Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry
This journal is © The Royal Society of Chemistry 2011

Supporting Information

Fayna Garcia-Martin, Hiroshi Hinou, Takahiko Matsushita, Shun Hayakawa and Shin-Ichiro Nishimura
Graduate School of Advanced Life Science, Hokkaido University, Sapporo, Japan, and Frontier Research Center for Post-genome Science and Technology, Hokkaido University, Sapporo, Japan
shin@glyco.sci.hokudai.ac.jp

Contents

page

- Characterization of H-Ser(Ac_3-α-GalNAc)-Thr(Ac_3-α-GalNAc)-Ala-NH_2 (**2**) S2~S5

- Characterization of MUC-1 fragments (**9**) and (**10**) S6~S10

- Characterization of tau glycopeptides (**11**), (**12**), and (**13**) S11~S21

Characterization of H-Ser(Ac$_3$-α-GalNAc)-Thr(Ac$_3$-α-GalNAc)-Ala-NH$_2$ (2)

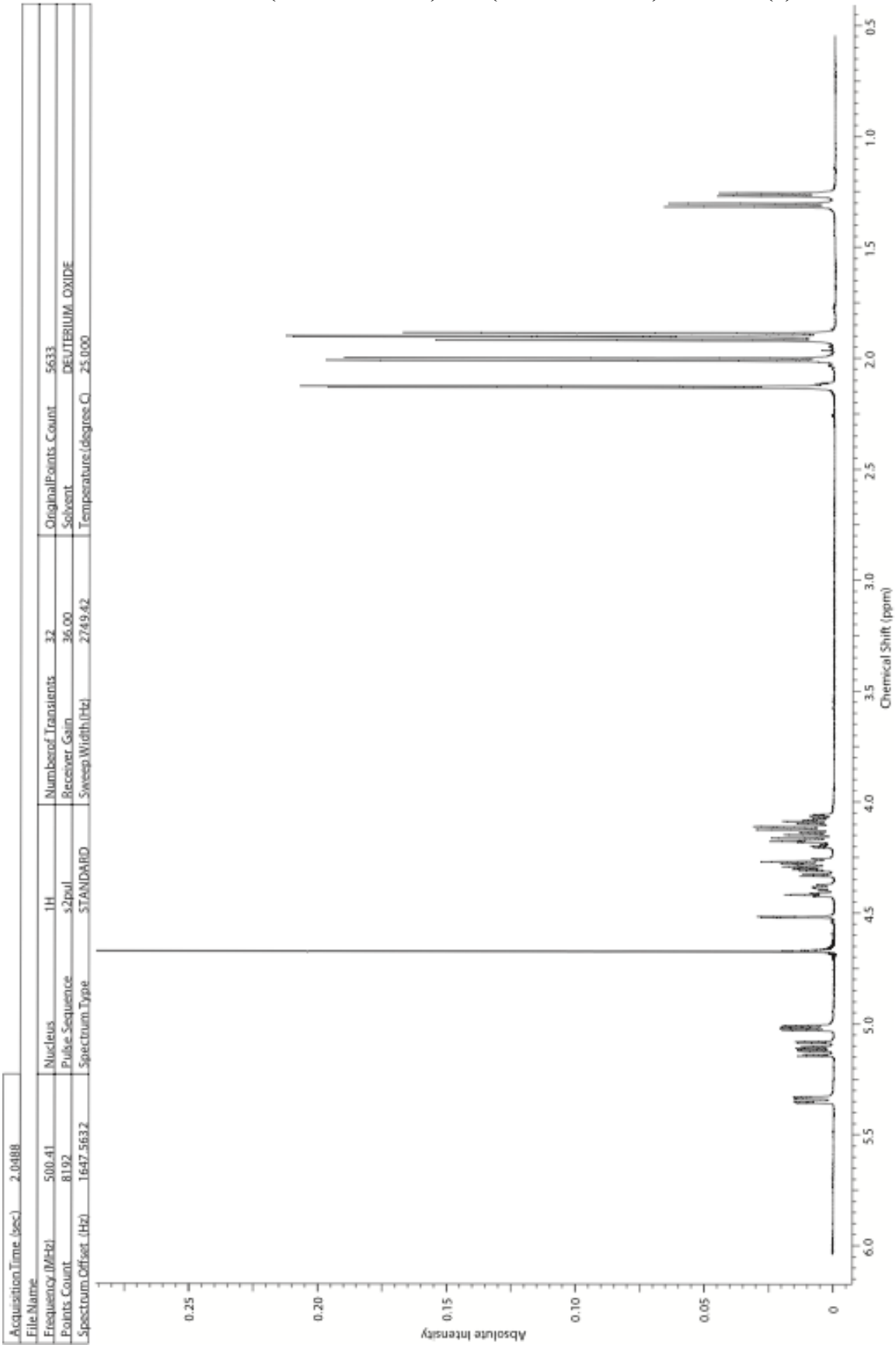


Figure S1. 500MHz ^{1}H spectrum of H-Ser(Ac$_3$-α-GalNAc)-Thr(Ac$_3$-α-GalNAc)-Ala-NH$_2$ (**2**) in D$_2$O

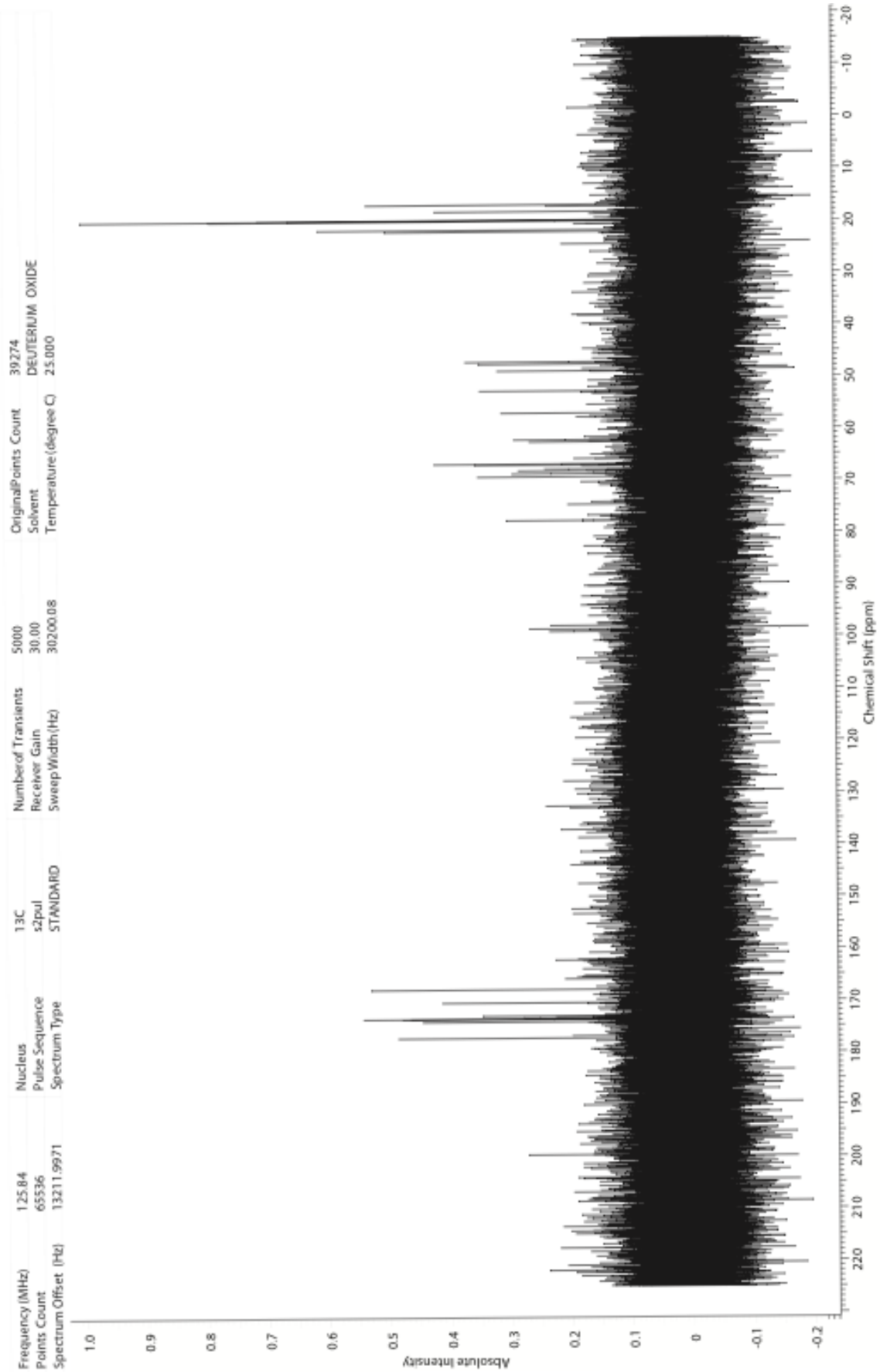


Figure S2. 125 MHz ^{13}C spectrum of H-Ser(Ac_3-α-GalNAc)-Thr(Ac_3-α-GalNAc)-Ala-NH_2 (**2**) in D_2O

