1 SUPPLEMENTAL FIGURES

2	Discovery of O-GlcNAc proteins in published large-scale proteome
3	data
4	
5	
6	Hannes Hahne ¹ , Amin Moghaddas Gholami ¹ , Bernhard Kuster ^{1, 2,*}
7	
8	
9	¹ Chair for Proteomics and Bioanalytics, Center of Life and Food Sciences Weihenstephan,
10	Technische Universität München, Emil-Erlenmeyer-Forum 5, 85354 Freising, Germany
11	² Center for Integrated Protein Science Munich, Emil-Erlenmeyer-Forum 5, 85354 Freising,
12	Germany
13	
14	
15	
16	* Corresponding author
17	Bernhard Kuster
18	e-Mail: kuster@tum.de
19	Tel. +49 8161 715696
20	Fax: +49 8161 715931
21	
22	Running Title: O-GlcNAc protein identification

1 Supplemental Figures

Name											
Title He	axNAc (C	CID and O	linked)								
Fullname N-	acetylg	lucosamin	e by CII)							
Delta Sp	ecificity	Ignore I	Massas	Misc	Reference						
Delta	sentercy	Ignorei	103303	MISC	Kereren	.03					
Monoisotopi	IC	203.079	9373								
Average		203.192	25								
Composition		HexNAc									
		Symbols	13C	*	1 💌	Add					
Delta Sp	ecificity	Ignore N	Masses	Misc	Referenc	es					
Specificity		Ignoro	I I I I I I I I I I I I I I I I I I I	11150	Increase						
Specificity	Site S	6 🔽			Position [Anywhere	~	Сору	Dele	ete H	lide (
Classification				_	Hidden 🗌			Group 0			
Notes		5.,,.								-	
Neutral loss	⊙ Sci	oring Os	atellite	OPer	ptide OF	Required Pe	eptide				
Composition						7		Symbols	13C	▶ 1	~
Neutral loss	-	nrina 🔘s	atellite	OPer	ntide OF	 Required Pe	entide				ſ
						203.1925	pudo				
Composition	HexN	Ac						Symbols	13C	▼ 1	~
Neutral loss	📀 Scr	oring 🔘 s	atellite	OPep	ptide 🔘 F	Required Pe	eptide				
	Monoi	sotopic: 2	21.089	937 A	verage: 2	221.2078					
Composition	HexN	Ac Water						Symbols	13C	▶ 1	~
Neutral loss						Required Pe	eptide				
2			203.079	373 A	verage: 2	203.1925					_
Composition		-		~	-			Symbols	13C	▶ 1	~
Neutral loss						Required Pe	eptide				
Composition			.21.009	937 H	veraye. z	221.2078		Symbols	130	∨ 1	~
New Neutr		_						-,	100		-
Delta Sp	ecificity	Ignore f	Masses	Misc	Referenc	es					
		Ignore M	Masses	Misc	Referenc	es					
	asses	Ignore M				es 649 Avera	ge 204.	1999 D	elete		
Ignore M	asses 1	Ignore f		otopic	204.086		ge 20 4.	1999 D	elete]	
Ignore M Ignore Mass	asses 1	Ignore M	Monois	otopic VAc e(-	204.086 1)	649 Avera		1999 D	elete]	
Ignore M Ignore Mass	asses	Ignore f	Monois H HexN Symbol	otopic VAc e(- s [130	204.086 1) C V	649 Avera	dd]	
Ignore M Ignore Mass Composition	asses	Ignore (Monoise H HexN Symbol Monoise	otopic VAc e(- s 130 otopic	204.086 1) C V	649 Avera 1 💌 🔺 084 Avera	dd]	
Ignore M Ignore Mass Composition Ignore Mass	asses	Ignore f	Monoise H HexN Symbol Monoise	otopic VAc e(- s 130 otopic VAc Wa	204.086 1) 2 V (186.076 ater(-1) e(649 Avera 1 💌 A 084 Avera	dd ge 186.				
Ignore M Ignore Mass Composition Ignore Mass	asses 1	Ignore (Monoise H HexN Symbol Monoise H HexN Symbol	otopic NAc e(- s 130 otopic NAc Wa s 130	204.086 1) 186.076 ater(-1) e(649 Avera 1 💌 A 084 Avera	dd ge 186 . dd	1846 🖸	elete		
Ignore M Ignore Mass Composition Ignore Mass Composition	asses 1 2 2	Ignore (Monoise H HexN Symbol Monoise Symbol Monoise	otopic NAC e(- s 130 otopic NAC Wa s 130 otopic	204.086 1) 186.076 ater(-1) e(649 Avera 1 Y A 084 Avera (-1) 1 Y A 519 Avera	dd ge 186 . dd	1846 🖸	elete]	
Ignore M Ignore Mass Composition Ignore Mass Ignore Mass	asses 1 2 2	Ignore (Monoise H HexN Symbol Monoise Symbol Monoise	otopic NAc e(- s 130 otopic NAc Wa s 130 otopic	204.086 1) 186.076 ater(-1) e(C V (168.065 ater(-2) e(649 Avera 1 A 084 Avera (-1) 1 A 519 Avera (-1)	dd ge 186 . dd	1846 🖸	elete		
