Journal of Proteomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Proteomics

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

O-GlcNAcylation site mapping by (azide-alkyne) click chemistry and mass spectrometry following intensive fractionation of skeletal muscle cells proteins

Barbara Deracinois[a], Luc Camoin[b], Matthias Lambert[a], Jean-Baptiste Boyer[b], Erwan Dupont[a], Bruno Bastide[a], Caroline Cieniewski-Bernard[a,*]

[a] Univ. Lille, EA 7369 - URePSSS - Unité de Recherche Pluridisciplinaire Sport Santé Société, F-59000 Lille, France
[b] Aix Marseille Univ, CNRS, INSERM, Institut Paoli-Calmettes, CRCM, Marseille Protéomique, Marseille, France

ARTICLE INFO

Keywords:
O-GlcNAcylation
Click chemistry
Mass spectrometry
Post-translational modifications
Sites localization
Skeletal muscle cells
Fractionation

ABSTRACT

The O-linked-*N*-acetyl-D-glucosaminylation (O-GlcNAcylation) modulates numerous aspects of cellular processes. Akin to phosphorylation, O-GlcNAcylation is highly dynamic, reversible, and responds rapidly to extracellular demand. Despite the absolute necessity to determine post-translational sites to fully understand the role of O-GlcNAcylation, it remains a high challenge for the major reason that unmodified proteins are in excess comparing to the O-GlcNAcylated ones. Based on a click chemistry approach, O-GlcNAcylated proteins were labelled with azido-GalNAc and coupled to agarose beads. The proteome extracted from C2C12 myotubes was submitted to an intensive fractionation prior to azide-alkyne click chemistry. This combination of fractionation and click chemistry is a powerful methodology to map O-GlcNAc sites; indeed, 342 proteins were identified through the identification of 620 peptides containing one or more O-GlcNAc sites. We localized O-GlcNAc sites on proteins involved in signalling pathways or in protein modification, as well as structural proteins. Considering the recent role of O-GlcNAcylation in the modulation of sarcomere morphometry and interaction between key structural protein, we focused on proteins involved in the cytoarchitecture of skeletal muscle cells. In particular, several O-GlcNAc sites were located into protein-protein interaction domains, suggesting that O-GlcNAcylation could be strongly involved in the organization and reorganization of sarcomere and myofibrils.

Significance: O-GlcNAcylation is an atypical glycosylation involved in the regulation of almost all if not all cellular processes, but its precise role remains sometimes obscure because of the ignorance of the O-GlcNAc site localization; thus, it remains indispensable to precisely map the O-GlcNAcylated sites to fully understand the role of O-GlcNAcylation on a given protein. For this purpose, we combined extensive fractionation of skeletal muscle cells proteome with click chemistry to map O-GlcNAc sites without an *a priori* consideration. A total of 620 peptides containing one or more O-GlcNAc sites were identified; interestingly, several of them belong to low expressed proteins, in particular proteins involved in signalling pathways. We also focused on structural proteins in view of recent data supporting the role of O-GlcNAcylation in the modulation of sarcomere cytoarchitecture; importantly, some of the O-GlcNAc sites were mapped into protein-protein interaction domains, reinforcing the involvement of O-GlcNAcylation in the organization and reorganization of sarcomere, and in larger extent, of myofibrils.

1. Introduction

The O-*N*-acetyl-β-D-glucosaminylation, termed O-GlcNAcylation, is an atypical glycosylation corresponding to the transfer of a unique monosaccharide, the *N*-acetyl-β-D-glucosamine, on the hydroxyl group of serine and threonine amino acids of nuclear, cytosolic and mitochondrial proteins [1, 2]. The O-GlcNAcylation has emerged as a key regulator of several cellular processes such as transcription, translation, regulation of signalling pathways, degradative processes, subcellular localization of targets, and so on [1, 3–7]. Because of its involvement in nearly all if not all cellular processes, O-GlcNAcylation is nowadays clearly associated with the aetiology of several acquired

* Corresponding author at: URePSSS Unité de Recherche Pluridisciplinaire Sport, Santé, Société, EA7369 Laboratoire Activité Physique, Muscle, Santé, Eurasport - 413, avenue Eugène Avinée, 59120 Loos, France.
E-mail address: caroline.cieniewski-bernard@univ-lille.fr (C. Cieniewski-Bernard).

https://doi.org/10.1016/j.jprot.2018.07.005
Received 8 February 2018; Received in revised form 13 June 2018; Accepted 3 July 2018

diseases, in particular diabetes, neuro-degenerative disorders, cardio-vascular diseases or cancer [8].

