Chapter 9

TUBEs-Mass Spectrometry for Identification and Analysis of the Ubiquitin-Proteome

Mikel Azkargorta, Iraide Escobes, Felix Elortza, Rune Matthiesen, and Manuel S. Rodríguez

Abstract

Mass spectrometry (MS) has become the method of choice for the large-scale analysis of protein ubiquitylation. There exist a number of proposed methods for mapping ubiquitin sites, each with different pros and cons. We present here a protocol for the MS analysis of the ubiquitin-proteome captured by TUBEs and subsequent data analysis. Using dedicated software and algorithms, specific information on the presence of ubiquitylated peptides can be obtained from the MS search results. In addition, a quantitative and functional analysis of the ubiquitylated proteins and their interacting partners helps to unravel the biological and molecular processes they are involved in.

Key words Ubiquitin, Tandem Ubiquitin Binding Entities (TUBEs), Posttranslational modification (PTM), Collision induced dissociation (CID), Mass spectrometry (MS), Gene Ontology (GO), Iodoacetamide (IAA), Chloroacetamide (CAA)

1 Introduction

Protein ubiquitylation is of paramount importance for the proper function and development of multiple cellular processes including proteolyisis, endocytosis, DNA repair, cellular localization, or activation of protein kinases [1, 2]. Its deregulation has been shown to be involved in a number of diseases, such as cancer, neurodegenerative and cardiovascular diseases, and immunological disorders, among others [3, 4].

The ubiquitin-proteome is integrated by the total ubiquitylated proteins present in the cell and their interacting partners (ubiquitin-interactome). The ubiquitin-interactome allows regulation and connection of ubiquitylated proteins with the effector functions. Large-scale analysis of protein ubiquitylation by MS has become one of the most valuable techniques to elucidate its role in physiology and pathology. However, the analysis of ubiquitylated proteins can be a daunting task because of their low

Rune Matthiesen (ed.), *Proteostasis: Methods and Protocols*, Methods in Molecular Biology, vol. 1449,
DOI 10.1007/978-1-4939-3756-1_9,

stoichiometry and their short life-span due to the action of deubiquitylation enzymes (DUBs) [5]. Therefore, protection and enrichment methods are mandatory for their analysis. A number of specific isolation methods have been developed in the last years, including the use of tandem ubiquitin-binding entities (TUBEs) [6–9], the expression of tagged ubiquitin molecules [10, 11], or even the development of anti-ubiquitin antibodies in order to pick ubiquitylated peptides [12–14].

The identification and analysis of ubiquitylated proteins by mass spectrometry involves the in vitro enzymatic digestion of the proteins of interest and the analysis of the generated peptides. Information on the mass of the peptides and their corresponding fragments is collected and contrasted with the information compiled in databases using dedicated software and algorithms (Fig. 1).

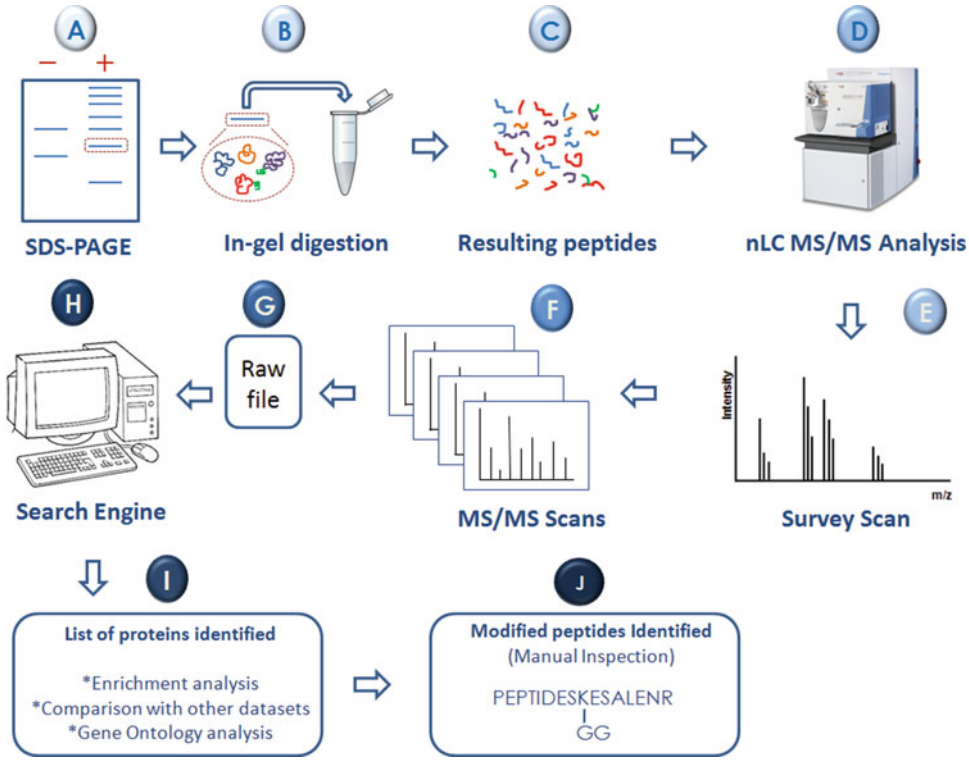

Fig. 1 Schematic overview of MS sample preparation. Processing starts from the SDS-PAGE run of the samples (**a**), followed by gel cutting and in-gel digestion of the obtained slices (**b**). The resulting peptides (**c**) are subjected to nLC-MS/MS analysis (**d**), where analysis of the peptides masses (**e**) and fragmentation patterns (**f**) is compiled. The RAW files containing this information (**g**) are loaded into the search engine (**h**), and a list of the identified proteins is obtained (**i**). Specific enrichment analysis can be carried out over this dataset, keeping only those proteins reliably enriched by TUBEs. Furthermore, this list can be compared with other datasets or subjected to functional analysis through the use of different bioinformatics tools, such as Gene Ontology (GO). Finally, modification site in a subset of identified proteins can be obtained (**j**). Manual inspection of spectra assignments is recommended in order to avoid false-positive assignments

