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

Barbara Deracinois[1], **Luc Camoin**[2], **Matthias Lambert**[1], **Jean-Baptiste Boyer**[2], **Erwan Dupont**[1], **Bruno Bastide**[1] **and Caroline Cieniewski-Bernard**[1, *]

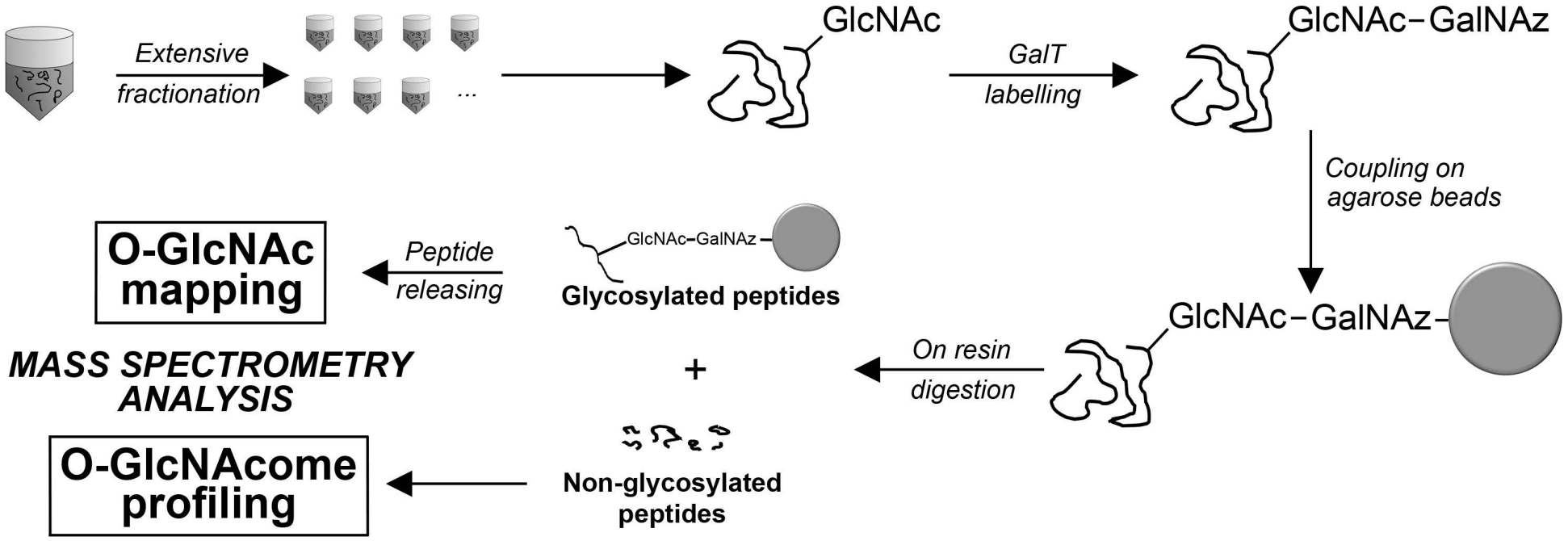
1 Univ. Lille, EA 7369 - URePSSS - Unité de Recherche Pluridisciplinaire Sport Santé Société, F-59000 Lille, France

2 Aix Marseille Univ, CNRS, INSERM, Institut Paoli-Calmettes, CRCM, Marseille Protéomique, Marseille, France

*To whom correspondence should be addressed:
Dr Caroline Cieniewski-Bernard
URePSSS Unité de Recherche Pluridisciplinaire Sport, Santé, Société
EA7369 Laboratoire Activité Physique, Muscle et Santé
Eurasport - 413, avenue Eugène Avinée
59120 Loos - France
caroline.cieniewski-bernard@univ-lille.fr

KEYWORDS

O-GlcNAcylation; click chemistry; mass spectrometry; post-translational modifications; sites localization; skeletal muscle cells; fractionation.


Extensive
fractionation
...
GlcNAc
GalT
labelling
GlcNAc–GalNAz
Coupling on
agarose beads
GlcNAc–GalNAz
On resin
digestion
+
GlcNAc–GalNAz
Glycosylated peptides
Peptide
releasing
O-GlcNAc
mapping
MASS SPECTROMETRY
ANALYSIS
Non-glycosylated
peptides
O-GlcNAcome
profiling

HIGHLIGHTS

- Combination of fractionation and click chemistry as a powerful methodology to map O-GlcNAc sites
- O-GlcNAc sites mapping on proteins involved in signalling pathways and in protein post-translational modifications
- O-GlcNAc sites mapping on structural proteins
- O-GlcNAcylation could by strongly involved in protein-protein interaction
- O-GlcNAcylation is a key regulator or sarcomeric organization and reorganization

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 organisation and reorganisation of sarcomere, and in larger extent, of myofibrils.

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 organisation and reorganisation 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 organisation and reorganisation of sarcomere, and in larger extent, of myofibrils.

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 diseases, in particular diabetes, neuro-degenerative disorders, cardiovascular 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.

EXPERIMENTAL SECTION

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.

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.

Cell harvesting and protein extraction

C2C12 myotubes were rinsed three times with cold PBS and scraped with 2x200 µl of cold lysis RipA buffer (10 mM Tris/HCl, pH 7.4; 150 mM NaCl; 1 mM EDTA; 1% TritonX-100; 0.5% sodium deoxycholate; 0.1% SDS) for whole extraction, or 2x150 µl of cold CMM buffer (50 mM Tris/HCl, pH 7.4; 2 mM EDTA; 5 mM EGTA; 5mM DTT; 0.05% saponin (w/v)) to get cytosol-enriched fraction, as previously described (56). Both lysis buffers were supplemented with 50µM PUGNAc (O-(2-acetamido-2-deoxy-D-glucopyranosilidene)amino-N-phenyl-carbamate), and proteases and phosphatases inhibitors. Collected materials were sonicated using Ultra-Sonic Cell Disruptor (Vibra Cell, 130 W ultrasonic processor); three pulses were performed at 70% of amplitude. Samples were then homogenized with gentle agitation for 1 h at 4 °C for RipA buffer in order to obtain the whole extract (WE), or centrifuged (14,000 g, 4°C, 30 min) after lysis with CMM buffer in order to obtain the enriched-cytosolic fraction in the supernatant. The residual pellet was then resuspended in the CMM buffer added with 1% Triton X-100, homogenised, centrifuged; the supernatant, corresponding to membrane-enriched fraction, was kept. The Triton-insoluble pellet, corresponding to the myofilament-enriched fraction, was resuspended in RipA buffer. The protein content of all fractions (WE, cytosolic,

membrane and myofilament fractions) was assayed using Bradford's method.

Fractionation of the cytosolic extract

Ammonium sulphate precipitation. One hundred milligrams of the protein extract were fractionated through 3-steps of ammonium sulphate (AS) precipitation. Briefly, extract was salted out with AS at 25% saturation for 2 h at 4°C and centrifuged (10000 g, 4°C, 15 min). The pellet, which corresponds to the fraction AS25, was stored, while the supernatant was salted out with 50% saturation in AS. The third step corresponded to a precipitation with AS at 75% saturation. The three fractions obtained corresponded to AS25, AS50 and AS75 fractions, respectively. The pellets were solubilized by UTCD buffer (4M urea; 2M thiourea; 2% CHAPS (w/v); 5mM DTT), desalted with Zeba Spin columns, and assayed using reducing agent and detergent compatible protein assay (RC DC™ protein assay).

Liquid phase IEF. Each fraction resulting from ammonium sulphate precipitation was submitted to liquid phase IEF using the MicroRotofor cell equipped with the PowerPac HV power supply, according to manufacturer (Bio-Rad). Protein samples (36 mg in total by fraction separated in 12 samples of 3 mg) were diluted at 1 mg/ml in UTCD buffer added with 3% ampholyte-IPG buffer pH 4-7, and loaded to pre-assembled focusing chamber of the MicroRotofor cell. The ion exchange membranes, separating the electrode reservoirs and the focusing chamber, were equilibrated in 0.1 M H_3PO_4 and 0.1 M NaOH for cation exchange membrane and anion exchange membrane, respectively. The focusing run was performed with constant power at 1W at 10°C. The separation typically occurred for 2.5 h and was terminated 30 minutes after the voltage stabilized. After focusing, 10 fractions of 200 µl (annotated F1 to F10 from the more acidic fraction to the more basic fraction) were quickly collected to avoid diffusion with the use of a vacuum. The pH of the individual fractions was measured with a pH meter, and fractions were kept at -20°C before further analysis.

One-dimensional polyacrylamide gel electrophoresis (1D-PAGE)

Twenty micrograms of each fraction (of 10 µl in case of IEF fractionation) were boiled in Laemmli buffer (62.5 mM Tris/HCl, pH 6.8; 10% glycerol; 2% SDS; 5% ß-mercaptoethanol; 0.02% bromophenol blue) and separated electrophoretically on 7.5% or on Any kD Mini-PROTEAN TGX Stain-Free™ (SF) Precast Gels (25 min, 300 V). The SF imaging was performed with ChemiDoc MP Imager and Image Lab 4.0.1 software (Bio-Rad); a 5-min activation time was used for the whole protein pattern imaging.

O-GlcNAc proteins enrichment

Click chemistry (azide/alkyne click reaction and enrichment) was performed on whole extract as well as on cytosol-enriched extract using the Click-iT™ O-GlcNAc Enzymatic Labelling System and the

Click-iT™ Protein Enrichment Kit according to the manufacturer's instructions and to the protocol described by Hahne and coworkers (55). The click chemistry protocol was also applied on each fraction issued from the MicroRotofor runs (themselves issued from ammonium sulphate precipitation of cytosol-enriched proteins); a total of 30 fractions were labelled. It is worth to note that for each sample, the same protein quantity (*i.e.* 2 mg) was used for the click chemistry-based enrichment.

After chloroform/methanol precipitation, performed at room temperature, the O-GlcNAc proteins were labelled overnight at 4°C with the Click-iT™ O-GlcNAc Enzymatic Labelling System. Briefly, Gal-T1 (Y289L) was incubated with proteins in labelling buffer (20 mM HEPES, pH 7.9; 50 mM NaCl; 2% NP-40; 5.5 mM $MnCl_2$; 25 µM UDP-GalNAz), according to manufacturer's recommendations. All reagents were provided in the kit, but the volume of each reagent was adjusted according to protein quantities. Reaction was performed at 4°C under gentle agitation for 20h, and then azide-labelled proteins were chloroform/methanol precipitated.

The azide-labelled proteins were then resuspended in urea lysis buffer, according to manufacturer's recommendations, following by an extensive homogenization. The proteins samples were then centrifuged at 10 000 rcf for 5 min in order to discard the insoluble material. The LacNAz-proteins were then precleared with agarose resin for 1 h, and the supernatant was then added with catalyst solution provided into the Click-iT™ Protein Enrichment Kit, containing copper (II) sulphate for a final concentration of 50 mM. The azide-labelled proteins were then incubated on an end-over-end rotator at room temperature for 18 h with 200 µL slurry of alkyne agarose resin according to the Click-iT™ Protein Enrichment Kit. After overnight click reaction, the supernatant was discarded and the resin was subjected to washes with 3 x 1.5 mL 10 mM diethylene triamine pentaacetic acid (DPTA) and 2 x 1.5 mL of MS grade water. Following reduction (10 mM dithiothreitol [DTT], 30 min, 55 °C) and alkylation (50 mM iodoacetamide [IAA], 60 min, RT), the resin was subjected to an extensive washing procedure performed in column as follow: 5 x 2 mL SDS wash buffer (100 mM Tris/HCl, pH 8; 1% SDS; 250 mM NaCl; 5 mM EDTA); 5 x 2 mL urea buffer (8 M urea; 100 mM Tris/HCl, pH 8); 10 x 2 mL 20% acetonitrile; 2 x 1 mL digestion buffer (100 mM Tris/HCl, pH 8.2; 2 mM $CaCl_2$; 10% ACN). Digestion of resin-bound proteins was performed o/n in 200 µL digestion buffer containing 1 µg of trypsin/Lys-C mix. Following the on-resin digestion, the remaining solution was discarded, and the resin washed with 500 µL of digestion buffer; both solutions, corresponding to non-retained peptides (NR peptides, *i.e.* the non-linked peptides), were pooled together and stored before desalting. The resin was then washed twice with 1.5 mL of MS grade water, following by 2 x 1.5 mL washes with dephosphorylation buffer (50 mM Tris/HCl, pH 7.6; 100 mM NaCl; 1 mM DTT; 10 mM $MgCl_2$; 1 mM $MnCl_2$).

Peptides-linked to agarose beads were submitted to dephosphorylation at 37 °C for 6 h in 400 µL of dephosphorylation buffer using 800 U λ phosphatase and 20 U calf intestine phosphatase. Following dephosphorylation, the resin was washed twice with 1.8 mL of H_2O, and the slurry volume was adjusted to 300 µL with H_2O before β-elimination with the GlycoProfile β-elimination kit. The β-elimination reaction was incubated on an end-over-end shaker with extensive mixing at 4 °C and quenched after 24

h with 1% TFA until pH6-8. Agarose beads were discarded, and the resulting solution contained the β-eliminated peptides which correspond to the initially O-GlcNAcylated peptides. The non-retained and the β-eliminated peptides were desalted with C18 reversed-phase columns and drained by vacuum concentrator before mass spectrometry analysis.

