Chapter 1

Introduction

The fundamental difference between lower and higher life forms is the ability for individual cells to interact and communicate with each other, forming a multicellular organism. The most complex organisms consist of a multitude of organs and tissues, which function in concert to support the body as a whole. However, while these differences are obvious at the surface, there is also a high degree of specialization within the individual cells that make up the entire organism. In general, the further up the evolutionary ladder, the more complex cells are. This complexity stems from increased cellular size, a wider range of cellular components, specialized cellular functionality, and increased ability to adapt to the environment. To fuel all of this potential, a larger and more complex genome is required.

The human genome contains over 20,000 genes [1], encoding proteins which possess widely different properties and functions. The structure, concentration and localization of these proteins are the main drive behind virtually all biological processes and molecular mechanisms. Depending on the cellular circumstances, different genes are transcribed and messenger ribonucleic acid (mRNA) is generated [2, 3]. Due to the effects of alternative promoters [4], alternative splicing [5], and editing of mRNA [6, 7], the 20,000 genes can be transcribed into well over 100,000 different transcripts [8, 9]. Subsequently, ribosomes read the nucleotide-based mRNA transcripts, and translate the contained information by coupling amino acids into proteins; essentially huge polypeptides [10, 11]. While this provides the basic framework for the regulation of all cellular functions, an additional level of complexity is required in order to provide all life with its ability to rapidly adapt to environmental changes.

Post-translational modifications

To further expand upon the functional repertoire of the proteome, many proteins are subject to post-translational modification (PTM) [12, 13]. These modifications can alter the biochemical properties of a protein, and regulate its biological activity (**Figure 1**)[14, 15]. For example, a protein may be able to interact with a partner protein after being modified by a PTM, and on the other hand, such a modification may abolish interaction between two proteins [16-18]. As a result, a protein's localization within the cell may be altered after attachment of a PTM [19, 20]. Conversely, a protein may also be subject to PTM depending on its localization in the cell [21]. The presence of a PTM can change the structural properties of a protein, altering its biological activity [22]. Additionally, certain PTMs can mark a protein for destruction by the proteasome [23]. Certain PTMs may sequentially modify target proteins, relying on each other's presence, which is commonly referred to as crosstalk [24]. Conversely, other PTMs may compete over the same modification site [25].

Mature proteins may be modified in numerous different ways, and many hundreds of different PTMs exist [26](pir.georgetown.edu/resid). Some of the most abundant and well-known modifications of proteins involve various biochemical functional groups, such as acetyl [27, 28], phosphate [29, 30], methyl [31, 32],

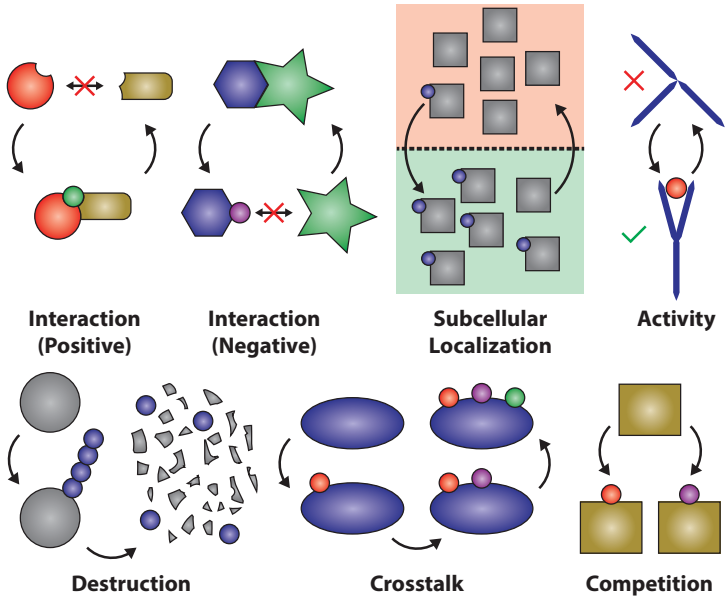


Figure 1. Schematic examples of PTMs influencing the functionality of proteins.
PTMs can potentiate an interaction between two proteins, or alternatively prevent or disrupt an interaction. Modification of a protein may predispose it to localization towards another part of the cell, or alternatively a protein may be modified depending on its subcellular localization. PTMs can directly affect the structure of a protein, and alter its activity. Proteins modified by PTMs may be marked for degradation. In case of crosstalk, different types of PTMs modify the same protein, in a sequential fashion. PTMs may also compete with each other for the same modification site, differentially regulating the protein they modify.

glycosyl [33, 34] and adenosine diphosphate (ADP) ribose [35, 36]. Additionally, proteins may be modified by covalent linkage of small modifier proteins, occurring primarily in eukaryotes. These include ubiquitin [37, 38] and other ubiquitin-like family members (Ubls) [39], which, in humans, include SUMO-1, SUMO-2, SUMO-3 [40-42], Neural precursor cell expressed developmentally down-regulated protein 8 (NEDD8) [43, 44], Ubiquitin D (FAT10) [45, 46], Interferon-induced 15 kDa protein (ISG15) [47, 48], Ubiquitin-fold modifier 1 (UFM1) [49, 50] and Ubiquitin-like protein FUBI (FAU) [51, 52]. Even though the Ubls often have a widely differing amino acid sequence, they all share similar structural properties (**Figure 2**).

Modification of proteins by virtually all modifiers is a reversible process, and dedicated enzymes exist which can reverse the modification process. Surprisingly, the amount of enzymes responsible for reversal of a PTM may be as numerous as the amount of enzymes involved in the initial modification, with over a dozen of demethylases being responsible for reversal of methylation [54], and almost one-hundred deubiquitylating enzymes taking charge of ubiquitylation removal [55, 56]. The cellular ability to edit proteins in such a fashion provides limitless potential for dynamic regulation. Ultimately, the amount of complexity and heterogeneity

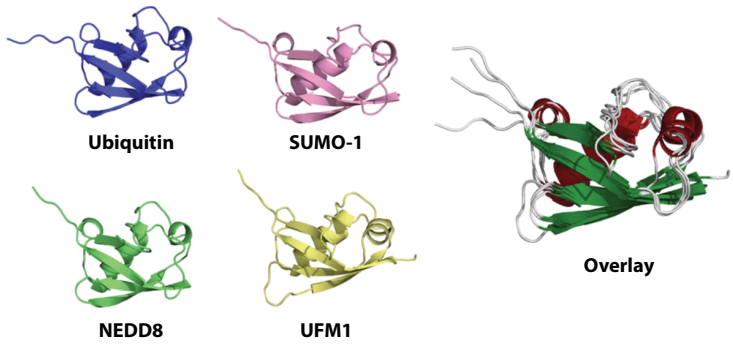


Figure 2. Structural comparison of various Ubl proteins.
Ubiquitin, SUMO-1, NEDD8 and UFM1 all share the same ubiquitin β-gasp fold, despite differences in amino acid sequence. When overlaying the 4 Ubls, their structural similarity becomes obvious. α-helices are indicated in red, β-sheets are indicated in green. Image adapted from Ha and Kim, 2008 [53].

within the cellular system is vastly expanded by the existence of PTMs, which allow for fine-tuning of many pivotal cellular processes (**Figure 3**).

Small ubiquitin-like modifier (SUMO)

Small ubiquitin-like modifiers (SUMOs) are a ubiquitin-like modifier, possessing the characteristic ubiquitin globular β-grasp fold [57], and have been reported to modify in the range of hundreds of proteins in mammalian cells [58, 59]. SUMO has been implicated in the regulation of many cellular functions, ranging from transcriptional regulation and chromatin remodeling to deoxyribonucleic acid (DNA) repair and control of cell cycle progression (**Figure 4**)[60-64].

Like other Ubls, SUMOs are covalently attached to lysines in target proteins through an isopeptide bond, chemically linking the carboxyl-terminal di-glycine to the ε-amino groups of lysines in target proteins (**Figure 5**). The enzymatic cascade responsible always involves the dimeric E1 activating enzymes SUMO-Activating Enzyme Subunit 1 and 2 (SAE1/2) and the E2 conjugation enzyme Ubiquitin Carrier Protein 9 (Ubc9) [66-70]. The E1 is responsible for activation of the SUMO carboxyl-terminus through means of adenosine triphosphate (ATP) hydrolysis, where SUMO is adenylated and coupled to SAE2 through a thioester bond. Ubc9 plays an important role in targeting of SUMOylation to the intended sites, and enables activated SUMO to reach and couple to its target. Whereas the presence of E1 and E2 is sufficient for SUMOylation of target proteins *in vitro* [71], a number of E3 ligases exist, which confer context-specificity and greatly enhance efficiency of the SUMO conjugation [60, 72-75]. Unlike other PTMs, SUMOs display an extent of specificity in their conjugation, and are often targeted to the canonical consensus motif [VIL]KxE [76, 77], although SUMOylation can occur on alternate or non-consensus motifs [78, 79].