Results provide direct information on the modified amino acids of the proteins through the identification of the GG ubiquitin signature. The tryptic cleavage of the ubiquitin sequence leaves a GG adduct attached to the substrate, increasing the mass of the peptide with 114.043 Da [15]. Due to incomplete digestion of the ubiquitin, trypsin may also leave a bigger tag comprised of LRGG, increasing the mass of the peptide with 383.228 Da [16] (Fig. 2). These mass shifts are indicative of ubiquitylation events, and are therefore used for the detection of the modified peptides in the data search step. However, identification of ubiquitylated peptides presents some drawbacks, since they are scarce in opposition to the non-modified tryptic peptides coming from the digestion of the modified protein, and are usually larger and get more charges than regular tryptic peptides [17], lowering the chances for their proper fragmentation and subsequent identification.

The experiments are typically done with at least three biological replicas, and negative controls are included for each pull down. For example, when using TUBEs pull down, beads cross-linked with GST can be used as a control. Using these negative controls, unspecifically enriched proteins can be discarded from the dataset and only those proteins more likely to be ubiquitylated (and their interacting partners) are considered for further analysis. As a starting point, this enriched dataset can be compared with other well-characterized datasets, and its functions can be outlined via a Gene Ontology term-enrichment analysis. Thus, a landscape of the molecular and biological processes these proteins are involved in can be obtained.

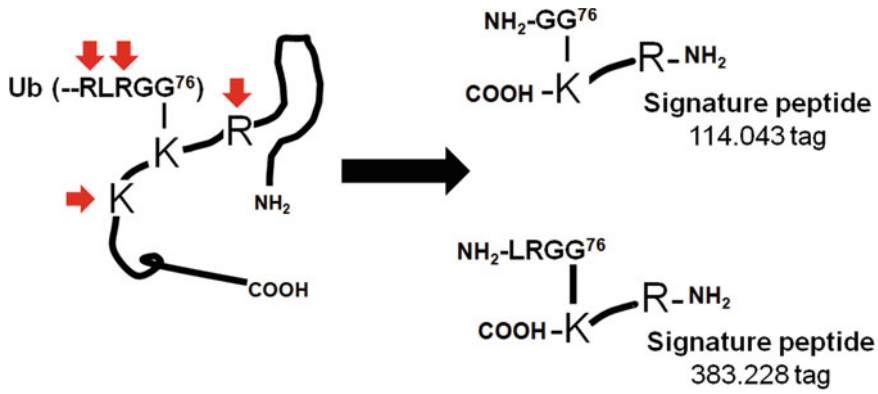

Fig. 2 Trypsin digestion of ubiquitylated proteins. Trypsin cleaves the protein sequence after K or R residues. The presence of an ubiquitin moiety attached to a K in the protein sequence usually hampers tryptic cleavage. Trypsin will cleave the available K and R in the sequence of both the protein and the ubiquitin attached to it, leaving a GG residue attached to the e-NH2 group of the ubiquitylated K. This tag can become an LRGG when trypsin fails to cut the ubiquitin sequence at the last R residue

2 Materials

2.1 Sypro Ruby Gel Staining and Image Acquisition

1. Fixing solution: 10 % Acetic acid and 30 % ethanol. Mix 100 mL acetic acid and 300 mL ethanol in a test tube, and make up to 1 L with Milli-Q water. Prepare it fresh.
2. Destaining solution: 7 % Acetic acid, 10 % ethanol. Mix 70 mL acetic acid and 100 mL ethanol in a test tube, and make up to 1 L with Milli-Q water. Prepare it fresh.
3. Wipe tissue: Precision Wipes Tissue Wipers.
4. Sypro Ruby protein gel stain, 1 L (Invitrogen).
5. Typhoon Trio Scanner, Variable Mode (GE Healthcare).

2.2 Digestion

1. 1 M Ammonium bicarbonate (AMBIC): Use Milli-Q water and store at −20 °C in 0.5 mL aliquots.
2. 100 mM AMBIC: Dilute a 0.5 mL aliquot of 1 M AMBIC to 5 mL with Milli-Q water. Prepare it fresh.
3. 50 mM AMBIC: Dilute a 0.5 mL aliquot of 1 M AMBIC to 10 mL with Milli-Q water. Prepare it fresh.
4. 1 M DTT stock: Dissolve DTT in Milli-Q water. Store at −20 °C in 10 μL aliquots.
5. 10 mM DTT: Dilute a 10 μL aliquot of 1 M DTT with 990 μL AMBIC 100 mM. Prepare it fresh.
6. 55 mM CAA: Dissolve CAA in AMBIC 100 mM. Prepare it fresh.
7. Trypsin Gold-Mass Spectrometry Grade 100 μg (Promega).
8. 1 μg/μL Trypsin Gold stock: Dissolve a vial of 100 μg trypsin gold in 0.1 ml of 50 mM acetic acid (*see* **Note 1**). Aliquots (15 μL) can be stored at −20 °C for at least 1 month.
9. 0.0125 μg/μL Trypsin Gold: Dilute the 15 μL aliquot of 1 μg/μL Trypsin Gold with 105 μL 50 mM AMBIC.
10. Trifluoroacetic acid (TFA) (Pierce).
11. 0.1 % TFA: Dissolve 0.1 mL TFA in 99.9 ml Milli-Q water.
12. Acetonitrile (ACN) (Symta).
13. Speed Vac: Rotational-Vacuum-Concentrator RVC 2–25 (Christ).
14. 50 mM Acetic acid: Dissolve 71.5 μL acetic acid in 24.28 mL Milli-Q water.
15. *E. coli* protein sample: ReadyPrep ™ *E. coli* Protein Sample 2.7 mg (Bio-Rad).

