Identification and Quantification of Glycoproteins Using Ion-Pairing Normal-phase Liquid Chromatography and Mass Spectrometry*⑤

Wen Ding‡§, Harald Nothaft‡¶, Christine M. Szymanski‡¶, and John Kelly‡

Glycoprotein structure determination and quantification by MS requires efficient isolation of glycopeptides from a proteolytic digest of complex protein mixtures. Here we describe that the use of acids as ion-pairing reagents in normal-phase chromatography (IP-NPLC) considerably increases the hydrophobicity differences between non-glycopeptides and glycopeptides, thereby resulting in the reproducible isolation of *N*-linked high mannose type and sialylated glycopeptides from the tryptic digest of a ribonuclease B and fetuin mixture. The elution order of non-glycopeptides relative to glycopeptides in IP-NPLC is predictable by their hydrophobicity values calculated using the Wimley-White water/octanol hydrophobicity scale. *O*-linked glycopeptides can be efficiently isolated from fetuin tryptic digests using IP-NPLC when *N*-glycans are first removed with PNGase. IP-NPLC recovers close to 100% of bacterial *N*-linked glycopeptides modified with non-sialylated heptasaccharides from tryptic digests of periplasmic protein extracts from *Campylobacter jejuni* 11168 and its *pglD* mutant. Label-free nano-flow reversed-phase LC-MS is used for quantification of differentially expressed glycopeptides from the *C. jejuni* wild-type and *pglD* mutant followed by identification of these glycoproteins using multiple stage tandem MS. This method further confirms the acetyltransferase activity of PglD and demonstrates for the first time that heptasaccharides containing monoacetylated bacillosamine are transferred to proteins in both the wild-type and mutant strains. We believe that IP-NPLC will be a useful tool for quantitative glycoproteomics. *Molecular & Cellular Proteomics 8:2170–2185, 2009.*

Protein glycosylation is a biologically significant and complex post-translational modification, involved in cell-cell and receptor-ligand interactions (1–4). In fact, clinical biomarkers and therapeutic targets are often glycoproteins (5–9). Comprehensive glycoprotein characterization, involving glycosylation site identification, glycan structure determination, site occupancy, and glycan isoform distribution, is a technical challenge particularly for quantitative profiling of complex protein mixtures (10–13). Both *N*- and *O*-glycans are structurally heterogeneous (*i.e.* a single site may have different glycans attached or be only partially occupied). Therefore, the MS[1] signals from glycopeptides originating from a glycoprotein are often weaker than from non-glycopeptides. In addition, the ionization efficiency of glycopeptides is low compared with that of non-glycopeptides and is often suppressed in the presence of non-glycopeptides (11–13). When the MS signals of glycopeptides are relatively high in simple protein digests then diagnostic sugar oxonium ion fragments produced by, for example, front-end collisional activation can be used to detect them. However, when peptides and glycopeptides co-elute, parent ion scanning is required to selectively detect the glycopeptides (14). This can be problematic in terms of sensitivity, especially for detecting glycopeptides in digests of complex protein extracts.

Isolation of glycopeptides from proteolytic digests of complex protein mixtures can greatly enhance the MS signals of glycopeptides using reversed-phase LC-ESI-MS (RPLC-ESI-MS) or MALDI-MS (15–24). Hydrazide chemistry is used to isolate, identify, and quantify *N*-linked glycopeptides effectively, but this method involves lengthy chemical procedures and does not preserve the glycan moieties thereby losing valuable information on glycan structure and site occupancy (15–17). Capturing glycopeptides with lectins has been widely used, but restricted specificities and unspecific binding are major drawbacks of this method (18–21). Under reversed-phase LC conditions, glycopeptides from tryptic digests of gel-separated glycoproteins have been enriched using graphite powder medium (22). In this case, however, a second digestion with proteinase K is required for trimming down the peptide moieties of tryptic glycopeptides so that the glycopeptides (typically <5 amino acid residues) essentially resemble the glycans with respect to hydrophilicity for subsequent separation. Moreover, the short peptide sequences of the proteinase K digest are often inadequate for *de novo* sequencing of the glycopeptides.

[1] The abbreviations used are: MS, mass spectrometry; IPR, ion-pairing reagents; IP-NPLC, ion-pairing normal-phase liquid chromatography; RT, retention time; wt, wild-type; TFA, trifluoroacetic acetic acid; BPC, base peak chromatograms; UDP, uridine diphosphate; I.D., internal diameter.

Glycopeptide enrichment under normal-phase LC (NPLC) conditions has been demonstrated using various hydrophilic media and different capture and elution conditions (23–28). NPLC allows either direct enrichment of peptides modified by various N-linked glycan structures using a ZIC®-HILIC column (23–27) or targeting sialylated glycopeptides using a titanium dioxide micro-column (28). However, NPLC is neither effective for enriching less hydrophilic glycopeptides, *e.g.* the five high mannose type glycopeptides modified by 7–11 monosaccharide units from a tryptic digest of ribonuclease b (RNase B), nor for enriching O-linked glycopeptides of bovine fetuin using a ZIC-HILIC column (23). The use of Sepharose medium for enriching glycopeptides yielded only modest recovery of glycopeptides (28). In addition, binding of hydrophilic non-glycopeptides with these hydrophilic media contaminates the enriched glycopeptides (23, 28).

We have recently developed an ion-pairing normal-phase LC (IP-NPLC) method to enrich glycopeptides from complex tryptic digests using Sepharose medium and salts or bases as ion-pairing reagents (29). Though reasonably effective the technique still left room for significant improvement. For example, the method demonstrated relatively modest glycopeptide selectivity, providing only 16% recovery for high mannose type glycopeptides (29). Here we report on a new IP-NPLC method using acids as ion-pairing reagents and polyhydroxyethyl aspartamide (A) as the stationary phase for the effective isolation of tryptic glycopeptides. The method was developed and evaluated using a tryptic digest of RNase B and fetuin mixture. In addition, we demonstrate that O-linked glycopeptides can be effectively isolated from a fetuin tryptic digest by IP-NPLC after removal of the N-linked glycans by PNGase F.

The new IP-NPLC method was used to enrich N-linked glycopeptides from the tryptic digests of protein extracts of wild-type (wt) and *PgID* mutant strains of *Campylobacter jejuni* NCTC 11168. *C. jejuni* has a unique N-glycosylation system that glycosylates periplasmic and inner membrane proteins containing the extended N-linked sequon, D/E-X-N-X-S/T, where X is any amino acid other than proline (30–32). The N-linked glycan of *C. jejuni* has been previously determined to be GalNAc-α1,4-GalNAc-α1,4-[Glcβ1,3]-GalNAc-α1,4-GalNAc-α1,4-GalNAc-α1,3-Bac-β1 (BacGalNAc$_5$Glc residue mass: 1406 Da), where Bac is 2,4-diacetamido-2,4,6-trideoxyglucopyranose (30). In addition, the glycan structure of *C. jejuni* is conserved, unlike in eukaryotic systems (30–32). IP-NPLC recovered close to 100% of the bacterial N-linked glycopeptides with virtually no contamination of non-glycopeptides. Furthermore, we demonstrate for the first time that acetylation of bacillosamine is incomplete in the wt using IP-NPLC and label-free MS.

EXPERIMENTAL PROCEDURES

Materials and Reagents—Bovine RNase B, bovine fetuin, dithiothreitol, iodoacetamide, and the ion-pairing reagents were acquired from Sigma. Modified trypsin was purchased from Promega. The polyhydroxyethyl A™ Javelin® guard column (1 cm $\times$ 1 mm I.D., 5 μm) and the ZIC®-HILIC guard column (0.5 cm $\times$ 1 mm I.D., 5 μm) were purchased from Nest Group (Southborough, MA). The Luna NH$_2$ column (1 cm $\times$ 0.3 mm I.D., 5 μm) was purchased from Bodman Industries (Aston, PA). PNGase F was purchased from Roche Applied Sciences (Mannheim, Germany).

Preparation of Periplasmic Protein Extracts of C. jejuni 11168 (wt) and the Isogenic pglD Mutant for IP-NPLC-MS Analysis—*C. jejuni* glycoprotein extracts were prepared from 4 liters of culture as described (30). Proteins were quantified spectrophotometrically (Nanodrop ND-1000 Spectrophotometer; Thermo Fisher Scientific), adjusted to a concentration of 20 μg/μl using pure water (Milli-Q system, Millipore Corp.) and were either processed immediately or stored at -20 °C.

Tryptic Digestion of the Standard Glycoproteins and Complex Protein Mixtures and PNGase F Digestion of Fetuin—Bovine RNase B or fetuin at 1 mg/ml was reduced, alkylated, and digested with trypsin, as described previously (29). A 100-μl solution of periplasmic protein extracts of *C. jejuni* 11168 or the *pglD* mutant at 320 μg/μl in 50 mM NH$_4$HCO$_3$ was reduced with 8 mM dithiothreitol at 37 °C for 1 h and alkylated with 100 mM iodoacetamide at 37 °C for 30 min. The reagents used for reduction and alkylation were removed by centrifugal ultrafiltration (3000 MWCO) until the samples were at pH 5–6. After the addition of 60 μg of trypsin, the protein solution (120 μl with 50 mM NH$_4$HCO$_3$) was incubated at 37 °C for 16 h. For the O-glycopeptide isolation experiments, half a unit of PNGase F was added to 40 μl of 69 pmol/μl of a fetuin tryptic digest with 50 mM NH$_4$HCO$_3$ and then incubated at 37 °C for 14 h.

IP-NPLC Separation of Glycopeptides—All IP-NPLC experiments were performed either on a CapLC™ capillary LC system or nanoAcquity UPLC® system coupled to a Q-TOF-2™ hybrid quadrupole/TOF mass spectrometer (Waters). The NPLC and IP-NPLC experiments were performed as follows: 1) The sample was suspended in 7.5 μl of 80% ACN + 20% H$_2$O (pH 4.7), if not otherwise stated, with/without addition of ion-pairing reagents. All pHs in this report were measured in ACN/H$_2$O with electrodes calibrated in water (33). 2) A gradient of 15% to 30% solvent B (100% HPLC grade H$_2$O) for 5 min, then 30% to 50% solvent B for 5 min was used at a flow rate of 12 μl/min. The column was equilibrated at 15% solvent B for 4 min after each gradient. Solvent A is 100% ACN. The flow was split after the column so that ~400 nL/min was directed to the ESI source to obtain NPLC-ESI-MS or IP-NPLC-ESI-MS chromatograms for each sample. The remainder of the column eluate was directed to a fraction collector or discarded. 3) The peptides eluted from the column between ~1-x min are referred to below as the "non-glycopeptide fractions" whereas the peptides eluted between ~ x-13 min are referred to as the "glycopeptide fraction". The time x refers to the elution time of the least hydrophilic glycopeptide observed from IP-NPLC-MS of the RNase B tryptic digest and was used in all IP-NPLC experiments as a reference time for determination of the starting elution time of glycopeptides. If not otherwise stated, the IP-NPLC column used in this study was polyhydroxyethyl A, which was used for months without significant deterioration of performance.

