Published in final edited form as:

Methods Enzymol. 2019 ; 626: 323–346. doi:10.1016/bs.mie.2019.06.025.

Analysis of the topology of ubiquitin chains

Lucia Geis-Asteggianti[†], Amanda E. Lee^{†,‡}, Catherine Fenselau^{*}

Department of Chemistry and Biochemistry, University of Maryland, College Park, MD, United States

Abstract

The small protein ubiquitin and its multiple polymers are encountered free in cells and as post-translational modifications on all proteins. Different polyubiquitin three dimensional structures are shown to correlate uniquely with different cellular functions as part of the diverse ubiquitin signaling. At the same time, this multiplicity of structures provides serious challenges to the analytical biochemist. Globally applicable strategies are presented here for the analyses of polyubiquitins and of ubiquitinated proteins, which take advantage of the speed, specificity and sensitivity of top-down tandem mass spectrometry. Particular attention is given to the supervised interpretation of fragmentation as revealed in the MS/MS spectra of these branched proteins. The strategy is compatible with any MS activation technology, is applicable to all polyubiquitin linkage and chain types, can be extended to ubiquitin-like proteins, and will be compatible with and enhanced by continuing advances in LC-MS/MS instrumentation and interpretation software.

1. Introduction

Ubiquitin (Ub) is a small protein conserved in eukaryotes throughout the evolutionary chain, which has been shown to be required for cell survival and proliferation (Varshavsky, 2012). In prokaryotic cells the ubiquitin-like protein This has been shown to have similar functions (Xu et al., 2015). Other ubiquitin-like proteins have also been characterized in a range of eukaryotic cell types, including Rub1 (Nedd8 in mammals) (Jones et al., 2008) and SUMO (Bruderer et al., 2011), which have functions complementary to Ub. Ub occurs free as a monomer and as polymers (polyUb) of various topologies. Topology refers to the spatial arrangement of the Ub subunits and is determined by the identities of the linkages between ubiquitin subunits (Varadan et al., 2004; Varadan, Walker, Pickart, & Fushman, 2002). Both monomers and polymers control the fate of proteins to which they become covalently attached. Protein (poly) ubiquitination has been linked to many cellular processes, including protein degradation (Chau et al., 1989), endocytosis (Boname et al., 2010; Goto et al., 2010), DNA repair (Chen & Chen, 2013) and tumor morphology (Mourtzoukou, Drikos, Goutas, & Vlachodimitropoulos, 2018). Numerous other functions have been linked to Ub and are explored in a review by Swatek and Komander (2016).

^{*}Corresponding author: fenselau@umd.edu.

[†]These authors contributed equally.

[‡]Current address: Thermo Fisher Scientific, West Palm Beach, FL, United States.

A key finding is that these different cellular responses to ubiquitin signaling are directed primarily by the different structures that polyUb can take— i.e., the chain topology and linkages present (Akutsu, Dikic, & Bremm, 2016). Ub chains are connected through the formation of isopeptide bonds between the carboxyl terminus (G76) of one Ub moiety and the ϵ -amine of a lysine (K) residue in another. As a post-translational modification (PTM) monoubiquitination (monoUb) is defined as Ub attached to a substrate protein through the addition of a single Ub (Fig. 1A) or multiple Ubs (Fig. 1B). A polyUb chain not bonded to a substrate protein is known as an unanchored chain (Fig. 1C), while a chain attached to a substrate protein is anchored (Fig. 1D). Ub chains can comprise the same positional linkages (homotypic) or linkages at multiple sites (heterotypic). Polymers can be linear, or branched as shown in (Fig. 1D). For example, tetrameric Ub chains can exist theoretically in 819 different isomeric structures. PolyUb PTMs can exist with many potential topologies, some with proven and others with potential abilities to affect different functions within the cell. Full characterization of Ub chains is imperative to understand the roles this PTM family plays in diverse biological processes. Thus far, characterization of the ubiquitinome has been limited (Nguyen, Dobrzy ki, Fey, & Kholodenko, 2014). Existing methods have not allowed for facile identification of structures of polyUb PTMs. PolyUb chains can also be modified, which mediates their activities and further confounds characterization. For example, phosphorylation of PolyUb chains leads to a genuine Parkin receptor, involved in the depolarization of mitochondria (Okatsu et al., 2015).

Since the first analysis of a histone H2B proteoform modified by an ubiquitin moiety (Thorne, Sautiere, Briand, & Crane-Robinson, 1987), the positions of most ubiquitination events have been characterized by recognition of the glycylglycine (GG) tag left behind after digestion with trypsin. This 114 Da mass tag is compatible with peptide mass spectrometry (MS) and the strategy has significantly widened our knowledge of the extent of protein modification in vivo. However, the tag offers no information about the structure of a monoUb or *polyUb* PTM. Immunoprecipitation has been used in conjunction with GG tag analysis to characterize linkages (e.g., in western blot experiments) and/or to isolate proteins modified with the targeted linkage for further analysis. Many of the biological functions listed above were correlated with polyUb topology using such strategies. Reiterative analyses are required when the polyUb has multiple types of linkages and the sequence of linkages may not be defined (Newton et al., 2008). It has been shown that some antibodies do not isolate the same cohort reproducibly (Gilda et al., 2015). The reiterative use of selective deubiquitinases (DUBs) has been proposed for analysis of chains, including those with heterotypic linkages, without mass spectrometry (Hospenthal, Freund, & Komander, 2013; Hospenthal, Mevissen, & Komander, 2015). DUBs have not been identified yet for some less common linkages.

Here, we describe a strategy for analysis of polyUb chains that is universally applicable to all linkages. The methodology presented is potentially applicable to all chain types, and it can easily integrate future advances in analytical instrumentation and software. As currently practiced, enrichment of the sample is beneficial.

2. Protocol for sample analysis

The protocol below details the analysis of polyubiquitin chains, ubiquitinated proteins and rubylated proteins by top-down tandem MS. The strategy has been established utilizing 5 synthetic ubiquitin conjugates, 3 rubylated conjugates and 20 synthetic ubiquitin chains (7 dimers, 6 trimers, 6 tetramers and 1 pentamer). These proteins were obtained by different synthetic approaches, such as using linkage-specific enzymes or a nonenzymatic chain assembly strategy. More information on the synthesis of the polyubiquitin chains, ubiquitinated conjugates and rubylated conjugates can be found elsewhere (Castañeda, Liu, Chaturvedi, et al., 2011; Castañeda, Liu, Kashyap, et al., 2011; Chen et al., 2018; Dixon, Castaneda, Kashyap, Wang, & Fushman, 2013; Gomes et al., 2019; Nakasone, Livnat-Levanon, Glickman, Cohen, & Fushman, 2013; Varadan et al., 2002, 2004). In all cases protein purity was evaluated by SDS-PAGE. The syntheses yields ranged between 9% and 15%, providing protein quantities of 0.3–1 mg. All samples were lyophilized and reconstituted in water:acetonitrile (97.5, 2.5) with 0.1% formic acid to a final concentration of 30 pg/mL prior to LC-MS/MS analysis.