Ignore M Ignore Mass Composition Ignore Mass Ignore Mass	asses 1 2 3	Ignore (Monoise H HexN Symbol Monoise Symbol Monoise H HexN Symbol	otopic NAC e(- s 130 otopic NAC Wa s 130 otopic NAC Wa s 130	204.086 1) 186.076 ater(-1) et 168.065 ater(-2) et C V (649 Avera 1 A 084 Avera (-1) 1 A 519 Avera (-1)	dd ge 186. dd ge 168. dd	1846 D	elete		
Ignore M Ignore Mass Composition Ignore Mass Composition Ignore Mass Composition	asses 1 2 3 3 4	Ignore I	Monoisu H HexN Symbol Monoisu H HexN Symbol Monoisu H HexN Symbol	otopic NAC e(- s 130 otopic NAC Wa s 130 otopic NAC Wa s 130 otopic	204.086 1) 186.076 ater(-1) et 168.065 ater(-2) et C V (649 Avera 1 V A 084 Avera (-1) 1 V A 519 Avera (-1) 1 V A 519 Avera	dd ge 186. dd ge 168. dd	1846 D	elete		
Ignore M Ignore Mass Composition Ignore Mass Composition Ignore Mass Composition Ignore Mass	asses 1 2 3 3 4	Ignore I	Monoisu H HexN Symbol Monoisu H HexN Symbol Monoisu H HexN Symbol	otopic NAc e(- s 130 otopic NAc Wa s 130 otopic NAc Wa s 130 otopic (10) N	204.086 1) 186.076 ater(-1) et 168.065 ater(-2) et 144.065 O(3) e(-1)	649 Avera 1 V A 084 Avera (-1) 1 V A 519 Avera (-1) 1 V A 519 Avera	dd ge 186. dd ge 168. dd ge 144.	1846 D	elete		
Ignore M Ignore Mass Composition Ignore Mass Composition Ignore Mass Composition Ignore Mass	asses 1 5 2 5 3 5 4	Ignore I	Monois H HexN Symbol Monois H HexN Symbol Monois Symbol Monois C(6) H(Symbol	otopic NAC e(- s 130 otopic NAC Wa s 130 otopic s 130 otopic (10) N s 130	204.086 1) 186.076 ater(-1) et 168.065 ater(-2) et 144.065 0(3) et-1 2	649 Avera 1 V A 084 Avera (-1) 1 V A 519 Avera (-1) 1 V A 519 Avera	dd ge 186. dd ge 168. dd ge 144. dd	1846 D 1694 D 1480 D	elete elete		
Ignore Mass Composition Ignore Mass Composition Ignore Mass Composition Ignore Mass Composition	asses 1 2 3 3 4 5 5	Ignore I	Monoisu H HexN Symbol Monoisu H HexN Symbol Monoisu C(6) H(Symbol Monoisu	otopic VAc e(- s 130 otopic VAc Wa s 130 otopic s 130 otopic (10) N s 130 otopic	204.086 1) 186.076 ater(-1) et 168.065 ater(-2) et 144.065 0(3) et-1 2	649 Avera 1 × A 084 Avera (-1) 1 × A 519 Avera (-1) 1 × A 519 Avera) 1 × A	dd ge 186. dd ge 168. dd ge 144. dd	1846 D 1694 D 1480 D	elete elete		
Ignore M Ignore Mass Composition Ignore Mass Composition Ignore Mass Composition Ignore Mass	asses 1 2 3 3 4 5 5	Ignore I	Monoise H HexN Symbol Monoise H HexN Symbol Monoise C(6) H(Symbol Monoise C(7) H(otopic s 13C otopic vAc Was s 13C otopic s 13C otopic s 13C otopic (10) N s 13C otopic (8) N C	204.086 1) 186.076 ater(-1) et ater(-2) et 168.065 ater(-2) et 144.065 O(3) e(-1) 138.054 O(2) e(-1)	649 Avera 1 × A 084 Avera (-1) 1 × A 519 Avera (-1) 1 × A 519 Avera) 1 × A	dd ge 186. ge 168. dd ge 168. dd ge 144. dd ge 138.	1846 D 1694 D 1480 D	elete elete		
Ignore Mass Composition Ignore Mass Composition Ignore Mass Composition Ignore Mass Composition Ignore Mass	asses 1 2 3 3 4 5 5	Ignore (Monoisu H HexN Symbol Monoisu H HexN Symbol Monoisu C(6) H(Symbol Monoisu C(7) H(Symbol	otopic NAc e(- S 130 otopic s 130 otopic s 130 otopic (10) N S 130 otopic (8) N C S 130 otopic s 130	204.086 1) 186.076 ater(-1) e(168.065 ater(-2) e(144.065 0(3) e(-1) 138.054 0(2) e(-1) 0)	649 Avera 1 V A 084 Avera (-1) 1 V A 519 Avera (-1) 1 V A 519 Avera) 1 V A 955 Avera	dd ge 186. ge 168. dd ge 168. dd ge 144. ge 138.	1846 D 1694 D 1480 D	elete elete elete		

