

OPEN ACCESS

Repository of the Max Delbrück Center for Molecular Medicine (MDC) in the Helmholtz Association

<http://edoc.mdc-berlin.de/15132>

Proteomic techniques to probe the ubiquitin landscape

Beaudette, P. and Popp, O. and Dittmar, G.

This is the final version of the manuscript. It is the peer reviewed version of the following article:

Beaudette, P., Popp, O. and Dittmar, G. (2016), Proteomic techniques to probe the ubiquitin landscape. *Proteomics*, 16: 273–287. doi:10.1002/pmic.201500290

which has been published in final form in:

Proteomics
2016 JAN; 16(2): 273-287
2015 DEC 15 (first published online)
doi: [10.1002/pmic.201500290](https://doi.org/10.1002/pmic.201500290)

Publisher: [Wiley-Blackwell](http://www.wiley-blackwell.com)

Copyright © 2015 John Wiley & Sons Ltd

This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#).

Proteomic techniques to probe the ubiquitin landscape

Patrick Beaudette, Oliver Popp & Gunnar Dittmar

Mass-spectrometry, Max-Delbrück Center for molecular medicine, Robert-Rössle Strasse 10, 13125 Berlin, Germany

Corresponding author:

Gunnar Dittmar

Max-Delbrück-Centrum for Molecular Medicine

Robert-Rössle-Str. 10, D-13092 Berlin, Germany

Tel.: +49 30 9406 2642

Fax.: +49 30 9406 49194

Email: gdittmar@mdc-berlin.de

Abbreviations

E1 ubiquitin activating enzyme

E2 ubiquitin conjugating enzyme

E3 ubiquitin ligase

E4 ubiquitin chain assembly factor

PTM post-translational modification

Probing the ubiquitin landscape

UPS ubiquitin-proteasome system

HECT Homologous to the E6-AP Carboxyl Terminus

SRM Selected reaction monitoring

MS mass spectrometry

Keywords

Mass spectrometry, Proteomics, Ubiquitin, Ubiquitin-chain, Ubiquitination

Total number of words

6820

Abstract

Protein ubiquitination is a powerful modulator of cellular functions. Classically linked to the degradation of proteins, it also plays a role in intracellular localization, DNA damage response, vesicle fusion events, and the immune and transcriptional responses. Ubiquitin is versatile and can code for several distinct signals, either by adding a single ubiquitin or forming a chain of ubiquitins on the target protein. The enzymatic cascade associated with the cellular process determines the nature of the modification.

Numerous efforts have been made for the identification of ubiquitin acceptor sites in the target proteins using genetic, biochemical or mass-spectrometry based proteomic methods, such as affinity-based enrichment of ubiquitinated proteins, and antibody-based enrichment of modified peptides. Modern instrumentation enables quantitative mass-spectrometry strategies to identify and characterize hundreds of ubiquitin substrates in a single analysis making it the dominant method for ubiquitin site detection. Characterization of the inter-ubiquitin connectivity in ubiquitin polymers has also moved into focus, with the field of targeted proteomics techniques proving invaluable for identifying and quantifying linkage types found in such polyubiquitin chains.

This review seeks to provide an overview of the many mass-spectrometry based proteomics techniques available for exploring this dynamic field.

Word count abstract: 193

Introduction

The ubiquitin-proteasome system

The flow of information, as formulated by Crick 60 years ago, describes how the genetic instructions encoded in the DNA is first transcribed into RNA followed by translation into proteins[1]. This central dogma of biology is complemented by the final step of protein destruction, as formulated by Schoenheimer, erasing the information[2,3]. Protein destruction is tightly controlled and important for regulated growth. Failure of this regulation can lead to degeneration of cells, ultimately leading to cancer or proteostatic diseases such as Parkinson's and Alzheimer's [4-8].

The most important system for selective protein degradation is the ubiquitin-proteasome system (UPS). The UPS is the major proteolytic system in eukaryotes, with critical functions in cell cycle control, apoptosis, inflammation, transcription, signal transduction, protein quality control, and many other biological processes. The system utilizes the small protein ubiquitin as a covalent modifier. Proteins selected for destruction are first identified by a cascade of enzymes and labeled with ubiquitin. In a second step, the ubiquitin label is recognized by the proteasome, a 2-MDa protease complex. The proteasome unfolds the target protein and threads the polypeptide chain into the destruction chamber, where three proteolytic subunits with different specificities hydrolyze the proteins into smaller peptides. During this process the intact ubiquitin molecules are cleaved from target proteins by three proteasome-associated deubiquitinases and recycled for further rounds of substrate labeling[9].

Ubiquitin itself is a small protein (76 amino acids, 8.6 kDa) found in all eukaryotes and highly conserved between yeast and human. Eukaryotic genomes usually contain several ubiquitin

Probing the ubiquitin landscape

genes in different genetic locations (four in humans and yeast). Ubiquitin is expressed as a precursor protein, either as a polyubiquitin head-to-tail fusion or as N-terminal fusions with ribosomal proteins (Table 1). The gene products are co-translationally processed and mature ubiquitin is released by proteolysis [10].

Mature ubiquitin is transferred to the selected target protein by an enzymatic cascade consisting of three enzymes. The first step of the cascade is the activation of ubiquitin by the ubiquitin-activating enzyme, E1, a step that requires ATP hydrolysis. The C-terminus of ubiquitin forms a thioester with the cysteine side-chain thiol at the active center of the E1 enzyme. The activated ubiquitin is transferred to a ubiquitin-conjugating enzyme, E2. This enzyme, in concert with a ubiquitin ligase, E3, catalyzes the transfer of the activated ubiquitin to an ϵ -amino-group of a lysine side-chain, either in the target protein or on a previously conjugated ubiquitin molecule. The specificity of the E2 enzyme can be modulated via interaction with different E3 proteins, expanding the conjugation possibilities of the system. Finally, ubiquitin-chain formation can be modulated by a ubiquitin chain assembly factor, E4. An exception to the E2/E3 pairs are the HECT-domain containing E3 proteins, which accept an activated ubiquitin on their own active center and transfer the ubiquitin without the aid of an E2 to the target protein[11]. Similarly to other post-translational modifications (PTM), ubiquitination is reversible. While a cascade of several enzymes is necessary for the transfer of ubiquitin to its targets, the removal of ubiquitin is catalyzed by a group of monomeric deubiquitinases (DUBs), ubiquitin-specific cysteine proteases (Figure 1) [12].

Ubiquitin signals

Ubiquitin can code for several distinct signals, by adding just a single ubiquitin or forming a chain of ubiquitins on the target protein. An interesting example is the DNA-binding protein

Probing the ubiquitin landscape

PCNA, which can either be mono- or polyubiquitinated depending on the molecular function[13]. Mono-ubiquitination was first determined to be essential for receptor internalization and intra-cellular transport processes, while the first reported role of the ubiquitin chain was its function in proteasomal degradation of substrates[14-16].

The polyubiquitin chain itself can code several different protein fates depending on the properties of the chain. Each of the seven lysine amino groups of ubiquitin as well as the N-terminus can act as an acceptor site for the extension of the chain, the simplest case being formation of homotypic chains utilizing one ubiquitin lysine type for extension. Depending on the chain architecture different functions have been associated with polyubiquitination. Chains with a lysine 48 (K48) linkage as well as K11-linked chains have been associated with protein degradation at the proteasome. K63-linked chains are generally associated with non-degradative functions, e.g. they can be induced by DNA damage and are associated with a stabilization of the attached substrates. Linear chains using a head-to-tail conjugation at the N-terminus are associated with essential steps in the NF- κ B activation[17-21]. Much less is known about chains with mixed linkage types or branching points, [22], formed when two lysines within a single ubiquitin molecule are used to extend the ubiquitin chain.

In addition to mono- and polyubiquitination of lysine, the occasional modification of the N-terminus of proteins or the modification of other amino acids has been reported. In the absence of all internal lysine residues the N-terminus of proteins like MyoD can be the acceptor site for polyubiquitination[23]. The Coscoy group reported polyubiquitin chains covalently linked to cysteine residues in a MHC-I protein, catalyzed by a viral E3 ligase[24]. Similarly, recent work showed that polyubiquitin chains linked to a substrates cysteine are necessary for cargo translocation across the peroxisomal membranes[25-27] or can be formed during the process of chain-assembly for the formation of polyubiquitin chains [28-30]. Other

Probing the ubiquitin landscape

groups reported the occasional modification of threonines and serine hydroxyl groups by polyubiquitin chains[29,31]. Ubiquitination on serine, threonine and cysteine has to be analyzed with modified proteomic methods because of the chemical nature of the ester linkage, which is subject to hydrolysis during the typical workflow of a proteomic experiment.

Ubiquitin-like modifiers

Ubiquitin has a number of close relatives sharing sequence and structural homology. These proteins fall into two categories, small proteins that are acting as post-translational modifiers like ubiquitin, and those that have a ubiquitin-like domain that is not processed or conjugated. The category of the small ubiquitin-like modifiers comprises the SUMO, Nedd8, Urm1, Apg8 and Apg12 in lower eukaryotes and additionally ISG15, Fat10, Mnsf1 and Ufm1 in higher eukaryotes[32]. Each of these modifiers comes with its own conjugation cascade that activates, conjugates and removes the respective ubiquitin-like modifier[33]. One major difference between modification by ubiquitin and its relatives is, with the exception of SUMO, none form polymeric structures on the substrate, with poly-SUMO chains playing a small role in overall SUMO signaling. For further details we refer to a number of excellent reviews on this subject[33-38].

For the detection of ubiquitination sites the similarity between the C-terminus of ubiquitin and ubiquitin-like molecules is of particular interest. The C-terminus is for all of those modifiers the site of conjugation to other proteins. Here the C-terminal glycine is conjugated via its carboxy-group to the amino group of the lysine side-chain or the N-terminus. Some of the ubiquitin-like modifiers share significant homology at the C-terminal site of conjugation (Figure 2) making the distinction of the target sites extremely complicated by standard proteomic workflows.

Probing the ubiquitin landscape

The following sections will shed light on recent and current techniques used in unraveling the complexity of ubiquitination.

