

Published in final edited form as:

Nat Cell Biol. 2005 August ; 7(8): 750–757.

## Weighing in on Ubiquitin: The Expanding Role of Mass Spectrometry-based Proteomics

Donald S. Kirkpatrick<sup>1</sup>, Carilee Denison<sup>1</sup>, and Steven P. Gygi<sup>1,2</sup>

<sup>1</sup>Department of Cell Biology and <sup>2</sup>Taplin Biological Mass Spectrometry Facility, Harvard Medical School, Boston, Massachusetts 02115, USA

### Abstract

Mass spectrometry-based proteomics has become an essential tool for qualitative and quantitative analysis of cellular systems. The biochemical complexity and functional diversity of the ubiquitin system are well suited to proteomic studies. This review summarizes advances involving the identification of ubiquitinated proteins, the elucidation of ubiquitin-modification sites, and the determination of poly-ubiquitin chain linkages, while offering a perspective on the application of emerging technologies to mechanistic and functional studies of protein ubiquitination. [DK1]

### Introduction

The ubiquitin system is a well characterized pathway involved in regulating nearly every cellular process in eukaryotes<sup>1–3</sup>. In a series of elegant experiments, Hershko, Ciechanover, Rose and colleagues posited that the ATP-dependent modification of protein substrates by ubiquitin (termed APF-1 at the time) targeted them for degradation<sup>4, 5</sup>. This hypothesis was substantiated when Varshavsky and colleagues demonstrated a role for ubiquitin in protein turnover within cells<sup>6, 7</sup>. Since that time, ubiquitination has emerged as a central regulatory mechanism controlling not only protein stability, but also localization, interactions, and functional activity for a vast number of protein substrates (Fig. 1). [DK2]

The hallmark of the ubiquitin system is post-translational modification of protein substrates by ubiquitin, a highly conserved 76 amino acid polypeptide. The C-terminal glycine of ubiquitin is covalently linked through an isopeptide bond to the side chain of lysine(s) within the substrate. Substrates can be modified, either by a mono-ubiquitin, multiple mono-ubiquitins (multiubiquitination), or a poly-ubiquitin chain(s) (polyubiquitination) [DK4] (Fig. 1). In addition to ubiquitin, an entire family of ubiquitin-like (Ubl) proteins has been discovered. These proteins share significant sequence homology to ubiquitin, and in many cases, form covalent post-translational modifications by similar mechanisms. However, while ubiquitin often acts as a degradation signal, Ubl modifications appear to modulate exclusively non-proteasomal endpoints. Among these Ubl proteins are SUMO (Small Ubiquitin-like MOdifier)<sup>8</sup>, ISG15 (Interferon-Stimulated Gene of 15 kDa)<sup>9</sup>, and NEDD8 (NEural precursor cell-expressed and Developmentally Down-regulated gene)<sup>10, 11</sup>.

Due to improvements in instrument sensitivity, mass accuracy, peptide fragmentation, and database searching, mass spectrometry (MS)-based proteomics is becoming a mature platform for the systematic characterization of both the ubiquitin and Ubl systems. This review will focus on the existing and future applications of MS-based proteomics in such analyses. We

will discuss the use of shotgun sequencing to identify protein substrates, and ubiquitin/Ubl modification sites, as well as enzymes involved in conjugation (ligases), deconjugation (DUBs), and ubiquitin-dependent proteolysis (26S proteasome). Ongoing work in this area seeks to develop a more complete understanding of the structure and function of poly-ubiquitin chains, and to define the role and specificity of ubiquitin-binding proteins. Since these and other mechanistic studies require carefully controlled comparisons between samples, pertinent subtractive/differential display approaches will be discussed. Finally, the focus will shift to the looming impact of stable isotope based quantitative approaches and their unequaled potential for further elucidation of the ubiquitin system.

## Shotgun sequencing by MS as a biological assay

The term “shotgun sequencing,” coined by Yates and colleagues, refers to the automated identification and cataloging of proteins directly from complex mixtures<sup>12</sup>. At the heart of shotgun sequencing is the acquisition of thousands of tandem mass (MS/MS) spectra. In this approach, proteins are enzymatically digested into peptides, separated via reversed-phase chromatography and analyzed automatically by a mass spectrometer. Since each MS/MS spectrum represents the measurement of fragment ions produced from a single peptide (Fig. 2A), peptides are “sequenced” by the correlation of each MS/MS spectrum against a sequence database using software<sup>13, 14</sup>. Complex samples containing hundreds of proteins can be sequenced during a single analysis. This approach can be used as a biological assay to probe specific cell states by collating lists of identified peptides (cataloging proteomics), enumerating differences in peptide composition between samples (subtractive proteomics), or by comparing protein profiles between cell states using stable isotope labeling (quantitative proteomics) [DK6](Fig. 2B).

When more complex mixtures (e.g., cell lysates) are analyzed, additional separation is required for maximal protein coverage. To date, reports identifying more than 1,000 proteins have used either gel-based separation of proteins (GeLC-MS)<sup>15, 16</sup> or multi-dimensional chromatography of peptides<sup>17, 18</sup> prior to tandem mass spectrometry [DK9] (Fig. 2C). Multi-dimensional chromatography, performed following enzymatic digestion, typically employs a step of strong cation exchange chromatography prior to reversed phase separation. Two high-profile studies of the malaria parasite, *Plasmodium falciparum*, provide examples of how these two different approaches can be used for large scale proteomic analyses<sup>15, 17</sup>. While generating an enormous amount of data, a significant amount of analysis time is required to process and analyze each sample individually. The original multi-dimensional approach, multidimensional protein identification technology (MUDPIT), partially circumvents this problem by coupling the two chromatographic steps online with detection via mass spectrometry in an automated fashion, effectively eliminating many of the intermediate sample handling steps<sup>19</sup>. An advantage of multi-dimensional chromatography in ubiquitin analyses is that since proteins in the sample are digested together, rather than separated by molecular weight as in GeLC-MS approaches, maximum sensitivity is obtained for each individual substrate. The loss of sensitivity resulting from molecular weight separation can be partially overcome in GeLC-MS, since it is possible to load as much as 10 mg protein lysate on a preparative SDS-PAGE gel under optimized conditions.

## Identification of substrates for ubiquitin and Ubl proteins

Initial contributions of MS-based proteomics to the study of ubiquitin and Ubl proteins have demonstrated that as many as one thousand ubiquitinated proteins can be identified within a single experiment<sup>18</sup>. Typically, substrates are purified via an N-terminal epitope tag [DK10] fused to ubiquitin, digested using trypsin, and analyzed by proteome-scale shotgun sequencing. To

date, this basic approach [DK11] has been used to identify both ubiquitin- and SUMO- substrates *en masse*.

The first reported large scale analysis of ubiquitinated proteins using shotgun sequencing identified 1075 candidate substrates from yeast expressing epitope-tagged ubiquitin<sup>18</sup>. In subsequent studies, similar approaches were used to characterize defined subsets of ubiquitinated proteins<sup>20, 21</sup>, as well as proteins modified by SUMO<sup>22–26</sup>. These approaches have been extended into mammalian systems for both ubiquitin<sup>27</sup> and SUMO<sup>28–31</sup>. In addition, a transgenic mouse expressing (His)<sub>6</sub>-ubiquitin has been described that may be useful for isolating conjugates from mammalian tissues<sup>32</sup>. For large-scale studies, yeast offer a distinct advantage over mammalian systems since the multiple genes encoding ubiquitin can be genetically inactivated prior to introduction of epitope-tagged ubiquitin, making it the sole form of ubiquitin within cells<sup>33</sup>. Nontagging strategies for enriching targets, such as using ubiquitin-binding proteins<sup>21, 34</sup>, may help overcome this difficulty by avoiding the use of epitope-tags completely, allowing for large scale analysis of ubiquitinated proteins from untransfected cells, animal tissues, or possibly even clinical specimens.

