Supplement to "Cysteine S-linked N-acetylglucosamine (S-GlcNAcylation), a new post-translational modification in mammals."

Jason C. Maynard, Alma L Burlingame, Katalin F. Medzihradszky Department of Pharmaceutical Chemistry, School of Pharmacy, University of California San Francisco, 600 16th Street Genentech Hall, N474A, Box 2240, San Francisco, California 94158-2517, United States

# Legend

- We listed the UniProt accession number of the S-GlcNAcylated proteins, their name, the species they were isolated from, the glycopeptide sequence and the sequence position of the modified Cys-residue
- Both the beam-type CID spectrum (HCD) (upper panel) and the ETD data (lower panel) are presented for each S-GlcNAcylated peptide
- In most database searches an unspecified 203-203.1Da modification was permitted on Cys residues thus, the exact mass of the modification was reported by Protein Prospector and this information is listed in S Table 1. Here the modification is shown according to the final assignment, i.e. GlcNAc. However, in some later searches a user-specified exactly defined modification was introduced: Cgly that stands for a HexNAc on Cys.
- For the O-GlcNAcylated peptides the Prospector output was copied, where "HexNAc' was assigned as the modification. Since isomeric GalNAc and GlcNAc both may derivatize peptides this generic assignment is used in database searches.
- The measured precursor ion masses were included.
- In the HCD spectra we assigned just enough fragment ions to indicate that the sequence identification is reliable.
- In the ETD spectra we labeled the precursor ion and its charge-reduced forms (\*), it was also indicated (\*) if the charge reduced form of a coeluting ion of different charge in the precursor selection window was detected. The characteristic side-chain loss from the charge-reduced form(s) of the precursor ion is pointed out.
- Protein Prospector output, i.e. the masses used in the database search and their assignments, including the mass measurement error, is presented in the Tables. These assignments were manually implemented to indicate the w ion formation from the Cys-residues.

### S Figure 1A

Bassoon 088737

Mouse

Cys-1775

#### LDFGQGSGSPVC(GlcNAc)LAQVK



S Figure 1B

*In vitro* GlcNAcylated LDFGQGSGSPVC(GlcNAc)LAQVK

The modified peptide eluted slightly earlier than the unmodified sequence; fragment ions making the site-assignment unambiguous are labeled











088737

#### MC(GlcNAc)AALNSMDQYGGR



Bassoon	088737
Cys-2489	Mouse

## QKAPFPAT(GlcNAc)C(GlcNAc)EAPSR



## S Figure 4B

## QKAPFPAT(GlcNAc)C(GlcNAc)EAPSR

#### - Peak Matches

414.5430	430.6040	442.5660	543.4830	559.6430	579.5750	612.5930	613.5750	639.1620	647.6000
z4(0.32) z+14(-0.69)	y4(0.36)	c4(0.29)	z <sub>5</sub> (0.22) z+1 <sub>5</sub> (-0.79)	у <sub>5</sub> (0.36)			w6		
658.6090	686.6670	736.1520	757.7700	802.5220	822.7940	828.4890	829.3730	832.9140	837.0300
	c <sub>6</sub> (0.27)		c7(0.33)					c <sub>12</sub> <sup>+2</sup> (0.019) c <sup>-1</sup> 12 <sup>+2</sup> (0.52)	
837.9480	846.3930	849.2250	854.0380	883.4770	920.2940	927.7070	935.3730	938.5850	948.7850
		z <sub>6</sub> (-0.13)		z+1 <sub>13</sub> <sup>+2</sup> (0.058) z <sub>13</sub> <sup>+2</sup> (0.56)					
1060.7300	1224.7560	1274.0190	1367.8360	1496.9520	1636.8970	1674.0060	1708.0170	1751.2130	1752.0460
c-1 <sub>8</sub> (0.18)	z <sub>8</sub> (0.24) z+1 <sub>8</sub> (-0.77)		c <sub>9</sub> (0.19)	c <sub>10</sub> (0.26)	z <sub>12</sub> (0.17)			c-1 <sub>13</sub> (0.41) c <sub>13</sub> (-0.60)	c <sub>13</sub> (0.23)



