

An Application of Electrostatic Repulsion Hydrophilic Interaction Chromatography in Phospho- and Glycoproteome Profiling of Epicardial Adipose Tissue in Obesity Mouse

Trang Huyen Tran^a, Injae Hwang^c, Jong-Moon Park^a, Jae Bum Kim^c, and Hookeun Lee^{a,b,*}

^aLee Gil Ya Cancer and Diabetes Institute, Gachon University, 7-45 Songdo-dong, Yeonsu-ku, Incheon 406-840, Korea

^bCollege of Pharmacy, Gachon University, 7-45 Songdo-dong, Yeonsu-ku, Incheon 406-840, Korea

^cInstitute of Molecular Biology and Genetics School of Biological Sciences, Seoul National University, San 56-1, Sillim-Dong, Kwanak-Gu Seoul 151-742, Korea

Received May 7, 2012; Revised May 19, 2012; Accepted May 20, 2012

First published on the web June 28, 2012; DOI: 10.5478/MSL.2012.3.2.39

Abstract: Phosphorylation and glycosylation are two of the most important and widespread post-translational modifications (PTMs) in an organism. Proteomics analysis of the PTMs has been challenged by low stoichiometry of the modified proteins and suppression effects by high abundance proteins, typically non-functional house-keeping proteins. In this study, a novel method was applied for not only isolating PTM peptides from intact peptides but also concurrently characterizing of glyco- and phosphoproteome using electrostatic repulsion hydrophilic interaction chromatography (ERLIC) packed with silica coated by cross-linked polyethyleneimine. For 2 mg tryptic digest of mouse proteome of epicardial adipose tissue with fat diet, 802 N-glycosylated peptides of 316 glycoproteins and 159 phosphorylated peptides of 75 phosphoproteins were identified using HPLC chip/quadrupole time-of-flight (Q-TOF) tandem mass spectrometer.

Key words: ERLIC, Phosphoproteome enrichment, Glycoproteome enrichment, Simultaneous enrichment, PTMs

Introduction

It is estimated that about 50% of all mammalian proteins are carrying glycan residues,¹ and 30% of all proteins are phosphorylated.² Protein post-translational modifications (PTMs) have been known as directly involved in protein function and cellular activities,³⁻⁵ therefore the identification and characterization of these PTMs from the same sample are the prerequisites for biological discovery. Due to the frequently low stoichiometry of PTMs and the ion suppression effect from unmodified peptides of high abundance, phosphopeptides and glycopeptides have to be isolated before MS analysis to minimize such suppression.⁶⁻¹⁰

So far, for protein fractionation to reduce sample complexity and increase the sensitivity, strong cation exchange (SCX)^{8,9} and anion-exchange (AEX) chromatography¹⁰ are the two most common types of chromatography. Recently, a novel mode of chromatography termed electrostatic repulsion hydrophilic interaction chromatography (ERLIC) was first introduced by Alpert for separation of biomolecules and phosphopeptide enrichment.¹¹ It has been extended to the fractionation of N-linked glycopeptides from complex samples based on negatively charged sialyl

group in glycopeptides.¹² However, it might be able to enrich other hydrophilic glycopeptides when an oligosaccharide side chain is a large and hydrophilic domain that causes a significant increase in retention time of the peptide in any hydrophilic interaction liquid chromatography-based mode of ERLIC. Based on using AEX column set up at the ERLIC mode, the combination of electrostatic attraction and hydrophilic interaction suffices to pull phosphorylated tryptic peptides and many glycopeptides away from the unmodified peptides. The simultaneous characterization of the glycoproteome and phosphoproteome of mouse brain membrane has been achieved with ERLIC.¹³ The method was developed in optimization of mobile phases and gradient for comprehensive profiling proteome along with PTMs study.¹⁴

In this work, ERLIC was applied to simultaneously characterize phospho- and glyco-proteome of epicardial adipose tissue in obesity mouse.

Experimental

ERLIC-system

PolyWAXTM packed with silica coated by cross-linked polyethyleneimine (200 × 4.6 mm, 5 μm particle diameter, 300 Å⁰ pore diameter, item 204WX0503) was a product of PolyLC Inc. (Columbia, MD). The mobile phases which made by acetonitrile (ACN) and formic acid (FA) as described with 80% ACN, 0.1% FA as buffer A and 10%

*Reprint requests to Dr. Hookeun Lee
E-mail: hkleee@gachon.ac.kr

ACN, 2% FA as buffer B to establish 140 minutes gradient for ERLIC.¹⁴ The detector (Waters 484) was set up at wavelength 280 nm.

