

Dissertation zur Erlangung des Doktorgrades  
der Fakultät für Chemie und Pharmazie  
der Ludwig-Maximilians-Universität München

# **Proteomics of SUMO, the small ubiquitin-like modifier**

**Ivan Matic**

**aus**

**Šibenik, Kroatien**



**2009**

## **Erklärung**

Diese Dissertation wurde im Sinne von §13 Abs. 3 der Promotionsordnung vom  
29. Januar 1998 von Herrn Prof. Matthias Mann betreut.

## **Ehrenwörtliche Versicherung**

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfe erarbeitet.

München, am 02/04/2009

---

(Ivan Matić)

Dissertation eingereicht am 02/04/2009

Hon. - Prof. Dr. M. Mann (Erstgutachter)

Hon. - Prof. Dr. R. Fässler (Zweitgutachter)

Univ. - Prof. Dr. P. Cramer (Vorsitzender)

Univ. - Prof. Dr. Ch. Wahl-Schott

Univ. - Prof. Dr. M. Biel

Univ. - Prof. Dr. A. Vollmar

Mündliche Prüfung am 13/05/2009

# Table of Contents

Summary .....	1
1 Introduction: sumoylation in signaling pathways.....	5
1.1 Posttranslational modifications as regulators of cellular pathways .....	5
1.2 Ubiquitin and ubiquitin-like proteins .....	6
1.3 SUMO proteins .....	9
1.4 Functional consequences of SUMO modification .....	13
1.5 Interplay between different PTMs.....	15
2 Introduction: mass spectrometry-based proteomics .....	18
2.1 Mass spectrometry.....	18
2.2 Mass spectrometry-based proteomics .....	20
2.3 Quantitative proteomics .....	24
2.4 Computational MS-based proteomics .....	26
2.5 Quantitative proteomics program MaxQuant.....	29
2.6 Analysis of PTMs by mass spectrometry.....	30
2.6.1 Identification of sumoylation sites by mass spectrometry .....	31
3 Evidence for SUMO polymerization <i>in vivo</i> .....	33
4 Phosphorylation of SUMO-1 .....	47
5 Cross-talk between SUMO-2/3 and the ubiquitin-proteasome system .....	56
6 Profiling of the SUMO substrates proteome after heat shock .....	73
7 Conclusion and perspectives .....	99
8 Bibliography .....	101
9 Acknowledgments .....	112
10 Curriculum vitae .....	114



# Summary

This thesis investigates different aspects of protein sumoylation by qualitative and quantitative mass spectrometry. SUMO, a small ubiquitin-like modifier, is a highly versatile protein modifier involved in a number of biological pathways, but many aspects of sumoylation are currently unknown including most cellular substrates and its interplay with other post-translational modifications. Novel mass spectrometric methods are developed in this thesis to characterize the primary structure of the protein SUMO and direct evidence of sumoylation, ubiquitination and phosphorylation sites on SUMO are provided. The application of SILAC-based quantitative proteomics allows the identification of novel SUMO substrates and quantitative ‘systems-wide’ profiling of the SUMO substrate proteome upon perturbation of cellular systems.

The first project studies SUMO polymerization by a novel mass spectrometric strategy. Mapping sumoylation sites by mass spectrometry is technically challenging because the fragmentation of the long SUMO tryptic peptide conjugated to the target lysine produces complex overlapping MS/MS spectra. To overcome this problem we developed a mass spectrometric strategy based on the transfer of *in vitro* data to the more complex *in vivo* data and very high resolution mass spectrometry. In this way I provided the first direct evidence that SUMO-1,-2 and -3 form mixed polymers in cells. Importantly, SUMO-1 modifies SUMO-2 and SUMO-3 and since it does not contain an internal sumoylation consensus site, it is a potential terminator of poly-SUMO-2/3 chains (Matic et al, 2008b).

The second project investigates the cross-talk between sumoylation and ubiquitination. We found that a subset of SUMO-2 conjugates is subsequently ubiquitinated and

degraded. SILAC-based quantitative proteomics enabled the identification of 73 proteins modified by SUMO-2 that accumulate after treatment with proteasome inhibitor. In addition 40 proteins were desumoylated probably because of a lack of free SUMO-2 recycled from ubiquitinated proteins (Schimmel et al, 2008).

The third project raises the question if SUMO proteins can be targeted by post-translational modifications (PTMs) other than SUMO or ubiquitin. We found that serine 2, the N-terminal residue of SUMO-1 after the removal of the initiator methionine, is N-acetylated and phosphorylated *in vivo*. The unambiguous identification of the phosphopeptide was achieved by measuring its precursor and fragment ions with very high accuracy and by using the recently introduced higher collision energy (HCD) fragmentation in addition to standard peptide fragmentation. This phosphoserine is conserved through evolution as we report the same residue to be phosphorylated in yeast and *Drosophila* SUMO. This study raises important biological questions. Could serine 2 of SUMO-3 also be target of phosphorylation, thus constitute the first functional difference between SUMO-3 and SUMO-2, which contains an alanine in this position? Is the flexible SUMO N-terminal arm, target of sumoylation and phosphorylation, analogous to long unfolded histones tails, in which recruitment of proteins is regulated by PTMs? More generally, are SUMO proteins, as both “modified” and “modifying players”, central nodes in the PTMs-based signaling (Matic et al, 2008a)?

In the fourth project, I studied the global profile of the SUMO-2 substrate proteome during heat shock (HS) and heat shock recovery response by SILAC-based quantitative proteomics. Using tandem affinity purification, high accuracy mass spectrometry and novel quantitative proteomics algorithms collectively termed MaxQuant, we have detected more than 750 sumoylated proteins and quantified changes in the SUMO substrates proteome in response to HS. Notably, the patterns of sumoylation show clearly that proteins whose sumoylation is increased upon HS, lose the modification after HS recovery; conversely, the HS-induced desumoylation is not entirely recovered. In

response to HS SUMO polymerized into polySUMO chains and redistributed between a wide variety of proteins. This first systems-wide analysis of a ubiquitin-like modifier substrate proteome shows that SUMO modification plays a larger role than previously considered in the regulation of stress response. Furthermore, the functions of the substrate proteins implicate sumoylation in the control of the cell cycle, RNA and DNA metabolism, transcription and apoptosis.

Studies in this thesis have profound implications for different aspects of the emerging SUMO field and have established methods which will be useful for future directed as well as large-scale investigations of sumoylation. Direct identification of phosphorylation, ubiquitination and sumoylation sites on SUMO proteins extends the concept of modification of a protein modifier. The quantitative proteomics part of the thesis will be the basis for future studies quantitatively monitoring global changes in SUMO substrates proteomes in response to different cell stimuli.

## Publications from this thesis

(\* first author or shared first author)

1. Filip Golebiowski, **Ivan Matic\***, Michael H. Tatham, Christian Cole, Akihiro Nakamura, Geoffrey J. Barton, Matthias Mann & Ronald T. Hay. System-wide changes in the SUMO-2 proteome in response to heat stress. (*under review*).
2. Jürgen Cox, **Ivan Matic**, Maximiliane Hilger, Nagarjuna Nagaraj, Matthias Selbach, Jesper V. Olsen and Matthias Mann. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics (*accepted in Nature Protocols*)
3. Andersen JS, **Matic I**, Vertegaal AC (2009) Identification of SUMO target proteins by quantitative proteomics. *Methods in molecular biology (Clifton, NJ)* **497**: 19-31
4. Schimmel J, Larsen KM, **Matic I\***, van Hagen M, Cox J, Mann M, Andersen JS, Vertegaal AC (2008) The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle. *Mol Cell Proteomics* 7: 2107-2122
5. **Matic I\***, Macek B, Hilger M, Walther TC & Mann M (2008) Phosphorylation of SUMO-1 occurs *in vivo* and is conserved through evolution. *J Proteome Res* 7, 4050-4057.
6. **Matic I\***, van Hagen M, Schimmel J, Macek B, Ogg SC, Tatham MH, Hay RT, Lamond AI, Mann M & Vertegaal AC (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.

Impact factors (ISI Web of knowledge):

Mol Cell Proteomics: 9.4

J Proteome Res: 5.7

# 1 Introduction: sumoylation in signaling pathways

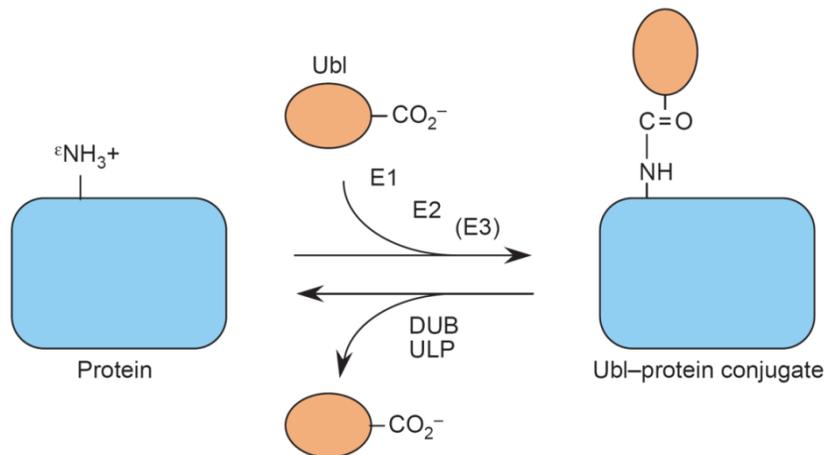
## 1.1 Posttranslational modifications as regulators of cellular pathways

The complexity and diversity of a proteome are greatly increased by reversible covalent post-translational modifications (PTMs), which compensate for the surprisingly low number of genes in vertebrate genomes. Most proteins undergo some form of PTMs, which can alter their physicochemical properties and conformation. PTMs are particularly suitable for prompt cellular response to external and internal factors since their kinetics are much faster than the regulation of protein expression levels. An intricate interplay of these modifications regulates fundamental protein properties, such as stability, localization, activity and interaction with other proteins.

Phosphorylation is one of the most common and well-studied reversible PTMs, and represents an important mechanism for the regulation of protein function (Johnson & Barford, 1993). It principally occurs on serine, threonine and tyrosine residues in eukaryotes, whereas it also targets histidine, arginine or lysine side-chains in prokaryotic proteins. The importance of phosphorylation is highlighted by the fact that a significant part of the human proteome is involved in phosphorylation or dephosphorylation: there are more than 520 kinases and more than 120 phosphatases (Manning et al, 2002).

## 1.2 Ubiquitin and ubiquitin-like proteins

PTMs include reversible covalent modifications not only by small chemical entities, such as phosphorylation, sulfation, acetylation and methylation, but also of entire small proteins. Ubiquitin (Ub) was described in 1975 as the first example of a protein that can be covalently attached to other proteins. It is a globular 76 amino acids long polypeptide that seems to be present in all eukaryotes and that is one of the most conserved proteins throughout the phylogenetic tree (Glickman & Ciechanover, 2002). It is perfectly identical in amino acid sequence from arthropods to mammals and only four amino acids are different among animals, plants and yeast (Catic & Ploegh, 2005). The presence of several Ub genes and of recycling mechanisms ensures that high levels of ubiquitin are always present, mostly in conjugated form. Its abundance and ubiquitous presence in cells give rise to its name. Ubiquitination is characterized by its diversity of conjugation products: Ub can be conjugated to target proteins as monoubiquitination (addition of a single ubiquitin residue to a single target lysine), multiubiquitination (attachment of several single Ub molecules to different lysines) or polyubiquitination (Ub polymers linked to a single lysine) (Haglund & Dikic, 2005). Polymerization on lysine 48 is the best studied one and represents a signal for proteasome-dependent protein degradation: a series of four ubiquitin molecules linked through lysine 48 is the minimal signal for proteasomal recognition (Voges et al, 1999). Chains formed via lysine 63 are involved in processes different from proteasomal degradation, such as cell signaling (Krappmann & Scheidereit, 2005), and DNA repair (Hoegge et al, 2002). Lysine 48-Ub polymers and lysine 63-chains adopt two different conformations. Polyubiquitinations formed through lysine 48 have a compact conformation, whereas lysine 63-polymers are extended (Varadan et al, 2002). All the remaining lysine residues on Ub, which in total has 7 lysines, are likewise used to form polymers, although their functional roles are less clear (Peng et al, 2003). Ub chains can also be formed by using different lysines for conjugation with the consequent creation of bifurcations of the polymer (Ikeda & Dikic, 2008).



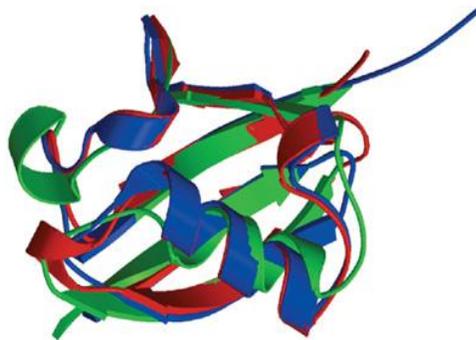
**Figure 1.1:** Conjugation/deconjugation of Ub/Ubls

Ubiquitin and ubiquitin-like proteins are covalently attached to substrates via an isopeptide bond that forms between the C-terminus of SUMO and a lysine  $\epsilon$ -amino group on the target protein. Reproduced from Mark Hochstrasser, *Nature Cell Biology* 2000.

Since the discovery of Ub, several other protein modifiers have been defined. Due to their similarity to Ub, they were named ubiquitin-like proteins (Ubls). ISG15 (Interferon-stimulated gene-15) was the first member of the Ubl family to be identified only 4 years after the discovery of Ub (Farrell et al, 1979). Successively, other Ubls were discovered: NEDD8 (neural-precursor-cell-expressed, developmentally downregulated protein 8), SUMO (small ubiquitin like modifier)-1,-2,-3, FAT10, FUB1 (Fau ubiquitin-like protein 1), UBL5, URM1 (ubiquitin related modifier 1), ATG8 (autophagy-related ubiquitin-like modifier) and ATG12 (Kerscher et al, 2006). Like ubiquitin, Ubls are small protein modifiers that covalently attach to other proteins through an isopeptide bond: the C-terminal glycine of these modifiers is linked to the  $\epsilon$ -amino group of lysine of a substrate protein (Fig.1.1) (Welchman et al, 2005). Although primary sequence similarity between them can be low, ubiquitin and Ubls share a similar conjugation mechanism (Fig. 1.1) and almost identical three-dimensional structure, the ubiquitin superfold, which is a  $\beta$ -grasp fold: 5  $\beta$  strands are folded around an  $\alpha$  helix (Fig. 1.2) (Hochstrasser, 2000).

Ub and Ubl conjugation pathways are more complex than those of most other PTMs. Protein methylation, acetylation, glycosylation and phosphorylation require the action of single enzymes, such as kinases for phosphorylation. In contrast, attachment of Ub and Ubls requires several reactions and is usually catalyzed by at least three groups of enzymes, which are part of a common enzymatic cascade that works sequentially (Fig. 1.1). First, the Ub/Ubls are activated by a specific activating enzyme (E1). The energy of the whole conjugation reaction is provided in this step by the E1-mediated hydrolysis of ATP to AMP and pyrophosphate. The activated Ub/Ubl is then transferred to a conjugating-enzyme (E2). Ub has several E2 enzymes whereas there is only one E2 in the SUMO cascade. In the last step, the Ub/Ubl is attached to its target with the aid of a ligase (E3) (Hershko & Ciechanover, 1998). This process is reversible and the removal of Ub/Ubl is mediated by specific deconjugating enzymes (Welchman et al, 2005).

A recent study showed for the first time that bacteria have an ubiquitin-like pathway previously believed only to exist in eukaryotes. The protein modifier Pup is involved in proteasome degradation in *Mycobacterium tuberculosis* by conjugation to proteasome targets, a process called pupylation (Pearce et al, 2008).



**Figure 1.2:** Conserved ubiquitin superfold

Ubiquitin and its kin are related by a common conserved three-dimensional structure. The figure shows the superimposition of ubiquitin (blue), SUMO-1 (green) and NEDD8 (red). Reproduced from Rebecca L. Welchman, Colin Gordon and R. John Mayer, *Nature Reviews Molecular Cell Biology*, 2005.

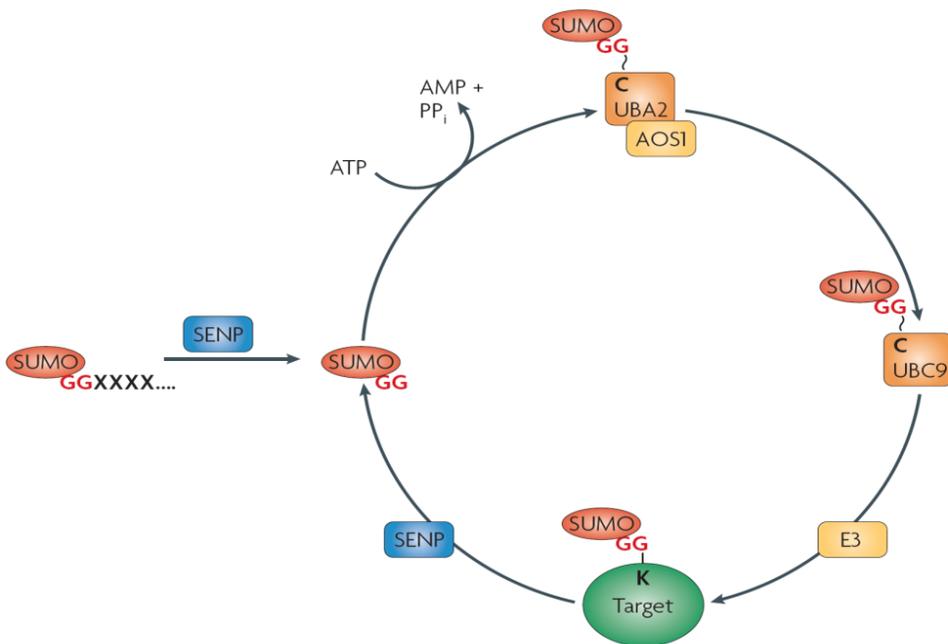
### 1.3 SUMO proteins

The Small ubiquitin-like modifier (SUMO) is the Ubl that seems to modify the largest pool of proteins and it shares approximately 20% sequence identity with ubiquitin (Johnson, 2004). Despite the similarity in the three-dimensional structures between ubiquitin and SUMO, surface charges are very different. In contrast to ubiquitin, whose surface electrostatic potential is positive, the corresponding region of SUMO is negatively charged. Depending on which organism the protein comes from, most SUMOs are around 100 residues long. Similarly to ubiquitin, SUMO is present in all eukaryotes, but absent in prokaryotes and archaea. Contrary to ubiquitin, which is a single conserved protein modifier, in some eukaryotes SUMO represents a family of paralogous proteins. Lower eukaryotes, such as *Saccharomyces cerevisiae* and *Drosophila melanogaster*, have a single SUMO protein, called Smt3 in these organisms, whereas several SUMO paralogues are expressed in vertebrates and plants. In humans there are three SUMO isoforms (SUMO-1,-2 and -3). SUMO-2 and -3, which differ between them only by three amino acids in their conjugated forms, constitute a subfamily distinct from SUMO-1, and are only 50% identical in sequence to SUMO-1 (Hay, 2005). In addition SUMO-2 and SUMO-3 isoforms have different cellular localization from SUMO-1 (Zhang et al, 2008). Whereas there is a large reservoir of unconjugated SUMO-2/3, rapidly available for conjugation after stress conditions, such as heat shock and osmotic stress, almost all SUMO-1 is attached to target proteins and is generally much less responsive to stress (Saitoh & Hinchey, 2000).

SUMO is synthesized as a precursor, whose C-terminal extension is not present in the mature form. SENP (sentrin/SUMO-specific) proteases process the newly synthesized SUMO into its conjugatable functional form by removing the additional amino acids and exposing its C-terminal diglycine motif. Mature SUMO is then activated in an ATP-dependent reaction by formation of a thioester linkage between its last glycine and the SUMO-activating E1, consisting of the heterodimer Aos1-Uba2. The Aos1 subunit uses ATP to adenylate the SUMO C-terminus, which is then followed by release of AMP and

the parallel formation of a thioester bond between SUMO and the catalytic cysteine in the Uba2 subunit. In the second reaction SUMO is transesterified to cysteine 93 of the single conjugating enzyme Ubc9. In the final step, SUMO is covalently attached to a substrate protein through the formation of an isopeptide bond between a lysine on the target protein and the C-terminal glycine of SUMO (Johnson, 2004). Differently from the ubiquitin E2 enzymes, Ubc9 interacts directly with substrates via their SUMO consensus motif, which usually resides in an unstructured part of the protein (Bernier-Villamor et al, 2002). Many, but not all, SUMO target lysines are within this Ubc9 binding motif, which has the consensus sequence  $\Psi\text{KX(E/D)}$ , where  $\Psi$  is a hydrophobic amino acid, K the acceptor lysine for conjugation and X is any amino acid (Rodriguez et al, 2001). In some substrates phosphorylation of a nearby serine (the extended motif is  $\Psi\text{KXEXXpSP}$ ) or the presence of negatively charged amino acids enhance this motif (Hietakangas et al, 2006; Yang et al, 2006). Because of the direct interaction between Ubc9 and the consensus motif, E1 and E2 enzymes are sufficient for *in vitro* sumoylation of target proteins. However the low affinity of Ubc9 for the  $\Psi\text{KX(E/D)}$  motif, with  $K_d$  in the mM range, accounts for slow reaction rates in *in vitro* conjugation assay in the absence of E3 enzymes (Lin et al, 2002). Consequently Ubc9 alone might not be sufficient for efficient *in vivo* attachment, which would require the additional aid of an E3 ligase. Most SUMO E3 ligases belong to the PIAS (protein inhibitor of activated STAT) family of proteins (Johnson & Gupta, 2001; Sachdev et al, 2001; Takahashi et al, 2001). These proteins have a motif that is similar to the RING motif in one class of Ub E3 enzymes, called SP-RING motif (Hochstrasser, 2001). They recognize Ubc9 and increase SUMO modification rates by stabilizing the binding of Ubc9 to the target protein. PIAS1, a member of the PIAS family, enhances the SUMO conjugation of the tumor suppressor p53 (Kahyo et al, 2001). RanBP2 is a SUMO E3 ligase different from any ubiquitin specific E3 enzyme (Pichler et al, 2004). A short flexible domain of RanBP2 promotes SUMO conjugation by positioning SUMO on the Ubc9 surface in a conformation that facilitates sumoylation of a target lysine (Reverter & Lima, 2005). The polycomb protein Pc2 is another type of SUMO ligase (Wotton & Merrill, 2007).

Sumoylation is reversed by removal of SUMO from its substrates by a number of deconjugating enzymes (SENPs) (Fig. 1.3). SENPs are cysteine proteases able to process SUMO precursors by exposing the C-terminal diglycine as the first step in the conjugation cascade (see above) and to deconjugate SUMO from target lysines. SENPs catalyze the removal of SUMO both directly from its substrates and from other SUMO moieties in the depolymerization of SUMO chains (Hay, 2007).



**Figure 1.3:** Reversible attachment of SUMO to target proteins

SUMO proteins are proteolytically processed by SUMO-specific isopeptidases, called SENPs (Sentrin-specific proteases) in mammals, which shorten SUMO C-terminal tail and reveal the diglycines motif necessary for the conjugation to target lysines. Mature SUMO is activated by the E1 enzyme consisting of the Aos-Ubs2 heterodimer and is then transferred to the E2 enzyme Ubc9. Finally, SUMO is conjugated to a lysine in the substrate. Reproduced from Ruth Geiss-Friedlander & Frauke Melchior, Nature Reviews Molecular Cell Biology 2007.

Intriguingly, all SUMO proteins have a long, charged and highly flexible N-terminal protrusion that is not present in ubiquitin and other Ubls (Fig. 1.4) (Bayer et al, 1998). The sequence of this extension is less conserved through evolution than the protein core. The three amino acid residues different between the otherwise identical SUMO-2 and 3 are all located in this N-terminal arm. In addition the protrusions of SUMO-2 and 3 contain the SUMO consensus motif around lysine 11, which is used for the formation of SUMO polymers (Tatham et al, 2001). Since there is no such consensus sequence on SUMO-1, it is thought not be able to be sumoylated *in vivo*, although SUMO-1 chains have been generated *in vitro* (Cooper et al, 2005; Pedrioli et al, 2006).

Ubiquitin	MQIFVKTLTG-KTITLVEPSDTIENVK
SUMO-1	MSDQEAKPSTEDLGDKKEG-EYIKLKVISQD-SSEIHFVKVMTTHLKKLK
SUMO-2	MAD-E-KPK-E-- <u>G</u> <b>V</b> <u>K</u> TEN-DHINLKVAGQDGS-VVQFKIKRHTPLSKLM
SUMO-3	MSE-E-KPK-E-- <u>G</u> <b>V</b> <u>K</u> TENNDHINLKVAGQDGS-VVQFKIKRHTPLSKLM
Ubiquitin	AKIQDKEGIPPDQQRLLIFAGKQLEDGRTLSDYNIQKESTLHLVLRRL <b>RGG</b>
SUMO-1	ESYCQRQGVPMNSLRFLFEGQRIADNHTP <b>K</b> ELGMEEEDVIEVYQEQTGG
SUMO-2	KAYCERQGLMSRQIR <b>F</b> FDGQPINETDTPAQLEMEDED <b>T</b> IDVFQQQTGG
SUMO-3	KAYCERQGLMSRQIR <b>F</b> FDGQPINETDTPAQLEMEDED <b>T</b> IDVFQQQTGG

**Figure 1.4:** Sequences of human SUMOs and ubiquitin  
 Alignment of human SUMOs and ubiquitin. Sites of the C-terminal trypsin cleavage sites are indicated in orange. The SUMO conjugation consensus sequence on SUMO-2 and -3 is underlined and the target lysine is in red. The tryptic peptides that remain conjugated to the substrates are indicated in blue. Reproduced from Ivan Matic et al, Molecular and Cellular Proteomics, 2008.

One of the consequences of ubiquitin modification is the creation of an additional noncovalent interaction surface on the substrate protein. The newly attached ubiquitin is recognized by an interaction partner through its ubiquitin binding domain (Kirkin & Dikic, 2007). Similarly to ubiquitin, SUMO-recognizing proteins have a SUMO interaction motif (SIM), which however, is a short amino acidic stretch and not an entire domain, as is the case for ubiquitin. This motif is composed of a series of hydrophobic residues, namely (V/I) (V/I)X(V/I/L), surrounded by acidic and hydrophilic amino acids and generates low-affinity interactions between a SIM-containing protein and SUMO (Kerscher, 2007). SIM forms a  $\beta$ -strand that interacts in a parallel or anti-parallel orientation with the  $\beta$ 2-strand of SUMO (Hecker et al, 2006; Song et al, 2004). SIMs are present not only in proteins that are recruited by a sumoylated target, but also in some SUMO E3 enzymes, where the motif is essential for their catalytic activity (Palvimo, 2007).

#### **1.4 Functional consequences of SUMO modification**

With the identification of new SUMO targets, it is becoming increasingly clear that, similarly to other PTMs, virtually all fundamental cellular processes are regulated at some level by sumoylation. From a molecular perspective sumoylation acts by regulating the interactions of its substrates with other cellular components. SUMO can mask or create a binding site for protein-protein interaction or result in a conformational change of the modified substrate. Components of the SUMO system have been shown to play critical roles in regulation of gene expression (Girdwood et al, 2004), RNA metabolism, DNA replication and repair (Baek, 2006), nucleo-cytoplasmic transport, neuronal survival (Lieberman, 2004), cancer development and cell cycle regulation, including roles in division, mitotic chromosome structure, cell cycle progression, kinetochore

function and cytokinesis (Dasso, 2008; Geiss-Friedlander & Melchior, 2007; Meulmeester & Melchior, 2008). Cellular stresses, such as heat shock, hibernation and osmotic stress induce SUMO conjugation (Kurepa et al, 2003; Saitoh & Hinchey, 2000). In these cases sumoylation is induced globally with many target proteins being affected, although the cellular mechanisms responsible for this global change of the SUMO conjugation pattern remain to be discovered (Tempe et al, 2008).

A puzzling aspect of sumoylation is that for most SUMO substrates only a small proportion of the cellular pool is modified by SUMO at any one time, even though the functional consequences of their sumoylation are dramatic. This enigma remains to be resolved, although some explanations have been suggested (Geiss-Friedlander & Melchior, 2007; Hay, 2007). SUMO attachment to a substrate may be responsible for its association to or dissociation from a protein complex. Once the targeted protein is included into or released from the complex, SUMO may not be necessary for the subsequent functions of the protein and it can be detached (Hay, 2005). In this view, a large percentage of SUMO substrates can be functionally active, even if, because of their rapid conjugation/deconjugation cycle, a small fraction of the protein is detectably modified.

The first identified role of SUMO was the modification of GTP-activating protein RanGAP1 (Mahajan et al, 1997; Matunis et al, 1996), which stimulates the GTPase activity of Ran. RanGAP1 is the protein with the highest proportion of SUMO modification, and this is almost exclusively with SUMO-1 modification. Upon sumoylation, RanGAP1 binds to RanBP2, a component of the nuclear pore and a SUMO E3 as mentioned above, and consequently localize to the nuclear envelop.

SUMO has an important role in the regulation of gene expression, by modifying transcription factors which constitute the largest functional category of SUMO targets (Girdwood et al, 2004). An example of SUMO-mediated activation of transcription is the sumoylation of heat shock transcription factors (HSF) 1 and 2, whose DNA binding and transcriptional activities are increased after sumoylation (Goodson et al., 2001; Hong et al., 2001). In the majority of cases SUMO inhibits transcription (Gill, 2005). Sumoylated

lysines can be localized in repression domains in transcription factors. For example glucocorticoid, progesterone, androgen and mineralcorticoid receptors have a ‘transcriptional synergy control domain’, whose sumoylation reduces the transcriptional synergy of multiple transcription factors bound to the same promoter (Hay, 2005). Another mechanism by which sumoylation reduces transcription is the SUMO-dependent recruitment of proteins to particular nuclear multiprotein structures, such as PML (promyelocytic leukaemia) bodies, which make them unavailable for activation of transcription. PML is a tumor suppressor known for its association with acute promyelocytic leukemia when fused with the retinoic acid receptor (Salomoni & Pandolfi, 2002). Sumoylated PML is responsible for the formation of PML bodies by acting as a scaffold, which recruits other SUMO-modified proteins (Ishov et al, 1999; Lallemand-Breitenbach et al, 2001; Zhong et al, 2000). The transcriptional repression in PML bodies is achieved through the recruitment of proteins repressing transcription, such as the transcription factor Sp3 (Ross et al, 2002). In other cases, transcriptional repression results from SUMO-dependent recruitment of histone deacetylases with consequent chromatin inactivation at certain promoters (Gill, 2005).

## **1.5 Interplay between different PTMs**

Protein functions are not regulated separately by different PTMs, but rather by an intricate crosstalk, in which PTMs may have an agonistic or antagonistic effect on each other (Hunter, 2007).

A connection between the ubiquitin-proteasome pathway and sumoylation was recently uncovered by studying the substrate-specificity of RNF4, a ubiquitin E3 ligase. RNF4 only conjugates ubiquitin onto PML when PML is conjugated to a SUMO polymer on lysine 160 (Lallemand-Breitenbach et al, 2008; Tatham et al, 2008). This interaction is

mediated by the four SIMs present on RNF4. PML is also an example for crosstalk between phosphorylation and sumoylation: it has been reported that arsenic-induced phosphorylation of PML facilitates its SUMO conjugation (Hayakawa & Privalsky, 2004).

Another mechanism in which phosphorylation regulates SUMO conjugation is by extending the conventional sumoylation consensus site. The phosphorylation-dependent sumoylation motif (PDSM), which is present in heat shock factor-1 among others, consists of the classical consensus motif followed by a phosphorylated serine and proline:  $\Psi$ KXEXXpSP (Hietakangas et al, 2006). By adding a negative charge in proximity to the acceptor lysine, phosphorylation facilitates SUMO conjugation. In contrast, for some other proteins, such as NF- $\kappa$ B, phosphorylation has a negative effect on sumoylation (Karin & Ben-Neriah, 2000).