MS Analysis of Tryptic Peptide Mixtures from Complex Protein Extracts—The tryptic digest of periplasmic protein extracts of the *C. jejuni* 11168 and the *pglD* mutant (40 μg of each sample suspended in 7.5 μl of 80% ACN + 20% H$_2$O + 1% HCl + 10 mM NH$_4$HCO$_3$) were repetitively subjected to IP-NPLC for off-line isolation of glycopeptides. The glycopeptide and non-glycopeptide fractions (~90 μl for each fraction) were collected and dried to ~1 μl, which was then resuspended in 100 μl of 0.1% formic acid (aq) and analyzed by label-free nano-flow RPLC-ESI-MS (nanoRPLC-ESI-MS) using a nanoAcquity UPLC system coupled to a Q-TOF Ultima™ hybrid quadrupole/TOF mass spectrometer (Waters). The peptides were first
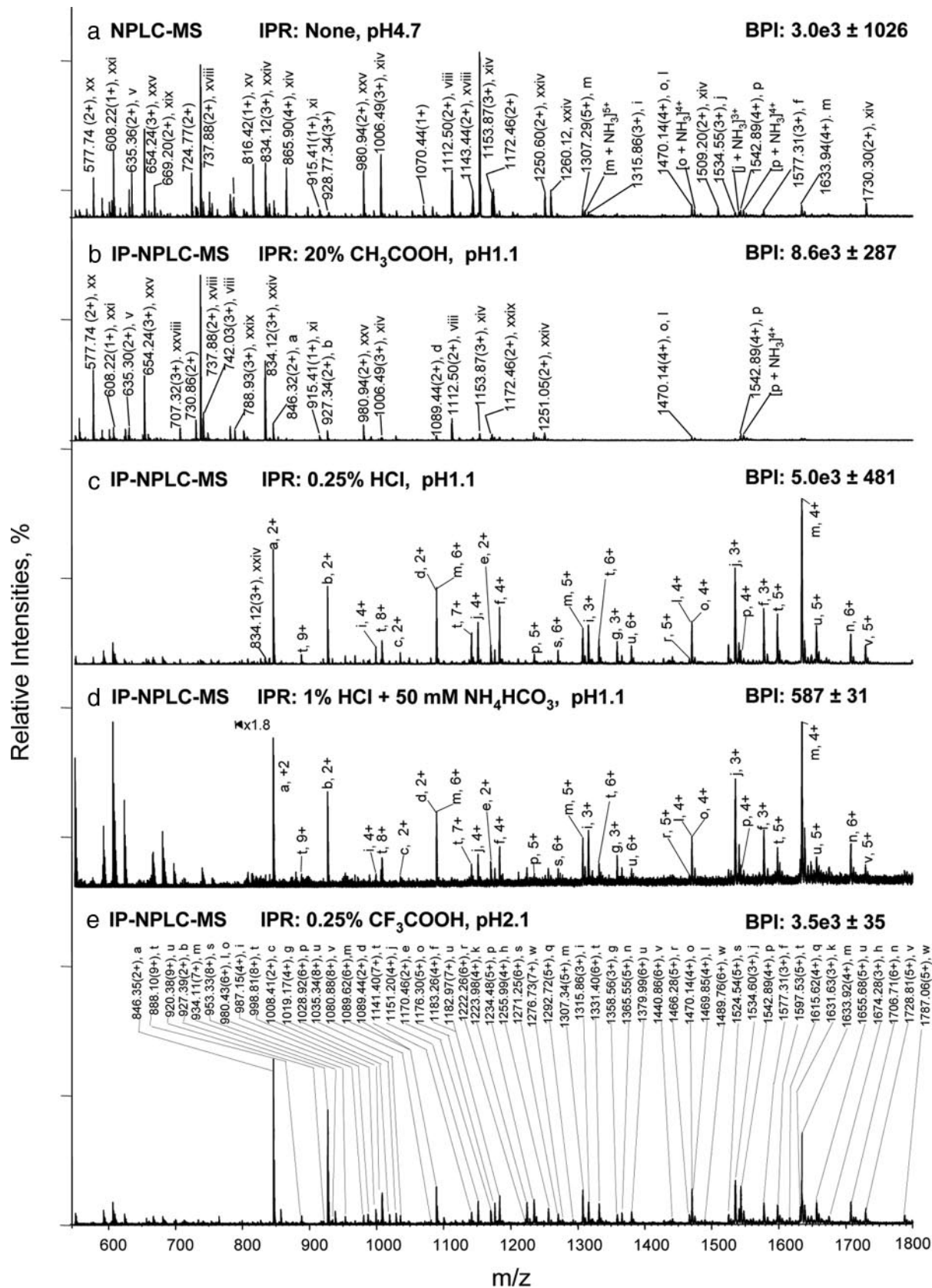
List of identified non-glycopeptides, their calculated hydrophobicity values (ΔG_{octw}, kcal/mol), and IP-NPLC retention times from the RNase B and fetuin tryptic digest

Sequences in italic letters are from RNase B; Glu^0, Asp^0, His^+, Lys^+, Arg^+, N-T NH_3^+, and C-T COO^- are used for calculation of the charged non-glycopeptides (pH 1.1) using MPExTotalizer(35, 37). A value of 4.1 kcal/mol was further added to the calculated value for peptide xiv of fetuin due to likely salt bridge formation of this peptide (36). A value of 2 kcal/mol was further added to the calculated values for peptides xxiv–xxv due to losses of NH_3 at C-terminal, respectively (35, 38). The tryptic digest of RNase B and fetuin mixture was injected with 0.25% HCl in 80% ACN/20% H_2O for IP-NPLC-MS. ^{Cam}C, carboxyamidomethylcysteine.

Peptide no.	Sequnece of identified non-glycopeptides	Calculated hydrophobicity values	IP-NPLC retention time (RT)	Error of RT
		ΔG_{octw}, kcal/mol	min	%
i	(K)VWPR	−8.0	3.8	42.7
ii	(R)AQFVPLPVSVSVEFAVAATD^{cam}CIAK(E)	−8.5	4.2	36.4
iii	*(K)FER(Q)*	−8.8	5.0	0.6
iv	(R)YFK(I)	−8.9	4.3	3.5
v	(K)QDGQFSVLFTK(B)	−10.1	5.1	5.8
vi	(K)^{cam}CNLLAEK(Q)	−10.3	5.2	8.1
vii	(R)L^{cam}CPGR(I)	−10.4	5.5	5.2
viii	*(K)HIIVACEGNPYVPVHFDASV(−)*	−10.7	5.8	2.9
viv	R)VVHAVEVALATFNAESNGSYLQLVEISR(A)	−10.7	5.4	6.9
x	(K)QYGF^{cam}CK(G)	−10.9	5.5	1.7
xi	*(K)YPN^{cam}CAYK(T)*	−11.4	5.7	4.6
xii	(K)HLPR(G)	−11.6	6.2	11.5
xiii	(R)AHYDLR(H)	−11.7	6.2	11.5
xiv	(−)IPLDPVAGYKEPA^{cam}CDDPDTEQAALAAVDYINK(H)	−11.9	6.0	0.6
xv	(K)ALGGEDVR(V)	−12.0	5.3	1.7
xvi	(K)GSVIQK(A)	−12.2	5.9	3.5
xvii	*(K)NVACK(N)*	−12.2	6.2	0.6
xviii	(K)TPIVGQPSIPGGPVR(V)	−12.7	5.3	1.7
xix	(K)^{cam}CDSSPDSAEDVR(K)	−13.3	6.7	3.5
xx	(K)HTLNQIDSVK(V)	−13.6	6.6	0.6
xxi	*(R)ETGSSK(Y)*	−13.8	6.9	0.6
xxii	*(K)TTQANK(H)*	−14.0	7.0	1.7
xxiii	*(K)NGQTN^{cam}CYQSYSTMSITD^{cam}CR(E)*	−14.1	7.0	0.0
xxiv	*(R)^{cam}CKPVNTFVHESLADVQAV^{cam}CSQK(N)*	−14.9	6.5	0.6
xxv	(R)pyroQQTQHAVEGD^{cam}CDIHVLK(Q)	−15.9	6.9	2.9
xxvi	(K)^{cam}CDSSPDSAEDVRK(L)	−16.1	7.4	0.6
xxvii	(R)QQTQHAVEGD^{cam}CDIHVLK(Q)	−17.9	7.5	0.6
xxviii	(R)HTFSGVASVESSSGEAFHVGK(T)	−18.4	7.1	0.6
xxix	*(R)QHMDSSTSAASSSNY^{cam}CNQMMK(S)*	−18.6	7.7	0.6

loaded onto a 180 μm I.D. × 20 mm 5-μm symmetry® C18 trap (Waters), then eluted to a 100 μm I.D. × 10 cm 1.7-μm BEH130C18 column (Waters) using a linear gradient from 0% to 36% solvent B (ACN + 0.1% formic acid) in 36 min, 36–90% solvent B for 2 min. Solvent A was 0.2% formic acid in water. The peak areas (signal/noise ≥ 3) from extracted ion chromatograms of label-free nanoR-PLC-ESI-MS of the glycopeptides isolated by IP-NPLC were used for differential expression analysis of periplasmic glycoproteins between

the wt and mutant strains. NanoRPLC-ESI-MS/MS analyses were also performed with data-dependent analysis with a survey intensity threshold of 20 for the total digest and the glycopeptide and non-glycopeptide fractions of the periplasmic protein extract of the *pglD* mutant of *C. jejuni* 11168.

The glycopeptide fractions were iteratively analyzed with retention time segments of peptides by nanoRPLC-ESI-MS[2] and MS[3] using an Ettan MDLC system (Amersham Biosciences AB, Uppsala, Sweden)

Fig. 1. **Isolation of glycopeptides from a RNase B and fetuin tryptic digest using IP-NPLC.** Total ion mass spectrum from NPLC-ESI-MS of peptides in the glycopeptide fraction from the tryptic digest injected with the following ion-pairing reagents (*IPR*): *a*, none in 80% ACN/20% H_2O; (pH 4.7); *b*, 20% acetic acid in 80% ACN (pH 1.1); *c*, 0.25% HCl in 80% ACN/20% H_2O (pH 1.1); *d*, 50 mM NH_4HCO_3 + 1% HCl in 80% ACN/20% H_2O (pH 1.1); and *e*, 0.25% TFA in 85% ACN/15% H_2O (pH 2.1). *BPI*, base peak intensities. Throughout this report, the total ion mass spectrum for NPLC-ESI-MS of the glycopeptide fraction refers to the summed MS scans across the NPLC retention times between the most hydrophobic high mannose type glycopeptide (*a*) from RNase B (6.5–9.0 min depending on the ion-pairing reagents used (Fig. 2, *2nd top panel*) and the most hydrophilic sialylated glycopeptide from fetuin detected (earlier than 13 min). Non-glycopeptides eluted earlier than the high mannose type glycopeptide (*a*) of RNase B when HCl or TFA were used as ion-pairing reagents. The tryptic digest contains 14.2 pmol of RNase B tryptic digest + 13.8 pmol of fetuin tryptic digest. The solvent compositions used to dissolve the different IPR, CH_3COOH, HCl, and TFA with/without NH_4HCO_3 are as indicated in each panel of the figure. The pH values listed are those of the samples. All abundant ions that are not labeled in Fig. 1 are singly charged contaminants.

TABLE II

List of identified glycopeptides and corresponding glycoforms for the N- and O-linked oligosaccharides from the RNase B and fetuin tryptic digest

	Protein name	Sequence of identified peptides	Oligosacchride composition	Glycosylation site	Glycopeptide mass	
					Calculated	Measured
a[a]	**RNase B**	(R)**NLT**K(D)	GlcNAc$_2$ Man$_5$	Asn-60	1690.70	1690.70
b	"	"	GlcNAc$_2$ Man$_6$	"	1852.76	1852.78
c	"	"	GlcNAc$_2$ Man$_7$	"	2014.81	2014.82
d	"	"	GlcNAc$_2$ Man$_8$	"	2176.86	2176.86
e	"	"	Glc$_2$ NAc$_2$ Man$_9$	"	2338.91	2338.92
f	**Bovine fetuin**	(K)KL^{Cam} CPD^{Cam} CPLLAPL**NDS**R(V)	Hex$_6$ HexNAc$_5$ SA$_3$	Asn-138	4728.93	4728.93
g	"	"	Hex$_5$ HexNAc$_4$ SA$_2$	"	4072.70	4072.68
h	"	"	Hex$_6$ HexNAc$_5$ SA$_4$	"	5020.02	5019.96
i	"	(K)L^{Cam}CPD^{Cam} CPLLAPL**NDS**R(V)	Hex$_5$ HexNAc$_4$ SA$_2$	"	3944.61	3944.58
j	"	"	Hex$_6$ HexNAc$_5$ SA$_3$	"	4600.84	4600.80
k	"	"	Hex$_6$ HexNAc$_5$ SA$_4$	"	4891.93	4891.89
l	"	(R)RPTGEVYDIEIDTLETT^{Cam} CHVLDPTPLA **N**^{Cam} **CS**VR(Q)	Hex$_5$ HexNAc$_4$ SA$_2$	Asn-81	5875.54	5875.40
m	"	"	Hex$_6$ HexNAc$_5$ SA$_3$	"	6531.77	6531.68
n	"	"	Hex$_6$ HexNAc$_5$ SA$_4$	"	6822.86	6822.88
o	"	(R)VVHAVEVALATFNAES**NGS**YLQLVEISR(A)	Hex$_6$ HexNAc$_5$ SA$_3$	Asn-158	5876.57	5876.56
p	"	"	Hex$_6$ HexNAc$_5$ SA$_4$	"	6167.66	6167.56
q	"	"	Hex$_6$ HexNAc$_5$ SA$_5$	"	6458.75	6458.65
r	"	(R)VT^{Cam} CTLFQTQPVIPQPQPDGAEAEAP**S**AVP DAAGPT**TPS**AAGPPVASVVVGPSVVAVPLPLHR(A)	Hex$_2$ HexNAc$_2$ SA$_2$	T228–288	7326.60	7326.40
s	"	"	Hex$_2$ HexNAc$_2$ SA$_3$	"	7617.70	7617.70
t	"	"	Hex$_3$ HexNAc$_3$ SA$_3$	"	7982.81	7982.64
u	"	"	Hex$_3$ HexNAc$_3$ SA$_4$	"	8273.91	8273.79
v	"	"	Hex$_4$ HexNAc$_4$ SA$_4$	"	8639.04	8639.16
w	"	"	Hex$_4$ HexNAc$_4$ SA$_5$	"	8930.13	8930.43

[a] The 23 glycopeptides detected by IP-NPLC-MS of ~300 fmol of the RNase B and fetuin tryptic digest have been assigned a letter code (a–w) in order to facilitate identification. The consensus sequence of *N*-glycosylation and possible *O*-glycosylation sites are highlighted (12, 13). ^{Cam}C, carboxyamidomethylcysteine.

coupled to an LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific) to acquire MS^2 and MS^3 mass spectra for determination of the glycan composition of the glycopeptides and for protein identification. The peptides were first loaded onto a 300 μm I.D. $\times$ 5 mm C18 PepMap100^{TM} trap (LC Packings, San Francisco, CA), then eluted off to a 75 μm I.D. x 5 cm C18 Picofrit^{TM} column (New Objective, Woburn, MA) using the same gradient as described above.