2

Figure S1. HexNAc modification definition for Mascot. In the standard setting, Mascot provides the proper composition and exact mass for HexNAc. However, the standard modification definition should be extended as depicted in the "Specificity" (B) and in the "Ignore Masses" tab (C). To achieve proper matching of sequence-specific peptide fragments, it is important to set the 1 HexNAc fragment neutral loss as "scoring" as depicted in panel B. These fragment neutral loss 2 peaks are then considered for scoring and significantly improve the matching of sequence-3 specific ions and the Mascot score. Otherwise, these fragment neutral loss peaks would be 4 treated as noise and reduce the Mascot ion score. The "peptide" neutral loss setting matches the 5 neutral loss from the intact peptide, which is then not treated as noise peak for scoring. As 6 depicted in panel C, the HexNAc oxonium ion as well as its fragments should be included in the 7 "Ignore Masses" list. A detailed explanation of the modification definitions can also be found at the Mascot website (www.matrixscience.com/help/pt_mods_help.html). 8

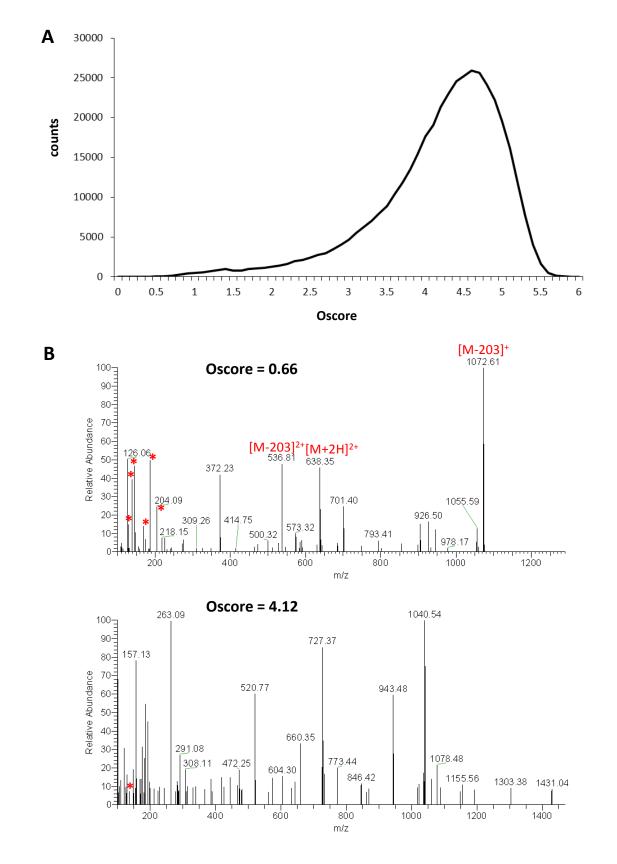


Figure S2. Typical Oscore distribution and representative spectra for low and high Oscores. (A)
The Oscore distribution is based on 414,599 spectra (complete Geiger *et al.* data set), for which
an Oscore could be calculated. (B) The upper spectrum represents a typical O-GlcNAc case with

a low Oscore, where most of the abundant fragment ions point to the O-GlcNAc modification
(indicated in red). The spectrum below represents a typical "high Oscore" case and only contains
one weak fragment ion (m/z 126.05) diagnostic for O-GlcNAc.

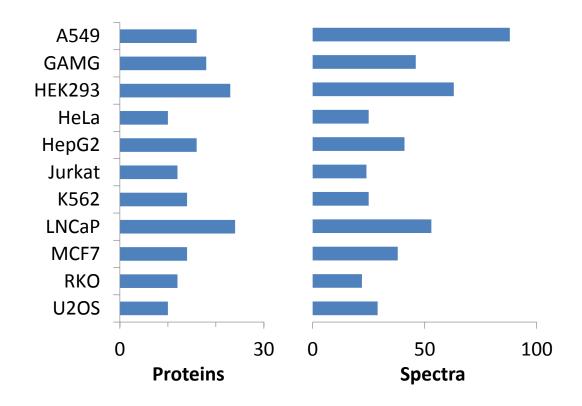




Figure S3. O-GlcNAc protein identifications and O-GlcNAc spectra across all eleven cell lines
analyzed by Geiger *et al.* (Geiger, T., Wehner, A., Schaab, C., Cox, J., and Mann, M.
Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying
expression of most proteins. Mol. Cell. Proteomics 11, M111 014050).