PTM crosstalk can also be achieved by competition of different PTMs for the same target residue: a lysine can be alternatively modified by methylation, acetylation, ubiquitination or sumoylation (Anckar & Sistonen, 2007; Ulrich, 2005). As an example the same lysine in PCNA (proliferating cell nuclear antigen) can be targeted by monoubiquitination, lysine 63-polyubiquitination or sumoylation. PCNA is responsible for loading DNA polymerases on the DNA assuring its processivity. It regulates essential replication-related functions, which include the response to DNA repair (Moldovan et al, 2007). PCNA becomes modified after DNA damage either by a single ubiquitin molecule or by a lysine 63 polyubiquitin chain (Hoegge et al, 2002). Monoubiquitinated PCNA recruits translesion synthesis (TLS) polymerases, such as polymerases eta and iota, responsible for error-prone bypass of the DNA damage lesion. By incorporating correct or incorrect nucleotides at the lesion site, TLS polymerases assure the progression of the replication fork during DNA synthesis (Bienko et al, 2005). Polyubiquitination of PCNA is responsible for the error-free bypass, although the mechanism is still unknown. Sumoylated PCNA is not involved in DNA damage response and it associates with the Srs2 helicase to suppress recombination during normal replication (Pfander et al, 2005). Another example of a protein whose ubiquitination is opposed by sumoylation on the

same lysine is given by I $\kappa$ B. NF- $\kappa$ B is inactivated by I $\kappa$ B, which, by masking the nuclear localization signal of NF- $\kappa$ B, keeps it in the cytoplasm and thereby inactive form. NF- $\kappa$ B is activated by ubiquitination and proteasome-mediated degradation of I $\kappa$ B. SUMO competes with ubiquitin for the same lysine on I $\kappa$ B, thereby inhibiting I $\kappa$ B degradation. (Desterro et al, 1998; Perkins, 2007).

In some cases, a combination of mechanisms is used to finely modulate the function of a protein by a dynamic interplay between different PTMs. For example, sumoylation of lysine 403 in the transcription factor MEF2A (myocyte-specific enhancer factor-2A) reduces its transcriptional activity, whereas acetylation of the same site results in a transcriptionally active form. In addition to the mutual action of these two PTMs, phosphorylation at serine 408 enhances SUMO modification of MEF2A, and therefore promotes its repressed state. In contrast, its dephosphorylation leads to a switch of lysine 403 from sumoylation to acetylation (Shalizi et al, 2006).

The interplay between ubiquitin and NEDD8 is exemplified by the Mdm2 E3 ligase, which is a central component of ubiquitin and NEDD8 conjugation. Mdm2 not only promotes the ubiquitination and degradation of p53 via the proteasome, but also inhibits its transcriptional activity by mediating NEDD8 modification (Xirodimas et al, 2004).

A new concept of crosstalk between different PTMs, which is specific for protein modifiers, is recently emerging: the protein nature of Ub and Ubls makes them potential targets of different PTMs. So far, SUMO-1 has been shown to be directly targeted by two PTMs, namely removal of the initial methionine and N-terminal acetylation, which however are not considered to be regulatory PTMs (Lallemand-Breitenbach et al, 2008). The question if an Ubl can be modified by a PTM not belonging to the Ubl family, however, has remained open until now.

## 2 Introduction: mass spectrometry-based proteomics

### 2.1 Mass spectrometry

Mass spectrometry is a technique used to measure mass-to-charge ( $m/z$ ) ratios of electrically charged molecules. Generally, a mass spectrometer (MS) consists of an ion source for sample introduction, transfer to the gas phase and ionization, a mass analyzer to separate ions according to their mass to charge ( $m/z$ ) ratios and a detector, which records a signal from the separated ions and produces a mass spectrum.

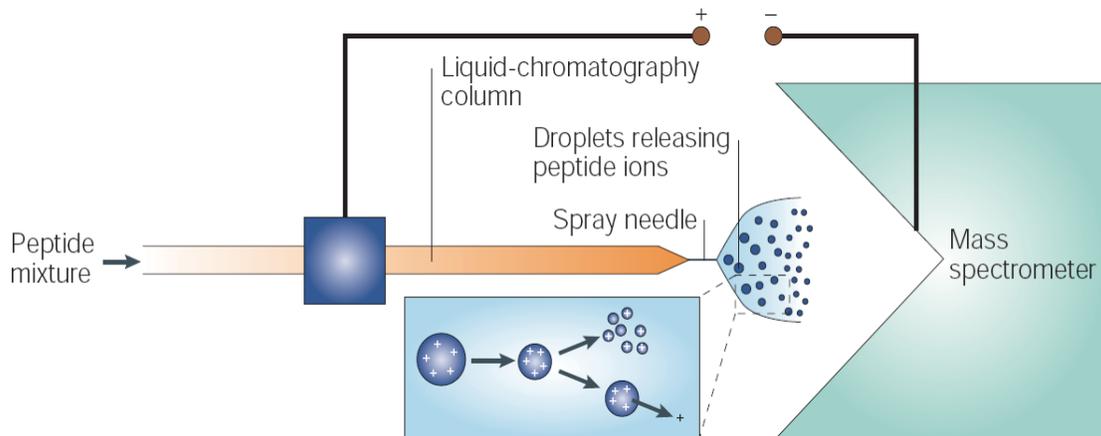


**Figure 2.1: Mass spectrometer**

The sample is injected through a column inlet into the mass spectrometer. At the end of the inlet the molecules of the sample are ionized by the ion source. The mass analyzer separates the ions by their  $m/z$  values by applying electric and/or magnetic fields. The ions abundance is quantified by a detector.

Historically, mass spectrometry was limited to the analysis of small molecules and therefore a tool for chemists but not biochemists. Macromolecules, such as peptides, are fragile and fragment when conventional ionization methods are used. In this perspective, the most important contribution to the establishment of mass spectrometry as a tool for biology was probably the development of soft ionization techniques, namely matrix assisted laser desorption/ionization (MALDI) (Karas & Hillenkamp, 1988) and electrospray ionization (ESI) (Fenn et al, 1989). In the MALDI approach, peptides or proteins are homogeneously mixed with an organic matrix and the ionization is achieved by a laser pulse. The laser energy is absorbed by the matrix, which ionizes and transfers charge to the biomolecules while protecting them from being destroyed by absorbing the laser irradiation. MALDI produces mainly singly-charged ions.

In ESI, a liquid containing the biomolecules to be studied is pumped through a capillary column, which ends in an orifice of very small diameter. A high electric potential is applied to the liquid, which is thereby electrostatically dispersed into small, highly charged droplets containing the peptides. As the volatile solvent evaporates, the droplet size decreases and the charge concentration increases. When the coulombic repulsion is larger than the droplet's surface tension, the droplets explodes, producing smaller and less highly charged droplets. Eventually, ions are fully desolvated by repeated evaporation and disintegration cycles of droplets and/or desorption of ions out of the droplets due to the high electrical field on the droplet surface. In contrast to MALDI, this ionization technique typically produces ions in multiple charged states (Steen & Mann, 2004).



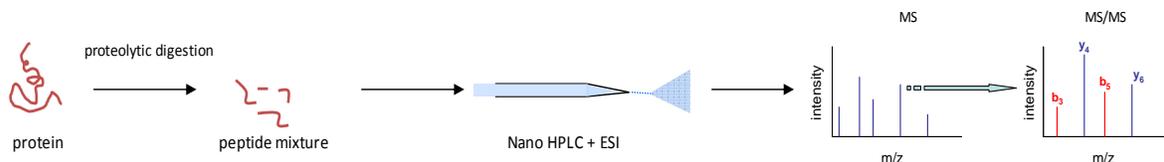
### Figure 2.2: Electrospray

Desalted and concentrated peptides are injected into a capillary chromatographic column. At the end of the column the peptides are ionized by electrospray and directly transferred into the mass spectrometer. Reproduced from Hanno Steen and Matthias Mann, *Nature Reviews Molecular Cell Biology*, 2004.

## 2.2 Mass spectrometry-based proteomics

Proteomics is defined as the large-scale study of proteins with the ultimate goal to identify and analyze all the proteins expressed by a cell or tissue (Aebersold & Mann, 2003). Before the proteomics era the application of mass spectrometry was limited to the characterization of individual peptides or proteins. Due to a number of recent technological and methodological advances, MS-based proteomics has established itself as the leading high-throughput proteomics technology and become sufficiently mature to allow large-scale studies of proteomes and sub-proteomes (Cox & Mann, 2007). The sequencing of the human genome and of other genomes has also contributed to the success of MS-based proteomics by providing protein databases, against which the experimental spectra can be matched. One of the most powerful and popular MS-based proteomics formats is the combination of liquid chromatography and tandem mass

spectrometry (LC-MS/MS), which now allows acquisition of thousands spectra in a few hours. The typical experiment starts with protein extract preparation from a biological source, such as cell culture. After one or more fractionation steps (i.e. SDS-PAGE) proteins are digested with a sequence-specific proteolytic enzyme, usually trypsin. The resulting peptide mixture is subjected to LC-MS/MS analysis. Peptides are separated by nano-flow high-performance liquid chromatography (HPLC), in which the small inner diameter of the column allows very low flow rates and high peptide concentrations and consequently a good MS signal even with few  $\mu\text{g}$  of total peptide. At the end of the column the solution containing the peptides flows through a needle and the peptides are sprayed and ionized into the mass analyzer. ESI has the great advantage over MALDI of allowing a direct on-line coupling of the MS to the HPLC (Fig. 2.2). As peptides elute from the chromatographic column, they are subjected to automated MS and MS/MS analysis: an MS spectrum is followed by a specified number of MS/MS spectra, in which fragment ions of the most abundant peptides determined in the MS scan are acquired. Then the cycle repeats with another measurement in the MS mode. The MS spectrum assigns each peptide ion an  $m/z$  value and intensity. Since this information is not sufficient by itself for a confident identification, each peptide is isolated and broken apart in the mass spectrometer to produce product ions, whose  $m/z$  values are then recorded in the MS/MS spectrum (Fig. 2.3). The peptide is identified by searching its fragmentation spectrum against an appropriate protein sequence database.



**Figure 2.3: typical LC-MS/MS workflow: from proteins to MS/MS spectra**

Proteins are digested and the generated peptides are separated by nano-flow HPLC. Eluting peptides are ionized by electrospray and analyzed by mass spectrometry. After the acquisition of precursor ions, the most intense ions are fragmented and recorded in the MS/MS mode.

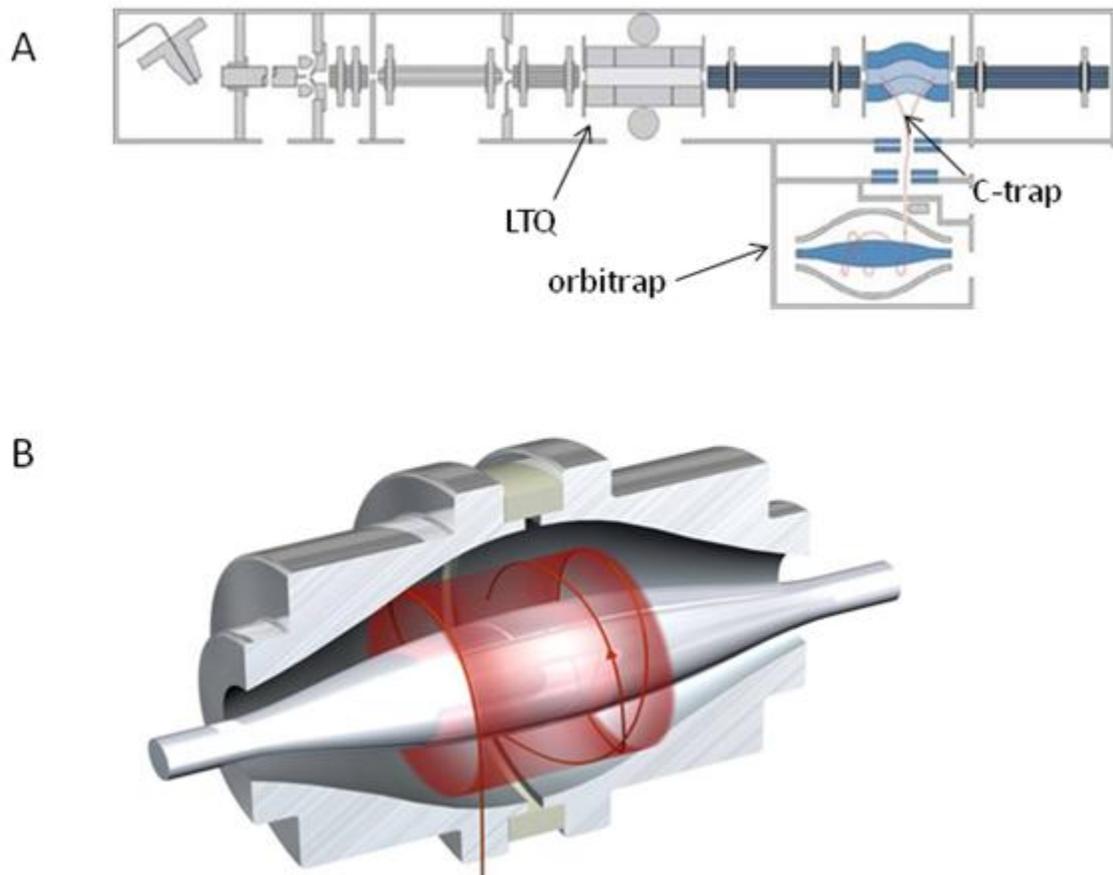
The analyzer used for the studies presented in this thesis is the hybrid linear ion trap (LTQ)-Orbitrap mass spectrometer (Scigelova & Makarov, 2006). In the most commonly used mode, the precursor ions are measured in the orbitrap and fragmented in the LTQ by collision induced fragmentation (CID) whereas the resulting fragment ions are acquired in the LTQ analyzer (Olsen et al, 2005). This configuration takes full advantage of both mass spectrometric systems, by combining the high sensitivity and scanning speed of the LTQ with the very high resolution of the orbitrap, which together with its internal calibration option produces unprecedented mass accuracy. In addition specific application, such as unambiguous identification of particularly interesting peptides, greatly benefit from the high resolution and high accuracy measurement of MS/MS spectra in the orbitrap, albeit at the cost of a slower scanning rate, and from different complementary fragmentation techniques, such as HCD (Olsen et al, 2007) and ETD (Coon et al, 2005). Peptides ionized by ESI are entrained by the gas flow due to the vacuum in the instrument. They enter the MS through a transfer capillary and are guided by electric fields until they reach the ion trapping region of the LTQ, where they are stored. From here the ions can be axially transferred and trapped in the C-trap (Fig. 2.4A), In the C-trap the ion population is compacted into a small cloud and injected onto the orbitrap, where it is electrostatically trapped and starts to orbit around the central electrode (Fig. 2.4B). The attraction to the central electrode is counterbalanced by centrifugal forces, which maintain the ions in orbit. The circular movement around the inner electrode is combined with harmonic oscillations along the central spindle at a frequency only dependent on the  $m/z$  value (Makarov, 2000). The motion along the axis of rotation is independent of rotational motion and can be used for calculation of  $m/z$  ratios:

$$\omega = \sqrt{\frac{k}{m/z}}$$

where  $\omega$  is the axial oscillation frequency and  $k$  the instrumental constant.

The oscillating ions induce an image current between the two outer orbitrap electrodes, which is then converted into an  $m/z$  spectrum using a mathematical operation called

Fourier Transformation. The harmonic nature of the oscillations is responsible for the high performance of the Orbitrap mass spectrometer with respect to the resolution and mass accuracy. The Orbitrap has a high dynamic range partly because the ions are shielded from each other by the central electrode, minimizing space charge.



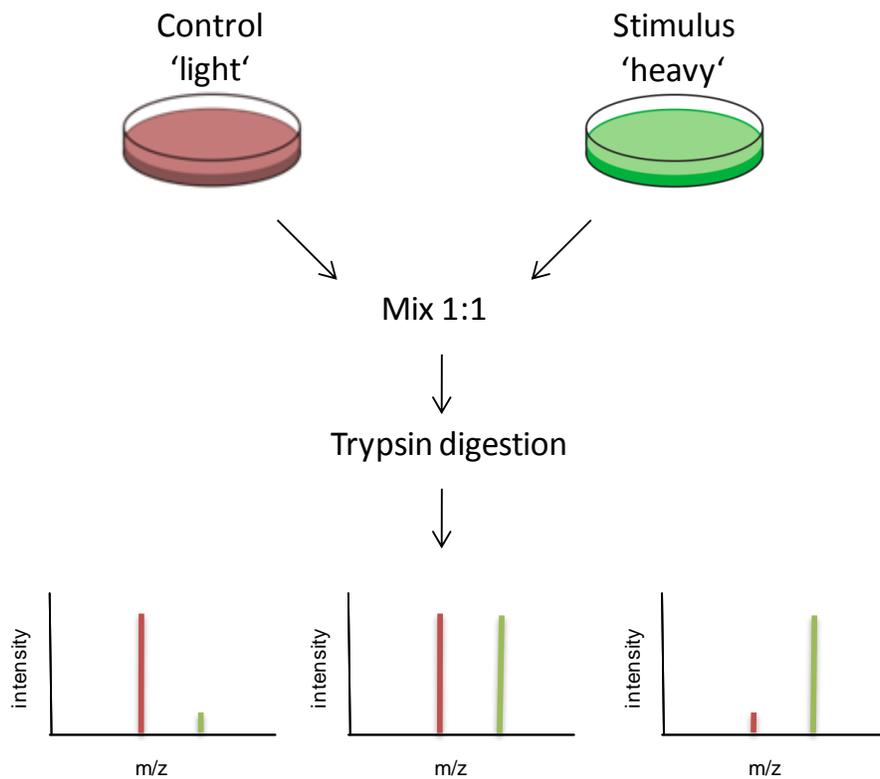
**Figure 2.4: LTQ-Orbitrap mass spectrometer**

(A) The LTQ-Orbitrap consists of the linear ion trap coupled to a C-trap for intermediate storage of ions. From the C-trap the ions enter the orbitrap. (B) The orbitrap cell consists of a coaxial central spindle-shaped electrode and an outer barrel-like electrode. The ions orbit around the inner electrode and oscillate along its long axis. Reproduced from Michaela Scigelova and Alexander Makarov, Proteomics, 2006.

## 2.3 Quantitative proteomics

Mass spectrometry is not inherently quantitative, due to different digestion and mass spectrometric behaviors of different peptide ions, including solubility, accessibility to the protease, ionization efficiency and detector response. The most accurate way to obtain quantitative information with MS involves the use of stable, non-radioactive isotopes (Ong & Mann, 2005). Peptides that differ only in mass as a result of different isotopic composition will behave identically in a proteomics experiment apart from being distinguishably detectable by MS because of their different mass. Quantitation can be either absolute or relative. The popular AQUA (Absolute Quantitation) approach (Gerber et al, 2003) consists in adding to a sample a known amount of synthetically produced, isotopically labeled peptide that mimics a proteolytic peptide present in the peptide mixture.

Relative quantification at the protein levels of two cell populations can also be achieved using stable isotopes. Stable isotope labeling by amino acids in cell culture, SILAC (Ong et al, 2002), uses the metabolic labeling of proteins with either normal or heavy isotope variants of amino acids. Generally, the method relies on growing two populations of cells; one in a medium containing the natural ('light') form of a particular amino acid, and the other in a medium that contains the isotopically labeled ('heavy') analogue of the same amino acid. When, in place of the natural form of an amino-acid, the 'heavy' form is supplied to cells in culture, its incorporation into a peptide leads to a defined mass shift compared to the 'light' peptide. This shift is easily detectable by modern MS analyzers. Typically, arginine and lysine are replaced by their  $^{13}\text{C}$  and/or  $^{15}\text{N}$  carbon labeled forms, leading to C-terminal labeling of tryptic peptides. Since the only difference between the labeled amino acid and its natural isotope counterpart is the mass, the cells grown in the heavy isotope behave exactly like the control cells cultivated in the presence of normal amino acid (Mann, 2006).



### Figure 2.5: SILAC approach

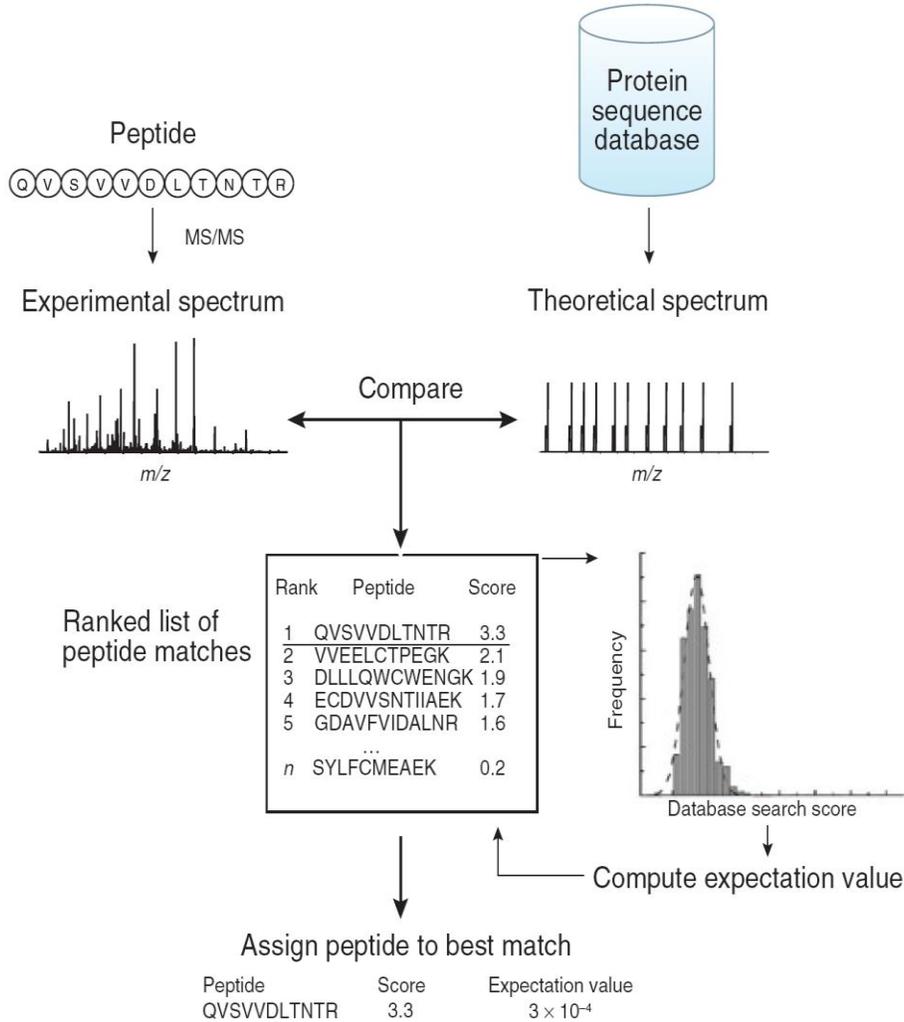
Two cell populations are grown in media containing respectively 'light' and 'heavy' isotopes of particular amino acids (i.e. arginine and lysine). One cell population is perturbed (i.e. by stimulation with a growth factor such as EGF) when the other is the control. By measuring the relative intensities of the SILAC peaks by MS one obtains the ratio of peptides and consequently of the proteins from which they are derived. Three possible scenarios are illustrated: upon stimulation a peptide can be down regulated (left spectrum), unchanged (middle spectrum) or up regulated (right spectrum).

## 2.4 Computational MS-based proteomics

Once the precursor and fragment ion spectra have been acquired, the next crucial step in the MS-based proteomics workflow is the identification of peptides and proteins from the raw spectra. The most commonly used approach to automatically assign peptide sequences to MS/MS spectra is through database searching (Colinge & Bennett, 2007). The masses from both MS and MS/MS scans are submitted to a database search engine, such as Mascot (Perkins et al, 1999) or SEQUEST (Eng et al, 1994). Measured spectra are compared against theoretical spectra from candidate sequences with matching mass obtained after *in silico* digestion of protein sequences using the enzymatic cleavage rules. For each theoretical peptide the mass of its precursor and fragment ions are calculated and a theoretical fragmentation spectrum is obtained. If the measured precursor mass is equal to its theoretical counterpart within the maximum allowed mass deviation, considering allowed types of PTMs and enzyme constraints, the acquired MS/MS spectrum is matched against the theoretical fragmentation patterns. This procedure is repeated for each peptide from the protein sequence database.

For each experimental fragmentation spectrum, the software produces a list of peptide sequences and assigns to each peptide sequence a score according to the quality of the match with the observed fragmentation pattern. As it is a measure of similarity between the calculated spectrum and the observed spectrum, this probability score indicates the likelihood that the match is correct and, therefore, that the corresponding peptide was correctly identified. Since the peptide sequence with the best matching score has the highest probability to be assigned correctly to the MS/MS spectrum, usually only this match is used for successive statistical analyses. Due to non-ideal and random matching between theoretical and experimental spectra, database searching programs inevitably produce both correct and incorrect identifications. The score given by the search software is not a good indication of the confidence of identification (i.e. to distinguish true from false positives): it is a property of a single match between fragmentation spectrum and peptide sequence and therefore does not consider factors affecting the whole population,

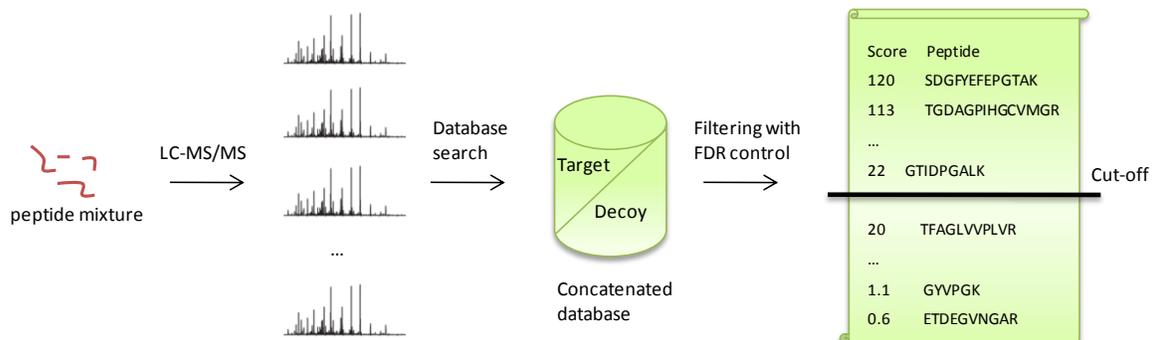
which include the size of the searched database and performance of the instrument (Nesvizhskii et al, 2007).



**Figure 2.6: Workflow of a typical database search tool**

An experimental fragmentation spectrum is matched against theoretical spectra derived from a protein sequence database. The database searching engine scores the similarity between the spectra. In addition the score can be converted to an expectation value, which is defined as the expected number of peptides that match randomly to the observed spectrum and have scores equal or higher than the reported score. Reproduced from Alexey I Nesvizhskii, Olga Vitek and Ruedi Aebersold, Nature Methods, 2007.

Global strategies, which study the distribution of the whole population of peptide scores, have been recently introduced as a better way to estimate the ambiguity of peptide identification by calculating the rate of false positives. Among these the target-decoy search approach is usually considered the best way to control false positives in large-scale proteomics studies (Elias & Gygi, 2007). The first step in this strategy is the creation of a target-decoy database: the target database contains peptide sequences representative of the analyzed peptide mixture. The decoy database is obtained by reversing or randomizing the protein sequences of the original target database. By assuming that incorrect identification of peptides are equally likely in the target and decoy database, it is possible to obtain an estimate of the number of false positives by doubling the number of hits found in the decoy portion of the database, which are incorrect identifications by definition. In the second step, the list of identified peptides is filtered according to user-specified criteria and the False Discovery Rate (FDR) is estimated from the number of decoy hits. In contrast to the search engine score or expectation value, FDR, instead of controlling the probability that any peptide is a false positive, estimates the expected proportion of false positives in the whole population.



### Figure 2.7: Target-decoy search approach

Experimental spectra are searched against a concatenated target/decoy database. The FDR of this list is estimated by counting the number of matches from the decoy database and confident identifications are separated by a user-specified cut-off.

## 2.5 MaxQuant algorithms for high confidence identification and quantitation

The introduction of high resolution, high performance mass spectrometers, such as the LTQ-Orbitrap (Scigelova & Makarov, 2006), has in principle allowed high mass accuracy of in the low or sub p.p.m. (parts per million) range to be used in peptide identification. Such a high mass accuracy restricts the searchable space of a database and thus decreases the false positive identification rate by rejecting many false positives and improves the confidence of peptide and protein identification (Zubarev & Mann, 2007). However, in practice the potential of the highly accurate masses was not fully realized in peptide identification. The novel data processing software MaxQuant (Cox & Mann, 2008a) was specifically developed to take full advantage of high resolution MS analyzers. It further improves estimated peptide mass accuracy with linear and non-linear mass recalibration and integration of multiple mass measurements over a liquid chromatographic peak. Most importantly, it introduced the concept of individualized mass accuracy depending on the signal of each peptide. Furthermore, by specifying the number of arginine and lysine residues in a SILAC experiment, MaxQuant uses these additional criteria to restrict the database search space. The software has additional features, which make it an ideal solution for the quantitative analysis of a large number of raw data, such as a method for a three dimensional (time,  $m/z$  and abundance) identification of MS peaks, methods for computing statistics at peptide and protein levels and statistically robust method for quantifying proteins.

To estimate the FDR, MaxQuant divides the forward peptide sequences from the reverse hits. A histogram of the distribution of hits is calculated separately for the two lists as a function of Mascot score and sequence length. From the separate Bayesian probability densities for true positives and false positives it is possible to determine the probability of being a false hit, given the Mascot score and the length of the peptide. This quantity, called posterior error probability (PEP), is a property of a single peptide and represents the probability that the individual peptide is incorrectly identified. Similarly to the approach described above, filtering by FDR is achieved by sorting all the identified

peptides according to their PEP, starting with the lowest. Peptides are accepted along the list until a certain FDR is reached.

## **2.6 Analysis of PTMs by mass spectrometry**

One of the main applications of mass spectrometry is the direct mapping and quantitation of PTMs of proteins. These modifications result in mass changes that can be detected during MS analysis (Witze et al, 2007). Although all kinds of PTMs can potentially be studied by MS, phosphoproteomics has been particularly successful due both to the functional importance of phosphorylation and the availability of powerful methods for detecting protein phosphorylation. Phosphopeptides are detected by the increase in the peptide mass of 80 Da, which corresponds to the addition of a phospho group, and the precise modification sites are mapped from mass shifts of the fragments in the MS/MS spectra. Large-scale *in vivo* identification of phosphorylation sites is possible mainly because of the development of methods for enrichment of phosphopeptides. The simplest and most powerful methods to enrich phosphopeptides, including titanium dioxide chromatography (Larsen et al, 2005; Pinkse et al, 2004), strong cation exchange chromatography (SCX) and immobilized metal affinity chromatography (IMAC) (Andersson & Porath, 1986), are based on the electrostatic interactions between a positively charged matrix and the negatively charged phosphate group. Building on these methods and employing triple encoding SILAC, more than 6,000 phosphopeptides were identified and quantified in a time dependent manner in HeLa cells after stimulation with EGF (Olsen et al, 2006).

