Uncovering Ubiquitin and Ubiquitin-like Signaling Networks

Alfred C. O. Vertegaal*

Department of Molecular Cell Biology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

1. INTRODUCTION

Proteomes are significantly more complex than genomes and transcriptomes due to protein processing and extensive post-translational modification (PTM) of proteins. Hundreds of different modifications exist. Release 66 of the RESID database1 (http://www.ebi.ac.uk/RESID/) contains 559 different modifications, including small chemical modifications such as phosphorylation, acetylation, and methylation and modification by small proteins, including ubiquitin and ubiquitin-like (UBL) proteins that are covalently coupled to proteins to regulate their activity. A wide variety of cellular processes are regulated by these reversible modifications, including transcription, replication, cell-cycle progression, and responses to DNA damage.

Protein modifications have been studied for many years at the level of single target proteins, but currently available technologies enable proteome-wide studies of these modifications by mass spectrometry (MS).2,3 Powerful proteomics tools are available to study phosphorylation and acetylation at a systems-wide level in a site-specific manner. It is more challenging to study ubiquitin targets and targets for ubiquitin-like proteins at a proteome-wide level in a site-specific manner due to the relatively large size of these modifications, but hundreds of potential target proteins have been uncovered over the past eight years, mainly in a non-site-specific manner. This review is focused on uncovering signaling networks for ubiquitin and ubiquitin-like proteins by mass spectrometry and highlights the site-specific studies published in 2010 and 2011. Site-specific methodologies will likely have a major impact on the ubiquitin field in the near future. The methodology, results, challenges, pitfalls, crosstalk with other PTMs, and future directions are discussed in this review.

Received: May 27, 2011
Published: October 17, 2011
1.1. Ubiquitin and Ubiquitin-like Proteins

Ubiquitin was first discovered in the mid-1970s, and the 2004 Nobel Prize in Chemistry was awarded for this finding. Ubiquitin is a 76 amino acid protein that is highly conserved from yeast to plants and mammals. Many ubiquitin-like proteins have been uncovered, including Nedd8, small ubiquitin-like modifier 1 (SUMO-1), SUMO-2, SUMO-3, PUBL, HUBL, ISG15, FAT10, URM1, UFM1, Atg12, and Atg8. Ubiquitin-like proteins are also found in prokaryotes and archaea; PUPs are prokaryotic ubiquitin-like proteins, and SAMPs are ubiquitin-like small archaeal modifier proteins. Despite limited sequence homology of some family members with ubiquitin, all ubiquitin family members display structural homology via the characteristic β-grasp ubiquitin fold.4–9

These small proteins are covalently coupled to target proteins via isopeptide bonds between C-terminal diglycine motifs and ε-amino groups in lysines of target proteins using an enzymatic cascade that consists of an E1 enzyme,10 an activator of ubiquitin and UBLs, an E2 enzyme,11,12 and a ligase, known as an E3 enzyme13 (Figure 1). Humans express 8 E1 enzymes10 (including 1 dedicated to ubiquitin, 1 shared between ubiquitin and the UBL FAT10, and 6 dedicated to other UBLs) and 35 active E2 enzymes (including 28 dedicated to ubiquitin, 3 shared between ubiquitin and the UBL ISG15, 3 dedicated to other UBLs, and 1 putative E2).15 Ubiquitin E3 enzymes are subdivided into HECT-type E3 enzymes (homology to E6AP carboxyl terminus)14 and RING-type E3 enzymes (really interesting new gene).15 HECT-type E3 enzymes form thioesters with ubiquitin, whereas RING-type E3 enzymes lack catalytic cysteines. Over 600 human genes encode components of RING-based E3 ligases.15

Ubiquitylation is a reversible process; nearly 100 different deubiquitylating enzymes (Dubs) are responsible for removing ubiquitin from target proteins in humans and for disassembling ubiquitin chains.16–19 These proteases belong to six different families, including five families of cysteine proteases: ubiquitin-specific proteases (USPs), herpesvirus tegument USPs (htUSPs), ubiquitin C-terminal hydrolases (UCHs), ovarian tumor proteases (OTUs), and the family of Josephins. The sixth family of ubiquitin proteases are the JAB1/MPN/MOV34 proteases (JAMMs), which are Zn²⁺ metalloproteases.

Seven internal lysines are present in ubiquitin that are used for chain formation and different chain topologies, including linear ubiquitin chains that can encode different signals20 and can be processed by specific proteases.21 Single ubiquitin moieties attached to target proteins (monoubiquitins) are also widely used as PTMs.22,23 In addition to covalent interactions, noncovalent interactions contribute to the complexity of signaling. At least 20 different types of domains have been identified in ubiquitin binding proteins (UBPs) that interact with ubiquitin in a noncovalent manner24,25 to regulate the fate of ubiquitylated proteins.

1.2. Quantitative Proteomics

Methods such as stable isotope labeling by amino acids in cell culture (SILAC),26,27 isobaric tags for relative and absolute quantification (iTRAQ),28 and absolute quantification (AQUA)29,30 are powerful tools that enable quantitative proteomics studies. Quantification can even be performed label-free.31 SILAC technology employs stable isotopic variants of amino acids for metabolic labeling of endogenous proteins and subsequent quantification.32 Labeling can also be performed postharvesting using chemical methods such as iTRAQ.28 Labeled marker peptides are used for AQUA, and these peptides are spiked at known concentrations to enable absolute quantifications.29,30 Quantitative proteomics is particularly useful for studying ubiquitylation dynamics in response to different stimuli. Furthermore, control cell populations can be included in experiments to distinguish between ubiquitin target proteins and contaminants. Cell lines can be used that express tagged ubiquitin or tagged UBL proteins. In SILAC experiments, these cell lines and control
Figure 2. Ubiquitin purification strategies. Three strategies are available to purify ubiquitin. The first strategy employs epitope-tagged ubiquitin that is attached to target proteins. These tags are used to purify ubiquitin. The second strategy makes use of anti-ubiquitin antibodies to purify endogenous ubiquitin. The third strategy uses ubiquitin traps that consist of epitope-tagged proteins containing one or more ubiquitin interaction motifs (UIMs) to enrich for endogenous ubiquitin.

1.3. Setting the Scenery: Mass Spectrometry Based Investigation of Phosphorylation and Acetylation

Mass spectrometry based proteomics studies of phosphorylation and acetylation are particularly successful and provide useful frameworks for proteomics studies of other PTMs, including ubiquitinylation. Critical tools are available that enable the purification of modified peptides and proteins, and these modifications can be identified in a site-specific manner. Phosphopeptides are purified using Fe\(^{3+}\) immobilized metal affinity chromatography (IMAC) and TiO\(_2\). The copurification of acidic peptides can be circumvented by including 2,5-dihydroxybenzoic acid (DHB) in the buffers as a competitive agent. Phosphorylated tyrosine containing peptides and proteins can be enriched using a specific antibody directed against phosphotyrosines. Alternatively, Src homology 2 (SH2) domains that normally bind phosphotyrosines in a sequence-specific manner can be used as traps to enrich tyrosine-phosphorylated proteins.

In 2006, a key paper was published on phosphorylation dynamics in response to epidermal growth factor (EGF). This study combined SILAC technology with phosphopeptide enrichment to identify EGF-regulated phosphorylation events at an until then unprecedented scale of 6600 events. Different clusters of kinetic profiles were identified, including signal initiators, intermediate stimulators, late stimulators, terminal effectors, and early and late negative regulators. The relative frequency of phosphorylation of tyrosine, threonine, and serine residues was 1.8%, 11.8%, and 86.4%, respectively, compared to initial estimates of 0.05%, 10%, and 90%, respectively, by Hunter and Seyton, thus a marked increase in tyrosine phosphorylation. Another example of the power of proteomics was the identification of 3067 phosphorylation sites during embryonic stem cell differentiation. More recently, over 20,000 phosphorylation events were uncovered during cell-cycle progression and nearly 36,000 phosphorylation sites were detected in different mouse organs, revealing tissue-specific phosphorylation patterns.

Studying acetylation is more challenging. The available polyclonal antibody is a limited resource and has been used to identify over 3600 lysine acetylation sites in 1750 proteins. SILAC was employed to identify changes in the acetylated proteome in response to the deacetylase inhibitors suberoylanilide hydroxamic acid and MS-275. This study showed that lysine acetylation preferentially targets large protein complexes involved in many different cellular processes, including cell-cycle progression, chromatin remodeling, pre-mRNA splicing, and nuclear transport. Monoclonal antibodies that recognize acetylated lysines are also currently available.

2. Ubiquitin and Ubiquitin-Like Protein Purification Approaches

Major approaches to purify ubiquitin and ubiquitin-like proteins include the use of epitope-tagged forms of these small protein modifiers, specific traps, and specific antibodies that enable the purification of endogenous ubiquitin and ubiquitin-like proteins (Figure 2). These approaches can be regarded as complementary methods with specific advantages and disadvantages that are discussed below. Importantly, purification methods have to deal with the challenge of Dubs and UBL proteases to avoid deconjugation from target proteins.

