Supporting Information

Exploration of Human Xylosyltransferase for Chemoenzymatic Synthesis of Proteoglycan Linkage Region

Jia Gao,[a] Po-han Lin,[a] Setare Tahmasebi Nick,[a] Kunli Liu,[a] Kefei Yu,[b] Erhard Hohenester,[c] Xuefei Huang[a,d*]

[a]Department of Chemistry and Institute for Quantitative Health Science and Engineering, Michigan State University, East Lansing, MI 48824 USA

[b]Department of Microbiology & Molecular Genetics, Michigan State University, East Lansing, MI 48824 USA

[c]Department of Life Sciences, Imperial College London, London SW7 2AZ, UK

[d]Department of Biomedical Engineering, Michigan State University, East Lansing, Michigan 48824, USA

*Email: huangxu2@msu.edu

Supporting Information

Table of Contents

Materials. Signal peptide-His6-XT-I gBlocks gene was purchased from Integrated DNA Technologies (Coralville, IA). FreeStyle 293 Expression Medium and Coomassie Brilliant Blue G-250 were purchased from Thermo Fischer Scientific (Waltham, MA). Nickel columns and Nickel resins were purchased from Bio-rad (Hercules, California). SDS-PAGE gels, 10x Tris/Glycine/SDS electrophoresis buffer, prestained protein ladder, sample loading buffer, and Coomassie Blue R-250 were purchased from Bio-rad (Hercules, California). Tris-HCl buffer was purchased from MilliporeSigma (St. Louis, MO). UDP-xylose was purchased from Complex Carbohydrate Research Center (Athens, Georgia). Amino acid building blocks were purchased from Chem-Impex International, Inc (Wood Dale, IL). Cy5-alkyne was purchased from MilliporeSigma (St. Louis, MO). Glycosyltransferase Activity Kit was purchased from R&D Systems. All other chemical reagents were purchased from commercial sources and used without additional purifications unless otherwise noted.

General Information. High-performance liquid chromatography was carried out with LC-8A Solvent Pumps, DGU-14A Degasser, SPD-10A UV-Vis Detector, SCL-10A System Controller (Shimadzu Corporation, JP) and Vydac 218TP 10 μm C18 Preparative HPLC column (HICHROM Limited, VWR, UK) or 20RBAX 300SB-C18 Analytical HPLC column (Agilent Technologies, CA) using HPLC-grade acetonitrile (EMD Millipore Corporation, MA) and Milli-Q water (EMD Millipore Corporation, MA). A variety of eluting gradients were set up on LabSolutions software (Shimadzu Corporation, JP)). The dual-wavelength UV detector was set at 220 nm and 254 nm for monitoring the absorbance from amide and Fmoc-, correspondingly. 3D structure of glycopeptide compounds was prepared with Maestro software. Docking simulations were acquired with AutoDock Vina and UCSF Chimera (UCSF, CA). Enzymatic activity was quantified by absorbance at 620 nm using a SpectraMax M3 96-well plate reader (Molecular Devices, CA). NMR data were obtained with DirectDrive2 500 MHz (Agilent, CA) at ambient temperature.

XT-I Expression, Purification and Characterization. Expi293F cells were grown in FreeStyle™ 293 Expression Medium on a platform shaker in humidified 37 °C CO_2 (5%) incubator with rotation at 150 rpm. When the cell density reached between 4 x 10^5 and 3 x 10^6 cells/ml, cells were split to a density of 1 x 10^6 cells/ml and cultured overnight in the same condition. Cells were then transfected with His6-XT-I gene 24 hours after they were split. Before transfection, cells were harvested by centrifugation at 1200 rpm for 10 min at room temperature and re-suspended in fresh pre-warmed media. To transfect the cells, a final concentration of 2-3 μg/ml of the XT-I gene and 9 μg/ml of PEI were added. PEI stock solution was prepared at the concentration of 1 mg/ml in a buffer containing 25 mM HEPES and 150 mM NaCl, pH 7.4. The flask was returned to the shaker platform in the incubator. Cells were diluted 1:1 with pre-warmed media supplemented with valproic acid (VPA) to a final concentration of 2.2 mM. Four to six days after

the transfection, cells were harvested. Clarified lysate was purified by nickel column (Cytiva, MA) (a. washing buffer: 20 mM Tris, 0.5 M NaCl and 40 mM imidazole; b. eluting buffer: 20 mM Tris, 0.5 M NaCl and 40-250 mM imidazole). Protein purity was confirmed with SDS-PAGE gel electrophoresis and the concentration and expression yield were determined by standard Bradford assay.

CD44 Expression, Purification and Characterization. The His tagged human $CD44_{20\text{-}178}$ protein was cloned and expressed following a revised literature procedure.[1] The His tagged human $CD44_{20\text{-}178}$ gene was purchased from GenScript, and was cloned into pET-28a(+) vector at NdeI/XhoI site. The resulting plasmids were transformed into the *E. coli* strain Rosetta 2(DE3)pLysS for bacterial expression. During inoculation, *E. coli* was grown in 37 °C incubator and shaken at 270 RPM overnight. Optical density was periodically checked until it reached mid-log value between 0.6 and 0.9. IPTG was then add to the media for induction. It was allowed to grow for another 4 hours at 270 RPM and 37 °C, followed by centrifugation and sonication to gather inclusion bodies. The inclusion bodies were subject to repeated centrifugation at 9000 g for 20 mins to remove cell debris. Denatured human CD44 were then refolded in buffer containing 250mM L-arginine and 2mM reduced and 1mM oxidized glutathione. The solution was slowly stirred at 4°C followed by ultrafiltration to concentrate it down. Protein purity was confirmed with SDS-PAGE gel electrophoresis. The concentration and expression yield were determined by the standard Bradford assay. The sequence of the expressed protein was the following: MGSSHHHHHHSSGLVPRGSHMAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPT MAQMEKALSIGFETCRYGFIEGHVVIPRIHPNSICAANNTGVYILTSNTSQYDTYCFNASAPPEED CTSVTDLPNAFDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDV.

General Procedure for Automated Solid-Phase Peptide Substrate Synthesis. All the peptides were synthesized on a Liberty Blue™ Automated Microwave Peptide Synthesizer following the standard Fmoc-based solid-phase peptide synthesis protocol. The Cl-MPA ProTide resins were purchased from CEM Corporation. The Liberty Blue software (CEM Corporation, NC) was used to program the synthesis, including resin swelling, amino acid loading, couplings and Fmoc- removal. Commercially available *N,N*-dimethylformamide (DMF) from Fischer Chemical was supplied to the synthesis module as reaction and washing solvent. Peptide synthesis was enabled by sequential couplings of Fmoc-amino acid (purchased from Chem-Impex), which was preactivated by *N*,*N*-diisopropylcarbodiimide (DIC), Oxyma Pure and *N*,*N*-diisopropylethyl amine (DIPEA) at 50 °C for 10 min, and deprotections with 20% piperidine in DMF at 60 °C for 4 min. In-between each coupling/deprotection step, resin-bound peptide was thoroughly washed with DMF. Resin-bound peptides were cleaved off the solid support with a cocktail solution of trifluoroacetic acid (TFA), triisopropylsilane (TIPS) and water (TFA/TIPS/H_2O, 95:2.5:2.5). The crude peptides were then

purified with reverse-phase C18 preparative HPLC. Compound purity was confirmed by C18 analytical HPLC analysis.

General Procedure for XT-I-Catalyzed Glycosylation. The 10x 2-(*N*-morpholino)ethanesulfonic acid (MES) reaction buffer for XT-I-catalyzed glycosylation was prepared in advance following the recipe of 250 mM MES, 250 mM KCl, 50 mM KF, 50 mM $MgCl_2$, 50 mM $MnCl_2$. The pH of the 10x reaction buffer was adjusted to 6.5 by adding concentrated NaOH solution. A solution of 1 mM peptide substrate and 1.1-3.0 mM UDP-xylose (1.1-3.0 equiv. per glycosylation site, depending on peptide acceptors) was made with the reaction buffer. The addition of XT-I enzyme (0.02 mol%) initiated the glycosylation. The reaction solution was kept at 37 °C overnight. The reaction progress was monitored with LC-MS. After the reaction, the enzyme was deactivated and precipitated out of the reaction mixture by addition of ethanol. The mixture was centrifuged and the supernatant was loaded onto a G-10 size exclusion column for purification.

Procedure for XT-I-Catalyzed Transfer of UDP-6-Azidoglucose to Yield Compound 12. A solution of 0.5 mM peptide **11** and 2.5 mM UDP-6-azidoglucose (5 equiv. per glycosylation site) was made with the MES reaction buffer. The addition of XT-I enzyme (0.1 mol%) initiated the glycosylation. The reaction solution was kept at 37 °C overnight. The reaction progress was monitored with LC-MS and the formation of glycopeptide **12** was confirmed. After the reaction, the enzyme was deactivated and precipitated out of the reaction mixture by addition of ethanol. The mixture was centrifuged and the supernatant was carried over without further purification. (ESI-MS: $C_{54}H_{87}N_{19}O_{26}^{2-}$, calcd: 708.8038, obsd: 708.7978 (8.5 ppm))

Procedure for Copper (I)-Catalyzed Azide-Alkyne Cycloaddition to Yield Compound 13. To a solution of azide-tagged glycopeptide **12** (100 μM), $CuSO_4$ (20 mM), (tris-hydroxypropyltriazolylmethylamine) (THPTA) ligand (10 mM), aminoguanidine (100 mM), Cy5-alkyne (1 mM), and Na ascorbate (100 mM) were added. The reaction tube was attached to a 20 round-per-minute (rpm) end-over-end rotator. The reaction was allowed to proceed for 2 hours at room temperature. The formation of Cy5 conjugated glycopeptide **13** was confirmed using LC-MS (ESI-MS: $C_{91}H_{131}N_{22}O_{36}S_3^{2-}$, calcd: 734.6092, obsd: 734.5991 (13.7 ppm))

General Procedure for One-Pot Two-Enzyme (OP2E) Glycosylation. The 10x MES reaction solution for XT-I and β4GalT7 OP2E glycosylation was prepared following the recipe of 225 mM MES, 125 mM KCl, 25 mM KF, 25 mM $MgCl_2$, 75 mM $MnCl_2$. A solution of 1 mM peptide and 1.5-3.0 mM UDP-xylose (1.5-3.0 equiv. per glycosylation site, depending on peptide acceptors) and 2.0 mM UDP-galactose (2.0 equiv. per glycosylation site) was made with the reaction buffer. XT-I enzyme (0.05 mol%) and β4GalT7

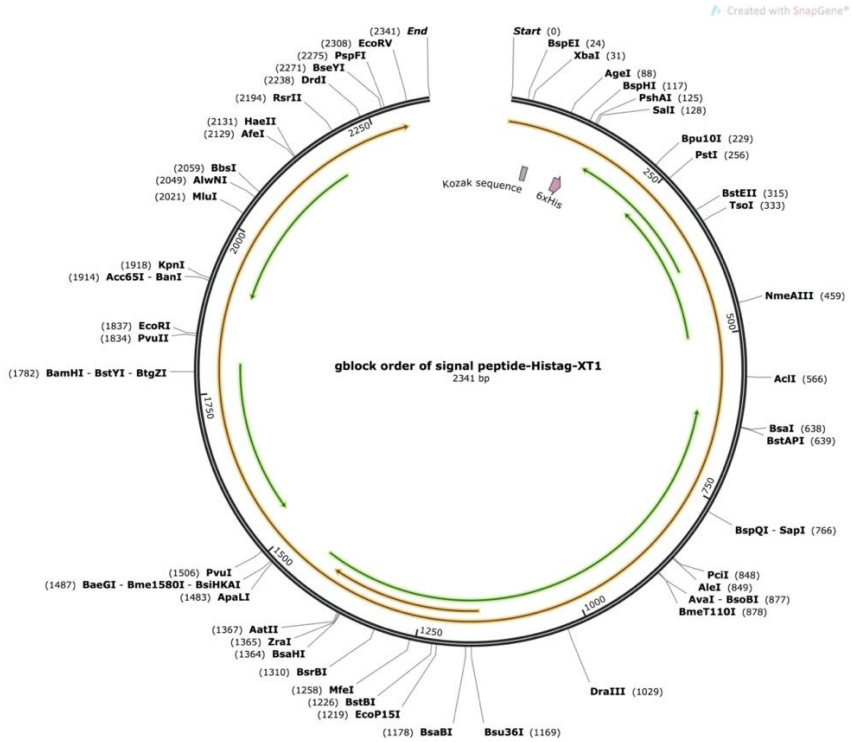
enzyme (0.5 mol%) were added to initiate the glycosylation reactions. The reaction tube was kept in an incubator at 37 °C overnight. The reaction progress was monitored via LC-MS. Upon reaction completion, the reaction mixture was directly injected into HPLC, and the reaction yield was quantified from peak areas of HPLC chromatograms. The identities of all the glycopeptide products were confirmed by HRMS analysis as well as HPLC retention time as compared to the reported compounds.[2]

Phosphatase-Coupled Enzymatic Kinetic Assay. The kinetic assay protocol follows the general assay conditions reported by R&D Systems Inc. with modifications.[3]

1) Prepare 30 µL reaction solutions of UDP-xylose, peptide acceptor and XT-I enzyme in the 96-well plate;
2) Cover the plate with a plate sealer, and incubate the plate at 37 °C for 20 min;
3) Quickly add 12 µL 10x phosphatase assay buffer, 3 µL $MnCl_2$ solution (100 mM), 3 µL MilliQ water and 2 µL coupling phosphatase 1 (20 ng/µL), to a total volume of 50 µL;
4) Cover the plate with a plate sealer again, and incubate at 37 °C for 20 min;
5) Add 30 µL of Malachite Green Reagent A to each well. Mix gently by tapping the plate;
6) Add 100 µL of deionized or distilled water to each well;
7) Add 30 µL of Malachite Green Reagent B to each well. Mix gently by tapping the plate;
8) Incubate the plate for 5 minutes at room temperature to have consistent color development;
9) Determine the optical density of each well using a microplate reader set to 620 nm, and adjust the OD by subtracting the reading of the negative control;
10) Calculate product formation using the conversion factor determined from the phosphate standard curve.

Supporting Figure and Table

Signal peptide-His6-XT-I Sequence



5’AAGACACCGGGACCGATCCAGCCTCCGGACTCTAGAGCCGCCACCATGGGTTGGAGTTGTATCATCCTTTTCCTG
GTAGCTACCGCAACCGGTGTTCATTCACATCACCACCATCATCATGACGTAAGTCGACCTCCTCACGCAAGAAAGA
CGGGTGGCTCTAGCCCGGAGACTAAGTATGACCAGCCGCCGAAGTGCGACATTAGCGGTAAAGAAGCGATCTCTG
CCCTGAGCCGGGCAAAATCAAAACACTGCAGACAGGAGATTGGTGAGACGTATTGCCGACACAAACTGGGGCTCC
TCATGCCAGAGAAGGTAACCAGATTTTGTCCGCTGGAGGGGAAGGCCAACAAAAACGTCCAATGGGACGAGGATA
GCGTCGAATACATGCCTGCGAATCCCGTCAGGATCGCGTTTGTCCTTGTCGTTCATGGCCGAGCGAGCAGACAACT
TCAGCGCATGTTTAAAGCAATCTACCACAAAGACCATTTCTATTATATTCATGTCGATAAGCGGTCAAACTACCTG
CACCGGCAGGTACTCCAGGTTTCACGCCAATACTCCAACGTTCGCGTAACTCCATGGCGGATGGCCACGATCTGGG
GTGGGGCTTCACTCCTCTCAACGTATTTGCAGAGCATGCGAGACCTTCTGGAAATGACTGACTGGCCATGGGACTT
TTTCATCAATTTGAGCGCAGCCGACTATCCAATCCGAACCAATGATCAGCTTGTAGCATTTCTGAGTCGCTATAGG
GACATGAATTTCCTGAAGAGCCATGGGCGGGATAACGCGCGGTTCATACGAAAGCAAGGGCTGGATAGGCTGTTT
CTTGAATGCGACGCACACATGTGGAGGCTTGGGGATAGAAGGATTCCCGAGGGGATCGCCGTGGATGGAGGAAGC
GACTGGTTCCTTCTGAATCGACGGTTTGTCGAGTATGTCACGTTCAGCACGGATGATTTGGTCACGAAAATGAAAC
AATTCTACAGTTATACGCTCCTGCCCGCTGAGAGCTTCTTCCACACGGTGTTGGAAAACTCCCCGCATTGTGATACA
ATGGTTGATAATAATTTGAGGATTACAAATTGGAATCGAAAACTTGGGTGCAAATGTCAGTATAAGCATATAGTGG
ACTGGTGTGGATGTTCTCCTAATGACTTTAAACCTCAGGATTTTCATCGATTCCAGCAGACAGCACGGCCTACTTTT
TTTGCGCGAAAATTCGAAGCAGTCGTCAATCAAGAGATTATCGGACAATTGGATTACTACCTGTATGGAAACTATC
CTGCCGGTACGCCTGGGCTCCGCTCCTATTGGGAGAATGTCTATGATGAACCTGACGGAATACATTCCCTTAGTGA
CGTCACCCTCACTCTTTATCATAGTTTTGCACGCTTGGGTCTGAGACGGGCCGAAACTTCTCTTCATACAGACGGCG
AAAACAGTTGTCGCTATTACCCGATGGGCCACCCCGCATCAGTGCACCTTTATTTCCTGGCCGATCGATTCCAGGG
GTTTCTGATCAAGCATCATGCGACAAACCTCGCAGTGAGCAAATTGGAAACTCTTGAAACCTGGGTGATGCCCAAA
AAAGTGTTCAAAATCGCTAGTCCTCCCTCCGACTTTGGTAGGTTGCAGTTCTCCGAAGTAGGGACAGATTGGGACG
CGAAGGAGAGACTGTTTCGGAACTTCGGCGGGTTGTTGGGACCGATGGATGAGCCAGTTGGCATGCAAAAGTGGG
GCAAAGGGCCTAACGTCACTGTAACAGTGATCTGGGTGGATCCAGTCAACGTCATCGCCGCAACTTACGATATACT
GATTGAGAGTACAGCTGAATTCACCCACTATAAACCGCCCTTGAACCTTCCCCTGCGACCTGGAGTGTGGACCGTT
AAGATTCTTCACCACTGGGTACCTGTGGCGGAGACGAAATTTTTGGTGGCCCCGTTGACTTTTTCCAATCGACAAC
CTATAAAGCCTGAAGAGGCCCTTAAACTGCACAACGGTCCACTGCGAAACGCGTATATGGAACAGTCTTTCCAGTC

TCTGAACCCTGTACTTAGTCTTCCAATAAATCCGGCCCAAGTTGAGCAAGCCCGGCGGAATGCCGCTTCCACTGGA
ACAGCGCTCGAAGGATGGCTTGATAGCCTGGTTGGAGGTATGTGGACAGCCATGGACATCTGCGCCACCGGACCG
ACCGCGTGTCCGGTGATGCAAACTTGTTCTCAGACTGCGTGGTCTAGCTTCTCACCTGATCCAAAGTCCGAGCTGG
GCGCAGTGAAACCCGACGGTAGACTTAGGTGATATCTCGACAATCAACCTCTGGATTACAAAATTT 3'

Figure S1. XT-I gene sequence.

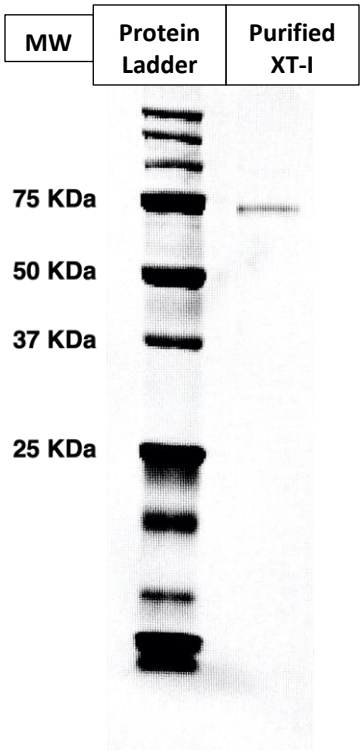


Figure S2. SDS-PAGE gel of purified XT-I.

