Title

Electron transfer dissociation (ETD): The mass spectrometric breakthrough essential for O-GlcNAc protein site assignments—a study of the O-GlcNAcylated protein Host Cell Factor C1

Journal

Proteomics, 13(6)

ISSN

1615-9853

Authors

Myers, Samuel A
Daou, Salima
Affar, El Bachir
et al.

Publication Date

2013-03-01

DOI

10.1002/pmic.201200332

Peer reviewed

Published in final edited form as:
Proteomics. 2013 March ; 13(6): 982–991. doi:10.1002/pmic.201200332.

Electron transfer dissociation (ETD): The mass spectrometric breakthrough essential for O-GlcNAc protein site assignments – A study of the *O*-GlcNAcylated protein Host Cell Factor C1

Samuel A. Myers[1], **Salima Daou**[2], **El Bachir Affar**[2], and **AL Burlingame**[1]

[1]Department of Pharmaceutical Chemistry, UCSF, San Francisco, CA 94158-2517

[2]Maisonneuve-Rosemont Hospital Research Center, Department of Medicine, University of Montréal, Montréal PQ H1T 2M4, Canada

Abstract

The development of electron-based, unimolecular dissociation mass spectrometric methods, i.e. electron capture and electron transfer dissociation (ECD and ETD, respectively), has greatly increased the speed and reliability of labile post-translational modification (PTM) site assignment. The field of intracellular *O*-GlcNAc (*O*-linked *N*-acetylglucosamine) signaling has especially advanced with the advent of ETD mass spectrometry. Only within the last five years have proteomic-scale experiments utilizing ETD allowed the assignment of hundreds of *O*-GlcNAc sites within cells and subcellular structures. Our ability to identify and unambiguously assign the site of *O*-GlcNAc modifications using ETD is rapidly increasing our understanding of this regulatory glycosylation and its potential interaction with other PTMs. Here, we discuss the advantages of using ETD, complimented with collisional-activation mass spectrometry (CID/CAD), in a study of *O*-GlcNAc modified peptides of the extensively *O*-GlcNAcylated protein Host Cell Factor C1 (HCF-1). HCF-1 is a transcriptional co-regulator, forms a stable complex with *O*-GlcNAc transferase and is involved in control of cell cycle progression. ETD, along with higher energy collisional dissociation (HCD) mass spectrometry, was employed to assign the PTMs of the HCF-1 protein isolated from HEK293T cells. These include nineteen sites of *O*-GlcNAcylation, two sites of phosphorylation and two sites bearing dimethylarginine, and showcase the residue-specific, PTM complexity of this regulator of cell proliferation.

Keywords

Electron Transfer Dissociation; *O*-GlcNAc; Host Cell Factor C1; Post-translational modifications

Introduction

While the post-translational modification (PTM) of nuclear and cytosolic proteins by *O*-linked *N*-acetylglucosamine of serine and threonine residues (*O*-GlcNAc) was discovered

Corresponding Author Prof. A.L. Burlingame, 600 16th St. MC2240, San Francisco, CA 94158, alb@cgl.ucsf.edu, Fax (415) 502-1655.

The authors declare no conflicts of interest.

almost three decades ago, our knowledge of its biological functions has lagged far behind that of serine and threonine phosphorylation [1–4]. Development of robust phosphopeptide enrichment strategies and electrospray tandem mass spectrometry using collision-induced dissociation (CID/CAD) have enabled the accelerated assignment of thousands of phosphorylation sites [5]. In contrast, *O*-GlcNAc enrichment combined with conventional mass spectrometric methods have not proven to be as effective for high throughput site identification. While methods for *O*-GlcNAc peptide enrichment have improved, the major hurdle to identifying residues modified by *O*-GlcNAc with mass spectrometry has been the lability of the *O*-glycosidic bond upon collisional activation [6].

However, with the discovery of electron capture dissociation (ECD) by gas-phase polyprotonated peptide species and recognition that this internal energy deposition process was non-ergodic, it became clear that peptide backbone bond cleavage(s) occurred rapidly without vibronic energy randomization to side chain residues or to any covalently linked modifications. It was quickly shown by McLafferty and collegues that γ-carboxyglutamic acid could be observed intact, shifted appropriately in peptide sequence ion series [7]. Zubarev and colleagues then established that *O*-linked hexosamine modified peptides from mucin derived synthetic glycopeptides were stable in the sequence ion series as well [8]. ECD was soon followed by the development of the electron transfer process using energetically suitable anion-radical electron donor species [9, 10]. This new technology was then implemented on the LTQ-Orbitrap platform by Coon and co-workers, setting the stage for highly sensitive, rapid electron-transfer dissociation (ETD) analysis of complex mixtures [11]. Studies from our laboratory then showed that 58 *O*-GlcNAc modified peptides could be assigned by interpretation of ETD spectra recorded during in a single 60-minute capillary UPLC analysis coupled with an LTQ Orbitrap XL [12].

With a reliable mass spectrometry strategy for high sensitivity *O*-GlcNAc peptide site assignment in hand, there is the additional need for methodologies that will enable the enrichment of cytosolic and nuclear components bearing this sub-stoichiometric PTM. Two primary strategies are presently being developed and utilized to enrich for *O*-GlcNAc at the peptide level: one based on enzymatic and chemical derivitization of *O*-GlcNAc peptides, where the most recent iteration utilizes a cleavable, biotinylated reagent [13, 14]. The other enriches native *O*-GlcNAc modified peptides using lectin weak affinity chromatography (LWAC), specifically wheat germ agglutinin (Figure 1, [12, 15, 16]).

Overall, these recent studies have established that a wide variety of proteins, involved in regulating most cellular processes, are modified by *O*-GlcNAc. Thus far the most extensive studies have focused on assignment of *O*-GlcNAc sites in cerebrocortical brain tissue and the murine synapse. Alfaro and co-workers reported the identification of sites on 195 proteins from cerebrocortical tissue using the chemical/enzymatic photocleavable *O*-GlcNAc enrichment strategy [13]. This study also described evidence of *O*-GlcNAc on extracellular domains of membrane proteins. Recent studies from our laboratory have extended earlier work on the postsynaptic density pseudo-organelle [12] to synaptosomes and have reported over 1700 *O*-GlcNAc sites together with 16,500 phosphorylation sites from the *same* biological sample [17]. This study revealed 439 peptides containing multiple sites of *O*-GlcNAcylation. Both *O*-GlcNAc and phosphorylation sites were found to cluster

preferentially on unstructured or disordered regions, but were found not to co-cluster on the same residues. Kinases were the class of proteins modified most extensively, suggesting cross talk at the level of regulation of enzyme activity [18, 19].

