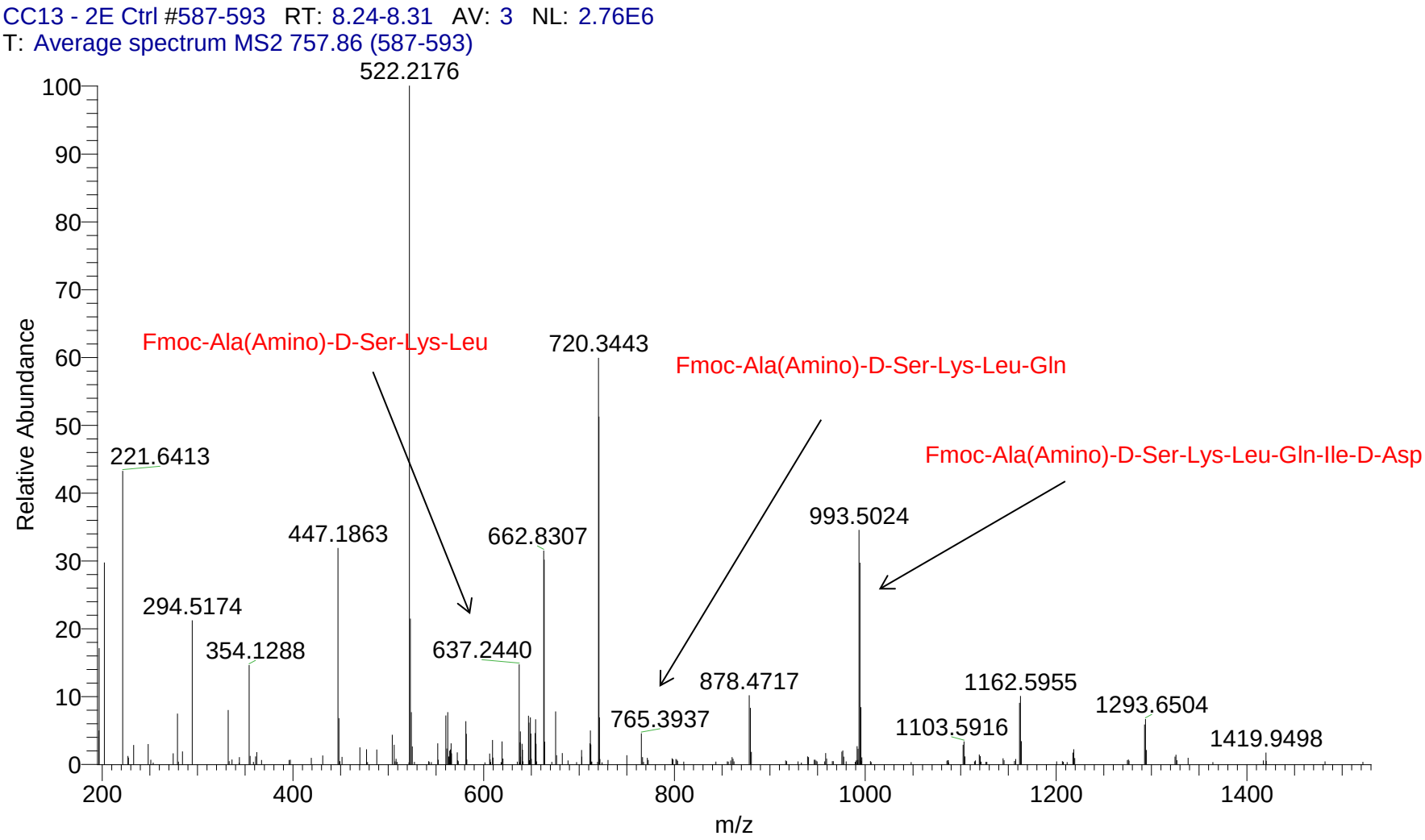


Figure S26: MS MS of the peptide substrate Fmoc-Ala(Amino)-D-Ser-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH (8).

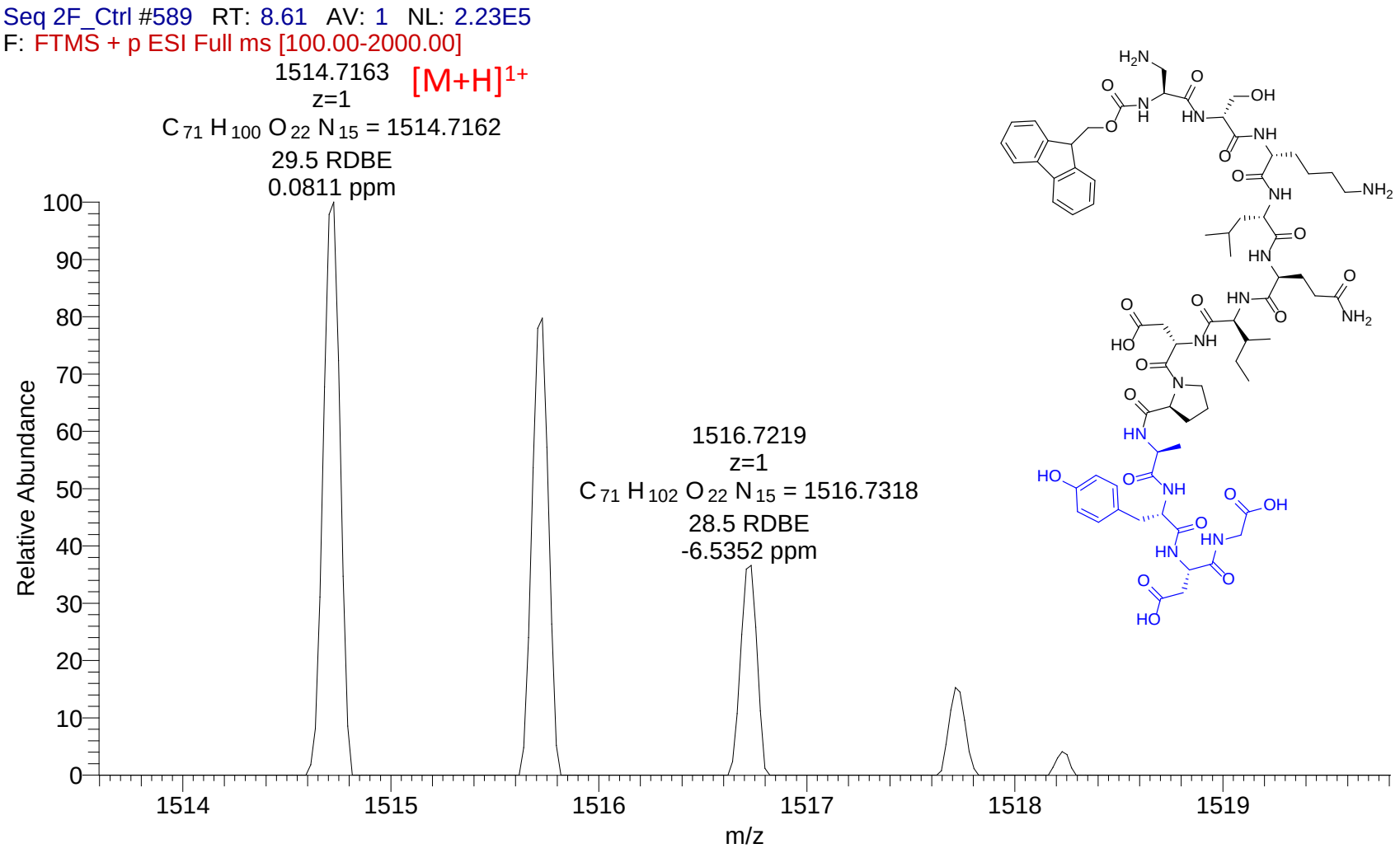


Figure S27: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-D-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH (9).

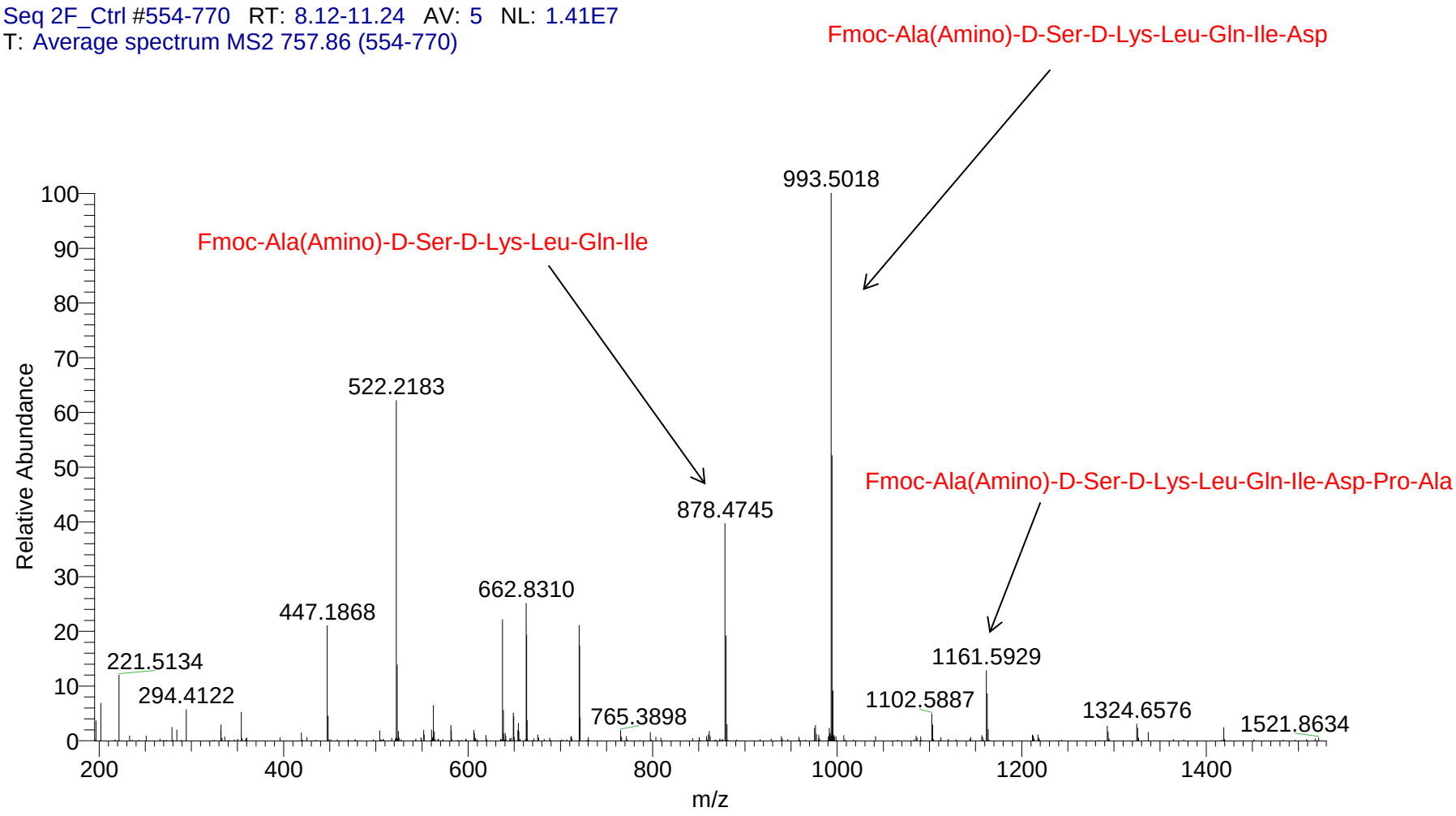


Figure S28: MS MS of the peptide substrate Fmoc-Ala(Amino)-D-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-Ala-Tyr-Asp-Gly-OH (9).

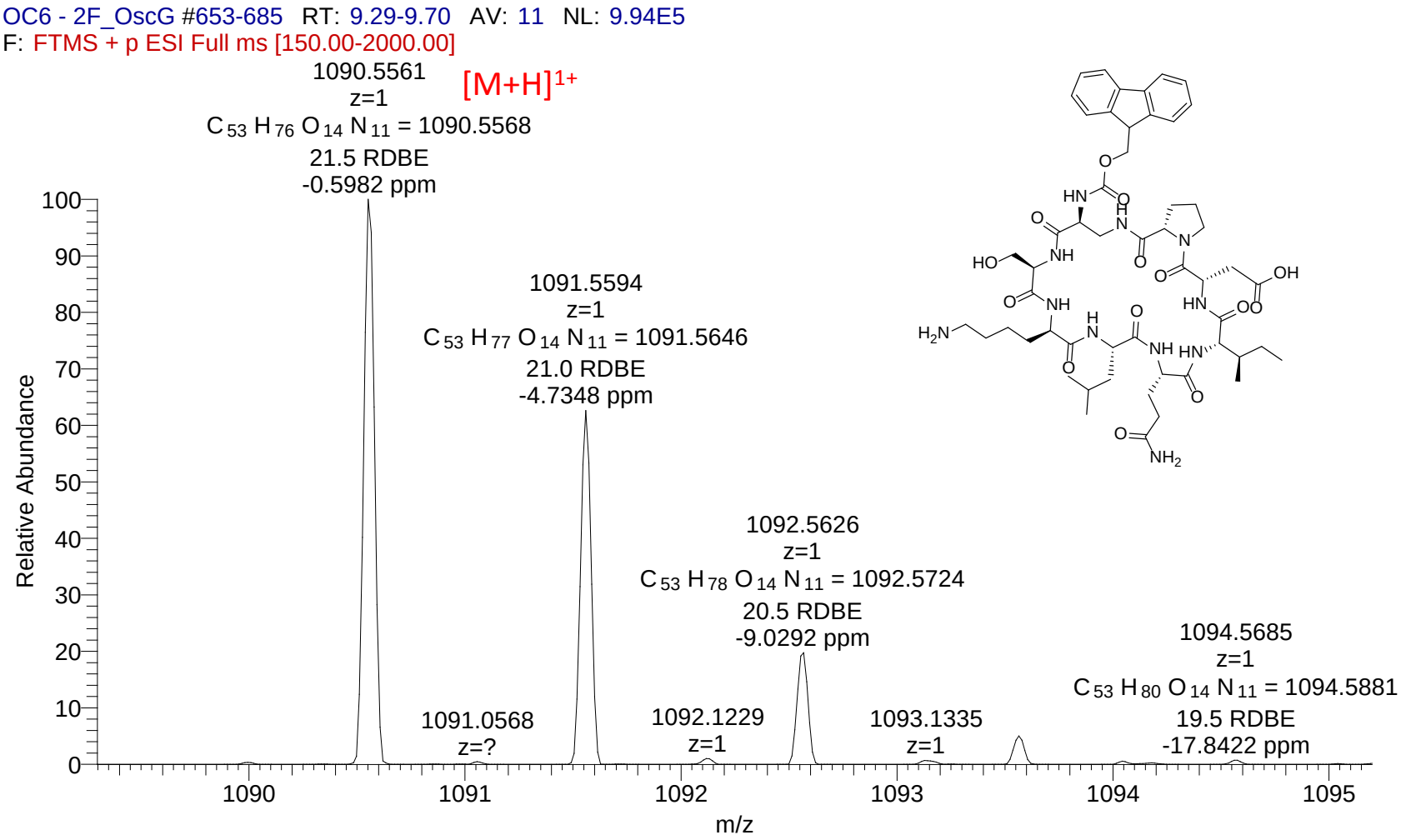


Figure S29: HR LCMS of the cyclic peptide product cyclo-[Dpr(Fmoc)-D-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-] (9) generated with $OscG_{mac}$.

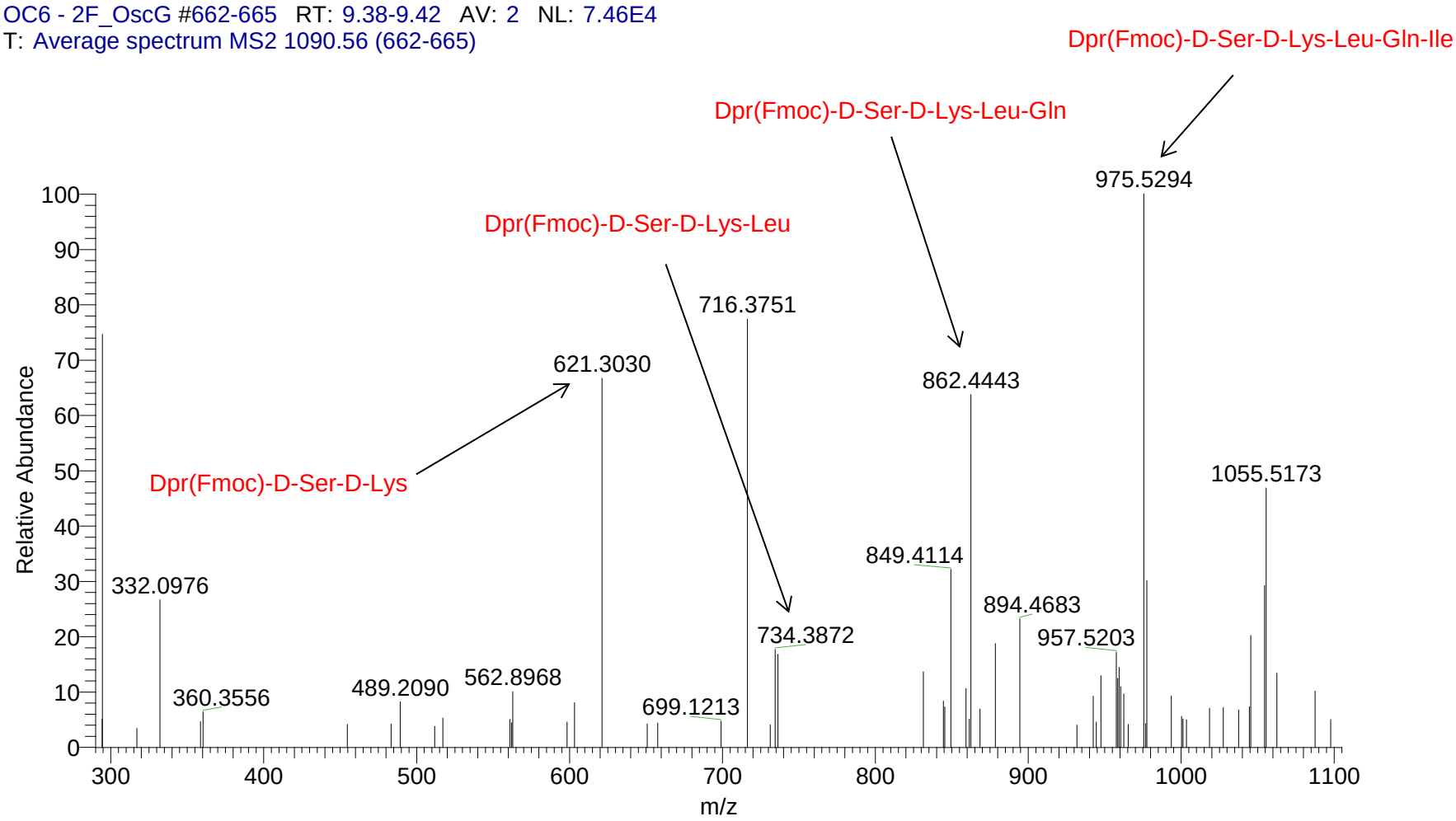


Figure S30: MS MS of the cyclic peptide product cyclo-[Dpr(Fmoc)-D-Ser-D-Lys-Leu-Gln-Ile-Asp-Pro-] (9).