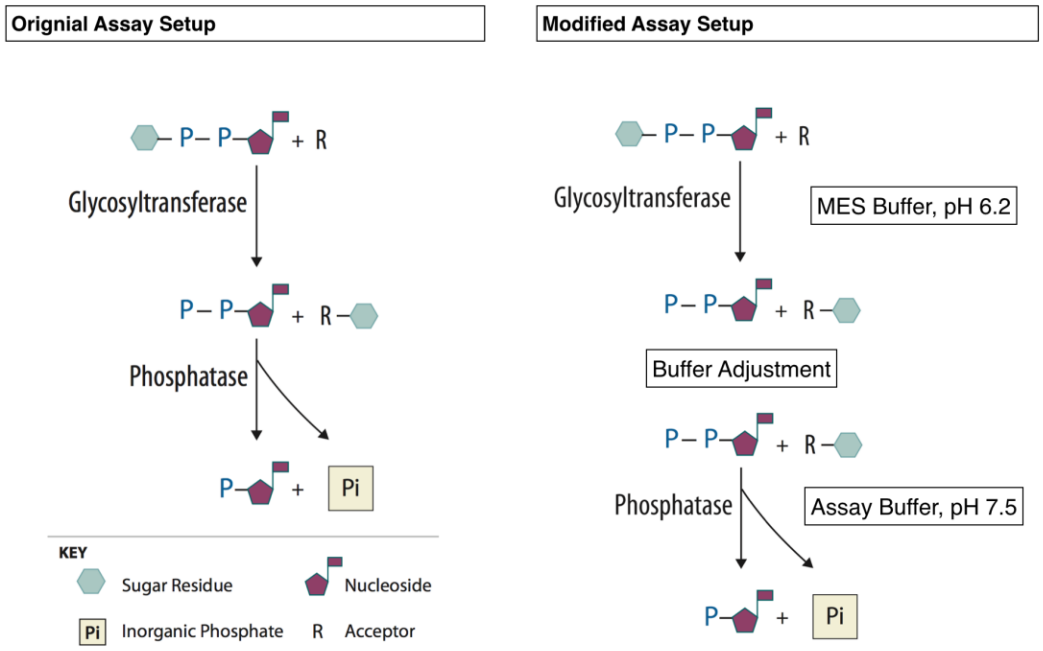


Figure S3. Schematic demonstrations of the original and the modified kinetic assay set-up.[3]

Phosphate Standard

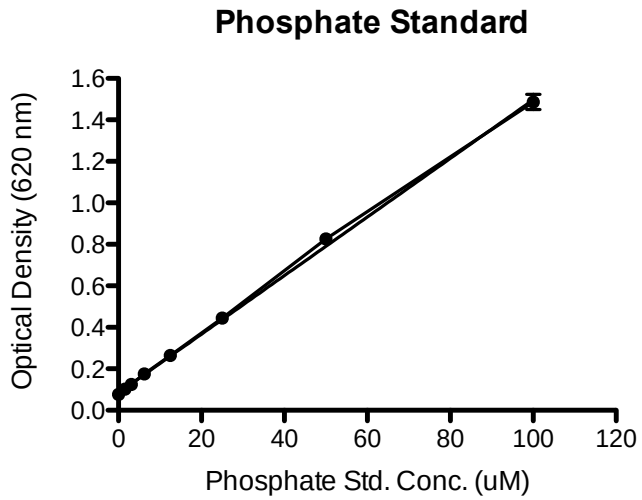


Figure S4. Phosphate conversion factor measurement. Conversion Factor = 3541 pmol/OD (Plot is displayed as mean ± S.D. of two replicates, phosphate standard volume = 50 µL).

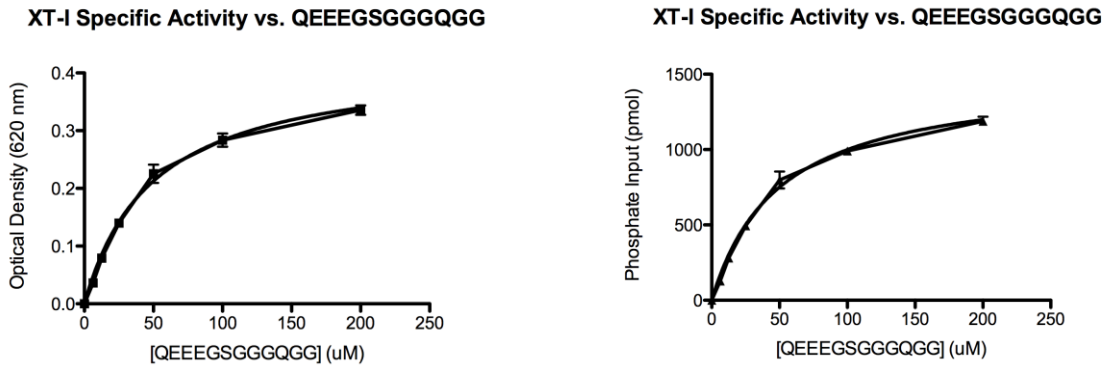


Figure S5. Phosphatase-coupled assay result of QEEEGSGGGQGG **1**. k_{cat} = 28 min^{-1}, K_m = 49.8 μM, k_{cat}/K_m = 562 mM^{-1} min^{-1}.

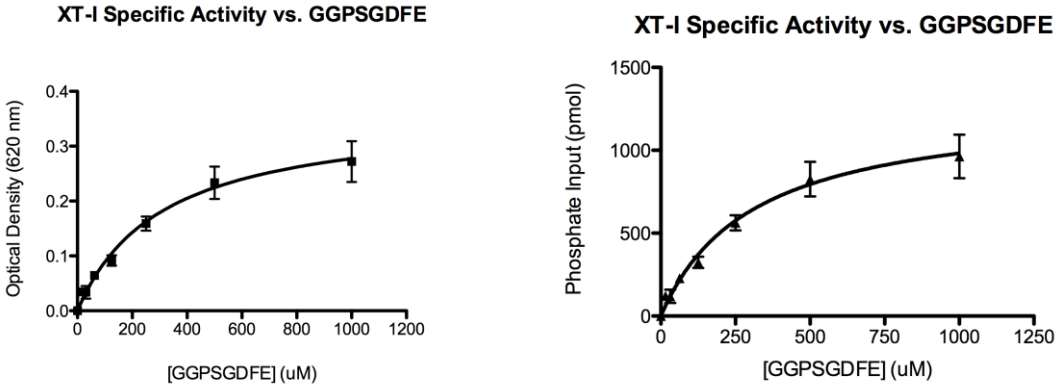


Figure S6. Phosphatase-coupled assay result of GGPSGDFE **3**. k_{cat} = 3 min^{-1}, K_m = 308.0 μM, k_{cat}/K_m = 10 mM^{-1} min^{-1}.

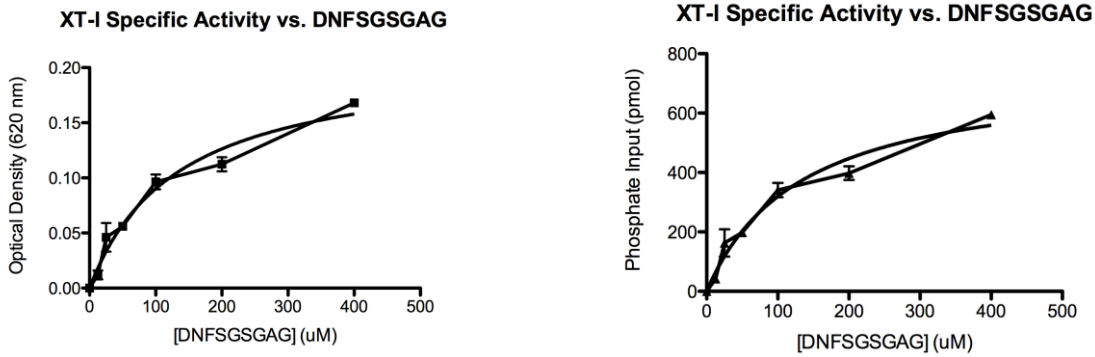


Figure S7. Phosphatase-coupled assay result of DNFSGSGAG **4**. k_{cat} = 16 min^{-1}, K_m = 133.8 μM, k_{cat}/K_m = 120 mM^{-1} min^{-1}.

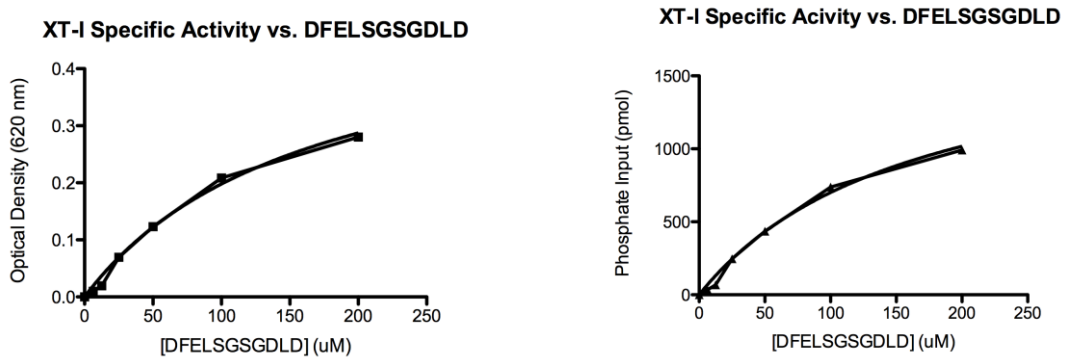


Figure S8. Phosphatase-coupled assay result of DFELSGSGDLD **5**. k_{cat} = 15 min^{-1}, K_m = 164.4 μM, k_{cat}/K_m = 91 mM^{-1} min^{-1}.

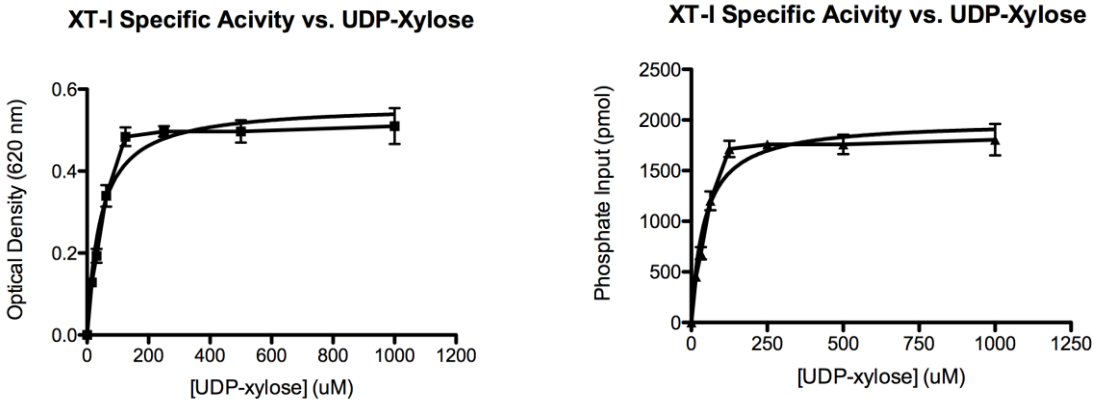


Figure S9. Phosphatase-coupled assay result of UDP-xylose. k_{cat} = 13 min^{-1}, K_m = 43.4 μM, k_{cat}/K_m = 266 mM^{-1} min^{-1}.

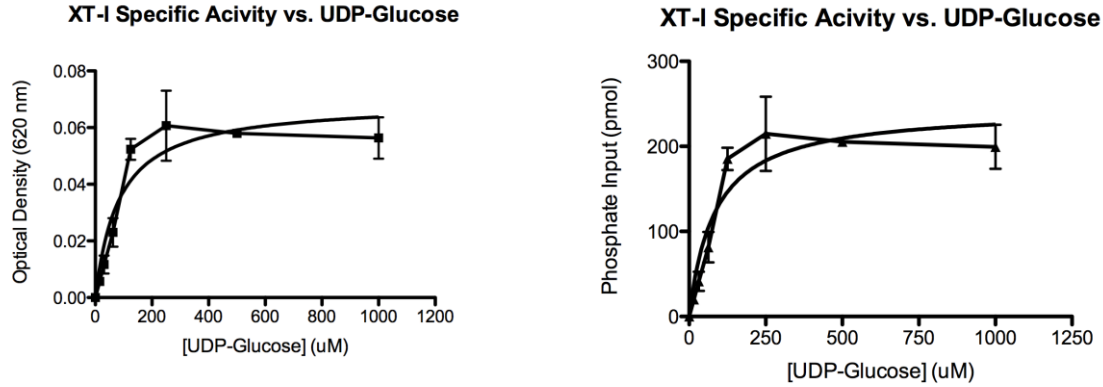


Figure S10. Phosphatase-coupled assay result of UDP-glucose. k_{cat} = 2 min^{-1}, K_m = 84.0 μM, k_{cat}/K_m = 20 mM^{-1} min^{-1}.

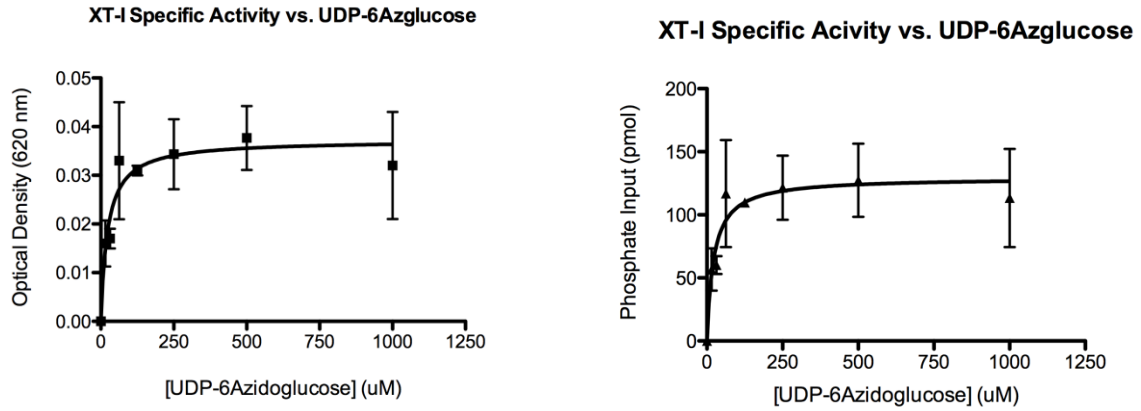


Figure S11. Phosphatase-coupled assay result of UDP-6-azido-glucose. k_{cat} = 1 min^{-1}, K_m = 23.4 μM, k_{cat}/K_m = 39 mM^{-1} min^{-1}.

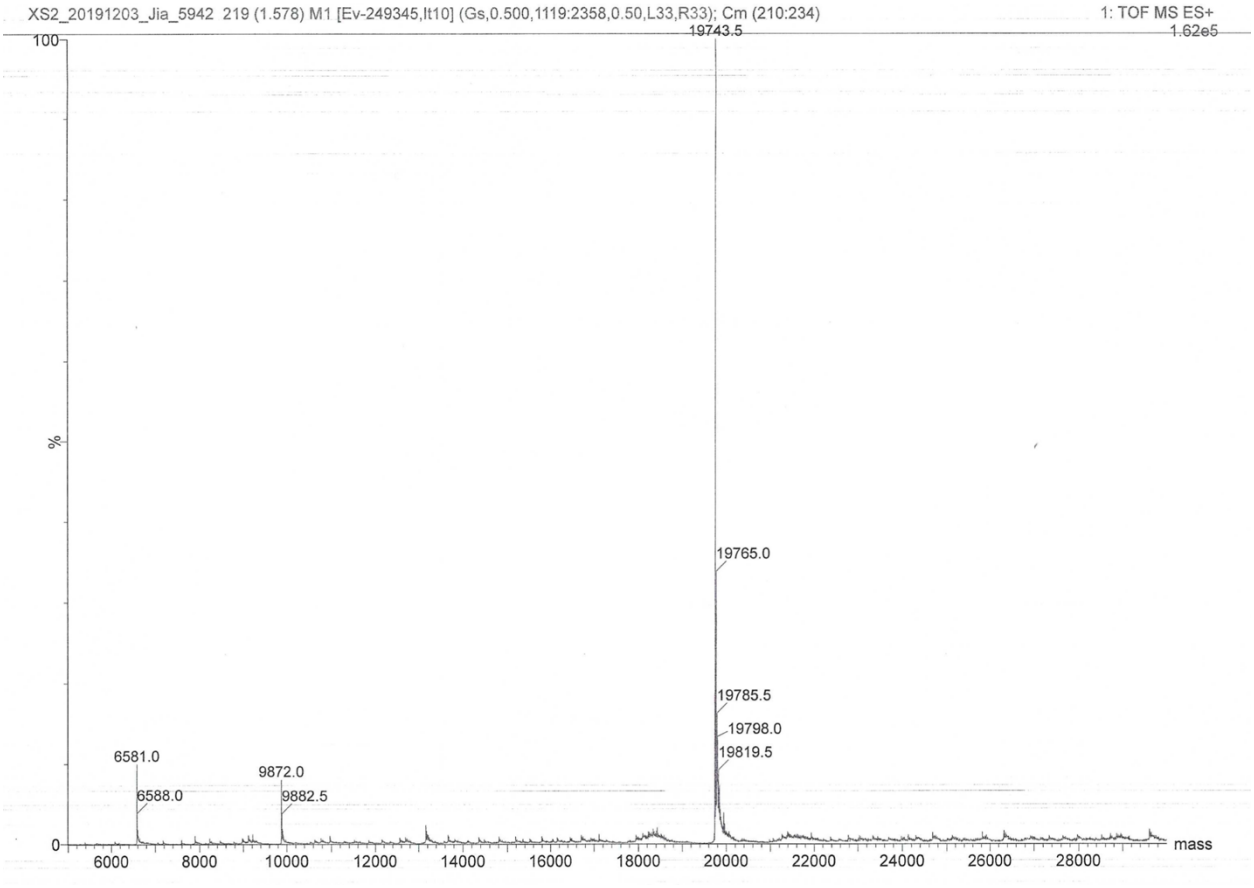


Figure S12. ESI-MS of recombinant polyhistagged $hCD44_{20\text{-}178}$.