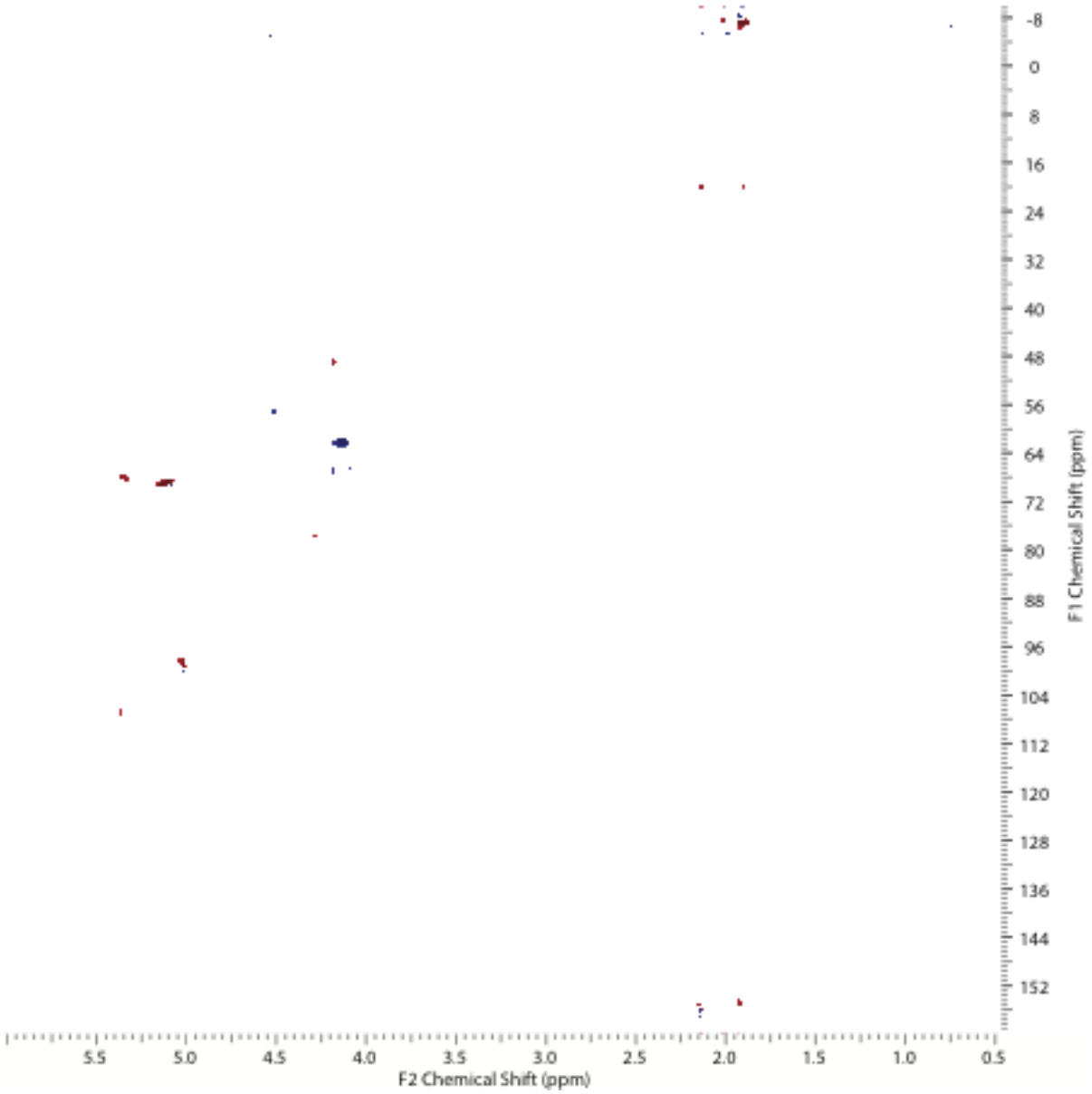


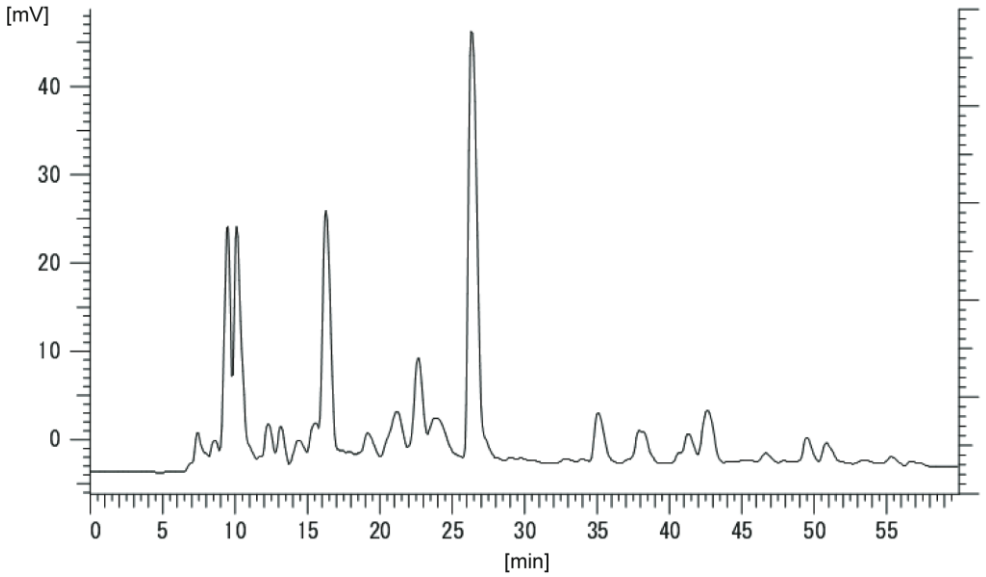
Figure S3. HSQC spectrum of H-Ser(Ac_3-α-GalNAc)-Thr(Ac_3-α-GalNAc)-Ala-NH_2 (**2**) in D_2O

Table S1. 500MHz ^{1}H and 125 MHz ^{13}C NMR Chemical Shifts (ppm) of the H-Ser(Ac_3-α-GalNAc)-Thr(Ac_3-α-GalNAc)-Ala-NH_2 (**2**) in D_2O.