Identifying ubiquitination sites

Ubiquitination site identification was at the center of numerous biological experiments. Initially in order to reveal the precise site of ubiquitination, mutagenesis of the putative target residues was performed thus uncovering the site by an exclusion principle. Typically, lysine residues are mutated into arginine, as it also has a basic side-chain but is non-nucleophilic and therefore cannot be ubiquitinated. This approach assumes no significant biological effect on the cell other than abolished ubiquitination is produced by the mutation. However, changing the amino acid sequence could have an impact on the structure of the molecule. Some ubiquitination reactions have a promiscuous nature and modify any lysine within a certain region, making it almost impossible to identify the primary ubiquitination site with genetic methods. Detection was usually achieved using western blot techniques employing ubiquitin-specific antibodies, with the corresponding limitations in specificity and sensitivity of the antibody and the detection reaction. A bottom-up proteomics approach is advantageous as peptides can be accurately identified and quantified, and the exact ubiquitination site determined with high fidelity. Identification of proteins is unbiased and not limited by the availability of antibodies, resulting in the determination of many more substrates. Furthermore, the organism or cell culture can be kept under physiological conditions for the experimental approach and subtle changes introduced by a specific treatment can be addressed. MS-techniques not only allow for a qualitative determination of the site but also enable researchers to quantify differences in the levels of ubiquitination among variable conditions. The following sections will deal with different proteomic techniques that have been successfully used to identify a plethora of

Probing the ubiquitin landscape

ubiquitination sites in a single experiment or allow the precise quantification of ubiquitin in a certain cellular system.

MS methods

Discovery Proteomics

In the most common ‘bottom-up’ proteomic workflows[39], the specific proteolytic activity of trypsin is utilized, cleaving after lysine and arginine residues to create predictable peptide sequences well-suited for LC-MS/MS analysis. The C-terminus of ubiquitin ends in the amino acid sequence –RGG, therefore treatment with trypsin leaves a glycine dipeptide as an additional stub on the side-chain of modified lysines[40]. Using MS it is possible to identify this additional mass shift in the MS/MS spectrum of the peptide. Modified peptides will preferentially have higher charge states due to the presence of a mid-sequence amine at the end of the diglycine stub, so during the MS measurement the selection of peptides for MS/MS can be restricted to higher charge states, excluding doubly-charged species with little loss[41]. The homology of the C-terminus of ubiquitin and ubiquitin-like proteins creates an issue of ambiguous identifications. The tryptic digest of Nedd8, ISG15 and ubiquitin-modified peptides all generate a glycine dipeptide on the lysine acceptor site, making them indistinguishable by standard MS based methods (Figure 2).

Bottom-up MS strategies have been key for identifying new substrates of ubiquitin and the specific residues that are modified. The review by Mann and Aebersold offers a comprehensive and elegant introduction to the fundamentals of the field[39]. In the ensuing decade, bottom-up MS proteomics has fully emerged as one of the most powerful analytical techniques in protein science, capable of identifying thousands of proteins and post-translational modifications over

Probing the ubiquitin landscape

the course of a few hours. The identification of ubiquitin in samples can be confusing, since ubiquitin is annotated in the databases as fusion proteins with ribosomal subunits (Table 1). To clarify this, the use of a specialized database containing the sequences of mature ubiquitin and mature ribosomal subunits can be helpful.

Identification of primary ubiquitination sites is the initial goal, often followed up by experiments probing how this modification is modulated under certain conditions. Stable isotope-labeling is widely used in quantitative proteomics as a means to compare protein/peptide levels amongst two or more samples. Differential isotopic composition creates no physicochemical difference between the samples, such as chromatographic elution time or ionization potential, but the small mass difference is readily distinguished by the mass spectrometer. Therefore two identical peptides arising from different samples can be combined and measured in a single run and their relative abundance measured. Introducing the label can be achieved through either metabolic or chemical means. A popular metabolic labeling technique called SILAC (stable isotope labeling by amino acids in cell culture) involves growing an organism in media supplemented with isotopically-coded amino acids, mostly lysine and arginine to conform with the trypsin-based bottom-up approach[42]. Exploiting the organism's own metabolic machinery to label proteins minimizes error by reducing sample handling steps. The number of isotope combinations that can be multiplexed together is restricted, with three being the practical maximum to avoid interference between the isotope distributions. Every different label permits parallel comparison to another experimental condition but also adds an additional fold of complexity to the sample; mixing an isotopically "light" Lysine ($^{12}\text{C}_6^{14}\text{N}_2$)/Arginine ($^{12}\text{C}_6^{14}\text{N}_4$) proteome with a "heavy" Lysine ($^{13}\text{C}_6^{15}\text{N}_2$)/Arginine ($^{13}\text{C}_6^{15}\text{N}_4$) proteome effectively doubles the number of unique analytes seen by the mass spectrometer, reducing the possibility of

Probing the ubiquitin landscape

observing low-level analytes such as post-translationally modified peptides due to ionization suppression effects.

Chemical labeling strategies are more flexible as they can be applied to any type of sample, not just cells grown in culture, and most commonly target the modification of protein or peptide primary amine groups. Dimethylation of primary amine groups via reductive methylation with formaldehyde offers a cheap and effective labeling strategy. The various isotopic combinations allow the parallel comparison of three different conditions[43]. In the case of ubiquitin proteomics, fragmentation of a dimethylated GG peptide (methyl groups added to the native N-terminus and the distal end of the GG ubiquitin remnant) yields characteristic diagnostic ions that can be utilized to help eliminate false positives[44,45]. Both SILAC and dimethylation increase sample complexity and compare analyte intensities at the MS1 level. A more sophisticated chemical labeling approach uses reagents such as iTRAQ[46] and TMT[47], which permit a greater number of channels to be monitored simultaneously. Sample complexity is not increased by labeling because of the isobaric nature of the label. This means that differential tags have the same mass, such that differentially labeled peptides derived from different experiments will all have the same MS1 mass, as opposed to being split across multiple signals such as with SILAC or dimethylation, diluting the intensity. Upon MS/MS fragmentation these peptides release unique mass reporter ions; the relative abundance of these reporter ions in the MS/MS spectrum is the basis for quantitation between samples. A recent development in labeling called neutron encoding (NeuCode) combines the metabolic labeling of SILAC with an ability to multiplex many channels, currently up to 18 when used with differential dimethylation[48]. It relies on the small mass defects (6 mDa) that can be generated between different isotopologues of lysine, creating a nominally isobaric tag that can be resolved by a high-resolution FT-MS[48].

Targeted proteomics

When the large-scale identification of proteins is not the objective, but rather the robust measurement of only a subset of proteins, targeted proteomics techniques are used. Selected Reaction Monitoring (SRM) MS is a targeted technique frequently applied to the quantification of the different polyubiquitin chain linkages, for example. While traditional shotgun proteomics aspires to sequence every peptide present in a trypsin-digested proteome, SRM focuses on a pre-selected group of tryptic peptides, detecting them based upon a characteristic conversion of the intact peptide into a fragment ion upon high-energy gas-phase collisions within the mass spectrometer called a transition (Figure 5a). By comparing this signal with that of a stable-isotope (^{13}C , ^{15}N , ^2H) labeled version of the peptide that was spiked into a cell lysate at a known concentration, one can determine the concentration of the peptide in a complex mixture, and even calculate the copy numbers of the protein per cell[49,50]. A recently developed technique, Parallel Reaction Monitoring (PRM) was adapted to the quadrupole-Orbitrap instruments[51] and has also been extended to the quantitative analysis of polyubiquitin chain linkages[52]. The Orbitrap mass analyzer allows parallel monitoring of all transitions while the significant increase in resolution and mass accuracy over triple quadrupole instruments offers improved selectivity in complex matrices.

Pitfalls of ubiquitin site identification

Identifying new ubiquitination sites in a protein poses a challenge on several levels[53]. As with other post-translational modifications, the non-modified peptide usually occurs in large excess compared to the modified peptide. An enrichment step, either of a specific target or a more general enrichment of the modified proteins to reduce the background and boost sensitivity for low-level analytes is often required[41,54]. In the case of ubiquitination the large variety of very

Probing the ubiquitin landscape

active, non-specific DUBs amplifies the challenge, requiring addition of a general inhibitor for such enzymes to prevent reversal of the target PTM. Since most of the DUBs have a cysteine in their active center, treatment with an alkylating chemical agent inactivates these efficiently. A number of compounds have been used, but the most common ones are *N*-ethylmaleimide (NEM) and iodoacetamide (IAA). The use of either compound is associated with additional challenges for the mass spectrometric analysis. NEM can undergo a hydrolyzation step to produce an additional side-product during the alkylation reaction, splitting the signal across two peaks in the respective spectra, as well as modifying amino acids other than cysteine[55]. The reactivity of NEM towards cysteine residues is much lower than IAA, requiring higher concentrations to get a complete inhibition of the DUBs [56]. Off-target alkylation of lysines by IAA can lead to a mass-shift indistinguishable from the 114.043 Da mass shift introduced by Gly-Gly addition [57]. The modification has different chemical properties and can be resolved chromatographically as a peak doublet[57]; it is also not well-recognized by the K-GG peptide antibody[58,59]. The detailed analysis of MS/MS spectra of these chemical artifacts shows a higher tendency for a -57 Da or a -144 Da neutral loss, which can be used as an additional quality criterion[58]. IAA alkylation of lysine is temperature-dependent with modification only occurring at higher incubation temperatures[56], and so can be largely avoided.

Chloroacetamide has successfully been used as an alternative to these two compounds,[57], its' lower reactivity making it more more specific to cysteine thiol-groups. However, in terms of DUB inhibition to best preserve ubiquitination sites, iodoacetamide is more effective than chloroacetamide[56]

Another challenge is posed by the occurrence of false positive interpretations of the spectra by current analytical software like Mascot[60], Sequest[61] or MaxQuant[62]. While the software treats ubiquitination as a normal mass shift of one of the amino acids, some of the spectra are

Probing the ubiquitin landscape

annotated as being modified on a C-terminal lysine. This is a very unlikely modification, since trypsin cannot cleave adjacent to a modified lysine (Figure 3). These peptides are most likely false positives or peptides that carry an asparagine residue (apparent mass shift of 114.043 Da) at the C-terminus[63] and should be discarded as false-positive ubiquitination sites.

Identification of sites in specific proteins

For many biological questions the identification of ubiquitination sites in a specific protein is the primary goal. In order to obtain good identification data the main key is sufficient sequence coverage of proteolytically-generated peptides. As previously mentioned, it is often necessary to enrich for the proteins of interest to overcome the analytical challenge where the PTM represents a small fraction of the total population of a low-abundance protein. For many targets of the ubiquitin system, several ubiquitination sites can be present in a single substrate. These can have different signaling properties, as has been shown for IKK γ [64], or can act simply as alternative sites that trigger the same event[65-67].

