Glycomic and sialoproteomic data of gastric carcinoma cells overexpressing ST3GAL4

Mereiter, Stefan; Magalhães, Ana; Adamczyk, Barbara; Jin, Chunsheng; Almeida, Andreia; Drici, Lylia; Vea, Maria Ibanez; Larsen, Martin Røssel; Kolarich, Daniel; Karlsson, Niclas G; Reis, Celso A

Published in:
Data in Brief

DOI:
10.1016/j.dib.2016.03.022

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Data in Brief 7 (2016) 814–833

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Glycomic and sialoproteomic data of gastric carcinoma cells overexpressing ST3GAL4

Stefan Mereiter [a,b,c], Ana Magalhães [a,b], Barbara Adamczyk [d], Chunsheng Jin [d], Andreia Almeida [e,f], Lylia Drici [g], Maria Ibáñez-Vea [g], Martin R. Larsen [g], Daniel Kolarich [e], Niclas G. Karlsson [d], Celso A. Reis [a,b,c,h,*]

[a] *I3S - Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal*
[b] *Institute of Molecular Pathology and Immunology of the University of Porto - IPATIMUP, Porto, Portugal*
[c] *Institute of Biomedical Sciences of Abel Salazar - ICBAS, University of Porto, Portugal*
[d] *Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Sweden*
[e] *Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, 14424 Potsdam, Germany*
[f] *Free University Berlin, Berlin, Germany*
[g] *Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark*
[h] *Medical Faculty, University of Porto, Portugal*

ARTICLE INFO

Article history:
Received 22 December 2015
Received in revised form
22 February 2016
Accepted 4 March 2016
Available online 14 March 2016

Keywords:
N-glycome
O-glycome
Gastric cancer
Sialyltransferase
Sialoproteome

ABSTRACT

Gastric carcinoma MKN45 cells stably transfected with the full-length *ST3GAL4* gene were characterised by glycomic and sialoproteomic analysis. Complementary strategies were applied to assess the glycomic alterations induced by ST3GAL4 overexpression. The *N*- and *O*-glycome data were generated in two parallel structural analyzes, based on PGC-ESI-MS/MS. Data on glycan structure identification and relative abundance in ST3GAL4 overexpressing cells and respective mock control are presented. The sialoproteomic analysis based on titanium-dioxide enrichment of sialopeptides with subsequent LC-MS/MS identification was performed. This analysis identified 47 proteins with significantly increased sialylation. The data in this article is associated with the research article published in Biochim Biophys Acta "Glycomic analysis of gastric carcinoma cells discloses

DOI of original article: http://dx.doi.org/10.1016/j.bbagen.2015.12.016
* Corresponding author at: I3S - Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal.
E-mail address: celsor@ipatimup.pt (C.A. Reis).

http://dx.doi.org/10.1016/j.dib.2016.03.022

glycans as modulators of RON receptor tyrosine kinase activation in cancer" [1].

Specifications Table

Subject area	*Biology*
More specific subject area	*Glycobiology in cancer*
Type of data	*Tables*
How data was acquired	*N-and O-glycome analysed by PGC-LC ESI MS/MS performed on an LTQ ion trap mass spectrometer and PGC-nanoLC ESI MS/MS performed on an amaZon ETD Speed ion trap. Sialoproteome performed using Easy-nLC II system and Orbitrap Fusion Tribrid system.*
Data format	*Analyzed*
Experimental factors	*MKN45 cells stably transfected with the full-length ST3GAL4 gene or empty vector (mock).*
Experimental features	*MKN45 cells grown in vitro in RPMI medium with 10% FBS.*
Data source location	*Not applicable*
Data accessibility	*Data is within this article.*

Value of the data

- Data shows the *N*- and *O*-glycome of the MKN45 gastric carcinoma cells and ST3GAL4 sialyltransferase overexpressing cells.
- Data provides a list of glycoproteins with altered sialylated *N*-glycans in ST3GAL4 overexpressing cells, including several cancer associated proteins.
- These data are valuable as a source for novel biomarkers for gastric cancer.

1. Data

This data article includes the *N*- and *O*-glycome of MKN45 cells and assesses the glycosylation alterations induced by ST3GAL4 overexpression. In addition, we provide the data on the proteins with significant increased sialylated *N*-glycans upon ST3GAL4 overexpression.

2. Experimental design, materials and methods

Glycomic and sialoproteomic analyses were performed as described below comparing MKN45 cells stably transfected with the full-length *ST3GAL4* gene or empty vector (mock). Glycome data and statistical evaluation are shown in Tables 1, 2, 3 and 4. Sialoproteome data and statistical evaluation are shown in Tables 5 and 6.

3. *N*-glycomic strategy I: sample preparation and PGC LC-ESI-MS/MS

Frozen cell pellets (10^7 cells) of mock or *ST3GAL4* transfected MKN45 cells [2] were directly resuspended in 7 M urea, 2 M thiourea, 40 mM Tris, 2% CHAPS, 10 mM DTT and 1% protease inhibitor (Sigma-Aldrich, St. Louis, MO). The cell membranes were disrupted by 10 times 10 s sonication with 16 amplitudes and 1 min on ice in between, and subsequent shaking at 4 °C overnight. To reduce the viscosity of the lysates, the DNA was degraded by adding 1 μl benzonase® nuclease (250 units, Sigma-Aldrich) and 30 min incubation at 37 °C. In order to impair refolding of proteins, 25 mM iodoacetamide were added for alkylation during 1 h in the dark. The lysates were centrifuged for 30 min at 14,000 rpm and the supernatants were transferred to a fresh tube.

Then, solubilized proteins were concentrated by adding 150 μl of supernatant on a 10 kDa cut-off spinfilter (PALL, Port Washington, NY), spinning down for 5 min with 12,000xg and washing 3 times with 100 μl 50 mM NH_4HCO_3, pH 8.4. *N*-linked oligosaccharides were released in the spinfilter using 20 μl 50 mM NH_4HCO_3 and PNGase F (5 mU, Prozyme, Hayward, CA) with incubation at 37 °C overnight. Subsequently, the *N*-glycans were collected by washing 3 times with 20 μl H_2O and dried in Speedvac. Reactions were quenched with 1 μl of glacial acetic acid and *N*-glycan samples were desalted and dried as previously described [3]. *N*-glycan samples were subjected to LC-ESI-MS/MS analysis using a 10 cmx250 μm I.D. column, prepared in-house, containing 5 μm porous graphitized carbon (PGC) particles (Thermo Scientific, Waltham, MA). Glycans were eluted using a linear gradient from 0% to 40% acetonitrile in 10 mM NH_4HCO_3 over 40 min at a flow rate of 10 μl/min. The eluted *N*-glycans were detected using a LTQ ion trap mass spectrometer (Thermo Scientific) in negative-ion mode with an electrospray voltage of 3.5 kV, capillary voltage of −33.0 V and capillary temperature of 300 °C. Air was used as a sheath gas and mass ranges were defined dependent on the specific structure to be analyzed. The data were processed using the Xcalibur software (version 2.0.7, Thermo Scientific) and manually interpreted from their MS/MS spectra.

Optionally prior to analysis *N*-glycan were digested in 50 mM sodium phosphate, pH 6.0 at 37 °C overnight using sialidase S (4 mU, ProZyme) that releases α2-3 linked non-reducing terminal sialic acids

Table 1
Identified *N*-glycan structures I.