Residues	**1H**	**shift (ppm)**	**13C**	**shift (ppm)**
Ala	αH	1.31	αC	18.3
	βH	4.18	βC	48.7
Thr	αH	4.52	αC	56.9
	βH	4.28	βC	77.7
	γH	1.26	γC	18.3
Ser	αH	4.42	αC	45.2
	βH	4.07	βC	66.5
		4.15		
GalNAc I (Thr)				
	H-1	5.02	C-1	98.1
	H-2	4.30	C-2	67.1
	H-3	5.10	C-3	68.5
	H-4	5.34	C-4	67.8
	H-5	4.38	C-5	65.8
	H-6	4.10	C-6	62.2
		4.16		
	C*H*$_3$CO	1.89	*C*H$_3$CO	19.6
		1.91		22.0
		2.01		19.7
		2.13		19.9
GalNAc II (Ser)				
	H-1	5.01	C-1	98.7
	H-2	4.29	C-2	67.1
	H-3	5.13	C-3	69.1
	H-4	5.35	C-4	67.5
	H-5	4.26	C-5	66.6
	H-6	4.07	C-6	61.9
		4.12		
	C*H*$_3$CO	1.89	*C*H$_3$CO	19.6
		1.91		22.0
		2.01		19.7
		2.13		19.9

Characterization of mucin glycopeptides

(A)



(B)

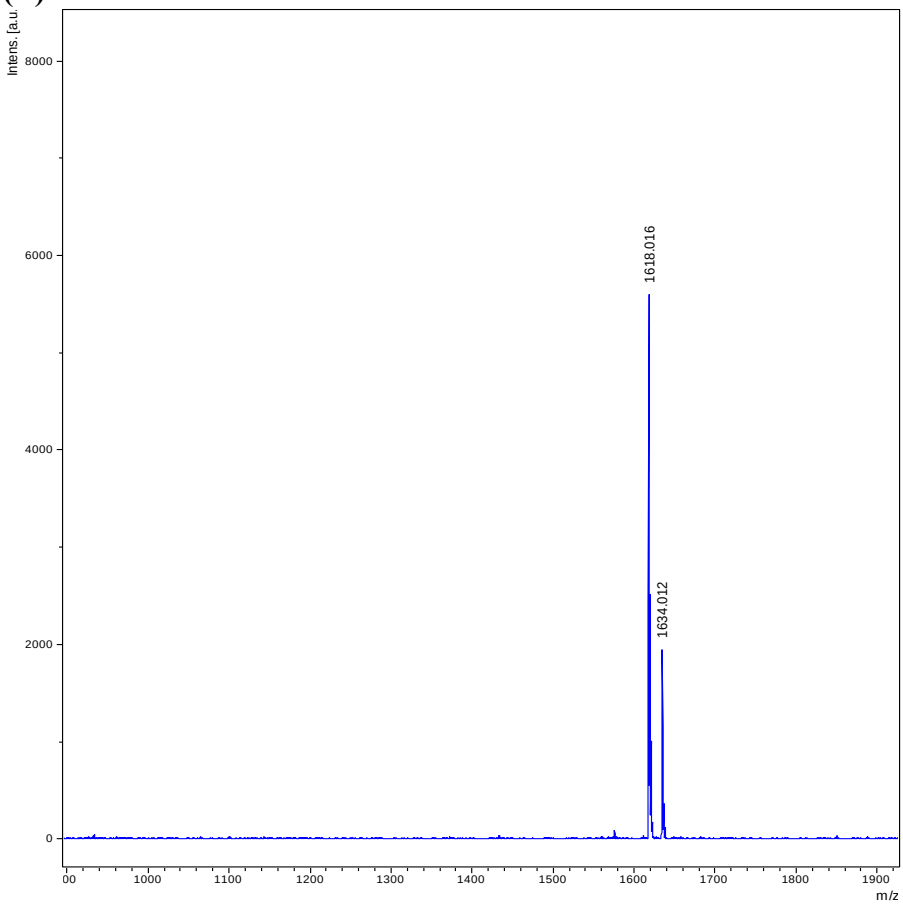


Figure S4. Characterization of Thr(core1)-Thr(core1)-Val (**9**). **(A)** RP-HPLC (Linear gradient from 0.4:0.6 to 0.6:0.4 of CH_3CN (0.1% TFA) in H_2O (0.1% TFA) for 50 min). **(B)** MALDI-TOF MS of isolated compound ((*m/z* calcd. 1594. 6; found 1618.0 $[M+Na]^+$, 1634.0 $[M+K]^+$)

Table S2. 500MHz ^{1}H and 125 MHz ^{13}C NMR Chemical Shifts (ppm) of Thr(core1)-Thr(core1)-Val (**9**) in D_2O.

residues	nuclei	ppm
Thr1	NH	8.435
	Hα	4.587
	Hβ	4.281
	γCH3	1.260
	Cα	57.714
	Cβ	76.866
	Cγ	18.408
Thr2	NH	8.806
	Hα	4.599
	Hβ	4.228
	γCH3	1.171
	Cα	56.526
	Cβ	77.261
	Cγ	18.498
Val3	NH	8.227
	Hα	4.002
	Hβ	1.921
	γCH3	0.866
	Cα	59.051
	Cβ	30.228
	Cγ	18.147

residues	nuclei	ppm
GalNAc (Thr1)	H-1	4.374
	H-2	4.153
	H-3	3.954
	H-4	4.122
	H-5	3.996
	H-6a	3.665
	H-6b	3.665
	C-1	104.633
	C-2	48.407
	C-3	77.089
	C-4	68.871
	C-5	71.009
	C-6	61.149
	CH3	1.933
Gal (Thr1)	H-1	4.760
	H-2	3.418
	H-3	3.527
	H-4	3.821
	H-5	3.561
	H-6a	3.665
	H-6b	3.665
	C-1	99.165
	C-2	70.484
	C-3	72.501
	C-4	68.563
	C-5	75.047
	C-6	61.149

residues	nuclei	ppm
GalNAc (Thr1)	H-1	4.374
	H-2	4.153
	H-3	3.954
	H-4	4.122
	H-5	3.996
	H-6a	3.665
	H-6b	3.665
	C-1	104.633
	C-2	48.407
	C-3	77.089
	C-4	68.871
	C-5	71.009
	C-6	61.149
	CH3	1.949
Gal (Thr1)	H-1	4.760
	H-2	3.418
	H-3	3.527
	H-4	3.821
	H-5	3.561
	H-6a	3.665
	H-6b	3.665
	C-1	99.165
	C-2	70.484
	C-3	72.501
	C-4	68.563
	C-5	75.047
	C-6	61.149