S Figure 6

Piccolo

Q9QYX7

AVC(Carbamidomethyl)C(GlcNAc)DMVYKLPFGR Cys-2870 Mouse



#### Ankyrin-2 Q8C8R3

Cys-2906

Mouse

#### IQTDTC(GlcNAc)HSTVVHSPEVYSVIIR



S Figure 7B

## IQTDTC(GlcNAc)HSTVVHSPEVYSVIIR

#### - Peak Matches

259.3170	360.4030	385.4340	475.3530	484.4330	571.4500	576.3280	643.0950	673.9720	703.5850
c <sub>2</sub> (0.14)	c <sub>3</sub> (0.18)	z <sub>3</sub> (0.17)	c4(0.10)	z4(0.096)	z5(0.081)	c <sub>5</sub> (0.029) c <sub>14</sub> <sup>+3</sup> (0.057) c-1 <sub>14</sub> <sup>+3</sup> (0.39)	z+1 <sub>11</sub> <sup>+2</sup> (0.24) z <sub>11</sub> <sup>+2</sup> (0.75)		c <sub>11</sub> <sup>+2</sup> (0.25) c-1 <sub>11</sub> <sup>+2</sup> (0.76)
741.5560	772.1180	791.8960	813.5690	835.7440	853.8260	861.8640	864.2790	867.9370	870.1260
z <sub>13</sub> <sup>+2</sup> (0.14) z+1 <sub>13</sub> <sup>+2</sup> (-0.36)	c <sub>12</sub> <sup>+2</sup> (0.26) c-1 <sub>12</sub> <sup>+2</sup> (0.76)	z <sub>14</sub> <sup>+2</sup> (-0.043) z+1 <sub>14</sub> <sup>+2</sup> (-0.55)		z+1 <sub>15</sub> <sup>+2</sup> (-0.22) z <sub>15</sub> <sup>+2</sup> (0.29)	z+1 <sub>21</sub> <sup>+3</sup> (0.078) z <sub>21</sub> <sup>+3</sup> (0.41)		c <sub>14</sub> +2(0.38)		
875.3980	883.3300	904.4260	928.9490	962.5170	1007.3400	1010.6990	1059.3710	1059.9840	1103.2620
		z+1 <sub>16</sub> <sup>+2</sup> (-0.063) z <sub>16</sub> <sup>+2</sup> (0.44)	c <sub>15</sub> +2(0.53)	z <sub>8</sub> (-0.026)			c <sub>17</sub> <sup>+2</sup> (-0.12) c-1 <sub>17</sub> <sup>+2</sup> (0.39)	c <sub>17</sub> +2(0.49)	c <sub>18</sub> +2(0.26) c-1 <sub>18</sub> +2(0.76)
1106.5210	1116.6750	1146.3870	1153.7160	1165.0410	1166.0590	1265.8390	1284.6590	1317.7700	1338.5170
c <sub>8</sub> (0.052)		z <sub>10</sub> (-0.24)		z <sub>19</sub> +2(-0.021) z+1 <sub>19</sub> +2(-0.52)	z+1 <sub>19</sub> <sup>+2</sup> (0.49)	c <sub>21</sub> <sup>+2</sup> (0.22) c-1 <sub>21</sub> <sup>+2</sup> (0.72)	z+1 <sub>11</sub> (-0.036)		