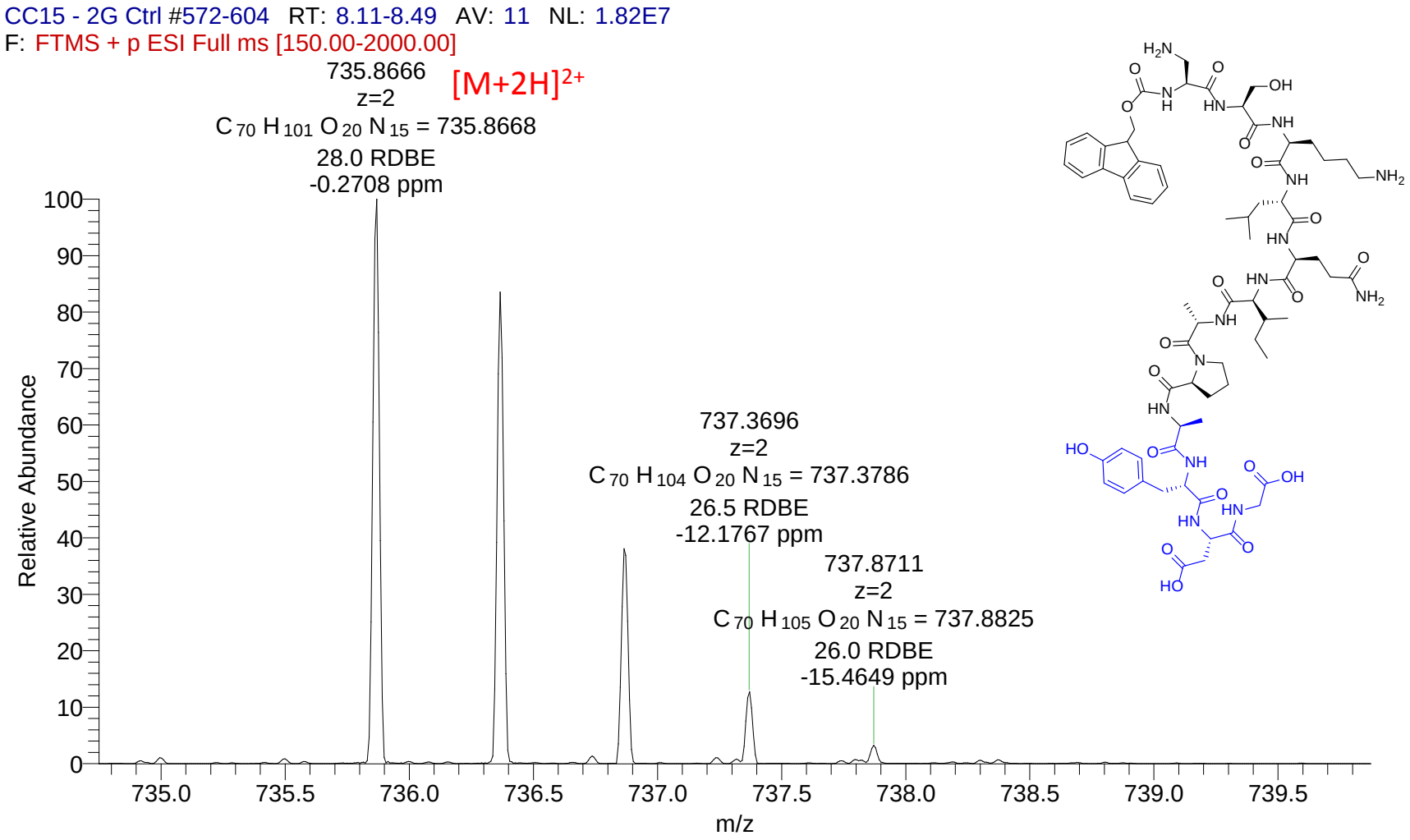


Figure S31: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-Ala-Pro-Ala-Tyr-Asp-Gly-OH (10).

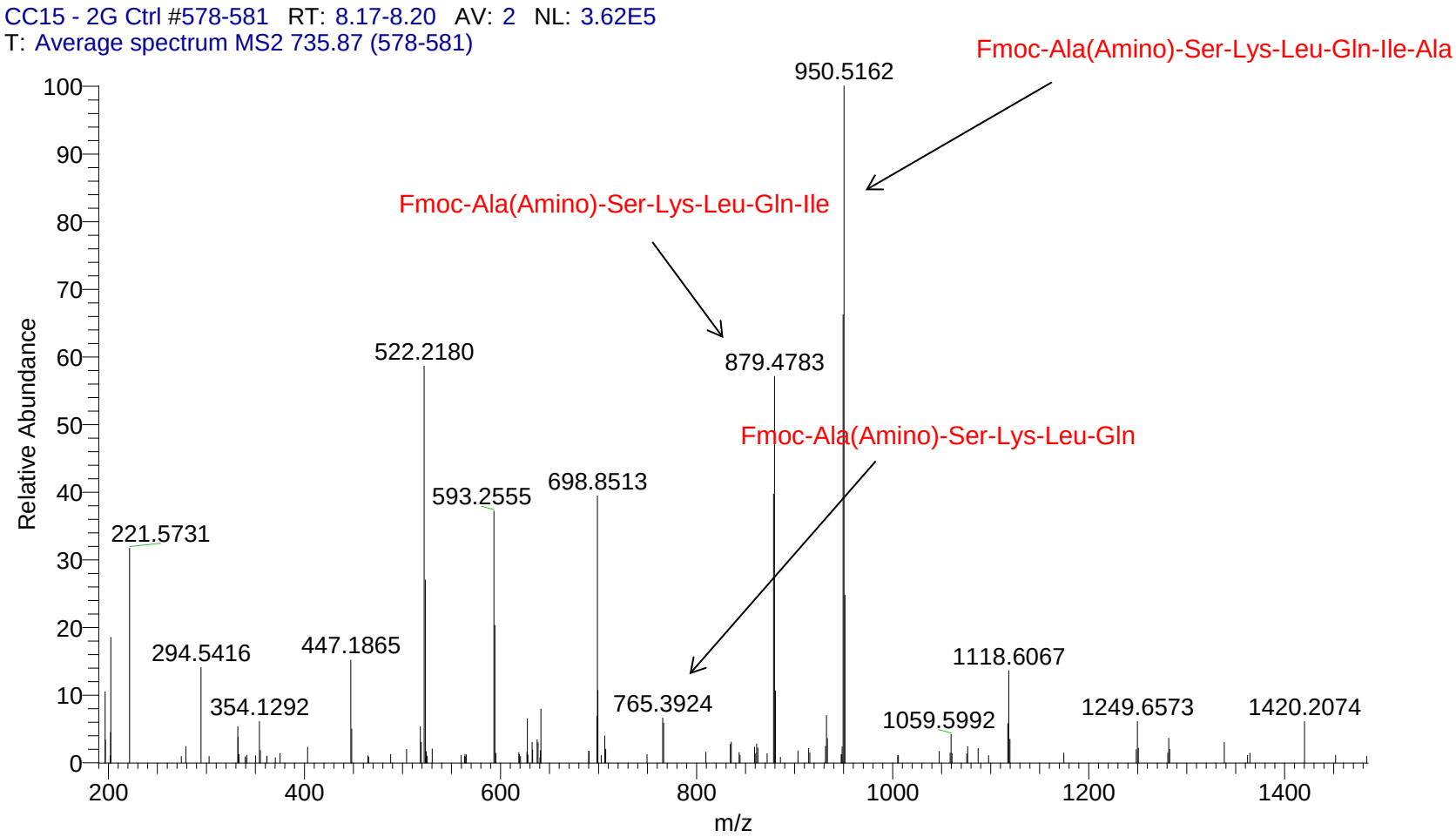


Figure S32: MS MS of the peptide substrate Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-Ala-Pro-Ala-Tyr-Asp-Gly-OH (10).

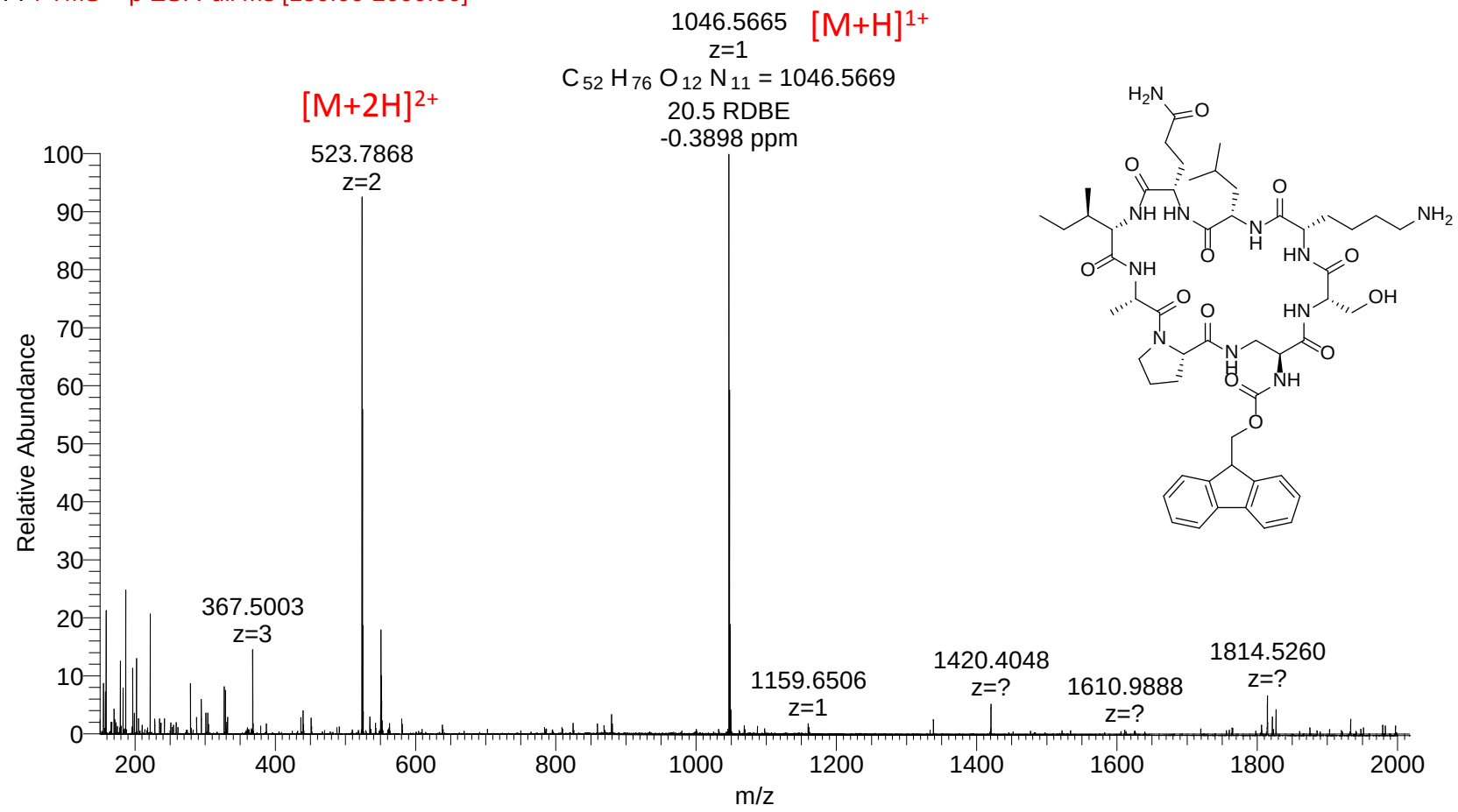


Figure S33: HR LCMS of the cyclic peptide product cyclo-[Dpr(Fmoc)-Ser-Lys-Leu-Gln-Ile-Asp-Pro-] (10) generated with $OscG_{mac}$.

OC7 - 2G_OscG #749-752 RT: 10.67-10.70 AV: 2 NL: 1.17E5
T: Average spectrum MS2 523.79 (749-752)

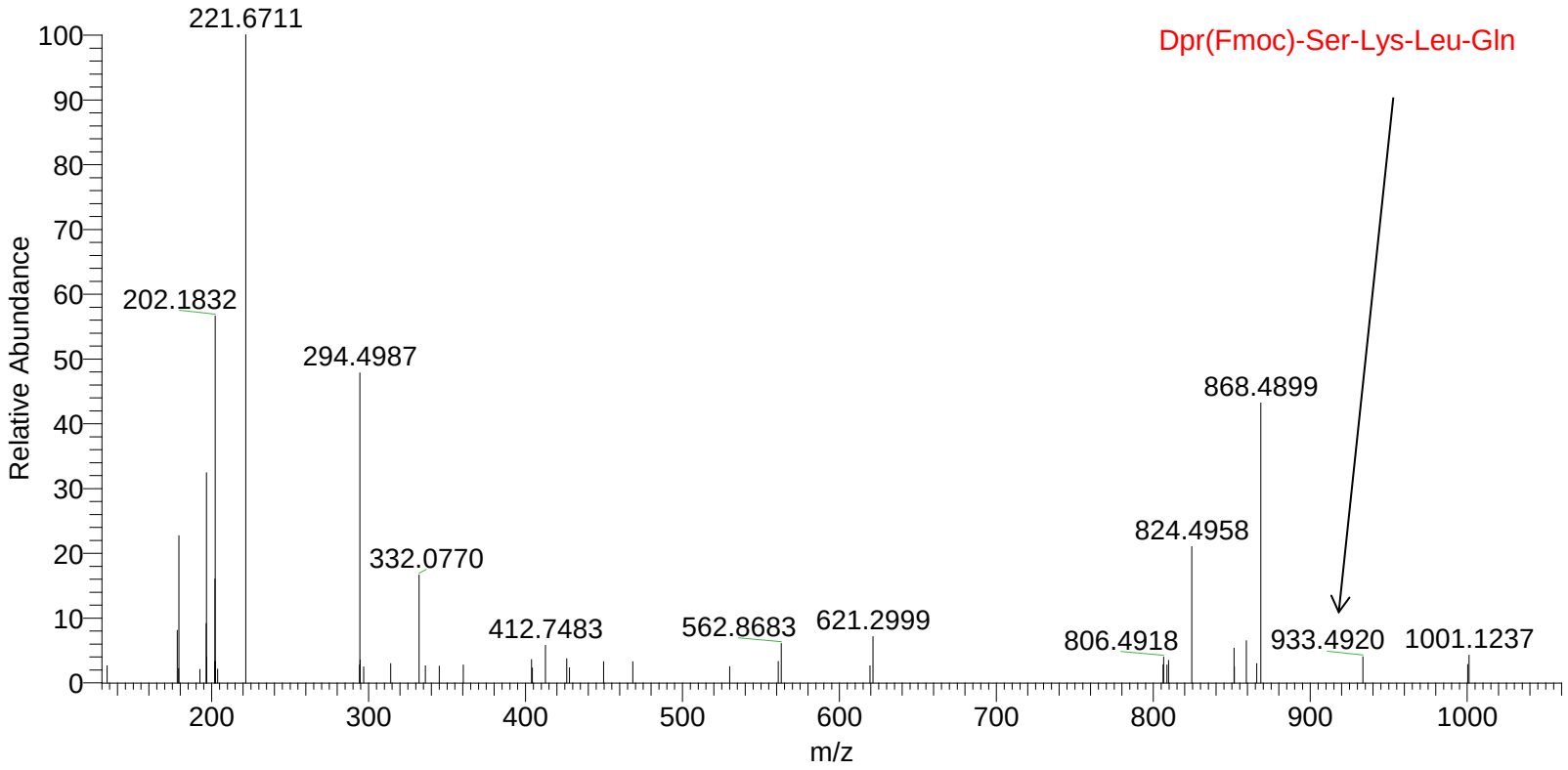


Figure S34: MS MS of the cyclic peptide product cyclo-[Dpr(Fmoc)-Ser-Lys-Leu-Gln-Ile-Asp-Pro-] (10).