2.1. Epitope-Tagged Ubiquitin and Ubiquitin-like Proteins

Short histidine stretches (hexamer to octamer) have traditionally been very useful epitope tags for ubiquitin and UBL enrichment since they interact with metal ions, including Ni\(^{2+}\) and Co\(^{2+}\), in strongly denaturing buffers such as 8 M urea and 6 M guanidinium hydrochloride that inactivate Dubs and UBL proteases. In addition, this methodology also prevents the copurification of noncovalent ubiquitin- and UBL-interacting proteins. Subsequent analysis of purified conjugates by immunoblotting has extensively been used in the field to show the modification of exogenous and endogenous target proteins. These tags can also be used for proteomics approaches; however, proteins that contain endogenous histidine stretches will be copuriified.

Other useful epitope tags include the Myc-tag, HA-tag, and FLAG-tag, but the purification of tagged proteins is more challenging in denaturing buffers. Lysates can be prepared in sodium dodecyl sulfate (SDS)-containing buffers to inactivate proteases, but need to be diluted in milder buffers to avoid the
denaturation of the anti-tag antibodies. This could cause refolding of proteins and potentially allows noncovalent interactions to occur. Nevertheless, purified fractions are relatively pure. Tandem affinity purification (TAP) strategies include the traditional TAP-tag consisting of a protein A domain and a calmodulin binding domain separated by a tobacco etch virus (TEV) cleavage site.50,51 Alternatively, smaller tandem tags could be employed. Tandem tag strategies are used to obtain very pure fractions, but the final yields can be suboptimal. Lysates are also prepared in SDS-containing buffers, and dilution is needed to allow purification.

Another interesting strategy is the use of peptide stretches at the N-terminus of ubiquitin and UBL proteins that can be biotinylated and subsequently purified using avidin or streptavidin resins.52,53 Since the biotin—streptavidin interaction is very strong, denaturing conditions can be used for purification. However, endogenous proteins are frequently biotinylated in cells and are copurified. A major advantage of the use of epitope-tagged ubiquitin and UBL proteins is that these methods are unlikely to discriminate between different types of ubiquitin chains and between mono- and polyubiquitination.

2.2. Traps Based on Ubiquitin- and Ubiquitin-like Binding Domains

As stated earlier, the ubiquitin system includes an extensive family of noncovalent ubiquitin interactors that bind ubiquitylated proteins via at least 20 different types of ubiquitin interaction motifs.24,25,54—56 These ubiquitin binding domains are used in the field to create traps to purify endogenous ubiquitin and UBL proteins, enabling the use of primary cells and tissues and patient material that cannot be studied by epitope tagging based methods. Capturing the full ubiquitylated proteome by ubiquitin traps is challenging due to potential preferences for specific types of chains and preferences for polyubiquitination versus monoubiquitination.

2.3. Antibody-Based Purification of Ubiquitin and Ubiquitin-like Proteins

Similar to the purification of proteins that are acetylated on lysine residues or phosphorylated on tyrosine residues, specific antibodies can be used to purify ubiquitin and UBL proteins. Since ubiquitin is highly conserved from yeast to humans, recombinant ubiquitin is a suboptimal antigen, and it is therefore difficult to generate high-affinity anti-ubiquitin antibodies. Consequently, this results in a relatively low yield of immunoprecipitates. A tool that enables site-specific identification of ubiquitin target proteins is a monoclonal antibody that enriches for peptides containing lysine residues modified by diglycine.57

2.4. Challenges and Pitfalls

Currently available methodologies harbor the risk of copurifying noncovalent ubiquitin interactors, particularly when using partly renatured lysates. Furthermore, it is important to discriminate between specifically purified proteins and contaminants. Quantitative proteomics experiments that include control pools of cells could be used to discriminate between ubiquitylated proteins and contaminants. When using affinity tags, the size of the tag could potentially influence the conjugation and deconjugation rates of the tagged proteins. Similar to acetylation and phosphorylation, site-specific detection of modified residues in target proteins would be the most rigorous method to unequivocally identify modified proteins, but this is still challenging.

A critical aspect of the methodology is the inactivation of ubiquitin and UBL proteases. Most of these proteases are cysteine proteases; therefore, iodoacetamide has extensively been used in the past. However, the use of iodoacetamide is not recommended since iodoacetamide adducts on lysine residues have exactly the same mass shift as the diglycine tryptic fragments of ubiquitin: 114.043 Da. These artifacts are particularly efficiently produced at higher temperatures, but are not detected at 21 °C58 and can also be avoided by using chloroacetamide instead of iodoacetamide.59

Unfortunately, the mass of the diglycine fragment of 114.043 Da is also indistinguishable from the mass of Asn, and this could lead to false-positive assignments.60 In contrast, the monoisotopic mass of 79.966 Da for phosphorylation is unique, thus avoiding confusion with other amino acids. In addition to the diglycine fragment, a larger ubiquitylation remnant of 383.228 Da corresponding to LRGG has been observed.60,61 It could be helpful to use this larger fragment also for database searches to identify ubiquitylation events. One group has reported that trypsin also directly cleaves C-terminal of ubiquitylated lysines,61 but this was not confirmed by others.60,62

2.5. Summary

Three different methodologies are used in the field to purify ubiquitin and UBL target proteins, including protein tagging, specific traps, and direct immunoprecipitation. Tagged forms of ubiquitin and UBLs are extensively used. They enable the purification of target proteins in an unbiased manner and are unlikely to discriminate between modification via single ubiquitin or UBL entities or via different polymeric forms. Care should be taken to limit overexpression of these endogenous modifiers. Histidine tags and biotinylated tags enable the purification under fully denaturing conditions. These conditions will inactivate ubiquitin and UBL proteases, and the copurification of noncovalent interactors is avoided. However, proteins that contain endogenous histidine stretches and endogenous biotinylated proteins will be copurified during these procedures. Epitope tags such as HA, FLAG, and Myc also enable efficient purification, but are not compatible with fully denaturing conditions, thus allowing the copurification of noncovalent ubiquitin and UBL interactors. Tandem tags are successfully employed to obtain very pure fractions, but large tags can affect the conjugation and deconjugation rates of ubiquitin family members.

Specific traps consisting of ubiquitin and UBL binding domains enable the purification of target proteins for endogenous ubiquitin and UBLs. Multimerized forms of these binding domains are very potent capturing reagents, and hundreds of target proteins can simultaneously be purified. However, these traps are likely to discriminate between monomodification and modification by specific forms of polyubiquitination.

Endogenous ubiquitin and UBL proteins are also purified by immunoprecipitation using antibodies against ubiquitin and ubiquitin family members. Several antibodies are available that recognize ubiquitin polymers in a linkage-dependent manner. However, due to the strong conservation of ubiquitin, it is difficult to obtain high-affinity antibiubiquitin antibodies.

Complementary use of the available methodologies will yield the most reliable data sets. Identification of the ubiquitin and UBL acceptor lysines in target proteins provides the most reliable evidence for modification and is important for functional follow-up studies. However, it is currently challenging to obtain a good coverage of ubiquitin and UBL sites at a systems-wide level.
The ultimate goal of ubiquitin and UBL target proteomics would be to identify dynamically regulated endogenous ubiquitin and UBL target proteins at a systems-wide level in a site-specific and quantitative manner.

3. Ubiquitin Proteomics

3.1. Proteomic Studies Employing Tagged Ubiquitin

In 2003, a landmark paper on ubiquitin proteomics in Saccharomyces cerevisiae was published by the laboratory of Gygi, highlighting the impact of proteomics on the ubiquitin field. A yeast strain engineered to express His6-tagged ubiquitin as the only source of ubiquitin was used for large-scale purification of ubiquitylated proteins under denaturing conditions. Shotgun sequencing identified 1075 potential ubiquitin substrates, indicating a broad impact of ubiquitin on all cellular processes. Similar purification approaches were used in subsequent studies using different affinity tags.

An elegant virtual immunoblot method was later developed to validate this extensive list of potential ubiquitin substrates. Purified ubiquitylated proteins and total cell lysates were size-separated by one-dimensional SDS—polyacrylamide gel electrophoresis (PAGE), and the gel was sliced in small gel bands and analyzed by mass spectrometry. The distributions of the proteins in the gel slices in both samples were compared to verify shifts in the gel position that could be accounted for by ubiquitylation. Despite the use of denaturing conditions, only ~30% of the potential ubiquitin conjugates were confirmed by this method, indicating the copurification of large amounts of nonmodified proteins with His6 ubiquitin and illustrating the problem of distinguishing between covalently modified ubiquitin target proteins and contaminants.

3.2. Ubiquitin Binding Domains

Over 250 proteins have been identified that contain ubiquitin binding domains (UBDs). Recombinant UBD-containing proteins can be used to identify the ubiquitylated proteins that are regulated by these proteins. Controls that contain mutations in the UBD domain can be included to distinguish ubiquitin-dependent interactions from ubiquitin-independent interactions.

