Supplementary Information for:

Genetically-encoded fragment-based (GE-FBD) discovery of glycopeptide ligands for antibodies related to mycobacterial infections

Ying Chou, Elena N. Kitova, Maju Joe, Richard M Brunton, Todd L Lowary, John S. Klassen, Ratmir Derda*

*Corresponding author: ratmir@ualberta.ca

Table of Contents (click on the topic)

Abbreviation

Ara	Arabinose
Ara_6	Methyl β-D-arabinofuranosyl-(1→2)-α-D-arabinofuranosyl-(1→5)-[β-D-arabinofuranosyl-(1→2)-α-D-arabinofuranosyl-(1→3)]-α-D-arabinofuranosyl-(1→5)-α-D-arabinofuranoside
bp	Base pair
BSA	Bovine serum albumin
DIC	*N*,*N'*-Diisopropylcarbodiimide
DIPEA	*N*,*N*-Diisopropylethylamine
DMF	Dimethyl formamide
dNTP	Deoxynucleotide
dsDNA	Double-stranded DNA
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
eq.	Equivalent
ESI	Electrospray ionization
FA	Formic acid
F_{ab}	Antigen-binding fragment
Fmoc	Fluorenylmethyloxycarbonyl
FTS	Fluorescein-5-thiosemicarbazide
Glc	Glucose
GSH	Glutathione
h	Hour
HBTU	2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOAt	1-Hydroxy-7-azabenzotriazole
HPLC	High-performance liquid chromatography
kDa	Kilo dalton
MALDI	Matrix-assisted laser desorption ionization
MeCN	Acetonitrile
MWCO	Molecular weight cut-off
min	Minute
MOPS	3-Morpholinopropane-1-sulfonic acid
MQ	Milli-Q
NMR	Nuclear magnetic resonance
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PFU	Plaque forming unit
rha	Rhamnose
RP	Reversed phase
rpm	Revolutions per minute
RT	Room temperature

s	Second
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
ssDNA	Single-stranded DNA
tBu	*tert*-Butyl
TFA	Trifluoroacetic acid
TRITC	tetramethylrhodamine
TIPS	Triisopropylsilane
UPLC	Ultra performance liquid chromatography

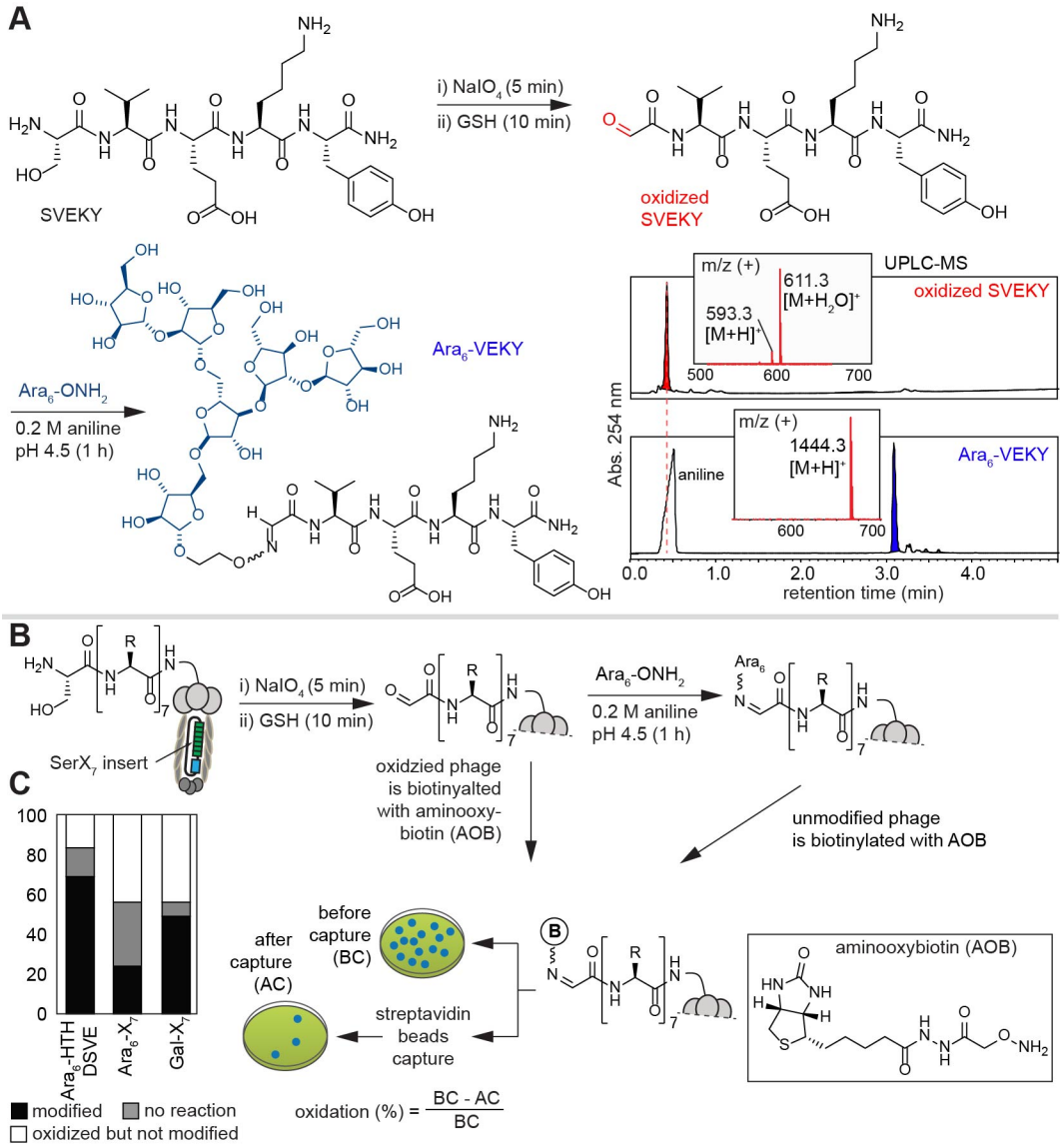


Figure S1. Validation of chemical modifications on peptide, phage clones and libraries.
(A) LC-MS confirms oxidation and oxime ligation reactions on the peptide SVEKY. (B) Oxidation and oxime ligation reaction conditions and quantification method on phage libraries. (C) The conversion of the oxidation and oxime ligation reactions on monoclonal phage and phage libraries as determined by direct biotin capture (oxidation) and differential biotin capture (ligation).

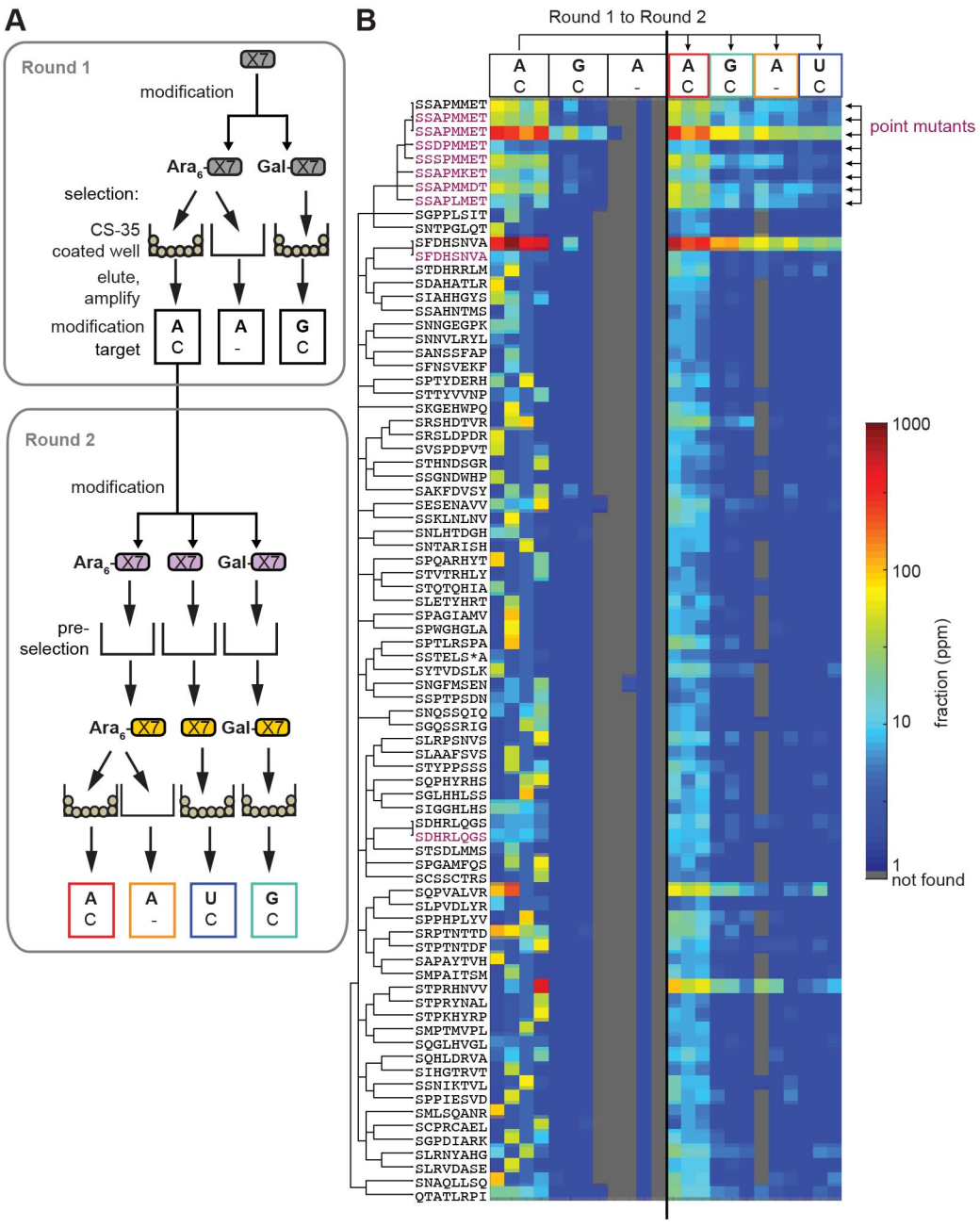


Figure S2. Detailed selection scheme and post-selection analysis.

(A) Schemes of panning. (B) Heat map showing 80 putative hits discovered from two rounds of panning. Sequences of putative hits were enriched greater or equal to 4-fold ($R \geq 4$, $p < 0.05$) when compared to control experiments (GC from RI, A-from RI, GC from RII, A- from RII, and UC from RII).

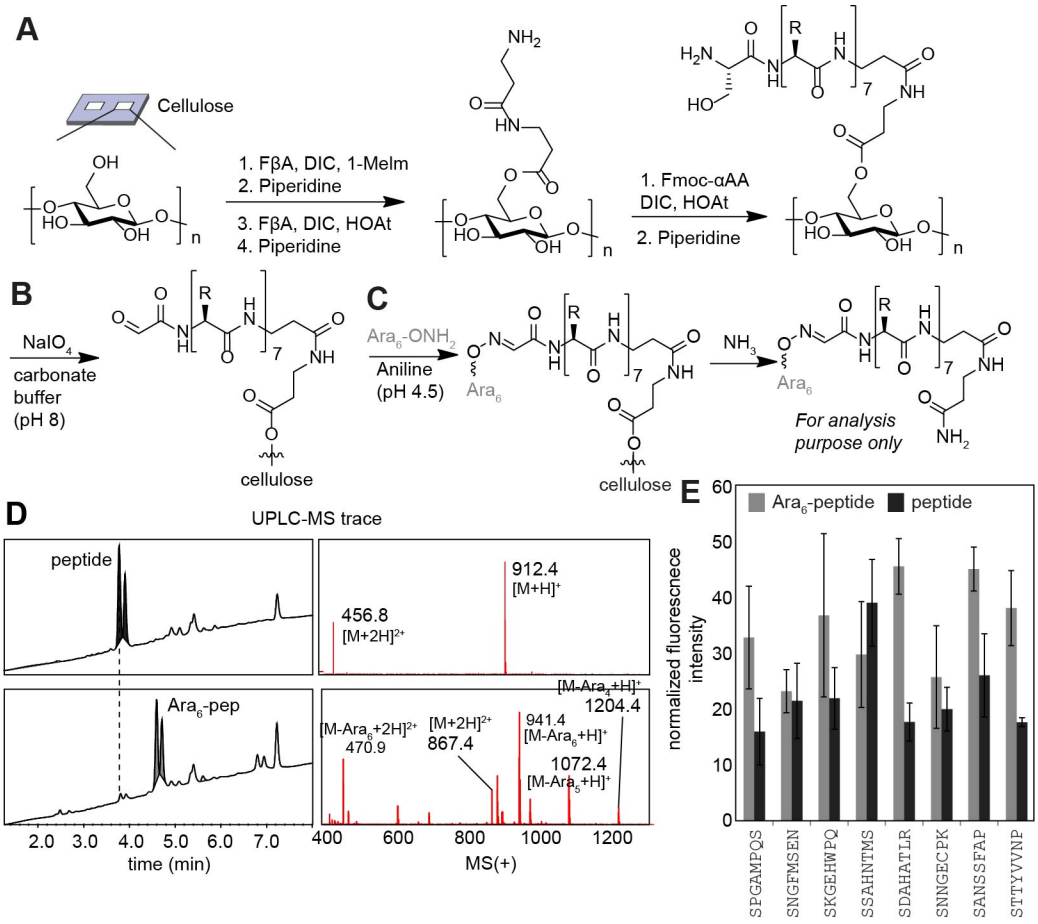


Figure S3. Binding of glycopeptides to CS-35 on paper-based assay.
(A) SyntArray to synthesize peptides on Teflon-patterned paper. (B) Subsequent oxidation and (C) oxime ligation yielded glycopeptide arrays. (D) UPLC-MS trace to confirm the formation of Ara_6-peptide (SGPPLSIT-βA-βA) on paper arrays. Only doubly charged species and fragmented singly charged species were observed in MS. (E) The results of paper-based assay showing binding of fluorescently labeled CS-35 to peptides and Ara_6-peptides. The "hits" were defined as Ara_6-peptides that bound significantly more antibody than aglycone-peptide. Fluorescence was obtained using excitation wavelength at 532 nm and long-pass emission filter LPG (≥575 nm).

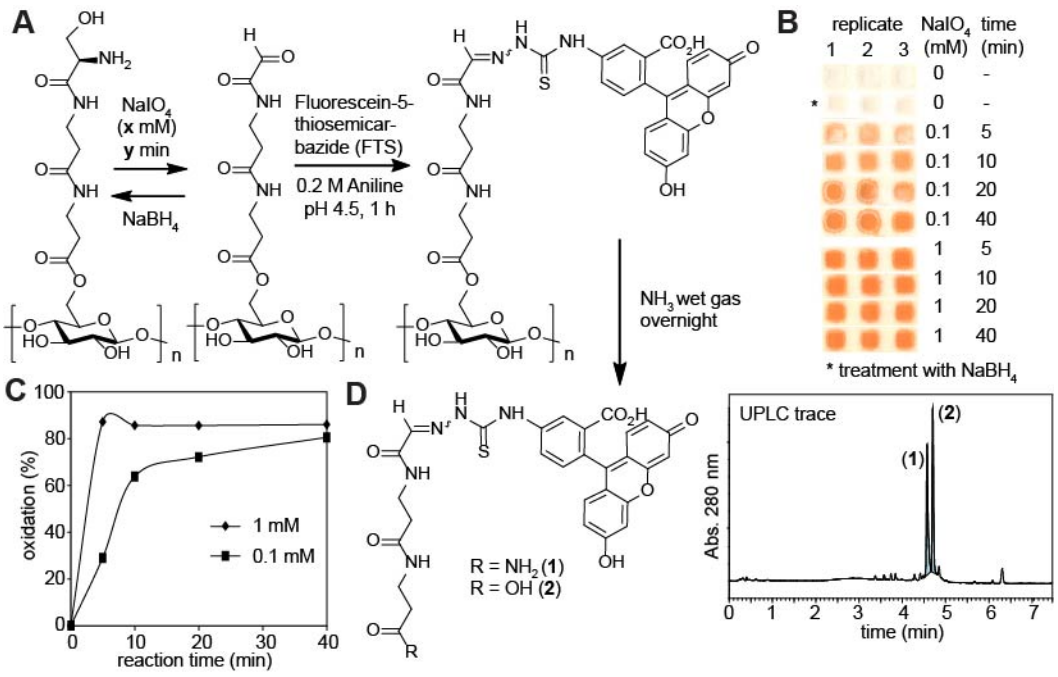


Figure S4. Optimization of oxidation conditions for peptides immobilized on paper.
(A) Modification of peptides on paper array via aldehyde formation and subsequent ligation to N-terminal aldehyde. (B) Treatment of peptides with different concentrations of $NaIO_4$ for specified time length, followed by reaction with FTS resulted in the formation of FTS-conjugate (orange). (C) Oxidation went near completion in 5 min using 1 mM $NaIO_4$. (D) Structures and LC-MS validation of FTS-βAla-βAla-conjugates formed in the reaction on paper.

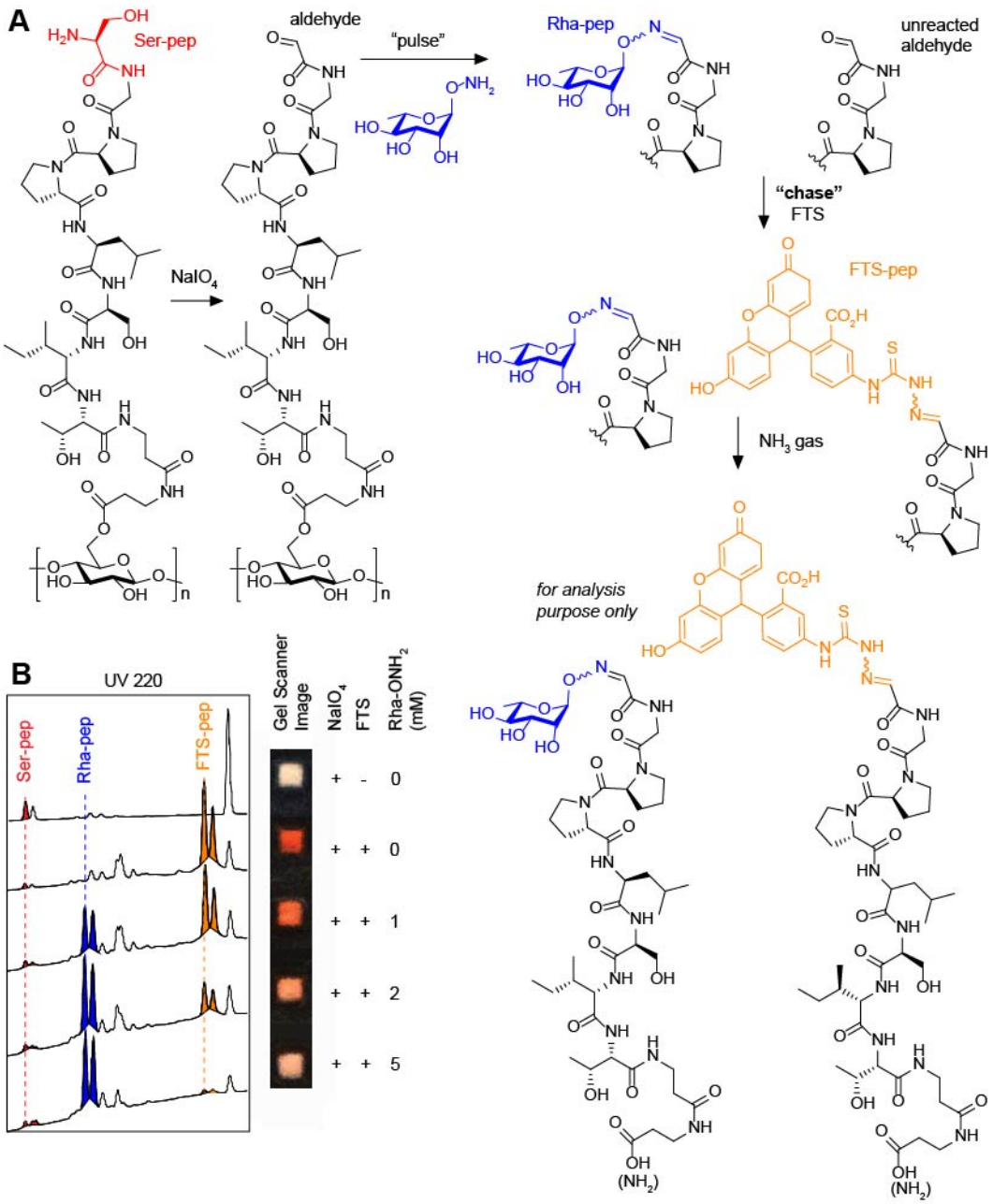


Figure S5. Optimization of glycopeptide formation on paper.

(A) N-terminal serine is oxidized, followed by the addition of increasing concentration of aminooxy derivative (Rha-ONH2). FTS is then added to react with remaining oxidized but un-ligated peptides. (B) FTS readily reacts with oxidized peptides, forming FTS-peptide, and can be visually detected by the orange color on paper. Both colorimetric assay and LCMS confirmed formation of Rha-peptide.

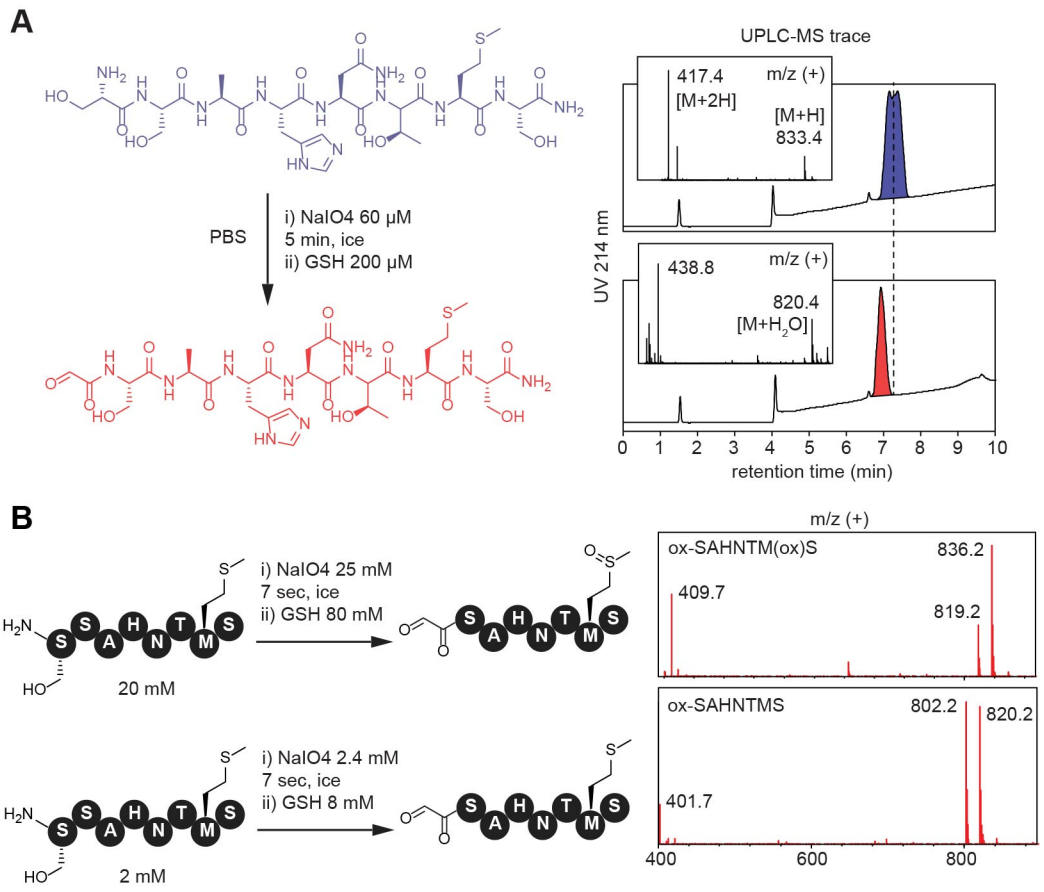


Figure S6. Selective oxidation of methionine using kinetic control.

(A) Reaction condition similar to phage reaction yields peptides that are oxidized only at N-terminal serine but not methionine. (B) At 25 mM of $NaIO_4$, methione and N-terminal serine are both oxidized. (C) At 2.4 mM of $NaIO_4$, methionine is not oxidized, but N-terminal serine is oxidized.

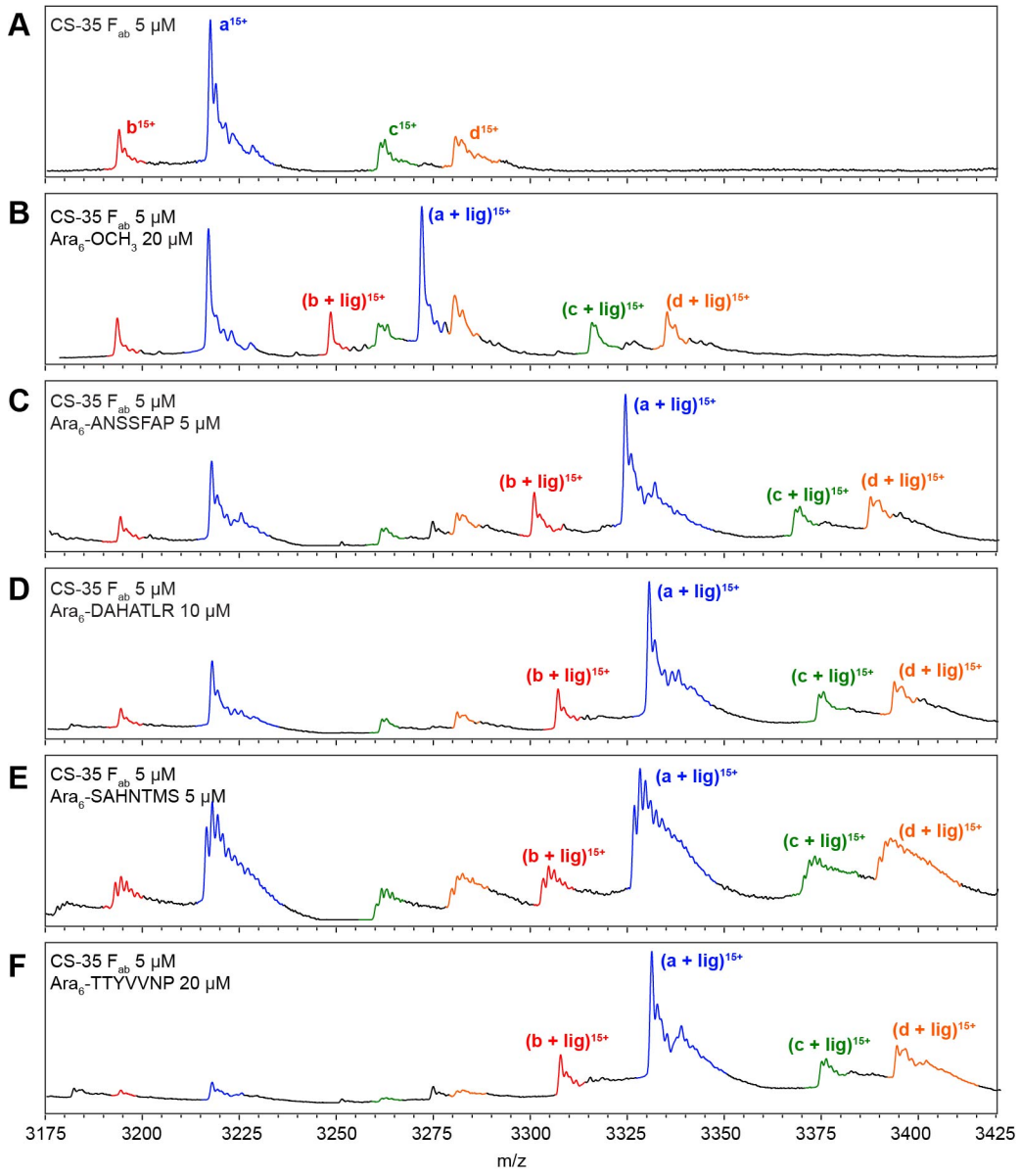


Figure S7. ESI-MS spectra of Ara_6-glycopeptides and parent glycan Ara_6 to CS-35 F_{ab}.
ESI-MS spectra of: (A) CS-35 F_{ab} labeled with four glycoform species a, b, c, d within the charge state 15; (B) parent glycan Ara_6 binding to CS-35 F_{ab}; (C)-(F) Ara_6-containing glycopeptides binding to CS-35 F_{ab}. The ligand-F_{ab} complex forming from each glycoform species is labeled correspondingly.

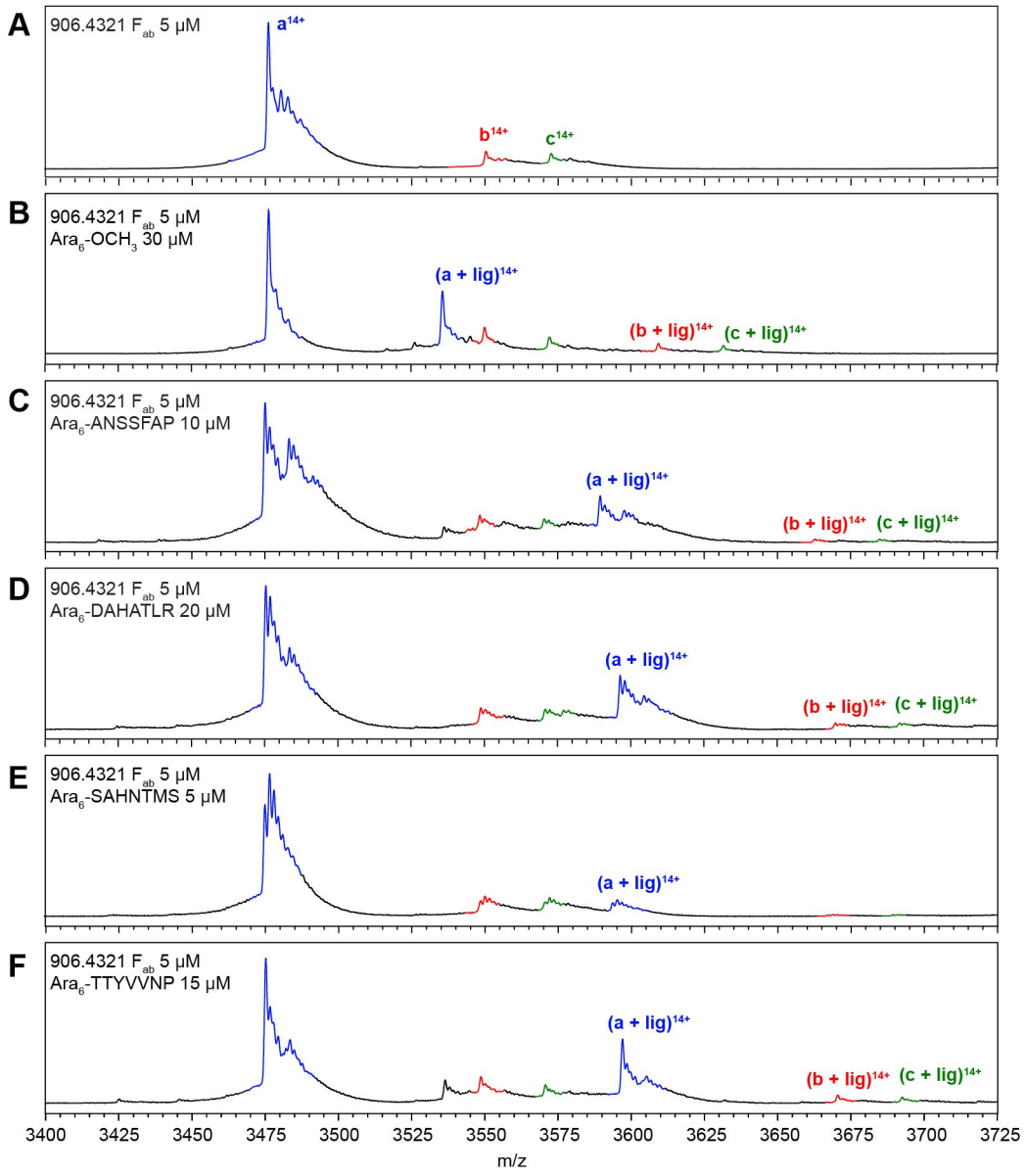


Figure S8. ESI-MS spectra of Ara$_6$-glycopeptides and parent glycan Ara$_6$ to 906.4321 F$_{ab}$.
ESI-MS spectra of: (A) 906.4321F$_{ab}$ labeled with three glycoform species a, b, c within the charge state 14; (B) parent glycan Ara$_6$ binding to 906.4321 F$_{ab}$; (C)-(F) Ara$_6$-containing glycopeptides binding to 906.4321 F$_{ab}$. The ligand-F$_{ab}$ complex forming from each glycoform species is labeled correspondingly.

Table S1. K_D of ligands (1)-(13) to CS-35 and 906.4321 F_{ab} measured by ESI-MS

	K_D CS-35 F_{ab} (μM)	K_D 906.4321 F_{ab} (μM)
SANSSFAP (**1**)	372 ± 107	710 ± 344
Glc-ANSSFAP (**2**)	N/A*	630 ± 852
Ara_6-ANSSFAP (**3**)	1.4 ± 1.0	20 ± 3
STTYVVNP (**4**)	463 ± 282	2090 ± 1402
Glc-TTYVVNP (**5**)	N/A*	N/A*
Ara_6-TTYVVNP (**6**)	1.9 ± 1.2	33 ± 4.7
SSAHNTMS (**7**)	312 ± 114	1082 ± 732
Glc-SAHNTMS (**8**)	N/A*	83.3*
Ara_6-SAHNTMS (**9**)	3.2 ± 1.0	47 ± 13
SDAHATLR (**10**)	388 ± 146	1012 ± 310
Glc-DAHATLR (**11**)	N/A*	N/A*
Ara_6-DAHATLR (**12**)	7.0 ± 6.2	59 ± 8.3
Ara_6 (**13**)	19 ± 9.1	52 ± 31

***** K_D of ligands could not be measured due to overlap of some of the ions of F_{ab} with ions of F_{ab}-ligand species.