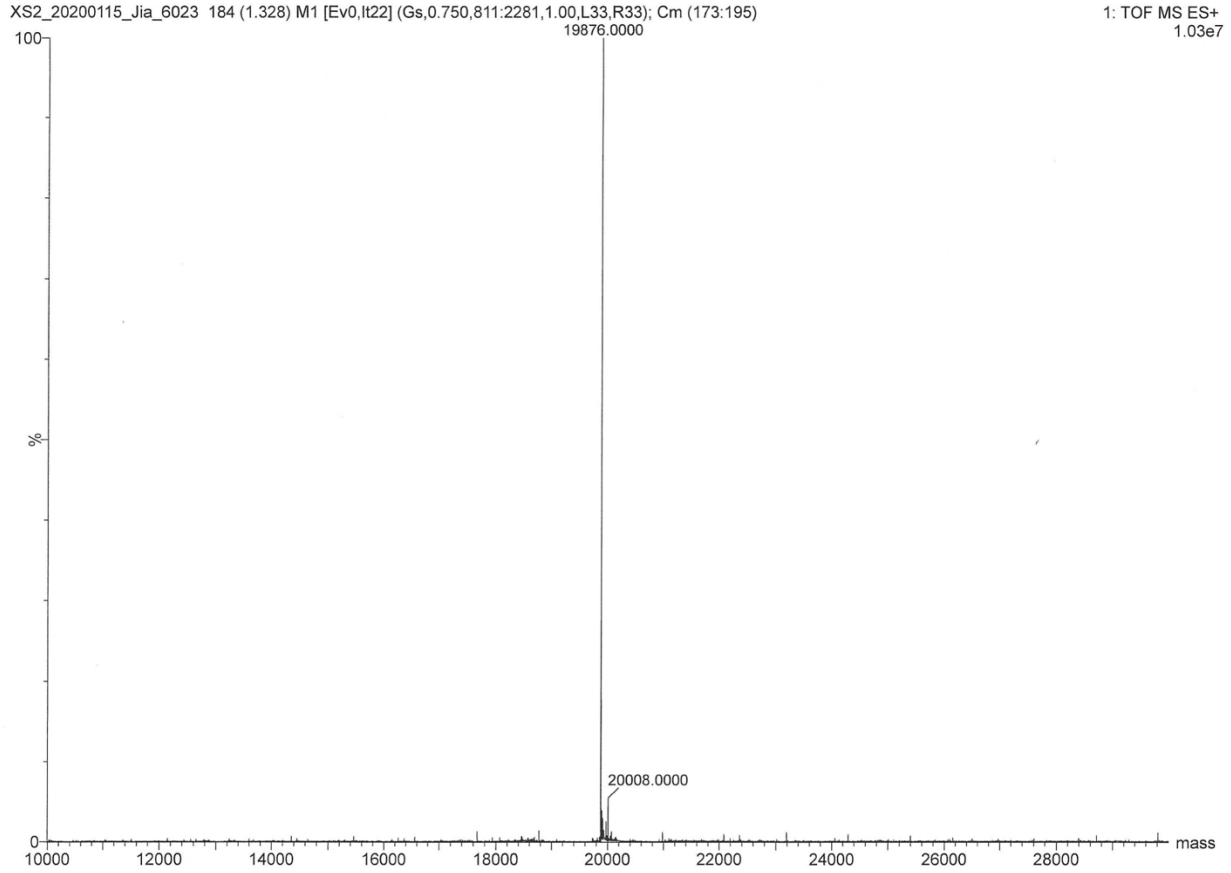


Figure S13. ESI-MS of recombinant polyhistagged $hCD44_{20\text{-}178}$ following the treatment with XT-1. The mass shift to 19,876 suggests the addition of one xylose unit to CD44 forming CD44 (O-Xyl).

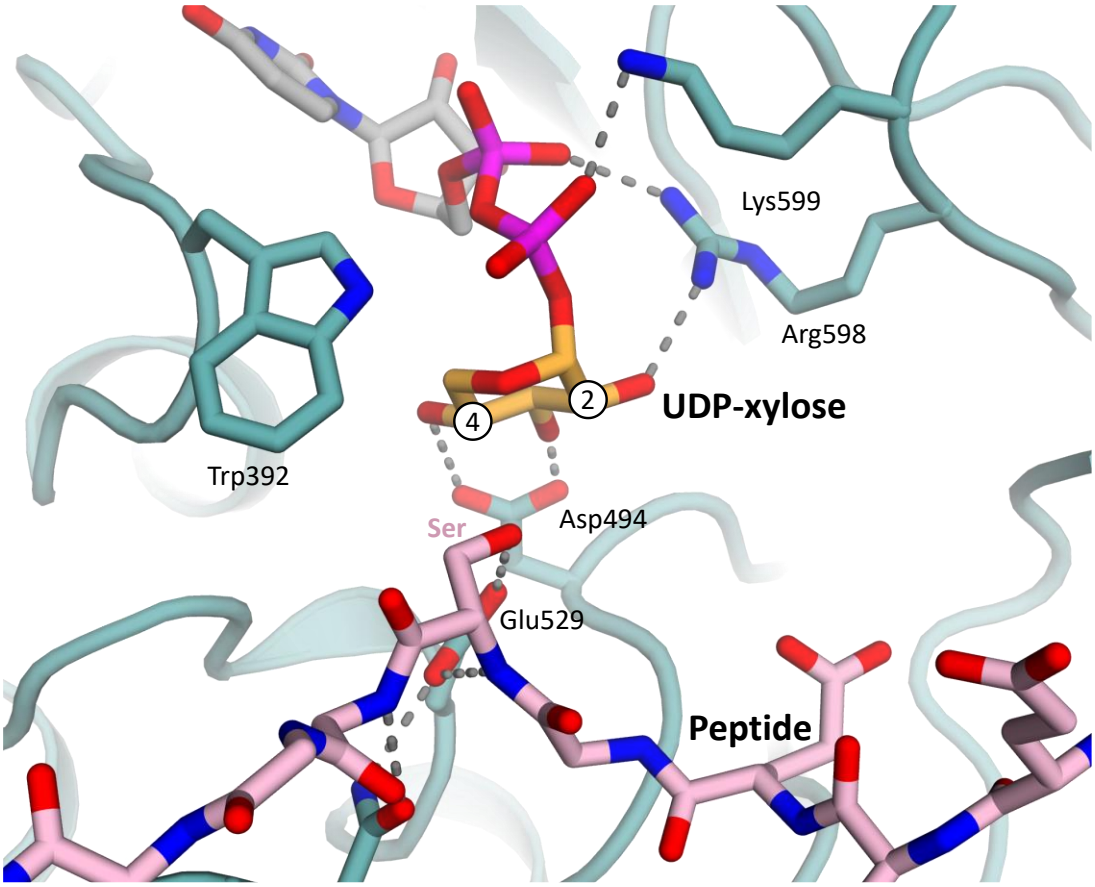


Figure S14. Structure of the active site of XT-I bound with UDP-Xyl and the peptide acceptor derived from the crystal structure (PDB code: 6EJ7). The 2-OH and 4-OH of UDP-Xyl have been labeled with the numbers 2 and 4 in circles. The key residues in the active site interacting with the UDP-Xyl have been highlighted. The structure 6EJ7 is a ternary complex of XT-I, UDP-Xyl and the acceptor peptide with a Ser-to-Ala mutation (to prevent Xyl transfer occurring in the crystal). To generate this figure, the serine was inserted back into the peptide acceptor to demonstrate the geometry of the acceptor complex.

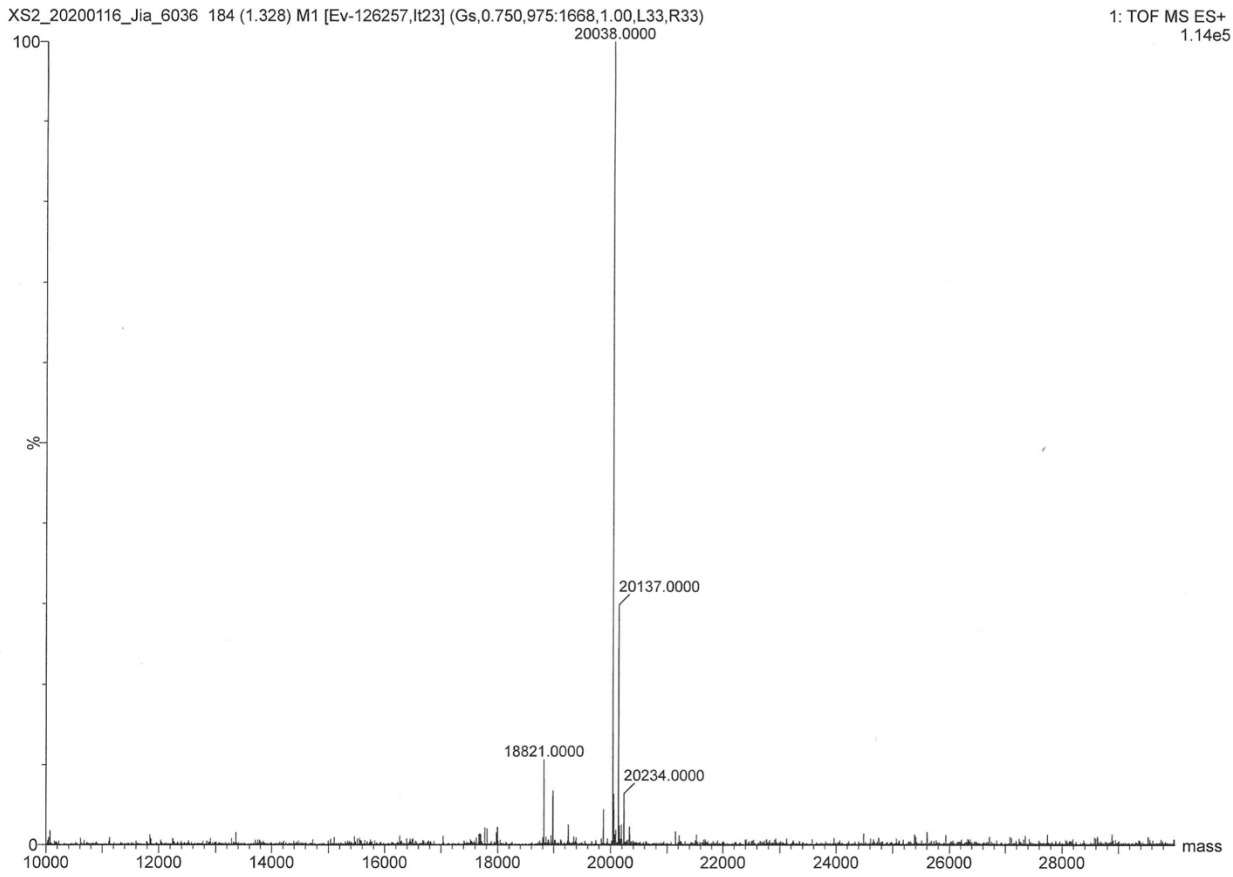
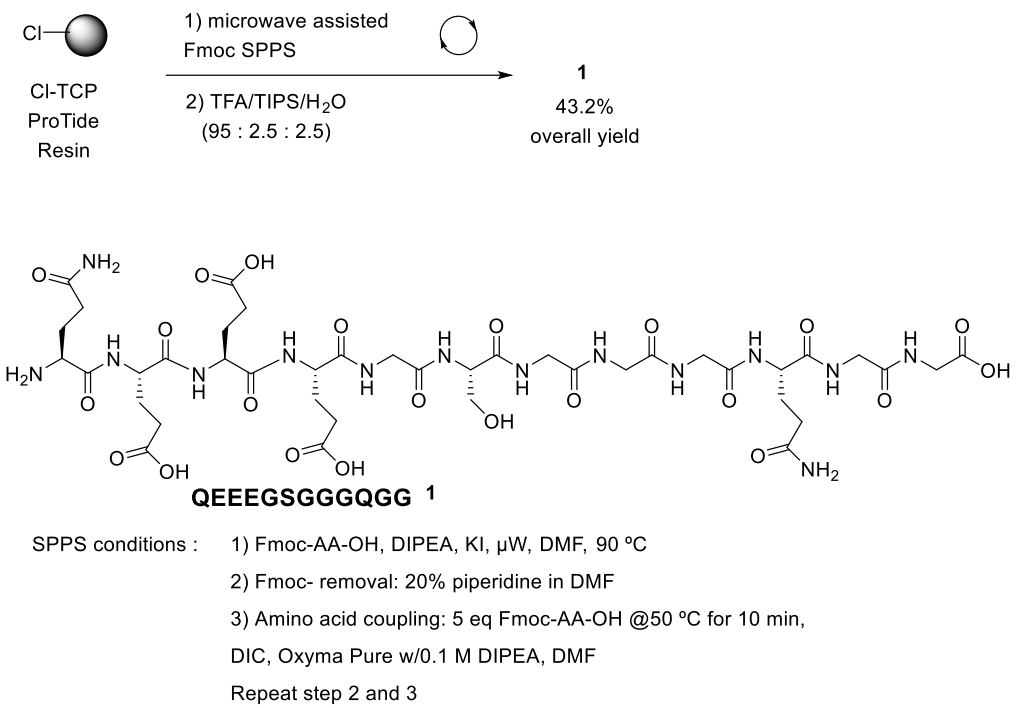


Figure S15. ESI-MS of recombinant polyhistagged $hCD44_{20\text{-}178}$ following the treatment with XT-1 and β4GalT7. The mass shift to 20,038 suggests the addition of one xylose unit and one galactose unit to the CD44 forming CD44 (O-Xyl-Gal).

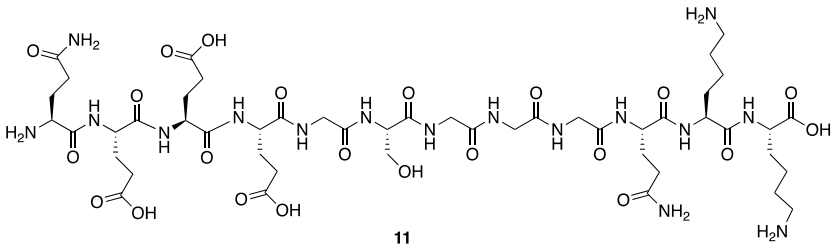


Scheme S1. General procedure for SPPS synthesis of bikunin peptide (QEEEGSGGGQGG) **1**.

Table S1. Summary of synthesized peptide acceptors and the corresponding yields.

Sequence	**SPPS Yield (%)**
QEEEGSGGGQGG **1**	43.2
GGPSGDFE **3**	47.7
DNFSGSGAG **4**	61.7
DFELSGSGDLD **5**	38.2
DLYSGSGSGYFE **6**	33.1
QEEEGSGGGQKK **11**	47.8

Scheme S2. XT-I catalyzed transfer of non-native 6AzGlc to bikunin peptide **11**, followed by incorporation of sulfo-Cy5 fluorescent dye *via* CuAAC reaction.

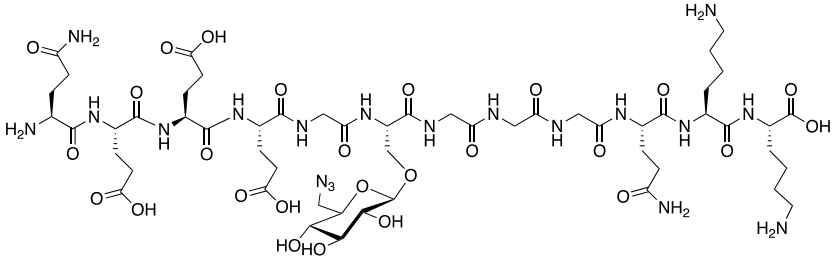


11

ESI-MS: $C_{48}H_{82}N_{16}O_{22}^{2+}$,
calcd: 617.2892, obsd: 617.2892

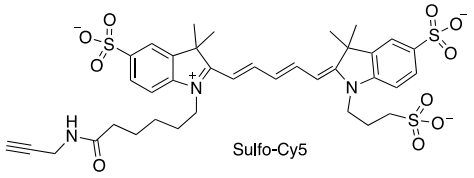
UDP-6AzGlc, XT-I

MES Buffer, $MnCl_2$, pH 6.5, 37 °C



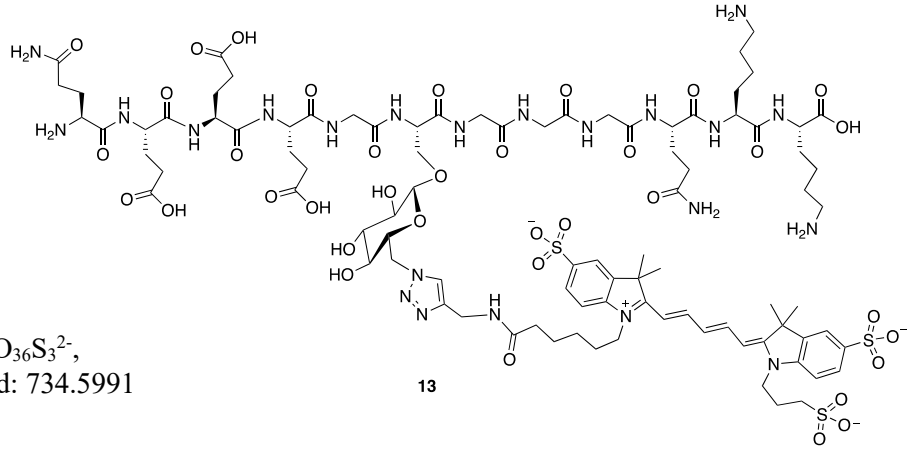
12

ESI-MS: $C_{54}H_{87}N_{19}O_{26}^{2-}$,
calcd: 708.8038, obsd: 708.7978



Sulfo-Cy5

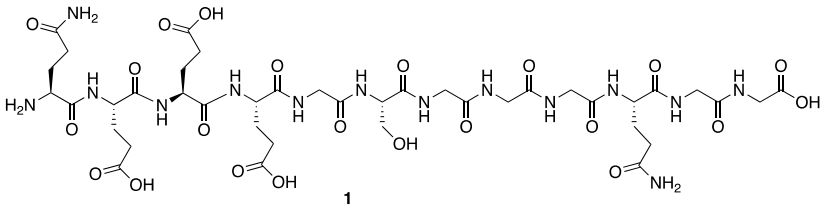
$CuSO_4$, THPTA, Aminoguanidine, Na Ascorbate



13

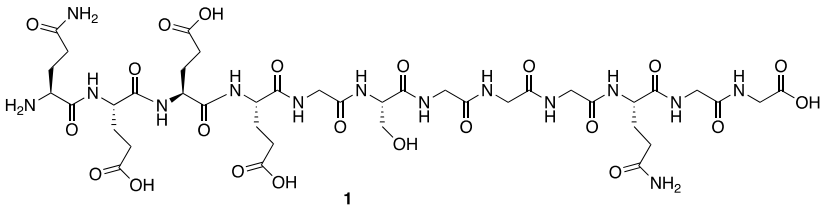
ESI-MS: $C_{91}H_{131}N_{22}O_{36}S_3^{2-}$,
calcd: 734.6092, obsd: 734.5991

Product Characterization



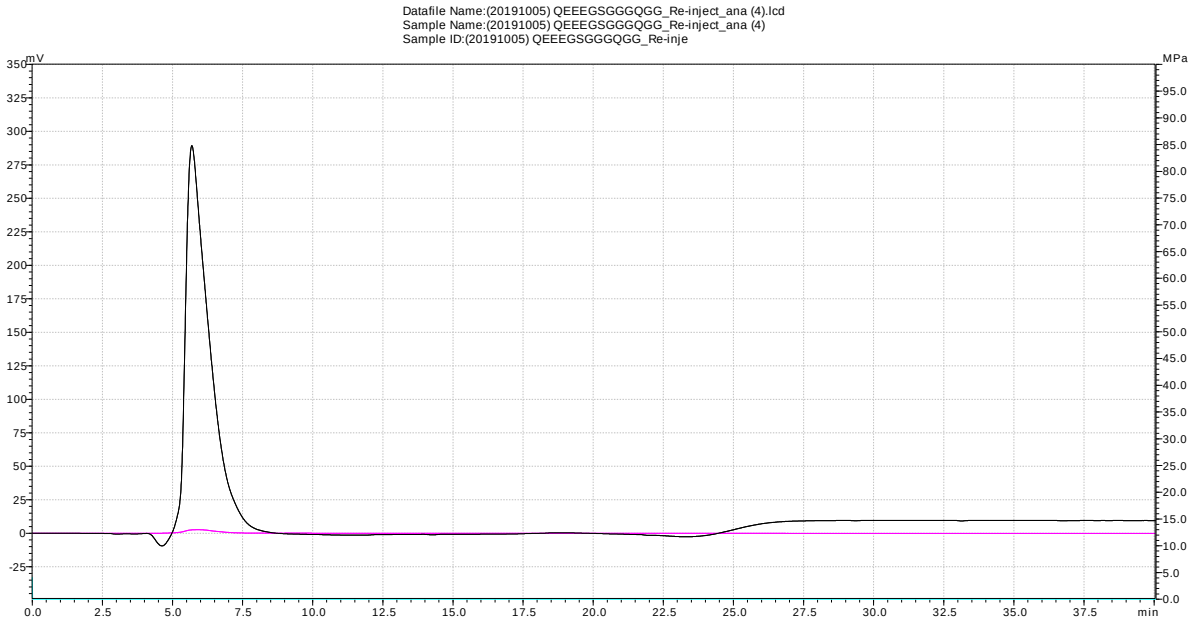
1

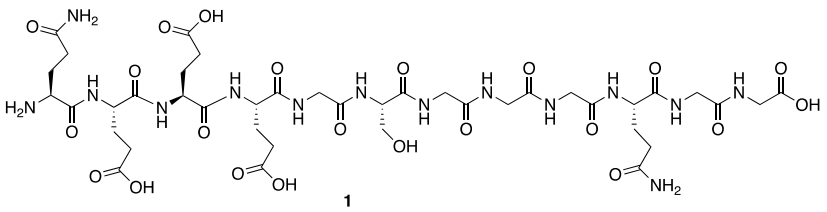
The purity of peptide **1** was verified with analytical C-18 HPLC (water, 0.1% trifluoroacetic acid). $[\alpha]_D^{20}$= +28 (c 0.1, H_2O). ^{1}H-NMR (500 MHz, D_2O) δ 1.45 – 1.52 (5 H, m), 1.62-1.58 (10 H, m), 1.76 – 1.90 (2 H, m), 1.91 – 2.06 (2 H, m), 2.14 – 2.47 (5 H, m), 2.93 – 3.05 (9 H, m), 3.65 – 3.99 (5 H, m), 4.15 – 4.37 (2 H, m); ^{13}C-NMR (125 MHz, D_2O) 14.4, 16.9, 17.7, 19.7, 20.4, 21.1, 21.2, 21.7, 22.0, 22.0, 23.2, 24.4, 24.5, 25.2, 26.0, 26.2, 26.4, 27.2, 27.4, 28.7, 29.7, 29.9, 29.9, 30.7, 31.5, 40.0, 41.3, 42.1, 42.1, 42.1, 42.3, 42.3, 42.6, 43.3, 44.3, 44.3, 44.6, 45.3, 46.6, 52.1, 52.9, 52.9, 53.1, 55.6, 55.7, 56.7, 60.9, 60.9. ESI-MS: $C_{40}H_{63}N_{14}O_{22}$ $[M+H]^+$ calcd: 1091.4236, obsd: 1091.4216 (1.8 ppm).