### 2.6.1 Identification of sumoylation sites by mass spectrometry

More recently, similar proteomic approaches have been applied to identify peptides modified by ubiquitin and Ubls. During tryptic digestion of the ubiquitinated protein, there is no cleavage at the modified lysine and the mass of the ubiquitinated peptide is increased by the two glycines, which remain of the cleaved ubiquitin (Kirkpatrick et al, 2005). This predictable mass shift of the precursor ion, combined with the fact that the diglycine remains attached to the lysine during fragmentation, allows an easy detection of ubiquitinated peptides. However, the diglycine signature tag is not unique for ubiquitin: the ubiquitin-like proteins NEDD8 and ISG15 share the same Gly-Gly mass shift. In addition, iodoacetamide, the standard alkylating agent used in proteomics to block cysteines, can form a 2-acetamideacetamide adduct to lysines with the same atomic composition of two glycines (Nielsen et al, 2008).

Mapping of sumoylation sites poses additional challenges for detection with mass spectrometry due to the sequence of SUMO. Trypsin-digested SUMOs release larger signature tags, such as 19 and 32 amino acids respectively for mammalian SUMO-1 and SUMO-2/3, which produce many fragment ions during MS/MS fragmentation (Fig. 1.4). Therefore, although it makes the SUMO-cross-linked peptides unambiguous as compared to ubiquitin, it also makes identification challenging because the long modifying SUMO peptide leads to complex MS/MS fragmentation patterns, which are not interpretable by conventional automated database searching engines. SUMmOn (Pedrioli et al, 2006), a pattern recognition tool, in combination with low resolution mass spectrometry, has been successful in detecting peptide modified by SUMO *in vitro*, although its utility has not been tested with *in vivo* samples, which represents further challenges due to the much higher complexity of the peptide mixture and very low abundance of SUMO conjugates.

As discussed above, the success of phosphoproteomics is due to the intrinsic chemical nature of phosphopeptides, namely their strong negative charge derived from the acidic phosphate group. A similarly powerful enrichment strategy as is routinely used for phosphorylation is not currently available for peptides modified by ubiquitin, SUMO or

other Ubls. Thus large-scale application of mass spectrometry-based proteomics to the Ub/Ubl field currently relies on purification at the protein level, not the peptide level, of Ub/Ubl covalent conjugates. This is limiting for two reasons: first, it is necessary to use cell line expressing a non-endogenous tagged, often overexpressed, Ub/Ubl; secondly no information can be obtained about the modification site, which is especially critical when a protein is modified on multiple sites.

Therefore, although mass spectrometry is well suited to the analysis of sumoylation, further methodological advances would improve its applicability to the SUMO field.

### **3 Evidence for SUMO polymerization *in vivo***

**PAPER: *In vivo* identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an *in vitro* to *in vivo* strategy.**

The following article was published in the January 2008 issue of Molecular & Cellular Proteomics, pages 132-44.

It provides the first direct evidence that SUMO polymers can be formed *in vivo*. Of particular interest to the SUMO field is the evidence of the formation of mixed chains formed by SUMO-1 and SUMO-2/3. From a methodological perspective, the article introduces a novel mass spectrometric approach for the *in vivo* identification of particularly interesting peptides by transferring information obtained more easily *in vitro*.

## *In Vivo* Identification of Human Small Ubiquitin-like Modifier Polymerization Sites by High Accuracy Mass Spectrometry and an *in Vitro* to *in Vivo* Strategy<sup>§</sup>

Ivan Matic<sup>‡§</sup>, Martijn van Hagen<sup>§¶</sup>, Joost Schimmel<sup>¶</sup>, Boris Macek<sup>‡</sup>, Stephen C. Ogg<sup>||</sup>, Michael H. Tatham<sup>\*\*</sup>, Ronald T. Hay<sup>\*\*</sup>, Angus I. Lamond<sup>\*\*††</sup>, Matthias Mann<sup>‡§§</sup>, and Alfred C. O. Vertegaal<sup>¶¶</sup>

The length and precise linkage of polyubiquitin chains is important for their biological activity. Although other ubiquitin-like proteins have the potential to form polymeric chains their identification *in vivo* is challenging and their functional role is unclear. Vertebrates express three small ubiquitin-like modifiers, SUMO-1, SUMO-2, and SUMO-3. Mature SUMO-2 and SUMO-3 are nearly identical and contain an internal consensus site for sumoylation that is missing in SUMO-1. Combining state-of-the-art mass spectrometry with an “*in vitro* to *in vivo*” strategy for post-translational modifications, we provide direct evidence that SUMO-1, SUMO-2, and SUMO-3 form mixed chains in cells via the internal consensus sites for sumoylation in SUMO-2 and SUMO-3. *In vitro*, the chain length of SUMO polymers could be influenced by changing the relative amounts of SUMO-1 and SUMO-2. The developed methodology is generic and can be adapted for the identification of other sumoylation sites in complex samples. *Molecular & Cellular Proteomics* 7:132–144, 2008.

The ubiquitin family (1, 2) includes small ubiquitin-like modifiers (SUMOs)<sup>1</sup> that are similar in structure to ubiquitin (3). In contrast to the well known role of ubiquitin in protein degradation by the proteasome, SUMO conjugation does not directly target proteins for destruction (4–6). In general, sumoy-

lation regulates the function of target proteins by affecting protein-protein interactions, which can result in altered sub-cellular localization and activity. Sumoylation is essential for the viability of eukaryotic cells (7–12).

A significant number of target proteins have been identified for Smt3, the yeast SUMO family member, and for mammalian SUMOs (13). These proteomics studies have highlighted the broad cellular impact of SUMOs on processes including transcription, replication, RNA processing, translation, signaling, and transport.

Conjugation of SUMOs to target proteins, analogous to the ubiquitin system, involves E1, E2, and E3 enzymes (1, 2, 4–6). The E1 enzyme is a dimer that consists of SAE1 and SAE2, and in contrast to the large set of E2 enzymes involved in ubiquitination, a single E2 enzyme, Ubc9, is responsible for sumoylation. In addition, E3 enzymes, including protein inhibitor of activated signal transducer and activator of transcription family members and RanBP2, can enhance the sumoylation of target proteins but are not strictly required *in vitro* (14–17). Sumoylation is reversible; SUMOs can be removed from target proteins by specific SUMO proteases (4–6, 11). These proteases are also responsible for the maturation of SUMO precursors, a process that exposes the carboxyl-terminal diglycine motif that is characteristic for ubiquitin-like proteins and required for conjugation to target proteins.

Ubiquitin is able to form chains on target proteins via all seven internal lysines (18, 19). Ubiquitin chains were initially discovered by studying the role of ubiquitin in targeting protein substrates for proteolysis. These chains are Lys-48-linked polymers that mark target proteins for proteasome-mediated destruction (20). Structurally different ubiquitin chains can also play other roles in cells that are unrelated to protein degradation (18, 21). For example, Lys-63-linked chains are involved in translation, protein kinase activation, vesicle trafficking, and DNA repair. An interesting NMR study has revealed that the conformation of a Lys-63-linked ubiquitin dimer is distinct from a Lys-48-linked dimer (22). Yeast cells that express a K63R mutant of ubiquitin are compromised in DNA repair, but proteolysis is not affected in these cells (23).

From the <sup>‡</sup>Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany, <sup>¶</sup>Department of Molecular Cell Biology, Leiden University Medical Center, 2300 RC Leiden, the Netherlands, <sup>||</sup>Centre for Molecular Medicine, 61 Biopolis Drive (Proteos), Singapore 138673, Singapore, and <sup>\*\*</sup>Wellcome Trust Biocentre, University of Dundee, Dundee DD1 5EH, United Kingdom

Received, April 17, 2007, and in revised form, September 10, 2007  
Published, MCP Papers in Press, October 15, 2007, DOI 10.1074/mcp.M700173-MCP200

<sup>1</sup> The abbreviations used are: SUMO, small ubiquitin-like modifier; E1, SUMO-activating enzyme; E2, SUMO protein carrier protein; E3, SUMO ligase; HIF, hypoxia-inducible factor; LTQ, linear quadrupole ion trap; LDS, lithium dodecyl sulfate; RanBP2, Ran-binding protein 2; SAE, SUMO-activating enzyme; Ubc9, ubiquitin-conjugating enzyme 9; aa, amino acids; HRP, horseradish peroxidase; GPMW, General Protein/Mass Analysis for Windows.

Evidence for SUMO Polymerization *in Vivo*

In contrast to the extensive amount of data on ubiquitin chain formation (19, 24, 25), very little is known about multimerization of ubiquitin-like proteins. The single SUMO family member in *Saccharomyces cerevisiae*, Smt3, has been shown to form chains, but these chains are not required for viability (26). A yeast strain in which wild-type Smt3 was replaced by a lysine-deficient Smt3 mutant was viable; it had no obvious growth defects or stress sensitivities. The amount of Smt3 chains in yeast is limited due to the activity of the Smt3 protease Ulp2 (26). Interestingly Smt3 chains accumulate during meiosis (27). Here we investigated the polymerization of the three mammalian SUMOs by mass spectrometry. Trypsin digestion of ubiquitinated proteins produces diglycine-modified lysines, which are easily detected in MS and MS/MS spectra because of their predictable mass shift. In contrast, it is technically challenging to map attachment sites for human SUMO family members due to the fact that the long SUMO tryptic peptides attached to modified lysines substantially increase the mass of the peptide and also fragment during MS/MS. The resulting fragmentation patterns are very complex and not readily interpretable with currently available software for analyzing MS/MS spectra. Recently an automated pattern recognition tool (29) has been developed to overcome this limitation, but further work is needed to test its utility *in vivo*. Mutational strategies where trypsin cleavage sites are introduced close to the SUMO carboxyl-terminal diglycine (28, 30) simplify the mass spectrometric analysis but suffer from the use of non-physiological modifiers. Here we developed an alternative mass spectrometric strategy based on high resolution MS and the transfer of *in vitro* MS data to the *in vivo* data generated from very small sample amounts and high sample complexity. We used this strategy to identify conjugation sites for human SUMO family members and to unambiguously detect SUMO branched peptides. This approach allowed us to map the internal lysines that are used for SUMO chain formation and to demonstrate the ability of SUMOs to form chains *in vivo*.

## EXPERIMENTAL PROCEDURES

**Plasmids, Proteins, and Antibodies**—SUMO-1 and SUMO-2 proteins were produced in *Escherichia coli* and purified as described previously (31). GST-SUMO-1, GST-SAE2-SAE1, GST-Ubc9, and control GST were produced in *E. coli* and purified as described previously (31, 32). The GST tag was removed from the E2 by thrombin cleavage to increase the enzymatic activity. T7-HIF-1 $\alpha$ -His<sub>6</sub> (aa 373–605) was produced in *E. coli* and purified as described previously (33).

Peptide antibody AV-SM23-0100 against SUMO-2/3 was generated in a rabbit using the peptide MEDEDITIDVQQQTG (Eurogentec) (34). Monoclonal antibody 21C7 against SUMO-1 was obtained from Zymed Laboratories Inc., and monoclonal antibody 610958 against hypoxia-inducible factor-1 $\alpha$  (HIF-1 $\alpha$ ) was obtained from BD Biosciences. Anti-T7 antibody coupled to HRP was obtained from Novagen (1:5000). Secondary antibodies used were anti-rabbit HRP and anti-mouse HRP (1:5000, Pierce) and Texas Red-conjugated anti-rabbit and fluorescein isothiocyanate-conjugated anti-mouse (1:350, Jackson ImmunoResearch Laboratories).

**Electrophoresis, Silver Staining, and Immunoblotting**—Protein

samples were size-fractionated on Novex 4–12% 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol gradient gels using 4-morpholinepropanesulfonic acid buffer (Invitrogen). Total protein was visualized by silver staining. For immunoblotting experiments, size-fractionated proteins were subsequently transferred onto Hybond-C extra membranes (Amersham Biosciences) using a submarine system (Invitrogen). The membranes were incubated with specific antibodies as indicated. Bound antibodies were detected via chemiluminescence with ECL Plus (Amersham Biosciences).

**Cell Culture**—HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% FCS and 100 units/ml penicillin and streptomycin (Invitrogen). HIF-1 $\alpha$  was stabilized by 0.9 mM CoCl<sub>2</sub> (Sigma) treatment for 3 h. HeLa cells stably expressing His<sub>6</sub>-SUMO-2 were described previously (34).

**Purification of GST-SUMO Conjugates, His<sub>6</sub>-SUMO Conjugates, and Endogenous SUMO-2/3 Conjugates**—GST-SUMO-1 conjugates were obtained by incubating 20  $\mu$ g of GST-SUMO-1 or control GST with 100  $\mu$ l of HeLa nuclear extract (CILBiotech) in a buffer containing 1.5 mM ATP, 5 mM creatine phosphate (Sigma), 5 mM DTT, and 2 mM MgCl<sub>2</sub> for 2.5 h at 30 °C. GST-SUMO-1 conjugates were bound to 30  $\mu$ l of glutathione beads (GE Healthcare) for 1 h at 4 °C. Beads were successively washed with conjugation buffer, PBS, PBS containing 0.1% Triton X-100, and PBS only at 4 °C. Bound proteins were eluted successively in 8 M urea, pH 7, and NuPage LDS protein sample buffer (Invitrogen).

His<sub>6</sub>-SUMO-2 conjugates were purified essentially as described previously (34). Endogenous SUMO-2/3 conjugates were purified from HeLa cells lysed in 2% SDS, 50 mM Tris-HCl, pH 7.5, and 10 mM iodoacetamide supplemented with protease inhibitor mixture 1873580 (Roche Diagnostics GmbH) (35). Lysates were sonicated and diluted 20-fold in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5 mM  $\beta$ -mercaptoethanol, and 0.5% Nonidet P-40 supplemented with protease inhibitor mixture. Immunoprecipitations were performed with antibody AV-SM23-0100 or preimmune serum covalently cross-linked to protein G-Sepharose beads (GE Healthcare) for 3 h at room temperature. After extensive washing, bound proteins were eluted in NuPage LDS protein sample buffer (Invitrogen).

**In Vitro Sumoylation**—SUMO polymer formation described in Fig. 2A was carried out in 10- $\mu$ l volumes containing 120 ng of SAE1/2, 2 mM ATP, 0.6 units·ml<sup>-1</sup> inorganic pyrophosphatase, 10 mM creatine phosphate, 3.5 units·ml<sup>-1</sup> creatine kinase (Sigma), 5 mM MgCl<sub>2</sub>, 50 mM Tris-HCl, pH 7.5, 800 ng of Ubc9, protease inhibitor mixture, and the amounts of Ubc9, SUMO-1, and/or SUMO-2 indicated in the figure. Experiments described in Fig. 2B were carried out in 5- $\mu$ l volumes and contained the ATP regeneration mixture, 60 ng of SAE1/2, 400 ng of Ubc9, and the indicated amounts of SUMO-1 and/or SUMO-2. For mass spectrometric analysis, a similar experiment without protease inhibitors was carried out using 2  $\mu$ g of SUMO-1, 2  $\mu$ g of SUMO-2, 480 ng of SAE1/2, and 4  $\mu$ g of Ubc9 in a total volume of 40  $\mu$ l. Assays were incubated for 3 h at 37 °C before either endopeptidase Lys-C and trypsin digestion and mass spectrometric analysis or addition of SDS sample buffer for immunoblotting analysis. Aliquots representing 6% of the reaction mixtures were loaded on the gel. 5  $\mu$ g of recombinant T7-HIF-1 $\alpha$ -His<sub>6</sub> (aa 373–605) was sumoylated *in vitro* and subsequently purified in 8 M urea on Talon beads for mass spectrometric analysis.

**Mass Spectrometry and Data Analysis**—Mass spectrometric analysis was performed by nanoscale LC-MS/MS using a linear ion trap-Fourier transform-ion cyclotron resonance mass spectrometer (LTQ-FT-ICR, Thermo Fisher Scientific, Bremen, Germany) or an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with a nano-electrospray ion source (Proxeon Biosystems, Odense, Denmark) and coupled to an Agilent 1100 nano-HPLC system (Agilent Technologies) fitted with an in-house made 75- $\mu$ m reverse phase C<sub>18</sub>

Evidence for SUMO Polymerization *in Vivo*

column as described previously (36, 37). In-solution digestion was performed essentially as described previously (38). The 50-kDa band from a silver-stained gel containing GST-SUMO-1 conjugates (Fig. 4A) was excised, cut into 1-mm<sup>3</sup> cubes, and subjected to in-gel digestion according to Olsen *et al.* (37). The resulting peptides were desalted on RP-C<sub>18</sub> stop and go extraction tips (39). Peptides were eluted with a 140-min linear gradient of 98% solvent A (0.5% acetic acid in H<sub>2</sub>O) to 50% solvent B (80% acetonitrile and 0.5% acetic acid in H<sub>2</sub>O).

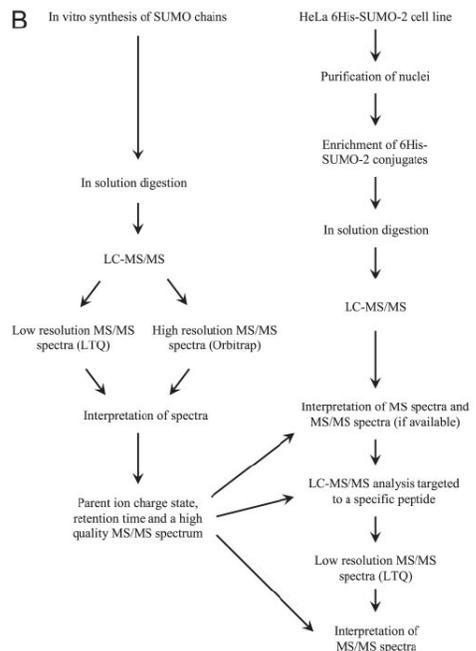
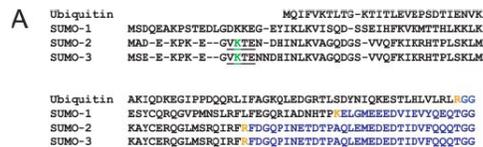
Data were acquired in the data-dependent mode: in the case of the LTQ-FT-ICR instrument, full scan spectra ( $m/z$  300–1800,  $R = 50,000$ , and ion accumulation to a target value of 3,000,000) were acquired in the ICR cell. The three most intense ions were sequentially isolated for accurate mass measurements by selected ion monitoring scans with 10-Da mass range,  $R = 50,000$ , and a target accumulation value of 50,000 and fragmented in the linear ion trap by collisionally induced dissociation followed by MS<sup>3</sup> analysis of the most intense product ion in the MS/MS scan.

In the case of the LTQ-Orbitrap, the precursor ion spectra were acquired in the orbitrap analyzer ( $m/z$  300–1600,  $R = 60,000$ , and ion accumulation to a target value of 1,000,000), and the five most intense ions were fragmented and recorded in the ion trap. In a separate experiment, peptides derived from the digestion of *in vitro* produced SUMO polymers were fragmented in the linear ion trap, and the fragment ions were recorded in the orbitrap ( $R = 15,000$ ). The lock mass option enabled accurate mass measurement in both MS and orbitrap MS/MS mode as described previously (37). Target ions already selected for MS/MS were dynamically excluded for 30 s. The detection and fragmentation of the SUMO-1/SUMO-2 parent ion, derived from the digestion of the His<sub>6</sub>-SUMO-2 and GST-SUMO-1 conjugates, were obtained with the selected ion monitoring mode with 10-Da mass range in which the dynamic exclusion option was not active and only ions with charge state equal to or larger than 4+ or unassigned were fragmented and recorded in the LTQ.

All spectra were acquired in the profile mode. The monoisotopic  $m/z$  values for the SUMO-1/SUMO-2, SUMO-1/SUMO-3, SUMO-2/SUMO-2, and SUMO-2/SUMO-3 branched precursor peptides were calculated with GPMW software (Lighthouse Data, Hanstholm, Denmark) and used to search for the corresponding ions with Xcalibur software (Thermo Fisher Scientific). Assignment was confirmed by manually interpreting all MS/MS spectra. The corresponding "virtual" peptides were fragmented *in silico* with GPMW, and the resulting  $m/z$  values were used to manually assign fragment ions to the peaks in the experimental fragmentation spectra. All reported MS/MS spectra were manually validated. Only branched peptides having an extensive coverage of  $\gamma$  ions were considered. The peptides modified by SUMO-2, containing two prolines, were required to show pronounced cleavage amino-terminal to the proline residue. Parent ion charge and retention time were derived from a pilot LC-MS run of simple peptide mixtures from an *in vitro* sumoylation reaction and used together with precursor  $m/z$  values and fragmentation spectra to search for the branched peptides in a more complex mixture.

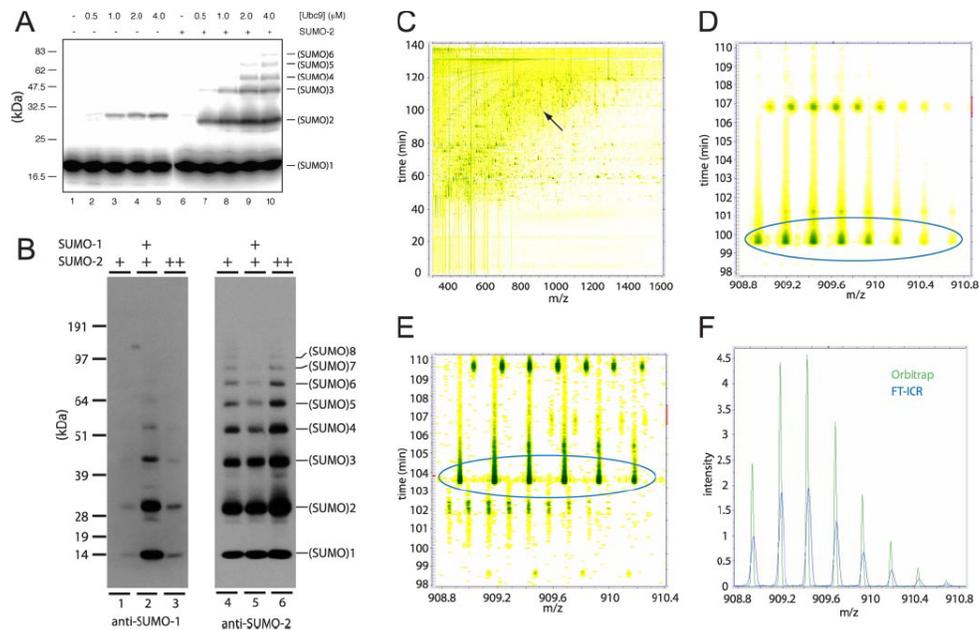
**Visualization of MS Data**—The LC-MS runs were visualized by using the Viewer tool of our in-house quantitative proteomics processing pipeline (40). The resulting two-dimensional LC-MS plots show the peptide  $m/z$  values ( $x$  axis) along the retention time axis. All the precursor isotope peaks were incorporated into the plots, and their signal intensities are color-coded with white representing the lowest intensities and green representing the highest intensities. The "full range" mode visualizes the peaks present in all MS spectra acquired during a 140-min chromatographic gradient. The enlarged view of specific SUMO branched peptides was obtained by selecting the corresponding  $m/z$  and time range and was used for a visual comparison of their abundances between different samples.

**Microscopy and Image Analysis**—HeLa cells were grown on glass



**FIG. 1. SUMO-2 and SUMO-3 contain internal sumoylation sites.** **A**, humans express three different SUMO family members, SUMO-1, SUMO-2, and SUMO-3. Lysines 11 of SUMO-2 and SUMO-3 (green) are situated within the underlined sumoylation motif  $\psi$ KX(E/D) where  $\psi$  is Val, Leu, Ile, Phe, or Met. SUMO-1 lacks a consensus sumoylation motif. These polymerization sites are located within the flexible amino-terminal extensions that are absent in ubiquitin. Sites of the carboxyl-terminal trypsin cleavage sites are indicated in orange. The peptides that remain conjugated to the targets after trypsin cleavage are indicated in blue. Note that the carboxyl-terminal SUMO-2 and SUMO-3 tryptic fragments are identical. **B**, experimental work flow. SUMO polymers were generated *in vitro* and analyzed by mass spectrometry. The obtained information was subsequently used to study SUMO multimerization in purified fractions from cells.

coverslips and fixed for 10 min in 3.7% paraformaldehyde in 37 °C PHEM buffer (60 mM PIPES, 25 mM HEPES, 10 mM EGTA, and 2 mM MgCl<sub>2</sub>, pH 6.9) (41). Subsequent manipulations were carried out at room temperature. Permeabilization was carried out for 20 min in PBS containing 0.5% Triton X-100. Cells were incubated with primary antibodies AV-SM23-0100 against SUMO-2/3 (1:2000) and 21C7

Evidence for SUMO Polymerization *in Vivo*

**Fig. 2. SUMO-1 limits the chain length of unanchored SUMO-2 polymers.** *A*, *in vitro* SUMO-1 and SUMO-2 heteroconjugate formation. *In vitro* SUMO conjugation reactions were set up containing recombinant SUMO-1 labeled with  $^{125}\text{I}$  in the absence (lanes 1–5) and presence (lanes 6–10) of unlabeled SUMO-2 and the indicated amounts of Ubc9. Reactions were incubated for 3 h at 37 °C before addition of SDS sample buffer to halt reaction progress. Samples were fractionated by SDS-PAGE. Dried gels were subjected to phosphorimaging to detect radiolabeled species. The positions of the SUMO conjugates are indicated. *B*, *in vitro* SUMO conjugation reactions were set up containing unlabeled SUMO-1 and/or unlabeled SUMO-2. Samples contained 3  $\mu\text{g}$  of SUMO-2 (lanes 1 and 4), 3  $\mu\text{g}$  of SUMO-2 plus 3  $\mu\text{g}$  of SUMO-1 (lanes 2 and 5), or 6  $\mu\text{g}$  of SUMO-2 (lanes 3 and 6). Reactions were incubated for 3 h at 37 °C before addition of LDS sample buffer to halt reaction progress. Samples were size-fractionated by SDS-PAGE, transferred to membranes, and probed using antibody 21C7 to detect SUMO-1 (lanes 1–3) or AV-SM23-0100 to detect SUMO-2 (lanes 4–6). *C–F*, mixed chains of SUMO-1 and SUMO-2 were generated *in vitro*, digested in solution with endopeptidase Lys-C and trypsin, and analyzed by mass spectrometry. *C*, the LC-MS/MS analysis using LTQ-FT-ICR is represented two-dimensionally. A very similar visualization was obtained when the LTQ-Orbitrap was used (data not shown). Peptide intensities were color-coded with *white* representing the lowest intensities and *green* representing the highest intensities. The region that contains the SUMO-1/SUMO-2 branched peptide is indicated by an arrow. Isotopic distributions of the SUMO-1/SUMO-2 branched peptide (*ellipse*) analyzed on the LTQ-FT-ICR (*D*) and the LTQ-Orbitrap (*E*) were visualized by enlarging the corresponding region. *F*, MS scans with highest intensities (*D* and *E*) were directly superimposed. LTQ-FT-ICR peaks are wider and of lower intensity, indicating a better resolution and mass accuracy for the LTQ-Orbitrap at this *m/z* value.

against SUMO-1 (1:50), washed, and incubated with secondary antibodies. DNA was stained with 0.3  $\mu\text{g}/\text{ml}$  4',6-diamidino-2-phenylindole (Sigma). After washing, cells were mounted in Vectashield (Vector Laboratories).

Three-dimensional images and sections were recorded on a Zeiss Axiovert S100 2TV DeltaVision Restoration microscope (Applied Precision) using a Zeiss Plan-Achromat 100  $\times$  1.40-numerical aperture objective and a CCD-1300-Y/HS camera (Roper Scientific). Images were captured and processed by constrained iterative deconvolution using SoftWorx (Applied Precision). Images presented here are maximal intensity projections of Z stacks.

## RESULTS

**SUMO Chain Formation *in Vitro***—Human SUMO-2 and SUMO-3 contain an internal consensus site for sumoylation

that is absent from SUMO-1 (Fig. 1A). This allows SUMO-2 and SUMO-3 to polymerize *in vitro* (31). To investigate SUMO multimerization in cells by mass spectrometry, we developed a novel strategy (Fig. 1B). Our approach is based on the idea that peptide parent ion charge state, retention time, and a high quality MS/MS spectrum can be more easily, sensitively, and accurately obtained by LC-MS/MS analysis of a low complexity mixture, such as that resulting from an *in vitro* sumoylation assay. This information can be used subsequently to formulate a modification-specific MS method for the analysis of a complex protein mixture, such as partially purified cell lysates. The spectra from the complex mixture, usually of lower quality, can be matched to the high quality



spectra derived from analysis of *in vitro* samples, allowing unambiguous characterization of peptides.

The formation of polySUMO chains was analyzed *in vitro* using  $^{125}\text{I}$ -labeled SUMO-1, unlabeled SUMO-2, the E1 and E2 enzymes, and an ATP regeneration mixture (Fig. 2A). When large amounts of E2 enzyme were used, SUMO-1 was able to form homodimers in the absence of SUMO-2. In the presence of SUMO-2, extensive mixed chain formation was observed. To study the effect of SUMO-1 on SUMO-2 chain length, a second set of *in vitro* sumoylation assays was performed with 3  $\mu\text{g}$  of SUMO-2 and either with or without 3  $\mu\text{g}$  of SUMO-1 (Fig. 2B, lanes 1, 2, 4, and 5). We noticed that SUMO-2 polymer formation decreased when SUMO-1 was added to the reaction, indicating that SUMO-1 limits SUMO-2 polymer length *in vitro*. This was not due to a lack of SUMO conjugation capacity in the mixture because adding extra SUMO-2 instead of SUMO-1 resulted in increased SUMO-2 polymer formation (Fig. 2B, lane 6). Note that the SUMO-1 antibody cross-reacted to a small extent with the relatively large amounts of SUMO-2 that were used in this assay.

Next we performed *in vitro* sumoylation assays that contained equal amounts of SUMO-1 and SUMO-2. We digested the entire reaction mixture with endopeptidase Lys-C and trypsin and analyzed the resulting peptides by LC-MS/MS on hybrid linear ion trap-Fourier transform mass spectrometers (LTQ-FT-ICR and LTQ-Orbitrap). The complexity of the peptide mixture is depicted in Fig. 2C. The most abundant peptides were isolated in the instrument and fragmented by CID, and the fragment ions were acquired in the LTQ when LTQ-FT-ICR was used and in the orbitrap when the LTQ-Orbitrap was used. In the case of the LTQ-FT-ICR mass spectrometer, the most intense product ions in the MS/MS scans were further fragmented and analyzed in the LTQ ( $\text{MS}^3$ ) to confirm the identity of the peptide (data not shown). Both instruments have very high resolution for MS spectra, and the MS/MS analysis in the orbitrap also results in very high accuracy fragmentation data albeit at the cost of using more ions. The accurate  $m/z$  values of the SUMO-1/SUMO-2 and SUMO-2/SUMO-2 branched precursor peptides were calculated and used to search for the corresponding ions. To compare the resolution of the two instruments we plotted the chromatographic regions of the 4+ charged SUMO-1/SUMO-2 precursor ions (Fig. 2, D and E) and directly superimposed the MS spectra (Fig. 2F). Because under given conditions for these

large peptides the LTQ-Orbitrap performed better in terms of resolution and mass accuracy, we decided to use this instrument for subsequent experiments.

To interpret the MS/MS spectra, we needed to calculate the  $m/z$  of the fragment ions; this is not trivial for cross-linked peptides because they can have up to four different fragmentation series (for an introduction to peptide sequencing see Ref. 42). We noticed that only a few peaks could be assigned if the modifying peptides of SUMO-1 and SUMO-2 were considered as indivisible modifications. Constructing a "virtual peptide" consisting of the entire modifying peptide joined with the carboxyl terminus of the modified peptide (Fig. 3A) allowed the assignment of the majority of peaks (except the ones arising from the three amino-terminal amino acids of the modified SUMO-2 peptide) and the identification of Lys-11 in SUMO-2 as the major internal SUMO acceptor site (Fig. 3, B and C, and supplemental Fig. S3). The isopeptide bond is chemically identical to the peptide bond, therefore no mass difference between the virtual and "real" peptides is observed. It is important to note that almost all fragment ions are composed of amino acids derived from both the modifying and modified peptides; this explains why it is not possible to interpret the MS spectra by separately fragmenting the two parts of the branched peptides *in silico*. By using our targeted approach we also found that Lys-5 of SUMO-2 can be modified by both SUMO-1 and SUMO-2 *in vitro*, but we did not detect the corresponding peptides *in vivo* (supplemental Figs. S1 and S2) (43).