S Figure 9

Trafficking kinesin-binding protein 1 Q6PD31

Mouse

**Cys-695** 

VTPSLNSAPAPAC(GlcNAc)SSTSHLK

V20141119-21 #3454 RT: 34.05 AV T: FTMS + c NSI d Full ms2 724.70@ (2+)y18 204.09 100-(2+)y12G 884.44 80-**PSLNSA** 138.06 **PSLNSAPA** 201.1=b2 (2+)y18G 60y12 1198.59 186.08 (2+)y10 (2+)y10G 40y10 **PSLN** 515.75 y7 570.29 617.29 1030.50 y10G 657.65 701.34 738.38 770.41 20-412.22 894.43 985.98 875.93 228.13 280.17 590.29 1233.58 1269.62 499.25 934.96 323.95 420.22 1065.49 1101.51 1180.58 273.12 384.22 811.38 بلابها 0-600 1000 100 200 300 400 500 700 800 900 1100 1200 m/z V20141119-21 #3455 RT: 34.06 AV: 1 NL: 9.53E3 \*\* T: ITMS + c NSI d sa Full ms2 724.70@etd66.67 [50.00-2000.00] x5 724.74 100-\*\* (2+)-SGIcNAc 724.3612(3+) 80-1086.29 60-968.58 1029.20 40-469.28 884.43 381.36 20-929.86 1429.68 1289.70 514.43 1544.76 1658.64 657.34 744.43 570.30 1121.55 1705.71 1858.91 1929.80 809.41 1228.70 245.30 315.34 437.28 1407.51 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500 1600 1700 1800 1900 Peak Matches 315.3430 381.3560 468.3720 469.2780 568.3470 570.3030 202.1490 245.2990 514.4250 628.4820 z+1<sub>2</sub>(0.11) c<sub>3</sub>(0.14) z<sub>3</sub>(0.12) z<sub>4</sub>(0.10) z+1<sub>4</sub>(0.0011) c-1<sub>5</sub>(0.11) y9<sup>+2</sup>(-0.42) z+15(-0.022) c-1<sub>6</sub>(0.13) 656.4650 657.3350 716.4700 743.4990 744.4290 770.5110 629.3780 725.8880 809.4090 884.4300 c<sub>6</sub>(0.016) z<sub>6</sub>(0.12) z7(0.12) z+17(0.040) b<sub>8</sub>(0.11) z+16(-0.022) c7(0.076) 0.054); 1029.2020 964.8200 968.5840 1021.6740 1028.2050 929.8600 941.8340 960.7070 982.9330 985.3320 z+117<sup>+2</sup>(-0.088)  $c_{19}^{+2}(0.18)$ z<sub>19</sub>+2(-0.29) c18<sup>+2</sup>(-0.14) z+1<sub>19</sub>+2(0.20) y<sub>18</sub>+2(-0.65 z17<sup>+2</sup>(0.42) c-118<sup>+2</sup>(0.37) c-119<sup>+2</sup>(0.68) z+119<sup>+2</sup>(-0.79) z<sub>19</sub>+2(0.71) 1037.8650 1052.1530 1058.8110 1066.1390 1079.4190 1288.6590 1428.5890 1456.6470 1544.7640 1658.6420 c11(-0.42)

z<sub>11</sub>(0.059)

y<sub>8</sub>(0.65)

c-1<sub>11</sub>(0.59)

z13(-0.043)

c-113(-0.11)

z+114(0.034)

z+1<sub>15</sub>(-0.13)

S Figure 10A

MLX-interacting protein Q2VPU4

Cys-579

Mouse

#### NAC(GlcNAc)IAPAAFSGQPQK



S Figure 10B

## NAC(GlcNAc)IAPAAFSGQPQK

#### [\_] Peak Matches

322.6670	326.0350	372.6160	373.5250	391.2560	485.6090	507.5850	509.6740	526.7310	542.7060
		y <sub>3</sub> (0.39)			z+14(0.34)	c-1 <sub>3</sub> (-0.62)	c <sub>3</sub> (0.46)		z+1 <sub>5</sub> (0.41)
559.8100	595.8510	612.8390	626.7150	628.6710	629.7490	776.7390	847.8870	854.8150	917.8200
				z <sub>6</sub> (0.35) z+1 <sub>6</sub> (-0.65)	z+1 <sub>6</sub> (0.42)	z+17(0.35)	z+1 <sub>8</sub> (0.46)		z9(0.36) z+19(-0.65)
1085.9550	1198.9600	1250.3300	1269.0430	1279.0000	1448.0080	1454.0630	1469.9520	1470.9810	1488.4730
z <sub>11</sub> (0.40)	z <sub>12</sub> (0.33)				c-1 <sub>13</sub> (0.32)				
z+1 <sub>11</sub> (-0.60)	z+1 <sub>12</sub> (-0.68)		W12		c <sub>13</sub> (-0.68)				
1502.9210	1522.2060	1533.8910	1576.0370	1577.0140	1641.9130				
			z <sub>14</sub> (0.27)						
			c-1 <sub>14</sub> (0.30)	z+1 <sub>14</sub> (0.24)					
			c <sub>14</sub> (-0.71)	c <sub>14</sub> (0.26)					
			z+1 <sub>14</sub> (-0.74)						