While identification of cellular substrates is optimal, several reports demonstrate that *in vitro* systems can be used effectively to identify targets for both ubiquitin and Ubl proteins<sup>35–37</sup>. *In vitro* systems have been particularly useful in focused studies characterizing substrates for ubiquitin ligases such as the tumor suppressor BRCA1<sup>38, 39</sup> and erythrocyte spectrin<sup>40</sup>, as well as the SUMOylation of CENP-C by Ubc9<sup>41</sup>. [DK12]

While cataloging proteomics often offers insights into the breadth of ubiquitin- and Ubl- pathways, comparative biology requires more stringent subtractive approaches<sup>42</sup>. Subtractive proteomics takes advantage of the fact that a mass spectrometer is a concentration-sensitive detector. Since the number of peptides identified from a given protein within a mixture is roughly proportional to its abundance, comparing the number of peptides identified for a single protein in parallel samples can be semiquantitative. For example, ubiquitinated proteins were classified as substrates of the endoplasmic-reticulum associated degradation (ERAD) pathway if they displayed a ‘down-up-down’ profile for the number of peptides identified during analysis of wild type, *npl4-1* mutant, and *npl4-1Ubc7/Δ* double mutant strains<sup>20</sup>. This profile was predicted based upon observations that *npl4-1* mutants displayed increased levels of ubiquitinated proteins which were reversed in the absence of the E2 Ubc7. In another example, subtractive proteomics was used to identify ubiquitin-substrates whose degradation relied upon the ubiquitin-binding protein Rpn10<sup>21</sup>.

Comparisons of enriched pools [DK14] of ubiquitinated proteins from wild type and mutant cells, cells in different metabolic states, or cells grown under different conditions, have the potential to globally define ligase-substrate-DUB relationships. While subtractive approaches have proven useful, our belief is that differences between samples are best quantified using stable isotope labeling. Useful stable isotopes that are readily distinguishable by MS include deuterium, <sup>13</sup>C, and <sup>15</sup>N, which each provide a mass increment of ~1 Da/atom. Two common approaches used for incorporation of stable isotopes are metabolic labeling<sup>43–45</sup> and post-harvest derivatization<sup>46</sup>. Both methods facilitate the comparison of proteins between two or more samples following differential labeling (Fig. 3). Quantification in both strategies relies on the fact that unlabeled and labeled versions of a peptide are chemically identical and co-elute during reverse-phase chromatography, so long as <sup>13</sup>C and <sup>15</sup>N isotopes are used for labeling. A peptide derived from a protein present at 2-fold higher concentration in one sample than the other is represented by a pair of peptides with peak intensities differing by 2-fold (Fig. 3). Since the large-scale characterization of ubiquitinated proteins is often partially obstructed by huge excesses of peptides from ubiquitin, any method which is transparent to ubiquitin peptides offers a unique analytical benefit. Since the isotope coded affinity tagging (ICAT)

strategy is based upon enrichment of cysteine containing peptides, a residue which is completely lacking within ubiquitin, the quantification of ubiquitinated proteins using ICAT offers this benefit (Fig. 3). In any case, by adapting workflows for shotgun proteomics (Fig. 2B), thousands of peptides and their corresponding isotopic pairs can be sequenced and quantified from a single sample.

Although mass spectrometry offers a powerful tool for identifying ubiquitin- and Ubl-substrates, a number of unresolved issues remain. Despite many advances, MS data is inherently biased toward more abundant substrates. The effects of epitope tags on ubiquitin and Ubl proteins remain incompletely understood, including whether purification biases exist and whether ubiquitin pathway enzymes [DK15] utilize tagged and wild-type ubiquitin with equal efficiency. Ongoing work seeks to determine if ubiquitin-binding proteins or ubiquitin antibodies may work efficiently as affinity reagents [DK16] in order to lessen the need for epitope tagged ubiquitin.

Validation of database matched proteins as true substrates remains a major endeavor since ubiquitin/Ubl- and substrate-associated proteins can co-purify with true targets, particularly when affinity purifications are performed solely under non-denaturing conditions. Several non MS-based approaches may be useful for validating conjugates and identifying low abundance substrates. Genetic methodologies, including two-hybrid screens and high-copy suppressor screens, have been used to identify SUMO substrates<sup>25</sup>. High-throughput immunoprecipitation and western blot analysis of epitope-tagged proteins, recently used to identify direct and indirect SUMO targets<sup>47</sup>, may be a feasible approach for systematic validation for substrates of ubiquitin and other UbIs identified by shotgun sequencing.

In efforts to increase the stringency of ubiquitin/Ubl substrates identified by shotgun sequencing, a general trend towards double-affinity purification procedures has emerged<sup>21, 22, 24, 25</sup>. Several groups studying both ubiquitin and SUMO have demonstrated that these approaches can be useful for identifying low-abundance conjugates while minimizing the number of false positives. Two papers using a single-step versus a double-affinity purification provide a relevant basis for comparing and contrasting the enrichment and analytical aspects of these methods [DK17]<sup>23, 24</sup>. While generating largely overlapping lists, an optimized MUDPIT analysis<sup>23</sup> successfully identified more proteins than double-affinity purification coupled to GeLC-MS<sup>24</sup>. One advantage of the double-affinity approach was that it acted as independent confirmation that identified proteins were true substrates, even in the absence of follow up studies. Interestingly, both protocols have issues with false negatives. Despite identifying more proteins, the incidence of false negatives following single-step purification may actually underestimate the number of modified proteins, since many proteins are ruled out as potential substrates based on their identification in the negative control sample. By contrast, stringency inherent to the multi-step purification results in the loss of some true substrates prior to MS analysis.

## Characterization of enzymes regulating ubiquitin and Ubl systems

In addition to analyses focused on substrate identification, MS-based proteomics has been useful in characterizing many enzymatic components of the ubiquitin system. Tandem mass spectrometry [DK18] has been used to identify novel members of the yeast anaphase promoting complex (APC)<sup>48, 49</sup> and to characterize a network of Skp-Cullin-F<sub>box</sub> (SCF) ubiquitin-ligases<sup>50, 51</sup>. Enzymes regulating ubiquitination of substrates such as p53<sup>52-54</sup>, IκB<sup>55</sup>, histones<sup>56, 57</sup>, c-Jun<sup>58</sup>, and the EGF receptor<sup>59</sup> have also been studied. Comprehensive analyses of the intact 26S proteasome have yielded a number of novel proteasome-associated proteins including both the APC and SCF E3 ligases<sup>60</sup>, while helping demonstrate that the Rpn11 subunit acts as a deubiquitinating enzyme for incoming proteasome

substrates<sup>61</sup>. In a series of reports, activity-based probes, coupled with MS, have been used to characterize known DUB enzymes, identify a novel family of DUBs, and to perform expression profiling of DUBs in various cells and tissues<sup>62–64</sup>. The versatility of these probes was further extended to the characterization of deconjugating<sub>[DK19]</sub> enzymes for SUMO, Nedd8 and ISG15<sup>65</sup>.

## Identification of ubiquitin and Ub-like modification sites

Large scale analysis of post-translational modifications, particularly protein phosphorylation, remains a fruitful area of research for MS-based proteomics<sup>66, 67</sup>. By using mass spectrometry coupled to phosphopeptide enrichment strategies such as immobilized metal affinity chromatography<sup>68</sup>, strong cation exchange chromatography<sup>69</sup>, or peptide immunoprecipitation<sup>70</sup>, thousands of modification sites have been identified. Recently, these approaches have been modified in order to identify ubiquitination sites. Unlike with phosphorylation, the identification of consensus ubiquitination sites based upon primary sequence seems unlikely. Rather, we feel that the combination of MS approaches with three dimensional protein structures will be useful in identifying exposed surfaces on which protein modification is occurring at a number of similarly positioned lysines. As the number of identified ubiquitination sites increases, one related benefit may be the identification of recognition motifs for enzymes such as ligases or DUBs. Since only a handful of *bona fide* ubiquitination sites are currently known, significant efforts are needed in this area. <sub>[DK20]</sub>

A number of studies have demonstrated that precise ubiquitination sites can be identified using MS, by taking advantage of the fact that isopeptide-linked ubiquitin is cleaved by trypsin at the junction between Arg-74 and Gly-75, producing a –GG signature peptide (Fig. 4)<sup>18, 20, 39, 71–73</sup>. Two large scale studies have been performed to identify ubiquitination sites from yeast<sup>18, 20</sup>. In more focused studies, modification sites on Gpa1<sup>71</sup>,  $\gamma$ -tubulin<sup>39</sup>, TRAF-1<sup>74</sup>, Smad4<sup>75</sup>, and Met4<sup>73</sup> have been identified using similar methods. Another noteworthy report used MS to characterize the N-terminal ubiquitination of both p21 and ERK3<sup>72</sup>.

