Talanta 170 (2017) 509–513

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Quantification of N-glycosylation site occupancy status based on labeling/label-free strategies with LC-MS/MS

Shu Zhang[a,1], Wei Li[b,1], Haojie Lu[b,c,*], Yinkun Liu[a]

[a] Liver Cancer Institute, Zhongshan Hospital, Fudan University, Key Laboratory of Carcinogenesis and Cancer Invasion, Ministry of Education, Shanghai 200032, China
[b] Institutes of Biomedical Sciences and Department of Chemistry, Fudan University, Shanghai 200032, China
[c] Key Laboratory of Glycoconjugates Research, Ministry of Public Health, Fudan University, Shanghai 200032, China

ARTICLE INFO

Keywords:
N-glycosylation
Macroheterogeneity
Site occupancy
Quantification
Mass spectrometry

ABSTRACT

Protein N-glycosylation plays important roles in physiological and pathological processes. Characterizing the site-specific N-glycosylation including N-glycan macroheterogeneity (glycosylation site occupancy) and microheterogeneity (site-specific glycan structure) is important for understanding of glycoprotein biosynthesis and function. N-Glycan macroheterogeneity is a physiological property of glycoprotein and the technical obstacles have restricted research into the regulation and functions of this heterogeneity. Quantification of N-glycosylation site occupancy would uncover the critical role of macroheterogeneity in a variety of biological properties. Liquid chromatography (LC)- mass spectrometry (MS)-based quantification is emerging as a powerful tool for glycosylation characterization. This review summarizes the labeling and label-free quantitative MS approaches for quantifying N-glycosylation site occupancy, including its quantification for target glycoproteins in recent years.

1. Introduction

Protein glycosylation is a common post-translational modification (PTM), which plays an important role in many physiological and pathological processes, including protein folding and trafficking, cell-cell and cell-matrix interaction, cellular differentiation, fertilization, immune response, the initiation and metastasis of tumors [1,2]. It is well-known that N-glycosylation and O-glycosylation are the two main forms of protein glycosylation. N-glycans are attached to asparagine residue with the sequence of Asn-X-Ser/Thr (where X represents any amino acid except Pro) by an N-glycosidic bond, GlcNAc β1-Asn. All N-linked glycans have the common core pentasaccharide, Man3GlcNAc2 (Fig. 1 and Table 1).

Accurate characterization of glycoproteins with multiple N-glycosylation sites and assessment of the N-glycan macroheterogeneity (glycosylation site occupancy) and microheterogeneity (site-specific glycan structure) are needed for understanding of the functions of glycans. Glycosylation site occupancy reflects the fraction of proteins that are glycosylated and thus reveals site availability, enzyme kinetics, and substrate concentrations in the ER (endoplasmic reticulum) [3]. The important cellular functions of glycoproteins are influenced by site occupancy, and defects in site occupancy may be detrimental to physiological functions [4]. Microheterogeneity arises from differences in monosaccharide composition, linkage between monosaccharides, anomeric state, branching structures, and other substitutions [5]. In this review, we focus on quantification of N-glycosylation site occupancy, which could be physiological, affected by diseases or modified by heterologous protein overexpression.

N-glycosylation site occupancy is the proportion of a specific glycosylation site that is actually glycosylated. Alternatively, a given site is modified to any extent [6], which means quantification of site-specific glycosylated level to its parent protein level. LC-MS has become widely used for site occupancy analyses in complex or clinical samples. Developments in LC-MS-based technologies have considerably advanced quantification of N-glycosylation site occupancy, in which labeling and label-free methods are commonly applied [7]. Labeling methodologies such as SILAC (Stable isotope labeling by amino acids in cell culture), dimethylation, TMT (Tandem Mass Tags) and iTRAQ (Isobaric tags for absolute and relative quantification) are based on similar principle, the corresponding 'heavy' and 'light' isotopes or isobaric tags are introduced into different samples, respectively. The quantitative results could be obtained by comparing the abundance of the isotopologues [8]. Label-free methods have the advantage of experimental simplicity in sample preparation. Its quantification is

* Corresponding author at: Institutes of Biomedical Sciences and Department of Chemistry, Fudan University, Shanghai 200032, China.
E-mail address: luhaojie@fudan.edu.cn (H. Lu).
[1] Shu Zhang and Wei Li contributed equally to this work.

http://dx.doi.org/10.1016/j.talanta.2017.04.053
Received 6 February 2017; Received in revised form 18 April 2017; Accepted 21 April 2017
Available online 23 April 2017
0039-9140/