Mass spectrometry

Samples were reconstituted with 0.1% trifluoroacetic acid in 4% acetonitrile and analysed by liquid chromatography (LC)-tandem mass spectrometry (MS/MS) using a nanoLC Ultimate 3000 chromatography system (Dionex, Sunnyvale, CA) coupled to an LTQ-Orbitrap-Velos or a Q-Exactive mass spectrometer (Thermo Electron, Bremen, Germany). Peptides were separated on a Dionex Acclaim PepMap RSLC C18 column. First, peptides were concentrated and purified on a pre-column from Dionex (C18 PepMap100, 2 cm x 100 µm I.D, 100 Å pore size, 5 µm particle size) in solvent A (0.1% formic acid in 2% acetonitrile) for 5 min at 5µL/min. In the second step, peptides were separated on a reverse phase column from Dionex (C18 PepMap100, 15 cm x 75 µm I.D, 100 Å pore size, 2 µm particle size) at 300 nL/min flow rate. After column equilibration using 4% of solvent B (20% water - 80% acetonitrile - 0.1% formic acid), peptides were eluted from the analytical column by a two-steps linear gradient (4-20% acetonitrile/H_2O; 0.1 % formic acid for 90 min and 20-45% acetonitrile/H_2O; 0.1 % formic acid during 30 min for the LTQ-Orbitrap Velos and 4-25% acetonitrile/H2O; 0.1% formic acid for 40 min and 25-50% acetonitrile/H2O; 0.1% formic acid during 10 min for the Q-Exactive). For peptide ionisation in the nanospray source, spray voltage was set between 1.4 and 1.5 kV and the capillary temperature at 275 °C. The LTQ-Orbitrap-Velos was set up in data dependent mode to switch consistently between MS and MS/MS. MS spectra were acquired with the Orbitrap in the range of m/z 400-1700 at a FWHM resolution of 60 000 measured at 400 m/z and an automatic gain control (AGC) target value of 1e6. For internal mass calibration the 445.120025 ion was used as lock mass. The 20 abundant precursor ions were selected and collision-induced dissociation fragmentation with 35% normalized collision energy (NCE), and an AGC target value of 1e4 was performed in the ion trap on the 20 most intense precursor ions measured to have maximum sensitivity and yield a maximum amount of MS/MS data. The signal threshold for an MS/MS event was set to 500 counts. Charge state screening was enabled to exclude precursors with 0 and 1 charge states. Dynamic exclusion was enabled with a repeat count of 1, exclusion list size 500 and exclusion duration of 30 s. For Q-Exactive runs, survey MS scan were acquired in the Orbitrap on the 300-1700 m/z range with a resolution of 70 000 (AGC target at 1e6), the 12 most intense ions per survey scan (Intensity threshold 1e5) were selected for HCD fragmentation (AGC target 5e5, NCE 25%) and resulting fragments were analysed at a resolution of 17500 in the Orbitrap. Fragmented precursor ions were dynamically excluded for 25 s.

Protein identification

The acquired raw LC Orbitrap MS data were processed using Proteome Discoverer 1.4.1.14 (Thermo Fisher Scientific). This software was used to search data *via* in-house Mascot server (version 2.3.0; Matrix Science, London, U.K.) and SEQUEST HT against the Uniprot mouse reference proteome database (version 09.08.2013, 50,850 entries). The following parameters were used for searches: (i) trypsin; (ii) two missed cleavages were allowed; (iii) monoisotopic precursor tolerance of 8 ppm, followed by 0.8 Da for fragment ions from MS/MS; and (iv) cysteine carbamidomethylation (+57.0215) and methionine oxidation (+15.995) as variable modifications. The false discovery rate (FDR) was processed the built-in Percolator and was set to 5% (57). Protein identification from ß-elimination experiments was performed similarly as above for data processing with the exception of the following modifications. Dehydration of Ser and Thr (-18.011) as well as ß-elimination of Cys (-33.988 Da) was added as variable modification. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (58) partner repository with the dataset identifier PXD004860.

Bioinformatics resources and protein lists

Protein lists were compared using nwCompare software (59). All identified proteins were converted into gene names and classified into families and subfamilies of shared function, which were then categorized by molecular function and biological process ontology terms in the Protein Analysis Through Evolutionary Relationships (PANTHER) classification system (www.pantherdb.org) (60-62).

RESULTS

Fractionation of cytosol-enriched fraction from C2C12 myotubes

The workflow applied in our study was presented on Fig.1. Thus, whole proteins were extracted from C2C12 differentiated myotubes (Whole Extract, WE), or submitted to successive fractionation steps as described on Fig.1; the protein profiles of the resulting fractions were shown on Fig.2. Briefly, a differential protocol of extraction was applied to recover a cytosol-enriched fraction (CYT), a membrane-enriched fraction (MB) and a myofilament-enriched fraction (MYO); these fractions contained 64.2%, 18.9% and 16.9% of WE proteins, respectively (Fig.2A). It is noteworthy that the profile of cytosol-enriched fraction was quite similar than those corresponding to whole extract (Fig.2A). This fraction was then submitted to two successive fractionation protocols, while the membrane- and the myofilament-enriched fractions were discarded from the following analysis for the main reason that these fractions were poorly resolubilized following the chloroform/methanol precipitation.

The cytosol-enriched fraction was firstly partitioned through ammonium sulphate (AS) precipitation. Four fractions were obtained, annotated as AS25, AS50, AS75 and AS100 according to the final percentage of ammonium sulphate, as indicated on Fig.2B. The protein profile of each fraction was shown on Fig.2B; 37.9%, 43.9%, 17.4% and 0.8% of the non-fractionated CYT proteins were obtained for AS25, AS50, AS75 and AS100 fractions, respectively. Because of the poor protein content on AS100 fraction, it was discarded from the resting analysis. An additional step of fractionation was done on the AS25, AS50 and AS75 fractions, based on the separation of proteins according to their isoelectric point, using the MicroRotofor apparatus. Ten fractions were obtained in each case, containing on average 1.7%, 3.6%, 4.7%, 5%, 7%, 10.6%, 9.9%, 9.8%, 16.5% and 31.2% from F1 (the more acidic fraction) to F10 (the more basic fraction). The corresponding pH for each fraction was indicated on supplemented Table I. As shown on Fig.2C, the fractions F1 to F3 (pH 4.42 to 5.19) contained very few proteins comparing with the following fractions. In contrast, the neutral fractions (F7 to F10, pH 6.18 to 8.01) contained together approximately 59.7%, 68.5% and 73.8% of the AS25, AS50 and AS75 fractions, respectively.

At least, a total of 30 fractions were obtained for cytosol-enriched fractions (fCYT). Each of them, as well as non-fractionated cytosol-enriched fraction (CYT) and whole extract (WE), were submitted to click chemistry protocol in order to map the O-GlcNAc sites. These fractions, submitted to labelling-coupling protocol and analysed on mass spectrometry, were indicated in bold italic on Fig.1.

Efficiency of the O-GlcNAcylation mapping after extensive fractionation

The improvement of the fractionation on O-GlcNAcylation mapping was demonstrated in table I, which indicated the number of the identified peptides and the resulting identified proteins. Data corresponded to the number of peptides which were not linked to agarose beads (corresponding to non-O-GlcNAcylated peptides, but which belong to the O-GlcNAcylated proteins coupled to agarose beads), as well as the number of dehydrated peptides which were linked to agarose beads (corresponding to O-GlcNAcylated peptides covalently linked to agarose beads and released from beads by beta-elimination). Data were presented for the non-fractionated C2C12 extract (WE, whole extract), the non-fractionated cytosol-enriched extract (CYT), and the fractionated cytosol-enriched extract (CYT-AS-IEF combined fractions, corresponding to fCYT fraction). Based on these data, we determined the efficiency factor as the ratio of identified peptides/proteins after extensive fractionation (fCYT fraction) comparing with the non-fractionated cytosol-enriched extract (CYT).

We firstly analysed the non-linked peptides in order to identify the O-GlcNAcylated proteins linked to agarose beads, independently of their O-GlcNAc sites. In the non-fractionated whole extract, 554 proteins were identified, while 449 were identified in the cytosol-enriched fraction. Once the fractionation protocol applied, the number of identifications was increased by a factor 3 comparing with CYT fraction since 1362 proteins were identified in fCYT fraction. It is worth to note that 14573 peptides (containing 9.3% of dehydrated peptides) led to the identification of these 1362 O-GlcNAcylated proteins, while “only” 5540 peptides (containing 6.2% of dehydrated peptides) were identified in the non-fractionated cytosol-enriched fraction. Thus, three times more peptides were identified when extensive fractionation was applied on CYT fraction. Interestingly, we identified 620 contained one or more dehydrated serine or threonine (corresponding to O-GlcNAc sites) derived from the beta-elimination (BE) of peptides covalently linked to agarose beads; in parallel, 311 and 142 dehydrated peptides were identified in WE and CYT fractions. Thus, extensive fractionation of cytosol enriched-fraction increased the number of identified peptides bearing one or more O-GlcNAc site(s) by a factor 4.4 compared with non-fractionated cytosol-enriched fraction. It should be mentioned that among the total number of peptides identified in the beta-eliminated peptides fractions, some of them were not dehydrated, suggesting that a few peptides were non-specifically retained on agarose beads despite the pre-clearing step with agarose beads to avoid the non-specific retention of proteins on agarose beads, and despite the extensive washing steps applied to abrogate protein-protein interaction and so to eliminate the proteins which were not covalently linked to agarose beads. These non-dehydrated peptides were discarded from the rest of the analysis.

All data files corresponding to the identification of proteins and the identified peptides were presented as supplemental data (Supplemented Table II to VII, for WE_BE, WE_NR, CYT_BE, CYT_NR, fCYT_BE and fCYT_NR, respectively); all mass spectrometry data were deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository (58, 63) with the dataset identifier PXD004860.

Global analysis of O-GlcNAcylated proteins and focus on particular protein classes

As indicated in Table I, 342 proteins were identified consecutively to the identification of peptides released from agarose beads by beta-elimination, *i.e.* threefold than those identified in non-fractionated CYT fractions. We classified these proteins using the PANTHER classification system (Protein Analysis THrough Evolutionary Relationships, http://www.pantherdb.org/) (64, 65). Thus, proteins were classified according to their molecular function (Fig.3A) or the protein class (Fig.3B). Among the classified proteins according to their molecular function, 48% had a binding activity, 28.5% a catalytic activity, and 11.5% a structural molecule activity; the other proteins, representing less than 10%, had transporter, receptor, signal transducer, antioxidant or translation regulator activities (6%, 3.5%, 1.5%; 0.5% and 0.5%, respectively). Proteins were also classified according to the class they belong: nucleic acid metabolism (36%), enzymatic activities (30.5%), cytoskeletal proteins and chaperones (11.8%), signalling proteins (8.4%), transporter and binding proteins (5.4%), cell adhesion molecule (3%) or structural proteins (2.5%); 2.5% of proteins (others) were not assigned to the classes of proteins described just above.

We focused on proteins identified through the identification of beta-eliminated peptides containing one or several dehydrated serine and threonine. We mapped the O-GlcNAcylation site(s) (indicated in bold and underlined in peptide sequence) within different classes of proteins, in particular those involved in cytoskeleton and sarcomeric organization (Table II) or involved in signalling pathways, in protein degradation and in protein post-translational modifications (Table III). Table II combined (i) cytoskeleton proteins and proteins involved in cytoskeleton organization; (ii) actin cytoskeleton organization and reorganization; (iii) proteins involved in sarcomeric organization, motor activity and contractile fibers proteins; (iv) microtubules cytoskeleton organization and reorganization; (v) chaperones; (vi) intermediate filament proteins; and (vii) cell adhesion molecules. Table III combined (i) proteins involved in signalling or apoptotic pathways; (ii) proteins involved in protein modification (such as phosphorylation, ubiquitination, SUMOylation); and (iii) proteins involved in protein degradation or autophagy. Classification was done according to gene ontology annotation indicated in UniProtKB (http://www.uniprot.org) or GeneCards database (http://www.genecards.org/). We indicated the fraction (WE, CYT, fCYT) in which the peptides were identified. As shown on these tables, some peptides were identified in all fractions; in contrast, a large number of them were only identified when cytosolic-enriched extract was extensively fractionated. We also observed that in some case, within the same peptide, serine or threonine residues were differently dehydrated (for example, this was particularly observed for the Nascent polypeptide-associated complex subunit alpha, in Table III).

DISCUSSION

Mapping O-GlcNAc sites is a crucial point to characterize the role of O-GlcNAcylation on a given protein; indeed, the identification of the O-GlcNAc site within a specific domain of a protein could bring arguments to make hypothesis and determine strategies to clearly demonstrate and understand the role of the monosaccharide in the modulation of the function of a protein of interest. Strategies based on copper-catalysed azide/alkyne click chemistry emerged during the last decade; Hahne and collaborators improve and validate this method permitting the linkage of O-GlcNAcylated proteins to agarose beads to enrich the modified proteins, to identify them, and to map the O-GlcNAc sites (55). Beside the method, we opted for the enzymatic labelling of O-GlcNAcylated proteins rather than metabolic labelling since GlcNAz incorporation preferentially occurs in complex glycans (66). We added several steps of fractionation prior to the click chemistry. It is worth to note that we opted for fractionation according to ionic strength (using ammonium sulphate precipitation) or isoelectric point; of course, it could be envisaged to apply other protocols of fractionation to yield complementary data to complete the non-exhaustive list of O-GlcNAcylated peptides. From the analysis of linked peptides on agarose beads and released from beads by beta-elimination, we identified 342 proteins in the fractionated subproteome, corresponding to a 3.5-fold increase of identified proteins compared with a non-fractionated cytosol-enriched fraction. Indeed, through the fractionation protocol, based on the physicochemical properties of proteins, the sample is made less complex. Thus, there is an enrichment of certain proteins that could not have been identified in a complex mixture. As consequence, the number of identified peptides/proteins increased after fractionation. We also identified 620 peptides containing one or several O-GlcNAcylated sites. Over increasing the number of identified peptides/proteins, the fractionation protocol performed prior to click chemistry led to the mapping of O-GlcNAc sites on numerous signalling proteins, such as proteins involved in MAPK pathway, including the TGF-beta pathway. In addition, we identified several proteins involved in the ubiquitination process, in particular several E3-ubiquitin ligases, as well as proteins responsible of deubiquitination. Thus, our data suggest that modulation of ubiquitination through O-GlcNAcylation could be involved in the modulation of degradative process (and so in the regulation of protein homeostasis essential for muscle healthcare), as well as intracellular processes modulated by ubiquitination.