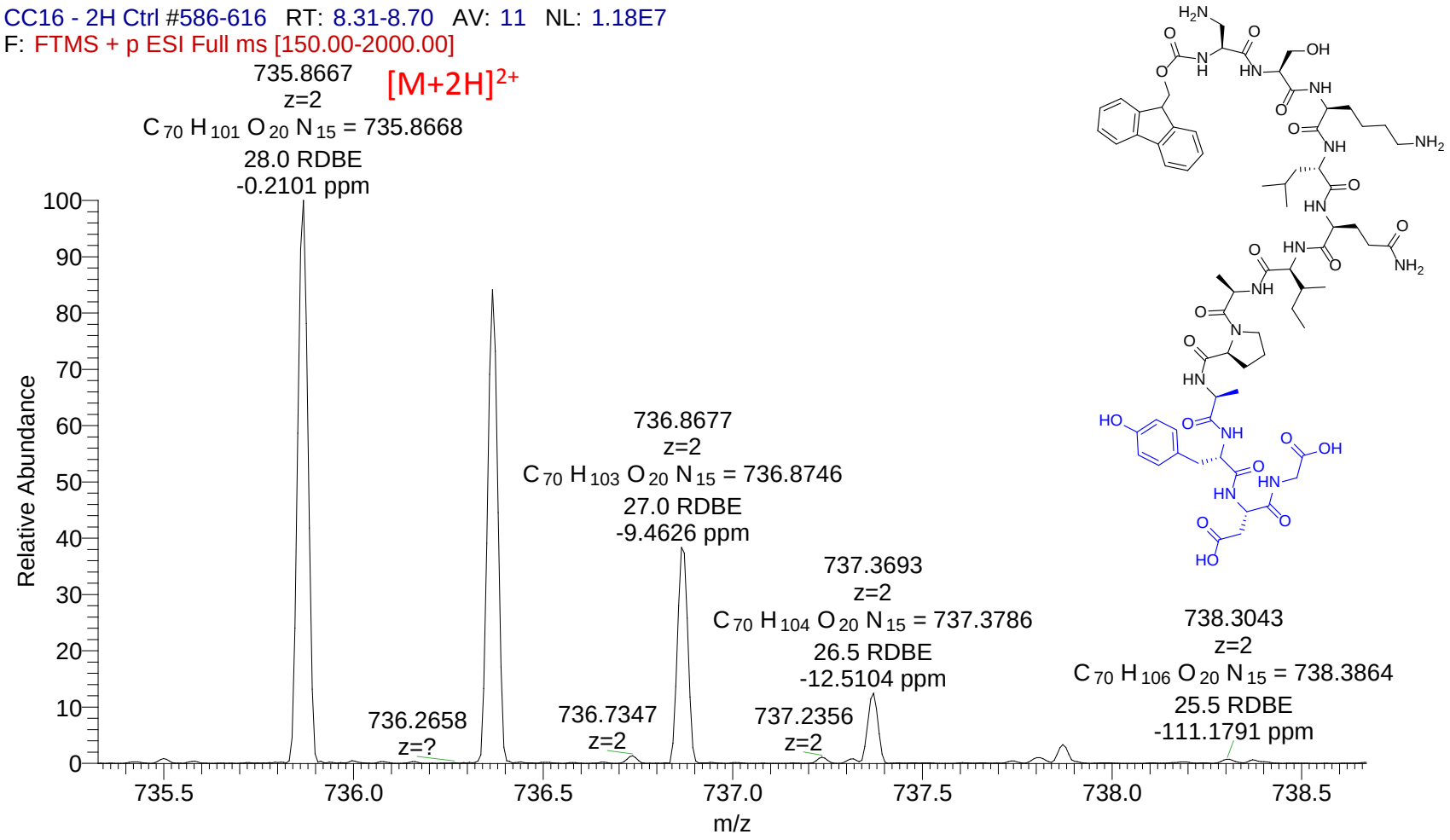


Figure S35: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-D-Ala-Pro-Ala-Tyr-Asp-Gly-OH (11).

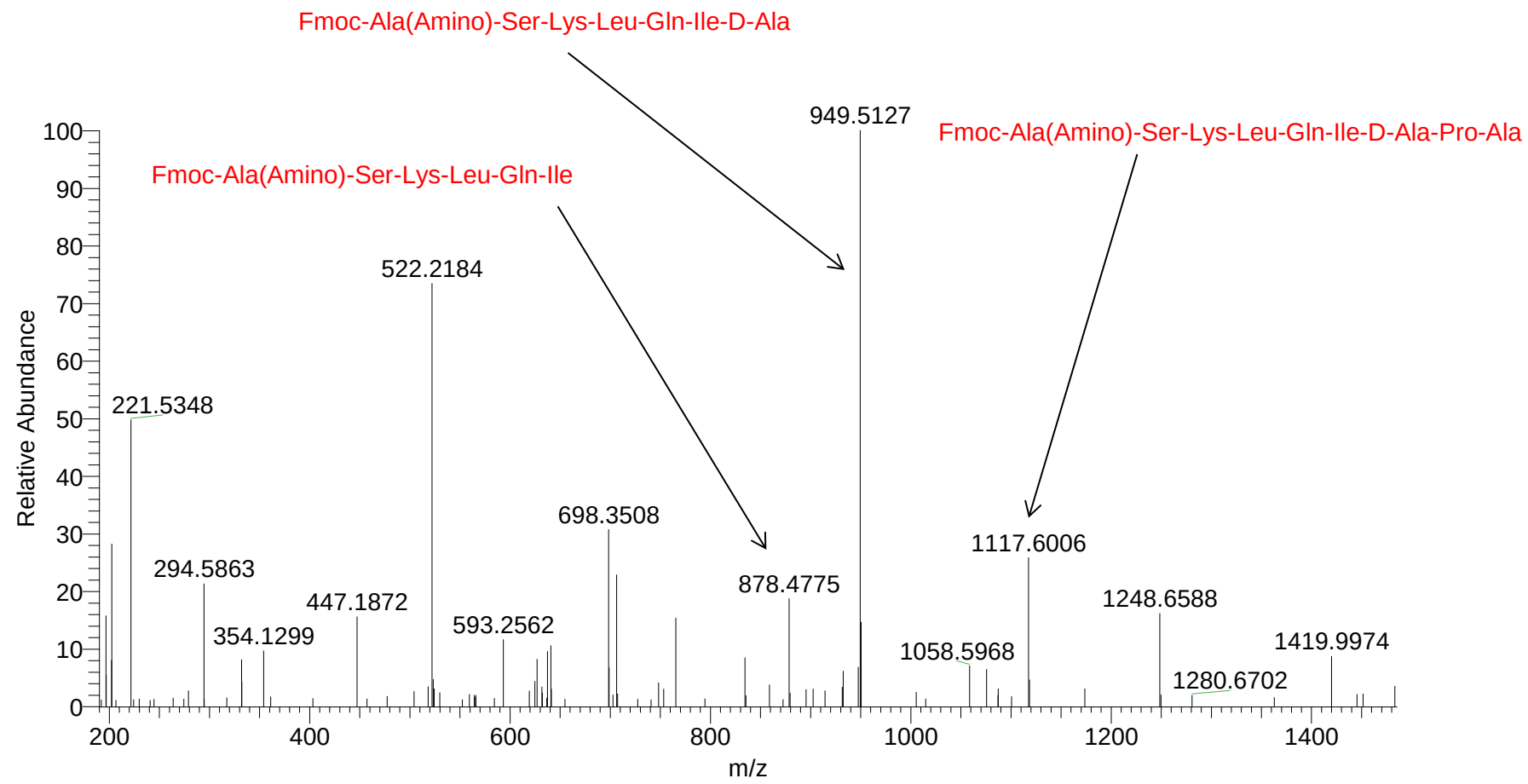


Figure S36: MS MS of the peptide substrate Fmoc-Ala(Amino)-Ser-Lys-Leu-Gln-Ile-D-Ala-Pro-Ala-Tyr-Asp-Gly-OH (11).

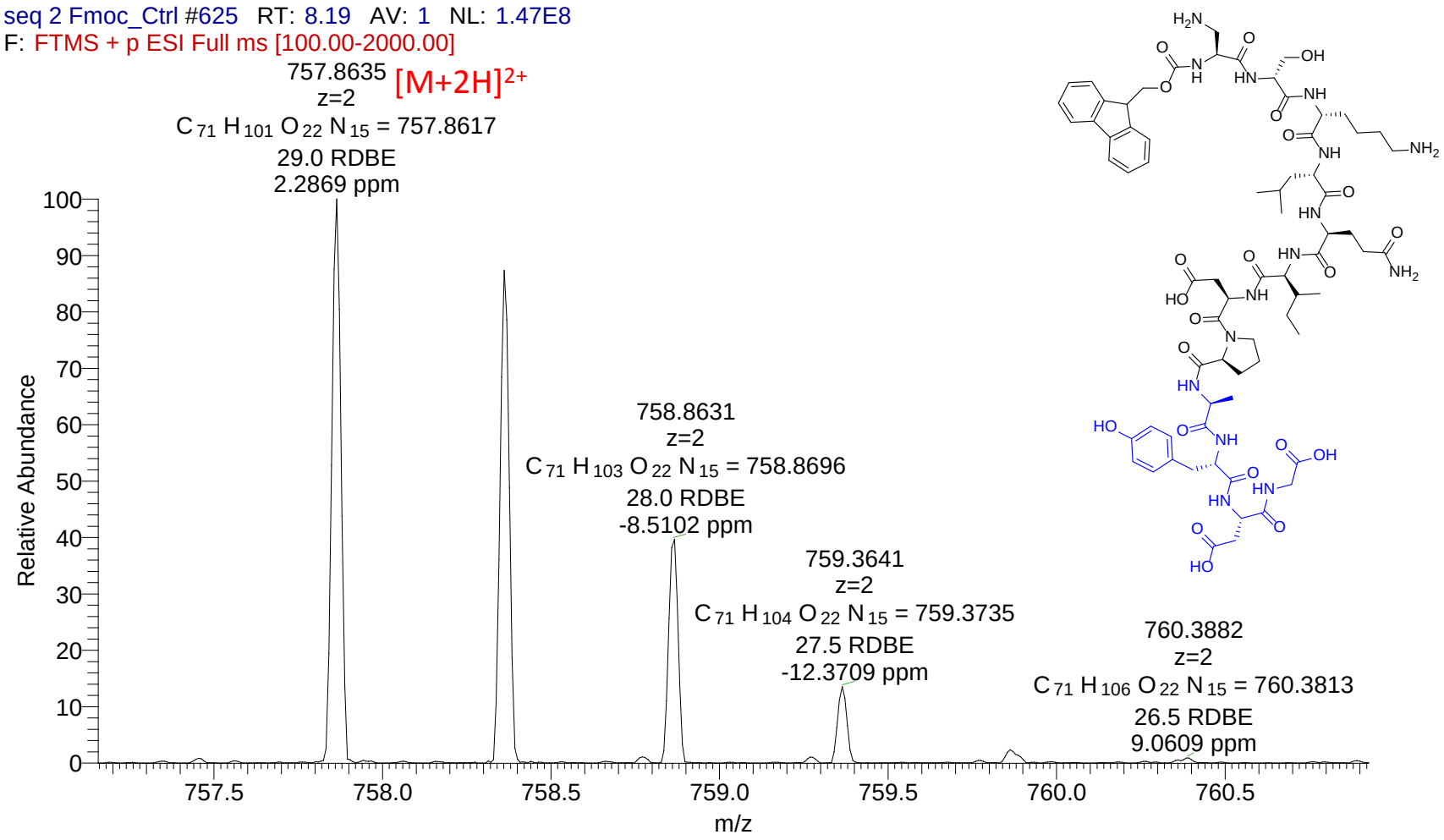


Figure S37: HR LCMS of the peptide substrate Fmoc-Ala(Amino)-D-Ser-D-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH (12).

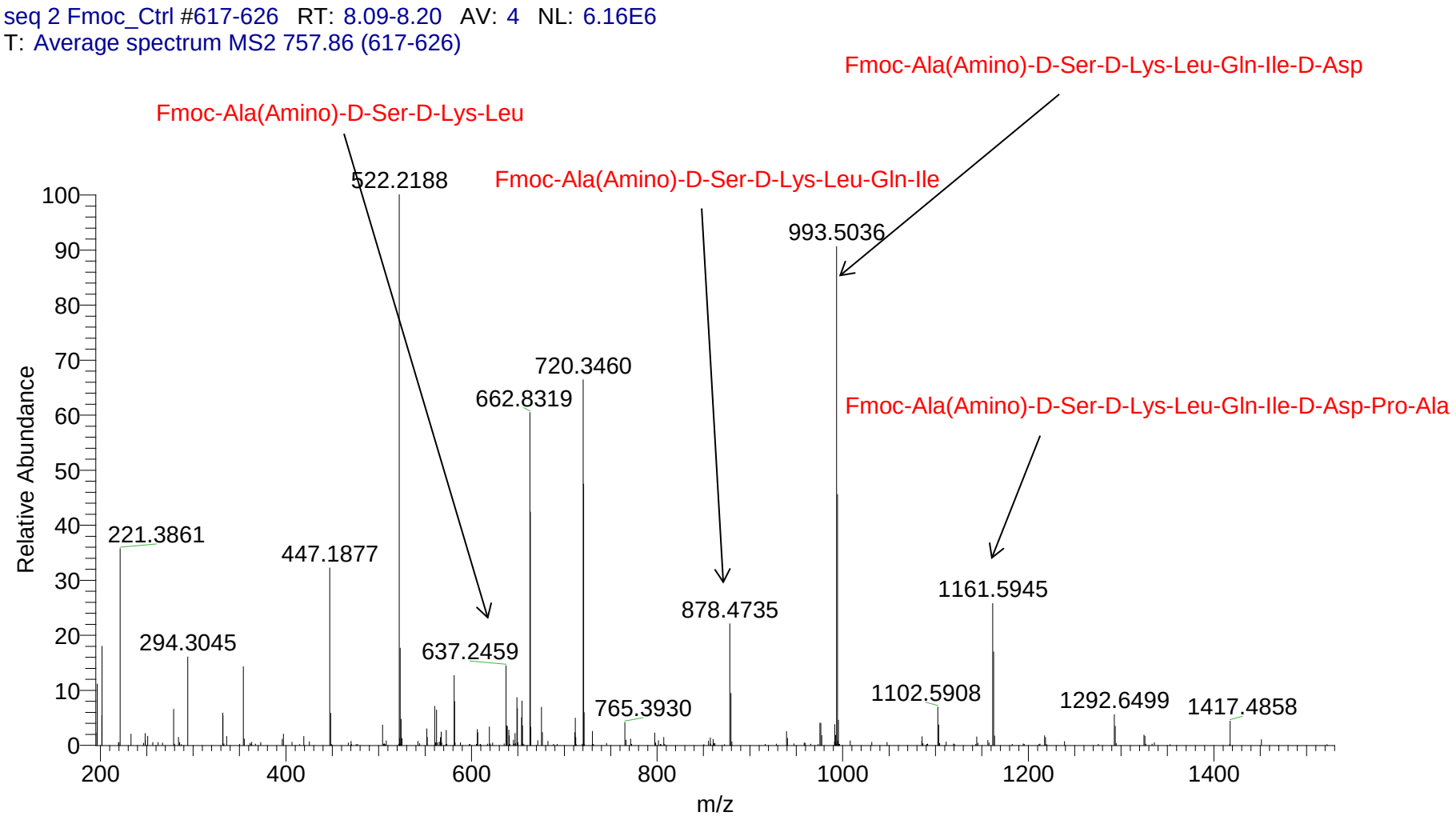


Figure S38: MS MS of the peptide substrate Fmoc-Ala(Amino)-D-Ser-D-Lys-Leu-Gln-Ile-D-Asp-Pro-Ala-Tyr-Asp-Gly-OH (12).

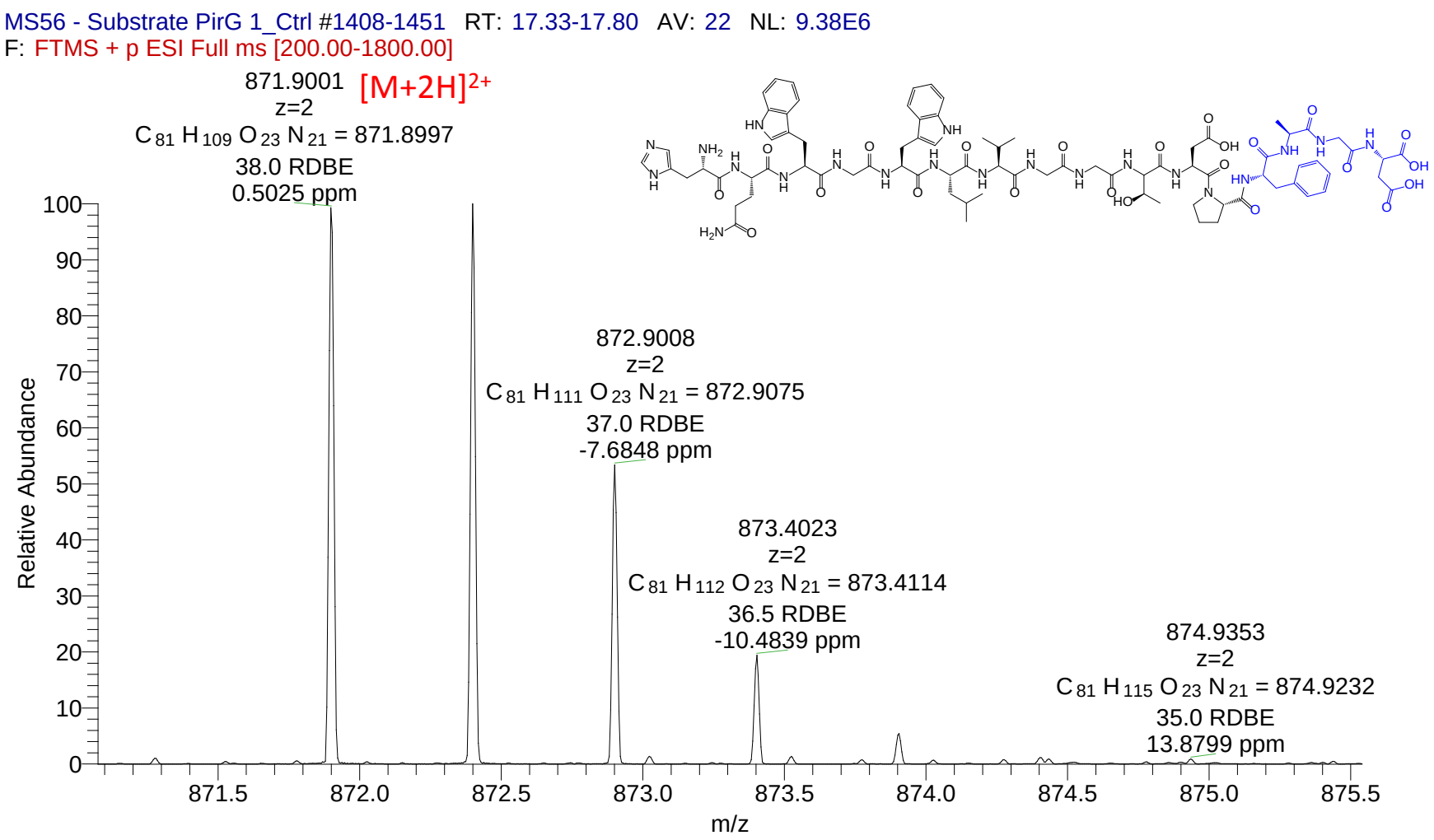


Figure S39: HR LCMS of the peptide substrate H-His-Gln-Trp-Gly-Trp-Leu-Val-Gly-Gly-Thr-Asp-Pro-Phe-Ala-Gly-Asp-OH (13).

