Electronic Supplementary Material (ESI) for Analyst

Supporting Information

Boronic acid-functionalized detonation nanodiamond for specific enrichment of glycopeptides in glycoproteome analysis

Guobin Xu[1], Wei Zhang[1], Liming Wei*, Haojie Lu, Pengyuan Yang*

Institutes of Biomedical Sciences & Department of Chemistry, Fudan University, Shanghai 200032, China

*Corresponding author, weiliming@fudan.edu.cn， pyyang@fudan.edu.cn

Tel: -86-21-5423-7961 Fax: -86-21-5423-7961

[1] These authors contribute equally to this work.

Table S1. Detailed information of the observed glycopeptides obtained from tryptic bovine fetuin [1]

Peak number	Observed m/z	Theoretical peptide mass	Theoretical peptides sequence	Theoretical glycan mass	Glycan structure
No. 1	1742.0	1742.0	^{145}LC^{a}PDC^{a}PLLAPL*N#*DSR^{159}		
No. 2	1945.2	1742.0	^{145}LC^{a}PDC^{a}PLLAPL*N#*DSR^{159}	203.2	$HexNAc_1$
No. 3	2837.0	1742.0	^{145}LC^{a}PDC^{a}PLLAPL*N#*DSR^{159}	1095.9	Hex_3 $HexNAc_3$
No. 4	2999.3	1742.0	^{145}LC^{a}PDC^{a}PLLAPL*N#*DSR^{159}	1258.0	Hex_4 $HexNAc_3$
No. 5	3364.6	1742.0	^{145}LC^{a}PDC^{a}PLLAPL*N#*DSR^{159}	1623.4	Hex_5 $HexNAc_4$
No. 6	3729.6	1742.0	^{145}LC^{a}PDC^{a}PLLAPL*N#*DSR^{159}	1988.7	Hex_6 $HexNAc_5$
No. 7	3017.3		^{160}VVHAVEVALATFNAES*N#*GSY LQLVEISR^{187}		
No. 8	3219.8	3017.3	^{160}VVHAVEVALATFNAES*N#*GSY LQLVEISR^{187}	203.2	$HexNAc_1$
No. 9	3857.5	3670.8	^{72}RPTGEVYDIEIDTLETTCHV LDPTPLA*N#*C^{a}SVR^{103}	185.2	$HexNAc_1$-H_2O
No. 10	4274.6	3017.3	^{160}VVHAVEVALATFNAES*N#*GSY LQLVEISR^{187}	1258.0	Hex_4 $HexNAc_3$
No. 11	4639.7	3017.3	^{160}VVHAVEVALATFNAES*N#*GSY	1623.4	Hex_5

			LQLVEISR[187]		$HexNAc_4$
No. 12	4842.7	3017.3	[160]VVHAVEVALATFNAES*N#*GSY LQLVEISR[187]	1826.6	Hex_5 $HexNAc_5$
No. 13	5004.6	3017.3	[160]VVHAVEVALATFNAES*N#*GSY LQLVEISR[187]	1988.7	Hex_6 $HexNAc_5$
No. 14	5294.5	3670.8	[72]RPTGEVYDIEIDTLETTCHV LDPTPLA*N#*C[a]SVR[103]	1623.4	Hex_5 $HexNAc_4$
No. 15	5659.5	3670.8	[72]RPTGEVYDIEIDTLETTCHV LDPTPLA*N#*C[a]SVR[103]	1988.7	Hex_6 $HexNAc_5$
No. 16	2755.5	2755.5	[79]DIEIDTLETTCHVLDPTPLA *N#*CSVR[103]	-	-

Table S2. Detailed information of the observed glycopeptides obtained from tryptic horseradish peroxidase (HRP) [2]