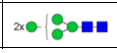
m/z		Composition	Proposed structure	Mock replicates (n=3)					ST3GAL4 replicates (n=3)					
z=1	z=2			%	%	%	Avg	SD	%	%	%	Avg	SD	p-value
708	/	HexNAc1Hex3		0.04	0.02	0.2	0.09	0.10	0.08	0.08	0.03	0.06	0.03	0.3636
749	/	HexNAc2Hex2		0.36	0.3	0.3	0.32	0.03	0.3	0.29	0.23	0.27	0.04	0.0955
895	/	HexNAc2Hex2Fuc1		1.18	0.95	1.58	1.24	0.32	0.75	0.69	1.38	0.94	0.38	0.181
911	/	HexNAc2Hex3		0.17	0.27	1.27	0.57	0.61	0.32	0.39	0.37	0.36	0.04	0.3054
911	/	HexNAc2Hex3		0.02	0.03	0.45	0.17	0.25	0.07	0.09	0.13	0.10	0.03	0.3356
911	/	HexNAc2Hex3		0.1	0.09	0.04	0.08	0.03	0.09	0.09	0.06	0.08	0.02	0.4421
1057	/	HexNAc2Hex3Fuc1		0.96	0.84	2.28	1.36	0.80	0.75	0.65	0.98	0.79	0.17	0.1719
1073	/	HexNAc2Hex4		0.06	0.12	0.68	0.29	0.34	0.11	0.13	0.12	0.12	0.01	0.2437
1073	/	HexNAc2Hex4		0.3	0.25	0.18	0.24	0.06	0.23	0.3	0.13	0.22	0.09	0.3604
1202	/	2HexNAc3Hex1Sia		2.29	2.24	6.23	3.59	2.29	1.94	3.09	3.34	2.79	0.75	0.3077
1202	/	HexNAc2Hex3Sia1		0.29	0.21	0.69	0.40	0.26	0.18	0.28	0.48	0.31	0.15	0.33
1235	/	HexNAc2Hex5		0.41	0.45	0.24	0.37	0.11	0.41	0.55	0.24	0.40	0.16	0.3896

Table 1 (*continued*)

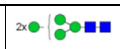
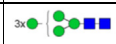
1235	/	HexNAc2Hex5		0.2	0.17	0.09	0.15	0.06	0.17	0.21	0.07	0.15	0.07	0.4765
1235	/	HexNAc2Hex5		1.53	1.25	2.26	1.68	0.52	1.81	1.8	2.55	2.05	0.43	0.1974
1364	681	HexNAc2Hex4Sia1		1.2	1.28	2.29	1.59	0.61	0.85	1.28	1.25	1.13	0.24	0.1592
1397	698	HexNAc2Hex6		0.26	0.24	0.28	0.26	0.02	0.3	0.33	0.25	0.29	0.04	0.1463
1397	698	HexNAc2Hex6		9.32	8.64	7.54	8.50	0.90	7.52	8.27	6.52	7.44	0.88	0.1082
1422	711	HexNAc3Hex4Fuc1		na	na	1.29	1.29	/	na	na	0.49	0.49	/	n/a
1422	711	HexNAc3Hex4Fuc1		na	na	0.41	0.41	/	na	na	0.22	0.22	/	n/a
1438	719	HexNAc3Hex5		na	na	0.31	0.31	/	na	na	0.2	0.20	/	n/a
1559	779	HexNAc2Hex7		0.25	0.24	0.17	0.22	0.04	0.32	0.37	0.19	0.29	0.09	0.1542
1559	779	HexNAc2Hex7		6.14	6.24	4.61	5.66	0.91	4.72	5.47	3.41	4.53	1.04	0.116
1559	779	HexNAc2Hex7		3.71	3.39	3.28	3.46	0.22	3.15	3.45	2.79	3.13	0.33	0.1173
1567	783	HexNAc3Hex3Sia1		na	na	0.6	0.60	/	0	0	0	0.00	0.00	n/a
1567	783	HexNAc3Hex4Sia1		na	na	0.7	0.70	/	na	na	0.63	0.63	/	n/a
1584	792	HexNAc3Hex5Fuc1		na	na	0.63	0.63	/	na	na	0.57	0.57	/	n/a
1600	800	HexNAc3Hex6		na	na	0.52	0.52	/	na	na	0.59	0.59	/	n/a
1713	856	HexNAc3Hex4Sia1Fuc1		7.34	5.42	8.25	7.00	1.44	4.71	3.79	6.94	5.15	1.62	0.1067
1721	860	HexNAc2Hex8		9.65	9.88	6.89	8.81	1.66	8.53	10.15	5.82	8.17	2.19	0.3544
1729	864	HexNAc3Hex5Sia1		0.95	1.59	3.19	1.91	1.15	1.19	1.48	1.77	1.48	0.29	0.2944
1729	864	HexNAc3Hex5Sia1		0.96	1	1.36	1.11	0.22	1.02	0.89	1.15	1.02	0.13	0.2978
1729	864	HexNAc3Hex5Sia1		na	na	0.8	0.80	#DIV/0!	na	na	1.09	1.09	/	n/a
1729	864	HexNAc3Hex5Sia1		0	0	0	0.00	0.00	na	na	0.58	0.58	/	n/a
1746	873	HexNAc3Hex6Fuc1		na	na	0.41	0.41	/	na	na	0.32	0.32	/	n/a
1787	893	HexNAc4Hex5Fuc1		0.39	0.41	1.7	0.83	0.75	0.38	0.27	0.63	0.43	0.18	0.2247

Table 1 (*continued*)

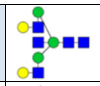
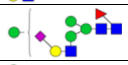
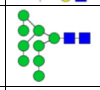
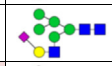
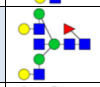
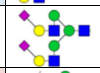
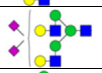
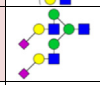
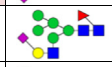
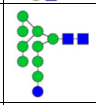
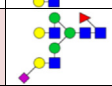
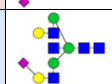
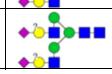
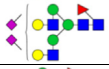
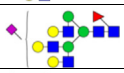
/	922	HexNAc5Hex5		0.06	0.1	0.09	0.08	0.02	0	0.01	0	0.00	0.01	0.0081 **
/	937	HexNAc3Hex5Sia1Fuc1		1.74	1.45	1.48	1.56	0.16	1.83	1.3	2.04	1.72	0.38	0.2703
/	941	HexNAc2Hex9		3.31	3.36	2.63	3.10	0.41	3.32	3.74	2.36	3.14	0.71	0.4687
/	945	HexNAc3Hex6Sia1		1.44	1.31	1.12	1.29	0.16	1.51	1.32	1.62	1.48	0.15	0.1024
/	945	HexNAc3Hex6Sia1		0.04	0.03	0	0.02	0.02	0.3	0.27	0.35	0.31	0.04	0.0009 ***
/	966	HexNAc4Hex5Sia1		0.87	0.78	0.71	0.79	0.08	0.98	0.96	0.73	0.89	0.14	0.1706
/	995	HexNAc5Hex5Fuc1		na	na	0.74	0.74	/	na	na	0.15	0.15	/	n/a
/	1010	HexNAc3Hex5Sia2		7.58	9.96	7.37	8.30	1.44	5.23	7.11	6.89	6.41	1.03	0.0723
/	1010	HexNAc3Hex5Sia2		3.17	3.43	1.89	2.83	0.82	5.67	6.23	6.69	6.20	0.51	0.0034 **
/	1010	HexNAc3Hex5Sia2		1.24	1.1	0.25	0.86	0.54	3.07	2.89	4.24	3.40	0.73	0.0052 **
/	1018	HexNAc3Hex6Sia1Fuc1		1.04	0.82	0.86	0.91	0.12	0.96	0.98	1.07	1.00	0.06	0.1464
/	1022	HexNAc2Hex10		0.37	0.38	0.38	0.38	0.01	0.39	0.37	0.39	0.38	0.01	0.2191
/	1039	HexNAc4Hex5Sia1Fuc1		3.73	3.32	5.22	4.09	1.00	3.86	3.45	5.57	4.29	1.12	0.4133
/	1039	HexNAc4Hex5Sia1Fuc1		0.19	0.06	0.55	0.27	0.25	0.97	0.62	1.51	1.03	0.45	0.0389 *
/	1067	HexNAc5Hex5Sia1		0.45	0.56	na	0.51	0.08	0.1	0.04	na	0.07	0.04	0.0189 *
/	1111	HexNAc4Hex5Sia2		1.78	2.37	na	2.08	0.42	0.09	0.15	na	0.12	0.04	0.0463 *
/	1111	HexNAc4Hex5Sia2		5.39	5.23	na	5.31	0.11	5.49	7.31	na	6.40	1.29	0.2209
/	1111	HexNAc4Hex5Sia2		1.74	1.92	na	1.83	0.13	2.18	1.4	na	1.79	0.55	0.4677