2.3 MS Analysis

1. Formic acid (FA) (Pierce).
2. 0.1 % FA: Dissolve 0.1 mL FA in 99.9 ml Milli-Q water.
3. Acetonitrile (ACN) (Symta).
4. Vials combination package (glass vial Type I) (Waters).

5. NanoAcquity UPLC System (Waters)
6. BEH C18 nanoACQUITY Column, 1.7 μm, 75 μm × 200 mm (Waters).
7. Symmetry C18 Trap Column, 5 μm, 180 μm × 20 mm (Waters).
8. Stainless steel emitters (Thermo Scientific).
9. Mass spectrometer for large-scale proteomics such as LTQ Orbitrap XL ETD Mass Spectrometer (Thermo Scientific).
10. Ultrasonic cleaning bath "Ultrasons" 6 L (J.P. Selecta).

3 Methods

Prepare all solutions using Milli-Q water (8 MΩ cm at 25 °C) and analytical grade reagents. Proceed with great care in order to avoid keratin contamination (*see* **Note 2**). Prepare and store all reagents at 4 °C (unless indicated otherwise) and follow all waste disposal regulations when disposing waste material.

3.1 Gel Staining and Image Acquisition

1. Following electrophoresis, pry the gel plates open with the use of a spatula. The gel remains on one of the glass plates. Rinse the gel with water and transfer carefully to a glass petri dish.
2. Fix the protein by the addition of 100 mL of fixing solution. Incubate for 30 min under gentle agitation, discard the solution and add 100 mL of SYPRO RUBY. Incubate overnight under agitation and in the dark (*see* **Note 3**).
3. Add 100 mL destaining solution and incubate under agitation for 30 min. Repeat this operation once. Replace the solution by 100 mL Milli-Q water, incubate for 10 min and replace the solution by 100 mL of fresh Milli-Q water and proceed to the image acquisition. Sypro Ruby images are acquired in the Typhoon Trio scanner-Variable Mode imager (GE Healthcare) using the program Typhoon scanner control v 5.0 (*see* **Note 4**).
4. Wash the scanner surface with ethanol and dry it with a wipe tissue before acquisition. Add Milli-Q water over the scanner´s surface and put the gel over the water carefully.
5. The parameters used for the acquisition are the following:
   - Acquisition mode: Fluorescence
   - Setup:
     - 610 BP 30 Deep Purple, Sypro Ruby
     - PMT: 535 V (*see* **Note 5**).
     - Laser: Blue (488) (*see* **Note 6**).
     - Sensitivity: Normal
   - Orientation: R
   - Focal plate: Platen

6. Acquire a preliminary image with a pixel size of 1000 μm and check that the selected area and laser voltage are suitable for the acquisition (*see* **Note 7**). Once the parameters are fine-tuned, acquire the image with a pixel size of 100 μm.
7. Save the image in TiFF format.
8. Once the image is acquired, remove the gel from the scanner and carefully bring it back to the Petri dish. Gels can be stored in Milli-Q water at 4 °C for at least 1 month. Clean the scanner surface with ethanol, and dry it with a wipe tissue.

3.2 Gel Cut and Digestion of the Gel Slices

Great care must be taken to avoid keratin contamination of the samples during the digestion step (*see* **Note 2**). The use of CAA instead of IAA as alkylating agent is recommended in order to avoid the generation of ubiquitylation-false positives [18].

1. Prepare a template for cutting the gel (*see* **Note 8**). Try to isolate clear bands in independent slices and keep the total number of slices limited to 10 (*see* **Note 9**) (Fig. 3a, b).
2. Print the gel image with the template at 100% of the gel image size. Put this image behind a clean glass plate, and put the gel in the upper part of this glass so that it fits the image behind it. Cut the gel following the template (*see* **Note 10**) (Fig. 3c).
3. Cut each gel slice into small pieces of approximately 1 mm^3 with a clean scalpel, and put them in a new identified Eppendorf tube with Milli-Q water (*see* **Note 11**).

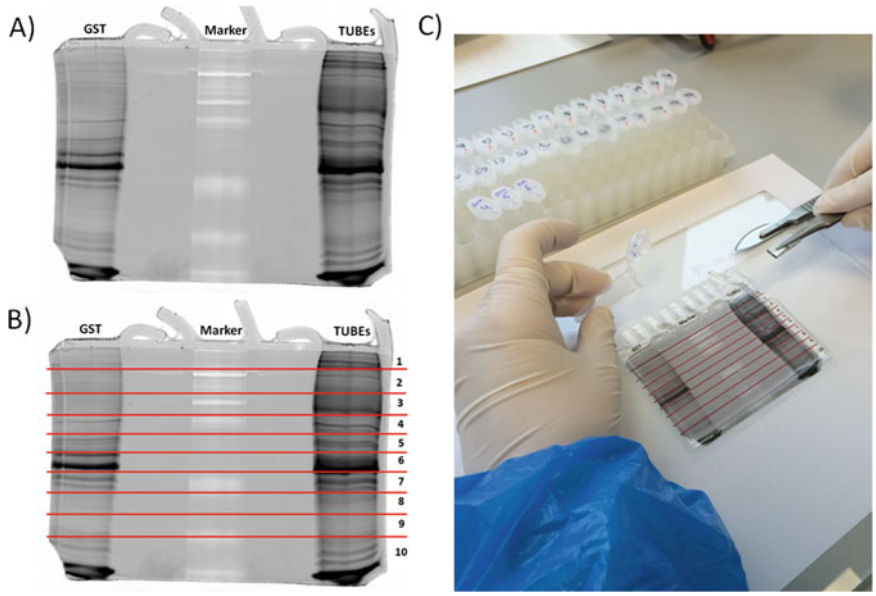

Fig. 3 Template and gel cutting. Once the gel image is acquired (**a**), a template for gel cutting can be created using different programs, such as Microsoft Powerpoint (**b**). Try to follow the band pattern of the gel and cut both the control and the sample following the same criteria to avoid variations in the pattern. Then, the template is printed at the gel size and placed behind the gel that is going to be cut (**c**). Cut the gel following the template and keep gel pieces in individual Eppendorf tubes