Standard proteins

Six purified proteins purchased from Sigma Aldrich including bovine serum albumin (BSA), α -amylase, β -casein, β -lactoglobulin, transferrin, myoglobin were used as the standard, with 500 μ g each set. Samples were dissolved in 6 M urea, reduced with 5 mM tris (2-carboxyethyl) phosphine (TCEP) (Thermo Scientific) at 37 °C and alkylated with 15 mM iodoacetamide in the dark. For digestion, sequence-grade trypsin (Promega) was added at ratio of 1:25 (trypsin/protein) then incubated at 37 °C overnight. The sample was cleaned up by a C18 spin column (Harvard Apparatus) and eluted by 80% ACN in 0.1% FA which can be immediately injected into ERLIC system.

ERLIC in phospho- and glyco-proteome study of epicardial adipose tissue from obesity mouse

Mouse was fed with fat diet for sixteen weeks, and then the epicardial adipose tissue was collected. Proteins were extracted and assayed by 2-D Quan Kit (GE Healthcare Life Sciences). 4 mg of protein was reduced, alkylated and digested into peptides mixture, which then cleaned up by SepPak cartridges. 2 mg of digest reconstituted in 200 μ l of buffer A was loaded into a PolyWAX column. Fractions were collected at 1-min intervals. The consecutive fractions were combined, and finally a total of eight fractions were then dried, reconstituted into 30 μ l of 25 mM NH_4HCO_3 , and treated with 3 μ l of 500 units/ μ l peptide-N-glycosidase F (New England Biolabs) at 37 °C overnight. All samples were dried in a SpeedVac.

Enrichment of N-Linked Glycopeptides of epicardial adipose tissue from obesity mouse using hydrazide chemistry

To determine the efficiency of ERLIC in glycopeptides capturing, commercial hydrazide bead (Biorad) was tested with 2 mg of the same sample to compare. The procedure was followed the widely-used protocol for solid phase extraction of N-glycosylated peptides.¹⁵

Mass spectrometric analysis

Each dried fraction was reconstituted in 20 μ l of 0.1% FA and analyzed three times using the 6520 Accurate-Mass Q-TOF LC/MS (Agilent Technologies, DE) coupled with the 1200 HPLC unit (Agilent). For each analysis, 5 μ l of the samples was injected and separated in a capillary column packed with C18 at a flow rate of 0.3 μ l min^{-1} . Mobile phase A (0.1% FA in H_2O) and mobile phase B (90% ACN, 0.1% FA in H_2O) were used to establish the 30 min gradient of 5 to 45% B, 0–15 min; 45 to 90% B shocking, 15–15.01 min; 90% B, 15.01–20 min; 90 to 5% B, 20–20.01 min; and 5% B for 20.01–30 min. Peptide were then analyzed by an

electrospray potential of 1.8 kV. A drying gas flow of 5 L min^{-1} , and the fragmentor of 175 V were used. The q-TOF was set to perform data acquisition in the positive mode that the m/z range of 100–3000 was used in the MS/MS scan.

Database searching

Raw data from Q-TOF were converted to mzXML format and searched against the Mouse IPI (version 3.81). The database search was performed using both X!Tandem, Sorcerer (SEQUEST) with precursor and peptide mass tolerance of 50 ppm and 20 ppm respectively. The results were combined by iProphet. Carbamidomethylation (C) was set as a static modification, and oxidation (M), carbamylation at any residues, deamidation (N) after PNGase F deglycosylation for glycopeptides and phosphorylation (S, T and Y) for phosphopeptides were set as variable modifications. Peptides identified with a consensus N-X-S/T (with X is not Proline) and a modification of deamidation of Asparagine were regarded as N-linked glycopeptides. PeptideProphet and ProteinProphet were used to evaluate the quality of peptide and protein identification with error rate less than 1% and 5% respectively.

Protein classification and functional annotation using MetaCore

MetaCore version 6.4 (GeneGo Inc. St. Joseph, MI, USA) (www.genego.com) was used to classify and annotate the identified.

Results and Discussion

ERLIC Chromatogram

According to the principle of peptide fractionation by ERLIC, unmodified peptides hypothetically eluted in early fractions whereas glycopeptides retained in column and come out later.