Database Searching for Glycoprotein Identification—The peaklist files of MS^3 spectra of the *N*-linked glycopeptides of *C. jejuni* proteins were generated using Xcalibur 2.0.6 and were searched against the NCBI *C. jejuni* database (2007.09.05) with 12283 entries using the MASCOT^{TM} search engine (version 2.2.0) (Matrix Science) for protein identification. The database searches contained the following variable modifications: glycosylated asparagine by bacillosamine (2,4-diacetamido-2,4,6-trideoxyglucopyranose with a formula of $C_{10}H_{16}N_2O_4$) (+228 Da) (30), and oxidized methionines (+16 Da). The variable substitution of alanine by aspartic acid (+ 44 Da) was used as well because our *de novo* sequencing of MS^3 spectra of *C. jejuni* peptides indicated that alanine can be mutated to aspartic acid. The mass tolerance for precursor ions is $\pm$ 1.5 Da and the mass tolerance for fragment ions is $\pm$ 1.0 Da with no enzyme specified. The cut-off ions score of 40 above which ion scores indicate identity and 0.00% false positive rate were used for accepting individual MS^3 spectra. In addition, all spectral matches were verified manually. Proteins not matched to *C. jejuni* 11168 were excluded.

RESULTS

IP-NPLC for Isolating Tryptic Glycopeptides of Standard Glycoproteins Using Acids as Ion-pairing Reagents—A tryptic digest of an RNase B and fetuin mixture was chosen to illustrate the use of IP-NPLC for glycopeptide isolation (Fig. 1). The tryptic digestion of RNase B yields five *N*-linked glyco-

peptides: NLTK-GlcNAc$_2$Man$_{5-9}$, where Man is mannose, and GlcNAc is *N*-acetylglucosamine, with each of the five high mannose type glycans attached to Asn-60 (34). Fetuin has three *N*-linked sites (Asn-81, Asn1–38, and Asn-158) and four *O*-linked sites (Ser-253, Thr-262, Ser-264, and Ser-323), which have sialylated glycans attached and low *O*-glycosylation occupancy (12–13). Table I shows the non-glycopeptides (i-xxix) sequences of RNase B and fetuin identified from IP-NPLC-MS analysis representing 94% of the amino acid sequence of RNase B and 63% of the amino acid sequence of fetuin. The mass spectral peaks (i–xxix) labeled in Fig. 1 correspond to non-glycopeptides sequences listed in Table I. Hydrophobicity values, free energies of transfer ΔG_{octw} (kcal/mol) from n-octanol to water (octw), of non-glycopeptides (i–xxix) were calculated at the sample pH throughout this report using MPExTotalizer based on the Wimley-White water/octanol free energy scale for the 20 amino acids (a more negative energy value indicates a more hydrophilic peptide, Table I) (35–37). Table II lists the isolated glycopeptides (a–w) and their glycoforms from the standard tryptic digest using IP-NPLC. The mass spectral peaks (a–w) labeled in Fig. 1 correspond to the glycopeptides (a–w) listed in Table II.

When no acid was added to the sample (pH 4.7) prior to sample injection, the glycopeptides from RNase B (a–e) were not detected at all, and the fetuin glycopeptides were only observed with low MS signals from the glycopeptide fraction (*e.g.* the glycopeptide ion (*p*) of fetuin at *m/z* 1542.89(4+)),
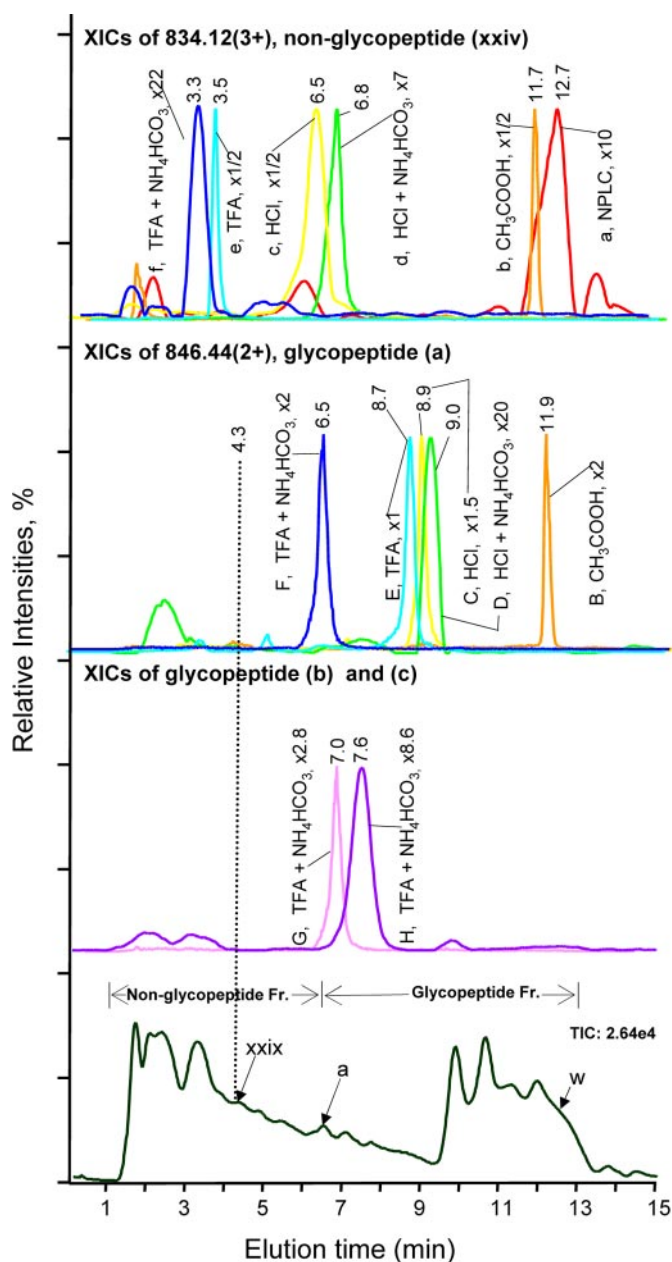
FIG. 2. **Effect of acids as ion-pairing reagents on peptide retention in IP-NPLC.** *Top panel*, overlaid extracted ion chromatograms (*XIC*s) of *m/z* 834.12 (3+) of the hydrophilic non-glycopeptide (*xxiv*), CKPVNTFVHESLADVQAVCSQK, from the RNase B and fetuin tryptic digest injected with the following ion-pairing reagents (*IPR*): (*a*) none; (*b*) 20% CH$_3$COOH; (*c*) 0.25% HCl; (*d*) 1% HCl + 50 mM NH$_4$HCO$_3$; (*e*) 0.25% TFA; and (*f*) 1% TFA + 50 mM NH$_4$HCO$_3$. *Second top panel*, overlaid XICs of *m/z* 846.44 (2+) of the least hydrophilic glycopeptide (*a*), NLTK-GlcNAc$_2$Man$_5$, from the same tryptic digest injected with the following IPRs: (*B*) 20% CH$_3$COOH; (*C*) 0.25% HCl; (*D*) 1%HCl + 50 mM NH$_4$HCO$_3$; (*E*) 0.25% TFA; and (*F*) 1% TFA + 50 mM NH$_4$HCO$_3$; *Second bottom panel*, (*G*) XIC of *m/z* 927.48 (2+) of the glycopeptide (*b*), NLTK-GlcNAc$_2$Man$_6$; (*H*) XIC of *m/z* 1008.43 (2+) of the glycopeptide (*c*), NLTK-GlcNAc$_2$Man$_7$. IPR used for *G* and *H*: 1% TFA + 50 mM NH$_4$HCO$_3$. *Bottom panel*, total ion chromatogram from IP-NPLC-ESI-MS of the standard tryptic digest injected with 1% TFA + 50 mM NH$_4$HCO$_3$. The magnification factors (*e.g.* x1) of the base peak

because of the signal suppression of co-eluting non-glyco-peptides (v, viii, xi, xiv, xviii-xxi, xxiv, and xxv) (Fig. 1*a*). Upon the addition of 20% acetic acid to the sample (pH 1.1), non-glycopeptides (v, viii, xi, xiv, xviii-xxii, xxiv, xxv, and xxix) still co-eluted with the glycopeptides (Fig. 1*b*) and dominated the total ion mass spectrum from NPLC-ESI-MS of the glycopeptide fraction.

In contrast, when 0.25% HCl was added to the sample (pH 1.1), the non-glycopeptides (i–xxix) were almost entirely removed, and the spectrum of the glycopeptide fraction was dominated by glycopeptide ions from the RNase B and fetuin digest (Fig. 1*c* and Table II). 0.25% HCl was insuffi-cient for removal of non-glycopeptides from the glycopep-tide fraction if the sample also contained 50 mM ammonium biocarbonate, a common reagent used in tryptic digestion. For example, relatively intense ion signals were observed for non-glycopeptides (v, viii, xviii-xx, xxiv, xxv, and xxviii) in the MS spectrum of the glycopeptide fraction when 50 mM ammonium bicarbonate was added to the sample (supple-mental Fig. S1a). However, increasing the HCl concentra-tion to 1% restored the separation of glycopeptides and non-glycopeptides (Fig. 1*d*).

When 0.25% trifluoroacetic acetic acid (TFA), a common ion-pairing reagent used to increase peptide retention in RPLC, was added to the sample, a similar degree of separa-tion between glycopeptides and non-glycopeptides was ob-tained when compared with 0.25% HCl (Fig. 1*d*). Again, when the tryptic digest contained 50 mM NH$_4$HCO$_3$, the TFA con-centration needed to be increased to 1% to eliminate the interference from non-glycopeptides (supplemental Fig. S1b). A total of 23 glycopeptides with five high mannose type glycopeptides from RNase B (a–e, Asn-60; Table II), 13 sia-lylated *N*-linked glycopeptides (f–q, Asn-38, 81, and 158), and six sialylated *O*-linked glycopeptides from fetuin (r–w, T228–288) were detected. The IP-NPLC-MS of the RNase B and fetuin tryptic digest were reproducible (relative standard de-viation (RSD) < 10% on base peak ion intensities; RSD ≤ 0.2% on the retention times of base peak ions (*n* = 3) (Fig. 1 (*c–e*) and supplemental Table S1)) with recovery of high man-nose type glycopeptides close to 100% (data not shown).