Peak number	**Observed m/z**	**Theoretical peptide mass**	**Theoretical peptides sequence**	**Theoretical glycan mass**	**Glycan structure**
No. 1	2069.6	898.5	[42]P*N#*VSNIVR[49]	1171.3	$Hex_3HexNAc_2dHex_1Pent_1$
No. 2	2533.7	2182.9	[295]SFA*N#*STQTFFNAFVEAMDR[313]	349.5	$HexNAc_1Pent_1$
No. 3	2591.9	1420.6	[225]PTL*N#*TTYLQTLR[236]	1171.3	$Hex_3HexNAc_2dHex_1Pent_1$
No. 4	2706.0	2501.2	[272]GLIQSDQELFSSP*N#*ATDTIP LVR[294]	203.2	$HexNAc_1$
No. 5	2851.7	2501.2	[272]GLIQSDQELFSSP*N#*ATDTIP LVR[294]	349.5	$HexNAc_1Pent_1$
No. 6	3339.7	2168.4	[31]QLTPTFYDNSCP*N#*VSNIVR[49]	1171.3	$Hex_3HexNAc_2dHex_1Pent_1$
No. 7	3355.5	2182.9	[295]SFA*N#*STQTFFNAFVEAMDR[313]	1171.3	$Hex_3HexNAc_2dHex_1Pent_1$
No. 8	3527.9	2501.2	[272]GLIQSDQELFSSP*N#*ATDTIP LVR[294]	1026.6	$Hex_3HexNAc_2dHex_1$
No. 9	3673.1	2501.2	[272]GLIQSDQELFSSP*N#*ATDTIP LVR[294]	1171.3	$Hex_3HexNAc_2dHex_1Pent_1$
No. 10	3750.5	2725.2	[69]LHFHDCFVNGCDASILLD*N#*T TSFR[92]	1026.6	$Hex_3HexNAc_2dHex_1$
No. 11	3896.1	2725.2	[69]LHFHDCFVNGCDASILLD*N#*T TSFR[92]	1171.3	$Hex_3HexNAc_2dHex_1Pent_1$
No. 12	4164.6	2642.3	[214]LY*N#*FSNTGLPDPTL*N*TTYLQT LR[236]	1522.2	$Hex_3HexNAc_3dHex_1Pent_2$
No. 13	4223.6	3071.8	[31]QLTPTFYDNSCP*N#*VSNIVR[49]	1171.3	$Hex_3HexNAc_2dHex_1P$

			^{115}AAVESACPR^{123}		ent$_1$
No. 14	4840.1	2642.3	^{214}LY*N#*FSNTGLPDPTLNTTYLQT LR^{236}	2197.6	Hex$_6$HexNAc$_4$dHex$_2$P ent$_1$
No. 15	4986.1	2642.3	^{214}LY*N#*FSNTGLPDPTLNTTYLQT LR^{236}	2343.7	Hex$_6$HexNAc$_4$dHex$_2$P ent$_2$

Fig. S1 MALDI mass spectra of glycopeptides ^{160}VVHAVEVALATFNAESN#GSY LQLVEISR^{187} consists of partially ^{18}O labeled N#. The mass spectrum of this glycopeptide not labeled by ^{18}O (a) or 100% labeled by ^{18}O (b). The isotope distribution of this peptide is 3017 (61.6%): 3018 (100%): 3019 (86.1%): 3020 (51.9%). Define x= the intensity of this peptide from the first part (the release of the glycans with PNGaseF in H_2O) AND y= the intensity of this peptide from the second part (the capture of the glycopeptides using dND-p-APBA and eluting them and then releasing the glycans with PNGaseF in $H_2{}^{18}O$). So y/x represents the recovery of glycopeptides. Based on the first and the third peak, we got a group of equations: 0.62 x= 5694.1 and 0.86 x+ 0.62 y= 11007.8, educed y/x = 54.6%. As the same, based on the second and the fourth peak, we got another group of equations: x= 8054.9 and 0.52 x+ y= 11415.7, educed y/x = 89.6%. In sum, the average recovery of glycopeptides is 72.2% [3].