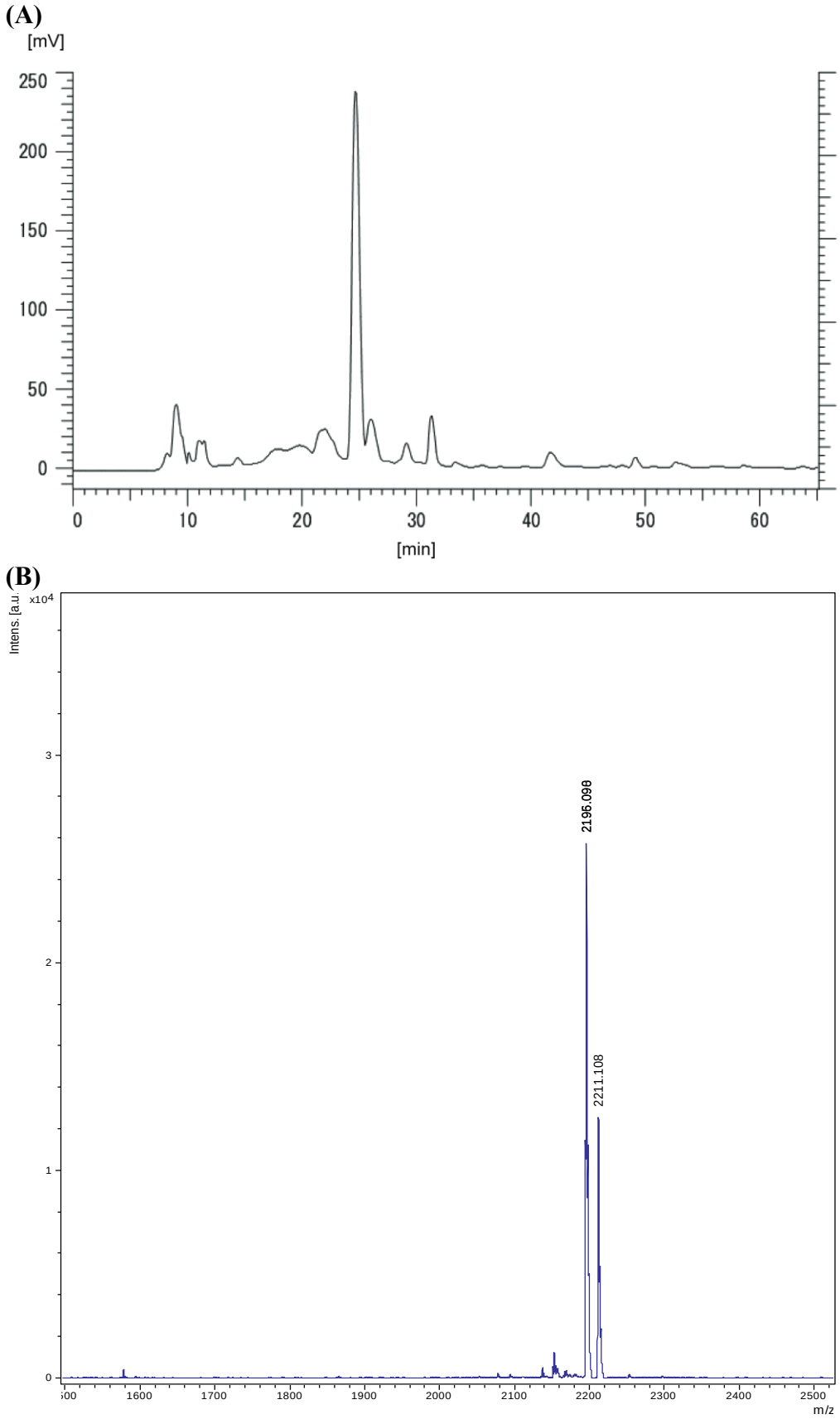


Figure S5. Characterization of Ac-TTS<u>TT</u>VSTT-NH_2 (**10**). **(A)** RP-HPLC (Linear gradient from 0.3:0.7 to 0.6:0.4 of CH_3CN (0.1% TFA) in H_2O (0.1% TFA) for 50 min). **(B)** MALDI-TOF MS of isolated compound ((*m/z* calcd. 2172.8; found 2196.1 $[M+Na]^+$, 2211.1 $[M+K]^+$)

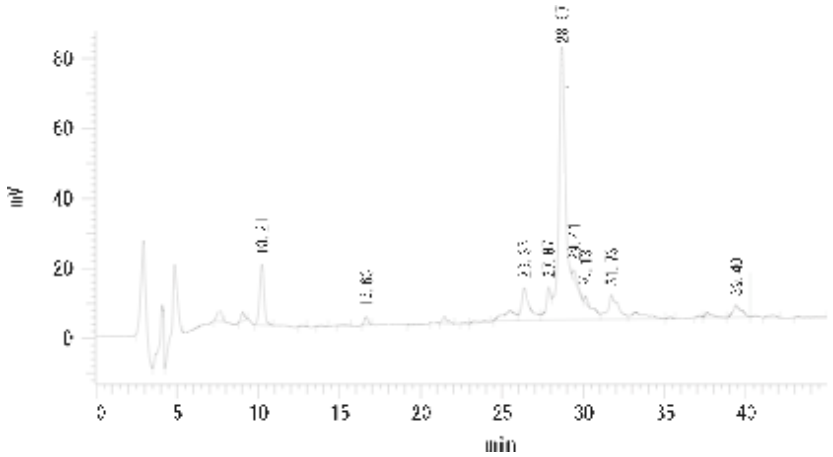
Table S3. 500MHz ^{1}H and 125 MHz ^{13}C NMR Chemical Shifts (ppm) of Ac-TTSTTVSTT-NH_2 (**10**) in D_2O.

residues	nuclei	ppm	residues	nuclei	ppm
Thr1	NH	8.161	Val6	NH	8.186
	Hα	4.298		Hα	4.140
	Hβ	4.127		Hβ	1.978
	γCH3	1.125		γCH3	0.819
	Cα	59.150		Cα	58.240
	Cβ	66.944		Cβ	30.659
	Cγ	18.740		Cγ	17.798
Thr2	NH	8.149	Ser7	NH	8.375
	Hα	4.349		Hα	4.386
	Hβ	4.159		Hβ	3.827
	γCH3	1.134			3.736
	Cα	58.937		Cα	55.425
	Cβ	66.944		Cβ	61.061
	Cγ	18.740			61.039
Ser3	NH	8.326	Thr8	NH	8.356
	Hα	4.592		Hα	4.338
	Hβ	3.811		Hβ	4.232
	Cα	55.425		γCH3	1.144
	Cβ	61.061		Cα	56.937
				Cβ	66.732
Thr4	NH	8.632		Cγ	18.740
	Hα	4.676			
	Hβ	4.317	Thr9	NH	8.008
	γCH3	1.222		Hα	4.238
	Cα	57.123		Hβ	4.172
	Cβ	77.262		γCH3	1.109
	Cγ	18.040		Cα	58.706
				Cβ	66.907
Thr5	NH	8.744		Cγ	18.740
	Hα	4.572			
	Hβ	4.231			
	γCH3	1.149			
	Cα	56.310			
	Cβ	76.769			
	Cγ	18.740			

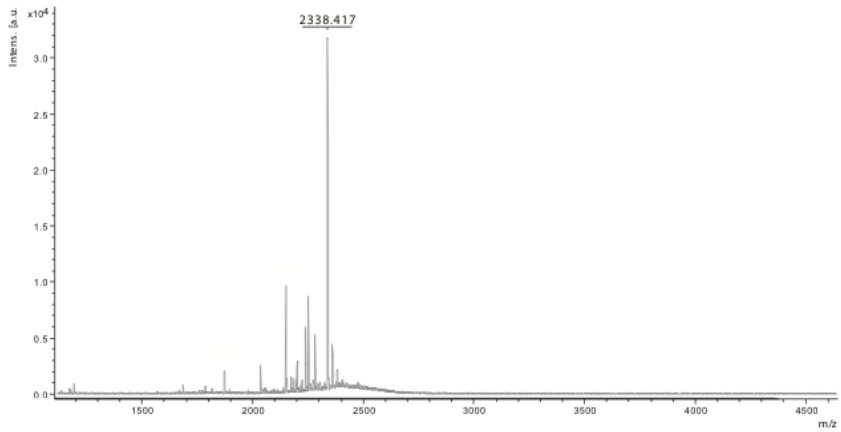
residues	nuclei		ppm			residues	nuclei		ppm
GalNAc (Thr4)	H-1		4.740			GalNAc (Thr5)	H-1		4.740
	H-2		4.155				H-2		4.155
	H-3		3.943				H-3		3.943
	H-4		4.112				H-4		4.112
	H-5		3.986				H-5		3.986
	H-6a		3.659				H-6a		3.659
	H-6b		3.659				H-6b		3.659
	C-1		99.149				C-1		99.149
	C-2		48.258				C-2		48.258
	C-3		77.205				C-3		77.205
	C-4		68.837				C-4		68.837
	C-5		70.912				C-5		70.912
	C-6		61.063				C-6		61.063
	CH3		1.923				CH3		1.941
Gal (Thr4)	H-1		4.343			Gal (Thr5)	H-1		4.343
	H-2		3.406				H-2		3.406
	H-3		3.508				H-3		3.508
	H-4		3.801				H-4		3.801
	H-5		3.539				H-5		3.539
	H-6a		3.651				H-6a		3.651
	H-6b		3.651				H-6b		3.651
	C-1		104.770				C-1		104.770
	C-2		70.628				C-2		70.628
	C-3		72.461				C-3		72.461
	C-4		68.525				C-4		68.525
	C-5		74.979				C-5		74.979
	C-6		61.063				C-6		61.063

Characterization of tau glycopeptides.