The O-GlcNAcylated proteins bear similarities with the phosphorylated ones, in particular the reversibility of both processes since the phosphate and the GlcNAc moieties could be added and removed several times along the protein lifetime, and their turn-over is shorter than the protein backbone turn-over [9]. The O-GlcNAcylation rapidly emerged as a major cellular mechanism which could compete with phosphorylation in terms of modified proteins and their importance in cellular physiology. But in contrast of the plethora of kinases and phosphatases responsible of the phosphorylation/dephosphorylation process on specific proteins, a unique couple of antagonist enzymes (OGT/OGA) is involved in the O-GlcNAcylation process. While kinases recognize a consensus sequence, phosphorylation sites are easily predictable from primary sequence of a protein. However, no consensus sequence was clearly defined for OGT, but it appears that peptidic sequences modified by O-GlcNAcylation are enriched in small amino acids, with a proximal proline residue; these sequences also present preferential secondary structures such as loop and disorganised regions instead of α-helix and β-strand [10–15]. The O-GlcNAcylated sites could also correspond to phosphorylated ones; thus, many proteins are modified by both O-GlcNAc and phosphates groups, and these two post-translational modifications could compete to the same or to neighbouring sites [4, 16].

Despite the crucial role of O-GlcNAcylation in numerous cellular processes, the precise localization of O-GlcNAcylated sites remains an indispensable prerequisite for the fine understanding of its biological function. However, mapping the O-GlcNAcylated sites remains laborious but challenging, because of (i) the low stoichiometry of O-GlcNAcylation; (ii) the ion suppression of the modified peptide by the unmodified peptides present in large excess and (iii) the labile β bond between serine or threonine and the O-GlcNAc moiety which is broken during the CID (Collision-Induced Dissociation) fragmentation process, leading to loss of site information during mass spectrometry analysis [17, 18].

Therefore, new strategies such as enrichment of O-GlcNAc modified proteins and the use of other fragmentation processes like ECD (Electron Capture Dissociation), ETD (Electron Transfer Dissociation), or HCD (High-energy Collisional Dissociation), able to limit the O-GlcNAc loss during the fragmentation, have been developed [19–21]. Thus, enrichment step of O-GlcNAcylated proteins or tryptic peptides by immuno-affinity or affinity chromatography (in particular LWAC (lectin weak affinity chromatography)) were extensively used to identify O-GlcNAc sites [21–28], or combination of both approaches [29]. Moreover, chemical-based enrichment using BEMAD (β-Elimination followed by Michael Addition of Dithiothreitol (DTT)) was applied to O-GlcNAcylated proteins to overcome the O-GlcNAc loss during fragmentation, since labile peptide-GlcNAc bonds were substituted by no-labile peptide-DTT bonds [30–32]. New strategies based on a chemoenzymatic labelling with galactosyltransferase ($GalT^{Y289L}$) appeared during the last decade. Azido-modified *N*-acetyl-galactosamine is transferred on the hydroxyl moiety of a GlcNAc acceptor allowing the fixation of aminooxy-biotin, and in consequence, the purification of O-GlcNAc proteins or peptides with streptavidin beads [33–36]. Nevertheless, biotin fragmentation occurs during MS/MS fragmentation, leading to spectral analysis difficulties despite the stable GlcNAc (or more specifically LacNAc-biotin)-peptide bond [17]. Moreover, the hard condition of elution, necessary to cut the biotin-streptavidin bond, can lead to peptide modifications and to low yield of the O-GlcNAc-proteins/peptides purification [37]. To overcome these problems, UV cleavable biotin can be used, this approach presenting the advantage to obtain positive charged peptides, leading to enhanced ETD fragmentation and O-GlcNAc site localization [38–41]. It is worth to note that BEMAD and purification of biotin-labelled O-LacNAc peptides could be paired to localize O-GlcNAc sites [42–46].