Despite these successes, the identification of ubiquitinated lysines has proven to be difficult for many proteins. Whereas site-directed mutagenesis of ubiquitinated residues in Gpa1 and TRAF1 nearly eliminated substrate modification<sup>71, 74</sup>, in many cases, downstream function requires that only one of a subset of lysines be ubiquitinated<sup>76, 77</sup>. When identifying ubiquitination sites on a protein, it is common to find that each individual lysine is modified in only a fraction of the sample. This multiple lysine effect <sub>[DK21]</sub> decreases the abundance of each individual –GG signature peptide. Since many proteins are heterogeneously poly-ubiquitinated, indicated by ubiquitin smears which extend >100 kD on an SDS-PAGE gel, common methods such as gel band analysis must be optimized for each individual substrate. In several cases, this problem has been overcome using multi-dimensional chromatography approaches coupled to MS, in which proteins are digested in solution without SDS-PAGE separation<sup>18, 20, 73</sup>. Additionally, because of the missed lysine cleavage at the ubiquitin modification site, some –GG signature peptides become too large for standard analyses, and require alternate digestion strategies.

In attempts to circumvent some of these complications, several groups have sought to develop new techniques for detecting precise ubiquitination sites. The high mass accuracy of Fourier-transform ion-cyclotron resonance mass spectrometry has been beneficial for identifying ubiquitination sites<sup>78</sup>. Alternate enzymes can be used to digest substrates and generate significantly longer signature peptide (e.g. STLHLVLRRLRGG for GluC)<sup>79</sup>, reminiscent of the lysine modifications encountered following tryptic digestion of several Ubl proteins (SUMO-2, SUMO-3, Hub1). In contrast, if –GG signature peptides lack basic residues and are effectively too long for analysis, the use of multiple enzymes may be beneficial<sup>73</sup>. One disadvantage of

using trypsin is that it forms identical –GG signature peptides for ubiquitin, as well as the Ubl proteins NEDD8 and ISG15. Alternate enzymes which produce slightly larger signatures easily differentiate between the three. These alternate or multiple enzyme strategies are particularly effective when focusing on a single protein, in experiments where the approach can be directly tailored based upon the amino acid sequence of the substrate.

Another approach involves amino-terminal labeling of –GG signature peptides either by sulfonation<sup>80</sup> or modification by fluoruous affinity tags<sup>81</sup>. One potential benefit of such a post-digest modification is that since –GG signature peptides effectively have two amino-termini, and typical tryptic peptides only one, enrichment of multiply modified peptides may be an effective strategy for concentrating –GG signature peptides<sup>81</sup>. Additionally, strategies developed for large-scale phosphopeptide identification such as peptide immunoprecipitation<sup>70</sup> or ion exchange chromatography<sup>69</sup> may also be useful for capturing –GG signature peptides. While further work is needed to test the utility of newly developed methods in the large scale identification of ubiquitination sites from *in vivo* substrates, work involving phosphopeptides suggests that prefractionation of –GG signature peptides will be a key component of any successful strategy.

## Defining the structure of poly-ubiquitin chains

The formation of a poly-ubiquitin chain provides an opportunity for increased regulatory complexity within many ubiquitin-dependent processes<sup>82</sup> (Fig. 1). Much of the current understanding of the function and frequency of various poly-ubiquitin linkages is derived either from experiments performed using lysine-to-arginine mutant forms of ubiquitin<sup>33, 83, 84</sup> or studies performed with synthetic poly-ubiquitin chains<sup>85–87</sup>. Additional reports suggest that antibodies with preferences for and/or against specific forms of ubiquitin may be used to differentiate between monoubiquitin and poly-ubiquitin in samples<sup>88–90</sup>. [DK22] The indirect nature of these molecular approaches relies upon a number of assumptions. While preventing the formation of certain ubiquitin-ubiquitin linkages, substitution of highly conserved lysines within the ubiquitin sequence could also have uncharacterized effects on protein-protein interactions, enzyme activities, or poly-ubiquitin chain formation within a complex system. Synthetic chains are a proven biochemical tool, but may not be representative of all substrates carrying a similar modification. Despite the proven utility of these molecular strategies, there remains a need for a direct, quantitative method to analyze poly-ubiquitin chain composition. Since trypsin digestion of a poly-ubiquitin chain produces unique –GG signature peptides for each possible linkage, analysis of these peptides can allow direct measurement of linkage types and frequencies (Fig. 4).

Large-scale shotgun sequencing studies have identified –GG signature peptides from the various types of poly-ubiquitin chains. For example, in yeast, it was shown that all seven lysines in ubiquitin can participate in poly-ubiquitin chain formation *in vivo* [DK23]<sup>18</sup>. Furthermore, in the same study, a branched ubiquitin peptide was also identified in which a single ubiquitin was simultaneously modified at two adjacent residues (K29 and K33). These observations suggest that chain formation through alternative lysines within ubiquitin may provide new levels of regulatory complexity. Analysis of ubiquitinated Met4 purified from yeast showed it to be modified by K48-linked chains, despite the fact that the substrate was stable and not being targeted for degradation<sup>73</sup>. Subsequent reports have also used mass spectrometry to examine poly-ubiquitin linkages formed during *in vitro* reactions. In two such examples, the yeast Ufd2 ligase complex was shown to synthesize both K48 and K63 chains<sup>91</sup>, while the BRCA1/BARD1 heterodimer generated K6 linkages<sup>92, 93</sup>. In all of these studies which utilize either shotgun sequencing or more focused MS/MS approaches to identify the presence of specific poly-ubiquitin chain linkages, caution should be used. Because of peptide specific chromatographic properties, as is the case with the diminutive K29 branched peptide, or MS

insensitivity, as with the K11 branched peptide from yeast, absence of evidence should not necessarily be construed as evidence of absence.

Many labs are involved in understanding the biological consequences of protein polyubiquitination through noncanonical ubiquitin-ubiquitin linkages. Whereas K48 and K63 linked chains have been tied to many effects within cells, little is known about chains formed through K6, K11, K29 and others. Recently, MS was used to demonstrate the effect of biotinylation and/or mutation of K6 in processes such as ubiquitin-conjugation and proteasomal degradation<sup>94</sup>. Another group purified and identified linkage specific binding proteins associated with K29-linked chains, including two proteins from the ubiquitin pathway, Ubp14/ Isopeptidase T and Ufd3<sup>95</sup>. The extension of these studies to other chain linkages should further elucidate the mechanisms through which poly-ubiquitin chains dictate biological functions.

MS-based proteomics [DK24] has begun to make progress towards a direct, quantitative method for analyzing poly-ubiquitin chains (D.K. and S.G. manuscript in preparation). The absolute quantification (AQUA) method<sup>96</sup> may be useful in revalidating many basic tenets within the ubiquitin field, including discoveries made using mutant ubiquitin in genetic and biochemical analyses. For example, the *in vivo* role of specific poly-ubiquitin linkages, including the highly abundant K48-linkage, on chain formation, processing, and substrate degradation are not fully understood. The modification of substrates by multiple mono-ubiquitins [DK25] has been reported to have distinct effects on the localization of proteins such as p53<sup>97</sup> and receptor tyrosine kinases<sup>90</sup>, while oligo-ubiquitination (short poly-ubiquitin chains, e.g. Ub3) was recently shown to be a regulated intermediate in protein degradation<sup>98</sup>. In both cases, since AQUA would facilitate measurements of frequency for monoubiquitination and various poly-ubiquitin chain linkages, it may be useful in testing the generality of these observations and the possibility of extending them into a broader biological context.