Peptide identification is unambiguous in these experiments for four reasons. (i) The enzymatic digestion of the *in vitro* sumoylation reaction produces a simple mixture of peptides (Fig. 2C) that decreases the possibility that two peptides share the same  $m/z$  value. (ii) Precursor and fragment ions analyzed on the orbitrap have very high mass accuracy in the low or sub-ppm range. (iii) No SUMO-1/SUMO-2 or SUMO-2/SUMO-2 branched peptides were detected when only SUMO-1 was used in the *in vitro* reaction. The SUMO-2/SUMO-2 ion, but not the SUMO-1/SUMO-2 ion, was detected if we used only SUMO-2 in the mixture (supplemental Fig. S4). (iv) Unique characteristics of SUMO branched peptides, such as high charge state and complex fragmentation patterns, are not shared by non-sumoylated peptides.

**SUMO Chain Formation in Nuclear Extracts**—To study

Fig. 3. The SUMO-1/SUMO-2 branched peptide has a unique fragmentation pattern. Mapping sumoylation sites by mass spectrometry is challenging due to the complex MS/MS spectra that are generated. The interpretation of MS/MS spectra resulting from the fragmentation of the modified peptide is possible if the cross-linked peptide is "reversed" (A). The virtual peptide consists of the entire modifying peptide joined with the carboxyl terminus of the modified peptide. B and C, SUMO-1 is conjugated to lysine 11 of SUMO-2 *in vitro*. MS/MS fragmentation spectrum of a tryptic peptide consisting of aa 79–97 of SUMO-1 and 8–20 of SUMO-2 analyzed in the LTQ (B) or analyzed in the orbitrap (C). B, precursor ion mass was measured in the FT-ICR analyzer ( $m/z$  908.9266 (4+); mass deviation,  $-1.76$  ppm), and the peptide was fragmented and acquired in the LTQ mass spectrometer. C, precursor ion mass was measured in the orbitrap mass spectrometer ( $m/z$  908.9282 (4+); mass deviation, 0.03 ppm), and the peptide was fragmented in the LTQ and analyzed in the orbitrap. The insets are magnifications of the most abundant fragment ion at  $m/z$  1068.5 (3+). Note the much higher resolution and the isotope spacing of the orbitrap (C) compared with the LTQ (B). The low resolution of the LTQ does not allow an unambiguous assignment of the charge state.

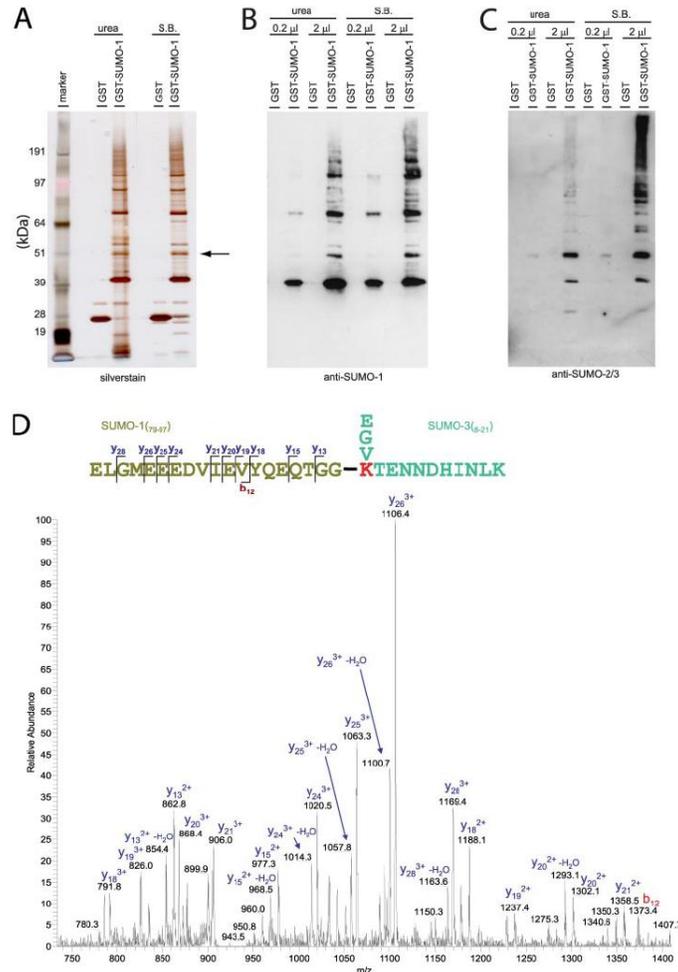
Evidence for SUMO Polymerization *in Vivo*

FIG. 4. GST-SUMO-1 is conjugated to endogenous SUMO-2 and SUMO-3 in a nuclear extract. A–C, GST-SUMO-1 or control GST was added to HeLa nuclear extracts in the presence of ATP and incubated at 30 °C for 2.5 h prior to purification on glutathione beads. The purified proteins were eluted in 8 M urea, and proteins that were still bound to the glutathione beads after urea elution were subsequently eluted in LDS protein sample buffer (S.B.). Purified fractions were size-separated by SDS-PAGE and analyzed by silver staining (A) or by immunoblotting using antibody 21C7 to detect SUMO-1 (B) or AV-SM23-0100 to detect SUMO-2 and SUMO-3 (C). D, the GST-SUMO-1-enriched fraction was digested in solution with endopeptidase Lys-C and trypsin. GST-SUMO-1 was conjugated to lysine 11 of endogenous SUMO-3. Shown is the MS/MS fragmentation spectrum of a tryptic peptide consisting of aa 79–97 of SUMO-1 and aa 8–21 of SUMO-3. Precursor ion mass was measured in the orbitrap mass spectrometer ( $m/z$  937.4390 (4+); mass deviation,  $-0.15$  ppm), and the peptide was fragmented and acquired in the LTQ mass spectrometer.

SUMO chain formation under more physiologically relevant conditions, an assay was developed that utilizes HeLa nuclear extracts as a source of E1, E2, and E3 enzymes, SUMO target

proteins, and SUMO proteases to allow dynamic sumoylation and desumoylation cycles. HeLa nuclear extracts contain significant amounts of SUMO-2/3 but hardly any detectable

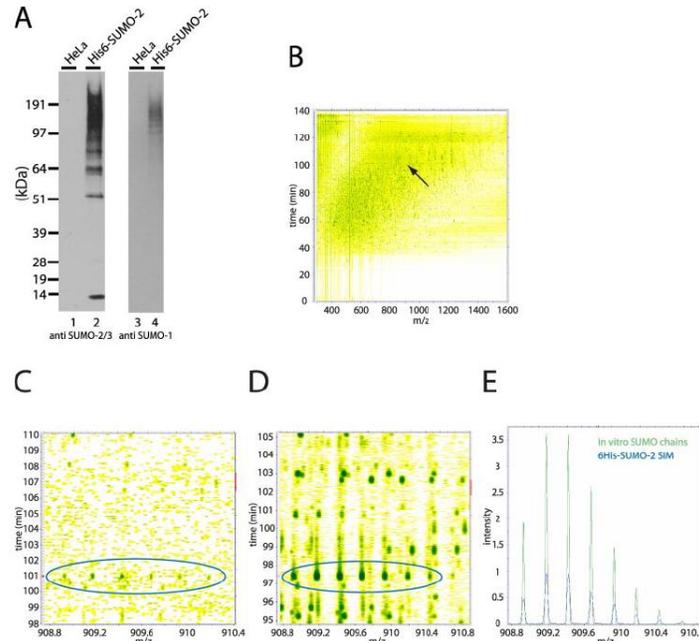


FIG. 5. SUMO polymers detected in extracts from HeLa cells. A, SUMO-2 conjugates were purified from nuclei of HeLa<sup>His6</sup>-SUMO-2 cells, separated by SDS-PAGE, transferred to membranes, and probed using antibody AV-SM23-0100 to detect SUMO-2/3 (lanes 1 and 2) or 21C7 to detect SUMO-1 (lanes 3 and 4). B–E, purified His<sub>6</sub>-SUMO-2 conjugates were digested in solution by endopeptidase Lys-C and trypsin and analyzed in the LTQ-Orbitrap. B, the LC-MS/MS analysis is represented two-dimensionally, and the region that contains the SUMO-1/SUMO-2 branched peptide is indicated by an arrow. Isotopic distributions of the SUMO-1/SUMO-2 branched peptide (*ellipse*) analyzed by the full scan method (C) or by the selected ion monitoring method (D) were visualized by enlarging the corresponding regions. E, the MS spectra of the SUMO-1/SUMO-2 branched peptide from the analysis of the *in vitro* and *in vivo* sample were directly superimposed.

amounts of SUMO-1 (data not shown). Adding ATP to this reaction mixture is sufficient to stimulate the conjugation of GST-SUMO-1 to target proteins. GST-SUMO-1 conjugates were purified from the reaction mixture and analyzed by silver staining (Fig. 4A) and by immunoblotting using an antibody that detects SUMO-1 (Fig. 4B). Many different high molecular weight SUMO-1 conjugates were detected in the GST-SUMO-1-purified fraction. Interestingly immunoblotting results indicated the presence of SUMO-2 and/or SUMO-3 in the GST-SUMO-1-enriched fraction (Fig. 4C). The most prominent band detected by this antibody suggested that SUMO-2 and/or SUMO-3 are direct targets for GST-SUMO-1. Note also that the SUMO-2/3 antibody detected the relatively large amount of monomeric GST-SUMO-1 due to apparent cross-reactivity, which was rather surprising because recombinant SUMO-1 and control GST were not detected by the antibody (44). A potential explanation could be that the antibody also recognizes an epitope that consists of a carboxyl-terminal fragment of GST and an amino-terminal fragment of SUMO-1.

To investigate SUMO-SUMO conjugates present under these conditions, the 50-kDa band recognized by both antibodies was excised from the silver-stained gel (Fig. 4A, arrow), digested with trypsin, and analyzed by mass spectrometry. The MS/MS spectrum that was obtained showed conjugation of GST-SUMO-1 to the internal sumoylation site of endogenous SUMO-2 (supplemental Fig. S5). In addition purified GST-SUMO-1 conjugates were digested in solution with endopeptidase Lys-C and trypsin and analyzed by mass spectrometry. The mass spectrometric analysis of the in-solution digestion showed that Lys-11 of endogenous SUMO-3 can be modified by SUMO-1 (Fig. 4D), whereas the SUMO-1/SUMO-2 peptide was not selected for sequencing by the software.

**SUMO Chains Purified from Cells**—Next we addressed whether SUMO chain formation occurs *in vivo* in cultured mammalian cells. Due to the activity of SUMO proteases and the lack of specific inhibitors of these proteases, SUMO conjugates are best preserved in denaturing buffers. To purify

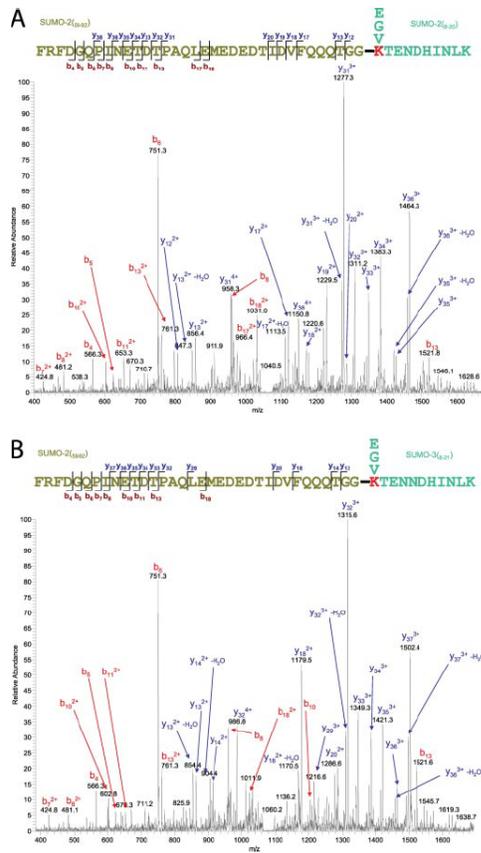
Evidence for SUMO Polymerization *in Vivo*

SUMO conjugates, we made use of our previously published stable cell line that expresses low levels of His<sub>6</sub>-SUMO-2 (34). Nuclei were prepared from this stable cell line, and His<sub>6</sub>-SUMO-2 conjugates were purified by immobilized metal affinity chromatography. The purified fraction was size-separated on a gradient gel, transferred to a membrane, and probed with antibodies to detect SUMO-2/3 and SUMO-1. SUMO-1 was detected in the His<sub>6</sub>-SUMO-2-enriched fraction (Fig. 5A, lane 4). The SUMO-1 signal in lane 4 was not due to cross-reactivity of the antibody with SUMO-2 because monomeric SUMO-2 was not detected by the SUMO-1 antibody.

The purified fraction was digested in solution with endopeptidase Lys-C and trypsin, and SUMO-SUMO conjugates in the purified fraction were studied by mass spectrometry (Fig. 5, B-E, and supplemental Fig. S7). The complexity of the peptide mixture is depicted in Fig. 5B. The intensity of the SUMO-1/SUMO-2 precursor ion was very low, almost indistinguishable from background (Fig. 5C), and as a result was not automatically selected for sequencing. However, the very low mass deviation of 0.8 ppm, agreement in retention time, and complete superimposition of the precursor isotopic distribution of this sample and of the *in vitro* sumoylation reaction indicated that the ion signal was derived from the SUMO-1/SUMO-2 peptide. To unequivocally confirm this finding we used a peptide-specific method based on the information obtained from the *in vitro* experiment (Fig. 1B and "Experimental Procedures"). Only a 10-Da range around the ion of interest was monitored in the MS mode (selected ion monitoring), dynamic exclusion of sequenced ions was disabled, and the branched peptide was repetitively fragmented and sequenced. As expected, the intensity of the precursor isotope cluster was much higher compared with the full scan detection (Fig. 5, C and D). The MS/MS spectrum confirmed the presence of the SUMO-1/SUMO-2 branched peptide in the purified fraction (supplemental Fig. S7). Note the striking similarity in MS/MS patterns in Fig. 3, B and C, and supplemental Fig. S7. Furthermore we obtained MS/MS data that confirmed the presence of SUMO-2/SUMO-2 and SUMO-2/SUMO-3 polymers (Fig. 6, A and B). Thus, we conclude that SUMOs form polymers in cultured mammalian cells.

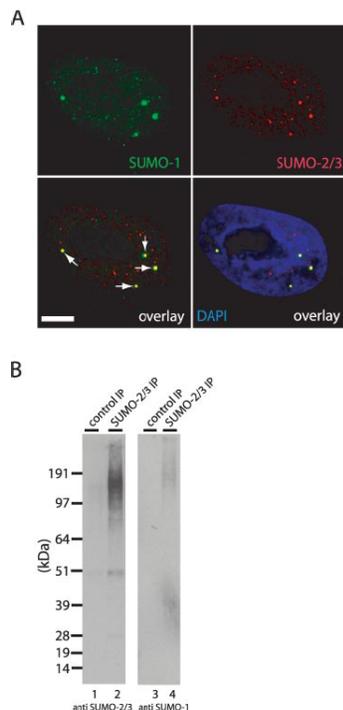
Using the same narrow mass range approach we were also able to obtain a high quality fragmentation spectrum of the SUMO-1/SUMO-2 branched peptide in the GST-SUMO-1 sample digested in solution (supplemental Fig. S6). Note that this peptide was not sequenced when a full range acquisition method was used.

Subsequently endogenous SUMOs were studied. First, the subcellular localization of SUMO-1 and SUMO-2/3 was determined by immunostaining (Fig. 7A). SUMOs are nuclear proteins that are present throughout the nucleoplasm and also accumulate in nuclear bodies (34, 45, 46). These nuclear bodies are the most prominent sites of colocalization of SUMO-1 and SUMO-2/3 and therefore could be sites where mixed SUMO chains are present in cells.



**Fig. 6. SUMO polymers detected in extracts from HeLa cells.** Purified His<sub>6</sub>-SUMO-2 conjugates were digested in solution by endopeptidase Lys-C and trypsin and analyzed in the LTQ-Orbitrap. **A**, SUMO-2 is conjugated to lysine 11 of another molecule of SUMO-2. Shown is the MS/MS fragmentation spectrum of a tryptic peptide consisting of aa 59–92 of SUMO-2 and aa 8–20 of another molecule of SUMO-2. Precursor ion mass was measured in the orbitrap mass spectrometer ( $m/z$  1070.7000 (5+); mass deviation,  $-1.03$  ppm), and the peptide was fragmented and acquired in the LTQ mass spectrometer. **B**, SUMO-2 is conjugated to lysine 11 of SUMO-3. Shown is the MS/MS fragmentation spectrum of a tryptic peptide consisting of aa 59–92 of SUMO-2 and aa 8–21 of SUMO-3. Precursor ion mass was measured in the orbitrap mass spectrometer ( $m/z$  1093.5092 (5+); mass deviation,  $-0.45$  ppm), and the peptide was fragmented and acquired in the LTQ mass spectrometer.

Second, HeLa cell lysates were prepared in a denaturing buffer, and these lysates were subsequently diluted with a milder buffer to allow the immunoprecipitation of endogenous

Evidence for SUMO Polymerization *in Vivo*

**FIG. 7. Endogenous SUMO polymers detected in extracts from HeLa cells.** A, SUMO-2/3 and SUMO-1 colocalize in nuclear bodies. HeLa cells were fixed, permeabilized, and stained with 4',6-diamidino-2-phenylindole (DAPI) to detect DNA (blue). Immunofluorescence was performed to detect SUMO-1 (green) and SUMO-2/3 (red). Nuclear bodies are enriched for both SUMO-1 and SUMO-2 as indicated by the arrows. Scale bar, 5  $\mu$ m. B, endogenous SUMO-2/3 conjugates were immunoprecipitated from total HeLa lysates, and a control immunoprecipitation (IP) was performed using preimmune serum. Proteins were separated by SDS-PAGE, transferred to membranes, and probed using antibody AV-SM23-0100 to detect SUMO-2/3 (lanes 1 and 2) or 21C7 to detect SUMO-1 (lanes 3 and 4).

SUMO-2/3 conjugates. Endogenous SUMO-1 in this immunoprecipitate could be detected by immunoblotting (Fig. 7B, lane 4) in line with our previous observations (34). To investigate whether endogenous SUMOs can form chains, we digested the SUMO-2/3 immunoprecipitate and analyzed the resulting peptide mixture by mass spectrometry. The very low peptide signals in the mass spectrometric analysis confirmed that the efficiency of the immunoprecipitation was very low (data not shown). The precursor ion of the endogenous SUMO-2/SUMO-2 conjugate and/or SUMO-3/SUMO-2 conjugate (these precursor ions are indistinguishable; Fig. 1A) was detected and sequenced (supplemental Fig. S8) showing

that endogenous SUMOs polymerize. We were not able to detect the SUMO-1/SUMO-2 peptide. This may be because this peptide is below the detection limit of our instrumentation and does not in itself prove that SUMO-1 cannot modify endogenous SUMO-2 *in vivo*.

**HIF-1 $\alpha$  Is Conjugated to SUMO Chains**—Our data clearly show that SUMO chain formation occurs in cells. However, little is known about target proteins that can be conjugated to these SUMO chains, although it has been shown that HDAC4 is attached to a SUMO-2 dimer in cells (31). Nearly all the target proteins for SUMOs reported to date are detected in mono- or disumoylated forms. This could reflect either the limited sensitivity of the detection methods used and/or the relatively low abundance of the corresponding sumoylated proteins in cells. To identify target proteins that could be conjugated to SUMO chains, we performed immunoblotting experiments using antibodies directed against previously published SUMO targets. In this screen for proteins that are conjugated to a relatively large number of SUMO molecules, we found that the transcriptional regulator HIF-1 $\alpha$  is conjugated to up to seven SUMO molecules in cells (Fig. 8A). HIF-1 $\alpha$  contains three sumoylation consensus sites, although only two appear to be utilized, *i.e.* Lys-391 and Lys-477 (47).

To investigate the molecular composition of the SUMO chains attached to HIF-1 $\alpha$  by mass spectrometry, an *in vitro* conjugation of recombinant HIF-1 $\alpha$  to a mixture of SUMO-1 and SUMO-2 was performed, and after sumoylation, HIF-1 $\alpha$  was purified from the reaction mixture. Sumoylation was confirmed by immunoblotting (Fig. 8B), and the remaining amount of protein was digested with trypsin and analyzed by mass spectrometry for the presence of SUMO/SUMO chimeric peptides. SUMO-1/SUMO-2 chimeric peptides (supplemental Fig. S9) and SUMO-2/SUMO-2 chimeric peptides (supplemental Fig. S10) were found by mass spectrometry. Thus, HIF-1 $\alpha$  is conjugated to SUMO polymers *in vitro*. In addition, we confirmed the conjugation of SUMO-2 to Lys-391 of HIF-1 $\alpha$  (supplemental Figs. S11 and S12) (47).

## DISCUSSION

One of the key features of ubiquitin is its ability to form chains (18). Here we show that the ubiquitin family members SUMO-1, SUMO-2, and SUMO-3 are also able to form multimers *in vitro* and *in vivo*. SUMO molecules are linked via internal sumoylation sites present in SUMO-2 and SUMO-3. SUMO-1 can also be incorporated in these chains; however, the absence of an internal consensus sumoylation site in SUMO-1 appears to limit further elongation of the chains.

MS/MS has been successfully applied to the identification of many types of post-translational modifications because of its unique advantage in providing direct evidence of the modified peptides. However, until now very few peptides modified by mammalian SUMO have been identified because the lack of an Arg or a Lys in the proximity of the carboxyl terminus leads to the uninterpretability of fragment ions spectra in a

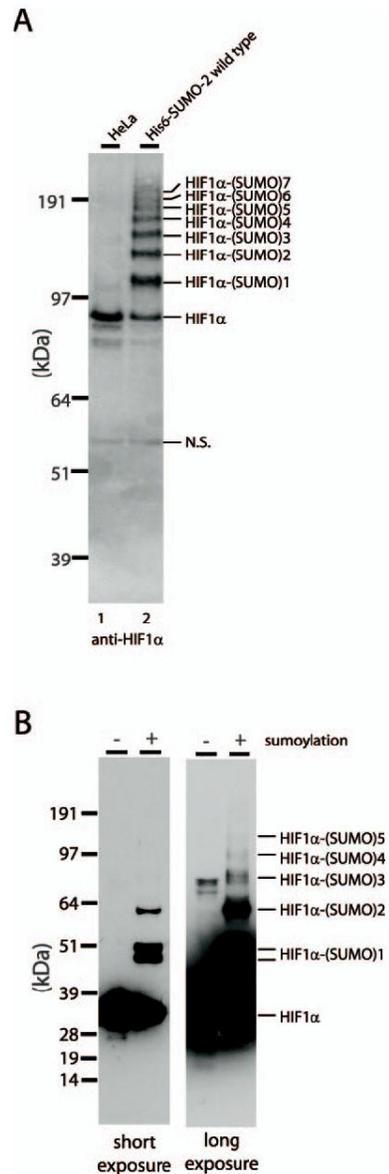
Evidence for SUMO Polymerization *in Vivo*

FIG. 8. SUMO chain formation on HIF-1 $\alpha$ . A, His<sub>6</sub>-SUMO conjugates were purified from nuclei of cells stably expressing His<sub>6</sub>-SUMO-2 and from nuclei of control HeLa cells, separated by SDS-

straightforward way. This has limited progress in the field toward understanding the biological roles of this class of modification. Different approaches have been proposed such as mutational strategies (28, 30) to obtain a short indivisible sequence linked to the modified lysine after digestion or a pattern recognition tool (28, 29) to interpret the complex overlapping MS/MS spectra. The mutational method has the advantage that standard sequencing software can be used successfully, but the disadvantage of working with a SUMO variant is that the conjugation efficiency and other properties could differ from the wild type in cells. The pattern recognition software has been successful in detecting peptides modified by wild-type SUMO *in vitro* even when using relatively low accuracy mass spectrometers. The successful identification of the mammalian SUMO modification sites mapped in this study using mass spectrometry has been achieved by first analyzing proteins sumoylated *in vitro*. The low abundance of SUMO conjugates in cells and the complexity of the *in vivo* samples represents a further challenge. In our experience not more than 20% of peptides present in a complex mixture are sequenced in an LC-MS/MS experiment. The low intensity precursor ions generally are not fragmented when a standard sequencing method is used.

By taking advantage of the fact that SUMO-2 and SUMO-3 have a known internal consensus sumoylation site, we applied a targeted approach for the specific and direct detection of the SUMO modification at this site. The same can be done with other specific proteins especially if consensus sumoylation sites are present.

It has been shown that chain formation of the single SUMO family member in *S. cerevisiae* is limited due to the activity of the SUMO protease Ulp2 (26). It is likely that SUMO proteases also play a role in mammalian cells to limit steady state levels of SUMO chain length (48). In addition, we have shown here that the presence of three mammalian SUMO family members in principle allows an alternative mechanism to regulate SUMO chain length.

The most obvious way for forming SUMO chains would be the sequential addition of SUMO-1, SUMO-2, or SUMO-3 to a pre-existing SUMO-2 or SUMO-3 molecule on a target protein. It will be important to determine in the future whether unanchored SUMO chains are formed *in vivo*. In this regard, our *in vitro* data demonstrate that in principle the conjugation machinery can effectively link together unanchored SUMO molecules in the absence of other protein targets even in the absence of E3 factors. Interestingly SUMO chain formation *in*

PAGE, transferred to a membrane, and probed with an antibody to detect HIF-1 $\alpha$ . B, 5  $\mu$ g of recombinant HIF-1 $\alpha$  was conjugated *in vitro* to a mixture of 2  $\mu$ g of SUMO-1 and 2  $\mu$ g of SUMO-2. HIF-1 $\alpha$  was subsequently purified from the reaction mixture on Talon beads in 8 M urea, and a small aliquot of the purified fraction and control HIF-1 $\alpha$  were size-fractionated by SDS-PAGE, transferred to a membrane, and probed using anti-T7 antibody to detect HIF-1 $\alpha$ . N.S., nonspecific.

*in vitro* was recently found to be enhanced by noncovalent interaction between Ubc9 and SUMO (43, 49).

Adding E3 ligases to *in vitro* sumoylation reactions has been reported previously to increase SUMO chain formation. A small fragment of RanBP2 was reported to be hypermodified by SUMO-1 chains *in vitro* (16). The SUMO-1 lysines 6, 16, 17, 37, 39, and 46 were all shown to function as acceptor sites for other SUMO-1 molecules in this case (29, 50). In addition, the small RanBP2 fragment is able to multimerize SUMO-2, a molecular event involving the internal sumoylation consensus site and two non-consensus sites (50). The yeast protein inhibitor of activated signal transducer and activator of transcription family member Siz1 was reported to enhance the multimerization of Smt3 (*i.e.* yeast SUMO) in an *in vitro* system in the presence and in the absence of septin target proteins (15). Most likely, E3 ligases will also play a role in SUMO polymerization in cells.

Ubiquitin chain formation is best known for its role in targeting conjugated proteins to the proteasome for degradation, a process that involves chain formation via Lys-48 or Lys-29 (18). Ubiquitin also forms chains via other internal lysines such as Lys-63, and for yeast ubiquitin, it has even been reported that chain formation can occur via all seven internal lysines (19). The functional relevance of SUMO chain formation remains to be determined. So far, there is little evidence for a role of sumoylation in protein degradation, although sumoylation has been linked to the degradation of DNA topoisomerase II $\beta$  by a catalytic inhibitor (51), and SUMO modification of the promyelocytic leukemia protein-retinoic acid receptor  $\alpha$  fusion protein is required for arsenic-induced, proteasome-dependent degradation (52).

**Acknowledgments**—We thank A. Groot for the HIF-1 $\alpha$  expression construct. Juergen Cox is acknowledged for visualization of the MS data.

\* This work was supported in part by the Netherlands Organisation for Scientific Research (NWO) (to A. C. O. V.) as part of the Innovative Research Incentives Scheme, by The European Community (RUBICON, VI Framework) (to M. M.), a grant from the Association for International Cancer Research (to R. T. H.), and by a Wellcome Trust Programme grant (to A. I. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ The on-line version of this article (available at <http://www.mcponline.org>) contains supplemental materials.

§ Both authors contributed equally to this work.

‡‡ A Wellcome Trust Principal Research Fellow.

§§ To whom correspondence may be addressed. Tel.: 49-89-8578-2557; Fax: 49-89-8578-2219; E-mail: mmann@biochem.mpg.de.

¶¶ To whom correspondence may be addressed. Tel.: 31-71-526-9621; Fax: 31-7-526-8270; E-mail: vertegaal@lumc.nl.