4. Discard Milli-Q water and add 50 µL of 50 mM AMBIC, vortex, incubate for 5 min, and discard supernatant. Repeat the procedure with 100 µL ACN (*see* **Note 12**).
5. Add 100 µL of a solution containing 10 mM DTT in 100 mM AMBIC, and incubate for 20 min at 56 °C under agitation. Discard the solution and add 55 mM CAA in 100 mM AMBIC (*see* **Note 13**). Incubate at room temperature for 30 min in the dark.
6. Wash the gel pieces adding ACN, vortex, incubate for 5 min, and discard supernatant.
7. Cover the gel with 50 µL 0.0125 µg/µL trypsin in 50 mM AMBIC. Allow the gel pieces to swell in ice for 30 min. If the gel dries out add more trypsin, and cover the gel gently.
8. Discard the trypsin supernatant and add 50 µL of AMBIC. Incubate at 37 °C overnight.
9. Add 100 µL of ACN, vortex, and incubate for 5–10 min. Put supernatants in a new microtube (one microtube per gel slice) (*see* **Note 14**).
10. Add 50 µL of 0.1% TFA in water, vortex, and incubate for 5–10 min. Add 100 µL of ACN, vortex, and incubate for 5–10 min (*see* **Note 14**). Add the supernatants corresponding to each sample in the previously identified microtubes, and dry vacuum them in the Speed-Vac (*see* **Notes 15** and **16**).

3.3 MS Analysis of the Samples

1. Resuspend the samples in 10 µL 0.1% FA, sonicate 5 min in the ultrasonic cleaning bath.
2. Put the resuspended samples on a vial and load the sample into the mass spectrometer (*see* **Note 17**).
3. Peptides are separated using a BEH130 C18 column, 75 µm × 200 mm, 1.7 µm coupled to a Symmetry 300 C18 UPLC Trap column, 180 µm × 20 mm, 5 µm (Waters) on a nanoACQUITY UPLC system (Waters).
4. The recommended chromatographic gradient includes the following steps (*see* **Note 18**):

Time (min)	A%	B%	Flow
0	97	3	0.3 mL/min
60	60	40	
61	15	85	
70	15	85	
72	97	3	
90	97	3	

A: FA 0.1% in H_2O
B: FA 0.1% in ACN

5. The MS acquisition method in the LTQ Orbitrap XL ETD includes the following parameters (*see* **Note 19**):
   - Full MS survey spectra (m/z 400 – 2000) are acquired in the orbitrap with 30,000 resolution at m/z 400.
   - Fragmentation of the six most intense precursors, with charge states equal to or greater than 2, by CID in the linear ion trap (*see* **Note 20**). Analyzed peptides are excluded from further analysis during 30 s using dynamic exclusion lists.

3.4 MS Data Analysis

As mentioned, protein identification and peptide modification assignment are carried out by searching the acquired peptide spectra in databases, such as UniProt or NCBI. Database search is carried out using different search engines, that is, algorithms that attempt to identify peptide sequences from the fragment ion spectra of the peptides in the dataset [19]. Mascot, Sequest, OMSSA, or VEMS, among others, are examples of search engines used for protein identification [20, 21]. Typical ubiquitylation tags (114.043 Da for GG, 383.228 Da for LRGG) must be considered when searching the spectra in order to find modified peptides. Furthermore, the tag LRGG gives the diagnostic ions 270.1925 (b2) and 384.2354 (b4) in MSMS which can be used to fish out potential MSMS spectra from ubiquitin modified peptides. The diagnostic ions are especially useful if MSMS spectra with high mass accuracy is available (<10 ppm).

1. Once the acquisition has ended, the generated unprocessed data are loaded into the search engine in order to identify the detected peptides and proteins.
2. Recommended search parameters include the following:
   - Carbamidomethylation of cysteines as fixed modification. Oxidation of methionines, GG (+114.043 Da) and LRGG (+383.228) modification of lysines, and protein N-terminal acetylation as variable modifications (*see* **Note 21**).
   - Peptide mass tolerance of 10 ppm and 0.5 Da fragment mass tolerance, and four missed cleavages allowed (*see* **Note 22**).
3. A decoy search is recommended in order to estimate the false discovery rate (FDR) for the samples. Once identified, selected proteins can be subjected to the functional analysis step (*see* **Note 23**).
4. Information on the presence of ubiquitylated peptides among the identified proteins can specifically be obtained by looking for those peptides carrying typical ubiquitin-modification (GG +114.043, LRGG +328.228) (*see* **Note 24**) (Fig. 4).