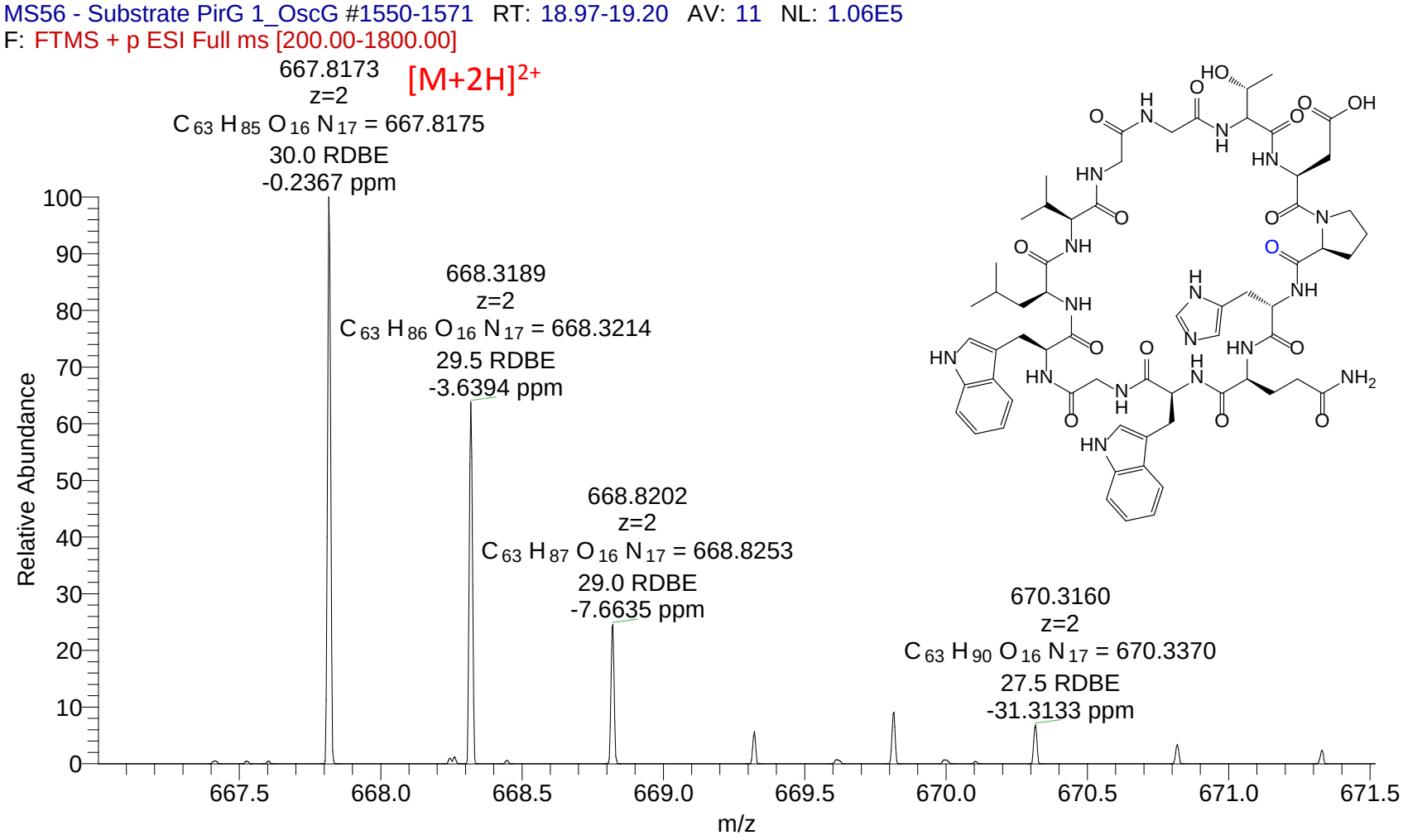


Figure S40: HR LCMS of the cyclic peptide product cyclo-[His-Gln-Trp-Gly-Trp-Leu-Val-Gly-Gly-Thr-Asp-Pro-] (13) generated with $OscG_{mac}$.

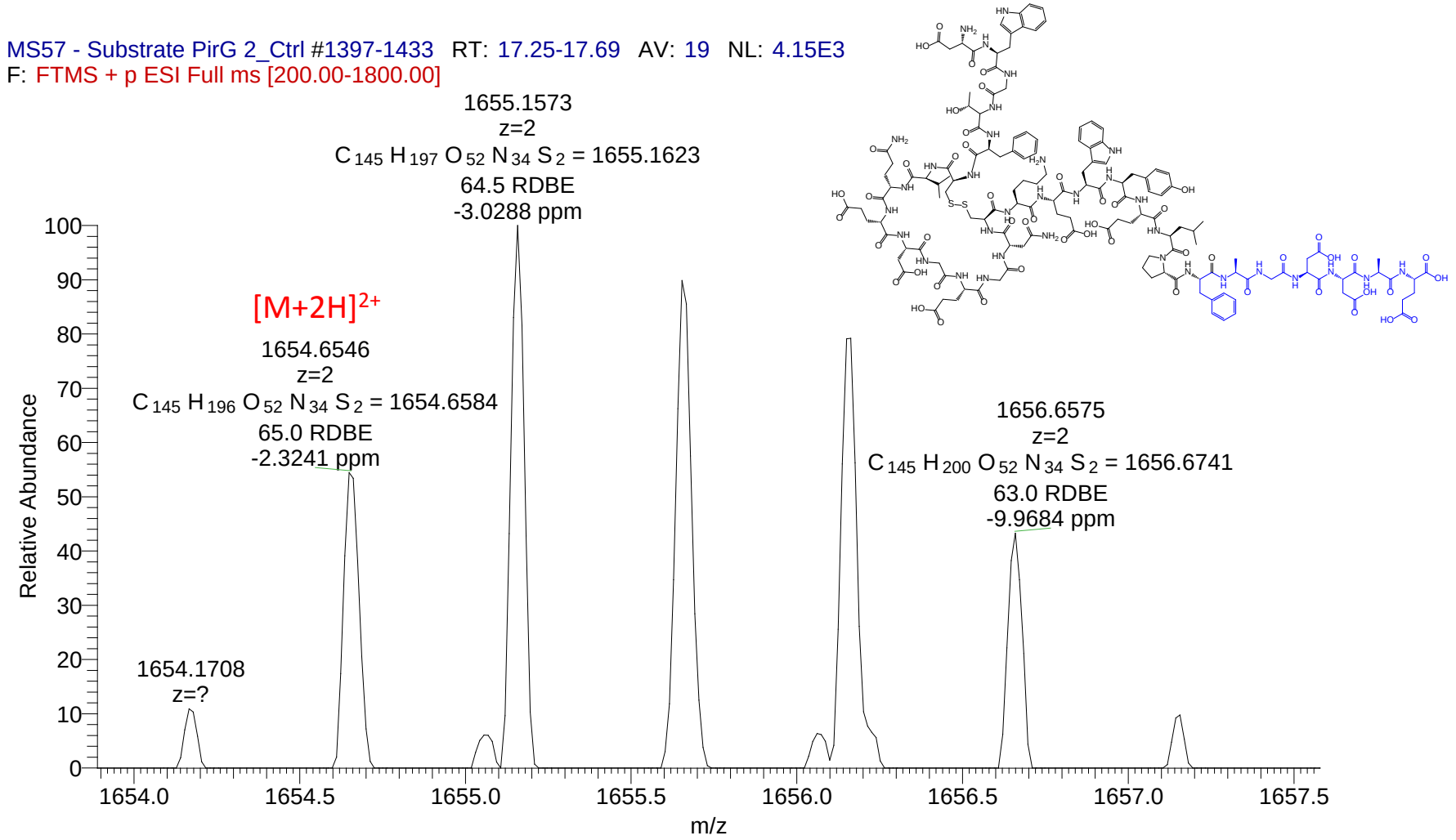


Figure S41: HR LCMS of the peptide substrate H-Asp-Trp-Gly-Thr-Phe-Cys-Val-Gln-Glu-Asp-Gly-Glu-Gly-Asn-Cys-Lys-Glu-Trp-Tyr-Glu-Leu-Pro-Phe-Ala-Gly-Asp-Asp-Ala-Glu-OH (14).

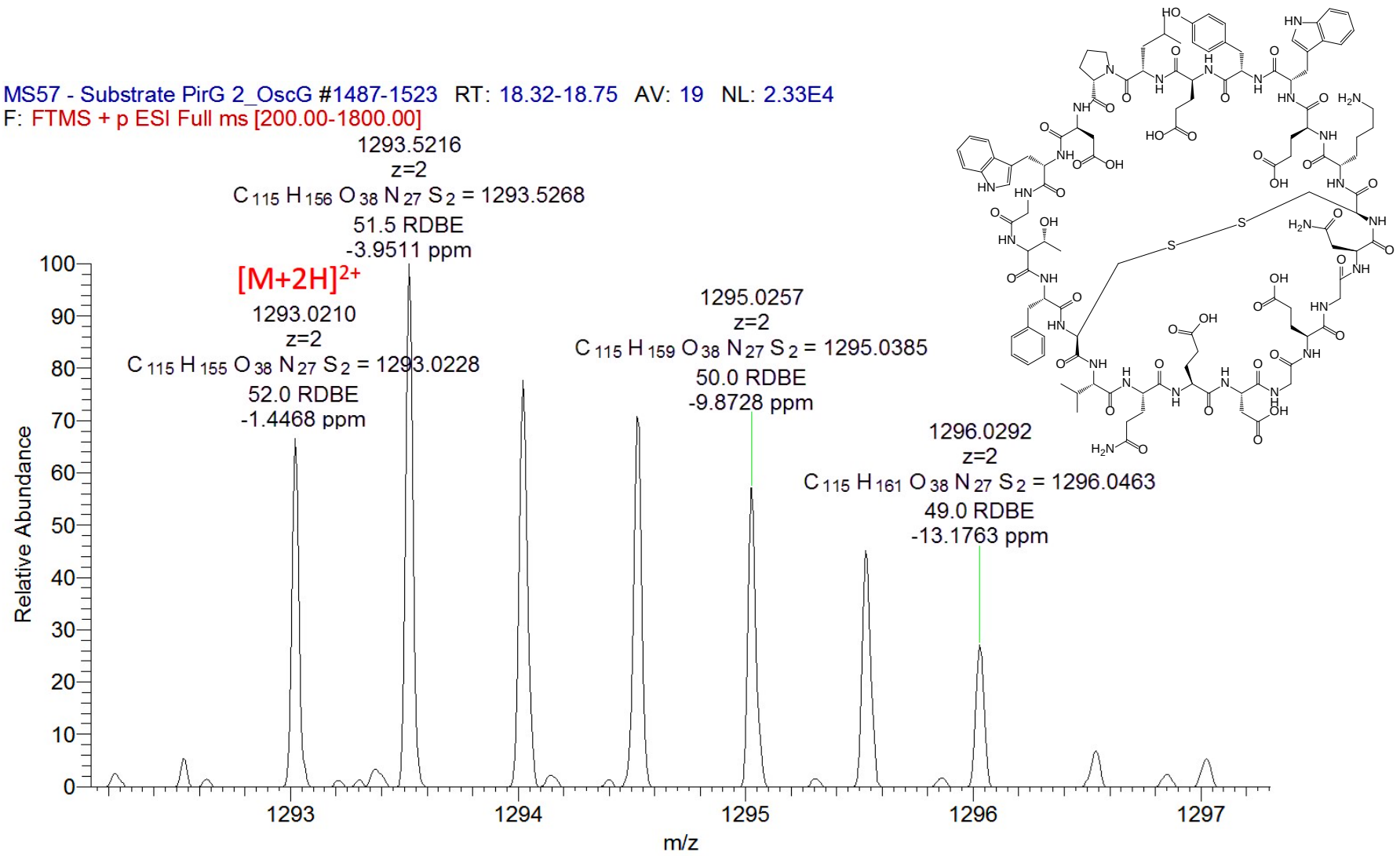


Figure S42: HR LCMS of the cyclic peptide product cyclo-[Asp-Trp-Gly-Thr-Phe-Cys-Val-Gln-Glu-Asp-Gly-Glu-Gly-Asn-Cys-Lys-Glu-Trp-Tyr-Glu-Leu-Pro-] (disulfide bond) (14) generated with $OscG_{mac}$.

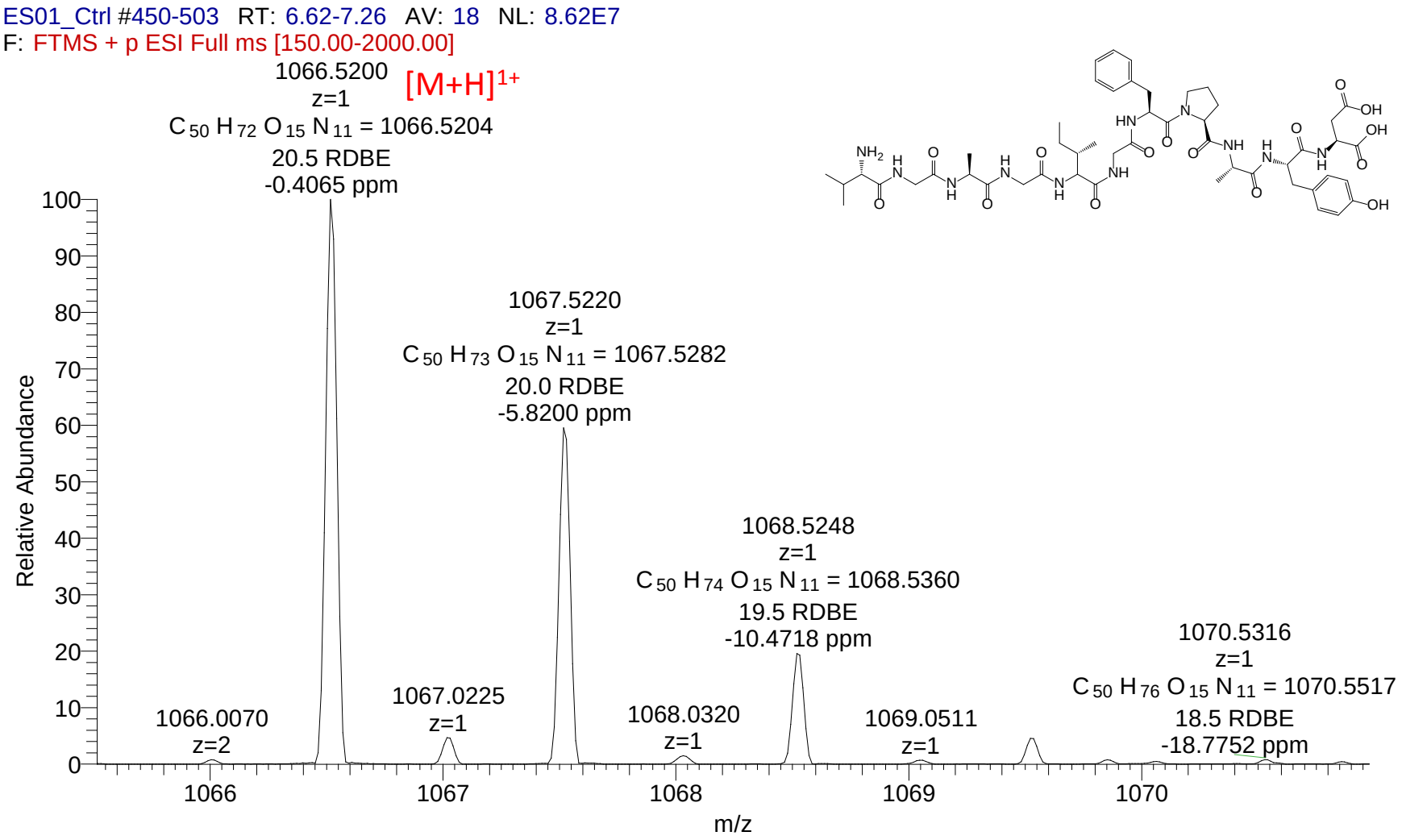


Figure S43: HR LCMS of the peptide substrate H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Pro-Ala-Tyr-Asp-OH (15).