*Effect of Ion-pairing on Peptide Retention in IP-NPLC—*The effect of different acids on retention time shift of a non-glycopeptide (*e.g.* xxiv) and that of a glycopeptide (*e.g.* a) is demonstrated (Fig. 2). When there were no acids added to the sample, broad chromatographic peaks and peak tailing were

intensity (*BPI*) of the extracted ions were relative to the BPI of (*E*) for normalization. The elution time of the least hydrophilic tryptic glyco-peptide (*a*) of RNase B, *i.e.* the reference time determining the starting elution time of glycopeptide fraction, varies from 6.5 min to 11.9 min depending on the IPR used (*Second top panel*). The *broken line* at 4.3 min indicates the elution time of the most hydrophilic tryptic non-glycopeptide (*xxix*) present in the sample. w, glycopeptide (w) in Table II; Fr., fraction.
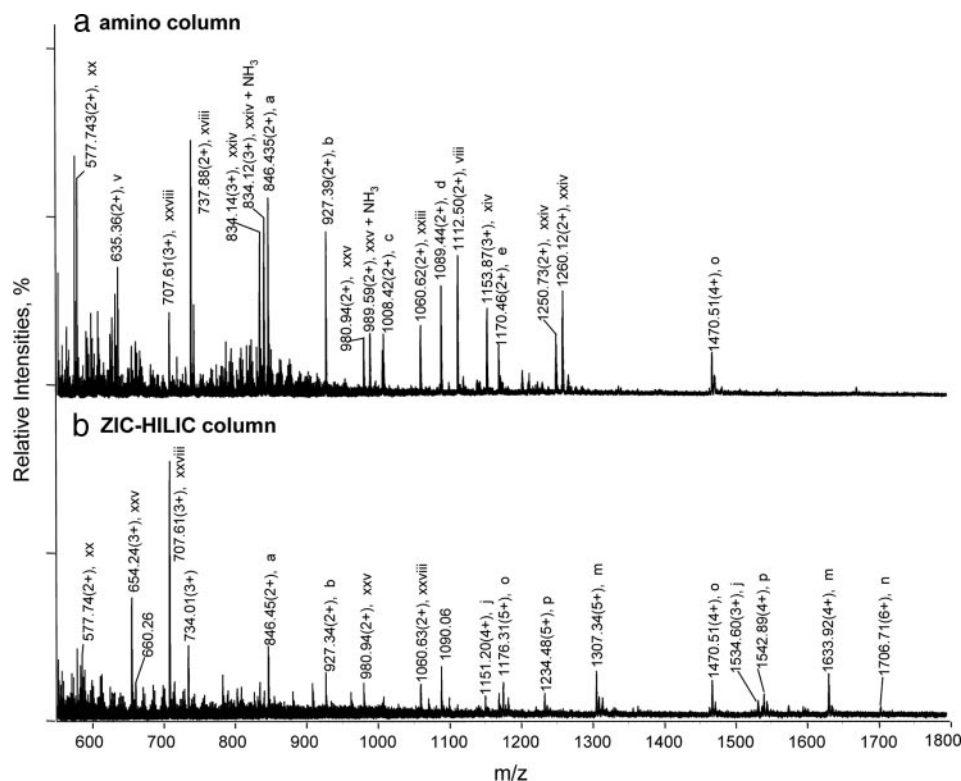
FIG. 3. **Effect of stationary phase charges on glycopeptide isolation using IP-NPLC.** Total ion mass spectrum from NPLC-ESI-MS of peptides in the glycopeptide fraction from the RNase B and fetuin tryptic digest injected with 1% TFA + 50 mM NH$_4$HCO$_3$ using different hydrophilic columns: (*a*) amino; (*b*) ZIC-HILIC. All other abundant ions that are not labeled in Fig. 3 are singly charged contaminants. Ammonium adducts and minor peptide peaks are not labeled.

observed for the non-glycopeptide (xxiv) at 12.7 min (Fig. 2, *red peak a*) whereas the glycopeptide (a) was not detected. When the sample was injected with 20% CH$_3$COOH (pH 1.1), the peptide peak shape was dramatically improved (Fig. 2, *orange peak b*) but resulted in co-elution of glyco-peptide (a) and non-glycopeptide (xxiv) at 11.9 min (*orange peaks B* and *b*). In contrast, when the samples were injected with HCl or TFA with/without 50 mM NH$_4$HCO$_3$, retention time differences of 2.2 min – 5.2 min between the non-glycopeptide (xxiv) and glycopeptide (a) were observed, therefore allowing removal of non-glycopeptide (xxiv) from the glycopeptide fraction. Similar retention time shifts were found for other peptides (Table S1).

IP-NPLC for Isolating Tryptic Glycopeptides on Different Hydrophilic Media—IP-NPLC experiments were carried out on a ZIC-HILIC and an amino column using the same exper-imental conditions as for the polyhydroxyethyl A column ex-cept that a flow rate of 1.1 μl/min was used for the 0.3 mm I.D. amino column. The three stationary phases are different in that the polyhydroxyethyl A is almost neutral, the ZIC-HILIC medium is zwitterionic, and the amino column is positively charged. The addition of TFA to the sample improved isolation selectivity of all three columns. However, isolation selectivity decreased considerably as the surface charges increased (Fig. 1e and Fig. 3). The neutral polyhydroxyethyl A column

was superior to the other two columns which also yielded much wider chromatographic peaks (data not shown). The results indicate that ionic interactions between peptides and the stationary phases are detrimental to glycopeptide selec-tivity. The polyhydroxyethyl A medium retains solutes almost solely through hydrophilic interactions, and it has only a slight positive charge on its surface under ~pH 2. Therefore, inter-ferences arising from ion-exchange processes in IP-NPLC using this column were absent.

Prediction of the Elution Order of Non-glycopeptides in IP-NPLC Using Calculated Hydrophobicity Values—We plot-ted the calculated hydrophobicity values of the 29 tryptic non-glycopeptides of RNase B and fetuin (Table I) against their IP-NPLC elution times (Fig. 4) with the addition of 0.25% HCl to the sample. The linearity of the plot ($r^2 = 0.84$) and the reproducibility of retention times of IP-NPLC suggest that co-elution of a non-glycopeptide with glycopeptides occur-ring in IP-NPLC can be estimated using their calculated hydrophobicity values. For example, the most hydrophilic non-glycopeptide (xxix), QHMDSSTSAASSSNYCNQMMK, of RNase B has a hydrophobicity value of −18.6 kcal/mol and eluted at 7.7 min, later than all the other tryptic non-glyco-peptides in the digest (i-xxviii) (Table I and supplemental Table S1). In contrast, the most hydrophobic glycopeptide (a) from RNase B, NLTK-GlcNAc$_2$Man$_5$, eluted at 8.9 min (Fig. 2, *peak*
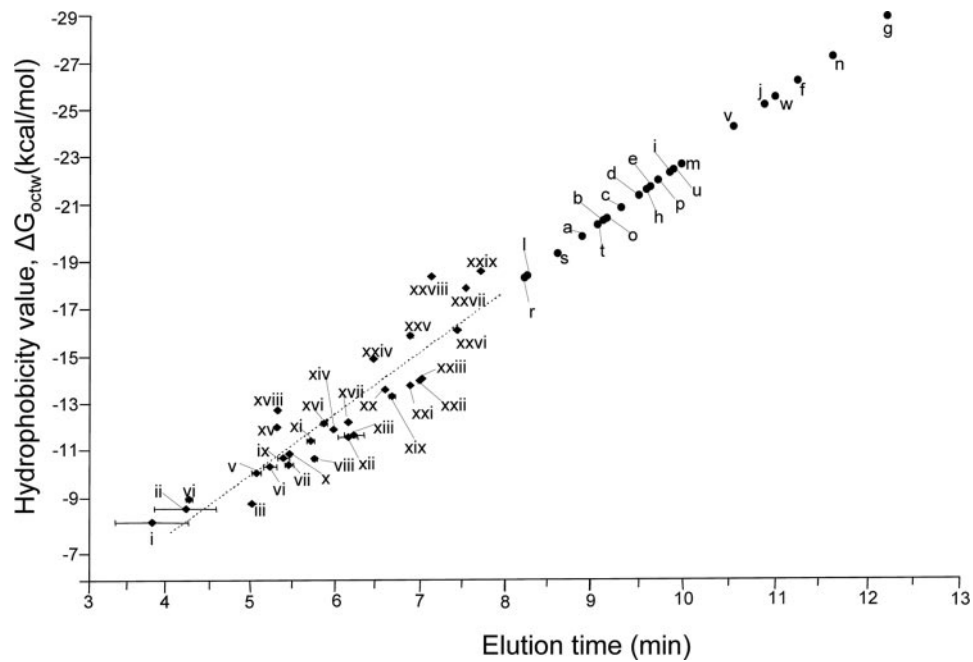
FIG. 4. **Plot of the hydrophobicity values of the non-glycopeptides of the RNase B and fetuin tryptic digest *versus* their elution times in IP-NPLC.** The sample was injected with 0.25% HCl (pH 1.1) as the ion-pairing reagent. The calculated hydrophobicity values (ΔG_{octw}, kcal/mol) are listed in Table I. The *broken line* indicates the predicted hydrophobicity values based on linear regression. The retention time (RT) errors were calculated based on triplicate IP-NPLC-MS analyses. The hydrophobicity values of glycopeptides (a–w), which can not be calculated using MPExTotalizer because of the presence of glycans, were projected based on their retention times and the linear regression equation for non-glycopeptides (i–xxix) (ΔG_{octw} = 2.6 RT - 3.1). The retention time shifts and ion intensity changes of the glycopeptides (a–w) and non-glycopeptides (i–xxix) with/without ion-pairing reagents are provided in supplemental Table S1.
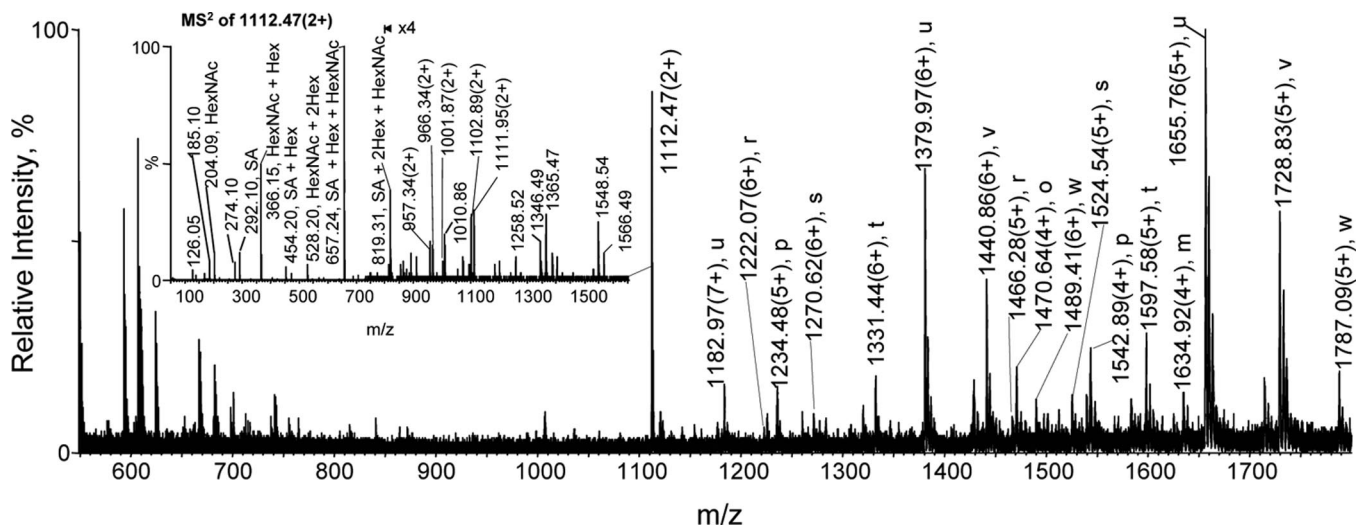

FIG. 5. **Isolation of sialylated *O*-linked glycopeptides from a fetuin tryptic digest using IP-NPLC with *N*-glycans first removed with PNGase F.** Total ion mass spectrum from IP-NPLC-MS of peptides in the glycopeptide fraction from a 13.8 pmol fetuin tryptic digest *N*-glycan deglycosylated with PNGase F. Sample was injected with 1% HCl + 50 mM NH_4HCO_3 as the ion-pairing reagent. Ammonium adducts of glycopeptides are not labeled. *Inset*, MS[2] spectrum of the ion at *m/z* 1111.95 (2+).

C). Therefore, non-glycopeptides with hydrophobicity values higher (more positive) than −18.6 kcal/mol are likely to elute earlier than glycopeptides modified by high mannose type heptasaccharides in IP-NPLC. The occurrence of tryptic peptides of common proteins having hydrophobicity values lower than −18.6 kcal/mol is low (data not shown) suggesting that

IP-NPLC is capable of isolating glycopeptides cleanly from tryptic digests of complex protein mixtures.

