

Integrated mass spectrometry-based analysis of plasma glycoproteins and their glycan modifications

Hong Wang^{1,2}, Chee-Hong Wong¹, Alice Chin¹, Ayumu Taguchi¹, Allen Taylor¹, Samir Hanash¹, Sadanori Sekiya³, Hidenori Takahashi³, Masaki Murase³, Shigeki Kajihara³, Shinichi Iwamoto³, Koichi Tanaka³

¹ Fred Hutchinson Cancer Research Center, Seattle, WA. ² Ventana Medical Systems, Inc, Oro Valley, AZ. ³ SHIMADZU CORPORATION, Kyoto, Japan.

1. Introduction

Glycosylation is the most widespread and complex form of protein post-translational modification, as more than 50% of human proteins are glycosylated. Glycosylation plays a key role in protein folding, stability, activity, trafficking, molecular recognition and immunogenicity. In cancer, structural changes of oligosaccharide occur as part of initial oncogenic transformation, as well as part of key events in promoting tumor cell invasion and metastasis. An immune response may occur in cancer resulting from aberrant glycosylation. Therefore, profiling glycoproteins in disease states may have relevance to the development of biomarkers as well as therapeutics that target proteins with aberrant glycosylation.

We have previously implemented a MS-based strategy for in-depth analysis of the plasma proteome^{1,2)}.

In this study, we present a methodology with LC-MALDI-DIT MS/MS in which the analyses of protein sequence, glycosylation sites and glycan compositions are integrated. The presented method has been applied to the analysis of plasma glycoproteins.

2. Objective

To develop a platform for plasma glycoproteomics study based on mass spectrometry, allowing for the investigation of glycoproteins for their overall level as well as changes in their glycan structures associated with biological processes and disease states.

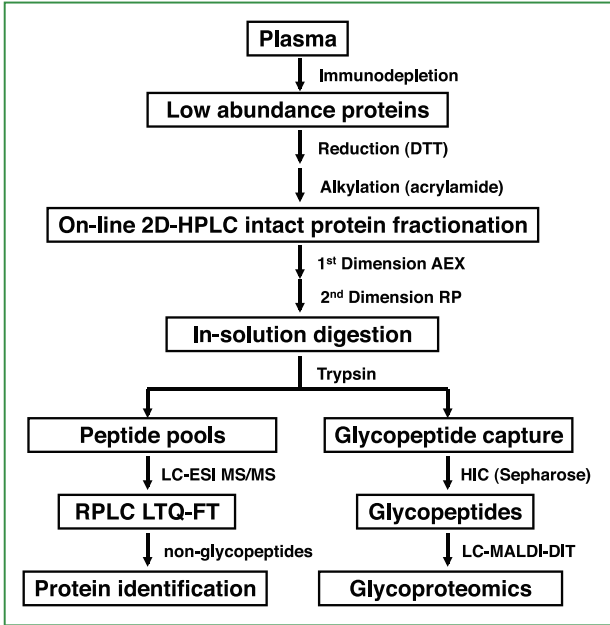


Figure 1 An integrated strategy for profiling human plasma glycoproteomics

3. Materials and Methods

Human plasma samples were processed by immunodepletion chromatography followed by an online, intact protein-based 2D-HPLC separation procedure according to a protocol published by our group (Fig.1)^{1,3)}. The low-abundance protein fraction (flow-through fraction) was fractionated by an online 2D-HPLC (anion exchange chromatography as the first dimension and reversed-phase chromatography as the second dimension) before protein digestion for LC-MS/MS analysis. Glycopeptides from the digested protein fractions are enriched by hydrophilic interaction chromatography⁴⁾, then spotted onto MALDI plate with an online reversed-phase liquid chromatography (RPLC) AccuSpot instrument (Shimadzu) for offline LC-MALDI-DIT MS/MS analysis (Fig.2). MS and MS/MS data were acquired by in-house built MALDI-DIT MS. Acquired spectra are automatically processed through an in-house built intact glycopeptide analysis pipeline (IGAP) to obtain integrated information pertaining to glycoprotein amino acid sequences, glycosylation sites and glycan compositions.