Enrichment by ubiquitin pull-down

Ubiquitin-specific antibodies

Antibodies recognizing ubiquitin have been available for a long time[68] and although they were able to recognize ubiquitin, cross-reactivity was high. The development of monoclonal antibodies recognizing polyubiquitin but not monoubiquitin facilitated the targeted analysis of polyubiquitinated material[69]. One of these monoclonal antibodies (FK2) was selected because of its high specificity toward polyubiquitinated proteins[70] and was first introduced in proteomic studies of the ubiquitinome[71] (Figure 4C). This study identified 670 proteins that were enriched under highly denaturing conditions, but lacked the detection of GG-peptides

Probing the ubiquitin landscape

that would boost confidence to make them *bona fide* identifications. The same strategy was later used to characterize the EGF-induced response of the ubiquitin network[72]. A detailed comparison of different antibody-based ubiquitinome analyses found that the use of the FK2 antibody yielded similar results as the use of TUBE-enrichment strategies[73] (see below).

Using epitope-tagged ubiquitin

With the continuous development and refinement of various molecular biology tools, the study of ubiquitination has been pushed forward in the past two decades. The preparation of ubiquitin-specific antibodies has been hampered by low affinity and high background, so the introduction of N-terminally epitope-tagged ubiquitin as a probe was quickly identified as a useful alternative for ubiquitination studies[74]. Tagged ubiquitin allows for a more robust purification[75], as it can be performed under strongly denaturing conditions[76]. Cells transformed with tagged-ubiquitin formed the basis for several studies in different organisms that have been undertaken since, with the first large-scale study performed in yeast[54]. All four copies of the ubiquitin gene were replaced by a ubiquitin gene carrying a 6-His tag at the N-terminus. The His-tag was used to enrich for ubiquitinated material via Ni^{2+} -chelate chromatography, that was subsequently digested by trypsin (Figure 4A) and peptides identified using MS. In total this led to the identification of 110 proteins with a ubiquitination site. Although the purification was done under denaturing conditions, a significant number of proteins identified had no diglycine-modified peptide detected (Figure 4). Experiments in yeast have shown that the N-terminus of ubiquitin can also be modified by different epitope tags. These fusion-proteins can still support growth where all additional copies of wild-type ubiquitin have been removed. It is even possible to transfer an N-terminal GST-ubiquitin-fusion, a 26 kDa epitope tag, to a substrate (MHC Class I Heavy Chain) by the ubiquitin conjugation

Probing the ubiquitin landscape

system[77]. Alternative epitope-tags include myc-, Flag, HA-, His-FLAG and biotin-6His double-tag, expressed either ectopically[78-81] or stably integrated [82].

In a different approach, a lysine-free ubiquitin mutant was created for the analysis. Here all seven lysines of ubiquitin were mutated to arginines, creating a ubiquitin variant that can be conjugated to a target but is resistant to digestion by a lysine-specific protease. Following such a digest, the intact mutant ubiquitins conjugated to remnant of substrates can be isolated by size-selection. This led to the identification of 1392 ubiquitination sites in human cells[83].

bio-Ubiquitin

The interaction of biotin with avidin or streptavidin is one of the strongest non-covalent interactions in nature. The bacterial BirA ligase recognizes a specific sequence (avi-tag) and covalently adds a biotin moiety[84]. By fusing the avi-tag to the N-terminus of ubiquitin with the simultaneous expression of the birA-ligase, ubiquitin can be biotinylated *in vivo*, facilitating the purification of ubiquitinated proteins under stringent wash conditions (Figure 4B). Recently the Mayor laboratory constructed a vector allowing expression of a penta-ubiquitin-birA-ligase fusion (hexa-ubiquitin in the transgenic mouse). Like wild-type ubiquitin, this version of ubiquitin is processed co-translationally, releasing the birA-ubiquitin and the BirA-ligase. The birA-ubiquitin is then incorporated into the ubiquitin-chains allowing the pull-down of biotin-ubiquitinated substrates using streptavidin or avidin-beads. Coupled to mass-spectrometric analysis, this study identified several hundred ubiquitinated substrates[85-87].

Ubiquitin site identification using peptide-specific antibodies

Rather than enrichment at the protein level, a peptide-based technique has been developed exploiting the characteristic branch structure of ubiquitinated peptides after a tryptic digest.

Probing the ubiquitin landscape

Specific antibodies that recognize lysine residues carrying a glycine dipeptide on the ϵ -amino group have been raised and allow the enrichment for ubiquitinated peptides from complex proteomic samples[81,88] (Figure 4E). This technique was used in three landmark papers from the Gygi, Mann and Elledge laboratories to identify the largest set, to date, of ubiquitin substrates[41,89,90]. Although the identification rate was initially in the range of 300 to 800 ubiquitination sites[81,88] employment of peptide fractionation, such as isoelectric focusing or strong cation exchange chromatography, improved the identification rates to 10000 – 20000 ubiquitination sites in one study[41,91-93]. Using β -interferon stimulation to stimulate ISG15 expression as a positive control, Kim *et al.* estimated the contribution of ISGylation (Figure 2) of substrates to the total number of di-glycine modified peptides, and concluded that a negligible portion of the total modified peptide population can be attributed to ISG15 modification[89]. In the same study, a general DUB was used to cleave all ubiquitin moieties from their substrates leaving only the sites modified by Nedd8. Once again the group concluded that the fraction of proteins modified by Nedd8 was minor[89].

Probing the Ubiquitin chain topology

Chain-specific antibodies

In 2008 new ubiquitin chain-specific antibodies have been introduced, allowing for the detection of specific chain-types in western-blot[94]. Although these have been successfully used for the detection of specific ubiquitin chains in western blot analysis, the antibody's performance in large-scale proteomic studies has been mediocre, although it was used successfully for the optimization of the ubiquitin chain quantification by SRM[95]. This is probably due to the relatively low affinity of these antibodies and their high background in pull-down experiments.

Probing the ubiquitin landscape

Ubiquitin chain quantifications

Ubiquitin chain quantification represents a critical complement to primary ubiquitin site identifications. Each linkage has a unique structure, with particular linkages associated with particular functions. Knowledge of the character of a polyubiquitin chain is essential to understanding the purpose of an ubiquitination event on a specific target protein as well as exploring how the global linkage landscape is altered by perturbing the system.

Assessing linkage types can be challenging by traditional molecular biology methods, but represents an excellent application of MS-based proteomics. The method takes advantage of the generation of unique peptides during the tryptic digest of polyubiquitin chains. Depending on the type of chain present in the sample, the diglycine moiety can sit on any of the seven lysine residues of ubiquitin. Each of these peptides is proteotypic and can be quantified as a proxy for frequency of ubiquitin molecules incorporated into a certain chain type. By using heavy-labeled reference peptides it is possible to use absolute quantification for the type and number of chains in the samples.

Depending on the experimental question being addressed, analysis may be applied to whole cell lysates to study global effects[6,52,96,97], though protein purification or enrichment is necessary to study the linkage types present in an isolated system[21,98]. Care must be taken in interpreting the results as the background for an IP can easily number in the hundreds of proteins. A drawback of a bottom-up proteomics approach to polyubiquitin is the loss of linkage context that occurs during the digestion process, especially in complex biological systems. Once reduced to its component parts, it is impossible to know whether signals for multiple linkage types arose from a single polyubiquitin chain on a single target or from multiple sources. Examples in the literature point to the formation of branched chains (K48, K63 or K48,

Probing the ubiquitin landscape

K11) [22,99]. Unfortunately, the evidence for these branched chain formation is destroyed by a complete enzymatic digestion.

Tandem Ubiquitin Binding Entity (TUBE) based enrichment

The ubiquitin-associated (UBA) domain is a ubiquitin binding domain that is present in a number of different proteins of the ubiquitin pathway[100]. Prominent examples are the proteasomal shuttling factors Rad23 and Dsk2, which contain a ubiquitin-like domain and one or two UBA-domains [101-103]. Other proteins containing UBA domains are E3 ligases, deubiquitinating enzymes and E2s. These domains are well-conserved across evolution and consist of a sequence of 40-50 amino acids forming an α -helical structure. Depending on the protein, several of these domains can be arranged in multimers. The specificity of the domain can vary with preferences for mono-ubiquitination or towards homotypic chains like K48, K63 or linear[104-108]. A large study by Raasi and coworkers pinpointed the specificity of 30 of these domains[109]. Although the ubiquitin-chain is recognized by Rad23 and transported to the proteasome, the chain appears to be protected and disassembly and degradation of the substrate is delayed[110].

As UBA domains have high affinity for polyubiquitin chains, they have been used as tools for enriching polyubiquitinated material[111]. By combining two UBA domains the affinity of the probe can be increased, leading to the TUBE probes, also referred as ubiquitin-traps. During the purification of the ubiquitin chains the chain itself is protected against rapid disassembly by DUBs. By combining several UBA domains in one TUBE the specificity of the chain-enrichment can be modulated[112,113]. The technique has been applied to study specific E3 ligase substrates *in vivo* by overexpressing them together with a FLAG-tagged, trypsin-resistant TUBE (TR-TUBE)[114]. An interesting alternative for monitoring ubiquitin chains *in vivo* has

Probing the ubiquitin landscape

been developed by combining the TUBE-probe with fluorescent probes, allowing imaging of ubiquitin chains in certain parts of the cell[115]. In their recent study, Yoshida *et al.* [114] aimed to identify substrates of the ubiquitin ligase FBXO21. To achieve this, the E3 was overexpressed followed by enrichment using FLAG-tagged TR-TUBES, which - according to the authors - protected against deubiquitination. The authors also mentioned that trypsin can be hindered in accessing the ubiquitin chains in the presence of the TUBES, which was overcome by completely denaturing the sample. After tryptic digest another level of enrichment was achieved by using α -GlyGly-Lys antibodies to pull-out ubiquitinated peptides, which were identified by MS. This identified eight ubiquitination sites in substrates of Skp2 and another six in substrates of FBXO21[114].