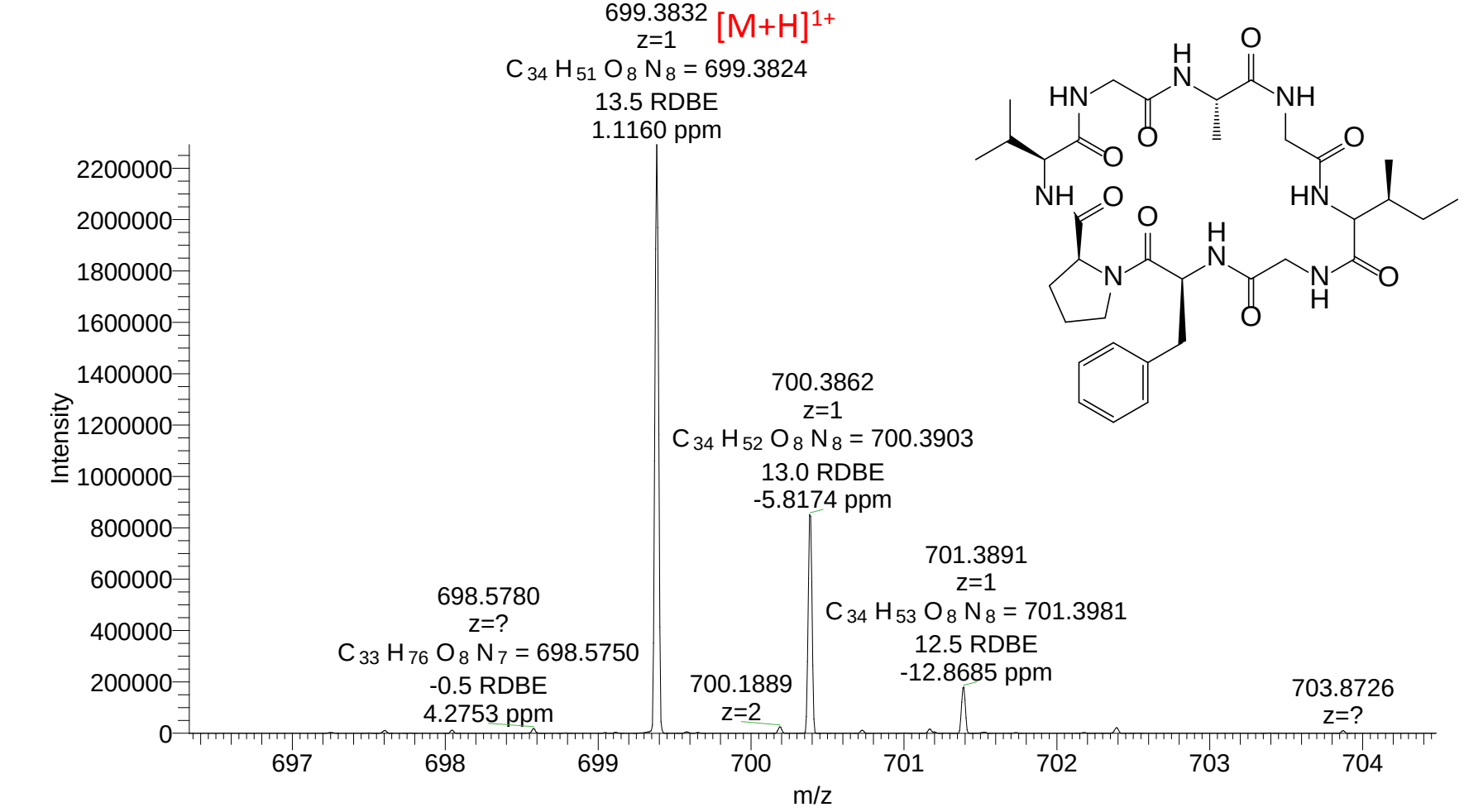


Figure S44: HR LCMS of the cyclic peptide product cyclo-[Val-Gly-Ala-Gly-Ile-Gly-Phe-Pro-] (15) generated with $OscG_{mac}$.

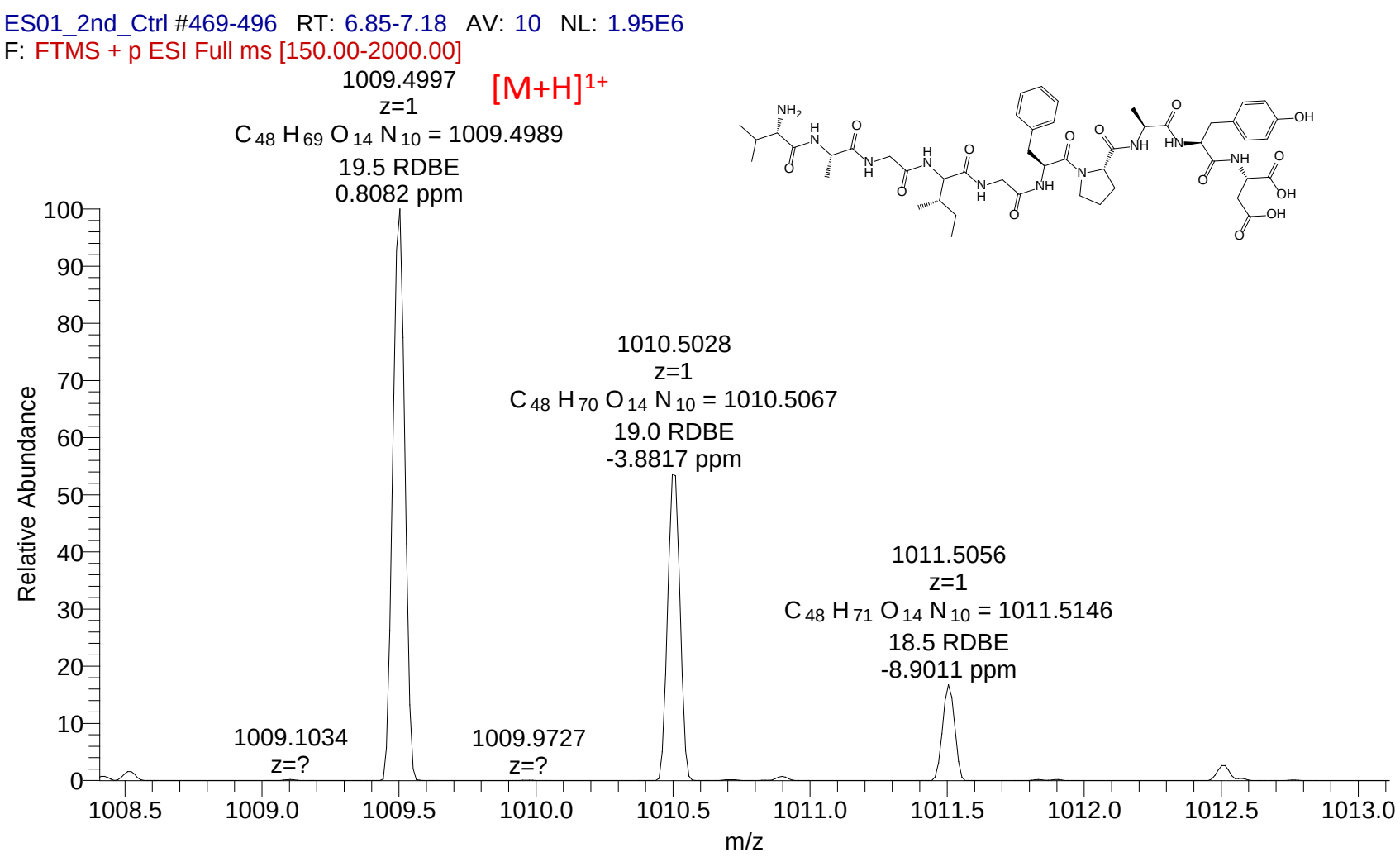


Figure S45: HR LCMS of the peptide substrate H-Val-Ala-Gly-Ile-Gly-Phe-Pro-Ala-Tyr-Asp-OH (16).

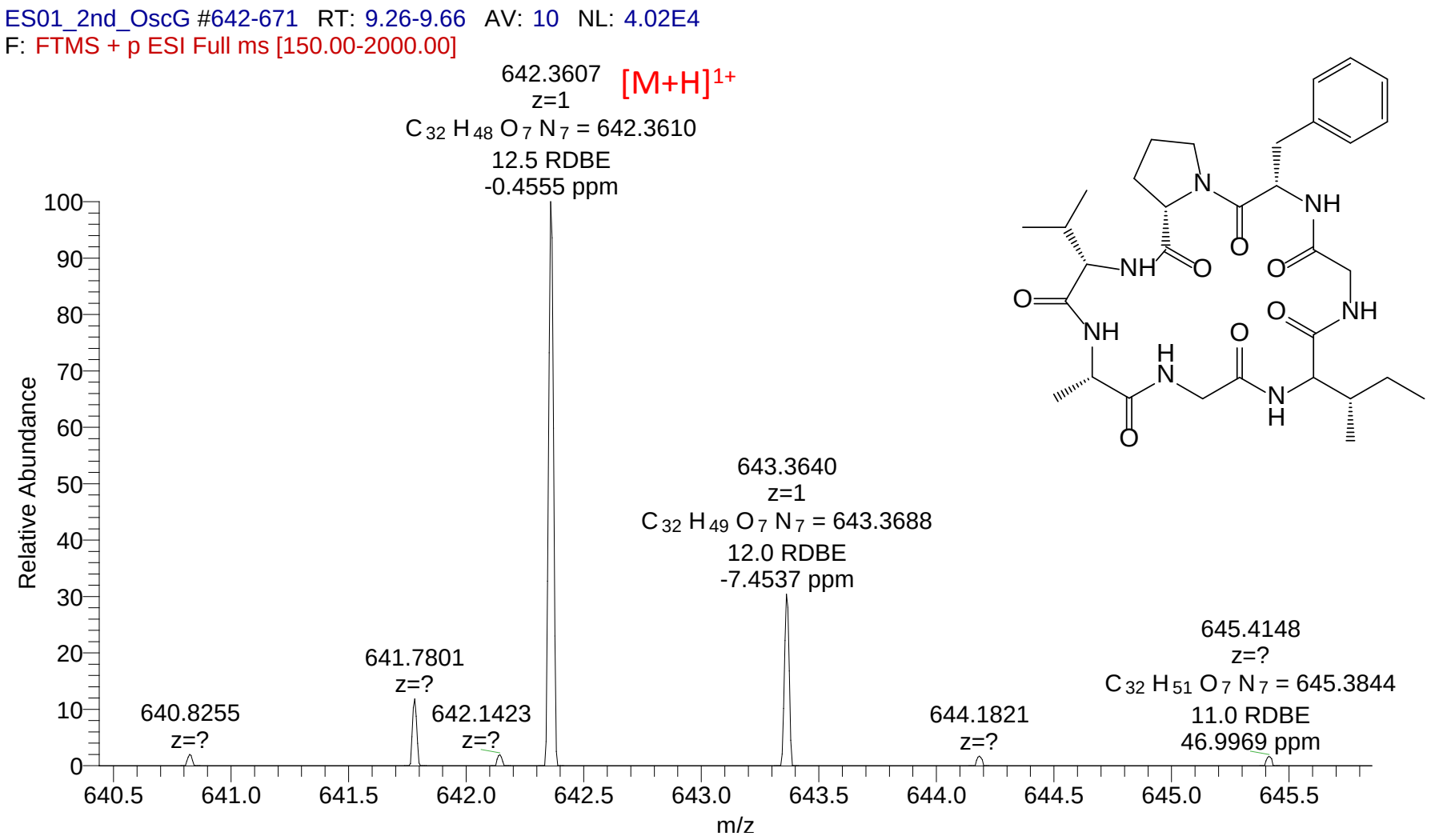


Figure S46: HR LCMS of the cyclic peptide product cyclo-[Val-Ala-Gly-Ile-Gly-Phe-Pro-] (16) generated with $OscG_{mac}$.

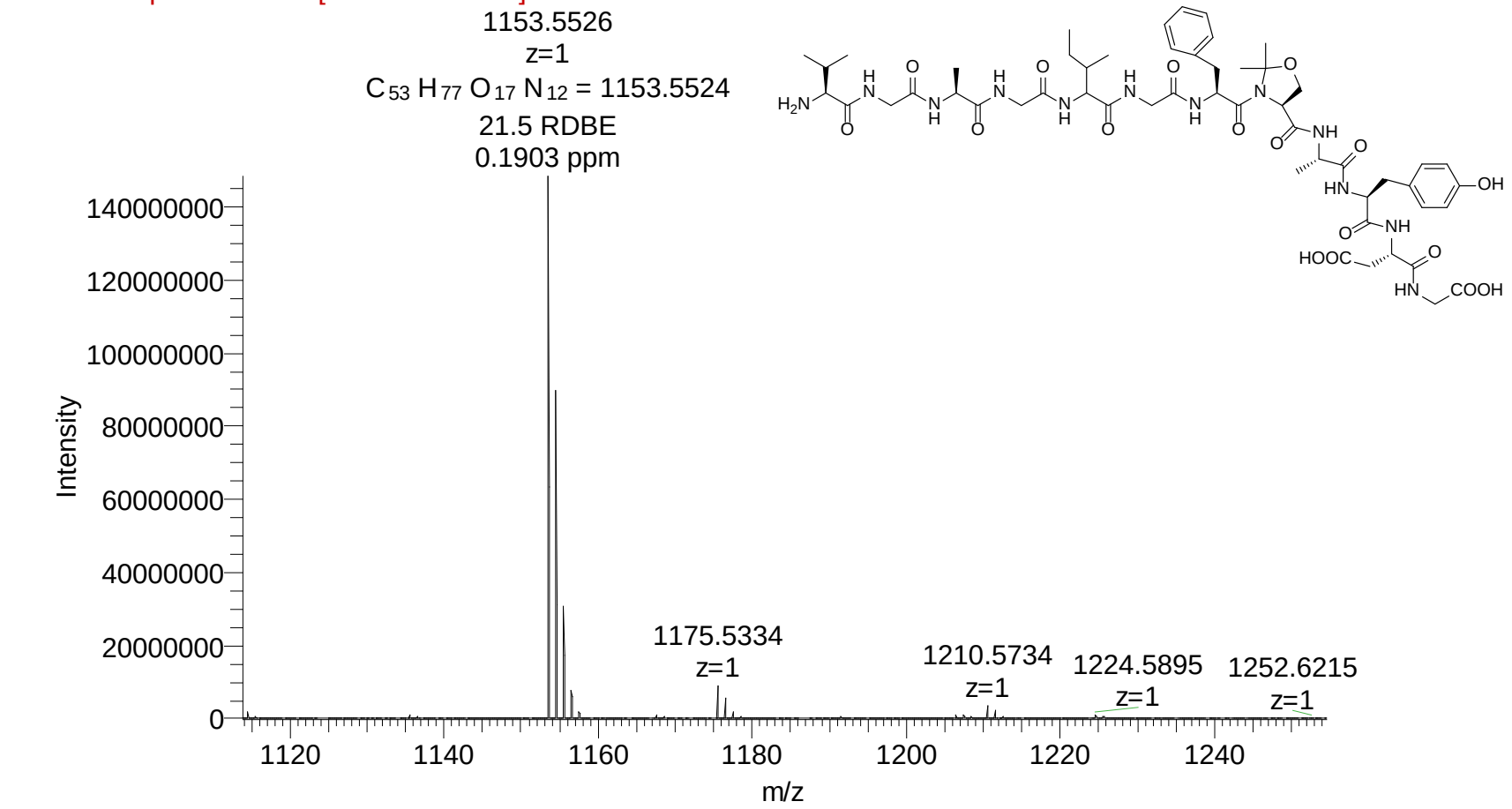


Figure S47: HR LCMS of the peptide substrate H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Ser(ψ^{Me,Me}Pro)-Ala-Tyr-Asp-Gly-OH (17).

FE09 #748-1027 RT: 12.14-16.08 AV: 4 NL: 3.01E4
T: Average spectrum MS2 1153.55 (748-1027)

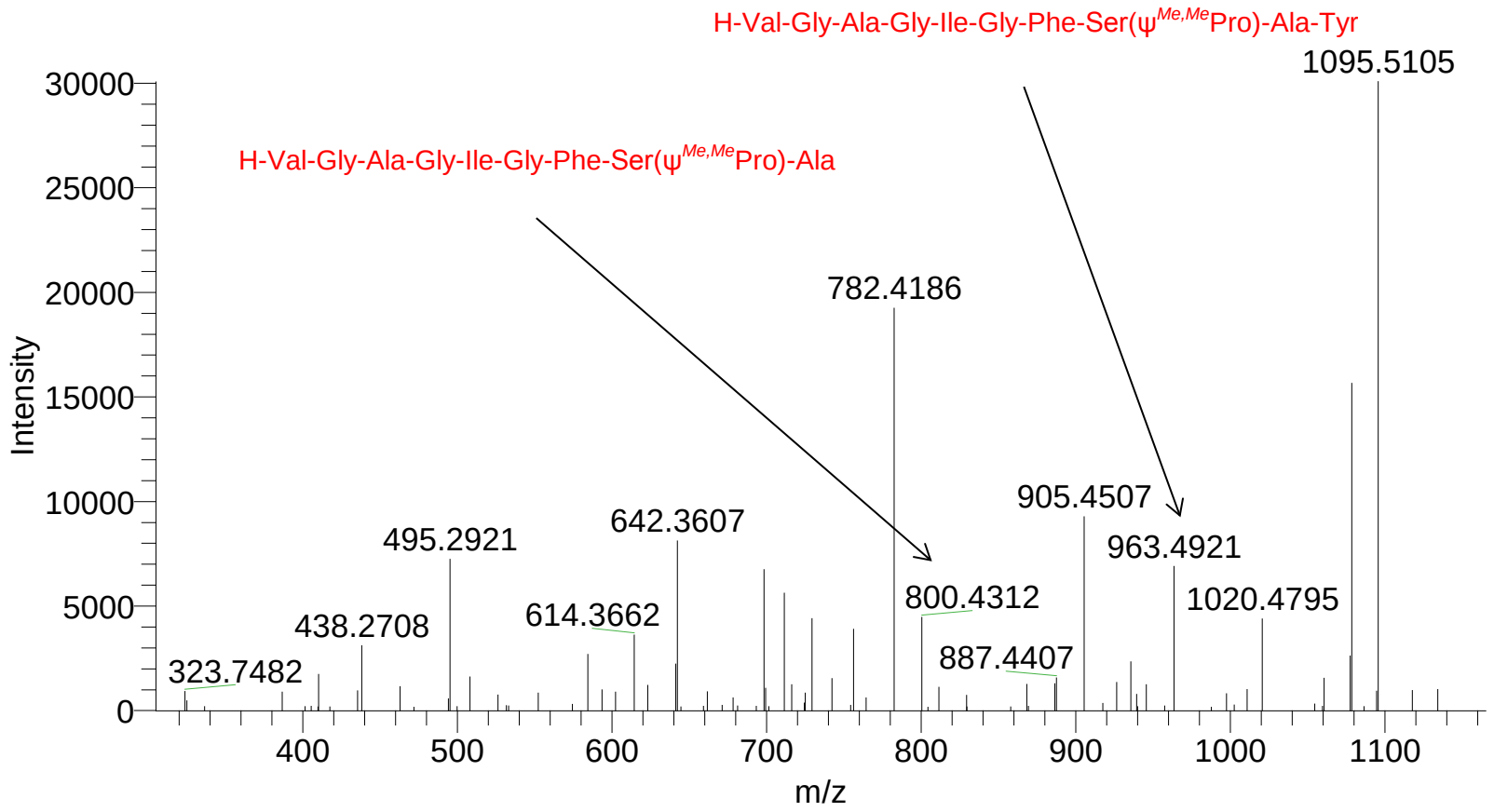


Figure S48: MS MS of the peptide substrate H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Ser(ψ^{Me,Me}Pro)-Ala-Tyr-Asp-Gly-OH (17).