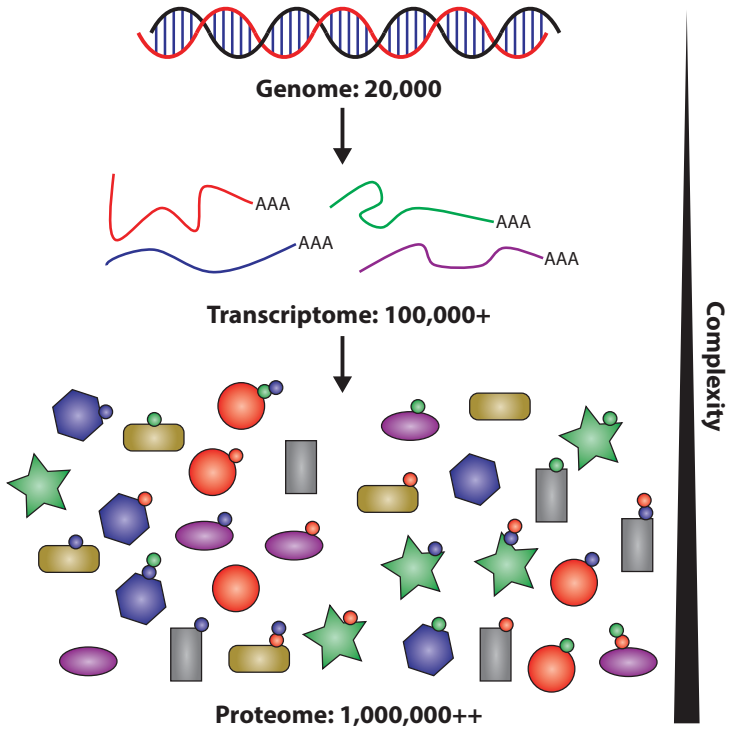


Figure 3. Complexity of the proteome.
The human genome includes 20,000 genes, which are transcribed into over 100,000 different transcripts. The transition from the genome to the transcriptome significantly increases cellular complexity. After translation into proteins, PTMs inflate the system's complexity by another order of magnitude, with over a million of differentially modified proteins existing within the cell.

SUMOylation of proteins is a reversible process, since SUMO-specific proteases are capable of efficiently removing SUMO from target proteins [80, 81]. SUMO-specific proteases are also essential for the maturation of SUMO, cleaving additional residues off the SUMO carboxyl-terminus and exposing the di-glycine. All in all, the conjugation and deconjugation of SUMO allows cells to rapidly and dynamically respond to a wide range of cellular stresses and growth conditions [82, 83].

In humans, SUMO is often classified into two families, SUMO-1 and SUMO-2/3. Mature SUMO-2 and SUMO-3 are virtually identical [84], only differing by three amino acids in the amino-terminus, including a serine in SUMO-2, although there is no evidence for amino-terminal phosphorylation of SUMO-2. No functional difference between the two has ever been reported, and no antibody exists that can differentiate between them. Conversely, SUMO-1 is only 47% similar to SUMO-2/3, and can be classed as a truly separate PTM. Surprisingly, all forms of SUMO are conjugated by the same machinery, and while an overlap exists between their targets

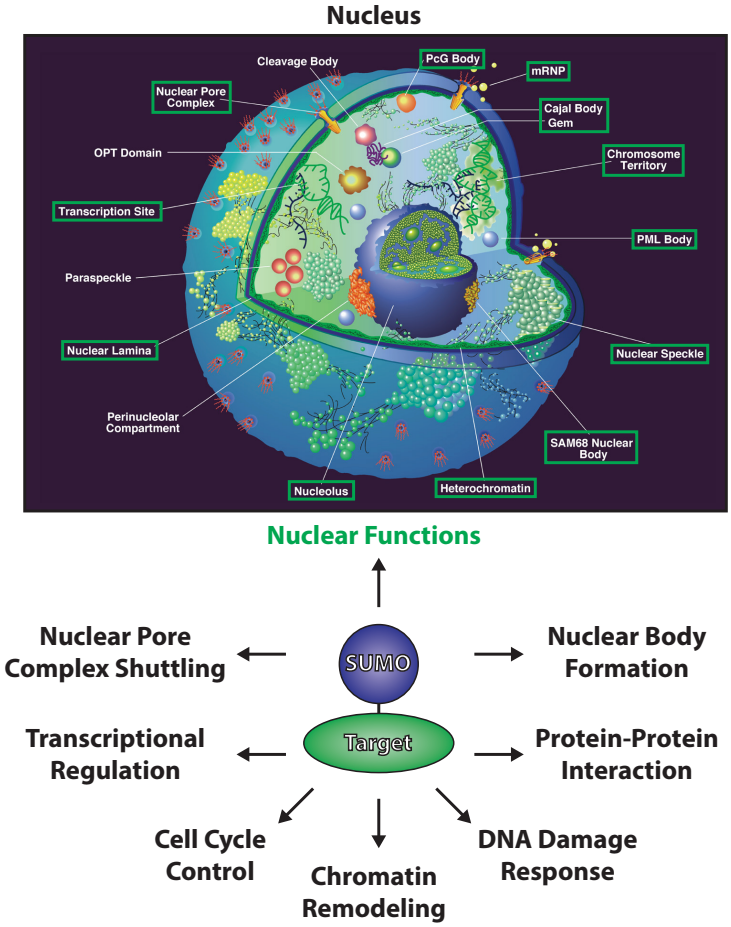


Figure 4. An overview of SUMO's involvement in key cellular processes.
SUMOylation of proteins has widespread regulatory roles within the cell, including nuclear pore complex shuttling, transcriptional regulation, cell cycle control, chromatin remodeling, the DNA damage response, protein-protein interactions, formation of nuclear bodies, and various other nuclear functions (SUMO is involved at least in the processes boxed in green). Nucleus image adapted from Spector, 2001 [65].

of conjugation, there are also differences [85, 86]. Whereas SUMO-1 is predominantly conjugated to Ran GTPase-Activating Protein 1 (RanGAP1), SUMO-2/3 is more dynamic and much more abundant [87], with the free unconjugated pool of SUMO-2/3 functioning as a reservoir to allow the cell to respond to environmental changes and cellular stresses [58, 88, 89]. In fact, it could be argued that one of the main purposes of SUMO-2/3 is to provide the cell with this adaptability mechanism. SUMO, like ubiquitin, is able to modify itself, forming polymeric chains [90-92]. The formation of these SUMO chains is substantiated by cellular stress conditions such as heat shock [58]. In humans, SUMO-2/3 is primarily modified on lysine-11,

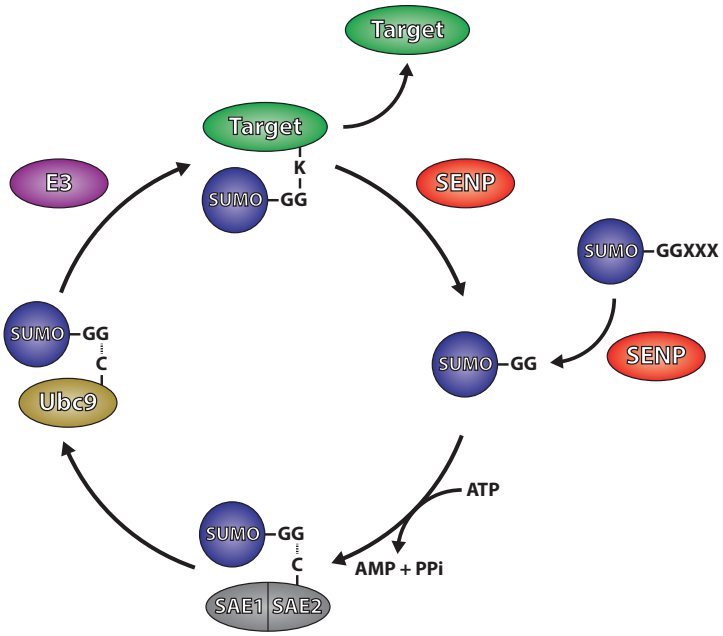


Figure 5. The SUMO cycle.
Precursor SUMO is matured by exposure of its di-glycine motif through the action of SUMO-specific proteases. Through an enzymatic cascade of the activating E1 enzyme SAE1/2, the conjugating E2 enzyme Ubc9, and optionally the involvement of a catalytic E3 enzyme, SUMO is conjugated to a lysine residue within the target protein. SUMOylation is a reversible process, since SUMO-specific proteases may remove SUMO from a protein, freeing up the SUMO for re-conjugation.

although other internal lysine residues may be targeted for chain formation *in vitro* [93]. With SUMO-1 lacking an internal consensus motif, it is considered to be inefficiently SUMOylated, and thus serves as a chain terminator [90, 91, 93]. Polymeric SUMO chains are important in replication, proteasomal degradation, and during the cell cycle [91]. In yeast, SUMO chains are important for maintenance of higher-order chromatin structure [94].

SUMO is indispensable for eukaryotic life

All eukaryotes express at least one family member of SUMO, with most mammals expressing three copies, and up to eight copies exist in *Arabidopsis thaliana* [95]. While in some cases there is functional redundancy between multiple SUMO copies, the process of SUMOylation is essential for nearly all eukaryotic life [60, 96]. Only some forms of yeast [97] and fungi [98] are able to survive without SUMO, while still displaying severe growth defects. The conjugating enzyme Ubc9 is essential for viability of higher eukaryotic, with its depletion leading to chromosomal damage and induction of apoptosis [99]. Furthermore, the importance of the SUMO pathway was clearly demonstrated by mice deficient for Ubc9, which perish at the

early post-implantation stage due to aggravated defects in chromosome condensation and segregation [100]. In *Arabidopsis thaliana*, knockdown of the activating enzyme SAE2, the SUMO Conjugating Enzyme 1 (SCE1), or double knockdown of SUMO1 and SUMO2 are all embryonic lethal, further underlining the importance of SUMO [101].