In addition to the assignment of *O*-GlcNAc sites at the proteomic scale now possible with ETD mass spectrometry, ETD has also greatly facilitated protein-specific studies of single *O*-GlcNAcylation events. Housley *et al.* employed ETD to identify sites of *O*-GlcNAcylation on the nutrient-responsive transcription factor FOXO1, and provided insight to how the sole, mammalian *O*-GlcNAc transferase (OGT) targets its substrates [20]. More recently, Rexach and colleagues were able to identify the *O*-GlcNAc sites on the transcription factor CREB from analysis of a chymotryptic peptide and were able to show a single *O*-GlcNAcylation site can influence neurite outgrowth and development of long-term memory in mice [21]. These types of site-specific *O*-GlcNAc studies will continue to contribute to our knowledge of the function(s) of protein *O*-GlcNAcylation.

One of the most remarkable and heavily *O*-GlcNAc modified proteins is Host Cell Factor C1 (HCF-1). HCF-1 is a large, chromatin associated scaffolding protein that plays an important role in cell proliferation [22–24]. HCF-1 has been repeatedly described as *O*-GlcNAcylated [12, 15, 16, 25, 26] and to form a stable complex with OGT [12, 15, 25–27, 28]. Interestingly, it was recently shown that OGT promotes the proteolytic maturation of HCF-1, which is critical for normal HCF-1 function [29, 30]. However, the complex relationship between OGT and HCF-1 is not well understood. *O*-GlcNAc modification is necessary for HCF-1 maturation, but might also be involved in regulating HCF-1 function after its proteolytic cleavage. Therefore, to understand the role of the *O*-GlcNAc modification of HCF-1, one must be able to identify the precise sites of modification so that biochemical and cell biological studies can take place. Here, we describe mass spectrometric methods used to identify the sites of *O*-GlcNAcylation and other PTM's on HCF-1 and illustrate ways to increase confidence in both glycopeptide identification and site assignment.

Materials and Methods

Purification and preparation of HCF-1

HEK293T (293T) cells were maintained in DMEM with 5 % FBS. The cells were plated at 70 % confluency (10 × 10 cm dishes) and transfected with 7 μg of pGCN-HA-HCF-1 FL using Polyethylenimine (2 mg/ml) in serum free media. Next day, the transfected cells were plated into 15X15cm dishes. Three days post-transfection, cells were harvested for HCF-1 purification. Total cell extracts were prepared using the lysis buffer (50 mM Tris-HCl, pH 7.3; 5 mM EDTA; 300 mM Nacl; 10 mM NAF; 1% NP-40; 1 mM phenylmethylsulfonyl fluoride (PMSF); 1 mM dithiothreitol; protease inhibitors cocktail (Sigma), and 5 μM PUGNAC (TRC)). Following centrifugation at 15000 rpm/30 min, the supernatant was incubated overnight at 4 °C with 200 μl of anti-HA-agarose beads (Sigma) (100 μl packed beads for 10 ml extract). The beads were washed several times with the lysis buffer and then transferred into a 800 μl chromatography columns (Biorad) for elution. Bound proteins were eluted with 200 μg/ml of HA peptide (sigma). The purified proteins were precipitated with cold TCA (30%), washed with cold acetone and resuspended in SDS-PAGE sample buffer

for western blotting and Coomassie staining. Anti-HCF-1 antibodies, a generous gift from Winship Herr's laboratory, and the anti-*O*-GlcNAc antibody RL-2 (Santa Cruz Biotechnology) were used as previously described [29]. Following gel destaining, HCF-1 bands were cuts from the gel, reduced with 2 mM TCEP (Thermo) for 60 minutes at 56°C, alkylated with iodoacetamide (Sigma) for 30 minutes at room temperature in the dark and digested in-gel with 1:100 trypsin or GluC (Roche). Peptides were extracted with 5% formic acid and 50% acetonitrile, concentrated using C18 Ziptips (Millipore) and vacuum centrifugation followed by LC-MS/MS analysis.

LC MS/MS analysis

Chromatography was performed on a Nanoacquity HPLC (Waters) at 600 nl/min with a BEH130 C18 75 μM ID ×150mm column (Waters). A 120-minute gradient from 2% solvent A (0.1% formic acid) to 35% solvent B (0.1% formic acid in acetonitrile) was used. Mass spectrometry was performed on an LTQ-Orbitrap Velos equipped with ETD (Thermo). Data dependent analysis selected the three most highly abundant, multiply charged ions within a 3 Da isolation window for subsequent HCD and ETD. Precursor scans and HCD product ions were measured in the Orbitrap at a resolution 30,000 and 7,500, respectively. For ions measured in the Orbitrap, one microscan of 250 ms injection time was used. ETD product ions were measured in the ion trap with one microscan, allowing 100 ms for ion injection time. Normalized activation energy and activation time for HCD was set to 30 and 30 msec, respectively. ETD activation time was charge state dependent, where doubly charged precursors reacted for 100 ms, triply charged for 66.6 msec, and so on. Automatic gain control for precursor ions was set at 1e6, 5e4 MS/MS scans and 1e6 for the ETD reagent, fluoranthene. Supplemental activation was enabled for ETD. Dynamic exclusion was set for 45 seconds.

Data analysis

Raw data was converted to peaklists using in-house software called PAVA [31]. HCD and ETD peaklists were searched separately using Protein Prospector v 5.10.0. against the Swissprot database with a concatenated, decoy database (21 March, 2012) where 36,775 entries were searched. Only human and mouse genomes were searched. Precursor mass tolerance was set to 10 ppm, where fragment ion error was allowed at 20 ppm and 0.6 Da for HCD and ETD, respectively. Cysteine residues were assumed to be carbamidomethylated, variable modifications considered were *N*-terminal acetylation, *N*-terminal pyroglutamine conversion and methionine oxidation. Trypsin was allowed one missed cleavage, while GluC was allowed up to two. Proteins identified from this analysis were searched again allowing for HexNAc (neutral loss for HCD data) and phosphorylation of serine and threonines, acetylation, methylation and carbamidomethylation of lysines though only the former two for arginine residues. SLIP scoring was reported for all modification site localization [32] unless the modification site was ambiguous (a false localization rate of less than 5%), then manual interpretation was employed [33].

Results

HCF-1 purification yielded several polypeptides that were readily visible with silver stain and Coomassie blue. Immunodetection of *O*-GlcNAc using the RL2 antibody confirmed that HCF-1 is strongly *O*-GlcNAcylated (Figure 2a).

HCF-1 was then subjected to proteolytic digestion using both GluC and trypsin. Both peptide digests were analyzed using capillary UPLC coupled to an ESI LTQ Velos Orbitrap mass spectrometer using both HCD and ETD. The results from this combined mass spectral characterization of the trypsin and GluC digests provided nearly complete coverage of the HCF-1 protein sequence. GluC was important in providing peptides that covered the latter C-terminal portion of the protein (Figure 2b and Supplemental figure 1).