The chromatography system was tested first with 500 μ g peptide mixture of 6 purified proteins, signal was detected at 280 nm in 140 minutes of gradient. Chromatogram generated as Figure 1a demonstrated that ERLIC is capable in peptide fractionation and ready to be applied on biological samples. A 2 mg protein extracted from epicardial adipose tissue of 16-week high fat diet mouse was digested and peptides were injected into ERLIC. In each ERLIC run, a total of 140 fractions were collected, but only eight fractions were finally submitted for LC-MS/MS analysis by combining twenty (early stage) or ten (late stage) consecutive fractions (Figure 1b) as we anticipated the peptide complexity to be substantially reduced such that it was below the detection sensitivity of one-dimensional LC-MS/MS.

Identification and characterization of subproteome

Protein identification was achieved by using two search engines, SEQUEST and Tandem. We unambiguously identified 159 unique phosphoproteins matched with 75

phosphoproteins. Meanwhile, totally 802 unique peptides were detected with N-linked glycosylation sites validated by consensus sequence N-X-S/T, assigned to 313 glycoproteins. In addition, only 219 unique glycopeptides and 96 glycoproteins were recovered from the hydrazide resin bead. Certainly, ERLIC showed the better performance in glycoproteome enrichment much more than the common widely-used hydrazide approach.

Figure 2 presented the distribution of subproteome through fractions. It clearly showed that, phosphopeptides and glycopeptides were eluted later with fewer identification in the earlier fractions (less than 1% in fraction 1 and 2) and more in the later fractions (70–90 % specificity in fraction

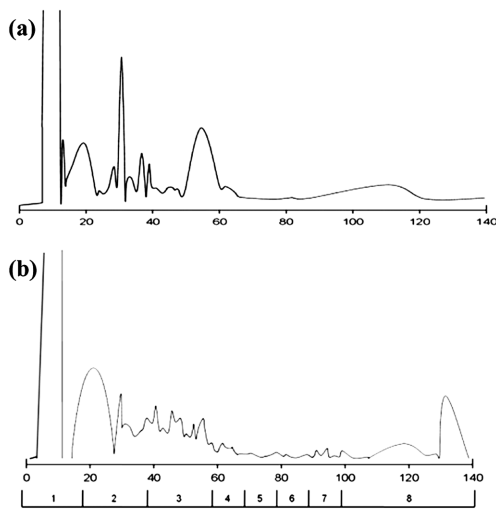


Figure 1. ERLIC chromatogram of purified proteins (a) and 2 mg tryptic digest of obesity mouse epicardial adipose tissue (b) with 140- minute gradient, detected at 280 nm.

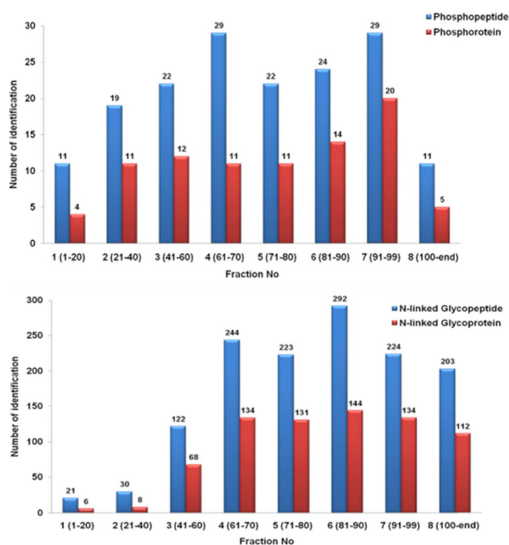


Figure 2. Fraction distribution of Phospho- (upper) and Glyco- (lower) Proteome.

3–8). The results confirmed that these two modified peptides had an elution profile different from intact peptides under two-buffer elution system.

In addition, the identified were categorized according to their subcellular locations using GeneGO program. The venn diagram in Figure 3 showed that most of phosphoproteins exist in cytoplasm (67.8%) and nucleus (11.86%), whereas higher abundance glycoproteins distributed at cytoplasm (39.92%), extracellular membrane (34.48%) and extracellular region (18.58%). Study of both subproteomes might suggest some insights about correlation between intra- and extracellular signaling pathways.