Their specificity with respect to ubiquitin binding, however, has been challenged by several studies showing that UBA domains interact with UBL domains and other proteins. Some pull-down experiments using UBA domains found unspecific recovery of ubiquitin chains, likely due to interactions of multiple UBAs that overwhelmed the specificity of a single UBA[112,116-118].

UbiCRest

The Ubiquitin Chain Restriction (UbiCRest) assay (Figure 4F) introduced by the Komander group provides a means of systematically probing a polyubiquitin chain with a toolbox of DUBs of known specificity and using gel-based analysis to determine the linkage varieties present as well as the architecture of a heterotypic polyubiquitin chain[119]. Assumptions are made about the activity of the DUBs, the specificity of which was largely tested using di-ubiquitin, and which may have altered specificities under different concentrations and incubation times or may utilize additional binding domains or interaction partners. In addition, the enrichment

Probing the ubiquitin landscape

strategy used to isolate a polyubiquitinated substrate may induce changes in the structure that result in resistance to proteolysis by DUBs.

Post-translational modifications on ubiquitin

Although ubiquitin is a post-translational modifier it can itself be the target of post-translational modifications. Phosphorylation on ubiquitin, at residue S57, was initially reported in the first large-scale study in yeast[54]. A number of other large-scale studies extended the list of modification to phosphorylations on serines, threonine and tyrosines [120-122], acetylation on all internal lysines[123] and the modification of lysine 11 by SUMO[124], another ubiquitin-like modification. Serine 65 has been shown to be phosphorylated by the PINK1 kinase[125], and is induced after mitochondrial depolarization[126]. The inclusion of phosphorylated ubiquitin into the ubiquitin chain alters the structure, generating a different signal[125]. This can change the polymerization of the chain as well as the overall amount of ubiquitination in the cell, while the overall degradation rates decrease[127].

Top-down based analysis

Recent approaches using limited trypsination and so-called middle-down proteomics exploit the fact that non-denatured ubiquitin has only a single cleavage site exposed, at R74, generating longer stretches of the ubiquitin chain with the branching points preserved [128,129].

While middle-down proteomics approaches have had some success in characterizing polyubiquitin chains with respect to their length and architecture[128,129], the field of top-down proteomics holds significant promise for characterizing the combinatorial modifications occurring at the intact protein level, comprised of co-occurring PTMs that combine to generate

Probing the ubiquitin landscape

a particular functional state of a given protein. The top-down methodology skips the protease digestion step that reduces whole proteins to peptides and instead seeks to measure what has been termed 'proteoforms', defined by Kelleher *et al.* [130] as "the specific molecular form of the protein resulting from combinations of genetic variation, alternative splicing, and post-translational modifications." PTMs such as phosphorylation and ubiquitination can have a significant impact on a protein's function, and bottom-up proteomics have augmented protein databases with the annotation of tens of thousands of modified sites. PTM analysis via top-down MS gives researchers the opportunity to learn what PTMs are occurring simultaneously over the entire length of the polypeptide chain and what the additive/combinatorial effect of multiple PTMs occurring in parallel or the temporal relations between them.

Native MS is an even purer form of top-down analysis, whereby non-denaturing conditions are used to isolate the protein analytes, preserving non-covalent protein-protein interactions and permitting study of homo- or heteromeric complexes. A tremendous technical benefit of native MS is that folded proteins have a lower distribution of charge states, owing to the reduced surface exposure of ionizable groups. This boosts the sensitivity for a signal that would be otherwise spread across many more channels, such as with ubiquitin where in the native state only three charge states exist, compared to eight for the unfolded molecule.

Outlook

A full understanding of the ubiquitination status of any particular target protein needs to take into account not only the site of modification but also the structure of the attached polyubiquitin chain. Several methods might lead to this final goal. For small numbers of proteins modern structure determination techniques such as crystallography or NMR might be applicable, though the usual problems of gathering enough material of adequate purity will

Probing the ubiquitin landscape

present the usual stumbling blocks. The rapidly developing field of top-down MS might lead the charge, which gains momentum as benchtop instrumentation increases in capability through ever-climbing resolution and sensitivity, and the enrichment and separation strategies become more refined.

For example, the multiple forms of a protein of several hundred kDa with a conjugated tetraubiquitin can be readily resolved, and one might use a toolbox of specific DUBs[119] to probe the connectivity present in a stepwise manner. This represents a true frontier of ubiquitin proteomics. With the preservation of the rich, detailed information intrinsic to these polymeric structures, we can begin to understand more precisely how a target protein is modified and for what downstream purpose.

Aside from MS, new methods for the analysis of this important post-translational modification are appearing on the horizon. The recent report from the Meller laboratory describes the use of a nano-pore for the detection of differentially-linked ubiquitin chains[131].

With the number of identified ubiquitination sites increasing rapidly, questions about the details in signaling can be addressed in depth. Bioinformatic studies are already using the extensive data collected in the large scale studies[132,133]. New studies are now starting to combine the information of ubiquitination with other post-translational modifications like phosphorylation in order to construct more complex networks[134]. New methods for multiplexing samples like the NeuCode technique that is based on differences in the nuclear binding energy allow higher degrees of multiplexing while not increasing the complexity of the sample[48,135]. This will not only allow larger ubiquitination based data sets, but also allows for multiplexing for top-down applications as well[136].

Acknowledgements

We thank Daniel Finley, Ernst Jarosch and Ugo Mayor for reading the manuscript, commenting and the fruitful discussions. PB has been supported by the virtual Helmholtz institute HICE.

References

- [1] Crick, F., Central dogma of molecular biology. *Nature* 1970, 227, 561–563.
- [2] Schoenheimer, R., Rittenberg, D., DEUTERIUM AS AN INDICATOR IN THE STUDY OF INTERMEDIARY METABOLISM. *Science* 1935, 82, 156–157.
- [3] Schoenheimer, R., Ratner, S., Rittenberg, D., Studies in Protein Metabolism VII. The Metabolism of Tyrosine. *Journal of Biological Chemistry* 1939.
- [4] Hoeller, D., Hecker, C.-M., Dikic, I., Ubiquitin and ubiquitin-like proteins in cancer pathogenesis. *Nat Rev Cancer* 2006, 6, 776–788.
- [5] Drews, O., Taegtmeyer, H., Targeting the ubiquitin-proteasome system in heart disease: the basis for new therapeutic strategies. *Antioxid. Redox Signal.* 2014, 21, 2322–2343.
- [6] Dammer, E.B., Na, C.H., Xu, P., Seyfried, N.T., et al., Polyubiquitin linkage profiles in three models of proteolytic stress suggest the etiology of Alzheimer disease. *J Biol Chem* 2011, 286, 10457–10465.
- [7] Bennett, E.J., Shaler, T.A., Woodman, B., Ryu, K.-Y., et al., Global changes to the ubiquitin system in Huntington's disease. *Nature* 2007, 448, 704–708.
- [8] Ciechanover, A., Brundin, P., The ubiquitin proteasome system in neurodegenerative diseases: sometimes the chicken, sometimes the egg. *Neuron* 2003, 40, 427–446.
- [9] Kulathu, Y., Komander, D., Atypical ubiquitylation - the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. *Nat Rev Mol Cell Biol* 2012, 13, 508–523.
- [10] Ozkaynak, E., Finley, D., Solomon, M.J., Varshavsky, A., The yeast ubiquitin genes: a family of natural gene fusions. *EMBO J* 1987, 6, 1429–1439.
- [11] Dittmar, G., Sommer, T., *Ubiquitination*, Springer, 2006.
- [12] Reyes-Turcu, F.E., Ventii, K.H., Wilkinson, K.D., Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annu. Rev. Biochem.* 2009, 78, 363–397.
- [13] Hoege, C., Pfander, B., Moldovan, G.-L., Pyrowolakis, G., Jentsch, S., RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* 2002, 419, 135–141.
- [14] Terrell, J., Shih, S., Dunn, R., Hicke, L., A function for monoubiquitination in the internalization of a G protein-coupled receptor. *Molecular Cell* 1998, 1, 193–202.
- [15] Shih, S.C., Katzmann, D.J., Schnell, J.D., Sutanto, M., et al., Epsins and Vps27p/Hrs

Probing the ubiquitin landscape

- contain ubiquitin-binding domains that function in receptor endocytosis. *Nat Cell Biol* 2002, 4, 389–393.
- [16] Ciehanover, A., Hod, Y., Hershko, A., A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochem Biophys Res Commun* 1978, 81, 1100–1105.
- [17] Tokunaga, F., Sakata, S.-I., Saeki, Y., Satomi, Y., et al., Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation. *Nat Cell Biol* 2009, 11, 123–132.
- [18] Iwai, K., Tokunaga, F., Linear polyubiquitination: a new regulator of NF-kappaB activation. *EMBO Rep* 2009, 10, 706–713.
- [19] Ikeda, F., Deribe, Y.L., Skånland, S.S., Stieglitz, B., et al., SHARPIN forms a linear ubiquitin ligase complex regulating NF-κB activity and apoptosis. *Nature* 2011, 471, 637–641.
- [20] Gerlach, B., Cordier, S.M., Schmukle, A.C., Emmerich, C.H., et al., Linear ubiquitination prevents inflammation and regulates immune signalling. *Nature* 2011, 471, 591–596.
- [21] Müller-Rischart, A.K., Pilsl, A., Beaudette, P., Patra, M., et al., The E3 Ligase Parkin Maintains Mitochondrial Integrity by Increasing Linear Ubiquitination of NEMO. *Molecular Cell* 2013, 49, 908–921.
- [22] Meyer, H.-J., Rape, M., Enhanced protein degradation by branched ubiquitin chains. *Cell* 2014, 157, 910–921.
- [23] Breitschopf, K., A novel site for ubiquitination: the N-terminal residue, and not internal lysines of MyoD, is essential for conjugation and degradation of the protein. *EMBO J* 1998, 17, 5964–5973.
- [24] Cadwell, K., Ubiquitination on Nonlysine Residues by a Viral E3 Ubiquitin Ligase. *Science* 2005, 309, 127–130.
- [25] Williams, C., van den Berg, M., Sprenger, R.R., Distel, B., A conserved cysteine is essential for Pex4p-dependent ubiquitination of the peroxisomal import receptor Pex5p. *J Biol Chem* 2007, 282, 22534–22543.
- [26] Hensel, A., Beck, S., Magraoui, El, F., Platta, H.W., et al., Cysteine-dependent ubiquitination of Pex18p is linked to cargo translocation across the peroxisomal membrane. *J Biol Chem* 2011, 286, 43495–43505.
- [27] Wang, X., Herr, R.A., Hansen, T.H., Ubiquitination of Substrates by Esterification. *Traffic* 2012, 13, 19–24.
- [28] Ravid, T., Hochstrasser, M., Autoregulation of an E2 enzyme by ubiquitin-chain assembly on its catalytic residue. *Nat Cell Biol* 2007, 9, 422–427.
- [29] Wang, X., Herr, R.A., Chua, W.J., Lybarger, L., et al., Ubiquitination of serine, threonine, or lysine residues on the cytoplasmic tail can induce ERAD of MHC-I by viral E3 ligase mK3. *J Cell Biol* 2007, 177, 613–624.
- [30] Shimizu, Y., Okuda-Shimizu, Y., Hendershot, L.M., Ubiquitylation of an ERAD substrate occurs on multiple types of amino acids. *Molecular Cell* 2010, 40, 917–926.
- [31] Boban, M., Ljungdahl, P.O., Foisner, R., Atypical ubiquitylation in yeast targets lysine-less Asi2 for proteasomal degradation. *J Biol Chem* 2015, 290, 2489–2495.
- [32] Jentsch, S., Pyrowolakis, G., Ubiquitin and its kin: how close are the family ties? *Trends Cell Biol* 2000, 10, 335–342.
- [33] Hochstrasser, M., Origin and function of ubiquitin-like proteins. *Nature* 2009, 458, 422–429.
- [34] Enchev, R.I., Schulman, B.A., Peter, M., Protein neddylation: beyond cullin-RING ligases. *Nat Rev Mol Cell Biol* 2014, 16, 30–44.