(A)



(B)



(C)

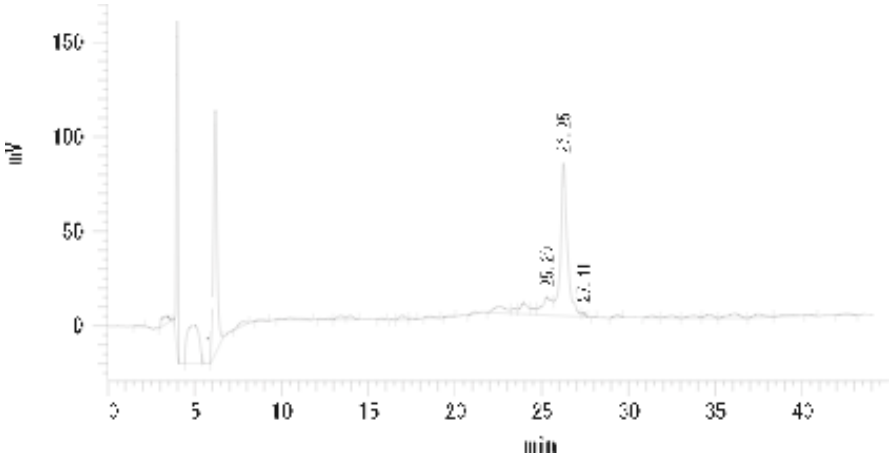


Figure S6. Characterization of His[388] to His[407] (**11**): NH_2 HGAEIVYK**S**(βGlcNAc)PVVSGDTSPRH-NH_2 **(A)** RP-HPLC (Linear gradient from 0.3:9.7 to 2.5:7.5 of CH_3CN (0.1% TFA) in H_2O (0.1% TFA) for 40 min, 1.0 mL/min), (B) MALDI-TOFMS, and (C) Analytical RP-HPLC of isolated glycopeptide, same analytical conditions as described above.

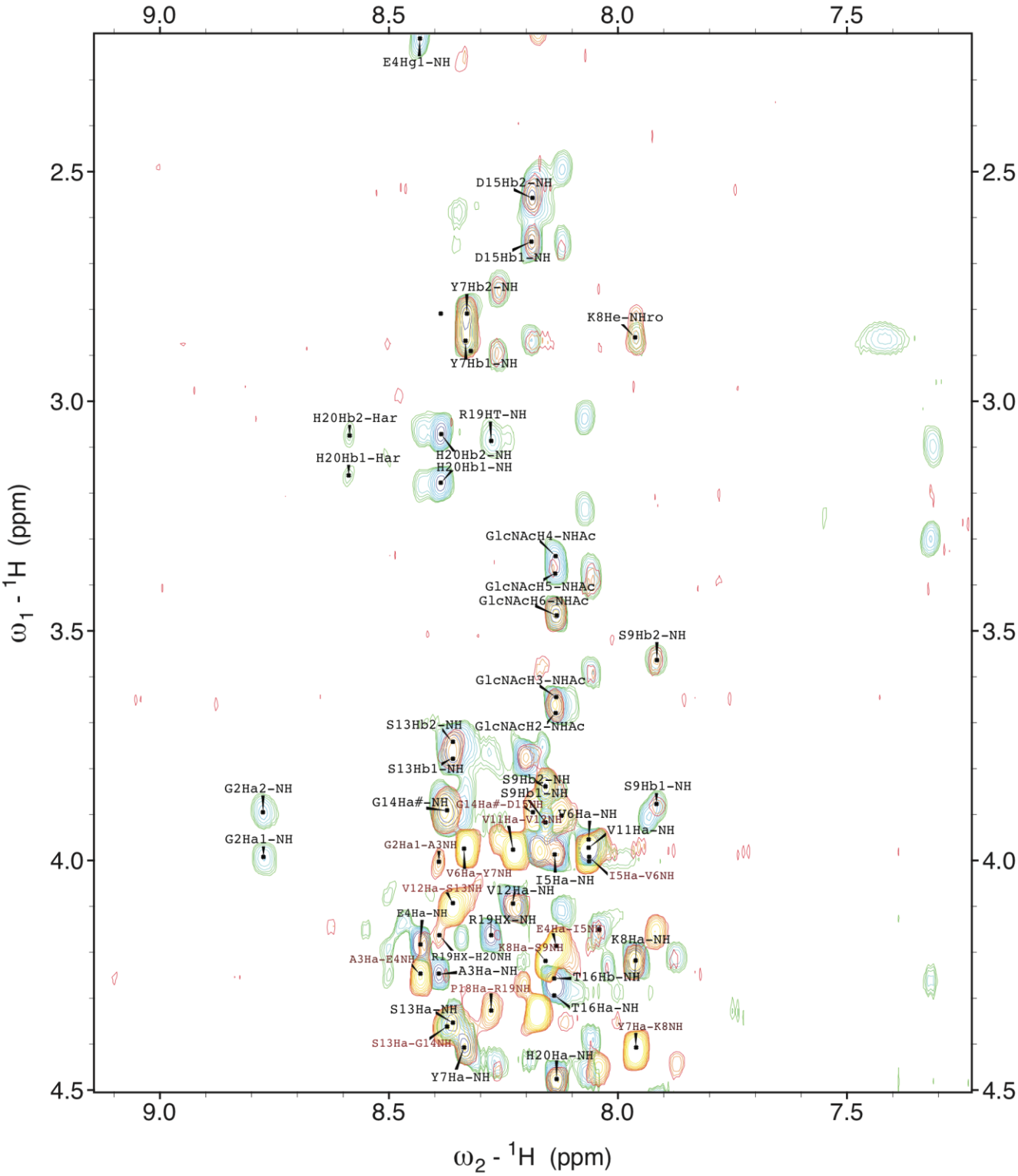


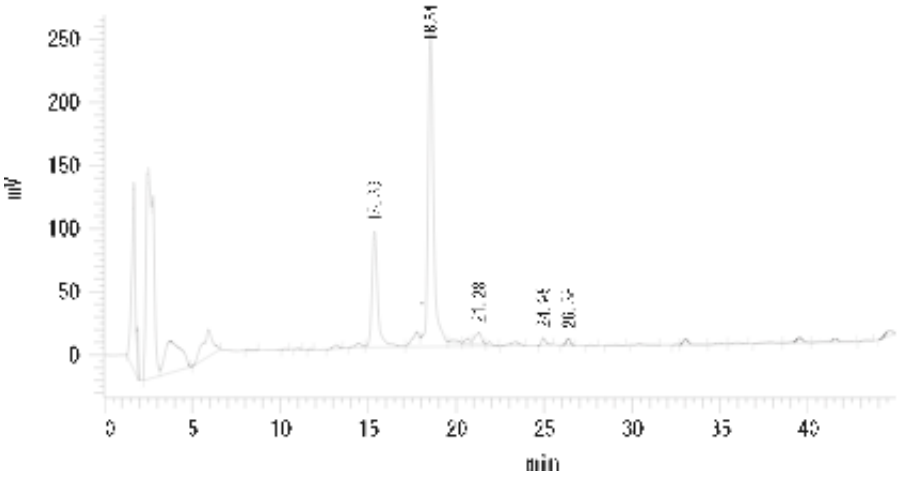
Figure S7. NMR fingerprint region of glycopeptide **11**. Overlay of TOCSY (blue-green) and NOESY (red-yellow).

Table S4. 600 MHz 1H-NMR chemical shifts (ppm) of amino acids moieties of compound **11** at 300 K.