2.1 Liquid chromatography

The LC system used is an Ultimate 3000 ultra-high-performance liquid chromatograph (Thermo Fisher, San Jose, CA) controlled by Chromeleon Xpress software. Sample desalting, concentration and separation was achieved using a PepSwift RP-4H monolith trap (100 μ m x 5 mm) and a ProSwift RP-4H monolith analytical column (200 μ m x 25 cm), respectively.

2.1.1. Mobile phases:

- a. Prepare 1L of mobile phase A by combining water-acetonitrile in a (97.5,2.5) ratio. Add formic acid (>99% purity) to reach 0.1% (v/v).
- b. Prepare 1L of mobile phase B by combining water-acetonitrile in a (25,75) ratio. Add formic acid to reach 0.1% (v/v).

Note: All mobile phases are prepared using LC-MS grade reagents.

2.1.2. Purge the system with the prepared mobile phases.

2.1.3. Set the autosampler temperature to 4°C.

2.1.4. Condition the trap and analytical column: Set the loading flow rate to 5 μ L/min and the analytical column flow rate to 1.5 μ L/min. Set the column temperature to 35 °C. Flush the system with mobile phase A for 30 min.

Note: The column temperature can be increased to improve separation.

2.1.5. Reconstitute the polyubiquitin sample into mobile phase A to reach a concentration of at least 30 μ g/mL.

2.1.6. Load the sample:

- a. Load 3 μL of sample (equivalent to at least 90 ng) into the trap column.
- b. Desalt and concentrate the sample by flushing the trap with mobile phase A for 5 min at a flow rate of 5 $\mu\text{L}/\text{min}$.

2.1.7. Separate the concentrated sample under a linear gradient from 5% to 55% mobile phase B over 20min at a flow rate of 1.5 $\mu\text{L}/\text{min}$. Note that longer gradients may be necessary depending on the sample complexity.

2.2 Tandem mass spectrometry

Tandem mass spectra were acquired using an orbitrap Fusion Lumos tribrid mass spectrometer (Thermo Fisher) controlled by Xcalibur 2.3 software. Activation of precursor ions for fragmentation was achieved by combining electron transfer dissociation (ETD) and collision induced dissociation (CID)—termed ETciD—or combining ETD and higher-energy CID (HCD)—termed EThcD.

2.2.1. Set ion routing multipole (IRM) pressure to 0.01 or 0.03mTorr and perform instrument calibration using Thermo Fisher positive ion calibration solution. Higher fragment ion density was observed when using ETciD with IRM pressure of 0.01mTorr or EThcD with IRM pressure of 0.03mTorr.

2.2.2. Set in-source fragmentation energy to 10V.

2.2.3. Set mass resolution to 120,000 at 200 m/z for both precursor and fragment ions.

Note: For ubiquitin conjugates or polyubiquitin chains of large mass (>50kDa), precursor ion mass resolution is set to 15,000.

2.2.4. Set number of microscans to 20 for both MS and MS/MS acquisition.

2.2.5. Set data-dependent mode for MS/MS acquisition to select the top 2 most abundant m/z once in a 30s period prior to exclusion for 60s.

Note: For improved sensitivity a precursor ion m/z list can be included, which will be given preference when selecting ions.

2.2.6. Set precursor ion fragmentation:

- a. For ETciD: set ETD reaction time to 3 s after which supplemental CID activation of 10% normalized collision energy is applied.
- b. For EThcD: set ETD reaction time to 6 s and supplemental CID activation of 10% normalized collision energy.

3. Data processing methods

The identification of substrate proteins by top-down MS was performed by database search using ProSightPD node 1.1 (Proteinaceous, Inc., <http://proteinaceous.net>) integrated into Proteome Discoverer (PD) 2.2 (Thermo Fisher). Note that other freely available top-down MS software such as TopPIC Suite (Indiana University-Purdue University Indianapolis,

<http://proteomics.informatics.iupui.edu/software/toppic/index.html>) or MSPathFinder (Pacific Northwestern National Laboratory, <https://omics.pnl.gov/software/mspathfinder>) can also be used. Data processing of ubiquitin chains was carried out using Xtract 3.0 inside Xcalibur 3.0 (Thermo Fisher) software for mass spectra deconvolution, and ProSight Lite (Northwestern University, <http://prosightlite.northwestern.edu>) for aiding tandem mass spectra graphical interpretation.

3.1 Recognition and identification of proteins modified by ubiquitin

The strategy below is focused on proteins modified by ubiquitin, however, it can also be applied to recognize proteins modified by other small protein modifiers such as SUMO or Rub1/Nedd8 as recently reported by Gomes et al. (2019).

3.1.1. Select the following ProSightPD nodes:

- a. Spectrum Selector
- b. Top-Down High/High cRAWler. Alternatively, the Top- Down Low/High cRAWler node can be selected if a precursor ion mass resolution of 15,000 was used.
- c. Absolute Mass Search.

3.1.2. Define and load an appropriate protein sequence database.

3.1.3. Set precursor mass tolerance to 9000Da to accommodate for monoubiquitination.

Note: A wider precursor mass tolerance window can be set when searching for proteins carrying a polyubiquitin chain.

3.1.4. Set fragment ion mass tolerance to 4ppm.

3.1.5. Add static modifications (if needed).

3.1.6. Set a protein identification FDR threshold 1% or E-value threshold 1×10^{-4} .

3.1.7. Perform search and inspect the candidate proteins list by checking the precursor ion's retention time and charge assignment. For conjugates, ubiquitin is expected to be identified alongside the substrate protein

3.1.8. Compare the identified ubiquitinated protein's observed intact mass and theoretical mass. Mass differences <0.5 Da are considered acceptable, when all other post-translational or chemical modifications are taken into account.

3.1.9. Visualize matching fragment ions. In the case of samples analyzed with ProSightPD, fragment ions are directly visualized in ProSight Lite, a tool integrated into the software. When other software are utilized, obtain the list of fragment ion decharged masses and continue to Section 3.3.