SUMOylation has become increasingly implicated as a major player in carcinogenesis [102-104], and various key factors involved in cancer are known to be functionally SUMOylated, such as Hypoxia-Inducible Factor 1-alpha (HIF1α) [83], Hypoxia-Inducible Factor 2-alpha (HIF2α) [105], and Cellular Tumor Antigen p53 (p53) [106]. SAE1 and SAE2 were identified in a screen for Myc Proto-Oncogene Protein (Myc) synthetic lethal genes, showing that Myc-driven tumors are reliant on SUMOylation [107]. Thus, SUMO has been gaining popularity as a therapeutic target, with its clinical involvement ranging from cancer to Alzheimer's disease [60, 108, 109].

SUMO, the guardian of the nucleus

When comparing the machinery involved in the conjugation of SUMO to that of ubiquitin, which employs hundreds of enzymes for effective conjugation [110, 111], it becomes clear that SUMO employs only a small subset of enzymes in order to modify its targets. Furthermore, whereas ubiquitin, acetylation and phosphorylation modify proteins all throughout the cell, SUMOylation is an event that occurs predominantly in the nucleus [112, 113]. Within the nucleus, there is a further enrichment of SUMOylation that occurs at the chromatin [114-116] and in nuclear bodies [117, 118]. These bodies are chiefly Promyelocytic Leukemia (PML) bodies, and SUMO plays a critical role in the regulation of this subcellular domain [117, 119, 120], which are clusters containing many factors important in the cellular response to DNA damage and various stresses [121, 122].

As SUMO is intrinsically focused in the nucleus and around the DNA, it provides an effective method for regulation of proteins that are involved in the cellular response to DNA damage. Not surprisingly, the coordination of the DNA damage response by SUMO has been subject of extensive study over the last decade [61, 63, 123]. There have been multiple reports on the regulation of single DNA damage response proteins by SUMOylation, including G/T mismatch-specific thymine DNA glycosylase (TDG) [22], Proliferating cellular nuclear antigen 1 (PCNA) [124, 125], Breast cancer type 1 susceptibility protein (BRCA1) [126, 127], and Mediator of DNA damage checkpoint protein 1 (MDC1) [128-130]. Entire functional clusters of proteins may also be modified in concert by SUMOylation, for example in order to orchestrate an efficient response to DNA damage [21, 131].

SUMO and ubiquitin in concert

While one PTM by itself is innately interesting, the combination of two or more PTMs acting together becomes highly intriguing. Crosstalk between PTMs allows

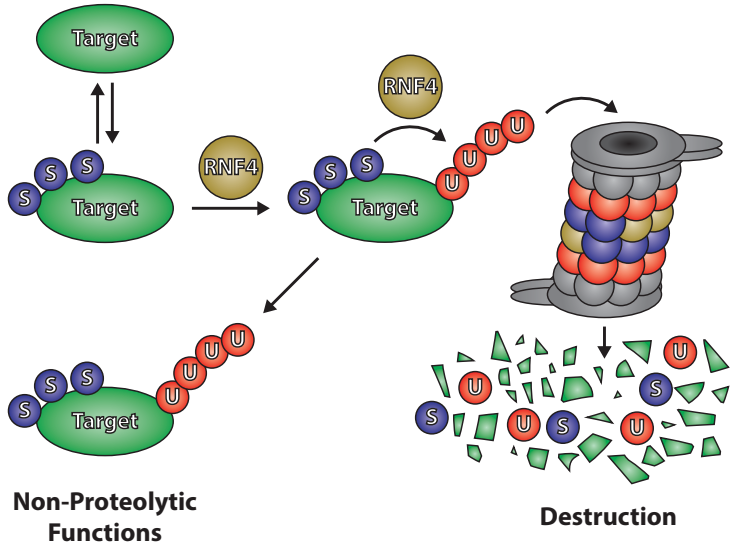


Figure 6. SUMO-Targeted Ubiquitin Ligases.
STUbLs, such as RNF4, recognize poly-SUMOylated target proteins, and may subsequently ubiquitylate these proteins. One common result of poly-ubiquitylation of proteins is targeting to the proteasome and subsequent proteolytic destruction. Alternatively, the combined SUMOylated and ubiquitylation of a protein may have non-proteolytic functions.

virtually limitless combinations in regulatory potential, but moreover allows for multiple cellular functionalities to be directly connected to each other. One such example are the SUMO-targeted Ubiquitin Ligases (STUbLs), which are a subset of ubiquitin E3 ligases that specifically recognize and ubiquitylate SUMOylated proteins [132-135].

In yeast, the E3 ubiquitin-protein ligase complex SLX5-SLX8 subunit SLX5b (Slx5) and E3 ubiquitin-protein ligase complex SLX5-SLX8 subunit SLX8 (Slx8) proteins form a heterodimeric ubiquitin ligase, which specifically recognizes SUMOylated proteins. Deletion of either of these proteins in yeast leads to accumulation of SUMOylated proteins [136, 137], and hypersensitivity to certain types of DNA damage as well as accumulation of spontaneous damage during replication [138, 139]. The specific recognition of SUMOylated proteins occurs through SUMO interaction motifs (SIMs), which are present in Slx5, whereas interaction between Slx5 and Slx8 occurs through a RING-RING (Really Interesting New Gene) interaction [133]. Canonical SIMs are short hydrophobic sequences containing or being flanked by an acidic residue, which allow for interaction with the hydrophobic pocket and a basic surface on SUMO, respectively [140-142]. The overall strongest SIM has been defined as [VILFY]-[VI]-D-L-T [141].

In humans, the main STUbL is RING Finger Protein 4 (RNF4) [143], which has been identified to play a pivotal role in arsenic-induced degradation of PML

[144-147](**Figure 6**). Like in yeast, RNF4 has been implicated to have functions in the DNA damage response [148], where RNF4 promotes efficient DNA repair by regulating the turnover of repair proteins [128-130, 149]. A more recently discovered STUbL is RING Finger protein 111 (RNF111) [141], which has been linked to important non-proteolytic roles in the DNA damage response [150](**Figure 6**). RNF111 has additionally been implicated to increase neddylation at DNA damage sites [151], although overexpressed NEDD8 has been known to mimic ubiquitin due to extremely high fold similarity.

Whereas SUMO and ubiquitin may act together in a sequential fashion, there is also the intriguing possibility of an inversed regulation. Instead of ubiquitin being conjugated to SUMOylated proteins, there could also be an opposing mechanism, where ubiquitylation of SUMOylated proteins is removed. There are examples in the literature of ubiquitin ligases interacting and functioning together with ubiquitin-specific proteases, where the ligases are protected by these proteases from being proteolytically degraded through auto-ubiquitylation [152, 153]. Ubiquitin chain remodeling has also been implicated, where interacting ligases and proteases oppose each other's chain-editing functions [154].

Thus, it is not unimaginable that STUbLs have associated ubiquitin-specific proteases that could protect them from auto-degradation. Alternatively, such proteases could counterbalance the function of the STUbLs by reverse-editing SUMO-ubiquitin chains or removing ubiquitin from poly-SUMOylated targets, thereby protecting these proteins from proteolytic degradation. Such an association, however, has yet to be discovered.

Elucidating networks of PTMs at the system-wide level

One of the most effective ways to study how PTMs affect the entire proteome, is through system-wide proteomics [155-157], analyzing the PTM straight from complex samples derived from cultured cells or mammalian tissues. In the last years, there have been massive advances in the field of mass spectrometry and supporting bio-informatics [158, 159], allowing for large-scale analysis of PTMs at the system-wide level [160, 161]. Modern high-resolution and high-throughput mass spectrometers allow for identification of multiple peptides per second, even at very low abundance and from complex samples, while generating spectra detailed enough to confidently quantify modified peptides [162-164]. Combined with optimized methodologies and carefully refined purification methods in order to enrich modified proteins or peptides [165], this has greatly accelerated the understanding of the modified proteome. Additionally, quantitative approaches such as Stable Isotope Labeling of Amino Acids in Culture (SILAC) [166, 167] and Isobaric Tag for Relative and Absolute Quantitation (iTRAQ) [168, 169] have made it feasible to efficiently monitor the dynamics of PTMs. Several landmark papers have been published on several major PTMs, including acetylation [170], phosphorylation [171, 172], methylation [173], glycosylation [174] and ubiquitylation [175-179].

These studies identified many thousands of sites, and after these landmark papers, additional studies have greatly expanded upon the number of known sites. PhosphoSitePlus (PSP; PhosphoSitePlus®, www.phosphosite.org, [180]) is one of the major databases keeping track of all known modification sites. As of this moment, there are over 200,000 known phosphorylation sites, over 50,000 ubiquitylation sites, nearly 25,000 acetylation sites, and 7,500 methylation sites. Strikingly, there are only just over 700 known SUMOylation sites, with most of these sites originating from low-throughput mutagenesis approaches. Only around 150 SUMO sites were discovered by tandem mass spectrometry (MS/MS) methodology, with the majority being identified in two reported studies (**Figure 7**)[79, 181].

SUMOylation and mass spectrometry, a great challenge

The stark contrast between the amount of modification sites known for SUMOylation and other PTMs may be attributed to several causes. Firstly, SUMOylation occurs at a relatively low stoichiometry, and additionally in low abundant proteins [73]. Granted, there are several notable SUMO target proteins such as PML [19] and RanGAP1 [40, 41] which are an exception to the rule, and can be found to occur in predominantly SUMOylated form. However, the majority of known SUMO target proteins cannot readily be visualized in SUMOylated form, and pre-enrichment of SUMOylated proteins is required in order to study many of these proteins.