Probing the ubiquitin landscape

- [35] Vertegaal, A.C.O., Uncovering ubiquitin and ubiquitin-like signaling networks. *Chem Rev* 2011, 111, 7923–7940.
- [36] Daniel, J., Liebau, E., The ufm1 cascade. *Cells* 2014, 3, 627–638.
- [37] Streich, F.C., Lima, C.D., Structural and functional insights to ubiquitin-like protein conjugation. *Annu Rev Biophys* 2014, 43, 357–379.
- [38] Schwartz, D.C., Hochstrasser, M., A superfamily of protein tags: ubiquitin, SUMO and related modifiers. *Trends Biochem Sci* 2003, 28, 321–328.
- [39] Aebersold, R., Mann, M., Mass spectrometry-based proteomics. *Nature* 2003, 422, 198–207.
- [40] Peng, J., Gygi, S.P., Proteomics: the move to mixtures. *J Mass Spectrom* 2001, 36, 1083–1091.
- [41] Wagner, S.A., Beli, P., Weinert, B.T., Nielsen, M.L., et al., A Proteome-wide, Quantitative Survey of In Vivo Ubiquitylation Sites Reveals Widespread Regulatory Roles. *Mol Cell Proteomics* 2011, 10, M111.013284.
- [42] Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., 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.
- [43] Hsu, J.-L., Huang, S.-Y., Chow, N.-H., Chen, S.-H., Stable-isotope dimethyl labeling for quantitative proteomics. *Anal Chem* 2003, 75, 6843–6852.
- [44] Chicooree, N., Connolly, Y., Tan, C.-T., Malliri, A., et al., Enhanced detection of ubiquitin isopeptides using reductive methylation. *J Am Soc Mass Spectrom* 2013, 24, 421–430.
- [45] Griffiths, J.R., Chicooree, N., Connolly, Y., Neffling, M., et al., Mass spectral enhanced detection of UbIs using SWATH acquisition: MEDUSA--simultaneous quantification of SUMO and ubiquitin-derived isopeptides. *J Am Soc Mass Spectrom* 2014, 25, 767–777.
- [46] Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., et al., Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* 2003, 75, 1895–1904.
- [47] Dayon, L., Hainard, A., Licker, V., Turck, N., et al., Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal Chem* 2008, 80, 2921–2931.
- [48] Hebert, A.S., Merrill, A.E., Bailey, D.J., Still, A.J., et al., Neutron-encoded mass signatures for multiplexed proteome quantification. *Nat Methods* 2013, 10, 332–334.
- [49] Picotti, P., Bodenmiller, B., Mueller, L.N., Domon, B., Aebersold, R., Full Dynamic Range Proteome Analysis of *S. cerevisiae* by Targeted Proteomics. *Cell* 2009, 138, 795–806.
- [50] Picotti, P., Lam, H., Campbell, D., Deutsch, E.W., et al., A database of mass spectrometric assays for the yeast proteome. *Nat Methods* 2008, 5, 913–914.
- [51] Michalski, A., Damoc, E., Hauschild, J.-P., Lange, O., et al., Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics* 2011, 10, M111.011015.
- [52] Tsuchiya, H., Tanaka, K., Saeki, Y., The parallel reaction monitoring method contributes to a highly sensitive polyubiquitin chain quantification. *Biochem Biophys Res Commun* 2013, 436, 223–229.
- [53] Seyfried, N.T., Xu, P., Duong, D.M., Cheng, D., et al., Systematic approach for validating the ubiquitinated proteome. *Anal Chem* 2008, 80, 4161–4169.
- [54] Peng, J., Schwartz, D., Elias, J.E., Thoreen, C.C., et al., A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* 2003, 21, 921–926.

Probing the ubiquitin landscape

- [55] Brewer, C., Riehm, J., Evidence for possible nonspecific reactions between N-ethylmaleimide and proteins* 1. *Anal. Biochem.* 1967.
- [56] Xu, P., Duong, D.M., Seyfried, N.T., Cheng, D., et al., Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell* 2009, 137, 133–145.
- [57] Nielsen, M.L., Vermeulen, M., Bonaldi, T., Cox, J., et al., Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nat Methods* 2008, 5, 459–460.
- [58] Bustos, D., Bakalarski, C.E., Yang, Y., Peng, J., Kirkpatrick, D.S., Characterizing ubiquitination sites by peptide-based immunoaffinity enrichment. *Mol Cell Proteomics* 2012, 11, 1529–1540.
- [59] Na, C.H., Jones, D.R., Yang, Y., Wang, X., et al., Synaptic protein ubiquitination in rat brain revealed by antibody-based ubiquitome analysis. *J Proteome Res* 2012, 11, 4722–4732.
- [60] Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999, 20, 3551–3567.
- [61] Eng, J., McCormack, A., Yates, J., An approach to correlate tandem mass spectral data of peptides with amino acid ... *Society for Mass Spectrometry* 1994.
- [62] Cox, J.U.R., Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008, 26, 1367–1372.
- [63] Shi, Y., Xu, P., Qin, J., Ubiquitinated proteome: ready for global? *Mol Cell Proteomics* 2011, 10, R110.006882.
- [64] Hinz, M., Stilmann, M., Arslan, S.C., Khanna, K.K., et al., A cytoplasmic ATM-TRAF6-clAP1 module links nuclear DNA damage signaling to ubiquitin-mediated NF- κ B activation. *Molecular Cell* 2010, 40, 63–74.
- [65] Rodriguez, M.S., Desterro, J.M.P., Lain, S., Lane, D.P., Hay, R.T., Multiple C-Terminal Lysine Residues Target p53 for Ubiquitin-Proteasome-Mediated Degradation. *Mol Cell Biol* 2000, 20, 8458–8467.
- [66] Nakamura, S., Roth, J.A., Mukhopadhyay, T., Multiple Lysine Mutations in the C-Terminal Domain of p53 Interfere with MDM2-Dependent Protein Degradation and Ubiquitination. *Mol Cell Biol* 2000, 20, 9391–9398.
- [67] Haglund, K., Sigismund, S., Polo, S., Szymkiewicz, I., et al., Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat Cell Biol* 2003, 5, 461–466.
- [68] Hershko, A., Eytan, E., Ciechanover, A., Haas, A.L., Immunochemical analysis of the turnover of ubiquitin-protein conjugates in intact cells. Relationship to the breakdown of abnormal proteins. *J Biol Chem* 1982, 257, 13964–13970.
- [69] Fujimuro, M., Sawada, H., Yokosawa, H., Production and characterization of monoclonal antibodies specific to multi-ubiquitin chains of polyubiquitinated proteins. *FEBS Lett* 1994, 349, 173–180.
- [70] Takada, K., Nasu, H., Hibi, N., Tsukada, Y., et al., Immunoassay for the quantification of intracellular multi-ubiquitin chains. *Eur. J. Biochem.* 1995, 233, 42–47.
- [71] Matsumoto, M., Hatakeyama, S., Oyamada, K., Oda, Y., et al., Large-scale analysis of the human ubiquitin-related proteome. *Proteomics* 2005, 5, 4145–4151.
- [72] Argenzio, E., Bange, T., Oldrini, B., Bianchi, F., et al., Proteomic snapshot of the EGF-induced ubiquitin network. *Mol Syst Biol* 2011, 7, 462–462.
- [73] Schwertman, P., Bezstarosti, K., Laffeber, C., Vermeulen, W., et al., An immunoaffinity purification method for the proteomic analysis of ubiquitinated protein