Table 1 (*continued*)

/	1111	HexNAc4Hex5Sia2		0.18	0.2	na	0.19	0.01	1.41	0.98	na	1.20	0.30	0.0668
/	1140	HexNAc5Hex5Sia1Fuc1		2.01	2.54	na	2.28	0.37	0.33	0.24	na	0.29	0.06	0.0385 *
/	1184	HexNAc4Hex5Sia2Fuc1		9.37	9	7.91	8.76	0.76	7.41	5.95	7.95	7.10	1.03	0.0475 *
/	1184	HexNAc4Hex5Sia2Fuc1		2.52	2.81	2.25	2.53	0.28	7.47	4.87	7.03	6.46	1.39	0.0175 *
/	1184	HexNAc4Hex5Sia2Fuc1		0.38	0.54	0.23	0.38	0.16	6.22	4.33	5.88	5.48	1.01	0.0057 **
/	1213	HexNAc5Hex5Sia1Fuc2		1.27	1.06	0.85	1.06	0.21	0.12	0.1	0	0.07	0.06	0.0048 **
/	1213	HexNAc5Hex5Sia2		1.08	1.13	0.98	1.06	0.08	0.17	0.18	0	0.12	0.10	0.0002 ***
/	1221	HexNAc5Hex6Sia1Fuc1		0.96	1.02	na	0.99	0.04	0.99	0.75	na	0.87	0.17	0.2474
/	1286	HexNAc5Hex5Sia2Fuc1		na	na	2.76	2.76	/	0	0	0	0.00	0.00	n/a
			Total:	100	100	100			100	100	100			

Structures are represented by the mass to charge ratio (m/z) in which they were identified and quantified, by monosaccharide composition and proposed structure based on MS/MS analyses. The relative quantities were determined by base-peak intensity of extracted ion chromatograms. The average value (Avg) and standard deviation (SD) of triplicates are shown, as well as the p-value (t-test; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$). Increased or decreased relative abundance are shown in red or blue, respectively. Structures marked as not analyzed (na) were either not detected or overlapped with other structures in given sample, precluding their quantification. Unknown linkage is represented by "?".

(recombinant sialidase from *Streptococcus pneumoniae*, expressed in *Escherichia coli*) and sialidase A (5 mU, ProZyme) that releases α2-3/6/8 linked non-reducing terminal sialic acid (recombinant gene from *Arthrobacter ureafaciens*, expressed in *Escherichia coli*) to confirm sialic acid linkage.

All analyzes were performed in three independent replicates and results were subjected to statistical analyses (Average, standard deviation and unpaired t-test) Table 1.

4. *N*-glycomic strategy II: sample preparation and PGC nanoLC-ESI MS/MS

Frozen cell pellets (10^7 cells) of mock or *ST3GAL4* transfected MKN45 cells were directly resuspended in 2 mL of lysis buffer (50 mM Tris–HCl, 100 mM NaCl, 1 mM EDTA and protease inhibitor at pH 7.4) and stored on ice for 20 min. The cells were lysed using a Polytron homogenizer for at least three times for 10 s in a cold room. Cellular debris and unlysed cells were sedimented by centrifugation at 2000*g* for 20 min at 4 °C. The supernatant was collected and the pellets resuspended in 1 mL of lysis buffer and centrifuged again at 2000*g* for 20 min at 4 °C. All the supernatants were combined, diluted in 20 mM Tris–HCl (pH 7.4) and ultracentrifuged at 120,000*g* for 90 min at 4 °C. The supernatant was separated from the pellet containing the cell membrane proteins. The membrane proteins were resuspended with 150 μL of 100 mM ammonium bicarbonate buffer and lyophilized overnight. The dried samples were solubilized in 50 μL of 8 M urea and 10 μL aliquots were dot-blotted onto PVDF membranes as described previously [4]. *N*- and *O*-glycan release as well as

Table 2
Identified *N*-glycan structures II.

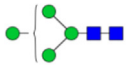
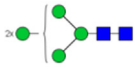
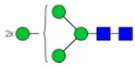
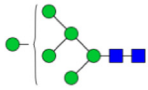
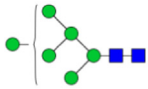
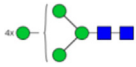
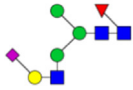
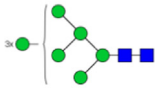
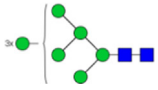
[M-H]-	z	Composition	Proposed structure	mock	ST3GAL4	ST3GAL4/mock
1073.44	1	HexNAc2Hex4		0.26	0.23	0.91
1235.48	1	HexNAc2Hex5		0.35	0.31	0.87
1235.5	1	HexNAc2Hex5		0.47	0.59	1.25
1235.54	1	HexNAc2Hex5		1.21	1.30	1.08
1397.52	1	HexNAc2Hex6		0.37	0.31	0.83
1397.56	1	HexNAc2Hex6		4.06	2.67	0.66
1559.64	2	HexNAc2Hex7		2.18	1.72	0.79
1559.69	2	HexNAc2Hex7		4.66	3.36	0.72
1713.82	2	HexNAc3Hex4Sia1Fuc1		1.61	1.52	0.94
1721.69	2	HexNAc2Hex8		0.06	0.11	1.84
1721.72	2	HexNAc2Hex8		0.21	0.16	0.80

Table 2 (*continued*)