Secondly, SUMOylation is a reversible process, and there are several highly efficient SUMO-specific proteases that can cleave SUMO from its target proteins [81]. Under normal cellular growth conditions, the activity of these proteases is controlled. However, when processing cells or tissues for analysis, essentially mimicking an *in vitro* system, these proteases are given free reign, and swiftly remove all SUMO [182]. Strikingly, these proteases remain active in most standard lysis buffer conditions, and remain functional at freezing temperatures. There are no known effective and targeted inhibitors for SUMO-specific proteases, and many broad-spectrum protease inhibitor cocktails have zero effect. Mostly, copious amounts of acetamide are added during the lysis in order to alkylate the active cysteine of the SUMO proteases [59], yielding partial protection of the SUMOylated proteins. In order to completely counteract the activity of SUMO proteases, preparation of the sample has to be performed under highly denaturing conditions, such as high concentrations of sodium dodecyl sulphate (SDS), urea, or guanidine. While successful in preserving the SUMO on the target proteins, lysis in harsh conditions greatly complicates subsequent purification of the SUMO when using conventional approaches such as immunoprecipitation.

Thirdly, while SUMO-specific antibodies exist, and these have been applied to identify SUMOylated proteins [86], these antibodies are required in great quantities and are not always cost-effective. Furthermore, due to the aforementioned reasons, purification of SUMOylated proteins is a complicated procedure when using antibodies, because they will likewise be denatured by the stringent buffer

conditions. In order to gain respectable yields, a large amount of material – in the order of many billions of cells from culture – has to be processed for a singular sample. Conversely, only a few thousands of cells are needed in order to detect virtually any protein by various biochemical techniques such as Enzyme-Linked Immunosorbent Assay (ELISA) [183] or immunoblotting [184].

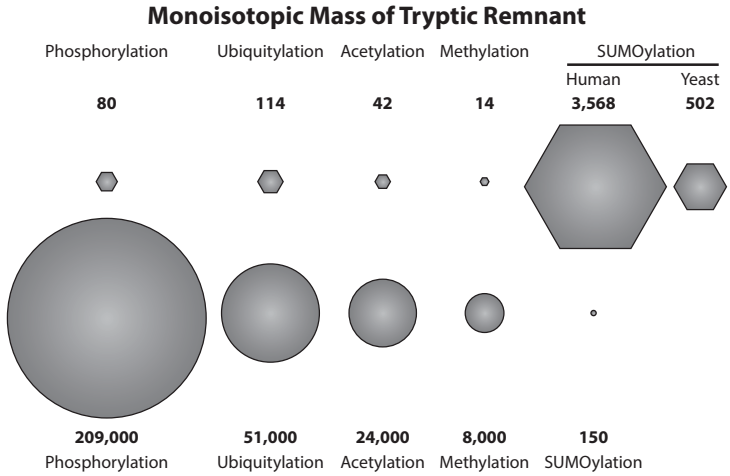


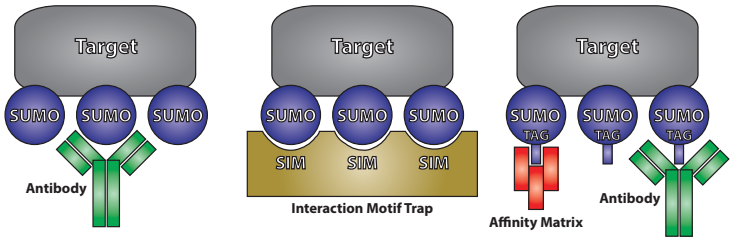
Figure 7. A system-wide proteomics overview of all major PTMs, as compared to SUMOylation. Phosphorylation, ubiquitylation, acetylation and methylation are readily detectable as modifications on peptides, due to a modest monoisotopic mass increase on the tryptic peptides. Conversely, human SUMO-1 and SUMO-2/3 yield a tryptic remnant with masses exceeding 3,000 Dalton, preventing efficient identification of modification sites by MS/MS. Yeast SUMO (Smt3) yields a tryptic remnant with a mass of 502 Dalton, which is more suitable for proteomic analysis. Not surprisingly, the amount of MS/MS-identified SUMOylation sites pales in comparison to the other major PTMs. Based on the amount of known SUMOylated proteins, many more SUMOylation sites exist, but have yet to be pinpointed by MS/MS. Monoisotopic masses and known modification sites are drawn to scale.

Fourthly, mass spectrometric PTM analysis is most frequently and routinely performed on peptide mixtures generated after digestion of proteins by trypsin [185, 186]. As trypsin specifically cleaves carboxyl-terminal of arginines and lysines, tryptic peptides have a length dependent on the arginine and lysine content of the protein. For ubiquitin-like modifiers, besides the tryptic digestion pattern of the modified protein, the digestion pattern of the modifier itself also plays an important factor. For ubiquitin a tryptic digest results in a remnant di-glycine on the modified lysine in the target peptide, which may be readily detected by mass spectrometry [187], and antibodies have been developed to specifically recognize peptides modified by di-glycine [179, 188]. In case of SUMO-2/3, the tryptic remnant is 32 amino acids long [79, 189], which is too bulky and prohibits efficient analysis by current mass spectrometry (**Figure 7**). Additional enzymes would have to be

employed to shorten this remnant [189, 190], which in turn would also further shorten the target peptide, often leading to ambiguous peptide identifications. Also, these more exotic proteomics-grade enzymes are prohibitively expensive, and often not stable under partially denaturing conditions, whereas commonly applied enzymes such as Lysyl Endopeptidase (Lys-C) and trypsin function efficiently in 8 M and 2 M urea, respectively.

SUMO proteomics, the current state of affairs

In spite of the difficulties in studying SUMOylation at the site-specific level, several advances have been made at the protein level in a system-wide manner. Most commonly, SUMO is fused with an amino-terminal epitope tag or tandem tags; such as His_6 [85, 191], His_6-FLAG [192], His_6-HA [193], Myc [194], FLAG-TEV [195], or protA-TEV-CBP [58], and overexpressed in a cultured cell line or model organism (**Figure 8**). Overexpression of SUMO, in combination with the epitope tags that allow for more efficient purification, has allowed for identification of many hundreds of SUMOylated proteins [58, 196, 197]. Often, these approaches employ SILAC, in order to extract quantitative information about the changes in SUMOylation in response to several treatments [58, 64, 89, 198]. Furthermore, several approaches have proven successful in identifying proteins putatively modified by endogenous SUMOylation, through the use of antibodies [86] or SUMO interacting motif (SIM) traps [59](**Figure 8**).



	IP	Trap	Tag-PD	Tag-IP
SUMO Expression	Endogenous	Endogenous	Exogenous	Exogenous
Purification Bias	Poly	Poly Only	None	Poly
Buffer Conditions	Mild	Mild	Harsh	Intermediate
Yield	Low	Low	High	High
Background	Medium	High	Medium	Low

Figure 8. An overview of the various methodologies used to study SUMOylation.
SUMOylated proteins may be directly purified through immunoprecipitation (IP), using of antibodies directed against SUMO. Poly-SUMOylated proteins have been captured using SIM-based traps. Both these methodologies allow study of endogenously SUMOylated proteins. More efficient purification methods involve pulldown (PD) using affinity matrices and immunoprecipitation using antibodies targeted against overexpressed epitope-tagged SUMO, providing a much higher yield and less background interference.

When comparing the different SUMO purification methodologies, there is a clear distinction between endogenous versus exogenous approaches. SUMOylated proteins may be directly purified using immunoprecipitation, through use of antibodies directed against SUMO. While allowing studying of proteins modified by endogenous SUMO, the yield of this approach is relatively low and mild buffer conditions have to be used, giving free reign to SUMO proteases. This necessitates the use of large quantities of antibody, and large amounts of starting material, while having to cope with a relatively high amount of background. Poly-SUMOylated proteins have been captured using SIM-based traps, also allowing the study of endogenously SUMOylated proteins. However, this methodology cannot detect mono-SUMOylated proteins efficiently, suffers from high background interference due to non-specific interactions with the SIMs, as well as some of the same drawbacks from immunoprecipitation.

Epitope-tagged SUMO can be purified far more efficiently, due to specialized methodology and commercially available purification tools. Histidine and biotin tags allow for pulldown of SUMOylated proteins using affinity matrices, which remain functional under the harshest of conditions, completely inactivating SUMO proteases. The interaction between the affinity matrices and the tags are among the strongest non-covalent interactions known, allowing for rigorous washing procedures. The abundance of matrix material allows for total and complete purification of all SUMO without any bias. As a slight drawback, the chemical interactivity of the affinity matrices may result in some background binding due to high histidine content of non-related proteins or endogenous protein biotinylation. Alternatively, antibodies directed against common epitope tags such as HA or FLAG can be used for immunoprecipitation of tagged SUMO. The quality of the antibodies against epitope tags is very high and commercially coupled antibody matrices are very robust and of high quality. Compared to antibodies used against SUMO itself, this allows for the use of more stringent buffer conditions, and results in a much higher yield combined with less background. Regardless, there are still limits to buffer conditions during the immunoprecipitation, allowing for some co-purification of non-specific proteins. Ultimately, the main drawback of epitope tag approaches is the requirement for a model system which allows exogenous expression of the tagged SUMO.