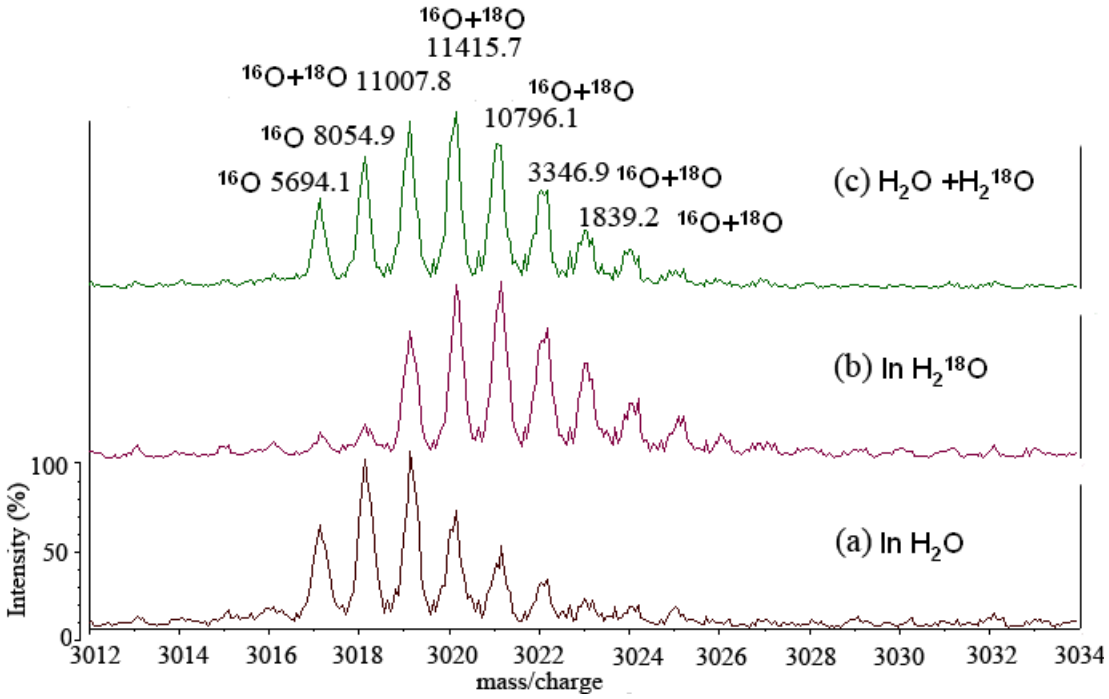


Fig. S2 RPLC chromatogram of the digest of mouse liver extract. One fraction is indicated with arrow at 10–14 min (1#). Mobile phase A consisted of 0.1% FA in water and mobile phase B consisted of 0.1% FA in ACN (all v/v). The flow rate was 0.8 mL/min. The separation was performed using gradient conditions as follows: 0% B increased up to 50% B in 30 min, then linearly increased to 90% B in 60 min, maintained at 90% B for 10 min for washing the column, and then ramped down to 0% B for equilibrium. Effluents were started to collect after 10-min gradient elution, and 25 RPLC fractions were collected every 2 min (1.6 mL) continuously regardless of the eluting profile shape. On column UV detection for RPLC was carried out at 215 nm using a SPD-20A UV Detector (Shimadzu Co.), and then all fractions dried by lyophilization.

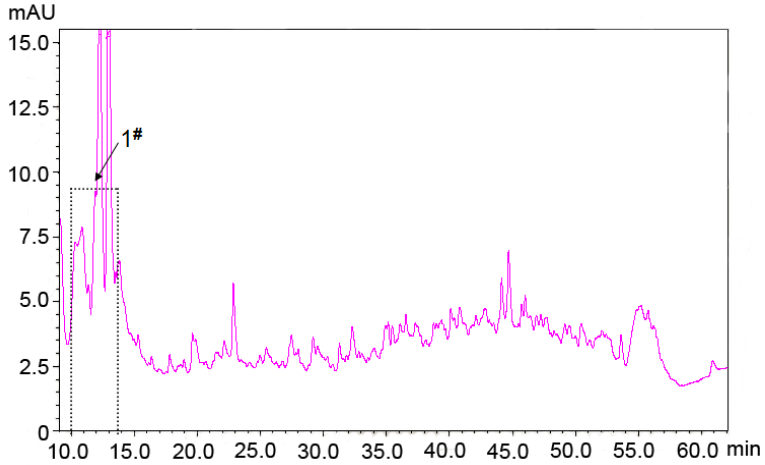


Fig. S3 MS/MS spectra of de-glycosylated peptides, the de-glycosylated peptide of corticosteroid 11-beta-dehydrogenase isozyme 1 QSN#GSIAVISSLAGK （(M+2H)$^{2+}$，m/z=943.83），#-labeled to the amino acid of N-glycosylated modification.