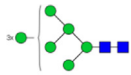
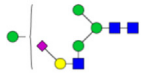
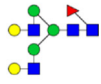
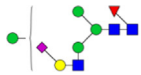
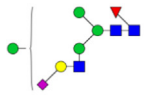
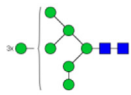
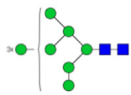
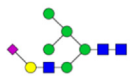
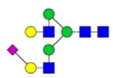
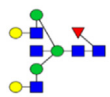
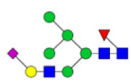
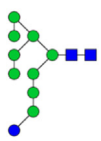
1721.78	2	HexNAc2Hex8		14.36	12.98	0.90
1729.76	2	HexNAc3Hex5Sia1		0.43	0.41	0.95
1787.79	2	HexNAc4Hex5Fuc1		0.87	0.69	0.80
1875.78	2	HexNAc3Hex5Sia1Fuc1		2.99	2.51	0.84
1875.89	2	HexNAc3Hex5Sia1Fuc1		0.00	0.63	/
1883.74	2	HexNAc2Hex9		11.94	11.97	1.00
1883.75	2	HexNAc2Hex9		2.40	2.19	0.91
1891.78	2	HexNAc3Hex6Sia1		2.26	2.11	0.93
1932.80	2	HexNAc4Hex5Sia1		0.97	0.65	0.67
1990.84	2	HexNAc5Hex5Fuc1		0.85	0.00	0.00
2037.83	2	HexNAc3Hex6Sia1Fuc1		1.23	0.90	0.73
2045.73	2	HexNAc1Hex10		2.05	2.16	1.05
2053.70	2	HexNAc3Hex7Sia1Fuc1		0.49	0.28	0.57

Table 2 (*continued*)

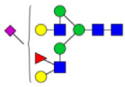
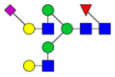
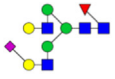
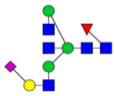
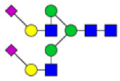
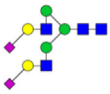
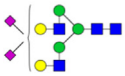
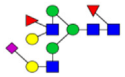
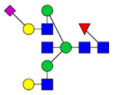
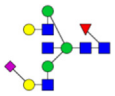
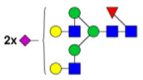
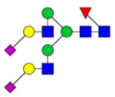
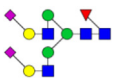
2078.80	2	HexNAc4Hex5Sia1Fuc1		0.91	1.20	1.32
2078.85	2	HexNAc4Hex5Sia1Fuc1		1.48	0.58	0.39
2078.90	2	HexNAc4Hex5Sia1Fuc1		11.24	7.82	0.70
2119.83	2	HexNAc5Hex4Sia1Fuc1		1.07	0.00	0.00
2223.85	2	HexNAc4Hex5Sia2		1.38	1.62	1.17
2223.87	2	HexNAc4Hex5Sia2		0.00	0.45	/
2223.89	2	HexNAc4Hex5Sia2		0.34	0.72	2.12
2224.92	2	HexNAc4Hex5Sia1Fuc2		5.64	3.56	0.63
2281.82	2	HexNAc5Hex5Sia1Fuc1		0.28	0.00	0.00
2281.89	2	HexNAc5Hex5Sia1Fuc1		1.97	0.32	0.16
2369.88	2	HexNAc4Hex5Sia2Fuc1		1.31	6.45	4.93
2369.91	2	HexNAc4Hex5Sia2Fuc1		0.00	5.26	/
2369.97	2	HexNAc4Hex5Sia2Fuc1		8.23	8.53	1.04

Table 2 (*continued*)

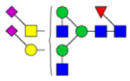
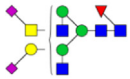
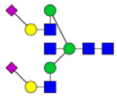
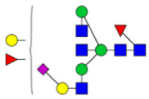
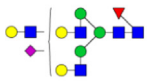
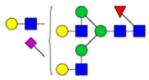
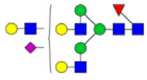
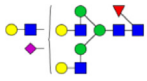
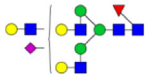
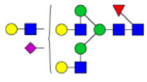
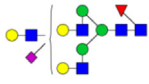
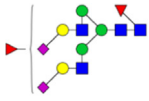
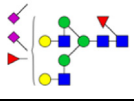
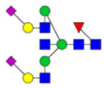
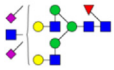
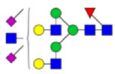
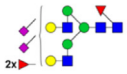
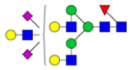
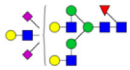
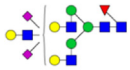
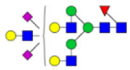
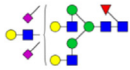
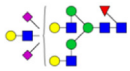
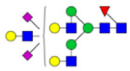
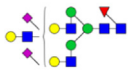
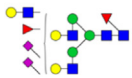
2410.85	2	HexNAc5Hex4Sia2Fuc1		0.16	0.14	0.90
2410.91	2	HexNAc5Hex4Sia2Fuc1		0.00	0.49	/
2426.89	2	HexNAc5Hex5Sia2		0.35	0.00	0.00
2427.87	2	HexNAc5Hex5Sia1Fuc2		1.29	0.10	0.07
2443.86	2	HexNAc5Hex6Sia1Fuc1		0.17	0.35	2.05
2443.87	2	HexNAc5Hex6Sia1Fuc1		0.00	0.10	/
2443.89	2	HexNAc5Hex6Sia1Fuc1		0.41	0.15	0.38
2443.90	2	HexNAc5Hex6Sia1Fuc1		0.54	0.46	0.84
2443.91	2	HexNAc5Hex6Sia1Fuc1		1.21	1.58	1.31
2443.95	2	HexNAc5Hex6Sia1Fuc1		1.12	0.00	0.00
2444.17	2	HexNAc5Hex6Sia1Fuc1		0.00	0.84	/
2515.91	2	HexNAc4Hex5Sia2Fuc2		0.00	0.55	/
2516.01	2	HexNAc4Hex5Sia2Fuc2		0.00	0.91	/

Table 2 (*continued*)

2572.95	2	HexNAc5Hex5Sia2Fuc1		1.81	0.25	0.14
2572.97	2	HexNAc5Hex5Sia2Fuc1		0.00	1.16	/
2572.97	2	HexNAc5Hex5Sia2Fuc1		0.00	1.37	/
2662.01	2	HexNAc4Hex5Sia2Fuc3		0.00	0.33	/
2734.91	2	HexNAc5Hex6Sia2Fuc1		0.72	0.72	0.99
2734.96	2	HexNAc5Hex6Sia2Fuc1		0.33	0.40	1.21
2734.96	2	HexNAc5Hex6Sia2Fuc1		0.38	0.67	1.74
2734.99	2	HexNAc5Hex6Sia2Fuc1		0.44	0.18	0.40
2735.00	2	HexNAc5Hex6Sia2Fuc1		0.00	0.45	/
2735.03	2	HexNAc5Hex6Sia2Fuc1		0.51	0.40	0.79
2735.04	2	HexNAc5Hex6Sia2Fuc1		0.00	0.60	/
2735.08	2	HexNAc5Hex6Sia2Fuc1		0.31	0.00	0.00
2880.98	2	HexNAc5Hex6Sia2Fuc2		0.00	0.48	/
3025.85	2	HexNAc5Hex6Sia3Fuc1		0.00	0.78	/
3026.03	2	HexNAc5Hex6Sia3Fuc1		0.14	0.27	1.97

Structures are represented by their [M–H] value and charge state in which they were identified and quantified, by monosaccharide composition and proposed structure based on MS/MS analyzes. The relative quantities were determined by base-peak intensity of extracted ion chromatograms. The ratio of glycan abundance in ST3GAL4 transfected cells relative to mock transfected is shown on the right column. Increases or decreases larger than 2 fold in glycan abundance are highlighted red or blue, respectively.

PGC-nanoLC ESI MS/MS analysis on an amaZon ETD Speed ion trap (Bruker, Bremen, Germany) were performed as described in detail previously [4,5] (Table 2).