1

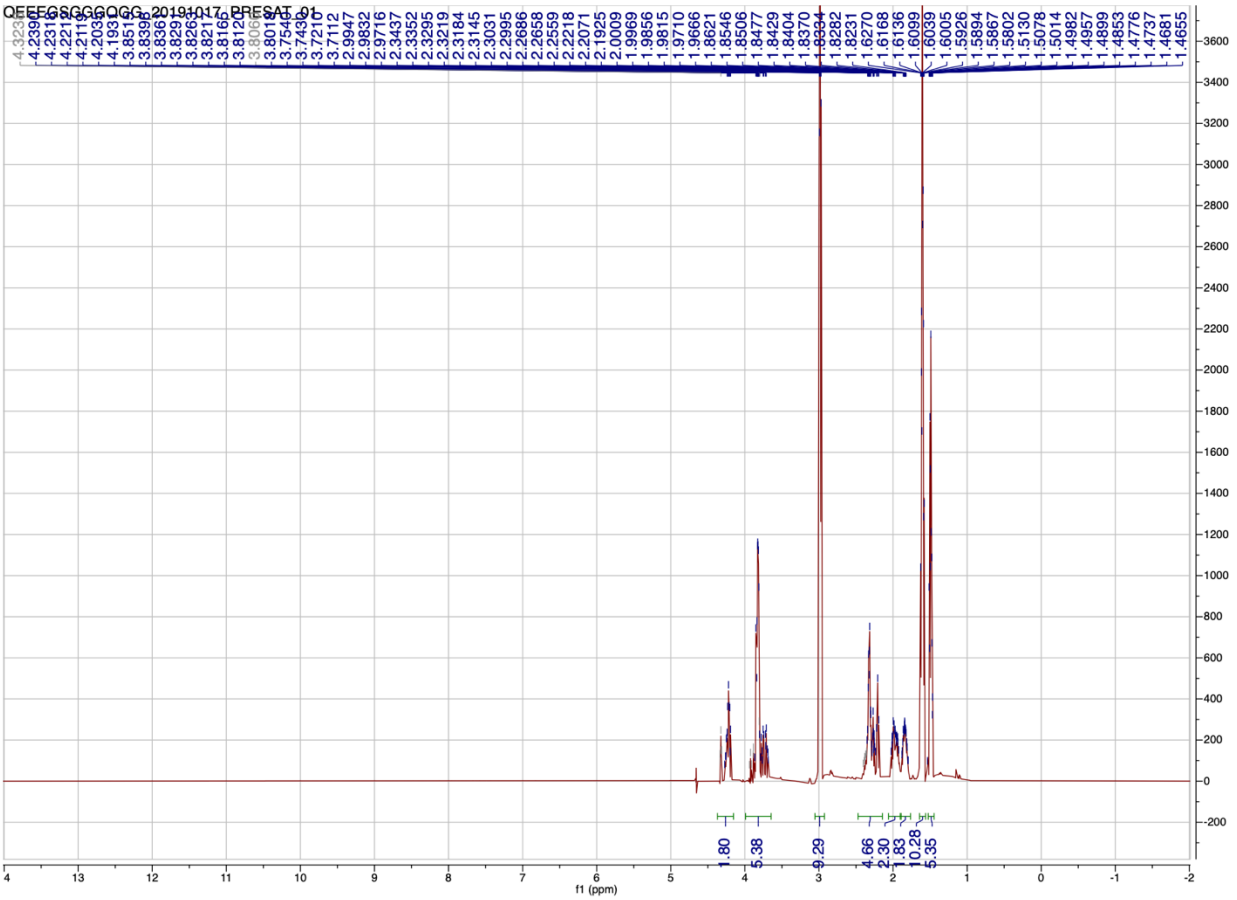
HPLC

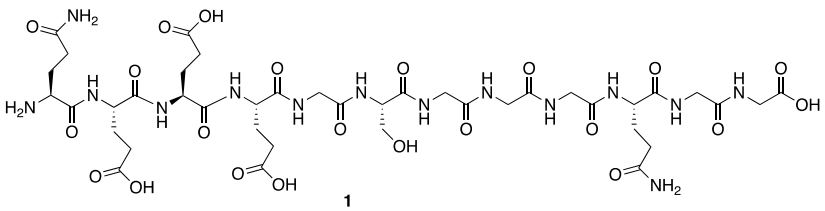

Datafile Name:(20191005) QEEEGSGGGQGG_Re-inject_ana (4).lcd
Sample Name:(20191005) QEEEGSGGGQGG_Re-inject_ana (4)
Sample ID:(20191005) QEEEGSGGGQGG_Re-inje
mV
MPa
min



1

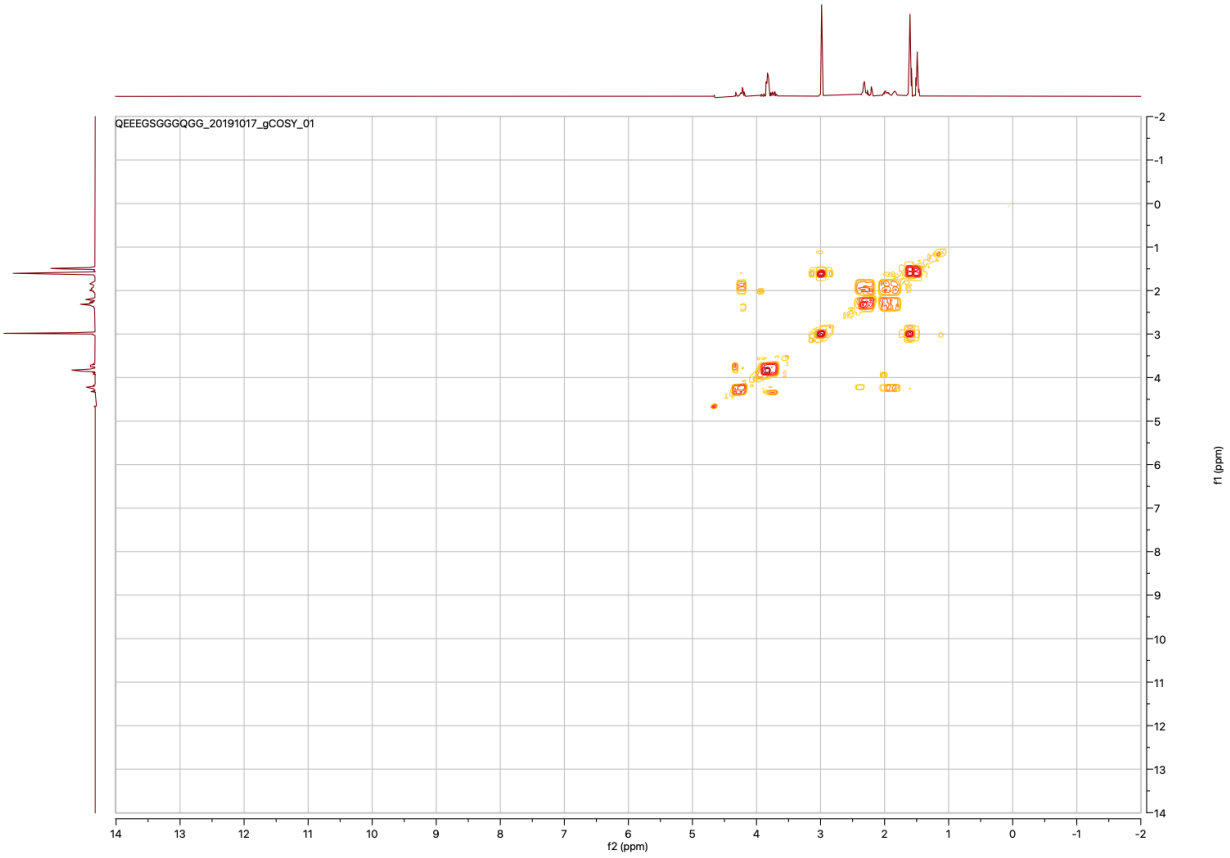
^{1}H-NMR (500 MHz, D_2O)

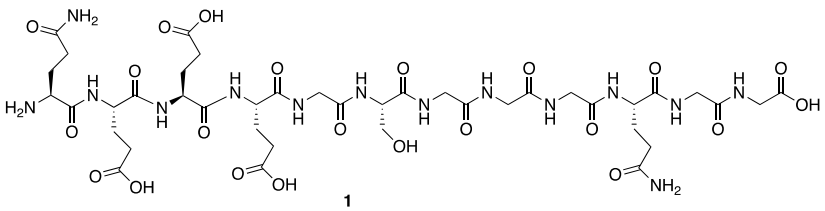




1

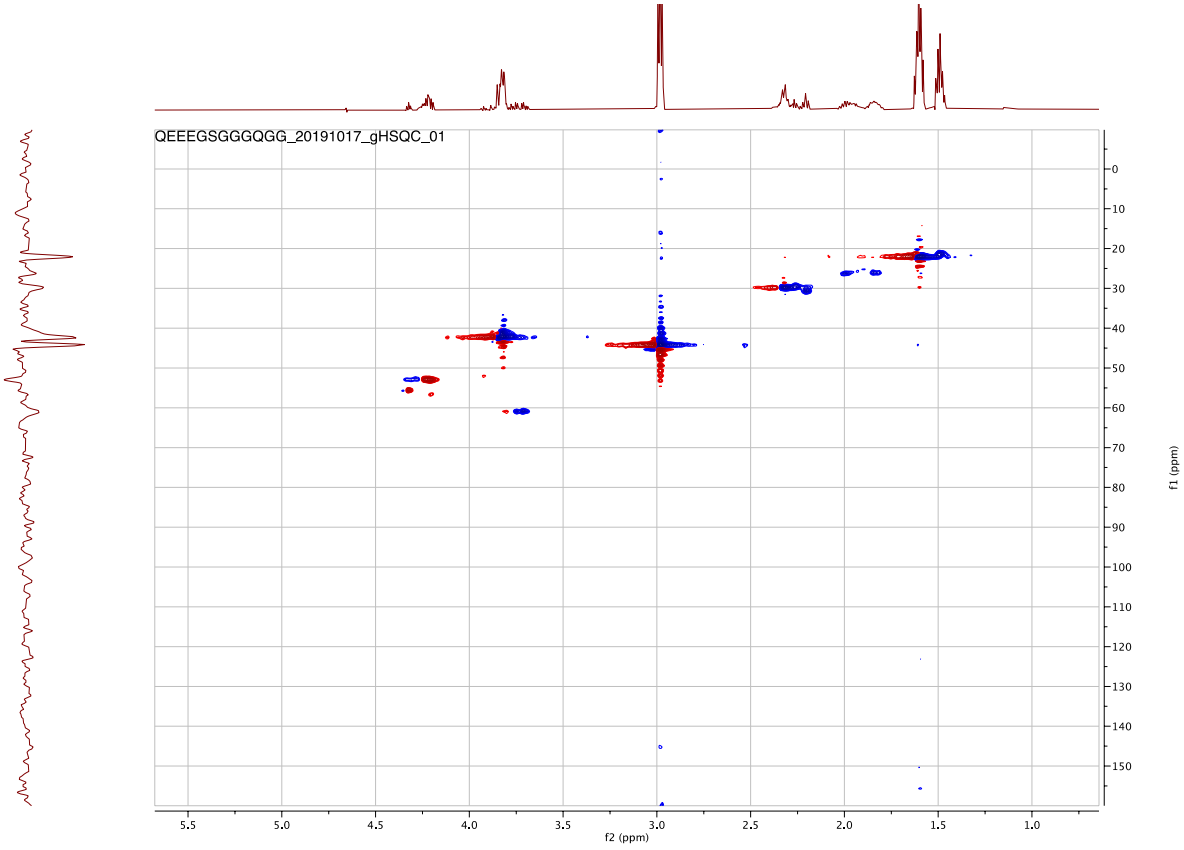
COSY (500 MHz, D_2O)

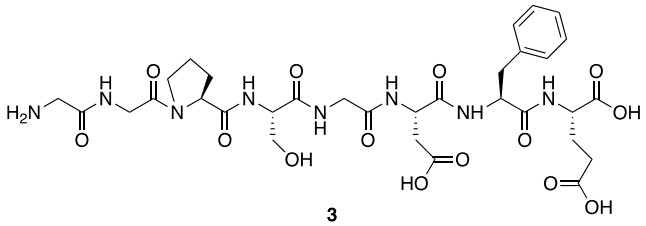




1

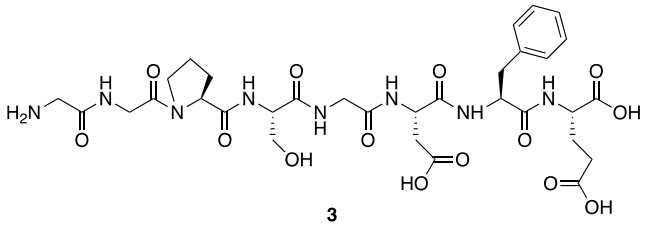
HSQC (500 MHz, D_2O)





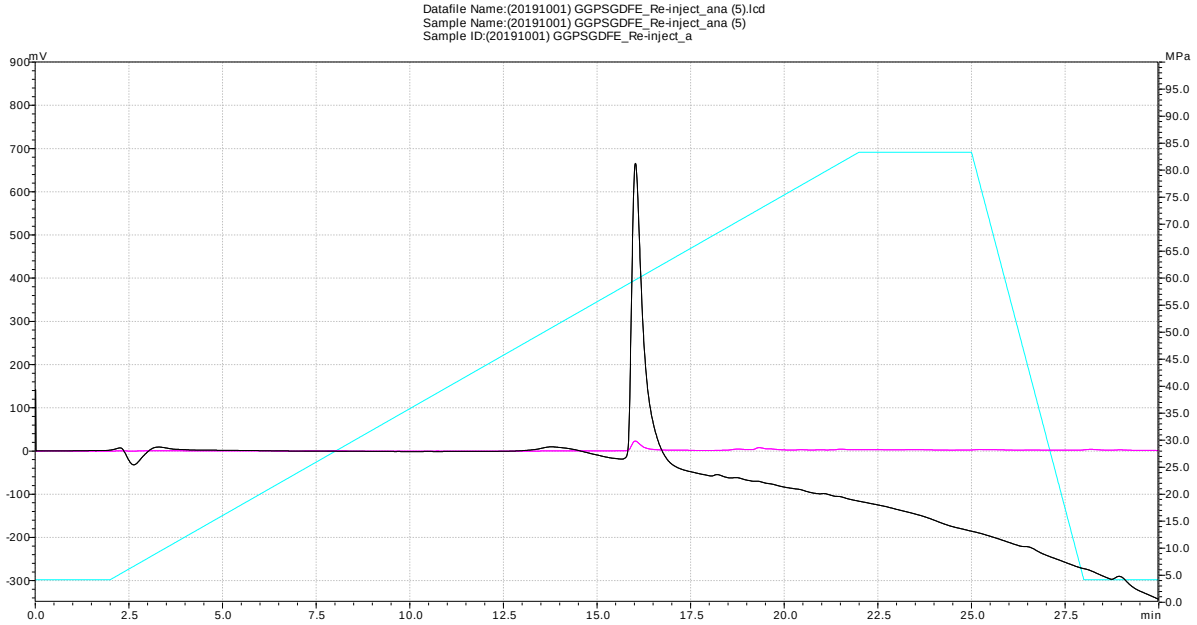
3

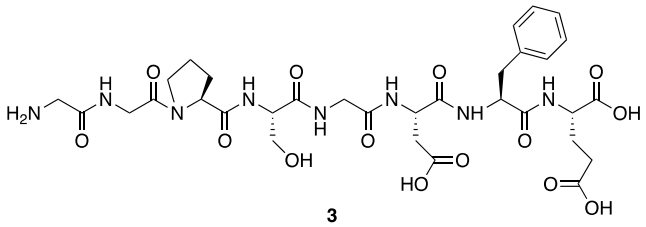
The purity of peptide **3** was verified with analytical C-18 HPLC (5-100% acetonitrile/water, 0.1% trifluoroacetic acid). $[\alpha]_D^{20}$= -111 (c 0.1, H_2O). ^1H NMR (500 MHz, D_2O) δ 1.72 – 1.90 (4 H, m), 1.94 – 2.05 (1 H, m), 2.11-2.15 (1 H, m), 2.22-2.25 (2 H, t, *J* = 7.3 Hz), 2.54-2.58 (1 H, m), 2.63 – 2.72 (1 H, m), 2.83 – 2.92 (1 H, m), 2.99-3.03 (1 H, m), 3.42 – 3.53 (2 H, m), 3.64 – 3.85 (6 H, m), 3.92 – 4.09 (2 H, m), 4.19-4.23 (1 H, m), 4.26 – 4.34 (2 H, m), 4.44-4.48 (1 H, m), 4.51-4.55 (1 H, m), 7.07 – 7.11 (2 H, m), 7.11 – 7.17 (1 H, m), 7.19-7.21 (2 H, m); ^{13}C NMR (125 MHz, D_2O) δ 24.3, 25.7, 29.3, 29.7, 35.0, 36.6, 40.2, 40.5, 41.6, 42.4, 47.0, 49.8, 51.9, 54.8, 55.6, 60.5, 60.9, 127.0, 128.6, 129.1, 136.1, 167.4, 169.1, 170.8, 171.7, 172.2, 172.5, 173.8, 174.5, 174.5, 177.0. ESI-MS: $C_{32}H_{45}N_8O_{14}$ $[M+H]^+$ calcd: 765.3050, obsd: 755.3022 (3.7 ppm).