3.2 Deconvolution of tandem mass spectra

The analysis of ubiquitin chains or ubiquitinated proteins of known identity do not require performing a database search. The decharged intact protein observed mass and fragment ion masses can be readily obtained by deconvolution of isotopically resolved MS and MS/MS spectra using the Xtract 3.0 algorithm inside Xcalibur Qual Browser or FreeStyle (Thermo Fisher).

3.2.1. Open .RAW data file in Xcalibur Qual Browser or FreeStyle.

3.2.2. Select the most intense precursor ion charge state.

3.2.3. Obtain the observed mass of the polyubiquitin chain or ubiquitinated protein using Xtract 3.0. Set S/N threshold to 3, fit factor to 44, reminder to 25% and maximum charge to 50.

3.2.4. Compare to the theoretical mass. Mass differences <0.5Da are considered acceptable.

3.2.5. Select 3–5 of the precursor charge states that were fragmented:

- a. Obtain the fragment ion mass list using Xtract as stated in step 3.2.3.
- b. Export the fragment ions “Exact Mass” list to a .csv file. This file will contain the masses and intensities of the fragment ions.

3.2.6. Combine all fragment ion masses into one list.

Note: For precursor ions acquired at medium mass resolution that produce isotopically unresolved spectra, algorithms such as ReSpect (Positive Probability Ltd.) or UniDec (Marty et al., 2015) can be used to determine the observed mass of the intact protein.

3.3 Tandem mass spectra visualization

3.3.1. In ProSight Lite: Open “Modify Experimental Data”:

- a. Add the ubiquitin chain or ubiquitin conjugated protein *observed mass*.
- b. Paste the list of fragment ion masses.
- c. Select ETciD (or EThcD) as fragmentation method, hence, b, c, y and z ions will be matched.
- d. Set fragment ion mass tolerance from 5 to 10 ppm.

3.3.2. Open “Modify Candidate Sequence” and add the protein sequence of monoubiquitin or the substrate protein as template for data interpretation.

4. Data interpretation strategies

4.1 Procedure for interpreting spectral data from monoubiquitinated proteins

A five step strategy was implemented to characterize monoubiquitinated proteins (Chen et al., 2018). Additionally, Gomes et al. (2019) recently applied this approach to characterize rubylated (Rub-1 modified) proteins. This strategy is summarized in Fig. 2A and

exemplified by the characterization of the synthetically produced fluorescently labeled and chemically modified Ub-⁹PTEN_(5–21) conjugate.

Step 1: *Identify intact proteins.* Based on the list of candidate proteins generated after database search, we define putative branched proteins by the simultaneous identification of substrate proteins with observed large mass additions (≈ 9000 Da) and a common modifier protein (such as Ub, Rub1/NEDD8 or SUMO). The summed mass for each combination of substrate (anchor) and modifier proteins are then calculated and compared to the observed intact mass. The correct monoubiquitinated protein's theoretical mass should closely match the observed intact mass. In the example shown in Fig. 2B, the combination of synthetic PTEN peptide (5–21) and Ub modified in two sites by Alloc (*N*-(allyloxycarbonyloxy)succinimide) match the intact observed mass. Alloc are protective groups added to Lys residues during synthesis, which had not been completely removed (Chen et al., 2018).

Step 2: *Define anchor and modifier(s) protein.* The deconvoluted tandem MS mass list from the identified monoubiquitinated protein is visualized using ProSight Lite utilizing the sequence of the substrate or modifier protein as template. We define the modifier protein as the sequence template that presents only b and c ions, since y and z ions will not match the template until the mass of the substrate protein (anchor) is added to its C-terminus. The anchor protein shows b, c, y and z matching ions, however, these ions do not explain all protein residues since areas with gaps in fragmentation are present. These gap areas indicate the location of linkage sites or modifications. Hence, in the example displayed in Fig. 2C, the Ub is defined as the modifier protein and the PTEN peptide is defined as the anchor with a fragmentation gap indicating K₉ as the linkage site.

Step 3: *Determine linkage sites.* Ubiquitin modifies the anchor protein by isopeptide linkage to a Lys residue on the anchor protein. Hence, in Step 3 we perform mass addition trials by adding the modifier protein mass onto each Lys residue present in the fragmentation gap areas. The correct linkage site provides the highest fragmentation density. This increase in fragmentation density can be clearly observed in Fig. 2D after linkage is established.

Step 4: *Identify and locate other modifications.* Both the modifier and anchor proteins may also carry other post-translational or chemical modifications. The presence of additional modifications is indicated by the remaining mass difference between the observed and theoretical mass after linkage site determination and by the presence of area(s) with gaps in fragmentation. Based on the remaining mass difference, the sample's origin and preparation, we can hypothesize the type of modification(s) present and determine its identity and location by performing mass addition trials on residues located in the fragmentation gap area(s). As exemplified in Fig. 2E, the presence of other modifications is supported by the identification of Ub carrying two Alloc groups in Step 1 and by the clear fragmentation gap observed in Step 2. Hence, mass addition trials are performed to determine which 2 out of the 3 Lys residues (29, 33 and 48), present in the fragmentation gap area, are

modified by Alloc. We determined that modifications in K₂₉ and K₄₈ provide the largest number of matching fragment ions.

Step 5: *Final visualization*. The most likely final structure is that which provides the highest number of fragment ions. Fig. 2F displays the final structure of the Ub-⁹PTEN₍₅₋₂₁₎ conjugate.

4.2 Protocol for data interpretation of ubiquitin chains

Polyubiquitin chains can have diverse structures that determine the varied biological functions, known and yet to explore, of the ubiquitin signaling pathway. Investigating this structural diversity entails defining the topology of these chains. For proteins with linear sequences, protein topology describes the protein fold by defining the orientation and connectivity of its secondary structures (Rawlings, Taylor, Nyakairu, Fox, & Sternberg, 1985). In the case of polyubiquitin chains, where the protein sequence is no longer linear, topology describes the arrangement of the Ub subunits which is determined by their connectivity (linkage sites). Characterizing ubiquitin chains is very challenging due to the varied topologies Ub chains can take and the added complexity of 7 possible lysine sites for each isopeptide linkage. In order to discuss interpreting the acquired spectra we need to first define the nomenclature used in this protocol (Lee et al., 2016, 2016): (i) polyubiquitins exist in unanchored and anchored form when their C-terminus is free or attached to a substrate protein (S), respectively; (ii) the Ub moiety with a C-terminus that is free or attached to S is termed proximal (P); (iii) intermediate Ub moieties attached directly to P are termed α_n -endo; (iv) intermediate Ub moieties attached to α_n -endo Ub are termed β_n -endo and so on (e.g., an unbranched pentamer has α -, β - and γ -endo Ubs); and (v) the Ub moiety that is only attached to one other Ub by an isopeptide linkage through its C-terminus is termed distal (D). The subscript n is added to the endo Ub nomenclature to state the number of Ub moieties attached to its C-terminus (e.g., for an unanchored poly- ubiquitin a β_2 -endo Ub has two Ubs attached to its C-terminus: α_1 -endo and P). The data interpretation strategy has six general steps (Lee, Geis-Asteggianti, Dixon, Miller, et al., 2016) and is summarized in Fig. 3. Additionally, a step by step characterization of the unanchored branched tetramer [Ub-⁶³Ub,Ub]-^{6,63}Ub, previously reported in Lee, Geis-Asteggianti, Dixon, Miller, et al. (2016), is detailed in Fig. 4 and will be used to provide examples throughout the protocol.