5. *O*-glycomic strategy I: sample preparation and LC-ESI-MS/MS

After the removal of *N*-glycan, as previously described in "1. *N*-glycomic strategy I: Sample preparation and LC-ESI-MS/MS", the *O*-linked glycans were released from retained glycoproteins in spinfilter using reductive β-elimination (0.5 M $NaBH_4$, 50 mM NaOH at 50 °C, 16 h). Reactions were quenched with 1 μl of glacial acetic acid and glycan samples were desalted and dried as previously described [3]. Glycans were subjected to LC-ESI-MS/MS analysis using a 10 cmx250 μm I.D. column, prepared in-house, containing 5 μm porous graphitized carbon (PGC) particles (Thermo Scientific. Waltham. MA). Glycans were eluted using a linear gradient from 0% to 40% acetonitrile in 10 mM NH_4HCO_3 over 40 min at a flow rate of 10 μl/min. The eluted *O*-glycans were detected using a LTQ ion trap mass spectrometer (Thermo Scientific) in negative-ion mode with an electrospray voltage of 3.5 kV, capillary voltage of −33.0 V and capillary temperature of 300 °C. Air was used as a sheath gas and mass ranges were defined dependent on the specific structure to be analyzed. The data were processed using the Xcalibur software (version 2.0.7. Thermo Scientific) and manually interpreted from their MS/MS spectra (Table 3).

6. *O*-glycomic strategy II: sample preparation and PGC nanoLC-ESI MS/MS

Sample preparation and analysis were performed as previously described in Section 2. "*N*-glycomic strategy II: Sample preparation and PGC nanoLC-ESI MS/MS" (Table 4).

7. Sialoproteomic analysis

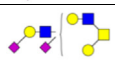
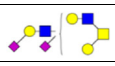
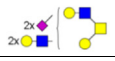
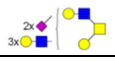
7.1. Cell lysis, protein digestion and iTRAQ labeling

Cell pellets were redissolved in ice-cold Na_2CO_3 buffer (0.1 M, pH 11) supplemented with protease inhibitor (Roche complete EDTA free), PhosSTOP phosphatase inhibitor cocktail (Roche) and 10 mM sodium pervanadate on ice. The suspensions were tip probe sonicated for 20 s (amplitude = 50%) twice and incubated at 4 °C for 1 h. The lysates were then centrifuged at 100,000 × *g* for 90 min at 4 °C to enrich membrane proteins (pellet). The pellets were washed with 50 mM triethylammonium

Table 3
Identified *O*-glycan structures I.

m/z		Composition	Proposed structure	Mock replicates (n=3)					ST3GAL4 replicates (n=3)					p-value
z=1	z=2			%	%	%	Avg	SD	%	%	%	Avg	SD	
384	/	HexNAc1Hex1		3.52	0.91	0.95	1.79	1.50	2.65	1.65	1.8	2.03	0.54	0.407
513	/	HexNAc1Sia1		1.52	0.83	0.93	1.09	0.37	0.74	1.39	0.73	0.95	0.38	0.336
587	/	HexNAc2Hex1		1.62	0	0	0.54	0.94	1.34	0	0	0.45	0.77	0.450
587	/	HexNAc2Hex1		0.94	0	0	0.31	0.54	0.6	0	0	0.20	0.35	0.390
675	/	HexNAc1Hex1Sia1		4.79	2.22	1.79	2.93	1.62	3.77	2.13	2.41	2.77	0.88	0.444
675	/	HexNAc1Hex1Sia1		6.13	8.1	5.79	6.67	1.25	9.5	8.99	9.9	9.46	0.46	0.024 *

Table 3 (*continued*)

749	/	HexNAc2Hex2		3.59	0.84	0.6	1.68	1.66	2.9	0.38	0.84	1.37	1.34	0.409
749	/	HexNAc2Hex2		1.65	0.88	0.34	0.96	0.66	0.65	0.5	0.57	0.57	0.08	0.210
878	/	HexNAc2Hex1Sia1		6.38	3.46	2.16	4.00	2.16	6.05	4.44	3.43	4.64	1.32	0.344
895	/	HexNAc2Hex2Fuc1		0.74	0	0	0.25	0.43	0.35	0	0	0.12	0.20	0.334
966	/	HexNAc1Hex1Sia2		3.77	11.25	9.05	8.02	3.84	6.32	9.04	9.39	8.25	1.68	0.466
1040	/	HexNAc2Hex2Sia1		8.09	2.82	1.2	4.04	3.60	9.11	3.04	4.93	5.69	3.11	0.290
1040	/	HexNAc2Hex2Sia1		11.86	3.85	2.05	5.92	5.22	11.98	5.77	6.99	8.25	3.29	0.278
1114	/	HexNAc3Hex3		1.77	0	0	0.59	1.02	0.27	0	0	0.09	0.16	0.244
1114	/	HexNAc3Hex3		0.87	0	0	0.29	0.50	0.15	0	0	0.05	0.09	0.248
1186	/	HexNAc2Hex2Sia1Fuc1		0.27	0	0	0.09	0.16	0.37	0	0	0.12	0.21	0.419
1186	/	HexNAc2Hex2Sia1Fuc1		2.26	0	0	0.75	1.30	1.14	0	0	0.38	0.66	0.344
1331	665	HexNAc2Hex2Sia2		22.75	15.58	15.81	18.05	4.07	41.62	41.5	42.3	41.81	0.43	0.005 **
1477	738	HexNAc2Hex2Sia2Fuc1		0	0	0	0.00	0.00	0.48	0	0	0.16	0.28	0.211
1696	848	HexNAc3Hex3Sia2		0	8.78	11.34	6.71	5.95	0	3.99	4.64	2.88	2.51	0.061
1696	848	HexNAc3Hex3Sia2		0	7.09	10.78	5.96	5.48	0	8.65	6.74	5.13	4.54	0.314
/	1030	HexNAc4Hex4Sia2		9.38	18.42	22.07	16.62	6.53	0	8.51	5.34	4.62	4.30	0.033 *
/	1213	HexNAc5Hex5Sia2		8.09	14.96	15.15	12.73	4.02	0	0	0	0.00	0.00	0.016 *
			Total:	100	100	100			100	100	100			

Structures are represented by the mass to charge ratio (m/z) in which they were identified and quantified, by monosaccharide composition and proposed structure based on MS/MS analyzes. The relative quantities were determined by base-peak intensity of extracted ion chromatograms. The average value (Avg) and standard deviation (SD) of triplicates are shown, as well as the p-value (t-test; $^{*}p < 0.05$; $^{**}p < 0.01$). Increased or decreased relative abundance are shown in red or blue, respectively. Unknown linkage is represented by "?".

bicarbonate (TEAB) to remove any remaining soluble protein. Membrane fraction was resuspended directly in 6 M urea and 2 M thiourea, reduced in 10 mM DTT for 30 min and then alkylated in 20 mM IAA for 30 min at room temperature in the dark.

Samples were incubated with endoproteinase Lys-C (Wako, Osaka, Japan) for 2 h (1:100 w/w). Following the incubation, the samples were diluted 8 times with 50 mM TEAB (pH 8) and trypsin was added at a ratio of 1:50 (w/w) and left overnight at room temperature. Trypsin digestion was stopped by the addition of 2% formic acid and then the samples were centrifuged at 14,000 × g for 10 min to

precipitate any lipids present in the sample. The supernatant was purified using in-house packed staged tips with a mixture of Poros R2 and Oligo R3 reversed phase resins (Applied Biosystem, Foster City, CA, USA). Briefly, a small plug of C18 material (3 M Empore) was inserted in the end of a P200 tips, followed by packing of the stage tip with the resins (resuspended in 100% ACN) by applying gentle air pressure. The acidified samples were loaded onto the micro-column after equilibration of the column with 0.1% trifluoroacetic acid (TFA), washed twice with 0.1% TFA and peptides were eluted with 60% ACN/0.1% TFA. A small amount of purified peptides (1 μl) from each sample was subjected to Qubit assay to determine the concentration, while the remaining samples were dried by vacuum centrifugation. Afterwards, peptides were redissolved in dissolution buffer and a total of 150 μg for each condition was labeled with 4-plex iTRAQ™ (Applied Biosystems, Foster City, CA) as described by the manufacturer. After labeling, the samples were mixed 1:1:1:1 and lyophilized by vacuum centrifugation.