#### REFERENCES

- Welchman, R. L., Gordon, C., and Mayer, R. J. (2005) Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat. Rev. Mol. Cell Biol.* **6**, 599–609
- Pickart, C. M., and Eddins, M. J. (2004) Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta* **1695**, 55–72
- Bayer, P., Arndt, A., Metzger, S., Mahajan, R., Melchior, F., Jaenicke, R., and Becker, J. (1998) Structure determination of the small ubiquitin-related modifier SUMO-1. *J. Mol. Biol.* **280**, 275–286
- Gill, G. (2004) SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? *Genes Dev.* **18**, 2046–2059
- Hay, R. T. (2005) SUMO: a history of modification. *Mol. Cell* **18**, 1–12
- Johnson, E. S. (2004) Protein modification by SUMO. *Annu. Rev. Biochem.* **73**, 355–382
- Hayashi, T., Seki, M., Maeda, D., Wang, W., Kawabe, Y., Seki, T., Saitoh, H., Fukagawa, T., Yagi, H., and Enomoto, T. (2002) Ubc9 is essential for viability of higher eukaryotic cells. *Exp. Cell Res.* **280**, 212–221
- Johnson, E. S., Schwienhorst, I., Dohmen, R. J., and Blobel, G. (1997) The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *EMBO J.* **16**, 5509–5519
- Jones, D., Crowe, E., Stevens, T. A., and Candido, E. P. (2002) Functional and phylogenetic analysis of the ubiquitylation system in *Caenorhabditis elegans*: ubiquitin-conjugating enzymes, ubiquitin-activating enzymes, and ubiquitin-like proteins. *Genome Biol.* **3**, RESEARCH0002
- Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P., and Ahringer, J. (2003) Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231–237
- Li, S. J., and Hochstrasser, M. (1999) A new protease required for cell-cycle progression in yeast. *Nature* **398**, 246–251
- Nacerddine, K., Lehembre, F., Bhaumik, M., Artus, J., Cohen-Tannoudji, M., Babinet, C., Pandolfi, P. P., and Dejean, A. (2005) The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev. Cell* **9**, 769–779
- Xu, P., and Peng, J. (2006) Dissecting the ubiquitin pathway by mass spectrometry. *Biochim. Biophys. Acta* **1764**, 1940–1947
- Hochstrasser, M. (2001) SP-RING for SUMO: new functions bloom for a ubiquitin-like protein. *Cell* **107**, 5–8
- Johnson, E. S., and Gupta, A. A. (2001) An E3-like factor that promotes SUMO conjugation to the yeast septins. *Cell* **106**, 735–744
- Pichler, A., Gast, A., Seeler, J. S., Dejean, A., and Melchior, F. (2002) The nucleoporin RanBP2 has SUMO1 E3 ligase activity. *Cell* **108**, 109–120
- Sachdev, S., Bruhn, L., Sieber, H., Pichler, A., Melchior, F., and Grosschedl, R. (2001) PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. *Genes Dev.* **15**, 3068–3103
- Pickart, C. M., and Fushman, D. (2004) Polyubiquitin chains: polymeric protein signals. *Curr. Opin. Chem. Biol.* **8**, 610–616
- Peng, J., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D., and Gygi, S. P. (2003) A proteomics approach to understanding protein ubiquitination. *Nat. Biotechnol.* **21**, 921–926
- Chau, V., Tobias, J. W., Bachmair, A., Marriot, D., Ecker, D. J., Gonda, D. K., and Varshavsky, A. (1989) A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **243**, 1576–1583
- Sun, L., and Chen, Z. J. (2004) The novel functions of ubiquitination in signaling. *Curr. Opin. Cell Biol.* **16**, 119–126
- Varadan, R., Assfalg, M., Haririnia, A., Raasi, S., Pickart, C., and Fushman, D. (2004) Solution conformation of Lys63-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling. *J. Biol. Chem.* **279**, 7055–7063
- Spence, J., Sadis, S., Haas, A. L., and Finley, D. (1995) A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. *Mol. Cell Biol.* **15**, 1265–1273
- Hoeller, D., Crosetto, N., Blagoev, B., Raiborg, C., Tikkanen, R., Wagner, S., Kowanez, K., Breitling, R., Mann, M., Stenmark, H., and Dikic, I. (2006) Regulation of ubiquitin-binding proteins by monoubiquitination. *Nat. Cell Biol.* **8**, 163–169
- Kirkpatrick, D. S., Weldon, S. F., Tsapralis, G., Liebler, D. C., and Gandolfi, A. J. (2005) Proteomic identification of ubiquitinated proteins from human cells expressing His-tagged ubiquitin. *Proteomics* **5**, 2104–2111
- Bylebyl, G. R., Belichenko, I., and Johnson, E. S. (2003) The SUMO isopeptidase Ulp2 prevents accumulation of SUMO chains in yeast. *J. Biol. Chem.* **278**, 44113–44120
- Cheng, C. H., Lo, Y. H., Liang, S. S., Ti, S. C., Lin, F. M., Yeh, C. H., Huang, H. Y., and Wang, T. F. (2006) SUMO modifications control assembly of synaptonemal complex and polycomplex in meiosis of *Saccharomyces*

Evidence for SUMO Polymerization *in Vivo*

- cerevisiae. *Genes Dev.* **20**, 2067–2081
28. 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
29. Pedrioli, P. G., Raught, B., Zhang, X. D., Rogers, R., Aitchison, J., Matunis, M., and Aebersold, R. (2006) Automated identification of SUMOylation sites using mass spectrometry and SUMOn pattern recognition software. *Nat. Methods* **3**, 533–539
30. 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
31. Tatham, M. H., Jaffray, E., Vaughan, O. A., Desterro, J. M., Botting, C. H., Naismith, J. H., and Hay, R. T. (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
32. Mohrmann, L., Kal, A. J., and Verrijzer, C. P. (2002) Characterization of the extended Myb-like DNA-binding domain of trithorax group protein Zeste. *J. Biol. Chem.* **277**, 47385–47392
33. Groot, A. J., Verheesen, P., Westerlaken, E. J., Gort, E. H., van der Groep, P., Bovenschen, N., van der Wall, E., van Diest, P. J., and Shvarts, A. (2006) Identification by phage display of single-domain antibody fragments specific for the ODD domain in hypoxia-inducible factor 1 $\alpha$ . *Lab. Invest.* **86**, 345–356
34. Vertegaal, A. C., Ogg, S. C., Jaffray, E., Rodriguez, M. S., Hay, R. T., Andersen, J. S., Mann, M., and Lamond, A. I. (2004) A proteomic study of SUMO-2 target proteins. *J. Biol. Chem.* **279**, 33791–33798
35. Jaffray, E. G., and Hay, R. T. (2006) Detection of modification by ubiquitin-like proteins. *Methods* **38**, 35–38
36. Olsen, J. V., Ong, S. E., and Mann, M. (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **3**, 608–614
37. Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4**, 2010–2021
38. Foster, L. J., de Hoog, C. L., and Mann, M. (2003) Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 5813–5818
39. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75**, 663–670
40. Cox, J., and Mann, M. (2007) Is proteomics the new genomics? *Cell* **130**, 395–398
41. Trinkle-Mulcahy, L., Andrews, P. D., Wickramasinghe, S., Sleeman, J., Prescott, A., Lam, Y. W., Lyon, C., Swedlow, J. R., and Lamond, A. I. (2003) Time-lapse imaging reveals dynamic relocalization of PP1 $\gamma$  throughout the mammalian cell cycle. *Mol. Biol. Cell* **14**, 107–117
42. Steen, H., and Mann, M. (2004) The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* **5**, 699–711
43. Knipscheer, P., van Dijk, W. J., Olsen, J. V., Mann, M., and Sixma, T. K. (2007) Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation. *EMBO J.* **26**, 2797–2807
44. Vertegaal, A. C., Andersen, J. S., Ogg, S. C., Hay, R. T., Mann, M., and Lamond, A. I. (2006) Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. *Mol. Cell. Proteomics* **5**, 2298–2310
45. Mahajan, R., Gerace, L., and Melchior, F. (1998) Molecular characterization of the SUMO-1 modification of RanGAP1 and its role in nuclear envelope association. *J. Cell Biol.* **140**, 259–270
46. 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
47. Bae, S. H., Jeong, J. W., Park, J. A., Kim, S. H., Bae, M. K., Choi, S. J., and Kim, K. W. (2004) Sumoylation increases HIF-1 $\alpha$  stability and its transcriptional activity. *Biochem. Biophys. Res. Commun.* **324**, 394–400
48. Mukhopadhyay, D., Ayaydin, F., Kelli, N., Tan, S. H., Anan, T., Kametaka, A., Azuma, Y., Wilkinson, K. D., and Dasso, M. (2006) SUSP1 antagonizes formation of highly SUMO2/3-conjugated species. *J. Cell Biol.* **174**, 939–949
49. Capili, A. D., and Lima, C. D. (2007) Structure and analysis of a complex between SUMO and Ubc9 illustrates features of a conserved E2-Ubl interaction. *J. Mol. Biol.* **369**, 608–618
50. Cooper, H. J., Tatham, M. H., Jaffray, E., Heath, J. K., Lam, T. T., Marshall, A. G., and Hay, R. T. (2005) Fourier transform ion cyclotron resonance mass spectrometry for the analysis of small ubiquitin-like modifier (SUMO) modification: identification of lysines in RanBP2 and SUMO targeted for modification during the E3 autoSUMOylation reaction. *Anal. Chem.* **77**, 6310–6319
51. Isik, S., Sano, K., Tsutsui, K., Seki, M., Enomoto, T., Saitoh, H., and Tsutsui, K. (2003) The SUMO pathway is required for selective degradation of DNA topoisomerase II $\beta$  induced by a catalytic inhibitor ICRF-193(1). *FEBS Lett.* **546**, 374–378
52. Lallemand-Breitenbach, V., Zhu, J., Puvion, F., Koken, M., Honore, N., Doubekovsky, A., Duprez, E., Pandolfi, P. P., Puvion, E., Freemont, P., and de The, H. (2001) Role of promyelocytic leukemia (PML) sumolation in nuclear body formation, 11S proteasome recruitment, and As<sub>2</sub>O<sub>3</sub>-induced PML or PML/retinoic acid receptor  $\alpha$  degradation. *J. Exp. Med.* **193**, 1361–1371

# 4 Phosphorylation of SUMO-1

**PAPER: Phosphorylation of SUMO-1 occurs *in vivo* and is conserved through evolution.**

The following article was published in the September 2008 issue of Journal of Proteome Research, pages 4050-4057.

It shows for the first time that a protein modifier can be targeted by a regulatory post-translational modification not-belonging to the ubiquitin-like protein family, raising the concept of a 'modified modifier'.

## Phosphorylation of SUMO-1 Occurs *in Vivo* and Is Conserved through Evolution

Ivan Matic,<sup>†</sup> Boris Macek,<sup>†</sup> Maximiliane Hilger,<sup>†</sup> Tobias C. Walther,<sup>‡</sup> and Matthias Mann<sup>\*†</sup>

Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany, and Organelle Architecture and Dynamics, Max Planck Institute for Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany

Received May 20, 2008

Protein dynamics is regulated by an elaborate interplay between different post-translational modifications. Ubiquitin and ubiquitin-like proteins (Ubls) are small proteins that are covalently conjugated to target proteins with important functional consequences. One such modifier is SUMO, which mainly modifies nuclear proteins. SUMO contains a unique N-terminal arm not present in ubiquitin and other Ubls, which functions in the formation of SUMO polymers. Here, we unambiguously show that serine 2 of the endogenous SUMO-1 N-terminal protrusion is phosphorylated *in vivo* using very high mass accuracy mass spectrometry at both the MS and the MS/MS level and complementary fragmentation techniques. Strikingly, we detected the same phosphorylation in yeast, *Drosophila* and human cells, suggesting an evolutionary conserved function for this modification. The nearly identical human SUMO-2 and SUMO-3 isoforms differ in serine 2; thus, only SUMO-3 could be phosphorylated at this position. Our finding that SUMO can be modified may point to an additional level of complexity through modifying a protein-modifier.

**Keywords:** SUMO-1 • Smt3 • phosphorylation • mass spectrometry • HCD • higher energy dissociation • evolution • conservation

### Introduction

Post-translational modifications (PTMs) are essential in virtually every cellular process and often function by modification-dependent protein interactions.<sup>1,2</sup> Ubiquitin and ubiquitin like proteins (Ubls) covalently modify target proteins by attaching their C-termini to specific substrate lysines. They are related by a common tertiary structure, called the ubiquitin superfold.<sup>3</sup> SUMO modification regulates the function of target proteins by modulating protein interactions, intracellular transport, half-life and activity.<sup>4</sup> There are three distinct human SUMO isoforms (SUMO-1, -2, and -3), whereas a single SUMO, called Smt3, is present in yeast and *Drosophila*. SUMO-2 and SUMO-3 are almost identical to each other and 50% identical to SUMO-1.<sup>5</sup> In contrast to ubiquitin and other Ubls, SUMO proteins contain an unstructured and flexible N-terminal segment that protrudes from the protein core.<sup>6</sup> No functional differences have yet been demonstrated between SUMO-2 and SUMO-3, which differ in just three amino acids, all of them localized on the N-terminal protrusion.

Sumoylation is mediated by an enzymatic mechanism analogous to ubiquitylation involving three enzymes, namely, E1, E2 and E3. SUMO-specific proteases catalyze SUMO precursor processing and deconjugation of SUMO from modified substrates.<sup>7</sup>

Members of the ubiquitin and Ubl family are not only PTMs, but also proteins that can be targeted by PTMs, raising the concept of "modification of a modification". All seven internal lysines of ubiquitin can be ubiquitylated,<sup>8</sup> of which at least three have different functional consequences. K11 of SUMO-2 and SUMO-3 is modified *in vivo* by SUMO-1 and SUMO-2/3.<sup>9,10</sup> Recently, it was shown that K11, K32 and K41 of SUMO-2 can be modified by ubiquitin.<sup>11,12</sup> This raises the question if SUMO can also be targeted by PTMs not belonging to the ubiquitin/Ubl family.

Here, we use a targeted, mass spectrometric approach to characterize the primary structure of SUMO-1. We report that the SUMO-1 N-terminal arm is phosphorylated and investigate the evolutionary conservation of this PTM.

### Materials and Methods

**Cell Culture and Protein Extraction.** HeLa-S3 suspension cells were grown in RPMI medium containing 10% FBS and 100 units/mL penicillin/streptomycin (Invitrogen), harvested by centrifugation for 5 min at 400g, washed with cold PBS and lysed under non-denaturing conditions with lysis buffer (50 mM HEPES, pH 7.5, 10% glycerol, 150 mM NaCl, 1% Triton X-100, 1 mM EDTA, 1 mM EFTA, 5 mM  $\beta$ -glycerophosphate, 1 mM

\* To whom correspondence should be addressed. Prof. Dr. Matthias Mann, Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany. Tel, +49 89-8578-2557; fax, +49 89-8578-2219; e-mail, mmann@biochem.mpg.de.

<sup>†</sup> Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry.

<sup>‡</sup> Organelle Architecture and Dynamics, Max Planck Institute for Biochemistry.

**Table 1.** Detected Phosphorylation Sites from *S. cerevisiae* and *Drosophila melanogaster* Smt3 and Human SUMO-1

protein	phosphopeptide sequence	charge	<i>m/z</i>	mass deviation [ppm]	Mascot score	intensity	acquisition method
SUMO-1	Ac-pSDQEAKPSTEDLGDK	2	871.3618	-0.06	36.93	3.30 × 10 <sup>5</sup>	Full range Top5
	Ac-SDQEAKPSTEDLGDK	2	831.3787	-1.05	16.39	1.99 × 10 <sup>5</sup>	Full range Top5
	Ac-pSDQEAKPSTEDLGDKK	2	935.4093	-0.19	35.93	1.04 × 10 <sup>6</sup>	Full range Top5
	Ac-SDQEAKPSTEDLGDKK	2	895.4262	-0.27	45.45	4.21 × 10 <sup>5</sup>	Full range Top5
	Ac-pSDQEAKPSTEDLGDKK	3	623.9420	0.05	46.94	6.37 × 10 <sup>5</sup>	Full range Top5
	Ac-SDQEAKPSTEDLGDKK	3	597.2865	-0.07	48.02	2.40 × 10 <sup>6</sup>	Full range Top5
(from His-SUMO-2)	Ac-pSDQEAKPSTEDLGDKK	2	935.4093	0.05	29.78	2.19 × 10 <sup>6</sup>	SIM CID/HCD
	Ac-pSDQEAKPSTEDLGDKKEGEYIK	3	863.7250	0.43	24.76	1.46 × 10 <sup>6</sup>	Full range Top5
	Ac-SDQEAKPSTEDLGDKKEGEYIK	3	837.0695	0.31	41.87	1.44 × 10 <sup>6</sup>	Full range Top5
Yeast Smt3	pSDSEVNQEAK	2	593.7348	-0.47	38.73	8.85 × 10 <sup>4</sup>	Full range Top5
	Ac-pSDSEVNQEAKPEVKPEVK	3	712.3350	-0.5	34.52	2.13 × 10 <sup>5</sup>	Full range Top5
	Ac-pSDSEVNQEAKPEVKPEVK	2	1067.9988	-0.52	59.79	2.14 × 10 <sup>5</sup>	Full range Top5
Drosophila Smt3	Ac-pSDEKKGGETEHINLK	2	903.9171	-0.77	28.19	7.88 × 10 <sup>5</sup>	
	Ac-pSDEKKGGETEHINLK	3	602.9472	-1.76	21.36	2.71 × 10 <sup>6</sup>	

NaVO<sub>4</sub>, 5 mM NaF, complete protease inhibitor mixture (Roche)). The cell lysate was centrifuged for 1 h at 60 000g at 4 °C and the supernatant was collected.

The YAL6B yeast strain<sup>13</sup> was grown in YPD liquid medium to log-phase (OD<sub>600</sub> 0.8), harvested by centrifugation for 5 min at 4000g, washed with cold water and frozen in liquid nitrogen. The cells were resuspended in lysis buffer (50 mM HEPES, pH 6.8, 150 mM KOAc, 2 mM MgOAc, 5 mM β-glycerophosphate, 1 mM NaVO<sub>4</sub>, 5 mM NaF, complete protease inhibitor mixture (Roche)) and lysed under non-denaturing conditions using glass beads. The yeast lysate was centrifuged at 10 000g for 10 min at 4 °C and the supernatant was collected.

**Electrophoresis, Coomassie Staining, and Protein Digestion.** Proteins were separated by one-dimensional SDS-PAGE using Novex 4–12 % precast gels and MES SDS running buffer (Invitrogen) according to the manufacturer's instructions. The gel was stained with Coomassie blue using Colloidal Blue Staining Kit (Invitrogen). The protein bands in the approximate molecular weight range 5–20 kDa (Supplementary Figure 1) were excised from the gel and subjected to in-gel digestion with trypsin (Promega), essentially as described in Shevchenko et al.<sup>14</sup> The resulting peptide mixture from HeLa cells was cleaned, desalted, concentrated and enriched using self-made StageT-ips.<sup>15</sup> Phosphopeptides from the yeast mixture were enriched by titanium dioxide (TiO<sub>2</sub>) chromatography, as described previously.<sup>16</sup>

**Mass Spectrometry.** LC-MS/MS analysis was performed by a linear ion trap-orbitrap hybrid mass spectrometer (LTQ-Orbitrap, Thermo Fisher Scientific) equipped with a nano-electrospray ion source (Proxeon Biosystems) and coupled with an Agilent 1200 nano-HPLC system (Agilent Technologies). The digested peptides were separated on an in-house packed 75 μm reversed-phase C18 column.<sup>17</sup> Survey scan acquisition was performed in the orbitrap (*R* = 60 000 ion accumulation to a target value of 1 000 000). The five most intense ions were fragmented and acquired in the LTQ using collisionally induced dissociation (CID) ("Full range Top5" method).<sup>18</sup> In the second measurement, the phosphorylated and N-terminally acetylated peptide of SUMO-1 was analyzed in the selected-ion monitoring (SIM) mode with a scanning range of *m/z* 932.4–938.4 (*R* = 60 000 ion accumulation to a target value of 100 000). The two most intense ions were isolated and fragmented in the LTQ using CID and in the C-trap by higher energy CID (HCD). The acquisition of the resulting MS/MS spectra was performed in the orbitrap analyzer (*R* = 30 000, ion accumulation to a target value of 75 000) ("SIM CID/HCD" method). Multistage activa-

tion was enabled on all CID MS/MS spectra and the lock mass option was enabled in all measurements to further increase mass accuracy, as described previously.<sup>18</sup>

**Data Processing and Analysis.** Raw data files were analyzed with the quantitative proteomics software MaxQuant (version 1.0.9.9),<sup>19</sup> which in this case was used for peak list generation, identification and filtering, and Mascot search engine (version 2.2.0; Mascot Science). MS/MS spectra were searched against the human International Protein Index (IPI) database (Version 3.37),<sup>20</sup> the *Drosophila* proteome database (FlyBase 5.4) or a yeast *Saccharomyces cerevisiae* ORF protein database. All three databases contained the forward and reversed sequences and were supplemented with the most commonly observed contaminants. For database search, enzyme specificity was set to trypsin allowing up to three missed cleavages. Carbamidomethylation (C) was set as fixed modification, while Oxidation (M), N-Acetyl (Protein) and phospho (STY) were searched as variable modifications. Initial maximum mass deviation was set to 7 ppm and the fragment ion mass tolerance to 0.5 Da.

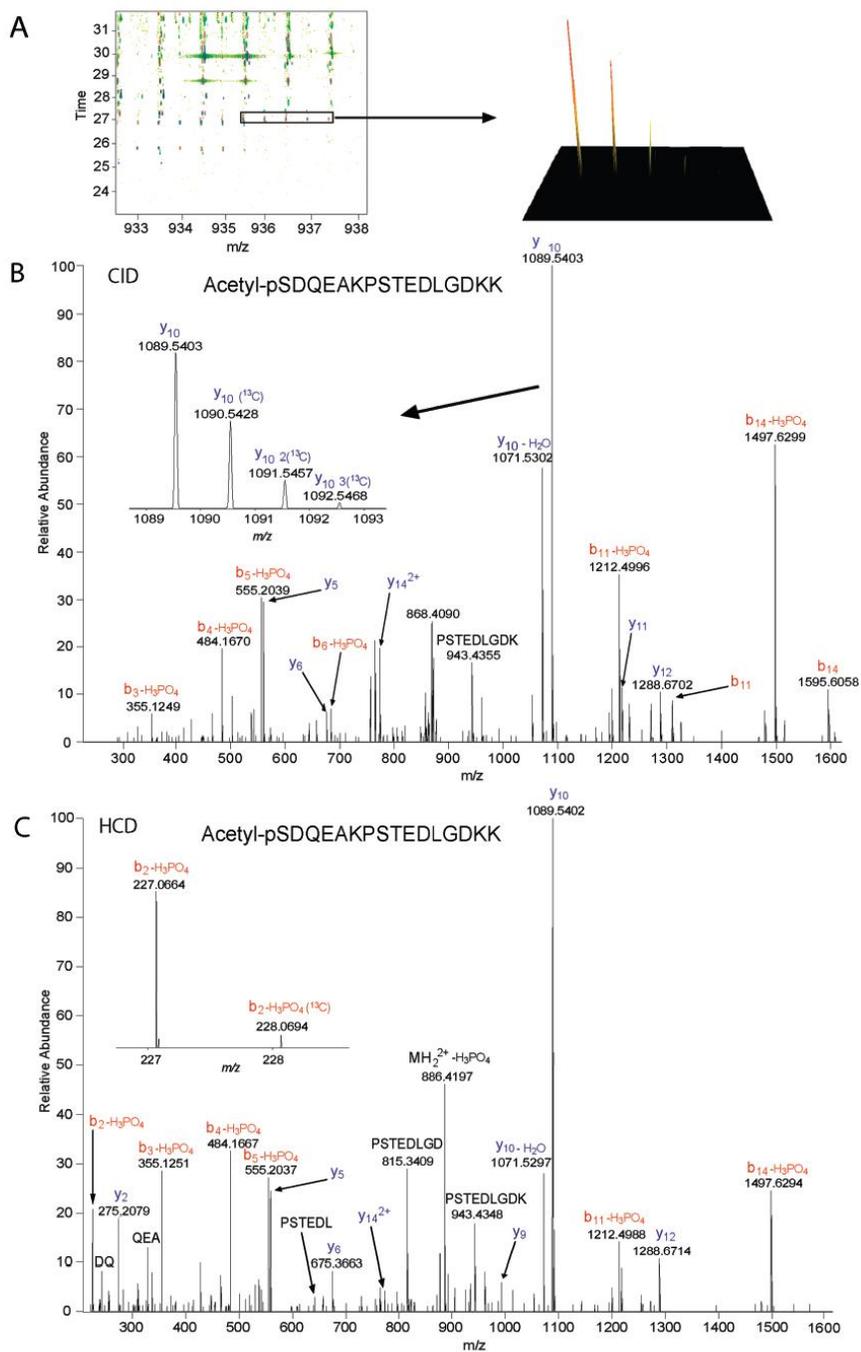
The Viewer tool of MaxQuant was used for a 2D and 3D visualization of LC-MS runs. The signal intensities are color-coded, with white and black representing the lowest intensities and green and red the highest intensities, respectively, in the 2D and 3D view.

Multiple sequence alignment was performed using the ClustalX program.<sup>21</sup>

## Results and Discussion

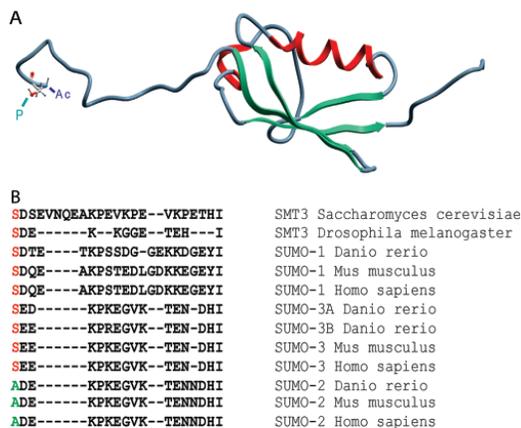
**Phosphorylation of Endogenous Human SUMO-1 *in Vivo*.** Recent mass spectrometric identification of sumoylated and ubiquitinated SUMO peptides<sup>10–12</sup> has induced us to investigate the possibility that SUMO may be targeted by an additional major regulatory PTM, phosphorylation.

SUMO is usually prepared for MS analysis from stable cell lines expressing a tagged form of the protein, such as the His<sub>6</sub>-SUMO-2 HeLa cell line.<sup>10</sup> These tagged forms have the advantage of good purification yield, but also the drawback that tagged SUMO expressed at nonendogenous levels may behave differently from its endogenous counterpart. To avoid this and still obtain sufficient material for analysis, we decided to lyse cells under non-denaturing conditions. Under these conditions, cellular SUMO is rapidly deconjugated from target proteins due to the strong activity of the desumoylation machinery.<sup>22</sup> Therefore, previously conjugated SUMO is enriched in the region of the gel whose molecular weight corresponds to that



**Figure 1.** Serine 2 of endogenous human SUMO-1 is phosphorylated *in vivo*. HeLa total cell lysate was size-fractionated by SDS-PAGE. (A) Enlarged 2D view of the MS data from the *in gel*-digested band is shown and the precursor ion isotopic distribution is represented three-dimensionally. Precursor ion was measured in the orbitrap mass spectrometer ( $m/z$  935,4093 (2+); mass deviation, 0.05 ppm), and the peptide was fragmented by CID in the LTQ (B) and by HCD in the C-trap (C). In both cases, the fragment ions were analyzed in the orbitrap analyzer. Insets, magnifications of the most intense fragment ion (B) and of the low mass region (C).

## Phosphorylation of SUMO-1



**Figure 2.** Position of the N-acetylated and phosphorylated residue and its evolutionary conservation. (A) SUMO-1 structure (1A5R)<sup>6</sup> was visualized using ICM Browser Pro (Molsoft, Redmond, WA; version 3.5-1m). (B) Sequence alignment of the N-terminal arm of SUMO family.

of free SUMO. We lysed HeLa cells, and after separation by SDS-PAGE, excised and in-gel digested the band in the approximate range 5–20 kDa (Supplementary Figure 1). Dephosphorylation and protein degradation were prevented by adding phosphatase and protease inhibitors. The resulting peptides were analyzed by LC-MS/MS on a hybrid linear ion trap-orbitrap mass spectrometer (LTQ-Orbitrap). We employed the standard method used in our laboratory, which consists of a high resolution scan in the orbitrap concurrent with up to five low resolution MS/MS scans in the linear ion trap ("Full range Top5" method, see Materials and Methods).

We found that the N-terminal peptide always lacked the initial methionine and that it was always N-acetylated in HeLa cells. This has also very recently been observed by others.<sup>23</sup> Unexpectedly, however, this peptide occurred in two forms; nonphosphorylated or phosphorylated at the N-terminal serine (Ser2). The charge state of the nonphosphopeptide was predominantly triply charged, whereas the phosphopeptide was predominantly doubly charged, presumably because of the negatively charged phosphogroup (Table 1). To obtain a rough measure of the stoichiometry of phosphorylation, we integrated the ion current of the two peptide forms. Often, phosphopeptides are assumed to ionize less well than nonphosphopeptides, although this is not necessarily true.<sup>24</sup> Even assuming equal ionization efficiency, our data is consistent with 35–40 % of Ser2 being phosphorylated. Thus, our data indicates that a large proportion of SUMO is phosphorylated *in vivo*.

Because SUMO is the object of intense biological interest and because no phosphorylation site has yet been described, we next wanted to confirm the identity of the phosphopeptide beyond doubt. The identified phosphopeptides of SUMO-1 (Table 1, Supplementary Figures 2–4) had very low mass deviations from the corresponding calculated values (low ppm range or below). However, their MS/MS spectra had low resolution and low mass accuracy (for a direct comparison of resolutions between orbitrap and LTQ see Matic et al.<sup>10</sup>). Unambiguous identification of this phosphopeptide was achieved using the capability of the LTQ-Orbitrap mass spectrometer to obtain very high mass accuracy at both MS and

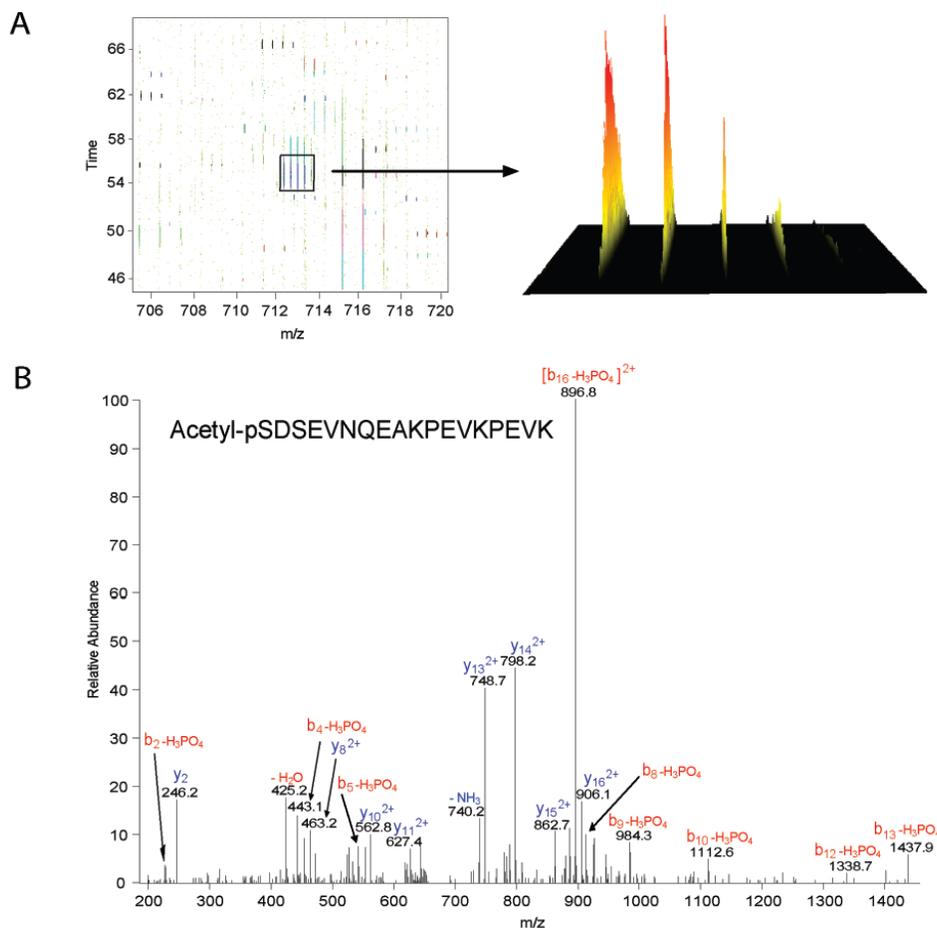
MS/MS level.<sup>18</sup> On the basis of the information from the standard, "Full range Top5" method, we devised a targeted peptide-specific method, similarly to the approach used previously for the identification of the SUMO-1/SUMO-2 branched peptide.<sup>10</sup> A narrow mass range of 6 Th around the  $m/z$  value of the precursor ion of interest was monitored in the survey scan ("SIM CID/HCD" method). The two most abundant peaks were isolated and fragmented both in the LTQ by CID and in the C-trap by higher energy dissociation (HCD) (Figure 1). In the HCD method, the C-trap, which is normally used for storage of ions prior to injection in the orbitrap, is used to produce highly efficient quadrupole-like fragmentation.<sup>25</sup> The fragment ions were acquired in the orbitrap producing high-resolution and high-accuracy MS/MS spectra. Note the similarity of MS/MS profiles in Figure 1, panels B and C. These high resolution spectra verified the identity of the b and y ions, crucial for localizing the modification, where b-ions were observed with loss of phosphoric acid ( $H_3PO_4$ ). In particular, the  $b_3$  to  $b_6$  ion series in Figure 1A and the  $b_2$  ion in Figure 1C, all with loss of phosphoric acid, as well as the nonmodified  $y_{10}$  and  $y_{12}$  ions localize the phosphorylation to the very N-terminus of the peptide. The fragmentation pathway obtained by HCD fragmentation in the C-trap also leads to a pronounced internal fragmentation (Figure 1C), giving rise to the fragment ions in the low mass part of the spectrum that were absent in CID fragmentation (Figure 1B). These internal fragments were not phosphorylated. Low resolution fragmentation spectra of the peptide with and without missed cleavage further confirmed the assignments (Table 1 and Supplemental Figures 2–4). In addition, we also found Ser2 phosphorylation of SUMO-1 in our previous global study of phosphorylation dynamics during EGF signaling in HeLa cells.<sup>16</sup> Thus, by multiple lines of mass spectrometric evidence, we unambiguously confirmed that the endogenous human SUMO-1 is phosphorylated in HeLa cells.