Probing the ubiquitin landscape

- complexes. *Anal. Biochem.* 2013, 440, 227–236.
- [74] Ellison, M.J., Hochstrasser, M., Epitope-tagged ubiquitin. A new probe for analyzing ubiquitin function. *J Biol Chem* 1991, 266, 21150–21157.
- [75] Spence, J., Sadis, S., Haas, A.L., Finley, D., A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. *Mol Cell Biol* 1995, 15, 1265–1273.
- [76] Spence, J., Gali, R.R., Dittmar, G., Sherman, F., et al., Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* 2000, 102, 67–76.
- [77] Flierman, D., Ye, Y., Dai, M., Chau, V., Rapoport, T.A., Polyubiquitin Serves as a Recognition Signal, Rather than a Ratcheting Molecule, during Retrotranslocation of Proteins across the Endoplasmic Reticulum Membrane. *J Biol Chem* 2003, 278, 34774–34782.
- [78] Fang, N.N., Ng, A.H.M., Measday, V., Mayor, T., Hul5 HECT ubiquitin ligase plays a major role in the ubiquitylation and turnover of cytosolic misfolded proteins. *Nat Cell Biol* 2011.
- [79] Tagwerker, C., Flick, K., Cui, M., Guerrero, C., et al., A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivocross-linking. *Mol Cell Proteomics* 2006, 5, 737–748.
- [80] Kim, J.Y., Anderson, E.D., Huynh, W., Dey, A., Ozato, K., Proteomic Survey of Ubiquitin-Linked Nuclear Proteins in Interferon-Stimulated Macrophages. *Journal of Interferon & Cytokine Research* 2011, 31, 619–628.
- [81] Danielsen, J.M.R., Sylvestersen, K.B., Bekker-Jensen, S., Szklarczyk, D., et al., Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level. *Mol Cell Proteomics* 2011, 10, M110.003590–M110.003590.
- [82] Akimov, V., Henningsen, J., Hallenborg, P., Rigbolt, K.T.G., et al., StUbEx: Stable tagged ubiquitin exchange system for the global investigation of cellular ubiquitination. *J Proteome Res* 2014, 13, 4192–4204.
- [83] Oshikawa, K., Matsumoto, M., Oyamada, K., Nakayama, K.I., Proteome-wide identification of ubiquitylation sites by conjugation of engineered lysine-less ubiquitin. *J Proteome Res* 2012, 11, 796–807.
- [84] Beckett, D., Kovaleva, E., Schatz, P.J., A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Science* 1999.
- [85] Franco, M., Seyfried, N.T., Brand, A.H., Peng, J., Mayor, U., A novel strategy to isolate ubiquitin conjugates reveals wide role for ubiquitination during neural development. *Mol Cell Proteomics* 2011, 10, M110.002188.
- [86] Lectez, B., Migotti, R., Lee, S.Y., Ramirez, J., et al., Ubiquitin profiling in liver using a transgenic mouse with biotinylated ubiquitin. *J Proteome Res* 2014.
- [87] Min, M., Mayor, U., Dittmar, G., Lindon, C., Using in vivo-biotinylated ubiquitin to describe a mitotic exit ubiquitome from human cells. *Mol Cell Proteomics* 2014.
- [88] Xu, G., Paige, J.S., Jaffrey, S.R., Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat Biotechnol* 2010, 28, 868–873.
- [89] Kim, W., Bennett, E.J., Huttlin, E.L., Guo, A., et al., Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome. *Molecular Cell* 2011, 44, 325–340.
- [90] Emanuele, M.J., Elia, A.E.H., Xu, Q., Thoma, C.R., et al., Global identification of modular cullin-RING ligase substrates. *Cell* 2011, 147, 459–474.
- [91] Kim, G.-W., Yang, X.-J., Comprehensive lysine acetylomes emerging from bacteria to humans. *Trends Biochem Sci* 2011, 36, 211–220.
- [92] Udeshi, N.D., Mani, D.R., Eisenhaure, T., Mertins, P., et al., Methods for quantification of in vivo changes in protein ubiquitination following proteasome and

Probing the ubiquitin landscape

- deubiquitinase inhibition. *Mol Cell Proteomics* 2012.
- [93] Udeshi, N.D., Svinkina, T., Mertins, P., Kuhn, E., et al., Refined preparation and use of anti-diglycine remnant (K- ϵ -GG) antibody enables routine quantification of 10,000s of ubiquitination sites in single proteomics experiments. *Mol Cell Proteomics* 2013, 12, 825–831.
- [94] Newton, K., Matsumoto, M.L., Wertz, I.E., Kirkpatrick, D.S., et al., Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. *Cell* 2008, 134, 668–678.
- [95] Phu, L., Izrael-Tomasevic, A., Matsumoto, M.L., Bustos, D., et al., Improved Quantitative Mass Spectrometry Methods for Characterizing Complex Ubiquitin Signals. *Mol Cell Proteomics* 2011, 10, M110.003756–M110.003756.
- [96] Mirzaei, H., Rogers, R.S., Grimes, B., Eng, J., et al., Characterizing the connectivity of poly-ubiquitin chains by selected reaction monitoring mass spectrometry. *Mol. BioSyst.* 2010, 6, 2004–2014.
- [97] Ziv, I., Matiuhin, Y., Kirkpatrick, D.S., Erpapazoglou, Z., et al., A perturbed ubiquitin landscape distinguishes between ubiquitin in trafficking and in proteolysis. *Mol Cell Proteomics* 2011, 10, M111.009753.
- [98] Bagola, K., Delbrück, von, M., Dittmar, G., Scheffner, M., et al., Ubiquitin binding by a CUE domain regulates ubiquitin chain formation by ERAD E3 ligases. *Molecular Cell* 2013, 50, 528–539.
- [99] Nakasone, M.A., Livnat-Levanon, N., Glickman, M.H., Cohen, R.E., Fushman, D., Mixed-linkage ubiquitin chains send mixed messages. *Structure* 2013, 21, 727–740.
- [100] Hofmann, K., Bucher, P., The UBA domain: a sequence motif present in multiple enzyme classes of the ubiquitination pathway. *Trends Biochem Sci* 1996, 21, 172–173.
- [101] Schaubert, C., Chen, L., Tongaonkar, P., Vega, I., et al., Rad23 links DNA repair to the ubiquitin/proteasome pathway. *Nature* 1998, 391, 715–718.
- [102] Chen, L., Madura, K., Rad23 promotes the targeting of proteolytic substrates to the proteasome. *Mol Cell Biol* 2002, 22, 4902–4913.
- [103] Elsasser, S., Gali, R.R., Schwickart, M., Larsen, C.N., et al., Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat Cell Biol* 2002, 4, 725–730.
- [104] Watkins, J.F., Sung, P., Prakash, L., Prakash, S., The *Saccharomyces cerevisiae* DNA repair gene RAD23 encodes a nuclear protein containing a ubiquitin-like domain required for biological function. *Mol Cell Biol* 1993, 13, 7757–7765.
- [105] Bertolaet, B.L., Clarke, D.J., Wolff, M., Watson, M.H., et al., UBA domains of DNA damage-inducible proteins interact with ubiquitin. *Nat. Struct. Biol.* 2001, 8, 417–422.
- [106] Wilkinson, C.R., Seeger, M., Hartmann-Petersen, R., Stone, M., et al., Proteins containing the UBA domain are able to bind to multi-ubiquitin chains. *Nat Cell Biol* 2001, 3, 939–943.
- [107] Zhang, D., Raasi, S., Fushman, D., Affinity makes the difference: nonselective interaction of the UBA domain of Ubiquilin-1 with monomeric ubiquitin and polyubiquitin chains. *Journal of Molecular Biology* 2008, 377, 162–180.
- [108] Rahighi, S., Ikeda, F., Kawasaki, M., Akutsu, M., et al., Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. *Cell* 2009, 136, 1098–1109.
- [109] Raasi, S., Varadan, R., Fushman, D., Pickart, C.M., Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nat Struct Mol Biol* 2005, 12, 708–714.
- [110] Raasi, S., Pickart, C.M., Rad23 ubiquitin-associated domains (UBA) inhibit 26 S proteasome-catalyzed proteolysis by sequestering lysine 48-linked polyubiquitin

Probing the ubiquitin landscape

- chains. *J Biol Chem* 2003, 278, 8951–8959.
- [111] Shi, Y., Chan, D.W., Jung, S.Y., Malovannaya, A., et al., A data set of human endogenous protein ubiquitination sites. *Mol Cell Proteomics* 2011, 10, M110.002089.
- [112] Hjerpe, R., Rodríguez, M.S., Efficient approaches for characterizing ubiquitinated proteins. *Biochem Soc Trans* 2008, 36, 823–827.
- [113] Hjerpe, R., Aillet, F., Lopitz-Otsoa, F., Lang, V., et al., Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *EMBO Rep* 2009, 10, 1250–1258.
- [114] Yoshida, Y., Saeki, Y., Murakami, A., Kawawaki, J., et al., A comprehensive method for detecting ubiquitinated substrates using TR-TUBE. *Proc Natl Acad Sci USA* 2015, 112, 4630–4635.
- [115] van Wijk, S.J.L., Fiškin, E., Dikic, I., Selective monitoring of ubiquitin signals with genetically encoded ubiquitin chain-specific sensors. *Nat Protoc* 2013, 8, 1449–1458.
- [116] Lowe, E.D., Hasan, N., Trempe, J.F., Fonso, L., et al., Structures of the Dsk2 UBL and UBA domains and their complex. *Acta Crystallogr. D Biol. Crystallogr.* 2006, 62, 177–188.
- [117] Seibenhener, M.L., Babu, J.R., Geetha, T., Wong, H.C., et al., Sequestosome 1/p62 is a polyubiquitin chain binding protein involved in ubiquitin proteasome degradation. *Mol Cell Biol* 2004, 24, 8055–8068.
- [118] Layfield, R., Tooth, D., Landon, M., Dawson, S., et al., Purification of poly-ubiquitinated proteins by S5a-affinity chromatography. *Proteomics* 2001, 1, 773–777.
- [119] Hospenthal, M.K., Mevissen, T.E.T., Komander, D., Deubiquitinase-based analysis of ubiquitin chain architecture using Ubiquitin Chain Restriction (UbiCRest). *Nat Protoc* 2015, 10, 349–361.
- [120] Lundby, A., Secher, A., Lage, K., Nordsborg, N.B., et al., Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues. *Nat Commun* 2012, 3, 876–.
- [121] Lee, H.-J., Na, K., Kwon, M.-S., Kim, H., et al., Quantitative analysis of phosphopeptides in search of the disease biomarker from the hepatocellular carcinoma specimen. *Proteomics* 2009, 9, 3395–3408.
- [122] Kettenbach, A.N., Schweppe, D.K., Faherty, B.K., Pechenick, D., et al., Quantitative Phosphoproteomics Identifies Substrates and Functional Modules of Aurora and Polo-Like Kinase Activities in Mitotic Cells. *Science Signaling* 2011, 4, rs5–rs5.
- [123] Yang, L., Vaitheesvaran, B., Hartil, K., Robinson, A.J., et al., The Fasted/Fed Mouse Metabolic Acetylome: N6-Acetylation Differences Suggest Acetylation Coordinates Organ-Specific Fuel Switching. *J Proteome Res* 2011, 10, 4134–4149.
- [124] Galisson, F., Mahrouche, L., Courcelles, M., Bonneil, E., et al., A Novel Proteomics Approach to Identify SUMOylated Proteins and Their Modification Sites in Human Cells. *Mol Cell Proteomics* 2011, 10, M110.004796.
- [125] Wauer, T., Swatek, K.N., Wagstaff, J.L., Gladkova, C., et al., Ubiquitin Ser65 phosphorylation affects ubiquitin structure, chain assembly and hydrolysis. *EMBO J* 2015, 34, 307–325.
- [126] Koyano, F., Okatsu, K., Kosako, H., Tamura, Y., et al., Ubiquitin is phosphorylated by PINK1 to activate parkin. *Nature* 2014, 510, 162–166.
- [127] Swaney, D.L., Rodríguez-Mias, R.A., Villén, J., Phosphorylation of ubiquitin at Ser65 affects its polymerization, targets, and proteome-wide turnover. *EMBO Rep* 2015.
- [128] Xu, P., Peng, J., Characterization of polyubiquitin chain structure by middle-down