The same doubly *O*-GlcNAcylated tryptic peptide was analyzed using HCD and ETD energy deposition (Figure 2c and 2d). Collisional activation of this *O*-linked glycopeptide caused a prominent loss of GlcNAc due to the lability of the *O*-glycosidic bond(s) (Figure 2c). The ETD spectrum provided both c- and z-ions series that permitted both sequence determination of the peptide and the site assignments for both *O*-GlcNAc moieties through appropriate mass shifts (203.08 Da) of product sequence ions (Figure 2d). Note that the supplemental activation required to promote efficient product ion dissociation can cause sugar losses [34]. This example illustrates how HCD, while not providing site assigment, can be used to complement ETD in *O*-GlcNAc peptide analysis.

The efficiency of the electron transfer process for glycopeptides follows similar trends as has been described for unmodified peptides, where the greater the charge density, i.e. lower m/z, the more extensive the product ion series is obtained [35]. This phenomenon is especially relevant for glycopeptides, since the GlcNAc moiety itself adds a large, uncharged mass that acts to reduce the overall peptide charge density. This point is especially important since several commercially available bioinformatic, proteomic-searching software strategies do not adjust scoring algorithms accordingly. Without appropriate product ion scoring, some peptides remain unidentified leading to lower protein coverage than is actually present in the dataset. Protein Prospector and pFind are examples of software that adjust scoring based on charge state for ETD spectra [36–38]. The difference in ETD efficiency and resulting spectrum quality from the same doubly and triply charged ion is presented in Figure 3.

While the loss of the GlcNAc prevents reliable site assignment, there is utility in collisionally activating *O*-glycopeptides in beam-type collision instruments, especially HCD [26]. The sugar loss is detected as a singly charged GlcNAc oxonium ion at 204.087 m/z and often as GlcNAc fragments at 138.056, 144.066, 168.066 and 186.077 m/z. When measured with high mass accuracy, as done by the Orbitrap measuring HCD product ions, no amino acid composition can be mistaken for an elemental composition of 204.087m/z. The detection of GlcNAc diagnostic ions becomes useful when poor quality ETD spectra suggest an *O*-GlcNAc modified peptide. Identifying the GlcNAc ion and its fragments in HCD spectra from the same precursor of a suspect ETD *O*-GlcNAc modified peptide assignment provides certainty for the identification of an *O*-linked glycopeptide (Figure 4a). Though

there is a possibility of co-eluting/co-fragmented *O*-glycopeptides that can contribute to the HexNAc oxonium ion presence, an appropriate precursor isolation window can minimize such “contaminating ions.” This, again, illustrates the utility of ETD product ions retaining the HexNAc mass addition. While there are exceptions to every rule, the absence of a GlcNAc oxonium ion in collisional activation suggests a false, algorithm based, *O*-linked glycopeptide identification. Finally, another advantage of using HCD activation, that is, over ion trap collisional activation, is beam-type collisions in the HCD cell allow for secondary collisions, which can provide peptide sequence information (Figure 4b).

Nineteen sites of *O*-GlcNAcylation on HCF-1 were identified in this study, several of which were located in the *N*-terminal region (Table 1). Four of these *O*-GlcNAc sites reside in the proteolytic processing domain (PPD) of HCF-1, two of which were found in the conserved cleavage sequence [29, 30]. Two S/T phosphorylation sites, as well as two dimethylarginine sites (one of which was also identified at singly methylated) were identified in the PPD or basic region of HCF-1 [29, 30].

Discussion

HCF-1 is one of the most highly *O*-GlcNAcylated proteins found in cells. In this study we have identified 19 sites of *O*-GlcNAcylation on HCF-1 purified from HEK293T cells, six of which are novel sites (Table 1). Nearly 30 additional HCF-1 *O*-GlcNAc sites have previously been reported from different cell or tissue types in studies enriching for *O*-GlcNAcylated peptides that were not detected in this study (Supplemental Table 1). Combining the *O*-GlcNAc sites from this and other studies, including coverage of the C-terminal portion of HCF-1 using GluC digestion described here, we show that nearly all of the *O*-GlcNAc sites are located in the *N*-terminal half of the mature HCF-1 protein (Figure 2b). These *O*-GlcNAc region-specific assignments are consistent with previously published literature showing that OGT forms a stable complex with the *N*-terminal portion of HCF-1, including the PPD [28–30]. It has been proposed that *O*-GlcNAcylation of HCF-1 provides a signal required for its proteolytic maturation. Of note, OGT remains associated with HCF-1 polypeptides suggesting that these complexes may have other functions distinct from proteolytic maturation of HCF-1, such as modulation of epigenetic enzymes. HCF-1 is well known to associate with transcription regulatory proteins such as the Sin3a-HDAC complex (several components of which are *O*-GlcNAc modified) and Set/1Ash2 methyltransferase complex [6, 15, 28, 39, 40]. Thus, OGT may be recruited to HCF-1 to promote maturation, and allow the OGT-HCF-1 complex to modify transcriptional regulators.

Evidence that the OGT-HCF-1 complex recruits other transcriptional regulatory proteins is further implied by the methylation and dimethylation of arginine residues on HCF-1 (Table 1). CARM1, also known as protein arginine methyltransferase 4 (PRMT4), is known to mono- and asymmetrically dimethylate arginine residues on proteins involved in transcriptional regulation and mRNA stability [41–43]. Moreover, CARM1 interacts with, and is modified by, OGT [44]. While this study does not establish that CARM1 is responsible to HCF-1 arginine methylation, is it not unreasonable to suggest CARM1 may associate with the OGT-HCF-1 complex, and influence transcription.

The identification of phosphorylation within the heavily *O*-GlcNAcylated area of the basic region of HCF-1 lends evidence to a possible interaction between *O*-GlcNAc and phosphorylation. Both of the phosphorylated residues of HCF-1 are proline-directed, where a proline residue directly follows a phosphorylated serine or threonine. In fact, one phosphopeptide identified in this study contains the GSK3β phosphorylation motif, SXXXS, where the latter serine is phosphorylated. Inhibition of GSK3β with LiCl causes a decrease in *O*-GlcNAcylation of HCF-1 in COS7 cells [25]. Perhaps HCF-1 phosphorylation by GSK3β promotes OGT recruitment and *O*-GlcNAcylation, forming the mature OGT-HCF-1 complex as a mechanism to fine tune transcription through a wide reaching regulatory complex. It remains to be determined how these *O*-GlcNAcylation states of HCF-1, as well as the other post-translationally modified states, affect transcription and cell proliferation.