Whereas proteomic studies of SUMOylation over the last decade have provided insight into which proteins are subject to SUMOylation, and how they may dynamically be increased or decrease in SUMOylation upon cellular stresses, they fail to provide knowledge about the exact SUMO acceptor lysines. For efficient follow-up study, it is of paramount importance that the exact sites of SUMOylation are determined, in order to generate separation-of-function mutants which may then be used to assess the exact function of SUMO within the target proteins. Furthermore, while all SUMOylated proteins are identified by the presence of corresponding peptides in the purified fraction, there is a lack of direct evidence in

the form of a tryptic-remnant-modified peptide. Modified peptides serve as direct proof of modification, and greatly reduce the amount of false positive hits resulting from the digestion of background binders sticking to the purification matrices and plastic tube walls. Additionally, whereas an entire protein may be identified as being regulated by SUMOylation in response to a stress, there is a distinct possibility that multiple acceptor lysines within that same protein are differentially or even inversely regulated. Knowing in which region of a protein SUMOylation occurs may already provide clues as to the potential functionality of the SUMO modification. Lastly, the identified SUMOylation sites can be directly matched against other potential lysine modifications on the same sites, to investigate any PTM competition for the same lysines.

Some limited progress has been made in mapping SUMO acceptor lysines. In order to counter one of the largest issues with identification of SUMO sites – the oversized tryptic remnant – a SUMO mutant containing an additional arginine close to the carboxyl-terminus is commonly used, in order to generate a mass remnant which is small enough to identify reliably. Such approaches have identified 14 sites in HeLa cells [199], 17 sites in Human Embryonic Kidney 293 (HEK293) cells [200], and 17 sites in *Arabidopsis thaliana* [201]. One of the carboxyl-terminal arginine mutations used in SUMO-2 is Q87R, analogous to the yeast SUMO, Smt3, yielding the glutamine-glutamine-threonine-glycine-glycine (QQTGG) remnant. The other is T90R, analogous to ubiquitin, yielding the di-glycine remnant. These mutations do not significantly alter the behavior of SUMO [202, 203].

In order to enhance the efficiency of identification, we mutated every lysine within SUMO-2 to an arginine, in addition to the Q87R mutation. This grants the mutant SUMO-2 immunity to the endopeptidase Lys-C, which only cleaves carboxyl-terminal of lysines, whereas all other proteins in the sample will be readily digested. Subsequently, SUMOylated peptides can be purified using conventional approaches. Because enrichment of the SUMOylated peptides takes place after protein digestion, this allows for selection of peptides-of-interest only, greatly reducing the complexity of the sample (**Figure 9**). Using this approach, we identified

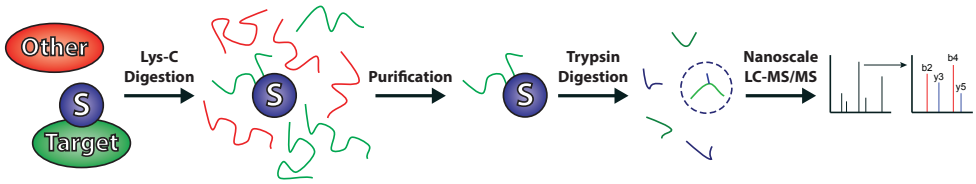


Figure 9. Site-specific identification of SUMOylation sites using lysine-deficient SUMO.
In order to enrich SUMOylated peptides, the entire total lysate is pre-digested with Lys-C, cleaving all proteins except the lysine-deficient SUMO. Subsequently SUMOylated peptides are enriched by His-pulldown, and digested with trypsin. Finally, peptides bearing the di-glycine or QQTGG remnant (depending on the SUMO mutant used) are analyzed using nanoscale liquid chromatography followed by tandem high-resolution mass spectrometry.

103 SUMOylation acceptor sites on endogenous proteins purified from a complex sample [79], making it the most comprehensive SUMOylation site study to date.

SUMO, the missing sites enigma

One of the greatest questions in the SUMO field is: where exactly are all the SUMOylated proteins being modified? Strikingly, SUMO is the only PTM where the amount of known modified proteins surpasses the amount of known sites of modification, especially when considering only evidence generated through system-wide proteomics approaches. Nearly one thousand SUMOylated proteins have been discovered through mass spectrometry and other screening methods, compared to a meager 150 SUMOylation sites. Assuming an average frequency of 2 to 3 modification sites per protein, one could argue that there could easily be a few thousand SUMOylation sites. Moreover, the large majority of all currently known SUMOylation sites have been mapped through low-throughput methodology, involving trial-and-error mutagenesis. In part, the success in identification of these sites may be attributed to the KxE consensus motif, which often allows researchers to successfully perform a so-called “intelligent guess”.

Looking at the 103 sites we identified by MS/MS [79], the overall amount of SUMOylation sites matching the KxE consensus was nearly 75%. This is consistent with the literature, in the sense that Ubc9 targets SUMO to KxE motifs [76, 77]. Thus, a high adherence to the KxE consensus motif may be applied as a quality control standard for SUMOylation site datasets. However, this dataset is just the tip of the iceberg; as it is likely limited to the most abundant SUMOylation sites, and pertains to one type of cells under standard growth conditions. There is little known about the global specificity certain E3 enzymes grant, and about the specificity of SUMO conjugation under cellular stresses. Heat shock and proteasome inhibition have been applied to study SUMO, and these treatments have been noted to lead to a large accumulation of SUMOylated proteins, in addition to a highly dynamic shuffling of SUMO between subsets of targets. Thus, with these dynamics likely carrying over at the site-specific level, it is feasible to assume that we are currently missing most pieces to the SUMO puzzle.

Finally, while approaches with mutant SUMO may yield qualitatively sound datasets, this methodology does not map truly endogenous SUMO sites, and is restricted to application in model systems. Ultimately, investigation of clinically relevant samples for aberrant SUMOylation would require an endogenous, quantitative, and site-specific methodology capable of identifying hundreds of modification sites. At the moment of writing this thesis, this goal may be years, if not decades, away from being achieved.

Reference List

1. Ota, T. et al. (2004) Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat. Genet.* 36, 40-45
2. Berger, S. L. (2007) The complex language of chromatin regulation during transcription. *Nature* 447, 407-412
3. Davidson, E. H., and Britten, R. J. (1973) Organization, transcription, and regulation in the animal genome. *Q. Rev. Biol.* 48, 565-613
4. Ayoubi, T. A., and Van de Ven, W. J. (1996) Regulation of gene expression by alternative promoters. *FASEB J.* 10, 453-460
5. Tischer, E. et al. (1991) The human gene for vascular endothelial growth factor. Multiple protein forms are encoded through alternative exon splicing. *J. Biol. Chem.* 266, 11947-11954
6. Hiesel, R., Wissinger, B., Schuster, W., and Brennicke, A. (1989) RNA editing in plant mitochondria. *Science* 246, 1632-1634
7. Covello, P. S., and Gray, M. W. (1989) RNA editing in plant mitochondria. *Nature* 341, 662-666
8. Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40, 1413-1415
9. Carninci, P. et al. (2005) The transcriptional landscape of the mammalian genome. *Science* 309, 1559-1563
10. Kozak, M. (1989) The scanning model for translation: an update. *J. Cell Biol.* 108, 229-241
11. Hershey, J. W. (1991) Translational control in mammalian cells. *Annu. Rev. Biochem.* 60, 717-755
12. Uy, R., and Wold, F. (1977) Posttranslational covalent modification of proteins. *Science* 198, 890-896
13. Wold, F. (1981) In vivo chemical modification of proteins (post-translational modification). *Annu. Rev. Biochem.* 50, 783-814
14. Chen, L. F., and Greene, W. C. (2003) Regulation of distinct biological activities of the NF-kappaB transcription factor complex by acetylation. *J. Mol. Med. (Berl)* 81, 549-557
15. Sakaguchi, K. et al. (1998) DNA damage activates p53 through a phosphorylation-acetylation cascade. *Genes Dev.* 12, 2831-2841
16. Gary, J. D., and Clarke, S. (1998) RNA and protein interactions modulated by protein arginine methylation. *Prog. Nucleic Acid Res. Mol. Biol.* 61, 65-131
17. Ducommun, B. et al. (1991) cdc2 phosphorylation is required for its interaction with cyclin. *EMBO J.* 10, 3311-3319
18. Hammer, G. D. et al. (1999) Phosphorylation of the nuclear receptor SF-1 modulates cofactor recruitment: integration of hormone signaling in reproduction and stress. *Mol. Cell* 3, 521-526
19. Muller, S., Matunis, M. J., and Dejean, A. (1998) Conjugation with the ubiquitin-related modifier SUMO-1 regulates the partitioning of PML within the nucleus. *EMBO J.* 17, 61-70
20. Meek, D. W., and Knippschild, U. (2003) Posttranslational modification of MDM2. *Mol. Cancer Res.* 1, 1017-1026
21. Psakhye, I., and Jentsch, S. (2012) Protein group modification and synergy in the SUMO pathway as exemplified in DNA repair. *Cell* 151, 807-820
22. Hardeland, U., Steinacher, R., Jiricny, J., and Schar, P. (2002) Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover.