About fifteen years ago, we attempted to cartography the O-GlcNAcylated proteins in skeletal muscle and we identified structural proteins, proteins involved in signalling pathways and contractile proteins as being O-GlcNAcylated (54). Five O-GlcNAc sites were mapped on actin and myosin using a BEMAD approach, and interestingly, some of them were located into or close to protein-protein interaction domain, suggesting that O-GlcNAcylation could play an important role in the modulation of sarcomeric protein interaction (32). In this way, we recently demonstrated that O-GlcNAcylation was a key modulator of sarcomere morphometry, in particular through the modulation of protein-protein

interaction within multiprotein complexes including key structural proteins such as desmin, αB-crystallin, α-actinin, filamin-C and moesin (53). In addition, we showed that, consecutively to global changes of O-GlcNAcylation level, the interaction between desmin and αB-crystallin was modulated; in this paper, we localized the O-GlcNAc sites of these two proteins. For desmin, a protein of intermediate filament essential for striated muscle cells integrity, GlcNAcylation site was localized on the serine 459, in the C-terminal domain or tail domain (67), suggesting a key role of this O-GlcNAcylation on the function of the protein. Indeed, this domain is involved in the head to tail longitudinal assembly of desmin tetramers, and controls the lateral packing as well as the elongation and stabilization of mature desmin filaments (68). In addition, this domain is also involved in the interaction between desmin and cytoskeletal protein partners, leading to the formation of a highly complex network of intermediate filaments (69). Moreover, the mutation of this site (corresponding to serine 460 for human desmin) leads to particular myofibrillar myopathies, *i.e.* desminopathies, characterized by the disintegration of desmin filament, the accumulation of insoluble protein aggregates (including some of the protein partners of desmin such as αB-crystallin), and at terms, to the destructuration of myofibrils (70). In transfected SW13 human cells and C2C12 murine cells, the mutation S460I led to short desmin filaments, with irregular diameter, and numerous aggregates randomly distributed in cytoplasm (70).

Desmin integrity is regulated by chaperones; among them, the αB-crystallin plays a crucial role in the localization or stabilization of cytoskeleton element (71), and in particular, desmin filaments (72). Thus, the αB-crystallin modulates the polymerization of desmin filaments and prevent their aggregation (73, 74); in addition, it decreases the desmin interfilament interaction *in vitro* (75). The interaction between desmin and its molecular chaperone involved three domains on αB-crystallin : the β3 and β8 sheets, and the 155-165 C-terminal residues (76). We have localized the O-GlcNAc moiety on the threonine 162, *i.e*. in one of the interaction domain of αB-crystallin with desmin. Interestingly, the mutation of the C-terminal domain of the chaperone is closely associated to an increase of association and aggregation of desmin filaments (76). It is worth to note that on αA-crystallin and on HSPB1, O-GlcNAc sites were also localized on the C-terminal domain, which corresponds to a flexible conserved domain common to sHSPs, essential to maintain sHSP solubility, stability and chaperone activity (77). All together, these data strongly support that O-GlcNAcylation may be a key regulator of desmin-αB-crystallin interaction. Interestingly, it was demonstrated that stress induced changes of O-GlcNAcylation level of αB-crystallin, associated to translocation of αB-crystallin to cytoskeleton (71).

The role of O-GlcNAcylation in the maintenance of sarcomere cytoarchitecture and in modulation of interactions between structural and/or sarcomeric proteins was also supported by other data gained from this study. Indeed, on another chaperone involved in the maintenance of sarcomere integrity and in the turn-over of Z-line proteins (78, 79), BAG3 (Bcl-2-associated athanogene 3), two out of the four O-GlcNAcylated sites are located into the PxxP domain, which is known to interact with SH3-containing proteins (80). Titin, essential for assembling and maintenance of sarcomere, providing to myofibril their

property of elasticity (81), is O-GlcNAcylated on T24950 and S24954, localized in the mouse domain Kelch-12 (corresponding to the human domain Ig-like 123), which is the interaction domain with protein partners such as myosin or My-BPC (82). Plectin, a large intermediate filament-associated protein acting as cytoskeletal crosslinker and signalling scaffold, is O-GlcNAcylated into the plakin domain repeat B5 (83, 84), located at the C-terminal domain of plectin and described to bind to various intermediate filament proteins (85).

To sum, while our data demonstrated that the combination of fractionation and click chemistry was a powerful methodology to map O-GlcNAc sites on several proteins involved in signalling pathways or in protein modification (in particular ubiquitination), as well as structural proteins, it brought strong arguments in favour of the essential role of O-GlcNAcylation for the sarcomeric cytoarchitecture of skeletal muscle cells. Thus, we identified herein several key structural proteins involved in the maintenance of filament networks permitting force development and providing muscle cells mechanical strength. Moreover, we localized numerous O-GlcNAcylated sites in protein-protein interaction domains. In addition, key structural proteins such as desmin, αB-crystallin, plectin, or BAG3 among others are O-GlcNAcylated; importantly, mutation that can occur at the O-GlcNAcylated sites of these proteins led to the sarcomere disorganisation and myofibrils destructuration because of the loss of protein-protein interactions; as consequence, proteins aggregate, resulting at terms to myofibrillar myopathies.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ACKNOWLEDGMENT

This research was funded by the French National Research Agency (ANR, Agence Nationale de la Recherche, Young Researchers Program, n°11JSV8 006 01). Barbara Deracinois is a recipient of a fellowship from ANR. The proteomic analyses were done using the mass spectrometry facility of Marseille Proteomics (http://map.univmed.fr/) supported by IBISA (Infrastructures Biologie Santé et Agronomie), the Cancéropôle PACA, the Provence-Alpes-Côte d'Azur Region, the Institut Paoli-Calmettes and the Centre de Recherche en Cancérologie de Marseille. The mass spectrometry proteomics data, including search result files have been deposited to the ProteomeXchange Consortium (www.proteomexchange.org) via the PRIDE partner repository with the dataset identifier PXD004860.

REFERENCES

1. Hart, G. W., Three Decades of Research on O-GlcNAcylation - A Major Nutrient Sensor That Regulates Signaling, Transcription and Cellular Metabolism. *Front Endocrinol. (Lausanne)* **2014,** 5, 183.
2. Torres, C. R.; Hart, G. W., Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *J. Biol. Chem* **1984,** 259, (5), 3308-3317.
3. Banerjee, P. S.; Ma, J.; Hart, G. W., Diabetes-associated dysregulation of O-GlcNAcylation in rat cardiac mitochondria. *Proc. Natl. Acad. Sci. U. S. A* **2015,** 112, (19), 6050-6055.
4. Hardiville, S.; Hart, G. W., Nutrient regulation of signaling, transcription, and cell physiology by O-GlcNAcylation. *Cell Metab* **2014,** 20, (2), 208-213.
5. Hart, G. W.; Housley, M. P.; Slawson, C., Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* **2007,** 446, (7139), 1017-1022.
6. Hart, G. W.; Slawson, C.; Ramirez-Correa, G.; Lagerlof, O., Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu. Rev. Biochem* **2011,** 80, 825-858.
7. Harwood, K. R.; Hanover, J. A., Nutrient-driven O-GlcNAc cycling - think globally but act locally. *J. Cell Sci* **2014,** 127, (Pt 9), 1857-1867.
8. Bond, M. R.; Hanover, J. A., A little sugar goes a long way: the cell biology of O-GlcNAc. *J Cell Biol* **2015,** 208, (7), 869-880.
9. Hart, G. W.; Akimoto, Y., The O-GlcNAc modification. In *Essentials of Glycobiology, 2nd edition*, 2009; Vol. Chapter 18, p Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
10. Chen, Y. X.; Du, J. T.; Zhou, L. X.; Liu, X. H.; Zhao, Y. F.; Nakanishi, H.; Li, Y. M., Alternative O-GlcNAcylation/O-phosphorylation of Ser16 induce different conformational disturbances to the N terminus of murine estrogen receptor beta. *Chem. Biol* **2006,** 13, (9), 937-944.
11. Nishikawa, I.; Nakajima, Y.; Ito, M.; Fukuchi, S.; Homma, K.; Nishikawa, K., Computational prediction of O-linked glycosylation sites that preferentially map on intrinsically disordered regions of extracellular proteins. *Int. J. Mol. Sci* **2010,** 11, (12), 4991-5008.
12. Xie, H.; Vucetic, S.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Uversky, V. N.; Obradovic, Z., Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J. Proteome. Res* **2007,** 6, (5), 1882-1898.
13. Xie, H.; Vucetic, S.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Obradovic, Z.; Uversky, V. N., Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J. Proteome. Res* **2007,** 6, (5), 1917-1932.
14. Yang, W. H.; Kim, J. E.; Nam, H. W.; Ju, J. W.; Kim, H. S.; Kim, Y. S.; Cho, J. W., Modification of p53 with O-linked N-acetylglucosamine regulates p53 activity and stability. *Nat. Cell Biol* **2006,** 8, (10), 1074-1083.
15. Lazarus, M. B.; Nam, Y.; Jiang, J.; Sliz, P.; Walker, S., Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature* **2011,** 469, (7331), 564-567.
16. Zeidan, Q.; Hart, G. W., The intersections between O-GlcNAcylation and phosphorylation: implications for multiple signaling pathways. *J. Cell Sci* **2010,** 123, (Pt 1), 13-22.
17. Hu, P.; Shimoji, S.; Hart, G. W., Site-specific interplay between O-GlcNAcylation and phosphorylation in cellular regulation. *FEBS Lett* **2010,** 584, (12), 2526-2538.
18. Ma, J.; Hart, G. W., O-GlcNAc profiling: from proteins to proteomes. *Clin. Proteomics* **2014,** 11, (1), 8.
19. Hahne, H.; Kuster, B., A novel two-stage tandem mass spectrometry approach and scoring scheme for the identification of O-GlcNAc modified peptides. *J. Am. Soc. Mass Spectrom* **2011,** 22, (5), 931-942.
20. Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A., Localization of O-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. *Anal. Chem* **1999,** 71, (20), 4431-4436.
21. Vosseller, K.; Trinidad, J. C.; Chalkley, R. J.; Specht, C. G.; Thalhammer, A.; Lynn, A. J.; Snedecor,

J. O.; Guan, S.; Medzihradszky, K. F.; Maltby, D. A.; Schoepfer, R.; Burlingame, A. L., O-linked N-acetylglucosamine proteomics of postsynaptic density preparations using lectin weak affinity chromatography and mass spectrometry. *Mol. Cell Proteomics* **2006,** 5, (5), 923-934.
22. Ma, Z. Y.; Skorobogatko, Y.; Vosseller, K., Tandem lectin weak affinity chromatography for glycoprotein enrichment. *Methods Mol. Biol* **2013,** 951, 21-31.
23. Nagel, A. K.; Schilling, M.; Comte-Walters, S.; Berkaw, M. N.; Ball, L. E., Identification of O-linked N-acetylglucosamine (O-GlcNAc)-modified osteoblast proteins by electron transfer dissociation tandem mass spectrometry reveals proteins critical for bone formation. *Mol Cell Proteomics* **2013,** 12, (4), 945-955.
24. Chalkley, R. J.; Thalhammer, A.; Schoepfer, R.; Burlingame, A. L., Identification of protein O-GlcNAcylation sites using electron transfer dissociation mass spectrometry on native peptides. *Proc. Natl. Acad. Sci. U. S. A* **2009,** 106, (22), 8894-8899.
25. Trinidad, J. C.; Barkan, D. T.; Gulledge, B. F.; Thalhammer, A.; Sali, A.; Schoepfer, R.; Burlingame, A. L., Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Mol. Cell Proteomics* **2012,** 11, (8), 215-229.
26. Zhao, P.; Viner, R.; Teo, C. F.; Boons, G. J.; Horn, D.; Wells, L., Combining high-energy C-trap dissociation and electron transfer dissociation for protein O-GlcNAc modification site assignment. *J Proteome. Res* **2011,** 10, (9), 4088-4104.
27. Fardini, Y.; Perez-Cervera, Y.; Camoin, L.; Pagesy, P.; Lefebvre, T.; Issad, T., Regulatory O-GlcNAcylation sites on FoxO1 are yet to be identified. *Biochem. Biophys. Res Commun* **2015,** 462, (2), 151-158.
28. Xu, S. L.; Chalkley, R. J.; Maynard, J. C.; Wang, W.; Ni, W.; Jiang, X.; Shin, K.; Cheng, L.; Savage, D.; Huhmer, A. F.; Burlingame, A. L.; Wang, Z. Y., Proteomic analysis reveals O-GlcNAc modification on proteins with key regulatory functions in Arabidopsis. *Proc Natl Acad Sci U S A* **2017,** 114, (8), E1536-E1543.
29. Lee, A.; Miller, D.; Henry, R.; Paruchuri, V. D.; O'Meally, R. N.; Boronina, T.; Cole, R. N.; Zachara, N. E., Combined Antibody/Lectin Enrichment Identifies Extensive Changes in the O-GlcNAc Sub-proteome upon Oxidative Stress. *J Proteome Res* **2016,** 15, (12), 4318-4336.
30. Vosseller, K.; Hansen, K. C.; Chalkley, R. J.; Trinidad, J. C.; Wells, L.; Hart, G. W.; Burlingame, A. L., Quantitative analysis of both protein expression and serine / threonine post-translational modifications through stable isotope labeling with dithiothreitol. *Proteomics* **2005,** 5, (2), 388-398.
31. Wells, L.; Vosseller, K.; Cole, R. N.; Cronshaw, J. M.; Matunis, M. J.; Hart, G. W., Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. *Mol Cell Proteomics* **2002,** 1, (10), 791-804.
32. Hedou, J.; Bastide, B.; Page, A.; Michalski, J. C.; Morelle, W., Mapping of O-linked beta-N-acetylglucosamine modification sites in key contractile proteins of rat skeletal muscle. *Proteomics* **2009,** 9, (8), 2139-2148.
33. Dehennaut, V.; Lefebvre, T.; Leroy, Y.; Vilain, J. P.; Michalski, J. C.; Bodart, J. F., Survey of O-GlcNAc level variations in Xenopus laevis from oogenesis to early development. *Glycoconj. J* **2009,** 26, (3), 301-311.
34. Khidekel, N.; Arndt, S.; Lamarre-Vincent, N.; Lippert, A.; Poulin-Kerstien, K. G.; Ramakrishnan, B.; Qasba, P. K.; Hsieh-Wilson, L. C., A chemoenzymatic approach toward the rapid and sensitive detection of O-GlcNAc posttranslational modifications. *J. Am. Chem. Soc* **2003,** 125, (52), 16162-16163.
35. Khidekel, N.; Ficarro, S. B.; Clark, P. M.; Bryan, M. C.; Swaney, D. L.; Rexach, J. E.; Sun, Y. E.; Coon, J. J.; Peters, E. C.; Hsieh-Wilson, L. C., Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics. *Nat. Chem. Biol* **2007,** 3, (6), 339-348.
36. Tai, H. C.; Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C., Parallel identification of O-GlcNAc-modified proteins from cell lysates. *J. Am. Chem. Soc* **2004,** 126, (34), 10500-10501.
37. Kim, E. J., Chemical arsenal for the study of O-GlcNAc. *Molecules* **2011,** 16, (3), 1987-2022.
38. Alfaro, J. F.; Gong, C. X.; Monroe, M. E.; Aldrich, J. T.; Clauss, T. R.; Purvine, S. O.; Wang, Z.; Camp, D. G.; Shabanowitz, J.; Stanley, P.; Hart, G. W.; Hunt, D. F.; Yang, F.; Smith, R. D., Tandem mass spectrometry identifies many mouse brain O-GlcNAcylated proteins including EGF domain-specific O-GlcNAc transferase targets. *Proc. Natl. Acad. Sci. U. S. A* **2012,** 109, (19), 7280-7285.
39. Wang, Z.; Udeshi, N. D.; O'Malley, M.; Shabanowitz, J.; Hunt, D. F.; Hart, G. W., Enrichment and