[DK26]

## Emerging concepts and concluding remarks

The 2004 Nobel Prize in chemistry was awarded to Hershko, Ciechanover and Rose in recognition of the central importance of ubiquitin in regulating protein degradation. Since that initial discovery, the ubiquitin system has matured from a biochemical explanation for how proteins are degraded to a ubiquitous regulatory network deserving of its name. Nonetheless, a number of fundamental questions remain as to the biochemical mechanisms and functional consequences of ubiquitination. Simple questions, such as which ligases are responsible for targeting individual substrates for degradation, are already being addressed using shotgun sequencing methods and stable isotope based quantitation. Subsequently, it will be important to validate recent advances in our understanding of ubiquitin receptors/ubiquitin binding factors<sup>98, 99</sup> in the format of large scale experiments looking at bulk substrate turnover. In the case of proteasome-independent signaling by ubiquitin, the continued identification of substrate modification sites and poly-ubiquitin chain linkages will be necessary to understand mechanistic specificity. In all cases, the coupling of MS-based approaches with genetic mutants<sup>33</sup>, small molecule perturbants<sup>100</sup>, or technologies such as siRNA, has the potential to define both substrate specific and global aspects of protein ubiquitination.

Arguably, the largest remaining void in the ubiquitin research field is our understanding of the structure and function of poly-ubiquitin chains. As was the case with ubiquitin function, the mechanisms of poly-ubiquitin chain formation are more complex than originally envisioned. Circumstantial evidence points toward the possibility that a proportion of poly-ubiquitin chains may be synthesized with mixed linkages, although neither the frequency nor relevance of such structures has been sufficiently addressed. Specifically, it is known that forked poly-ubiquitin

chains can form within cells [DK27]through adjacent lysines within ubiquitin (e.g., K29-33)<sup>18</sup>. Does poly-ubiquitin chain forking occur through non-adjacent lysines within ubiquitin, and if so, how abundant is it within cells? Furthermore, are these heterogeneous chains or forked structures esoteric, or rather, indicative of an additional level of regulatory complexity reminiscent of complex carbohydrate signaling? With regards to K48-linked chains, given their defined role in proteolysis, how can the high relative abundance of these chains within living cells be explained? Currently, it is only possible to speculate at models such as the inherent stability of some K48 chains, the presence of linkage specificity factors regulating assembly or disassembly of non-K48 linkages, and/or preferences of E2 enzymes for K48 synthesis. Only direct, quantitative analysis of poly-ubiquitin chains *in vitro* and *in vivo* will address these questions.

MS-based proteomics, and particularly the development of stable isotope based quantitation, has helped transition many questions within the ubiquitin field from the arena of theory and speculation into a landscape where they may be addressed by hypothesis-driven experimentation. Proteomic experiments will allow researchers to directly address numerous biochemical and functional aspects of the ubiquitin system which have been transparent[DK28] to existing molecular techniques. Only through the union of these two fields can the full biological scope of protein ubiquitination be revealed.

#### Acknowledgements

Work in the lab of S.P.G. is supported by National Institutes of Health grants HG00041 and GM67945.

#### References

1. Hershko A, Ciechanover A. The ubiquitin system. *Annu Rev Biochem* 1998;67:425–479. [PubMed: 9759494]
2. Pickart CM. Mechanisms underlying Ubiquitination. *Nat Rev Cell Mol Biol* 2001;70:503–533.
3. Finley D, Ciechanover A, Varshavsky A. Ubiquitin as a central cellular regulator. *Cell* 2004;116:S29–32. [PubMed: 15055578]
4. Ciechanover A, Heller H, Elias S, Haas AL, Hershko A. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proc Natl Acad Sci USA* 1980;77:1365–1368. [PubMed: 6769112]
5. Hershko A, Ciechanover A, Heller H, Haas AL, Rose IA. Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc Natl Acad Sci U S A* 1980;77:1783–1786. [PubMed: 6990414]
6. Finley D, Ciechanover A, Varshavsky A. Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85. *Cell* 1984;37:43–55. [PubMed: 6327059]
7. Ciechanover A, Finley D, Varshavsky A. Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. *Cell* 1984;37:57–66. [PubMed: 6327060]
8. Johnson ES. Protein modification by SUMO. *Annu Rev Biochem* 2004;73:355–382. [PubMed: 15189146]
9. Ritchie KJ, Zhang DE. ISG15: the immunological kin of ubiquitin. *Semin Cell Dev Biol* 2004;15:237–246. [PubMed: 15209384]
10. Chiba T, Tanaka K. Cullin-based ubiquitin ligase and its control by NEDD8-conjugating system. *Curr Protein Pept Sci* 2004;5:177–184. [PubMed: 15180522]
11. Pan ZQ, Kentsis A, Dias DC, Yamoah K, Wu K. Nedd8 on cullin: building an expressway to protein destruction. *Oncogene* 2004;23:1985–1997. [PubMed: 15021886]
12. Wolters DA, Washburn MP, Yates JR 3rd. An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* 2001;73:5683–5690. [PubMed: 11774908]
13. Eng J, McCormack AL, Yates JR III. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom* 1994;5