Step 1: *Determination of the number of Ub subunits*. The number of Ub monomers present in the chain can be estimated, if unanchored, by dividing the observed molecular mass by the mass of Ub and accounting for the loss of water as each isopeptide bond is formed. For anchored proteins, the molecular mass of S (and water) needs to be subtracted from the observed mass prior to estimating the number of Ub moieties present.

The subsequent steps entail the interrogation of each Ub moiety (from proximal to distal) in order to determine the chain topology. For this purpose all acquired fragment ions are matched using the monoubiquitin protein sequence as a template in ProSight Lite. The initially high density of fragments matched to the template supports the presence of polyubiquitin chains. At this stage, the b and c fragment ions observed are not informative as they can be redundant amongst all the Ub moieties present in the chain. Instead, y and z ions do provide information on chain topology.

Step 2: *Initial interrogation of P*. If the ubiquitin chain is unanchored, P will have a free C-terminus which would be supported by the presence of y and z ions in the vicinity of the C-terminus. If the ubiquitin chain is anchored, the mass of S needs first to be added to the C-terminus Gly₇₆ for y and z ions to match the template sequence. This step also allows the investigator to characterize a putative linkage site between P and other Ub subunit(s). This is achieved by following the sequence of y and z ions from the proximal Ub C-terminus to the Lys residue immediately preceding the last matched y and z ion. Y and z fragment ions that include one (or more) modified Lys will not match the monoubiquitin template. An example of a Step 2 outcome is shown in Fig. 4, where matching y/z ions are traced back to K₆₃ indicating it as one of P linkage sites.

Step 3: *Determination of the topology of P*. The topology of the proximal ubiquitin can be unbranched or branched depending if one or more Ub moieties are directly attached to P, respectively. Thus the existence of an α 1-endo Ub, which would only exist if P is unbranched, needs to be interrogated. In this case, the template sequence represents the α 1-endo Ub subunit that is attached by its C-terminus to P and by one or more of the 7 possible linkage sites (lysine 6, 11, 27, 29, 33, 48 and 63) to further Ub subunits. Thus, the mass of P (or PS, if anchored) is added to the template C-terminus and the mass corresponding to the other Ub subunits is added at the template N-terminus. Since the main interest here is to determine topology instead of linkage sites, the aggregate mass of all other subunits is added to the N-terminus. After the masses are added to the template, the presence of unique (non-redundant) fragment ions between K₆₃ and G₇₆ will support the existence of an α 1-endo Ub (see sequence highlighted green in Fig. 3—Step 3). If no fragment ions are observed, then an α 1-endo Ub is not present and the topology of the proximal Ub is branched. For trimeric Ub chains, this step defines the chain topology as “all branched” (if α 1-endo Ub is absent) or “all unbranched” (if present), allowing interpretation to move forward to Step 5 for linkage site determination. For exemplary branched tetramer [Ub-⁶³Ub,Ub]-^{6,63}Ub, the absence of matching fragmentation within the K₆₃ and G₇₆ region is expected and clearly shown in Fig. 4—Step 3.

Step 4: *Determination of the presence and topology of subsequent endo Ubs*. This step can include many sub-steps depending on the number of endo Ubs present beyond the previously interrogated α 1-endo Ub (i.e., one for tetramers, two for pentamers, etc.). Step 4.1 entails the interrogation of two possible endo Ubs (α_2 or β_2) following the same procedure introduced in Step 3. Depending on the topology of P:

4.1. For branched P, where more than one Ub subunit is attached to P, the existence of an α_2 -endo Ub needs to be interrogated. In this case the subscript n = 2 implies that a distal Ub is also attached to P, using a different Lys residue than that of the α_2 -endo Ub. As stated previously, interrogation entails the use of a monoubiquitin template representing an α_2 -endo Ub to which the aggregate mass of PD (or PSD, if anchored) is added to the template C-terminus and the aggregate mass of the other Ubs

is added to the N-terminus. The presence of unique fragment ions between K_{63} and G_{76} will support the existence of an α_1 -endo Ub. For tetrameric Ub chains, this step defines the chain topology as “all branched” (if α_2 -endo Ub is absent) or branched proximal (if present), allowing progression to Step 5 for linkage site determination. This observation is exemplified in Fig. 4—Step 4.1, where the presence of α_2 -endo Ub in the $[Ub-^{63}Ub, Ub]-^{6,63}Ub$ tetramer is supported by the fragment ions observed in the topology defining region (K_{63} to G_{76}). For larger polymers, depending whether an α_1 -endo Ub was present or not, then the existence of a β_3 -endo Ub or an α_3 -endo Ub needs to be interrogated in Step 4.2, respectively.

4.2. For unbranched P, the existence of a β_2 -endo Ub attached through its C-terminus to the previously defined α_1 -endo Ub needs to be interrogated. In this case, the aggregate mass of P, α_1 -endo and S (if anchored) needs to be added to the β_2 -endo Ub C-terminus and the aggregate mass of the other Ub subunits is added to its N-terminus. Similarly, unique fragment ions in the K_{63} and G_{76} region will support the existence of a β_2 -endo Ub. For tetrameric Ub chains, this step defines the chain topology as branched α_1 -endo Ub (if β_2 -endo Ub is absent) or “all unbranched” (if present). An example of deconvoluted mass spectra of the “all unbranched” $Ub-^{63}Ub-^{63}Ub-^{63}Ub$ is shown in Fig. 5, where products ions from the P, α_1 -endo, β_2 -endo and D subunits are observed. For larger polymers, depending if a β_2 -endo Ub was defined or not, then the existence of a γ -endo Ub or β_3 -endo Ub needs to be interrogated in a Step 4.2, respectively.

This procedure is then repeated in further sub-steps to interrogate subsequent possible endo Ubs (e.g., α_3 , γ , β_3 , etc.) until the topology of the chain is defined.