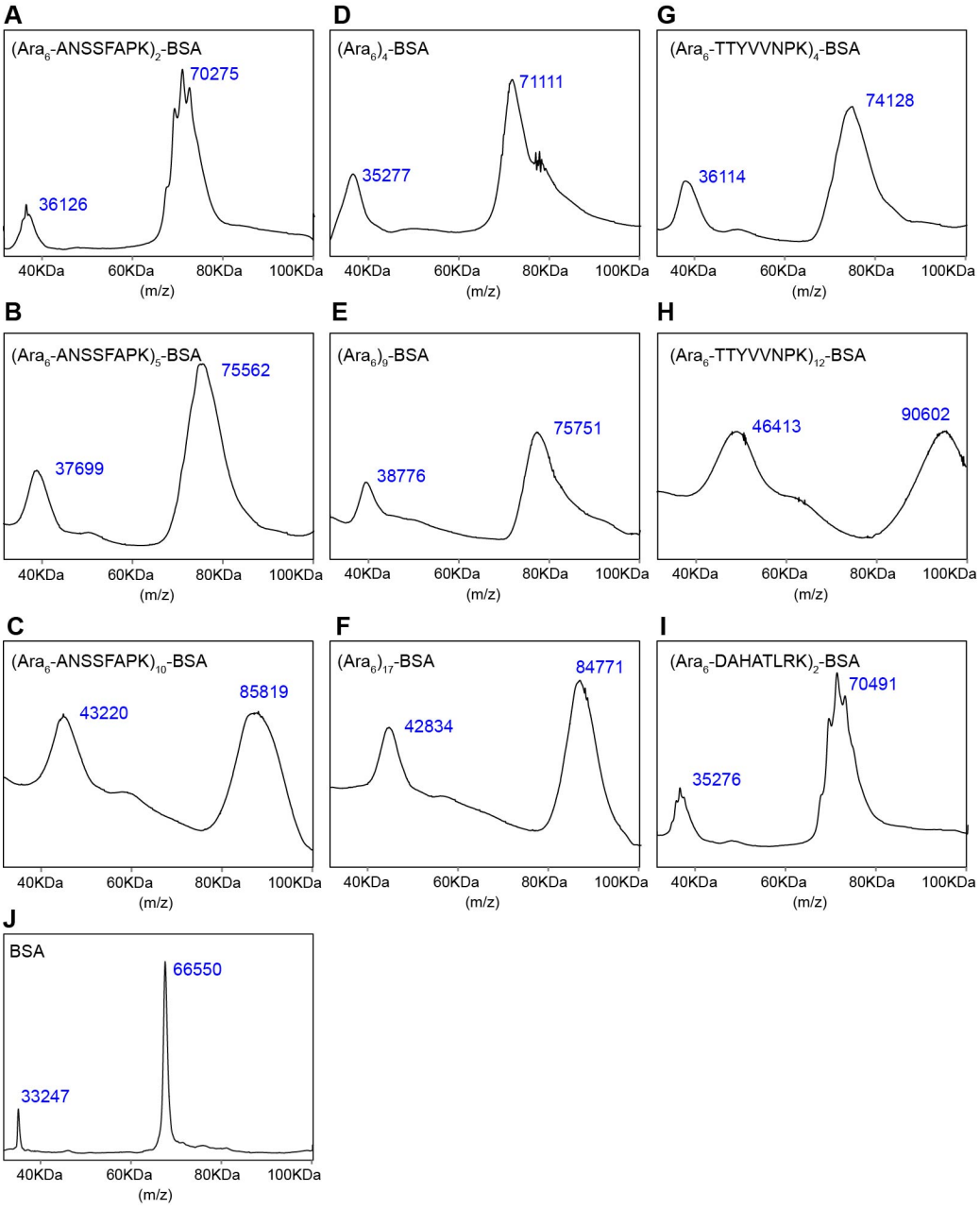


Figure S9. MALDI characterization of synthesized Ara_6-glycopeptide-BSA conjugates.
MALDI of: (A)-(I) Ara_6-glycopeptide-BSA conjugates with singly charged and doubly charged peaks; (J) BSA alone.

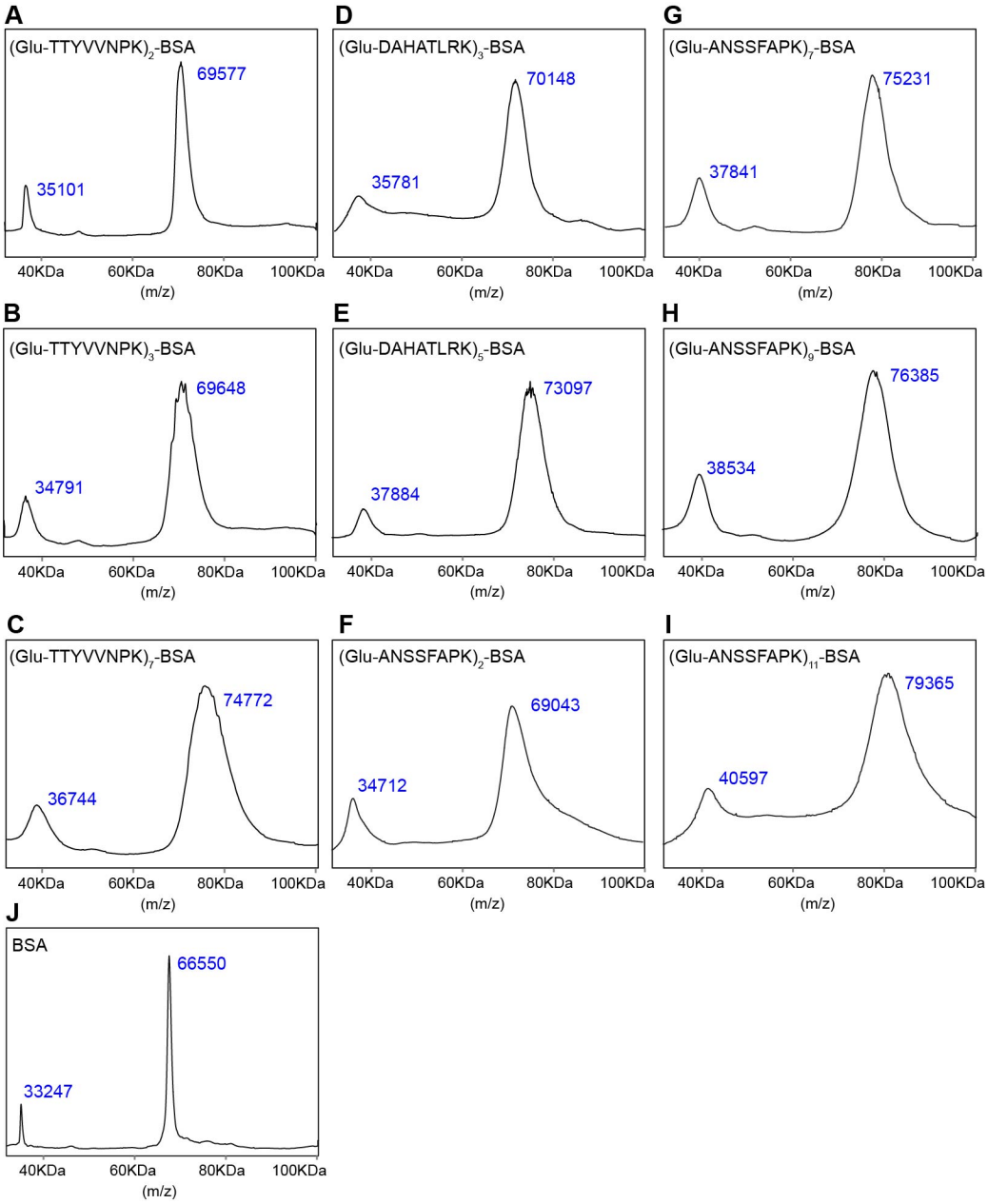


Figure S10. MALDI characterization of synthesized glucose-glycopeptide-BSA conjugates.
MALDI of: (A)-(I) glucose-glycopeptide-BSA conjugates with singly charged and doubly charged peaks; (J) BSA alone.

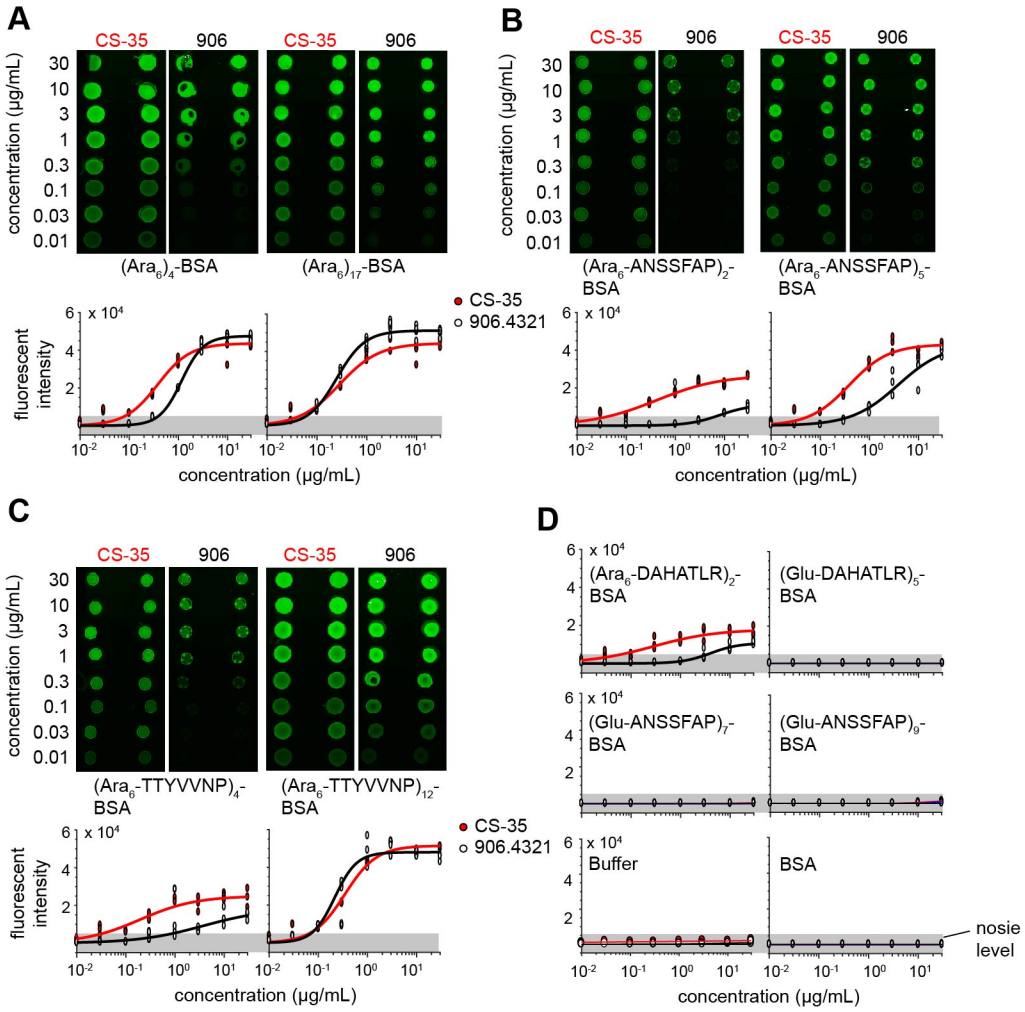


Figure S11. Ara_6-peptides binding in multivalent microarray assay.
Fluorescence images and dose-response curves of (A) $(Ara_6)_4$-BSA, $(Ara_6)_{17}$-BSA, (B) $(Ara_6$-ANSSFAP$)_2$-BSA, $(Ara_6$-ANSSFAP$)_{10}$-BSA, and (C) $(Ara_6$-TTYVVNP$)_4$-BSA, $(Ara_6$-TTYVVNP$)_{12}$-BSA binding to CS-35, CS-40, and 906.4321. (D) Dose-response curves of $(Ara_6$-DAHATLR$)_2$-BSA and glucose-containing ligands.

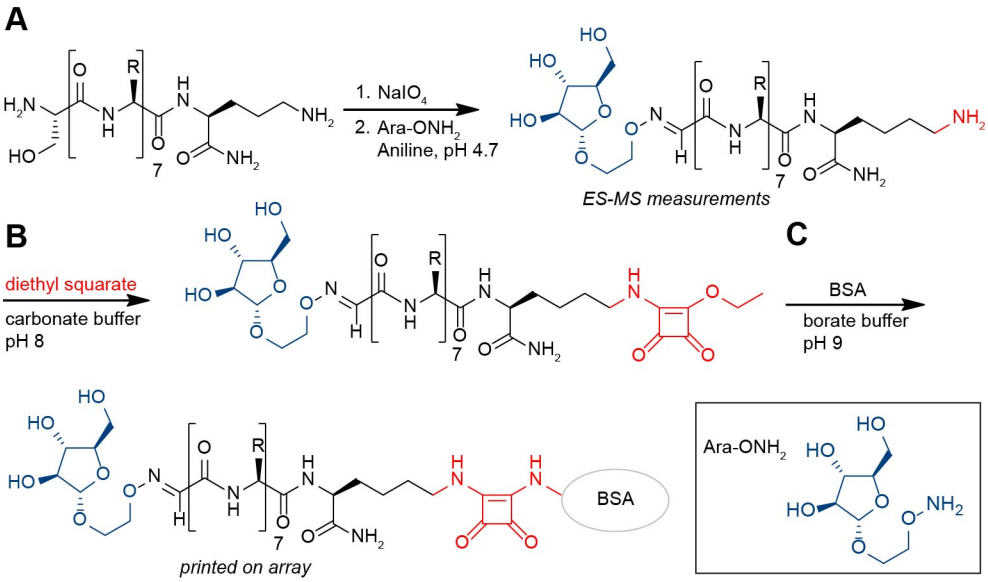


Figure S12. Synthesis scheme of Ara_1-peptides for ESI-MS measurements and microarray printing. (A) Synthesis of Ara_1-containing glycopeptides. BSA conjugates of glycopeptides were synthesized through (B) squaramide chemistry and (C) subsequent conjugation onto BSA.

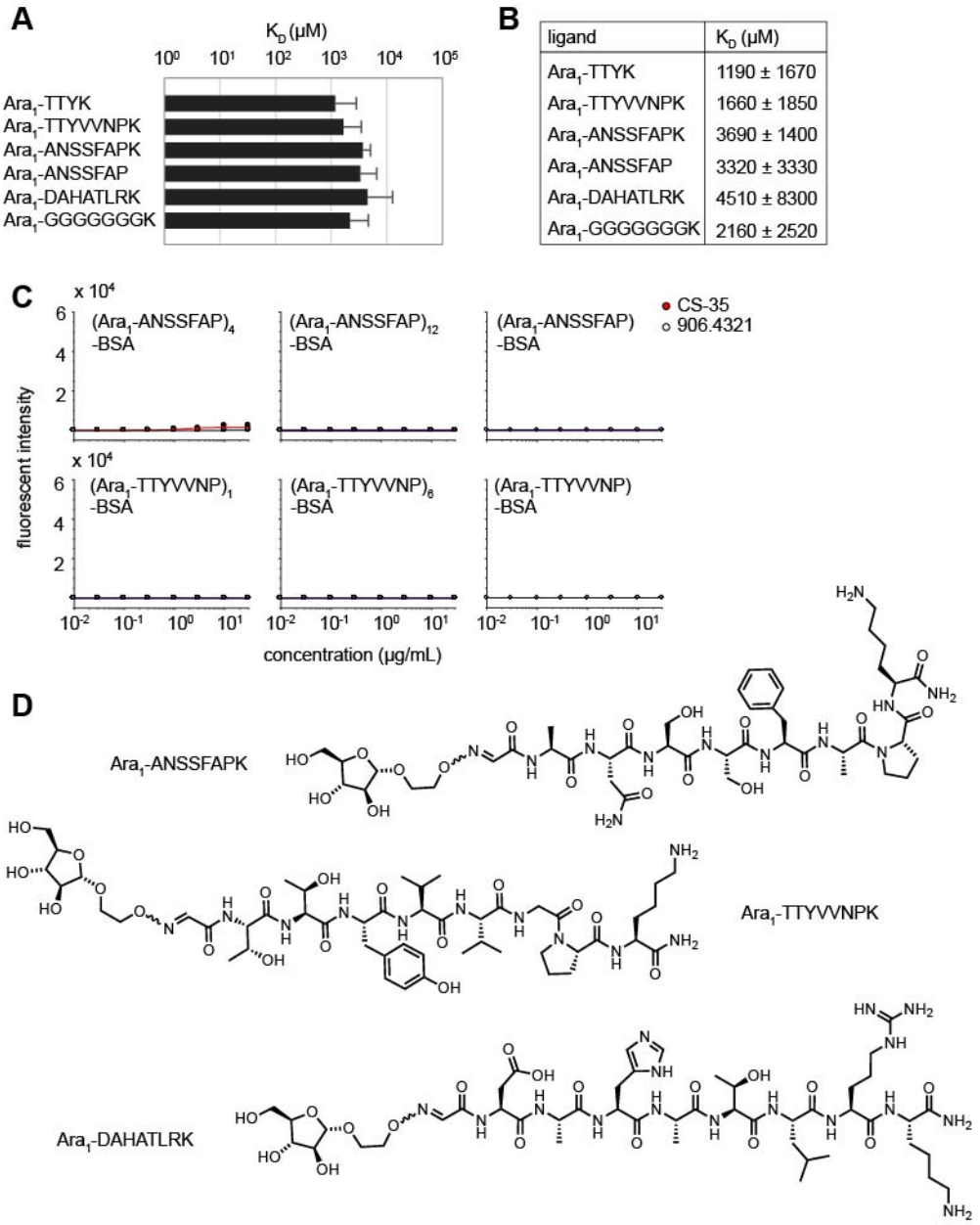


Figure S13. ESI-MS binding measurements and microarray testing of Ara_1-peptides.
(A) K_D of Ara-peptides to CS-35 measured by ESI-MS in bar graphs and in (B) table. (C) Dose-response curves of Ara-peptides binding to CS-35 and 906.4321. (D) Representative structures of Ara_1-peptides.

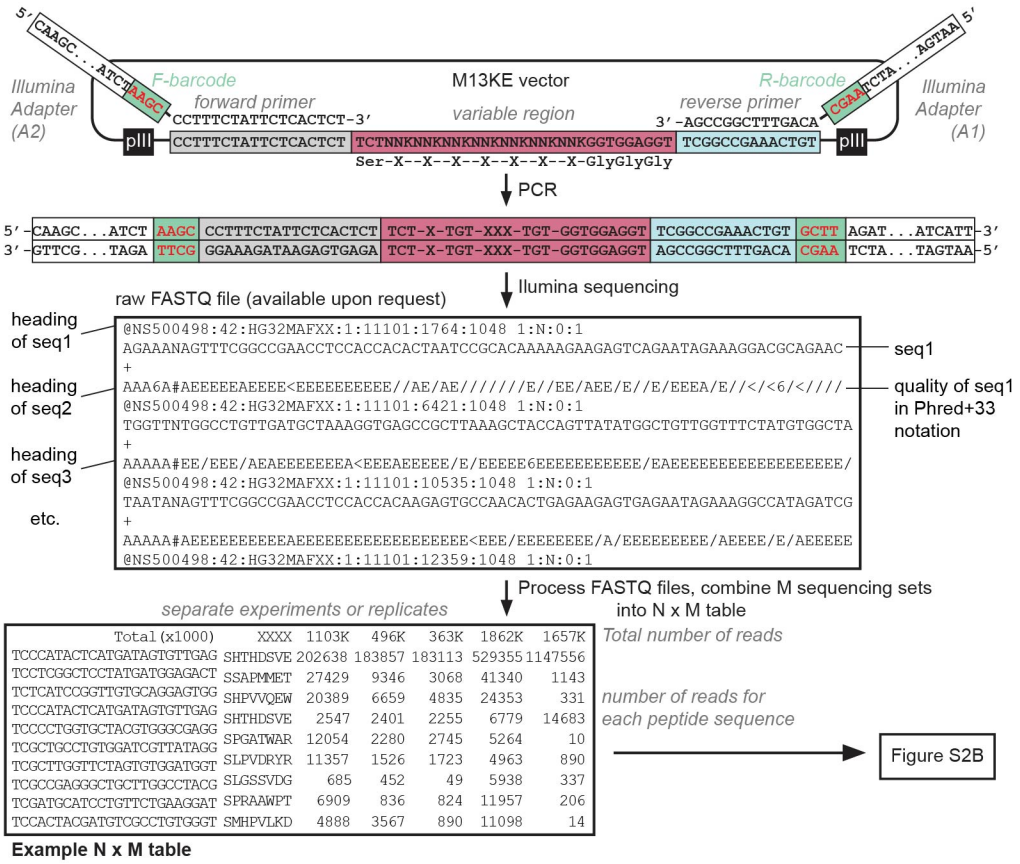
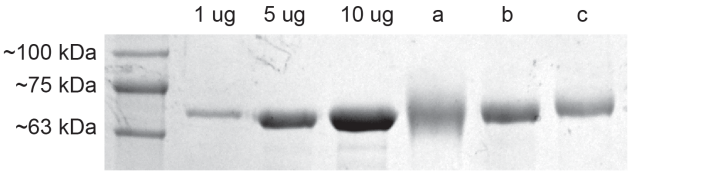
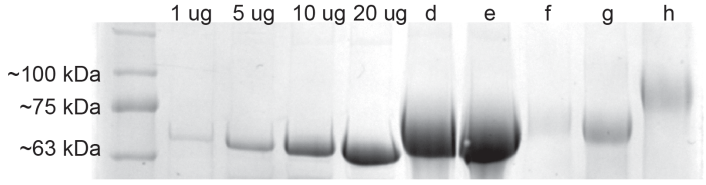
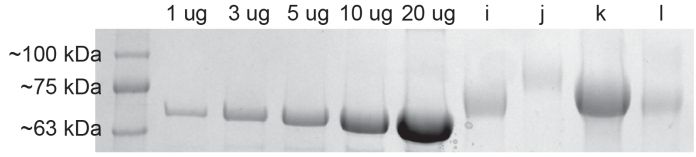


Figure S14. Analysis flow of the deep-sequencing data.

Ligand	# of conjugation	Amount (μg)	Gel position
Ara_6	4	432	e
Ara_6	9	328	d
Ara_6	17	241	k
Ara_6-ANSSFAPK squaramide	2	308	m
Ara_6-ANSSFAPK squaramide	5	36	i
Ara_6-ANSSFAPK squaramide	10	38	j
Ara_6-TTYVVNPK squaramide	4	218	o
Ara_6-TTYVVNPK squaramide	12	45	h
Ara_6-DAHATLRK squaramide	2	396	n
Glu-ANSSFAPK squaramide	2	82	b
Glu-ANSSFAPK squaramide	7	63	g
Glu-ANSSFAPK squaramide	11	50	l
Glu-TTYVVNPK squaramide	2	102	a
Glu-DAHATLRK squaramide	3	160	c
Glu-DAHATLRK squaramide	5	21	f







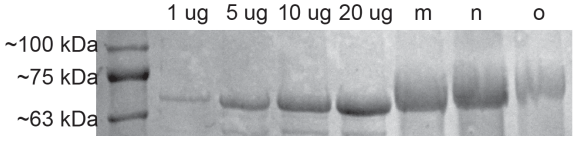


Figure S15. SDS-PAGE to quantify BSA-conjugates
Squaramide-BSA conjugates or Ara_6-BSA conjugates were quantified using SDS-PAGE with fixed amounts of BSA loadings (1 μg, 5 μg, 10 μg) as internal controls.

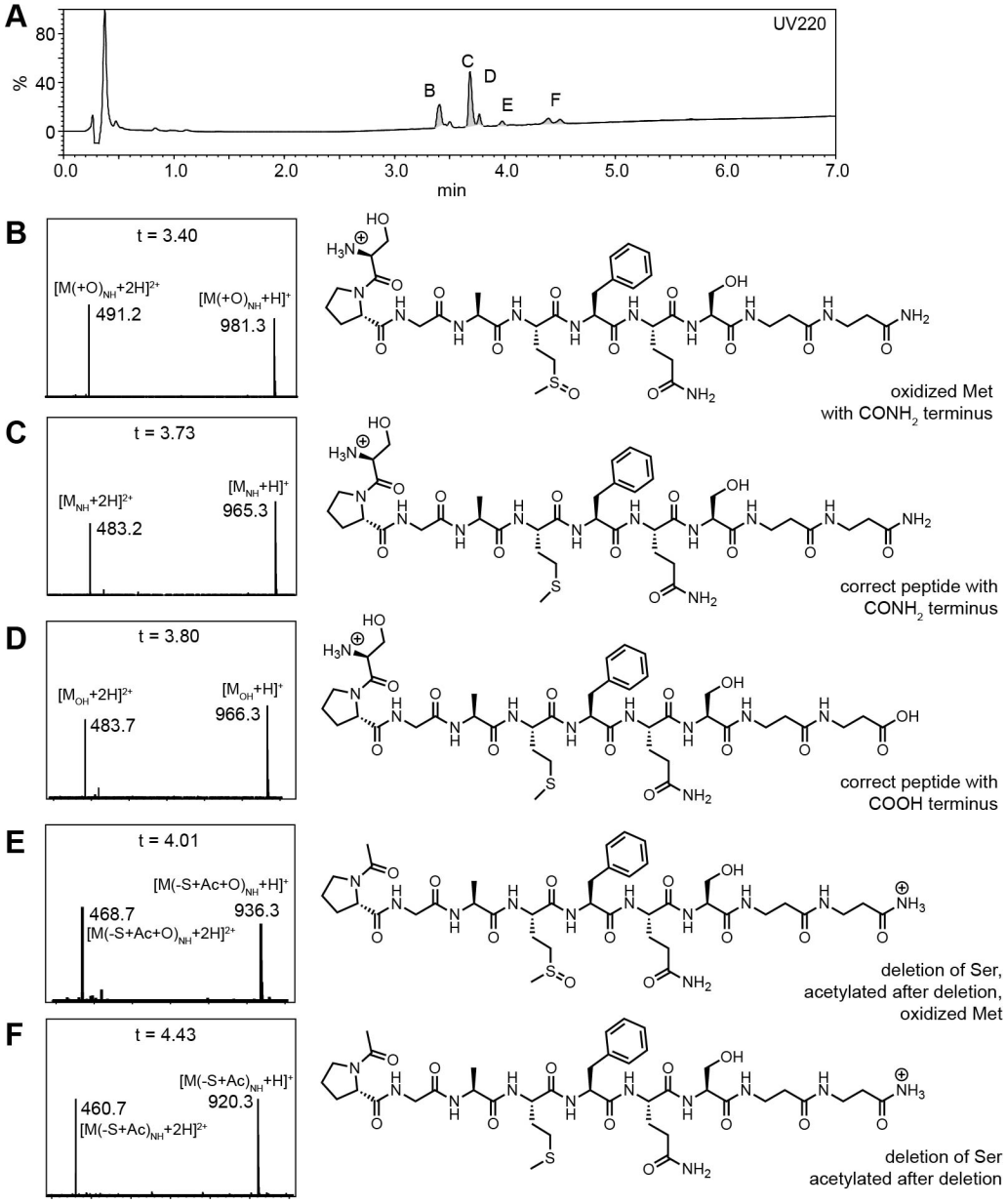


Figure S16. UPLC-MS to characterize peptide Ser-Pro-Ala-Gly-Ala-Met-Phe-Gln-Ser-βAla-βAla synthesized on paper.

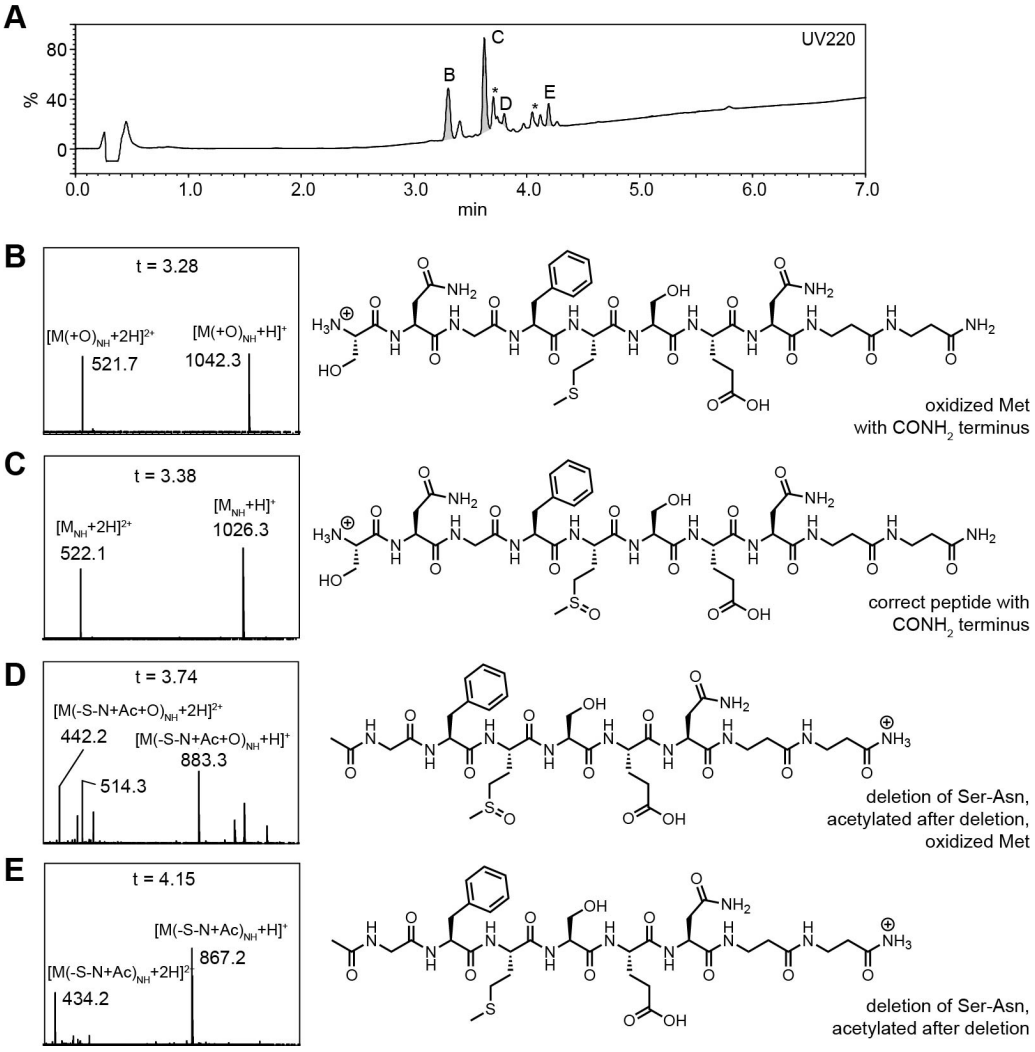


Figure S17. UPLC-MS to characterize peptide Ser-Asn-Gly-Phe-Met-Ser-Glu-Asn-βAla-βAla synthesized on paper.

(A) UPLC chromatogram of peptide synthesized on paper. (B)-(E) Identified peaks, associated masses, and the identification of the compound. *: unidentified peaks.

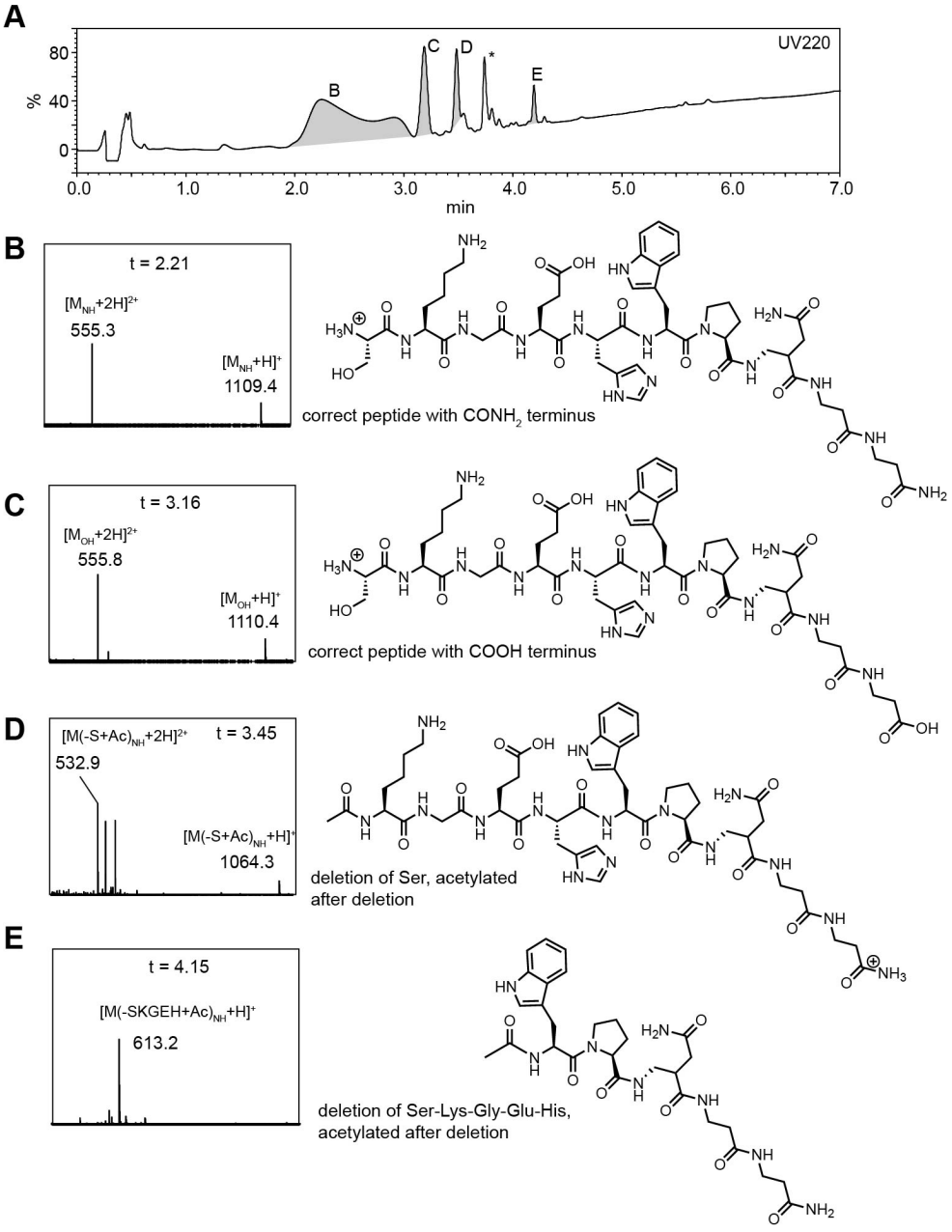


Figure S18. UPLC-MS to characterize peptide Ser-Lys-Gly-Glu-His-Trp-Pro-Gln-βAla-βAla synthesized on paper.