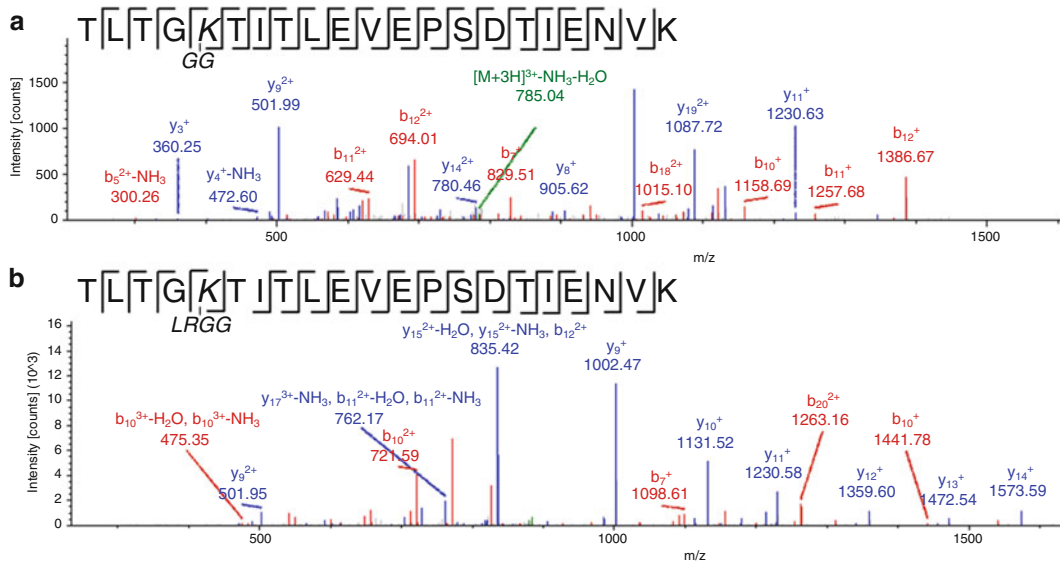

Fig. 4 Ubiquitylated peptide spectra examples. Two spectra for the ubiquitylated form of the peptide TLTGKTITLDVEPSDTIENVK, corresponding to K48 chains from polyubiquitin C (UBC_HUMAN) are provided, carrying the GG tag (**a**) or the LRGG tag (**b**) characteristic of ubiquitylation

3.5 Statistical Analysis

Once proteins are identified in both the TUBEs and negative controls (GST), it is necessary to discard unspecifically bound proteins and keep only the ubiquitylated proteins and their partners. In the example provided in this chapter, taken from Lopitz-Otsoa [22], a direct subtraction was performed. MCF7 cells treated with adriamycin were used for the characterization of the global ubiquitylation events in these cells, with the aim of pinpointing potential biomarkers and drug targets. Proteins identified in the GST controls were directly discarded from the dataset, giving a total of 643 proteins specifically bound to TUBEs. Of these, 269 were proteins consistently present in the replicates. This was the set of proteins considered as reliably enriched, and therefore further characterized in this work.

In addition to the direct subtraction of the identified proteins, relative quantitation of peptides and proteins can be carried out in order to make a more comprehensive enrichment analysis. The quantitative values can be obtained using experimental methods such as SILAC, stable isotope dimethyl labeling, tandem tags, area under the ion counts in the survey scans (XIC), or spectral counting. Matthiesen et al. [23] provides a review dealing with different MS-based quantitative methods. Then, multivariate analysis can be done in several software such as Excel, the statistical programming language R or Matlab. Below is a possible outline of the steps included in the multivariate analysis of the results.

1. Log transformation of the quantitative values. This will make the values normally distributed and lead to smaller *p*-values if a *t*-test is subsequently performed.
2. Optional normalization of the values across samples. For example, the R package "limma" supports a number of normalization procedures. References Kroll et al. and Bolstad et al. [24, 25] provide comprehensive overviews of different normalization procedures.
3. Subtraction of background values obtained from the control experiments, e.g., spectral counts from GST beads.
4. Calculate log ratios and *p* values to define difference in the level of ubiquitin modifications between experimental conditions (*see* **Note 25**). If the experiment has multiple conditions then ANOVA can be performed followed by a post-hoc test to lower the number of statistical tests.

3.6 Meta-Analysis

Frequently meta-analysis in proteomics starts out by comparing the identified proteins with proteins identified under different conditions or experimental settings. For example, in the work published by Lopitz-Otsoa et al. the set of enriched proteins was compared with the results obtained by other methods for the isolation and analysis of ubiquitylated proteins. Comparative analysis can also be done with well-known data databases such as UniProt or PhosphoSite. In this example, the Venn diagram in Fig. 5 compares TUBE-enriched proteins from Lopitz-Otsoa [22] against all O-GlcNAc-, SUMO-, and ubiquitin-annotated proteins in PhosphoSite (made with the R package "VennDiagram" which can plot Venn diagrams with up to five groups). The crosstalk between ubiquitin and different PTMs is the next level of complexity in the molecular regulation of multiple biological processes. SUMOylation and O-GlcNAcylation are known to be connected to ubiquitylation for the regulation of different functions [26–28]. Additionally, O-GlcNacylation and SUMOylation have been described to control transcription in the nucleus. In this context, the analysis of their correlation within the dataset enriched in the present analysis may be of great interest for a further characterization of these proteins.

Furthermore, proteins that are defined as significantly enriched in Subheading 3.5 above can be subjected to a functional enrichment analysis. A simple way to perform enrichment analysis is to submit the enriched protein IDs (*see* **Note 26**) to DAVID bioinformatics server [29, 30] and then export the result as text tables. The tab delimited text tables can then directly be imported in to R or Excel to produce summary graphics. For example, in Fig. 6 the ten most significant biological process categories were displayed for TUBE-enriched proteins and O-GlcNAc- and SUMO-annotated proteins in PhosphoSite.