Conclusions

Electron-based unimolecular dissociation mass spectrometric methods, i.e. ECD and ETD, have revolutionized *O*-GlcNAc analytical biochemistry. Glycopeptides previously unidentifiable are now readily identifiable on a chromatographic time-scale from purified proteins, or enriched peptides mixtures. ETD mass spectrometry provided this study the means to characterize the extent of *O*-GlcNAc site occupancy of HCF-1, a heavily modified protein involved in many aspects of transcriptional regulation and cell proliferation. This emerging technology will continue to allow us to further our fledgling understanding of *O*-GlcNAc biology and its interaction with other PTM systems, namely phosphorylation and arginine methylation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Winship Herr for his generous gift of HCF-1 construct and antibodies and Kati Medzihradszky and Barbara Panning for reading the manuscript. This work was supported by the Biomedical Technology Research Centers program of the NIH National Institute of General Medical Sciences, NIH NIGMS 8P41GM103481 and Howard Hughes Medical Institute. Work in E.B.A's laboratory was supported by grants from the Canadian Institutes for Health Research (CIHR) (MOP115132). E.B.A. is a scholar of the Canadian Institutes for Health Research (CIHR) and Le Fonds de la Recherche en Santé du Québec (FRSQ).

Abbreviations used

ECD	electron capture dissociation
ETD	electron transfer dissociation
HCD	higher energy collision induced dissociation
HCF-1	Host Cell Factor C1
***O*-GlcNAc**	*O*-linked *N*-acetylglucosamine

References

1. Torres CR, Hart GW. Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. J Biol Chem. 1984; 259(5):3308–3317. [PubMed: 6421821]
2. Hanover JA, Krause MW, Love DC. The hexosamine signaling pathway: O-GlcNAc cycling in feast or famine. Biochim Biophys Acta. 2005; 1800(2):80–95. [PubMed: 19647043]
3. Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu Rev Biochem. 2011; 80:825–858. [PubMed: 21391816]
4. Love DC, Krause MW, Hanover JA. O-GlcNAc cycling: emerging roles in development and epigenetics. Semin Cell Dev Biol. 2010; 21(6):646–654. [PubMed: 20488252]
5. Beltrao P, Albanese V, Kenner LR, Swaney DL, Burlingame A, Villen J, Lim WA, Fraser JS, Frydman J, Krogan NJ. Systematic functional prioritization of protein posttranslational modifications. Cell. 2012; 150(2):413–425. [PubMed: 22817900]
6. Chalkley RJ, Burlingame AL. Identification of novel sites of O-N-acetylglucosamine modification of serum response factor using quadrupole time-of-flight mass spectrometry. Mol Cell Proteomics. 2003; 2(3):182–190. [PubMed: 12684542]
7. Kelleher NL, Zubarev RA, Bush K, Furie B, Furie BC, McLafferty FW, Walsh CT. Localization of labile posttranslational modifications by electron capture dissociation: the case of gamma-carboxyglutamic acid. Anal Chem. 1999; 71(19):4250–4253. [PubMed: 10517147]
8. Mirgorodskaya E, Roepstorff P, Zubarev RA. Localization of O-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. Anal Chem. 1999; 71(20): 4431–4436. [PubMed: 10546526]
9. Huang TY, Emory JF, O'Hair RA, McLuckey SA. Electron-transfer reagent anion formation via electrospray ionization and collision-induced dissociation. Anal Chem. 2006; 78(21):7387–7391. [PubMed: 17073403]
10. Coon JJ, Ueberheide B, Syka JE, Dryhurst DD, Ausio J, Shabanowitz J, Hunt DF. Protein identification using sequential ion/ion reactions and tandem mass spectrometry. Proc Natl Acad Sci U S A. 2005; 102(27):9463–9468. [PubMed: 15983376]
11. McAlister GC, Berggren WT, Griep-Raming J, Horning S, Makarov A, Phanstiel D, Stafford G, Swaney DL, Syka JE, Zabrouskov V, et al. A proteomics grade electron transfer dissociation-enabled hybrid linear ion trap-orbitrap mass spectrometer. J Proteome Res. 2008; 7(8):3127–3136. [PubMed: 18613715]
12. Chalkley RJ, Thalhammer A, Schoepfer R, Burlingame AL. Identification of protein O-GlcNAcylation sites using electron transfer dissociation mass spectrometry on native peptides. Proc Natl Acad Sci U S A. 2009; 106(22):8894–8899. [PubMed: 19458039]
13. Alfaro JF, Gong CX, Monroe ME, Aldrich JT, Clauss TR, Purvine SO, Wang Z, Camp DG 2nd, Shabanowitz J, Stanley P, et al. Tandem mass spectrometry identifies many mouse brain O-GlcNAcylated proteins including EGF domain-specific O-GlcNAc transferase targets. Proc Natl Acad Sci U S A. 2012; 109(19):7280–7285. [PubMed: 22517741]
14. Wang Z, Udeshi ND, O'Malley M, Shabanowitz J, Hunt DF, Hart GW. Enrichment and site mapping of O-linked N-acetylglucosamine by a combination of chemical/enzymatic tagging, photochemical cleavage, and electron transfer dissociation mass spectrometry. Mol Cell Proteomics. 2010; 9(1):153–160. [PubMed: 19692427]
15. Myers SA, Panning B, Burlingame AL. Polycomb repressive complex 2 is necessary for the normal site-specific O-GlcNAc distribution in mouse embryonic stem cells. Proc Natl Acad Sci U S A. 2011; 108(23):9490–9495. [PubMed: 21606357]
16. Vosseller K, Trinidad JC, Chalkley RJ, Specht CG, Thalhammer A, Lynn AJ, Snedecor JO, Guan S, Medzihradszky KF, Maltby DA, et al. O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. Mol Cell Proteomics. 2006; 5(5):923–934. [PubMed: 16452088]