Simultaneous phosphorylated and N-linked glycosylated proteins

Notably, among identified results, there are four proteins listed with both phosphorylated and glycosylated sites, demonstrating the potential of ERLIC method for simultaneous characterization of post-translational modifications in a single protein, for example, alpha-2-hs-glycoproteins (Ahsg), Low-density lipoprotein receptor-related protein 2 (Lrp2), myosin-9 and serum deprivation response protein. Of these, phosphorylation and glycosylation occurs at the different sites of proteins results in different modified peptides observed. These results are in very good agreement with known facts regarding the identified may have

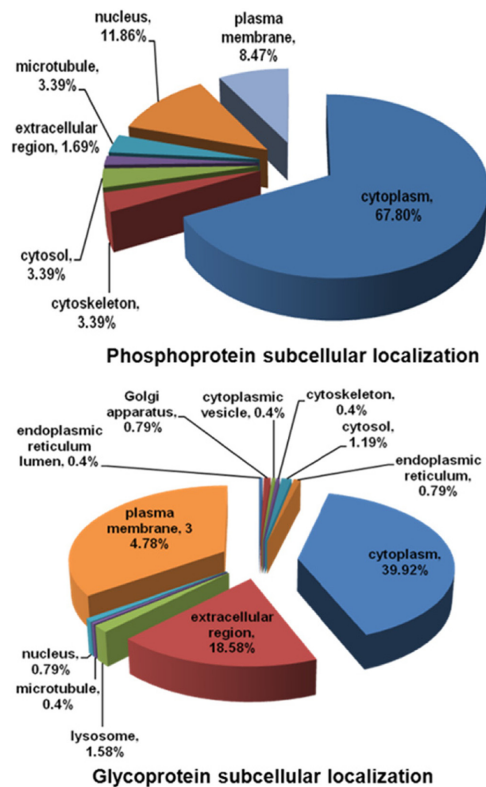


Figure 3. Subcellular localization of identified Phospho- (upper) and Glyco- (lower) proteins.

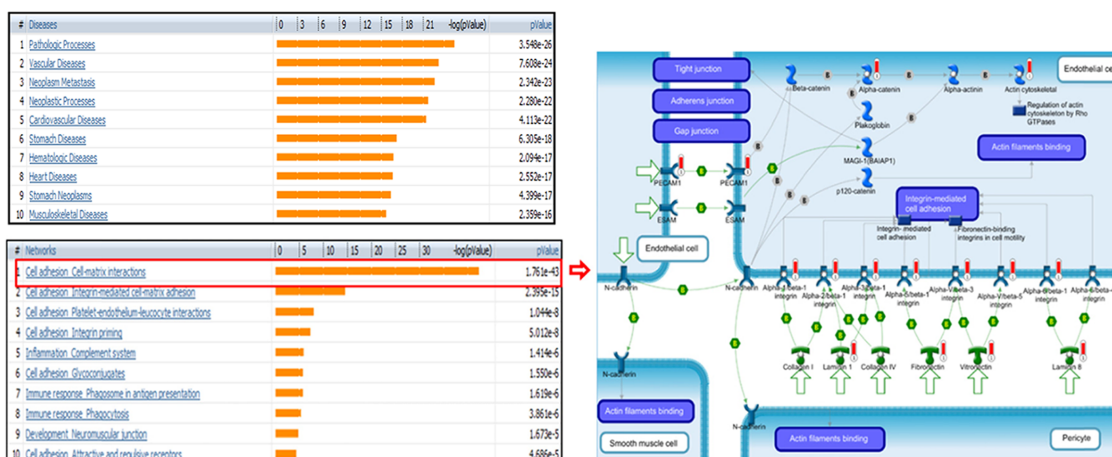


Figure 4. Top-scored diseases (upper) and GO processes (lower-left panel) related to the identified phospho- and glyco proteins. The top scored GO process Cell adhesion_Cell matrix interactions (right panel) highly correlative with the identified proteins (red thermometers).