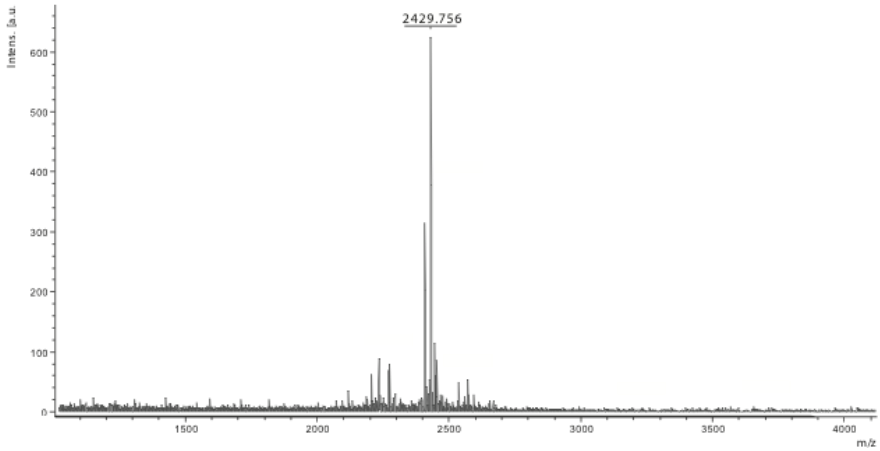
Residues	H type	shift (ppm)
His[1]	NH	8.34
	αH	4.17
	βH	3.70
		3.89
Gly[2]	NH	8.77
	αH	3.89
		4.00
Ala[3]	NH	8.39
	αH	4.25
	βH	1.28
Glu[4]	NH	8.43
	αH	4.18
	βH	1.82
		1.91
	γH	2.16
		2.21
Ile[5]	NH	3.98
	αH	8.14
	βH	1.64
	γCH_3	0.73
	γH	1.34
	δCH_3	0.59
Val[6]	NH	8.23
	αH	4.10
	βH	1.97
	γH	0.80
		0.85
Tyr[7]	NH	8.34
	αH	4.41
	βH	2.81
		2.87
Lys[8]	NH	7.96
	αH	4.22
	δCH_2	2.86
	βH	1.53
		1.63
	γCH_2	1.25
Ser[9]	NH	8.16
	αH	4.57
	βH	3.84
		3.92
GlcNAc I (Ser[9])	NH	8.14
	H-1	4.47
	H-2	3.66
	H-3	3.55
	H-4	3.33
	H-5	3.37
	H-6	3.46

residues	H type	shift (ppm)
Pro[10]	αH	4.32
	βH	1.78
		2.18
	γH	1.93
	δH	3.36
		3.40
Val[11]	NH	8.17
	αH	3.98
	βH	1.91
	γH	0.83
Val[12]	NH	8.06
	αH	3.97
	βH	1.86
	γH	0.77
Ser[13]	NH	8.36
	αH	4.36
	βH	3.76
		3.90
Gly[14]	NH	8.38
	αH	3.86
		3.92
Asp[15]	NH	8.19
	αH	4.60
	βH	2.56
		2.65
Thr[16]	NH	8.14
	αH	4.27
	βH	3.30
	γH	1.10
Ser[17]	NH	8.20
	αH	4.59
	βH	3.76
		3.80
Pro18	αH	4.33
	βH	1.76
		2.18
	γH	1.90
	δH	3.58
		3.74
Arg19	NH	8.28
	αH	4.16
	βH	1.51
	γH	1.63
		1.70
	δH	3.09
His[20]	NH	7.92
	αH	4.52
	βH	3.84
		3.91

(A)



(B)



(C)

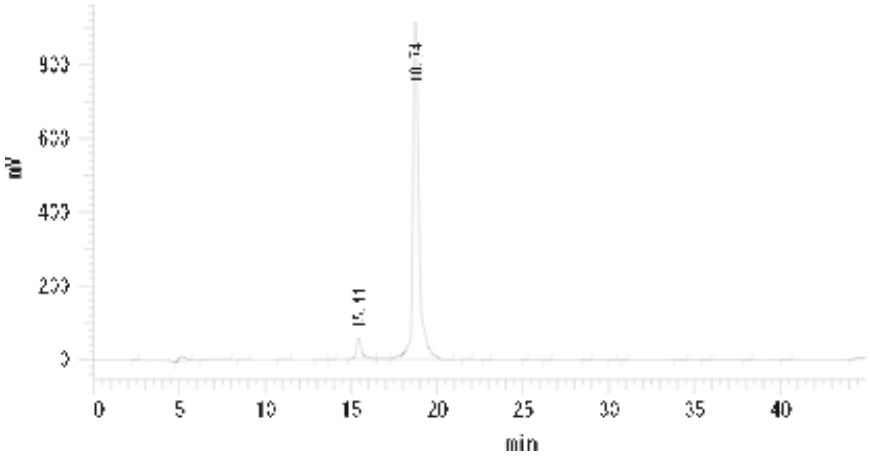


Figure S8. Characterization of Leu[408] to Ala[430] (**12**): NH_2-LSNVSSTGSIDMVDS(βGlcNAc)PQLATLAD-NH_2 (A) RP-HPLC (Linear gradient from 0.3:9.7 to 2.5:7.5 of CH_3CN (0.1% TFA) in H_2O (0.1% TFA) for 40 min, 1.0 mL/min), (B) MALDI-TOFMS, and (C) Analytical RP-HPLC of isolated glycopeptide, same analytical conditions as described above.

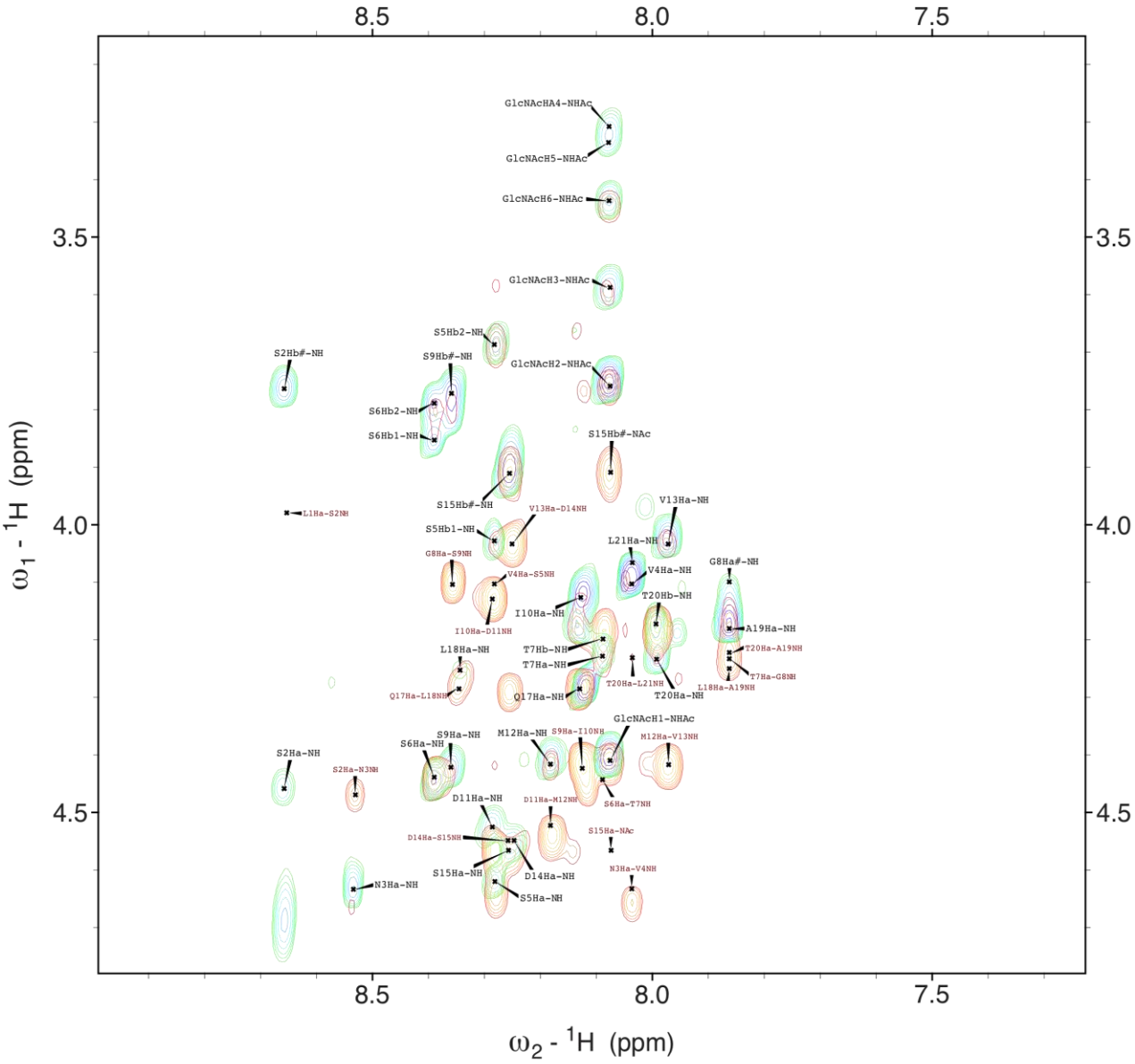


Figure S9. NMR fingerprint region of **12**. Overlay of TOCSY (blue-green) and NOESY (red-yellow) confirming the desired compound.