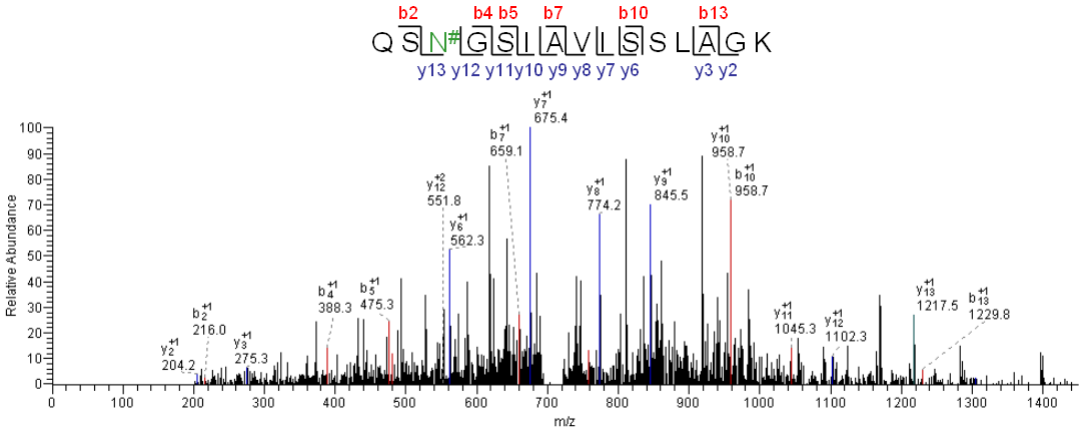
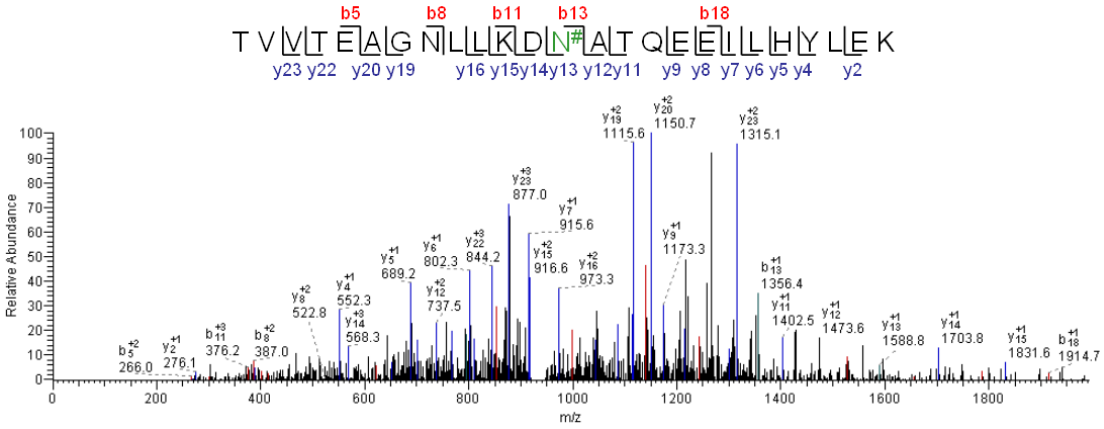


Table S3 *N*-glycosylation sites and glycoprotein identified from the RPLC fraction 1# with dND-p-APBA specific enrichment followed by LC-MS/MS

Protein IPI	Discription	site	peptides sequence	Swiss-prot note
IPI00776266.1	Protein	43	C.**N**#GSIPSDLAISGLLPYWMR.D	
IPI00110849.1	H-2 class II histocompatibility antigen, A-K alpha chain	105	K.RS**N**#STPATNEAPQATVFPKSPVLLGQPNT.L	
IPI00111908.8	Carbamoyl-phosphate synthase [ammonia]	861	K.ALEN**N**#MSLDEIVR.L	
IPI00115599.6	Corticosteroid 11-beta-dehydrogenase isozyme 1	162	K.QS**N**#GSIAVISSLAGK.M	potential
IPI00116591.1	Short-chain specific acyl-CoA dehydrogenase	180	R.EEGDSWVL**N**#GTK.A	
IPI00116591.1	Short-chain specific acyl-CoA dehydrogenase	159	K.IGCFALSEPGN#GSDAGAASTTAR.E	
IPI00116603.1	Ornithine carbamoyltransferase	240	K.LAEQYAKE**N**#GTKLSMTN.D	
IPI00117312.1	Aspartate aminotransferase	164	R.DVFLPKPSWG**N**#HTPIFR.D	
IPI00117857.2	Alpha-1-antitrypsin	70	S.**N**#TSNIFFSPVSIATAFAMLSLGSK.G	yes
IPI00117914.3	Arginase 1	60	R.DHGDLAFVDVP**N**#DSSFQIVK.N	
IPI00118625.1	Histidine ammonia-lyase	307	K.EGLALI**N**#GTQMITSLGCEALER.A	
IPI00118899.1	Alpha-actinin-4	680	R.ISIEM**N**#GTLEDQLSHLK.Q	
IPI00119114.2	Long-chain specific acyl-CoA dehydrogenase	198	R.SGSDWIL**N**#GSK.V	
IPI00120832.1	Major urinary protein 3	66	R.AFVE**N**#ITVLENSLVFK.F	yes
IPI00122048.2	Sodium/potassium-transporting ATPase subunit alpha-3	628	K.GVGIISEG**N**#ETVEDIAAR.L	
IPI00126253.1	LAG1 longevity assurance homolog 2	19	R.LWLPV**N**#LTWA.D	yes
IPI00131830.1	Serine protease inhibitor A3K	105	F.**N**#LTETPEADIHQGFGNLLQSLSQPEDQDQINIGNAMFIEK.D	potential
IPI00134691.3	UDP-glucuronosyltransferase 1-1	89	E.**N**#VTATLVELGR.T	potential
IPI00134691.3	UDP-glucuronosyltransferase 1-2	89	K.E**N**#VTATLVELGR.T	potential
IPI00135686.2	Peptidyl-prolyl cis-trans isomerase B	148	K.DT**N**#GSQFFITTVK.T	
IPI00154054.1	Acetyl-CoA acetyltransferase, mitochondrial	272	K.E**N**#GTITAANASTLNDGAAALVLMTAEAAQR.L	
IPI00154054.1	Acetyl-CoA acetyltransferase, mitochondrial	272	K.TVFQKE**N**#GTITAAN.A	
IPI00154054.1	Acetyl-CoA acetyltransferase, mitochondrial	272	K.TVFQKENGTITAA**N**#ASTLNDGAAALVLMTAEAAQR.L	