site mapping of O-linked N-acetylglucosamine by a combination of chemical/enzymatic tagging, photochemical cleavage, and electron transfer dissociation mass spectrometry. *Mol Cell Proteomics* **2010,** 9, (1), 153-160.
40. Wang, Z.; Udeshi, N. D.; Slawson, C.; Compton, P. D.; Sakabe, K.; Cheung, W. D.; Shabanowitz, J.; Hunt, D. F.; Hart, G. W., Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. *Sci. Signal* **2010,** 3, (104), ra2.
41. Wang, S.; Yang, F.; Petyuk, V. A.; Shukla, A. K.; Monroe, M. E.; Gritsenko, M. A.; Rodland, K. D.; Smith, R. D.; Qian, W. J.; Gong, C. X.; Liu, T., Quantitative proteomics identifies altered O-GlcNAcylation of structural, synaptic and memory-associated proteins in Alzheimer's disease. *J Pathol* **2017,** 243, (1), 78-88.
42. Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C., Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain. *Proc. Natl. Acad. Sci. U. S. A* **2004,** 101, (36), 13132-13137.
43. Zeidan, Q.; Wang, Z.; De, M. A.; Hart, G. W., O-GlcNAc cycling enzymes associate with the translational machinery and modify core ribosomal proteins. *Mol. Biol. Cell* **2010,** 21, (12), 1922-1936.
44. Sakabe, K.; Wang, Z.; Hart, G. W., Beta-N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc. Natl. Acad. Sci. U. S. A* **2010,** 107, (46), 19915-19920.
45. Ramirez-Correa, G. A.; Jin, W.; Wang, Z.; Zhong, X.; Gao, W. D.; Dias, W. B.; Vecoli, C.; Hart, G. W.; Murphy, A. M., O-linked GlcNAc modification of cardiac myofilament proteins: a novel regulator of myocardial contractile function. *Circ. Res* **2008,** 103, (12), 1354-1358.
46. Wang, Z.; Park, K.; Comer, F.; Hsieh-Wilson, L. C.; Saudek, C. D.; Hart, G. W., Site-specific GlcNAcylation of human erythrocyte proteins: potential biomarker(s) for diabetes. *Diabetes* **2009,** 58, (2), 309-317.
47. Cieniewski-Bernard, C.; Montel, V.; Stevens, L.; Bastide, B., O-GlcNAcylation, an original modulator of contractile activity in striated muscle. *J. Muscle Res. Cell Motil* **2009,** 30, (7-8), 281-287.
48. Cieniewski-Bernard, C.; Lambert, M.; Dupont, E.; Montel, V.; Stevens, L.; Bastide, B., O-GlcNAcylation, contractile protein modifications and calcium affinity in skeletal muscle. *Front Physiol* **2014,** 5, 421.
49. Hedou, J.; Cieniewski-Bernard, C.; Leroy, Y.; Michalski, J. C.; Mounier, Y.; Bastide, B., O-linked N-acetylglucosaminylation is involved in the Ca2+ activation properties of rat skeletal muscle. *J. Biol. Chem* **2007,** 282, (14), 10360-10369.
50. Cieniewski-Bernard, C.; Montel, V.; Berthoin, S.; Bastide, B., Increasing O-GlcNAcylation Level on Organ Culture of Soleus Modulates the Calcium Activation Parameters of Muscle Fibers. *PLoS. One* **2012,** 7, (10), e48218.
51. Stevens, L.; Bastide, B.; Hedou, J.; Cieniewski-Bernard, C.; Montel, V.; Cochon, L.; Dupont, E.; Mounier, Y., Potential regulation of human muscle plasticity by MLC2 post-translational modifications during bed rest and countermeasures. *Arch. Biochem. Biophys* **2013,** 540, (1-2), 125-132.
52. Cieniewski-Bernard, C.; Dupont, E.; Richard, E.; Bastide, B., Phospho-GlcNAc modulation of slow MLC2 during soleus atrophy through a multienzymatic and sarcomeric complex. *Pflugers Arch* **2014,** 466, (11), 2139-2151.
53. Lambert, M.; Richard, E.; Duban-Deweer, S.; Krzewinski, F.; Deracinois, B.; Dupont, E.; Bastide, B.; Cieniewski-Bernard, C., O-GlcNAcylation is a key modulator of skeletal muscle sarcomeric morphometry associated to modulation of protein-protein interactions. *Biochim. Biophys. Acta - General Subjects* **2016,** 1860, (9), 2017-2030.
54. Cieniewski-Bernard, C.; Bastide, B.; Lefebvre, T.; Lemoine, J.; Mounier, Y.; Michalski, J. C., Identification of O-linked N-acetylglucosamine proteins in rat skeletal muscle using two-dimensional gel electrophoresis and mass spectrometry. *Mol. Cell Proteomics* **2004,** 3, (6), 577-585.
55. Hahne, H.; Sobotzki, N.; Nyberg, T.; Helm, D.; Borodkin, V. S.; van Aalten, D. M.; Agnew, B.; Kuster, B., Proteome wide purification and identification of O-GlcNAc-modified proteins using click chemistry and mass spectrometry. *J. Proteome. Res* **2013,** 12, (2), 927-936.
56. Yin, X.; Cuello, F.; Mayr, U.; Hao, Z.; Hornshaw, M.; Ehler, E.; Avkiran, M.; Mayr, M., Proteomics analysis of the cardiac myofilament subproteome reveals dynamic alterations in phosphatase subunit

distribution. *Mol. Cell Proteomics* **2010,** 9, (3), 497-509.
57. Brosch, M.; Yu, L.; Hubbard, T.; Choudhary, J., Accurate and sensitive peptide identification with Mascot Percolator. *J Proteome Res* **2009,** 8, (6), 3176-81.
58. Vizcaino, J. A.; Csordas, A.; Del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; Xu, Q. W.; Wang, R.; Hermjakob, H., 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* **2016,** 44, (22), 11033.
59. Pont, F.; Fournie, J. J., Sorting protein lists with nwCompare: a simple and fast algorithm for n-way comparison of proteomic data files. *Proteomics* **2010,** 10, (5), 1091-4.
60. Thomas, P. D.; Campbell, M. J.; Kejariwal, A.; Mi, H.; Karlak, B.; Daverman, R.; Diemer, K.; Muruganujan, A.; Narechania, A., PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* **2003,** 13, (9), 2129-41.
61. Thomas, P. D.; Kejariwal, A.; Campbell, M. J.; Mi, H.; Diemer, K.; Guo, N.; Ladunga, I.; Ulitsky-Lazareva, B.; Muruganujan, A.; Rabkin, S.; Vandergriff, J. A.; Doremieux, O., PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res* **2003,** 31, (1), 334-41.
62. Mi, H.; Guo, N.; Kejariwal, A.; Thomas, P. D., PANTHER version 6: protein sequence and function evolution data with expanded representation of biological pathways. *Nucleic Acids Res* **2007,** 35, (Database issue), D247-52.
63. Vizcaino, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Rios, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; Binz, P. A.; Xenarios, I.; Eisenacher, M.; Mayer, G.; Gatto, L.; Campos, A.; Chalkley, R. J.; Kraus, H. J.; Albar, J. P.; Martinez-Bartolome, S.; Apweiler, R.; Omenn, G. S.; Martens, L.; Jones, A. R.; Hermjakob, H., ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* **2014,** 32, (3), 223-6.
64. Mi, H.; Muruganujan, A.; Casagrande, J. T.; Thomas, P. D., Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* **2013,** 8, (8), 1551-66.
65. Mi, H.; Poudel, S.; Muruganujan, A.; Casagrande, J. T.; Thomas, P. D., PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res* **2016,** 44, (D1), D336-42.
66. Cieniewski-Bernard, C.; Dupont, E.; Deracinois, B.; Lambert, M.; Bastide, B., Multiplexed Detection of O-GlcNAcome, Phosphoproteome, and Whole Proteome within the Same Gel. *Front Endocrinol. (Lausanne)* **2014,** 5, 184.
67. Paulin, D.; Li, Z., Desmin: a major intermediate filament protein essential for the structural integrity and function of muscle. *Exp. Cell Res* **2004,** 301, (1), 1-7.
68. Heimburg, T.; Schuenemann, J.; Weber, K.; Geisler, N., Specific recognition of coiled coils by infrared spectroscopy: analysis of the three structural domains of type III intermediate filament proteins. *Biochemistry* **1996,** 35, (5), 1375-82.
69. Hnia, K.; Ramspacher, C.; Vermot, J.; Laporte, J., Desmin in muscle and associated diseases: beyond the structural function. *Cell Tissue Res* **2014**.
70. Bar, H.; Goudeau, B.; Walde, S.; Casteras-Simon, M.; Mucke, N.; Shatunov, A.; Goldberg, Y. P.; Clarke, C.; Holton, J. L.; Eymard, B.; Katus, H. A.; Fardeau, M.; Goldfarb, L.; Vicart, P.; Herrmann, H., Conspicuous involvement of desmin tail mutations in diverse cardiac and skeletal myopathies. *Hum Mutat* **2007,** 28, (4), 374-86.
71. Krishnamoorthy, V.; Donofrio, A. J.; Martin, J. L., O-GlcNAcylation of alphaB-crystallin regulates its stress-induced translocation and cytoprotection. *Mol Cell Biochem* **2013,** 379, (1-2), 59-68.
72. Elliott, J. L.; Der, P. M.; Prescott, A. R.; Jansen, K. A.; Koenderink, G. H.; Quinlan, R. A., The specificity of the interaction between alphaB-crystallin and desmin filaments and its impact on filament aggregation and cell viability. *Philos. Trans. R. Soc. Lond B Biol. Sci* **2013,** 368, (1617), 20120375.
73. Nicholl, I. D.; Quinlan, R. A., Chaperone activity of alpha-crystallins modulates intermediate filament assembly. *EMBO J* **1994,** 13, (4), 945-53.
74. Costa, M. L.; Escaleira, R.; Cataldo, A.; Oliveira, F.; Mermelstein, C. S., Desmin: molecular interactions and putative functions of the muscle intermediate filament protein. *Braz. J Med. Biol Res* **2004,** 37, (12), 1819-1830.
75. Perng, M. D.; Cairns, L.; van den, I. J.; Prescott, A.; Hutcheson, A. M.; Quinlan, R. A., Intermediate filament interactions can be altered by HSP27 and alphaB-crystallin. *J. Cell Sci* **1999,** 112 (Pt 13),

2099-2112.
76. Houck, S. A.; Landsbury, A.; Clark, J. I.; Quinlan, R. A., Multiple sites in alphaB-crystallin modulate its interactions with desmin filaments assembled in vitro. *PLoS. One* **2011,** 6, (11), e25859.
77. Arrigo, A. P.; Simon, S.; Gibert, B.; Kretz-Remy, C.; Nivon, M.; Czekalla, A.; Guillet, D.; Moulin, M.; Diaz-Latoud, C.; Vicart, P., Hsp27 (HspB1) and alphaB-crystallin (HspB5) as therapeutic targets. *FEBS Lett* **2007,** 581, (19), 3665-3674.
78. Arndt, V.; Dick, N.; Tawo, R.; Dreiseidler, M.; Wenzel, D.; Hesse, M.; Furst, D. O.; Saftig, P.; Saint, R.; Fleischmann, B. K.; Hoch, M.; Hohfeld, J., Chaperone-assisted selective autophagy is essential for muscle maintenance. *Curr Biol* **2010,** 20, (2), 143-8.
79. Ulbricht, A.; Eppler, F. J.; Tapia, V. E.; van der Ven, P. F.; Hampe, N.; Hersch, N.; Vakeel, P.; Stadel, D.; Haas, A.; Saftig, P.; Behrends, C.; Furst, D. O.; Volkmer, R.; Hoffmann, B.; Kolanus, W.; Hohfeld, J., Cellular mechanotransduction relies on tension-induced and chaperone-assisted autophagy. *Curr Biol* **2013,** 23, (5), 430-5.
80. Rosati, A.; Graziano, V.; De Laurenzi, V.; Pascale, M.; Turco, M. C., BAG3: a multifaceted protein that regulates major cell pathways. *Cell Death Dis* **2011,** 2, e141.
81. Tskhovrebova, L.; Trinick, J., Roles of titin in the structure and elasticity of the sarcomere. *J. Biomed. Biotechnol* **2010,** 2010, 612482.
82. Kontrogianni-Konstantopoulos, A.; Ackermann, M. A.; Bowman, A. L.; Yap, S. V.; Bloch, R. J., Muscle giants: molecular scaffolds in sarcomerogenesis. *Physiol Rev* **2009,** 89, (4), 1217-1267.
83. Bouameur, J. E.; Favre, B.; Fontao, L.; Lingasamy, P.; Begre, N.; Borradori, L., Interaction of plectin with keratins 5 and 14: dependence on several plectin domains and keratin quaternary structure. *J Invest Dermatol* **2014,** 134, (11), 2776-2783.
84. Bouameur, J. E.; Schneider, Y.; Begre, N.; Hobbs, R. P.; Lingasamy, P.; Fontao, L.; Green, K. J.; Favre, B.; Borradori, L., Phosphorylation of serine 4,642 in the C-terminus of plectin by MNK2 and PKA modulates its interaction with intermediate filaments. *J Cell Sci* **2013,** 126, (Pt 18), 4195-207.
85. Favre, B.; Schneider, Y.; Lingasamy, P.; Bouameur, J. E.; Begre, N.; Gontier, Y.; Steiner-Champliaud, M. F.; Frias, M. A.; Borradori, L.; Fontao, L., Plectin interacts with the rod domain of type III intermediate filament proteins desmin and vimentin. *Eur J Cell Biol* **2011,** 90, (5), 390-400.