(A) UPLC chromatogram of peptide synthesized on paper. (B)-(E) Identified peaks, associated masses, and the identification of the compound. *: unidentified peaks.

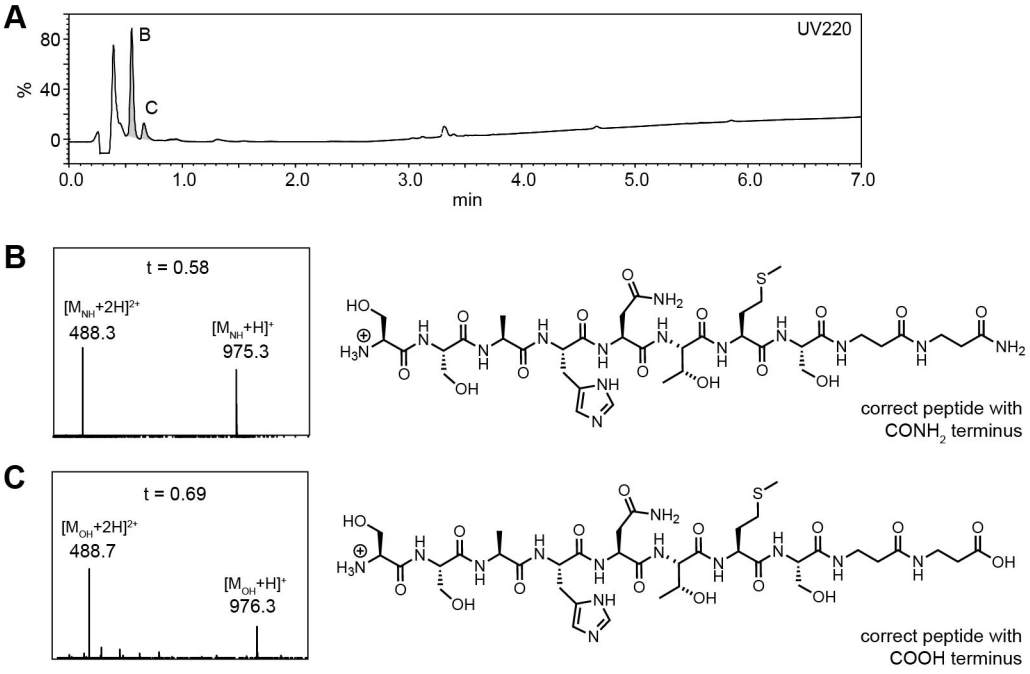


Figure S19. UPLC-MS to characterize peptide Ser-Ser-Ala-His-Asn-Thr-Met-Ser-βAla-βAla synthesized on paper.

(A) UPLC chromatogram of peptide synthesized on paper. (B)-(E) Identified peaks, associated masses, and the identification of the compound.

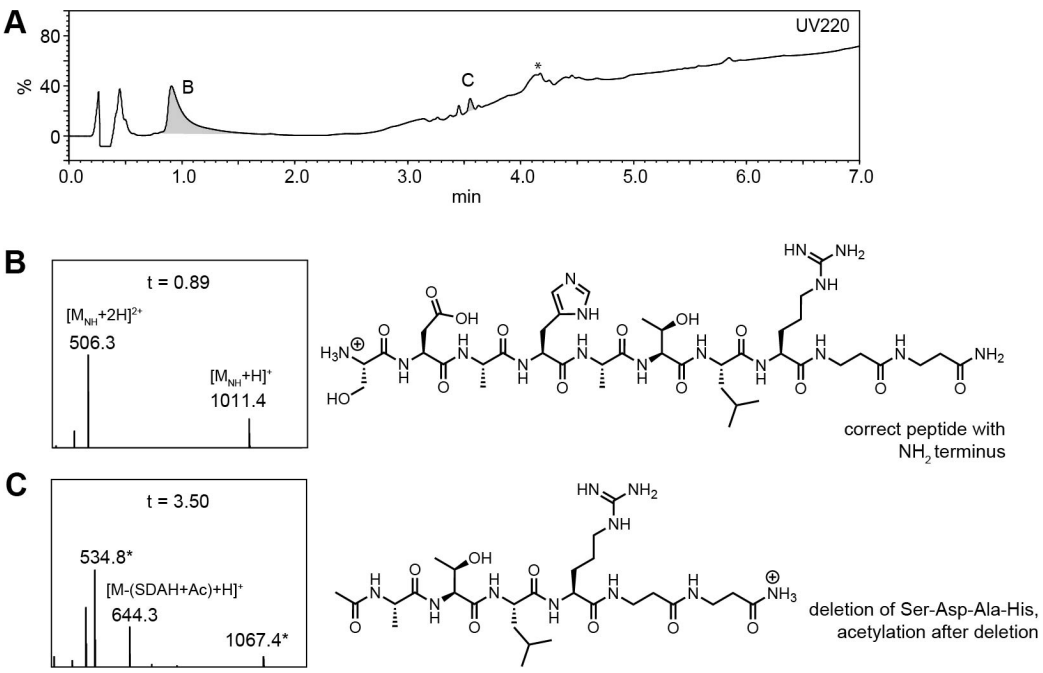


Figure S20. UPLC-MS to characterize peptide Ser-Asp-Ala-His-Ala-Thr-Leu-Arg-βAla-βAla synthesized on paper.

(A) UPLC chromatogram of peptide synthesized on paper. (B)-(E) Identified peaks, associated masses, and the identification of the compound. *: unidentified peaks.

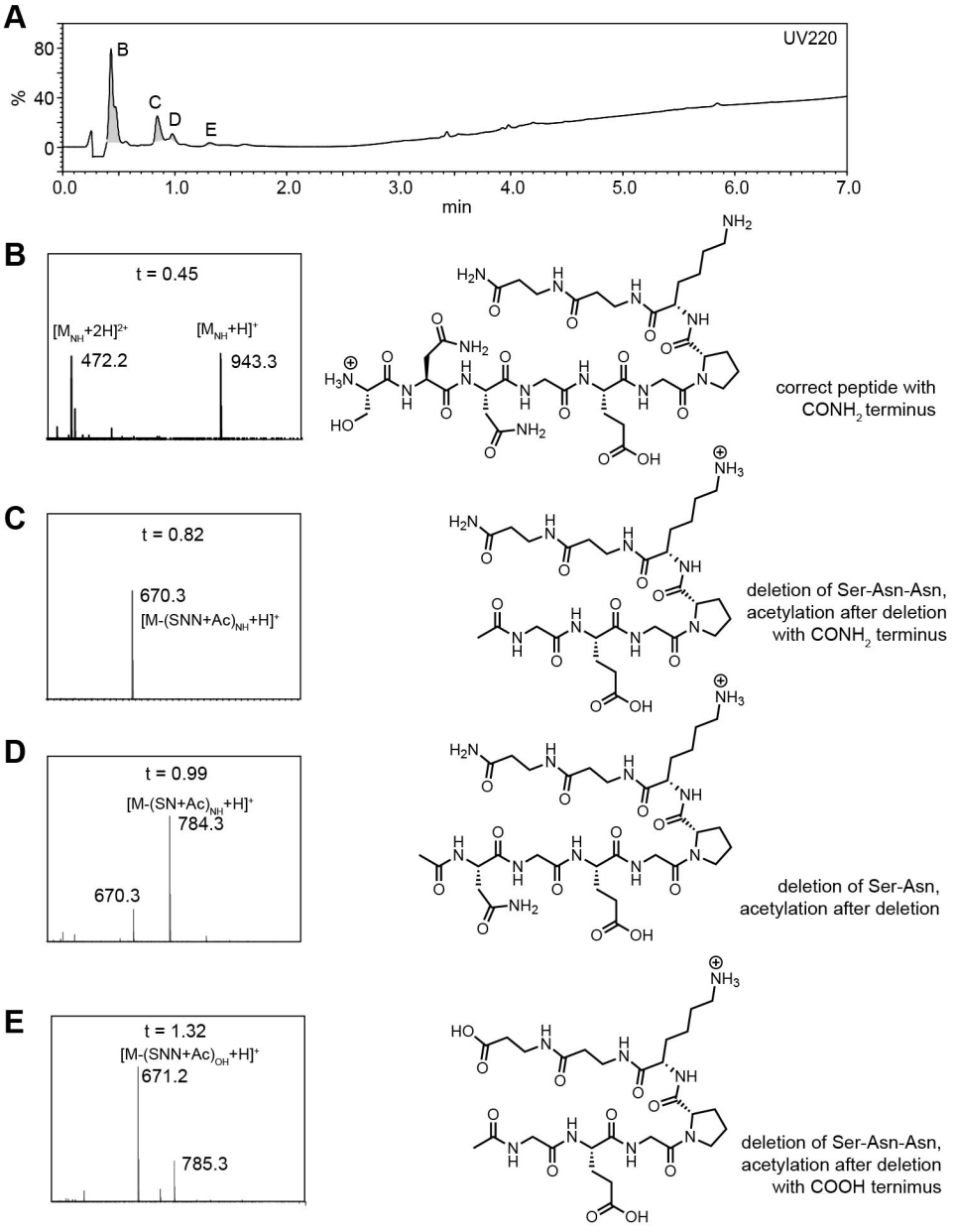


Figure S21. UPLC-MS to characterize peptide Ser-Asn-Asn-Gly-Glu-Gly-Pro-Lys-βAla-βAla synthesized on paper.

(A) UPLC chromatogram of peptide synthesized on paper. (B)-(E) Identified peaks, associated masses, and the identification of the compound.

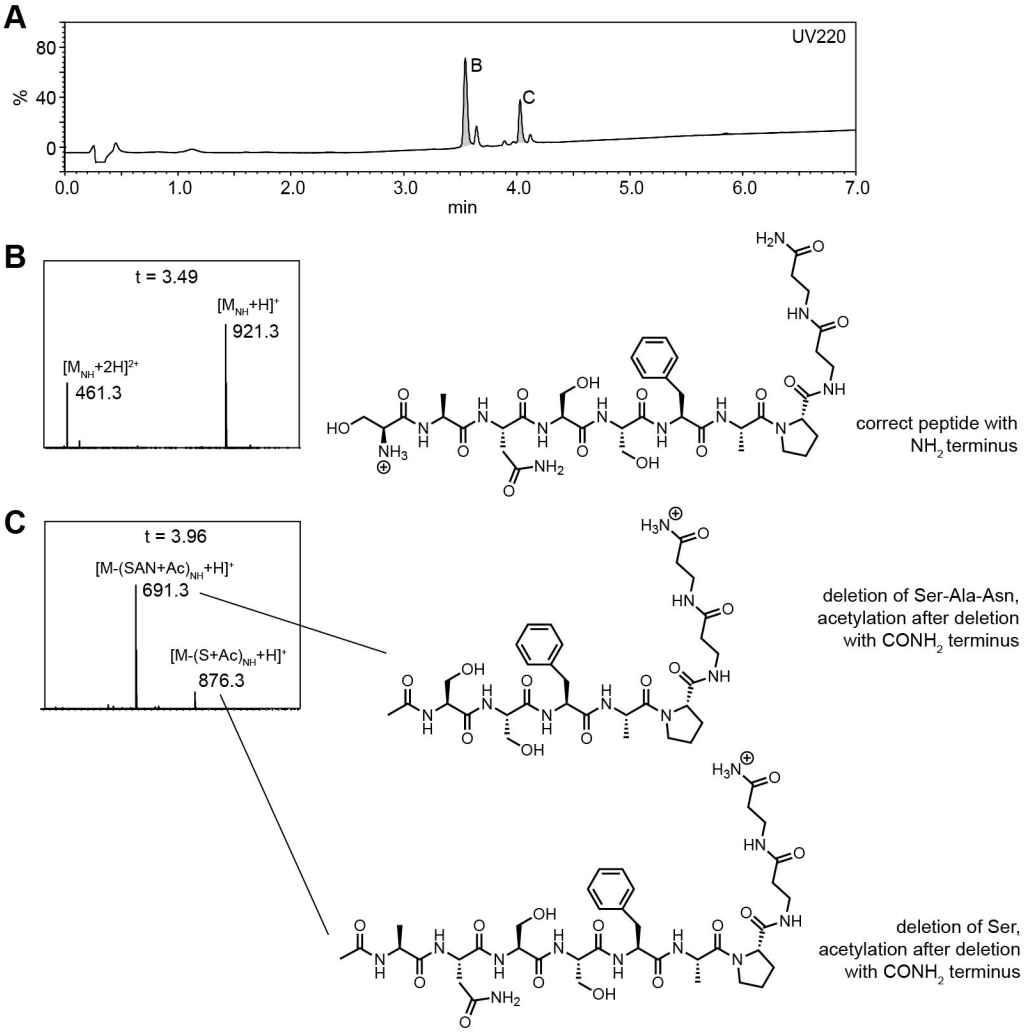


Figure S22. UPLC-MS to characterize peptide Ser-Ala-Asn-Ser-Ser-Phe-Ala-Pro-βAla-βAla synthesized on paper.

(A) UPLC chromatogram of peptide synthesized on paper. (B)-(E) Identified peaks, associated masses, and the identification of the compound.

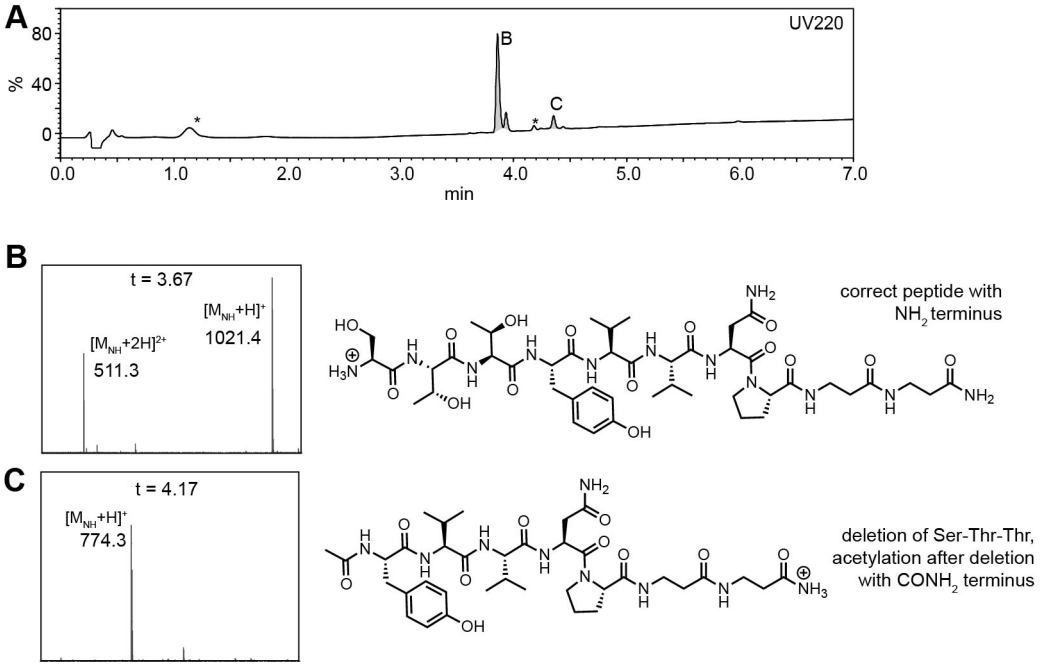


Figure S23. UPLC-MS to characterize peptide Ser-Thr-Thr-Tyr-Val-Val-Asn-Pro-βAla-βAla synthesized on paper.

(A) UPLC chromatogram of peptide synthesized on paper. (B)-(E) Identified peaks, associated masses, and the identification of the compound. *: unidentified peaks.

Materials and general information

PBS contains 10 mM sodium phosphate dibasic, 2 mM potassium phosphate monobasic, 137 mM sodium chloride and 2.7 mM potassium chloride with pH of 7.4 after preparation. MOPS buffer contains 50 mM MOPS, 150 mM NaCl, 2 mM $CaCl_2$ with pH adjusted to 7.4 after preparation. Borate buffer contains 500 mM boric acid with pH adjusted to 9.0 after preparation. Production of mAb IgG of CS-35, 906.4321 and F_{ab} of CS-35 and 906.4321 was previously reported.[1] Aminooxy-biotin (#10009350) was purchased from Cayman Chemical. 3,4-Diethoxy-3-cyclobutene-1,2-dione (#310778) was purchased from Sigma-Aldrich. Characterization of glycopeptides and squaramides were performed by UPLC-MS. Characterization of glycopeptide-BSA conjugates were performed by MALDI. Microarrays were printed at Engineering Arts LLC (Phoenix, Arizona) using Schott NEXTERION® Slide E and piezoelectric non-contact printing. All solutions used for phage work are sterilized either by autoclave or by filter sterilization (0.22 μm).

1. Solid-phase peptide synthesis

Peptides were synthesized on Prelude X (Protein Technologies Inc.). Standard Fmoc-protected amino acids, HBTU and Rink Amide AM resin were purchased from ChemPrep. Poly-Prep® chromatography columns (10 mL) were purchased from Bio-Rad.

2. RP-HPLC

For large scale synthesis, RP-HPLC were performed on Waters HPLC system equipped with a Waters 1525 EF binary pump, a Waters FlexInject manual injector (dual mode) and a Waters 2489 tunable UV detector. SymmetryPrep™ C18 semi-preparative column (19 × 50 mm, particle size 5 μm, pore size 100 Å) and XBridge BEH Amide OBD Prep column (19 × 250 mm, particle size 5 μm, pore size 130 Å) were used for purification of peptides and N-terminal oxidized peptides at a typical flow rate of 12 mL/min. For analytical run, Symmetry® C18 analytical column (4.6 × 75 mm, particle size 3.5 μm, pore size 100 Å) was used at a typical flow rate of 1 mL/min on a LC/MS system (Agilent 1100 LC/MSD). HPLC traces were monitored with UV detection at 220 nm and 254 nm.

3. Synthesis

The synthesis of Ara_6-OCH_3[2] were reported previously. ^{1}H NMR spectra of peptides and Ara_6-containing glycopeptides were acquired on Agilent/Varian VNMRS 700 MHz spectrometers in D_2O (referenced to external acetone at δ_H 2.225 ppm).

Detailed procedures for ligands search

1. Generation of Ara_6-X_7 and Gal-X_7 phage-displayed peptide library

N-SerX_7 phage-displayed peptide library (complexity: 3×10^8 pfu) was generated according to the referred protocol.[3] Prior to the chemical modification, the phage library (2×10^{12} pfu) was dialyzed extensively against 4 L PBS (4 °C, pH 7.4, 4X buffer changes over 24 h, 10K MWCO) to remove the storage buffer which contains 50% (v/v) glycerol. The phage library (100 μL, ~10^{11} pfu/mL) was oxidized with 0.006 mM sodium periodate (by adding 1 μL of 6 mM solution in MQ water) on ice for 5 min. The oxidation was quenched with 0.05 mM glutathione (by adding 1 μL of 50 mM solution in MQ water) at RT for 10 min. To monitor the oxidation, a small portion of the oxidized library was treated with aminooxy-biotin and captured with biotin-capture assay as described in a previously published method.[4] Typically, 40-60% of the fractions of phage library were successfully oxidized.

The oxidized library was distributed into two equal portions of 0.05 mL, and they were treated with 1 mM aminooxy-hexasaccharide (by adding 0.05 mL of 2 mM solution in 200 mM anilinium acetate buffer, pH 4.7) and 1 mM aminooxy-galactose (by adding 0.05 mL of 2 mM solution in 200 mM anilinium acetate buffer, pH 4.7) respectively. The reaction mixtures were incubated for 1 h at RT, after which, the excess of reagents was removed by two buffer exchanges in Amicon® Ultra-4 Centrifugal filter units (10K MWCO) using MOPS buffer. Each exchange was performed at on Heraeus™ Multifuge™ X3R Centrifuge, 4700 rpm for 15 min. The recovered phage, Ara_6-X_7 and Gal-X_7, were used for panning experiments. To quantify the reaction efficiency, right after the oxime ligation, a small portion of the library was treated with aminooxy-biotin and captured with biotin-capture assay as described in a previously published method.[4] Typically, 60-80% of the oxidized fractions of phage library were successfully modified with the reagents.

2. Selection of chemically-modified phage library against CS-35

First round of selection. 8 wells of a 96-well polystyrene plate were coated with a solution of CS-35 (100 μL, 40 μg/mL) in PBS overnight at 4 °C. After coating, the plate was rinsed with washing solution (10 × 300 μL, 0.1% (v/v) Tween-20 in MOPS buffer) using 405™ Touch Microplate Washer (BioTek). The selection of Ara_6-X_7 library against CS-35 (denote as screen **A1**) was performed in 4 replicates. The control selections, i.e., Gal-X_7 against CS-35 (**B1**), Ara_6-X_7 against blank wells (**C1**), were performed in parallel with screen **A1**. Specifically, the solutions of Ara_6-X_7, or Gal-X_7 were added into the corresponding wells (100 μL/well, ~10^9 pfu/well). After incubating for 1 h at RT, the unbound phage was rinsed with the washing solution (10 × 300 μL) using the plate washer. Phage remaining in the well was eluted for 9 min at RT by adding 200 μL of glycine elution buffer (0.2 M glycine-HCl, pH 2.2, 1 mg/ml BSA). The elution buffer was then transferred into a 1.5-mL microcentrifuge tube and immediately neutralized with 30 μL of 1 M Tris-HCl (pH 9.1). The recovered phage libraries–12 samples in total—were amplified for subsequent round of panning and for deep sequencing.

Second round of selection. Amplified phage recovered from screen **A1** in the first round was modified with Ara_6-ONH_2 or Gal-ONH_2 as described in "*Generation of Ara_6-X_7 and Gal-X_7 phage-displayed peptide library*." 12 wells of a 96-well polystyrene plate were coated with a solution of CS-35 (100 μL, 40 μg/mL) in PBS overnight at 4 °C. After coating, the plate was rinsed with washing solution (10 × 300 μL, 0.2% (v/v) Tween-20 in MOPS buffer) using the plate washer. The second round of selection of Ara_6-X_7 library against CS-35 (denote as screen **A2**) was performed in 3 replicates. The control selections, i.e.,

Gal-X_7 against CS-35 (**B2**), Ara_6-X_7 against blank wells (**C2**), and SX_7 against blank wells (**D2**), were performed in parallel with screen **A2**. Firstly, modified libraries (Ara_6-X_7, Gal-X_7) and unmodified library (SX_7) (100 μL ~10^9 pfu/well) were incubated with 12 blank wells for 1 h at RT. The supernatant from each blank well was transferred to the 9 wells that were coated with CS-35. After incubation for 1 h at RT, the unbound phage was rinsed with the washing solution (10 × 300 μL) using the plate washer. Phage remaining in the well was eluted for 9 min at RT by adding 200 μL of glycine elution buffer (0.2 M glycine-HCl, pH 2.2, 1 mg/ml BSA). The elution buffer was then transferred into a 1.5-mL microcentrifuge tube and immediately neutralized with 30 μL of 1 M Tris-HCl (pH 9.1). The recovered phage libraries—12 samples in total—were amplified for deep sequencing.

3. Phage Amplification and PCR of library DNA

The eluted phage from each round of selection was amplified separately by adding the eluate (24 × 0.23 mL) into 6 mL of ER2738 culture (1:100 dilution of overnight culture). The phage and bacterial mixtures were incubated for 4.5 h at 37 °C with vigorous shaking. The cultures were centrifuged (15 min, 4700 rpm) at 4 °C to pellet the bacterial cells. Aliquots of the supernatant from each replicate of screen **A1** (4 × 4.5 mL) were pooled into a sterilized centrifuge tube and precipitated with PEG/NaCl. This pooled library served as the input for the second round of selection. The remaining supernatant (24 × 1.5 mL) from each amplified eluate was poured into individual fresh tube. The ssDNA of the amplified phage was extracted using QIAprep spin M13 kit (Qiagen, #27704) according to manufacturer's instructions.

The ssDNA was then converted to Illumina-compatible dsDNA through PCR amplification using reverse barcoded primers and forward barcoded primers.[5] Forward (F) and reverse (R) primer sequences, 5' → 3':

```
F:
5'-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTXXXXCC
TTTCTATTCTCACTCT-3'

R:
5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTXXXXACAGT
TTCGGCCGA-3'
```

The XXXX in the primer sequence denotes four-nucleotide-long barcodes used to trace multiple samples in one Illumina sequencing experiment.

Briefly, the ssDNA (24 samples, 150 ng each) was amplified in a total volume of 50 μL with 1× Phusion® buffer (NEB), 1 mM each dNTPs, 0.5 μM forward barcoded primer, 0.5 μM reverse barcoded primer, and 0.5 μL Phusion® High-Fidelity DNA Polymerase (NEB). PCR was performed using the following thermo cycler program: a) 95 °C 30 s, b) 95 °C 10 s, c) 60.5 °C 15 s, d) 72 °C 30 s, e) repeat b)–d) for 25 cycles, f) 72 °C 5 min, g) 4 °C hold. The dsDNA fragments from the PCR were quantified by running at 2% (w/v) agarose gel in Tris-Borate-EDTA buffer at 100 volts for ~40 min using a low molecular weight DNA ladder as a standard (NEB, #N3233S). The dsDNA fragments (24 samples, 20 ng per sample) were pooled together and purified on E-Gel® SizeSelect™ 2% agarose gel (Invitrogen, #G6610-02). The desired band corresponding to ~200 bp with reference to the ladder was collected with RNAse-free water and the concentration was determined by Qubit® Fluorimeter (Invitrogen, #Q32851) using manufacturer's protocol. The purified dsDNA with Illumina compatible adapters was sequenced using the

Illumina HiSeq platform (The Donnelly Sequencing Centre at The Donnelley Centre for Cellular and Biomolecular Research, University of Toronto).

4. Analysis of the Deep-sequencing Data.

Analysis was adapted from Ng *et al.*[6] and Matochko *et al.*[7] with minor modifications and the MatLab scripts used to process raw FASTQ data were described in previous publications.[5] Briefly, the scripts use regular expressions to identify sequencing barcodes at the front and back end of the read (the `XXXX` sequence), and isolate the library region (crimson box, Figure S14), which is located between two constant adapter regions (grey box, Figure S14). Sequences and their copy numbers in each sequencing experiment were stored as N × M table, where N is the total number of unique sequences and M is total number of experiments and replicates. An example of N × M table is shown in Figure S14.

Identification of significantly enriched sequences was performed similar to procedure described in Ng *et al*[6] and Tjhung *et al*.[5] For each sequence, the average normalized frequency at which the sequence appeared in each sequencing set was calculated; then, the enrichment ratio was calculated using average frequency in test experiment divided by the average frequency in the control experiment. P-value between the replicates of test and those of control was calculated using a two-sided unequal variance t-test. Hit sequences were defined as those that had $p<0.05$ and ratio≥ 4 for all control experiments and were presented as heat map (Figure S2B).

Synthesis of ligands

1. Solid-phase peptide synthesis

Rink Amide AM resin (200 mg, 0.91 mmol g^{-1}, 0.18 mmol) was weighed into 40 mL reaction vessel. CH_2Cl_2 (5 mL) was added to the dried resin for swelling. After 10 min, the solvent was drained. The resin was washed with DMF (2 × 5 mL) and then deprotected with 20% (v/v) piperidine in DMF (5 mL) for 3 min. The deprotection was repeated for another 3 min using fresh 20% (v/v) piperidine in DMF (5 mL). The resin was washed with DMF (4 mL) for 1 min, followed by CH_2Cl_2 wash (4 × 4 mL). Fmoc-protected amino acid (1 mmol, 5.5 eq.) and HBTU (1 mmol, 5.5 eq.) dissolved in DMF (3 mL) were added to the resin, followed by DIPEA (2 mL, 2 mmol, 11 eq.). After 15 min of agitation with nitrogen and heating at 50 ^{o}C, the reagents were drained the resin was washed with DMF (6 × 4 mL). The Fmoc-deprotection, amide coupling, and washing steps were repeated consecutively to elongate the sequence up to Fmoc-Ser(tBu)-OH, the *N*-terminal residue. After final Fmoc-deprotection, the resin was washed with DMF (6 × 4 mL), followed by CH_2Cl_2 (5 × 4 mL). The resin was dried in reaction vessel and then transferred to Poly-Prep® chromatography column. A cleavage cocktail (2 mL) containing TFA/H_2O/phenol/TIPS [85/5/5/5 (v/v/w/v)] was added to the dried resin. The column was left on a rocker for 4 h to cleave the peptide. The flow through from the column was collected and the resin was rinsed with TFA (1 mL). The combined cleavage mixture was added dropwise to cold diethyl ether (20 mL) in a centrifuge tube. The mixture was incubated on ice for 30 min. The precipitates were centrifuged for 5 min at 3000 rpm. Supernatant was decanted and the precipitates were washed with cold diethyl ether (10 mL). The centrifugation and washing steps were repeated for another two cycles. The precipitates were air-dried. Typical yield: 50–150 mg.

2. Purification of crude peptide

Crude peptide (50 mg) was dissolved in MQ water (0.5 mL); and if peptide did not dissolve, acetonitrile was added in small aliquots (0.05 mL) until the solution appeared clear. The solution was injected into a semi-preparative RP-HPLC system with C18 column following the gradient attached to each HPLC trace of peptides (S46-S77) using solvent A (MQ water, 0.1% v/v TFA) and solvent B (MeCN, 0.1% v/v TFA) at a flow rate of 12 mL/min. The fractions corresponding to the main peak were collected. MeCN was removed by evaporation under reduced pressure. The aqueous solution was lyophilized to yield the peptide as white fluffy powder (10-30 mg).

3. Representative example of large scale synthesis of Glc-peptide and Ara-peptide conjugates

SDAHATLR (8 mg, 9 μmol, 1 eq.) was dissolved in 200 mM MOPS (0.5 mL, pH 7.0) in a 1.5-mL microcentrifuge tube containing sodium periodate (2.3 mg, 12 μmol, 1.2 eq.). The reaction mixture was incubated for 1 h at RT. Aminooxyglucose (2.7 mg, 13.5 μmol, 15 eq.) dissolved in 200 mM anilinium acetate (0.5 mL, pH 4.7) was added to the reaction mixture and the mixture was incubated for 1 h at RT, followed by injection into a RP-HPLC system using gradients described for each glycopeptide (S46-S77). Lyophilization yielded Glc-DAHATLR as white fluffy powder (11% isolated yield). The purity and identity of the product were confirmed with UPLC-MS.

4. Representative example of small scale synthesis of Ara_6-peptide conjugates

SANSSFAP (10 mg, 13 μmol, 1 eq.) was dissolved in 200 mM MOPS (0.5 mL, pH 7.0) in a 1.5-mL microcentrifuge tube containing sodium periodate (3.3 mg, 12 μmol, 1.2 eq.). The reaction mixture was incubated for 1 h at RT. The reaction mixture was injected to RP-HPLC and lyophilization yielded

glyoxyl-ANSSFAP as white fluffy powder. The purity and identity of the product were confirmed with UPLC-MS.

The glyoxyl-ANSSFAP (0.3 μmol, 1 eq.) was dissolved in 60 μL MQ water in a 1.5-mL microcentrifuge tube. Aminooxy-Ara_6 (0.3 mg, 0.36 μmol, 1.2 eq.) was dissolved in 18 μL MQ water and added to the microcentrifuge tube, followed by the addition of 500 mM anilinium acetate (19 μL, pH 4.7). The reaction mixture was incubated for 1 h at RT, then injected into a LC/MS system (Agilent 1100 LC/MSD) using symmetry C18 analytical column (WAT066224). Purification was carried out using gradients described for each Ara_6-peptide conjugate (S46-S77). Lyophilization yielded product as translucent film and the yield was determined by NMR using internal standard (S72-S76). The purity and identity of the product was determined with UPLC-MS.

5. NMR to determine yield of Ara_6-peptide conjugates

Each Ara_6-peptide conjugate was dissolved in 150 μL H_2O, 70 μL of which was added to 630 μL D_2O and a known amount of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) was spiked in the sample as an internal standard. The 1H spectra were collected on a 700 MHz, 4 channel Agilent instrument, fitted with a HCN Z-gradient Cold probe at 27 ^oC. VNMRJ 4.2A was used to acquire the spectra. Relaxation delay was set at 5 s, which included 2 s of presaturation of the residual water peak (HOD)). The acquisition time was 3 s and the spectral window was set to 12 ppm, 11 to -1 ppm. A reduced flip angle of 30 degrees was applied to ensure accurate integration.

6. Representative example of small scale synthesis of squaramides

Ara_6-ANSSFAPK (0.1 mg, 0.08 μmol, 1 eq.) was dissolved in 40 μL MQ water in a 0.6-mL microcentrifuge tube. 3,4-Diethoxy-3-cyclobutene-1,2-dione (0.02 mg, 0.16 μmol, 2 eq.) dissolved in 40 μL ethanol was added. Saturated $NaHCO_3$ was added to adjust the pH of the reaction mixture to ~8.0 as measured with universal pH-paper. Reaction was incubated for 1 h at RT, followed by purification on an analytical RP-HPLC system using gradients attached to each squaramide (S46-S77) to yield 67% (isolated yield) after lyophilization.

7. Synthesis of BSA-conjugates

BSA was dialyzed against MQ water (3 buffer changes over 12 h) and lyophilized prior to conjugation. To produce high density conjugates, lyophilized BSA (~200 μg, 1 eq.) was reacted with squaramide (50-100 μg, 15-25 eq.) in borate buffer (10 μL, 500 mM, pH 9.0) for 3 days at RT. The reaction mixture was subjected to dialysis against 2 L of MQ water (3 buffer changes over 8 h) at 4 ^oC. For medium and low density conjugates, lyophilized BSA (400-500 μg, 1 eq.) was reacted with squaramides (8-12 eq. or 4-7 eq. respectively) in borate buffer (10 μL, 500 mM, pH 9.0) for 3 days at RT. Conjugates were lyophilized after dialysis, and the conjugation efficiency was determined using MALDI (Figure S9-S10). The quantity of each conjugate was determined using SDS-PAGE gel with an internal standard BSA (Figure S15).

8. Synthesis of glycopeptides on paper.

Synthesis of peptides on paper (SyntArray) is reported by Deiss *et al.*[8] The identity of the synthesized peptides were characterized with UPLC-MS (Figure S16-S23). Prior to ligation of carbohydrates, solid-ink barriers were printed on the peptide-containing areas on the array. 5 min incubation of the array in 120 ^oC oven generated a wax-patterned paper array that contained barrier resistant to aqueous solutions.