We previously demonstrated that O-GlcNAcylation is an original and important regulator of skeletal muscle physiology [47, 48], in particular in the modulation of contractile activity [49–52] as well as the morphometry of the sarcomere [53], supported by the fact that numerous contractile and structural proteins are O-GlcNAcylated [50, 54]. Therefore, we developed an alternative strategy, specific, efficient and allowing purification of O-GlcNAc bearing proteins from skeletal muscle cells by the use, with minor changes, of the click chemistry methodology developed by Hahne and collaborators [55]. Briefly, O-GlcNAc bearing proteins were linked to agarose beads through azide-alkyne chemistry to enhance the enrichment of O-GlcNAcylated proteins. Proteolytic digestion of proteins linked to agarose beads and mass spectrometry analysis of resulting peptides provided the global identification of O-GlcNAcylated proteins. Analysis of released O-LacNAc peptides from agarose beads, resulting from a chemical cut-off of glycosidic bond, led to localization of the modified amino acids. To extensively map O-GlcNAc sites on proteins, we proposed herein an intensive fractionation of the muscle cell proteome according to solubility, hydrophobicity and isoelectric point of proteins prior to the click chemistry. Thus, the method of click chemistry was achieved (i) on whole proteome extracted from C2C12 differentiated myotubes, (ii) on a subproteome, the cytosol-enriched extract, and (iii) on the cytosol-enriched extract extensively fractionated. The non-glycosylated peptides, and the glycosylated peptides released by beta-elimination, were analysed on mass spectrometry. Through the analysis of peptides retained on agarose beads, we identified 342 O-GlcNAcylated proteins in the fractionated subproteome, corresponding to a 2-fold increase of the number of identified proteins from the whole extract, or a 3.5-fold increase of identified proteins from the non-fractionated subproteome, which reinforce the strength of the fractionation. Among these O-GlcNAcylated proteins, we also identified 620 peptides containing one or several dehydrated serine or threonine amino acids, corresponding so to O-GlcNAcylated sites.

2. Experimental section

2.1. Materials

Heat-inactivated horse serum (HI-HS), foetal bovine serum (FBS) and Dulbecco's Modified Eagle's Medium (DMEM) were purchased from GIBCO; antibiotic-antimycotic, GlycoProfile β-elimination kit and all analytical grade reagents from Sigma-Aldrich; anti-proteases (Complete EDTA-free) and anti-phosphatases (Phos-Stop) from Roche Diagnostic; Bradford and RC DC™ reagents protein assays, dithiothreitol, iodoacetamide, and stain-free gels from Bio-Rad; Zeba spin column from Thermo Fisher Scientific; IPG buffer from GE Healthcare; Click-It™ reagents from Molecular Probes; trypsin/Lys-C mix from Promega; λ phosphatase and calf intestine phosphatase from New England Biolabs; C18 reversed-phase columns from Grace.

2.2. Cell culture

Mouse C2C12 skeletal myoblasts were obtained from ATCC (American Type Culture Collection). Myoblasts were grown on 100 mm Petri Dishes in proliferation medium (DMEM supplemented with 10% FBS and 1% antibiotic-antimycotic) at 37 °C in a humidified atmosphere of 5% CO_2 still reaching 90–95% confluence. They were then induced to differentiate into myotubes by switching to differentiation medium (DMEM containing 2% HI-HS and 1% antibiotic-antimycotic). Medium was changed every two-days, and myotubes were maintained for 5 days until they were mature.

2.3. Cell harvesting and protein extraction

C2C12 myotubes were rinsed three times with cold PBS and scraped with 2 × 200 µL of cold lysis RipA buffer (10 mM Tris/HCl, pH 7.4; 150 mM NaCl; 1 mM EDTA; 1% Triton X-100; 0.5% sodium