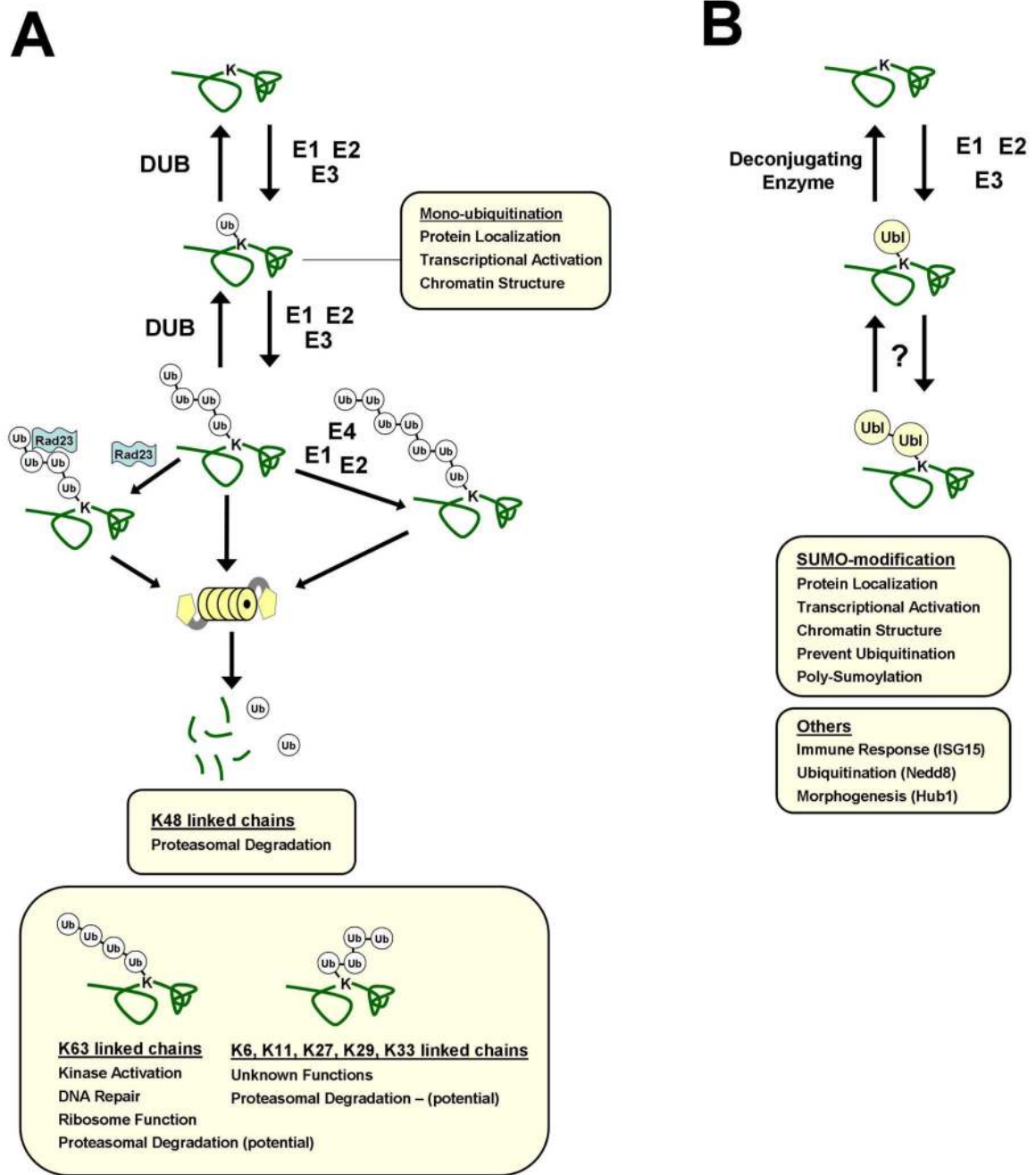


14. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999;20:3551–3567. [PubMed: 10612281]
15. Lasonder E, et al. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature* 2002;419:537–542. [PubMed: 12368870]
16. Everley PA, Krijgsveld J, Zetter BR, Gygi SP. Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol Cell Proteomics* 2004;3:729–735. [PubMed: 15102926]
17. Florens L, et al. A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* 2002;419:520–526. [PubMed: 12368866]
18. Peng J, et al. A proteomics approach to understanding protein ubiquitination. *Nat Biotech* 2003;21:921–926.
19. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol* 2001;19:242–247. [PubMed: 11231557]
20. Hitchcock AL, Auld K, Gygi SP, Silver PA. A subset of membrane-associated proteins is ubiquitinated in response to mutations in the endoplasmic reticulum degradation machinery. *Proc Natl Acad Sci USA* 2003;100:12735–12740. [PubMed: 14557538]
21. Mayor, T., Russell-Lipford, J., Graumann, J., Smith, G.T., & Deshaies, R.J. Analysis of poly-ubiquitin conjugates reveals that the Rpn10 substrate receptor contributes to the turnover of multiple proteasome targets. *Mol. Cell. Proteomics*, Epub (2005).
22. Zhou W, Ryan JJ, Zhou H. Global analyses of sumoylated proteins in *Saccharomyces cerevisiae*. Induction of protein sumoylation by cellular stresses. *J Biol Chem* 2004;279:32262–32268. [PubMed: 15166219]
23. Wohlschlegel JA, Johnson ES, Reed SI, Yates 3rd JR. Global analysis of protein sumoylation in *Saccharomyces cerevisiae*. *J Biol Chem* 2004;279:45662–45668. [PubMed: 15326169]
24. Denison C, et al. Proteomic insights into protein sumoylation. *Mol Cell Proteomics* 2005;4:246–254. [PubMed: 15542864]
25. Hannich JT, et al. Defining the SUMO-modified proteome by multiple approaches in *Saccharomyces cerevisiae*. *J Biol Chem* 2005;280:4102–4110. [PubMed: 15590687]
26. Panse VG, Hardeland U, Werner T, Kuster B, Hurt E. A proteome-wide approach identifies sumoylated substrate proteins in yeast. *J Biol Chem* 2004;279:41346–41351. [PubMed: 15292183]
27. Kirkpatrick DS, Weldon SF, Tsapralis G, Liebler DC, Gandolfi AJ. Proteomic identification of ubiquitinated proteins from human cells expressing His-tagged ubiquitin. *Proteomics* 2005;5:2104–2111. [PubMed: 15852347]
28. Vertegaal AC, et al. A proteomic study of SUMO-2 target proteins. *J Biol Chem* 2004;279:33791–33798. [PubMed: 15175327]
29. Manza LL, et al. Global shifts in protein sumoylation in response to electrophile and oxidative stress. *Chem Res Tox* 2004;17:1706–1715.
30. Zhao Y, Kwon SW, Anselmo A, Kaur K, White MA. Broad spectrum identification of cellular small ubiquitin-related modifier (SUMO) substrate proteins. *J Biol Chem* 2004;279:20999–21002. [PubMed: 15016812]
31. Rosas-Acosta G, Russell WK, Deyrieux A, Russell DH, Wilson VG. A universal strategy for proteomic studies of SUMO and other ubiquitin-like modifiers. *Mol Cell Proteomics* 2005;4:56–72. [PubMed: 15576338]
32. Tsigotis M, et al. Analysis of ubiquitination in vivo using a transgenic mouse model. *Biotechniques* 2001;31:120–126. [PubMed: 11464505]128, 130
33. Finley D, et al. Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant. *Mol Cell Biol* 1994;14:5501–5509. [PubMed: 8035826]
34. Weekes J, et al. Hyperubiquitination of proteins in dilated cardiomyopathy. *Proteomics* 2003;3:208–216. [PubMed: 12601813]
35. Gururaja T, Li W, Noble WS, Payan DG, Anderson DC. Multiple functional categories of proteins identified in an in vitro cellular ubiquitin affinity extract using shotgun peptide sequencing. *J Prot Res* 2003;2:394–404.

36. Li T, et al. Sumoylation of heterogeneous ribonucleoproteins, zinc finger proteins, and nuclear pore complex proteins: a proteomic analysis. *Proc Natl Acad Sci USA* 2004;101:8551–8556. [PubMed: 15161980]
37. Gocke CB, Yu H, Kang J. Systematic identification and analysis of mammalian small ubiquitin-like modifier substrates. *J Biol Chem* 2005;280:5004–5012. [PubMed: 15561718]
38. Sato K, et al. Nucleophosmin/B23 is a candidate substrate for the BRCA1-BARD1 ubiquitin ligase. *J Biol Chem* 2004;279:30919–30922. [PubMed: 15184379]
39. Starita LM, et al. BRCA1-dependent ubiquitination of gamma-tubulin regulates centrosome number. *Mol Cell Biol* 2004;24:8457–8466. [PubMed: 15367667]
40. Chang TL, Cubillos FF, Kakhniashvili DG, Goodman SR. Ankyrin is a target of spectrin's E2/E3 ubiquitin-conjugating/ligating activity. *Cell Mol Biol (Noisy-le-grand)* 2004;50:59–66. [PubMed: 15040428]
41. Chung DL, et al. In vitro modification of centromere protein CENP-C fragments by small ubiquitin-like modifier (SUMO) protein: definitive identification of the modification sites by tandem mass spectrometry analysis of isopeptides. *J Biol Chem* 2004;279:39653–39662. [PubMed: 15272016]
42. Schirmer EC, Florens L, Guan T, Yates JR 3rd, Gerace L. Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science* 2003;301:1380–1382. [PubMed: 12958361]
43. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc Natl Acad Sci USA* 1999;96:6591–6596. [PubMed: 10359756]
44. Ong SE, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002;1:376–386. [PubMed: 12118079]
45. Andersen JS, et al. Nucleolar proteome dynamics. *Nature* 2005;433:77–83. [PubMed: 15635413]
46. Gygi SP, et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotech* 1999;17:994–999.
47. Wykoff DD, O'Shea EK. Identification of sumoylated proteins by systematic immunoprecipitation of the budding yeast proteome. *Mol Cell Proteomics* 2005;4:73–83. [PubMed: 15596868]
48. Zachariae W, et al. Mass spectrometric analysis of the anaphase-promoting complex from yeast: identification of a subunit related to cullins. *Science* 1998;279:1216–1219. [PubMed: 9469814]
49. Yoon HJ, et al. Proteomics analysis identifies new components of the fission and budding yeast anaphase-promoting complexes. *Curr Biol* 2002;12:2048–2054. [PubMed: 12477395]
50. Seol JH, et al. Cdc53/cullin and the essential Hrt1 RING-H2 subunit of SCF define a ubiquitin ligase module that activates the E2 enzyme Cdc34. *Genes Dev* 1999;13:1614–1626. [PubMed: 10385629]
51. Seol JH, Shevchenko A, Deshaies RJ. Skp1 forms multiple protein complexes, including RAVE, a regulator of V-ATPase assembly. *Nat Cell Biol* 2001;3:384–391. [PubMed: 11283612]
52. Li M, et al. Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* 2002;416:648–653. [PubMed: 11923872]
53. Rajendra R, et al. Topors functions as an E3 ubiquitin ligase with specific E2 enzymes and ubiquitinates p53. *J Biol Chem* 2004;279:36440–36444. [PubMed: 15247280]
54. Harada JN, Shevchenko A, Pallas DC, Berk AJ. Analysis of the adenovirus E1B-55K-anchored proteome reveals its link to ubiquitination machinery. *J Virol* 2002;76:9194–9206. [PubMed: 12186903]
55. Yaron A, et al. Identification of the receptor component of the IkappaB alpha-ubiquitin ligase. *Nature* 1998;396:590–594. [PubMed: 9859996]
56. Liu Z, Oughtred R, Wing SS. Characterization of E3Histone, a novel testis ubiquitin protein ligase which ubiquitinates histones. *Mol Cell Biol* 2005;25:2819–2831. [PubMed: 15767685]
57. Wang H, et al. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 2004;431:873–878. [PubMed: 15386022]
58. Wertz IE, et al. Human De-etiolated-1 regulates c-Jun by assembling a CUL4A ubiquitin ligase. *Science* 2004;303:1371–1374. [PubMed: 14739464]
59. Feshchenko EA, et al. TULA: an SH3- and UBA-containing protein that binds to c-Cbl and ubiquitin. *Oncogene* 2004;23:4690–4706. [PubMed: 15107835]