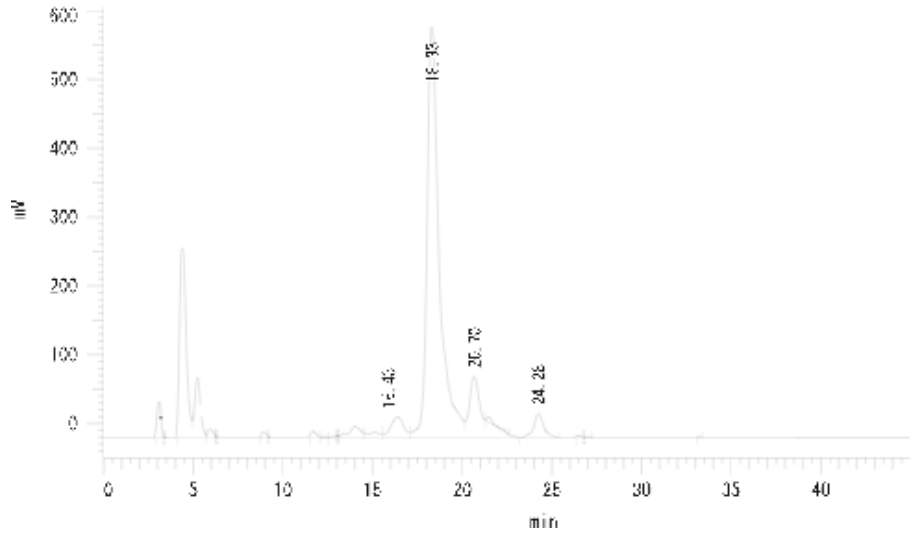
Table S5. 600 MHz 1H-NMR chemical shifts (ppm) of amino acids moieties of compound **12** at 300 K

residues	H type	shift (ppm)
Leu[1]	αH	3.98
	βH	1.64
	γH	1.57
	δH	0.86
Ser[2]	NH	8.66
	αH	4.46
	βH	3.76
		3.85
Asn[3]	NH	
	αH	4.63
	βH	2.69
		2.73
Val[4]	NH	8.04
	αH	4.01
	βH	1.99
	γH	0.82
Ser[5]	NH	8.28
	αH	4.62
	βH	4.03
		3.69
Ser[6]	NH	8.39
	αH	4.44
	βH	3.79
		3.85
Thr[7]	NH	8.09
	αH	4.23
	βH	4.20
	γH	1.31
Gly[8]	NH	7.86
	αH	4.11
		4.19
Ser[9]	NH	8.36
	αH	4.42
	βH	3.76
		3.81
Ile[10]	NH	8.13
	αH	4.13
	βH	1.77
	γCH_3	1.11
	γH	1.30
	δCH_3	0.79
Asp[11]	NH	8.29
	αH	4.53
	βH	2.53
		2.65
GlcNAc I (Ser[15])	NH	8.08
	H-1	4.41
	H-2	3.76
	H-3	3.59
	H-4	3.31

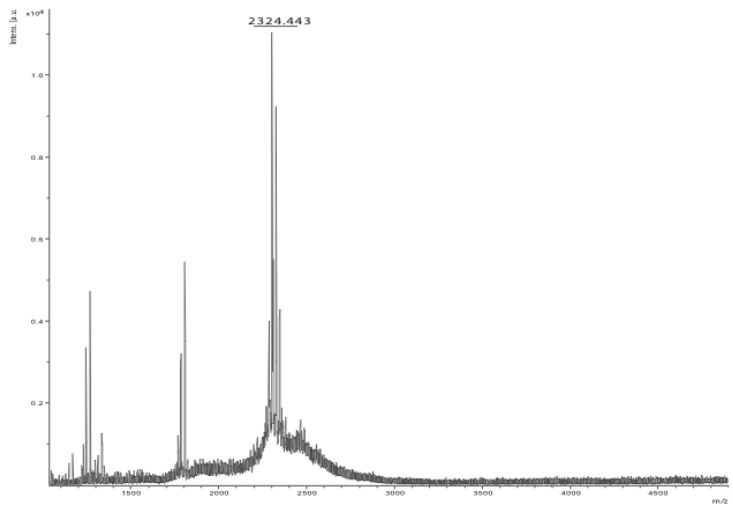
residues	H type	shift (ppm)
Met[12]	NH	8.18
	αH	4.42
	βH	1.90
		2.00
	γH	2.41
		2.49
	δH	3.58
Val[13]	NH	7.97
	αH	4.03
	βH	1.98
	γH	0.81
Asp[14]	NH	8.25
	αH	4.55
	βH	2.53
		2.65
Ser[15]	NH	8.26
	αH	4.56
	βH	3.87
		3.93
Pro[16]	αH	4.20
	βH	1.81
		2.21
	γH	1.88
	δH	3.59
		3.65
Gln[17]	NH	8.13
	αH	4.29
	βH	1.95
		2.02
	γH	2.30
Leu[18]	NH	8.34
	αH	4.25
	βH	1.54
		1.57
	γH	1.51
	δH	1.14
		1.32
Ala[19]	NH	7.86
	αH	4.18
	βH	1.11
Thr[20]	NH	7.99
	αH	4.23
	βH	4.17
	γH	1.53
Leu[21]	NH	8.03
	αH	4.07
	βH	1.31
	γH	1.27
	δH	1.19
Ala[22]	NH	7.96
	αH	4.19
	βH	1.10
Asp[23]	NH	8.09
	αH	4.29

H-5	3.33		βH	3.59
H-6	3.44			3.65

(A)



(B)



(C)

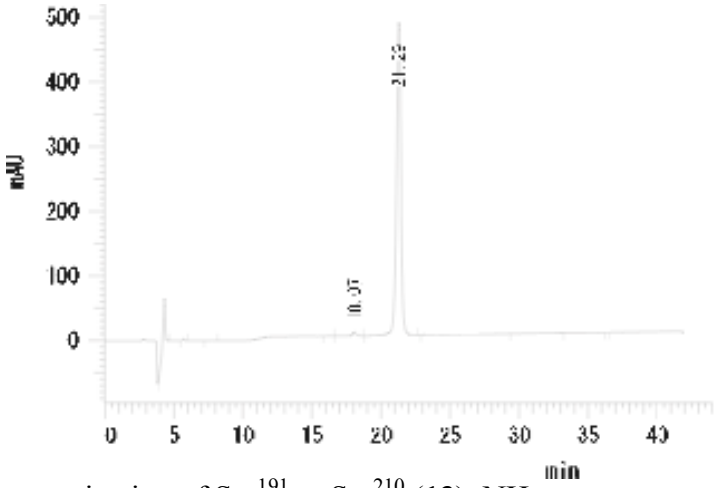


Figure S10. Characterization of Ser[191] to Ser[210] (**13**): NH_2, SGDRSGYSS(βGlcNAc)PGS(βGlcNAc)PGTPGSRS-NH_2 (A) RP-HPLC (Linear gradient from 0.3:9.7 to 2.5:7.5 of CH_3CN (0.1% TFA) in H_2O (0.1% TFA) for 40 min, 1.0 mL/min), (B) MALDI-TOFMS, and (C) Analytical RP-HPLC of isolated glycopeptide, same analytical conditions as described above.