**Phosphorylation of Endogenous Human SUMO-1 in His<sub>6</sub>-SUMO-2 Conjugates.** We next asked if SUMO-1 could be phosphorylated when conjugated to target proteins. For this purpose, we interrogated the raw spectra from the His<sub>6</sub>-SUMO-2 fraction, in which we had previously detected mixed SUMO-branched peptides.<sup>10</sup> That fraction did not contain any free SUMO-1 as determined by Western blotting. We reanalyzed the data with phosphorylation as a variable modification and detected the phosphorylation on SUMO-1 (Supplementary Figure 8). This shows that SUMO-1 conjugated to target proteins (likely SUMO-2 in this case) can be phosphorylated.

Identically to the in-gel digested fraction, Ser2 was always N-terminally acetylated. The nonphosphorylated and phosphorylated forms of the peptide have approximately equal integrated ion current (Table 1), which confirms that a large fraction of endogenous SUMO-1 is phosphorylated.

**Evolutionary Conservation of SUMO Phosphorylation.** The identified peptides indicate that the N-terminal tail of SUMO-1 undergoes three co- or post-translational modifications: removal of the starting methionine residue, N-terminal acetylation and phosphorylation of the first amino acid (Ser2) of the mature SUMO-1. Since Ser2 is in a highly flexible and accessible part of SUMO-1, this phosphoserine is a good candidate for the regulation of protein binding (Figure 2A).

To investigate evolutionary conservation of the Ser2, we aligned the amino acid sequences of SUMOs of five different eukaryotic species (Figure 2B). The alignment shows that Ser2 of human SUMO-1 is conserved in the single SUMO forms present in *S. cerevisiae* and *D. melanogaster* (both



**Figure 3.** Serine 2 of endogenous yeast Smt3 is phosphorylated *in vivo*. Yeast total cell lysate was separated by SDS-PAGE, the band was digested in gel, and the phosphopeptides were enriched by  $TiO_2$  chromatography. The precursor ion, shown in a 2D and 3D representation (A), was analyzed in the orbitrap mass spectrometer ( $m/z$  712.3350 (3+); mass deviation,  $-0.58$  ppm), and subsequently fragmented by CID in the LTQ. The resulting fragmentation spectra were acquired in the LTQ analyzer (B).

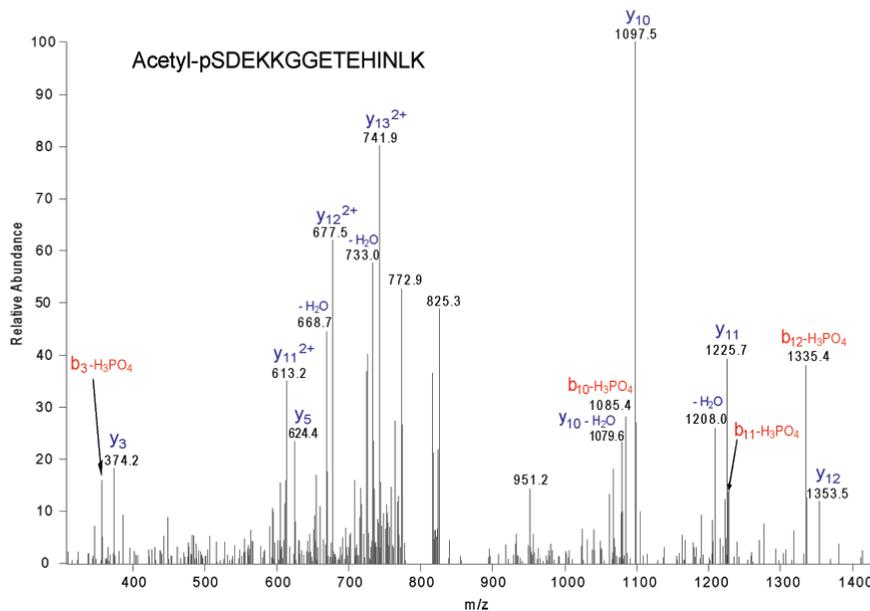
called Smt3). In vertebrates, which have three SUMO paralogs, position 2 is serine in SUMO-1 and SUMO-3, whereas it is alanine in SUMO-2.

It is interesting to note that the amino acid at position 2 is one of only three amino acids that are different between human SUMO-2 and SUMO-3. SUMO-2/3 are so far functionally indistinguishable, but their sequences are quite different from SUMO-1. Thus, SUMO-1 and SUMO-3, but not SUMO-2, can in principle be phosphorylated at their N-terminus.

Despite different strategies, including purification of nuclei, in-solution and in-gel digestion followed by  $TiO_2$  enrichment and the targeted mass spectrometric approach, in this study, we were not able to detect the SUMO-3 phosphoserine-containing peptide. Digestion with trypsin could produce a peptide which is too small to be detected by MS (Figure 3B). However, the phosphopeptide was not identified even after digestion with endoprotease Asp-N, which should produce a longer and more easily detectable peptide. Our failure may be

due to the low stoichiometry of the modification and technical limitations of our approach and does not prove that SUMO-3 cannot be phosphorylated.

**Phosphorylation of Endogenous *S. cerevisiae* and *D. melanogaster* Smt3 *in Vivo*.** As mentioned above, Ser2 is evolutionary conserved in eukaryotic species. To test whether N-terminal acetylation and phosphorylation of Ser2 are also conserved, we enriched yeast Smt3 as described for mammalian cells above and digested the proteins in gel with trypsin. Phosphopeptides were enriched by  $TiO_2$  chromatography and analyzed with the LTQ-Orbitrap using the standard acquisition method. The identified peptide of interest lacks the first methionine and Ser2 is N-acetylated and phosphorylated (Figure 3, Table 1 and Supplementary Figures 5 and 6), just as in human SUMO-1. The presence of the  $b_2$  ion with loss of phosphoric acid and of the unmodified double charged  $y_{16}$  ion (Figure 3B) excludes the possibility that the phosphorylation is localized on the nonconserved Ser4. As in the human case, the N-terminal



**Figure 4.** Serine 2 of endogenous *Drosophila* Smt3 is phosphorylated *in vivo*. Proteins were separated by SDS-PAGE and digested in gel. Phosphopeptides were enriched using TiO<sub>2</sub> beads. The precursor ion was analyzed in the orbitrap mass spectrometer ( $m/z$  903.9171 (2+); mass deviation,  $-0.77$  ppm), and subsequently fragmented and acquired in the LTQ.

peptide of yeast Smt3 was always lacking the initial methionine, but contrary to what we observed in HeLa cells, we also detected a peptide in which Ser2 is phosphorylated, but not N-acetylated (Table 1, Supplementary Figure 5).

We next interrogated data from an ongoing *Drosophila* phosphoproteome project (Hilger et al, manuscript in preparation). Proteins extracted from Schneider SL2 cells were in-gel digested and phosphopeptides were enriched by TiO<sub>2</sub> chromatography. As shown in Figure 4, Table 1 and Supplementary Figure 7, the N-terminal peptide is acetylated and phosphorylated in this species as well.

## Conclusions

The ubiquitin and ubiquitin like proteins are unique in that they are both post-translational modifiers but can also themselves be modified. Thus, they have a “modifying” mode, but also a “modified” mode. This latter property has only been explored in depth for ubiquitin and SUMO chain formation.<sup>3,9,10</sup> Here, we have shown that a member of the Ubl family, SUMO-1, is the target of phosphorylation. We used state-of-the-art high-resolution mass spectrometry at the MS and MS/MS levels to establish the presence of this modification beyond doubt. Two factors argue that this modification may be of functional significance rather than ‘background phosphorylation’. First, the stoichiometry of the modification is high—the signal of the phosphopeptide was almost as large as that of the unphosphorylated peptide in HeLa cells. It appears unlikely that a nonfunctional site, especially a highly conserved site potentially important for protein–protein interaction, would be phosphorylated at such high levels. Secondly, we found this modification over an evolutionary distance of at least one billion years

and in organisms as different as a unicellular eukaryote, an insect and human.

Interestingly, a SUMO polymerization site resides within the amino-terminal arm.<sup>10</sup> Here, we have established that two PTMs not belonging to the ubl/ubiquitin family, namely, phosphorylation and N-terminal acetylation, contribute to the SUMO “modified” mode. The unique feature of the SUMO proteins compared to other Ubls is the long and very flexible N-terminal protrusion. It has been suggested that this unstructured tail of the protein might provide an additional surface for protein–protein interaction.<sup>6</sup> In this perspective, phosphorylation of this extreme N-terminal serine could modulate the ability of proteins to interact with SUMO.

Interestingly, similar amino-terminal tail domains are present in histones and are essential for chromatin dynamics. Modifications of histone tails regulate gene expression by mediating the recruitment of proteins to chromatin.<sup>26</sup> The N-terminal arms of both SUMOs and histones protrude from the protein core, contain many charged amino acids and different PTMs sites. Analogously to SUMO, the first serine of histones H2A and H4 are N-acetylated and phosphorylated.<sup>27</sup> Thus, one can speculate that the N-terminal extensions of SUMO may recruit interaction partners to sumoylated proteins or SUMO itself in a similar way that histones amino terminal tails recruit proteins to chromatin.

SUMO-2 and SUMO-3 have yet to be functionally differentiated. All three residues different between these two paralogs are localized on the N-terminal protrusion and serine is conserved from SUMO-1 to SUMO-3, but not SUMO-2. The slight difference in the amino acid sequences between SUMO-2 and SUMO-3 tails and the hypothetical SUMO-3 specific phosphorylation could lead to a different recruitment of

interacting proteins. If regulated Ser2 phosphorylation on SUMO-3 can be detected, this would be the first functional difference between SUMO-2 and SUMO-3. Thus, the functional meaning of the phosphorylation on SUMO Ser2 and the investigation of interaction partners merits further studies.

**Abbreviations:** PTM, post-translational modification; Ubls, ubiquitin-like proteins; SUMO, small ubiquitin-like modifier; E1, SUMO-activating enzyme, E2, SUMO-protein carrier protein; E3, SUMO ligase; LC-MS, liquid chromatography-mass spectrometry; MS/MS, tandem mass spectrometry; SIM, selected ion monitoring; CID, collision-induced dissociation; HCD, higher-energy C-trap dissociation; LTQ, linear quadrupole ion trap; ppm, parts per million.

**Acknowledgment.** We thank Florian Fröhlich and Michael Rehman for advice, helpful discussion and help with multiple sequence analysis. We thank other members of our department for fruitful discussions. This work was supported by The European Community (RUBICON, VI Framework).

**Supporting Information Available:** Supplementary Figure 1, gel region around free SUMO-1 or Smt3 (approximately 5–20 kDa) was digested in gel and analyzed by LC-MS/MS; Supplementary Figure 2, MS/MS fragmentation spectrum of the human SUMO-1 peptide Acetyl-pSDQEAKPSTEDLGDK. Precursor ion mass was measured in the orbitrap analyzer ( $m/z$  871.3618 (2+); mass deviation,  $-0.06$  ppm) and the peptide was fragmented by CID and acquired in the LTQ mass spectrometer; Supplementary Figure 3, MS/MS fragmentation spectrum of the human SUMO-1 peptide Acetyl-pSDQEAKPSTEDLGDK. Precursor ion mass was measured in the orbitrap analyzer ( $m/z$  935.4093 (2+); mass deviation,  $-0.19$  ppm) and the peptide was fragmented by CID and acquired in the LTQ mass spectrometer; Supplementary Figure 4, MS/MS fragmentation spectrum of the human SUMO-1 peptide Acetyl-pSDQEAKPSTEDLGDK. Precursor ion mass was measured in the orbitrap analyzer ( $m/z$  623.9420 (3+); mass deviation,  $0.05$  ppm) and the peptide was fragmented by CID and acquired in the LTQ mass spectrometer; Supplementary Figure 5, MS/MS fragmentation spectrum of the yeast Smt3 peptide pSDSEVNQEAK. Precursor ion mass was measured in the orbitrap analyzer ( $m/z$  593.7348 (2+); mass deviation,  $-0.47$  ppm) and the peptide was fragmented by CID and acquired in the LTQ mass spectrometer; Supplementary Figure 6, MS/MS fragmentation spectrum of the yeast Smt3 peptide Acetyl-pSDSEVNQEAKPEVKPEVK. Precursor ion mass was measured in the orbitrap analyzer ( $m/z$  1067.9988 (2+); mass deviation,  $-0.52$  ppm) and the peptide was fragmented by CID and acquired in the LTQ mass spectrometer; Supplementary Figure 7, MS/MS fragmentation spectrum of the *Drosophila* Smt3 peptide Acetyl-pSDEKKGGETEHLNK. Precursor ion mass was measured in the orbitrap analyzer ( $m/z$  602.9472 (3+); mass deviation,  $-1.76$  ppm) and the MS/MS spectra were acquired in the LTQ mass spectrometer; Supplementary Figure 8, His<sub>6</sub>-SUMO-2 MS/MS data from Matic et al.<sup>10</sup> were analyzed with MaxQuant. MS/MS fragmentation spectrum of the human SUMO-1 peptide Acetyl-pSDQEAKPSTEDLGDKKEGEYIK. Precursor ion mass was measured in the orbitrap analyzer ( $m/z$  863.7250 (2+); mass deviation,  $0.44$  ppm) and the peptide was fragmented by CID and acquired in the LTQ mass spectrometer. Note that the multistage activation was not enabled and that the main

fragmentation ion is the neutral loss of the precursor ion. The enlarged left and right regions of the MS/MS spectrum show the presence of other fragment ions. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (1) Krishna, R. G.; Wold, F. Post-translational modification of proteins. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1993**, *67*, 265–98.
- (2) Seet, B. T.; Dikic, I.; Zhou, M. M.; Pawson, T. Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* **2006**, *7* (7), 473–83.
- (3) Welchman, R. L.; Gordon, C.; Mayer, R. J. Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat. Rev. Mol. Cell Biol.* **2005**, *6* (8), 599–609.
- (4) Hay, R. T. SUMO: a history of modification. *Mol. Cell* **2005**, *18* (1), 1–12.
- (5) Johnson, E. S. Protein modification by SUMO. *Annu. Rev. Biochem.* **2004**, *73*, 355–82.
- (6) Bayer, P.; Arndt, A.; Metzger, S.; Mahajan, R.; Melchior, F.; Jaenicke, R.; Becker, J. Structure determination of the small ubiquitin-related modifier SUMO-1. *J. Mol. Biol.* **1998**, *280* (2), 275–86.
- (7) Bossis, G.; Melchior, F. SUMO: regulating the regulator. *Cell Div.* **2006**, *1*, 13.
- (8) Pickart, C. M.; Fushman, D. Polyubiquitin chains: polymeric protein signals. *Curr. Opin. Chem. Biol.* **2004**, *8* (6), 610–6.
- (9) Tatham, M. H.; Jaffray, E.; Vaughan, O. A.; Desterro, J. M.; Botting, C. H.; Naismith, J. H.; Hay, R. T. Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *J. Biol. Chem.* **2001**, *276* (38), 35368–74.
- (10) Matic, I.; van Hagen, M.; Schimmel, J.; Macek, B.; Ogg, S. C.; Tatham, M. H.; Hay, R. T.; Lamond, A. I.; Mann, M.; Vertegaal, A. C. 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* **2008**, *7* (1), 132–44.
- (11) Tatham, M. H.; Geoffroy, M.-C.; Shen, L.; Plechanovova, A.; Hattersley, N.; Jaffray, E. G.; Palvimo, J. J.; Hay, R. T. RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat. Cell Biol.* **2008**, *10* (5), 538–46.
- (12) Schimmel, J.; Larsen, K. M.; Matic, I.; van Hagen, M.; Cox, J.; Mann, M.; Andersen, J. S.; Vertegaal, A. C. The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle. *Mol. Cell. Proteomics* **2008**.
- (13) Gruhler, A.; Olsen, J. V.; Mohammed, S.; Mortensen, P.; Faergeman, N. J.; Mann, M.; Jensen, O. N. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics* **2005**, *4* (3), 310–27.
- (14) Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **2006**, *1* (6), 2856–60.
- (15) Rappsilber, J.; Mann, M.; Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2007**, *2* (8), 1896–906.
- (16) Olsen, J. V.; Blagoev, B.; Gnäd, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **2006**, *127* (3), 635–48.
- (17) Olsen, J. V.; Ong, S. E.; Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **2004**, *3* (6), 608–14.
- (18) Olsen, J. V.; de Godoy, L. M.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **2005**, *4* (12), 2010–21.
- (19) Cox, J.; Mann, M. Is proteomics the new genomics. *Cell* **2007**, *130* (3), 395–8.
- (20) Kersey, P. J.; Duarte, J.; Williams, A.; Karavidopoulou, Y.; Birney, E.; Apweiler, R. The International Protein Index: an integrated database for proteomics experiments. *Proteomics* **2004**, *4* (7), 1985–8.
- (21) Thompson, J. D.; Gibson, T. J.; Plewniak, F.; Jeanmougin, F.; Higgins, D. G. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **1997**, *25* (24), 4876–82.
- (22) Mahajan, R.; Delphin, C.; Guan, T.; Gerace, L.; Melchior, F. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **1997**, *88* (1), 97–107.
- (23) Lallemand-Breitenbach, V.; Jeanne, M.; Benhenda, S.; Nasr, R.; Lei, M.; Peres, L.; Zhou, J.; Zhu, J.; Raught, B.; de The, H. Arsenic

### Phosphorylation of SUMO-1

- degrades PML or PML-RAR[alpha] through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat. Cell Biol.* 2008, 10 (5), 547–55.
- (24) Steen, H.; Jebaranthirajah, J. A.; Rush, J.; Morrice, N.; Kirschner, M. W. Phosphorylation analysis by mass spectrometry: myths, facts, and the consequences for qualitative and quantitative measurements. *Mol. Cell. Proteomics* 2006, 5 (1), 172–81.
- (25) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* 2007, 4 (9), 709–12.
- (26) Iizuka, M.; Smith, M. M. Functional consequences of histone modifications. *Curr. Opin. Genet. Dev.* 2003, 13 (2), 154–60.
- (27) Barber, C. M.; Turner, F. B.; Wang, Y.; Hagstrom, K.; Taverna, S. D.; Mollah, S.; Ueberheide, B.; Meyer, B. J.; Hunt, D. F.; Cheung, P.; Allis, C. D. The enhancement of histone H4 and H2A serine 1 phosphorylation during mitosis and S-phase is evolutionarily conserved. *Chromosoma* 2004, 112 (7), 360–71.

PR800368M

### research articles

## **5 Cross-talk between SUMO-2/3 and the ubiquitin-proteasome system**

**PAPER:** The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle.

The following article was published in the November issue of Molecular & Cellular Proteomics, pages 2107-2122.

This study identifies proteins that are targeted by SUMO-2 and consequently ubiquitinated for proteasomal degradation.



## The Ubiquitin-Proteasome System Is a Key Component of the SUMO-2/3 Cycle\*<sup>§</sup>

Joost Schimmel<sup>‡§</sup>, Katja M. Larsen<sup>§¶</sup>, Ivan Matic<sup>§||</sup>, Martijn van Hagen<sup>‡</sup>, Jürgen Cox<sup>||</sup>, Matthias Mann<sup>||</sup>, Jens S. Andersen<sup>¶\*\*</sup>, and Alfred C. O. Vertegaal<sup>‡ ††</sup>

Many proteins are regulated by a variety of post-translational modifications, and orchestration of these modifications is frequently required for full control of activity. Currently little is known about the combinatorial activity of different post-translational modifications. Here we show that extensive cross-talk exists between sumoylation and ubiquitination. We found that a subset of SUMO-2-conjugated proteins is subsequently ubiquitinated and degraded by the proteasome. In a screen for preferential SUMO-1 or SUMO-2 target proteins, we found that ubiquitin accumulated in purified SUMO-2 conjugates but not in SUMO-1 conjugates. Upon inhibition of the proteasome, the amount of ubiquitin in purified SUMO-2 conjugates increased. In addition, we found that endogenous SUMO-2/3 conjugates, but not endogenous SUMO-1 conjugates, accumulated in response to proteasome inhibitors. Quantitative proteomics experiments enabled the identification of 73 SUMO-2-conjugated proteins that accumulated in cells treated with proteasome inhibitors. Cross-talk between SUMO-2/3 and the ubiquitin-proteasome system controls many target proteins that regulate all aspects of nucleic acid metabolism. Surprisingly the relative abundance of 40 SUMO-2-conjugated proteins was reduced by proteasome inhibitors possibly because of a lack of recycled SUMO-2. We conclude that SUMO-2/3 conjugation and the ubiquitin-proteasome system are tightly integrated and act in a cooperative manner. *Molecular & Cellular Proteomics* 7:2107–2122, 2008.

The ubiquitin-proteasome system plays a key role in virtually all cellular processes by tightly regulating the degradation of a large set of proteins (1). Proteins are targeted for degradation by lysine 48-linked polyubiquitin chains that are covalently conjugated to lysines in target proteins. Ubiquitination furthermore regulates target proteins in a degradation-

independent manner, e.g. monoubiquitination is important for endocytosis (2). A significant part of the human genome encodes components of the ubiquitin-proteasome system, including E1,<sup>1</sup> E2, and hundreds of E3 enzymes that mediate the conjugation of target proteins to ubiquitin and ubiquitin proteases that remove ubiquitins from target proteins.

The ubiquitin family comprises ubiquitin-like proteins NEDD8, ISG15, SUMO-1, -2, -3, FAT10, FUB1, UBL5, URM1, ATG8, and ATG12 (3, 4). These proteins share the three-dimensional structure of ubiquitin and are also conjugated to target proteins. Like ubiquitination, sumoylation is essential for eukaryotic life (5). The largest functional group of SUMO targets are transcription factors (6), and in general, sumoylation inhibits their transcriptional activity (7). Sumoylation also regulates other cellular processes including DNA repair, RNA metabolism, protein transport, translation, and replication (8–10). Whereas mature SUMO-2 and SUMO-3 are nearly identical (~95% identity), they differ significantly from SUMO-1 (~50% identity). Previously we have shown that SUMO-1 and SUMO-2 are conjugated to preferential sets of target proteins (6). Furthermore SUMO-2 and SUMO-3 contain an internal sumoylation site that is used for SUMO chain formation *in vivo* (11, 12).

Sumoylation is not linked to the degradation of target proteins, although some exceptions have been reported. The SUMO-accepting lysine 160 in PML and in the oncogenic PML-retinoic acid receptor  $\alpha$  protein is required for the degradation of this fusion protein upon arsenic trioxide treatment (13). Furthermore it has been published that sumoylation might be important for the degradation of DNA topoisomerase II $\beta$  in response to a catalytic inhibitor (14).

<sup>1</sup> The abbreviations used are: E1, SUMO-activating enzyme; E2, SUMO protein carrier protein; E3, SUMO ligase; aa, amino acids; Arg0, [<sup>12</sup>C<sub>6</sub>,<sup>14</sup>N<sub>4</sub>]arginine; Arg6, [<sup>13</sup>C<sub>6</sub>,<sup>14</sup>N<sub>4</sub>]arginine; Arg10, [<sup>13</sup>C<sub>6</sub>,<sup>15</sup>N<sub>4</sub>]arginine; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; hnRNP M, heterogeneous nuclear ribonucleoprotein M; LTQ, linear quadrupole ion trap; Lys0, [<sup>12</sup>C<sub>6</sub>,<sup>14</sup>N<sub>2</sub>]lysine; Lys4, [<sup>2</sup>H<sub>4</sub>,<sup>12</sup>C<sub>6</sub>,<sup>14</sup>N<sub>2</sub>]lysine; Lys8, [<sup>13</sup>C<sub>6</sub>,<sup>15</sup>N<sub>2</sub>]lysine; MCM-7, minichromosome maintenance protein 7; PIAS, protein inhibitor of activated signal transducer and activator of transcription; PML, promyelocytic leukemia protein; RanGAP1, Ran GTPase-activating protein 1; RNF4, RING finger protein 4; SAFB, scaffold attachment factor B; SART, squamous cell carcinoma antigen recognized by T-cells; SILAC, stable isotope labeling by amino acids in cell culture; SUMO, small ubiquitin-like modifier; HRP, horseradish peroxidase.

From the <sup>‡</sup>Department of Molecular Cell Biology, Leiden University Medical Center, 2300 RC Leiden, the Netherlands, <sup>¶</sup>Center for Experimental Bioinformatics (CEBI), Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230, Odense M, Denmark, and <sup>||</sup>Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany

Received, January 17, 2008, and in revised form, June 18, 2008  
Published, MCP Papers in Press, June 18, 2008, DOI 10.1074/mcp.M800025-MCP200

## Cross-talk between SUMO-2/3 and the Proteasome

Different kinds of cross-talk between ubiquitination and sumoylation have been reported recently (15). SUMO and ubiquitin were shown to counteract each other by competing for the same acceptor lysine in I $\kappa$ B $\alpha$  (16). NF- $\kappa$ B signaling is furthermore affected by the ubiquitination and sumoylation of NF- $\kappa$ B essential modulator (NEMO)/I $\kappa$ B kinase  $\gamma$ , a structural component of the I $\kappa$ B kinase complex (17). In this case, SUMO-1 and ubiquitin do not directly compete for the same acceptor lysine but are conjugated in a sequential manner in response to genotoxic stress. Proliferating cell nuclear antigen is also modified by SUMO and ubiquitin on the same acceptor lysine (Lys-164) (15, 18). Sumoylation enables the interaction between proliferating cell nuclear antigen and the helicase Srs2, whereas monoubiquitination enables translesion synthesis by Pol $\eta$ , a DNA damage-tolerant polymerase, and polyubiquitination is needed for DNA repair.

In a screen for preferential SUMO-1 and preferential SUMO-2 conjugates, we found that ubiquitin specifically co-enriched with SUMO-2. The amount of ubiquitin in SUMO-2-purified fractions significantly increased upon inhibition of the proteasome. Quantitative proteomics enabled us to study the cross-talk between sumoylation and the ubiquitin-proteasome system at the target protein level. We show here that the conjugation of a large set of target proteins to SUMO-2 is tightly connected to the ubiquitin-proteasome system and conclude that the ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle in cells.

### EXPERIMENTAL PROCEDURES

**Cell Culture and Transfections**—HeLa cells stably expressing His $_6$ -SUMO-1 or His $_6$ -SUMO-2 were described previously (6). HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% FCS and 100 units/ml penicillin and streptomycin (Invitrogen). Stable isotope labeling was carried out essentially as described previously (6, 19) using [ $^{12}\text{C}_6$ ,  $^{14}\text{N}_4$ ]arginine (referred to as Arg0), [ $^{13}\text{C}_6$ ,  $^{14}\text{N}_4$ ]arginine (referred to as Arg6), [ $^{13}\text{C}_6$ ,  $^{15}\text{N}_4$ ]arginine (referred to as Arg10), [ $^{12}\text{C}_6$ ,  $^{14}\text{N}_2$ ]lysine (referred to as Lys0), [ $^{2}\text{H}_4$ ,  $^{12}\text{C}_6$ ,  $^{14}\text{N}_2$ ]lysine (referred to as Lys4), or [ $^{13}\text{C}_6$ ,  $^{15}\text{N}_2$ ]lysine (referred to as Lys8) as indicated. Transfections were carried out using 25-kDa linear polyethyleneimine (Brunschwig-Chemie) essentially as described before (20).

**Plasmids, Mutagenesis, Antibodies, Protein Electrophoresis, and Immunoblotting**—The plasmids encoding His $_6$ -ubiquitin wild type or 7KR and the plasmid encoding wild-type His $_6$ -SUMO-2 were described previously (21, 22). The His $_6$ -SUMO-2 K11R, E13A, K32R, K32R,K34R,K41R,K44R, and allKR plasmids were generated by site-directed mutagenesis using the QuikChange II kit according to the instructions of the manufacturer (Stratagene). Mutants were confirmed by DNA sequencing.

The amino acid sequence of the mature protein that we refer to as SUMO-2 is MSEEKPKGKVTENDHINLKVAGDQGSVVQFKIKRHTPLSKLKMAYCERQGLSMRQIRFRFDGQPINETDTPAQLEMEDED-IDVFQQQTGG (12). Peptide antibody AV-SM23-0100 against SUMO-2/3 was generated in rabbit using the peptide MEDEDIDVFQQQTG (Eurogentec) (6, 22). Peptide antibody 1607 against SART1 was also generated in rabbit by Eurogentec using peptides CSLSIEET-NKLRAKLGKPLEV and CNLDEEKQQQDFASSTT as described previously (6). Monoclonal antibodies 21C7 against SUMO-1 and 19C7 against RanGAP1 were obtained from Zymed Laboratories Inc..

Monoclonal antibodies HIS-1 against polyhistidine, R-3902 against hnRNP M, and M-7931 against MCM-7 were obtained from Sigma. Monoclonal antibody SC-8017 against ubiquitin and polyclonal antibody SC-8152 against PIAS1 were obtained from Santa Cruz Biotechnology. Monoclonal antibody ab8060 against SAFB was obtained from Abcam. This antibody also recognizes SAFB2. Secondary antibodies used were anti-rabbit HRP and anti-mouse HRP (1:5000; Pierce) and anti-goat HRP (1:5000; Sigma).

Protein samples were size-fractionated on Novex 4–12% Bis-Tris gradient gels using MOPS buffer (Invitrogen). For immunoblotting experiments, size-fractionated proteins were subsequently transferred onto Hybond-C Extra membranes (Amersham Biosciences) using a submarine system (Invitrogen). The membranes were incubated with specific antibodies as indicated. Bound antibodies were detected via chemiluminescence with ECL Plus (Amersham Biosciences).

**Purification of His $_6$ -SUMO- and His $_6$ -Ubiquitin-conjugated Proteins**—His $_6$ -SUMO conjugates and His $_6$ -ubiquitin conjugates were purified essentially as described previously (23). Cells were scraped in ice-cold PBS. Two small aliquots of each sample were lysed in lithium dodecyl sulfate protein sample buffer (Invitrogen) as input control or in 8 M urea, 100 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 10 mM Tris/HCl, pH 7.0 to determine the protein concentration. The remaining cells were solubilized in lysis buffer (6 M guanidinium-HCl, 100 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 10 mM Tris/HCl, pH 8.0, 20 mM imidazole, 10 mM  $\beta$  mercaptoethanol) and sonicated to reduce the viscosity. His $_6$ -SUMO conjugates or His $_6$ -ubiquitin conjugates were enriched on nickel-nitrilotriacetic acid-agarose beads (Qiagen) and washed using wash Buffers A–D (Buffer A: 6 M guanidinium-HCl, 100 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 10 mM Tris/HCl, pH 8.0, 0.2% Triton-X-100; Buffer B: 8 M urea, 100 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 10 mM Tris/HCl, pH 8.0, 0.2% Triton-X-100; Buffer C: 8 M urea, 100 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 10 mM Tris/HCl, pH 6.3, 0.2% Triton-X-100; Buffer D: 8 M urea, 100 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 10 mM Tris/HCl, pH 6.3, 0.1% Triton-X-100). These wash buffers also contained 10 mM  $\beta$  mercaptoethanol. Samples were eluted in 6.4 M urea, 80 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 8 mM Tris/HCl, pH 7.0, 200 mM imidazole.

For the experiment described in Fig. 1, A–C, a previously described method was used (6). For the experiments described in Figs. 3, A and B, and 6, A and B, His $_6$ -SUMO conjugates were immunoprecipitated using monoclonal antibody HIS-1 (Sigma) as described previously (23). For the experiment described in Fig. 7A HeLa cells were transfected with the His $_6$ -ubiquitin 7KR encoding plasmid. Cells were lysed in 6 M guanidinium-HCl, 100 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 10 mM Tris/HCl, pH 8.0, and proteins were digested with endopeptidase Lys-C. His $_6$ -ubiquitin conjugates were subsequently purified on Talon beads (BD Biosciences), washed two times with lysis buffer and four times with 8 M urea, 100 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 10 mM Tris/HCl, pH 8.0, and eluted in 6.4 M urea, 80 mM Na $_2$ HPO $_4$ /Na $_2$ HPO $_4$ , 8 mM Tris/HCl, pH 7.0, 200 mM imidazole. His $_6$ -ubiquitin and conjugated peptides were digested with trypsin in solution and identified by mass spectrometry.