To oxidize the peptides, arrays were first submerged in carbonate buffer (10 mL, pH = 8.0) to wash residual TFA from the peptide synthesis. The array was then immersed in 5 mM sodium periodate in carbonate buffer (5 mL, pH 8.0) for 30 min at RT in the dark. Arrays were then transferred to a container with 200 mM anilinium acetate buffer (5 mL, pH 4.7) and carefully dried with Kimwipe from the face of array that was not printed with peptides. Aminooxy-derivative was spotted (5 mM, 3 μL) on each peptide-containing area and the array was incubated in a humidity chamber for 1.5 h at RT. The array was then washed with MQ water, air-dried, and stored at -20 ^{o}C until use.

9. Optimization of oxidation and oxime ligations on paper.

The array was first oxidized using conditions described above, and then different concentrations of aminooxy-rhamnose (0 mM, 1 mM, 2 mM, 5 mM) were spotted. The reaction was incubated in a humidity chamber for 1.5 h at RT. The array was then submerged into fluorescein thiosemicarbazide (FTS) (1 mM in 200 mM anilinium acetate, pH 4.7) to convert all remaining un-ligated but oxidized peptides to FTS, giving rise to an orange color on paper. The relative amounts of rha-peptide and FTS-peptide were monitored using UPLC-MS to determine reaction completion.

10. Kinetic control of the oxidation of the peptide SSAHNTMS

We noted that the rate of Met oxidation is >10 times slower than the rate of cleavage of N-terminal Ser.[4, 9] Therefore, to minimize the oxidation of Met in the peptide SSAHNTMS during oxidative cleavage of Ser by $NaIO_4$, we kinetically controlled the oxidation with 60 mM $NaIO_4$ for 5 min at 4 ^{o}C. This condition did not produce any detectable amount of oxidized Met species. On the other hand, higher concentration (1 mM or 10 mM) of $NaIO_4$ used in a previous report[4] for a preparative synthesis of peptide glyoxals resulted in either partial or complete oxidation as confirmed by LC-MS (Figure S6).

Peptide SSAHNTMS (4 mg, 5 μmol, 1 eq.) was dissolved in 2.25 mL ice-cold PBS. Ice-cold solution of NaIO4 (15 μL, 400 mM, 1.2 eq.) was added for 7 sec, and the reaction was quenched by the addition of GSH (50 μL, 400 mM, pH 8.0, 4 eq.). Reaction was vortexed and immediately injected into semi-preparative RP-HPLC with Amide column. The fraction containing the product was lyophilized to yield white solid (isolation yield: 10%).

Paper-based assay to determine ligands binding

1. Fluorescent labeling of mAb CS-35.

0.47 mg/mL of CS-35 was dialyzed against 4 L PBS (4 buffer changes over 24 h) in a 10K MWCO dialysis cassette. 25 mg/mL of TRITC dissolved in DMF was added to dialyzed CS-35 and the reaction was buffered with carbonate buffer to pH ~9 as measured by universal pH paper. Reaction mixture was incubated for 2 h at RT on a rocker, followed by a buffer exchange with PBS in Amicon Ultra-4 centrifugal filter units to remove excess TRITC. The concentration of conjugated antibody was determined from UV absorbance at 280 nm and 550 nm, taking into account the absorbance contribution from the covalently-bound TRITC dye at 550 nm and the correction factor (CF) adjusted for the amount of absorbance at 280 nm caused by the dye (CF=0.35) (ThermoScientific TR0031.7):

$$IgM\ (M) = \frac{A_{280} - 0.34 \times A_{550}}{\varepsilon_{280}}, \varepsilon_{280} = 21{,}000\ M^{-1}cm^{-1}$$

2. Paper-based fluorescence binding assay.

Glycopeptide or peptide arrays were submerged in blocking buffer (2% BSA, 0.4% Tween-20 in PBS) for 1 h at RT. In a small petri dish, the arrays were incubated with 5 μg/mL of TRITC-labeled CS-35 in blocking buffer (10 mL) for 30 min at 37 ^{o}C. The arrays were then washed with 0.4% Tween-20 in PBS (2 × 10 mL). Fluorescent images were obtained using Typhoon FLA 9500 (GE Healthcare Life Science), using λ_{ex} = 532 nm and long-pass emission filter LPG (≥575 nm).

ESI-MS binding measurements

The dissociation constants (K_D) for F_{ab} binding to ligands were determined using the direct ESI-MS binding assay. The sample solutions were prepared from aqueous stock solutions of F_{ab} and ligand with known concentrations. Aqueous ammonium acetate was added to the sample solutions to yield a final buffer concentration of ~200 mM. Direct ESI-MS binding assay was performed on a Synapt G2S Q-IMS-TOF mass spectrometer (Waters, Manchester, UK) equipped with nanoflow ESI (nanoESI) source. NanoESI was performed by applying a voltage of ~1 kV to a platinum wire inserted into the nanoESI tip, which was produced from a borosilicate glass capillary (1.0 mm o.d., 0.68 mm i.d.) pulled to ~5 μm o.d. using a P-1000 micropipette puller (Sutter Instruments, Novato, CA). The source temperature and gas flow rates were 60 ^{o}C and 2 mL/min, respectively. The cone, trap and transfer voltages were 30V, 5V, and 2V, respectively. MassLynx software (version 4.1) was used for data acquisition and processing. The values of K_D for binding of a monovalent protein (P) to monovalent ligand (L) were calculated from the abundance ratio of L-bound to free P ions (i.e., R), measured by ESI-MS:

$$R = \frac{Ab(PL)}{Ab(P)} = \frac{[PL]}{[P]} \qquad (\mathbf{1})$$

$$K_D = \frac{[P][L]}{[PL]} = \frac{[L]_0 - \frac{R}{R+1}[P]_0}{R} \qquad (\mathbf{2})$$

where $[P]_0$ and $[L]_0$ are initial concentrations of protein and ligand, correspondingly.[10]

Two batches of CS-35 Fab were used for the binding measurements of all ligands. All ligands (Table S1, **1**-**13**) were tested against the first batch, while ligands **3**, **6**, **9**, **12**, **13** were tested additionally using the second batch. Previously reported K_D for CS-35 F_{ab} to Ara_6-OCH_3 (ligand **13**) was 1.6×10^5 M^{-1}. However, using the first batch of CS-35 F_{ab}, the K_D we obtained drifted from literature value, possibly due to prolonged storage of this batch. Representative mass spectra for aqueous solutions of the first batch of CS-35 F_{ab} were provided in Figure S7-S8. The four distinct F_{ab} species (**a**, **b**, **c**, **d**) resulted from digestion of the IgG mAb with papain were detected at charge states ranging from +13 to +15. The masses are: (**a**) 48.25 kDa, (**b**) 47.90 kDa, (**c**), 48.92 kDa, and (**d**) 49.20 kDa. The major glycoform **a** was used to calculate binding constants for all ligands. The second batch of CS-35 Fab yielded a different major species at mass of 47.90 kDa. This mass signal was used to calculate binding constants for ligands **3**, **6**, **9**, **12**, **13**. The binding constants given in Table S1 for ligands **3**, **6**, **9**, **12**, **13** were calculated by averaging the binding constants obtained from both batches of CS-35 F_{ab}.

For 906.4321 F_{ab}, three distinct species (**a**, **b**, **c**) were detected at charge states ranging from +13 to +15. The masses are: (**a**) 48.6 kDa, (**b**) 49.7 kDa, and (**c**) 50.0 kDa. The major glycoform **a** was used for calculating binding constants.

Glycopeptide microarray printing and analysis

1. Preparation of microarrays

Squaramide-BSA conjugates were aliquoted, lyophilized and sent to Engineering Arts LLC (Phoenix, Arizona) for piezoelectric non-contact printing. Briefly, the conjugates spotting solutions were prepared at 0.1 mg/mL in buffer (1:10 PBS + 0.005% v/v Triton-X100). Each antigen was microarrayed in triplicate spots onto epoxysilane-coated glass slides (Schott, NEXTRERION® Slide E) with 360 picoliters per spot producing 140-160 μm diameter spots. The spot-to-spot spacing was 0.5 mm (column pitch) and 0.4 mm (row pitch). On each epoxy slide, 24 replicate arrays were spotted to provide 72 identical spots for each conjugate and control. The resulting arrays were sealed in a slide box and stored at -20 ^{o}C until use.

2. Testing antibody binding to glycopeptides on microarray

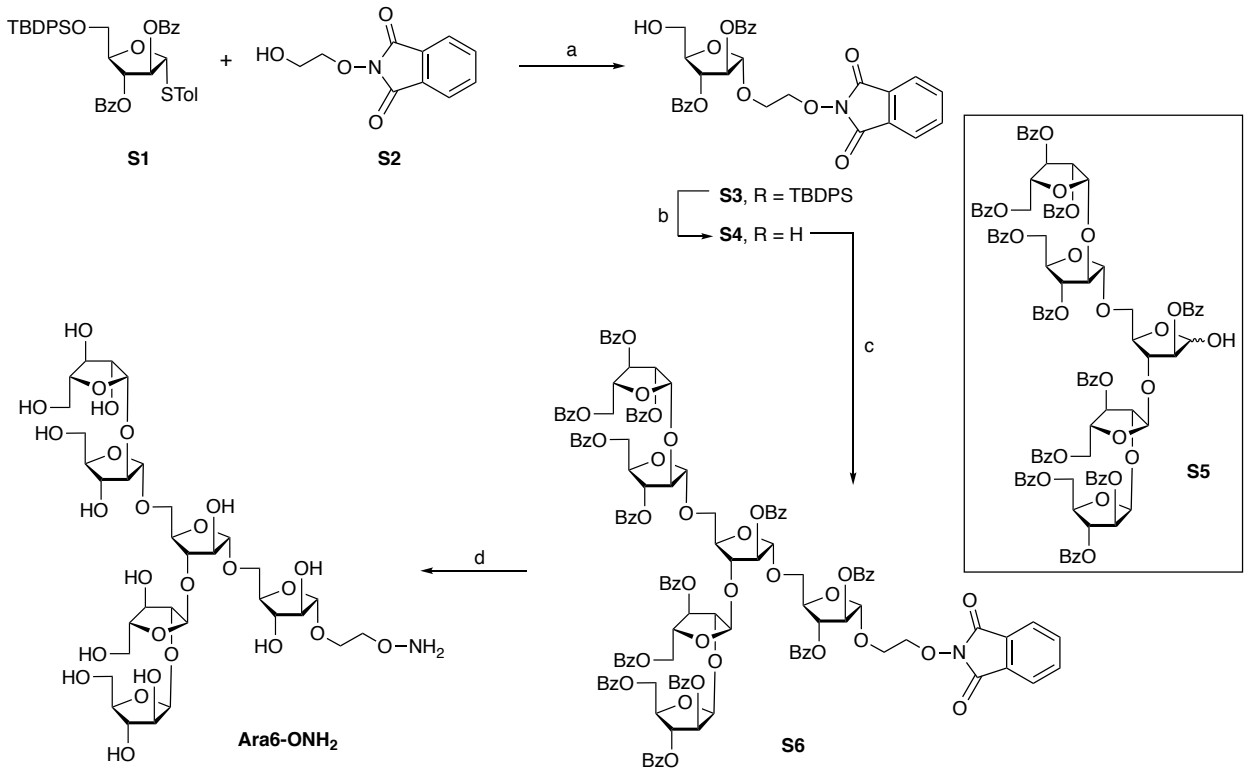
Array was equilibrated to RT prior to testing. The slide was attached to ProPlate® multi-array slide system, 24 chambers (Grace Bio-Labs, 246824). Each chamber on the slide was washed with washing buffer (2 × 200 μL, 0.05% v/v Tween-20 in PBS), followed by 1 × rinse with PBS. The chambers were then blocked with blocking buffer (200 μL, 1% w/v BSA in PBS), sealed with plastic cover, and incubated for 1.5 h at 37 ^{o}C. Blocking buffer was then aspirated, followed by the addition of varying concentrations (30, 10, 3, 1, 0.3, 0.1, 0.03, 0.01 μg/mL) of primary monoclonal antibodies (CS-35, 906.4321 IgG) pre-saturated with blocking buffer (30 min at 37 ^{o}C). The antibody-ligand solution was incubated for 1 h at 37 ^{o}C. After rinsing with washing buffer (10 × 200 μL) and followed by PBS (2 × 200 μL), goat anti-mouse IgG pre-saturated with blocking buffer at 1:500 dilution was added to each chamber and incubated for 1 h at RT. After rinsing with washing buffer (10 × 200 μL), the slide was detached from the gasket and the array was submerged in 50 mL of washing buffer for 5 min. The array was then transferred to 50 mL MQ water and submerged for 5 min. Array was dried using a minicentrifuge, scanned immediate, and then stored at room temperature in a microarray slide storage box.

3. Array scanning and data extraction

Arrays were scanned on Molecular Devices GenePix 4000B (Molecular Biology Service Unit, University of Alberta) using 532 nm laser and the photomultiplier tube (PMT) voltage setting at 350 V. The location of each spot on the array was outline using the automatic grid feature of the GenePix Pro 4.1 software and manually adjusted where necessary. The signal intensity from each spot was calculated after subtraction of the local background intensity. Dose-response curves were generated using MatLab script that fit the Hill equation to each conjugate binding to each monoclonal antibody.

General method for the synthesis of aminooxy derivatives

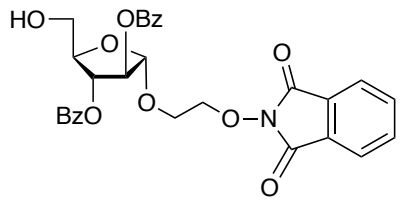
Reactions were carried out in oven-dried glassware. All reagents used were purchased from commercial sources and were used without further purification unless noted. Reaction solvents were purified by successive passage through columns of alumina and copper under nitrogen. Unless stated otherwise, all reactions were carried out at room temperature under a positive pressure of argon and were monitored by TLC on Silica Gel 60 F_{254} (0.25 mm, E. Merck). Spots were detected under UV light or by charring with acidified p–anisaldehyde solution in EtOH. Column chromatography was performed using Silicycle UltraPure silica gel (SiliaFlash® P60, 40–63 μm, Cat# R12030 B). The ratio between silica gel and crude product ranged from 100:1 to 20:1 (w/w). Optical rotations were measured at 22 ± 2 °C. ^{1}H NMR spectra were recorded at 600 MHz or 500 MHz and chemical shifts were referenced to TMS (0.0, $CDCl_3$). ^{1}H data were reported as though they were first order. ^{13}C NMR (APT) spectra were recorded at 125 MHz and ^{13}C chemical shifts were referenced to internal $CDCl_3$ (77.23, $CDCl_3$). Organic solutions were concentrated under vacuum at < 50 °C. Electrospray mass spectra were recorded on samples suspended in mixtures of THF with CH_3OH and added NaCl.



Scheme 1. Synthesis of Ara_6-ONH_2.

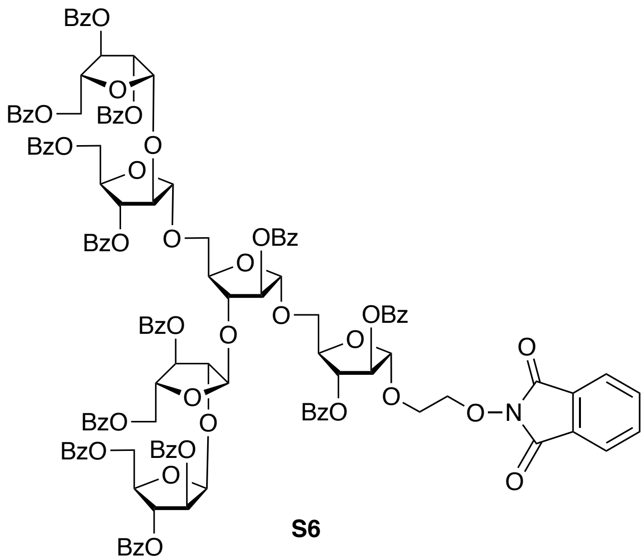
a) NIS, AgOTf, CH_2Cl_2; b) HF·pyridine, THF, pyridine, 87% over two steps; c) **S5**, CCl_3CN, DBU, CH_2Cl_2; then **S4**, TMSOTf, CH_2Cl_2; then BzCl, pyridine, 72%; d) H_2NNH_2, H_2O, 58%.

1. 2-*N*-(Phthalimido)-oxyethyl 2,3-di-*O*-benzoyl-α-D-arabinofuranoside (S4).



Thioglycoside **S1**[11] (0.57 g, 0.81 mmol), and alcohol **S2**[12] (0.18 g, 0.88 mmol) were dried under vacuum in the presence of P_2O_5 for 6 h before CH_2Cl_2 (16 mL) was added followed by powdered 4 Å molecular sieves (0.31 g). After stirring for 20 min, the reaction mixture was then cooled to 0 °C and N-iodosuccinimide (0.0.22 g, 0.98 mmol) and silver triflate (25 mg, 0.09 mmol) were added. After stirring the mixture for 50 min at 0 °C, Et_3N was added until the pH of the solution was slightly basic. The reaction mixture was diluted with CH_2Cl_2 (15 mL), filtered through Celite and the filtrate was washed with a saturated solution of $Na_2S_2O_3$ (20 mL), water (25 mL) and brine (15 mL). The organic layer was separated, dried (Na_2SO_4), filtered and concentrated to give monosaccharide **S3** as a syrup [R_f 0.43 (7:3 hexanes–EtOAc)]. The crude material was dried under vacuum for 3 h and then used in the next step without further purification. To an ice-cold solution **S3** in THF–pyridine (5:2, 21 mL) was added 70% HF·pyridine (0.4 mL) dropwise. The solution was then allowed to warm to rt and stirred overnight. The reaction mixture was then poured into a saturated aqueous $NaHCO_3$ solution (50 mL) and extracted with EtOAc (50 mL). The combined organic layer was washed with water (30 mL) and brine (25 mL). The separated organic layer was then dried (Na_2SO_4), filtered and concentrated to a syrupy residue that was purified chromatography (1:1 hexanes–EtOAc) to yield **S4** (0.385 g, 87% over two steps) as a thick syrup. R_f 0.34 (1:1 hexanes–EtOAc); $[\alpha]_D$ −37.9 (c = 0.2, $CHCl_3$); IR (cast film) 1732.5, 1790.3 cm^{-1}. ^1H NMR (500 MHz, $CDCl_3$, δ_H) 8.10–8.00 (m, 4H), 7.81–7.75 (m, 2 H), 7.75–7.68 (m, 2 H), 7.63–7.58 (m, 2 H), 7.52–7.43 (m, 4 H), 5.43–5.39 (m, 2 H), 5.33 (s, 1 H, H-1), 4.54–4.48 (m, 1 H), 4.47–4.39 (m, 2 H), 4.17–4.10 (m, 1 H), 4.05–3.92 (m, 3 H), 2.42 (app t, 1 H, J = 5.5, 6.4 Hz); ^{13}C NMR (125 MHz, $CDCl_3$, δ_C) 166.1(0) (C=O), 166.1(4) (C=O), 163.5 (C=O), 134.5 (Ar), 133.5(3) (Ar), 133.5 (Ar), 130.0 (Ar), 129.9 (Ar), 129.2 (Ar), 129.1 (Ar), 128.8 (Ar), 128.5(2) (Ar), 128.5(1) (Ar), 123.6 (Ar), 105.8 (C-1), 83.8 (C-3), 81.7 (C-2), 77.6 (C-4), 77.3, 65.8 (linker-$\underline{C}H_2$), 62.3. HRMS (ESI) calcd for $(M+Na)^+$ $C_{29}H_{25}NO_{10}Na$: 570.1371. Found: 570.1373.

2. 2-(*N*-Phthalimido-oxy)-ethyl 2,3,5-Tri-*O*-benzoyl-β-D-arabinofuranosyl-(1→2)-3,5-di-*O*-benzoyl-α-D-arabinofuranosyl-(1→5)-[2,3,5-tri-*O*-benzoyl-β-D-arabinofuranosyl-(1→2)-3,5-di-*O*-benzoyl-α-D-arabinofuranosyl-(1→3)]-2-*O*-benzoyl-α-D-arabinofuranosyl-(1→5)-2,3-di-*O*-benzoyl-α-D-arabinofuranoside (S6).

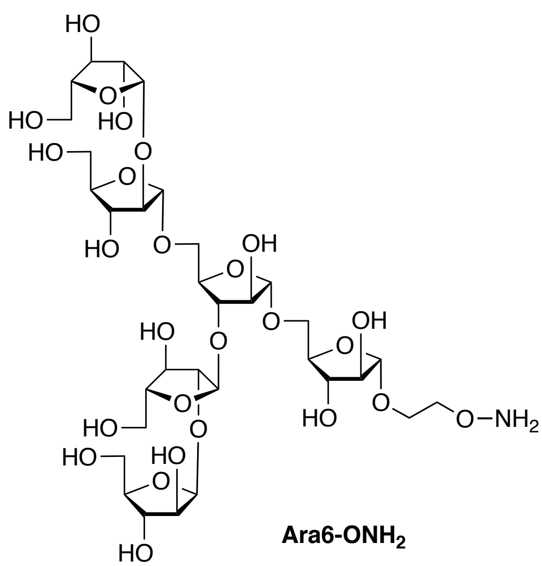


S6

To a solution of the pentasaccharide hemiacetal **S5**[11] (0.19 g, 0.1 mmol) and trichloroacetonitrile (0.2 mL, 2 mmol) in CH_2Cl_2 (5 mL) at 0 °C was added DBU (20 μL). The reaction mixture was stirred at 0 °C for 30 min and then warmed to rt over 30 min. The reaction mixture was directly concentrated to a syrupy residue that was quickly filtered through silica gel (3:2 hexanes–EtOAc containing about 0.1 % Et_3N). The fractions containing the trichloroacetimidate derivative were concentrated, dried under vacuum for 1.5 h and used immediately without further purification. This trichloroacetimidate derivative in CH_2Cl_2 (4 mL) was added to a solution of alcohol **S4** (0.07 g, 0.13 mmol) in CH_2Cl_2 (5 mL) containing 4 Å molecular sieves (0.19 g; stirred already for ~30 min) at −20 °C. A solution of TMSOTf (3 μL, 0.016 mmol) in CH_2Cl_2 (0.06 mL) was added dropwise over 5 min. The reaction mixture was then warmed to −9 °C over ~20 min and Et_3N (0.05 mL) was added. The solution was diluted with CH_2Cl_2 (15 mL) and filtered. The filtrate was concentrated to a syrup that was dried under vacuum for 3 h. As the acceptor alcohol **S4** had the same R_f as the hexasaccharide product, the mixture was benzoylated to ease purification. Thus, the crude mixture of **S4** and **S6** was dissolved in pyridine (4 mL), cooled to 0 °C under argon, and BzCl (0.5 mL, 4.3 mmol) was added dropwise. The reaction mixture was stirred overnight while warming to rt. Methanol (1 mL) was added, the mixture was stirred for 30 min, diluted with CH_2Cl_2 (20 mL), poured into satd aq $NaHCO_3$ soln (25 mL) and extracted with CH_2Cl_2 (20 mL). The combined CH_2Cl_2 layer was washed with water (25 mL), separated, dried (Na_2SO_4) and concentrated to a syrupy residue that was purified by chromatography (3:2 hexanes–EtOAc) to afford **S6** (0.175 g, 72% over two steps) as a thick syrup. R_f 0.17 (3:2 hexanes–EtOAc); IR (cast film) 2095.5, 1723.2 cm^{-1}; ^1H NMR (700 MHz, $CDCl_3$, δ_H) 8.05–7.98 (m, 6 H), 7.98–7.92 (m, 4 H), 7.92–7.84 (m, 8 H), 7.84–7.77 (m, 8 H), 7.71–7.67 (m, 2 H), 7.66–7.53 (m, 4 H), 7.56–7.47 (m, 2 H), 7.44–7.32 (m, 15 H), 7.31–7.14 (m, 18 H), 7.14–7.09 (m, 2 H), 5.92 (dd, 1 H, J = 5.6, 6.5 Hz), 5.87 (dd, 1 H, J = 5.3, 6.2 Hz), 5.69 (d, 2 H, J = 4.8 Hz, 2 × β Ara H-1), 5.50 (dd, 1H, J = 4.7, 6.2 Hz), 5.43 (d, 1 H, J = 4.9 Hz), 5.39–5.34 (m, 3 H), 5.32–5.30 (m, 2 H), 5.29 (dd, 1 H, J = 2.0, 4.4 Hz), 5.26 (d, 1 H, J = 1.2 Hz), 5.22 (s, 1 H), 5.07 (s, 1 H), 4.73 (ddd, 2 H, J

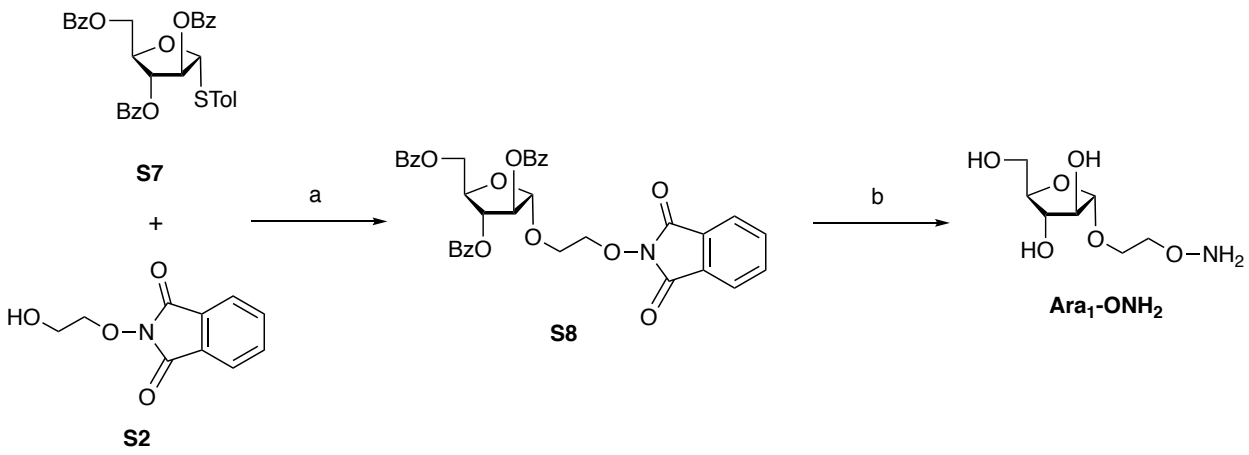
= 4.6, 11.7, 16.6 Hz), 4.63 (ddd, 2 H, *J* = 7.5, 11.7, 13.5 Hz), 4.56 (d, 1 H, *J* = 1.9 Hz), 4.50–4.30 (m, 10 H), 4.31–4.26 (m, 2 H), 4.11 (ddd, 2 H, *J* = 6.6, 11.6, 11.6 Hz), 4.07 (ddd, 1 H, *J* = 2.7, 5.3, 11.9 Hz), 4.02 (dd, 1 H, *J* = 5.6, 11.5 Hz), 3.95–3.88 (m, 3 H), 3.72 (dd, 1H, *J* = 1.9, 11.7 Hz); ^{13}C NMR (175 MHz, $CDCl_3$, δ_C) 165.9(3) (C=O), 165.9(1) (C=O), 165.8 (C=O), 165.7(0) (C=O), 165.7 (C=O), 165.6 (C=O), 165.5 (C=O), 165.4 (C=O), 165.3(3) (C=O), 165.3(1) (C=O), 165.1 (C=O), 163.4 (C=O), 133.1(0) (Ar), 133.1 (Ar), 132.8(1) (Ar), 132.8 (Ar), 129.9(2) (Ar), 129.8(6) (Ar), 129.8(1) (Ar), 129.8 (Ar), 129.7(3) (Ar), 129.7(1), 129.7 (Ar), 129.6(3) (Ar), 129.6 (Ar), 129.3 (Ar), 129.2 (Ar), 129.1(4) (Ar), 129.1 (Ar), 129.0 (Ar), 128.9 (Ar), 128.8(0) (Ar), 128.8 (Ar), 128.7 (Ar), 128.4(9) (Ar), 128.4(5) (Ar), 128.4(2) (Ar), 128.4 (Ar), 128.3(3) (Ar), 128.3 (Ar), 128.2(4) (Ar), 128.2 (Ar), 128.1(4) (Ar), 128.1 (Ar), 123.4, (Ar) 106.5 (C-1), 105.9 (C-1), 105.8 (C-1), 105.1 (C-1), 100.2 (C-1), 100.1 (C-1), 84.9, 84.8, 83.3, 81.8, 81.7, 81.6, 80.6(2), 80.6(0), 79.3, 79.1, 78.1, 78.0, 77.5, 76.7, 76.6, 76.4, 66.3, 65.7(9), 65.7(8), 65.6, 64.4, 64.2. HRMS (ESI) calcd. for $(M+NH_4)^+$ $C_{131}H_{113}N_2O_{41}$: 2369.6813. Found 2369.6841.

3. 2-(Aminooxy)ethyl β-D-arabinofuranosyl-(1→2)-α-D-arabinofuranosyl-(1→5)-[β-D-arabinofuranosyl-(1→2)-α-D-arabinofuranosyl-(1→3)]-α-D-arabinofuranosyl-(1→5)-α-D-arabinofuranoside (Ara$_6$-ONH$_2$).



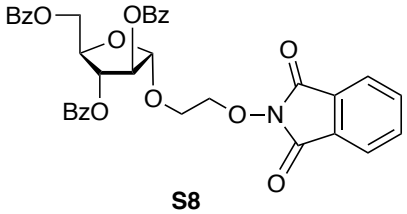
Ara6-ONH$_2$

To compound **S6** (34 mg, 0.014 mmol) was added hydrazine hydrate solution in water (64–65% solution, 1 mL) and the mixture was stirred at rt for 23 h. At this point, the reaction mass was sonicated (4 × 20 seconds) and the mixture was continued to stir at rt. After an additional 4 h, the solution was heated in an oil bath at 42 °C for 3 h and again left to stir at rt for another 15 h. The reaction mixture was directly concentrated on a rotary evaporator (<48 °C) to a syrupy residue that was purified on a G10 size exclusion chromatography (700 MW exclusion limit) using 0.6% HOAc in water as the eluant. The fractions containing the desired compound were pooled and lyophilized to give **Ara$_6$-ONH$_2$** (7.3 mg, 58%) as a foam. HRMS (ESI) calcd. for $(M+H)^+$ $C_{32}H_{56}NO_{26}$: 870.3084. Found: 870.3085.



Scheme 2. Synthesis of Ara_1-ONH_2.
a) NIS, AgOTf, CH_2Cl_2; 96%; b) H_2NNH_2, H_2O, 52%.

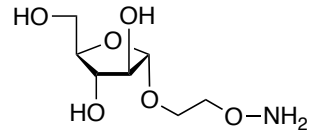
4. 2-*N*-(Phthalimido)-oxyethyl 2,3,5-tii-*O*-benzoyl-α-D-arabinofuranoside (S8).



S8

A suspension of thioglycoside donor **S7**[13] (5.7 g, 10.62 mmol), alcohol **S2**[12] (3.3 g, 15.93 mmol) and powdered 4 Å molecular sieves in dry CH_2Cl_2 (170 mL) was stirred at 0 °C for 30 min. N-iodosuccinimide (3.1 g, 13.81 mmol) and silver triflate (545 mg, 2.12 mmol) were added and then the reaction mixture warmed to rt stirred for 1 h before being neutralized with Et_3N. The solution was filtered through Celite, concentrated and the resulting residue was purified by chromatography (2:1 hexanes–EtOAc) to yield **S8** (6.62 g, 96%) as a white solid: R_f 0.35 (2:1 hexanes–EtOAc); $[\alpha]_D$ –9.4 (*c* = 1.0, $CHCl_3$); IR (cast film) 3063, 3002, 2945, 1790, 1731, 1451, 1270, 1111, 755, 712 cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$, δ_H) 8.06 (dtd, 6 H, *J* = 1.0, 8.0, 23.5 Hz, Ar), 7.80 (dd, 2 H, *J* = 3.0, 5.5 Hz, Ar), 7.71 (dd, 2 H, *J* = 3.0, 5.5 Hz, Ar), 7.61–7.64 (m, 3 H, Ar), 7.49–7.54 (m, 3 H, Ar), 7.43 (t, 2 H, *J* = 7.5 Hz, Ar), 7.31 (t, 1 H, *J* = 7.5 Hz, Ar), 5.57 (d, 1 H, *J* = 4.1 Hz, H-3), 5.42 (d, 1 H, *J* = 1.2 Hz, H-2), 5.40 (s, 1 H, H-1), 4.85–4.87 (m, 1 H, H-4), 4.68–4.71 (m, 2 H, H-5), 4.46–4.57 (m, 2 H, linker-C$\underline{H}_2$), 4.02–4.22 (m, 2 H, linker-C$\underline{H}_2$); ^{13}C NMR (125.7 MHz, $CDCl_3$, δ_C) 176.5 (<u>C</u>=O), 166.2 (<u>C</u>=O), 165.8 (<u>C</u>=O), 165.3 (<u>C</u>=O), 163.4 (<u>C</u>=O), 134.4 (Ar), 133.8 (Ar), 133.5 (Ar), 133.0 (Ar), 130.1 (Ar), 130.0 (Ar), 129.9 (Ar), 129.8 (Ar), 129.2 (Ar), 129.1 (Ar), 128.9 (Ar), 128.5 (Ar), 128.3 (Ar), 123.5 (Ar), 105.9 (C-1), 82.0 (C-3), 81.2 (C-2), 77.3 (C-4), 65.8 (linker-<u>C</u>H_2), 63.7 (C-5, linker-<u>C</u>H_2). HRMS (ESI) calcd for (M+Na) $C_{36}H_{29}NO_{11}$: 651.1741. Found: 651.1745.

5. 2-(Aminooxy)ethyl α-D-arabinofuranoside (Ara_1-ONH_2).