IP-NPLC for Isolating Sialylated O-linked Tryptic Glycopeptides of Fetuin—The *O*-linked glycopeptide ions from a fetuin tryptic digest (T228–288) were observed by direct IP-NPLC-MS analysis (glycopeptides r-w; Fig. 1*e*). However, the
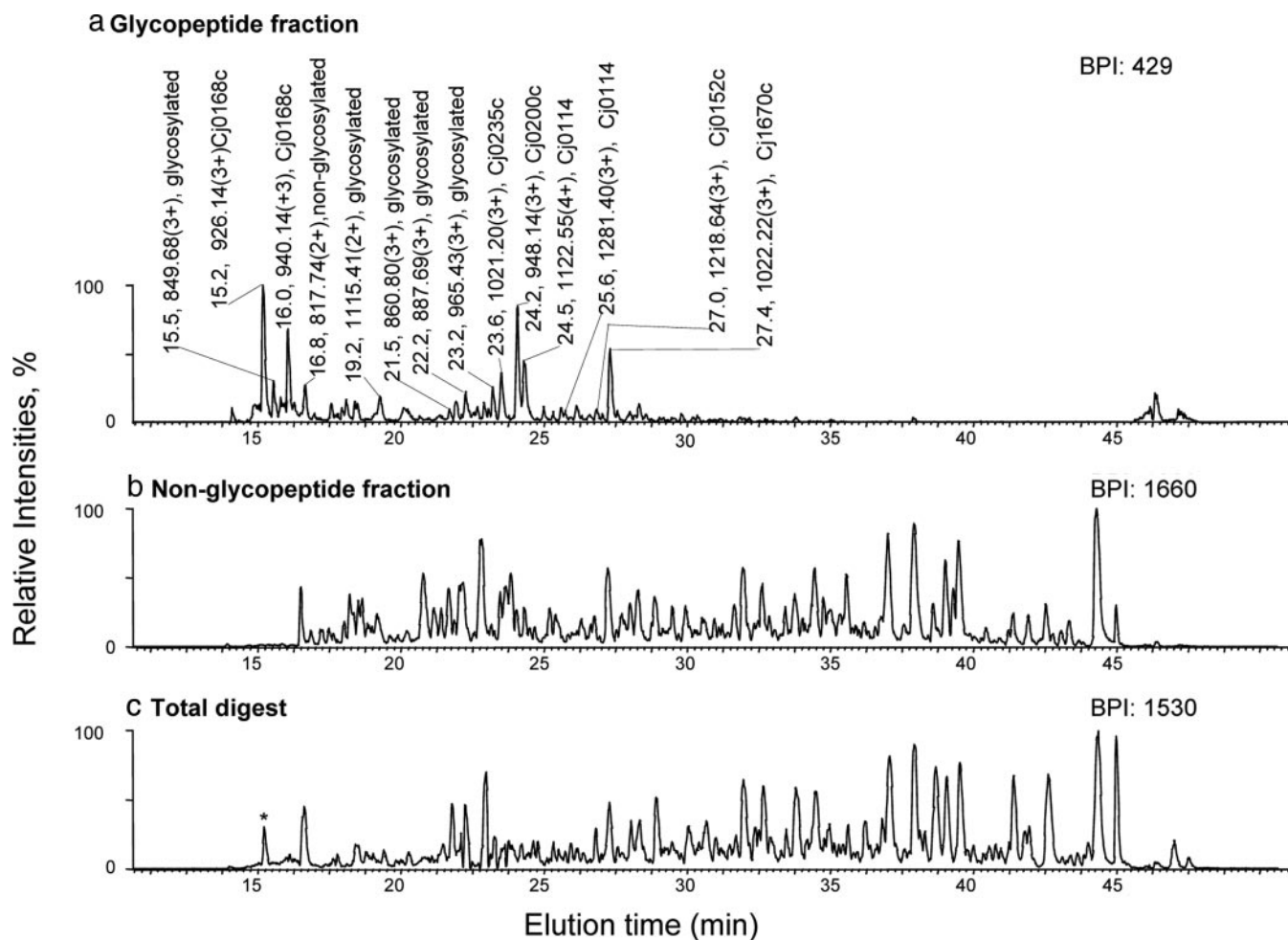
FIG. 6. **Analysis of glycopeptides isolated by IP-NPLC from a tryptic digest of periplasmic proteins extracted from *C. jejuni* 11168, *pglD* mutant.** Base peak intensity (*BPI*) chromatograms of nanoRPLC-MS of a tryptic digest of 4-µg periplasmic protein extracts from the *pglD* mutant of *C. jejuni* 11168 from the (*a*) glycopeptide fraction; (*b*) non-glycopeptide fraction; (*c*) total digest. The sample contained 1% HCl as the ion-pairing reagent. The gene numbers are provided in the peak labels for glycopeptides determined by MS[3] analysis. An example of the MS[2] and MS[3] spectra obtained for these glycopeptides is presented in Fig. 7. The amino acid sequence of the identified glycopeptides is provided in Table III. *Glycosylated*, base peak ions that are confirmed to be from glycopeptides by MS[2] analysis (MS[2] spectra containing the signature oxonium ions from these glycopeptide ions upon CID are not shown). The only glycopeptide detected from the total digest (*lower panel*) among the top 50 base peaks is marked by an *asterisk*.

MS signals of these *O*-linked glycopeptide ions were much lower compared with those of the *N*-linked glycopeptide ions, largely because of the low glycosylation occupancy (13). When the fetuin tryptic digest was treated with PNGase F to remove the *N*-linked glycans, the prominent ions detected in the IP-NPLC glycopeptide fraction were all from *O*-linked glycopeptides of fetuin (r–w) (Fig. 5). The ion at *m/z* 1112.3 (3+) is an unidentified glycopeptide (Fig. 5, *inset*).

Isolation of Glycopeptides from Complex Peptide Mixtures Using IP-NPLC for Quantitative Glycoprotein Profiling Using Label-free MS—Glycopeptides of a complex tryptic digest of total periplasmic protein extracts of the *C. jejuni* 11168 wt and the *pglD* mutant were isolated by IP-NPLC and subjected to nanoRPLC-MS analysis. The base peak chromatograms (BPCs) of the glycopeptide and non-glycopeptide IP-NPLC

fractions and the total digest are presented in Fig. 6. The BPCs of the non-glycopeptide fraction and total digest (Fig. 6, *b* and *c*) were very similar whereas that of the glycopeptide fraction (Fig. 6*a*) was far less complex because of removal of non-glycopeptides by IP-NPLC. MS/MS experiments determined that the most prominent ions (*m/z* 800–1800) from the BPC of the glycopeptide fraction were all from glycopeptides (Table III) whereas non-glycopeptide ions dominated the BPCs of the non-glycopeptide fraction and total digest.

To demonstrate that IP-NPLC can isolate glycopeptides reproducibly for quantitative analysis, the wt and *pglD* periplasmic extracts were fractionated in triplicate (40 µg each), and the glycopeptide fractions were subjected to label-free nanoRPLC-MS analyses (4 µg each). The peak areas of extracted ion chromatograms of the identified glycopeptides

TABLE III
Qualitative analysis of glycoproteins from the periplasmic protein extracts of C. jejuni 11168 (wt) and the isogenic pglD mutant

Accession number	Gene name	Protein name	Precursor[a]	z	Sequences of identified glycopeptides[b]	Oligosaccharide composition[c]	Ion score[d]	Sequence coverage	RT
			m/z					%	*min*
YP_002343627	Cj0168c	Putative periplasmic protein	940.14	3	(F)ANTPS**DVNQT**HTK(A)	BacGalNAc$_5$Glc		23	16.0
"	"	"	926.14	3	"	^{deAc}BacGalNAc$_5$Glc	79		15.2
NP_281430	Cj0235c	Protein export membrane protein	1035.21	3	(N)SIAPSAPQLPS**DVNSS**K(−)	BacGalNAc$_5$Glc		14	24.8
"	"	"	1021.20	3	"	^{deAc}BacGalNAc$_5$Glc	55	23.6	
YP_002343658	Cj0200c	Putative periplasmic protein	962.14	3	(K)LEGTIAQIY**DNNK** (T)	BacGalNAc$_5$Glc		11	25.8
"	"	"	948.14	3	"	^{deAc}BacGalNAc$_5$Glc	97		24.2
CAL34323	Cj0152c	Hypothetical protein	1231.31	3	(K)NISI**ENNISENNTT**LLDEEK(N)	BacGalNAc$_5$Glc		6	27.2
"	"	"	1218.64	3	"	^{deAc}BacGalNAc$_5$Glc	66		27.0
YP_002345038	Cj1670c	Putative membrane protein or CgpA[e]	1036.22	3	(K)T**DQNIT**LVAPPEFQK(E)	BacGalNAc$_5$Glc		7	29.2
"	"	"	1022.22	3	"	^{deAc}BacGalNAc$_5$Glc	57		27.4
YP_002343574	Cj0864	Putative periplasmic protein	1058.10	3	(K)**D^{oxi}MNVS**K(A)	BacGalNAc$_5$Glc	49	5	19.7
"	"	"	1037.10	3	"	^{deAc}BacGalNAc$_5$Glc			
YP_002344271	Cj0114	Putative periplasmic protein	1133.05	4	(K)TITPSVVVSTTDSNSTI**ENNNT**QNTQDDK(A)	BacGalNAc$_5$Glc		26	25.6
"	"	"	1122.55	4	"	^{deAc}BacGalNAc$_5$Glc	72		24.5
"	"	Putative periplasmic protein	1295.35	3	(R)LSQVEENNQNI**ENNFT**SEIQK(L)	BacGalNAc$_5$Glc	60		27.5
"	"	"	1281.40	3	"	^{deAc}BacGalNAc$_5$Glc			25.6
YP_002344190	Cj0783	Periplasmic nitrate reductase cytochrome c-type subunit	1692.12	2	(K)VNLV**EANFT**TLQPGESTR(F)	BacGalNAc$_5$Glc	41	10	29.7
"	"	"	1671.12	2	"	^{deAc}BacGalNAc$_5$Glc			

[a] The *m/z* and charge of precursors (MS[2] fragments) for acquisition of MS[3] spectra are provided in supplemental Fig. S2.
[b] The periplasmic protein, Cj0114, is identified by two peptides, and all other glycoproteins identified are single-peptide-based.
[c] Bac, 2,4-diacetamido-2,4,6-trideoxyglucopyronose; ^{deAc}Bac, 2-acetamido-4-amino-2,4,6-trideoxyglucopyronose, *i.e.* ^{deAc}Bac is mono-acetylated Bac at the C-2 position only (30, 39, 40). The consensus sequence of possible *N*-glycosylation sites for prokaryotic proteins are highlighted (32). ^{oxi}M, oxidized methionine; wt, wild-type; ″, the same as above.
[d] Ion scores are available only for precursors on which MS[3] were acquired and used for database searches with 0.00% false positive rate.
[e] According to Linton *et al.* (41).

listed in Table III were manually quantified for differential expression analysis between the wt and mutant strains. The relative standard deviation ($n = 3$) of the glycopeptide ions identified by this IP-NPLC label-free MS approach was between 3–45% irrespective of the peptide abundances and retention times, indicating that IP-NPLC was capable of isolating trace level glycopeptides reproducibly (Tables III and IV) (Fig. 6).

Multiple Stage Tandem MS of Differentially Expressed Tryptic Glycopeptides for Identification of Glycoproteins—The glycopeptide ions in the glycopeptide fractions of the wt and *pglD* mutant extracts were subjected to MS[2] and MS[3] analysis to determine their oligosaccharide composition and their amino acid sequences. The *N*-linked glycan of *C. jejuni* has been previously determined to be GalNAc-α1,4-GalNAc-α1,4-[Glcβ1,3]-GalNAc-α1,4-GalNAc-α1,4-GalNAc-α1,3-Bac-β1 (BacGalNAc$_5$Glc residue mass: 1406 Da), where Bac is 2,4-diacetamido-2,4,6-trideoxyglucopyranose (30). In addition, *in vitro* biosynthesis experiments have identified PglD as the acetyltransferase that modifies the uridine diphosphate (UDP)-2-acetamido-4-amino-2,4,6-trideoxyglycopyranose at the C-4 position to form UDP-2,4-diacetamido-2,4,6-trideoxyglucopyranose using acetyl-coenzyme A as the acetyl group donor (39).

As an example, the MS[2] spectra of equivalent glycopeptides from the wt and mutant strains are presented in Fig. 7, *a* and *b*, respectively. Both MS/MS spectra are dominated by ions associated with loss of sugar residues from the glycan modification. As expected, the oligosaccharide composition of the wt glycopeptide was determined to be BacGalNAc$_5$Glc (Fig. 7*a*). The MS[2] spectrum of the *pglD* glycopeptide ion (Fig. 7*b*) is very similar to that of the wt suggesting that oligosaccharide composition and order are alike in the two strains. However, the reducing end sugar is smaller by 42 Da in the mutant strain and has been annotated as ^{deAc}Bac (2-acetamido-4-amino-2,4,6-trideoxyglucopyranose). The glycan composition and structure of this glycopeptide ion are therefore assigned as GalNAc-α1,4-GalNAc-α1,4-[Glcβ1,3]-GalNAc-α1,4-GalNAc-α1,4-GalNAc-α1,3-^{deAc}Bac-β1 (residue mass: 1364 Da).

MS[3] analysis was performed on the doubly protonated MS[2] fragment ion at *m/z* 851.45, which corresponded to the unmodified peptide (Fig. 7*c*). Database searching matches the MS[3] fragment ions to TDQNITLVAPPEFQK from the *C. jejuni* periplasmic protein, Cj1670c (Table III) with an ion score of 57 as defined by the MASCOT™ algorithm. A total of 9 glycopeptides corresponding to eight periplasmic glycoproteins were identified by this approach (Table III). *N*-glycosylation was confirmed for Cj0200c, Cj0152c, Cj1670c, and Cj0114 (30, 32), and four novel glycopeptides from Cj0168c, Cj0235c, Cj0864, and Cj0783 were identified. The MS[2] and MS[3] spectra for these other glycopeptides are presented in supplemental Fig. S2.

TABLE IV

Quantitative analysis of glycoproteins from the periplasmic protein extracts of C. jejuni 11168 (wt) and the isogenic pglD mutant

wt, wild-type; Intensities, peak area; S. D., standard deviation is based on triplicate nanoRPLC-MS analyses of triplicate IP-NPLC glycopeptide fractions; nd, not detected; —, a ratio is not given because the peptides were not detected in wt or mutant; ″, the same as above.