IPI00154056.1	Lysosomal acid phosphatase	322	H.IFELYQEDNG**N#**FSVEMYFR.N	potential
IPI00265576.6	Simliar to calmodulin	181	D.G**N#**GTIDFPEFLTMMAR.K	
IPI00308885.6	60 kDa heat shock protein, mitochondrial	426	R.VTDAL**N#**ATR.A	
IPI00318595.1	Endoplasmic reticulum aminopeptidase 1	655	K.**N#**ETEIMPIFQALNELIPMYK.L	potential
IPI00321190.1	Sulfated glycoprotein 1	80	K.TVVTEAGNLLKD**N#**ATQEEILHYLEK.T	potential
IPI00330670.6	Tripartite motif-containing protein 69	458	R.RVGVYLDYEGGQVSFY**N#**AT.T	
IPI00336324.11	Malate dehydrogenase, cytoplasmic	26	Y.SLLYSIG**N#**GSVFGK.D	
IPI00379304.4	Peroxiredoxin 6	42	R.**N#**FTPVCTTELGR.A	
IPI00461964.3	Methylmalonate-semialdehyde dehydrogenase	445	K.IVNDNPYG**N#**GTAIFTTNGATAR.K	
IPI00461964.3	Methylmalonate-semialdehyde dehydrogenase	445	N.DNPYG**N#**GTAIFTTNGATAR.K	
IPI00462140.1	Keratin, type II cytoskeletal 1b	206	K.WELLQQV**N#**TSTR.T	
IPI00469218.1	Lysosomal membrane glycoprotein 1	296	R.L**N#**MTLPDALVPTFSIS.N	
IPI00554989.3	Peptidyl-prolyl cis-trans isomerase	108	K.HTGPGILSMANAGPNT**N#**GSQFFICTAK.T	
IPI00554989.3	Peptidyl-prolyl cis-trans isomerase	108	K.HTGPGILSMANAGPNT**N#**GSQFF.I	
IPI00555140.2	Phosphoglucomutase-1	132	K.F**N#**ISNGGPAPEAITDK.I	
IPI00624663.3	Alpha-2-macroglobulin	382	K.**N#**ITSVVSPLGYLSIFTTDEHGLA.N	potential
IPI00659793.3	Dehydrogenase E1 and transketolase domain containing 1	266	K.**N#**GSGLDWATAETLALGSLLAQGFNVR.L	

Fig. S4 MS/MS spectra of de-glycosylated peptides, the de-glycosylated peptide of Sulfated glycoprotein 1 TVVTEAGNLLKDN#ATQEEILHYLEK（（M+3H）$^{3+}$，m/z=716.89 ），#-labeled to the amino acid of N-glycosylated modification.



References for supporting information

[1] Y. Kanie, A. Enomoto, S. Goto, O. Kanie, Carbohydrate Res. 343 2008, 343, 758-768.

[2] Wuhrer M., Hokke C. H., Deelder A. M. Rapid Commun. Mass Spectrom. 2004, 18, 1741-1748.

[3] Xu Y., Wu Z., Zhang L., Lu H., Yang P. Webley P. A., Zhao D. Anal. Chem. 2009, 81, 503-508.