Moreover, recombinant UBDs are great tools for purifying ubiquitylated proteins in general, but they have intrinsic limitations since they will likely interact with subsets of ubiquitylated proteins. Furthermore, they usually have a relatively low affinity for ubiquitylated proteins. The S5a (Rpn10) subunit of the 26S proteasome is a useful tool and binds to polyubiquitin chains that consist of at least four ubiquitin moieties. Glutathione S-transferase (GST)—S5a was used by Tan et al. to purify ubiquitylated proteins from liver cells, and 83 potential ubiquitin targets were identified. In 17 substrates ubiquitylation sites were identified. Arabidopsis thaliana S5a coupled to GST and a double ubiquitin-associated domain (UBA) from AtUBP14 coupled to GST were used to purify and identify a total of 294 potential ubiquitin target proteins in Arabidopsis cell suspension cultures. An elegant study on S5a substrates in S. cerevisiae was published by Mayor et al. In this study, ubiquitin conjugates were purified in a two-step process using Rad23-Dsk2 resin and His6 ubiquitin purification using Ni2+ beads. Quantitative proteomics was employed using 15N labeling to compare the ubiquitin substrates in wild-type cells and cells deficient in S5a. Of the 225 ubiquitin—proteasome system substrates, 27% were degraded in an S5a-dependent manner.

Recently, multimerized UBDs were shown to display up to a 1000-fold increase in affinity for ubiquitin, and these so-called tandem ubiquitin binding entities (TUBEs) now enable the purification of ubiquitylated proteins from primary tissues. Conveniently, TUBEs protect ubiquitylated proteins from deubiquitylation by Dubs, even in mild buffers.

3.3. Anti-Ubiquitin Antibodies

Many different antibodies directed against ubiquitin have been produced, but due to virtually complete conservation of ubiquitin between different species, recombinant ubiquitin is a suboptimal antigen. One of the most frequently used anti-ubiquitin antibodies is the monoclonal antibody FK2. Lysates can be prepared in denaturing buffers to inactivate proteases and the proteasome, but buffers need to be diluted to milder conditions to enable immunopurifications. The use of partly renatured proteins increases the risk of copurifying ubiquitin interactors together with ubiquitin target proteins. Vasilescu et al. identified 70 proteins in FK2 immunoprecipitates from MCF-7 breast cancer cells. In a large study by Matsumoto et al., 345 distinct proteins were identified that were enriched by FK2 from HEK293T cells under stringent conditions. Eighteen ubiquitylation sites were found, including ubiquitylation sites on lysines 6, 11, 33, 48, and 63 of ubiquitin, indicating chain formation of ubiquitin via these residues.

Recently, a large quantitative ubiquitin proteomics project was published, describing how EGF affects the ubiquitin proteome. A total of 1175 proteins were identified in FK2 immunoprecipitates from HeLa cells. In addition, 582 proteins were identified in FLAG—His6 ubiquitin precipitates to reach a total nonredundant set of 1472 potential ubiquitin target proteins. SILAC technology was used to study EGF-induced alterations in protein ubiquitylation, and 265 EGF-responsive target proteins were identified, including endocytic proteins, ubiquitylating and deubiquitylating enzymes, transporters, and proteins involved in translation and transcription. The number of ubiquitylation sites identified in this study, 31, was still modest compared to the large total number of target proteins. Interestingly, EphA2 was identified as a novel, downstream ubiquitylated target of the epidermal growth factor receptor that is critical for proliferation and cell migration in response to EGF.

3.4. Site-Specific Ubiquitin Proteomics

Similar to phosphoproteomics and acetylation proteomics, site-specific identification of ubiquitin targets is the most reliable method for the detection of covalently modified proteins. In the study by Peng et al., 110 ubiquitylation sites were identified. Recently, a monoclonal antibody was generated that specifically recognizes diglycines attached to lysines corresponding to tryptic fragments of ubiquitin attached to tryptic target protein fragments (Figure 3). This antibody enabled the purification of ubiquitylated fragments from target proteins and the identification of 374 diglycine-modified lysines on 236 ubiquitylated proteins. Naturally occurring internal diglycines in proteins were not recognized by the antibody, indicating its specificity. This is an important step toward systems-wide identification of ubiquitin targets in a site-specific manner.

The majority of the identified target proteins are located in the cytoplasm, and 28.7% of the ubiquitin targets are nuclear. Furthermore, ubiquitylation sites were identified in proteins that localize in mitochondria, the endoplasmic reticulum, the Golgi, and the plasma membrane. Nearly half of the targets are involved...
in cellular metabolism and 13% in cell-cycle progression and apoptosis. Signaling proteins represent 9.1% of the targets, and 8.3% of the targets are involved in protein trafficking. Ubiquitylated targets furthermore include structural proteins (4.4%), components of the immune system (4.2%), and transporters of small molecules (2.3%).

This interesting data set was used to search for ubiquitylation motifs. However, only subtle enrichment for specific residues was found. KxLxD was proposed as a potential consensus ubiquitylation site, but this motif is only 1.8 times enriched among ubiquitylated lysines compared to lysines in general. Interestingly, ubiquitylated lysines are not necessarily situated on sites that are normally surface exposed since only 60% of the ubiquitylated lysines are exposed, compared to 45% for lysines in general.

This site-specific manner of ubiquitin target identification has two significant drawbacks since the tryptic fragments that remain attached to target proteins from the ubiquitin-like proteins Nedd8 and ISG15 are also diglycines and the monoclonal antibody that is being used to purify modified lysines is not able to discriminate between diglycines corresponding to ubiquitylation, Neddylation, or ISGylation. To circumvent this problem and to increase the yield of the procedure, His6-ubiquitin was employed for prepurification prior to immunoprecipitation of the ubiquitylation sites. Furthermore, the antibody fails to recognize ubiquitylated cysteines, and the antibody might not capture the full ubiquitylated proteome due to the nature of the antigen that was used to generate this antibody.

A TUBE purification strategy was used by Shi et al. to purify endogenous ubiquitin conjugates. They identified 293 ubiquitylation sites on 223 substrates, including 33 mitochondrial proteins, suggesting an important role for ubiquitin in mitochondria. KxLxD was identified as an infrequent putative ubiquitylation motif.

Recently, Nielsen and co-workers reported the currently largest ubiquitin proteomics study with a total of 5756 putative ubiquitin substrates. Proteins conjugated to Strep–HA–ubiquitin were purified from U2OS and HEK293T cells stably expressing tandem tagged ubiquitin. Mass spectrometry was performed using an LTQ Orbitrap Velos. Both precursor and fragment peptides were analyzed in the orbitrap, resulting in high mass accuracy and optimal spectrum quality. The increased fragmentation efficiency of higher energy collisional dissociation (HCD) enabled the identification of a total of 753 unique sites on 471 proteins. Thus, site-specific enrichment is not required to identify a large number of ubiquitylation sites.

Interestingly, N-terminal substrate ubiquitylation was not observed, indicating that this is a rare event. Alignments of protein regions flanking ubiquitylated lysines did not reveal unique sequence features. However, it remains to be determined whether specific enzymes recognize specific ubiquitylation sites. Ubiquitin preferably targeted lysine residues located in ordered helical regions that were surrounded by smaller and positively charged amino acids. Low conservation of ubiquitylation sites between eukaryotic species was observed, indicative of promiscuity of ubiquitylation at the site level. Ubiquitin was furthermore found to form mixed chains with the ubiquitin-like proteins SUMO-1, SUMO-2, SUMO-3, and Atg12.

In addition to these large-scale projects, mass spectrometry has also been extensively used to identify ubiquitylation sites on individual target proteins. A few examples are the ubiquitylation sites on the X-linked inhibitor of apoptosis protein, TRAF1, histone variant macro H2A1.2 and histone H2A, and checkpoint with forkhead-associated and RING protein Chfr.

### 3.5. Ubiquitin Chains

Ubiquitin contains seven internal lysines (K6, K11, K27, K29, K33, K48, and K63), and all lysines participate in ubiquitin chain formation as first demonstrated by mass spectrometry. E2 enzymes play an important role in chain assembly and regulate chain topology. Lysine 48-linked ubiquitin chains were previously identified as the ubiquitin polymer that is required to target proteins for degradation by the proteasome. Lysine 48 of ubiquitin is essential for cell viability since a K48R mutant of ubiquitin as the single source of ubiquitin in S. cerevisiae causes cell-cycle arrest. Subsequently, lysine 63-linked chains were shown to play a role in DNA repair, as activators of the NFκB regulating kinase complex IKK (IkB kinase), and in antiviral innate immunity. Lysine 6-linked ubiquitin chains are generated by the complex consisting of Breast cancer type 1 susceptibility protein (BRCA1) and BRCA1-associated RING domain protein 1 (BARD1), and lysine 11-linked ubiquitin chains are generated by the anaphase-promoting complex/cyclosome (APC/C) to orchestrate mitotic progression. Lysine 29-linked chains play a role in lysosomal degradation. Furthermore, the IKK component NEMO (NFκB essential modulator) is regulated via linear ubiquitin chains. Linear ubiquitin chains are assembled by the linear ubiquitin chain assembly complex (LUBAC), which contains the RING family members RNF31 and RNF54, and Sharpin.