FIGURE LEGENDS

Figure 1: Representative scheme of the workflow applied prior to the click chemistry. The protein fractions indicated in bold italic on figure 1 correspond to fraction submitted to azide-alkyne click chemistry and analysed on mass spectrometry.

Figure 2: Efficiency of the protein fractionation. (A) Protein profiles of the differential extraction. Twenty µg of proteins corresponding to whole extract (WE), cytosol-enriched fraction (CYT), membrane-enriched fraction (MB) and myofilament-enriched fraction (MYO) were separated on 7.5% Mini-PROTEAN TGX SF Precast Gels and visualized according to Stain-free technology. (B) Protein profiles resulting from ammonium sulphate (AS) precipitation of cytosol-enriched extract. Twenty µg of proteins corresponding to whole extract (WE), AS25, AS50, AS75 and AS100 (corresponding to final concentration of ammonium sulphate (AS)) were separated on AnykD Mini-PROTEAN TGX SF Precast Gels and visualized according to Stain-free technology. (C) Protein profiles of the 10 fractions resulting from MicroRotofor separation of proteins from AS25, AS50 and AS75 fractions, according to their isoelectric point on pH range of 4 to 7. Ten µl of each fraction were separated on AnykD Mini-PROTEAN TGX SF Precast Gels and visualized according to Stain-free technology. The percentage of proteins compared with the non-fractionated extract were indicated above each gel images. The pH of each fraction was indicated on supplemented table I.

Figure 3: Classification of the 342 proteins identified from dehydrated peptides released from agarose beads by beta-elimination. The classification was done using the PANTHER classification system (Protein Analysis THrough Evolutionary Relationships, http://www.pantherdb.org/). (A) Classification of proteins according to their molecular function. (B) Classification of proteins according to the protein classes.

SUPPLEMENTED TABLES

Table I: Measurement of pH of fractions resulting from isoelectric separation using MicroRotofor cell on pH 4-7 range. The pH of each fraction is indicated as value ± SD (standard deviation). The fraction F1 corresponds to the more acidic fraction, while the fraction F10 corresponds to the more basic one.

Table II, IV and VI: Data files corresponding to the identification of proteins from sequencing of peptides released from agarose beads by beta-elimination. These peptides correspond to O-

GlcNAcylated peptides (BE, beta-eliminated peptides). Table II corresponds to proteins identified from whole extract (WE), table IV to proteins identified from cytosol-enriched extract (CYT), and table VI to proteins identified from fractionated cytosol-enriched extract (fCYT).

Table III, V and VII: **Data files corresponding to the identification of proteins from sequencing of peptides released from agarose beads by trypsin.** These peptides correspond to non-O-GlcNAcylated peptides (NR, non-retained peptides), but they belong O-GlcNAcylated proteins covalently linked to agarose beads. Table III corresponds to proteins identified from whole extract (WE), table V to proteins identified from cytosol-enriched extract (CYT), and table VII to proteins identified from fractionated cytosol-enriched extract (fCYT).

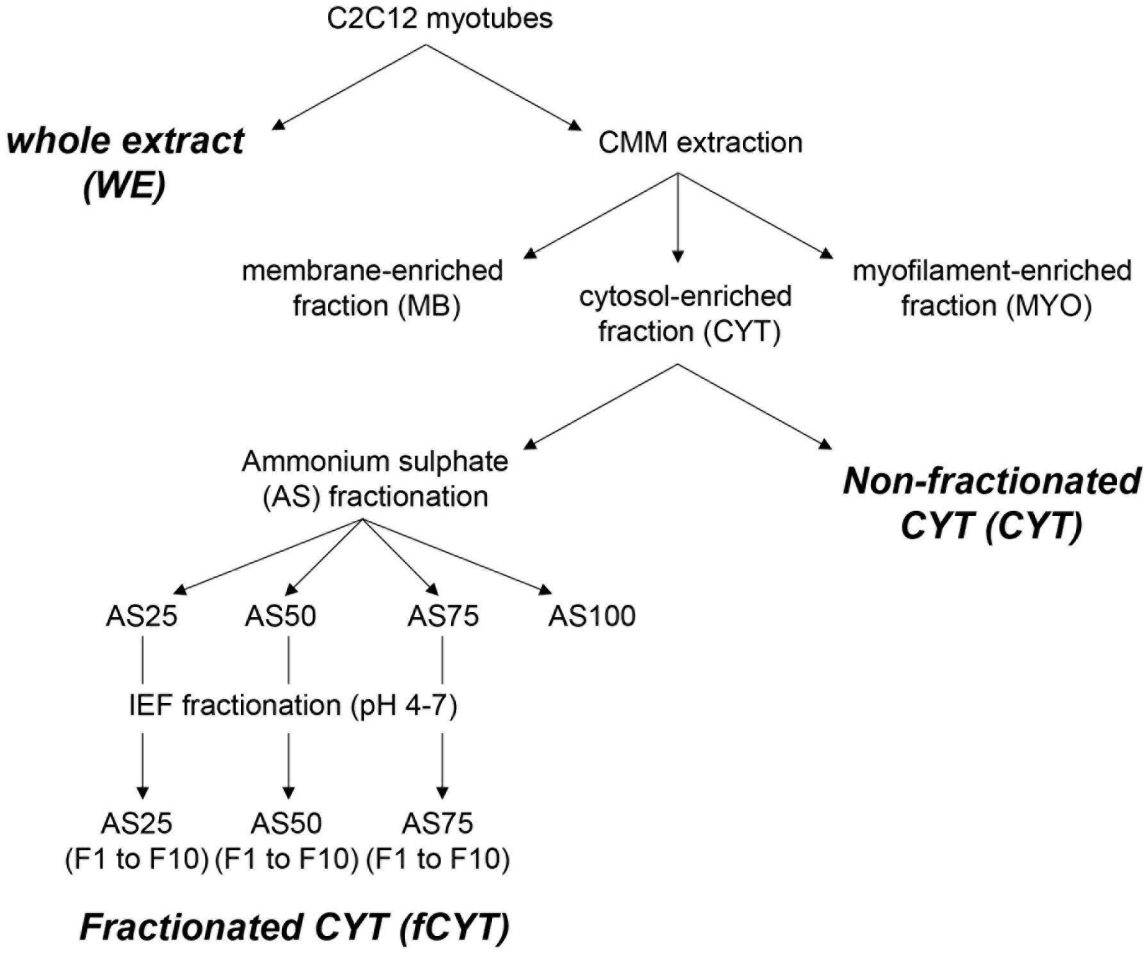

C2C12 myotubes
whole extract (WE)
CMM extraction
membrane-enriched fraction (MB)
cytosol-enriched fraction (CYT)
myofilament-enriched fraction (MYO)
Ammonium sulphate (AS) fractionation
Non-fractionated CYT (CYT)
AS25
AS50
AS75
AS100
IEF fractionation (pH 4-7)
AS25 (F1 to F10)
AS50 (F1 to F10)
AS75 (F1 to F10)
Fractionated CYT (fCYT)

Figure 1

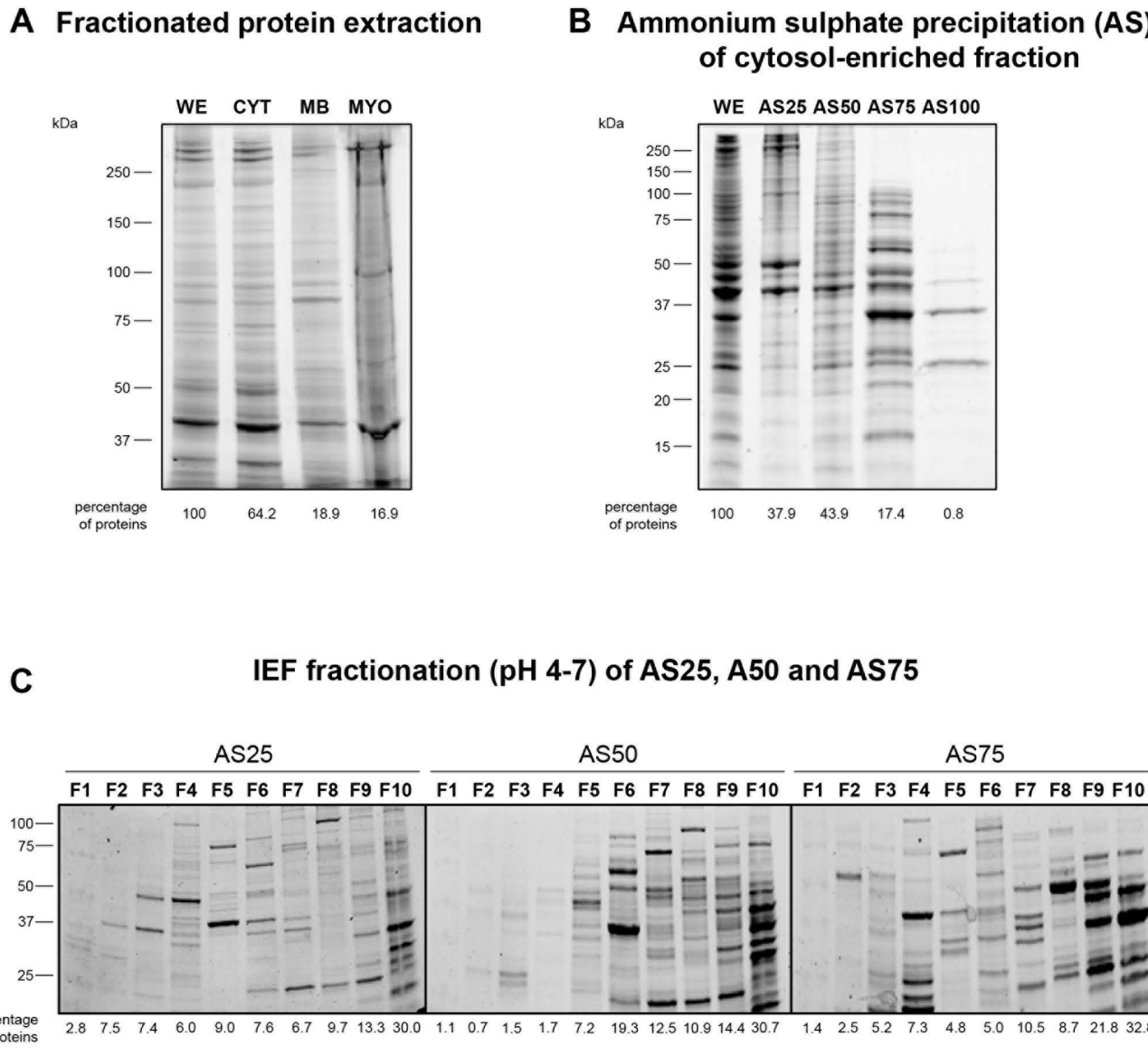
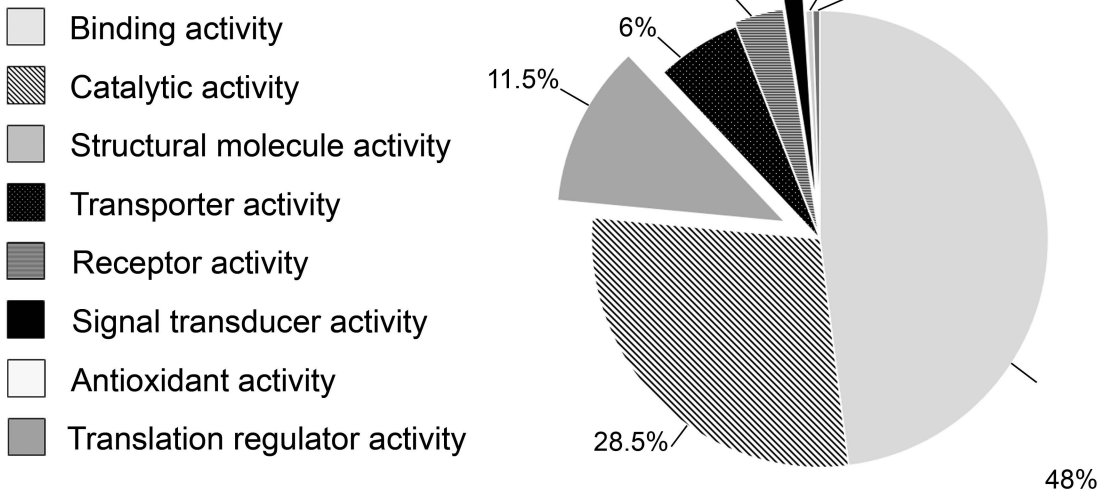