EMBO J. 21, 1456-1464

23. Hershko, A., and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.* 67, 425-479
24. Hunter, T. (2007) The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol. Cell* 28, 730-738
25. Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. *Nature* 403, 41-45
26. Garavelli, J. S. (2004) The RESID Database of Protein Modifications as a resource and annotation tool. *Proteomics.* 4, 1527-1533
27. Jenuwein, T., and Allis, C. D. (2001) Translating the histone code. *Science* 293, 1074-1080
28. Glozak, M. A., Sengupta, N., Zhang, X., and Seto, E. (2005) Acetylation and deacetylation of non-histone proteins. *Gene* 363, 15-23
29. Hunter, T. (1995) Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* 80, 225-236
30. Krebs, E. G., and Beavo, J. A. (1979) Phosphorylation-dephosphorylation of enzymes. *Annu. Rev. Biochem.* 48, 923-959
31. Martin, C., and Zhang, Y. (2005) The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 6, 838-849
32. Bedford, M. T., and Richard, S. (2005) Arginine methylation an emerging regulator of protein function. *Mol. Cell* 18, 263-272
33. Rudd, P. M., Elliott, T., Cresswell, P., Wilson, I. A., and Dwek, R. A. (2001) Glycosylation and the immune system. *Science* 291, 2370-2376
34. Brownlee, M. (1995) Advanced protein glycosylation in diabetes and aging. *Annu. Rev. Med.* 46, 223-234
35. Ueda, K., and Hayaishi, O. (1985) ADP-ribosylation. *Annu. Rev. Biochem.* 54, 73-100
36. D'Amours, D., Desnoyers, S., D'Silva, I., and Poirier, G. G. (1999) Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem. J.* 342 (Pt 2), 249-268
37. Hochstrasser, M. (1996) Ubiquitin-dependent protein degradation. *Annu. Rev. Genet.* 30, 405-439
38. Pickart, C. M. (2001) Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* 70, 503-533
39. Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu. Rev. Cell Dev. Biol.* 22, 159-180
40. Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997) A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* 88, 97-107
41. Matunis, M. J., Coutavas, E., and Blobel, G. (1996) A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J. Cell Biol.* 135, 1457-1470
42. Muller, S., Hoege, C., Pyrowolakis, G., and Jentsch, S. (2001) SUMO, ubiquitin's mysterious cousin. *Nat. Rev. Mol. Cell Biol.* 2, 202-210
43. Liakopoulos, D., Busgen, T., Brychzy, A., Jentsch, S., and Pause, A. (1999) Conjugation of the ubiquitin-like protein NEDD8 to cullin-2 is linked to von Hippel-Lindau tumor suppressor function. *Proc. Natl. Acad. Sci. U. S. A* 96, 5510-5515
44. Cope, G. A., and Deshaies, R. J. (2003) COP9 signalosome: a multifunctional regulator of SCF and other cullin-based ubiquitin ligases. *Cell* 114, 663-671
45. Raasi, S., Schmidtke, G., and Groettrup, M. (2001) The ubiquitin-like protein FAT10 forms covalent conjugates and induces apoptosis. *J. Biol. Chem.* 276, 35334-35343

46. Liu, Y. C. et al. (1999) A MHC-encoded ubiquitin-like protein (FAT10) binds noncovalently to the spindle assembly checkpoint protein MAD2. *Proc. Natl. Acad. Sci. U. S. A* 96, 4313-4318
47. Yuan, W., and Krug, R. M. (2001) Influenza B virus NS1 protein inhibits conjugation of the interferon (IFN)-induced ubiquitin-like ISG15 protein. *EMBO J.* 20, 362-371
48. D'Cunha, J., Knight E Jr, Haas, A. L., Truitt, R. L., and Borden, E. C. (1996) Immunoregulatory properties of ISG15, an interferon-induced cytokine. *Proc. Natl. Acad. Sci. U. S. A* 93, 211-215
49. Sasakawa, H. et al. (2006) Solution structure and dynamics of Ufm1, a ubiquitin-fold modifier 1. *Biochem. Biophys. Res. Commun.* 343, 21-26
50. Komatsu, M. et al. (2004) A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. *EMBO J.* 23, 1977-1986
51. Mourtada-Maarabouni, M., Kirkham, L., Farzaneh, F., and Williams, G. T. (2004) Regulation of apoptosis by fau revealed by functional expression cloning and antisense expression. *Oncogene* 23, 9419-9426
52. Michiels, L., Van der Rauwelaert, E., Van, H. F., Kas, K., and Merregaert, J. (1993) fau cDNA encodes a ubiquitin-like-S30 fusion protein and is expressed as an antisense sequence in the Finkel-Biskis-Reilly murine sarcoma virus. *Oncogene* 8, 2537-2546
53. Ha, B. H., and Kim, E. E. (2008) Structures of proteases for ubiqutin and ubiquitin-like modifiers. *BMB. Rep.* 41, 435-443
54. Mosammaparast, N., and Shi, Y. (2010) Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu. Rev. Biochem.* 79, 155-179
55. Nijman, S. M. et al. (2005) A genomic and functional inventory of deubiquitinating enzymes. *Cell* 123, 773-786
56. Komander, D., Clague, M. J., and Urbe, S. (2009) Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* 10, 550-563
57. Bayer, P. et al. (1998) Structure determination of the small ubiquitin-related modifier SUMO-1. *J. Mol. Biol.* 280, 275-286
58. Golebiowski, F. et al. (2009) System-wide changes to SUMO modifications in response to heat shock. *Sci. Signal.* 2, ra24
59. Bruderer, R. et al. (2011) Purification and identification of endogenous polySUMO conjugates. *EMBO Rep.* 12, 142-148
60. Flotho, A., and Melchior, F. (2013) Sumoylation: a regulatory protein modification in health and disease. *Annu. Rev. Biochem.* 82, 357-385
61. Ulrich, H. D., and Walden, H. (2010) Ubiquitin signalling in DNA replication and repair. *Nat. Rev. Mol. Cell Biol.* 11, 479-489
62. Hickey, C. M., Wilson, N. R., and Hochstrasser, M. (2012) Function and regulation of SUMO proteases. *Nat. Rev. Mol. Cell Biol.* 13, 755-766
63. Jackson, S. P., and Durocher, D. (2013) Regulation of DNA damage responses by ubiquitin and SUMO. *Mol. Cell* 49, 795-807
64. Vertegaal, A. C. (2011) Uncovering ubiquitin and ubiquitin-like signaling networks. *Chem. Rev.* 111, 7923-7940
65. Spector, D. L. (2001) Nuclear domains. *J. Cell Sci.* 114, 2891-2893
66. Schulman, B. A., and Harper, J. W. (2009) Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat. Rev. Mol. Cell Biol.* 10, 319-331
67. Bernier-Villamor, V., Sampson, D. A., Matunis, M. J., and Lima, C. D. (2002) Structural basis for E2-mediated SUMO

conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* 108, 345-356

68. Olsen, S. K., Capili, A. D., Lu, X., Tan, D. S., and Lima, C. D. (2010) Active site remodelling accompanies thioester bond formation in the SUMO E1. *Nature* 463, 906-912
69. Desterro, J. M., Rodriguez, M. S., Kemp, G. D., and Hay, R. T. (1999) Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1. *J. Biol. Chem.* 274, 10618-10624
70. Azuma, Y. et al. (2001) Expression and regulation of the mammalian SUMO-1 E1 enzyme. *FASEB J.* 15, 1825-1827
71. Okuma, T., Honda, R., Ichikawa, G., Tsumagari, N., and Yasuda, H. (1999) In vitro SUMO-1 modification requires two enzymatic steps, E1 and E2. *Biochem. Biophys. Res. Commun.* 254, 693-698
72. Nagy, V., and Dikic, I. (2010) Ubiquitin ligase complexes: from substrate selectivity to conjugational specificity. *Biol. Chem.* 391, 163-169
73. Hay, R. T. (2005) SUMO: a history of modification. *Mol. Cell* 18, 1-12
74. Johnson, E. S. (2004) Protein modification by SUMO. *Annu. Rev. Biochem.* 73, 355-382
75. Reverter, D., and Lima, C. D. (2005) Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature* 435, 687-692
76. Rodriguez, M. S., Dargemont, C., and Hay, R. T. (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J. Biol. Chem.* 276, 12654-12659
77. Sampson, D. A., Wang, M., and Matunis, M. J. (2001) The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. *J. Biol. Chem.* 276, 21664-21669
78. Gareau, J. R., and Lima, C. D. (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat. Rev. Mol. Cell Biol.* 11, 861-871
79. Matic, I. et al. (2010) Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. *Mol. Cell* 39, 641-652
80. Li, S. J., and Hochstrasser, M. (1999) A new protease required for cell-cycle progression in yeast. *Nature* 398, 246-251
81. Mukhopadhyay, D., and Dasso, M. (2007) Modification in reverse: the SUMO proteases. *Trends Biochem. Sci.* 32, 286-295
82. Di, B. A. et al. (2006) The SUMO-specific protease SENP5 is required for cell division. *Mol. Cell Biol.* 26, 4489-4498
83. Cheng, J., Kang, X., Zhang, S., and Yeh, E. T. (2007) SUMO-specific protease 1 is essential for stabilization of HIF1alpha during hypoxia. *Cell* 131, 584-595
84. Wang, Y., and Dasso, M. (2009) SUMOylation and deSUMOylation at a glance. *J. Cell Sci.* 122, 4249-4252
85. Vertegaal, A. C. et al. (2006) Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. *Mol. Cell Proteomics* 5, 2298-2310
86. Becker, J. et al. (2013) Detecting endogenous SUMO targets in mammalian cells and tissues. *Nat. Struct. Mol. Biol.* 20, 525-531
87. Saitoh, H., and Hinchey, J. (2000) Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J. Biol. Chem.* 275, 6252-6258
88. Kurepa, J. et al. (2003) The small ubiquitin-like modifier (SUMO) protein modification system in Arabidopsis. Accumulation of SUMO1 and -2 conjugates is