3

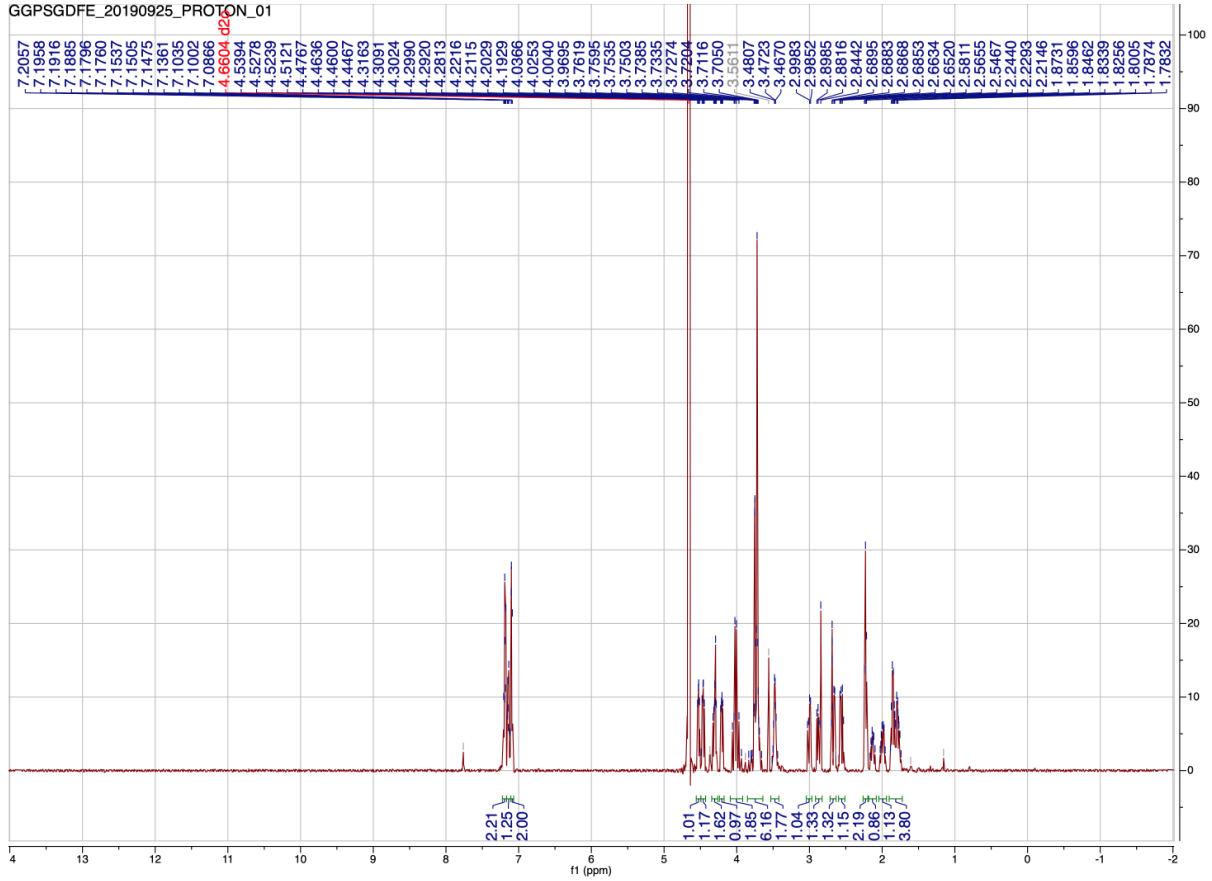
HPLC

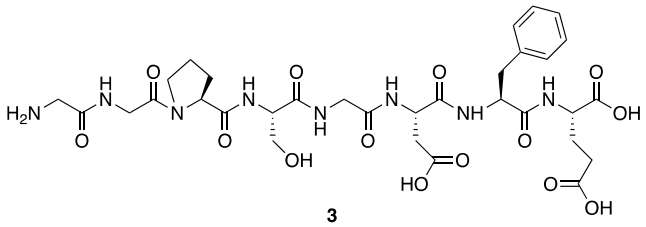

Datafile Name:(20191001) GGPSGDFE_Re-inject_ana (5).lcd
Sample Name:(20191001) GGPSGDFE_Re-inject_ana (5)
Sample ID:(20191001) GGPSGDFE_Re-inject_a
mV
MPa
min



3

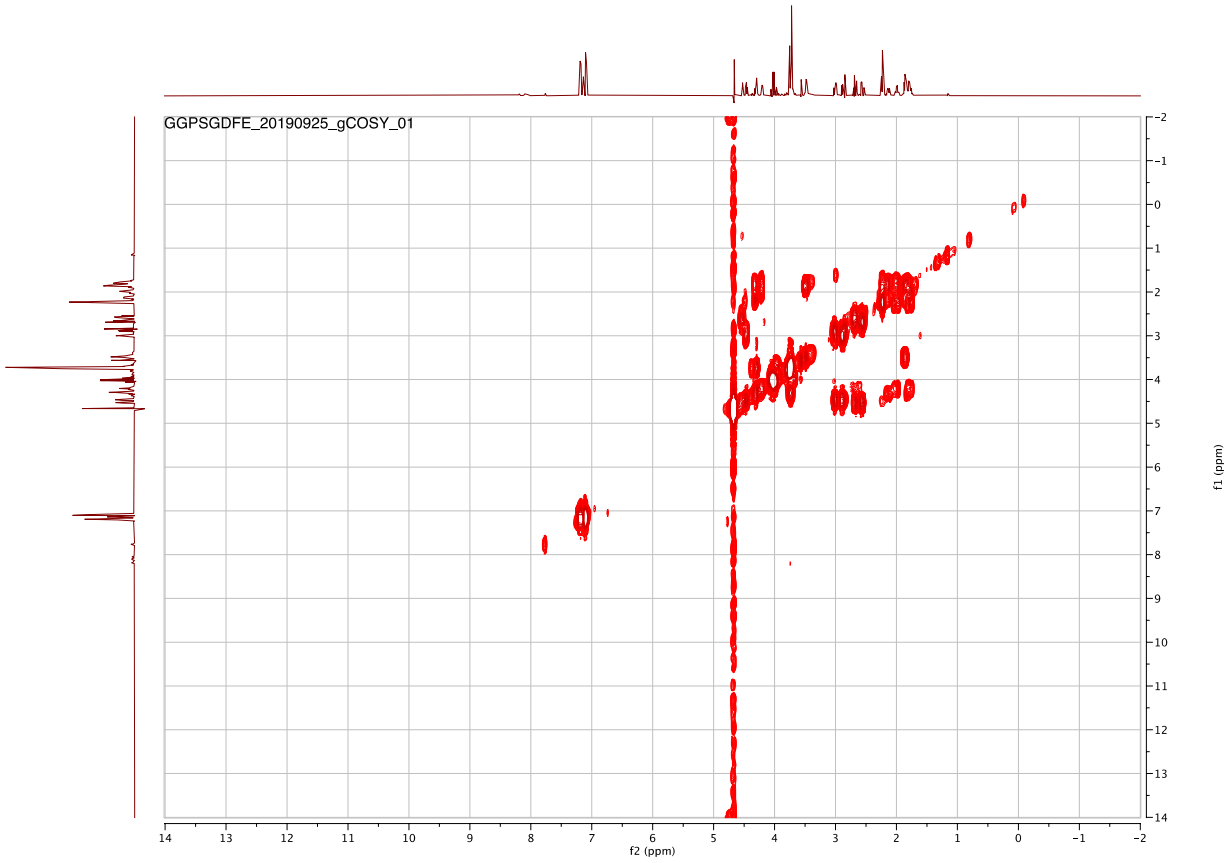
^{1}H-NMR (500 MHz, D_2O)

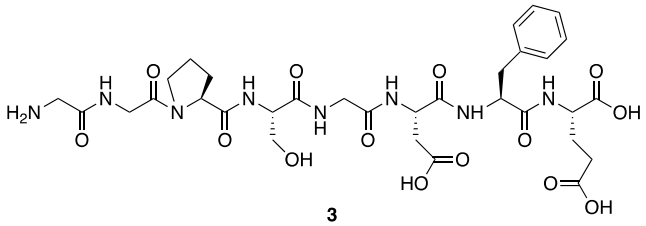




3

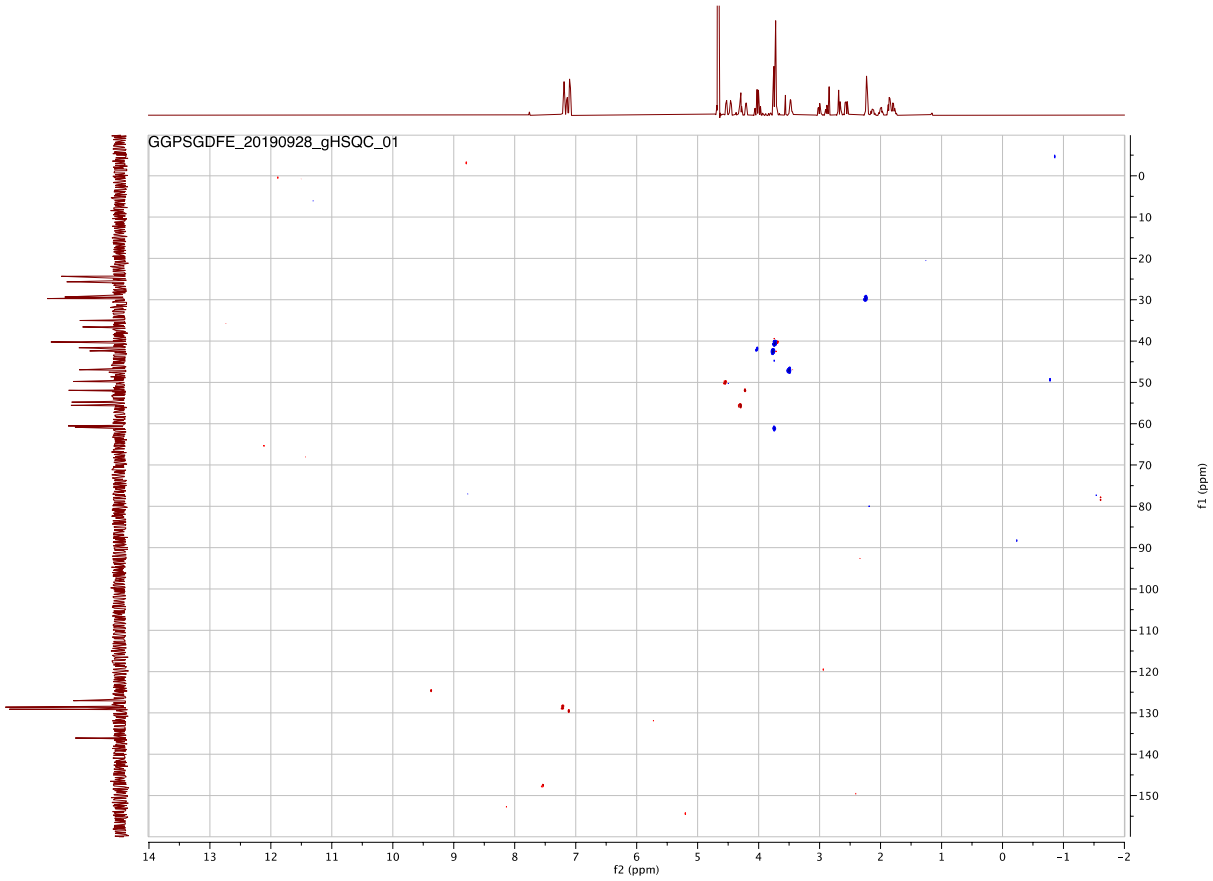
COSY (500 MHz, D_2O)

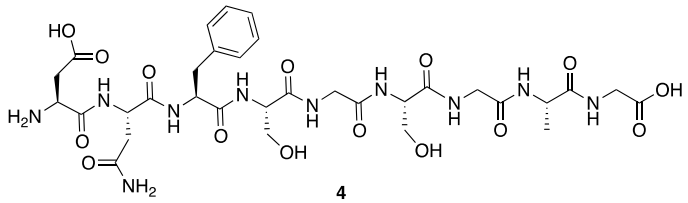




3

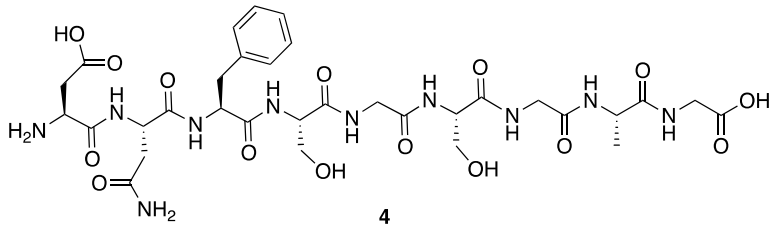
HSQC (500 MHz, D_2O)





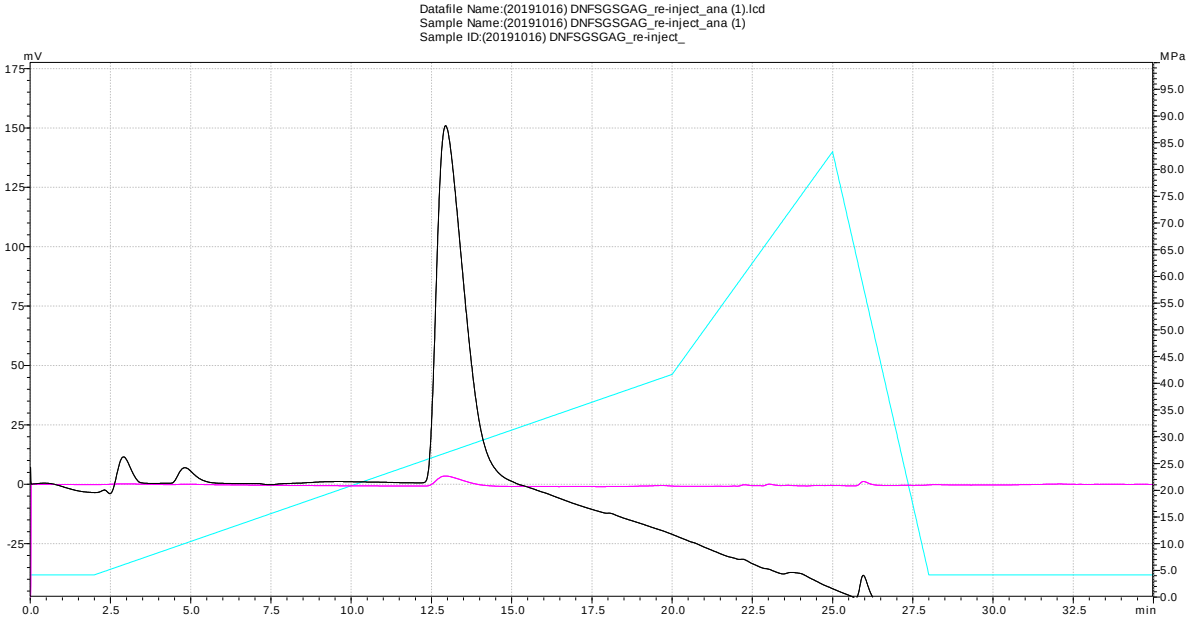
4

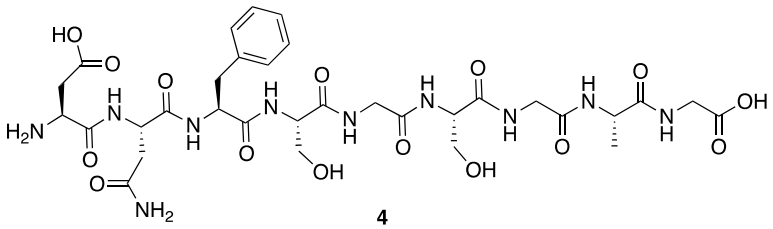
The purity of peptide **4** was verified with analytical C-18 HPLC (5-100% acetonitrile/water, 0.1% trifluoroacetic acid). $[\alpha]_D^{20}$= -329 (c 0.1, H_2O). ^{1}H NMR (500 MHz, D_2O) δ 1.22 (3 H, d, *J* = 7.2 Hz), 2.50-2.52 (1 H, m), 2.60-2.62 (1 H, m), 2.71 – 2.78 (2 H, m), 2.88-2.92 (1 H, m), 3.01-3.03 (1 H, m), 3.62 – 3.74 (4 H, m), 3.75 (1 H, d, *J* = 6.1 Hz), 3.78 (3 H, d, *J* = 12.8 Hz), 3.80 – 3.82 (2 H, m), 3.84-3.86 (1 H, m), 4.13 (1 H, t, *J* = 6.3 Hz), 4.20 (1 H, q, *J* = 7.2 Hz), 4.27 (1 H, t, *J* = 5.2 Hz), 4.32 (1 H, t, *J* = 5.1 Hz), 4.56-4.52 (2 H, m), 7.08 – 7.24 (5 H, m). ^{13}C NMR (125 MHz, D_2O) δ 16.5, 35.1, 35.9, 35.9, 36.7, 36.7, 41.2, 41.9, 41.9, 42.3, 42.3, 49.4, 49.6, 50.3, 55.0, 55.4, 55.6, 60.9, 60.9, 60.9, 127.1, 128.7, 129.1, 129.1. ESI-MS: $C_{32}H_{47}N_{10}O_{15}$ $[M+H]^+$ calcd: 811.3217, obsd: 811.3207 (1.2 ppm).



