Method Optimization for N-Glycosylation Analysis: Case Study for SynCAM 1

Edward Z. Voss*[1], Terence Wu[1], Michael L. Easterling[2], Mary LoPresti[1], Christopher J. Thompson[2], Thomas Biederer[3], Kenneth R. Williams[1], and TuKiet T. Lam[1]

[1]WM Keck Foundation Biotechnology Resource Laboratory and [3]Molecular Biophysics and Biochemistry , Yale University, New Haven CT; [2]Bruker Daltonics Inc., Billerica, MS

Overview

To optimize sample preparation strategy for detecting N-glycosylated proteins and to apply these parameters towards understanding the glycosylation of SynCAM 1.

Introduction

The evolving field of mass spectrometry and glycosylation has been reviewed [1,2]. Yet, despite great interest, limited efforts have been made to optimize the analysis techniques. These approaches include the use of lectin affinity columns, glycoprotein specific staining of 2D electrophoresis gels for glycoprotein purification followed by enzymatic digestion, and subsequent MALDI-Tof or FT-ICR MS analyses. Here we devised a plan to utilize several known glycosylated proteins in a set of different protocols to evaluate their effectiveness in maximizing the assignment of site(s) of glycosylation. We will then utilize our findings to characterize the heavily glycosylated SynCAM 1 which plays a critical role in synapse formation of neuronal network [3,4].

Instrumentation

Two FT-ICR MS system were utilized for this work. Our 9.4 Tesla Bruker Apex Qe FTMS system (Figure 1a) was utilized to obtain direct infusion MS data (i.e. MS and MS/MS of peaks of interests). A similar nESI source 12T FT-ICR MS instrument (located at Bruker Daltonics, Billerica, MA; Figure 1b) was utilized in conjunction with a Dionex Ultimate 3000 LC system for online data dependent ETD MS/MS analyses.

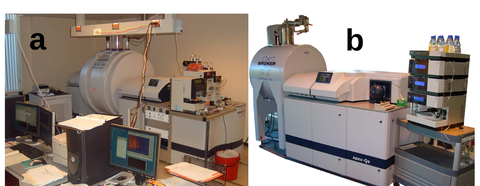

Figure 1. a) A 9.4Tesla Bruker Apex Qe Fourier Transform Ion Cyclotron Resonance Mass spectrometer in the Keck MS and Proteomics Resource located at the WM Keck Foundation Biotechnology Resource Laboratory; b)12T FT-ICR MS instrument located at Bruker Daltonics, Billerica, MA.

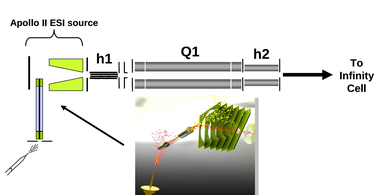

Figure 2. Ion optics of 9.4T FT-ICR MS showing the Apollo II ESI source which provided >10X improvements in spectral signals. Note that the generated ion beam exiting the glass capillary is 90º off axis from the remainder of the ion optics; this reduces that number of contaminations that enter into the mass spectrometer. Typical ion accumulation time ranges from 100 to 400 ms.

Samples

The SynCAM 1 extracellular domain were purified from stable HEK293 cell lines that express the protein and glycosylated it indistinguishably from brain tissue. Figure 3 shows the immunoblotting results of SynCAM 1-4 expressed in HEK293 cells. Table 1 summaries the samples utilized in our experiments for optimization and various N-linked glycosylation analyses.

From Table 1, briefly, 60ug of purified SynCAM 1 was utilized to generate 8 different sample sets for analysis (4 sets of varying enzyme digestion, 2 sets of glycans profiling from different deglycosylation method, and 2 sets of deglycosylated proteins). A set of standard proteins including Fetuin (The composition of bovine fetuin (weight % is polypeptide 74%, hexose 8.3%, hexosamines 5.5%, and sialic acid 8.7%), α1-Acid Glycoprotein (approximately 11-12% sialic acid, 13-17% neutral hexoses,12-15% hexosamine, and 0.7-1.5% fucose), and a mixture of ProteoProfile™ PTM Marker (Sigma) which contains the proteins and their PTM as shown in Table 2.

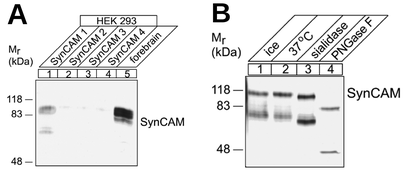

Figure 3. *SynCAM proteins are prominently expressed in brain.* **A**) HEK 293 cells were transfected with expression constructs encoding full length mouse SynCAM 1 – 4. Total cell lysates were prepared, and equal lysate fractions were analyzed by immunoblotting (lanes 1– 4). Twenty micrograms of adult rat forebrain proteins were analyzed as positive control (lane 5). **B**) SynCAM proteins are distinctly glycosylated and prominently expressed in brain during early postnatal development. SynCAM proteins are differentially glycosylated and equal protein amounts were subjected to enzymatic de-glycosylation at 37°C with sialicdase to remove sialic acids or PNGase F to remove N-linked carbohydrates. Control samples were incubated in parallel as indicated. The numbers on the left indicate positions of molecular weight markers. *Figure obtained with permission from*