**Mass Spectrometry and Data Analysis**—Mass spectrometric analysis was performed by nanoscale LC-MS/MS using a linear ion trap (LTQ) FT-ICR mass spectrometer (Thermo-Fisher Scientific, Bremen, Germany). Eluates were analyzed by one-dimensional gel electrophoresis. Gel lanes were cut in slices and subjected to in-gel digestion with Lys-C. The resulting peptides were extracted, concentrated, and then loaded onto a fused silica capillary with a 75- $\mu$ m inner diameter and an 8- $\mu$ m tip opening (New Objective, Woburn, MA) filled with Reprosil 3- $\mu$ m reverse phase material (Dr. Maisch, Ammerbuch, Germany). Peptides were eluted with a 140-min linear gradient of 95% buffer A (0.5% acetic acid in H $_2$ O) to 50% buffer B (80% acetonitrile, 0.5% acetic acid in H $_2$ O). The LTQ-FT-ICR instrument was operated in the data-dependent mode to acquire high resolution

## Cross-talk between SUMO-2/3 and the Proteasome

precursor ion spectra (from  $m/z$  300 to 1500,  $R = 50,000$ , and ion accumulation to a target value of 5,000,000) in the ICR cell. The three most intense ions were sequentially isolated for accurate mass measurements by selected ion monitoring scans (10-Da mass window,  $R = 50,000$ , and a target accumulation value of 90,000). The ions were simultaneously fragmented in the linear ion trap with a normalized collision energy setting of 27% and a target value of 10,000.

Peak list-generating software was DTA supercharge (release date, April 20, 2006). The combined peak list was searched in the International Protein Index database (release date, April 21, 2006; total of 66,279 sequences) using the Mascot program (Matrix Science, London, UK). The enzyme specificity was set to Lys-C, allowing for cleavage N-terminal of proline and between aspartic acid and proline. Cysteine carbamidomethylation was selected as a fixed modification, and methionine oxidation, protein N-acetylation, lysine- $d_4$ , and [ $^{13}C_6$ ,  $^{15}N_3$ ]lysine were searched as variable modifications. LTQ-FT-ICR data were searched with a peptide mass tolerance of 10 ppm and a fragment mass tolerance of 0.6 Da. Iterative calibration algorithms on the basis of identified peptides resulted in an average absolute peptide mass accuracy of better than 1 ppm. A maximum of one missed cleavage was allowed. Stringent criteria were required for protein identification based on the LTQ-FT-ICR data: at least two matching peptides per protein, a mass accuracy within 3 ppm, a Mascot score for individual ions of better than 20, and a delta score of better than 5.

Protein ratios were calculated for each peptide, and peptide ratios were averaged for all quantified peptides sequenced for each protein. MSQuant, an in-house-developed software program was used to extract information from the Mascot HTML database search files and to manually validate the certainty in peptide identification and in peptide abundance ratio. The quantitation was based on relative intensities from combined scans. The program is available as open source from SourceForge, Inc.

For the experiments described in Figs. 3, A and B, and 6, A and B, mass spectrometric analysis was performed by nanoscale LC-MS/MS using an LTQ-Orbitrap mass spectrometer, (Thermo Fisher Scientific) equipped with a nano-electrospray ion source (Proxeon Biosystems, Odense, Denmark) and coupled to an Agilent 1200 nano-HPLC system (Agilent Technologies) fitted with an in-house-made 75- $\mu$ m reverse phase  $C_{18}$  column as described previously (24). In-solution digestion was performed essentially as before (25). The resulting peptides were desalted on reverse phase  $C_{18}$  stop and go extraction (STAGE) tips (26). Peptides were eluted with a 140-min linear gradient of 98% solvent A (0.5% acetic acid in  $H_2O$ ) to 50% solvent B (80% acetonitrile, 0.5% acetic acid in  $H_2O$ ).

Data were acquired in the data-dependent mode: full scan spectra ( $m/z$  300–2000,  $R = 60,000$ , and ion accumulation to a target value of 1,000,000) were acquired in the orbitrap. The 10 most intense ions were fragmented and recorded in the ion trap as described before (24). Raw files were processed with our in-house quantitative proteomics software MaxQuant (version 1.0.7.5) that performs peak list generation, SILAC-based quantitation, false discovery rate determination, peptide to protein assembly, and data filtration essentially as described previously (27, 28). The quantitation was based on relative intensities from combined scans. The data were searched against a target/decoy human International Protein Index database (version 3.24) supplemented with frequently observed contaminants (total of 66,948 forward) using Mascot (Matrix Science, version 2.1.04). The enzyme specificity was set to trypsin, allowing for cleavage N-terminal of proline and between aspartic acid and proline. Cysteine carbamidomethylation was selected as a fixed modification, and methionine oxidation and protein N-acetylation and deamidation of asparagine and glutamine were searched as variable modifications. Spectra determined to be heavy labeled in the presearch MaxQuant

detection of SILAC pairs were searched with the fixed modifications Arg10 and Lys8; for MS/MS spectra with a SILAC state not determinable before the database search Arg10 and Lys8 were taken as variable modifications. Initial maximum allowed mass deviation was set to 7 ppm for peptide masses and 0.5 ppm for MS/MS peaks. The minimum peptide length was set to 6 amino acids, and a maximum of three missed cleavages and three labeled amino acids were allowed. Two proteins were grouped together if the peptide sequence set of one protein was equal to or a subset of the set of the second protein.

A false discovery rate of 1% at both the protein and peptide level was used. Peptide posterior error probabilities were calculated by deriving, with Bayes' theorem, the probability of a false identification for a top scoring peptide from Mascot score and peptide sequence length-dependent histograms. Protein group posterior error probabilities were calculated by multiplying the posterior error probabilities of the contained peptide sequences. Each distinct peptide sequence contributed only one factor, the posterior error probability of the MS/MS spectrum for a given peptide sequence with the best posterior error probability value. All MS/MS spectra associated with a peptide sequence, which might consist of resequencing events on the same peak, sequencing on different isotopic peaks in the same isotope pattern, sequencing of different charge states, or different SILAC or modification states, were used to calculate the protein false discovery rate only with one single peptide posterior error probability. Protein groups were then sorted by the posterior error probability, and a given false discovery rate was ensured by terminating a list of proteins so that a given percentage of reverse proteins were contained. The used false discovery rate of 1% ensured that at most 1% of the proteins were wrongly identified. Peptides that lost all proteins contained within a group after this procedure were also removed from the peptide list. In addition to the protein false discovery rate threshold, proteins were considered identified by at least two unique sequence peptides quantified with at least one quantifiable SILAC pair. No outliers were removed because of the use of median instead of average values.

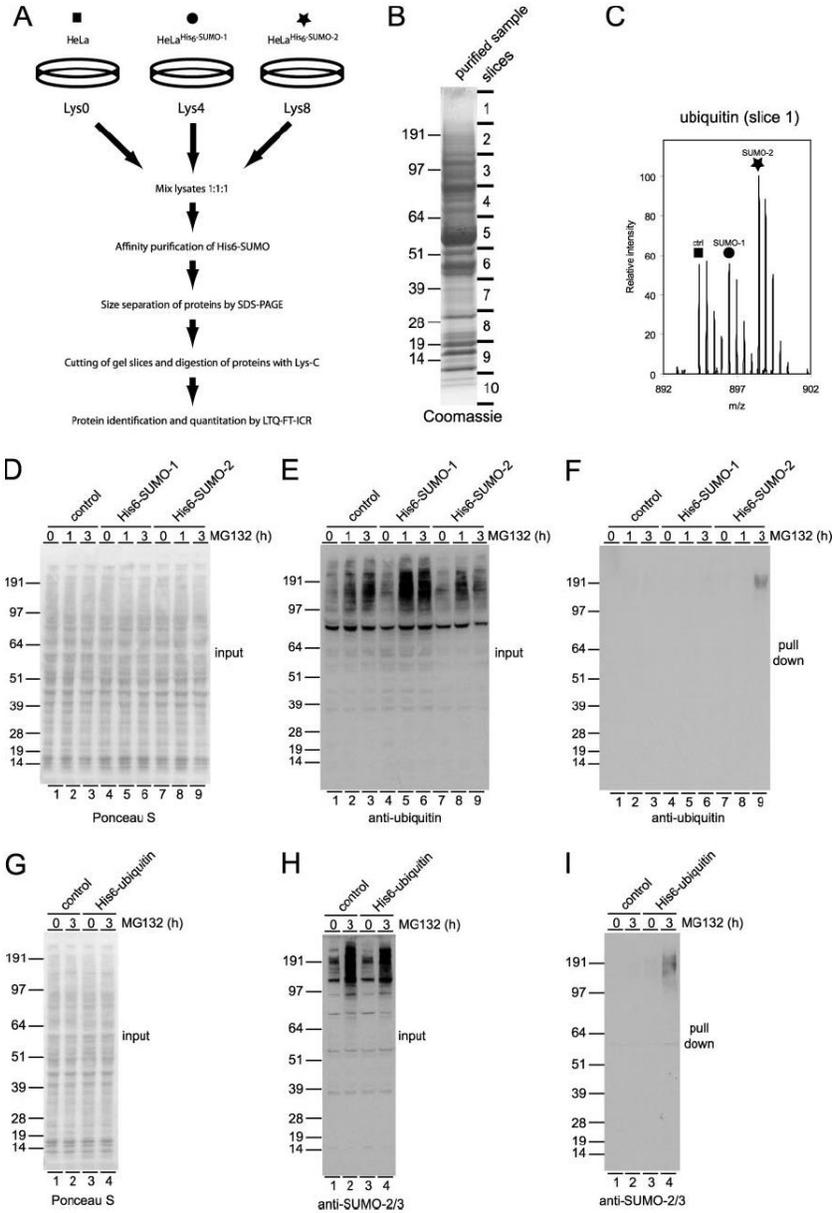
Significance of protein ratios was calculated in two different ways. Significance A was calculated by first estimating the variance of the distribution of all protein ratios in a non-parametric way and then reporting the error function for the z score corresponding to the given ratio. A robust and asymmetrical estimate of the standard deviation was obtained by calculating the 15.87, 50, and 84.13 percentiles  $r_{-1}$ ,  $r_0$ , and  $r_1$ , which correspond to  $1\sigma$  in each direction from the average.  $r_1 - r_0$  was defined as the right- and  $r_0 - r_{-1}$  was defined as left-sided robust standard deviations. In case of normally distributed data,  $r_1 - r_0$  and  $r_0 - r_{-1}$  would be equal to each other (conventional definition of standard deviation). The distance of a ratio  $r > r_0$  from the main distribution is measured in terms of the right standard deviation as follows.

$$z = \frac{r - r_0}{r_1 - r_0} \quad (\text{Eq. 1})$$

An analogous calculation is defined for  $r < r_0$ . Significance A is the value of the complementary error function for z above, which for a normal distribution corresponds to the probability of obtaining a value this large or larger by chance. Significance B was calculated using the same strategy, but in addition it is based on the dependence of the distribution on the summed protein intensity. We consider a protein as up-regulated if its significance B was below 0.001 and the ratio was higher than 1, down-regulated with significance B below 0.001 and ratio lower than 1, and not regulated if significance B was higher than 0.001.

Raw mass spectrometric files are stored at Tranche, a public repository for sharing scientific data. From the Tranche Website files

Cross-talk between SUMO-2/3 and the Proteasome



can be downloaded with the following hash: 51ebd8yrJ-ijxprwSrylms0w9M4Zhr6nwulAmeLgyGs4WgqWQ4e8vcMgaqspsuRcEwBJ41hpFV6ZILZBVLdEMWQJO5oAAAAAAAFJA==.

To address the reproducibility of the data, independent control immunoblotting experiments were performed (Fig. 4). We used antibodies directed against six different proteins identified by mass spectrometry and confirmed increases and decreases in sumoylation of these proteins upon inhibition of the proteasome, showing the reproducibility of the data.

## RESULTS

**Ubiquitin Co-purifies Preferentially with SUMO-2 Conjugates**—The quantitative proteomics experiment that we described previously to show that SUMO-1 and SUMO-2 are conjugated to preferential sets of target proteins (6) was repeated using lysine encoding instead of arginine encoding (Fig. 1A). Control HeLa cells were labeled with Lys0, HeLa cells stably expressing His<sub>6</sub>-SUMO-1 were labeled with Lys4, and HeLa cells stably expressing His<sub>6</sub>-SUMO-2 were labeled with Lys8. Cells were harvested, and nuclear lysates from the three different populations were mixed in a 1:1:1 ratio. His<sub>6</sub>-SUMO conjugates were subsequently purified and separated on a one-dimensional gel (Fig. 1B). The gel lane was cut in slices, and proteins were in-gel-digested by Lys-C and analyzed by mass spectrometry. Interestingly we identified two different ubiquitin peptides from the top slices of the gel lane that were preferentially enriched in the Lys8-encoded form, indicating that endogenous ubiquitin co-purified with large His<sub>6</sub>-SUMO-2 conjugates but not with large His<sub>6</sub>-SUMO-1 conjugates (Fig. 1C and supplemental Fig. S1). The amount of ubiquitin that co-purified with His<sub>6</sub>-SUMO-2 conjugates significantly increased upon inhibition of the proteasome (Fig. 1, D–F). The reverse experiment showed that endogenous SUMO-2/3 co-purified with His<sub>6</sub>-ubiquitin conjugates (Fig. 1, G–I).

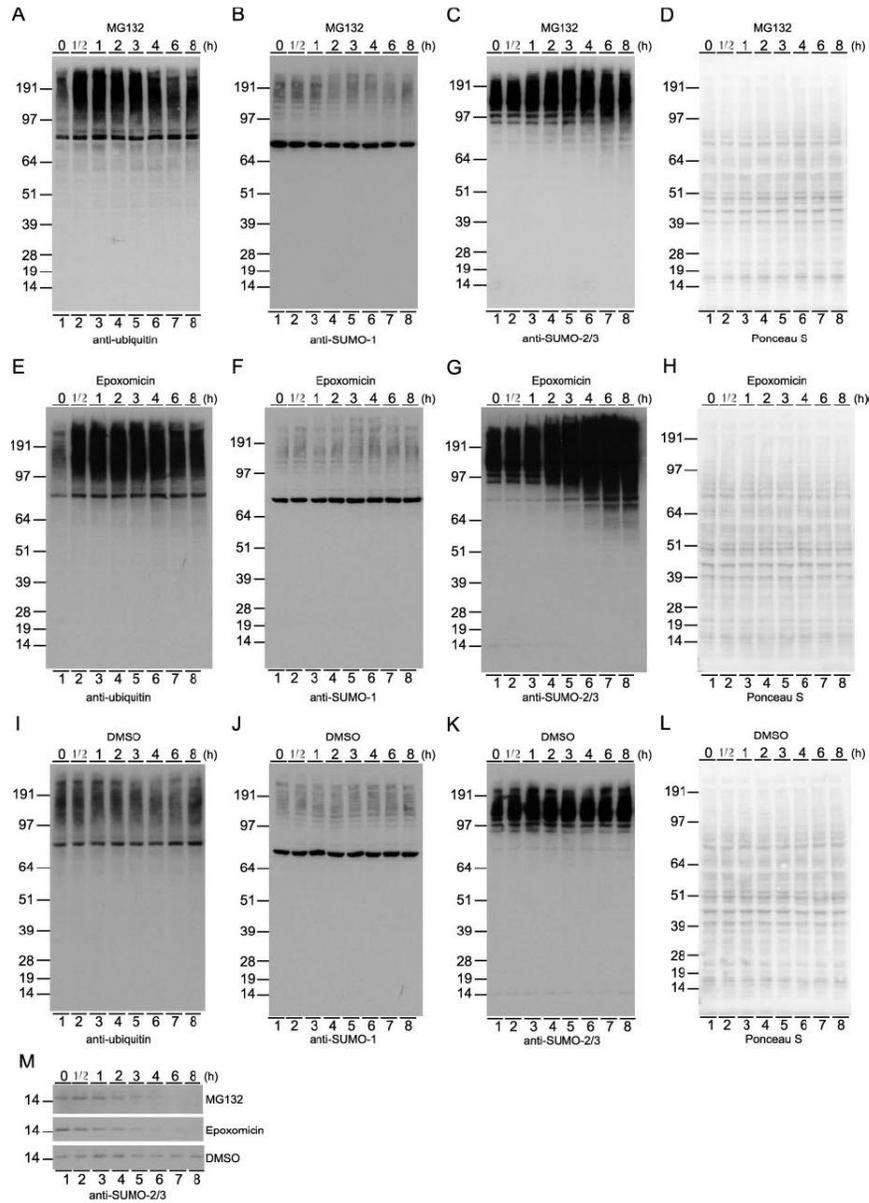
**Endogenous SUMO-2/3 Conjugates Accumulate in Cells Treated with Proteasome Inhibitors**—To study the effect of proteasome inhibition on endogenous SUMO-1 and endogenous

SUMO-2/3, HeLa cells were treated with MG132 or epoxomicin for up to 8 h (Fig. 2). These inhibitors caused rapid accumulation of ubiquitin in cells (Fig. 2, A and E) and in addition caused the accumulation of SUMO-2/3 conjugates albeit with slower kinetics (Fig. 2, C and G). Simultaneously the amount of non-conjugated SUMO-2/3 was significantly reduced by these inhibitors (Fig. 2M). The total amount of SUMO-1 conjugates in cells was not affected by MG132 or epoxomicin treatments, although the pattern of SUMO-1 conjugates slightly changed (Fig. 2, B and F). Control DMSO treatments did not affect ubiquitination or sumoylation (Fig. 2, I–K).

**Identification of SUMO-2 Target Proteins That Are Affected by the Proteasome Inhibitor MG132**—To identify individual SUMO-2 target proteins that are sensitive to proteasome inhibition, quantitative proteomics was used (Fig. 3A). Two pools of HeLa<sup>His<sub>6</sub>-SUMO-2</sup> cells were used for this experiment, the first pool of cells was labeled with Arg0 and Lys0 and treated with DMSO, and the second pool of cells was labeled with Arg10 and Lys8 and treated with MG132. It is important to note that this strategy was optimal for identifying SUMO-2 conjugates that were sensitive to proteasome inhibition, but it is likely that contaminating non-sumoylated proteins were co-purified. Especially the group of purified proteins with an unaltered ratio might contain a significant percentage of contaminants. Cells were harvested, and whole cell lysates were mixed in a 1:1 ratio. His<sub>6</sub>-SUMO-2 conjugates were subsequently purified, digested in solution with trypsin, and analyzed by mass spectrometry. A summary of the results is depicted in Fig. 3B. 847 proteins were identified by at least two unique peptides (supplemental Table 1 protein sheet), and in total 7643 peptides were identified by mass spectrometry (supplemental Table 1 peptide sheet). Interestingly two different subsets of MG132-sensitive SUMO-2 target proteins were identified; 73 target proteins showed a significant increase in sumoylation upon inhibition of the proteasome (sup-

**Fig. 1. SUMO-2 conjugates are enriched for ubiquitin.** A, a quantitative proteomics strategy to identify SUMO-1 and SUMO-2 conjugates. B, HeLa cells were labeled with Lys0, HeLa<sup>His<sub>6</sub>-SUMO-1</sup> cells were labeled with Lys4, and HeLa<sup>His<sub>6</sub>-SUMO-2</sup> cells were labeled with Lys8. Equal amounts of nuclear lysates from the three different populations were mixed, and proteins conjugated to His<sub>6</sub>-SUMO were purified. The SUMO-enriched fraction was separated by SDS-PAGE, proteins were visualized by Coomassie staining, the gel lane was cut in slices, and the proteins present in these slices were digested by Lys-C and identified by mass spectrometry. Peptide mass spectra were quantified to identify proteins potentially conjugated to SUMO-1 and/or SUMO-2. C, the His<sub>6</sub>-SUMO-2-purified fraction is specifically enriched for ubiquitin. Two different ubiquitin peptides (aa MQIFVK and TITLVEPSDTIENVK) were found to be enriched in His<sub>6</sub>-SUMO-2 conjugates but not in His<sub>6</sub>-SUMO-1 conjugates in the top part of the gel lane. The peptide mass spectra of the ubiquitin peptide TITLVEPSDTIENVK is shown. D–F, the proteasome inhibitor MG132 increases the amount of ubiquitin in His<sub>6</sub>-SUMO-2 conjugates. HeLa cells, HeLa<sup>His<sub>6</sub>-SUMO-1</sup> cells, and HeLa<sup>His<sub>6</sub>-SUMO-2</sup> cells were treated for 1 or 3 h with MG132 or were treated with DMSO for 3 h. Whole cell lysates were prepared, size-separated by SDS-PAGE, and transferred to a membrane. Total protein was visualized by Ponceau S staining (D), and the membrane was probed using antibody SC-8017 to detect ubiquitin (E). His<sub>6</sub>-SUMO conjugates from whole cell lysates were purified, size-separated by SDS-PAGE, transferred to a membrane, and probed to detect ubiquitin (F). G–I, the proteasome inhibitor MG132 increases the amount of SUMO-2/3 in His<sub>6</sub>-ubiquitin conjugates. HeLa cells were transfected with a plasmid that encodes His<sub>6</sub>-ubiquitin or with an empty control plasmid, and cells were subsequently treated with MG132 or DMSO for 3 h. Whole cell lysates were prepared, size-separated by SDS-PAGE, and transferred to a membrane. Total protein was visualized by Ponceau S staining (G), and the membrane was probed using antibody AV-SM23-0100 to detect SUMO-2/3 (H). His<sub>6</sub>-ubiquitin conjugates were purified from whole cell lysates, size-separated by SDS-PAGE, transferred to a membrane, and probed to detect SUMO-2/3 (I).

Cross-talk between SUMO-2/3 and the Proteasome

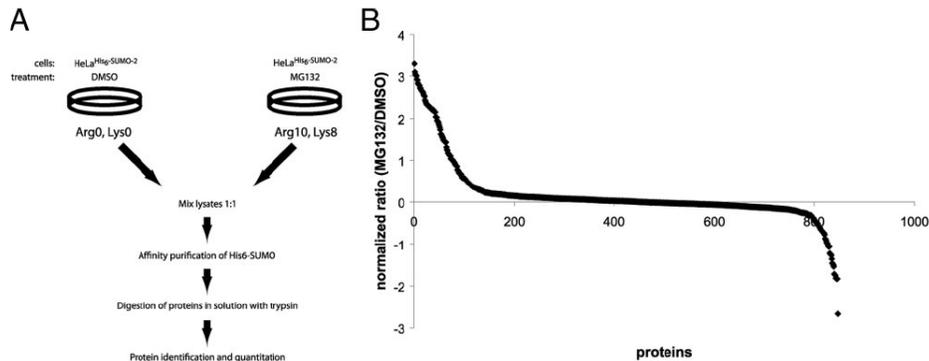


Downloaded from www.mcponline.org at MPI's MARTINSTRIED BIBLIOTHEK on January 7, 2009

ASBMB

Molecular & Cellular Proteomics





**Fig. 3. Dynamic alterations in the SUMO-2-conjugated proteome in response to the proteasome inhibitor MG132.** A, a quantitative proteomics strategy to study the effect of MG132 on the SUMO-2-conjugated proteome. HeLa<sup>His6-SUMO-2</sup> cells were labeled with Arg0 and Lys0 and treated with DMSO for 3 h, and a second pool of HeLa<sup>His6-SUMO-2</sup> cells was labeled with Arg10 and Lys8 and treated with MG132 for 3 h. Equal amounts of whole cell lysates from the two different populations were mixed, and proteins conjugated to His<sub>6</sub>-SUMO-2 were purified, digested by trypsin, and identified by mass spectrometry. Peptide mass spectra were quantified to identify MG132-induced changes in the SUMO-2-conjugated proteome. B, in total 847 proteins were identified by at least two unique peptides, including 73 proteins that were preferentially enriched upon MG132 treatment and 40 proteins that were significantly reduced in response to MG132. The natural logarithm of the SILAC ratio for each protein is depicted to visualize up-regulated and down-regulated proteins.

plemental Table 3), and 40 target proteins showed a significant decrease in sumoylation in response to MG132 (supplemental Table 4). The MG132-mediated increase in SUMO-2 conjugation was consistent with the immunoblot experiments (Fig. 2), whereas the MG132-mediated decrease in SUMO-2 conjugation of other target proteins was unexpected. The decrease in sumoylation of these target proteins might potentially be explained by a reduction in free SUMO-2 (Fig. 2M).

The largest functional group of MG132-regulated SUMO-2-conjugated proteins controls nucleic acid metabolism (supplemental Fig. S2). This group constitutes 47% of all MG132-up-regulated targets and 40% of all MG132-down-regulated targets and includes DNA repair factors, replication factors, helicases, basal transcription machinery components, transcription factors, chromatin modifiers, and RNA binding and processing factors (supplemental Fig. S2). In addition, both MG132-sensitive subsets contain a variety of other proteins, implicating that cross-talk between SUMO-2 and ubiquitin has a broad impact on cellular processes.

To confirm our findings independently, His<sub>6</sub>-SUMO-2 conjugates were purified from DMSO or MG132-treated cells, and immunoblotting experiments were carried out (Fig. 4). These experiments confirmed the accumulation of SUMO-2-conju-

gated forms of hnRNP M, MCM-7, and PIAS1 upon proteasome inhibition (Fig. 4, A-C) and a decrease of SUMO-2-conjugated forms of SAFB and SART1 (Fig. 4, E and F). We noticed a slight decrease in SUMO-2 conjugation of RanGAP1 in this experiment (Fig. 4D). Strikingly the total pools of these proteins were not affected by MG132, indicating that the ubiquitin-proteasome system specifically regulates SUMO-2-conjugated forms of these proteins (Fig. 4, A-F, *inputs*).

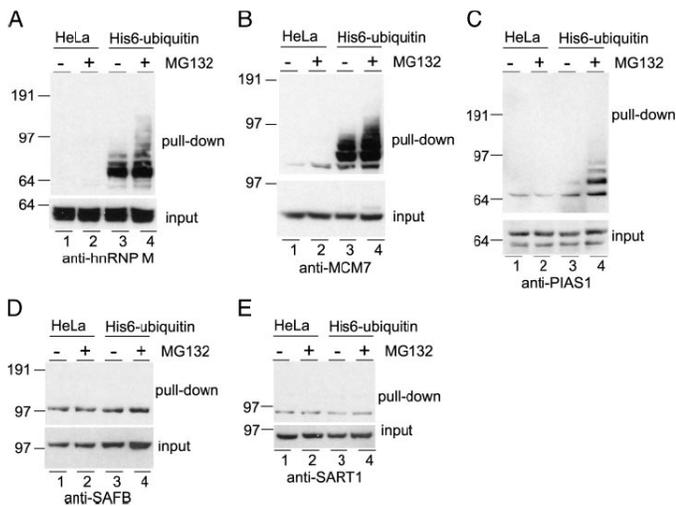
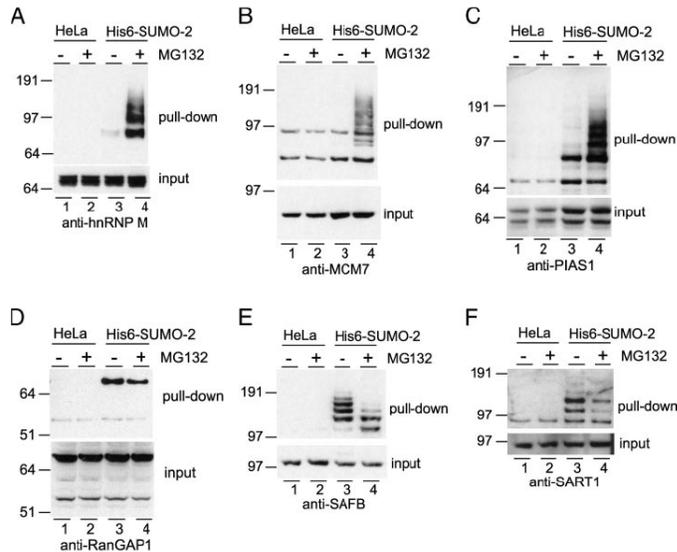
The most obvious explanation for our results would be that a subset of SUMO-2/3 target proteins is subsequently ubiquitinated and degraded by the proteasome. Inhibition of the proteasome system could then lead to proteins accumulating in the sumoylated and ubiquitinated form. To study the ubiquitination status of hnRNP M, MCM-7, PIAS1, SAFB, and SART1, similar experiments were carried out using His<sub>6</sub>-ubiquitin instead of His<sub>6</sub>-SUMO-2 (Fig. 5). As expected, hnRNP M, MCM-7, and PIAS1 were indeed ubiquitinated in an MG132-sensitive manner. PIAS1 (29) was also reported previously to be ubiquitinated. In contrast, ubiquitinated forms of SAFB or SART1 could not be detected in these experiments.

**SUMO-2/3 Chains Accumulate in Cells Treated with MG132**—Previously we have shown that SUMO-2 and SUMO-3 are able to multimerize in cells in a ubiquitin-like manner (11). To investigate whether SUMO-2/3 chains accu-

**Fig. 2. Endogenous SUMO-2/3 conjugates accumulate in cells treated with proteasome inhibitors.** A–M, HeLa cells were treated for the indicated periods of time with the proteasome inhibitors MG132 (A–D and M) or epoxomicin (E–H and M) or with DMSO (I–L and M). Whole cell extracts of HeLa cells were separated by SDS-PAGE, transferred to membranes, stained with Ponceau S to visualize total protein (D, H, and L), and probed using antibody SC-8017 to detect ubiquitin (A, E, and I), antibody 21C7 to detect SUMO-1 (B, F, and J), or antibody AV-SM23-0100 to detect SUMO-2/3 (C, G, K, and M).

## Cross-talk between SUMO-2/3 and the Proteasome

**FIG. 4. Dynamic alterations in the SUMO-2-conjugated proteome in response to the proteasome inhibitor MG132 detected by immunoblotting.** A–F, HeLa cells and HeLa<sup>His6-SUMO-2</sup> cells were treated with MG132 or DMSO for 3 h, and His<sub>6</sub>-SUMO-2 conjugates were purified from equal amounts of whole cell lysates. His<sub>6</sub>-SUMO-2 conjugates or equal amounts of whole cell extracts were size-separated by SDS-PAGE, transferred to membranes, and probed using antibodies to detect hnRNP M (A), MCM-7 (B), PIAS1 (C), RanGAP1 (D), SAFB (E), or SART1 (F). Note that the PIAS levels are slightly higher in the HeLa<sup>His6-SUMO-2</sup> cells compared with the HeLa cells.



**FIG. 5. hnRNP M, MCM7, and PIAS1 are conjugated to ubiquitin.** A–E, HeLa cells were transfected with a plasmid that encodes His<sub>6</sub>-ubiquitin or with an empty control plasmid, and cells were treated with MG132 or DMSO for 3 h. His<sub>6</sub>-ubiquitin conjugates were subsequently purified from equal amounts of whole cell lysates. His<sub>6</sub>-ubiquitin conjugates or equal amounts of whole cell extracts were size-separated by SDS-PAGE, transferred to membranes, and probed using antibodies to detect hnRNP M (A), MCM-7 (B), PIAS1 (C), SAFB (D), or SART1 (E).

ulated in cells treated with proteasome inhibitors, we searched for tryptic SUMO-SUMO peptides in purified His<sub>6</sub>-SUMO-2 conjugates from our quantitative proteomics experiments. Interestingly both SUMO-2-SUMO-2 and SUMO-2-SUMO-3 peptides were shown to accumulate upon MG132 treatment (Fig. 6, A and B).