4

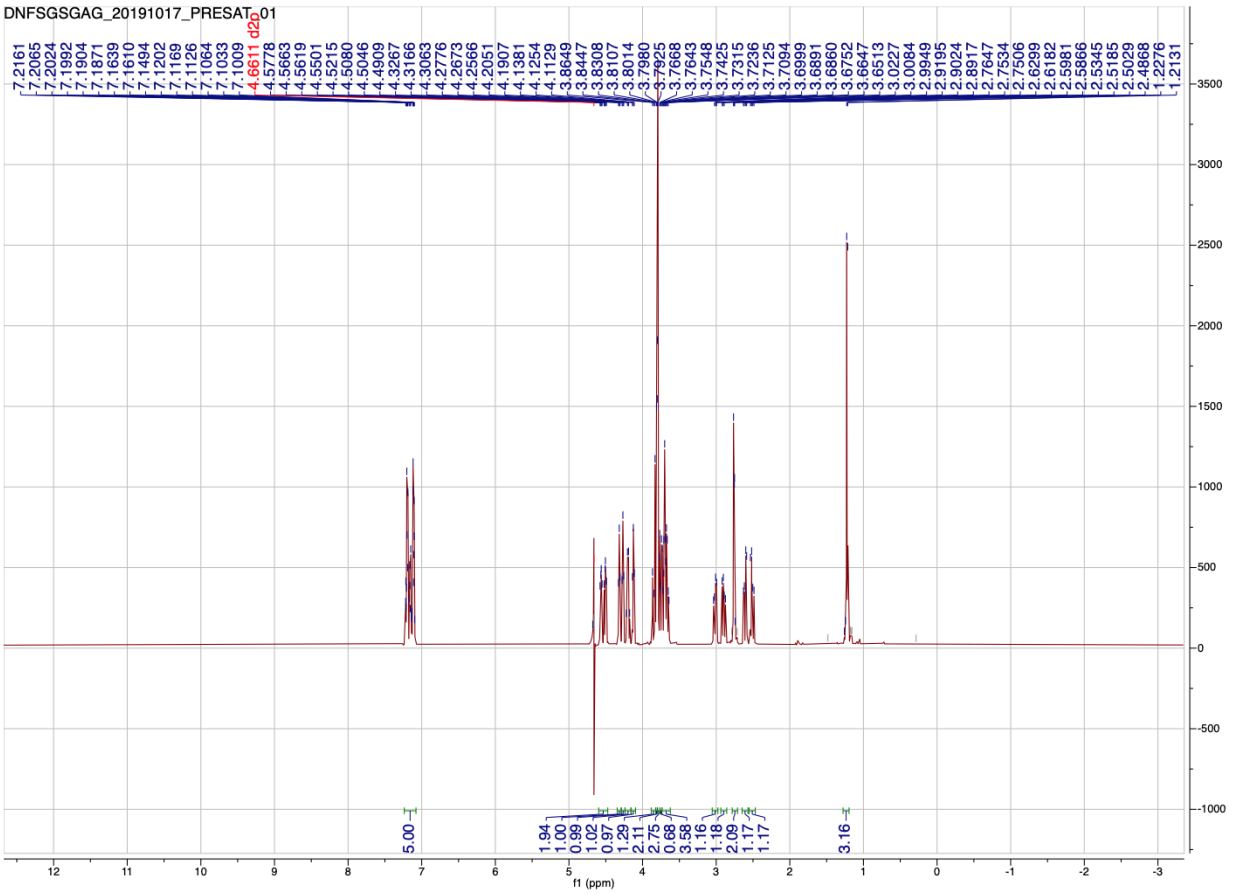
HPLC

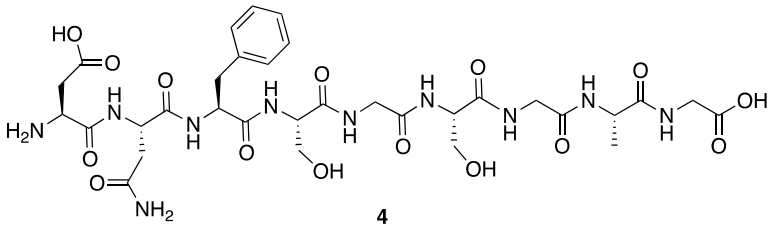

Datafile Name:(20191016) DNFSGSGAG_re-inject_ana (1).lcd
Sample Name:(20191016) DNFSGSGAG_re-inject_ana (1)
Sample ID:(20191016) DNFSGSGAG_re-inject_
mV
MPa
min



4

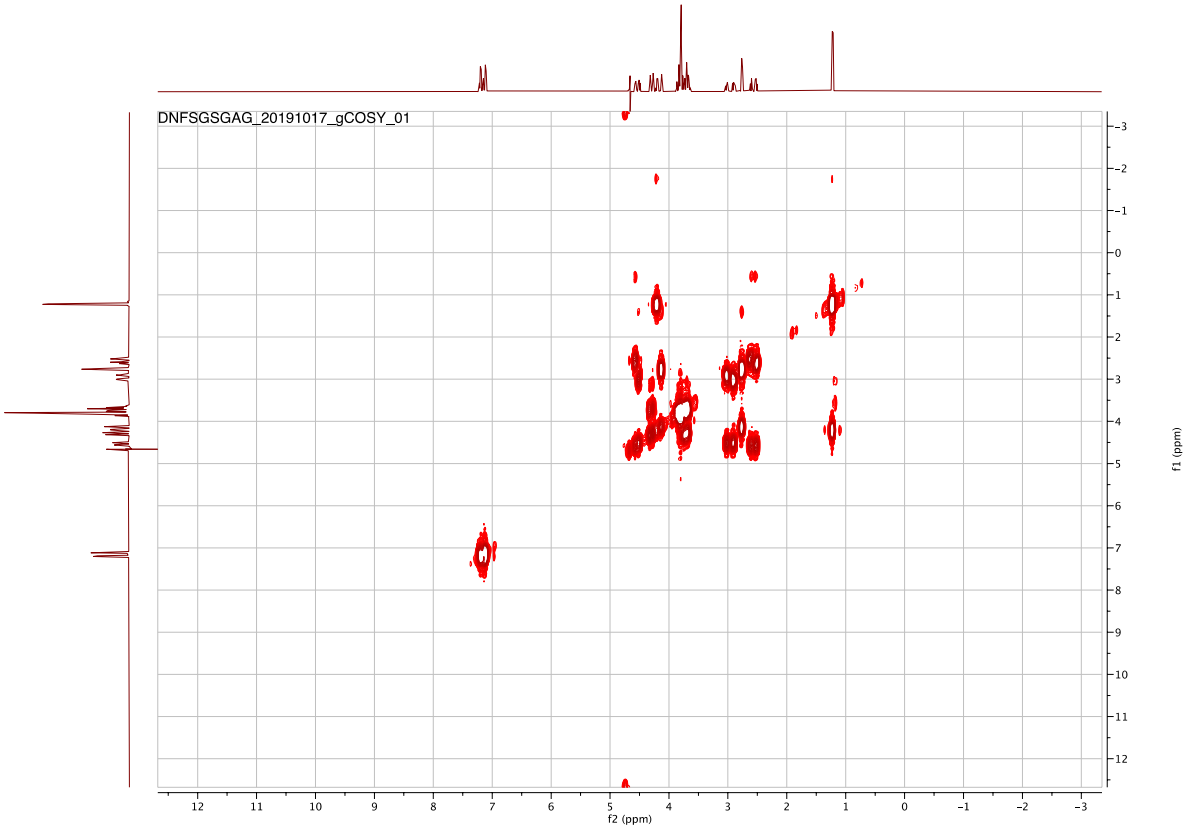
^{1}H-NMR (500 MHz, D_2O)

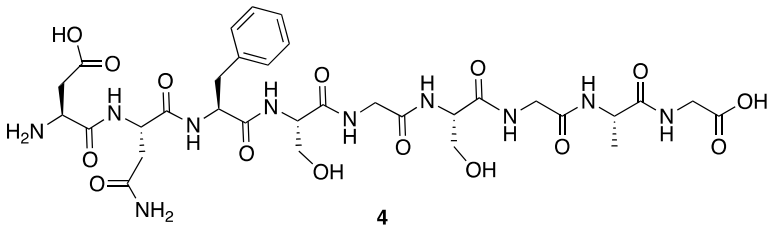




4

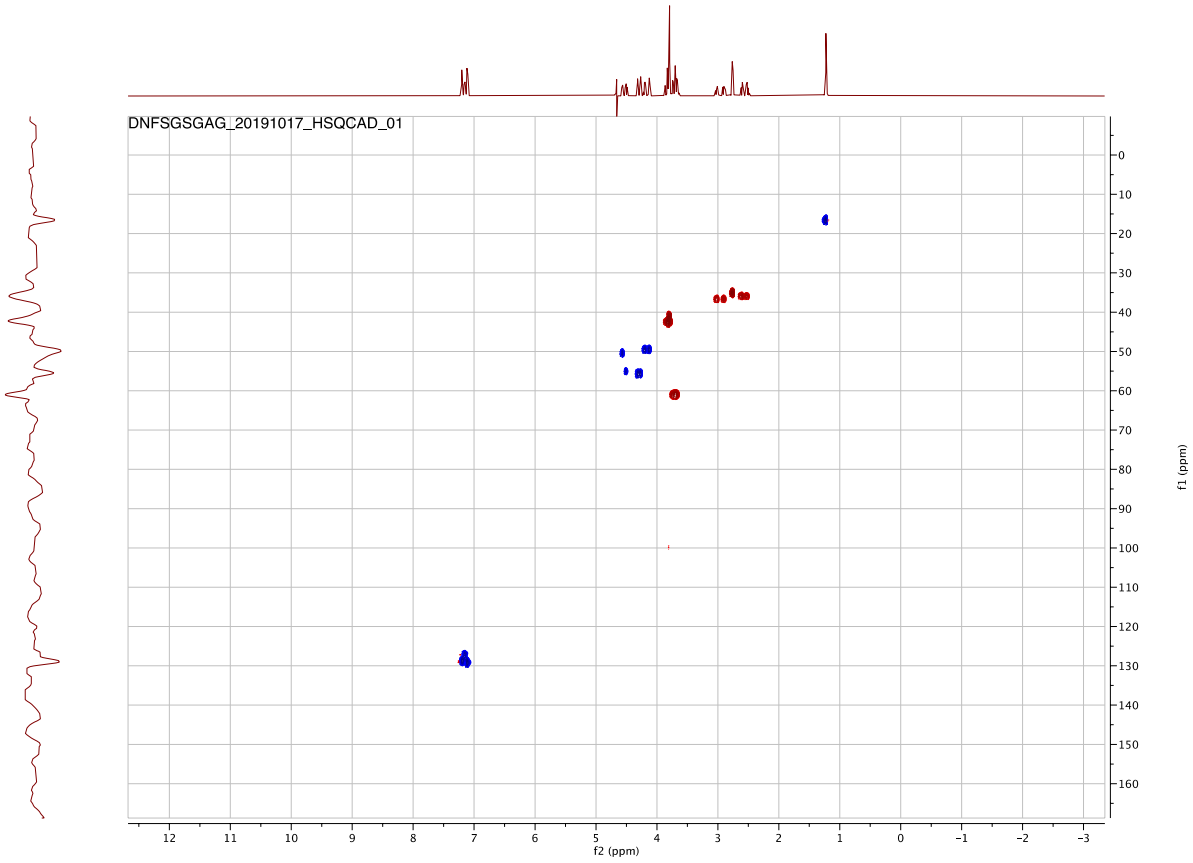
COSY (500MHz, D_2O)

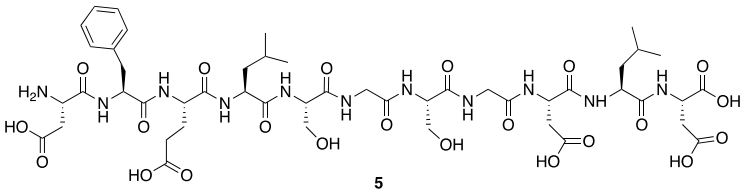




4

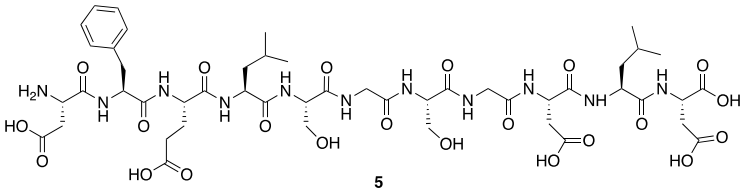
HSQC (500MHz, D_2O)





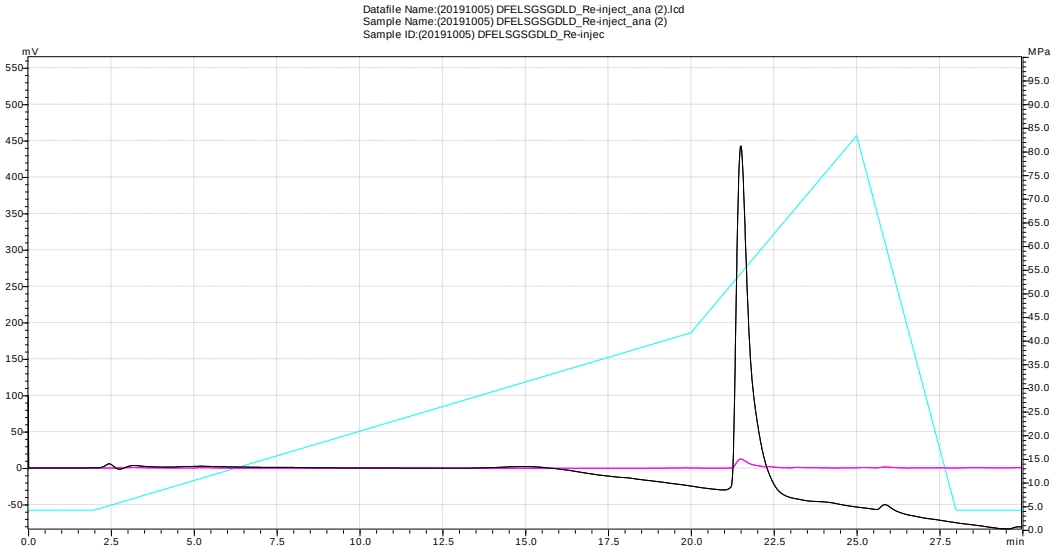
5

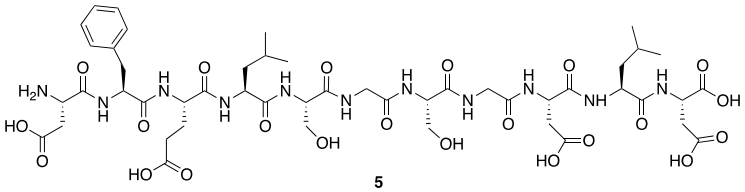
The purity of peptide **5** was verified with analytical C-18 HPLC (5-100% acetonitrile/water, 0.1% trifluoroacetic acid). $[\alpha]_D^{20} = +76$ (c 0.1, H_2O). ^1H NMR (500 MHz, D_2O) δ 0.67 – 0.72 (4 H, m), 0.74 – 0.85 (10 H, m), 1.16 (1 H, d, J = 1.2 Hz), 1.42 – 1.53 (6 H, m), 1.74 (1 H, m, 1.81 – 1.92 (1 H, m), 2.19 – 2.26 (2 H, m), 2.61 – 2.70 (1 H, m), 2.70 – 2.86 (5 H, m), 2.92 (2 H, t, J = 8.5 Hz), 3.66 – 3.77 (4 H, m), 3.77 – 3.92 (4 H, m), 4.08 – 4.24 (3 H, m), 4.26 – 4.32 (1 H, m), 4.45 – 4.52 (1 H, m), 4.55 (1 H, m), 7.06 – 7.12 (2 H, m), 7.12 – 7.18 (1 H, m), 7.18 – 7.24 (2 H, m). ^{13}C NMR (125 MHz, D_2O) δ 7.2, 20.5, 21.0, 21.6, 22.0, 29.6, 29.8, 30.2, 31.8, 34.5, 35.5, 36.5, 39.2, 42.3, 42.5, 42.5, 49.4, 49.8, 52.3, 55.6, 60.5, 60.7, 65.6, 80.0, 92.7, 121.6, 128.7, 129.1, 139.8. ESI-MS: $C_{48}H_{72}N_{11}O_{22}$ $[M+H]^+$ calcd: 1154.4848, obsd: 1154.4822 (2.3 ppm).



5

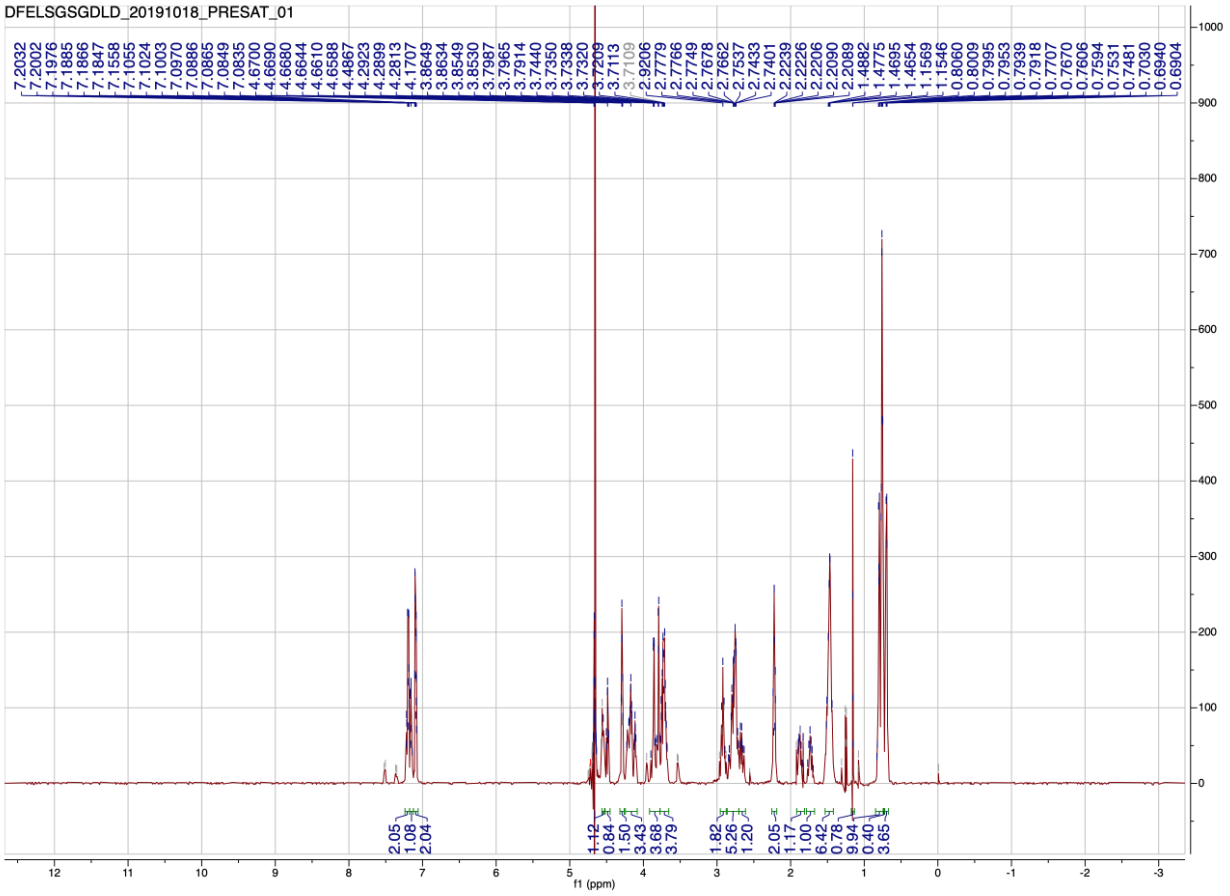
HPLC

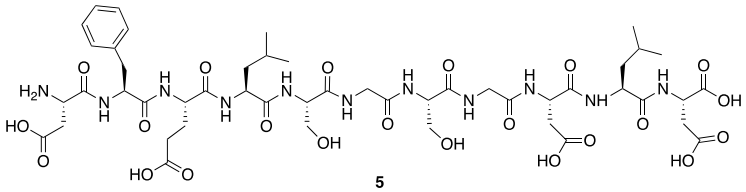

Datafile Name:(20191005) DFELSGSGDLD_Re-inject_ana (2).lcd
Sample Name:(20191005) DFELSGSGDLD_Re-inject_ana (2)
Sample ID:(20191005) DFELSGSGDLD_Re-injec
mV
MPa
min