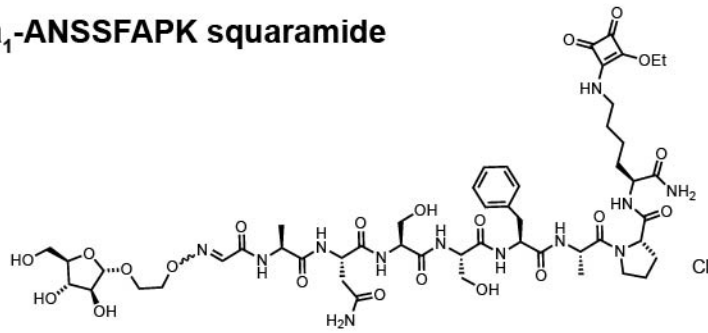


Ara_1-ONH_2

To a stirred solution of **S8** (3.0 g, 4.60 mmol) in CH_3OH (50 mL), hydrazine monohydrate (15 mL) was added dropwise at rt. The reaction mixture was left stirring overnight before filtering through Celite, diluting with water (20 mL), and concentrating. The compound was redissolved in water and concentrated (3 × 20 mL) before purification by chromatography (9:1 CH_2Cl_2–CH_3OH) to yield **Ara_1-ONH_2** (500.5 mg, 52%) as a yellow oil: R_f 0.22 (9:1 CH_2Cl_2–CH_3OH); $[\alpha]_D$ +5.6 (*c* = 1.8, $CHCl_3$); IR (cast film) 3304, 2921, 1593, 1459, 1314, 1219, 1033, 880 cm^{-1}; ^1H NMR (700 MHz, CD_3OD, δ_H) 4.91 (d, 1 H, *J* = 1.4 Hz, H-1), 4.00 (dd, 1 H, *J* = 1.7, 3.5 Hz, H-2), 3.96 (ddd, 1 H, *J* = 3.3, 5.5, 5.5 Hz, H-4), 3.81–3.89 (m, 4 H, linker-C$\underline{H}_2$), 3.75 (dd, 1 H, *J* = 3.3, 11.9 Hz, H-5), 3.62–3.66 (m, 2 H, H-3, H-5); ^{13}C NMR (125.7 MHz, CD_3OD, δ_C) 108.2 (C-1), 84.3 (C-3), 81.9 (C-2), 77.4 (C-4), 74.4 (linker-<u>C</u>H_2), 65.5 (C-5), 61.7 (linker-<u>C</u>H_2). HRMS (ESI) calcd for (M+Na) $C_7H_{15}NO_6$: 209.0899. Found: 209.0906.

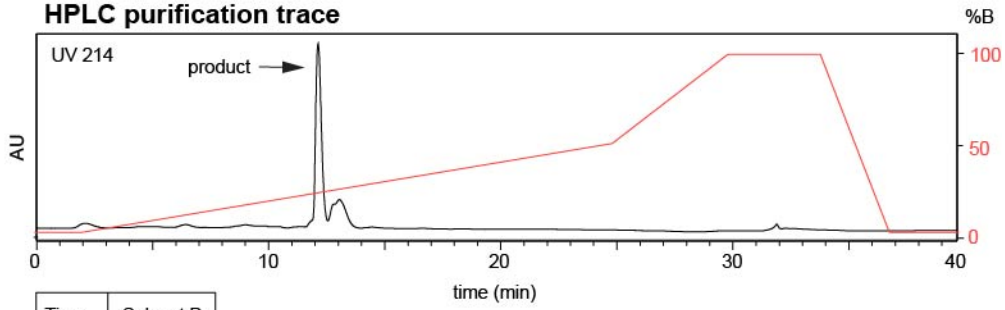
HPLC Purification profile and characterization of synthesized ligands

Ara_1-ANSSFAPK squaramide



Chemical Formula: $C_{51}H_{74}N_{12}O_{21}$
Exact Mass: 1190.51

HPLC purification trace

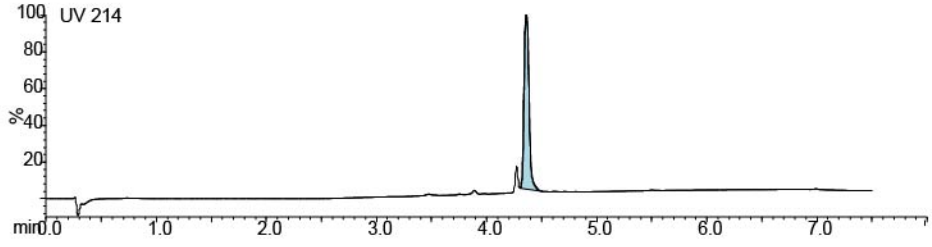


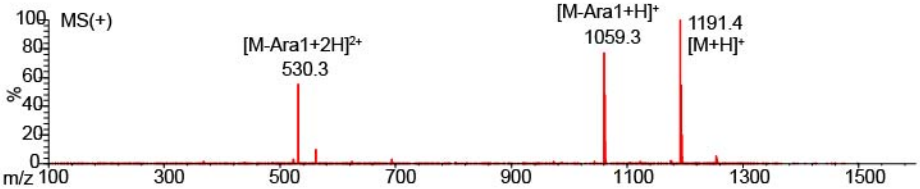
Time (min)	Solvent B (%)
0	2
2	2
25	50
30	100
34	100
37	2
40	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)

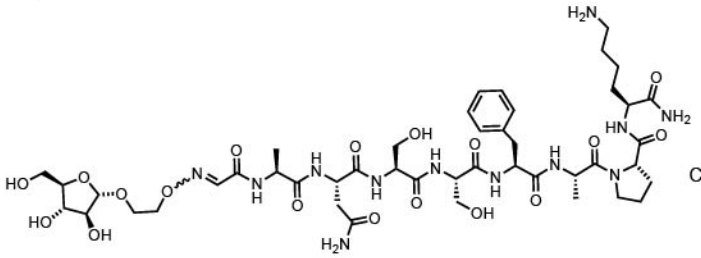
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace



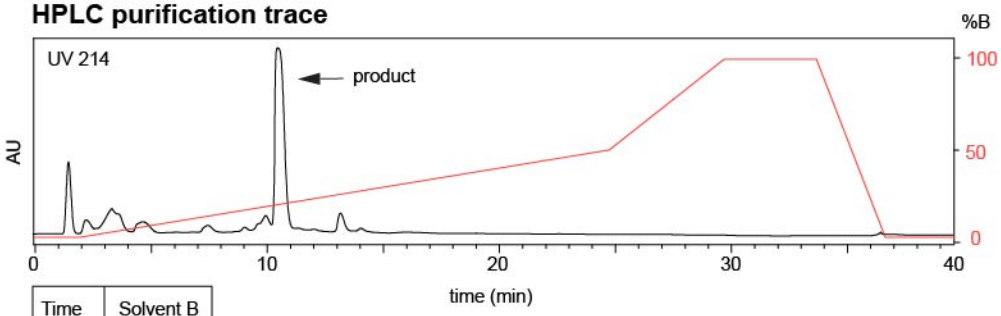


Ara_1-ANSSFAPK



Chemical Formula: $C_{45}H_{69}N_{11}O_{19}$
Exact Mass: 1067.48

HPLC purification trace

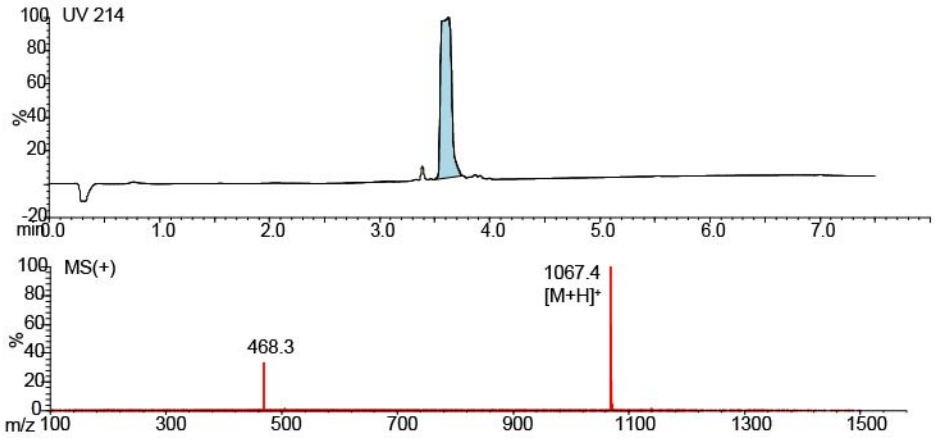


Time (min)	Solvent B (%)
0	2
2	2
25	50
30	100
34	100
37	2
40	2

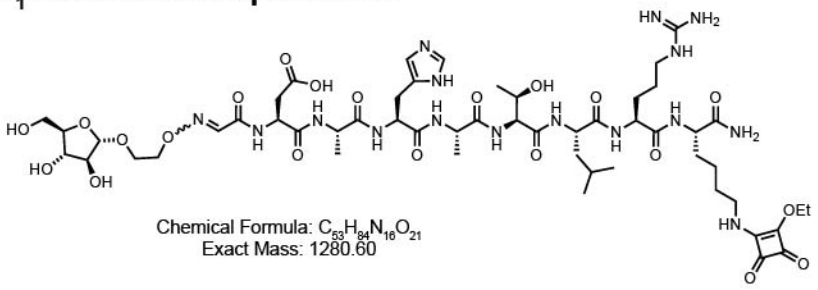
Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

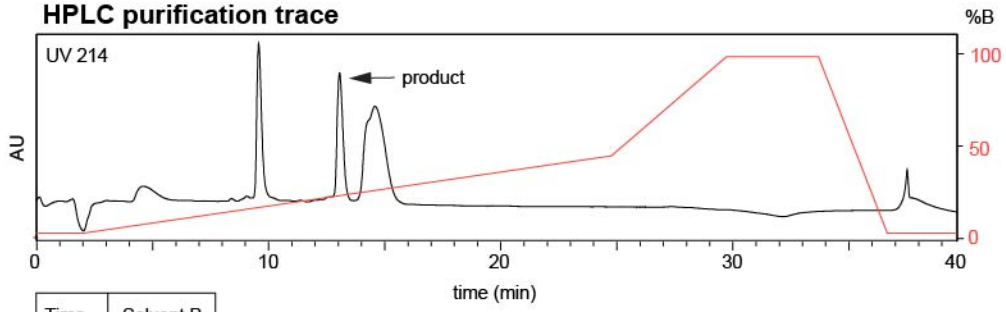


Ara_1-DAHATLRK squaramide



Chemical Formula: $C_{53}H_{84}N_{16}O_{21}$
Exact Mass: 1280.60

HPLC purification trace

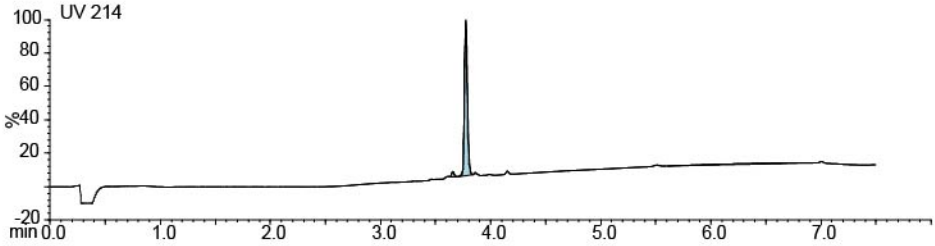


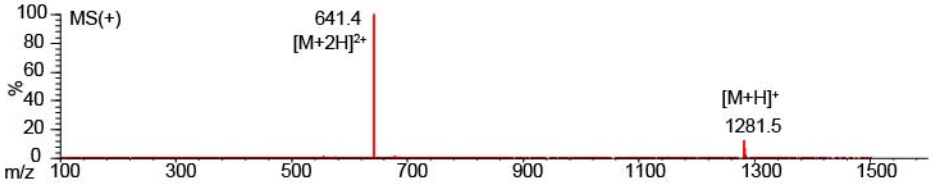
Time (min)	Solvent B (%)
0	2
2	2
25	45
30	100
34	100
37	2
40	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)

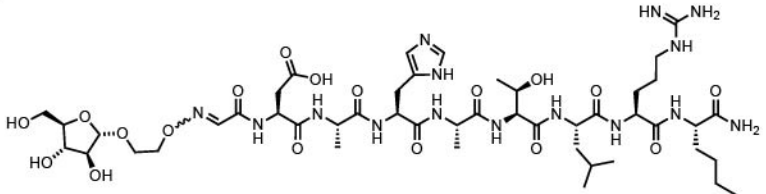
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace



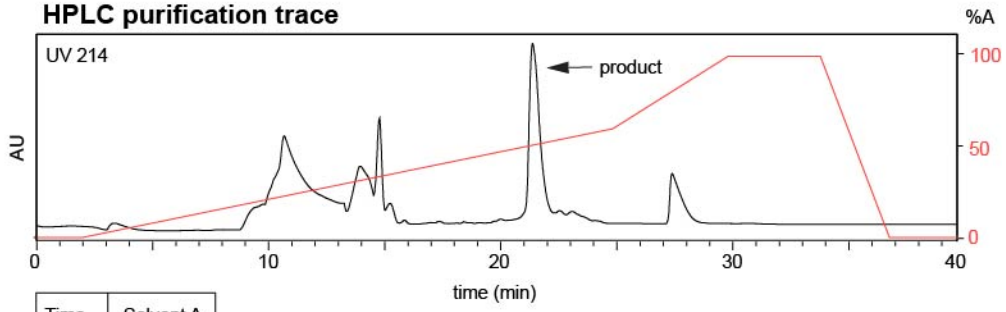


Ara_1-DAHATLRK



Chemical Formula: $C_{47}H_{80}N_{16}O_{18}$
Exact Mass: 1156.58

HPLC purification trace

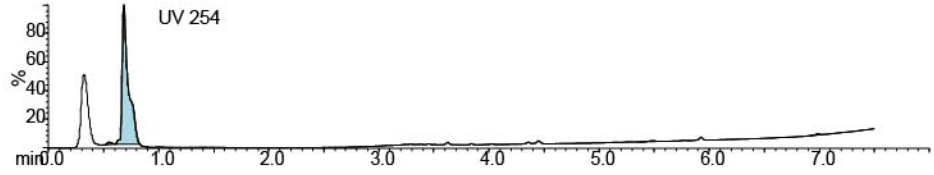


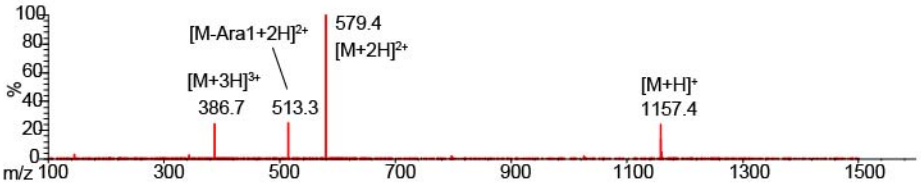
Time (min)	Solvent A (%)
0	0
2	0
25	60
30	100
34	100
37	0
40	0

Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

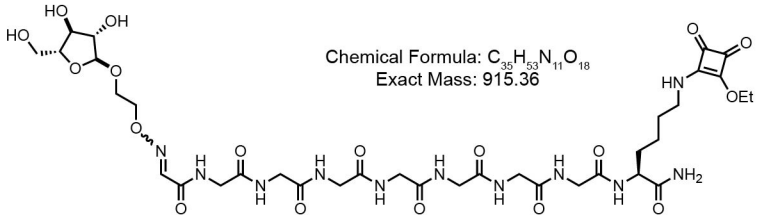
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

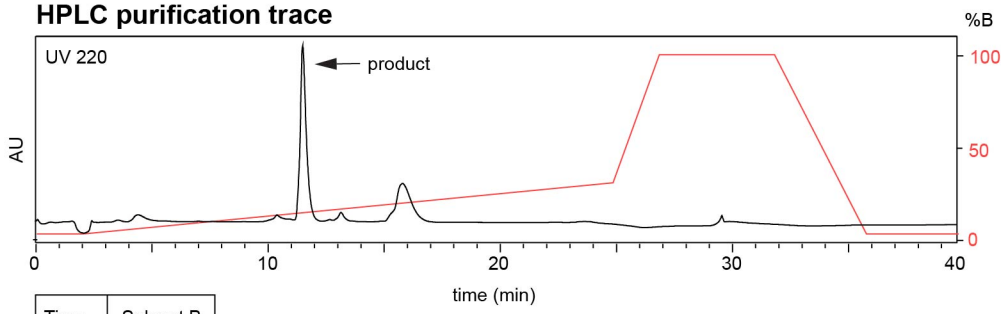




Ara_1-GGGGGGGK squaramide



HPLC purification trace

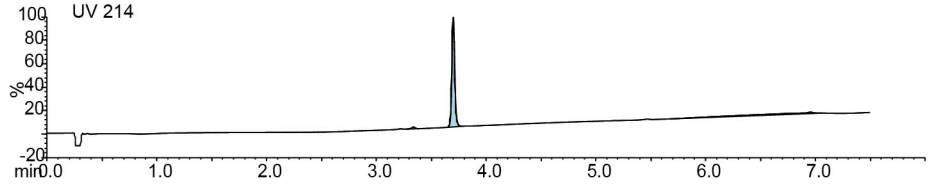


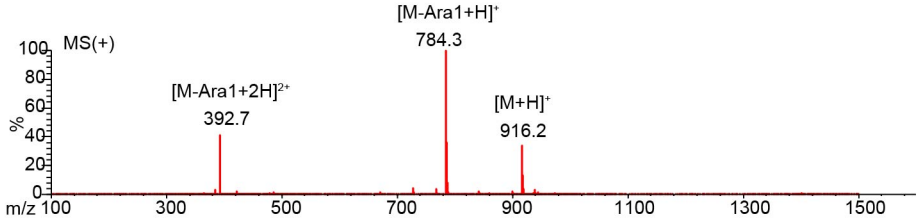
Time (min)	Solvent B (%)
0	2
2	2
25	30
27	100
32	100
36	2
40	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)

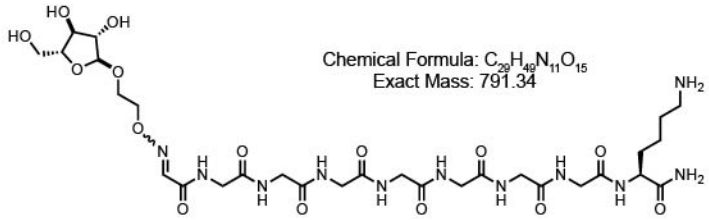
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

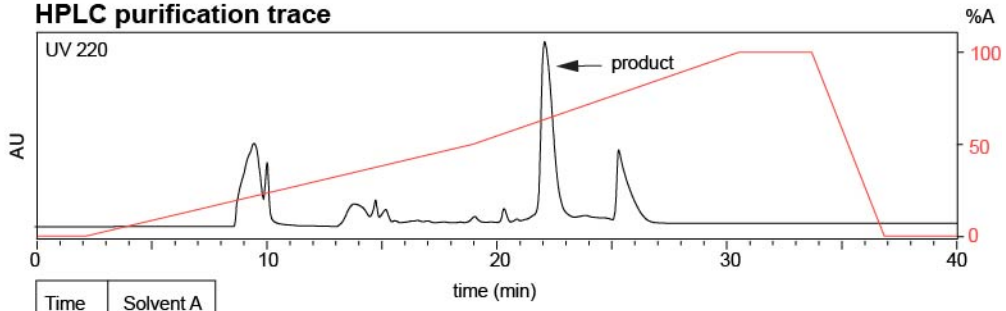




Ara_1-GGGGGGGK



HPLC purification trace

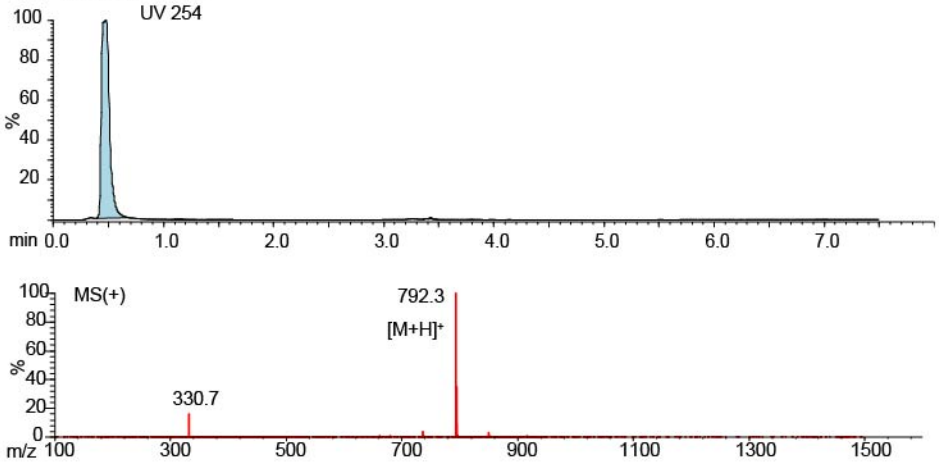


Time (min)	Solvent A (%)
0	0
2	0
18	50
29	100
32	100
35	2
38	2

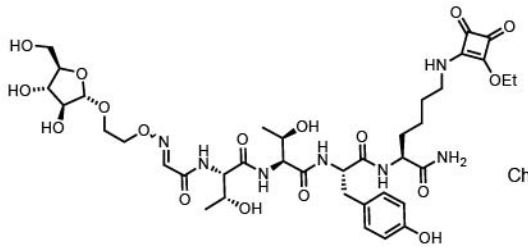
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

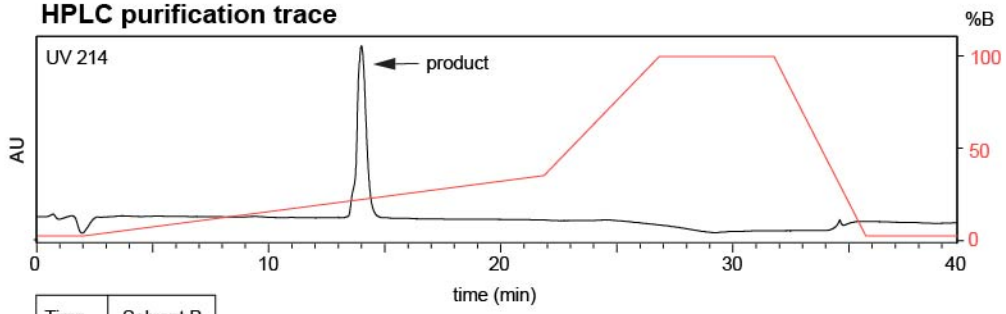


Ara_1-TTYK squaramide



Chemical Formula: $C_{38}H_{55}N_7O_{17}$
Exact Mass: 881.37

HPLC purification trace

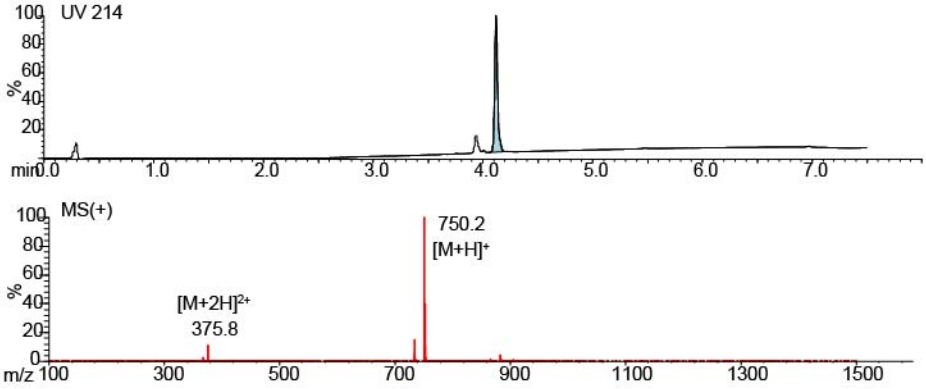


Time (min)	Solvent B (%)
0	2
2	2
22	35
27	100
32	100
36	2
40	2

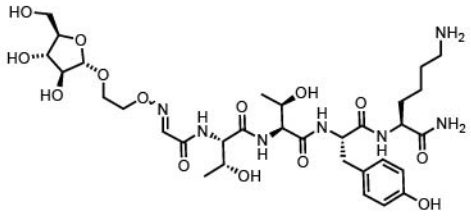
Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

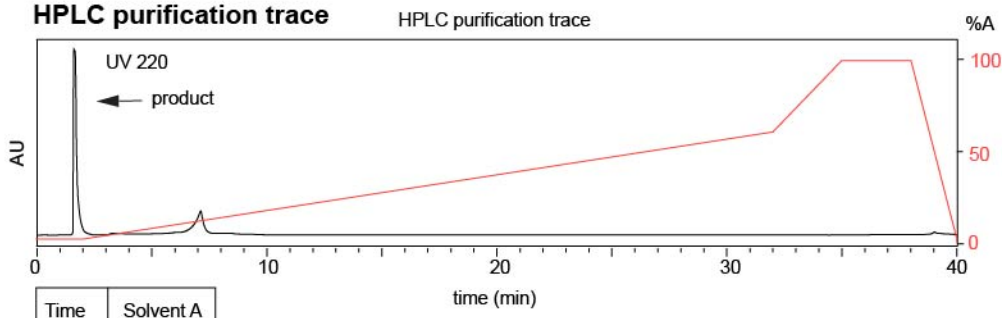


Ara_1-TTYK



Chemical Formula: $C_{32}H_{51}N_7O_{14}$
Exact Mass: 757.35

HPLC purification trace

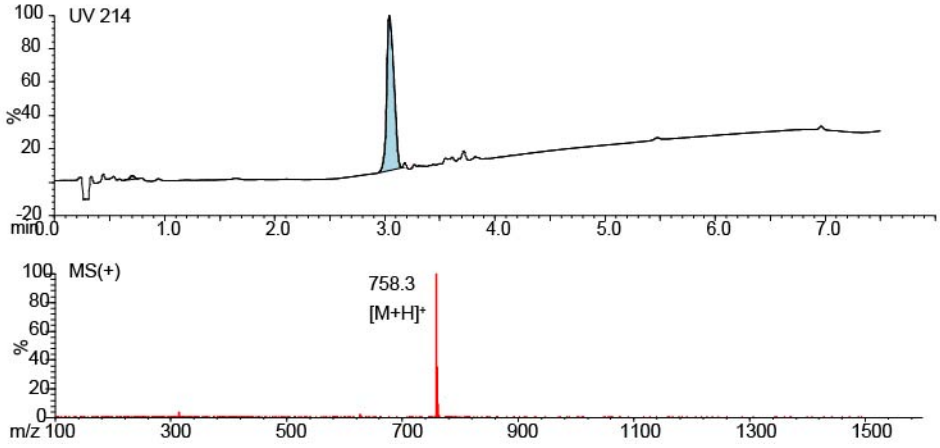


Time (min)	Solvent A (%)
0	0
2	0
32	60
35	100
38	100
41	0
46	0

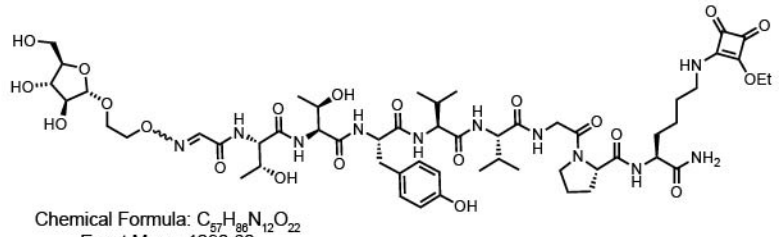
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

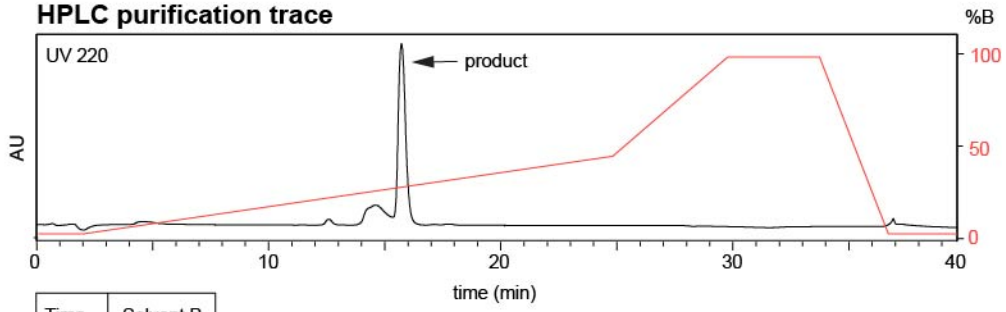


Ara_1-TTYVVNPK squaramide



Chemical Formula: $C_{57}H_{86}N_{12}O_{22}$
Exact Mass: 1290.60

HPLC purification trace

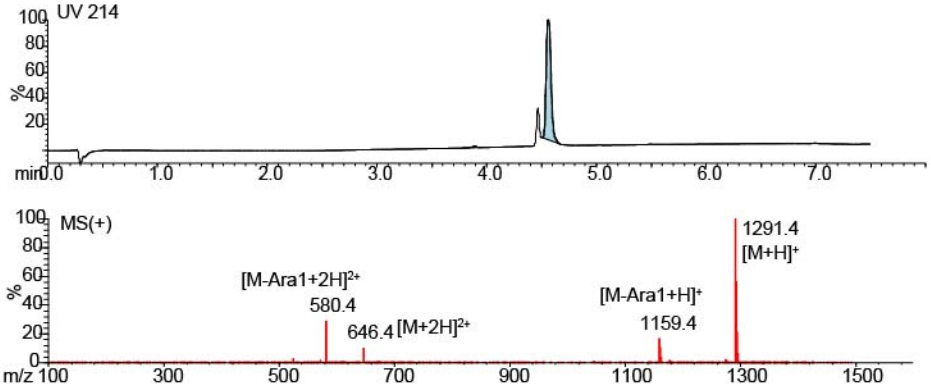


Time (min)	Solvent B (%)
0	2
2	2
25	45
30	100
34	100
37	2
40	2

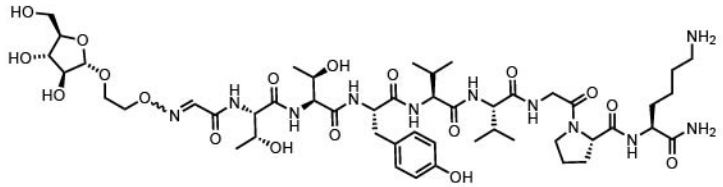
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

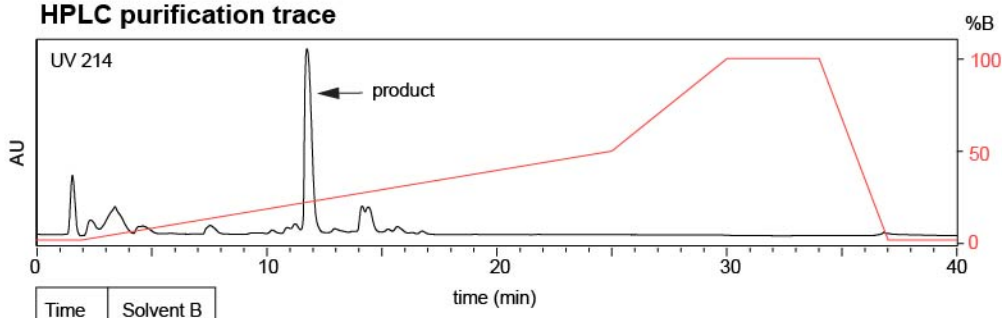


Ara_1-TTYVVNPK



Chemical Formula: $C_{51}H_{82}N_{12}O_{19}$
Exact Mass: 1166.58

HPLC purification trace

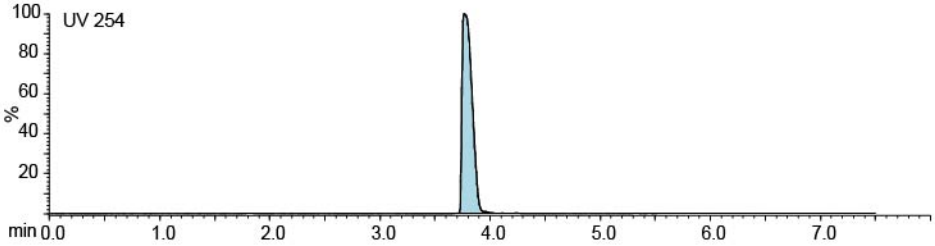


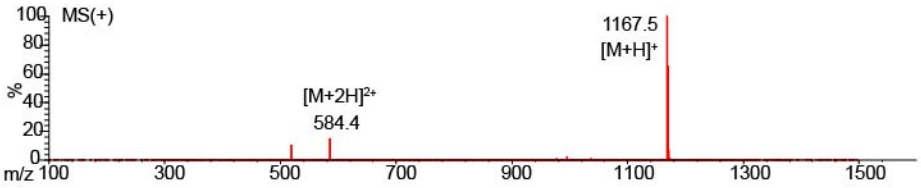
Time (min)	Solvent B (%)
0	2
2	2
25	50
30	100
34	100
37	2
40	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mmm)

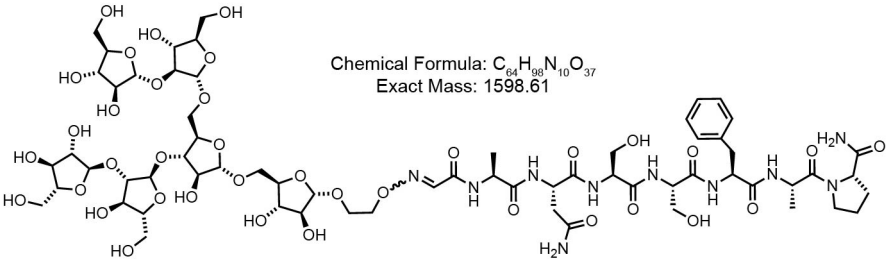
Solvent A: H_2O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