A Fractionated protein extraction
WE CYT MB MYO
kDa
250
150
100
75
50
37
percentage of proteins
100 64.2 18.9 16.9
B Ammonium sulphate precipitation (AS) of cytosol-enriched fraction
WE AS25 AS50 AS75 AS100
kDa
250
150
100
75
50
37
25
20
15
percentage of proteins
100 37.9 43.9 17.4 0.8
C IEF fractionation (pH 4-7) of AS25, A50 and AS75
AS25
AS50
AS75
kDa
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10
100
75
50
37
25
percentage of proteins
2.8 7.5 7.4 6.0 9.0 7.6 6.7 9.7 13.3 30.0
1.1 0.7 1.5 1.7 7.2 19.3 12.5 10.9 14.4 30.7
1.4 2.5 5.2 7.3 4.8 5.0 10.5 8.7 21.8 32.8

Figure 2

A. Molecular Function



B. Protein Class

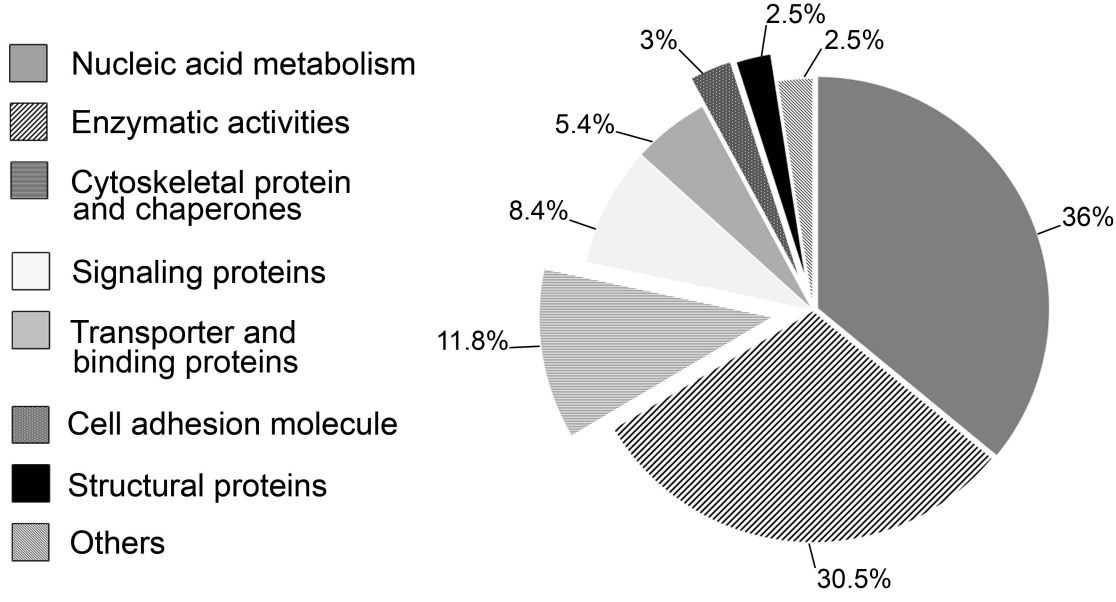


Figure 3

<u>Table I</u>: Overall evaluation of the O-GlcNAc mapping efficiency after extensive fractionation. The non-linked peptides on agarose beads and the peptides resulting from beta-elimination of peptides covalently linked to agarose beads were sequenced using mass spectrometry. The number of sequenced peptides and the corresponding number of identified proteins were indicated for each fraction (whole extract, WE; non-fractionated cytosol-enriched extract, CYT; fractionated cytosol-enriched extract, fCYT). For the non-linked peptides on agarose beads was also indicated the percentage of dehydrated peptides. The efficiency factor was determined as the ratio of sequenced peptides or identified proteins in fCYT fraction comparing the CYT fraction.

		Non fractionated whole extract (WE)	**Non-fractionated cytosol-enriched extract (CYT)**	**Fractionated cytosol-enriched extract (fCYT)**	**Efficiency factor: fCYT *vs* CYT**
Non-linked peptides on agarose beads (NR peptides)	**Number of peptides (% dehydrated peptides)**	5340 (7.1%)	5540 (6.2%)	14573 (9.3%)	**x 2.6**
	Number of identified proteins	554	449	1362	**x 3**
Linked peptides on agarose beads ; beta-eliminated	**Number of dehydrated peptides**	311	142	620	**x 4.4**
	Number of O-GlcNAcylated proteins	180	97	342	**x 3.5**

Table II: O-GlcNAcylation mapping on proteins involved in cytoskeleton and sarcomeric organization. Were indicated the accession number of the protein in UniProtKB database, the protein name (and [gene name]), the protein sequence (in bold and underlined were indicated the dehydrated amino acid), the O-GlcNAc site(s), and the fraction in which the peptides were identified (WE: whole extract; CYT: cytosol-enriched fraction; fCYT: fractionated cytosol-enriched fraction).

Accession number	Protein [gene name]	Peptide sequence	O-GlcNAc site(s)	Fraction of identified peptide(s)		
				WE	CYT	fCYT
Cytoskeleton proteins and proteins involved in cytoskeleton organization						
A2AHJ4	Bromodomain and WD repeat-containing protein 3 [Brwd3]	TCAPVAVLQGH**S**ASITSIQFCPSTKG**T**TR	S264 ; T279			X
H3BK97	FERM domain-containing protein 4A [Frmd4a] (fragment)	IISGSSGSLL**S**S	S386			X
D3Z589	LIM and calponin homology domains-containing protein 1 [Limch1]	FTATVET**T**IAR	T347			X
Q9QXD8	LIM domain-containing protein 1 [Limd1]	LAADGAAKPPLAVPTVAPGLAT**T**TAAAQPS YPSQEQR	T113			X
		GTTVSAPMVPS**S**ASQGACPK	S297			X
O70400	PDZ and LIM domain protein 1 [Pdlim1]	SAMPF**T**ASPAPSTR	T128		X	X
		SAMPFTASPAPS**T**R	T135			X
F7C957	PDZ and LIM domain protein 2 [Pdlim2] (fragment)	GGTPAFVPS**S**LSSQASLPTSR	S63			X
Q80TI1-2	Pleckstrin homology domain-containing family H member 1 (isoform 2) [Plekhh1]	HLADMMA**T**KWAALQGCSPPECIR	T511	X		
F7AWU3	Pleckstrin homology-like domain family B member 1 [Phldb1] (fragment)	NLAA**T**LQDIE**T**K	T112; T118		X	
Actin cytoskeleton organization and reorganization						
P60710	Actin, cytoplasmic 1 [Actb]	GYSF**T**TTAER	T201			X
P68033	Actin, alpha cardiac muscle 1 [Actc1]	VAPEEHP**T**LLTEAPLNPK	T105			X
		GY**S**FVTTAER	S201			X
E9Q9Q7	Actin-binding LIM protein 1 [Ablim1]	ST**S**QGSINSPVYSR	S110	X	X	X
		HSYTPT**T**SR	T128			X

D3YYR9	Actin-binding LIM protein 2 [Ablim2] (fragment)	TSSE**S**IVSVPASSTSGSPSR	S58			X
B1AXE9	Angiomotin [Amot] (fragment)	AHPPV**T**SAPLSPPQPNDLYK	T196			X
Q62418	Drebrin-like protein [Dbnl]	AMST**T**SVTSSQPGK	T279			X
Q3USJ8	FCH and double SH3 domains protein 2 [Fchsd2]	QLE**S**E**T**G**TT**EEHSLNK	S314; T316; T318; T319			X
D3Z4C0	FERM, RhoGEF and pleckstrin domain-containing protein 2 [Farp2]	CDGQVLL**T**QVWK	T71	X		X
Q76LL6	FH1/FH2 domain-containing protein 3 [Fhod3]	A**S**MCSGG**T**VGEQQGLDR	S345; T351			X
Q9CS72	Filamin-A-interacting protein 1 [Filip1]	**S**PRNHLSSRPGANKVTSTITITPVTTS**ST**R	S1112; S1139; T1140			X
		VTSTITITPV**T**TSSTR	T1136	X	X	X
		VTSTITITPV**TT**SSTR	T1136; T1137	X	X	X
E0CYM1	Filamin A-interacting protein 1-like [Filip1l]	V**T**SSITITPTATPLPR	T868		X	X
Q0GNC1-3	Inverted formin-2 (isoform 2) [Inf2]	ISDALLQL**T**CISCVR	T100		X	
Q9ERG0	LIM domain and actin-binding protein 1 [Lima1]	VST**T**ENSLVALSVPAEDDTCNSQVK	T326			X
P53668	LIM domain kinase 1 [Limk1]	MGEEG**S**ELPVCASCGQR	S20		X	X
G3X9H7	Metastasis suppressor 1, isoform CRA_e [Mtss1]	RPASTAGLPT**T**LGPAMVTPGVATIR	T580			X
Q3TN34	MICAL-like protein 2 [Micall2]	GFVQTELKPP**S**TSQVHVGSSAGPK	S314			X
		GFVQTELKPP**S**TSQVHVGS**S**AGPK	S314; S323			X
		VPTVVTVPT**S**K	S456			X
		VPTVVTVPT**S**KVPNVVTAPTSK	S456			X
B1ARU1	Microtubule-actin cross-linking factor 1 [Macf1]	GCD**T**ETDAD**S**L**S**HTLQPYK	T1912; S1918; S1920	X		
P59759	MKL/myocardin-like protein 2 [Mkl2]	V**S**ASPPPVTASTPAQFTSVSPAVPEFLK	S207		X	
		VSASPPPV**T**ASTPAQFTSVSPAVPEFLK	S214			X
		TPLTADQPPTR**S**TAPVLPTNTVSSAK	S245	X	X	X
F7D291	MTSS1-like protein [Mtss1l] (fragment)	HGEEV**S**PAASDLAMVLTR	S322			X
		HGEEV**S**PAASDLAMVL**T**R	S322; T333			X

Q04690-4	Neurofibromin (isoform 4) [Nf1]	GNSSMDSTAGC**S**GTPPICR	S676		X	
Q8CI51	PDZ and LIM domain protein 5 [Pdlim5]	EVVKPVPI**T**SPAVSK	T110	X		X
		EVVKPVPIT**S**PAVSK	S111		X	
		SSGTGASVGPPQPSDQDTLVQR	S381; S382		X	X
Q03173-4	Protein enabled homolog (isoform 3) [Enah]	VISAPVSDATPDYAVVTALPPT**S**TPPTPPLR	S292			X
		VISAPVSDATPDYAVVTALPPTS**T**PPTPPLR	T293	X	X	
		APST**S**TPEPTR	S701			X
		APSTS**T**PEPTR	T702		X	
D3YVS1	Smoothelin [Smtn]	RDG**S**SSSSTTTTTVQTK	S212	X		
		DGSSS**S**STTTTTVQTK	S215	X	X	
		RDGSSS**S**STTTTTVQTK	S215			X
		RDGSSSSST**T**TTTVQTK	T218		X	
E9Q6A3	Sorbin and SH3 domain-containing protein 1 [Sorbs1]	TPVDYIDLPY**S**SSPSR	S623			X
		TPVDYIDLPY**S**SSP**S**R	S623; S627			X
F6V513	Sorbin and SH3 domain-containing protein 2 [Sorbs2] (fragment)	ADLPGSSSTF**T**K	T148	X		
B2RXQ9	Sorbin and SH3 domain-containing protein 2 [Sorbs2]	ADLPGSS**S**TFTK	S368			X
Q62261	Spectrin beta chain, non-erythrocytic 1 [Sptbn1]	HDTSASTQSTPA**S**SR	S2322	X	X	X
Sarcomere organization, motor activity and contractile fibers proteins						
D3YUW7	Cingulin [Cgn]	GRPA**T**SR	T109			X
P70402	Myosin-binding protein H [Mybph]	A**T**SEASVSTPEETAPEPAK	T6			X
A2ABU4	Myomesin-3 [Myom3]	APGG**T**CR	T430		X	
Q5DTJ9	Myopalladin [Mypn]	IQKPNEVSSPPT**T**SAAIPPAAEAQPLAAQPR	T388		X	
		IQKPNEVSSPPT**TS**AAIPPAAEAQPLAAQPR	T388; S389			X
		IQKPNEVSSPPTT**S**AAIPPAAEAQPLAAQPR	S389			X
Q9JKS4-4	LIM domain-binding protein 3 [Ldb3] (isoform 4)	RPIPI**S**TTAPPIQSPLPVIPHQK	S90		X	
		RPIPIS**T**TAPPIQSPLPVIPHQK	T91			X
		SKRPIPIS**T**TAPPIQSPLPVIPHQK	T91			X
		SRPQASAYSPAAAASPAPSAH**T**SYSEGPAA PAPKPR	T277		X	