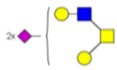
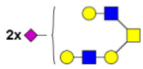
7.2. *Sialic acid containing glycopeptide enrichment by TiSH protocol*

The method used for sialylated glycopeptides enrichment is a modification of the TiSH protocol [6] described in [7,8]. Briefly, samples were resuspended in loading buffer (1 M glycolic acid, 80% ACN, 5%

Table 4
Identified *O*-glycan structures II.

[M-H]-	z	Composition	Type	Assigned structures	Mock	ST3GAL4	ST3GAL4/mock
675.30	1	HexNAc1Hex1Sia1	core 1		0.00	1.48	/
675.31	1	HexNAc1Hex1Sia1	core 1		0.00	1.17	/
749.40	1	HexNAc2Hex2	core 2		0.00	0.00	/
895.39	1	HexNAc2Hex2Fuc1	core 1 / 3	/	0.00	0.36	/
966.40	1	HexNAc1Hex1Sia2	core 1		0.00	0.09	/
1040.43	1	HexNAc2Hex2Sia1	core 1	/	18.81	17.40	0.93
1040.43	1	HexNAc2Hex2Sia1	core 2		2.33	1.44	0.62
1040.44	1	HexNAc2Hex2Sia1	core 2		6.65	7.91	1.19
1186.49	1	HexNAc2Hex2Sia1Fuc1	core 2 / 4	/	3.05	4.29	1.40
1186.52	1	HexNAc2Hex2Sia1Fuc1	core 2 / 4	/	0.51	0.15	0.30
1227.53	1	HexNAc3Hex1Sia1Fuc1	core 2 / 4	/	0.00	2.47	/
1331.49	1	HexNAc2Hex2Sia2	core 2		0.00	1.32	/
1331.52	1	HexNAc2Hex2Sia2	core 2	/	0.99	0.20	0.20

Table 4 (*continued*)

1331.57	1	HexNAc2Hex2Sia2	core 2		2.90	1.01	0.35
1389.56	1	HexNAc3Hex2Sia1Fuc1	core 2 / 4	/	26.90	45.91	1.71
1389.58	1	HexNAc3Hex2Sia1Fuc1	core 2 / 4	/	0.00	0.00	/
1405.55	1	HexNAc3Hex3Sia1	core 2 / 4	/	0.00	0.10	/
1405.55	1	HexNAc3Hex3Sia1	core 2 / 4	/	0.00	1.10	/
1405.58	1	HexNAc3Hex3Sia1	core 2 / 4	/	0.00	0.00	/
1405.59	1	HexNAc3Hex3Sia1	core 2 / 4	/	1.84	0.31	0.17
1477.56	1	HexNAc2Hex2Sia2Fuc1	core 2 / 4	/	3.01	0.57	0.19
1477.58	1	HexNAc2Hex2Sia2Fuc1	core 2 / 4	/	1.28	0.00	0.00
1477.60	1	HexNAc2Hex2Sia2Fuc1	core 2 / 4	/	1.49	0.17	0.11
1696.62	2	HexNAc3Hex3Sia2	core 2 / 4	/	0.00	2.17	/
1696.63	2	HexNAc3Hex3Sia2	core 2		1.81	0.00	0.00
1696.66	2	HexNAc3Hex3Sia2	core 2 / 4	/	0.60	0.00	0.00
2061.76	2	HexNAc4Hex4Sia2	core 2 / 4	/	1.46	0.00	0.00
2061.79	2	HexNAc4Hex4Sia2	core 2 / 4	/	1.61	0.44	0.27
2061.80	2	HexNAc4Hex4Sia2	core 2 / 4	/	10.19	4.86	0.48
				Total:	100	100	

Structures are represented by their [M–H] value and charge state in which they were identified and quantified, by monosaccharide composition, type of core and proposed structure based on MS/MS analyzes. The relative quantities were determined by base-peak intensity of extracted ion chromatograms. The ratio of glycan abundance in *ST3GAL4* transfected cells relative to mock transfected is shown on the right column. Increases or decreases greater than 1.5 fold in glycan abundance are highlighted red or blue, respectively.

TFA) and incubated with TiO_2 beads (GL Sciences, Japan, 10 μm; using a total of 0.6 mg TiO_2 beads per 100 μg of peptides). The supernatant containing the un-modified peptides was carefully separated. The TiO_2 beads were sequentially washed with loading buffer, washing buffer 1 (80% ACN, 1% TFA) and washing buffer 2 (20% ACN, 0.1% TFA), saving the washings with the previous supernatant. The bound peptides were eluted with 1.5% ammonium hydroxide by shaking for 15 min. The eluted fraction containing the phosphopeptides and sialylated glycopeptides was dried by vacuum centrifugation and subjected to an enzymatic deglycosylation in 20 mM TEAB buffer using 500 U of PNGase F (New England Biolabs, Ipswich, MA) and 0.1 U of Sialidase A (Prozyme, Hayward, CA) overnight at 37 °C.

Table 5
Proteins with increased *N*-glycan sialylation.