Gene name	Protein name	Strain	Count	Intensities in wt	% error	Intensities in mutant	% error	Ratio wt/mutant	% error
				(mean ± S. D.)		(mean ± S. D.)		(mean ± S. D.)	
Cj0168c	(YP_002343627) putative periplasmic protein	wt	1	16166 ± 503	3	1310 ± 157	12.0	12.34 ± 1.53	12.39
″	″	pglD	1	92 ± 18	19	1496 ± 250	16.7	0.06 ± 0.02	25.60
Cj0235c	(NP_281430) protein export membrane protein	wt	1	472 ± 26	6	27 ± 3	11.1	17.49 ± 2.17	12.41
″	″	pglD	1	70 ± 5	7	757 ± 126	16.6	0.09 ± 0.02	17.86
Cj0200c	(YP_002343658) putative periplasmic protein	wt	1	204 ± 87	43	416 ± 40	11.4	5.02 ± 2.22	44.33
″	″	pglD	1	204 ± 12	6	2033 ± 133	6.5	0.10 ± 0.01	8.78
Cj0152c	(CAL34323) hypothetical protein	wt	1	424 ± 103	24	nd		only in wt	—
″	″	pglD	1	nd		242 ± 81	33.3	only in mutant	—
Cj1670c	(YP_002345038) putative membrane protein or CgpA[c]	wt	1	224 ± 65	29	43 ± 4	10.1	5.22 ± 1.59	30.51
″	″	pglD	1	70 ± 5	7	1150 ± 114	9.9	0.06 ± 0.01	11.95
Cj0864	(YP_002343574) putative periplasmic protein	wt	1	2186 ± 914	42	186 ± 9	5.1	11.71 ± 4.93	42.09
″	″	pglD	1	nd		nd		not in both	—
Cj0114	(YP_002344271) putative periplasmic protein	wt	1	41 ± 5	11	nd		only in wt	—
″	″	pglD	1	121 ± 20	16	1393 ± 137	9.8	0.09 ± 0.02	18.91
Cj0114	(YP_002344271) putative periplasmic protein	wt	1	408 ± 167	41	32 ± 4	11.3	12.7 ± 5.42	42.42
″	″	pglD	1	115 ± 29	25	84 ± 16	19.3	1.37 ± 0.43	31.37
Cj0783	(YP_002344190) Periplasmic nitrate reductase cytochrome c-type subunit	wt	1	527 ± 9	16	nd		only in wt	—
″		pglD	1	nd		nd		nd in both	—

[a] According to Linton *et al.* (41).

[b] The putative periplasmic protein, Cj0114 is identified and quantified by two glycopeptides whose detailed information can be found in Table III.

Recovery and Selectivity of IP-NPLC for Isolating Tryptic Glycopeptides from Complex Peptide Mixtures—The nano-RPLC-MS extracted ion chromatograms for three glycopeptide ions from the *pglD* total digest as well as the corresponding glycopeptide and non-glycopeptide IP-NPLC fractions are presented in Fig. 8. The glycopeptide ions were detected in the total (Fig. 8, *c*, *f*, and *i*) and glycopeptide fractions (Fig. 8, *a*, *d*, and *g*) but were absent in the non-glycopeptide fraction (Fig. 8, *b*, *e*, and *h*). The glycopeptide ion intensities were comparable in the total and glycopeptide chromatograms. Similar results were obtained for all of the glycopeptides listed in Table III, regardless of their abundance, indicating that the recovery of these glycopeptides by IP-NPLC was close to 100%. Furthermore, only a few precursor ions from the non-glycopeptide fraction generated a fragment ion at *m/z* 204.09 upon CID, the majority of which were identified as non-glyco-peptides (>85%), indicating that virtually no glycopeptides were being lost in the non-glycopeptide fraction (Fig. 6*b*, Fig. 9, and supplemental Table S2).

The glycopeptide selectivity per nanoRPLC-MS analysis (scan range: *m/z* 800–1800) of the glycopeptide fraction (defined as the number of base peak glycopeptides determined by the *N*-consensus sequence or the signature oxonium ions of glycopeptides upon CID divided by the number of the prominent peptides identified from the glycopeptide fraction) in our approach was 90% (Fig. 6 and Table III). Non-glycopeptides rich in hydrophilic amino residues were present in the glycopeptide fraction, *e.g.* ALDNDNQEDQTESR from fla-

vodoxin (Cj1382c) eluting at 16.8 min (Fig. 6*a*). However, the occurrence of these peptides in a complex tryptic digest appeared low thus emphasizing the efficiency of our method.

DISCUSSION

In this report we describe a new IP-NPLC method for quantitatively isolating glycopeptides from complex tryptic digests using acids as ion-pairing reagents and polyhydroxyethyl A stationary phase. The addition of no acids or acetic acid to the sample led to co-elution of non-glycopeptides with glycopeptides from the RNase B and fetuin tryptic digest (Fig. 1*b* and Fig. 2). In contrast, the use of HCl or TFA resulted in separation of glycopeptides from non-glycopeptides (Figs. 1 and 2, supplemental Fig. S1b, and supplemental Table S1). It is apparent that the dramatic improvement in isolation of glycopeptides using IP-NPLC over NPLC was not simply because of the pH change in the sample because only hydrochloric acid but not acetic acid led to isolation of glycopeptides under the same sample pH (Fig. 1, *b* and *c*). Other ions present in the sample solution can diminish the effect of HCl or TFA on glycopeptide isolation selectivity, *e.g.* NH_4HCO_3 (supplemental Fig. S1a), but this can be overcome by increasing the concentration of the two acids (Fig. 1*d* and supplemental Fig. S1b). The results suggest that acidic anions present in the samples significantly affect the selectivity of NPLC for the isolation of glycopeptides.

A total of 18 different sialylated glycopeptides from fetuin, including 6 *O*-linked and 12 *N*-linked glycopeptides, and 5
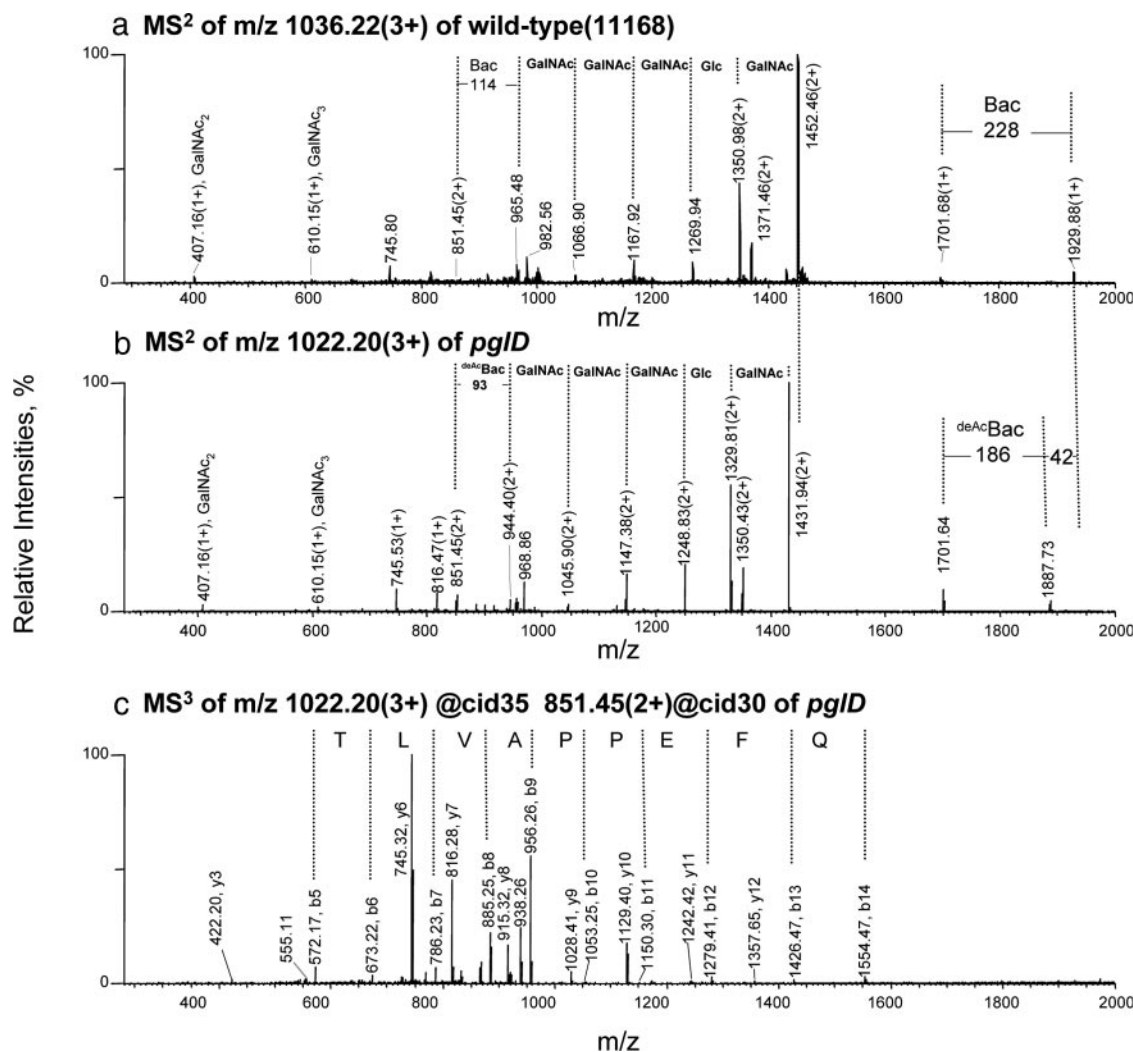
FIG. 7. **Multiple stage tandem MS analyses of differentially expressed tryptic glycopeptides of the wt and the *pglD* mutant of *C. jejuni*.** MS^2 spectrum of the glycopeptide ion at (*a*) *m/z* 1036.22 (3+) from the wt; (*b*) the equivalent ion at *m/z* 1022.20 (3+) from the *pglD mutant*; (*c*) MS^3 spectrum of the same ion at *m/z* 1022.20 (3+) from the *pglD mutant*. The MS^3 fragment ions were matched to TDQNITLVAPPEFQK of the *C. jejuni* periplasmic protein, Cj1670c by database searching using MASCOT^{TM} with an ion score of 57 (NCBI accession no. YP_002345038). *Bac*, bacillosamine (2,4-diacetamido-2,4,6-trideoxyglucopyranose); ^{deAc}*Bac*, monoacetylated bacillosamine at the C-2 position only (2-acetamido-4-amino-2,4,6-trideoxyglucopyranose) (30, 39, 40).

high mannose type *N*-linked glycopeptides from RNase B were isolated by IP-NPLC from a tryptic digest of a mixture of these two glycoproteins (Fig. 1*e* and Table II) (12, 13, 23). A typical *N*-linked carbohydrate derived from a eukaryotic host usually contains 7–25 monosaccharide units (O-linked carbohydrates have a similar size range but with less branching and may be as short as one monosaccharide unit) (4). Thus, NLTK-GlcNAc$_2$Man$_{5-9}$ glycopeptides from RNase B are modified with shorter high mannose type glycans and resemble non-glycopeptides in terms of hydrophilicity. Therefore, these glycopeptides have proven to be difficult to resolve from non-glycopeptides by other chromatographic methods (23, 25–26). In contrast, our IP-NPLC method was effective at resolving these less hydrophilic glycopeptides and the sialylated fetuin glycopeptides from the digest of RNase B and

fetuin mixture (Fig. 1, Table I, and supplemental Table S1). The method described here is more effective than other chromatographic methods in that it works well with a wide range of glycopeptide species (23, 25, 29).