S Figure 11

#### GLYDGPVC(GlcNAc)EVSVTPK

Cys-504

Rat





#### AGISTTSVC(GIcNAc)EGQIANPSPISR

Protein Fam222b D3Z

D3ZPE2

Cys-288 Rat





100-

80

60-

40-

20-

0-100

100-

80-

60-

40-

20-

#### LGPVYC(GlcNAc)QASFSGTNIIGNK



S Figure 14

Cys-579 Rat



S Figure 15

#### YEATS domain-containing protein 2 Q3TUF7



S Figure 16A

MLX-interacting protein Q2VPU4

Mouse

Cys-579

#### NAC(GlcNAc)IAPAAFSGQPQK



S Figure 16B

## NAC(GlcNAc)IAPAAFSGQPQK

## [\_] Peak Matches

312.5530	325.7990	366.5980	370.7740	372.6940	421.4680	454.5180	466.6710	483.6150	485.6000
				y <sub>3</sub> (0.47)				z4(-0.65)	z+14(0.33)
499.6160	514.7640	524.8190	529.7200	542.5330	557.6130	628.7810	629.6940	642.8330	643.7290
y4(-0.67)				z+1 <sub>5</sub> (0.24)	y <sub>5</sub> (0.31)	z <sub>6</sub> (0.46) z+1 <sub>6</sub> (-0.54)	z+1 <sub>6</sub> (0.37)		y <sub>6</sub> (-0.61)
775.7820	776.7470	846.7940	847.7910	854.8630	917.7180	918.7540	1085.8200	1198.9340	1222.9480
z <sub>7</sub> (0.40) z+1 <sub>7</sub> (-0.61)	z+17(0.35)	z <sub>8</sub> (0.37) z+1 <sub>8</sub> (-0.64)	z+1 <sub>8</sub> (0.36)		z <sub>9</sub> (0.26) z+1 <sub>9</sub> (-0.75)	z+1 <sub>9</sub> (0.29)	z <sub>11</sub> (0.27) z+1 <sub>11</sub> (-0.74)	z <sub>12</sub> (0.30) z+1 <sub>12</sub> (-0.71)	c-1 <sub>11</sub> (0.38) c <sub>11</sub> (-0.63)
1268.9910	1448.0450	1454.0830	1469.9460	1471.0500	1488.0740	1503.1710	1504.0430	1576.0690	1577.0370
w13	c-1 <sub>13</sub> (0.37) c <sub>13</sub> (-0.64)						z <sub>13</sub> (-0.68)	z <sub>14</sub> (0.31) c-1 <sub>14</sub> (0.33) c <sub>14</sub> (-0.68) z+1 <sub>14</sub> (-0.70)	z+1 <sub>14</sub> (0.27) c <sub>14</sub> (0.29)

S Figure 17

#### AGISTTSVC(GlcNAc)EGQIANPSPISR

Cys-288 Mouse



## LDFGQGSGSPVC(Cgly)LAQVK<sup>+3</sup>



## FPFGSSC(Cgly)TGTFHPAPSAPDK<sup>+3</sup>



S Figure 20

## QKAPFPATC(Cgly)EAPSR<sup>+3</sup>



## LDFGQGS(HexNAc)GSPVC(Carbamidomethyl)LAQVK<sup>+2</sup>



MH <sup>+1</sup> (av)	MH <sup>+1</sup> (mono)	MH <sup>+2</sup> (av)	MH <sup>+2</sup> (mono)
1967.2186	1965.9586	984.1130	983.4829