17. Trinidad JC, Barkan DT, Gulledge BF, Thalhammer A, Sali A, Schoepfer R, Burlingame AL. Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. Mol Cell Proteomics. 2012
18. Butkinaree C, Park K, Hart GW. O-linked beta-N-acetylglucosamine (O-GlcNAc): Extensive crosstalk with phosphorylation to regulate signaling and transcription in response to nutrients and stress. Biochim Biophys Acta. 2009; 1800(2):96–106. [PubMed: 19647786]
19. Wang Z, Udeshi ND, Slawson C, Compton PD, Sakabe K, Cheung WD, Shabanowitz J, Hunt DF, Hart GW. Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. Sci Signal. 3(104):ra2. [PubMed: 20068230]
20. Housley MP, Udeshi ND, Rodgers JT, Shabanowitz J, Puigserver P, Hunt DF, Hart GW. A PGC-1alpha-O-GlcNAc transferase complex regulates FoxO transcription factor activity in response to glucose. J Biol Chem. 2009; 284(8):5148–5157. [PubMed: 19103600]
21. Rexach JE, Clark PM, Mason DE, Neve RL, Peters EC, Hsieh-Wilson LC. Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. Nat Chem Biol. 2012; 8(3):253–261. [PubMed: 22267118]
22. Mangone M, Myers MP, Herr W. Role of the HCF-1 basic region in sustaining cell proliferation. PLoS One. 2010; 5(2):e9020. [PubMed: 20126307]
23. Wysocka J, Liu Y, Kobayashi R, Herr W. Developmental and cell-cycle regulation of Caenorhabditis elegans HCF phosphorylation. Biochemistry. 2001; 40(19):5786–5794. [PubMed: 11341844]
24. Wysocka J, Reilly PT, Herr W. Loss of HCF-1-chromatin association precedes temperature-induced growth arrest of tsBN67 cells. Mol Cell Biol. 2001; 21(11):3820–3829. [PubMed: 11340173]
25. Wang Z, Pandey A, Hart GW. Dynamic interplay between O-linked N-acetylglucosaminylation and glycogen synthase kinase-3-dependent phosphorylation. Mol Cell Proteomics. 2007; 6(8): 1365–1379. [PubMed: 17507370]
26. Zhao P, Viner R, Teo CF, Boons GJ, Horn D, Wells L. Combining high-energy C-trap dissociation and electron transfer dissociation for protein O-GlcNAc modification site assignment. J Proteome Res. 2011; 10(9):4088–4104. [PubMed: 21740066]
27. Mazars R, Gonzalez-de-Peredo A, Cayrol C, Lavigne AC, Vogel JL, Ortega N, Lacroix C, Gautier V, Huet G, Ray A, et al. The THAP-zinc finger protein THAP1 associates with coactivator HCF-1 and O-GlcNAc transferase: a link between DYT6 and DYT3 dystonias. J Biol Chem. 285(18): 13364–13371. [PubMed: 20200153]
28. Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W. Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. Genes Dev. 2003; 17(7):896–911. [PubMed: 12670868]
29. Daou S, Mashtalir N, Hammond-Martel I, Pak H, Yu H, Sui G, Vogel JL, Kristie TM, Affar el B. Crosstalk between O-GlcNAcylation and proteolytic cleavage regulates the host cell factor-1 maturation pathway. Proc Natl Acad Sci U S A. 2011; 108(7):2747–2752. [PubMed: 21285374]
30. Capotosti F, Guernier S, Lammers F, Waridel P, Cai Y, Jin J, Conaway JW, Conaway RC, Herr W. O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1. Cell. 2011; 144(3):376–388. [PubMed: 21295698]
31. Guan S, Price JC, Prusiner SB, Ghaemmaghami S, Burlingame AL. A data processing pipeline for mammalian proteome dynamics studies using stable isotope metabolic labeling. Mol Cell Proteomics. 2011; 10(12) M111 010728.
32. Baker PR, Trinidad JC, Chalkley RJ. Modification site localization scoring integrated into a search engine. Mol Cell Proteomics. 2011; 10(7) M111 008078.
33. Medzihradszky KF. Peptide sequence analysis. Methods Enzymol. 2005; 402:209–244. [PubMed: 16401511]
34. Swaney DL, McAlister GC, Wirtala M, Schwartz JC, Syka JE, Coon JJ. Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors. Anal Chem. 2007; 79(2):477–485. [PubMed: 17222010]
35. Swaney DL, McAlister GC, Coon JJ. Decision tree-driven tandem mass spectrometry for shotgun proteomics. Nat Methods. 2008; 5(11):959–964. [PubMed: 18931669]

36. Baker PR, Medzihradszky KF, Chalkley RJ. Improving software performance for peptide electron transfer dissociation data analysis by implementation of charge state- and sequence-dependent scoring. Mol Cell Proteomics. 2010; 9(9):1795–1803. [PubMed: 20513802]
37. Darula Z, Chalkley RJ, Lynn A, Baker PR, Medzihradszky KF. Improved identification of O-linked glycopeptides from ETD data with optimized scoring for different charge states and cleavage specificities. Amino Acids. 2011; 41(2):321–328. [PubMed: 20652609]
38. Sun RX, Dong MQ, Song CQ, Chi H, Yang B, Xiu LY, Tao L, Jing ZY, Liu C, Wang LH, et al. Improved peptide identification for proteomic analysis based on comprehensive characterization of electron transfer dissociation spectra. J Proteome Res. 2010; 9(12):6354–6367. [PubMed: 20883037]
39. Narayanan A, Ruyechan WT, Kristie TM. The coactivator host cell factor-1 mediates Set1 and MLL1 H3K4 trimethylation at herpesvirus immediate early promoters for initiation of infection. Proc Natl Acad Sci U S A. 2007; 104(26):10835–10840. [PubMed: 17578910]
40. Yokoyama A, Wang Z, Wysocka J, Sanyal M, Aufiero DJ, Kitabayashi I, Herr W, Cleary ML. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. Mol Cell Biol. 2004; 24(13):5639–5649. [PubMed: 15199122]
41. Bedford MT, Clarke SG. Protein arginine methylation in mammals: who, what, and why. Mol Cell. 2009; 33(1):1–13. [PubMed: 19150423]
42. Hassa PO, Covic M, Bedford MT, Hottiger MO. Protein arginine methyltransferase 1 coactivates NF-kappaB-dependent gene expression synergistically with CARM1 and PARP1. J Mol Biol. 2008; 377(3):668–678. [PubMed: 18280497]
43. Schurter BT, Koh SS, Chen D, Bunick GJ, Harp JM, Hanson BL, Henschen-Edman A, Mackay DR, Stallcup MR, Aswad DW. Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. Biochemistry. 2001; 40(19):5747–5756. [PubMed: 11341840]
44. Cheung WD, Sakabe K, Housley MP, Dias WB, Hart GW. O-linked beta-N-acetylglucosaminyltransferase substrate specificity is regulated by myosin phosphatase targeting and other interacting proteins. J Biol Chem. 2008; 283(49):33935–33941. [PubMed: 18840611]