The next step entails the determination of the lysine sites which participate in the formation of the Ub chain isopeptide linkages. The definition of the chain topology prior to linkage site determination is key in order to reduce the number of trials needed in Step 5.

Step 5: Determination of linkage sites. With the chain topology defined, P and endo Ubs are now interrogated in order to define their linkage sites. For each defined Ub subunit, the appropriate masses are first added to the template C-terminus and N-terminus, as done in earlier steps. In general, linkage sites are indicated by the absence of matching fragment ions due to the presence of a modified Lys residue, which has not been taken into account in the template. As the mass of the corresponding modification is moved onto a candidate Lys, watch for more fragment ions to be assigned on the template sequence. The correct linkage site will be that which provides the highest fragment ion density.

In the case of unbranched topology, linkage sites can be easily traced back from the Ub subunit C-terminus to the Lys located beyond the last matched fragment ions (b, c, y and z). In the case of branched topology, where more than one Ub is attached, only the sequence of

y and z ions is initially traced back from the subunit C-terminus to the first Lys located beyond the last matched fragment ion. Subsequently, the mass of one (or more) Ub subunit(s) is added onto that Lys, observing the appearance of new matching fragment ions to the putative sequence. These newly matched fragment ions (b, c, y and z) will provide information on the linkage location of the next Ub(s), by once again tracing fragmentation until matching stops and performing a new trial by adding the mass of one (or more) Ub(s) in the Lys preceding that location. Several trials may be needed in order to determine the structure that provides the highest fragment ion density. A clear example of this procedure is shown in Fig. 4—Step 5 for [Ub-⁶³Ub,Ub]-^{6,63}Ub. Until this step, P was defined as branched, hence, the tetramer is formed by two D, one α_2 -endo and a P Ub subunits. Additionally, K₆₃ was initially indicated as one of the linkage sites in Step 2. Therefore, two mass addition trials were performed where the mass of D or the summed mass of α_1 -endo and D were added to the K₆₃ site of the proximal Ub. The increase in fragmentation density supports the correct structure with only one D attached at K₆₃. The presence of a fragment ion at the N-terminus of K₁₁, indicates that the α_2 -endo and D subunits are attached to P by K₆.

Step 6: Final visualization. The last step entails the visualization of the interpreted data as a whole. The final image of each Ub subunit is then exported into a graphical editor, such as Inkscape or Adobe Illustrator, where lines that represent the linkages determined can be drawn between the Ub subunits. In this step, a visual comparison of the proposed correct structure and the other possible linkage sites can be performed. The correct structure is that which is supported by the presence of unique (non-redundant) fragment ions and that presents the highest fragment ion density as a whole.

5. Applications, limitations and outlook

Most studies of the ubiquitinome to date use affinity enrichments (e.g., anti-ubiquitin antibodies, anti-GG antibodies, and Tandem Ubiquitin Binding Entities) coupled to immunoblotting and/or MS analysis of peptides carrying a GG tag in the linking Lys residue after tryptic digestion (Emmerich & Cohen, 2015; Kim et al., 2011; Peng et al., 2003; Xu & Peng, 2006, 2008). Even though these approaches may provide information on the identity of the substrate proteins and their linkage sites, they do not offer knowledge of Ub chain size, topology and connectivity. Hospenthal et al. (2015) has utilized linkage-specific DUBs to elucidate polyubiquitin chains topology and linkage sites. This approach requires reiterative reactions and its efficacy depends on the capability of each enzyme to deubiquitinase specifically all possible ubiquitin chains including those of mixed linkages. The novel data interpretation strategies described in this chapter offer straightforward and effective characterization of ubiquitin conjugates and ubiquitin chains and is compatible with any MS activation technology that cleaves most or all peptide bonds in a protein. Additionally, these strategies make use of available and evolving bioinformatics tools for top-down MS, reducing the level of manual data interpretation needed and increasing throughput.

Our methodology was developed and thoroughly evaluated using 5 synthetic ubiquitin conjugates, 3 rubylated conjugates and 20 synthetic ubiquitin chains (7 dimers, 6 trimers, 6 tetramers and 1 pentamer). The ubiquitin conjugates studied were composed of ubiquitin carrying chemical modifications or mutations attached to the following anchor proteins: Rub1 (with T72R mutation), E2 enzyme UbcH5b (His-tagged) and a fluorescently labeled and chemically modified PTEN octadecapeptide (residues 5–21) (Chen et al., 2018). Additionally, the rubylated conjugates studied were Rub1 homodimers (with K4F and/or T72R mutations) and a rubylated ubiquitin (with Rub1 E63K and T72R mutation, and Ub lacking the C-terminal GG) (Gomes et al., 2019). The strategy described in this chapter was capable of fully characterizing these conjugates and of assigning chemical modifications that were introduced or inadvertently remained after synthesis. In the case of unanchored polyubiquitin chains, a diverse set of synthetic chains were prepared in order to cover various chain lengths, topologies and linkages (Lee, Geis-Asteggianti, Dixon, Kim, et al., 2016; Lee, Geis-Asteggianti, Dixon, Miller, et al., 2016). All ubiquitin chains evaluated presented an adequate number of matching fragment ions and were fully characterized, except for the tetrameric “all branched” chain [Ub]₃^{6,27,48}Ub. For this branched polyubiquitin, which has 3 distal Ub subunits attached to a proximal Ub, the fragment ion density in the proximal Ub was not sufficient to localize the linkage site of 2 out of the 3 distal Ubs attached. Based on the limited fragmentation obtained, we were able to localize the attachment of those 2 distal Ub to the sequence region containing the Lys residues 11, 27, 29, 33 and 48. Hence, efficient precursor ion activation that provides extensive fragmentation is key to achieve conclusive results.