		SRPQASAYSPAAAASPAPSAHTSY**S**EGPAAP APKPR	S280			X
E9Q3W4	Plectin [Plec]	**S**SSVG**SSS**SYPISSAGPR	S4086; S4091; S4092; S4093			X
		SS**S**VGSS**S**SYPIS**S**AGPR	S4088; S4093; S4099			X
P97306	SH3 and cysteine-rich domain-containing protein [Stac]	A**S**PGP**S**PIAIPGSPA**S**MP**T**K	S79; S83; S93; T96			X
Q91YE8-2	Synaptopodin-2 (isoform 2) [Synpo2]	V**S**SAPAMK	S990	X		
		VS**S**APAMK	S991		X	
D3YU08	Synaptopodin 2-like protein [Synpo2l]	VNEGLGSTSPAPSPFAAPPQGPTPLPSFTTV VP**S**HTPVSGASSSTQR	S299			X
		VNEGLGSTSPAPSPFAAPPQGPTPLPSFTTV VPSH**T**PVSGASSSTQR	T301	X		
		VNEGLGSTSPAPSPFAAPPQGPTPLPSFTTV VPSHTPV**S**GASSSTQR	S304		X	
		KVNEGLGSTSPAPSPFAAPPQGPTPLPSFTT VVPSHTPV**S**GASSSTQR	S304	X	X	
		KVNEGLGSTSPAPSPFAAPPQGPTPLPSFTT VVPSHTPV**S**GA**S**SSTQR	S304; S307			X
		KVNEGLGSTSPAPSPFAAPPQGPTPLPSFTT VVPSHTPVSGA**S**SSTQR	S307			X
E9Q8K5	Titin [Ttn]	RPGPPS**T**PEA**S**AITK	T24950; S24954			X
Q8BUJ6	Titin [Ttn] (Fragment)	TIV**S**TAQISETR	S209	X		
Microtubules cytoskeleton organization and reorganization						
A2AGT5-2	Cytoskeleton-associated protein 5 (isoform 2) [Ckap5]	TTAPGGSA**S**AGTK	S568	X	X	
F8WGN6	Kinesin-like protein KIF21A [Kif21a]	MTI**S**NMEADMNR	S930			X
Q6GQX2	Nck-associated protein 5-like [Nckap5l]	**TST**PQGPAFGGSR	T1269; S1270; T1271			X
P48678	Prelamin-A/C [Lmna]	AAGGAGAQVGGSI**S**SGSSASSVTVTR	S613	X	X	X
		AAGGAGAQVGGSI**SS**GSSASSVTVTR	S613 ; S614	X		

E9Q5G3	Protein Kif23 [Kif23]	SNSCS**S**ISVASCISEWEQK	S720	X		
Q8VHJ5	Serine/threonine-protein kinase MARK1 [Mark1]	GT**S**TGIISK	S649		X	
Chaperones						
P24622-2	Alpha-crystallin A chain (isoform 2) [Cryaa]	AIPV**S**REEKPSSAPSS	S185		X	
P23927	Alpha-crystallin B chain [Cryab]	TIPI**T**REEKPAVAAAPK	T162			X
Q9JLV1	BAG family molecular chaperone regulator 3 [Bag3]	QCGQMPATAT**T**AAAQPPTAHGPER	T163			X
		SQSPAASDCSSS**S**SSASLPSSGR	S189			X
		V**S**SAPIPCPSPSPAPSAVPSPPK	S372	X	X	
		VS**S**APIPCPSPSPAPSAVPSPPK	S373			X
P14602	Heat shock protein beta-1 [Hspb1]	AVTQSAEITIPV**T**FEAR	T188			X
P14602-3	Heat shock protein beta-1 (isoform C) [Hspb1]	AV**T**QSAEITIPVTFEAR	T144	X	X	
P63017	Heat shock cognate 71 kDa protein [Hspa8]	**S**TAGDTHLGGEDFDNR	S221			X
E9QNY8	Sacsin [Sacs]	LSV**S**VIDS**S**DEA**T**IR	S667; S972; T976			X
Intermediate filament proteins						
P31001	Desmin [Des]	DGEVV**S**EATQQQHEVL	S459			X
D6RH90	Intermediate filament family orphan 1 [Iffo1]	SPAGPLA**SS**AACHTSSSTSTSTAFSSSTR	S169; S170			X
		SPAGPLASSAACH**T**SSSTSTSTAFSSSTR	T175			X
Q8BXL9-2	Intermediate filament family orphan 1 (Isoform 2) [Iffo1]	SPAGPLA**S**SAACHTSSSTSTSTAFSSSTR	S169	X		
		SPAGPLA**SS**AACHTSSSTSTSTAFSSSTR	S169; S170	X	X	
		SPAGPLA**SS**AACHT**S**SSTSTSTAFSSSTR	S169; S170; S176		X	
		SPAGPLAS**S**AACHT**S**SSTSTSTAFSSSTR	S170; S176		X	
		SPAGPLASSAACH**T**SSSTSTSTAFSSSTR	T175	X	X	
Q9EPM5	Syncoilin [Sync]	SLEQADAPT**S**QAGGVEAQSPGTV	S457	X	X	X
P20152	Vimentin [Vim]	RMFGGSGT**S**SRPSSNR	S21			X
		SYV**T**TSTR	T32		X	
		SYVTT**S**TR	S34	X		X

		TY**S**LGSALRPSTSR	S39	X		
		TYSLGSALRPS**T**SR	T48		X	
		TYSLGSALRPST**S**R	S49			X
		SLY**S**SSPGGAYVTR	S54		X	
		SLYS**S**SPGGAYVTR	S55			X
Cell adhesion molecules						
P97326	Cadherin-6 [Cdh6]	IFNID**S**GNGSIF**TS**K	S435; T442; S443			X
O35161	Cadherin EGF LAG seven-pass G-type receptor 1 [Celsr1]	TALGE**ST**ASLDS**T**TR	S2773; T2774; T2780			X
		TALGES**T**ASLDS**TT**R	T2774; T2780; T2781		X	
Q62419	Endophilin-A2 [Sh3gl1]	I**T**ASSSFR	T284	X	X	X
		ASQLVSEKVGGAEG**T**K	T27		X	
E9Q616	Protein Ahnak [Ahnak]	LPSGSGPASPT**T**GSAVDIR	T220		X	X
		VEV**S**APDVSIEGSEGK	S787			X
		V**S**VATPDVSLEASEGAVK	S5166	X	X	X
		GPQV**S**SSLNLDTSK	S5376	X	X	X
		GGQTGLQGPGL**S**VSGPQGHLESESGK	S5414	X		
E9Q016	Ras-related protein Rab-40C [Rab40c]	**S**Y**S**LA**S**GAGGSGSKGN**S**LKR	S221; S223; S226; S237			X
P26039	Talin-1 [Tln1]	ALGDLISA**T**K	T2098		X	
Q62523	Zyxin [Zyx]	TPSSSQPPPQPQAKPQVQLHVQPQAKPHVQ PQPV**S**SANTQPR	S237			X
		TPSSSQPPPQPQAKPQVQLHVQPQAKPHVQ PQPVS**S**ANTQPR	S238		X	
		MVPPDAPS**S**VSTGSPQPPSFTYAQQK	S295		X	
		MVPPDAPSSV**S**TGSPQPPSFTYAQQK	S297	X		X

Table III: O-GlcNAcylation mapping on proteins involved in signaling pathways and in protein modifications. Were indicated the accession number of the protein in UniProtKB database, the protein name (and [gene name]), the protein sequence (in bold and underlined were indicated the dehydrated amino acid), the O-GlcNAc site(s), and the fraction in which the peptides were identified (WE: whole extract; CYT: cytosol-enriched fraction; fCYT: fractionated cytosol-enriched fraction).

Accession number	Protein [gene name]	Peptide sequence	O-GlcNAc site(s)	Fraction of identified peptide(s)		
				WE	CYT	fCYT
Proteins involved in signaling or apoptotic pathways						
P31230	Aminoacyl tRNA synthase complex-interacting multifunctional protein 1 [Aimp1]	L**ST**PLQ**T**NC**T**A**S**ESVVQSPSVATTASPATK	S83; T84; T88; T91; S93			X
		LSTPLQTNC**T**ASESVVQSPSVATTASPATK	T91	X	X	X
		LSTPLQTNC**T**A**S**ESVVQSPSVATTASPATK	T91; S93			X
		LSTPLQTNCTASE**S**VVQSPSVATTASPATK	S95		X	
F8WIT2	Annexin [Anxa6]	**T**NEQMHQLVAAYK	T123			X
Q8CG79	Apoptosis-stimulating of p53 protein 2 [Tp53bp2]	ENLPVSPDGNLPQQAV**S**APSR	S339	X	X	X
Q3TBU7	Arf-GAP domain and FG repeat-containing protein 2 [Agfg2] (fragment)	GVTYVDI**T**VGSFVCTTCSGLLR	T61			X
		TLLGDPVPSLSDPAS**T**SSQPGSQSQAR	T193			X
P47746	Cannabinoid receptor 1 [Cnr1]	**T**VFAFCSMLCLLN**ST**VNPIIYALR	T378; S391; T392		X	
Q80YE7-2	Death-associated protein kinase 1 (isoform 2) [Dapk1]	CCLLLDSVCSTIETVMAT**T**LPGLL**T**VKHYL SPQQLR	T1226 ; T1232		X	
P42128	Forkhead box protein K1 [Foxk1]	VCEVGPEEPAAAV**S**VAANAAPTPAASTTTS ASSSGEPEVK	S663			X
Q3UCQ1	Forkhead box protein K2 [Foxk2]	IIQT**S**QGTPVQTVTIVQQAPLGQHQLPIK	S540	X		X
D3Z7Q3	Guanine nucleotide-binding protein subunit gamma [Gng5]	QFCLQNAQHDPLL**T**GVSS**ST**NPFRPQK	T49		X	
Q5DU31	Interactor protein for cytohesin exchange factors 1 [Ipcef1]	**S**LEQA**S**L**S**PLGDR	S292 ; S297 ; S299		X	
E9Q1P8	Interferon regulatory factor 2-binding protein 2 [Irf2bp2]	AAATLAAVSG**T**PGLGAQPAELGTHK	T208			X
G3XA78	Interleukin-1 receptor accessory protein [Il1rap] (fragment)	CCV**T**YCEGE**S**HLR	T140; S146			X

D3Z6H5	Interleukin-12 receptor subunit beta-2 [Il12rb2]	SQNSHPI**SS**LQPR	S267; S268	X		
Q8BX02	KN motif and ankyrin repeat domain-containing protein 2 [Kank2]	EVEVAA**S**TAAGALAQR	S346			X
Q66L42	Mitogen-activated protein kinase kinase kinase 10 [Map3k10]	L**T**PMDCGGSSGSG**T**WSR	T519; T531			X
P70670	Nascent polypeptide-associated complex subunit alpha, muscle-specific form [Naca]	PGEATETVPATEQELPQPQAETAVLPM**S**SALK	S29			X
		SSSAPLSLVALAPHSVQK	S128		X	
		S**S**SAPLSLVALAPHSVQK	S129	X		X
		SSVCPPHPLTSPP**S**AAGAELGALTASIPPLEPK	S159			X
		TSTSQVPSQGTLNLK	T179	X		
		TSTSQVP**S**QGTLNLK	T179; S186	X		
		T**S**TSQVPSQGTLNLK	S180		X	
		T**S**T**S**QVPSQGTLNLK	S180; S182		X	
		T**S**TSQVPSQG**T**LNLK	S180; T189			X
		TS**T**SQVPSQGTLNLK	T181			X
		AFPSHLENPLASVQPGLMSCPQ**T**LSNTSPVK	T227	X	X	X
		GSNVALQPLVTQVPA**S**QK	S392			X
		EIPV**S**CIGATHHALDNPSAISVAPATHVPPPTSSGLVSSK	S403		X	X
		EIPV**S**CIGA**T**HHALDNPSAISVAPATHVPPPTSSGLVSSK	S403; T408		X	X
		EIPV**S**CIGATHHALDNPSAISVAPATHVPPP**T**SSGLVSSK	S403; T430		X	
		EIPV**S**CIGATHHALDNPSAISVAPATHVPPP**T**SSGLVS**S**K	S403; T430; S437			X
		EIPVSCIGA**T**HHALDNPSAISVAPATHVPPPTSSGLVSSK	T408		X	X
		EIPVSCIGA**T**HHALDNPSAISVAPATHVPPP**T**SSGLVSSK	T408; T430		X	
		EIPVSCIGA**T**HHALDNPSAISVAPATHVPPPTS**S**GLVSSK	T408; S432			X
		ESPS**S**QSASSLEVLSEDTVTK	S522		X	X

		TTGGPAPVVRPAIAGVA**TT**TSLR	T557; T558		X	
		TTGGPAPVVRPAIAGVAT**T**TSLR	T558	X	X	X
		TTGGPAPVVRPAIAGVAT**T**T**S**LR	T558; S560			X
		TTGGPAPVVRPAIAGVATT**T**SLR	T559	X		X
		ADSCVSPN**T**VSQPLKR	T580			X
		NTAPSTTSPLVPLA**S**EGCPVASSMALSPQNA SVSETALALSPEIPK	S615			X
		GTVVCLADSSLDTSVSA**S**K	S708			X
		LSPTPPS**S**K	S752		X	X
		GAPVPS**T**GAPPSPK	T760	X	X	X
		GAPIVP**T**ESSISSK	T774		X	X
		TPEV**T**ASR	T798		X	
		K**T**SATAVPK	T826		X	
		D**T**SATLSLK	T835		X	X
		E**T**LATSIPK	T886			X
		GPPA**T**LAETPTYPK	T972		X	X
		ETPAT**S**SEGVTAVPSEISPSPPTPASK	S1243		X	X
		DAPT**T**LAESPSSPK	T1485		X	X
		VPVPAETQEVAV**S**SR	S1541		X	X
		EFPASPSIKPVTT**S**LAQTAPPSLQK	S1613		X	X
		ENLAAPAVLPVS**S**K	S1645		X	X
		AAATETPIET**S**TAPSLEGAPK	S1691		X	
		AAATETPIETSTAP**S**LEGAPK	S1695			X
		AS**T**LPATTLPSLK	T1728			X
		EASVLSP**T**ATSSGK	T1746		X	
		EASVLSPTAT**S**SGK	S1749	X		X
		EASVLSPTATS**S**GK	S1750		X	
		DSHISPVSDAC**S**TGTTTPQASEK	S1764			X
Q01705-2	Neurogenic locus notch homolog protein 1 (isoform 2) [Notch1]	GLACG**S**K	S1856		X	
B1B0C9	Nik-related protein kinase [Nrk]	MFFASTL**S**NDH**S**R	S1432 ; S1436		X	
B8JJB2	PDZ and LIM domain protein 7 [Pdlim7] (fragment)	AQPVQ**S**KPQK	S89	X	X	