The retention time shifts of non-glycopeptides (Fig. 2 and supplemental Table S1) can be explained by ion-pairing processes that occur between the acidic anions and the counter ions of non-glycopeptides in solution (29, 43–45). The earlier elution of non-glycopeptides upon the addition of HCl or TFA to the sample rather than CH$_3$COOH can be attributed to formation of ion pairs $-NH_3^+CH_3COO^-$, $-NH_3^+Cl^-$, and $-NH_3^+CF_3COO^-$, which are progressively more hydrophobic ($-NH_3^+$ refers to the ionized amine groups in non-glycopeptides). The higher hydrophilicity of CH$_3$COO$^-$ or $-NH_3^+CH_3COO^-$ in IP-NPLC is likely because of the hydro-

FIG. 8. **Recovery and selectivity of IP-NPLC for isolating tryptic glyco-peptides from complex peptide mixtures of the *C. jejuni pgID* mutant.** Extracted ion chromatograms (*XIC*) of nanoRPLC-MS at *m/z* 1281.40 (3+) of a glycopeptide from the putative periplasmic protein, Cj0114, in the (*a*) Glycopeptide fraction; (*b*) Non-glycopeptide fraction; (*c*) Total digest. XICs of nanoRPLC-MS at *m/z* 1021.20 (3+) of a glycopeptide from protein export membrane protein, Cj0235c, in the (*d*) Glycopeptide fraction; (*e*) Non-glycopeptide fraction; (*f*) Total digest. XICs of nanoRPLC-MS at *m/z* 926.14 (3+) of a glycopeptide from the putative periplasmic protein, Cj0168c, in the (*g*) Glycopeptide fraction; (*h*) Non-glycopeptide fraction and (*i*) Total digest. The amino acid sequences of these glycopeptides are provided in Table III. The *insets* in (*d*), (*e*), and (*f*) show the combined mass spectra at *m/z* 1020.20 (3+) corresponding to the glycopeptide fraction, non-glycopeptide fraction, and total digest (23–24 min) for comparison. *PA*, peak area of the extracted ions; *Nd*, not detected; *Fr.*, fraction.
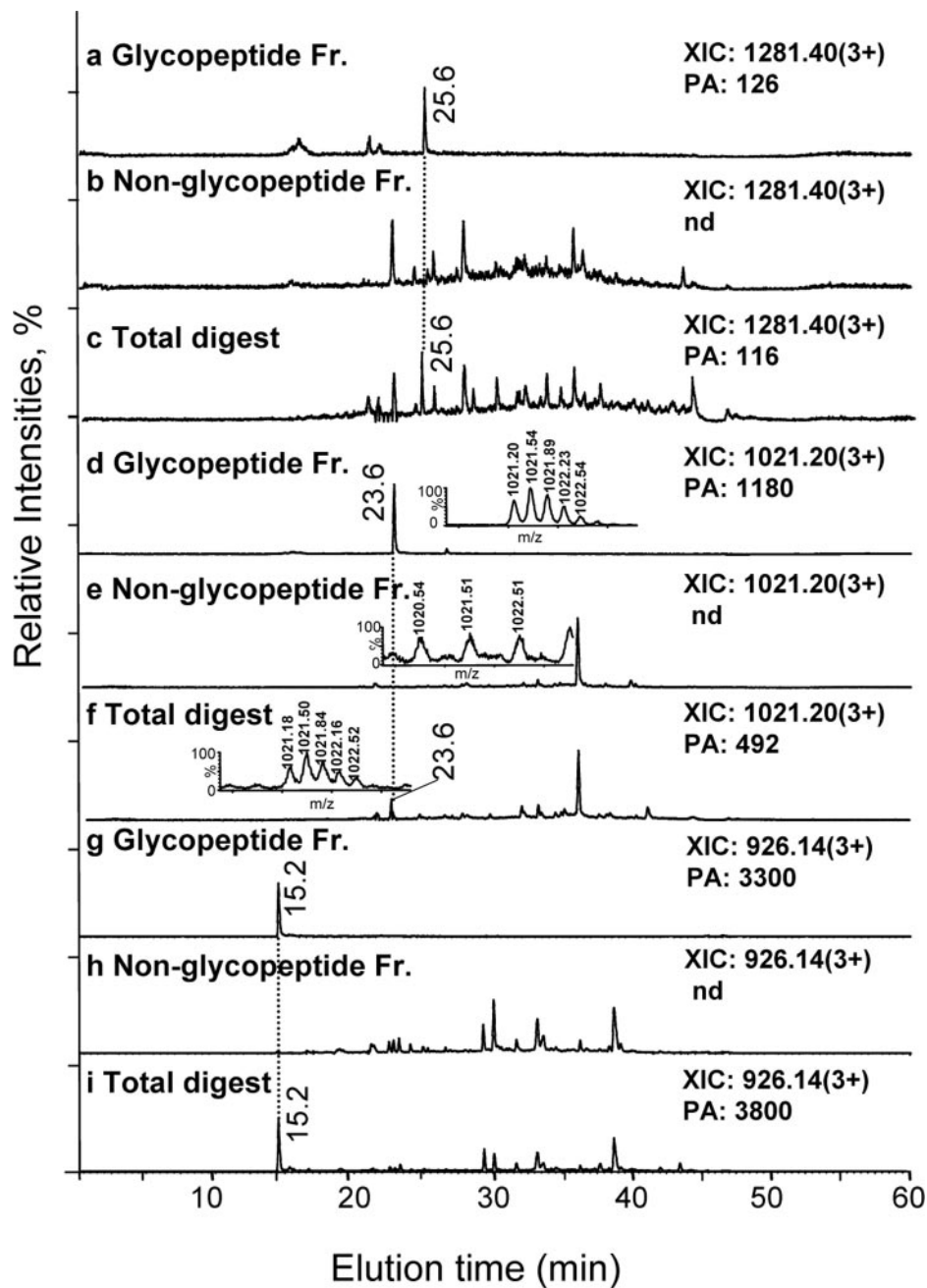
gen bonding ability of oxygen with the stationary phase whereas chlorine can not form strong hydrogen bonding. Covalently bound fluorine is not capable of hydrogen bonding, and it apparently dramatically weakened the hydrogen bonding ability of oxygen on CF_3COO^- (46). As a result, $-NH_3^+CF_3COO^-$ is more hydrophobic than $NH_3^+CH_3COO^-$ and $-NH_3^+Cl^-$.

The acetate anion formed relatively hydrophilic ion pairs with the non-glycopeptides allowing them to interact with the polyhydroxyethyl A medium and eluted with the glycopeptides (Fig. 2, *orange peak B* and *b*). In contrast, the ion pairs between the charged non-glycopeptides and the Cl^- or

CF_3COO^- were more hydrophobic and eluted earlier from the column (Fig. 2, *peak b versus peaks d–f*). In addition, improvement in chromatographic peak shapes was obtained using the three acids as ion-pairing reagents over NPLC alone (Fig. 2, *red peak a versus peaks b–f*) (43).

The retention times of glycopeptides were effected to a much lesser extent by ion pairing because of the strong non-ionic electrostatic interactions between their glycans and the hydrophilic stationary phase (Fig. 2, *peak C versus c*, *peak D versus d*, *peak E versus e*, *peak F versus f*). As a result, non-glycopeptides eluted much earlier than glycopeptides in IP-NPLC. Our data indicate that formation of ion-pair com-

FIG. 9. **NanoRPLC-MS^2 of the non-glycopeptide fraction of the *pglD* mutant using data dependent analysis: (*a*) Base peak intensity (BPI) chromatogram and (*b*) Extracted ion chromatogram (XIC) at *m/z* 204.09 of the non-glycopeptide fraction.** The ions labeled in (b) were all identified as non-glycopeptides, except the unidentified ion at *m/z* 897.42 (2+), and their sequences are provided in Table S2. The magnification factors in (a) (*e.g.* x80 for the base peak intensity (BPI) of the ion at *m/z* 851.51(2+)) of the nanoRPLC-MS^2 of the non-glycopeptide fraction were relative to the most abundant ion for comparison.
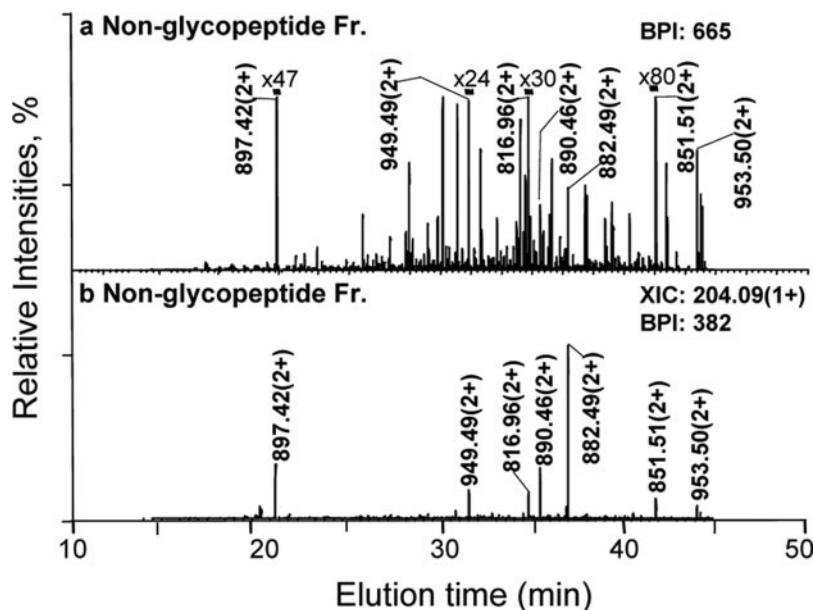
plexes between acidic anions and charged non-glycopeptides can increase the retention time differences between non-glycopeptides and glycopeptides, thereby leading to fractionation of a complex tryptic digest into a glycopeptide and a non-glycopeptide fraction (Figs. 1 and 6) (supplemental Table S1).

Ion pairs $-COO^-H^+$ between the hydronium ion (H_3O^+) and the negatively charged carboxyl group of non-glycopeptides ($-COO^-$) are likely to form in solution (47). In addition, Wimley *et al.* have determined that the neutral $-COOH$ terminus of peptides determined by pH, are considerably more hydrophobic than $-COO^-$ (35–37). Therefore, ion-pairing processes between hydronium ions and peptides likely contribute to the decreased retention described in Fig. 2. However, the decreased retention because of formation of hydronium ion pairs with peptides will be generally less significant compared with anions because hydrogen of an ion-pair is capable of hydrogen bonding with stationary phase whereas chlorine or covalently bound fluorine is not (46).

We demonstrate for the first time that the surface charges of the stationary phase decrease glycopeptide isolation selectivity using IP-NPLC. Furthermore, the elution order of non-glycopeptides in IP-NPLC can be predicted reasonably well using their hydrophobicity values (Fig. 4). Non-glycopeptides with a hydrophobicity value more positive than -18.6 kcal/mol (ΔG_{octw}) will likely elute in the non-glycopeptide fraction. In addition, we demonstrated that IP-NPLC can selectively isolate the *O*-linked glycopeptides from a fetuin digest after the *N*-glycans were removed with PNGase F (Fig. 5). This worked remarkably well and we intend to apply this methodology to the study of *O*-glycosylation in other biological systems.

We applied IP-NPLC to reproducibly and quantitatively isolate glycopeptides from tryptic digests of periplasmic proteins extracted from *C. jejuni* 11168 wt and *pglD* mutant for subsequent label-free nanoRPLC-MS analysis of glycoprotein expression (Tables III and IV). We chose to re-examine the *C. jejuni pglD* mutant with our new technique because in our previous study (48) we detected minor amounts of protein glycosylation in this mutant but because of sensitivity issues were unable to characterize these modifications any further. Four of the nine glycopeptides identified in this study have not been observed before. For example, the most abundant glycopeptide isolated from the wt strain is derived from the putative periplasmic protein, Cj0168c. This glycoprotein has a mass of only 5,813 Da (amino acid mass only) and therefore would not have been detected by the gel-based methods used in previous studies to resolve the glycoprotein mixtures prior to MS analysis (30). Remarkably, the reproducibility of this label-free MS approach using IP-NPLC is comparable with that obtained using hydrazide chemistry and stable isotope labeling (15). In addition, IP-NPLC does not involve lengthy chemical procedures as for hydrazide chemistry and also preserves the glycan information.

The presence of predominantly monoacetylated Bac residues on glycoproteins isolated from the C jejuni *pglD* mutant confirms that PglD is an acetyltransferase and that the oligosaccharyltransferase, PglB, is still capable of transferring ^{deAc}BacGalNAc$_5$Glc. However, the quantitative information on the differentially expressed *N*-linked glycoproteins further demonstrates the importance of Bac acetylation at *C*-2 for optimal PglB activity in *C. jejuni* (40). The quantitative proteomic data demonstrate for the first time that glycopeptides modified by ^{deAc}BacGalNAc$_5$Glc are also present in the wt strain, albeit at low abundance, indicating that acetylation of Bac is incomplete even in the presence of an active PglD enzyme. In addition, it is apparent that there is another

acetyltransferase complementing the PglD activity as we previously predicted (48) because glycopeptides modified by BacGalNAc$_5$Glc are also present in the mutant although at much lower abundance compared with the wt (Tables III and IV).

The glycopeptide selectivity of our approach reaches 90% based on the total number of glycopeptides and peptides identified in the *C. jejuni* glycopeptide fractions (Fig. 6 and Table III). Based on the retention time difference observed between the most hydrophilic non-glycopeptide (xxix) and the glycopeptides, NLTK-GlcNAc$_2$Man$_{5-7}$ in the RNase B and fetuin digest (Fig. 2), IP-NPLC may allow glycopeptides modified with just a trisaccharide similar to GlcNAc$_2$Man to be enriched from complex tryptic digests, although the glycopeptide selectivity would be compromised somewhat. Recovery of glycopeptides by IP-NPLC is close to 100% for the samples tested based on the absence of glycopeptide ion signals in the non-glycopeptide fraction (Figs. 8 and 9). This glycopeptide selectivity and recovery is comparable with that obtained with hydrazide chemistry capturing tryptic glycopeptides (16). We believe that this IP-NPLC method can be used for isolating glycopeptides at much lower levels because low abundance glycopeptides were effectively recovered from the tryptic digests of periplasmic protein extracts from *C. jejuni* (Fig. 6, *m/z* 1218.40(3+) at 25.6 min, Cj0114).