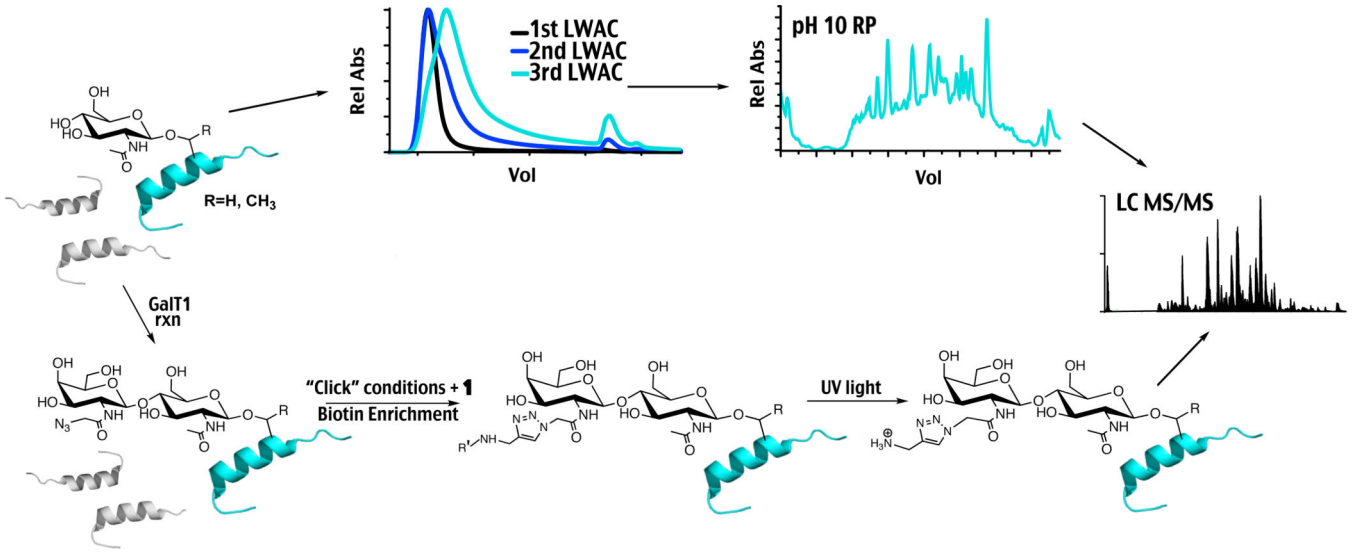

Figure 1.

Schematic of two widely used strategies for enrichment of *O*-GlcNAc modified peptides. The unmodified (grey peptides) and *O*-GlcNAc modified peptides (blue) can be enriched using single or multiple rounds of LWAC, fractionated by high pH reverse phase and analyzed using LC MS/MS (upper arm of diagram). Alternatively, *O*-GlcNAc modified peptides can be labeled with an azide containing galactosamine, conjugated to a photocleavable-biotin alkynyl linker (Reagent 1), enriched via biotin and cleaved off solid support with UV light (lower arm of diagram). The aminotriazole-GalNAc-*O*-GlcNAc peptides are then analyzed by LC MS/MS.

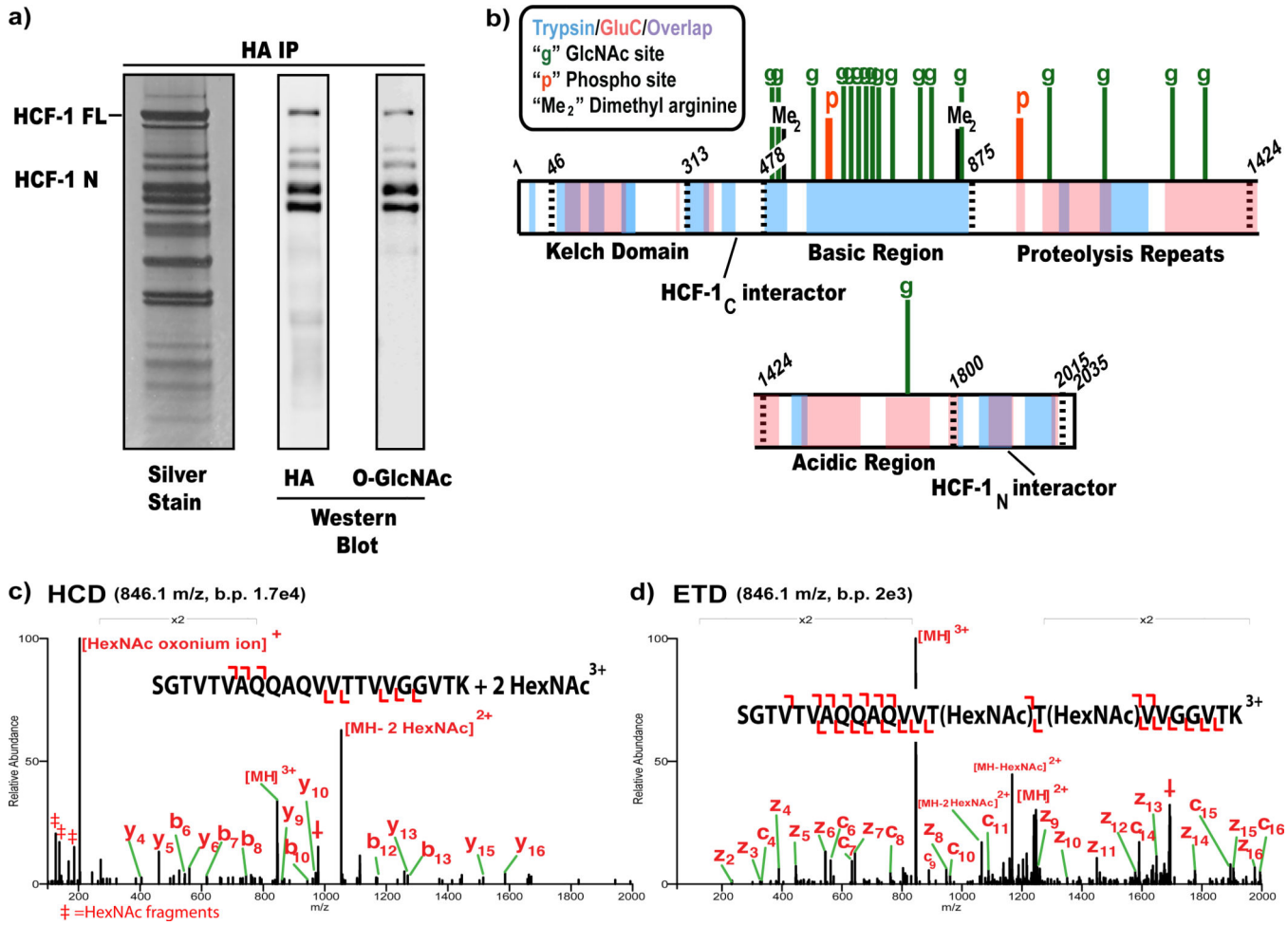

Figure 2.

Purification and characterization of HCF-1. a) HA tagged HCF-1 was purified form HEK293T cells. Eluates were silver stained or western blotted against HA and *O*-GlcNAc. b) Functional domains of HCF-1 are labeled (below diagram) where the landmark residue for the domains are numbered (above). Blue regions are trypsin-derived peptides, red are GluC derived and purple was covered by both enzymes. Green "g" signifies an *O*-GlcNAc site, while orange "p" phosphorylation and black, "Me_2", arginine dimethylation. c) HCD and d) ETD MS/MS spectra for a doubly *O*-GlcNAc modified HCF-1 tryptic peptide. "b.p" designates base peak intensity.