UPLC-MS trace

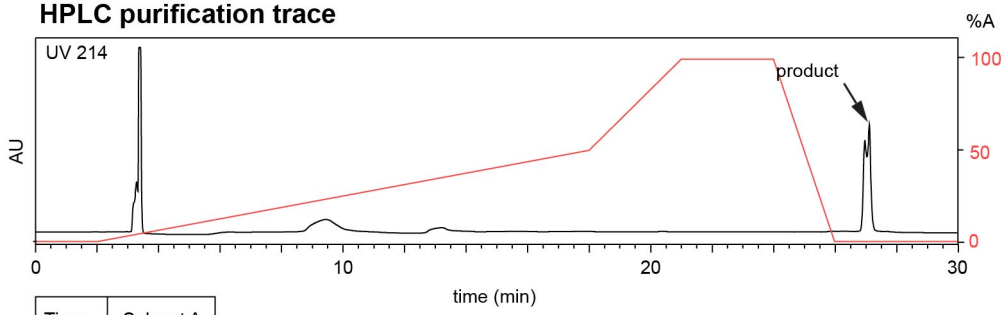




Ara_6-ANSSFAP



HPLC purification trace

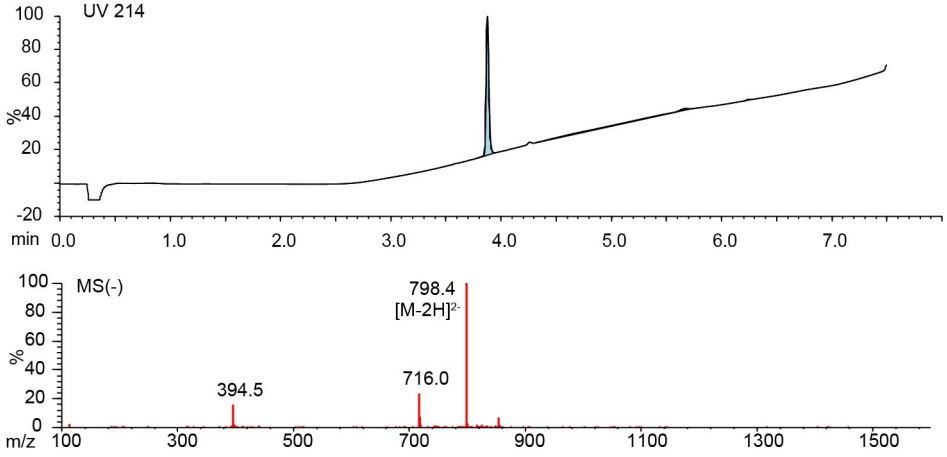


Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	2
30	2

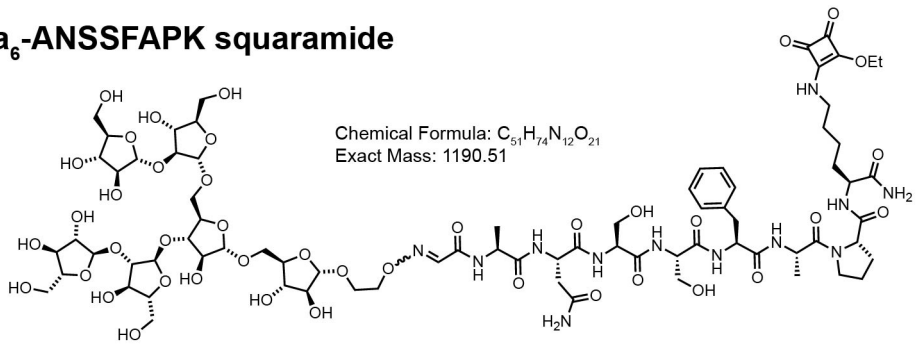
Flow rate: 1 mL/min
Amide Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

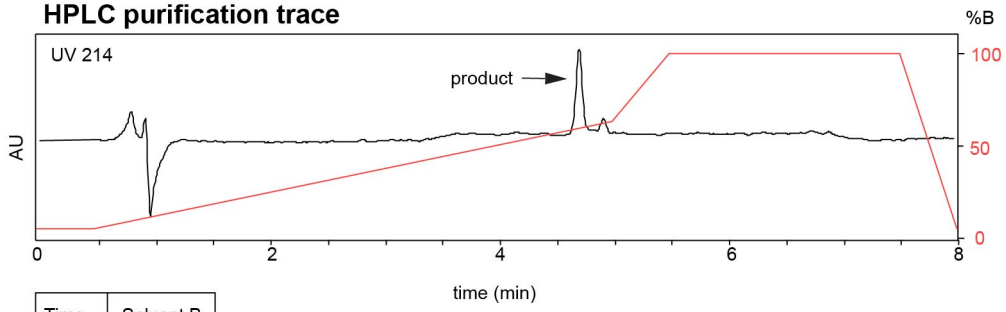
UPLC-MS trace



Ara_6-ANSSFAPK squaramide



HPLC purification trace

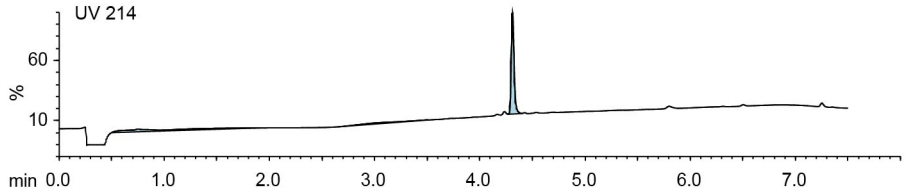


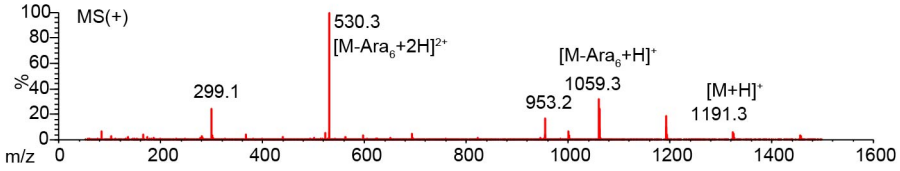
Time (min)	Solvent B (%)
0	5
0.5	5
5	60
5.5	95
7.5	95
8	5

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

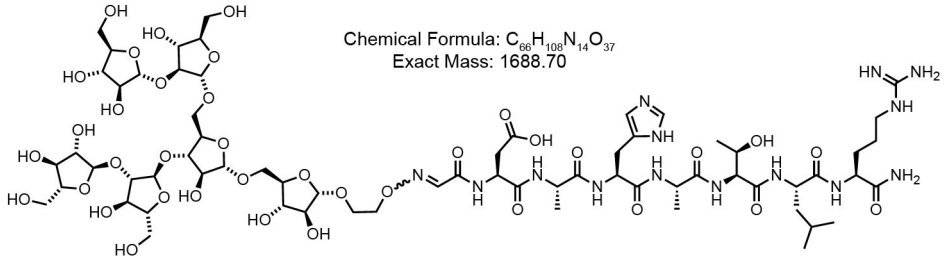
Solvent A: H_2O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

UPLC-MS trace

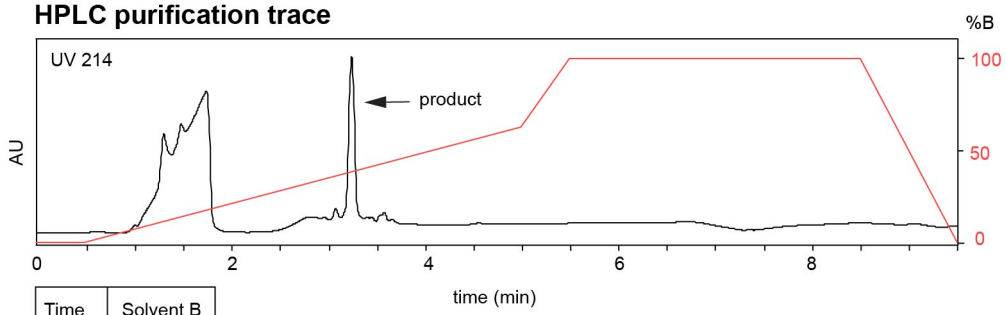




Ara_6-DAHATLR



HPLC purification trace

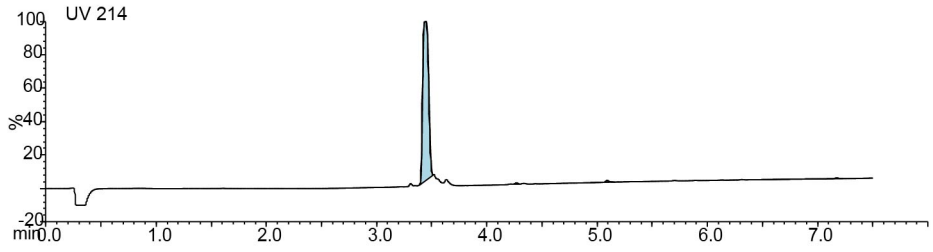


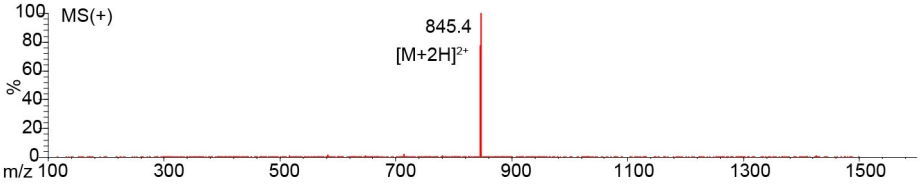
Time (min)	Solvent B (%)
0	1
0.5	1
5	60
5.5	95
8.5	95
9.5	1

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

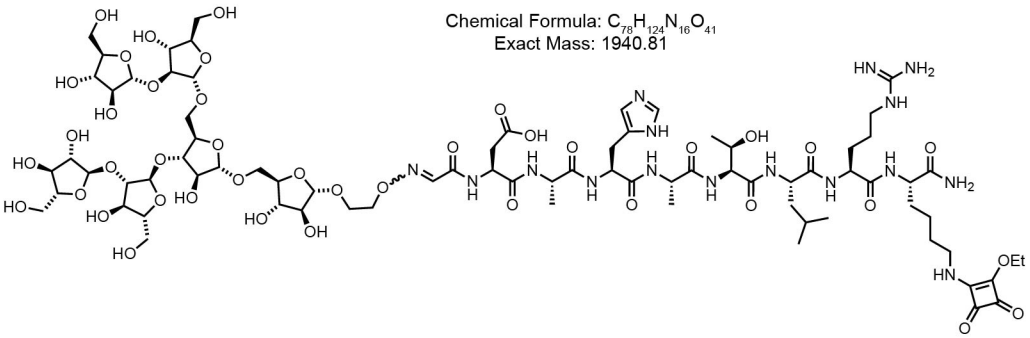
Solvent A: H_2O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

UPLC-MS trace



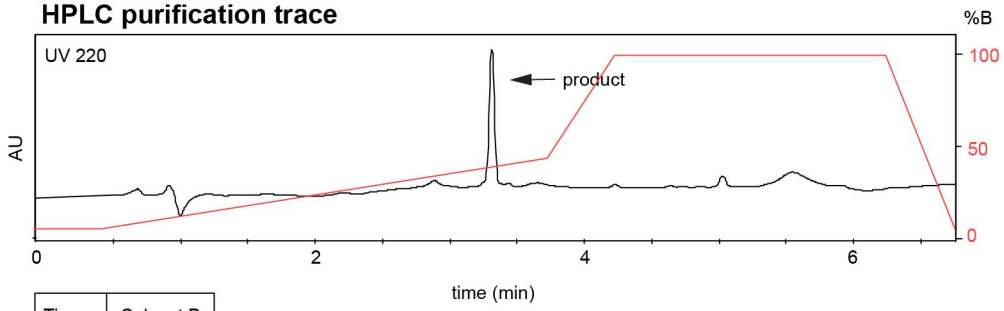


Ara_6-DAHATLRK squaramide



Chemical Formula: $C_{78}H_{124}N_{16}O_{41}$
Exact Mass: 1940.81

HPLC purification trace

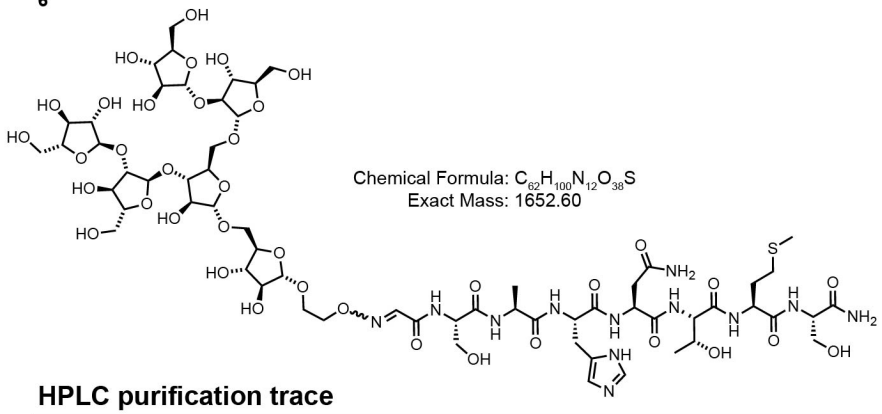


Time (min)	Solvent B (%)
0	5
0.5	5
3.8	41.6
4.3	95
6.3	95
6.8	5

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

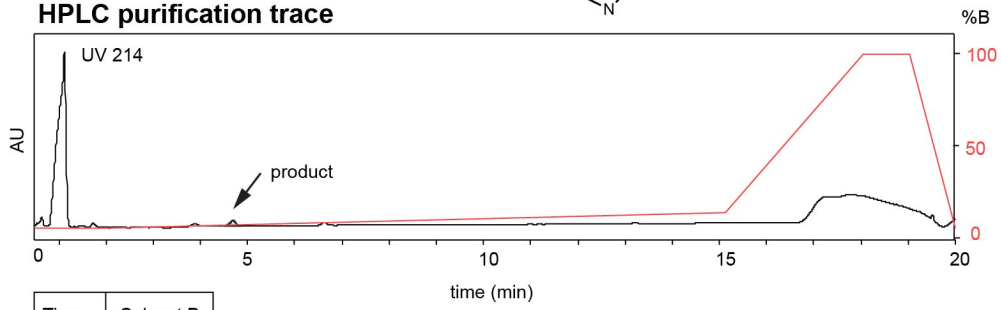
Solvent A: H_2O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

Ara_6-SAHNTMS



Chemical Formula: $C_{62}H_{100}N_{12}O_{38}S$
Exact Mass: 1652.60

HPLC purification trace

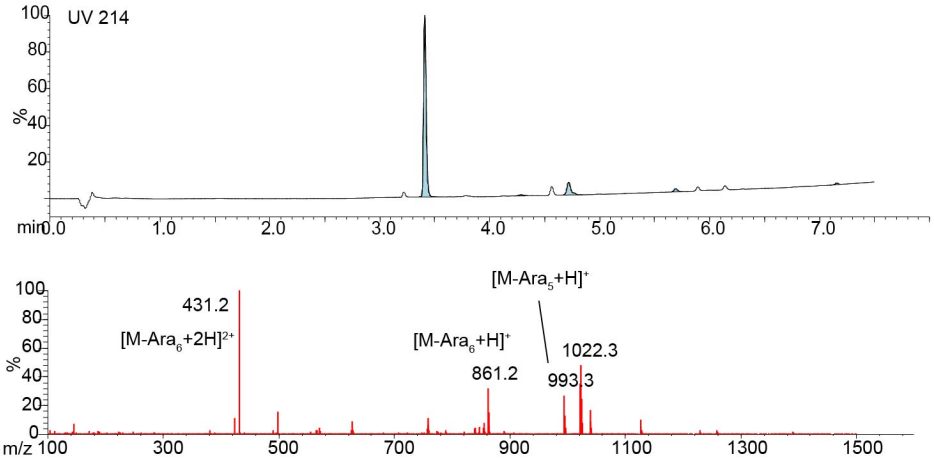


Time (min)	Solvent B (%)
0	5
1.5	5
15	13
18	95
19	95
20	5

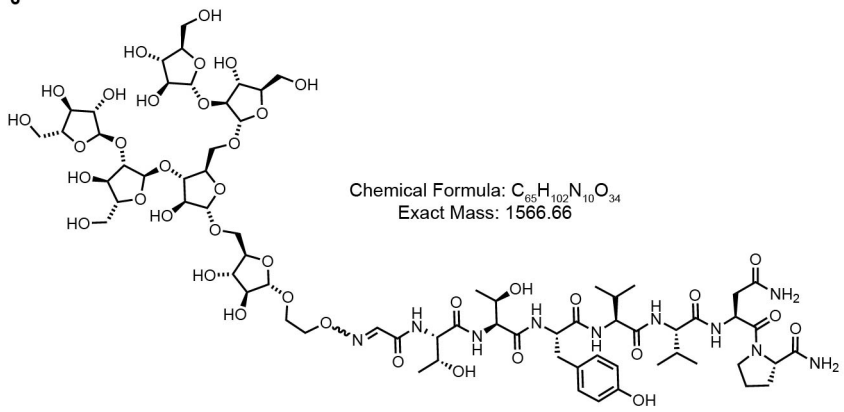
Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

Solvent A: H_2O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

UPLC-MS trace

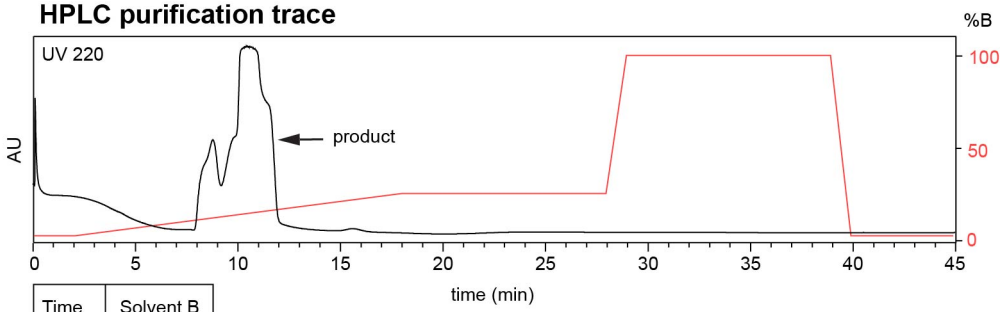


Ara_6-TTYVVNP



Chemical Formula: $C_{65}H_{102}N_{10}O_{34}$
Exact Mass: 1566.66

HPLC purification trace

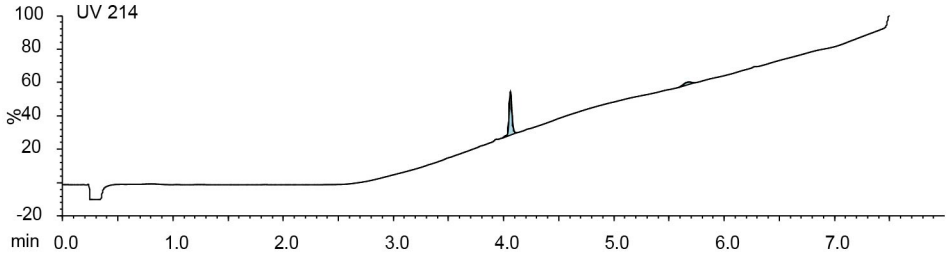


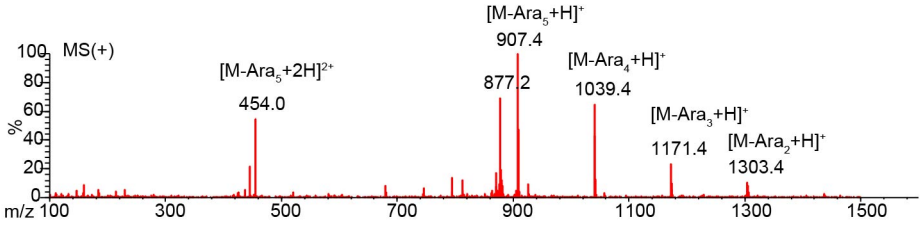
Time (min)	Solvent B (%)
0	2
2	2
18	25
28	25
29	100
39	100
40	2
45	2

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

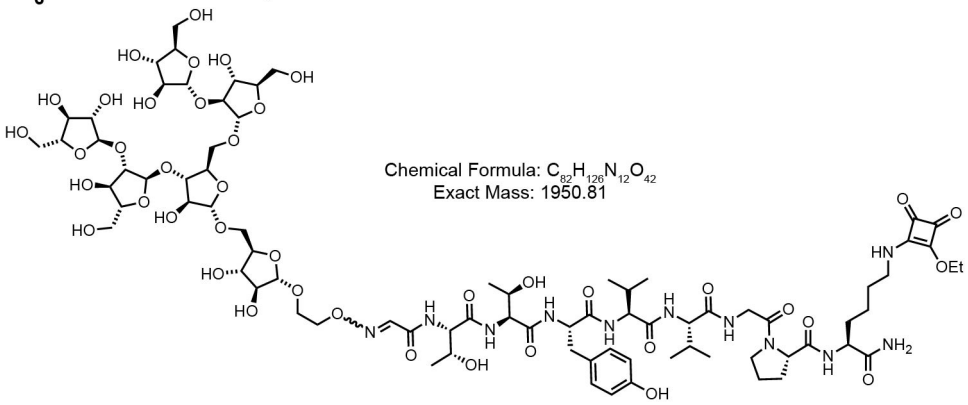
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace



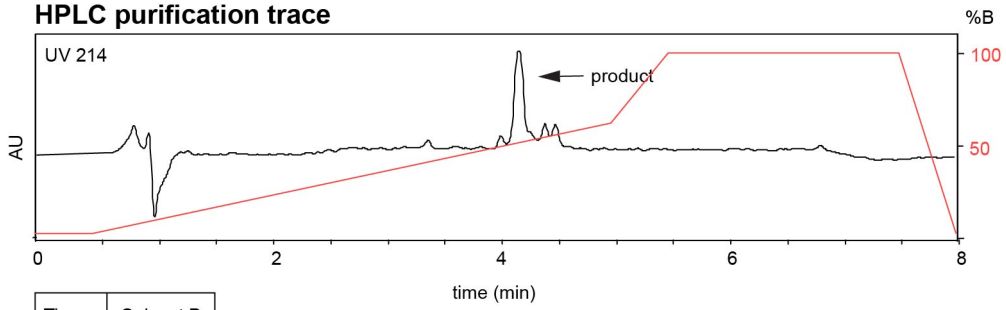


Ara_6-TTYVVNPK squaramide



Chemical Formula: $C_{82}H_{126}N_{12}O_{42}$
Exact Mass: 1950.81

HPLC purification trace

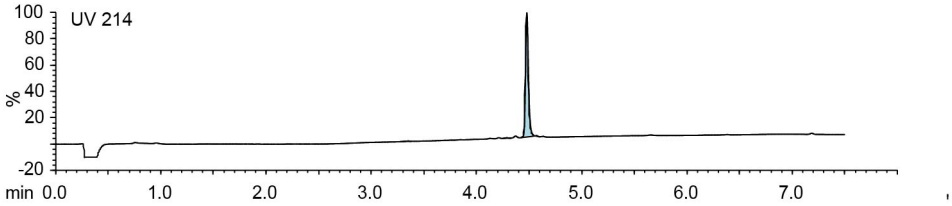


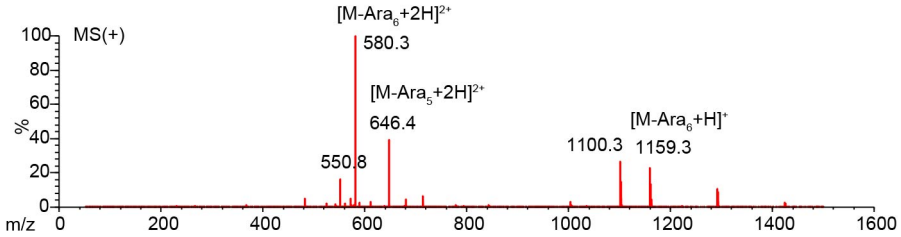
Time (min)	Solvent B (%)
0	5
0.5	5
5	60
5.5	95
7.5	95
8	5

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

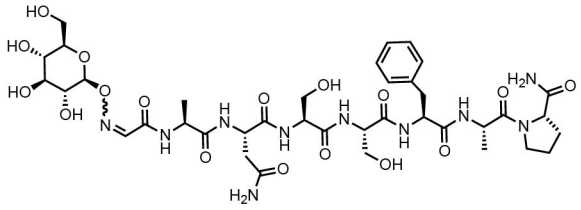
Solvent A: H_2O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

UPLC-MS trace



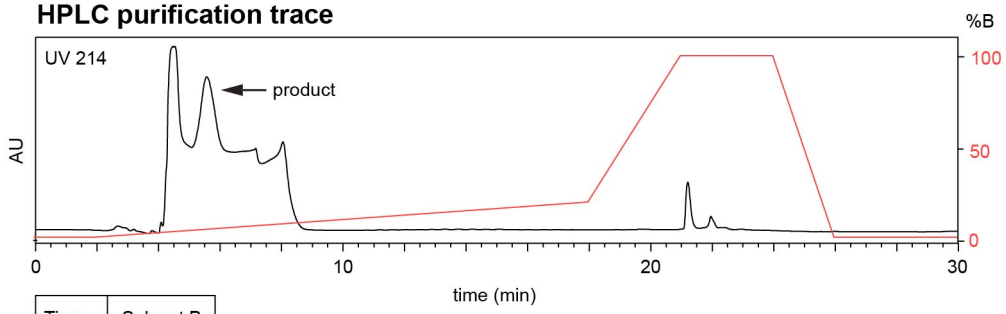


Glc-ANSSFAP



Chemical Formula: $C_{38}H_{56}N_{10}O_{17}$
Exact Mass: 924.38

HPLC purification trace

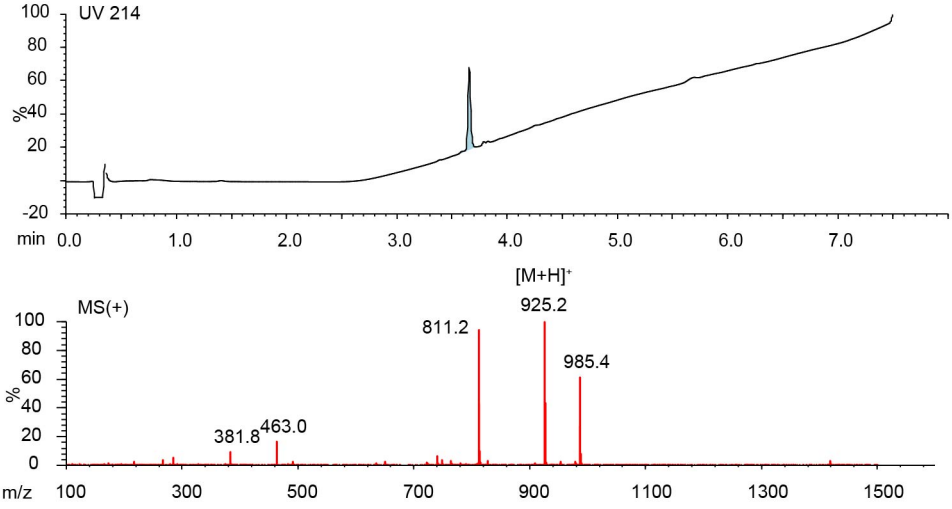


Time (min)	Solvent B (%)
0	2
2	2
18	20
21	100
24	100
26	2
30	2

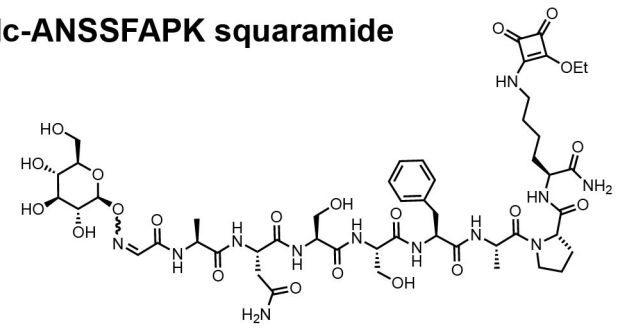
Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 5 µm, 4.6 mm X 150 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

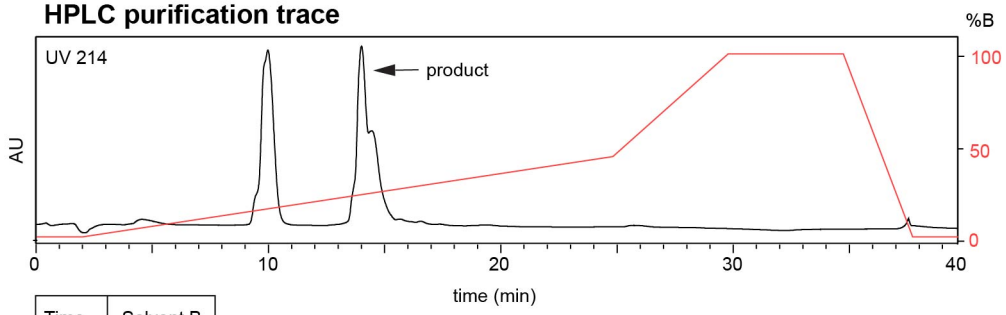


Glc-ANSSFAPK squaramide



Chemical Formula: $C_{50}H_{72}N_{12}O_{21}$
Exact Mass: 1176.49

HPLC purification trace

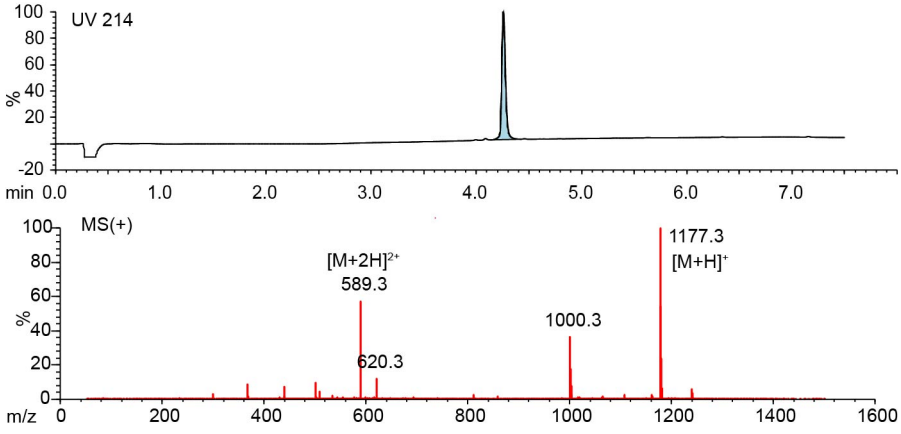


Time (min)	Solvent B (%)
0	2
2	2
22	35
27	100
32	100
36	2
40	2

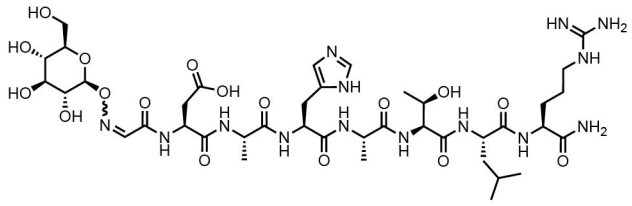
Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

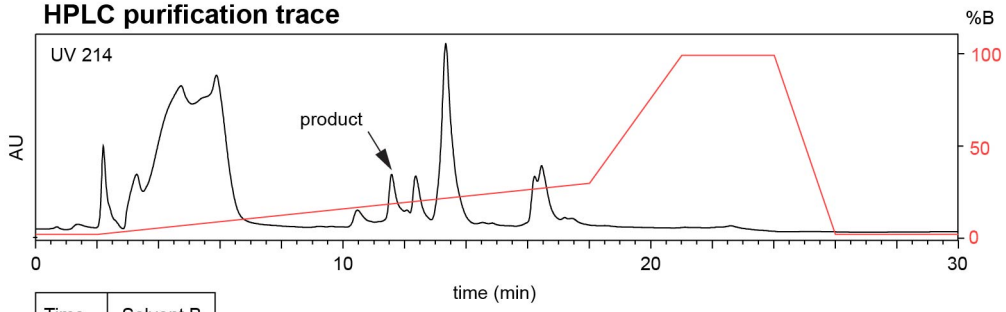


Glc-DAHATLR



Chemical Formula: $C_{40}H_{66}N_{14}O_{17}$
Exact Mass: 1014.47

HPLC purification trace

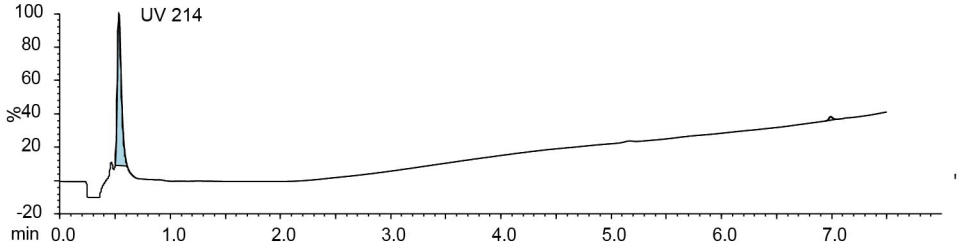


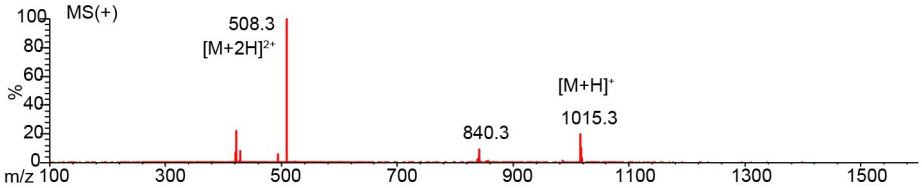
Time (min)	Solvent B (%)
0	2
2	2
18	30
21	100
24	100
26	2
30	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)

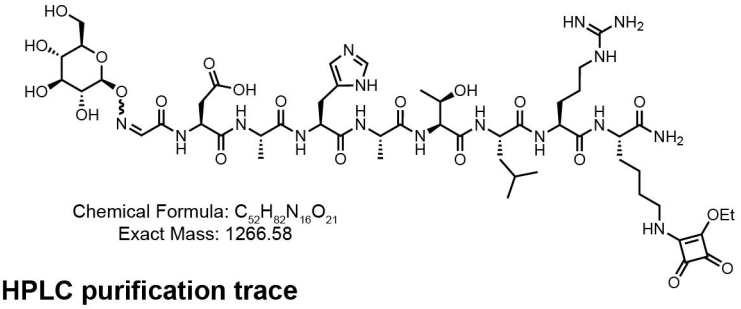
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace



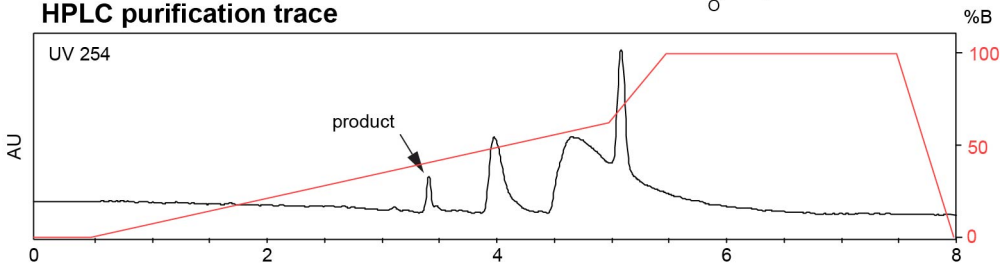


Glc-DAHATLRK squaramide



Chemical Formula: $C_{52}H_{82}N_{16}O_{21}$
Exact Mass: 1266.58

HPLC purification trace

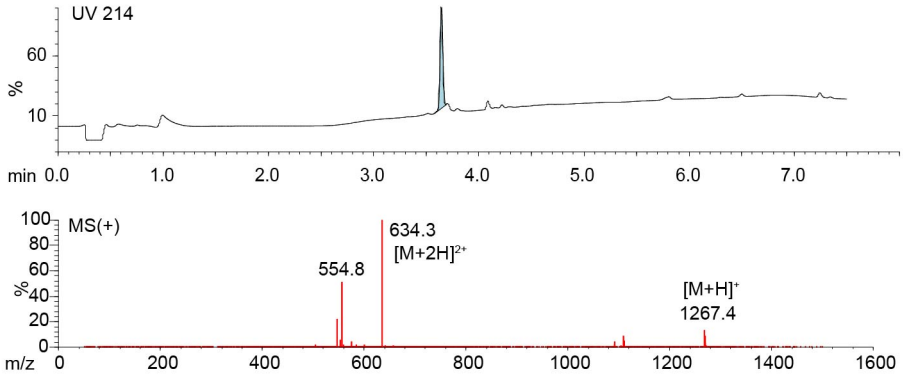


Time (min)	Solvent B (%)
0	2
0.5	2
5	60
5.5	95
7.5	95
8	2

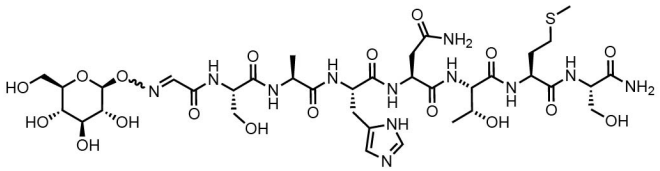
Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