Mass spectrometric evidence to support head-to-tail linear ubiquitin chains has recently been provided. Surprisingly, ubiquitin chain editing occurs on receptor-interacting serine/threonine-protein kinase 1 (RIP1) and Interleukin-1 receptor-associated kinase 1 (IRAK1) and plays a role in innate immune signaling. Complex ubiquitin chain topology on in vitro ubiquitylated cyclin B1 was previously reported, and mixed chain formation also occurs in cells. However, complex forked chains are poorly degraded by the proteasome, and the generation...
of these forked chains is counteracted by the S5a protein via its ubiquitin interaction motif (UIM) domain. Interestingly, all ubiquitin chains with the exception of lysine 63-linked chains were shown by AQUA to accumulate upon inhibition of the proteasome, indicating that they all play a role in protein degradation. Consistently, mutant ubiquitin with lysine 48 as the single lysine available for chain formation cannot support yeast viability. Interestingly, aberrancies in ubiquitin signaling have been linked to neurodegenerative diseases such as Huntington’s disease (HD), Parkinson’s disease, and Alzheimer’s disease. In affected neurons, ubiquitin accumulates in inclusion bodies also containing polyglutamine-expanded huntingtin protein characteristic for HD. Mass spectrometry based methods were used to reveal that polyubiquitin chains, including lysine 48-, lysine 63-, and lysine 11-linked chains, accumulated early in brains from a transgenic HD mouse model, brains from a knock-in mouse model of HD, and in brains from human HD patients and also in brains from Alzheimer’s disease patients. These results imply that accumulation of these ubiquitin chains could be used as biomarkers of the disease.

### 3.6. Wiring of the Ubiquitin System

A significant percentage of eukaryotic genomes are dedicated to the ubiquitin system, including nearly 100 different ubiquitin proteases, 250 ubiquitin receptors, and over 600 components of E3 ligases in humans. We are limited in our understanding of the wiring of this system, and unbiased proteome-wide studies are needed to establish the identity of the single targets or multiple targets that are being regulated by different proteases, E3 enzymes, and noncovalent ubiquitin binders. Quantitative proteomics methodologies are particularly powerful tools and can be used to identify these targets by comparing ubiquitylated proteins in the presence and absence of specific proteases, ligases, or receptors (Figure 4). Inactivation of specific proteases or ligases can be carried out by RNAi or by specific chemical inhibitors. Whether the identified proteins are direct targets or are indirectly regulated by these proteases and ligases can subsequently be established using in vitro assays containing recombinant enzymes and substrates.

The first study that reported on the identification of a specific ubiquitin ligase for a target protein by mass spectrometry was published in 1998 and identified the E3 ligase F-box and WD repeats protein β-TrCP as a regulator of the NFκB inhibitor IκBα. In vitro experiments confirmed the ubiquitylation of IκBα by β-TrCP.

Another important ubiquitin target is the tumor suppressor protein p53. Critical regulators are the RING finger proteins Mdm2 and p14ARF.
and MdmX/4.\textsuperscript{118,119} Later, a surprisingly wide array of E3 ligases were implicated in p53 regulation, including Cop1 (RNF200),\textsuperscript{120} Pih2 (RNF199),\textsuperscript{121} and ARF-binding protein 1 (ARF-BP1).\textsuperscript{122} To search for p53 regulators in an unbiased manner, Barton and colleagues employed a mouse model expressing a TAP-tagged p53 protein.\textsuperscript{123} Purified p53 protein complexes were analyzed by mass spectrometry and revealed MDM2 and a novel ubiquitin ligase, the TRIM24 (RNF82) protein, as p53 regulators.

The potential of these approaches is furthermore illustrated by a quantitative proteomics study of the orphan F box protein Mdm30p,\textsuperscript{124} a protein that regulates mitochondrial morphology. Ubiquitylated proteins were purified from yeast cells overexpressing Mdm30p and control cells, using a clever two-step ubiquitin affinity purification method.\textsuperscript{125} Cells that express both His\textsubscript{6}–ubiquitin and FLAG–ubiquitin were used, and poly- and multi-ubiquitylated proteins were selectively isolated by purifying His\textsubscript{6}–ubiquitin under denaturing conditions followed by an anti-FLAG immunoprecipitation. Eluates were analyzed by mass spectrometry to reveal the mitochondrial outer membrane protein Mdm34p as a target protein that is ubiquitylated by Mdm30p.

Other E3 ligase–target protein connections that were detected by mass spectrometry include BRCA1/BARD1 and nucleophosmin/B23,\textsuperscript{126} the cellular inhibitor of apoptosis 2 (cIAP2),\textsuperscript{127} and for the APC/C complex.\textsuperscript{134} The complex consisting of DNA damage-binding protein 1 (DDB1), Cullin 4 (Cul4), and RBX1/ROC1 (RNF75) is a smaller Cullin E3 complex which regulates DNA repair, replication, and transcription and can be employed by viruses. A novel family of WD40-repeat-containing proteins were identified as substrate-recruiting modules of this E3 by protein complex purification and mass spectrometry.\textsuperscript{142} These substrate recruiters directly bind to the DDB1 double-propeller fold, revealing a novel mechanism underlying the substrate recruitment by this important E3 complex.

Another interesting study involves the ATP-dependent chaperone p97, a critical regulator of endoplasmic reticulum associated protein degradation (ERAD).\textsuperscript{143} Proteomics was employed to reveal an important novel function for p97 in ubiquitin–proteasome-mediated turnover of soluble proteins.\textsuperscript{144} A set of eight ubiquitin-like module UBX-domain-containing proteins was identified by proteomics as p97 interactors, and subsequently, all 13 UBX proteins were shown to bind p97. UBX proteins are furthermore connected to at least 38 ubiquitin E3 ligases, and a subset of UBX proteins can directly bind to ubiquitylated proteins via UBX domains to connect p97 to ubiquitin-mediated protein degradation by the proteasome.

E3 ligases and ubiquitin proteases interact with partner proteins to carry out their functions,\textsuperscript{18} and these protein–protein interactions are efficiently revealed by mass spectrometric analysis of purified protein complexes. Recently, the protein product of the candidate oncogene TSPYL5 was identified as a partner and inhibitor of USP7, resulting in increased p53 ubiquitlation.\textsuperscript{145} In a systematic approach, 75 human Dubs were epitope tagged and purified from cells stably expressing low amounts of these proteases.\textsuperscript{146} A total of 774 candidate interacting proteins were identified, and connections with biological pathways were identified for many previously unstudied Dubs. Interestingly, this study revealed that 26 ubiquitin proteases interact with ubiquitin ligases to form complexes that enable tight regulation of ubiquitination levels. This was previously shown by others for individual USP–E3 pairs, including the USP7–MDM2 pair,\textsuperscript{147,148} the USP44–APC/C\textsuperscript{149} pair, and the USP19–KPC1 (RNF123) pair.\textsuperscript{150}

3.8. Summary

Tagged forms of ubiquitin, ubiquitin traps, and anti-ubiquitin antibodies are used in the field to purify ubiquitin conjugates. Advantages and disadvantages of these methodologies are summarized in section 2.5. Site-specific identification of ubiquitin conjugates was recently enabled by a monoclonal antibody directed against diglycines attached to the ε-amino groups of lysines.\textsuperscript{57} This antibody is unfortunately unable to discriminate between ubiquitylation sites and modification sites of the UBL proteins ISG15 and Nedd8. It is however not necessary to use this monoclonal antibody to obtain site-specific insights into protein ubiquitylation, since the largest site-specific study was performed by analyzing a complex mixture of ubiquitin conjugates.\textsuperscript{78} In total 753 ubiquitylation sites were identified in this study. Interestingly, no consensus site for ubiquitylation could be identified in this data set. Improvements in site-specific methodologies are expected to enable increases in the coverage of ubiquitylation sites. This is needed to study ubiquitylation dynamics in response to different stimuli and to identify specific ubiquitin target proteins that are regulated by the vast amounts of E2 enzymes, E3 enzymes, and ubiquitin proteases.

4. SMALL UBIQUITIN-LIKE MODIFIERS

4.1. Overview

SUMO-1 was identified as a RanGAP1-modifying protein that targets RanGAP1 to the nuclear pore.\textsuperscript{131–134} The SUMO family includes three members, SUMO-1, -2, and -3, which display the characteristic β-grasp ubiquitin fold, despite limited sequence homologies to ubiquitin of 20%, 16%, and 16%, respectively.\textsuperscript{4,5}
Mature SUMO-2 and SUMO-3 are virtually identical and differ ~50% from SUMO-1. SUMOs are frequently conjugated to target proteins via SUMOylation consensus sites and form chains via internal SUMOylation sites. SUMOylated proteins interact in a noncovalent manner with other proteins via SUMO interaction motifs (SIMs). The SUMO conjugation cascade consists of the heterodimeric SUMO E1, the single SUMO E2 Ube2I, previously known as Ubc9, and a limited number of E3 enzymes, including Siz/PIAS family members, RanBP2, Pc2, and Topors. SUMOylation is a reversible process; six SUMO proteases are known to remove SUMOs from target proteins and mediate the maturation of SUMO precursor proteins. Ube2I-deficient mouse embryos die at the early post-implantation stage due to nuclear and chromosomal aberrations, showing that SUMOylation is essential for eukaryotic viability.