5

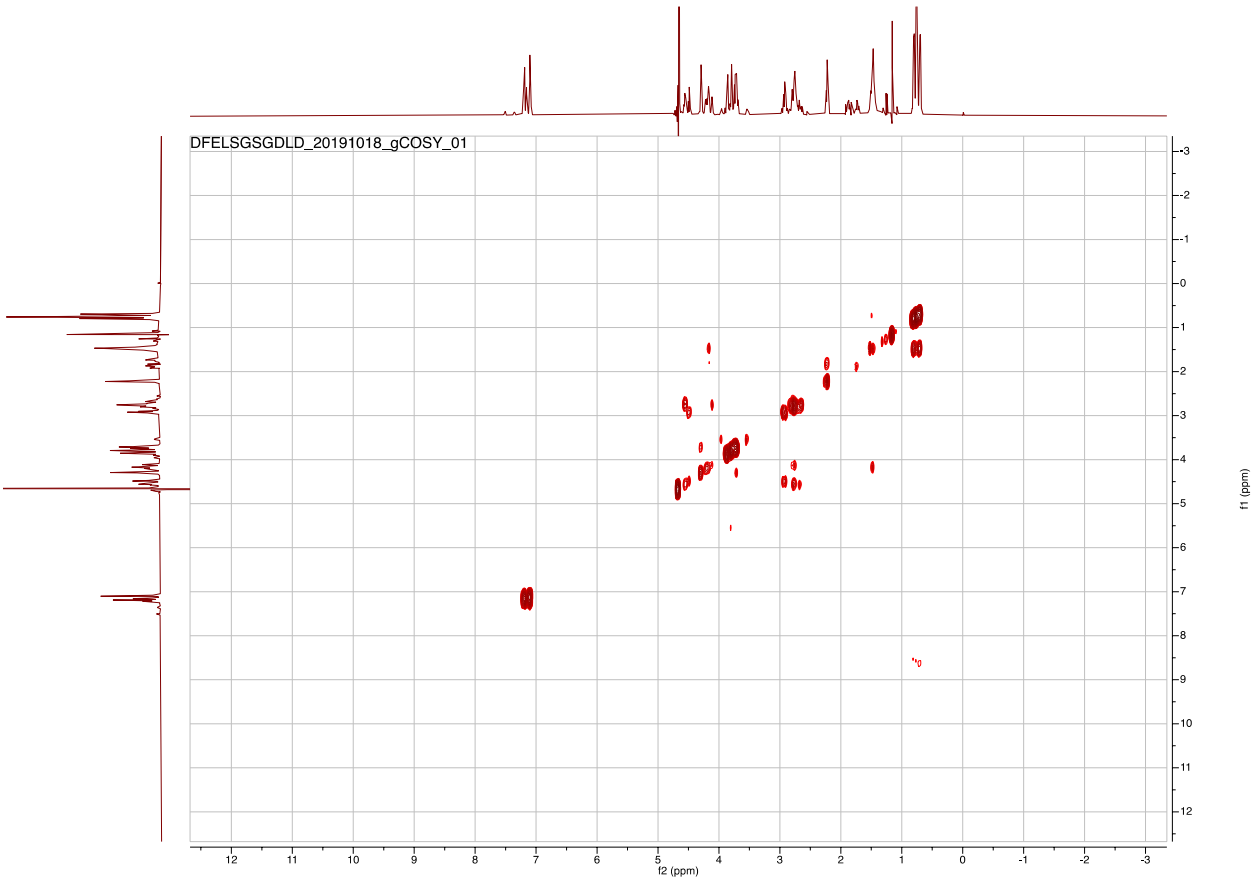
^{1}H-NMR (500 MHz, D_2O)

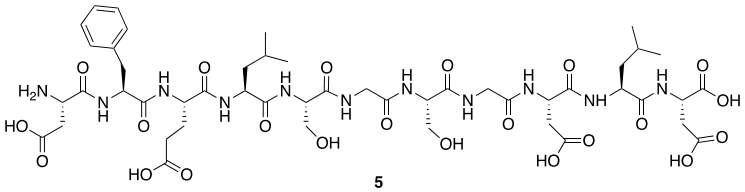




5

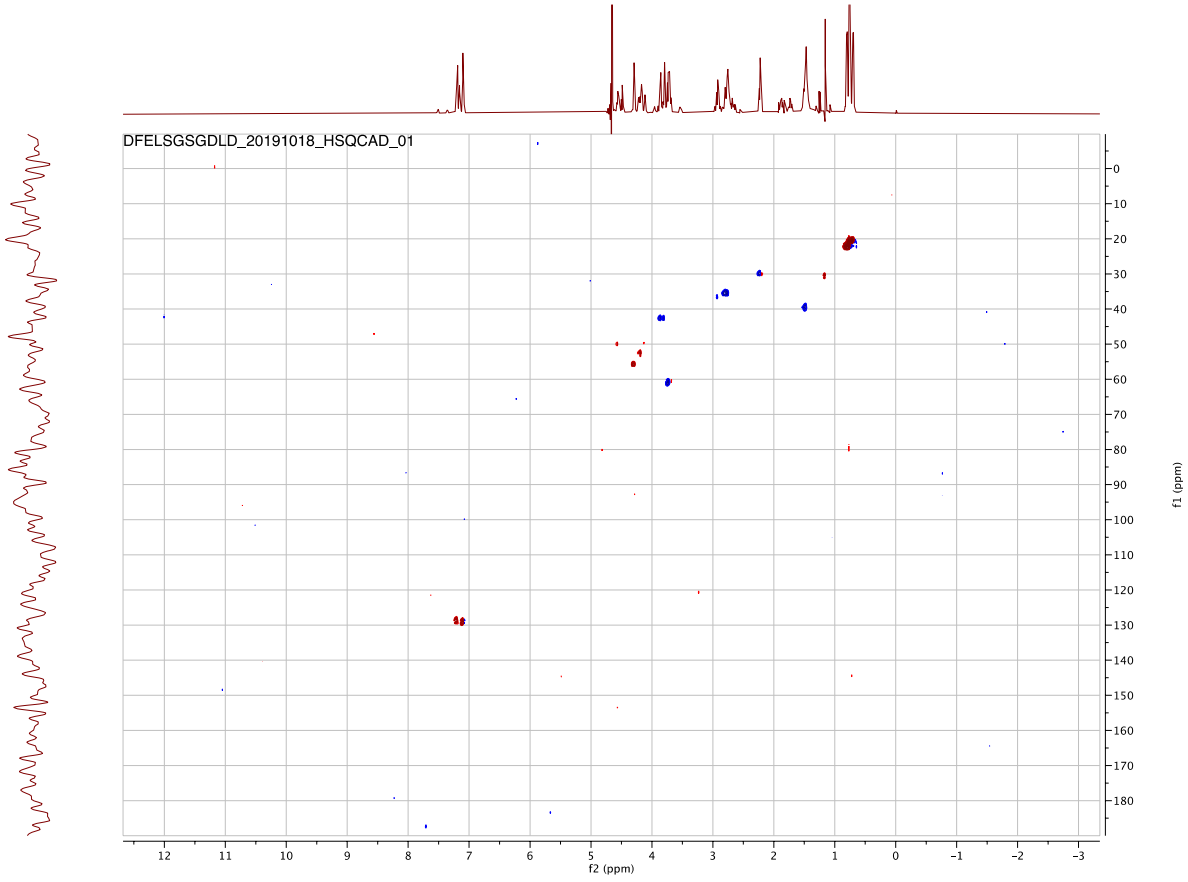
COSY (500 MHz, D_2O)

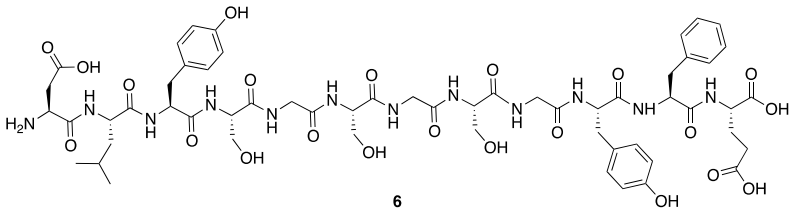




5

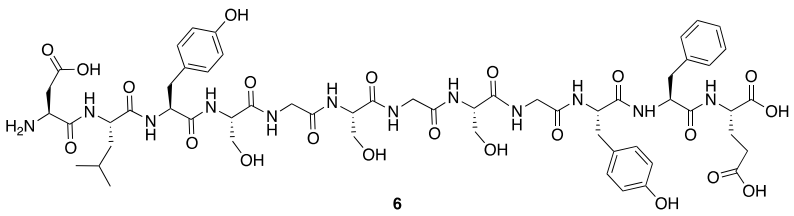
HSQC (500 MHz, D_2O)





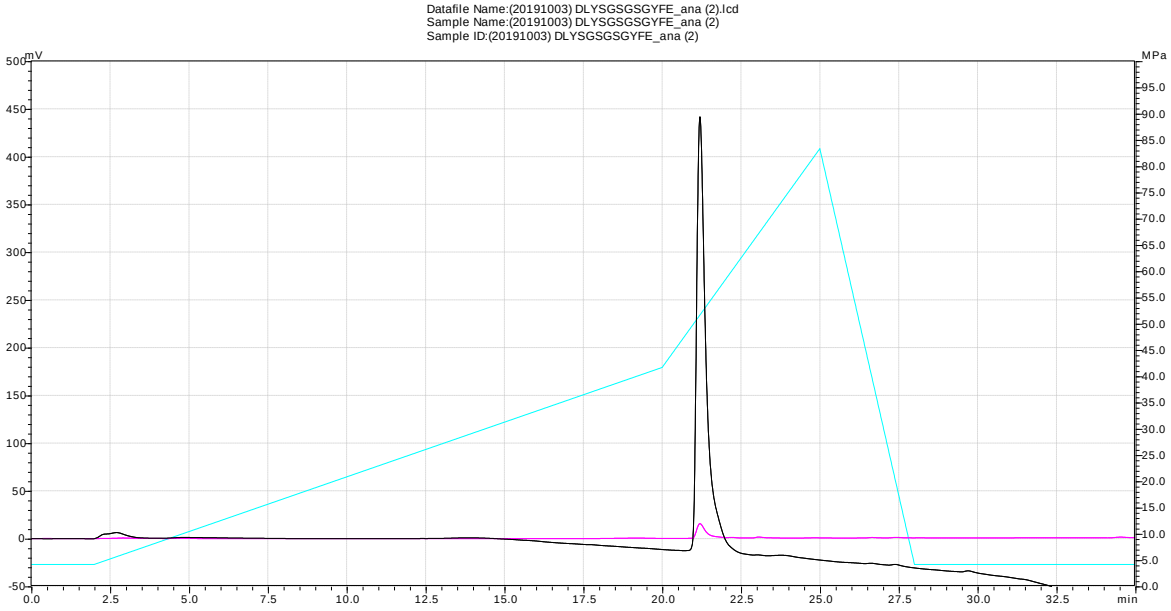
6

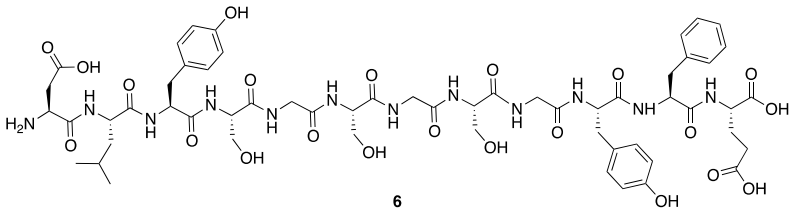
The purity of peptide **6** was verified with analytical C-18 HPLC (5-100% acetonitrile/water, 0.1% trifluoroacetic acid). $[\alpha]_D^{20}$= -36 (c 0.1, H_2O). ^{1}H NMR (500 MHz, D_2O) δ 0.66-0.70 (6 H, m), 1.23-1.27 (2 H, m), 1.29 – 1.36 (2 H, m), 1.72-1.76 (1 H, m), 1.93-1.97 (1 H, m), 2.20 (2 H, t, *J* = 7.4 Hz), 2.55 (1 H, s), 2.62 – 2.84 (7 H, m), 2.88-2.92 (2 H, m), 3.61 (2 H, dd, *J* = 11.9, 5.2 Hz), 3.64 – 3.71 (6 H, m), 3.71 – 3.84 (5 H, m), 3.88-3.92 (1 H, m), 4.08 – 4.16 (3 H, m), 4.22 – 4.35 (4 H, m), 4.40-4.43 (2 H, m), 6.57 – 6.67 (4 H, m), 6.79 – 6.85 (2 H, m), 6.90 – 6.96 (2 H, m), 7.01 – 7.07 (2 H, m), 7.08 – 7.21 (3 H, m); ^{13}C NMR (125 MHz, D_2O) δ 16.6, 20.7, 20.7, 21.9, 24.0, 25.7, 25.9, 27.7, 29.9, 32.2, 35.7, 36.0, 36.2, 37.0, 38.6, 39.5, 39.6, 42.3, 42.5, 49.6, 51.1, 51.3, 52.6, 52.9, 54.6, 54.7, 54.9, 55.1, 55.4, 55.6, 55.7, 55.9, 57.9, 58.1, 58.6, 59.9, 60.2, 61.1, 62.1, 62.7, 65.2, 88.8, 110.2, 111.2, 113.0, 114.3, 115.2, 115.3, 115.5, 117.0, 127.0, 128.4, 129.1, 130.3, 130.4, 130.8, 130.9, 135.9. ESI-MS: $C_{57}H_{77}N_{12}O_{22}$ $[M+H]^+$ calcd: 1281.5270, obsd: 1281.5206 (5.0 ppm)



6

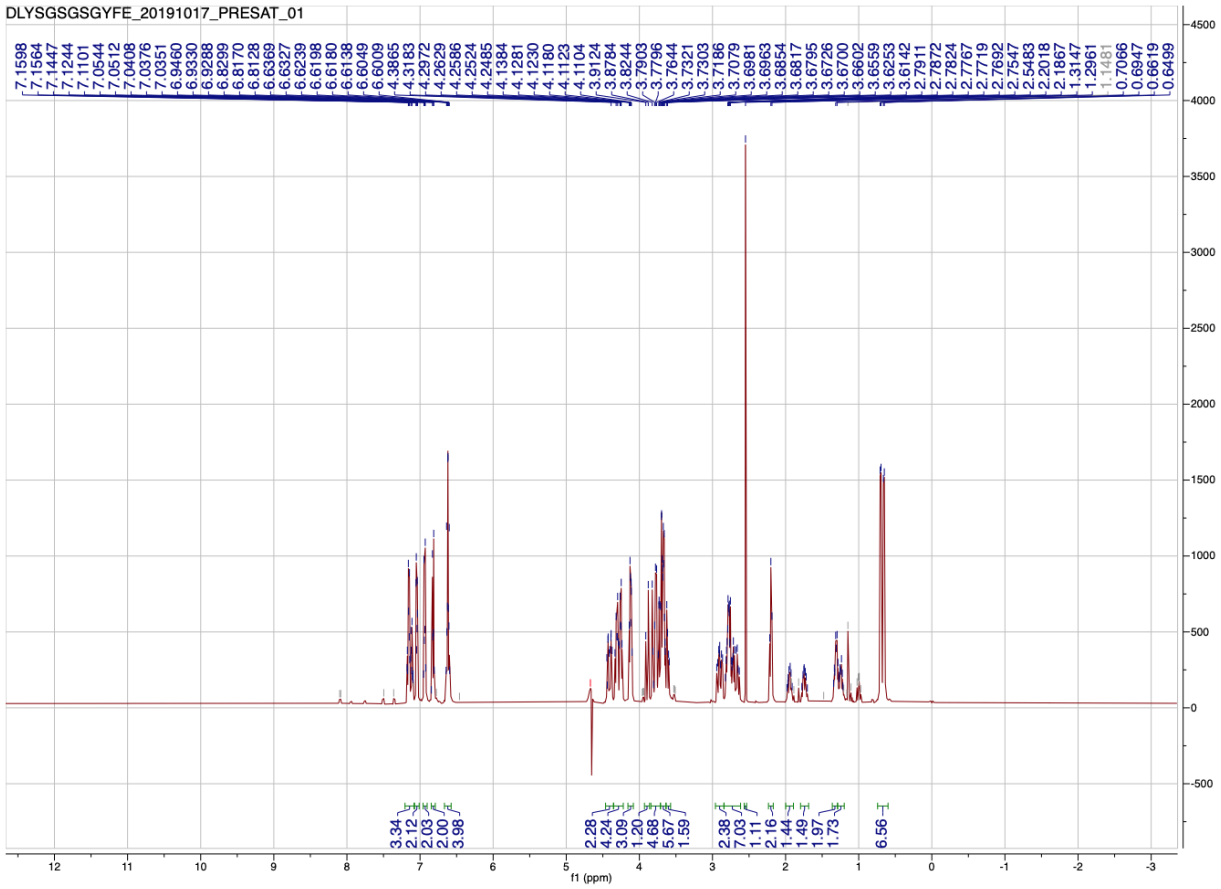
HPLC

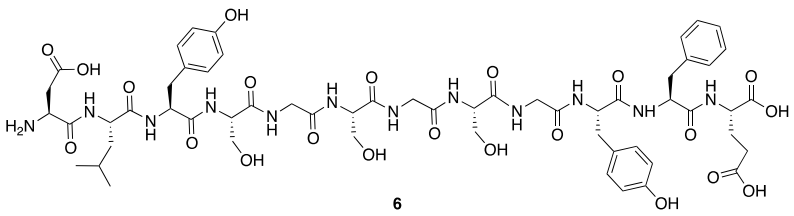

Datafile Name:(20191003) DLYSGSGSGYFE_ana (2).lcd
Sample Name:(20191003) DLYSGSGSGYFE_ana (2)
Sample ID:(20191003) DLYSGSGSGYFE_ana (2)
mV
MPa
min



6

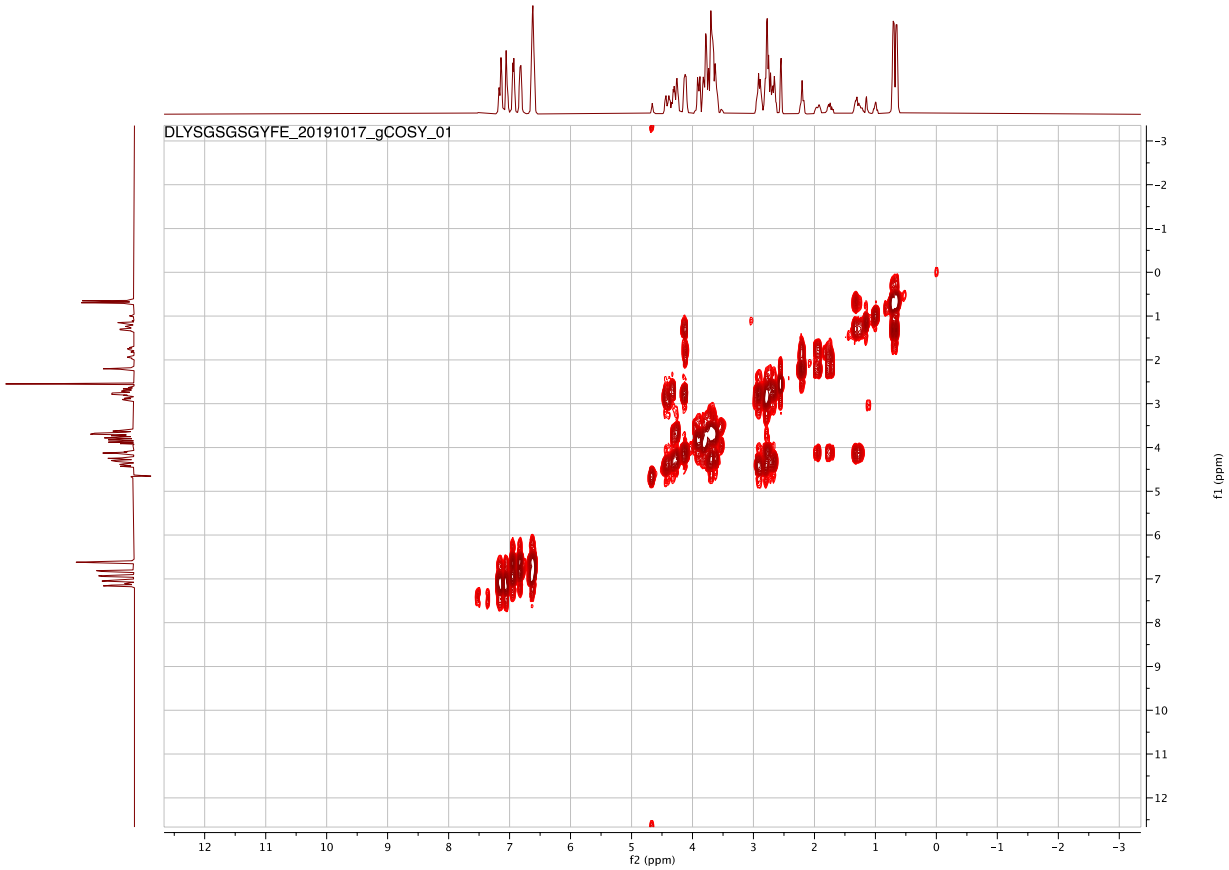
^{1}H-NMR (500 MHz, D_2O)