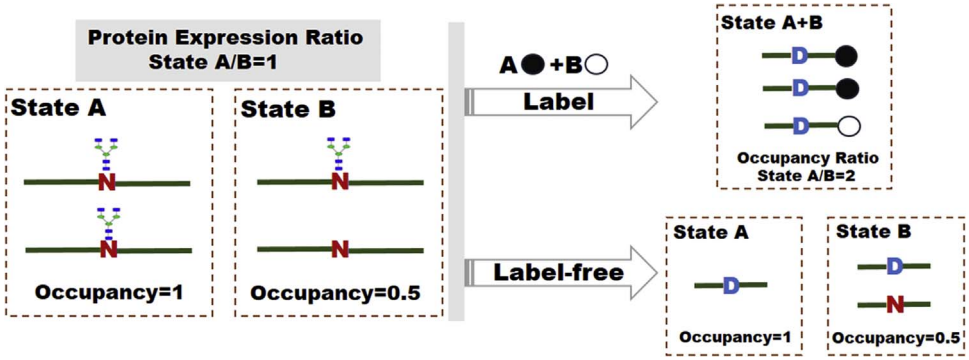

Fig. 1. MS-based labeling and label-free technologies for quantification of N-glycosylation site occupancy. In this schematic, protein expressions of one protein in state A and state B are equal hypothetically with different N-glycosylation site occupancies. Glycosidase treatment are often used to convert glycosylated forms (N-glycans attached to Asn) to their deglycosylated forms (Asp without N-glycans). Labeling quantification usually obtains occupancy ratio of this protein by comparing isotopically labeled forms of state A and B. Label-free quantification could obtain occupancy of this protein in state A or B by analyzing A or B independently. *N is the abbreviation of Asn, D is the abbreviation of Asp.

Table 1
Summary of progress in N-glycosylation site occupancy status with LC-MS/MS.

Quantification	Principle	Advantage	Disadvantage	Application examples (Target glycoprotein)
Labeling	Fold change of the deglycosylated peptide to the total protein levels or the average of all non-glycosylated peptides	Multiplexed analysis	Fold change often obtained without absolute occupancy	Hp
Label-free	The abundance of the deglycosylated peptide as a fraction of the sum of the deglycosylated peptide and unmodified versions of the same peptide	Simple workflow	Strict experimental pipeline for reproducibility	GM-CSF TNK-tPA CBG ccRCC hFXI ITIH4
MRM	The abundance of the deglycosylated peptide as a fraction of the sum of the deglycosylated peptide and unmodified versions of the same peptide	Absolute quantification of occupancy for target glycoprotein	Synthesis of isotopically labeled standard peptides	transferrin α_1-antitrypsin hPC Hp

either based on spectral counting or peptide-ion intensity [9]. We will give an overview of the quantification of N-glycosylation site occupancy using the two types of approaches. Apart from these, N-glycosylation site occupancy of target glycoproteins will be presented and discussed.

2. Labeling quantification of N-glycosylation site occupancy

Stable isotopes labeling methods in quantitative proteomics could help provide accurate information of N-glycosylation occupancy, mainly including enzyme labeling, metabolic labeling and chemical labeling. The significant advantage of labeling strategies is that simultaneous and efficient analysis of differentially labeled samples in a single experiment, which could remove signal variability associated with run-to-run inconsistencies [10].

2.1. Enzyme labeling

Trypsin catalyzed ^{18}O labeling leads to exchange of two ^{16}O atoms for two ^{18}O atoms at the C-terminal carboxyl group of tryptic peptides. A mass shift of 4 Da between singly charged, differentially labeled peptides could be observed in MS1 spectra [11–13]. PNGase F catalyzed ^{18}O labeling yields a deglycosylated protein whose N-glycosylation site Asn residues are converted to Asp labeled with one ^{18}O atom at their β-carboxyl group during this digestion [14]. The spectra of glycopeptides with N-glycan shifted 2 Da after the PNGase F digestions in $H_2{}^{16}O$ and $H_2{}^{18}O$, respectively [15]. Liu et al. [16] combined trypsin catalyzed ^{18}O labeling and PNGase F catalyzed ^{18}O labeling together. Thus, for ^{18}O labeling samples, three ^{18}O atoms were introduced into N-deglycopeptides and two ^{18}O atoms into non-glycosylated peptide. The relative quantities of N-glycosylation site occupancy change were obtained by measuring the intensity ratios of $^{18}O/^{16}O$ for deglycosylated (6 Da) and for non-glycosylated (4 Da) peptides, respectively. Using this strategy, occupancies of 31 N-glycosylation sites were quantified and 6 significant changes in ovarian cancer.

2.2. Metabolic labeling

SILAC/AACT is based on *in vivo* incorporation of specific amino acids into all proteins. Cell lines are grown in media lacking a standard essential amino acid but supplemented with a non-radioactive, isotopically labeled form of that amino acid [17]. Parker et al. [18] used SILAC (Arg^0/Lys^0 and Arg^{10}/Lys^8) to label differentiated 3T3-L1 adipocytes treated with and without TNF-alpha. After tryptic digestion, HILIC (hydrophilic interaction liquid chromatography) enrichment, PNGase F digestion of these membrane proteins and LC-MS/MS analysis, they normalized the fold change of the deglycosylated peptide to the total protein levels and found the relative N-glycan occupancy of adipocyte membrane proteins subjected to TNF-alpha-induced IR remained largely unchanged. Sun et al. [19] determined absolute N-glycosylation occupancy of OVCAR-3 cell proteins by TM (tunicamycin) to inhibit N-linked glycosylation and SILAC technology. Proteins from untreated cells grown in heavy SILAC medium (Arg^{10}/Lys^6) were mixed with the same amount of proteins from the TM-treated cells grown in light SILAC medium (Arg^0/Lys^0). After tryptic digestion, hydrazide chemistry enrichment and PNGase F digestion, the deglycosylated peptides and the remaining non-glycosylated peptides were analyzed by LC-MS/MS, respectively. By calculating three TM-treated/native ratios (deglycosylated peptide, unoccupied glycosite-containing peptide, and the total protein) at each partially glycosylated site, which