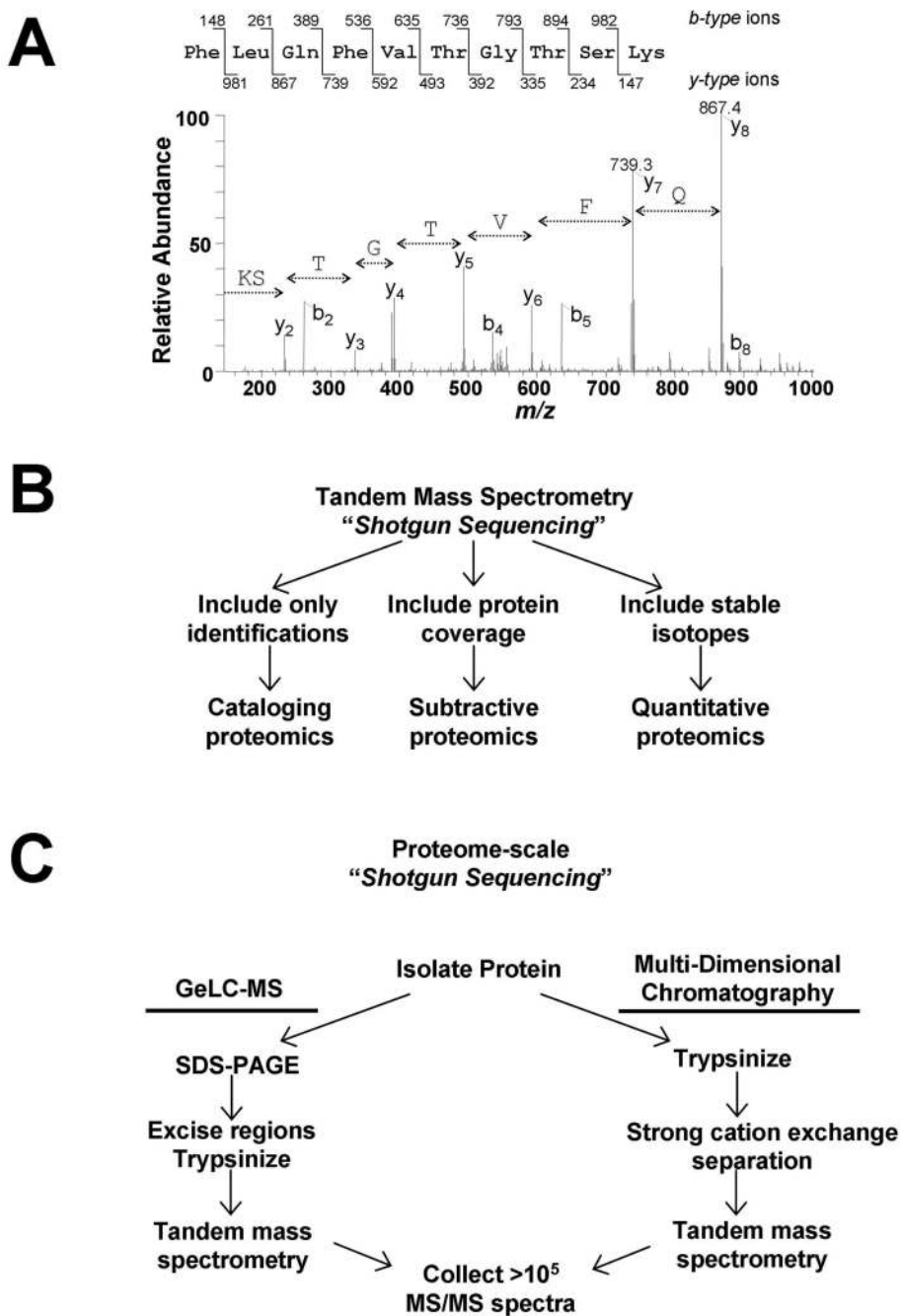
60. Verma R, et al. Proteasomal proteomics: identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes. *Mol Cell Biol* 2000;11:3425–3439.
61. Verma R, et al. Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* 2002;298:611–615. [PubMed: 12183636]
62. Borodovsky A, et al. Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family. *Chem Biol* 2002;9:1149–1159. [PubMed: 12401499]
63. Ovaa H, et al. Activity-based ubiquitin-specific protease (USP) profiling of virus-infected and malignant human cells. *Proc Natl Acad Sci USA* 2004;101:2253–2258. [PubMed: 14982996]
64. Hemelaar J, et al. Chemistry-based functional proteomics: mechanism-based activity-profiling tools for ubiquitin and ubiquitin-like specific proteases. *J Prot Res* 2004;3:268–276.
65. Hemelaar J, et al. Specific and covalent targeting of conjugating and deconjugating enzymes of ubiquitin-like proteins. *Mol Cell Biol* 2004;24:84–95. [PubMed: 14673145]
66. Mann M, Jensen ON. Proteomic analysis of post-translational modifications. *Nat Biotechnol* 2003;21:255–261. [PubMed: 12610572]
67. Chen WG, White FM. Proteomic analysis of cellular signaling. *Expert Rev Proteomics* 2004;1:89–100.
68. Ficarro SB, et al. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol* 2002;20:301–305. [PubMed: 11875433]
69. Beausoleil SA, et al. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci USA* 2004;101:12130–12135. [PubMed: 15302935]
70. Rush J, et al. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat Biotechnol* 2005;23:94–101. [PubMed: 15592455]
71. Marotti LA Jr, Newitt R, Wang Y, Aebersold R, Dohlman HG. Direct identification of a G protein ubiquitination site by mass spectrometry. *Biochem* 2002;41:5067–5074. [PubMed: 11955054]
72. Coulombe P, Rodier G, Bonneil E, Thibault P, Meloche S. N-Terminal ubiquitination of extracellular signal-regulated kinase 3 and p21 directs their degradation by the proteasome. *Mol Cell Biol* 2004;24:6140–6150. [PubMed: 15226418]
73. Flick K, et al. Proteolysis-independent regulation of the transcription factor Met4 by a single Lys 48-linked ubiquitin chain. *Nat Cell Biol* 2004;6:634–641. [PubMed: 15208638]
74. Lee JS, Hong US, Lee TH, Yoon SK, Yoon JB. Mass spectrometry analysis of tumor necrosis factor receptor-associated factor 1 ubiquitination mediated by cellular inhibitor of apoptosis 2. *Proteomics* 2004;4:3376–3382. [PubMed: 15468071]
75. Moren A, et al. Differential ubiquitination defines the functional status of the tumor suppressor Smad4. *J Biol Chem* 2003;278:33571–33582. [PubMed: 12794086]
76. Petroski MD, Deshaies RJ. Context of multiubiquitin chain attachment influences the rate of Sic1 degradation. *Mol Cell* 2003;11:1435–1444. [PubMed: 12820958]
77. Terrell J, Shih S, Dunn R, Hicke L. A function for monoubiquitination in the internalization of a G protein-coupled receptor. *Mol Cell* 1998;1:193–202. [PubMed: 9659916]
78. Cooper HJ, et al. Identification of sites of Ubiquitination in proteins: a fourier transform ion cyclotron resonance mass spectrometry approach. *Anal Chem* 2004;76:6982–6988. [PubMed: 15571350]
79. Warren MR, Parker CE, Mocanu V, Klapper D, Borchers CH. Electrospray ionization tandem mass spectrometry of model peptides reveals diagnostic fragment ions for protein ubiquitination. *Rapid Commun Mass Spectrom* 2005;19:429–437. [PubMed: 15655800]
80. Wang D, Cotter RJ. Approach for determining protein ubiquitination sites by MALDI-TOF mass spectrometry. *Anal Chem* 2005;77:1458–1466. [PubMed: 15732931]
81. Brittain SM, Ficarro SB, Brock A, Peters EC. Enrichment and analysis of peptide subsets using fluorinated affinity tags and mass spectrometry. *Nat Biotechnol* 2005;23:463–468. [PubMed: 15768030]
82. Pickart CM, Fushman D. Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol* 2004;8:610–616. [PubMed: 15556404]
83. Chau V, et al. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* 1989;243:1576–1583. [PubMed: 2538923]