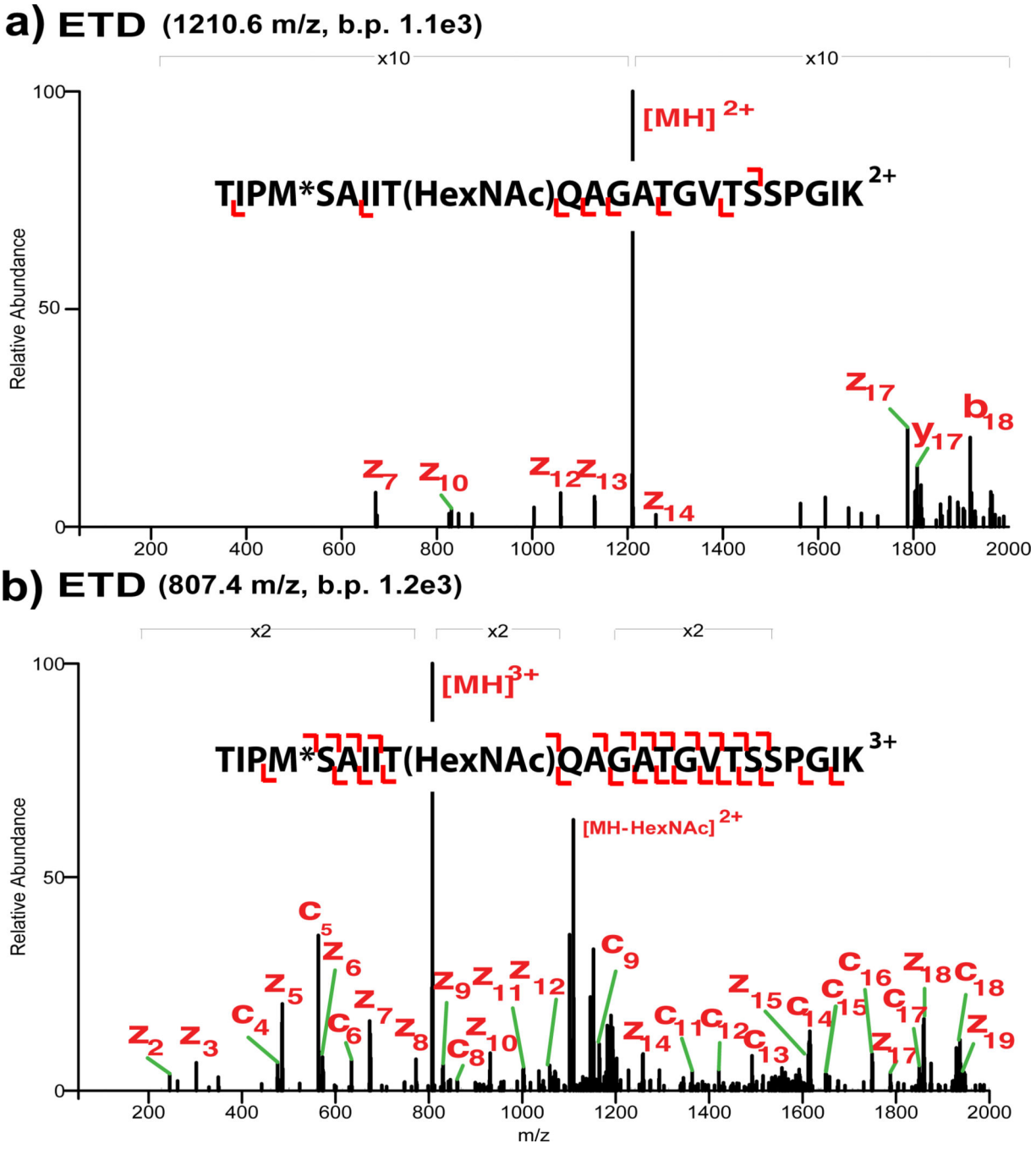

Figure 3.

ETD MS/MS spectra for the HCF-1 glycopeptide containing modification site T779 comparing ETD efficacies for the a) doubly charged and the b) triply charged precursors.

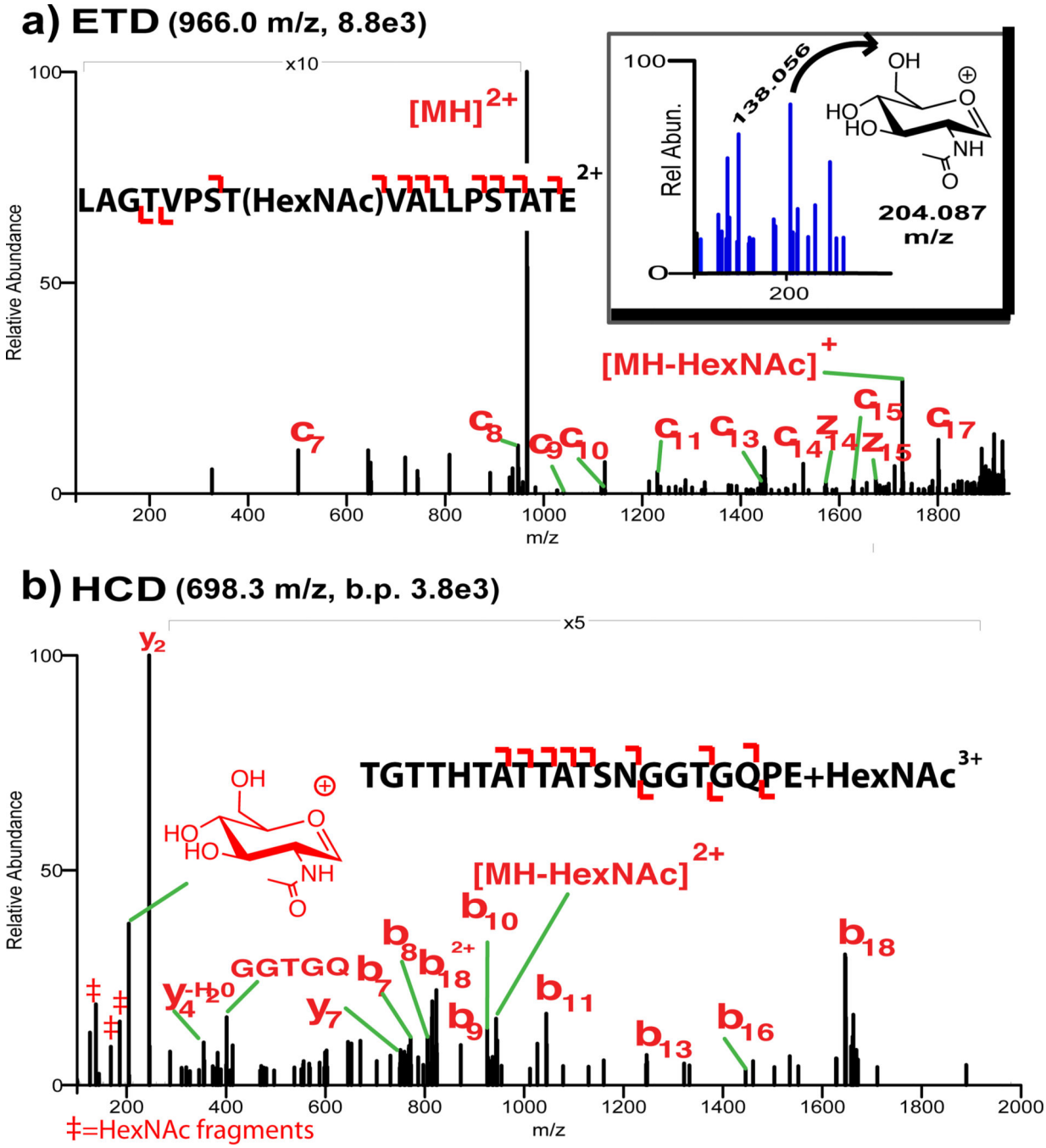

Figure 4.