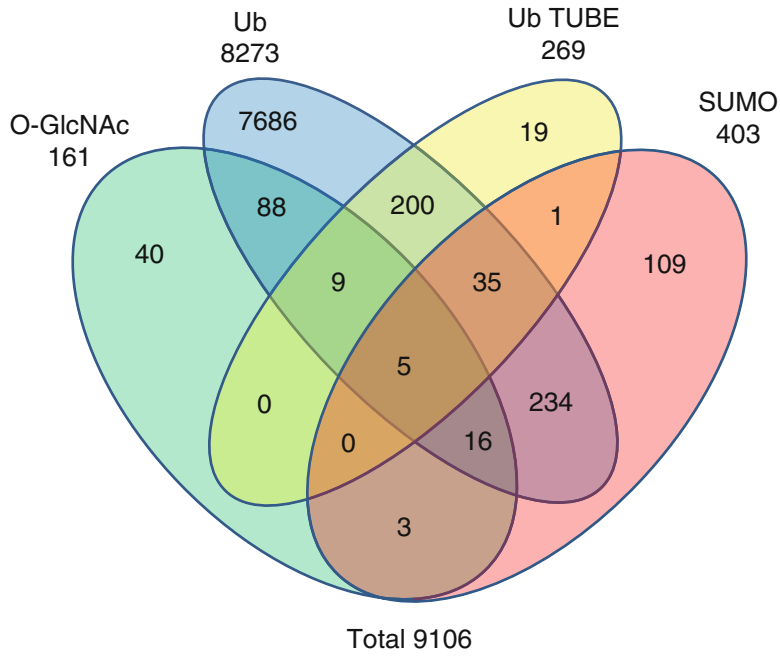

Fig. 5 Comparing all O-GlcNAc-, ubiquitin-, and SUMO-annotated proteins from PhosphoSite (Date 14-4-2015) with proteins reproducibly enriched by TUBEs and LC-MS identified by Lopitz-Otsoa et al. [22]

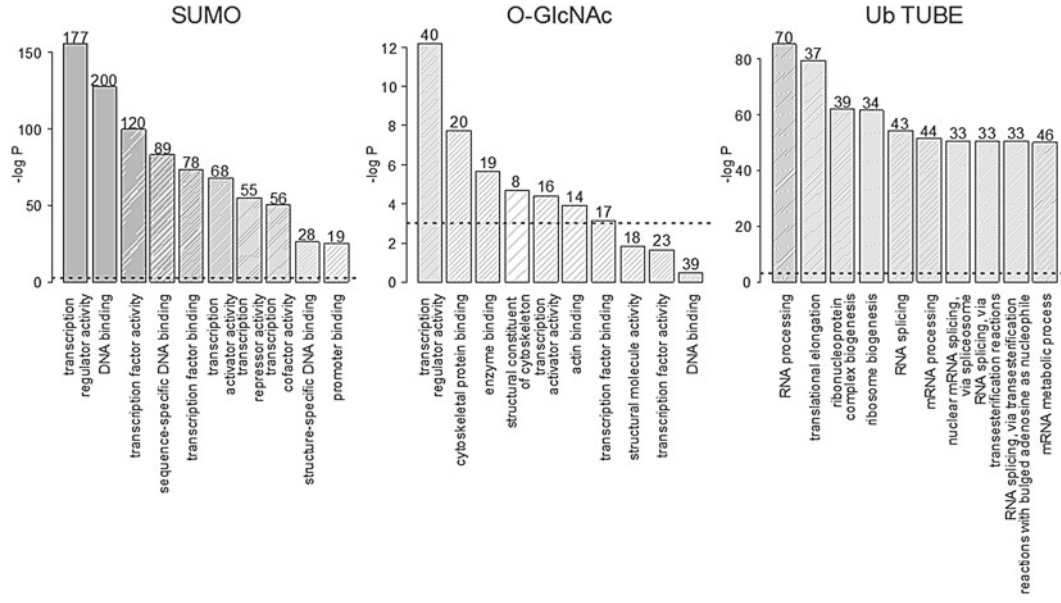

Fig. 6 Functional enrichment analysis by DAVID of the TUBEs-enriched proteins, and the SUMO- and O-GlcNAc-annotated proteins from PhosphoSite. Minus log of the FDR corrected significance of enrichment is indicated on the y-axis. The numbers on top of each bar indicate the number of proteins identified for each category

4 Notes

1. Resuspension in acid pH is necessary for the storage of trypsin, given that it prevents self-digestion events taking place in basic pH. However, if the whole aliquot is going to be used, the trypsin vial can be resuspended directly in 1 mL 50 mM AMBIC and then diluted to 8 mL with AMBIC 50 mM to achieve a final concentration of 0.0125 μg/μL. Use the trypsin immediately after resuspension and discard remnants, if any.
2. The use of a clean lab coat, disposable over-sleeves, and a cap is strongly recommended throughout the whole process in order to avoid keratin contamination. All digestion steps must be carried out in an isolated room. Clean all material with ethanol before use.
3. A few hours may be enough for protein detection, but overnight incubation of the gel in SYPRO is recommended for maximum sensitivity.
4. Switch the scanner on at least 30 min before image acquisition in order to warm up the system. The use of alternative up-to-date systems, such as the Versadoc Molecular Imager (Bio-Rad) is also a viable option for the image acquisition.
5. Laser gain values are illustrative. This value can be increased when the signal is weak, or decreased when saturated images are obtained, but it can be considered as a starting point.
6. The laser and filter setup used for this acquisition do not match with the default setup considered by the system. Therefore, a warning advice may appear when setting up the parameters. Ignore this advice and proceed with the acquisition.
7. Avoid saturation of the image and make sure that all the interesting parts of the gel are scanned before acquiring the image at high resolution. Low-resolution scans are much faster than high resolution, and therefore more suitable for the optimization of the image acquisition.
8. We used Microsoft Powerpoint, but programs intended for similar purposes can be used.
9. The number of slices may change depending on the pattern of the lane. However, keeping a reduced number of slices is a good idea in order not to increase too much the effort in the LC-MS side.
10. For a more dedicated cut, a UV transilluminator can be used. Otherwise, the entire gel lane can be cut in equal consecutive slices but the explained methodology is recommended.
11. Gel slices of a total *E. coli* extract processed in a similar way can be used in order to check the digestion process. These slices

should have a similar volume as the slices under analysis in the experiment.