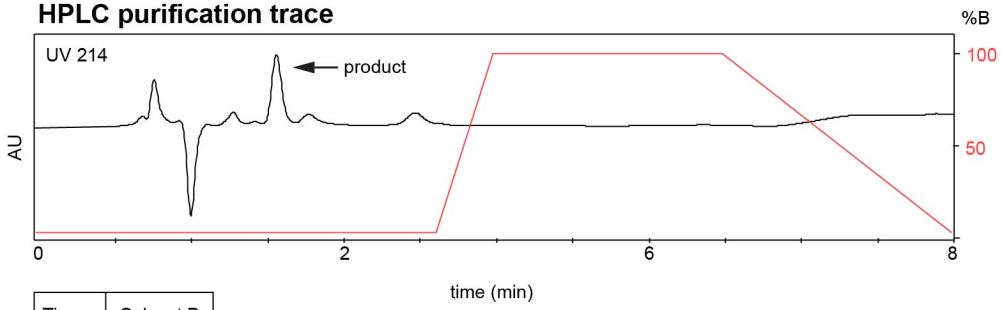


Glc-SAHNTMS



Chemical Formula: $C_{36}H_{58}N_{12}O_{18}S$
Exact Mass: 978.37

HPLC purification trace

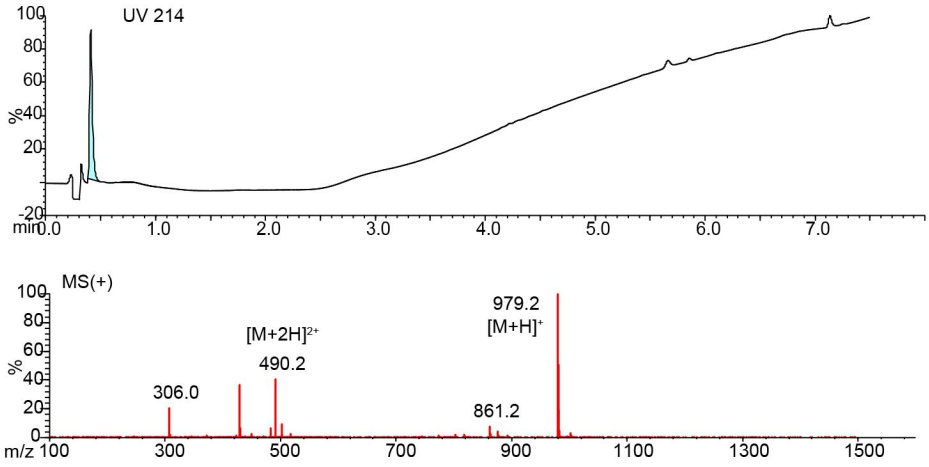


Time (min)	Solvent B (%)
0	2
3.5	2
4	90
6	90
8	2

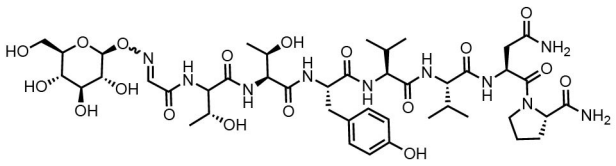
Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

Solvent A: H_2O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

UPLC-MS trace

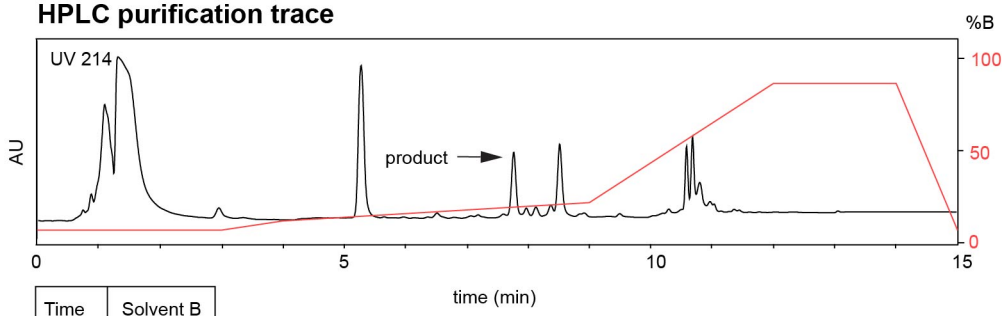


Glc-TTYVVNP



Chemical Formula: $C_{44}H_{68}N_{10}O_{18}$
Exact Mass: 1024.47

HPLC purification trace

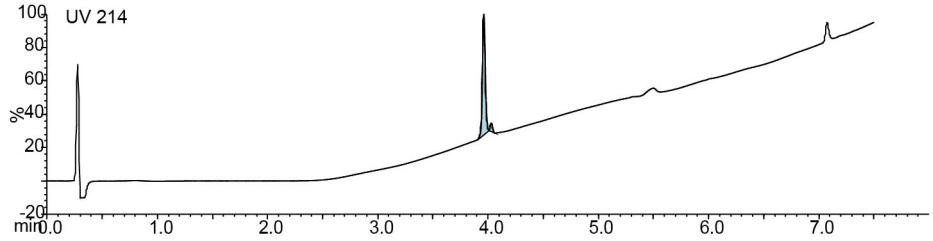


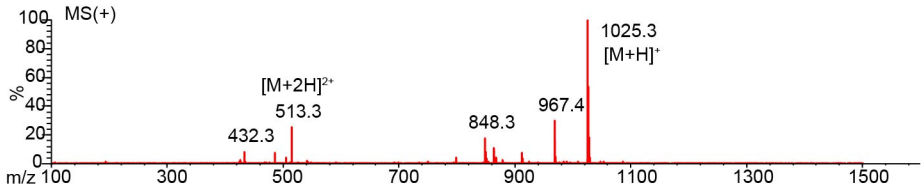
Time (min)	Solvent B (%)
0	5
3	5
4	10
9	20
12	85
24	85
15	5

Flow rate: 1 mL/min
Symmetry C18 Column
(100Å, 3.5 µm, 4.6 mm X 75 mm)

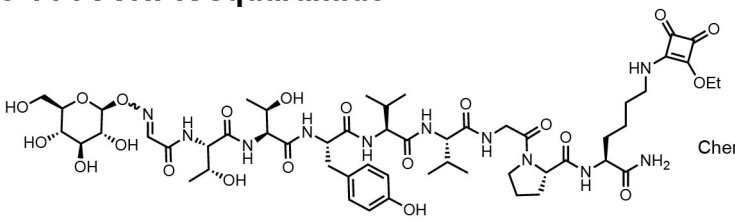
Solvent A: H_2O + 0.1% (v/v) FA
Solvent B: MeCN + 0.1% (v/v) FA

UPLC-MS trace



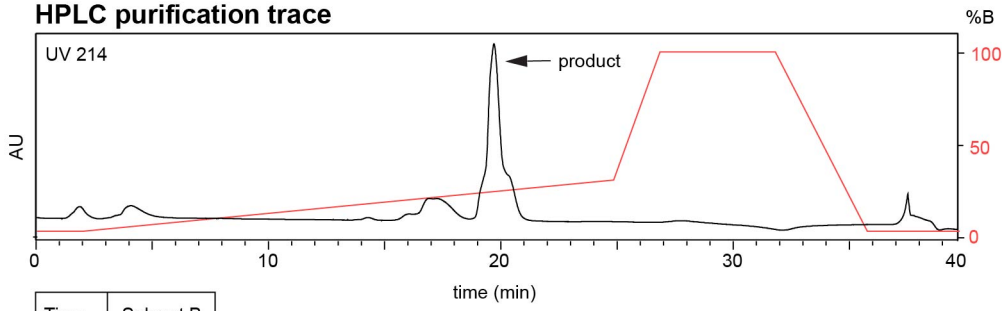


Glc-TTYVVNPK squaramide



Chemical Formula: $C_{56}H_{84}N_{12}O_{22}$
Exact Mass: 1276.58

HPLC purification trace

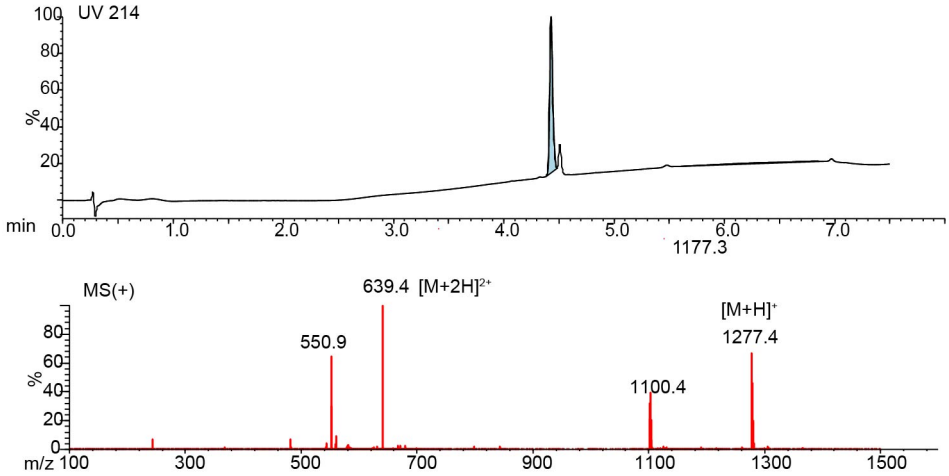


Time (min)	Solvent B (%)
0	2
2	2
25	30
27	100
32	100
36	2
40	2

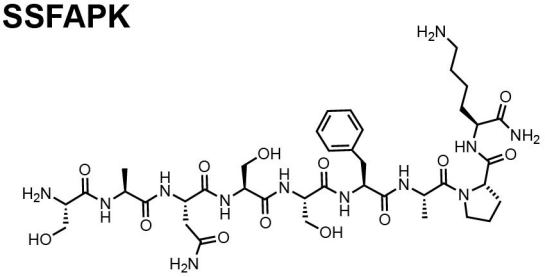
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

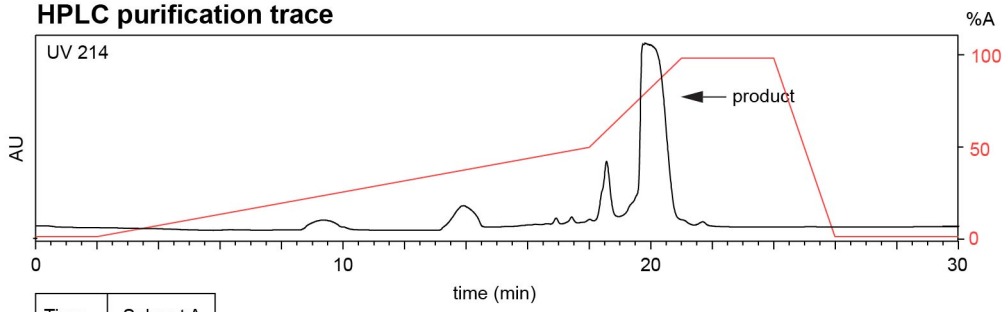


SANSSFAPK



Chemical Formula: $C_{39}H_{62}N_{12}O_{13}$
Exact Mass: 906.46

HPLC purification trace

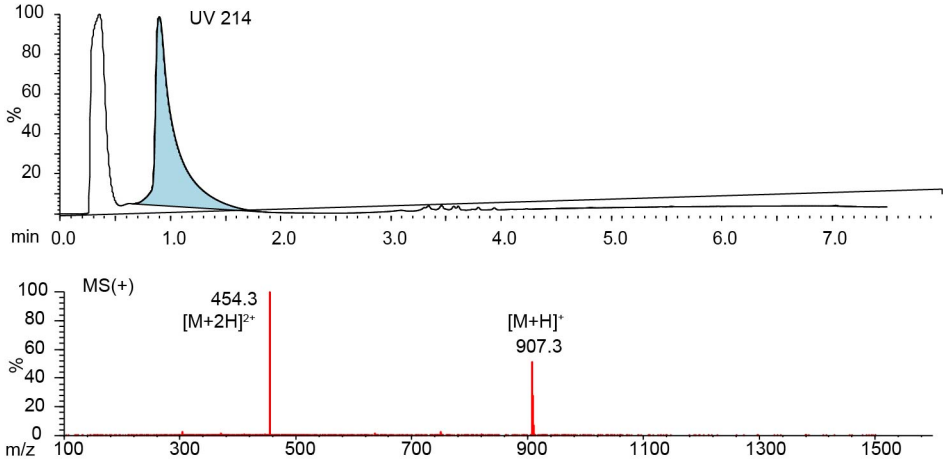


Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	0
30	0

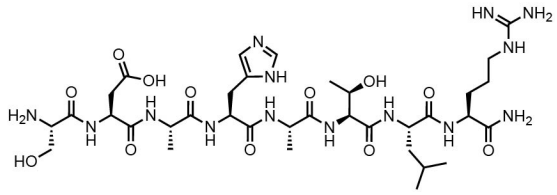
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

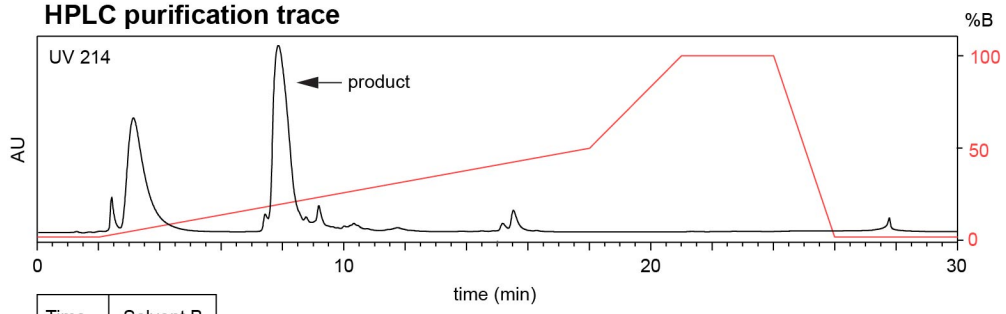


SDAHATLR



Chemical Formula: $C_{35}H_{60}N_{14}O_{12}$
Exact Mass: 868.45

HPLC purification trace

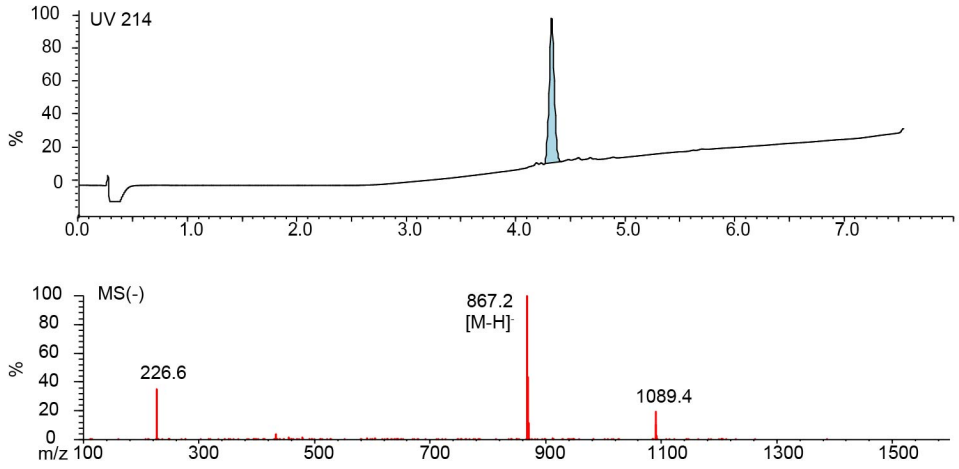


Time (min)	Solvent B (%)
0	2
2	2
18	50
21	100
24	100
26	2
30	2

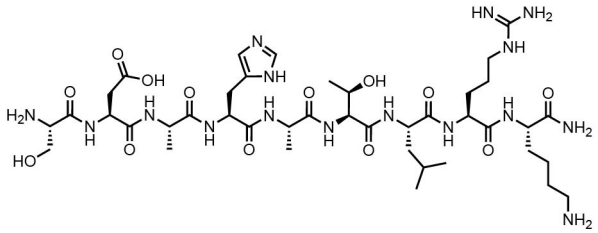
Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

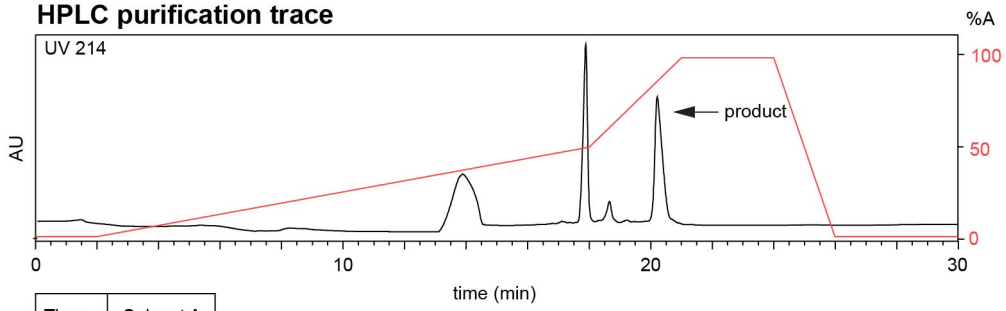


SDAHATLRK



Chemical Formula: $C_{41}H_{72}N_{16}O13$
Exact Mass: 996.55

HPLC purification trace

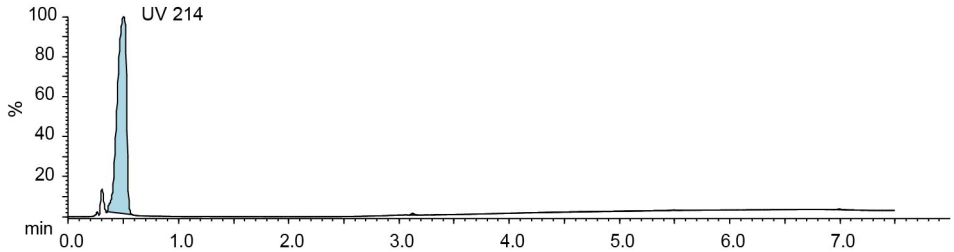


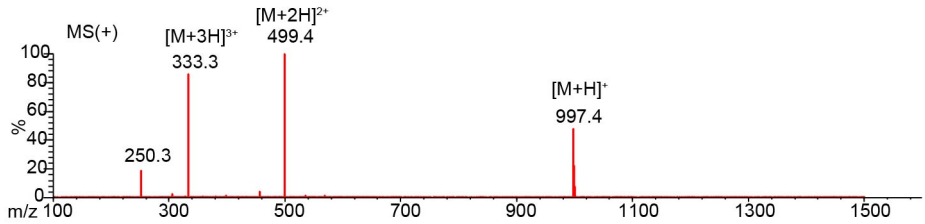
Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	0
30	0

Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

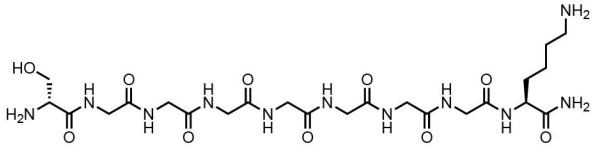
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace



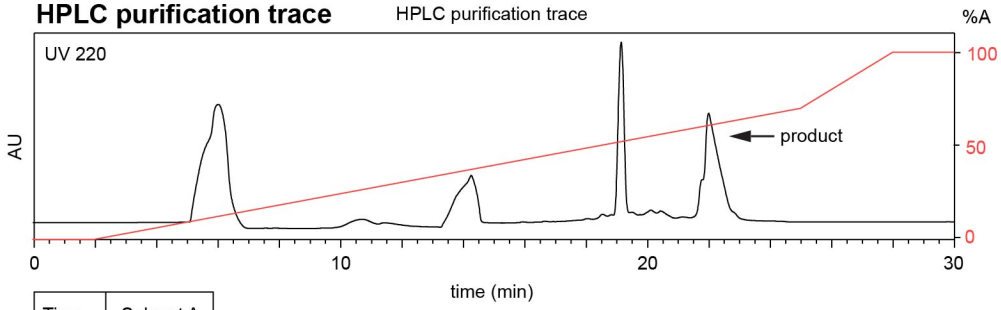


SGGGGGGGK



Chemical Formula: $C_{23}H_{41}N_{11}O_{10}$
Exact Mass: 631.30

HPLC purification trace

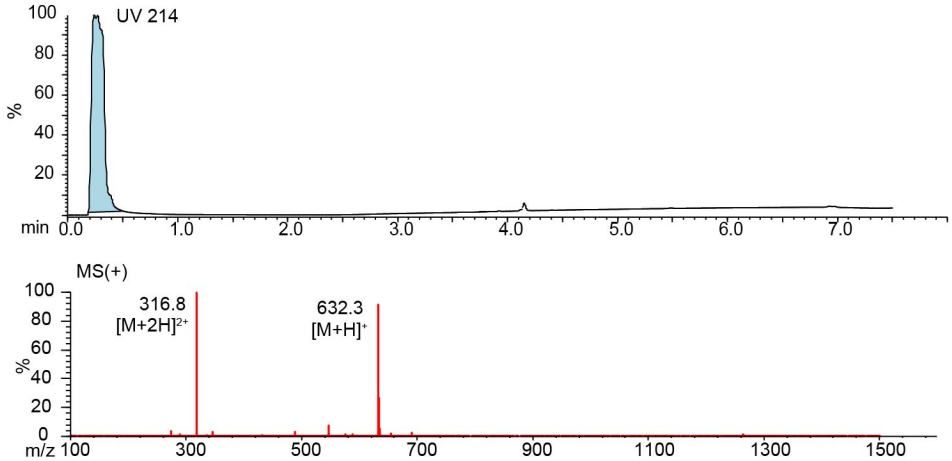


Time (min)	Solvent A (%)
0	0
2	0
25	70
28	100
30	100
34	0
38	0

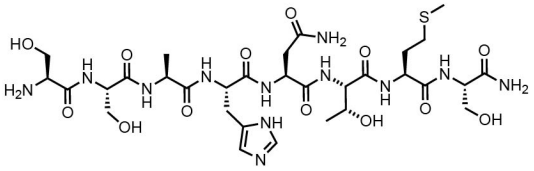
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

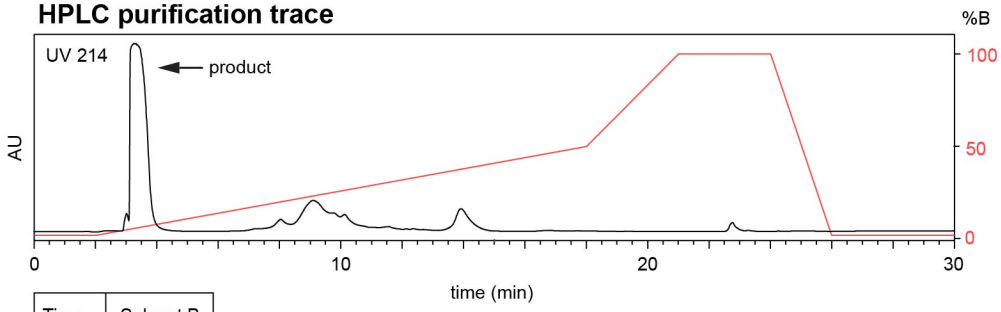


SSAHNTMS



Chemical Formula: $C_{31}H_{52}N_{12}O13S$
Exact Mass: 832.35

HPLC purification trace

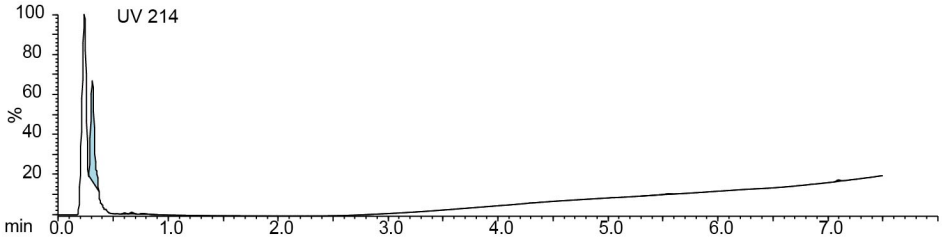


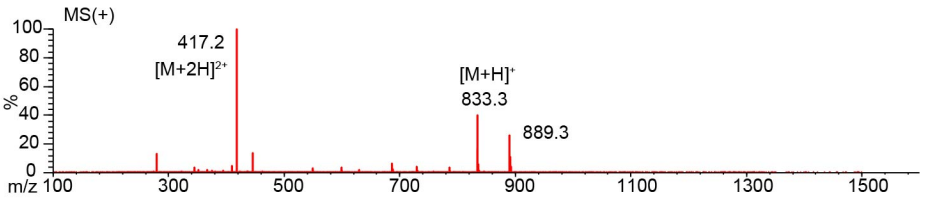
Time (min)	Solvent B (%)
0	2
2	2
18	50
21	100
24	100
26	2
30	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mmm)

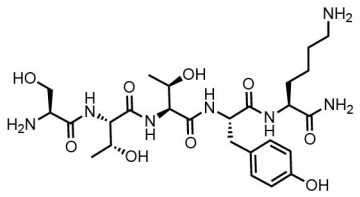
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace



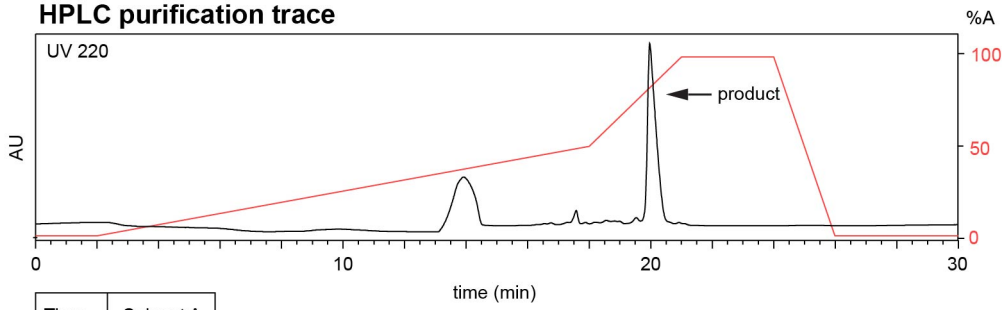


STTYK



Chemical Formula: $C_{26}H_{43}N_7O_9$
Exact Mass: 597.31

HPLC purification trace

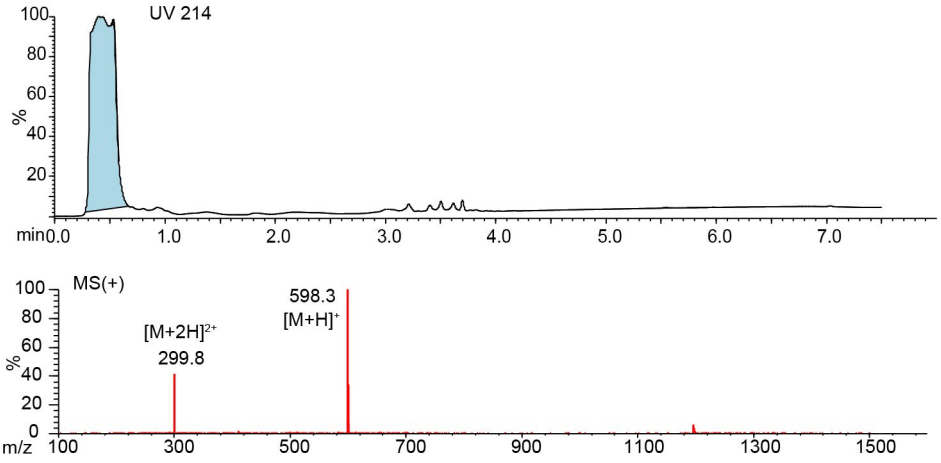


Time (min)	Solvent A (%)
0	0
2	0
18	50
21	100
24	100
26	0
30	0

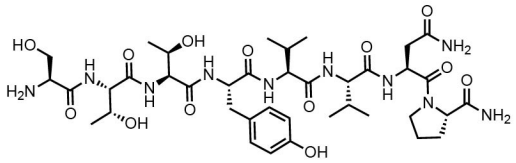
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace

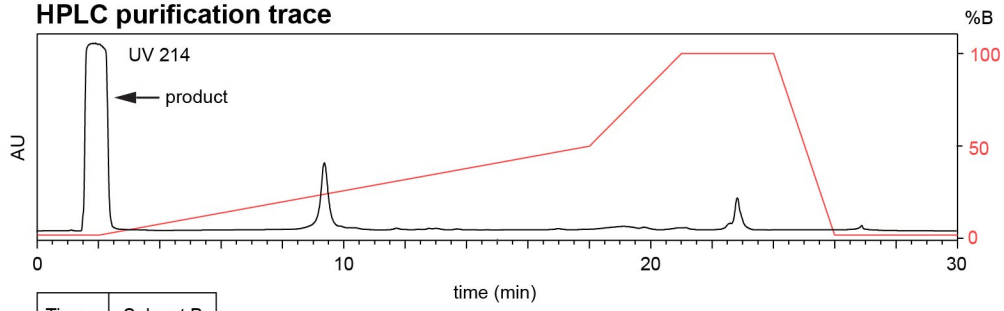


STTYVVNP



Chemical Formula: $C_{39}H_{62}N_{10}O_{13}$
Exact Mass: 878.45

HPLC purification trace

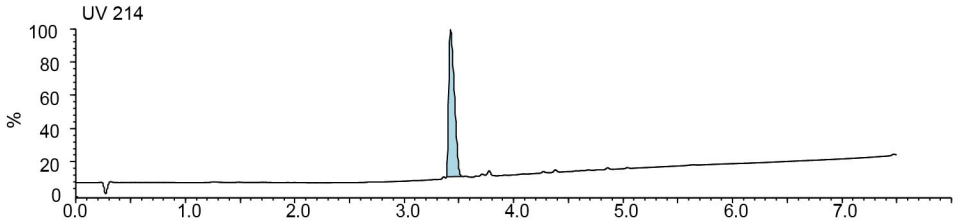


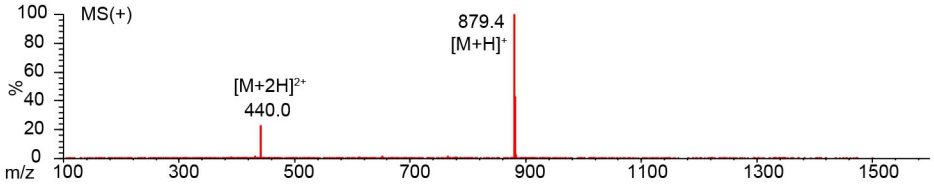
Time (min)	Solvent B (%)
0	2
2	2
18	50
21	100
24	100
26	2
30	2

Flow rate: 12 mL/min
Symmetry C18 Prep Column
(100Å, 5 µm, 19 mm X 50 mmm)

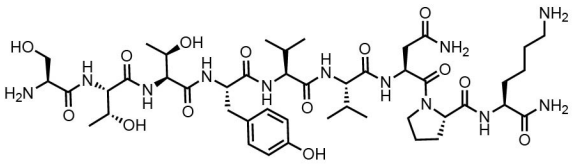
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