4.2. SUMO Proteomics

Similar to ubiquitylation, purification and identification of target proteins by mass spectrometry is critical for understanding protein SUMOylation. Epitope-tagged SUMOs are extensively used in the field to purify SUMO conjugates. Tags that have been used include protein A, biotinylated peptides, the His6-tag, the Myc-tag, the HA-tag, and the tandem tags His6-S tag, FLAG-TEV, His6-FLAG-HA, His6-MAHA, and protein A-TEV-CBP. Furthermore, SUMO conjugates have been purified using anti-SUMO antibodies or a SUMO affinity trap based on the N-terminal region from RNF4 that contains four SUMO Interaction Motifs (SIMs). Hundreds of potential SUMO target proteins have been identified in these studies.

The largest functional group of SUMO targets are transcriptional regulators including specific transcription factors, basal transcription factors, chromatin remodelers, coactivators, and corepressors. The emerging picture is that transcription factors are frequently inhibited by SUMOylation via recruitment of transcriptional repressors that contain SUMO interaction motifs. Furthermore, these proteomics studies have uncovered roles for SUMOs in many cellular processes, including DNA repair, RNA metabolism, protein transport, translation, and replication. Quantitative proteomics was used to show that different SUMO family members have overlapping and distinct sets of target proteins.

Most SUMO proteomics studies have been performed under regular cell culture conditions. Interestingly, SUMO-2 conjugates are strongly regulated by different stress conditions, including heat shock, and via crosstalk with the ubiquitin—proteasome system. SUMO proteases are very potent and are generally inhibited by employing denaturing buffers. These buffers also prevent the copurification of non-SUMOylated proteins that interact with SUMOylated proteins via SIMs. Several of the described tags require partial renaturation of the denatured proteins, and this could potentially lead to SUMO protease activity and copurification of non-SUMOylated proteins via SIMs.

4.3. Site-Specific SUMO Proteomics Approaches

The currently available methods to purify SUMOs harbor the risks of copurifying matrix-associated contaminants and SUMO binding proteins that interact noncovalently with SUMOylated proteins via SIMs. Site-specific techniques that enable direct mapping of SUMO acceptor lysines would greatly facilitate progress in the field. However, in contrast to phosphorylation and acetylation, it has been virtually impossible to directly identify SUMOylation sites in endogenous target proteins by mass spectrometry. This has hampered progress in the field since it is clear that a considerable percentage of SUMO targets are conjugated to SUMOs via nonconsensus sites for SUMOylation. More than 40% of the published yeast SUMO conjugation sites occur at nonconsensus lysine residues.

In contrast to ubiquitin, C-terminal tryptic fragments of SUMOs are very large and give rise to complex overlapping fragmentation spectra when analyzed by mass spectrometry. The C-terminal tryptic fragments of SUMO-2 and SUMO-3 are identical 32 amino acid peptides. The C-terminal tryptic fragment of SUMO-1 is considerably smaller and is 19 amino acids long. Several different approaches have been used for the identification of SUMOylation sites. Chymotrypsin treatment significantly reduces the size of the C-terminal SUMO fragment. The identification of SUMOylation sites is furthermore facilitated by introducing an artificial tryptic site through insertion of an arginine residue. It was shown that these mutations did not alter the behavior of SUMO-1. Related mutants were used in HeLa cells to identify 14 SUMO-1 conjugation sites in HEK293 cells to identify 17 SUMO-3 conjugation sites, and in A. thaliana to detect 17 SUMO-1 conjugation sites. A pattern recognition tool, SUMMOn, has been successful in detecting peptides modified by SUMO in vitro. We have used a targeted mass spectrometric approach combined with the linearization of the branched peptides to detect SUMO polymerization sites purified from cells. A database containing “linearized branched” peptides was employed by Hsiao et al. to detect SUMO-modified lysines, resulting in the identification of a single SUMO site in endogenous proteins purified directly from cells and 17 sites, including 8 sites...
on the SUMO E3 ligase RanBP2, after incubation of a cellular extract with ATP. Application of these strategies to in vivo samples has been limited by the very high complexity of the peptide mixture and very low abundance of SUMO conjugates. We have recently developed a novel method to purify and identify SUMO target proteins from cells in a site-specific manner \(157\) (Figure 5). The key innovation of our strategy is the use of a lysine-deficient SUMO mutant that is resistant to cleavage by the endopeptidase Lys-C, whereas virtually all other proteins contain lysines and are cleaved by Lys-C. After Lys-C treatment, intact SUMO mutants plus small fragments containing the modified lysines of target proteins are purified in a denaturing buffer, which excludes the copurification of SUMO-interacting proteins. This results in excellent purity of the final sample and a very low complexity due to removal of non-SUMOylated proteins and removal of non-SUMOylated tryptic fragments of target proteins. A disadvantage of lysine-deficient SUMOs is that these proteins no longer form SUMO chains. Therefore, the method is only suitable for studying mono-SUMOylation but not for studying poly-SUMOylation. However, the level of SUMO chains in cells under non-stress conditions appears to be very low.\(^{50,136}\) In addition, peak assignment was enabled by introducing an arginine at position 87 or 90, corresponding to arginines in \(S.\) cerevisiae SUMO (SMT3) or in ubiquitin, respectively (Figure 6). On the basis of the 103 SUMO conjugation sites that were identified (Figure 7), we have refined the SUMOylation consensus motif to [VILMFPC]KxE.\(^{157}\) Moreover, two novel SUMOylation consensus motifs were uncovered, including the inverted SUMOylation consensus motif [ED]xK[VILFP] and the hydrophobic-cluster SUMOylation motif (HCSM).  

**Figure 6.** Example of a SUMOylation site identified by mass spectrometry. Lysine 497 of nucleolar protein NOP5/58 was identified as a SUMO-2 acceptor site using the SUMO-2 mutant Q87R. The identified tryptic fragment contained a phosphorylated serine on position 502. The doubly modified peptide was identified by high-resolution tandem MS using higher energy collisional dissociation. Inset: Magnification of the low mass region showing the “QQ” signature fragment ions. Reprinted with permission from ref 157. Copyright 2010 Elsevier. 

### 4.4. Summary

Tagged forms of SUMO, SUMO traps, and anti-SUMO antibodies are used in the field to purify SUMO conjugates. Advantages and disadvantages of these methodologies are summarized in section 2.5. Site-specific identification of SUMOylation sites by mass spectrometry is very challenging due to the large C-terminal tryptic tags of mammalian SUMOs. The C-terminal tryptic tag of \(S.\) cerevisiae SUMO, SMT3, is much smaller and consists of the pentamer EQTGG. Introduction of arginines in the C-termini of mammalian SUMOs enables the mapping of SUMO acceptor lysines in target proteins. In the largest study, over 100 SUMOylation sites were identified in endogenous target proteins of SUMO-2.\(^{157}\) The most prominent SUMOylation consensus site in this data set was [VILMFPC]KxE. Site-specific methodologies will most likely be improved in the future to increase the coverage of SUMOylation sites. This will enable more detailed insight into SUMOylation dynamics in response to a wide variety of different stimuli. 

### 5. UBIQUITIN-LIKE PROTEIN NEDD8

#### 5.1. Overview

The Nedd8 protein has the highest homology with ubiquitin (58% identical) of all ubiquitin-like proteins and is conjugated to target proteins via a similar enzymatic cascade including the Nedd8-specific E1 complex UBE1C—NAE1 and the E2 enzyme UBE2M (Ubc12).\(^{196-198}\) This ubiquitin family member is an important regulator of Cullin-containing ubiquitin E3 ligases,\(^{199,200}\) including SCFs (Skp1, Cul-1, Roc1, and F-box proteins).\(^{197,201}\) Cullins are scaffold components of these multisubunit ubiquitin E3 ligase complexes and stimulate polyubiquitylation
and proteasomal degradation of important regulatory proteins such as cyclins.\(^{202-205}\) De-Neddylation of Cullins is mediated by the COP9 signalosome complex.\(^{206}\) Recently, the dynamic architecture of the Cullin—RING ubiquitin ligase network was elucidated using AQUA technology\(^{29,207}\) and that for Cul-1 by SILAC technology.\(^{26,27,208}\) Surprisingly, these studies revealed that Neddylation is not required for the interaction of Cullins and adapter proteins. Interestingly, Neddylation is deregulated in neurodegenerative diseases\(^{209,210}\) and cancer.\(^{210,211}\) Consequently, the Neddylation E1 inhibitor MLN4924 acts as a potent anticancer drug.\(^{212}\) MLN4924 functions as a cancer therapeutic because it inhibits Cullin—RING ubiquitin ligases by blocking Cullin Neddylation. The anticancer effect has been attributed to the deregulation of CDT1.\(^{213}\)