Optimization of instrumental parameters is thus crucial for successful top-down MS characterization, since it depends heavily on efficient precursor ion isolation, accumulation and fragmentation. Our method was evaluated using an orbitrap Fusion Lumos system, which offered enhanced sensitivity, several activation approaches (CID, HCD and ETD), and the combined activation by ETD and CID (ETciD) or HCD (EThcD) to achieve substantial peptide backbone fragmentation. In our experiments, the use of optimized ETciD and EThcD frequently achieved >70% peptide backbone cleavage, though as stated earlier lower percentages were observed for larger branched ubiquitin chains. The orbitrap Fusion Lumos also provides the high mass resolution and mass accuracy required for determining the substrate protein identity, its branching (linkage) site, and the identity and location of other post-translational or chemical modification(s). Nevertheless, this method is not restricted to hybrid orbitrap systems and can be also used with Fourier transformed ion cyclotron resonance and quadrupole time of flight (Q-TOF). Additionally, activation techniques such as ultraviolet photon dissociation (UVPD) (Fort et al., 2016) and Activated Ion-Electron Transfer Dissociation (AI-ETD) (Riley, Westphall, & Coon, 2017) have become commercially available since our method was developed. These activation techniques provide extensive fragmentation as demonstrated for heavily-modified histones (Greer & Brodbelt, 2018), homotypic polyubiquitins (Cannon, Martinez-Fonts, Robotham, Matouschek, & Brodbelt, 2015), monoclonal antibodies (Fornelli et al., 2018), and typical protein standards (Riley et al., 2017). The higher number of protein backbone cleavages, many of which are complementary to those obtained by the more traditionally used techniques (CID or HCD), allows for more reliable PTM assignment. Moreover, recent top-

down MS research has taken advantage of combining fragmentation techniques in order to increase sequence coverage, which can also prove advantageous to our strategy (Fornelli et al., 2018; Shliaha et al., 2018). If, however, only moderate fragmentation is achieved and linkage sites are not conclusively determined, middle-down approaches using limited proteolysis or microwave-assisted acid cleavage can help further discern the correct linkage sites.

It is also relevant to notice that this strategy was evaluated using high mass resolution for precursor ions, which offers a limited mass range of up to approximately 50kDa. However, depending on the identity of substrate protein and the modifier ubiquitin chain length, polyubiquitinated proteins may have masses larger than 50kDa. From the point of view of data processing, this challenge has been partially overcome using deconvolution algorithms that allow for the determination of the protein mass from isotopically unresolved spectra obtained at low to medium mass resolution. As stated in the protocols above, the database search software Proteome Discoverer 2.2 is capable of identifying intact proteins using precursor ion spectra acquired at low to medium mass resolution using the ReSpect algorithm integrated into the ProSightPD Top-Down Low/High cRAWler node. However, proteins modified by ubiquitin or ubiquitin-like proteins can have large mass additions requiring the use of a large precursor mass tolerance (>10kDa) during the search, which can reduce identification sensitivity. From the point of view of analysis of large molecules by top-down or native MS, hybrid instruments that use quadrupoles for ion transmission needed to be modified in order to allow for the transmission of high m/z ions. This feature has been available in Waters Q-TOF systems (m/z up to 32,000) for almost 20 years (Sobott, Hernández, McCammon, Tito, & Robinson, 2002) and it has been recently implemented in the Thermo Fisher Q-Exactive orbitrap systems Biopharma (m/z up to 8000) and UHMR (m/z up to 25,000). Note however that the Q-Exactive systems currently offer ion activation by HCD only, which may be not sufficient for our current approach.

This protocol is currently effective for enriched samples, such as those obtained by synthesis, recombinant production or immunoprecipitation (IP), since it requires moderate amounts of protein (i.e., ~90ng). The cellular Ub pool is highly dynamic and has been found to vary by cell type and in presence of stressors (Groothuis, Dantuma, Neefjes, & Salomons, 2006; Ordureau, Muench, & Harper, 2015). For example, Kaiser et al. (2011) reported that Ub (as free monomer) represents only 0.42% of the total protein weight in HEK293 cells, from which 65% corresponds to monoubiquitinated conjugates, 23% to free Ub and 11% to polyubiquitin chains. This suggests that ubiquitinated proteins are not very abundant in a cell lysate, therefore needing to be enriched. Preliminary efforts carried out in our laboratory to characterize polyubiquitin chains enriched from cells by IP have been hindered by poor recoveries and antibody cross-specificity, obtaining mixtures of ubiquitin and ubiquitinated conjugates. We have also observed that these complex mixtures do not separate well by reverse phase liquid chromatography, requiring slow gradients and showing poor peak shape. As a result, other fractionation technologies may be considered in future studies. Despite current limitations, our strategy can readily adapt to advances in MS technology and it can be extended to study other ubiquitin-like protein modifiers.

6. Summary

Ubiquitin is a dynamic post-translational modification that can be present as monomers or polymers in cells in both anchored (modifying substrate proteins) and unanchored forms. Ubiquitination plays a role in a myriad of biological functions depending on the three dimensional structure of the polymeric chain attached. Current proteomic and enzymatic approaches provide limited knowledge of ubiquitin conjugates, missing relevant information on chain topology and connectivity that is needed to better understand ubiquitin signaling. In this chapter we have introduced strategies to characterize ubiquitinated proteins and ubiquitin chains by top-down mass spectrometry. These methodologies are capable of providing a detailed view of the anchored or unanchored ubiquitin chain structure, determining also the identity of the modified substrate protein and the identity and location of other modifications that may be present.

Acknowledgment

This research was supported by NIH Grants GM021248, GM065334, and S10OD019938.

References

- Akutsu M, Dikic I, & Bremm A (2016). Ubiquitin diversity at a glance. *Journal of Cell Science*, 129, 875–880. [PubMed: 26906419]
- Boname JM, Thomas M, Stagg HR, Xu P, Peng J, & Lehner PJ (2010). Efficient internalization of MHC I requires lysine-11 and lysine-63 mixed linkage polyubiquitin chains. *Traffic*, 11, 210–220. [PubMed: 19948006]
- Bruderer R, Tatham MH, Plechanovova A, Matic I, Garg AK, & Hay, RT. (2011). Purification and identification of endogenous polySUMO conjugates. *EMBO Reports*, 12, 142–148. [PubMed: 21252943]
- Cannon JR, Martinez-Fonts K, Robotham SA, Matouschek A, & Brodbelt JS. 2015 Top-down 193-nm ultraviolet photodissociation mass spectrometry for simultaneous determination of polyubiquitin chain length and topology. *Analytical Chemistry*, 87, 1812–1820. [PubMed: 25559986]
- Castaneda C, Liu J, Chaturvedi A, Nowicka U, Cropp TA, & Fushman D (2011). Nonenzymatic assembly of natural polyubiquitin chains of any linkage composition and isotopic labeling scheme. *Journal of the American Chemical Society*, 133, 17855–17868. [PubMed: 21962295]
- Castañeda CA, Liu J, Kashyap TR, Singh RK, Fushman D, & Cropp TA (2011). Controlled enzymatic synthesis of natural-linkage, defined-length polyubiquitin chains using lysines with removable protecting groups. *Chemical Communications*, 47, 2026–2028. [PubMed: 21212884]
- Chau V, Tobias JW, Bachmair A, Marriott D, Ecker DJ, Gonda DK, et al. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science*, 243, 1576–1583. [PubMed: 2538923]
- Chen J, & Chen ZJ (2013). Regulation of NF- κ B by ubiquitination. *Current Opinion in Immunology*, 25, 4. [PubMed: 23312890]
- Chen D, Gomes F, Abeykoon D, Lemma B, Wang Y, Fushman D, et al. (2018). Top-down analysis of branched proteins using mass spectrometry. *Analytical Chemistry*, 90, 4032–4038. [PubMed: 29513006]
- Dixon EK, Castañeda CA, Kashyap TR, Wang Y, & Fushman D (2013). Nonenzymatic assembly of branched polyubiquitin chains for structural and biochemical studies. *Bioorganic & Medicinal Chemistry*, 21(12), 3421–3429. [PubMed: 23557636]
- Emmerich CH, & Cohen P (2015). Optimising methods for the preservation, capture and identification of ubiquitin chains and ubiquitylated proteins by immunoblotting. *Biochemical and Biophysical Research Communications*, 466, 1–14. [PubMed: 26325464]