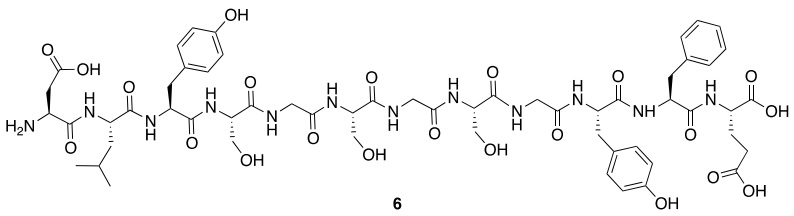




6

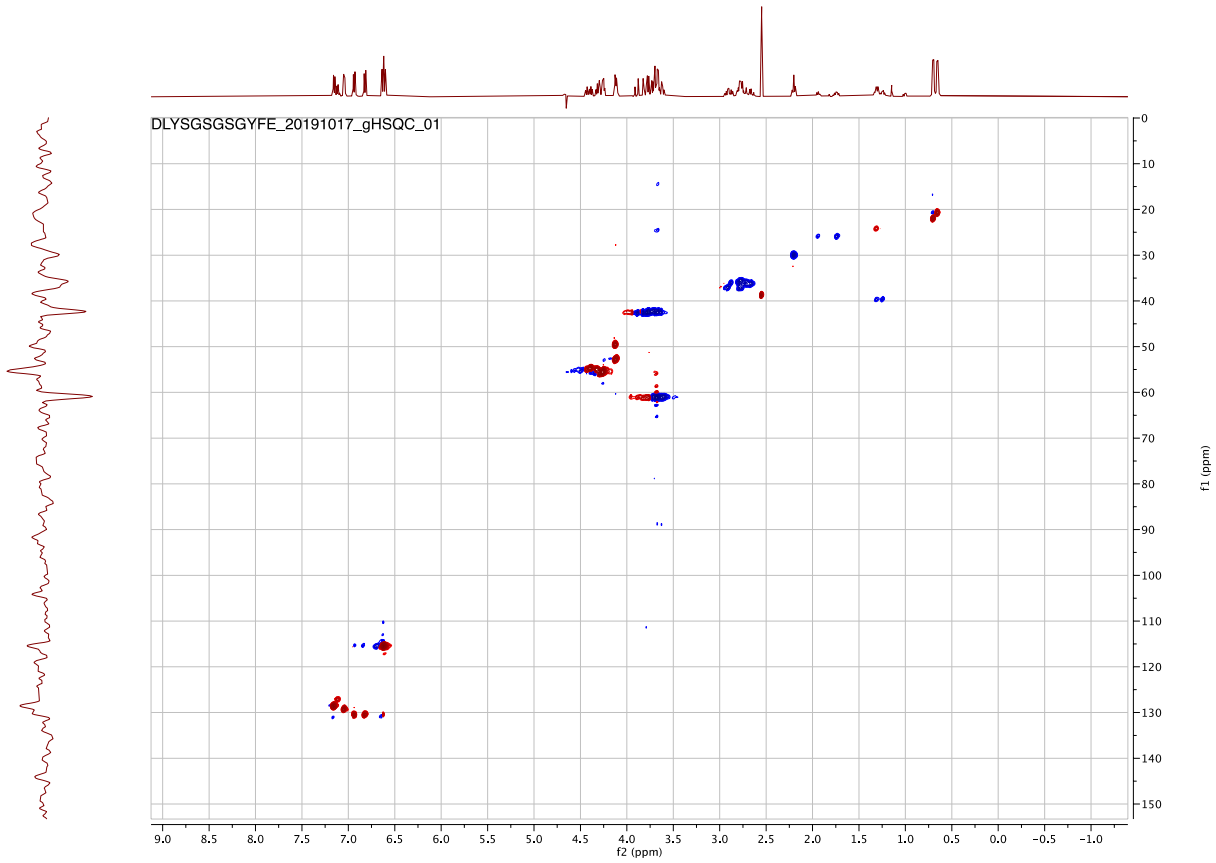
COSY (500 MHz, D_2O)

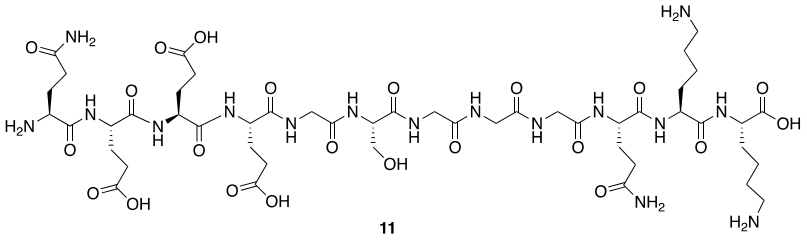




6

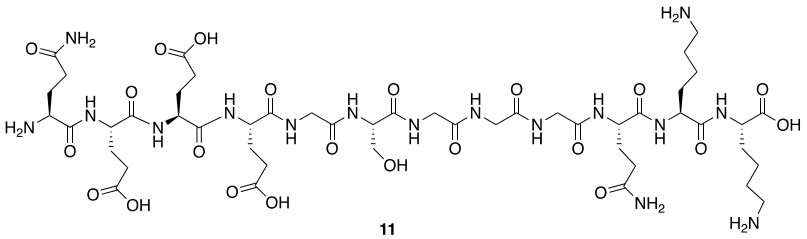
HSQC (500 MHz, D_2O)





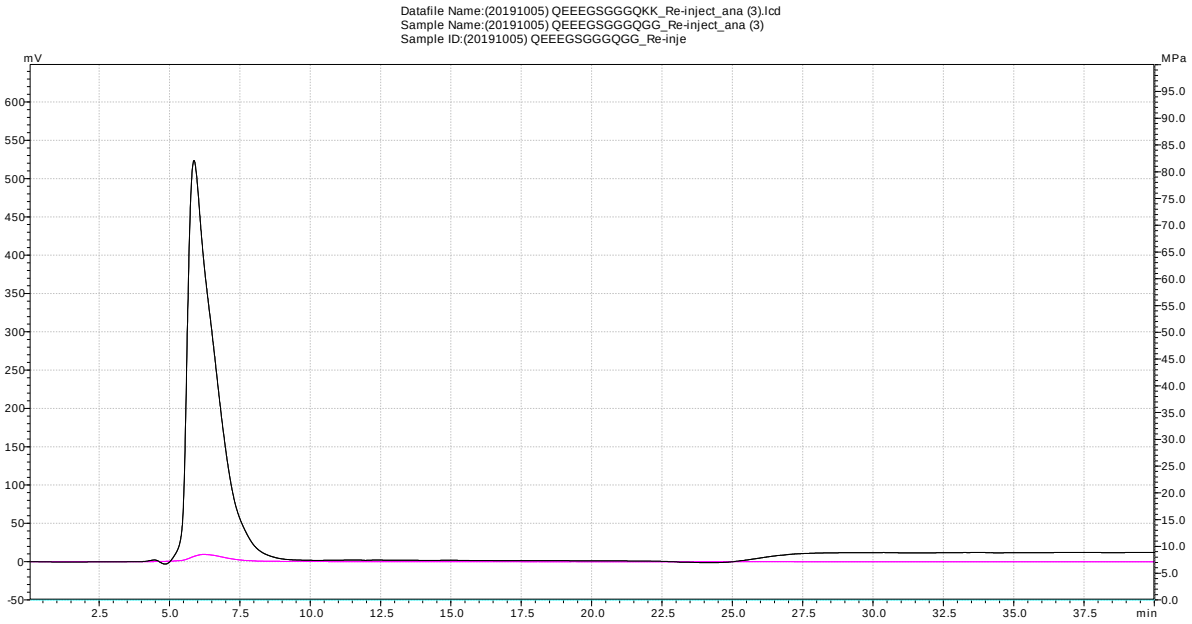
11

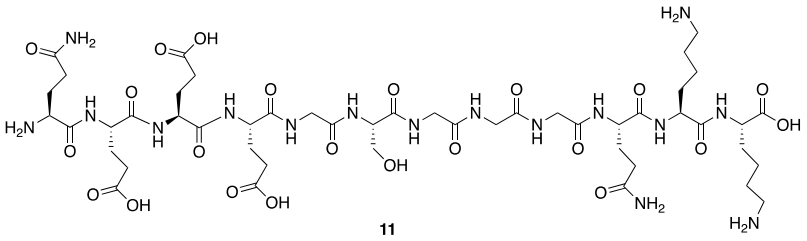
The purity of peptide **11** was verified with analytical C-18 HPLC (water, 0.1% trifluoroacetic acid). $[\alpha]_D^{20}$ = +31 (c 0.1, H_2O). ^{1}H NMR (500 MHz, D_2O) δ 1.27-1.31 (3 H, m), 1.45 – 1.55 (7 H, m), 1.56-1.64 (10 H, m), 1.70 – 2.07 (6 H, m), 2.16 – 2.41 (6 H, m), 2.83 (2 H, s), 2.92 – 3.08 (8 H, m), 3.67 – 3.81 (2 H, m), 3.81 – 3.98 (3 H, m), 4.08 – 4.35 (3 H, m); ^{13}C NMR (125 MHz, D_2O) δ 17.7, 19.7, 20.4, 21.1, 21.2, 21.4, 21.7, 22.0, 22.9, 23.2, 23.7, 24.5, 25.9, 26.0, 26.2, 26.4, 27.0, 27.2, 27.4, 28.7, 29.7, 29.9, 30.2, 30.7, 31.3, 31.8, 32.2, 34.0, 36.8, 38.1, 39.1, 39.3, 40.8, 42.1, 42.3, 42.6, 43.3, 43.5, 43.6, 44.1, 44.3, 44.6, 44.8, 45.3, 45.5, 46.6, 52.1, 52.9, 53.3, 54.1, 55.6, 56.7, 60.9. ESI-MS: $C_{48}H_{81}N_{16}O_{22}$ $[M+H]^+$ calcd: 1233.5706, obsd: 1233.5679 (2.2 ppm).



11

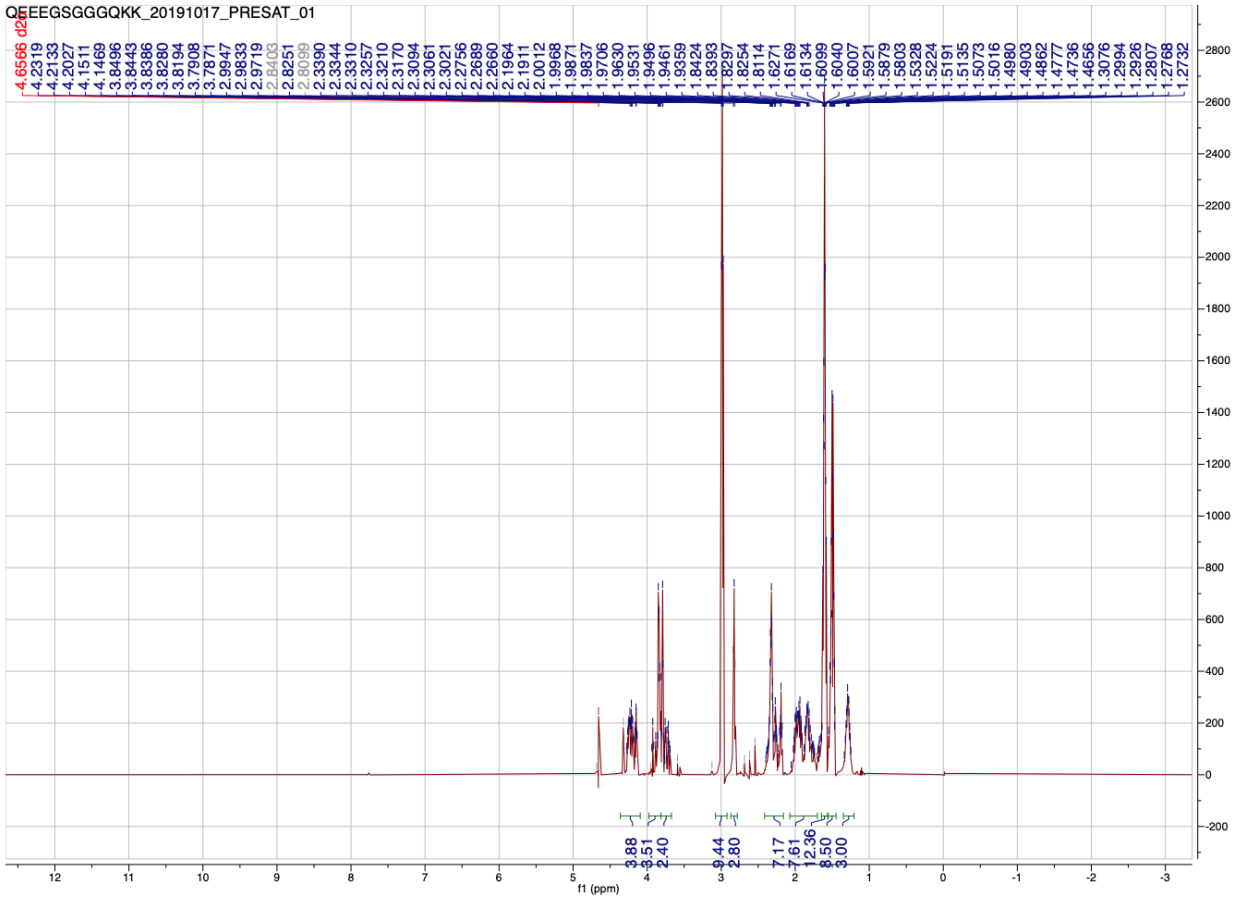
HPLC

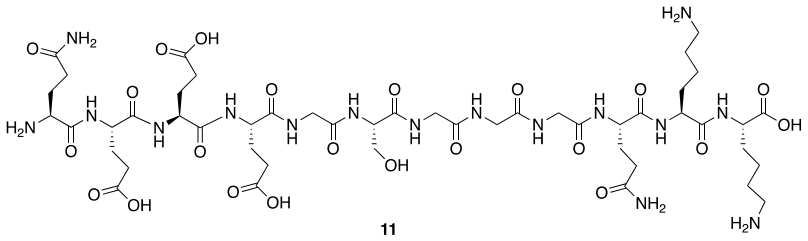

Datafile Name:(20191005) QEEEGSGGGQKK_Re-inject_ana (3).lcd
Sample Name:(20191005) QEEEGSGGGQGG_Re-inject_ana (3)
Sample ID:(20191005) QEEEGSGGGQGG_Re-inje
mV
MPa
min



11

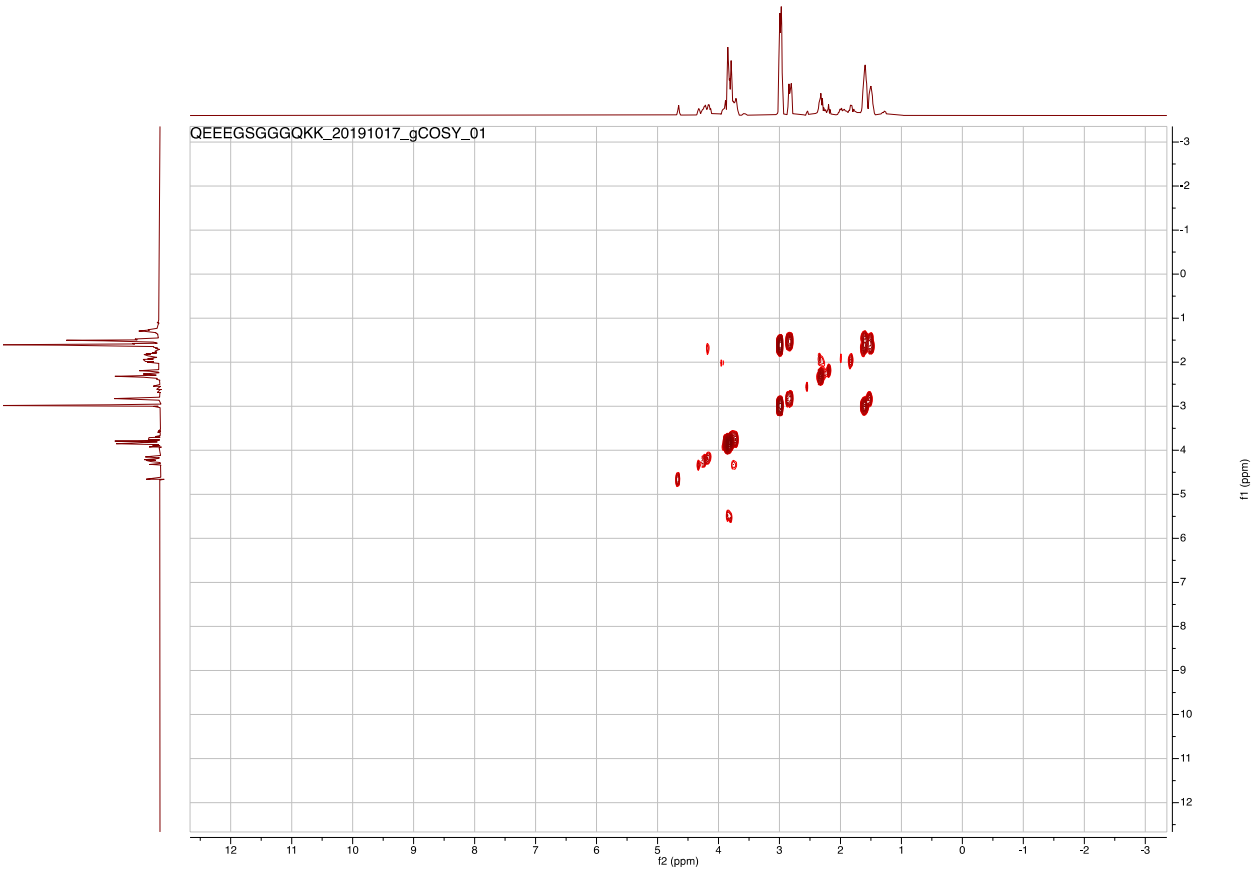
^{1}H-NMR (500 MHz, D_2O)

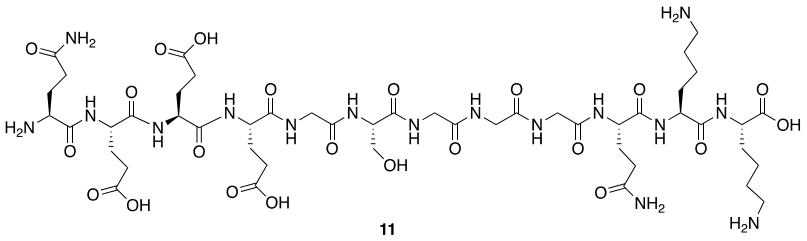




11

COSY (500 MHz, D_2O)





11

HSQC (500 MHz, D_2O)

References

1. Liu, L.-K.; Finzel, B. C., Fragment-Based Identification of an Inducible Binding Site on Cell Surface Receptor CD44 for the Design of Protein–Carbohydrate Interaction Inhibitors. *J. Med. Chem.* **2014,** *57* (6), 2714-2725.

2. Gao, J.; Lin, P.-H.; Nick, S. T.; Huang, J.; Tykesson, E.; Ellervik, U.; Li, J.; Huang, X., Chemoenzymatic Synthesis of Glycopeptides bearing Galactose-xylose Disaccharide from the Proteoglycan Linkage Region. *Org. Lett.* **2021**, *23* (5), 1738–1741.

3. Wu, Z. L.; Ethen, C. M.; Prather, B.; Machacek, M.; Jiang, W., Universal phosphatase-coupled glycosyltransferase assay. *Glycobiology* **2011,** *21* (6), 727-733.