12. The volumes provided are illustrative, and bigger or smaller volumes may be added. Use a volume that allows full coverage of the gel pieces.
13. Avoid the use of Iodoacetamide (IAA). IAA can artificially modify lysines with the addition of two acetamide moieties, which has the same molecular weight and chemical formula as the diglycine modification, and is completely undistinguishable by MS [18]. CAA does not provoke this effect, and therefore it is more suitable for this analysis.
14. Incubate the gel slices in ACN until they get white opaque.
15. Switch the Speed-Vac on at least 20 min before use in order to cold-up the trap.
16. Complete dryness is not recommended. Special care must be taken not overdrying the samples for a good sample recovery.
17. Sample resuspension and load may depend on the starting amount.
18. The columns and gradient are illustrative. However, a linear gradient followed by a washing step and an equilibration step are needed as part of the protocol. Length of each phase may depend on sample load, sample complexity, and/or column length among others. Adaptation to the system and optimization are therefore needed.
19. The parameters for the chromatography and the MS acquisition are illustrative. Dedicated methods and parameter optimization may be necessary for the acquisition with different equipment and samples.
20. The use of different fragmentation methods, such as ETD, has been shown to be a good alternative to CID, providing alternative fragmentation patterns. However, the method of choice for general purposes is still CID due to its ease of use and fragmentation capacity, and therefore its use is recommended.
21. If the presence of any other modification is suspected, consider it as part of the search, taking into account that it may increase search time and modify search space.
22. The parameters are typical for searching LTQ Orbitrap data. However, they should be adapted to the specific needs and characteristics of the equipment used. Four missed cleavages are allowed since the presence of ubiquitin moieties attached to the proteins is known to hamper tryptic cleavage of such residues, and therefore a high number of missed cleavages can be expected for highly modified peptides.
23. Decoy searches are recommended for complex samples. When sample complexity is low, however, its use is not recommended,

since FDR calculations might not be accurate. The threshold for protein selection may change depending on the sample. Selection of proteins with at least two peptides with a FDR < 5 % (or a Mascot *p*-value < 0.05 in the absence of an FDR estimation) or selection of the proteins with at least one peptide with a FDR < 1 % (or a Mascot *p*-value < 0.01 in the absence of an FDR estimation) are commonly used thresholds.

24. As mentioned, careful inspection of the spectra is recommended for avoiding false-positive assignments. The selection of spectra with the presence of fragments covering most of the peptide sequence and clearly assigning the modification site is recommended. Overcoming false positives is still one of the major issues when analyzing ubiquitylated proteins and peptides. The mass shift provoked by the GG addition is isobaric to many other chemical modifications, such as hydroxypropylation, asparagylation, or aspartylation, among others [17], and may therefore give rise to the detection of false positives. In addition, IAA can introduce false positives, as mentioned before. The use of high-accuracy mass spectrometers, such as the LTQ-Orbitrap XL ETD used in our approach, and the use of chloroacetamide (CAA) instead of IAA as alkylating agent during the protein digestion significantly reduce the number of false-positive assignments. However, most search engines lack robust enough tools for the unsupervised analysis and assignment of PTMs and therefore careful examination of the spectra is necessary to provide a reliable dataset of ubiquitylated peptides.
25. Published studies frequently do not provide *p*-values of enriched ubiquitin peptides because of high variance between samples. Frequently arbitrary thresholds are defined such as 1.5–2-fold enriched in a minimum number of biological replicas followed by for example validation by Western blot.
26. We find that DAVID provides a better mapping if the protein IDs and accession numbers are trimmed for version numbers. Furthermore, downloading the latest version of Gene Ontology annotation for the species of interest, and then manually map the enriched genes/proteins to Gene Ontology followed by enrichment statistics provides an even better mapping and more accurate results.

Acknowledgments

This work was funded by the Ministerio de Economia y Competitividad, Spain grant BFU2011-28536 (MSR), the Industry, Trade and Tourism Department of the Basque Government—SAIOTEK program, and Diputación Foral de Gipuzkoa (MSR). RM is funded by FCT investigator program 2012. MA, IE, and FE

are part of the Proteomics Platform at CIC bioGUNE, which is member of ProteoRed-ISCIII and CIBERehd and funded by Basque Government (ETORTEK Program) and Instituto de Salud Carlos III. The authors would like to acknowledge networking support by the Proteostasis COST Action (BM1307).