increased by stress. *J. Biol. Chem.* 278, 6862-6872
89. Schimmel, J. et al. (2008) The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle. *Mol. Cell Proteomics* 7, 2107-2122
90. Matic, I. et al. (2008) In vivo identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. *Mol. Cell Proteomics* 7, 132-144
91. Vertegaal, A. C. (2010) SUMO chains: polymeric signals. *Biochem. Soc. Trans.* 38, 46-49
92. Tatham, M. H. et al. (2001) Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *J. Biol. Chem.* 276, 35368-35374
93. Geoffroy, M. C., and Hay, R. T. (2009) An additional role for SUMO in ubiquitin-mediated proteolysis. *Nat. Rev. Mol. Cell Biol.* 10, 564-568
94. Srikumar, T. et al. (2013) Global analysis of SUMO chain function reveals multiple roles in chromatin regulation. *J. Cell Biol.* 201, 145-163
95. Miura, K., Jin, J. B., and Hasegawa, P. M. (2007) Sumoylation, a post-translational regulatory process in plants. *Curr. Opin. Plant Biol.* 10, 495-502
96. Geiss-Friedlander, R., and Melchior, F. (2007) Concepts in sumoylation: a decade on. *Nat. Rev. Mol. Cell Biol.* 8, 947-956
97. Tanaka, K. et al. (1999) Characterization of a fission yeast SUMO-1 homologue, pmt3p, required for multiple nuclear events, including the control of telomere length and chromosome segregation. *Mol. Cell Biol.* 19, 8660-8672
98. Wong, K. H. et al. (2008) Sumoylation in Aspergillus nidulans: sumO inactivation, overexpression and live-cell imaging. *Fungal. Genet. Biol.* 45, 728-737
99. Hayashi, T. et al. (2002) Ubc9 is essential for viability of higher eukaryotic cells. *Exp. Cell Res.* 280, 212-221
100. Nacerddine, K. et al. (2005) The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev. Cell* 9, 769-779
101. Saracco, S. A., Miller, M. J., Kurepa, J., and Vierstra, R. D. (2007) Genetic analysis of SUMOylation in Arabidopsis: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. *Plant Physiol* 145, 119-134
102. Bettermann, K., Benesch, M., Weis, S., and Haybaeck, J. (2012) SUMOylation in carcinogenesis. *Cancer Lett.* 316, 113-125
103. Mei, D. et al. (2013) Up-regulation of SUMO1 pseudogene 3 (SUMO1P3) in gastric cancer and its clinical association. *Med. Oncol.* 30, 709
104. Wang, Q. et al. (2013) SUMO-specific protease 1 promotes prostate cancer progression and metastasis. *Oncogene* 32, 2493-2498
105. van, H. M., Overmeer, R. M., Abolvardi, S. S., and Vertegaal, A. C. (2010) RNF4 and VHL regulate the proteasomal degradation of SUMO-conjugated Hypoxia-Inducible Factor-2alpha. *Nucleic Acids Res.* 38, 1922-1931
106. Rodriguez, M. S. et al. (1999) SUMO-1 modification activates the transcriptional response of p53. *EMBO J.* 18, 6455-6461
107. Kessler, J. D. et al. (2012) A SUMOylation-dependent transcriptional subprogram is required for Myc-driven tumorigenesis. *Science* 335, 348-353
108. Lee, L., Sakurai, M., Matsuzaki, S., Arancio, O., and Fraser, P. (2013) SUMO and Alzheimer's Disease. *Neuromolecular. Med.*
109. Lee, Y. J., and Hallenbeck, J. M. (2013) SUMO and Ischemic Tolerance. *Neuromolecular. Med.*

110. Deshaies, R. J., and Joazeiro, C. A. (2009) RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem.* 78, 399-434
111. Ardley, H. C., and Robinson, P. A. (2005) E3 ubiquitin ligases. *Essays Biochem.* 41, 15-30
112. Seeler, J. S., and Dejean, A. (2003) Nuclear and unclear functions of SUMO. *Nat. Rev. Mol. Cell Biol.* 4, 690-699
113. Kamitani, T., Nguyen, H. P., and Yeh, E. T. (1997) Preferential modification of nuclear proteins by a novel ubiquitin-like molecule. *J. Biol. Chem.* 272, 14001-14004
114. Shiio, Y., and Eisenman, R. N. (2003) Histone sumoylation is associated with transcriptional repression. *Proc. Natl. Acad. Sci. U. S. A* 100, 13225-13230
115. Stielow, B. et al. (2008) Identification of SUMO-dependent chromatin-associated transcriptional repression components by a genome-wide RNAi screen. *Mol. Cell* 29, 742-754
116. Uchimura, Y. et al. (2006) Involvement of SUMO modification in. *J. Biol. Chem.* 281, 23180-23190
117. Ishov, A. M. et al. (1999) PML is critical for ND10 formation and recruits the PML-interacting protein daxx to this nuclear structure when modified by SUMO-1. *J. Cell Biol.* 147, 221-234
118. Zhong, S. et al. (2000) Role of SUMO-1-modified PML in nuclear body formation. *Blood* 95, 2748-2752
119. Muller, S., and Dejean, A. (1999) Viral immediate-early proteins abrogate the modification by SUMO-1 of PML and Sp100 proteins, correlating with nuclear body disruption. *J. Virol.* 73, 5137-5143
120. Shen, T. H., Lin, H. K., Scaglioni, P. P., Yung, T. M., and Pandolfi, P. P. (2006) The mechanisms of PML-nuclear body formation. *Mol. Cell* 24, 331-339
121. Bernardi, R., and Pandolfi, P. P. (2003) Role of PML and the PML-nuclear body in the control of programmed cell death. *Oncogene* 22, 9048-9057
122. Dellaire, G., and Bazett-Jones, D. P. (2004) PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. *Bioessays* 26, 963-977
123. Bergink, S., and Jentsch, S. (2009) Principles of ubiquitin and SUMO modifications in DNA repair. *Nature* 458, 461-467
124. Stelter, P., and Ulrich, H. D. (2003) Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. *Nature* 425, 188-191
125. Hoege, C., Pfander, B., Moldovan, G. L., Pyrowolakis, G., and Jentsch, S. (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* 419, 135-141
126. Morris, J. R. et al. (2009) The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. *Nature* 462, 886-890
127. Galanty, Y. et al. (2009) Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. *Nature* 462, 935-939
128. Yin, Y. et al. (2012) SUMO-targeted ubiquitin E3 ligase RNF4 is required for the response of human cells to DNA damage. *Genes Dev.* 26, 1196-1208
129. Galanty, Y., Belotserkovskaya, R., Coates, J., and Jackson, S. P. (2012) RNF4, a SUMO-targeted ubiquitin E3 ligase, promotes DNA double-strand break repair. *Genes Dev.* 26, 1179-1195
130. Vyas, R. et al. (2013) RNF4 is required for DNA double-strand break repair in vivo. *Cell Death Differ.* 20, 490-502
131. Silver, H. R., Nissley, J. A., Reed, S. H., Hou, Y. M., and Johnson, E. S. (2011) A role for SUMO in nucleotide excision repair. *DNA*

Repair (Amst) 10, 1243-1251

132. Prudden, J. et al. (2007) SUMO-targeted ubiquitin ligases in genome stability. *EMBO J.* 26, 4089-4101
133. Perry, J. J., Tainer, J. A., and Boddy, M. N. (2008) A SIM-ultaneous role for SUMO and ubiquitin. *Trends Biochem. Sci.* 33, 201-208
134. Wang, Z., and Prelich, G. (2009) Quality control of a transcriptional regulator by SUMO-targeted degradation. *Mol. Cell Biol.* 29, 1694-1706
135. Nagai, S., Davoodi, N., and Gasser, S. M. (2011) Nuclear organization in genome stability: SUMO connections. *Cell Res.* 21, 474-485
136. Xie, Y. et al. (2007) The yeast Hex3.Slx8 heterodimer is a ubiquitin ligase stimulated by substrate sumoylation. *J. Biol. Chem.* 282, 34176-34184
137. Wang, Z., Jones, G. M., and Prelich, G. (2006) Genetic analysis connects SLX5 and SLX8 to the SUMO pathway in Saccharomyces cerevisiae. *Genetics* 172, 1499-1509
138. Zhang, C., Roberts, T. M., Yang, J., Desai, R., and Brown, G. W. (2006) Suppression of genomic instability by SLX5 and SLX8 in Saccharomyces cerevisiae. *DNA Repair (Amst)* 5, 336-346
139. Mullen, J. R., Kaliraman, V., Ibrahim, S. S., and Brill, S. J. (2001) Requirement for three novel protein complexes in the absence of the Sgs1 DNA helicase in Saccharomyces cerevisiae. *Genetics* 157, 103-118
140. Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P., and Dikic, I. (2006) Specification of S. *J. Biol. Chem.* 281, 16117-16127
141. Sun, H., and Hunter, T. (2012) Poly-small ubiquitin-like modifier (PolySUMO)-binding proteins identified through a string search. *J. Biol. Chem.* 287, 42071-42083
142. Song, J., Durrin, L. K., Wilkinson, T. A., Krontiris, T. G., and Chen, Y. (2004) Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proc. Natl. Acad. Sci. U. S. A* 101, 14373-14378
143. Sun, H., Leverson, J. D., and Hunter, T. (2007) Conserved function of RNF4 family proteins in eukaryotes: targeting a ubiquitin ligase to SUMOylated proteins. *EMBO J.* 26, 4102-4112
144. Tatham, M. H. et al. (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat. Cell Biol.* 10, 538-546
145. Weisshaar, S. R. et al. (2008) Arsenic trioxide stimulates SUMO-2/3 modification leading to RNF4-dependent proteolytic targeting of PML. *FEBS Lett.* 582, 3174-3178
146. Geoffroy, M. C., Jaffray, E. G., Walker, K. J., and Hay, R. T. (2010) Arsenic-induced SUMO-dependent recruitment of RNF4 into PML nuclear bodies. *Mol. Biol. Cell* 21, 4227-4239
147. Lallemand-Breitenbach, V. et al. (2008) Arsenic degrades PML or PML-RARalpha through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat. Cell Biol.* 10, 547-555
148. Kosoy, A., Calonge, T. M., Outwin, E. A., and O'Connell, M. J. (2007) Fission yeast Rnf4 homologs are required for DNA repair. *J. Biol. Chem.* 282, 20388-20394
149. Guzzo, C. M. et al. (2012) RNF4-dependent hybrid SUMO-ubiquitin chains are signals for RAP80 and thereby mediate the recruitment of BRCA1 to sites of DNA damage. *Sci. Signal.* 5, ra88
150. Poulsen, S. L. et al. (2013) RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response. *J. Cell Biol.* 201, 797-807
151. Ma, T. et al. (2013) RNF111-dependent neddylation activates DNA damage-induced ubiquitination. *Mol. Cell* 49, 897-907