Accession number	Protein name	Sequence	*N*-glycan site	ST3GAL4 /Mock	Limma&Rank (*p*-value)	Func. Grp.
O15031	Plexin-B2	ISVAGRNdeCSFQPER	N844	2.89	0.0342	RS
		ALSNdeISLR	N127	2.08	0.0358	
O60637	Tetraspanin-3	TYNdeGTNdePDAASR	N127	2.96	0.0358	RS
		TYNdeGTNPDAASR	N127	2.98	0.0433	
O75882	Attractin	GICNdeSSDVR	N300	2.35	0.0342	IR
P02786	Transferrin receptor protein 1	KDFEDLYTPVNdeGSIVIVR	N251	1.72	0.0433	TT
P05026	Sodium/potassium-transporting ATPase subunit beta-1	FKLEWLGNdeCSGLNDETYGYK	N158	2.79	0.0479	TT
P06213	Insulin receptor	HNdeLTITQGK	N445	2.45	0.0479	RS
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	TLTLFNdeVTR	N204	2.67	0.0342	AM
		TLTLFNVTRNdeDTASYK	N208	3.11	0.0342	
P06756	Integrin alpha-V	ISSLQTTEKNdeDTVAGQGER	N874	4.02	0.0342	AM
		NdeMTISR	N554	4.02	0.0342	
P07602	Prosaposin	TNdeSTFVQALVEHVKEECDR	N215	2.59	0.0479	O
P08962	CD63 antigen	CCGAANdeYTDWEK	N150	1.94	0.0394	RS
P10909	Clusterin	EIRHNdeSTGCLR	N291	2.12	0.0342	O
P11117	Lysosomal acid phosphatase	QTPEYQNdeESSR	N177	3.05	0.0383	E
P12821	Angiotensin-converting enzyme	IGLLDRVTNdeDTESDINYLLK	N445	2.62	0.0456	P
P13473	Lysosome-associated membrane glycoprotein 2	LNdeSSTIK	N275	3.09	0.0342	AM
		VASVINdeINPNdeTTHSTGSCR	N253	1.96	0.0358	
		VASVININPNdeTTHSTGSCR	N257	2.06	0.0456	
P13688	Carcinoembryonic antigen-related cell adhesion molecule 1	NdeQSLPSSER	N363	2.76	0.0479	AM
P13726	Tissue factor	RNdeNTFLSLR	N169	5.52	0.0342	P
		NdeNTFLSLR	N169	5.76	0.0358	
P18564	Integrin beta-6	EVEVNdeSSK	N463	3.71	0.0479	AM
P21589	5'-nucleotidase	LDNdeYSTQELGK	N333	1.67	0.0479	E
P26006	Integrin alpha-3	ELAVPDGYTNdeRTGAVYLCPLTAHK	N86	2.68	0.0433	AM
P30825	High affinity cationic amino acid transporter 1	LCLNNdeDTK	N235	3.48	0.0297	TT
		LCLNdeNdeDTK	N234	3.01	0.0342	
P35613	Basigin	ALMNdeGSESR	N268	2.91	0.0342	P
P42892	Endothelin-converting enzyme 1	NdeSSVEAFK	N632	1.89	0.0394	P
		DYYLNdeKTENEK	N270	5.64	0.0479	
P43007	Neutral amino acid transporter A	VVTQNdeSSSGNdeVTHEK	N201	2.58	0.0477	TT
P46059	Solute carrier family 15 member 1	NdeDSCPEVK	N562	4.66	0.0358	TT
P48960	CD97 antigen	WCPQNSSCVNdeATACR	N38	1.87	0.0479	AM
P54760	Ephrin type-B receptor 4	CAQLTVNdeLTRFPETVPR	N203	2.34	0.0342	RS
Q04912	Macrophage-stimulating protein receptor	LPEYVVRDPQGWVAGNdeLSAR	N841	3.26	0.0358	RS

Table 5 (*continued*)

Accession number	Protein name	Sequence	*N*-glycan site	ST3GAL4 /Mock	Limma&Rank (*p*-value)	Func. Grp.
Q08380	Galectin-3-binding protein	DAGVVCTNdeETR	N125	2.43	0.0358	AM
Q08722	Leukocyte surface antigen CD47	SDAVSHTGNdeYTCEVTELTREGETIIELK	N111	6.51	0.0056	AM
Q11206	ST3GAL4	LFGNdeYSR	N61	5.77	0.0088	O
Q12913	Receptor-type tyrosine-protein phosphatase eta	HGSNdeHTSTYDK	N582	3.15	0.0342	RS
		VSDNdeESSSNdeYTYK	N391	2.94	0.0342	
		VSDNESSSNdeYTYK	N396	2.74	0.0383	
		GPNdeGTEGASR	N525	2.76	0.0427	
		IHVAGETDSSNdeLNdeVSEPR	N411	1.69	0.0479	
		SNdeDTAASEYK	N142	2.77	0.0479	
Q13641	Trophoblast glycoprotein	CVNRNdeLTEVPTDLPAYVR	N81	1.85	0.0429	RS
Q14108	Lysosome membrane protein 2	ANIQFGDNdeGTTISAVSNdeK	N105	2.03	0.0394	T
		ANdeIQFGDNdeGTTISAVSNK	N99	1.78	0.0479	
Q15043	Zinc transporter ZIP14	ALLNHLDVGVGRGNdeVTQHVQGHR	N77	2.42	0.0479	TT
Q15758	Neutral amino acid transporter B(0)	NdeITGTR	N212	2.93	0.0479	TT
Q7Z7H5	Transmembrane emp24 domain-containing protein 4	QYGSEGRFTFTSHTPGDHQICLHSNdeSTR	N117	2.50	0.0479	T
Q8N271	Prominin-2	ILRNdeVSECFLAR	N725	3.11	0.0433	O
Q92542	Nicastrin	RPNdeQSQPLPPSSLQR	N417	1.66	0.0483	P
Q92673	Sortilin-related receptor	LTIVNdeSSVLDRPR	N871	1.92	0.0479	T
		SRNdeSTVEYTLNK	N1986	1.67	0.0479	
Q9BXB1	Leucine-rich repeat-containing G-protein coupled receptor 4	TLDLSYNNIRDLPSFNdeGCHALEEISLQR	N362	4.00	0.0358	RS
Q9BXS4	Transmembrane protein 59	LFSICQFVDDGIDLNdeRTK	N90	4.09	0.0342	T
Q9H330	Transmembrane protein 245	ILGDKVNdeNTAVIEK	N551	3.47	0.0433	O
Q9H5V8	CUB domain-containing protein 1	ASVSFLNFNdeLSNCERK	N270	3.33	0.0342	AM
		IGTFCSNdeGTVSR	N180	2.93	0.0358	
		NdeVSGFSIANR	N205	2.33	0.0358	
Q9HD43	Receptor-type tyrosine-protein phosphatase H	NdeATTAHNPVR	N203	6.86	0.0342	RS
		ETRNdeATTAPNPVR	N381	2.93	0.0479	
		NdeTTNTSVTAER	N434	2.72	0.0479	
Q9P2B2	Prostaglandin F2 receptor negative regulator	QRNdeNSWVK	N525	3.60	0.0429	RS
Q9UN76	Sodium- and chloride-dependent neutral and basic amino acid transporter B(0+)	SPIVTHCNdeVSTVNdeK	N174	2.44	0.0420	TT
Q9Y639	Neuroplastin	ANdeATIEVK	N229	2.59	0.0429	AM

List of significantly increased sialylated *N*-glycan modified peptides in the ST3GAL4 overexpressing cells compared to mock control shown with accession number, protein name, peptide sequence and the identified *N*-glycan site. The fold in increase, p-value and the protein's functional group are also presented. AM: Adhesion and migration; E: Metabolic enzyme; RS: Receptor and signaling; T: Trafficking; TT: Transmembrane transporter; P: Protease; O: Others.

Table 6
Proteins with decreased *N*-glycan sialylation.

Accession number	Protein name	Sequence	N-glycan site	ST3GalIV/Mock	Limma & rank (*p*-value)
O14672	Disintegrin and metalloproteinase domain-containing protein 10	NdeISQVLEK	N439	0.57	0.0149
P07602	Proactivator polypeptide	NdeSTKQEILAALEK	N426	0.57	0.0474

List of significantly decreased sialylated *N*-glycan modified peptides in the ST3GAL4 overexpressing cells compared to mock control shown with accession number, protein name, peptide sequence and the identified *N*-glycan site, fold in increase and the *p*-value.

To separate phosphorylated peptides and formerly glycosylated peptides, the samples were subjected to a second TiO_2 enrichment procedure to separate phosphorylated from deglycosylated peptides. The supernatant containing the deglycosylated peptides was saved and the beads were washed with 50% ACN, 0.1% TFA. The washing was added to the supernatant. The deglycosylated fraction was desalted on Oligo R3 staged tip column and dried prior to the HILIC fractionation [7]. All fractions were dried by vacuum centrifugation prior nLC-MS/MS analysis.