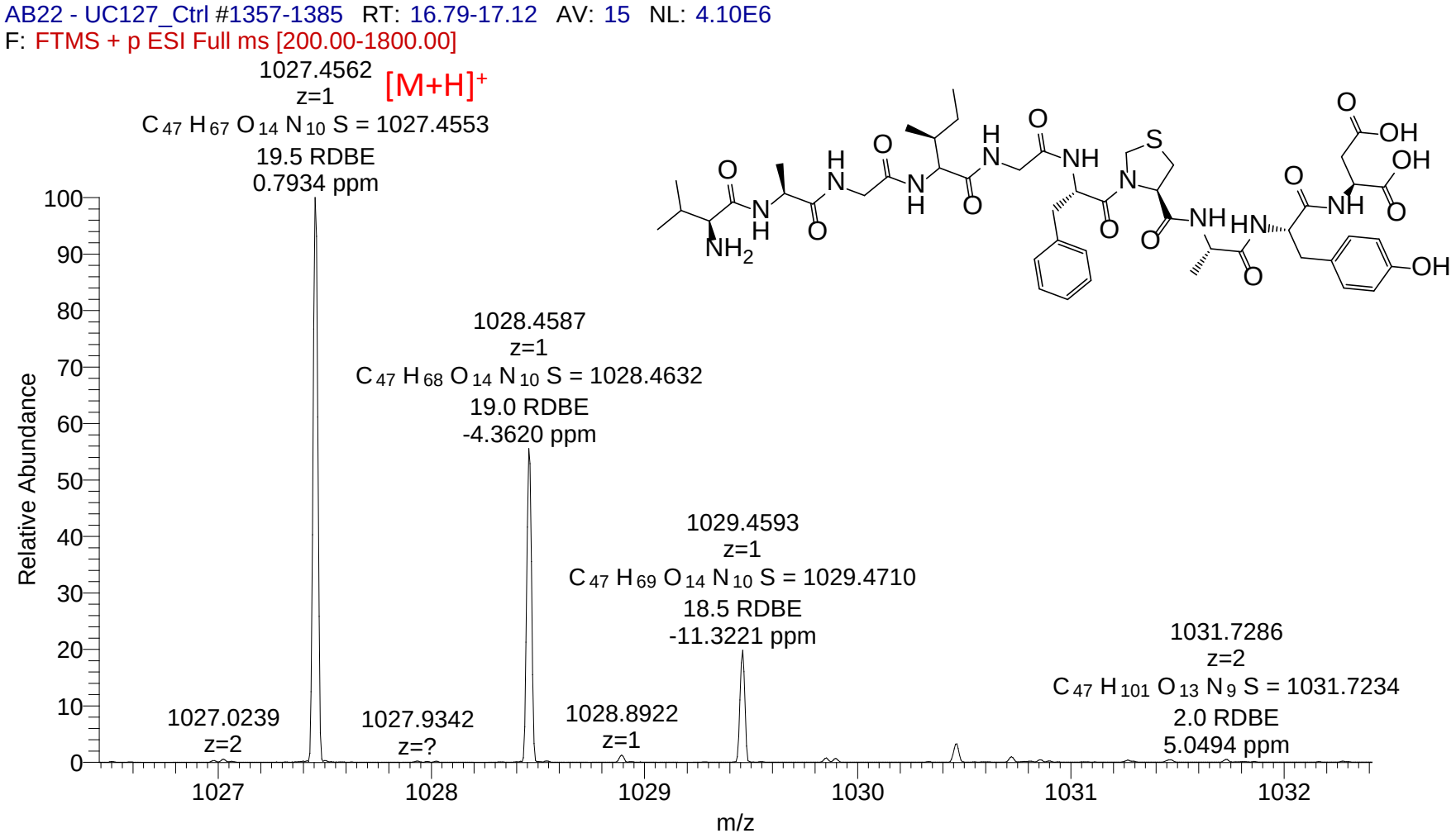


Figure S49: HR LCMS of the peptide substrate H-Val- Ala-Gly-Ile-Gly-Phe-Cys(ψPro)-Ala-Tyr-Asp-OH (18).

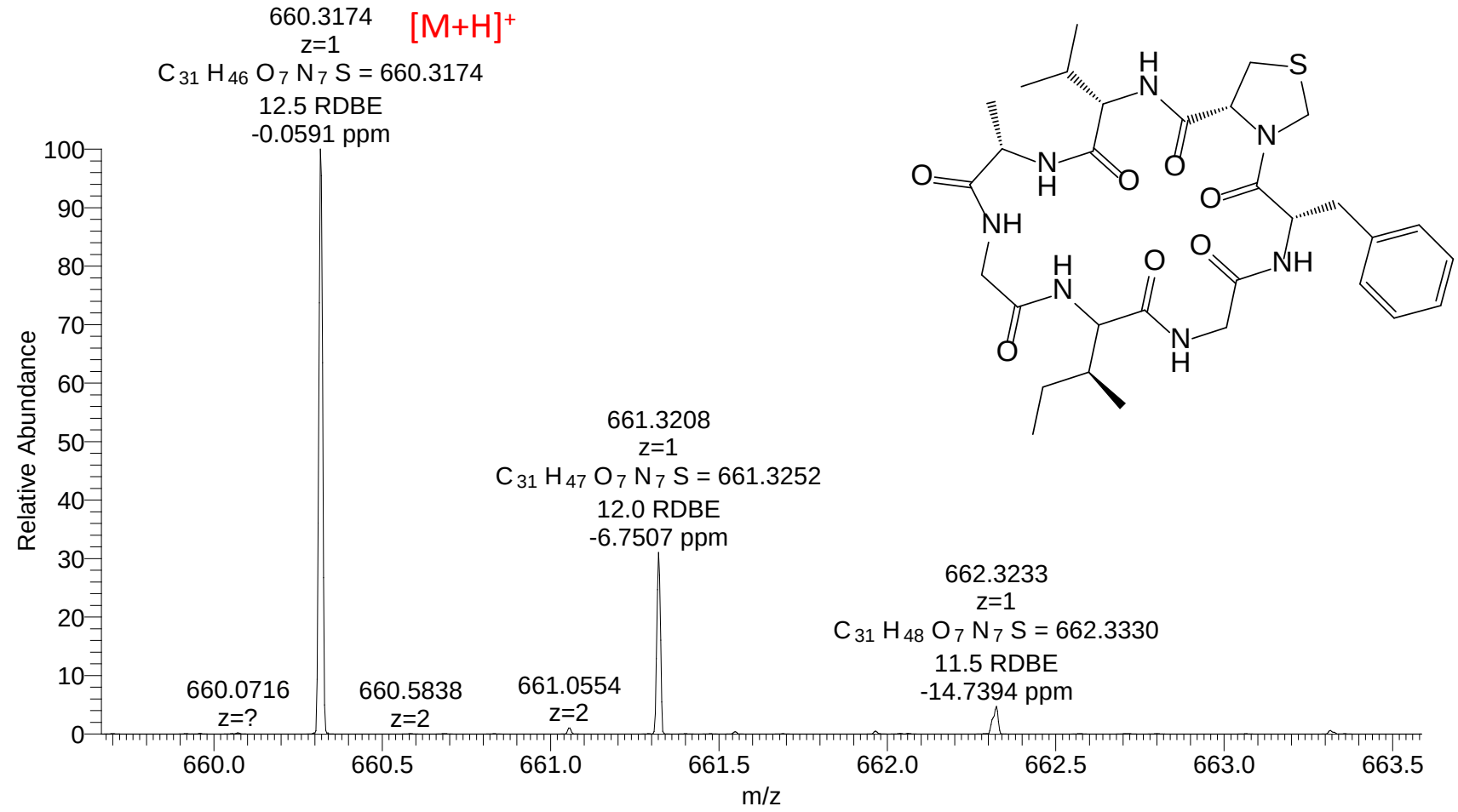


Figure S50: HR LCMS of the cyclic peptide product cyclo-[Val- Ala-Gly-Ile-Gly-Phe-Cys(ψPro)-] (18) generated with $OscG_{mac}$.

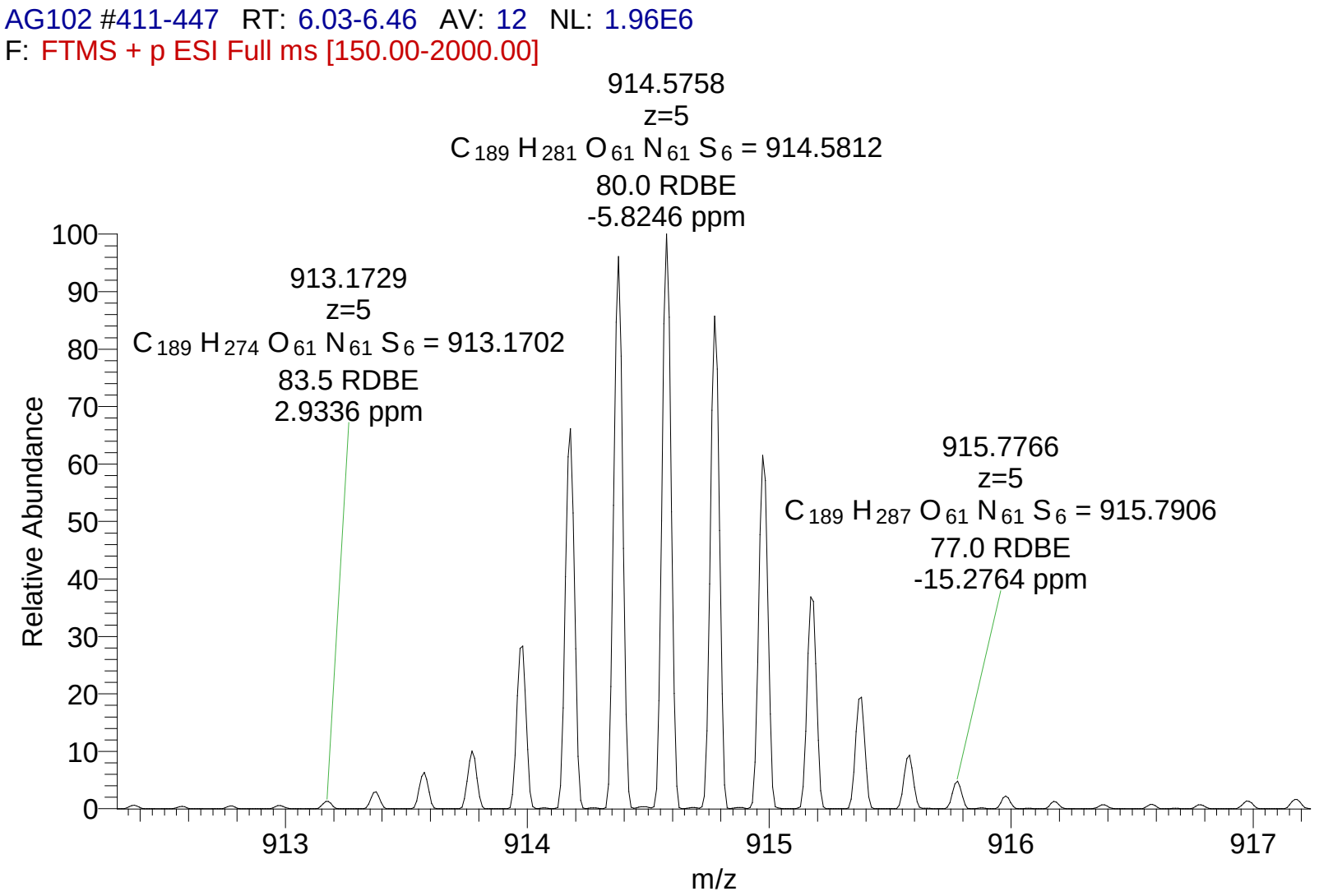


Figure S51: HR LCMS of the peptide substrate H-Gly-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Cys-Asn-Thr-Pro-Gly-Cys-Thr-Cys-Ser-Trp-Pro-Val-Cys-Thr-Arg-Asn-Gly-Leu-Pro-Ala-Tyr-Asp-Gly-Glu-Leu-Glu-His-His-His-His-His-His-OH (two disulfide bonds) (19).

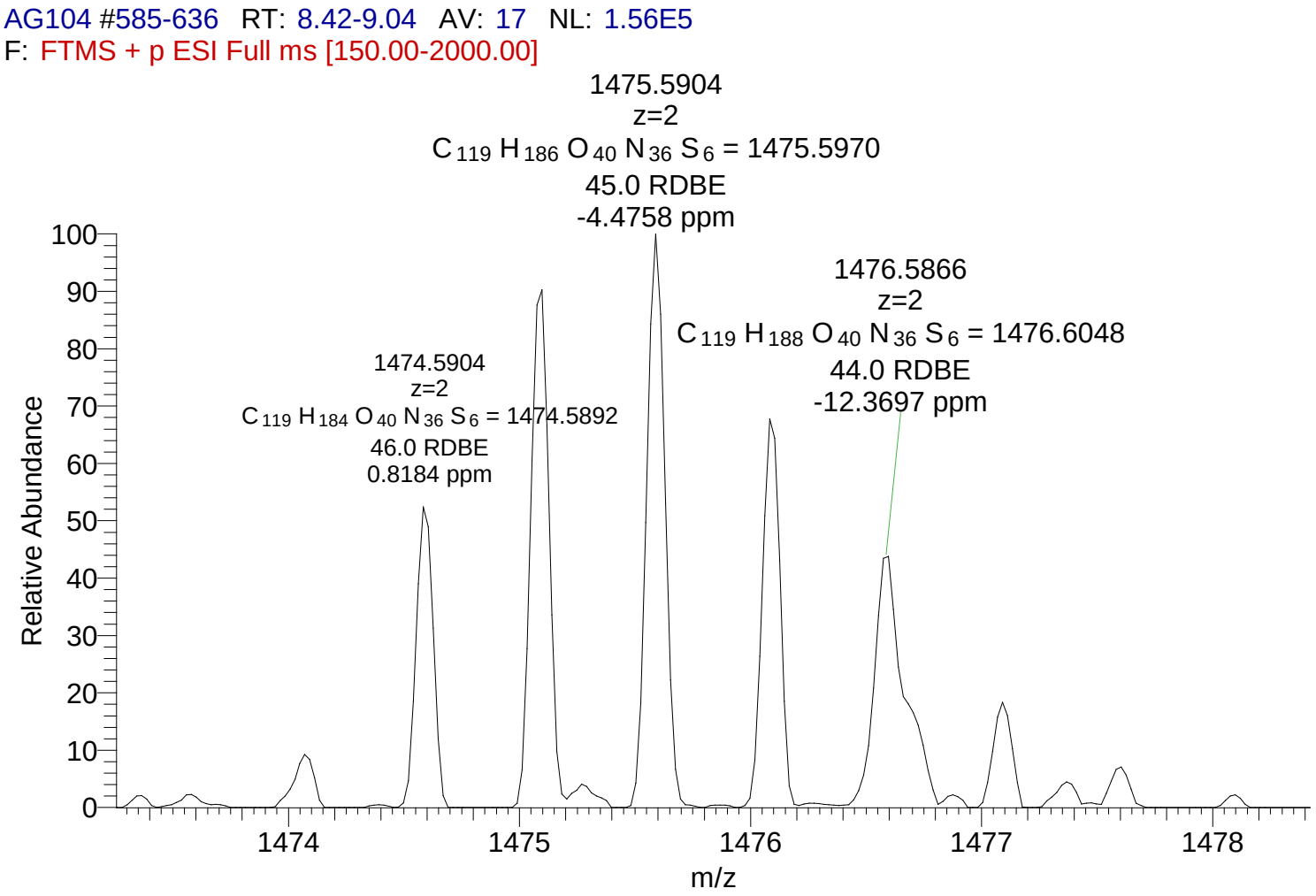


Figure S52: HR LCMS of the cyclic peptide product Cyclo-[Gly-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Cys-Asn-Thr-Pro-Gly-Cys-Thr-Cys-Ser-Trp-Pro-Val-Cys-Thr-Arg-Asn-Gly-Leu-Pro-] (two disulfide bonds) (19) generated with $OscG_{mac}$.