152. Canning, M., Boutell, C., Parkinson, J., and Everett, R. D. (2004) A RING finger ubiquitin ligase is protected from autocatalyzed ubiquitination and degradation by binding to ubiquitin-specific protease USP7. *J. Biol. Chem.* 279, 38160-38168
153. Scortegagna, M. et al. (2011) USP13 enzyme regulates Siah2 ligase stability and activity via noncatalytic ubiquitin-binding domains. *J. Biol. Chem.* 286, 27333-27341
154. Crosas, B. et al. (2006) Ubiquitin chains are remodeled at the proteasome by opposing ubiquitin ligase and deubiquitinating activities. *Cell* 127, 1401-1413
155. Witze, E. S., Old, W. M., Resing, K. A., and Ahn, N. G. (2007) Mapping protein post-translational modifications with mass spectrometry. *Nat. Methods* 4, 798-806
156. Pandey, A., and Mann, M. (2000) Proteomics to study genes and genomes. *Nature* 405, 837-846
157. Mann, M., and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications. *Nat. Biotechnol.* 21, 255-261
158. Cox, J. et al. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome. Res.* 10, 1794-1805
159. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367-1372
160. Olsen, J. V., and Mann, M. (2013) Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol. Cell Proteomics.*
161. Choudhary, C., and Mann, M. (2010) Decoding signalling networks by mass spectrometry-based proteomics. *Nat. Rev. Mol. Cell Biol.* 11, 427-439
162. Kelstrup, C. D., Young, C., Lavallee, R., Nielsen, M. L., and Olsen, J. V. (2012) Optimized Fast and Sensitive Acquisition Methods for Shotgun Proteomics on a Quadrupole Orbitrap Mass Spectrometer. *J. Proteome. Res.*
163. Michalski, A. et al. (2011) Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol. Cell Proteomics* 10, M111
164. Gallien, S. et al. (2012) Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Mol. Cell Proteomics.* 11, 1709-1723
165. Mertins, P. et al. (2013) Integrated proteomic analysis of post-translational modifications by serial enrichment. *Nat. Methods* 10, 634-637
166. Ong, S. E. et al. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell Proteomics.* 1, 376-386
167. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 1, 252-262
168. Aggarwal, K., Choe, L. H., and Lee, K. H. (2006) Shotgun proteomics using the iTRAQ isobaric tags. *Brief. Funct. Genomic. Proteomic.* 5, 112-120
169. Zhang, Y. et al. (2005) Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol. Cell Proteomics.* 4, 1240-1250
170. Choudhary, C. et al. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325, 834-840
171. Huttlin, E. L. et al. (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* 143, 1174-1189

172. Olsen, J. V. et al. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635-648
173. Guo, A. et al. (2013) Immunoaffinity Enrichment and Mass Spectrometry Analysis of Protein Methylation. *Mol. Cell Proteomics.*
174. Zielinska, D. F., Gnad, F., Wisniewski, J. R., and Mann, M. (2010) Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. *Cell* 141, 897-907
175. Kim, D. Y., Scalf, M., Smith, L. M., and Vierstra, R. D. (2013) Advanced proteomic analyses yield a deep catalog of ubiquitylation targets in Arabidopsis. *Plant Cell* 25, 1523-1540
176. Emanuele, M. J. et al. (2011) Global identification of modular cullin-RING ligase substrates. *Cell* 147, 459-474
177. Povlsen, L. K. et al. (2012) Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAF15 ubiquitylation in DNA-damage bypass. *Nat. Cell Biol.* 14, 1089-1098
178. Wagner, S. A. et al. (2011) A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Mol. Cell Proteomics* 10, M111
179. Kim, W. et al. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* 44, 325-340
180. Hornbeck, P. V. et al. (2012) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 40, D261-D270
181. Lamoliatte, F. et al. (2013) Targeted Identification of SUMOylation Sites in Human Proteins Using Affinity Enrichment and Paralog-specific Reporter Ions. *Mol. Cell Proteomics* 12, 2536-2550
182. Tatham, M. H., and Hay, R. T. (2009) FRET-based in vitro assays for the analysis of SUMO protease activities. *Methods Mol. Biol.* 497, 253-268
183. Mosmann, T. (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods* 65, 55-63
184. Burnette, W. N. (1981) "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* 112, 195-203
185. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* 4, 207-214
186. Domon, B., and Aebersold, R. (2006) Mass spectrometry and protein analysis. *Science* 312, 212-217
187. Denis, N. J., Vasilescu, J., Lambert, J. P., Smith, J. C., and Figeys, D. (2007) Tryptic digestion of ubiquitin standards reveals an improved strategy for identifying ubiquitinated proteins by mass spectrometry. *Proteomics.* 7, 868-874
188. Xu, G., Paige, J. S., and Jaffrey, S. R. (2010) Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat. Biotechnol.* 28, 868-873
189. Osula, O., Swatkoski, S., and Cotter, R. J. (2012) Identification of protein SUMOylation sites by mass spectrometry using combined microwave-assisted aspartic acid cleavage and tryptic digestion. *J. Mass Spectrom.* 47, 644-654
190. Azuma, Y., Arnaoutov, A., Anan, T., and Dasso, M. (2005) PIASy mediates SUMO-2 conjugation of Topoisomerase-II on mitotic chromosomes. *EMBO J.* 24, 2172-2182

191. Vertegaal, A. C. et al. (2004) A proteomic study of SUMO-2 target proteins. *J. Biol. Chem.* 279, 33791-33798
192. Denison, C. et al. (2005) A proteomic strategy for gaining insights into protein sumoylation in yeast. *Mol. Cell Proteomics* 4, 246-254
193. Ganesan, A. K. et al. (2007) Broad spectrum identification of SUMO substrates in melanoma cells. *Proteomics* 7, 2216-2221
194. Panse, V. G., Hardeland, U., Werner, T., Kuster, B., and Hurt, E. (2004) A proteome-wide approach identifies sumoylated substrate proteins in yeast. *J. Biol. Chem.* 279, 41346-41351
195. Hannich, J. T. et al. (2005) Defining the SUMO-modified proteome by multiple approaches in Saccharomyces cerevisiae. *J. Biol. Chem.* 280, 4102-4110
196. Tirard, M. et al. (2012) In vivo localization and identification of SUMOylated proteins in the brain of His6-HA-SUMO1 knock-in mice. *Proc. Natl. Acad. Sci. U. S. A* 109, 21122-21127
197. Miller, M. J., and Vierstra, R. D. (2011) Mass spectrometric identification of SUMO substrates provides insights into heat stress-induced SUMOylation in plants. *Plant Signal. Behav.* 6, 130-133
198. Yang, W. et al. (2012) Analysis of oxygen/glucose-deprivation-induced changes in SUMO3 conjugation using SILAC-based quantitative proteomics. *J. Proteome. Res.* 11, 1108-1117
199. Blomster, H. A. et al. (2010) In vivo identification of sumoylation sites by a signature tag and cysteine-targeted affinity purification. *J. Biol. Chem.* 285, 19324-19329
200. Galisson, F. et al. (2011) A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells. *Mol. Cell Proteomics* 10, M110
201. Miller, M. J., Barrett-Wilt, G. A., Hua, Z., and Vierstra, R. D. (2010) Proteomic analyses identify a diverse array of nuclear processes affected by small ubiquitin-like modifier conjugation in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A* 107, 16512-16517
202. Knuesel, M., Cheung, H. T., Hamady, M., Barthel, K. K., and Liu, X. (2005) A method of mapping protein sumoylation sites by mass spectrometry using a modified small ubiquitin-like modifier 1 (SUMO-1) and a computational program. *Mol. Cell Proteomics* 4, 1626-1636
203. Wohlschlegel, J. A., Johnson, E. S., Reed, S. I., and Yates, J. R., III. (2006) Improved identification of SUMO attachment sites using C-terminal SUMO mutants and tailored protease digestion strategies. *J. Proteome. Res.* 5, 761-770