5.2. Nedd8 Proteomics

Three proteomics studies were published on Nedd8 target proteins using TAP-tagged Nedd8,\(^{214}\) GST—Nedd8,\(^{215}\) or FLAG—HA—Nedd8.\(^{209,205}\) As expected, Cullins were detected as prominent Nedd8 targets; Xirodimas et al. identified Cul-1, -2, -3, -4A, -4B, and -5.\(^{216}\) Bennett et al. identified seven Cullins, including Cul-7\(^{207}\) and Jones et al. identified all eight Cullin family members including Parc.\(^{215}\) Jones et al. furthermore detected other components of Cullin-containing ubiquitin ligases, including 12 F-box proteins, RBX1/ROC1 (RNF75) and Skp1, and identified target proteins that play important roles in DNA repair, transcription, and replication.\(^{217}\) Bennett et al. also identified 12 F-box proteins, S-phase kinase-associated proteins 1 and 2 (Skp1 and Skp2) and several BTB-domain-containing proteins and WD40-repeat-containing proteins that were present in FLAG—HA—Nedd8 IPs in an MLN4924-sensitive manner.\(^{207}\) BTB-domain proteins act as substrate-specific adapters in Cul-3-containing E3 ligase complexes.\(^{217,218}\) Xirodimas et al. identified a subset of ribosomal proteins as novel Nedd8 targets and showed that lack of Neddylation caused ribosomal protein instability.\(^{214}\) Interestingly, Jones et al. furthermore revealed for the first time that Nedd8 forms chains via lysines 11, 22, 48, and 60.

6. UBIQUITIN-LIKE PROTEIN ISG15

6.1. Overview

ISG15 contains two ubiquitin-like domains that are respectively 29% and 36% identical to ubiquitin. This protein was identified more than 30 years ago as an interferon-\(\alpha\) and -\(\beta\)-regulated protein.\(^{219}\) ISG15 has antiviral activity against HIV, influenza, Sindbis, and Ebola.\(^{220-223}\) Conjugation of ISG15 to target proteins is delayed until 18–24 h after stimulation with interferon\(^{-}\) due to a delay in the induction of the E1, E2, and E3 enzymes.\(^{223}\) These enzymes have been identified as Ube1L,\(^{226}\) UbcH8/Ube2L6,\(^{227,228}\) and the HECT-domain ligase Herc5,\(^{225,229}\) respectively.

6.2. ISG15 Proteomics

More than 300 proteins have been identified as targets for ISG15 in proteomics studies.\(^{229-233}\) The identified proteins comprise a rather heterogeneous group containing targets from many cellular compartments with no obvious enrichment for specific functional groups. Interestingly, they include 12 interferon-induced proteins. Important novel insight was obtained in a recent proteomics study, showing that ISG15 conjugation is restricted to newly synthesized proteins and includes exogenously expressed foreign proteins.\(^{234}\) Thus, ISG15 targets newly translated viral proteins in infected cells, and this is a cellular defense mechanism against viral infection as ISG15 modification disrupts virus particle assembly. Association of the E3 ligase Herc5 with polyribosomes explains how newly synthesized proteins are selected. Consistently, ISGylation of the L1 capsid protein of papillomavirus reduces the infectivity of the produced viruses.\(^{234}\)

7. OTHER UBIQUITIN-LIKE PROTEINS

7.1. Overview

Target proteins for ubiquitin-related modifier 1, prokaryotic ubiquitin-like protein, and small archaeal modifier proteins have been identified and are described in sections 7.2—7.4. Currently, little is known about the target proteins for the ubiquitin-like proteins UBI1,\(^{235}\) FAT10,\(^{236}\) and UFM1.\(^{237}\) The autophagy-related ubiquitin-like proteins Atg8 and Atg12 are part of the autophagosome.\(^{238}\) Atg12 is conjugated to a single target protein, Atg5,\(^{239}\) by an E1 enzyme Atg7,\(^{240}\) and an E2 enzyme, Atg10.\(^{241}\) The target for Atg8 is phosphatidylethanolamine, and conjugation involves Atg7 and the E2 enzyme Atg3.\(^{242}\) Human Atg8 family members interact with at least 67 proteins, probably in a non-covalent manner.\(^{243}\) Interestingly, the ubiquitin-like protein Hub1\(^{244}\) is thought to bind proteins exclusively in a non-covalent manner,\(^{245,246}\) since it lacks the typical C-terminal diglycine motif used by ubiquitin and ubiquitin-like proteins for covalent conjugation to target proteins.

7.2. Ubiquitin-Related Modifier 1

Ubiquitin-related modifier 1 (Urm1) was discovered in 2000 by Furukawa et al. as a protein showing sequence similarity to the bacterial sulfur carriers ThiS and MosA.\(^{247}\) Characteristically, Urm1 is a small protein of 99 amino acids with a C-terminal diglycine motif.\(^{248,249}\) In a subsequent yeast two-hybrid screen for Urm1 interactors, the Uba4 protein, related to the ubiquitin-activating E1 enzyme Uba1, was identified. Immunoblotting experiments revealed Urm1 conjugates that were formed in a Uba4-dependent manner.\(^{250,251}\) Until recently, the only known Urm1 substrate was the peroxiredoxin Ahp1.\(^{252}\) Recently, HA—Urm1 conjugates were purified from HeLa cells, and 21 novel potential Urm1 target proteins were identified by mass
spectrometry. Four target proteins were subsequently confirmed, including adenyltransferase and sulfurtransferase MOCS3 and ATP-binding domain-containing protein 3 (ATPBD3), two proteins that function in the Urmylation pathway, and additionally USP15 and the nucleocytoplasmic shuttling factor cellular apoptosis susceptibility protein. This study also confirmed that Ahp1 is conjugated to Urm1 via internal lysine residues. Urm1 has a double function; it not only functions as a protein modifier but fittingly with its similarity to bacterial sulfur carriers also acts as a sulfur carrier in thiolation of eukaryotic tRNA.

7.3. Prokaryotic Ubiquitin-like Proteins

Searches have recently revealed a ubiquitin-like protein in prokaryotes named prokaryotic ubiquitin-like protein. PUP was identified in Mycobacterium tuberculosis. Its structure is intrinsically disordered, but an α-helix is formed in PUP upon binding to the proteasomal ATPase subunit Mpa, revealing binding-induced folding. Subsequently, PUP is degraded together with its target proteins. These aspects are mechanistically different from those of the eukaryotic ubiquitin—proteasome system. Conjugation of PUP to target proteins, PUPylation, also proceeds via distinct chemistry and involves a deamidase and a glutamine synthetase-like ligase. Recently, 243 candidate PUPylation targets were identified by mass spectrometry using the purification of His-tagged PUP in the model organism Mycobacterium smegmatis. For 41 of these protein targets, PUP acceptor lysines were identified by mass spectrometry. Half of these target proteins are involved in intermediary metabolism and respiration pathways. Other identified PUP targets are involved in lipid metabolism, virulence, adaptation, and detoxification.

7.4. Small Archaeal Modifier Proteins

Archaea have long been known to express proteasomes that are similar to eukaryotic proteasomes. Recently, two small archaeal modifier proteins, SAMP1 and SAMP2, were identified in the archaean Haloferax volcanii. These proteins contain the characteristic ubiquitin β-grasp fold and contain the C-terminal diglycine motif. SAMPs are conjugated to target proteins in H. volcanii. Whereas SAMP1 conjugates were regulated by the proteasome, SAMP2 conjugates were not degraded by the proteasome. FLAG–SAMP conjugates were immunopurified and identified by mass spectrometry, revealing that SAMPylation is part of the sulfur metabolism, stress responses, basic transcription, translation, and DNA replication. Isopeptide bonds between the ε-amino group of lysines in target proteins and the C-terminal glycine of SAMP2 were confirmed by mass spectrometry and also chain formation of SAMP2 via internal lysine 58. SAMPs are activated in an ATP-dependent manner by an E1-like enzyme termed E1-like SAMP activator (ELSA). SAMPs are members of a large superfamily that includes members from all major archaeal lineages, indicating that ubiquitin-like protein modification plays an important role in archaea.

8. CROSSTALK BETWEEN POST-TRANSLATIONAL MODIFICATIONS

The ultimate goal of PTM proteomics is to understand how different PTMs cooperatively regulate cellular processes. Intriguing examples of PTM crosstalk and competition have been published between ubiquitylation and other modifications (Figure 8). Several examples of cross-regulation between different modifications are described in the next section.