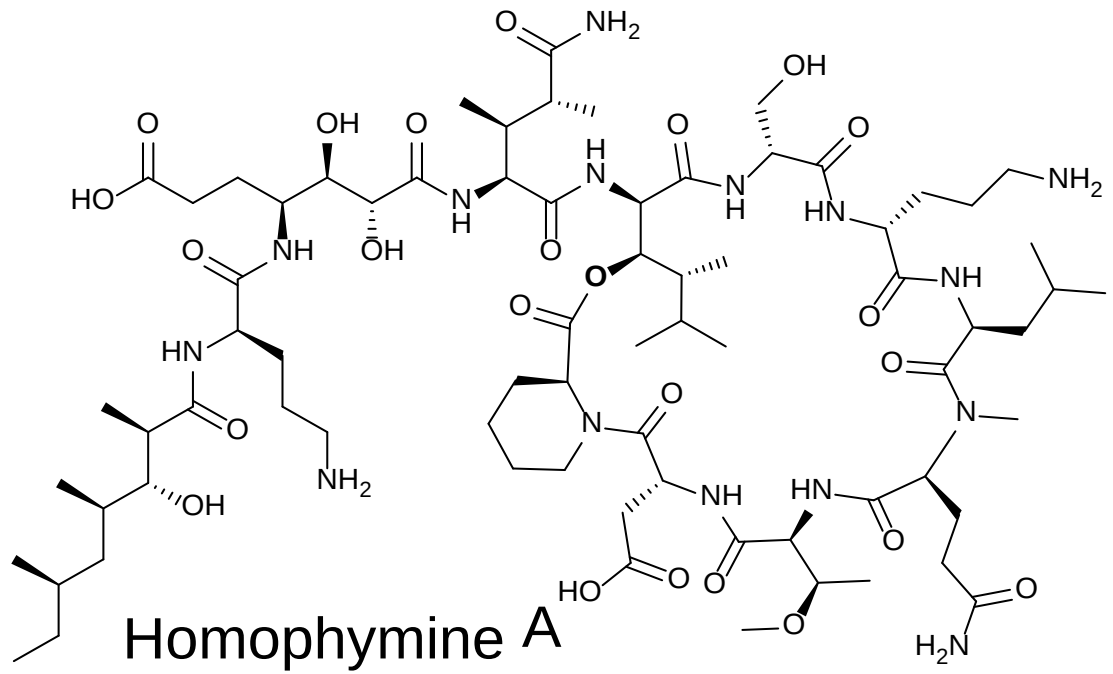


Figure S53: Chemical structure of the anti-HIV depsipeptide homophymine A.

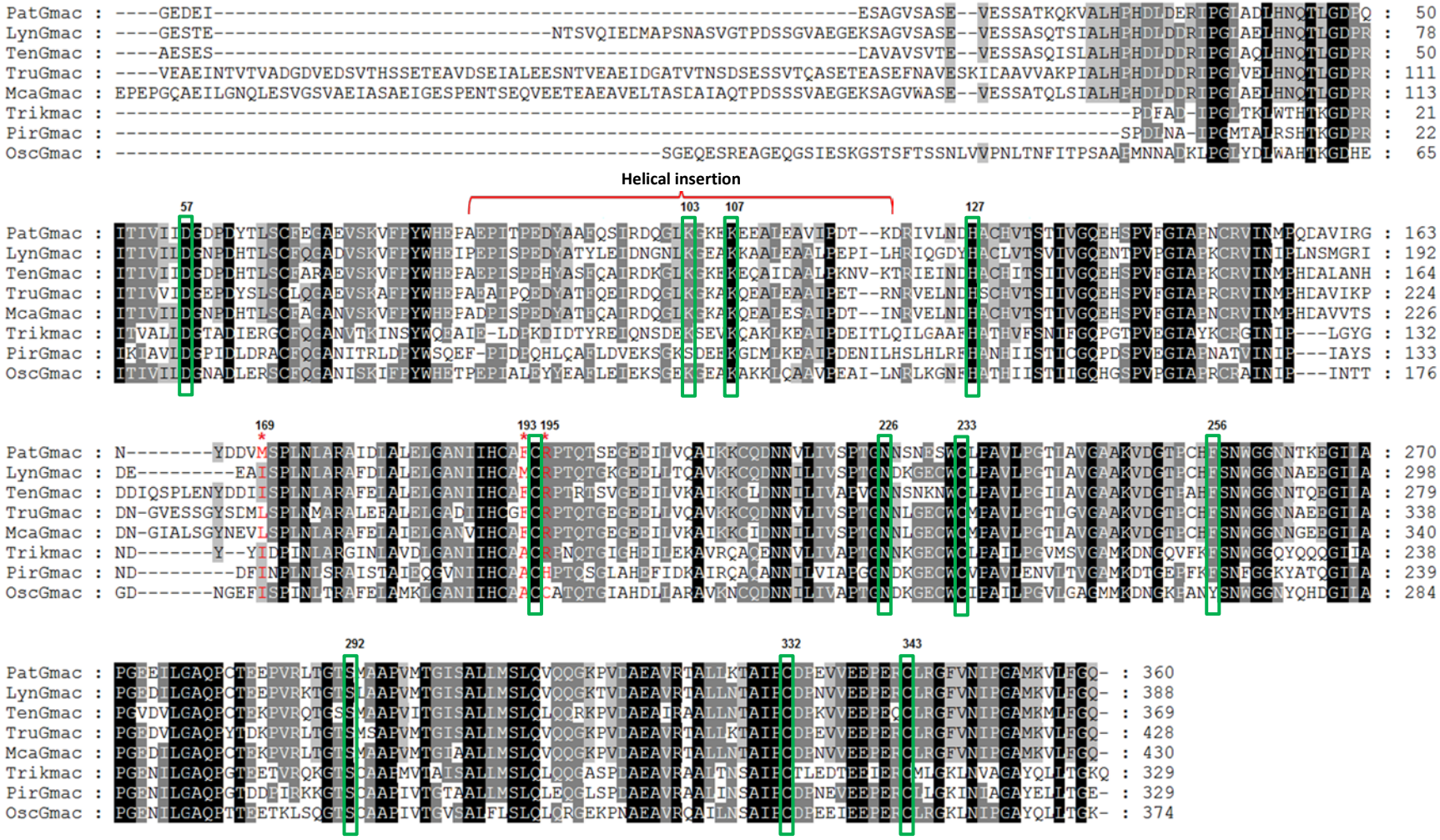


Figure S54: Alignment of the cyanobactin macrocylases $PatG_{mac}$, $LynG_{mac}$, $TenG_{mac}$, $TruG_{mac}$, $McaG_{mac}$, $TriK_{mac}$, $PirG_{mac}$ and $OscG_{mac}$. Sequence similarity is indicated by shading. Residues in $PatG_{mac}$ that are critical for macrocylization activity are in red and in green frames. Numbering on top of sequences is according to the $PatG_{mac}$ sequence. In $PatG_{mac}$, the side chains of the bulky residues in red protrude and

prevent the substrate from adopting the relaxed elongated conformation. In $OscG_{mac}$, the corresponding residues are much smaller and this may create enough space to allow binding and processing of longer substrates.

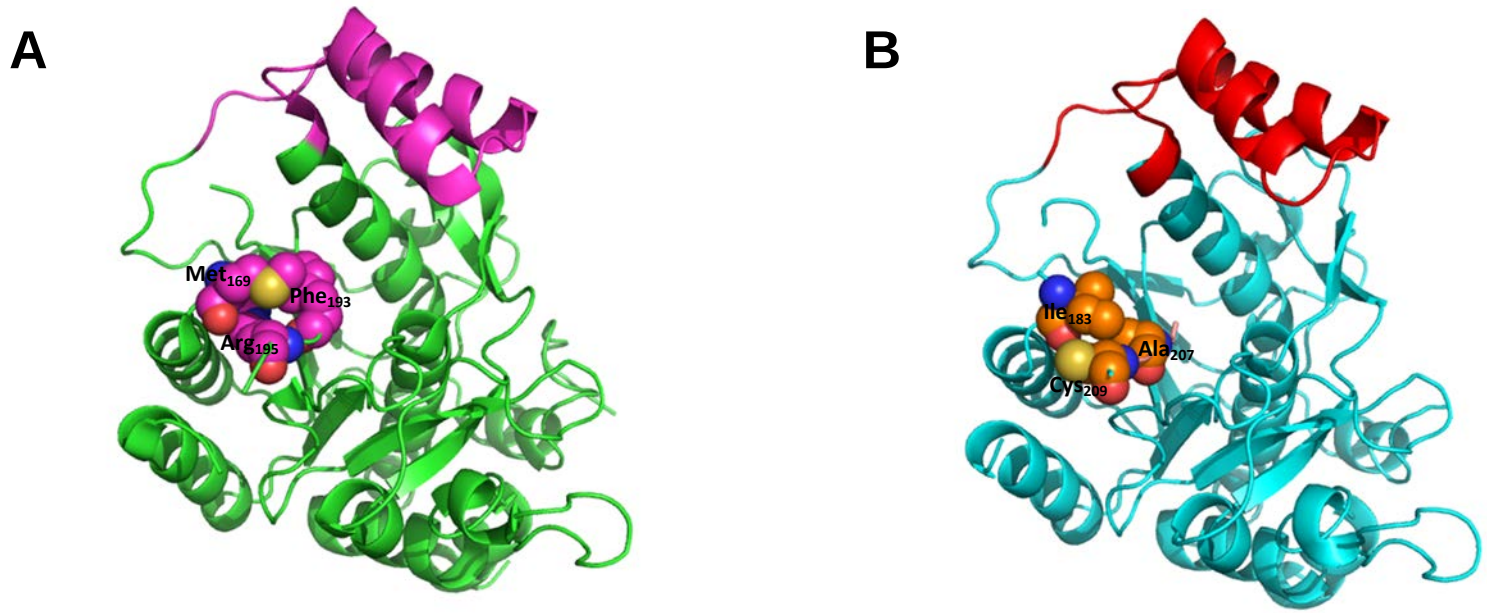


Figure S55: A) Crystal structure of $PatG_{mac}$ (PDB: 4AKS) showing the helical insertion (magenta) and the three bulky residues Met_{169}, Phe_{193} and Arg_{195} that protrude into the substrate binding groove and thus restrict the binding to substrates that adopt a bent-back conformation. B) Phyre 2 (Nat Protocols, 2015, 10, 845-858) Model of $OscG_{mac}$ showing the helical insertion (red) and the three residues Ile_{183}, Ala_{207} and Cys_{209} that correspond to the three residues highlighted above in $PatG_{mac}$.

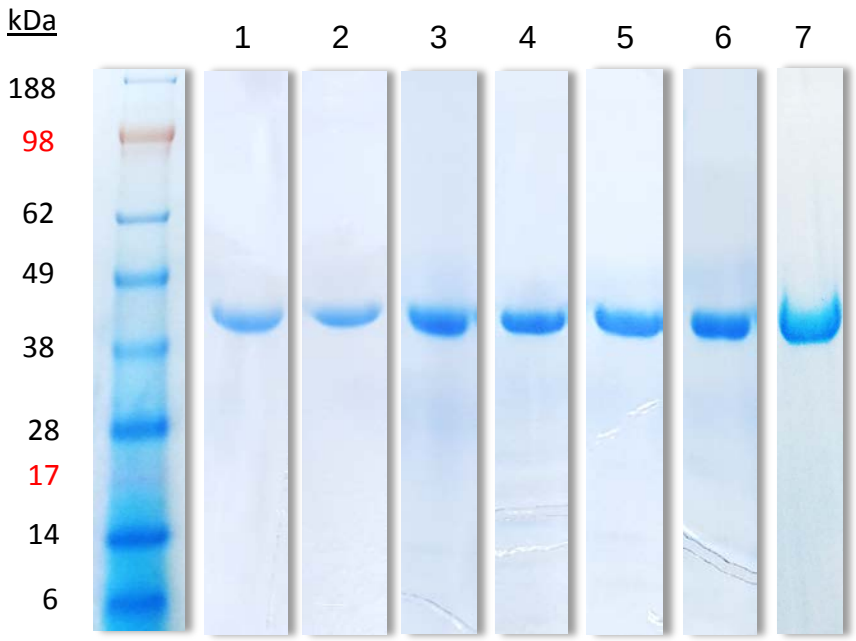


Figure S56: SDS-PAGE gel showing thepurified $PatG_{mac}$ mutants; 1 = M169I, 2 = F193A, 3 = R195C, 4 = M169I F193A, 5 = M169I R195C, 6 = F193A R195C, 7 = M169I F193A R195C against SeeBlue® Plus2 Pre-Stained protein standard (left lane).

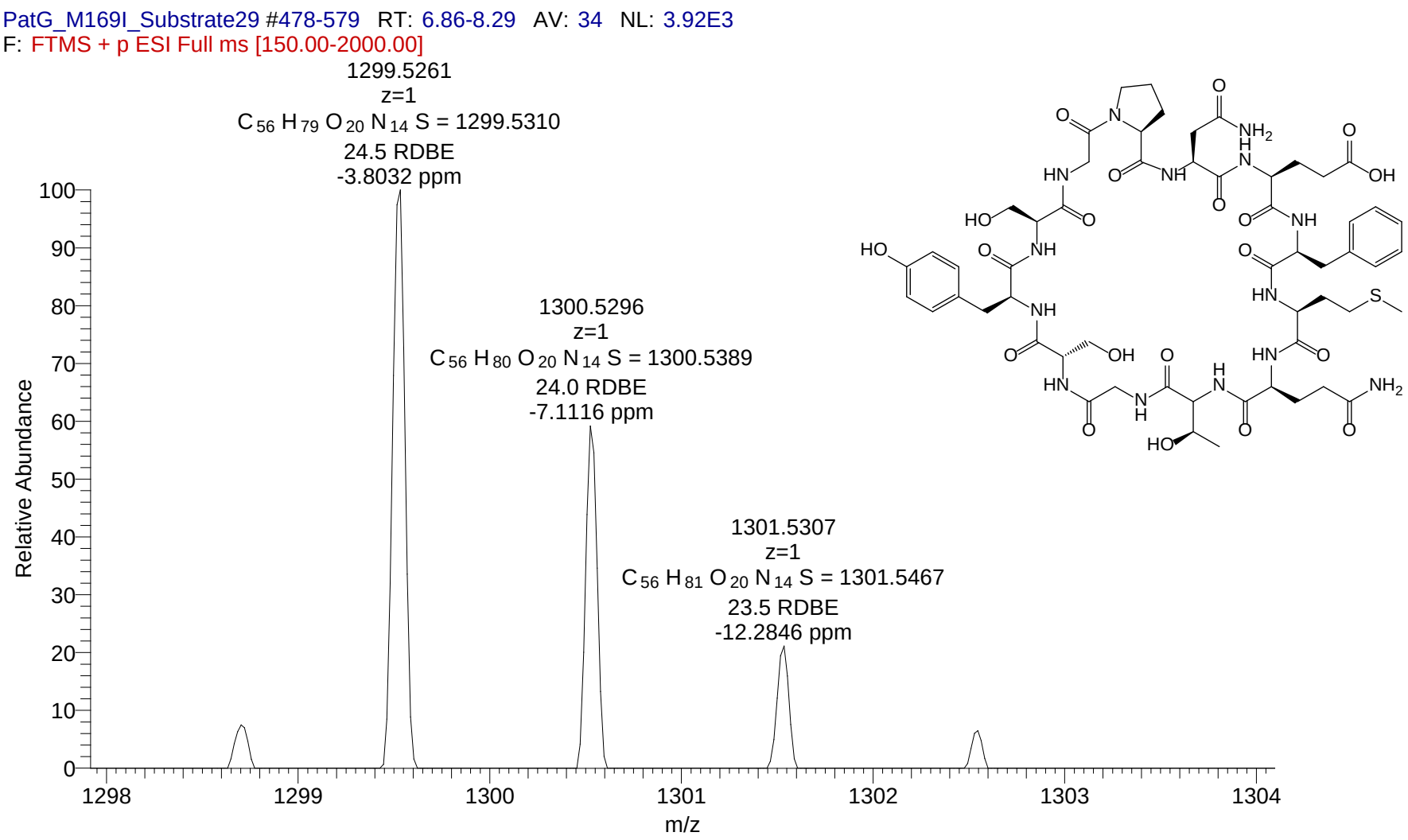


Figure S57: HR LCMS of the cyclic peptide product cyclo[-Asn-Glu-Phe-Met-Gln-Thr-Gly-Ser-Tyr-Ser-Gly-Pro] (1) generated with $PatG_{mac}$ mutant M169I.

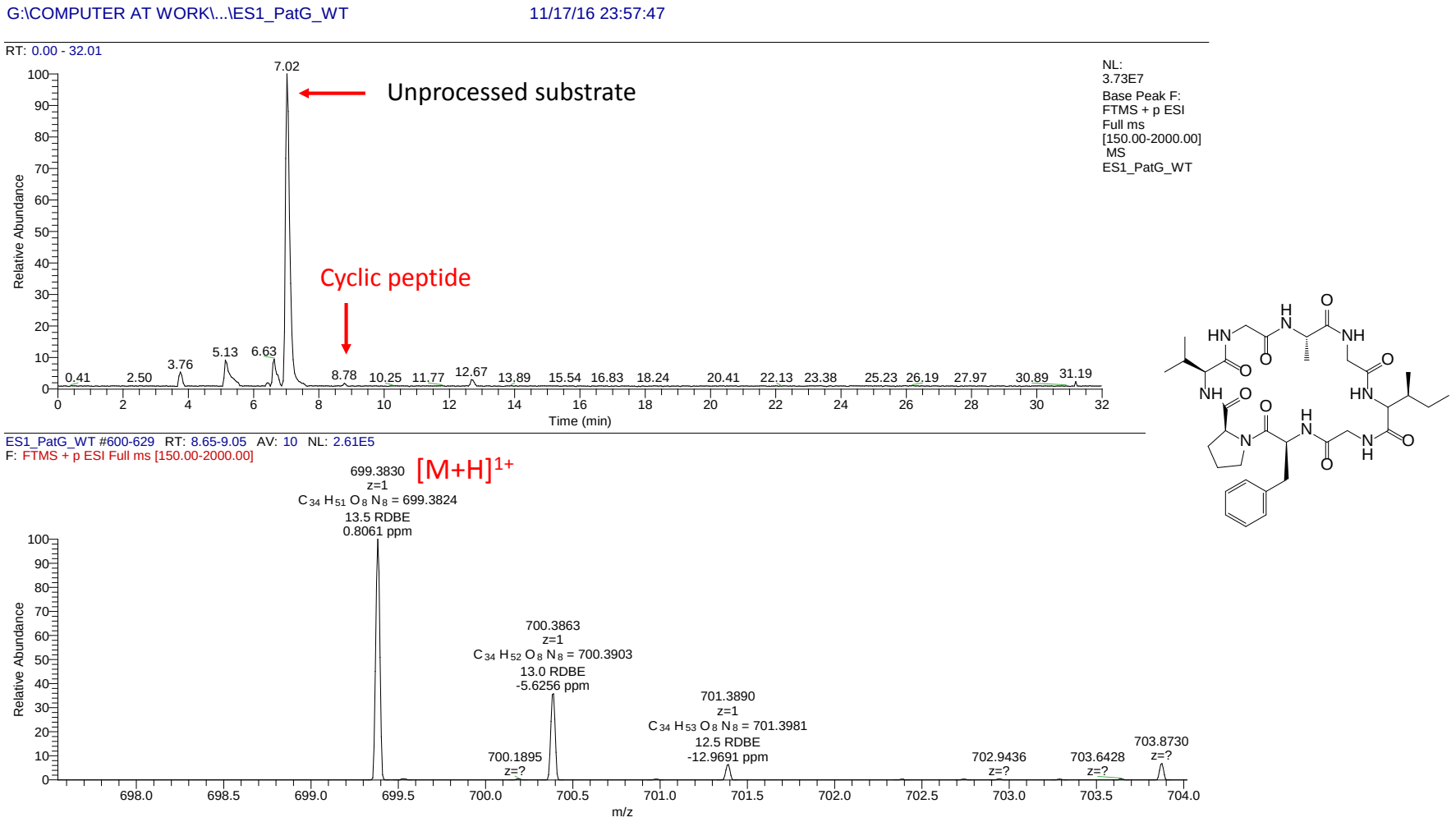


Figure S58: Processing of the peptide substrate H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Pro-Ala-Tyr-Asp-OH (16) with $PatG_{mac}$. Upper panel showing the total ion chromatogram and the lower panel is the MS data for the cyclic peptide.