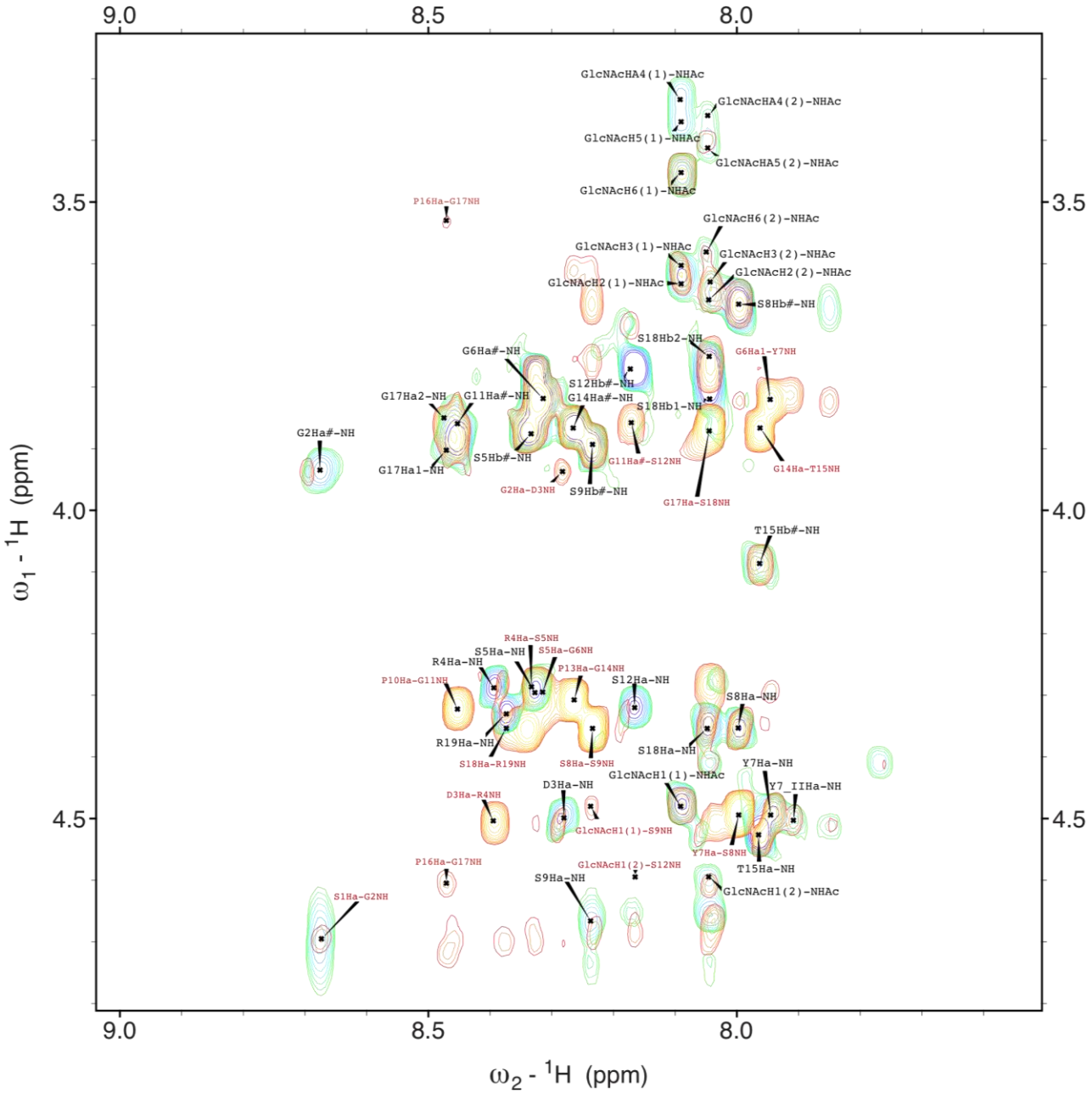


Figure S11. NMR fingerprint region of **13**. Overlay of TOCSY (blue-green) and NOESY (red-yellow) confirming the desired compound.

Table S6. 600 MHz 1H-NMR chemical shifts (ppm) of amino acids moieties of compound **13** at 300 K.

residues	H type	shift (ppm)
Ser[1]	αH	4.70
	βH	3.75
		3.63
Gly[2]	NH	8.28
	αH	3.93
Asp[3]	NH	8.28
	αH	4.50
	βH	2.56
		2.62
Arg[4]	NH	8.40
	αH	4.29
	βH	1.67
		1.82
	γH	1.54
	δH	3.06
Ser[5]	NH	8.33
	αH	4.30
	βH	3.78
		3.87
Gly[6]	NH	8.31
	αH	3.81
Tyr[7]	NH	7.94
	αH	4,49
	βH	2.86
		2.93
	2H	
	3H	
Ser[8]	NH	8.00
	αH	4.35
	βH	3.67
		3.79
Ser[9]	NH	8.24
	αH	4.67
	βH	3.87
		3.89
Pro[10]	αH	4.32
	βH	1.84
		2.23
	γH	1.91
	δH	3.60
		3.75
GlcNAc I (Ser[9])	NH	8.09
	H-1	4.48
	H-2	3.63
	H-3	3.60
	H-4	3.33
	H-5	3.37
	H-6	3.45

residues	H type	shift (ppm)
Gly[11]	NH	8.45
	αH	3.90
		3.86
Ser[12]	NH	8.17
	αH	4.32
	βH	3.75
		3.79
Pro[13]	αH	4.32
	βH	1.84
		2.23
	γH	1.91
	δH	3.60
		3.75
Gly[14]	NH	8.26
	αH	3.87
Thr[15]	NH	7.96
	αH	4.53
	βH	4.09
	γH	1.11
Pro[16]	αH	4.32
	βH	1.84
		2.23
	γH	1.91
	δH	3.60
		3.75
Gly[17]	NH	
	αH	4.50
Ser[18]	NH	8.05
	αH	4.35
	βH	3.75
		3.81
Arg[19]	NH	8.38
	αH	4.33
	βH	1.72
		1.87
	γH	1.58
	δH	3.11
Ser[20]	NH	
	αH	
	βH	
GlcNAc II (Ser[12])	NH	8.05
	H-1	4.60
	H-2	3.66
	H-3	3.63
	H-4	3.36
	H-5	3.41
	H-6	3.58

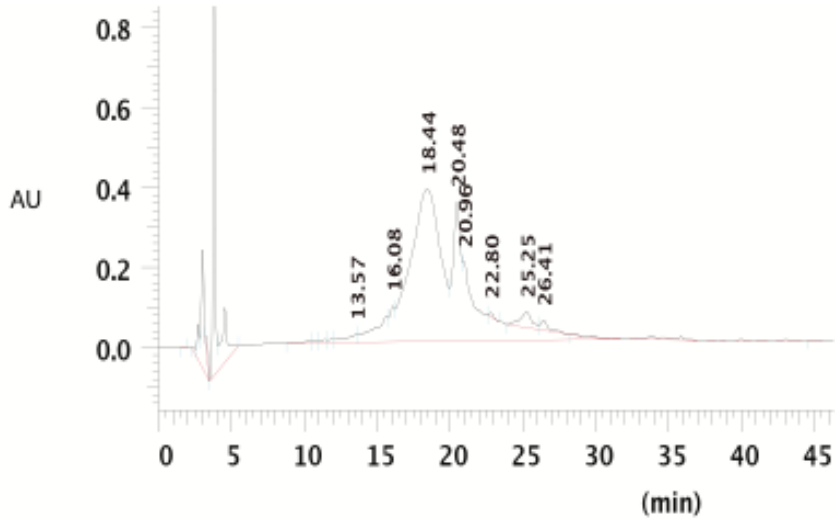


Figure S12. Purity of Ser[191] to Ser[210] (**13**) synthesized under a conventional HBTU conditions. RP-HPLC (Linear gradient from 0.5:9.5 to 2.5:7.5 of CH_3CN (0.1% TFA) in H_2O (0.1% TFA) for 40 min, 1.0 mL/min).