[_] Peak I	Peak Matches									
361.3270	478.4100	505.2610	517.7150	521.3960	543.3550	603.3420	610.3990	629.3110	663.4130	
					z+15(0.0049)					
711.6010	720.4410	744.3460	985.8490	986.5520	1043.6110	1310.5170	1332.6080	1333.6380	1390.5070	
			z9(0.32) z+19(-0.68)	z+19(0.018)	z+1 <sub>10</sub> (0.056)		z <sub>11</sub> (-0.051)	z+1 <sub>11</sub> (-0.029)	z+1 <sub>12</sub> (-0.18)	
1425.7370	1474.4510	1517.7350	1518.7530	1590.8830	1608.7190	1638.1760	1730.9030	1736.7500	1837.0530	
c <sub>12</sub> (0.11)		z <sub>13</sub> (-0.0040)	z+1 <sub>13</sub> (0.0062)	y <sub>14</sub> (0.10)	c-1 <sub>14</sub> (-0.026)			c-1 <sub>15</sub> (-0.053)	c <sub>16</sub> (0.17) z <sub>16</sub> (0.20)	
1859.8980	1876.9580	1897.0270								

#### FPFGS(HexNAc)SC(Carbamidomethyl)TGTFHPAPSAPDK<sup>+3</sup>



## FPFGSS(HexNAc)C(Carbamidomethyl)TGTFHPAPSAPDK<sup>+3</sup>



MH <sup>+1</sup> (av)	MH <sup>+1</sup> (mono)	MH <sup>+2</sup> (av)	MH <sup>+2</sup> (mono)	MH <sup>+3</sup> (av)	MH <sup>+3</sup> (mono)
2312.5351	2311.0336	1156.7713	1156.0204	771.5166	771.0160

[_] Peak I	Peak Matches										
316.2680	357.3900	366.3410	409.3500	414.3240	430.3440	465.2000	466.3310	501.4210	510.3470		
			c <sub>3</sub> (0.13)	z4(0.11)	y4(0.11)	c-14(-0.037)	c4(0.086)	z <sub>5</sub> (0.18)			
553.3380	597.2920	614.3770	640.4390	652.3400	669.3240	685.4680	772.2940	781.6040	809.5640		
c <sub>5</sub> (0.061)		y <sub>6</sub> (0.063)			z <sub>7</sub> (-0.0088)	y <sub>7</sub> (0.12)					
843.4030	903.4860	1003.3770	1026.1930	1033.6970	1046.6600	1050.5060	1055.8500	1069.6780	1091.4120		
c <sub>6</sub> (0.015)	z <sub>9</sub> (0.042)	c7(-0.042)	z <sub>18</sub> <sup>+2</sup> (0.24) z+1 <sub>18</sub> <sup>+2</sup> (-0.26) b <sub>18</sub> <sup>+2</sup> (0.74)	c-1 <sub>18</sub> <sup>+2</sup> (0.23) y <sub>18</sub> <sup>+2</sup> (-0.26) c <sub>18</sub> <sup>+2</sup> (-0.27)		z <sub>10</sub> (-0.0069)			c <sub>19</sub> +2(-0.069 c-1 <sub>19</sub> +2(0.43		
1104.7640	1111.4910	1112.1080	1115.7350	1128.4580	1149.5050	1161.5660	1309.6950	1409.6790	1899.8690		
c <sub>8</sub> (0.30)						c <sub>9</sub> (0.078)	z <sub>13</sub> (0.065)	c <sub>11</sub> (0.075)			

## MC(Carbamidomethyl)AALNS(HexNAc)MDQYGGR<sup>+2</sup>



#### Elemental Composition: C70 H114 N21 O27 S3

MH <sup>+1</sup> (av)	MH <sup>+1</sup> (mono)	MH+2(av)	MH <sup>+2</sup> (mono)
1778.0036	1776.7350	889.5055	888.8711