- mass spectrometry. *Anal Chem* 2008, 80, 3438–3444.
- [129] Valkevich, E.M., Sanchez, N.A., Ge, Y., Strieter, E.R., Middle-down mass spectrometry enables characterization of branched ubiquitin chains. *Biochemistry* 2014, 53, 4979–4989.
- [130] Siuti, N., Kelleher, N.L., Decoding protein modifications using top-down mass spectrometry. *Nat Methods* 2007, 4, 817–821.
- [131] Nir, I., Huttner, D., Meller, A., Direct Sensing and Discrimination among Ubiquitin and Ubiquitin Chains Using Solid-State Nanopores. *Biophysj* 2015, 108, 2340–2349.
- [132] Fu, C., Li, J., Wang, E., Signaling network analysis of ubiquitin-mediated proteins suggests correlations between the 26S proteasome and tumor progression. *Mol. BioSyst.* 2009, 5, 1809–1816.
- [133] Markson, G., Kiel, C., Hyde, R., Brown, S., et al., Analysis of the human E2 ubiquitin conjugating enzyme protein interaction network. *Genome Res* 2009, 19, 1905–1911.
- [134] Behrends, C., Sowa, M.E., Gygi, S.P., Harper, J.W., Network organization of the human autophagy system. *Nature* 2010, 466, 68–76.
- [135] Rose, C.M., Merrill, A.E., Bailey, D.J., Hebert, A.S., et al., Neutron encoded labeling for peptide identification. *Anal Chem* 2013, 85, 5129–5137.
- [136] Rhoads, T.W., Rose, C.M., Bailey, D.J., Riley, N.M., et al., Neutron-encoded mass signatures for quantitative top down proteomics. *Anal Chem* 2014.

Figure legends

Figure 1

The ubiquitin-driven destruction and signaling cycle. Ubiquitin (blue) is processed at the ribosome to its mature form. An E1 enzyme binds the mature ubiquitin while hydrolyzing one ATP to AMP. The bound ubiquitin is then transferred to an E2 enzyme in cooperation with the E3 ubiquitin ligase. The E3 ligase confers specificity, allowing only selected substrates (dark grey) to be ubiquitinated. The E2 and E3 combination is selective for both the residue of the substrate protein to be modified and, in the case of polyubiquitination, for the linkage type formed. The cycle of activation and transfer can be repeated, adding ubiquitin molecules to a growing chain on the target protein. If the substrate is modified with more than one ubiquitin moiety, the linkage type will define the regulatory outcome. K48-linked polyubiquitinated substrates are subjected to proteolysis by the proteasome (left panel), while other

Probing the ubiquitin landscape

polyubiquitin linkage types or mono-/multiple mono-ubiquitination lead to different regulatory outcomes. DUBs can edit the ubiquitin topology of the substrate by removing moieties, further modulating regulation of the substrate.

Figure 2

Alignment of the C-terminus of mature human ubiquitin and other ubiquitin-like modifiers

(UBLs). For all of these modifiers, conjugation to the substrate protein occurs at the C-terminus of the UBL. Depending on the sequence of the C-terminus, a tryptic digest leaves either a dipeptide or a short oligopeptide on the side-chain, a characteristic used for the identification of the modified residue. Trypsin cleaves C-terminal to basic lysine and arginine residues (indicated in light green). Ubiquitin, Nedd8 and ISG15 all have an arginine preceding the C-terminal glycine-glycine. Tryptic digestion, therefore, generates a glycine-dipeptide on the side-chain that is indistinguishable for all three proteins (dark green), while other UBLs generate different dipeptides or longer peptides that are more challenging to identify in proteomic screens.

Figure 3

Chemical structure of a ubiquitinated polypeptide. Trypsin recognizes lysine (light green) and arginine (dark green) and cleaves C-terminal to the peptide bond. The first lysine in this example polypeptide chain is modified by ubiquitin on the ϵ -amino-group. A tryptic digest will cleave C-terminal to the arginine in the ubiquitin on the side-chain, and it will not cleave after the modified lysine residue, due to steric bulk hindering access.

Figure 4

Enrichment strategies for ubiquitinated proteins. **A.** Expression of an N-terminal poly-histidine (His) tagged ubiquitin allows enrichment of ubiquitinated proteins using metal-chelate-affinity. **B.** Enrichment strategy using the BirA-tag and the co-expression of a specific ligase. The N-terminal bio-tag is recognized and *in vivo* biotinylated by the BirA-ligase. The biotinylated ubiquitin is enriched using avidin- or streptavidin-based affinity chromatography. **C.** Ubiquitin-chain specific antibodies recognize a particular polyubiquitin chain and allow the enrichment of associated substrate proteins. The specificity can be mediocre, however, and their usage is often hampered by high background. **D.** Tandem Ubiquitin Binding Entities (TUBEs) are based on concatenated ubiquitin-associated (UBA) domains. Immobilized TUBEs can be used for the chain-specific enrichment of ubiquitinated proteins. **E.** Diglycine (GG) remnant-targeted antibodies are designed to bind tryptic peptides containing the C-terminal GG motif of ubiquitin on a modified substrate peptide. These antibodies are used after the proteins have been converted into peptides and allow enrichment of ubiquitin-modified peptides on a large scale. **F.** Ubiquitin Chain RESTriction (UbiCRest) was designed to probe ubiquitin chains by using a set of DUBs with known specificity. After the enrichment of ubiquitinated proteins by other techniques, proteins are probed systematically using chain-specific DUBs.

Figure 5

Targeted proteomics approach using selected reaction monitoring to quantify ubiquitin and ubiquitin linkage types. **A.** Schematic representation of a triple-quadrupole (QQQ) mass spectrometer used for SRM analysis. A target precursor ion (e.g. a ubiquitin peptide) is selected in the first quadrupole (Q1) of the mass spectrometer and transmitted for

Probing the ubiquitin landscape

fragmentation in the Q2 (collision cell). In the Q3, specific fragment ions are selected, resulting in a measureable ion current. This approach is highly specific and allows reproducible quantification of a specific set of peptides. Absolute quantification occurs by comparison of the ion current trace to a spiked-in heavy-isotope analogue of the same peptide. **B.** The method can be applied to quantify linkage types by selecting precursor ion/fragment ion pairs specific to ubiquitin tryptic peptides bracketing one of the 7 lysine residues of ubiquitin. As an example, a K48 (left side) and K63 (right side) linked peptides are depicted. The method can be applied for all seven linkage types as well as stretches of native ubiquitin to quantify the overall abundance of ubiquitin in a sample.

Table 1 Ubiquitin genes

	Human		Yeast	
	Gene	Gene-symbol	Gene	Gene-symbol
Polyubiquitin	Polyubiquitin B	UBB	Polyubiquitin	Ubi4
	Polyubiquitin C	UBC		
Ribosomal fusions	Ubiquitin-60S ribosomal protein L40	UBA52	40S ribosomal protein S31 -	Rps27
	Ubiquitin-40S ribosomal	Rps27A	Ubiquitin-60S	Rpl40B