7.3. Sialic acid containing glycopeptide analysis by nLC-MS/MS

Samples were resuspended in 6 μL of 0.1% TFA for analysis. Peptides were loaded on an in-house packed Reprosil-Pur C18-AQ (2 cmx100 μm, 5 μm; Dr. Maisch GmbH, Germany) pre-column and separated on an in-house packed Reprosil-Pur C18-AQ (17 cmx75 μm, 3 μm; Dr. Maisch GmbH, Germany) column using an Easy-nLC II system (Thermo Scientific, Bremen, Germany) and eluted at a flow of 250 nL/min. Mobile phase was 95% acetonitrile (B) and water (A) both containing 0.1% formic acid. Depending on the samples, gradient was from 1% to 30% solvent B in 80 or 110 min, 30–50% B in 10 min, 50–100% B in 5 min and 8 min at 100% B. Mass spectrometric analyses were performed in an Orbitrap Fusion Tribrid system (Thermo Scientific, Bremen, Germany). MS scans (400–1200 m/z) were acquired in the orbitrap at a resolution of 120000 at 200 m/z for a AGC target of 5×10^5 ions and a maximum injection time of 60 ms. Data-dependent HCD MS/MS analysis at top speed of the most intense ions were performed at a resolution of 30000 at 200 m/z for a AGC target of 5×10^4 and a maximum injection time of 150 ms using the quadrupole to isolate the ions and an isolation window of 1.2 m/z, a NCE of 38% and a dynamic exclusion of 20 s.

The raw data were processed and quantified by Proteome Discoverer (version 1.4.1.14, Thermo Scientific) against SwissProt and Uniprot human reference databases by using Mascot (v2.3.02, Matrix Science Ltd, London, UK) and Sequest HT, respectively. Database searches were performed using the following parameters: precursor mass tolerance of 10 ppm, product ion mass tolerance of 0.02 Da, 1 missed cleavages for trypsin, carbamidomethylation of Cys and iTRAQ labeling on protein *N*-terminal and Lys as fixed modifications, and phosphorylation on S/T/Y and deamidation of Asn as dynamic modifications. The iTRAQ datasets were quantified using the centroid peak intensity with the "reporter ions quantifier" node. Only peptides with up to a *q*-value of 0.01 (Percolator), Mascot and Sequest HT rank 1, Sequest HT ΔCn of 0.1, cut-off value of Mascot score $\geq$ 18 and a cut-off value of XCorr score for charge states of +1, +2, +3, and +4 higher than 1.5, 2, 2.25 and 2.5, respectively, were considered for further analysis.

7.4. Data normalization and significance analysis

Three biological replicates were analyzed and submitted to the statistical analysis. The log2 values of the measured intensities were normalized by the median. Modified peptides were merged with the R Rollup function (http://www.omics.pnl.gov) allowing for one-hit-wonders and using the mean of the normalized intensities for each peptide. Quantification of proteins was obtained by merging the

un-modified peptides with the R Rollup function considering at least 2 unique peptides not allowing for one-hit-wonders and using the mean of the intensities. Then the mean over the experimental conditions for each peptide in each replicate was subtracted in order to merge data from different iTRAQ runs. Formerly sialylated glycopeptides containing the consensus motif for *N*-linked glycosylation (NXS/T/C; where X # P) were normalized based on the protein expression in each of the replicates. Significant up/down-regulations between experimental conditions were calculated allowing a false discovery rate of 0.05. Therefore, we applied combined limma and rank product tests [9], subsequently corrected for multiple testing according to Storey.

Since spontaneous deamidation is frequently observed for asparagine residues, especially when the C-terminal amino acid is glycine (NG), the sites with NGS/T/C are considered as only potential glycosylation. However, in order to reduce the contribution from spontaneous deamidation in the final list, we sort first for the *N*-linked consensus site (NXS/T/C) and then we filter for proteins that are membrane-associated in order to exclude intracellular proteins that are not *N*-linked glycosylated (Tables 5 and 6).

Acknowledgments

We acknowledge the support from the European Union, Seventh Framework Programme, Gastric Glyco Explorer Initial Training Network: Grant number 316929. IPATIMUP integrates the i3S Research Unit, which is partially supported by FCT, the Portuguese Foundation for Science and Technology. This work is funded by FEDER funds through the Operational Programme for Competitiveness Factors-COMPETE (FCOMP-01-0124-FEDER028188) and National Funds through the FCT-Foundation for Science and Technology, under the projects: PEst-C/SAU/LA0003/2013, PTDC/BBB-EBI/0786/2012, PTDC/BBB-EBI/0567/2014 (CR). This work was also supported by "Glycoproteomics" project Grant number PCIG09-GA-2011-293847 (to DK) and the Danish Natural Science Research Council and a generous Grant from the VILLUM Foundation to the VILLUM Center for Bioanalytical Sciences at the University of Southern Denmark (to MRL). AM acknowledges FCT, POPH (Programa Operacional Potencial Humano) and FSE (Fundo Social Europeu) (SFRH/BPD/75871/2011). The UPLC instrument was obtained with a grant from the Ingabritt and Arne Lundbergs Research Foundation. C.J. was supported by the Knut and Alice Wallenberg Foundation. The mass spectrometer (LTQ) was obtained by a grant from the Swedish Research Council (342-2004-4434).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2016.03.022.

References

[1] S. Mereiter, A. Magalhães, B. Adamczyk, C. Jin, A. Almeida, L. Drici, M. Ibáñez-Vea, C. Gomes, J.A. Ferreira, L.P. Afonso, L.L. Santos, M.R. Larsen, D. Kolarich, N.G. Karlsson, C.A. Reis, Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer, Biochim Biophys Acta, (in press).

[2] A.S. Carvalho, A. Harduin-Lepers, A. Magalhaes, E. Machado, N. Mendes, L.T. Costa, R. Matthiesen, R. Almeida, J. Costa, C. A. Reis, Differential expression of alpha-2,3-sialyltransferases and alpha-1,3/4-fucosyltransferases regulates the levels of sialyl lewis a and sialyl lewis x in gastrointestinal carcinoma cells, Int. J. Biochem. Cell Biol. 42 (2010) 80–89.

[3] B.L. Schulz, N.H. Packer, N.G. Karlsson, Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis, Anal. Chem. 74 (2002) 6088–6097.

[4] P.H. Jensen, N.G. Karlsson, D. Kolarich, N.H. Packer, Structural analysis of N-and O-glycans released from glycoproteins, Nat. Protoc. 7 (2012) 1299–1310.

[5] D. Kolarich, M. Windwarder, K. Alagesan, F. Altmann, Isomer-specific analysis of released N-glycans by LC-ESI MS/MS with porous graphitized carbon, Methods Mol. Biol. 1321 (2015) 427–435.

[6] K. Engholm-Keller, P. Birck, J. Storling, F. Pociot, T. Mandrup-Poulsen, M.R. Larsen, TiSH – a robust and sensitive global phosphoproteomics strategy employing a combination of TiO_2, SIMAC, and HILIC, J. Proteom. 75 (2012) 5749–5761.

[7] M.N. Melo-Braga, M. Ibanez-Vea, M.R. Larsen, K. Kulej, Comprehensive protocol to simultaneously study protein phosphorylation, acetylation, and N-linked sialylated glycosylation, Methods Mol. Biol. 1295 (2015) 275–292.
[8] M.R. Larsen, S.S. Jensen, L.A. Jakobsen, N.H. Heegaard, Exploring the sialiome using titanium dioxide chromatography and mass spectrometry, Mol. Cell. Proteom.: MCP 6 (2007) 1778–1787.
[9] V. Schwammle, I.R. Leon, O.N. Jensen, Assessment and improvement of statistical tools for comparative proteomics analysis of sparse data sets with few experimental replicates, Journal. Proteom. Res. 12 (2013) 3874–3883.