References

1. Deng L, Wang C, Spencer E, Yang L, Braun A, You J, Slaughter C, Pickart C, Chen ZJ (2000) Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. Cell 103(2):351–361
2. Wilkinson KD, Tashayev VL, O'Connor LB, Larsen CN, Kasperek E, Pickart CM (1995) Metabolism of the polyubiquitin degradation signal: structure, mechanism, and role of isopeptidase T. Biochemistry 34(44):14535–14546
3. Mata-Cantero L, Lobato-Gil S, Aillet F, Rodriguez MS (2015) The ubiquitin-proteasome system (UPS) as a cancer drug target: emerging mechanisms and therapeutics. In: Wondrak SL (ed) Stress response pathways in cancer. Springer Science + Business Media, Dordrecht, pp 225–264. doi:10.1007/978-94-017-9421-3_11
4. Xolalpa W, Perez-Galan P, Rodriguez MS, Roue G (2013) Targeting the ubiquitin proteasome system: beyond proteasome inhibition. Curr Pharm Des 19(22):4053–4093
5. Amerik AY, Hochstrasser M (2004) Mechanism and function of deubiquitinating enzymes. Biochim Biophys Acta 1695(1–3):189–207. doi:10.1016/j.bbamcr.2004.10.003
6. Aillet F, Lopitz-Otsoa F, Hjerpe R, Torres-Ramos M, Lang V, Rodriguez MS (2012) Isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. Methods Mol Biol 832:173–183. doi:10.1007/978-1-61779-474-2_12
7. Hjerpe R, Aillet F, Lopitz-Otsoa F, Lang V, England P, Rodriguez MS (2009) Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. EMBO Rep 10(11):1250–1258. doi:10.1038/embor.2009.192
8. Hjerpe R, Rodriguez MS (2008) Efficient approaches for characterizing ubiquitinated proteins. Biochem Soc Trans 36(Pt 5):823–827. doi:10.1042/BST0360823
9. Lopitz-Otsoa F, Rodriguez MS, Aillet F (2010) Properties of natural and artificial proteins displaying multiple ubiquitin-binding domains. Biochem Soc Trans 38(Pt 1):40–45. doi:10.1042/BST0380040
10. Franco M, Seyfried NT, Brand AH, Peng J, Mayor U (2011) A novel strategy to isolate ubiquitin conjugates reveals wide role for ubiquitination during neural development. Mol Cell Proteomics 10(5):M110 002188. Doi:10.1074/mcp.M110.002188
11. Lectez B, Migotti R, Lee SY, Ramirez J, Beraza N, Mansfield B, Sutherland JD, Martinez-Chantar ML, Dittmar G, Mayor U (2014) Ubiquitin profiling in liver using a transgenic mouse with biotinylated ubiquitin. J Proteome Res 13(6):3016–3026. doi:10.1021/pr5001913
12. Sowa ME, Bennett EJ, Gygi SP, Harper JW (2009) Defining the human deubiquitinating enzyme interaction landscape. Cell 138(2):389–403. doi:10.1016/j.cell.2009.04.042
13. Udeshi ND, Mertins P, Svinkina T, Carr SA (2013) Large-scale identification of ubiquitination sites by mass spectrometry. Nat Protoc 8(10):1950–1960. doi:10.1038/nprot.2013.120
14. Xu G, Paige JS, Jaffrey SR (2010) Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. Nat Biotechnol 28(8):868–873. doi:10.1038/nbt.1654
15. Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP (2003) A proteomics approach to understanding protein ubiquitination. Nat Biotechnol 21(8):921–926. doi:10.1038/nbt849
16. Shi Y, Xu P, Qin J (2011) Ubiquitinated proteome: ready for global? Mol Cell Proteomics 10(5):R110 006882. Doi:10.1074/mcp.R110.006882
17. Xu G, Jaffrey SR (2013) Proteomic identification of protein ubiquitination events. Biotechnol Genet Eng Rev 29:73–109. doi:10.1080/02648725.2013.801232
18. Nielsen ML, Vermeulen M, Bonaldi T, Cox J, Moroder L, Mann M (2008) Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. Nat Methods 5(6):459–460. doi:10.1038/nmeth0608-459
19. Shteynberg D, Nesvizhskii AI, Moritz RL, Deutsch EW (2013) Combining results of multiple search engines in proteomics. Mol Cell Proteomics 12(9):2383–2393. doi:10.1074/mcp.R113.027797

20. Carvalho AS, Ribeiro H, Voabil P, Penque D, Jensen ON, Molina H, Matthiesen R (2014) Global mass spectrometry and transcriptomics array based drug profiling provides novel insight into glucosamine induced endoplasmic reticulum stress. Mol Cell Proteomics 13(12):3294–3307. doi:10.1074/mcp.M113.034363
21. Eng JK, Searle BC, Clauser KR, Tabb DL (2011) A face in the crowd: recognizing peptides through database search. Mol Cell Proteomics 10 (11):R111 009522. Doi:10.1074/mcp.R111.009522
22. Lopitz-Otsoa F, Rodriguez-Suarez E, Aillet F, Casado-Vela J, Lang V, Matthiesen R, Elortza F, Rodriguez MS (2012) Integrative analysis of the ubiquitin proteome isolated using tandem ubiquitin binding entities (TUBEs). J Proteomics 75(10):2998–3014. doi:10.1016/j.jprot.2011.12.001
23. Matthiesen R, Carvalho AS (2013) Methods and algorithms for quantitative proteomics by mass spectrometry. Methods Mol Biol 1007:183–217. doi:10.1007/978-1-62703-392-3_8
24. Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19(2):185–193
25. Kroll TC, Wolfl S (2002) Ranking: a closer look on globalisation methods for normalisation of gene expression arrays. Nucleic Acids Res 30(11):e50
26. Aillet F, Lopitz-Otsoa F, Egana I, Hjerpe R, Fraser P, Hay RT, Rodriguez MS, Lang V (2012) Heterologous SUMO-2/3-ubiquitin chains optimize IkappaBalpha degradation and NF-kappaB activity. PLoS One 7(12):e51672. doi:10.1371/journal.pone.0051672
27. Guinez C, Mir AM, Dehennaut V, Cacan R, Harduin-Lepers A, Michalski JC, Lefebvre T (2008) Protein ubiquitination is modulated by O-GlcNAc glycosylation. FASEB J 22(8):2901–2911. doi:10.1096/fj.07-102509
28. Tatham MH, Geoffroy MC, Shen L, Plechanovova A, Hattersley N, Jaffray EG, Palvimo JJ, Hay RT (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. Nat Cell Biol 10(5):538–546. doi:10.1038/ncb1716
29. da Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4(1):44–57. doi:10.1038/nprot.2008.211
30. da Huang W, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37(1):1–13. doi:10.1093/nar/gkn923