84. Spence J, Sadis S, Haas AL, Finley D. A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. *Mol Cell Biol* 1995;15:1265–1273. [PubMed: 7862120]
85. Thrower JS, Hoffman L, Rechsteiner M, Pickart CM. Recognition of the polyubiquitin proteolytic signal. *EMBO J* 2000;19:94–102. [PubMed: 10619848]
86. Hofmann RM, Pickart CM. In vitro assembly and recognition of Lys-63 polyubiquitin chains. *J Biol Chem* 2001;276:27936–27943. [PubMed: 11369780]
87. Mastrandrea LD, You J, Niles EG, Pickart CM. E2/E3-mediated assembly of lysine 29-linked polyubiquitin chains. *J Biol Chem* 1999;274:27299–27306. [PubMed: 10480950]
88. Takada K, Hibi N, Tsukada Y, Shibasaki T, Ohkawa K. Ability of ubiquitin radioimmunoassay to discriminate between monoubiquitin and multi-ubiquitin chains. *Biochim Biophys Acta* 1996;1290:282–288. [PubMed: 8765132]
89. Takada K, et al. Immunoassay for the quantification of intracellular multi-ubiquitin chains. *Eur J Biochem* 1995;233:42–47. [PubMed: 7588772]
90. Haglund K, et al. Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat Cell Biol* 2003;5:461–466. [PubMed: 12717448]
91. Saeki Y, Tayama Y, Toh-e A, Yokosawa H. Definitive evidence for Ufd2-catalyzed elongation of the ubiquitin chain through Lys48 linkage. *Biochem Biophys Res Commun* 2004;320:840–845. [PubMed: 15240124]
92. Wu-Baer F, Lagrazon K, Yuan W, Baer R. The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. *J Biol Chem* 2003;278:34743–34746. [PubMed: 12890688]
93. Nishikawa H, et al. Mass spectrometric and mutational analyses reveal Lys6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase. *J Biol Chem* 2004;279:3916–3924. [PubMed: 14638690]
94. Shang, F. et al. Lys-6-modified ubiquitin inhibits ubiquitin-dependent protein degradation. *J. Biol. Chem.* (2005).
95. Russell NS, Wilkinson KD. Identification of a novel 29-linked polyubiquitin binding protein, Ufd3, using polyubiquitin chain analogues, *Biochemistry*. *Biochem* 2004;43:4844–4854. [PubMed: 15096053]
96. Kirkpatrick DS, Gerber SA, Gygi SP. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* 2005;35:265–273. [PubMed: 15722223]
97. Li M, et al. Mono- versus polyubiquitination: differential control of p53 fate by mdm2. *Science* 2003;302:1972–1975. [PubMed: 14671306]
98. Richly H, et al. A series of ubiquitin binding factors Cdc48/p97 to substrate multiubiquitylation and proteasomal targeting. *Cell* 2005;120:73–84. [PubMed: 15652483]
99. Verma R, Oania R, Graumann J, Deshaies RJ. Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system. *Cell* 2004;118:99–110. [PubMed: 15242647]
100. Verma R, et al. Ubistatins inhibit proteasome-dependent degradation by binding the ubiquitin chain. *Science* 2004;306:117–120. [PubMed: 15459393]



**Figure 1.** Mechanisms of protein modification by ubiquitin (a) and Ubl proteins (b). (a) Protein substrates are modified by ubiquitin and Ubl-proteins at a lysine (K) residue(s). (1) Ubiquitin attachment to a substrate is catalyzed by the coordinated actions of an E1<sub>[DK29]</sub> activating enzyme, E2 conjugating enzyme, and E3 ligase (2) Poly- ubiquitin chains can be formed through any of 7 lysine residues within ubiquitin. Deubiquitinating enzymes (DUBs) reverse ubiquitination and shorten poly-ubiquitin chains. (3) A subset of factors, termed E4 enzymes, can further lengthen poly- ubiquitin chains. (4, 5) Degradation of ubiquitinated proteins occurs through a mechanism largely dependent on ubiquitin-ubiquitin linkages formed through K48 of ubiquitin. Poly- ubiquitin chain binding proteins such as Rad23/Hhr23 and Dsk2/Ubiquilin

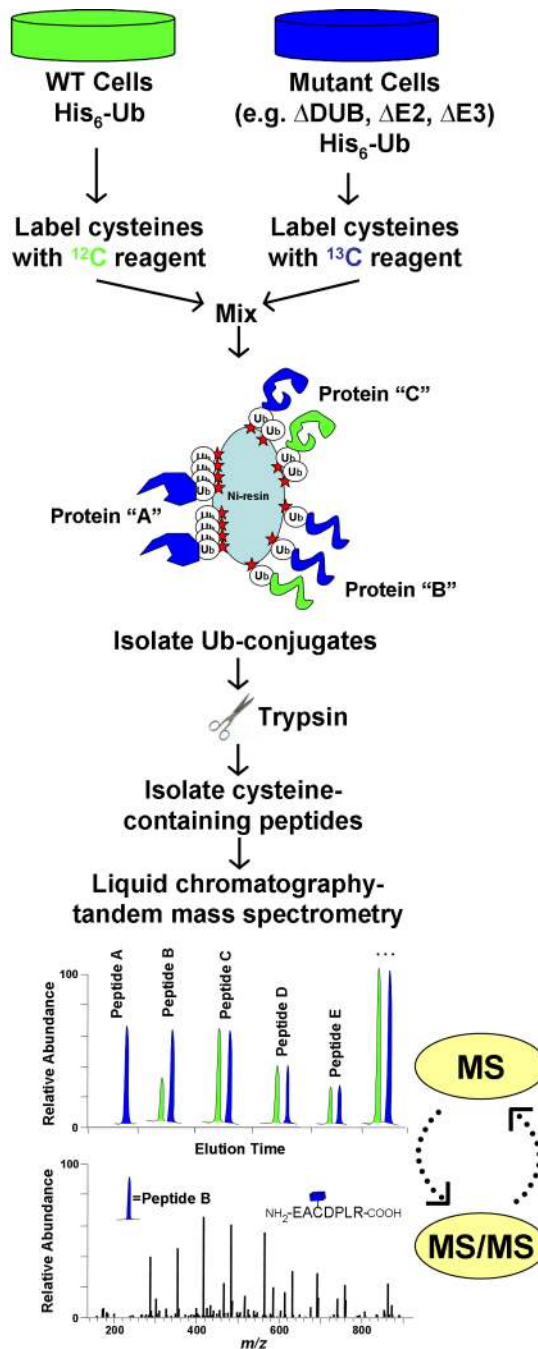
[<sub>DK30</sub>] facilitate recognition and degradation of ubiquitinated substrates by the 26S proteasome. Substrate modifications by mono-ubiquitin, K63-linked chains, or by Ubl-proteins (b) regulate a series of proteasome independent cellular processes. Biological significance of other ubiquitin linkages and many Ubl proteins remain poorly characterized. (b) [<sub>DK31</sub>](1) Ubl attachment occurs through a mechanism analogous to ubiquitination, involving E1, E2, and E3 enzymes. Ubl deconjugating enzymes remove Ubl modifications from substrates. (2) Poly-Ubl chains can be formed for SUMO. Poly-chain formation for other Ubls is currently under investigation. Modification by Ubl proteins can block ubiquitination sites, activate enzymes (including Ub E3 ligases), and affect protein localization.