a) GluC derived *O*-GlcNAcylated HCF-1 peptide lacking extensive z.-ions but containing c-ions allowing assignment of the modified residue. The insert, a zoom of the HCD spectrum from the same precursor, shows a prominent GlcNAc oxonium ion, providing further evidence of an O-linked hexosamine modified peptide. b) HCD spectrum of a GluC-derived glycopeptide. Sufficient b- and y-ions allow for peptide identification though no site-occupancy data is available. Prominent GlcNAc oxonium ion, and its fragments, provide evidence of O-GlcNAc modification.

Table 1

Modified HCF-1 peptides identified in this study.

m/z	z	ppm	Peptide sequence	PTM site of HCF-1 (P51610)	Prev. Id. or Novel	ETD ID/site info	HCD ID/site info
830.096	3	1.3	V**T**GPQA**T**TGTPLVTM*RPASQAGK	T490 & T495 mix	15, 17	yes/yes	yes/no
846.211	4	5.3	TQGVPAVLKV**T**GPQATTGTPLVTM*RPASQAGK	T490	17, 19	yes/yes	no/N.A.
931.767	4	3.3	IPPSSAPTVL**S**VPAGTTIVKTM*AVTPGTTTLPATVK	S569	Novel	yes/yes	yes/no
701.124	4	3.5	TAAAQVGT**S**VSSATNTSTRPIITVHK	S620	15, 17, 25	yes/yes	yes/no
701.124	4	4.1	TAAAQVGTSV**S**SATNTSTRPIITVHK	S622	13, 15, 17, 25	yes/yes	yes/no
701.124	4	3.5	TAAAQVGTSVS**S**ATNTSTRPIITVHK	S623	13, 15, 17, 25	yes/yes	yes/no
701.124	4	2.9	TAAAQVGTSVSSA**T**NTSTRPIITVHK	T625	25	yes/yes	yes/no
751.895	4	4.3	TAAAQVGTSV**S**S**A**T**N**T**S**TRPIITVHK	S620, ambiguous 2nd site.	15, 17	yes/yes	no/N.A.
778.428	3	5.8	SGTVTVAQQAQVV**T**TVVGGVTK	T651	13, 15, 17, 25	yes/yes	no/N.A.
846.120	3	4.2	SGTVTVAQQAQVV**TT**VVGGVTK	T651 & T652	13, 15, 17, 25	yes/yes	no/N.A.
1135.277	3	4.8	VM*SVVQTKPVQTSAVTGQASTGPVTQIIQTK	S685	13, 17	yes/yes	yes/no
984.136	5	−1.0	LV**T**SADGKPTTIITTTQASGAGTK‡PTILGISSVSPSTTK‡PGTTTIIK	T726	19	yes/yes	no/N.A.
807.428	3	4.2	TIPM*SAII**T**QAGATGVTSSPGIK	T779	13, 15, 17, 25	yes/yes	yes/no
731.730	3	2.9	SPITIIT**T**KVM*TSGTGAPAK	T801	13, 15, 16, 17, 25	yes/yes	no/N.A.
742.452	3	1.5	LVTPV**T**VSAVKPAVTTLVVK	T861	15, 17	yes/yes	yes/no
1211.590	3	0.3	TGTTN**T**A**T**TTVVANLGGHPQPTQVQFVCDRQE	T1030 or T1031	Novel (either)	yes/yes	no/N.A.
482.590	3	1.7	RACAAG**T**PAVIR	T1143	Novel	no/N.A.	no/N.A.
925.970	2	0.01	SLQGGSPSTTVTVT**A**LE	T1273	Novel	yes/yes	no/N.A.
698.319	3	0.005	TGTTHTAT**T**ATSNGGTGQPE	T1335	Novel	yes/yes	yes/no
966.018	2	3.5	LAGTVPS**T**VALLPSTATE	T1743	Novel	yes/yes	no/N.A.
767.453	4	1.5	TVPM*GGV**R(Me$_2$)**LVTPV**T**VSAVKPAVTTLVVK	Dimethyl@R855;HexNAc@T861	Novel	yes/yes	no/N.A.
767.077	3	3.4	VTGPQATTGTPLVTM***R(Me)**PASQAGK	Methyl@504	Novel	yes/yes	yes/yes
771.748	3	2.1	VTGPQATTGTPLVTM***R(Me$_2$)**PASQAGK	Dimethyl@504	Novel	yes/yes	yes/yes
716.685	4	5.1	TVPM*GGV**R(Me$_2$)**LVTPVTVSAVKPAVTTLVVK	Dimethyl@855	Novel	yes/yes	no/N.A.
756.352	2	3.4	VAS**S**PVM*VSNPATR	Phospho-S598	19	yes/yes	yes/yes

m/z	z	ppm	Peptide sequence	PTM site of HCF-1 (P51610)	Prev. Id. or Novel	ETD ID/site info	HCD ID/site info
977.983	2	0.003	AQPVHDLPVSILA**S**PTTE	Phospho-S984	Novel	yes/yes	no/N.A.

Residue numbering refers to Uniprot accession number P51610. Within the peptide sequence, bold black residues indicate unambiguous site assignment of HexNAc moiety, unless another modification is specified. Red residues indicate the site is ambiguous between one of the colored sites. The blue and green indicate a mixture of the two, singly HexNAc modified positional isomers is seen.

* indicates oxidation, and

‡ indicates lysine carbamidomethylation. Modified sites previously identified from studies reporting mass spectrometric statistics or annotated spectra are referenced, and correspond with the text. Otherwise, the identification is labeled “novel”. Peptides identified by ETD and/or HCD are designated as “ETD ID” and “HCD ID,” respectively. “Site info” refers to whether or not the dissociation method provided information on the site localization of the modification. Protein Prospector outputs for modified peptides are available for ETD (http://prospector.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=gjvopep7rv&search_name=msviewer) and HCD (http://prospector.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=vkwayutqbu&search_name=msviewer) at their respective links.