In comparison to other chromatographic approaches for enriching glycopeptides, this new IP-NPLC method offers the best glycopeptide selectivity and recovery for complex peptide mixtures to our knowledge (19, 22, 23, 25, 28–29). Nevertheless, there is still room for improvement as some non-glycopeptides were present in the glycopeptide fraction. The non-glycopeptides abundances, the magnitude of ion-pairing equilibrium constants under a given solvent composition (49), and the hydrophobicity of non-glycopeptides are likely the important factors concerning contamination. Further optimization of experimental conditions will improve glycopeptide selectivity and may allow the isolation of glycopeptides modified by glycans as short as trisaccharides. We believe that IP-NPLC will make a significant contribution to glycoproteomics research in the future.

REFERENCES

1. Jaeken, J., and Matthijs, G. (2001) Congenital disorders of glycosylation. *Annu. Rev. Genomics Hum. Genet.* **2,** 129–151
2. Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M., and Dwek, R. A. (2007) The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu. Rev. Immunol.* **25,** 21–50
3. Rudd, P. M., Elliott, T., Cresswell, P., Wilson, I. A., and Dwek, R. A. (2001) Glycosylation and the immune system. *Science* **291,** 2370–2376
4. Varki, A., Cummings, R., Esko, J., Freeze, H., Marth, J., Hindsgaul, O., Paulson, J., Lowe, J., Arbo, A., Manzi, A., and Powell, L. (1999) *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press, NY
5. Ferrara, N., and Kerbel, R. S. (2005) Angiogenesis as a therapeutic target. *Nature* **438,** 967–974
6. Burton, D. R., and Dwek, R. A. (2006) Sugar determines antibody activity. *Science* **313,** 627–628
7. Ferrari, L., Seregni, E., Martinetti, A., Van, Graafeiland, B., Nerini-Molteni, S., Botti, C., Artale, S., Cresta, S., and Bombardieri, E. (1998) Chromogranin A measurement in neuroendocrine tumors. *Int. J. Biol. Markers* **13,** 3–9
8. Urban, D., Myers, R., Manne, U., Weiss, H., Mohler, J., Perkins, D., Markiewicz, M., Lieberman, R., Kelloff, G., Marshall, M., and Grizzle, W. (1999) Molecular targets for chemoprevention of prostate cancer - evaluation of biomarker modulation by fenretinide in prostate cancer patients. *Eur. Urol.* **35,** 429–438
9. Huang, X., Ushijima, K., Komai, K., Takemoto, Y., Motoshima, S., Kamura, T., and Kohno, K. (2004) Co-expression of Y box-binding protein-1 and P-glycoprotein as a prognostic marker for survival in epithelial ovarian cancer. *Gynecol. Oncol.* **93,** 287–291
10. Domon, B., and Aebersold, R. (2006) Mass spectrometry and protein analysis. *Science* **312,** 212–217
11. Dell, A., and Morris, H. R. (2001) Glycoprotein structure determination by mass spectrometry. *Science* **291,** 2351–2356
12. Medzihradszky, K. F., Maltby, D. A., Hall, S. C., Settineri, C. A., and Burlingame, A. L. (1994) Characterization of protein *N*-glycosylation by reversed-phase microbore liquid chromatography/electrospray mass spectrometry, complementary mobile phases, and sequential exoglycosidase digestion. *J. Am. Soc. Mass Spectrom.* **5,** 350–358
13. Peterman, S. M., and Mulholland, J. J. (2006) A novel approach for identification and characterization of glycoproteins using a hybrid linear ion trap/FT-ICR mass spectrometer. *J. Am. Soc. Mass Spectrom.* **17,** 168–179
14. Carr, S. A., Huddleston, M. J., and Bean M. F. (1993) Selective identification and differentiation of *N*-and *O*-linked oligosaccharides in proteins by liquid chromatography-mass spectrometry. *Protein Sci.* **2,** 183–196
15. Zhang, H., Li, X. J., Martin, D. B., and Aebersold, R. (2003) Identification and quantification of *N*-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **21,** 660–666
16. Sun, B., Ranish, J. A., Utleg, A. G., White, J. T., Yan, X., Lin, B., and Hood, L. (2007) Shotgun glycopeptide capture approach coupled with mass spectrometry for comprehensive glycoproteomics. *Mol. Cell. Proteomics* **6**, 141–149
17. Hill, J. J., Moreno, M. J., Lam, J. C., Haqqani, A. S., and Kelly, J. F. (2009) Identification of secreted proteins regulated by cAMP in glioblastoma cells using glycopeptide capture and label-free quantification. *Proteomics* **9,** 535–549
18. Ghosh, D., Krokhin, O., Antonovici, M., Ens, W., Standing, K. G., Beavis, R. C., and Wilkins, J. A. (2004) Lectin affinity as an approach to the proteomic analysis of membrane glycoproteins. *J. Proteome Res.* **3,** 841–850
19. Zhao, J., Simeone, D. M., Heidt, D., Anderson, M. A., and Lubman, D. M. (2006) Comparative serum glycoproteomics using lectin selected sialic acid glycoproteins with mass spectrometric analysis: Application to pancreatic cancer serum. *J. Proteome Res.* **5,** 1792–1802
20. Durham, M., and Regnier, F. E. (2006) Targeted glycoproteomics: serial lectin affinity chromatography in the selection of O-glycosylation sites on proteins from the human blood proteome. *J. Chromatogr. A* **1132,** 165–173
21. Yang, Z., and Hancock, W. S. (2005) Monitoring glycosylation pattern changes of glycoproteins using multi-lectin affinity chromatography. *J. Chromatogr. A* **1070,** 57–64
22. Larsen, M. R., Højrup, P., and Roepstorff, P. (2005) Characterization of gel-separated glycoproteins using two-step proteolytic digestion combined with sequential microcolumns and mass spectrometry. *Mol. Cell. Proteomics* **4,** 107–119
23. Hägglund, P., Bunkenborg, J., Elortza, F., Jensen, O. N., and Roepstorff, P. (2004) A new strategy for identification of *N*-glycosylated proteins and

unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation. *J. Proteome Res.* **3,** 556–566

24. Hagglund, P., Matthiesen, R., Elortza, F., Højrup, P., Roepstorff, P., Jessen O. N., and Bunkenborg, J. (2007) An enzymatic deglycosylation scheme enabling identification of core fucosylated *N*-glycans and O-glycosylation site mapping of human plasma proteins. *J. Proteome Res.* **6,** 3021–3031

25. Thaysen-Andersen, M., Thøgersen, I. B., Nielsen, H. J., Lademann, U., Brünner, N., Enghild, J. J., and Højrup, P. (2007) Rapid and individual-specific glycoprofiling of the low abundance *N*-glycosylated protein tissue inhibitor of metalloproteinases-1. *Mol. Cell. Proteomics* **6,** 638–647

26. Thaysen-Andersen, M., and Hojrup, P. (2006) Enrichment and characterization of glycopeptides from gel-separated glycoproteins. *Am. Biotechnol. Lab.* **24,** 14–17

27. Larsen, M. R., Jensen, S. S., Jakobsen, L. A., and Heegaard, N. H. (2007) Exploring the sialiome using titanium dioxide chromatography and mass spectrometry. *Mol. Cell. Proteomics* **6,** 1778–1787

28. Wada, Y., Tajiri, M., and Yoshida, S. (2004) Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.* **76,** 6560–6565

29. Ding, W., Hill, J. J., and Kelly, J. (2007) Selective enrichment of glycopeptides from glycoprotein digests using ion-pairing normal-phase liquid chromatography. *Anal. Chem.* **79,** 8891–8899

30. Young, N. M., Brisson, J. R., Kelly, J., Watson, D. C., Tessier, L., Lanthier, P. H., Jarrell, H. C., Cadotte, N., St., Michael, F., Aberg, E., and Szymanski, C. M. (2002) Structure of the *N*-linked glycan present on multiple glycoproteins in the gram-negative bacterium *Campylobacter jejuni.* *J. Biol. Chem.* **277,** 42530–42539

31. Wacker, M., Linton, D., Hitchen, P. G. Nita-Lazar, M., Haslam, S. M., North, S. J., Panico, M., Morris, H. R., Dell, A., Wren, B. W., and Aebi, M. (2002) *N*-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli. Science* **298,** 1790–1793

32. Kowarik, M., Young, N. M., Numao, S., Schulz, B. L., Hug, I., Callewaert, N., Mills, D. C., Watson, D. C., Hernandez, M., Kelly, J. F., and Wacker, M. (2006) Definition of the bacterial *N*-glycosylation site consensus sequence. *EMBO J.* **25,** 1957–1966

33. Gagliardi, L. G., Castells, C. B., Clarra, R., Ràfols, C., Rosés, M., and Bosch, E. (2007) δ Conversion parameter between ph scales (^s_wpH and ^s_spH) in acetonitrile/water mixtures at various compositions and temperatures. *Anal. Chem.* **79,** 3180–3187

34. An, H. J., Peavy, T. R., Hedrick, J. L., and Lebrilla, C. B. (2003) Determination of *N*-glycosylation sites and site heterogeneity in glycoproteins. *Anal. Chem.* **75,** 5628–5637

35. Wimley, W. C., Creamer, T. P., and White, S. H. (1996) Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. *Biochemistry* **35,** 5109–5124

36. Wimley, W. C., Gawrisch, K., Creamer, T. P., and White, S. H. (1996) Direct measurement of salt-bridge solvation energies using a peptide model

system: Implications for protein stability. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 2985–2990

37. Jaysinghe, S., Hristova, K., Wimley, W. C., Snider, C., and White, S. H. (2006) www.blanco.biomol.uci.edu/mpex

38. Crabb, J. W., Carlson, A., Chen, Y., Goldflam, S., Intres, R., West, K. A., Hulmes, J. D., Kapron, J. T., Luck, L. A., Horwitz, J., and Bok, D. (1998) Structural and functional characterization of recombinant human cellular retinaldehyde-binding protein. *Protein Sci.* **7,** 746–757

39. Olivier, N. B., Chen, M. M., Behr, J. R., and Imperiali, B (2006) *In vitro* biosynthesis of UDP-*N,N'*-diacetylbacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. *Biochemistry*, **45,** 13659–13669

40. Wacker, M., Feldman, M. F., Callewaert, N., Kowarik, M., Clarke, B. R., Pohl, N. L., Hernandez, M., Vines, Enrique D., Valvano, Miguel A., Whitfield, C., and Aebi, M. (2006) Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 7088–7093

41. Linton, D., Allan, E., Karlyshev, Andrey V., Cronshaw Andrew D., and Wren, B. W. (2002) Identification of *N*-acetylgalactosamine-containing glycoproteins PEB3 and CgpA in *Campylobacter jejuni. Mol. Microbiol.* **43,** 497–508

42. McNaught, A. D., and Wilkinson, A. (1997) *IUPAC. Compendium of Chemical Terminology* (the "Gold Book") 2nd Ed., pp. 1126–1127, Blackwell Scientific Publications, Oxford, UK

43. Hancock, W. S., Bishop, C. A., Prestidge, R. L., Harding, D. R., and Hearn, M. T. (1978) Reversed-phase high-pressure liquid chromatography of peptides and proteins with ion-pairing reagents. *Science* **200,** 1168–1170

44. Winstein, S., and Robinso, G. C. (1958) Salt effects and ion pairs in solvolysis and related reactions. IX. The *fhreo-3-p-* Anisyl-2 -butyl System. *J. Am. Chem. Soc.* **80,** 169–181

45. Chadwick, S., Englich, U., Ruhlandt-Senge, K., Watson, C., Bruce, A. E., and Bruce, M. R. M. (2000) Formation of separated *versus* contact ion pairs in alkali metal thiolates and selenolates. *J. Chem. Soc. Dalton Trans.* **13,** 2167–2173

46. Dunitz, J. D., and Taylor, R. (1997) Organic fluorine hardly ever accepts hydrogen bonds. *Chem.* **3,** 89–98

47. Bonner, O. D. (1980) Unusual ion pairs and their relationship to acid ionization. *J. Solution Chem.* **9,** 877–884

48. Kelly, J. F., Jarrell, H., Millar L., Tessier, L., Fiori, L. M., Lau P. C., Allan, B., and Szymanski, C. S. (2006) Biosynthesis of the *N*-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. *J. Bacteriol.* **188,** 2427–2434

49. Katsuta, S., Ishitani, T., Suzuki, M., Ishii, Y., Kudo, Y., and Takeda, Y. (2004) Equilibrium study on ion-pair formation in water and distribution between water and *m*-Xylene of tetraalkylammonium picrates. *J. Solution Chem.* **33,** 437–451