**Figure 2.** Overview of shotgun sequencing from complex mixtures by mass spectrometry. **a)** Representative tandem mass (MS/MS) spectrum of a peptide from the protein Urebp1. Amino acid sequence for a single peptide can be deduced from the series of fragment ions present in the spectrum. **b)** Large-scale peptide detection via shotgun sequencing can be interpreted in three different ways to provide either lists (cataloging proteomics), differential identifications (subtractive proteomics), or abundance comparisons (quantitative proteomics). **c)** For proteome-scale analyses, huge numbers of MS/MS spectra are collected. Only two routes have proven successful for identifying thousands of proteins from a single sample. Both strategies utilize multiple<sub>[BK33]</sub> steps to fractionate the original sample prior to MS analysis. In one

strategy, SDS-PAGE separation followed by tandem MS analysis of many gel regions (sometimes called GeLC-MS) is used. When using multiple-dimensional chromatography, protein mixtures are directly proteolyzed, and the peptide mixture is separated first by strong cation exchange [DK34]chromatography. In both cases, the final step involves reversed-phase separation of peptides from multiple samples, followed by tandem MS analysis of multiple samples. Both techniques provide the opportunity to collect hundreds of thousands of MS/MS spectra from a single sample in less than 24 hrs.

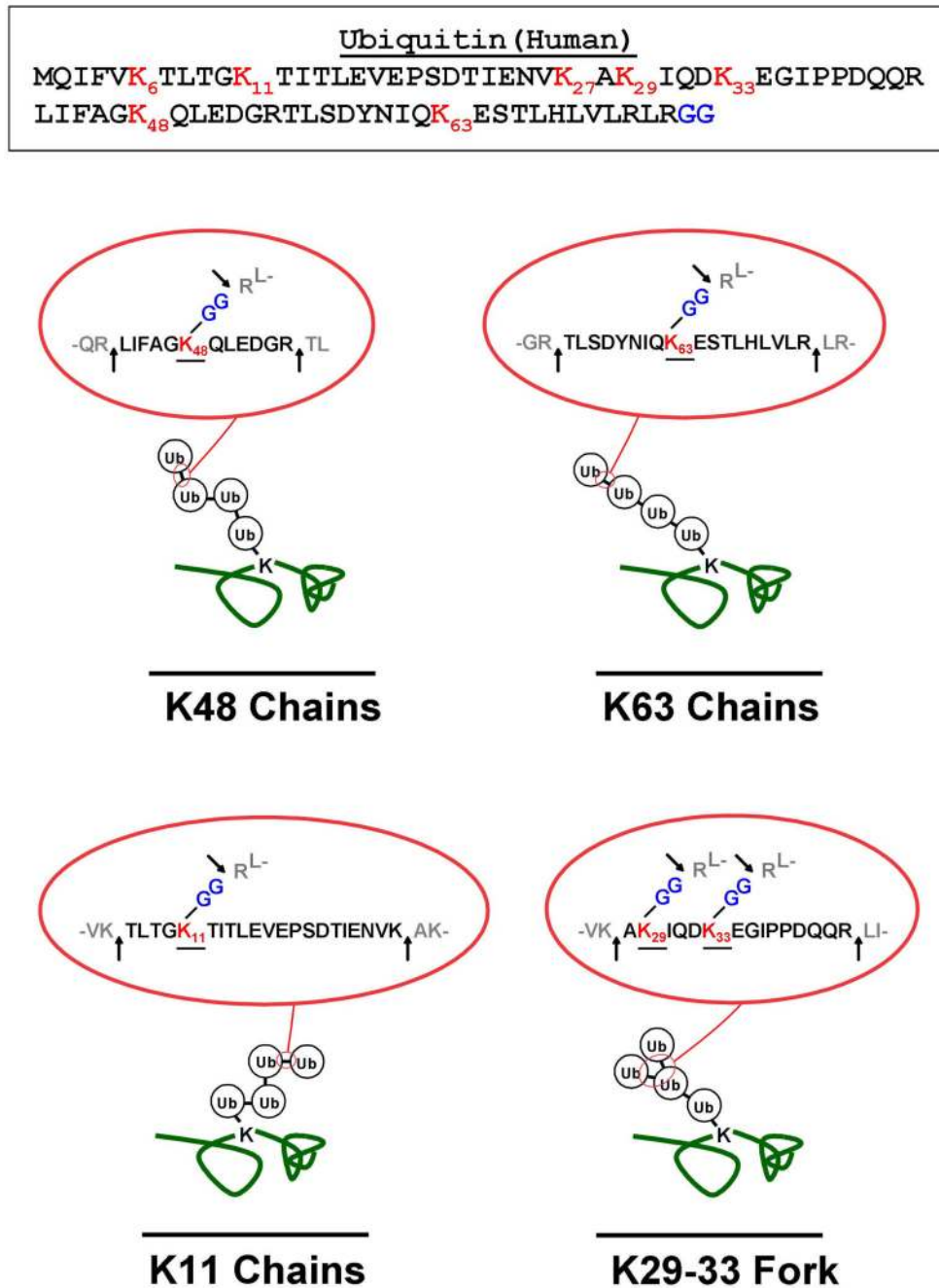




**Figure 3.**

Quantitative profiling of ubiquitin-conjugates using stable isotopes. The isotope coded affinity tag (ICAT) strategy is shown for comparing ubiquitinated proteins between wild type cells and mutant cells lacking a ubiquitin pathway enzyme (e.g. DUB, E2, E3). In an ICAT experiment, protein is harvested from two samples and differentially labeled at cysteinyl residues with either a <sup>12</sup>C- or <sup>13</sup>C-containing reagent. After labeled proteins are mixed, ubiquitinated-proteins are affinity-purified and digested into peptides. The ICAT label allows for further enrichment of cysteine-containing peptides, thus eliminating all peptides derived from ubiquitin. Since <sup>12</sup>C- and <sup>13</sup>C-containing peptides co-elute by reverse-phase chromatography, they can be simultaneously quantified during MS analysis. Rapid cycling between MS and MS/MS modes

allows for the acquisition of both sequence and abundance information for isotopic peptide pairs. In this example, protein “A” is exclusively ubiquitinated in mutant cells, while protein “B” is ubiquitination is increased in mutant cells. Protein “C” is ubiquitinated equally in both samples. Other potential strategies would utilize similar work-flows with minor modifications. For example, metabolic labeling (e.g. SILAC) involves incorporation of isotopes into living cells prior to harvest, and would not utilize the cysteine-enrichment step. A benefit of metabolic labeling would be quantification of non-cysteine containing peptides, which would include most –GG signature peptides.



**Figure 4.** Detecting unique diglycine (–GG) signature peptides for each poly- ubiquitin chain linkage. Each poly-ubiquitin chain conformation can be detected by monitoring a unique signature peptide containing a –GG modified lysine residue, produced by trypsin cleavage. The full amino acid sequence of human ubiquitin is shown at the top. Ubiquitin-ubiquitin linkages correspond to isopeptide bonds formed between the C-terminal glycine (blue) of one ubiquitin and the ε-amino group of a lysine residue (red) within the second. These linkages can be formed through any of seven lysine residues (K6, K11, K27, K29, K33, K48, and K63) (red). As an example, K48, K63, and K11 chains are demonstrated. Digestion of these linkages with trypsin

produces peptides with distinct amino acid sequences (see insets). Trypsin cleaves at lysine and arginine residues within both primary and branched ubiquitin molecules (see arrows in insets), but cannot cut at lysines modified by isopeptide linked ubiquitin (see underlined in insets). The resulting tryptic peptides contain a –GG modified lysine, bearing an additional mass of 114.04 Da, denoting the original position of the modification. Database searching algorithms can utilize both the missed cleavage and –GG modification as search criteria when assigning precise sites of ubiquitination. In the case of a forked poly-ubiquitin chain, as demonstrated through K29 and K33, two –GG modified lysines are detected on the same peptide.