Table 1: SynCAM 1 and standard samples utilized

Description	Working Conc (µg/µl)	Processes
SynCAM 1		
CNBR	0.476	no reduction/ no alkylation/C18 ZipTip
CNBR/Trypsin	0.476	reduction/alkylation/C18 ZipTip
LysC/Trypsin	0.278	reduction/alkylation/C18 ZipTip
Protease Type XIII	0.25	no reduction/ no alkylation/C18 ZipTip
PNGase treated (flowthrough)		Carbograph/reduction/alkylation/tryptic digestion/C18 ZipTip
EndoH treated (flowthrough)		Carbograph/no reduction/ no alkylation/ tryptic digestion/C18 ZipTip
PNGase treated (eluate)		Carbograph/C18 ZipTip
EndoH treated (eluate)		Carbograph/C18 ZipTip
Standard Proteins		
Fetuin	0.263	reduction/alkylation/ConA TopTip/C18 ZipTip
alpha-1-acid glycoprotein	0.263	reduction/alkylation/ConA TopTip/C18 ZipTip
protein standard mix:		
Albumin	0.263	reduction/alkylation/ConA TopTip/C18 ZipTip
beta-Casein	0.263	reduction/alkylation/ConA TopTip/C18 ZipTip
RNase B	0.263	reduction/alkylation/ConA TopTip/C18 ZipTip
Ovalbumin	0.263	reduction/alkylation/ConA TopTip/C18 ZipTip

Table 2: ProteoProfile™ PTM Marker (Sigma).

Protein	Carbohydrate (ng)	Phosphate (pmole)
Albumin	none	none
Ovalbumin	35	45
β-Casein	none	160
RNase B	200	none

Methods

All samples were processed as depicted in Figure 4. Starting 60ug SynCAM 1 purified sample was split into 6 equal aliquot (10ug each). 4 sets were targeted for enzyme digestion study (dark grey box in Figure 4); and 2 sets (light grey in Figure 4) were targeted for deglycosylation to remove glycans for profiling glycans and protein identification. For direct injection study on the 9.4T FT-ICR MS instrument, all samples were C18 ZipTip prior to infusion. For online LC MS analyses carried out on the 12T FT-ICR MS instrument (Bruker Daltonics, Billerica, MA), the LC separation was performed with a Dionex U3000 configured in capillary mode (100:1 split ratio), using a Higgins HAISIL 300 C18, 0.32x150mm, 5µ, 300Å column and Piccolo HAISIL 300 C18, 0.3x2.5mm, 5µm, 300Å guard column both heated to 40ºC. The flow rate was set to 5uL/min starting at 2%B (MeCN, 0.1% FA) and ramped to 35%B over 5 minutes followed by a slower ramp to 85%B over 40 minutes. The electrospray source was configured with a capillary (low flow) sprayer optimized for positive mode. Ions were detected in the 450-2500 m/z range over the 50 minutes LC run with 512K data points per MS scan. All data were processed utilizing DA Analysis Software v. 3.4, online GlycoMod (Expasy), and MASCOT search engine.

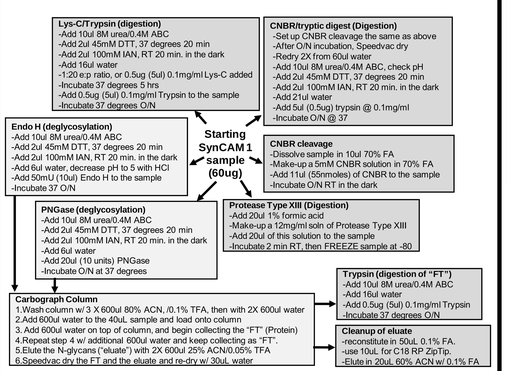

Figure 4. Schematic of how purified SynCAM 1 sample was split and process for method development. Dark gray shade boxes indicate enzyme study. Light grey shade indicate the deglycosylation study for looking at glycans and protein ID. All samples were stored at –80C after each procedure/process until MS analysis.

Standard proteins utilized (see Table 1) were all reduced, alkylated, and tryptic digested (similar to the procedure used in Figure 4). These were used to test for enrichment of glycopeptide strategy study. The lectin (ConA) based Top Tip utilized for the enrichment step is shown in Figure 5. The capacity for these tips is estimated at ~10-30ug. Eluate from the tips were then cleanup with either C18 ZipTip or by online LC separation for direct infusion and online LC MS analyses, respectively.

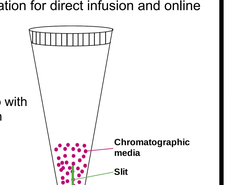

Figure 5: Top Tip (Glygen Corp.) is a pipette tip with a fine slit at the bottom (slit width: 1-2 µm which permits liquid to pass through but retains the chromatographic material (20-30 µm) in the tip.