- Fornelli L, Srzentic K, Huguet R, Mullen C, Sharma S, Zabrouskov V, et al. (2018). Accurate sequence analysis of a monoclonal antibody by top-down and middle-down Orbitrap mass spectrometry applying multiple ion activation techniques. *Analytical Chemistry*, 90, 8421–8429. [PubMed: 29894161]
- Fort KL, Dyachenko A, Potel CM, Corradini E, Marino F, Barendregt A, et al. (2016). Implementation of ultraviolet photodissociation on a benchtop Q exactive mass spectrometer and its application to phosphoproteomics. *Analytical Chemistry*, 88, 2303–2310. [PubMed: 26760441]
- Gilda JE, Ghosh R, Cheah JX, West TM, Bodine SC, & Gomes AV (2015). Western blotting inaccuracies with unverified antibodies: Need for a western blotting minimal reporting standard (WBMRS). *PLoS One*, 10, 1–18.
- Gomes F, Lemma B, Abeykoon D, Chen D, Wang Y, Fushman D, et al. (2019). Top-down analysis of novel synthetic branched proteins. *Journal of Mass Spectrometry*, 54, 19–25. [PubMed: 30347468]
- Goto E, Yamanaka Y, Ishikawa A, Aoki-Kawasumi M, Mito-Yoshida M, Ohmura- Hoshino M, et al. (2010). Contribution of K11-linked ubiquitination to MIR2- mediated MHC class I internalization. *Journal of Biological Chemistry*, 285, 35311–35319. [PubMed: 20833710]
- Greer SM, & Brodbelt JS (2018). Top-Down characterization of heavily modified histones using 193 nm ultraviolet photodissociation mass spectrometry. *Journal of Proteome Research*, 17, 1138–1145. [PubMed: 29343059]
- Groothuis TA, Dantuma NP, Neefjes J, & Salomons FA (2006). Ubiquitin crosstalk connecting cellular processes. *Cell Division*, 1, 21–27. [PubMed: 17007633]
- Hospenthal MK, Freund SMV, & Komander D (2013). Assembly, analysis and architecture of atypical ubiquitin chains. *Nature Structural & Molecular Biology*, 20, 555–565.
- Hospenthal MK, Mevissen TE, & Komander D (2015). Deubiquitinase-based analysis of ubiquitin chain architecture using ubiquitin chain restriction (UbiCRest). *Nature Protocols*, 10, 349–361. [PubMed: 25633630]
- Jones J, Wu K, Yang Y, Guerrero C, Nillegoda N, Pan Z-Q, et al. (2008). Targeted proteomic analysis of the ubiquitin-like modifier Nedd8 and associated proteins. *Journal of Proteome Research*, 7, 1274–1287. [PubMed: 18247557]
- Kaiser SE, Riley BE, Shaler TA, Trevino RS, Becker CH, Schulman H, et al. (2011). Protein standard absolute quantification (PSAQ) method for the measurement of cellular ubiquitin pools. *Nature Methods*, 8, 691–696. [PubMed: 21743460]
- Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, et al. (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular Cell*, 44, 325–340. [PubMed: 21906983]
- Lee AE, Geis-Asteggianti L, Dixon EK, Kim Y, Kashyap TR, Wang Y, et al. (2016). Preparing to read the ubiquitin code: Characterization of ubiquitin trimers by top-down mass spectrometry. *Journal of Mass Spectrometry*, 51, 315–321. [PubMed: 27041663]
- Lee AE, Geis-Asteggianti L, Dixon EK, Miller M, Wang Y, Fushman D, et al. (2016). Preparing to read the ubiquitin code: Top-down analysis of unanchored ubiquitin tetramers. *Journal of Mass Spectrometry*, 51, 629–637. [PubMed: 28239975]
- Marty MT, Baldwin AJ, Marklund EG, Hochberg GKA, Benesch JLP, & Robinson CV (2015). Bayesian deconvolution of mass and ion mobility spectra: From binary interactions to polydisperse ensembles. *Analytical Chemistry*, 87, 4370–4376. [PubMed: 25799115]
- Mourtzoukou D, Drikos I, Goutas N, & Vlachodimitropoulos D (2018). Review of the ubiquitin role in DNA repair and tumorigenesis, with emphasis in breast cancer treatment; current data and future options In Boutou E & Sturzbecher H-W (Eds.), *Ubiquitination governing DNA repair-implications in health and disease* (pp. 179–210). London: IntechOpen Limited.
- Nakasone MA, Livnat-Levanon N, Glickman MH, Cohen RE, & Fushman D (2013). Mixed-linkage ubiquitin chains send mixed messages. *Structure*, 21, 727–740. [PubMed: 23562397]
- Newton K, Matsumoto ML, Wertz IE, Kirkpatrick DS, Lill JR, Tan J, et al. (2008). Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. *Cell*, 134, 668–678. [PubMed: 18724939]