UPLC-MS trace



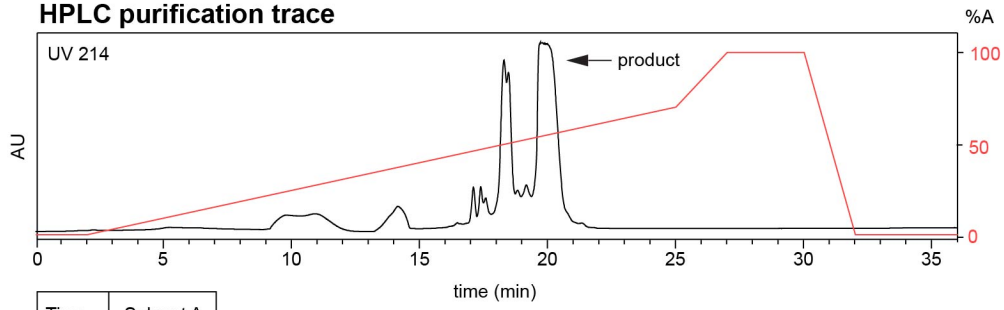


STTYVVNPK



Chemical Formula: $C_{45}H_{74}N_{12}O_{14}$
Exact Mass: 1006.54

HPLC purification trace

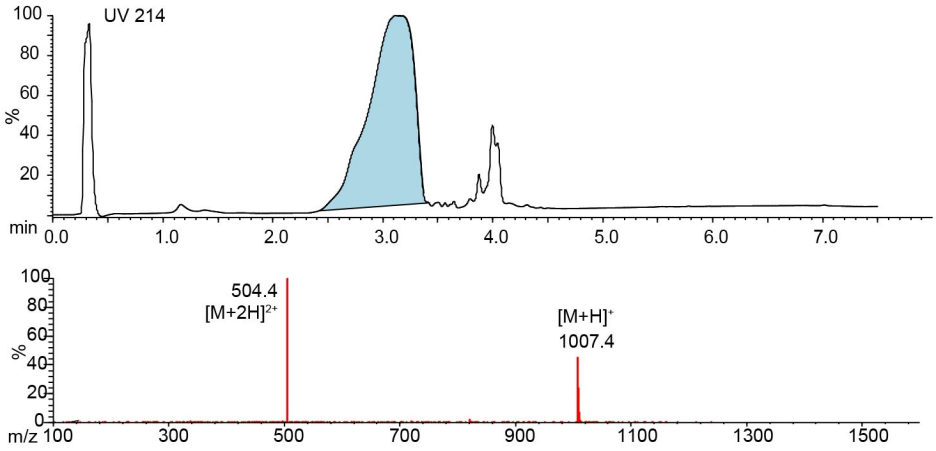


Time (min)	Solvent A (%)
0	0
2	0
25	70
27	100
30	100
32	0
36	0

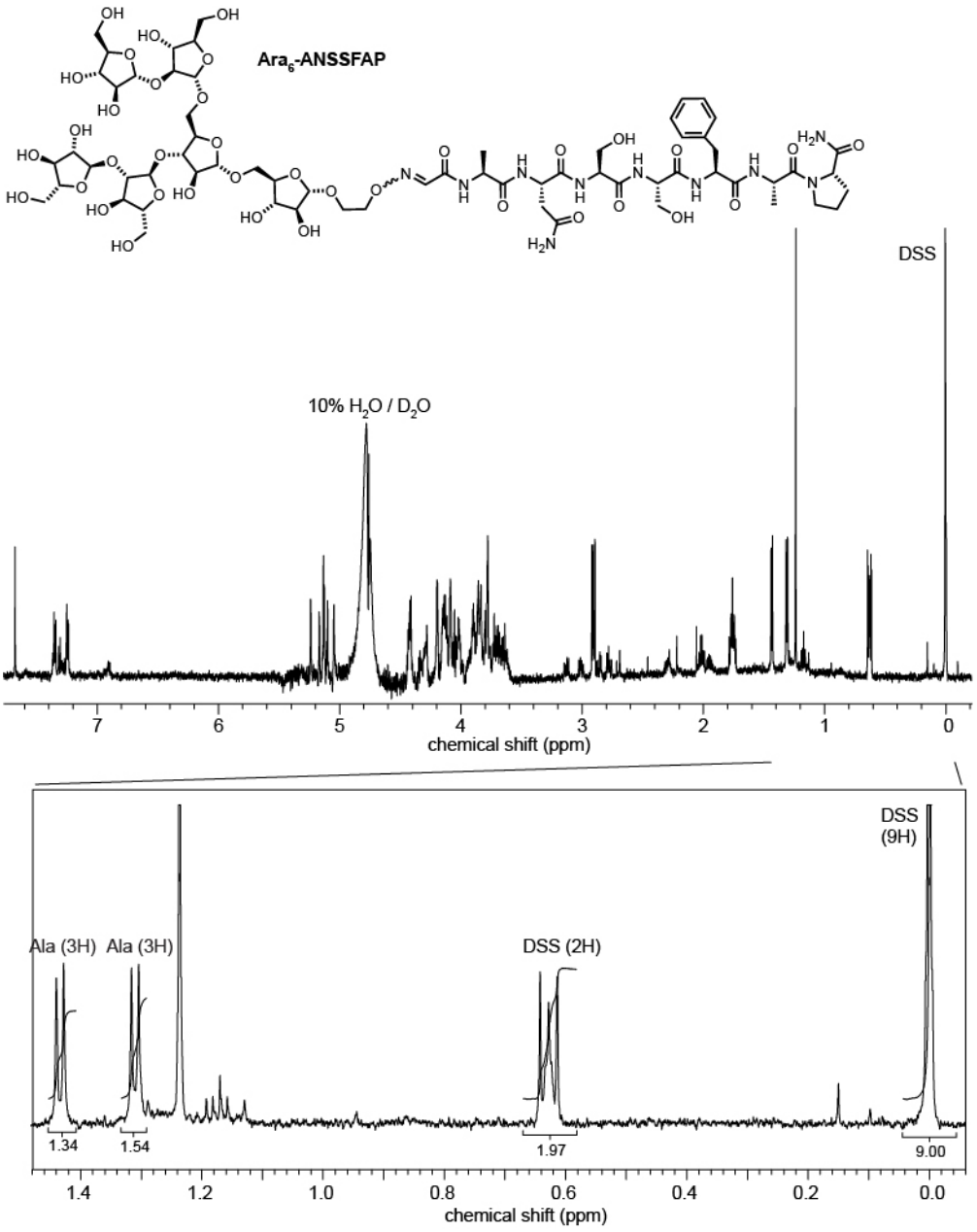
Flow rate: 12 mL/min
Amide Column
(130Å, 3.5 µm, 4.6 mm X 250 mm)

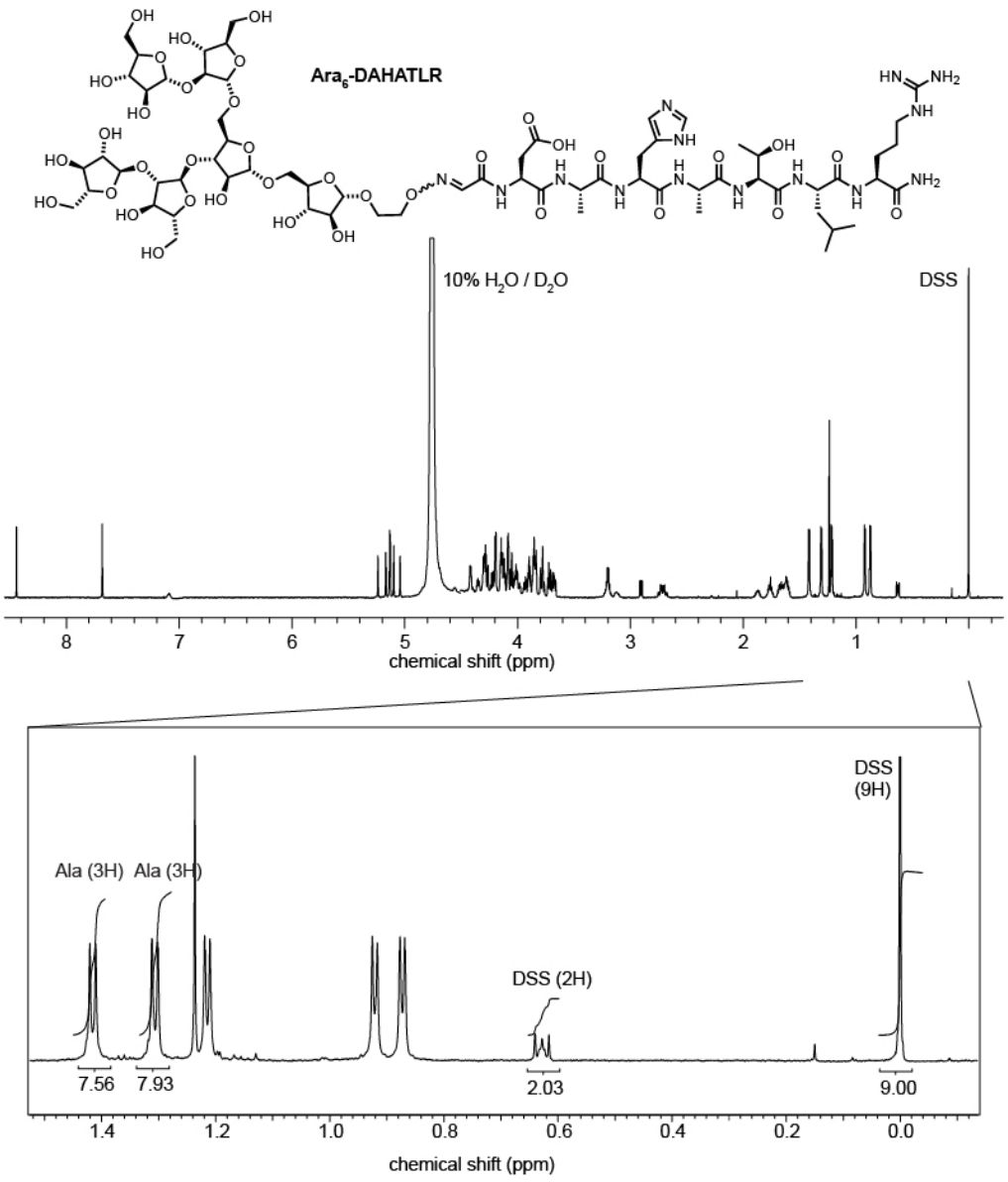
Solvent A: H_2O + 0.1% (v/v) TFA
Solvent B: MeCN + 0.1% (v/v) TFA

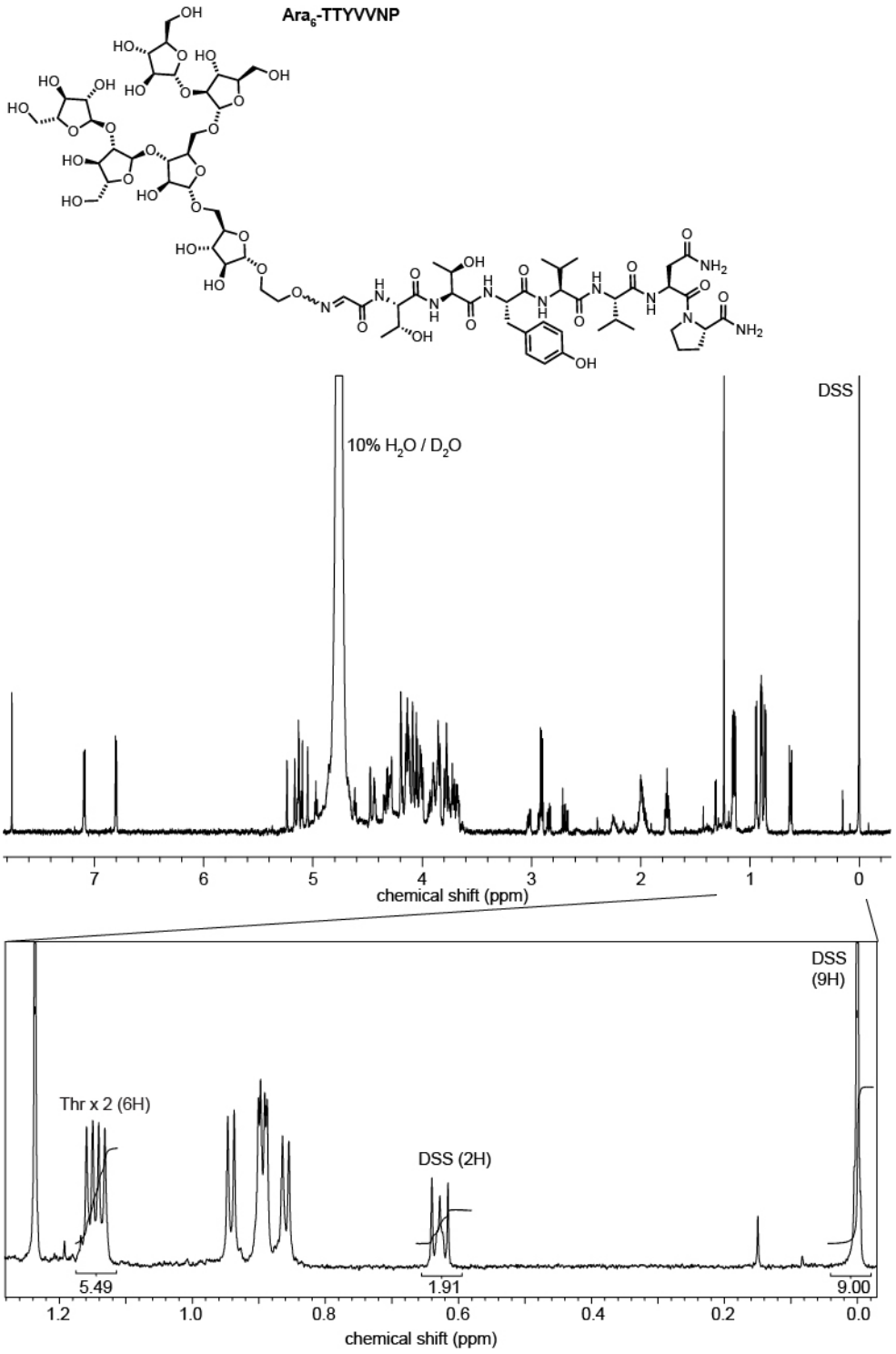
UPLC-MS trace

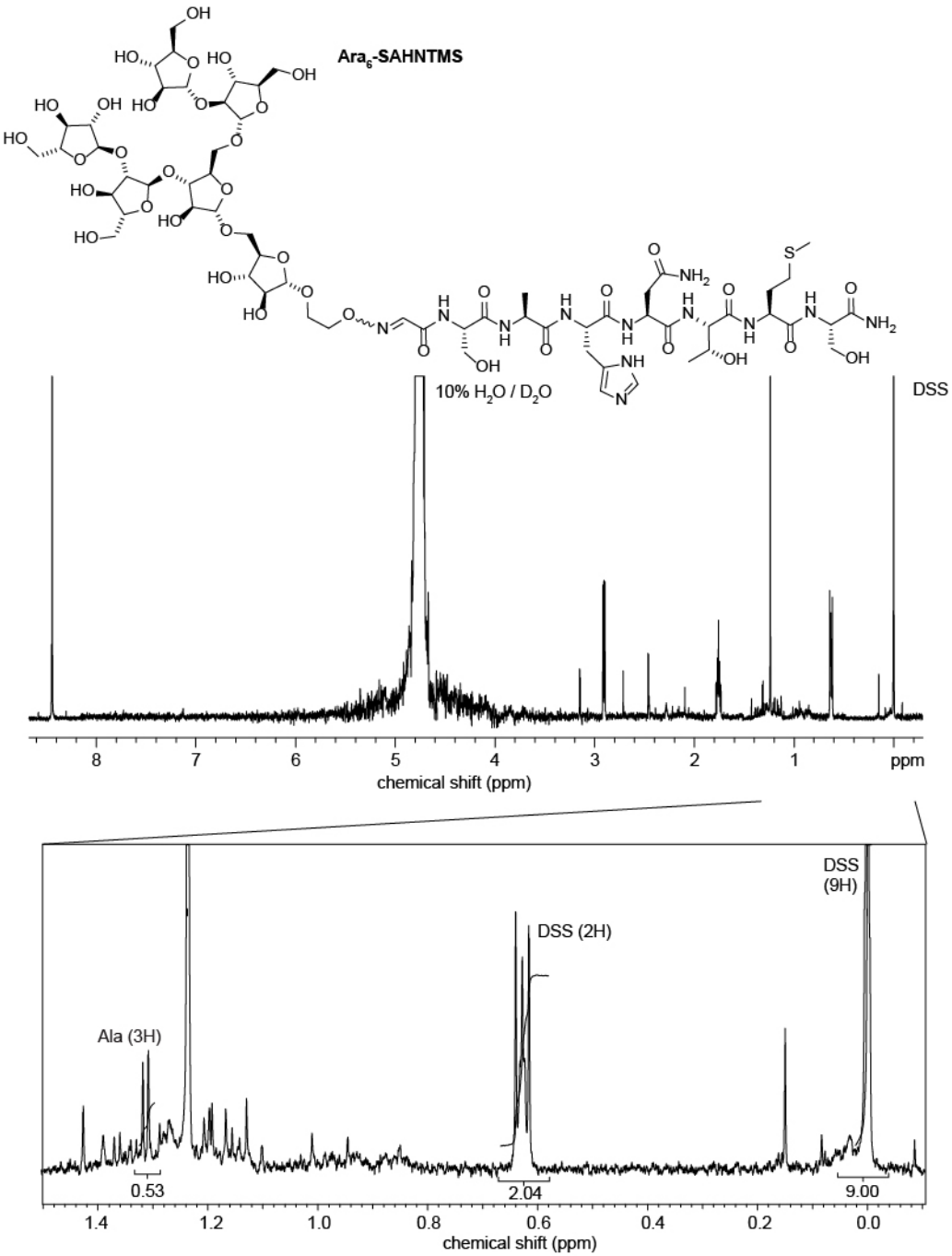


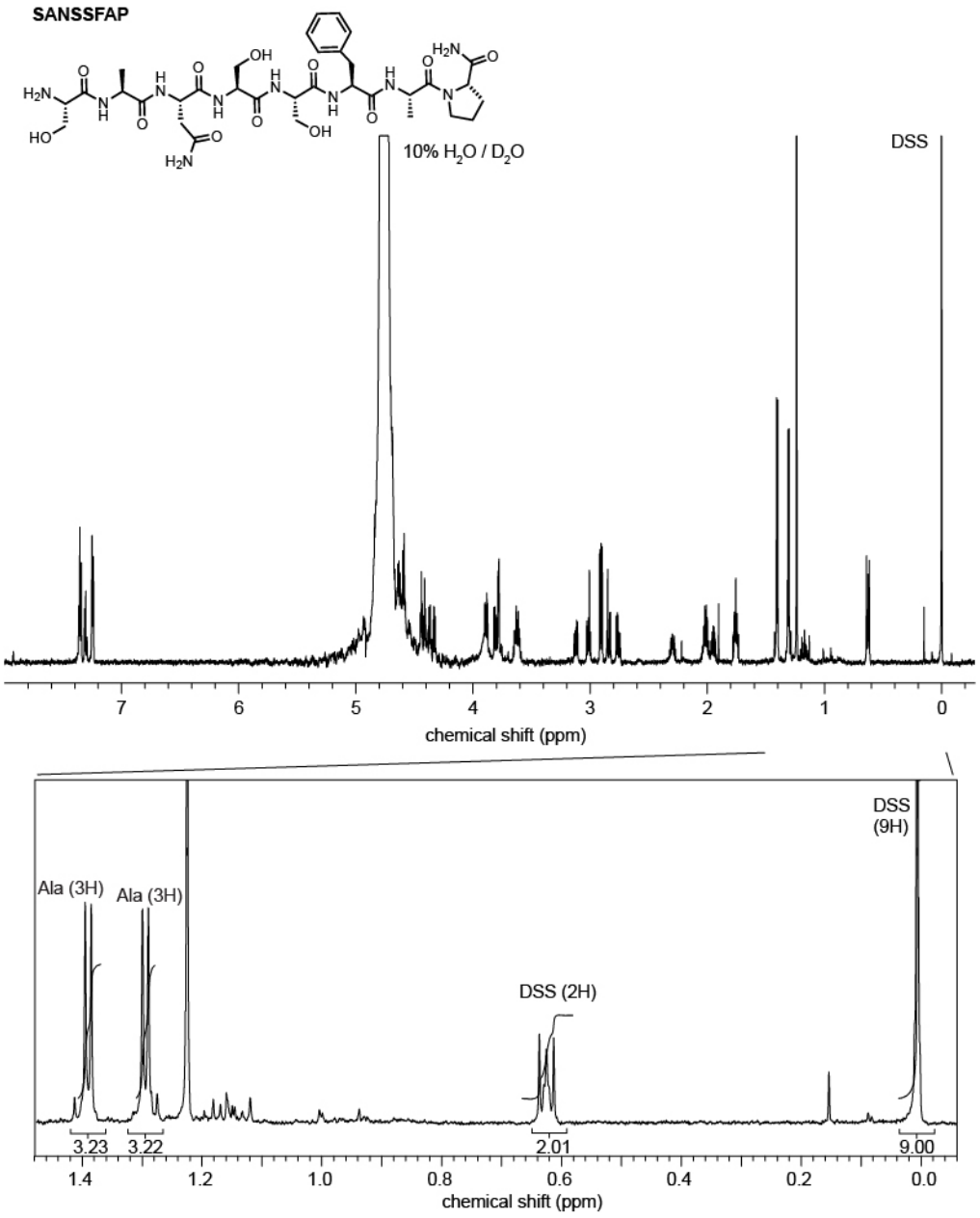
NMR spectra of Ara6-peptide conjugates and peptides

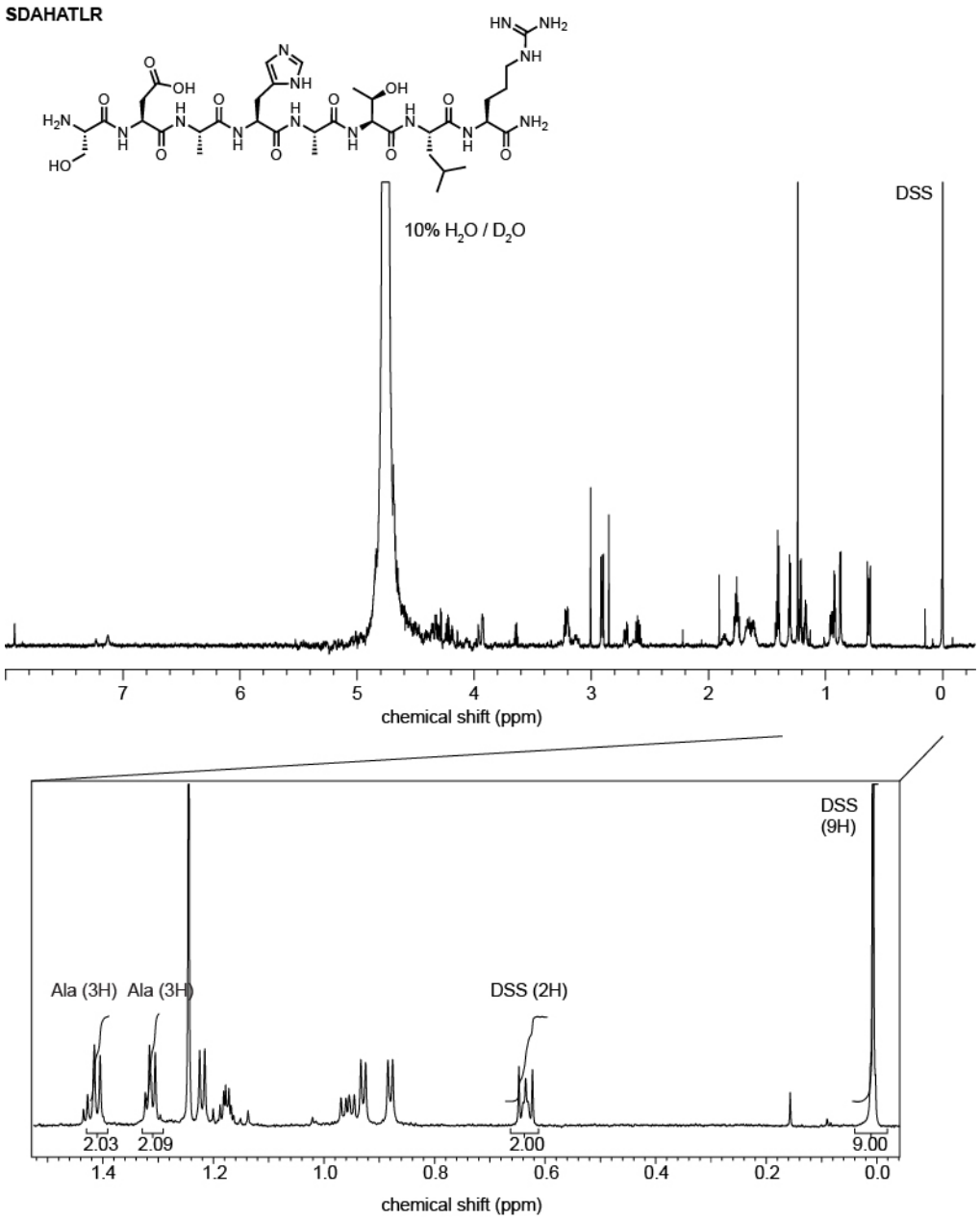



Ara6-DAHATLR
10% H2O / D2O
DSS
chemical shift (ppm)
DSS
(9H)
Ala (3H)
Ala (3H)
DSS (2H)
7.56
7.93
2.03
9.00
chemical shift (ppm)

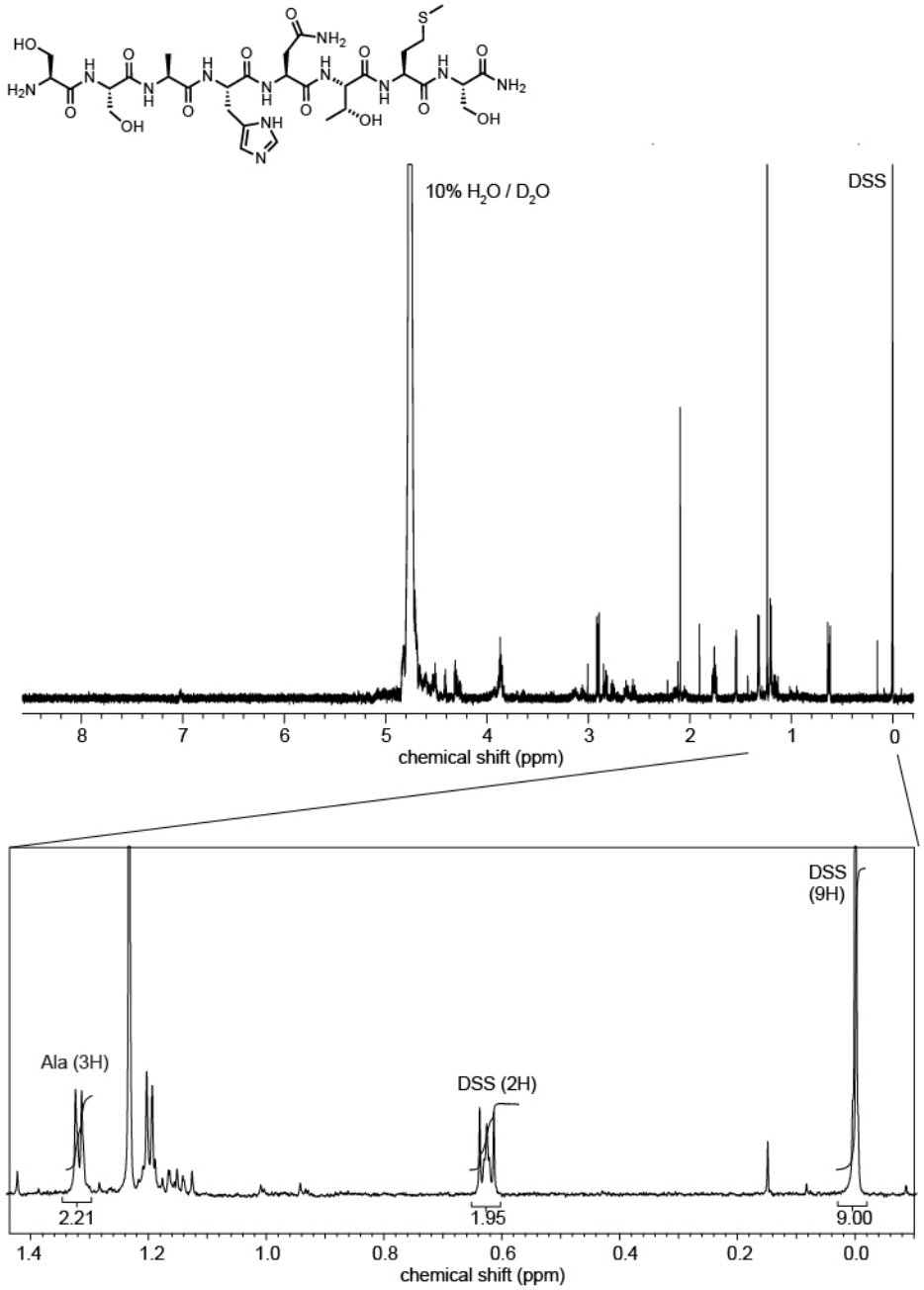

Ara6-TTYVVNP
DSS
10% H2O / D2O
chemical shift (ppm)
DSS
(9H)
Thr x 2 (6H)
DSS (2H)
5.49
1.91
9.00
chemical shift (ppm)


Ara6-SAHNTMS
10% H2O / D2O
DSS
chemical shift (ppm)
ppm
DSS (9H)
DSS (2H)
Ala (3H)
0.53
2.04
9.00
chemical shift (ppm)

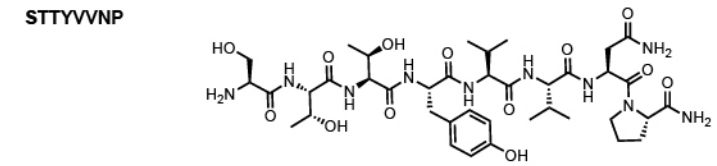

SANSSFAP
10% H2O / D2O
DSS
chemical shift (ppm)
Ala (3H)
Ala (3H)
DSS (2H)
DSS
(9H)
3.23
3.22
2.01
9.00
chemical shift (ppm)

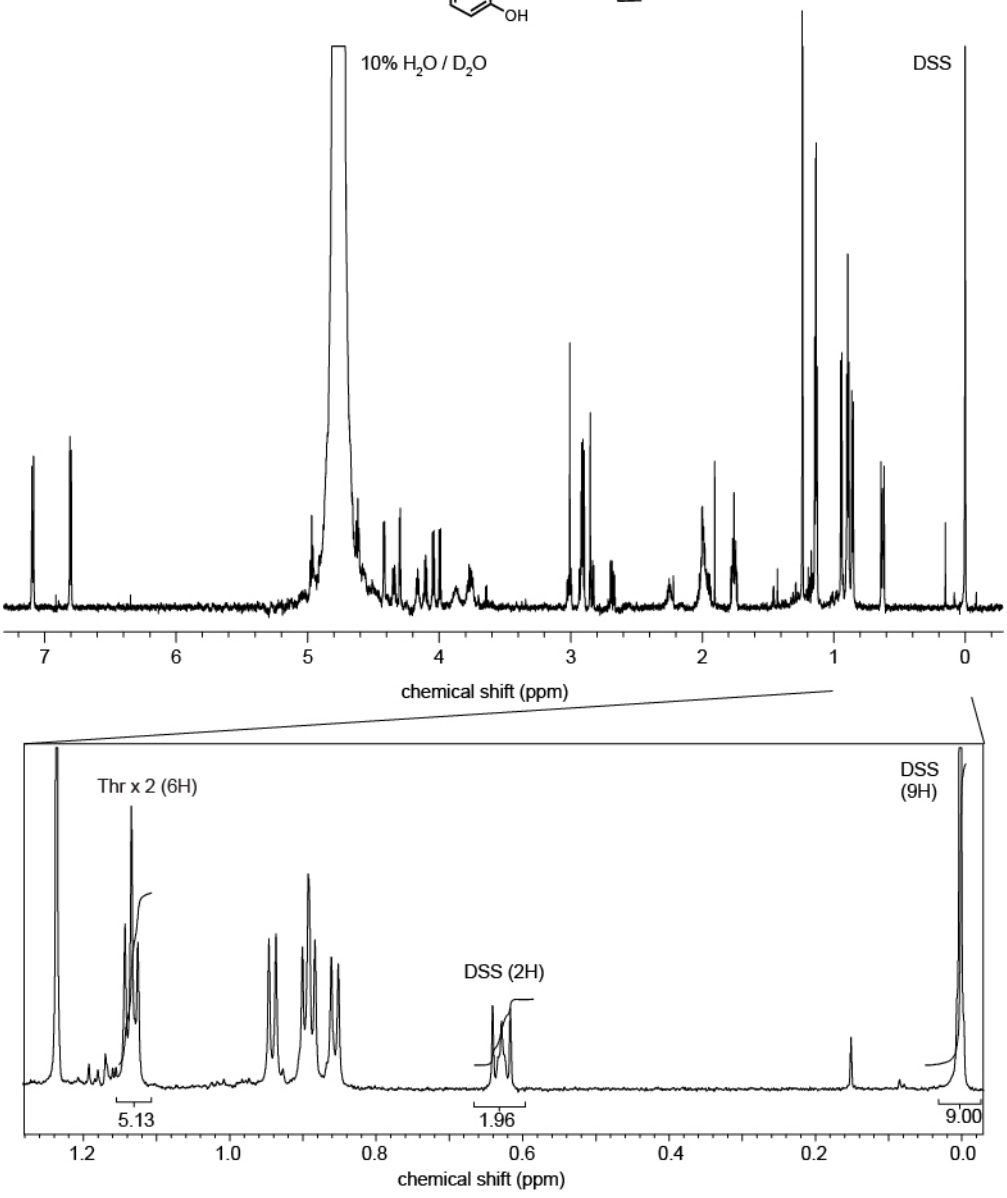

SDAHATLR
10% H2O / D2O
DSS
chemical shift (ppm)
DSS
(9H)
Ala (3H)
Ala (3H)
DSS (2H)
2.03
2.09
2.00
9.00
chemical shift (ppm)

SSAHNTMS


10% H2O / D2O
DSS
chemical shift (ppm)
DSS
(9H)
Ala (3H)
DSS (2H)
2.21
1.95
9.00
chemical shift (ppm)

STTYVVNP




10% H2O / D2O
DSS
chemical shift (ppm)
Thr x 2 (6H)
DSS
(9H)
DSS (2H)
5.13
1.96
9.00
chemical shift (ppm)

Supporting information references

1. C. Rademacher, G. K. Shoemaker, H.-S. Kim, R. B. Zheng, H. Taha, C. Liu, R. C. Nacario, D. C. Schriemer, J. S. Klassen, T. Peters and T. L. Lowary, *J. Am. Chem. Soc.,* **2007,** 10489-10502.

2. F. W. D'Souza and T. L. Lowary, *Org. Lett.,* **2000,** 1493-5.

3. K. A. Noren and C. J. Noren, *Methods,* **2001,** 169-178.

4. S. Ng, M. R. Jafari, W. L. Matochko and R. Derda, *ACS Chem. Biol.,* **2012,** 1482-1487.

5. K. F. Tjhung, P. I. Kitov, S. Ng, E. N. Kitova, L. Deng, J. S. Klassen and R. Derda, *J. Am. Chem. Soc.,* **2016,** 32-35.

6. S. Ng, K. Tjhung, B. Paschal, C. Noren and R. Derda, Chemical Posttranslational Modification of Phage-Displayed Peptides. In *Peptide Libraries*, Derda, R., Ed. Springer New York: 2015; Vol. 1248, pp 155-172.

7. W. L. Matochko, S. Cory Li, S. K. Y. Tang and R. Derda, *Nucleic Acids Res.,* **2013**.

8. F. Deiss, W. L. Matochko, N. Govindasamy, E. Y. Lin and R. Derda, *Angew. Chem. Int. Ed.,* **2014,** 6374-6377.

9. J. R. Clamp and L. Hough, *Biochem. J,* **1965,** 17-24.

10. E. N. Kitova, A. El-Hawiet, P. D. Schnier and J. S. Klassen, *J. Am. Soc. Mass. Spectrom.,* **2012,** 431-441.

11. M. Joe, Y. Bai, R. C. Nacario and T. L. Lowary, *J. Am. Chem. Soc.,* **2007,** 9885-9901.

12. S.-J. Wey, M. E. Augustyniak, E. D. Cochran, J. L. Ellis, X. Fang, D. S. Garvey, D. R. Janero, L. G. Letts, A. M. Martino, T. L. Melim, M. G. Murty, S. K. Richardson, J. D. Schroeder, W. M. Selig, A. M. Trocha, R. S. Wexler, D. V. Young, I. S. Zemtseva and B. M. Zifcak, *J. Med. Chem.,* **2007,** 6367-6382.

13. M. Joe, Y. Bai, L. Gandolfi-Donadío and T. Lowary, p-Tolyl 2,3,5-Tri-O-Benzoyl-1-Thio-α-d-Arabinofuranoside. In *Carbohydrate Chemistry*, CRC Press: 2011; pp 341-348.