Results

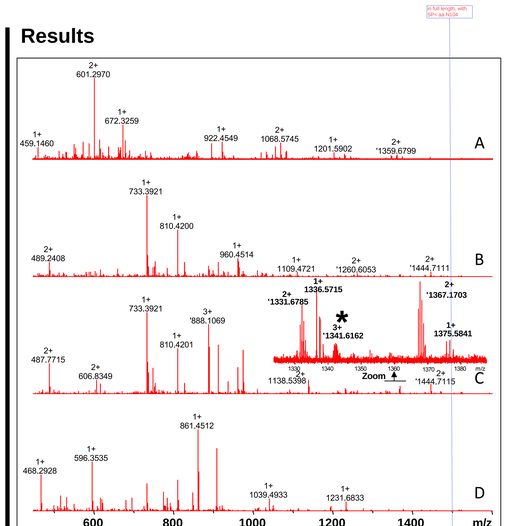

Figure 6. Broadband FT-ICR MS mass spectrum of four different enzymatic digestion of purified SynCAM-1 (non-deglycosylated). 10ug of SynCAM-1 was utilized for each digestion of A) CNBr, B) CNBr+Trypsin, C) Lys C+Trypsin, and D) Protease Type XIII. Digested samples were C18 RP ZipTip clean/desalted prior to collecting MS data. The * indicate one of the several peaks that will be targeted for further MS/MS fragmentation experiments. Based on FT-ICR accurate mass and GlycoMod [5] prediction, the * indicate the glycopeptide that corresponds to N63 of the extracellular domain portion of SynCAM-1.

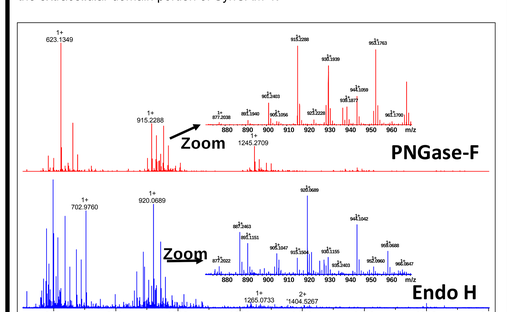

Figure 7. Glycan profiling by direct infusion analyses. Direct infusion of FT-ICR MS comparison spectrum of N-linked oligossacharides from PNGase (upper MS) and Endo H (lower MS) showing a multitude of positively charged ions observed.

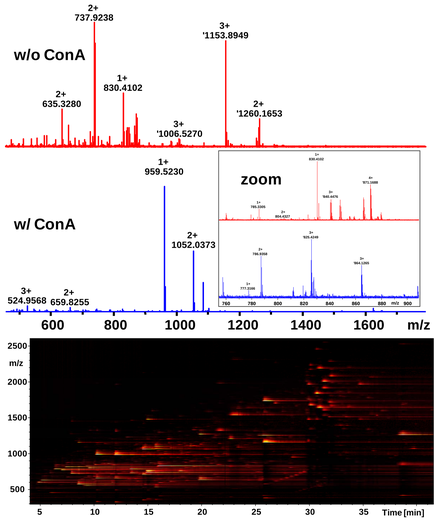

Figure 8. FT-ICR MS Spectral comparison for w/ (blue spectrum) and w/o (red spectrum) ConA enrichment for alpha-1-acid glycoprotein. The lower (yellow/red) ion map shows the complex spectral feature for on online LCMS profile of tryptic digest of alpha-1-acid glycoprotein.

Conclusion

- We have shown that in order to be effective at maximizing protein coverage for N-glycosylated protein a multi digestion scheme is necessary.
- Our predicted glycosylation site has been validated by site mutagenesis and additional biochemical and functional study.
- The Carbograph columns were effective in enriching for glycans and profiling glycans. This methodology is good for both purified and complex glycoprotein mixture.
- Lectin based affinity TopTip (ConA) is good for enriching for glycopeptide, however, there seem to be a very high back ground of non-specificity as observed that similar peaks are detected both in the flow through and bound fractions. A more thorough washing step is necessary to minimize this.
- FT-ICR MS high resolving power provides necessary separation for complex peptide and glycan mixture.

Acknowledgments

This project has been funded in part by NIH/NIDA (1 P30 DA 018343), Yale University, Bruker Daltonics Inc. in Billerica, MA, and the WM Keck Foundation Resource Laboratory in New Haven, CT. Our 9.4 T Apex Qe FTICR MS instrument was funded by NIH/NCRR 1 S10 RR17266-01.

References

[1] Morelle, W.; et al. (2006). *Proteomics* **6**(14): 3993-4015.
[2] David, J. H. (2009). *Mass Spectrometry Reviews* **28**(2): 273-361.
[3] Biederer, T.; et al. (2002) *Science* **297**, 1525-1531.
[4] Fogel, A.I.; et al. (2007) *J. Neuroscience* **27**, 12516-12530.
[5] Cooper, C.A.; et al. (2003) *Nucleic Acids Res.* **2003** January 1; 31(1): 511–513.