#### [\_] Peak Matches

411.6260	437.3740	529.5180	565.3280	567.3720	612.5470	636.5120	680.4120	811.4750	1100.5150
	z+1 <sub>4</sub> (0.16)		z+1 <sub>5</sub> (0.055)				z+1 <sub>6</sub> (0.11)	z+1 <sub>7</sub> (0.13)	z <sub>8</sub> (0.071)
1101.5260	1170.9160	1184.8830	1214.5880	1215.4930	1323.4250	1327.6050	1328.5370	1331.6040	1399.6640
z+1 <sub>8</sub> (0.074)			z9(0.10) c9(0.10)	z+19(-0.0016)		z <sub>10</sub> (0.034)	z+1 <sub>10</sub> (-0.042)		z+1 <sub>11</sub> (0.048)
1469.6840	1539.6730	1558.7920	1575.6380	1670.5870	1687.8580	1716.7740			
z <sub>12</sub> (0.039)	w13								

## S Figure 25

Estimating the relative amount of (1)FPFGSSC(Carbamidomethyl)T(GlcNAc)GTFHPAPSAPDK and (2)FPFGSSC(GlcNAc)TGTFHPAPSAPDK from the LC/MS data of a WGA-enriched but further not fractionated mouse synaptosome glycopeptide mixture





Host cell factor 1 P51610

#### RAC(GlcNAc)AAGTPAVIR

Cys-1139

Human





A) ETD spectrum of synthetic peptide QKAPFPAACEAPAR. ' $\blacklozenge$ ' labels the precursor ion and its charge-reduced form. The insert shows the bond cleavages detected.





B) HCD spectrum of synthetic peptide QKAPFPAACEAPAR. The insert shows the bond cleavages detected. y ions containing GlcNAc are labeled with a G.

#### S Figure 28A



A) ETD spectrum of synthetic peptide LDFGQGAGAPVCLAQVK. '◆' labels the precursor ion and its charge-reduced form. The insert shows the bond cleavages detected.



B) HCD spectrum of synthetic peptide QKAPFPAACEAPAR. The insert shows the bond cleavages detected. y ions containing GlcNAc are labeled with a G.

S Figure 28B

S Figure 29A





A) Mass spectrum across XIC peaks and extracted ion chromatograms of control peptide TAPTS(GlcNAc)TIAPG (modified and unmodified) after incubation with no OGA. Spectrum peaks are labeled to match XIC peaks. Peak areas are labeled above XIC peaks.

## S Figure 29B



B) Mass spectrum across XIC peaks and extracted ion chromatograms of control peptide TAPTS(GlcNAc)TIAPG (modified and unmodified) after incubation with OGA. Spectrum peaks are labeled to match XIC peaks. Peak areas are labeled above XIC peaks.

## S Figure 29C



C) Mass spectrum across XIC peaks and extracted ion chromatograms of control peptide QKAPFPATC(GlcNAc)EAPSR (modified and unmodified) after incubation with no OGA. Spectrum peaks are labeled to match XIC peaks. Peak areas are labeled above XIC peaks.

### S Figure 29D

V20160413-07 #1144-1635 RT: 12.22-17.31 AV: 127 NL: 1.69E5 T: FTMS + p NSI Full ms [350.00-1600.00]



D) Mass spectrum across XIC peaks and extracted ion chromatograms of control peptide QKAPFPATC(GlcNAc)EAPSR (modified and unmodified) after incubation with OGA. Spectrum peaks are labeled to match XIC peaks. Peak areas are labeled above XIC peaks.

-121

## S Figure 29E

## Ratio of TAPTSTIAPG and TAPTST(GlcNAc)IAPG





## Ratio of QKAPFPATCEAPSR and QKAPFPATC(GlcNAc)EAPSR



E and F) Relative XIC peak areas of modified and unmodified GlcNAc standard peptide (TAPTSTIAPG) and synthetic peptide (QKAPFPATCEAPSR) with and without incubation with OGA.