roles in physiological processes in obesity biospecimen.^{16,17}

Functional annotation of proteins involving adipose tissue of obesity sample

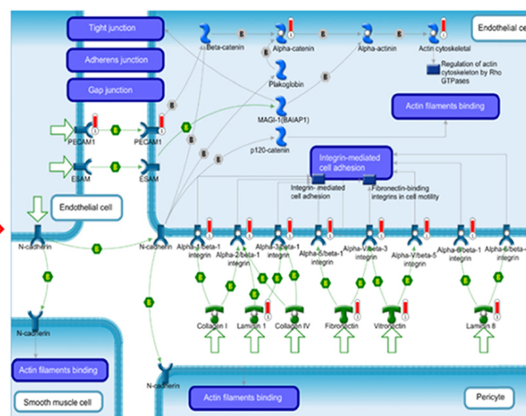
From the list of the identified, we performed functional studies of phospho- and glyco- proteins of epicardial adipose tissue from high fat diet mouse using MetaCore software. The top-scored diseases correlative with the identified showed the high relevance with obesity syndromes, significantly expressed with vascular diseases, cardiovascular diseases or heart diseases (Figure 4). In addition, the top-scored of GO process networks were built with the highest probability of Cell adhesion_Cell- matrix interactions which associated with pathological impairment of the structure and function of adipose cells in obesity phenotype, results in the initial stage of pathologic processes.¹⁸

Conclusions

ERLIC is a simple and robust approach for the recovery of not only glycoproteins but also phosphoproteins. The combination of both hydrophilic and electrostatic repulsion modes enabled this method to comprehensively identify PTMs from complex sample in one run experiment without any further clean up steps. Moreover, glycopeptides were collected from ERLIC fraction with their attached glycan, which could permit further glycan analysis. We expect that ERLIC could help to uncover more phospho- and glycoproteins and novel PTM sites from biomarker discovery.

Acknowledgements

This study was supported by grants of A11218-11-CP02 from the National Project for Personalized Genomic Medicine, Korea Health 21 R&D Project, Ministry of Health and



Welfare, Republic of Korea; R322011000102150 from the World Class University Project; 20110002217, 2011K00018, 20110021054 from the National Research Foundation.

References

- Cohen, P. *Trends Biochem. Sci.* **2000**, *25*, 596.
- Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochim. Biophys. Acta* **1999**, *1473*, 4.
- Dennis, J. W.; Granovsk, M.; Warren, C. E. *Bio Essays* **1999**, *21*, 412.
- Durand, G.; Seta, N. *Clin. Chem.* **2000**, *46*, 795.
- Cohen, P. *Eur. J. Biochem.* **2001**, *268*, 5001.
- Hood, L. *Mech. Ageing Dev.* **2003**, *124*, 9.
- Aebersold, R.; Cravatt, B. F. *Trends Biotechnol.* **2002**, *20*, 1.
- Ballif, B. A.; Villen, J.; Beausoleil, S. A.; Schwartz, D.; Gygi, S. P. *Mol. Cell Proteomics* **2004**, *3*, 1093.
- Lewandrowski, U.; Zahedi, R. P.; Moebius, J.; Walter, U.; Sickmann, A. *Mol. Cell Proteomics* **2007**, *6*, 1933.
- Gan, C. S.; Guo, T.; Zhang, H.; Lim, S. K.; Sze, S. K. *J. Proteome Res.* **2008**, *7*, 4869.
- Alpert, A. J. *Anal Chem.* **2008**, *80*, 62.
- Lewandrowski, U. L. K.; Zahedi, R.; Walter, D.; Sickmann, A. *Clin. Proteomics* **2008**, *4*, 25.
- Zhang, H.; Guo, T.; Li, X.; Datta, A.; Park, J. E.; Yang, J.; Lim, S. K.; Tam J. P.; Sze, S. K. *Mol. Cell Proteomic* **2010**, *9*, 635.
- Hao, P.; Guo, T.; Sze S. K. *Plos One* **2011**, *6*, 1.
- Tian, Y.; Zhou, Y.; Elliott, S.; Aebersold, R.; Zhang, H. *Nat. Protoc.* **2007**, *2*, 334.
- Ahmed, M.; Neville M. J.; Edelman, M. J.; Kessler, B. M.; Karpe, F. *Obesity (Silver Spring)* **2010**, *18*, 27.
- Schaub, N. P.; Jones, K. J.; Nyalwidhe, J. O.; Cazares, L. H.; Karbassi, I.D.; Semmes, O. J.; Feliberti, E. C.; Perry, R. R.; Drake, R. R. *J. Am. Coll. Surg.* **2009**, *208*, 970.
- Marecková, Z.; Heller, S.; Horký, K. *Vnitř. Lek* **1999**, *45*, 46.