- Nguyen LK, Dobrzynski M, Fey D, & Kholodenko BN (2014). Polyubiquitin chain assembly and organization determine the dynamics of protein activation and degradation. *Frontiers in Physiology*, 5, 4–13. [PubMed: 24478717]
- Okatsu K, Koyano F, Kimura M, Kosako H, Saeki Y, Tanaka K, et al. (2015). Phosphorylated ubiquitin chain is the genuine parkin receptor. *The Journal of Cellular Biology*, 209, 111–128.
- Ordureau A, Muench C, & Harper JW (2015). Quantifying ubiquitin signaling. *Molecular Cell*, 58, 660–676. [PubMed: 26000850]
- Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, et al. (2003). A proteomics approach to understanding protein ubiquitination. *Nature Biotechnology*, 21, 921–926.
- Rawlings CJ, Taylor WR, Nyakairu J, Fox J, & Sternberg MJ (1985). Reasoning about protein topology using the logic programming language PROLOG. *Journal of Molecular Graphics*, 3, 151–157.
- Riley NM, Westphall MS, & Coon JJ (2017). Activated ion-electron transfer dissociation enables comprehensive top-down protein fragmentation. *Journal of Proteome Research*, 16, 2653–2659. [PubMed: 28608681]
- Shliaha PV, Gibb S, Gorshkov V, Jespersen MS, Andersen GR, Bailey D, et al. (2018). Maximizing sequence coverage in top-down proteomics by automated multimodal gas-phase protein fragmentation. *Analytical Chemistry*, 90, 12519–12526. [PubMed: 30252444]
- Sobott F, Hernández H, McCammon MG, Tito MA, & Robinson CV (2002). A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Analytical Chemistry*, 74, 1402–1407. [PubMed: 11922310]
- Swatek KN, & Komander D (2016). Ubiquitin modifications. *Cell Research*, 26, 399–422. [PubMed: 27012465]
- Thorne AW, Sautiere P, Briand G, & Crane-Robinson C (1987). The structure of ubiquitinated histone H2B. *EMBO Journal*, 6, 1005–1010. [PubMed: 3036486]
- Varadan R, Assfalg M, Haririnia A, Raasi S, Pickart C, & Fushman D (2004). Solution conformation of Lys63-linked Di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling. *Journal of Biological Chemistry*, 279, 7055–7063. [PubMed: 14645257]
- Varadan R, Walker O, Pickart C, & Fushman D (2002). Structural properties of polyubiquitin chains in solution. *Journal of Molecular Biology*, 324, 637–647. [PubMed: 12460567]
- Varshavsky A (2012). The ubiquitin system, an immense realm. *Annual Review of Biochemistry*, 81, 167–176.
- Xu X, Niu Y, Liang K, Wang J, Li X, & Yang Y (2015). Heat shock transcription factor 532 is targeted for degradation via an ubiquitin-like protein ThiS in *Escherichia coli*. *Biochemical and Biophysical Research Communications*, 459, 240–245. [PubMed: 25721662]
- Xu P, & Peng J (2006). Dissecting the ubiquitin pathway by mass spectrometry. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1764, 1940–1947. [PubMed: 17055348]
- Xu P, & Peng J (2008). Characterization of polyubiquitin chain structure by middle-down mass spectrometry. *Analytical Chemistry*, 80, 3438–3444. [PubMed: 18351785]

Further reading

- Hjerpe R, Aillet F, Lopitz-Otsoa F, Lang V, England P, & Rodriguez MS (2009). Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitinbinding entities. *EMBO Reports*, 10, 1250–1258. [PubMed: 19798103]

Fig. 1. A ball-and-stick representation of different kinds of ubiquitin PTMs: (A) Substrate protein modified by a single monoUb. (B) Substrate protein modified by multiple monoUbs. (C) Unanchored and unbranched polyubiquitin chain. (D) Substrate protein modified by a branched polyubiquitin chain.

Fig. 2.

Interpretation of tandem mass spectra for the characterization of monoubiquitinated proteins.

(A) Scheme of the five step data analysis workflow. (B) Outcomes for each workflow step are exemplified by the characterization of an FITC-labeled and modified Ub-⁹PTEN₍₅₋₂₁₎ conjugate. *Adapted from* Chen, D., Gomes, F., Abeykoon, D., Lemma, B., Wang, Y., Fushman, D., et al. (2018). *Top-down analysis of branched proteins using mass spectrometry*. *Analytical Chemistry*, 90, 4032—4038.

Step 1: Determination of the number of Ub subunits

$$n_{\text{Ub subunits}} = \{[\text{Observed molecular mass} - \text{Substrate protein mass (if present)}] / \text{Ub mass}\}$$

Step 2: Initial interrogation of proximal Ub

Set template sequence and fragment ions mass list in ProSight Lite

Step 3: Determination of the topology of proximal Ub **Step 4: Determination of the presence and topology of subsequent endo Ub** **Step 5: Determination of linkage sites****Step 6: Final visualization****Reference code:**

● Proximal Ub ● α_1 α_1 -endo Ub ● α_2 α_2 -endo Ub ● β_2 β_2 -endo Ub ● β_3 β_3 -endo Ub ● γ γ -endo Ub ● Distal Ub

Fig. 3.

Data interpretation workflow of tandem mass spectra for the characterization of anchored and unanchored Ub chains. General steps are shown here for pentameric chains, though the method's logic can be easily extended to longer chains. *Adapted from* Lee, A. E., Geis-Asteggianti, L, Dixon, E. K, Kim, Y., Kashyap, T. R., Wang, Y., et al. (2016). *Preparing to read the ubiquitin code: Characterization of ubiquitin trimers by top-down mass spectrometry*. Journal of Mass Spectrometry, 51, 315–321 and Lee, A. E., Geis-Asteggianti, L, Dixon, E. K, Miller, M., Wang, Y., Fushman, D., et al. (2016). *Preparing to read the ubiquitin code: Top-down analysis of unanchored ubiquitin tetramers*. Journal of Mass Spectrometry, 51, 629–637.

Fig. 4. Characterization of the branched tetramer [Ub-⁶³Ub,Ub]-^{6,63}Ub showing the detailed outcomes for each data interpretation step. Adapted from Lee, A. E., Geis-Asteggianti, L., Dixon, E. K., Miller, M., Wang, Y., Fushman, D., et al. (2016). *Preparing to read the ubiquitin code: Top-down analysis of unanchored ubiquitin tetramers*. Journal of Mass Spectrometry, 51, 629–637.

Fig. 5.

Deconvoluted tandem mass spectra of the “all unbranched” Ub-⁶³Ub-⁶³Ub-⁶³Ub unanchored polyubiquitin chain. The combined product ion mass spectra acquired after fragmenting the five most abundant precursor m/z is shown. A fivefold magnification on the relative abundance axis was performed between 16,500 and 24,500 Da. The diagnostic product ions corresponding to the proximal, α_1 -endo and β_2 -endo Ub subunits are labeled in gray, pink and light green, respectively. *This figure was adapted from Lee, A. E., Geis-Asteggiane, L., Dixon, E. K., Miller, M., Wang, Y., Fushman, D., et al. (2016). Preparing to read the ubiquitin code: Top-down analysis of unanchored ubiquitin tetramers. Journal of Mass Spectrometry, 51, 629–637.*