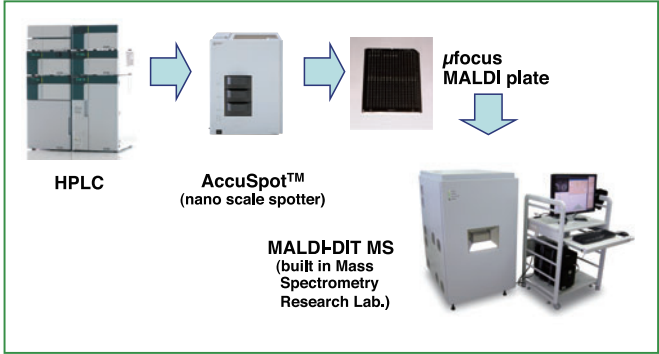


Figure 2 Offline LC-MALDI-DIT system

4. Results

As illustrated in Fig.1, LC-ESI LTQ-FT and LC-MALDI-DIT MS/MS were utilized for protein identification and glycoprotein characterization, respectively. By using LC-ESI LTQ-FT, 64 proteins were confidently identified and 30 of the identified proteins were annotated as glycoproteins from UniProt database entries. By using LC-MALDI-DIT MS/MS, 34 of the identified proteins were characterized as glycoproteins (Fig.3), and 24 of the characterized glycoproteins were annotated as glycoproteins from UniProt database entries. This results indicate that some peptides exhibit heterogeneity in their glycan attachments (Fig.4). For example, both 3Hex2GlcNAc1Fuc and 5Hex4GlcNAc are attached to the same complement factor I (CFI) peptide FLNN*GTCTAEGK, whereas both 3Hex2GlcNAc1Fuc and 5Hex4GlcNAc are attached to a C4B-binding protein β chain (C4BPB) peptide (EWDN*TTTECR). Glycosylation microheterogeneity at single glycosylation site is found with six different glycans (4Hex3GlcNAc, 5Hex4GlcNAc, 5Hex4GlcNAc1Sia, 6Hex5GlcNAc1Sia, 5Hex5GlcNAc2Sia1Fuc, and 6Hex5GlcNAc2Sia) attached to the same α -1-acid glycoprotein (ORM2) peptide (NEEYN*K).

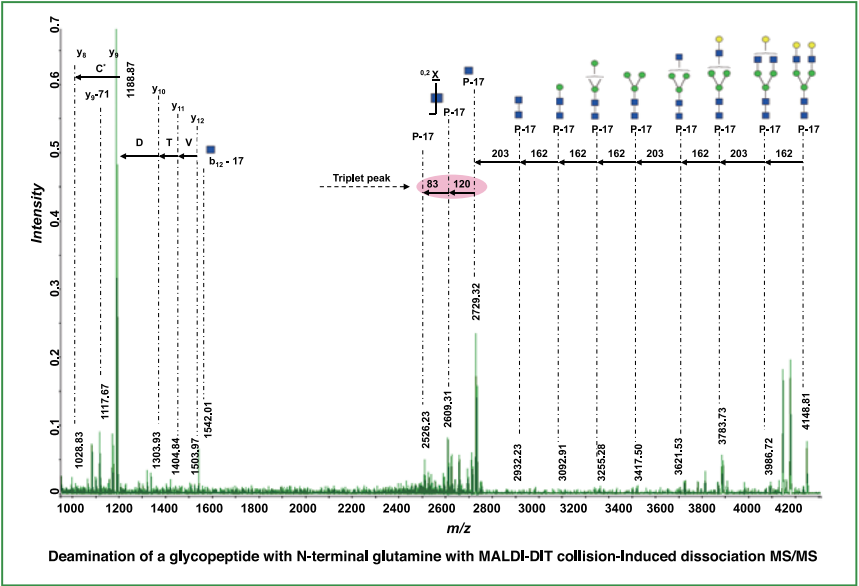


Figure 3 MS/MS data of glycopeptide

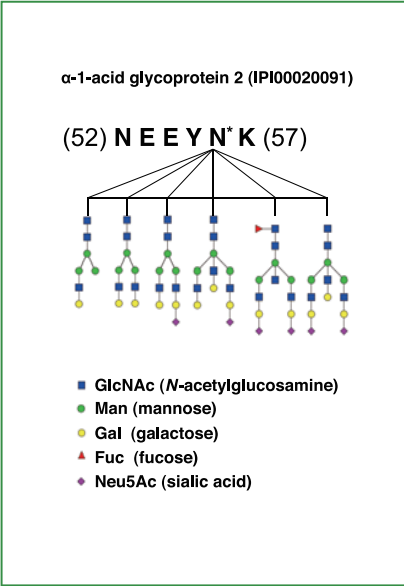


Figure 4 Glycosylation heterogeneity

5. Conclusion

The presented methodology provides an integrated mass spectrometry-based identification of glycopeptides in complex mixtures with simultaneous elucidation of peptide sequences, glycosylation sites and oligosaccharide compositions. Glycoprotein analysis has relevance to the role and extent of heterogeneity of proteins in health and in disease. The search for potential biomarkers representing modified glycoproteins in the blood would particularly benefit from this approach.

- References

1. Wang, H. *et al.* Intact-protein-based high-resolution three-dimensional quantitative analysis system for proteome profiling of biological fluids, *Mol. Cell. Proteomics*, 4, 618-625 (2005).
2. Fukuyama, Y. *et al.* Ionic liquid matrixes optimized for MALDI-MS of sulfated/sialylated/neutral oligosaccharides and glycopeptides. *Anal. Chem.* 80, 2171-2179 (2008).
3. Wang, H. *et al.* Integrated mass spectrometry-based analysis of plasma glycoproteins and their glycan modifications, *Nature Protocols*, 6, 253-269 (2011).
4. Wada, Y. *et al.* Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.* 76, 6560-6565 (2004).