8.1. Crosstalk between Phosphorylation and Ubiquitylation

A classical example of crosstalk between phosphorylation and ubiquitylation occurs in the NFκB signaling cascade.
transcription factor family consists of dimers of transcriptional activators that can be sequestered in the cytoplasm bound to members of the IκB family of inhibitors. Upon activation by TNFα, IL-1β, and a large number of other stimuli, the IKK kinase complex phosphorylates IκB on serines 32 and 36. This enables docking of the ubiquitin E3 ligase complex consisting of Skp1, Cul-1, Roc1, and βTrCP that ubiquitylates IκB on adjacent lysine residues 21 and 22, leading to the destruction of the inhibitor protein by the proteasome and translocation of the NFκB dimer to the nucleus to activate transcription. Interestingly, lysine 21 of IκBα can also be SUMOylated to protect the protein from degradation via the ubiquitin–proteasome pathway. βTrCP is furthermore responsible for the ubiquitylation of other proteins, including β-catenin, via its WD40 repeats, which bind the D- (phospho)SGXX(phospho) S degradation motif in target proteins. Many other SCF ligases are recruited to substrates in a phosphorylation-dependent manner. Nevertheless, we are still limited in our understanding of phosphorylation-dependent protein ubiquitylation.

Crosstalk furthermore includes phosphorylation of many components of the ubiquitylation machinery, including the E1 enzyme, Ube2_ J1, E3 ligases NEDD4-2, WWP1, UBR2, UBR5, MIB2, and MYCBP2, E4 B, and ubiquitin proteases USP31, USP39, USP42, UCH8, UCH10, UCH20, UCH24, UCH32, UCH34, UCH37, CYLD, FAF-X, FAF-Y, and BAP1. Vice versa, Xu et al. and Danielsen et al. described ubiquitylation of kinases and phosphatases, including cAMP-dependent protein kinase type Iα regulatory subunit, casein kinase II subunit α, serine/threonine-protein kinases 12, cAMP-dependent protein kinase type Iα, janus kinase 1, AXL receptor tyrosine kinase, serine/threonine-protein kinase SMG1, receptor-type tyrosine-protein phosphatase F, and serine/threonine protein phosphatase 2A 65 kDa regulatory subunit A and catalytic subunit.

8.2. Phosphorylation-Dependent SUMOylation

More recently, crosstalk between protein phosphorylation and SUMOylation has been uncovered. A phosphorylation-dependent SUMOylation motif (PDSM) was identified in heat shock factors HSFI and HSF4β, myocyte-specific enhancer factor 2, and erythroid transcription factor GATA-1. Recently, five proteins were identified by mass spectrometry that are simultaneously phosphorylated and SUMOylated with a preferred spacing between the SUMOylated lysine and the phosphorylated serine of four residues. Modification of NOP5/S8 by SUMOylation and phosphorylation is shown in Figure 6. Mutagenesis experiments confirmed that phosphorylation of the nucleolar protein NOP5/S8 enhanced SUMOylation. Mechanistically, this can be explained by increased binding of some phosphorylated proteins to the SUMO E2 enzyme Ube2I. Interestingly, negatively charged residues in the negatively charged amino acid-dependent SUMOylation motif (NDSM) can partly replace phosphorylated serines to enhance the SUMOylation levels of target proteins.

8.3. Competition between Different Lysine Modifications

Ubiquitin and UBL proteins are covalently linked to lysines in target proteins. Since lysines are also subjected to other modifications, including acetylation and methylation, this creates the potential for competition between different lysine modifications. Interestingly, four SUMOylated lysines in RanGAP1, bromodomain-containing protein 4, scaffold attachment factor B2, and Treacle protein TCOF1 were previously identified as acetylated lysines. D Dephosphorylation of a PDSM motif in the transcriptional regulator myocyte-specific enhancer factor 2A was shown to regulate a switch from SUMOylation to acetylation to control posttranslational differentiation. Competition between ubiquitylation and acetylation appears to be a very frequent event, with over 20% of the ubiquitin acceptor lysines also being reported as sites of acetylation; e.g., all the ubiquitin acceptor lysines in H2B, H3.1, and H4 were also reported to be acetylated.

8.4. Crosstalk between SUMOylation and the Ubiquitin–Proteasome System

Intriguingly, a subset of SUMO-2/3 conjugates are subsequently ubiquitylated and degraded. Our proteomics study showed that at least 10% of the SUMO-2 target proteins are regulated via this type of crosstalk, and this is critical for SUMO-2/3 recycling, indicating that the ubiquitin–proteasome system is an essential component of the SUMO cycle. Mechanistically, this involves specific ubiquitin ligases that contain SUMO interaction motifs that interact with SUMO in a noncovalent manner.

9. CONCLUSIONS AND FUTURE PERSPECTIVES

Post-translational modifications of proteins are extremely complex. Compared to the robust methodology available to study phosphorylation and acetylation, the ubiquitin field is still lagging behind. Site-specific identification of protein ubiquitylation and ubiquitin-like modification is still challenging. Currently, the field is aiming to develop novel methods to identify these modifications in a site-specific manner. These enabling techniques will open up unprecedented possibilities to decipher protein ubiquitylation and UBL modification on a proteome-wide scale. Furthermore, the field will benefit from improvements in peptide fragmentation enabled by HCD and improved sensitivity and mass accuracy of mass spectrometers and software to deal in an automated fashion with the assignment of modification sites in peptides and quantification to detect changes in modification levels.

The dissection of crosstalk between different modifications at a proteome-wide level will also be important. Searching MS/MS spectra for co-modifications, including a larger number of different modifications, is very challenging due to the steep increase in computation time. Systematic, unbiased, and proteome-wide studies on protein ubiquitylation and UBL conjugation will strongly improve our understanding of these modifications in the near future, particularly with respect to the identification of target proteins for the vast numbers of ligases and proteases in the system and to identify the targets that are dynamically regulated in response to a wide variety of stimuli. This is furthermore relevant for the detailed understanding of deregulation of ubiquitin signaling in human diseases and for the development of drugs that target components of ubiquitin signaling cascades.
Alfred Vertegaal received his Ph.D. from the Leiden University, The Netherlands, in 2001 on the basis of a research project supervised by Prof. Dr. Alex van der Eb and Dr. Alt Zantema. Subsequently, he joined the laboratory of Prof. Dr. Angus Lamond at the Wellcome Trust Biocentre in Dundee, U.K., where he initiated a project to purify and identify novel target proteins for small ubiquitin-like modifiers (SUMOs). He became a group leader in 2008 in the Department of Molecular Cell Biology at the Leiden University Medical Center. His group focuses on signal transduction via SUMOs.

ACKNOWLEDGMENT
This work was supported by The Netherlands Organisation for Scientific Research (NWO). I thank Peter ten Dijke, David Baker, Karolin Eifler, Ivo Hendriks, and Joost Schimmel for helpful discussions and suggestions.