F6T0L5	Plexin-D1 [Plxnd1] (fragment)	**S**PVSGGR	S190		X	
F8VQK3	Protein Gucy1a2 [Gucy1a2]	LCFDG**S**NP**S**NC**S**CL**T**FLIK	S279; S282; S285; T288			X
Q9Z1A1	Protein Tfg [Tfg]	LLDSLEPPGEPGPSTSIPENDTVDGREEKPA A**S**DSSGK	S151			X
		LLDSLEPPGEPGPSTSIPENDTVDGREEKPA ASDS**S**GK	S154		X	
P48614-2	Protein Wnt-10b (isoform short) [Wnt10b]	ESAF**S**F**S**MLAAGVMHAVATACSLGK	S120 ; S122		X	
P97855	Ras GTPase-activating protein-binding protein 1 [G3bp1]	NLPPSGAVPV**T**GTPPHVVK	T266		X	X
P97379	Ras GTPase-activating protein-binding protein 2 [G3bp2]	SATPPPAEPA**S**LPQEPPK	S235			X
P35279	Ras-related protein Rab-6A [Rab6a]	LVFLGEQ**S**VGKTSLI**T**R	S23; T31			X
F6TYF8	Ras-specific guanine nucleotide-releasing factor 2 [Rasgrf2] (fragment)	LSLT**S**SLNSR	S147			X
		LSLTS**S**LNSR	S148		X	
F6YRR2	Regulator of G-protein-signaling 12 [Rgs12] (fragment)	VCLLGGEAAF**S**GY**S**WVQGG	S99; S102			X
G5E8J9	SCY1-like protein 2 [Scyl2]	I**S**ASSTFTPVPSTGLGMMFSTPIDNTK	S740			X
		ISA**S**STFTPVPSTGLGMMFSTPIDNTK	S742			X
F6WNS2	Serine/threonine-protein kinase STK11 [Stk11] (fragment)	AVCVNGTEPQL**SS**K	S131; S132			X
P83741	Serine/threonine-protein kinase WNK1 [Wnk1]	QPIAVS**S**MPQQIGVPTSSLTQVVHSAGR	S1230		X	X
		DGTEGHV**T**AT**S**SGAGVVK	T1841; S1844			X
		DGTEGHVTA**TS**SGAGVVK	T1843; S1844	X		
		DGTEGHVTATS**S**GAGVVK	S1845	X		X
		FQVT**T**TANK	T1945	X	X	X
		FSAPGQLCVPMTSNLGGSTPI**S**AASATSLGH FTK	S2301		X	X
		KFSAPGQLCVPMTSNLGGSTPISAASA**T**SLG HFTK	T2306			X
E9Q6Q2	Serine/threonine-protein kinase WNK2 [Wnk2]	**S**AQCTAQPL**S**TGQGPCTPALEA**S**R	S1277; S1286; S1299			X
Q80TE4	Signal-induced proliferation-associated 1-like protein 2 [Sipa1l2]	LQEESQ**T**A**T**AQLR	T1702; T1704			X

B7ZC18	Signal transducer and activator of transcription 3 [Stat3]	FICVTPT**T**CSNTIDLPMSPR	T691		X	
Q62417-5	Sorbin and SH3 domain-containing protein 1 (isoform 5) [Sorbs1]	TPVDYIDLPY**S**SSPSR	S593	X	X	
		TPVDYIDLPY**S**SSP**S**R	S593; S597	X		
Q9DBG9	Tax1-binding protein 3 [Tax1bp3]	V**S**EGGPAEIAGLQIGDK	S61			X
Q8CF89	TGF-beta-activated kinase 1 and MAP3K7-binding protein 1 [Tab1]	VYPVSVPY**S**SAQSTSK	S393	X	X	X
		VYPVSVPY**SS**AQSTSK	S393; S394			X
D3Z216	TGF-beta-activated kinase 1 and MAP3K7-binding protein 2 [Tab2] (fragment)	NQPTLFIS**T**NSGPSAASR	T411			X
		VVV**T**QPNTK	T456		X	X
Q571K4	TGF-beta-activated kinase 1 and MAP3K7-binding protein 3 [Tab3]	NQHSLYTATTPP**S**SSPSR	S411	X		
		NQHSLYTATTPPS**S**SPSR	S412		X	
		NQHSLYTATTPPSS**S**PSR	S413			X
D3Z2W0	TNFAIP3-interacting protein 1 [Tnip1]	DSELSPP**T**SAPSLVSFDDLAELTGQDTK	T76			X
		VQVHPA**T**STAATTTATATTGNSMEKPEPASK	T103			X
		VQVHPA**TS**TAATTTATATTGNSMEKPEPASK	T103; S104			X
		VQVHPAT**S**TAATTTATATTGNSMEKPEPASK	S104			X
		VQVHPAT**ST**AATTTATATTGNSMEKPEPASK	S104; T105			X
P25446	Tumor necrosis factor receptor superfamily member 6 [Fas]	CA**S**CEHGTLEPC**T**ATSNTNCR	S144; S154			X
G3UYW7	Yorkie homolog [Yap1] (fragment)	TMT**T**NSSDPFLNSGTYHSR	T273			X
Proteins involved in protein modification (phosphorylation, ubiquitination, SUMOylation)						
Q8BZ25	Ankyrin repeat and protein kinase domain-containing protein 1 [Ankk1]	G**T**FLGI**T**HLLEYGADIHACNK	T679; T684			X
P63248	cAMP-dependent protein kinase inhibitor alpha [Pkia]	SSTEQSGEAQGEAAK**S**ES	S74			X
Q8BT14	CCR4-NOT transcription complex subunit 4 [Cnot4]	ELSVQDQPSLSPTSLQNAS**S**HTTTAK	S441			X
		SNPVIPI**S**SSNHSAR	S316			X
Q8BT14-2	CCR4-NOT transcription complex subunit 4 (isoform 2) [Cnot4]	SNPVIPI**S**SSNHSAR	S313	X		
		SNPVIPISS**S**NHSAR	S315		X	

Q9ERU9	E3 SUMO-protein ligase RanBP2 [Ranbp2]	SVFT**T**AASELANK	T1138	X		
B1AXR5	Perilipin-2 [Plin2] (fragment)	LPILNQPTSEIVA**S**AR	S107			X
		TVT**S**AAMTSALPIIQK	S59			X
Q9DBR7	Protein phosphatase 1 regulatory subunit 12A [Ppp1r12a]	TKPMASV**S**NAHTSSTQAAPAAVTAPTLSSNQGTPTSPVK	S381			X
		TKPMASVSNAHT**S**STQAAPAAVTAPTLSSNQGTPTSPVK	S386			X
		RQDDLISC**S**VPSTTSTPTVTSAAGLQR	S566			X
		RQDDLISCSVPS**T**TSTPTVTSAAGLQR	T570			X
		RQDDLISCSVPS**TT**STPTVTSAAGLQR	T570; T571			X
		RQDDLI**S**CSVPST**T**STPTVTSAAGLQR	S564; T571			X
Q8BG95	Protein phosphatase 1 regulatory subunit 12B [Ppp1r12b]	GSETPQ**T**IAPSTYTSTYLK	S542			X
Q8R4S0	Protein phosphatase 1 regulatory subunit 14C [Ppp1r14c]	GGTGGSPG**S**SSSSGSSREDSAPVTTVAAAGQVQQQQR	S36			X
G3UZ19	Serine/threonine-protein kinase BRSK2 [Brsk2] (fragment)	SISGA**S**SGLSTSPLSSPR	S229			X
Proteins involved in protein degradation or autophagy						
Q6ZPS6	Ankyrin repeat and IBR domain-containing protein 1 [Ankib1]	GHLFCWECLGEAHEPCDCQ**T**WK	T481	X		
E9QKG6	Ankyrin repeat domain-containing protein 17 [Ankrd17]	IGSSAPT**T**TAANSSLMGIK	T1570		X	
		IGSSAPT**TT**AANSSLMGIK	T1570; T1571	X	X	X
		MT**T**VALSSTSQTA**T**ALTVPAISSASTHK	T1584 ; T1595			X
		MTTVALS**S**TSQTATALTVPAISSASTHK	S1589		X	
		MTTVALS**S**TSQTATALTVPAI**S**SASTHK	S1589; S1603		X	
		MTTVALSSTSQ**T**ATALTVPAISSASTHK	T1593			X
		MTTVALSSTSQTATALTVPAISSAS**T**HK	T1607		X	
		EHYPV**S**SPSSPSPPAQPGGVSR	S1786		X	
		TPAPVQS**S**SASVLNVNHIK	S1942	X	X	X
Q8CDK2	Cytosolic carboxypeptidase 2 [Agbl2]	IQSKKPGFTA**S**C**S**PK	S814; S816			X
A2AKB9-4	DDB1- and CUL4-associated factor 10 (isoform 4) [Dcaf10]	LPGGPAVSPAERAGEFAAPGALELSAA**T**ASASQAKLSPS**SS**PR	T87; S99; S100		X	

Q8CDG3	Deubiquitinating protein VCIP135 [Vcpip1]	TEPPVFTAA**S**SNSELIR	S1075	X	X	X
A2AIR2	E3 ubiquitin-protein ligase BRE1A [Rnf20] (fragment)	SG**S**ALLQ**S**Q**S**STEDPK	S517; S522; S524		X	
Q4U2R1-2	E3 ubiquitin-protein ligase HERC2 (isoform 2) [Herc2]	**S**ARPTGR	S182		X	
Q69ZI1-3	E3 ubiquitin-protein ligase SH3RF1 (isoform 3) [Sh3rf1]	VSMS**T**AGQASR	T527	X		
Q5BJ29	F-box/LRR-repeat protein 7 [Fbxl7]	GSSSV**S**SDVSSSTDHTPTKAQR	S20		X	
A2A7H5	F-box only protein 44 [Fbxo44]	V**T**N**SS**ITIGPPLP	T244; S246; S247	X		
Q6P9N1	Hyccin [Fam126a]	AASSTSQ**S**GL**S**NSSHNC**S**NK	S362; S365; S372			X
Q8BX02-2	KN motif and ankyrin repeat domain-containing protein 2 (isoform 2) [Kank2]	EVEVAA**S**TAAGALAQR	S346	X		X
P0CG49	Polyubiquitin-B [Ubb]	E**S**TLHLVLR	S293	X	X	
Q9JKV1	Proteasomal ubiquitin receptor ADRM1 [Adrm1]	SQ**S**AAVTPSSSTSSAR	S213		X	X
		SQ**S**AAVTPSSS**T**SSAR	S213; T222			X
Q80TR8-4	Protein VPRBP (isoform 4) [Vprbp]	NCATFNPTDDLVLNDGVLWDVR**S**AQAIHK	S1247		X	
Q505D9	Tripartite motif-containing protein 67 [Trim67]	GAACSSLC**SSS**SSITCPQCHR	S147; S148; S149	X		
A2AMY5	Ubiquitin-associated protein 2 [Ubap2]	LKPPGLSPFPAAS**S**AQQNDTASPPATTAAW DLKPSAPQPSVLSR	S392			X
		SQPEPSPVLSQL**S**QR	S439			X
		E**S**TAGDGPSTVSR	S468	X	X	X
		SQHTVDT**T**SSVPAPK	T658		X	X
		DGNLANNPY**S**GDVTK	S876			X
		GV**S**V**S**SGTGLPDMTGSVYNK	S1017; S1019			X
		GVSV**S**SGTGLPDMTGSVYNK	S1019	X	X	X
		GVSV**SS**GTGLPDMTGSVYNK	S1019; S1020			X
Q80TQ2	Ubiquitin carboxyl-terminal hydrolase CYLD [Cyld]	V**T**GSTSDPG**S**R	T337; S345		X	
D6RHQ6	Ubiquitin carboxyl-terminal hydrolase 16 [Usp16]	GLSNLGN**T**CFFNAVMQNL**S**QTPVLR	T202; S213	X		

Q99MX1	Ubiquitin carboxyl-terminal hydrolase 26 [Usp26]	ACGL**T**LFK	T446	X		
A2AKB9	DDB1- and CUL4-associated factor 10 [Dcaf10]	LPGGPAVSPAERAGEFAAPGALELSAA**T**AS ASQAKLSPS**SS**PR	T87; S99; S100		X	
		TTSSSDLTTT**S**SSSGSR	S364			X
O70263	E3 ubiquitin-protein ligase LNX [Lnx1]	**S**AP**S**SVVLK	S581; S584			X
Q5SQY9	E3 ubiquitin-protein ligase TRIM37 [Trim37] (fragment)	WAEEVTQQLD**T**LQLCSL**T**K	T63; T70			X
E9QQ33	E3 ubiquitin-protein ligase SH3RF1 [Sh3rf1]	V**S**MSTAGQASR	S523			X
Q8VCK5	Kelch-like protein 20 [Klhl20]	C**T**NIRPGETGMDVT**S**R	T5; S18			X
Q80VP0	Tectonin beta-propeller repeat-containing protein 1 [Tecpr1]	GHE**TS**GNTDH**ST**ENACLTEGK	T451; S452; S458; S459			X
Q62381-2	Tolloid-like protein 1 (isoform 2) [Tll1]	IG**S**GLEQNN**T**MK	S120; T127		X	
F6UTU1	Ubinuclein-2 [Ubn2] (fragment)	TAPS**T**TTSSNYLAK	T94	X		
		L**T**NSSS**T**GTAGK	T237; T242	X		X
Q6VNB8	WD repeat and FYVE domain-containing protein 3 [Wdfy3]	DT**S**SQPSST**S**HRPR	S3283; S3290			X