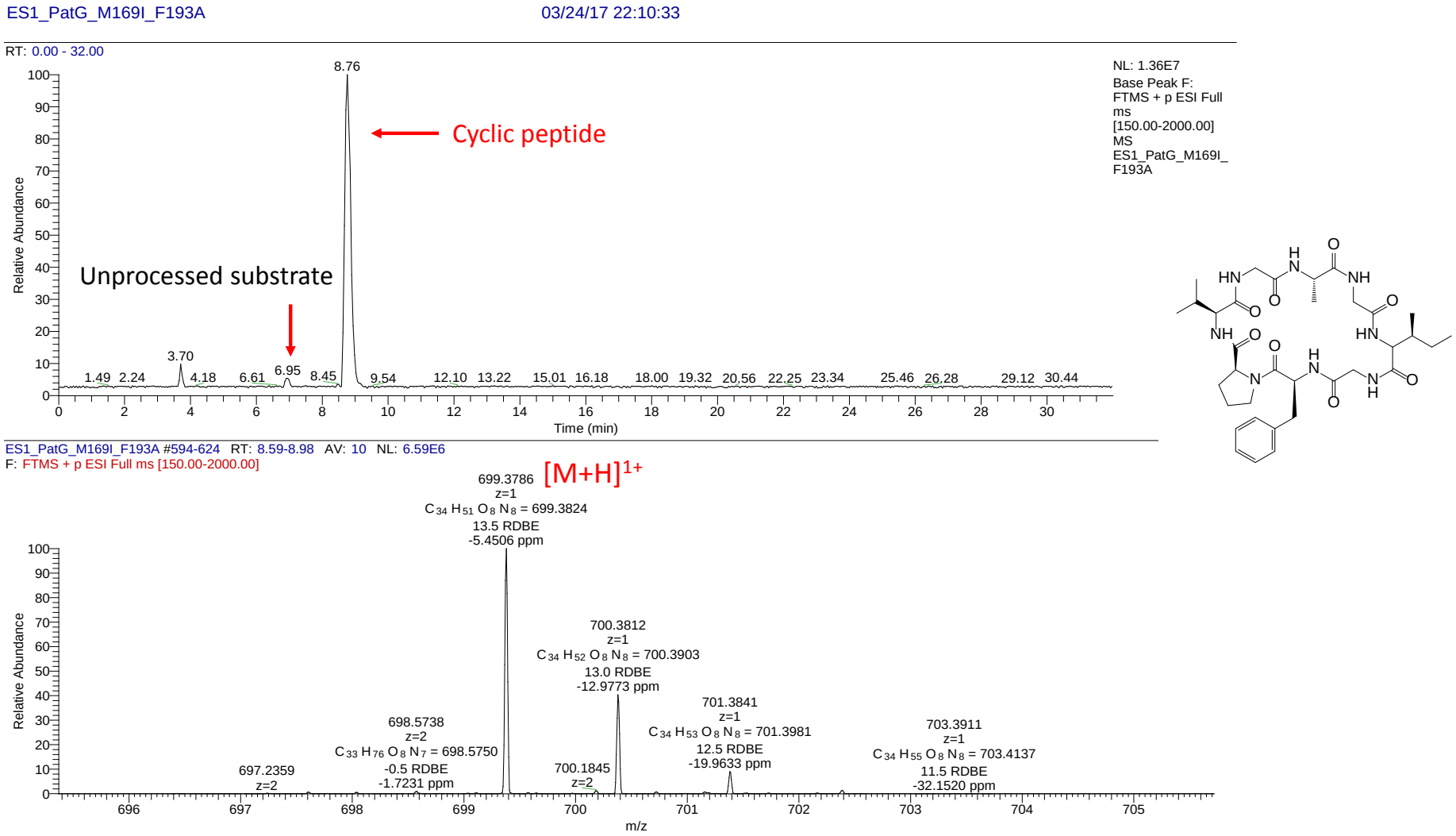


Figure S59: Processing of the peptide substrate H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Pro-Ala-Tyr-Asp-OH (16) with $PatG_{mac}$ mutant M169I F193A. Upper panel showing the total ion chromatogram and the lower panel is the MS data for the cyclic peptide.

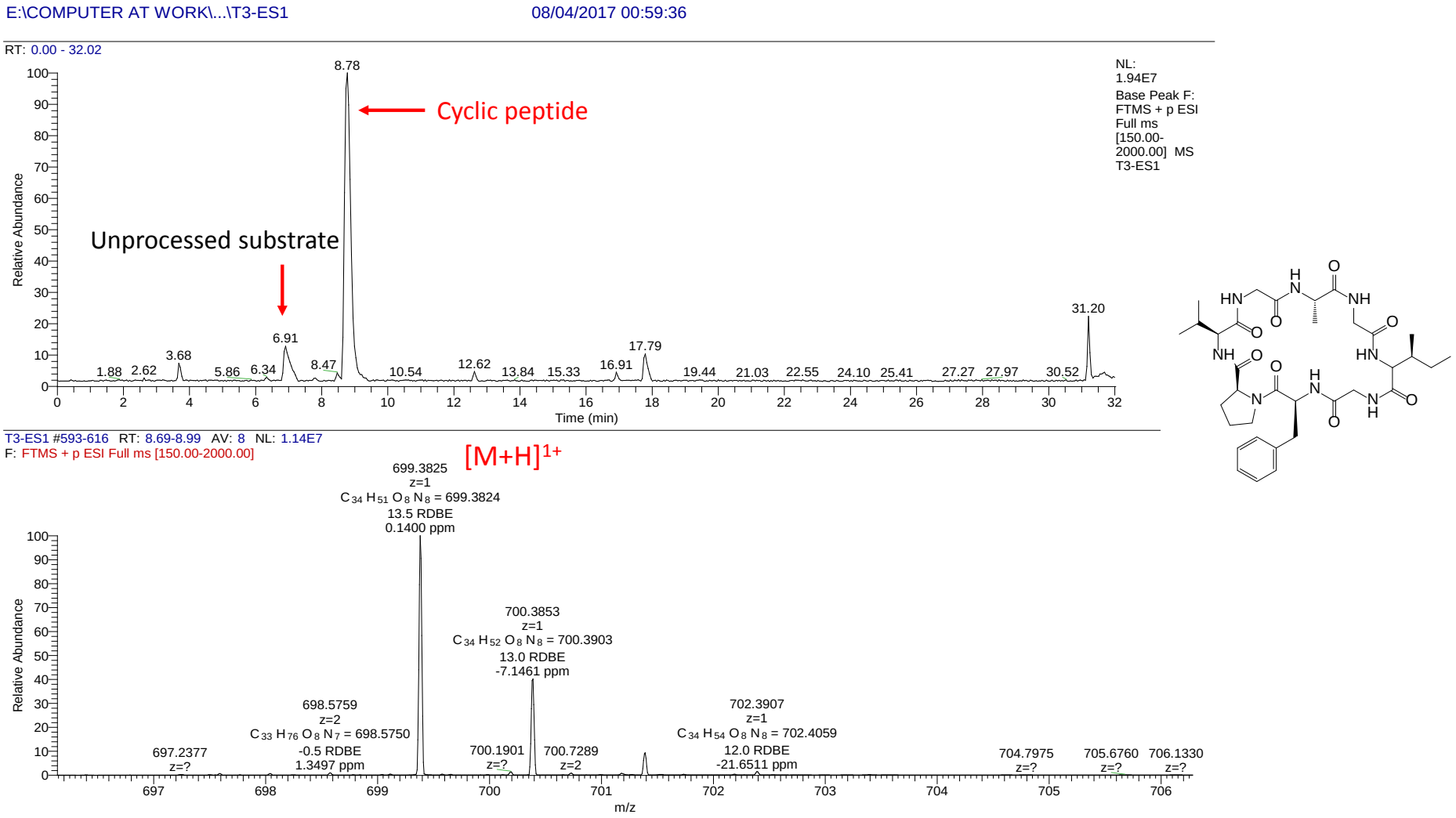


Figure S60: Processing of the peptide substrate H-Val-Gly-Ala-Gly-Ile-Gly-Phe-Pro-Ala-Tyr-Asp-OH (16) with $PatG_{mac}$ mutant M169I F193A R195C. Upper panel showing the total ion chromatogram and the lower panel is the MS data for the cyclic peptide.

Table S1. Predicted function of the proteins encoded in *osc* gene cluster (Fig. 2) from *Oscillatoria* sp. PCC 6506

Protein	Accession number	Length	Predicted function	Organism	Max identity
ORF	CBN56371	548	Hypothetical protein	*Microcystis aeruginosa* PCC 9809	60
OscC	CBN56372	75	Hypothetical protein	*Cyanothece* sp. PCC 7822	55
OscB	CBN56373	68	Hypothetical protein	*Nodularia spumigena* CCY 9414	79
OscA	CBN56374	659	N-terminal protease	*Anabaena* sp. 90	75
OscE1	CBN56375	83	Precursor peptide	*Nostoc spongiaeforme*	43
ORF	CBN56376	48	Transposase	*Microcystis aeruginosa* T1-4	75
OscF	CBN56377	310	Prenyltransferase	*Cyanothece* sp. PCC 7425	57
OscE2			Precursor peptide		
OscD	CBN56378	789	Heterocyclase	*Rivularia* sp. PCC 7116	66
OscG	CBN56379	1167	Oxidase/ C-terminal protease	*Rivularia* sp. PCC 7116	56

Table S2. Product ions of the protonated oscillacyclamide A or cyclo[-Trp-Gly-Asp-Lys-Tyr-Ala-Thz-Ser-Asn-Thr-Trp(Prenyl)-Gly-Asp-Thz-Val-Val-Thz] and the double protonated oscillacyclamide B or cyclo [-Gly-Gln-Phe-Gly-Ser-Thz-Asn-Thz-Ser-Val-Ser-Ser-Gly-Asp-Tyr-Trp-Ser-Thr-Phe-Asn-Leu-Gly-Pro-Leu-Thz]. Product ions 1 - 4 from the protonated oscillacyclamide A prove the macrocyclic nature of the peptide while ions 13 - 18 and 20 prove the presence of thiazoles and ions 22 - 23, 25 - 27, 29 - 30 and 32 prove the presence of the immonium ion forming amino acids in the molecule. Product ions 5 - 8 from the double protonated oscillacyclamide B prove the macrocyclic nature of the peptide while ions 13 - 14, 16 and 18 - 20 prove the presence of thiazoles and ions 22 - 24, 26 and 28 - 32 prove the presence of the immonium ion forming amino acids in the molecule.

No	Fragment sequence	Oscillacyclamide A			Oscillacyclamide B		
		$[M+H]^+$		Δ	$[M+H]^+$		Δ
		Calculated	Measured	(ppm)	Calculated	Measured	(ppm)
1	Val-Thz-Trp-Gly-Asp-Lys-Tyr	832.3447	832.3461	1.7	-	-	-
2	Val-Thz-Trp-Gly-Asp	541.1864	541.1879	2.7	-	-	-
3	Val-Thz-Trp-Gly	426.1594	426.1590	-1.1	-	-	-
4	Val-Thz-Trp	369.1380	369.1366	-3.8	-	-	-
5	Gly-Pro-Leu-Thz - Gly-Gln-Phe / Pro-Leu-Thz - Gly-Gln-Phe-Gly	-	-	-	683.2970	683.2976	0.8
6	Pro-Leu-Thz - Gly-Gln-Phe	-	-	-	626.2755	626.2763	1.1
7	Pro-Leu-Thz - Gly-Gln	-	-	-	479.2071	479.2071	-0.1
8	Gly-Pro-Leu-Thz / Pro-Leu-Thz - Gly	-	-	-	351.1485	351.1479	-2.0
9	[Thr-Trp(Prenyl)] − H_2O	338.1863	338.1861	-0.6	-	-	-
10	[Asn-Thr-Trp(Prenyl)] − H_2O	452.2292	452.2273	-4.3	-	-	-
11	Trp(Prenyl)-Gly-Asp-Thz	510.1806	510.1778	-5.5	-	-	-
12	Trp(Prenyl)	255.1492	255.1479	-5.3	-	-	-
13	Ala/Leu-Thz	155.0274	155.0275	0.6	197.0743	197.0758	7.3
14	[Ala/Leu –Thz] − NH_3	138.0008	138.0009	0.3	180.0478	180.0477	-0.6
15	Asp-Thz	199.0172	199.0160	-6.2	-	-	-
16	[Asp/Asn-Thz] − NH_3	181.9906	181.9906	-0.5	181.0066	181.0072	2.9
17	Val/Ser-Thz	183.0587	183.0599	6.5	171.0223	171.0225	1.1
18	[Val-Thz] − NH_3	166.0321	166.0320	-1.0			
19	[Ser-Thz] − H_2O	-	-	-	153.0117	153.0112	-3.7
20	Thiazole-CO	111.9852	111.9854	2.1	111.9852	111.9852	-0.1
	Immonium ions:						
21	Trp(prenyl)	227.1543	227.1532	-5.0	-	-	-
22	Trp	159.0917	159.0913	-2.6	159.0917	159.0918	0.5
23	Tyr	136.0757	136.0758	0.4	136.0757	136.0752	-4.0

24	Phe	-	-	-	120.0808	120.0805	-2.7
25	Lys	101.1073	101.1071	-2.7	-	-	-
26	Asp	88.0393	88.0391	-2.8	88.0393	88.0388	-6.2
27	Asn	87.0553	87.0549	-5.1	-	-	-
28	Leu/Ile	-	-	-	86.0964	86.0962	-3.3
29	Thr	74.0600	74.0596	-6.6	74.0600	74.0602	2.2
30	Val	72.0808	72.0814	7.9	72.0808	72.0819	14.8
31	Pro	-	-	-	70.0651	70.0660	11.7
32	Ser	60.0444	nd	-	60.0444	60.0452	12.7

Table S3. Some cyanobactin macrocyclases and their corresponding natural products:

Macrocyclase	Compound class	No of residues in the natural products	Natural recognition signal	Producer organism	Accession #	Ref.
$PatG_{mac}$	Patellamides, Lissoclinamides, Ulithiacyclamides	7, 8	Ala-Tyr-Asp	*Prochloron* sp.	AAY21156.1	S1
$McaG_{mac}$	Microcyclamides, Aerucyclamides	6, 7	Ala-Phe-Asp	*Microcystis aeruginosa* NIES-298 and PCC7806	CAO82089	S2
$TruG_{mac}$	Trunkamide, Patellins	6, 7, 8	Ser-Tyr-Asp	*Prochloron* sp.	ACA04494	S3
$TenG_{mac}$	Tenuecyclamides	6	Ala-Tyr-Asp	*Nostoc spongiaeforme* var. tenue str. Carmeli	ACA04486	S3
$LynG_{mac}$	Aesturamides	7	Ser-Tyr-Asp	*Lyngbya aestuarii* PCC 8106	WP_009787128	S3, S4
$TriK_{mac}$	Trichamide	11	Ser-Tyr-Asp	*Trichodesmium erythraeum* IMS101	YP_722058	S5
$PirG_{mac}$	Piricyclamides	9, 12, 14	Phe-Ala-Gly-Asp	*Microcystis aeruginosa* PCC7005	CCI37144	S6

References:

S1. E. W. Schmidt, J. T. Nelson, D. A. Rasko, S. Sudek, J. A. Eisen, M. G. Haygood, J. Ravel, *Proc. Natl. Acad. Sci. USA* **2005**, 102, 7315-7320.

S2. N. Ziemert, K. Ishida, P. Quillardet, C. Bouchier, C. Hertweck, N. T. de Marsac, E. Dittmann, *Appl. Environ. Microbiol.* **2008**, 74, 1791-1797.

S3. M. S. Donia, J. Ravel, E. W. Schmidt, *Nat. Chem. Biol.* **2008**, 4, 341-343.

S4. K. Sivonen, N. Leikoski, D. P. Fewer, J. Jokela, *Appl. Environ. Microbiol.* **2010**, 86, 1213-1225.

S5. S. Sudek, M. G. Haygood, D. T. A. Youssef, E. W. Schmidt, *Appl. Environ. Microbiol.* **2006**, 72, 4382-4387.

S6. N. Leikoski, D.P. Fewer, J. Jokela, M. Wahlsten, P. Alakoski, K. Sivonen, *PLoS One* **2012**, 7, e43002.