REFERENCES

(1) Garavelli, J. S. Proteomics 2004, 4, 1527.
(2) Aebersold, R.; Goodlett, D. R. Chem. Rev. 2001, 101, 269.
(3) Choudhary, C.; Mann, M. Nat. Rev. Mol. Cell Biol. 2010, 11, 427.
(4) Ding, H.; Xu, Y.; Chen, Q.; Dai, H.; Tang, Y.; Wu, J.; Shi, Y. Biochemistry 1998, 273, 34983.
(5) Namasivayam, J.; Wang, M.; Fu, Z.; Klein, J. M.; Haas, A. L.; Kim, J. J. Biol. Chem. 2005, 27, 27356.
(6) Ramelot, T. A.; Cort, J. R.; Yee, A. A.; Semesi, A.; Edwards, A. M.; Arrowsmith, C. H.; Kennedy, M. A. J. Struct. Funct. Genomics 2003, 4, 25.
(7) Xu, J.; Zhang, J.; Wang, L.; Zhou, J.; Huang, H.; Wu, J.; Zhong, Y.; Shi, Y. Proc. Natl. Acad. Sci. U.S.A. 2006, 103, 11625.
(8) Schulman, B. A.; Harper, J. W. Nat. Rev. Mol. Cell Biol. 2009, 10, 319.
(9) Wenzel, D. M.; Stoll, K. E.; Klevit, R. E. Biochem. J. 2010, 433, 31.
(10) van Wijk, S. J.; Timmers, H. T. FASEB J. 2010, 24, 981.
(11) Nagy, V.; Dikic, I. Mol. Biol. 2010, 391, 163.
(12) Rotin, D.; Kumar, S. Nat. Rev. Mol. Cell Biol. 2009, 10, 398.
(13) Deshaies, R. J.; Joao, C. A. Annu. Rev. Biochem. 2009, 78, 399.
(14) Komander, D.; Clague, M. J.; Urbe, S. Nat. Rev. Mol. Cell Biol. 2009, 10, 550.
(15) Nijman, S. M.; Luna-Vargas, M. P.; Velds, A.; Brummelkamp, T. R.; Dirac, A. M.; Sixma, T. K.; Bernards, R. Cell 2005, 123, 773.
(16) Ventii, K. H.; Wilkinson, K. D. Biochem. J. 2008, 1416, 161.
(17) Reyes-Turcu, F. E.; Ventii, K. H.; Wilkinson, K. D. Annu. Rev. Biochem. 2009, 78, 363.
(18) Ikeda, F.; Dikic, I. EMBO Rep. 2008, 6, 536.
(19) Cooper, E. M.; Cutcliffe, C.; Kristiansen, T. Z.; Pandey, A.; Pickart, C. M.; Cohen, R. E. EMBO J. 2009, 28, 621.
(20) Ziv, I.; Matiuhin, Y.; Kirkpatrick, D. S.; Erpazapoglou, Z.; Leon, S.; Pantazopoulou, M.; Kim, W.; Gygi, S. P.; Haggauer-Tsipas, R.; Reis, N.; Glickman, M. H.; Kleifeld, O. Mol. Cell. Proteomics 2010, 10, M110.009753.
(21) Sigismund, S.; Polos, S.; Di Fiore, P. P. Curr. Top. Microbiol. Immunol. 2004, 286, 149.
(22) Grabbe, C.; Dikic, I. Chem. Rev. 2009, 109, 1481.
(23) Winget, J. M.; Mayor, T. Mol. Cell 2010, 38, 627.
(24) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Mol. Cell. Proteomics 2002, I, 376.
(25) Ong, S. E.; Mann, M. Nat. Chem. Biol. 2005, 1, 252.
(26) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Mol. Cell. Proteomics 2004, 3, 1154.
(27) Choudhary, C.; Mann, M. Nat. Biotechnol. 2005, 23, 252.
(28) Ramelot, T. A.; Cort, J. R.; Yee, A. A.; Semesi, A.; Edwards, A. M.; Arrowsmith, C. H.; Kennedy, M. A. J. Struct. Funct. Genomics 2003, 4, 25.
(29) Ding, H.; Xu, Y.; Chen, Q.; Dai, H.; Tang, Y.; Wu, J.; Shi, Y. Biochemistry 1998, 273, 34983.
(30) Bier, J. M.; Stoll, K. E.; Klevit, R. E. Biochem. J. 2010, 433, 31.
(174) Hannich, J. T.; Lewis, A.; Kroetz, M. B.; Li, S. J.; Heide, H.; Emili, A.; Hochstrasser, M. J. Biol. Chem. 2005, 280, 4102.
(175) Denison, C.; Rudner, A. D.; Gerber, S. A.; Bakalarski, C. E.; Moazed, D.; Gygi, S. P. Mol. Cell Proteomics 2005, 4, 246.
(176) Ganesan, A. K.; Kho, Y.; Kim, S. C.; Chen, Y.; Zhao, Y.; White, M. A. Proteomics 2007, 7, 2216.
(177) Haas, H. H.; Meulmeester, E.; Frank, B. T.; Melchior, F.; Urlaub, H. Mol. Cell Proteomics 2009, 8, 2664.
(178) Grant, M. M. BMB Rep. 2010, 43, 720.
(179) Bruderer, R.; Tatham, M. H.; Plechanova, A.; Matic, I.; Garg, A. K.; Hay, R. T. EMBO Rep. 2011, 12, 142.
(180) Gill, G. Curr. Opin. Genet. Dev. 2005, 15, 536.
(181) Girdwood, D.; Bumpass, D.; Vaughan, O. A.; Thain, A.; Anderson, L. A.; Snowden, A. W.; Garcia-Wilson, E.; Perkins, N. D.; Hay, R. T. Mol. Cell 2003, 11, 1043.
(182) Girdwood, D. W.; Tatham, M. H.; Hay, R. T. Semin Cell Dev Biol. 2004, 15, 201.
(183) Kuo, H. Y.; Chang, C. C.; Jeng, J. C.; Hu, H. M.; Lin, D. Y.; Mau, G. G.; Kwok, R. P.; Shi, H. M. Proc Natl. Acad. Sci. U.S.A. 2005, 102, 16973.
(184) Lin, D. Y.; Fang, H. I.; Ma, A. H.; Huang, Y. S.; Pu, Y. S.; Jenster, G.; Kung, H. J.; Shi, H. M. Mol. Cell. Biol. 2004, 24, 10529.
(185) Lin, D. Y.; Huang, Y. S.; Jeng, J. C.; Kuo, H. Y.; Chang, C. C.; Chao, T. T.; Ho, C. C.; Chen, Y. C.; Lin, T. P.; Fang, H. I.; Hung, C. C.; Suen, C. S.; Hwang, M. J.; Chang, K. S.; Mau, G. G.; Shi, H. M. Mol. Cell. Biol. 2006, 24, 3411.
(186) Yang, S. H.; Sharrocks, A. D. Mol. Cell. Biol. 2004, 13, 611.
(187) Dou, H.; Huang, C.; Van Nguyen, N. T.; Lu, L. S.; Yeh, E. T. FEBS Lett. 2011, 585, 2891.
(188) Saitoh, H.; Hinchee, J. J. Biol. Chem. 2000, 275, 6252.
(189) Jeram, S. M.; Srikumar, T.; Pedrioli, P. G.; Raught, B. Proteomics 2009, 9, 922.
(190) Schimmelpfeng, J.; Balog, C. I.; Deelder, A. M.; Drijfhout, J. W.; Hensbergen, P. J.; Vertegaal, A. C. J. Proteomics 2010, 73, 1523.
(191) Ryu, H.; Furuta, M.; Kirkpatrick, D.; Gygi, S. P.; Azuma, Y. J. Cell Biol. 2010, 191, 783.
(192) Kruiswelm, M.; Cheung, H. T.; Hamady, M.; Barthel, K. K.; Liu, X. Mol. Cell Proteomics 2005, 4, 1626.
(193) Wohlshleger, J. A.; Johnson, E. S.; Reed, S. I.; Yates, J. R., III. J. Proteome Res. 2006, 5, 761.
(194) Blomster, H. A.; Imanishi, S. Y.; Siimes, J.; Kastu, J.; Morrice, N. A.; Eriksson, J. E.; Sistonen, L. J. Biol. Chem. 2010, 285, 19324.
(195) Pedrioli, P. G.; Raught, B.; Zhang, X. D.; Rogers, R.; Aitchison, J.; Matunis, M.; Aebi, M.; Rout, M. J. Cell Biol. 2003, 3, 553.
(196) Watson, I. R.; Irwin, M. S.; Ohm, M. Cancer Cell 2011, 19, 168.
(197) Merlet, J.; Burger, J.; Gomes, J. E.; Pintard, L. Cell. Mol. Life Sci. 2009, 66, 1924.
(198) Rabut, G.; Peter, M. EMBO Rep. 2008, 9, 969.
(199) Osaka, F.; Kawasaki, H.; Aida, N.; Saeki, M.; Chiba, T.; Kawashima, S.; Tanaka, K.; Kato, S. Genes Dev. 1998, 12, 2263.
(200) Hori, T.; Osaka, F.; Chiba, T.; Miyamoto, C.; Okabayashi, K.; Shimbara, N.; Kato, S.; Tanaka, K. Oncogene 1999, 18, 6829.
(201) Petroski, M. D.; Deshaies, R. J. Nat. Rev. Mol. Cell Biol. 2005, 6, 9.
(202) Tateishi, K.; Otama, M.; Tanaka, K.; Chiba, T. J. Cell Biol. 2001, 155, 571.
(203) Morimoto, M.; Nishida, T.; Honda, R.; Yasuda, H. Biochem. Biophys. Res. Commun. 2000, 270, 1093.
(204) Read, M. A.; Brownell, J. E.; Gladyshewa, T. B.; Hottelet, M.; Parent, L. A.; Coggins, M. B.; Pierce, J. W.; Podust, V. N.; Luo, R. S.; Chau, V.; Palombella, V. J. Mol. Cell. Biol. 2000, 20, 2326.
(205) Singer, J. D.; Gurian-West, M.; Clurman, B.; Roberts, J. M. Genes Dev. 1999, 13, 2375.
(206) Lyapina, S.; Cope, G.; Shevchenko, A.; Serino, G.; Tsuge, T.; Zhou, C.; Wolf, D. A.; Wei, N.; Shevchenko, A.; Deshaies, R. J. Science 2001, 292, 1382.
(207) Bennett, E. J.; Rush, J.; Gygi, S. P.; Harper, J. W. Cell 2010, 143, 951.
NOTE ADDED IN PROOF

After the acceptance of this paper, three landmark papers were published using monoclonal antibodies that recognize diglycine remnants as described in figure 3, and ~10000, ~11000, and ~19000 sites were identified respectively.287–289 These results indicate that the complexity of protein ubiquitylation could be comparable to the complexity of protein phosphorylation and that site-specific ubiquitylation studies at a proteome-wide level are now feasible. Wagner et al. discovered a non-proteasomal function for almost half of all identified diglycine sites.288 The paper by Kim et al. highlights that a very significant fraction of ubiquitin conjugates results from freshly translated proteins and that ubiquitylation is frequently a sub-stoichiometric event.289

Emanuele et al. and Kim et al. used this technology to identify substrates for Cullin-RING ligases.287,289 Their work underscores the diversity of ubiquitylation substrates, including those that are not degraded by the proteasome.288,289 The usage of diglycine-specific antibodies has been instrumental in understanding the complex web of ubiquitylation reactions and their roles in cellular processes.288,289 Therefore, the detection of diglycine residues not only aids in the identification of ubiquitylation sites but also provides insights into the specificity and regulation of ubiquitylation processes.288,289