Probing the ubiquitin landscape

protein S27a

ribosomal protein L40

Ubiquitin-60S

Rpl40A

ribosomal protein L40

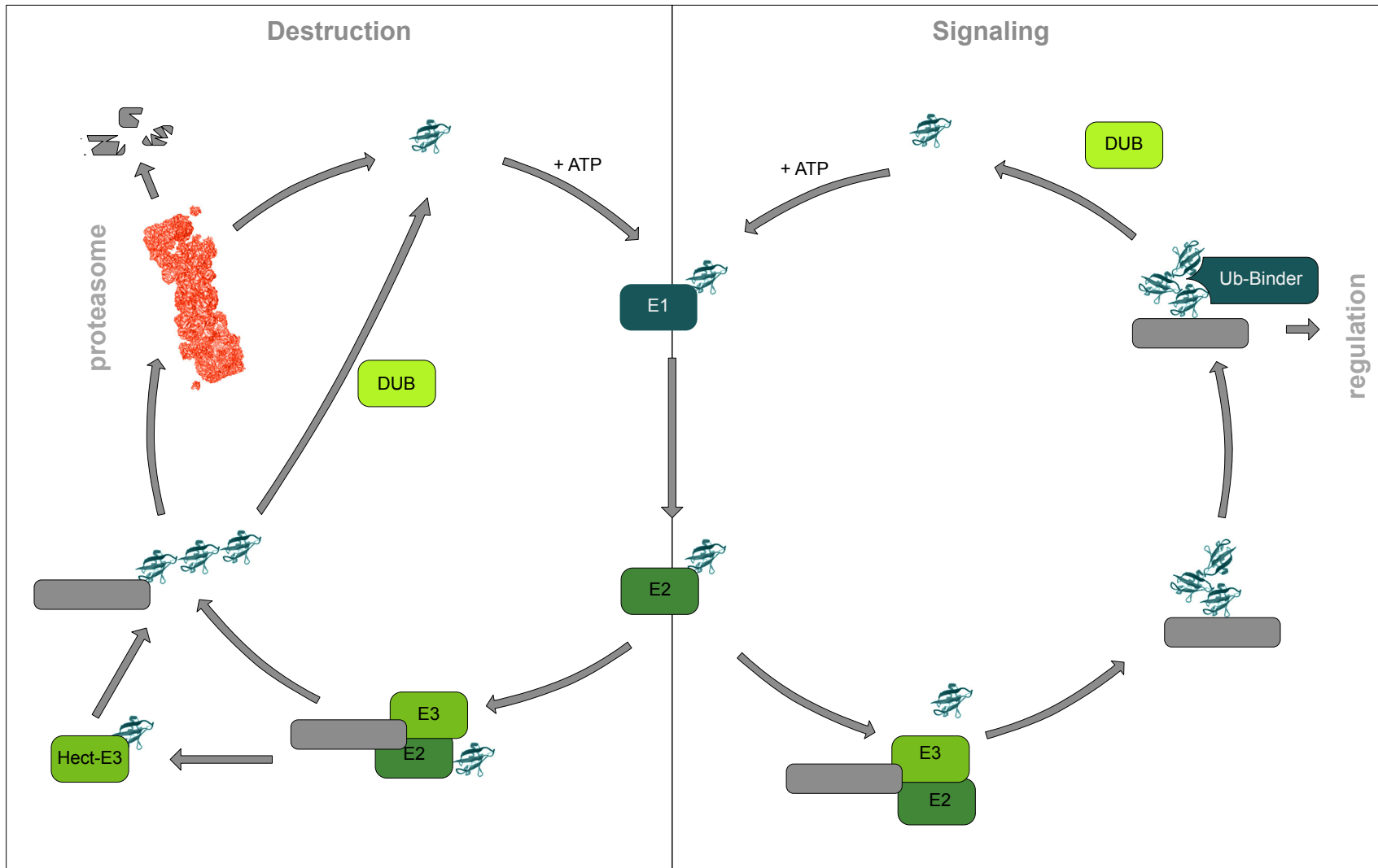
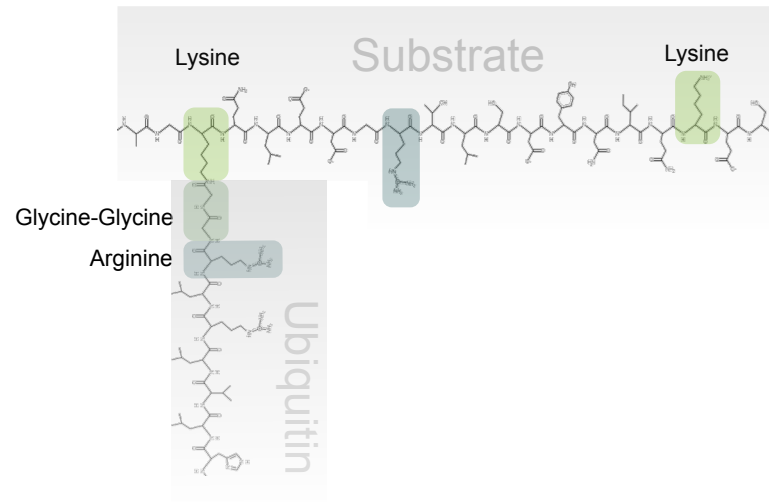


Figure 1

Figure 2

SP P0CG48 UBC_HUMAN	DQQR LIFAGK QLEDGRTLSDYNIQKESTLHLVLR L RGG
SP P05161 ISG15_HUMAN	DLFWLTFEGKPLEDQLPLGEYGLKPLSTVFMNLRL RGG
SP Q15843 NEDD8_HUMAN	QQRLIYSGKQMNDEKTAADYKILGGSVLHLVLAL RGG
SP P55854 SUM03_HUMAN	RQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQTGG
SP P61956 SUM02_HUMAN	RQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQTGG
SP P63165 SUM01_HUMAN	NSLRFLFEGQRIADNHTPKELGMEEEDVIEVYQEQTGG
SP O15205 UBD_HUMAN	ETQIVTCNGKRLEDGKMMADYGIRKGNLLFLACYCIGG
SP P61960 UFM1_HUMAN	TSAIITNDGIGINPAQTAGNVFLKHGSELRIIPDRVG
SP Q9BTM9 URM1_HUMAN	GILVLINDAD--WELLGELDYQLQDQDSVLFISTLHGG

Figure 3



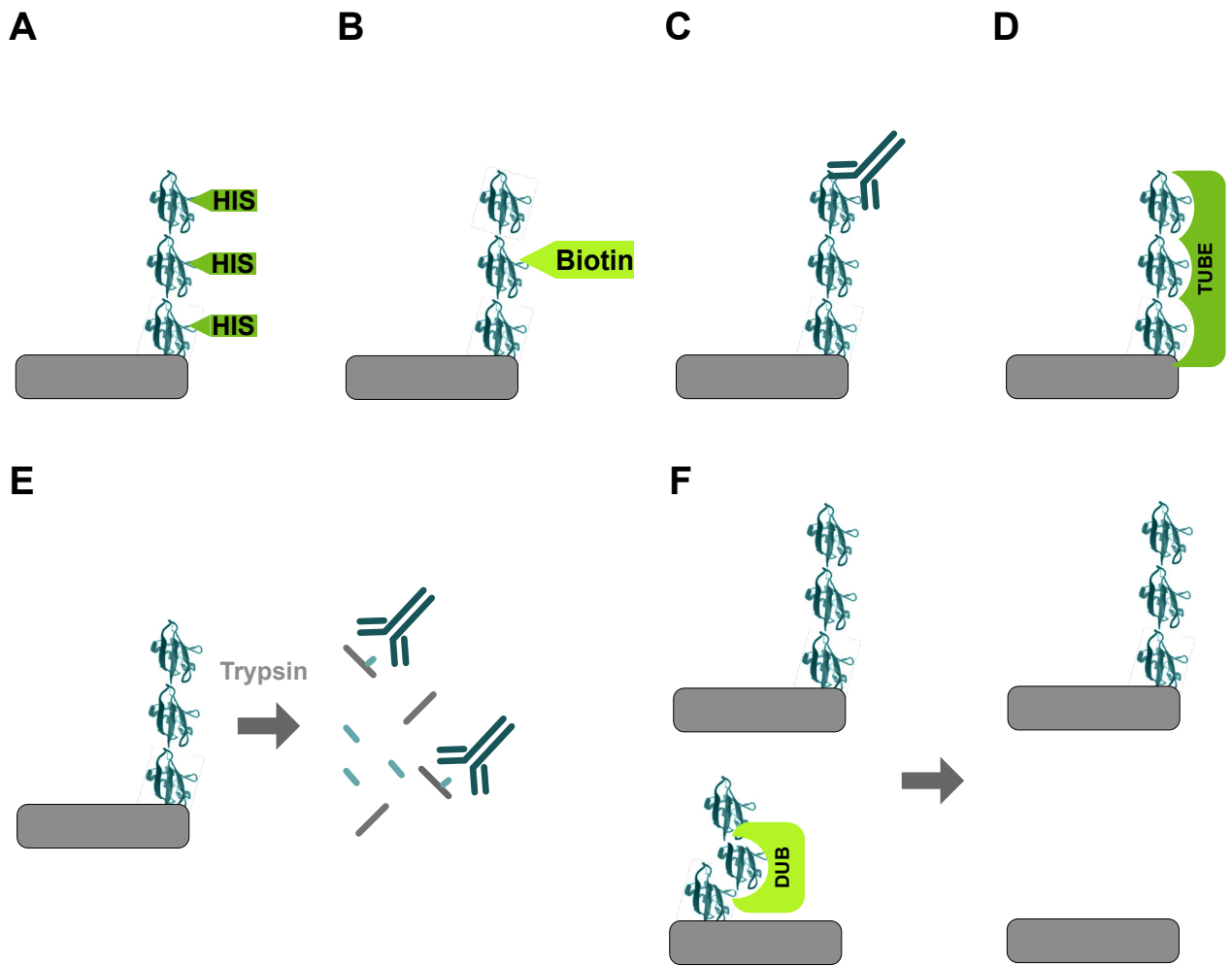


Figure 4

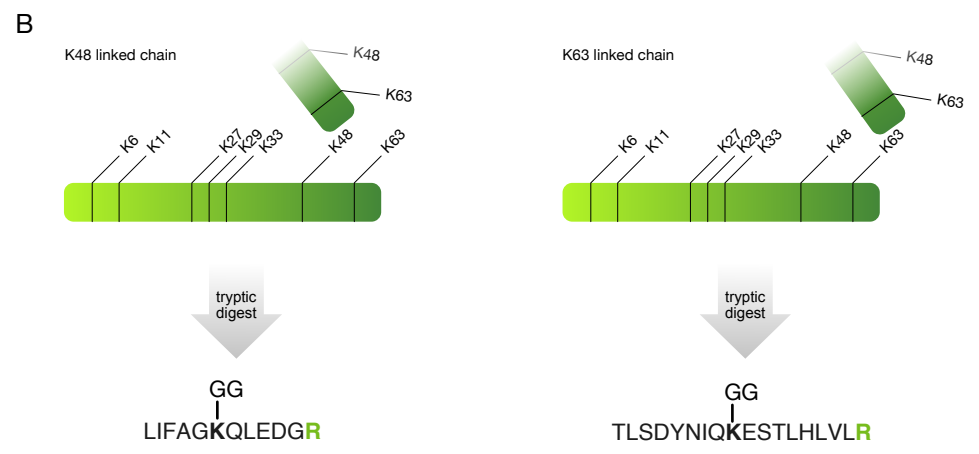
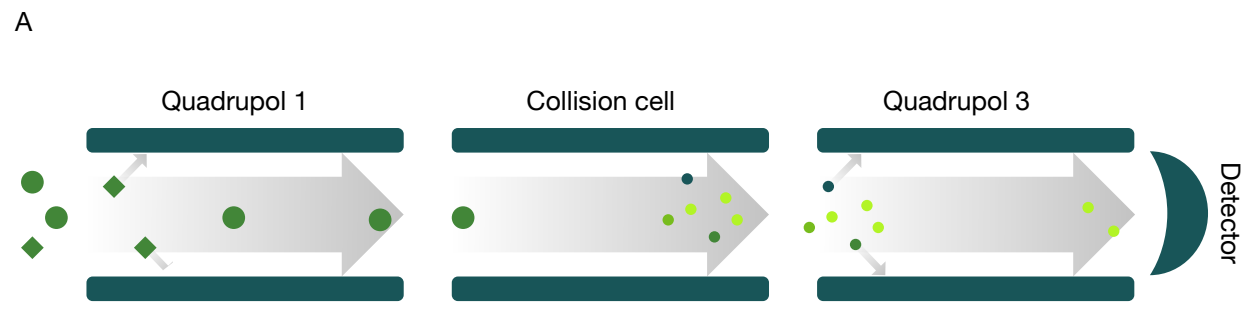


Figure 5