To determine whether these SUMO polymers are functionally important for the processing of SUMO-2 targets by the proteasome, plasmids were generated that encode SUMO-2 mutants reduced for chain formation. HeLa cells were transiently transfected and treated with DMSO or MG132, and proteins conjugated to wild-type or mutant

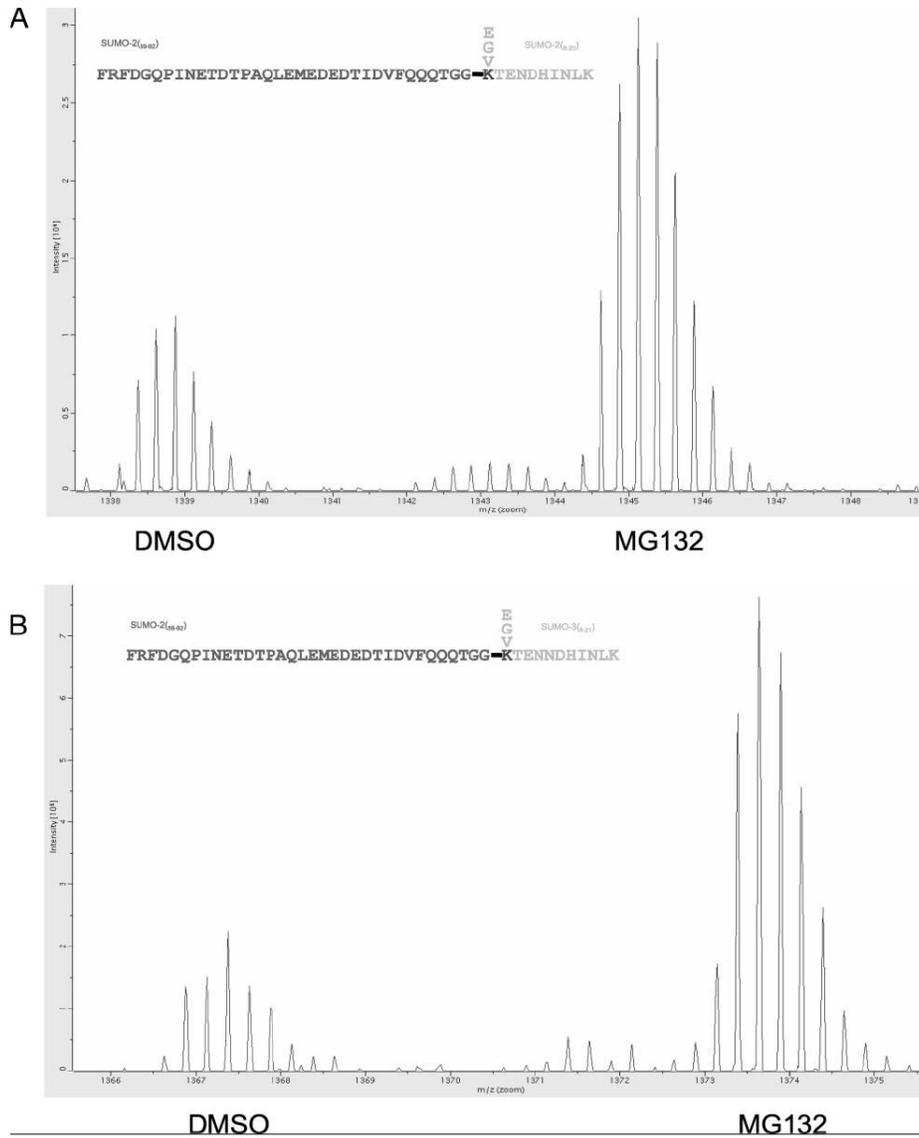


FIG. 6. SUMO chains accumulate in cells treated with MG132. A and B, a quantitative proteomics experiment was performed to identify changes in SUMO chains induced by MG132. HeLa<sup>His6</sup>-SUMO-2 cells were labeled with Arg0 and Lys0 and treated with DMSO for 3 h, and a second set of HeLa<sup>His6</sup>-SUMO-2 cells was labeled with Arg10 and Lys8 and treated with MG132 for 3 h. Equal amounts of whole cell lysates from the two different populations were mixed, and proteins conjugated to His<sub>6</sub>-SUMO-2 were purified, digested by trypsin in solution, and

## Cross-talk between SUMO-2/3 and the Proteasome

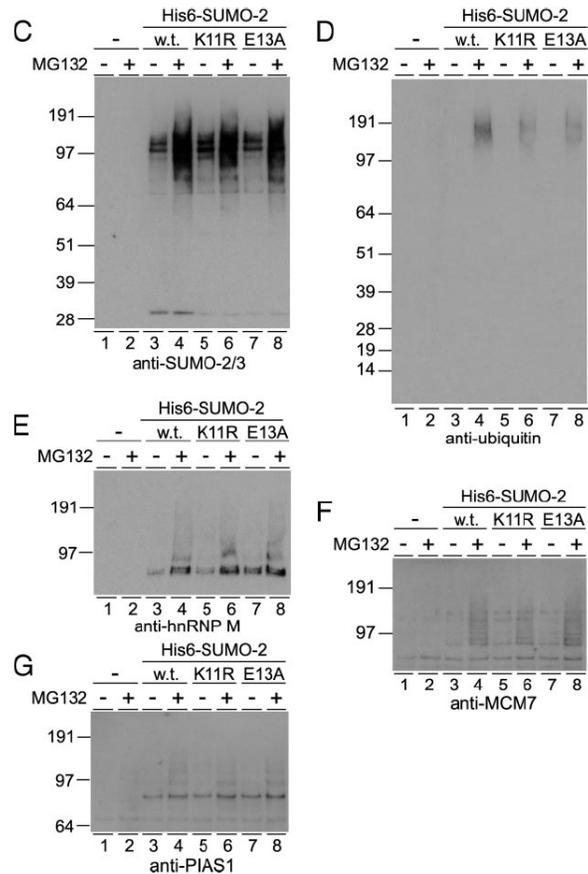


Fig. 6—continued

SUMOs were purified and analyzed by immunoblotting. The SUMO-2 mutants accumulated in MG132-treated cells similarly to wild-type SUMO-2 (Fig. 6C), and ubiquitin still copurified with the SUMO-2 mutants (Fig. 6D). Furthermore both wild-type and mutant SUMO-2-conjugated forms of hnRNP M, MCM-7, and PIAS1 accumulated upon inhibition of the proteasome (Fig. 6, E–G). We conclude that SUMO

chain formation is not required for the processing of SUMO-2 targets by the proteasome.

**Ubiquitination of SUMO-2/3**—The accumulation of ubiquitin in SUMO-2 conjugates could potentially be explained by the formation of mixed SUMO-2/ubiquitin chains and also by the conjugation of SUMO-2 and ubiquitin to independent lysines in target proteins. To test the formation of mixed

identified by mass spectrometry. A, MS spectrum of a tryptic peptide consisting of aa 59–92 of SUMO-2 and aa 8–20 of another molecule of SUMO-2 ( $m/z$  1338.1246 (4+); mass deviation,  $-1.21$  ppm). B, MS spectrum of a tryptic peptide consisting of aa 59–92 of SUMO-2 and aa 8–21 of SUMO-3 ( $m/z$  1366.6353 (4+); mass deviation,  $-1.65$  ppm). C–G, HeLa cells were transfected with plasmids that encode His<sub>6</sub>-tagged forms of wild type (w.t.) or K11R or E13A SUMO-2 mutants that are reduced for SUMO chain formation. Cells were treated with MG132 or DMSO for 5 h, and His<sub>6</sub>-SUMO-2 conjugates were purified from equal amounts of whole cell lysates. Purified fractions were size-separated by SDS-PAGE, transferred to membranes, and probed using antibodies to detect SUMO-2/3 (C), ubiquitin (D), hnRNP M (E), MCM-7 (F), or PIAS1 (G).

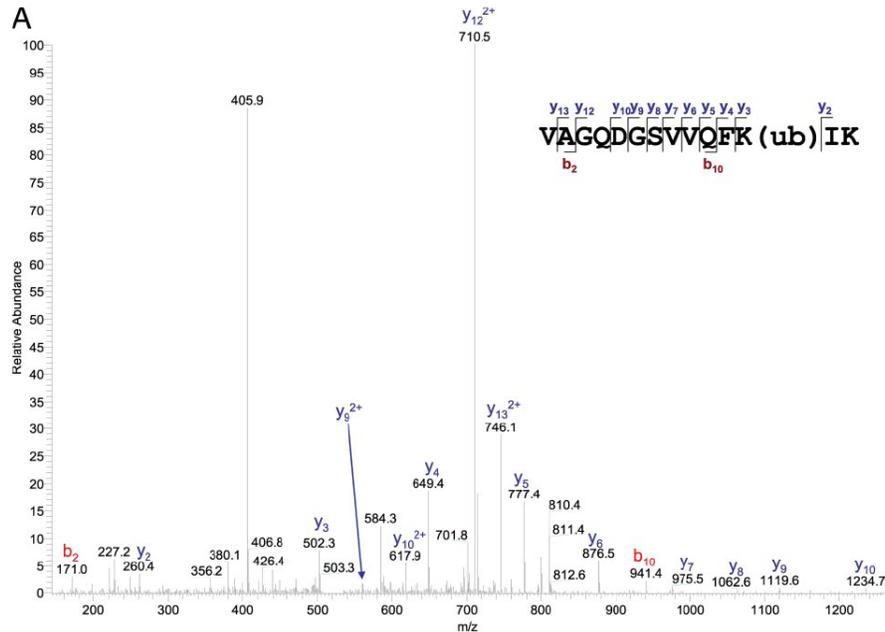


FIG. 7. SUMO-2 is ubiquitinated in cells. A, HeLa cells were transfected with a plasmid that encodes a His<sub>6</sub>-tagged lysine-deficient ubiquitin mutant. Cells were lysed, and His<sub>6</sub>-ubiquitin conjugates were purified and digested by trypsin. The digest was analyzed by mass spectrometry, and a peptide was identified that corresponds to ubiquitinated (ub) SUMO-2/3. The MS/MS fragmentation spectrum of this tryptic peptide consisting of aa 21–34 of SUMO-2 and the diglycine fragment of ubiquitin attached to lysine 32 of SUMO-2 is shown. This tryptic peptide is identical to a tryptic peptide consisting of aa 22–35 of SUMO-3 and the diglycine fragment of ubiquitin attached to lysine 33 of SUMO-3. Precursor ion mass was measured in the orbitrap mass spectrometer ( $m/z$  530.6264 (3+); mass deviation,  $-1.12$  ppm), and the peptide was fragmented and acquired in the LTQ mass spectrometer (Mascot score, 49.73; Mascot delta score, 3.28; posterior error probability,  $1.765 \times 10^{-23}$ ). B–F, His<sub>6</sub>-SUMO-2 plasmids were generated that encode K32R or K32R,K34R,K41R,K44R (4KR) mutants. HeLa cells were transfected with plasmids that encode His<sub>6</sub>-tagged forms of wild type (w.t.) or K32R or 4KR SUMO-2 mutants. Cells were treated with MG132 or DMSO for 5 h, and His<sub>6</sub>-SUMO-2 conjugates were purified from equal amounts of whole cell lysates. Purified fractions were size-separated by SDS-PAGE, transferred to membranes, and probed using antibodies to detect SUMO-2/3 (B), ubiquitin (C), hnRNP M (D), MCM-7 (E), or PIAS1 (F).

chains, His<sub>6</sub>-ubiquitin was purified from cells, digested, and analyzed by mass spectrometry to investigate the ubiquitination of endogenous SUMO-2/3. A high quality MS/MS spectrum was obtained showing the ubiquitination of endogenous SUMO-2 on lysine 32 or the ubiquitination of endogenous SUMO-3 on lysine 33 (Fig. 7A).

To determine whether mixed chains are functionally important for the processing of SUMO-2 targets by the proteasome, a plasmid was generated that encoded a SUMO-2 K32R mutant, and a second mutant was generated that lacked other adjacent lysines 34, 41, and 44 (designated 4KR). HeLa cells were transiently transfected and treated with DMSO or MG132, and proteins conjugated to wild-type or mutant SUMOs were purified and analyzed by immunoblotting. The SUMO-2 mutants accumulated in MG132-treated cells similarly to wild-type SUMO-2 (Fig. 7B), and ubiquitin co-purified

efficiently with the SUMO-2 mutants (Fig. 7C). Both wild-type and mutant SUMO-2-conjugated forms of hnRNP M, MCM-7, and PIAS1 all accumulated upon inhibition of the proteasome (Fig. 7, D–F). Moreover a lysine-deficient SUMO-2 mutant accumulated in MG132-treated cells similarly to wild-type SUMO-2, and both wild-type and mutant SUMO-2-conjugated forms of hnRNP M, MCM-7, and PIAS1 all accumulated upon inhibition of the proteasome (Fig. 8, A–E). We conclude that SUMO-2/ubiquitin mixed chain formation and SUMO-SUMO chain formation are not required for the processing of SUMO-2 targets by the proteasome. In principle, it is still possible that these SUMO-2 mutants form residual polymers with endogenous SUMO-2/3. However, dimers consisting of the lysine-deficient SUMO-2 mutant and endogenous SUMO-2/3 could not be detected by immunoblotting (Fig. 8A, lanes 5 and 6). Our data are compatible with consecutive sumoylation and ubiquiti-

## Cross-talk between SUMO-2/3 and the Proteasome

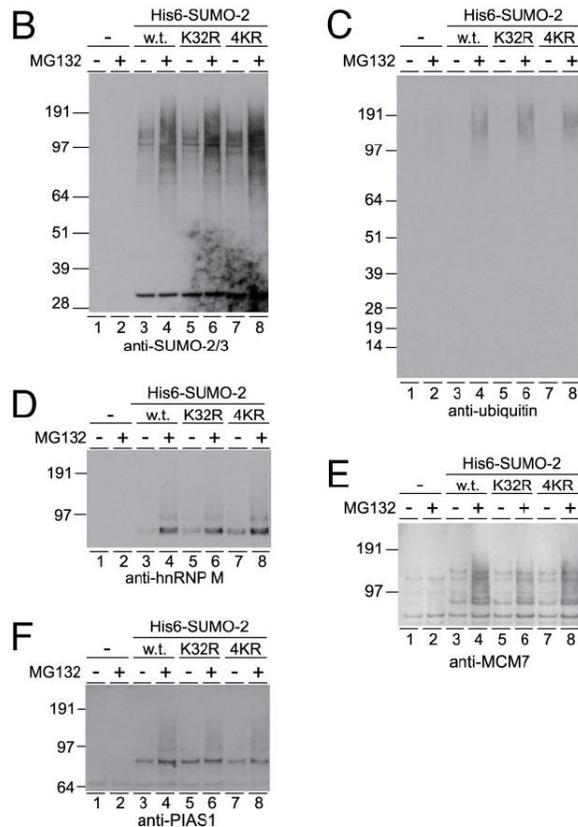


Fig. 7—continued

nation of a significant subset of SUMO-2 target proteins via independent lysines in these target proteins (Fig. 9). Upon ubiquitination, SUMO-2/3-conjugated proteins are degraded by the proteasome, enabling the recycling of SUMO-2/3.

## DISCUSSION

Ubiquitination and sumoylation are generally considered to be independent protein modifications. We have shown here that extensive cross-talk exists between ubiquitin and SUMO-2/3 conjugation of target proteins (Fig. 9). This cross-talk regulates selected SUMO-2/3 targets either directly or indirectly. A substantial percentage of SUMO-2/3 conjugates is directly regulated by subsequent ubiquitination and processing by the proteasome. A second subset of target proteins is indirectly dependent on the ubiquitin-proteasome system to provide a sufficiently large pool of free recycled SUMO-2/3 for subsequent rounds of conjugation.

The recent identification of the SLX5/8 complex in yeast and RNF4 in mammals (also known as small nuclear RING finger) provides mechanistic insight in the ubiquitination of sumoylated proteins (30–35). These proteins contain both a RING domain and SUMO interaction motifs and act as ubiquitin E3 ligases that are targeted to sumoylated proteins via these SUMO interaction motifs (31–33, 35–37). SLX5 and SLX8 are required for maintenance of the genome (31–35, 38–42). The identification of a large subset of nucleic acid-binding proteins that are sumoylated in a proteasome inhibitor-sensitive manner provides an important framework for further detailed analysis of the regulation of nucleic acid metabolism by SUMO-2/3 and the ubiquitin-proteasome system.

The precise roles of SUMO-SUMO chains as shown in Fig. 6, A and B, and of mixed SUMO-2/3 ubiquitin chains as shown in Fig. 7A is currently unclear. Because mutants that lack acceptor lysines still accumulate in cells treated with pro-

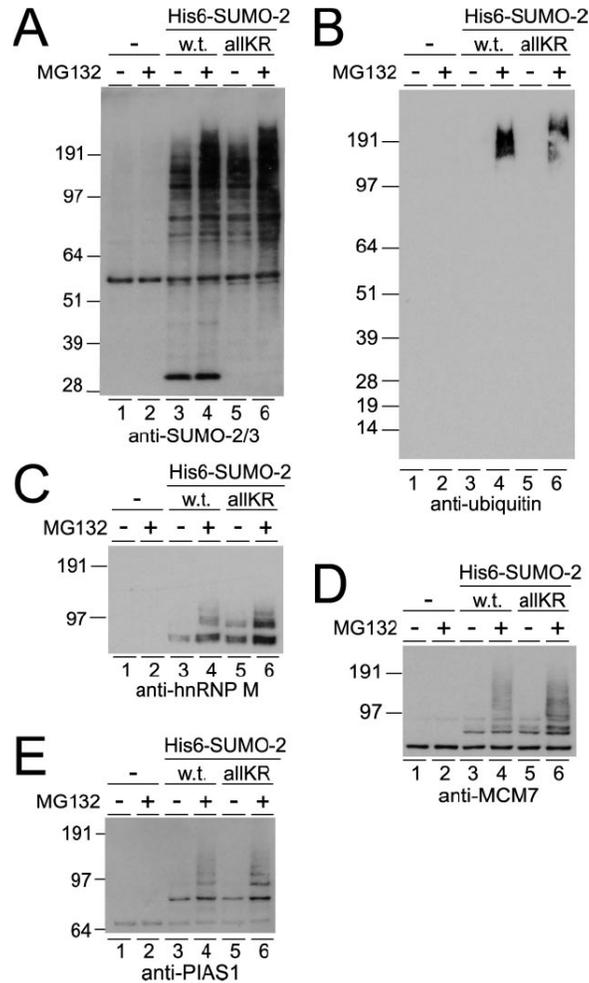


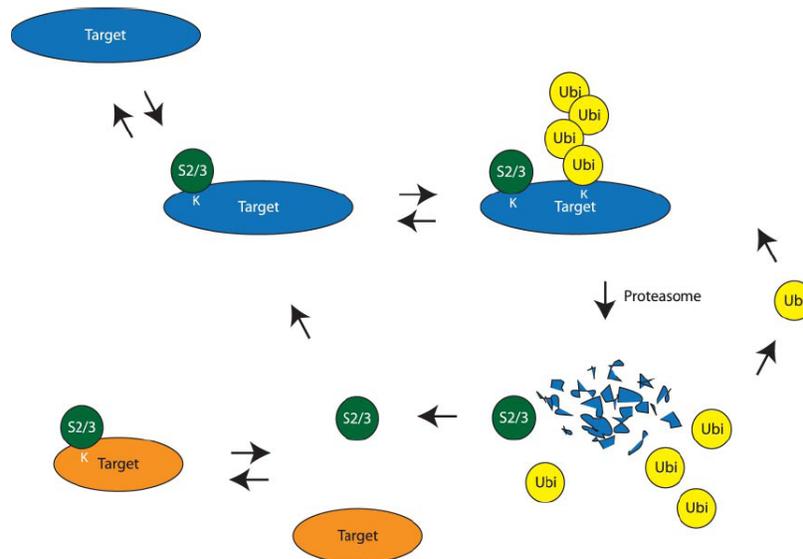
FIG. 8. A lysine-deficient SUMO-2 mutant is sensitive to proteasome inhibition. A-E, a His<sub>6</sub>-SUMO-2 plasmid was generated that encodes a lysine-deficient mutant (*allKR*). HeLa cells were transfected with plasmids that encode His<sub>6</sub>-tagged forms of wild type (*w.t.*) or the *allKR* SUMO-2 mutant. Cells were treated with MG132 or DMSO for 5 h, and His<sub>6</sub>-SUMO-2 conjugates were purified from equal amounts of whole cell lysates. Purified fractions were size-separated by SDS-PAGE, transferred to membranes, and probed using antibodies to detect SUMO-2/3 (A), ubiquitin (B), hnRNP M (C), MCM-7 (D), or PIAS1 (E).

teasome inhibitors and ubiquitin still co-purified similarly to wild-type SUMO forms, these chains are not required for targeting the majority of SUMO-2/3 targets to the proteasome. Nevertheless it is possible that certain proteins are targeted to the proteasome via these chains (36, 37). Alternatively these chains could regulate degradation-independent processes.

The conjugation of another subset of SUMO-2 target proteins was unaffected by the inhibition of the proteasome, including RanGAP1. Interestingly a very large fraction of RanGAP1 exists in a sumoylated form mainly conjugated to

SUMO-1 (6). It seems likely that sumoylated forms of RanGAP1 are very stable and have a very slow turnover compared with most other known target proteins. The half-life of the sumoylated forms of a subset of target proteins is regulated by ubiquitination and degradation by the proteasome as shown here but also via SUMO proteases (Sentrin-specific proteases) (43). Sumoylated RanGAP1 is localized at the cytoplasmic surface of the nuclear pore complex (44, 45); this might positively contribute to its stability because the majority of SUMO proteases are nuclear proteins (43).

### Cross-talk between SUMO-2/3 and the Proteasome



**FIG. 9. The ubiquitin-proteasome system is a component of the SUMO-2/3 cycle.** Our data indicate that the turnover of a subset of SUMO-2/3 (S2/3) conjugates is regulated by the ubiquitin-proteasome system. Target proteins that are conjugated to a single monomer of SUMO-2 or SUMO-3 or to multiple monomers can subsequently be ubiquitinated and degraded by the proteasome. This process enables the recycling of SUMOs to provide sufficient amounts of free SUMOs for new rounds of conjugation. In addition, mixed SUMO/ubiquitin (Ubi) chains and SUMO-SUMO chains are formed.

Although the sumoylated forms of hnRNP M, MCM-7, and PIAS1 and the ubiquitinated forms of these proteins are strongly stabilized by MG132, the total pools of these proteins were unaffected by this inhibitor. It is therefore likely that only a small fraction of these proteins are modified within the 3-h time frame of the experiment.

Transiently sumoylated proteins with short half-lives of the sumoylated forms because of the activity of SUMO proteases are likely to show the strongest decrease in sumoylation upon inhibition of the proteasome because of insufficient amounts of free SUMO-2/3. Therefore, the relative reduction in sumoylation of SUMO-2/3 target proteins upon inhibition of the proteasome is likely to reflect the relative instability of the sumoylated forms of these proteins.

Currently it is unclear why some SUMO-2/3 targets are subsequently ubiquitinated and degraded, whereas sumoylated forms of other SUMO-2/3 targets such as SART1 and SAFB are not subjected to detectable amounts of ubiquitination. Most likely, RNF4 plays a critical role and binds preferentially to a subset of SUMO-2/3 target proteins. Alternatively the localization of RNF4 in the cell might limit its activity to a subset of SUMO-2/3 target proteins in a manner similar to the SUMO proteases (43). It is interesting to note that RNF4 specifically localizes to PML bodies (46), sites

that are significantly enriched in SUMOs (22, 47). These nuclear bodies have also been shown to contain ubiquitin and proteasomal proteins (48), suggesting that the ubiquitination and degradation of sumoylated proteins can occur in PML bodies (36, 37).

The total levels of SUMO-1 were not affected by inhibition of the proteasome (Fig. 2) in contrast to a previous publication (49); nevertheless SUMO-1 accumulated in purified SUMO-2 in a manner similar to ubiquitin. Currently the role of SUMO-1 attached to SUMO-2 conjugates is unclear. Previously we have shown that SUMO-1 can directly be attached to lysine 11 of SUMO-2 (11). These mixed SUMO chains are unlikely to play a role in targeting proteins to the proteasome because SUMO-2 K11R or E13A mutants still accumulated upon proteasome inhibition, and ubiquitin efficiently co-purified with these mutants (Fig. 6).

In summary, we have shown here that the ubiquitin-proteasome system is an important component of the cellular SUMO-2/3 cycle that is required for the processing of a subset of SUMO-2/3 targets and the recycling of SUMO-2/3. The identification of other components of this pathway and the detailed dissection of the regulated target proteins will be an important step toward uncovering the connection between SUMO-2/3 and the ubiquitin-proteasome system in detail.



## Cross-talk between SUMO-2/3 and the Proteasome

\* This work was supported by the Netherlands Organisation for Scientific Research (NWO) (to A. C. O. V.) as part of the Innovational Research Incentives Scheme, by a generous grant from the Danish National Research Foundation (to J. S. A. and M. M.), and by a grant from The European Community (RUBICON, VI Framework) (to M. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ The on-line version of this article (available at <http://www.mcponline.org>) contains supplemental material.

¶ These authors contributed equally to this work.

\*\* To whom correspondence may be addressed. Tel.: 45-6550-2365; E-mail: jens.andersen@bmb.sdu.dk.

†† To whom correspondence may be addressed: Dept. of Molecular Cell Biology, Leiden University Medical Center, Postal zone S1-P, P. O. box 9600, 2300 RC Leiden, the Netherlands. Tel.: 31-71-5269621; Fax: 31-71-5268270; E-mail: vertegaal@lumc.nl.

## REFERENCES

- Goldberg, A. L. (2007) Functions of the proteasome: from protein degradation and immune surveillance to cancer therapy. *Biochem. Soc. Trans.* **35**, 12–17
- Mukhopadhyay, D., and Riezman, H. (2007) Proteasome-independent functions of ubiquitin in endocytosis and signaling. *Science* **315**, 201–205
- 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
- Welchman, R. L., Gordon, C., and Mayer, R. J. (2005) Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat. Rev. Mol. Cell Biol.* **6**, 599–609
- Nacerddine, K., Lehembre, F., Bhaumik, M., Artus, J., Cohen-Tannoudji, M., Babinet, C., Pandolfi, P. P., and Dejean, A. (2005) The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Dev. Cell* **9**, 769–779
- Vertegaal, A. C., Andersen, J. S., Ogg, S. C., Hay, R. T., Mann, M., and Lamond, A. I. (2006) Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. *Mol. Cell. Proteomics* **5**, 2298–2310
- Gill, G. (2005) Something about SUMO inhibits transcription. *Curr. Opin. Genet. Dev.* **15**, 536–541
- Wilson, V. G., and Heaton, P. R. (2008) Ubiquitin proteolytic system: focus on SUMO. *Expert Rev. Proteomics* **5**, 121–135
- Geiss-Friedlander, R., and Melchior, F. (2007) Concepts in sumoylation: a decade on. *Nat. Rev. Mol. Cell Biol.* **8**, 947–956
- Meulmeester, E., and Melchior, F. (2008) Cell biology: SUMO. *Nature* **452**, 709–711
- Matic, I., van Hagen, M., Schimmel, J., Macek, B., Ogg, S. C., Tatham, M. H., Hay, R. T., Lamond, A. I., Mann, M., and Vertegaal, A. C. (2008) In vivo identification of human SUMO polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. *Mol. Cell. Proteomics* **7**, 132–144
- Tatham, M. H., Jaffray, E., Vaughan, O. A., Desterro, J. M., Botting, C. H., Naismith, J. H., and Hay, R. T. (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
- Lallemand-Breitenbach, V., Zhu, J., Puvion, F., Koken, M., Honore, N., Doubaïkovsky, A., Duprez, E., Pandolfi, P. P., Puvion, E., Freemont, P., and de Thé, H. (2001) Role of promyelocytic leukemia (PML) sumoylation in nuclear body formation, 11S proteasome recruitment, and As2O3-induced PML or PML/retinoic acid receptor  $\alpha$  degradation. *J. Exp. Med.* **193**, 1361–1371
- Isik, S., Sano, K., Tsutsui, K., Seki, M., Enomoto, T., Saitoh, H., and Tsutsui, K. (2003) The SUMO pathway is required for selective degradation of DNA topoisomerase II $\beta$  induced by a catalytic inhibitor ICRF-193(1). *FEBS Lett.* **546**, 374–378
- Ulrich, H. D. (2005) Mutual interactions between the SUMO and ubiquitin systems: a plea of no contest. *Trends Cell Biol.* **15**, 525–532
- Desterro, J. M., Rodriguez, M. S., and Hay, R. T. (1998) SUMO-1 modification of I $\kappa$ B $\alpha$  inhibits NF- $\kappa$ B activation. *Mol. Cell* **2**, 233–239
- Huang, T. T., Wuerzberger-Davis, S. M., Wu, Z. H., and Miyamoto, S. (2003) Sequential modification of NEMO/I $\kappa$ B $\alpha$  by SUMO-1 and ubiquitin mediates NF- $\kappa$ B activation by genotoxic stress. *Cell* **115**, 565–576
- Moldovan, G. L., Pfander, B., and Jentsch, S. (2007) PCNA, the maestro of the replication fork. *Cell* **129**, 665–679
- Andersen, J. S., Lam, Y. W., Leung, A. K., Ong, S. E., Lyon, C. E., Lamond, A. I., and Mann, M. (2005) Nucleolar proteome dynamics. *Nature* **433**, 77–83
- Durocher, Y., Perret, S., and Kamen, A. (2002) High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res.* **30**, E9
- van der Horst, A., Vries-Smits, A. M., Brenkman, A. B., van Triest, M. H., van den Broek, N., Colland, F., Maurice, M. M., and Burgering, B. M. (2006) FOXO4 transcriptional activity is regulated by monoubiquitination and USP7/HAUSP. *Nat. Cell Biol.* **8**, 1064–1073
- Vertegaal, A. C., Ogg, S. C., Jaffray, E., Rodriguez, M. S., Hay, R. T., Andersen, J. S., Mann, M., and Lamond, A. I. (2004) A proteomic study of SUMO-2 target proteins. *J. Biol. Chem.* **279**, 33791–33798
- Jaffray, E. G., and Hay, R. T. (2006) Detection of modification by ubiquitin-like proteins. *Methods* **38**, 35–38
- Olsen, J. V., de Godoy, L. M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Hornung, S., and Mann, M. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics* **4**, 2010–2021
- Foster, L. J., de Hoog, C. L., and Mann, M. (2003) Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 5813–5818
- Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75**, 663–670
- Cox, J., and Mann, M. (2007) Is proteomics the new genomics? *Cell* **130**, 395–396
- Graumann, J., Hubner, N. C., Kim, J. B., Ko, K., Moser, M., Kumar, C., Cox, J., Schoeler, H., and Mann, M. (2008) Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol. Cell. Proteomics* **7**, 672–683
- Depaux, A., Regnier-Ricard, F., Germani, A., and Varin-Blank, N. (2007) A crosstalk between hSiah2 and Pias E3-ligases modulates Pias-dependent activation. *Oncogene* **26**, 6665–6676
- Wang, Z., Jones, G. M., and Prlich, G. (2006) Genetic analysis connects SLX5 and SLX8 to the SUMO pathway in *Saccharomyces cerevisiae*. *Genetics* **172**, 1499–1509
- Prudden, J., Pebernard, S., Raffa, G., Slavina, D. A., Perry, J. J., Tainer, J. A., McGowan, C. H., and Boddy, M. N. (2007) SUMO-targeted ubiquitin ligases in genome stability. *EMBO J.* **26**, 4089–4101
- Uzunova, K., Gottsche, K., Miteva, M., Weisshaar, S. R., Glanemann, C., Schnellhardt, M., Niessen, M., Scheel, H., Hofmann, K., Johnson, E. S., Praefcke, G. J., and Dohmen, R. J. (2007) Ubiquitin-dependent proteolytic control of SUMO conjugates. *J. Biol. Chem.* **282**, 34167–34175
- Sun, H., Levenson, 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
- 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
- Xie, Y., Kerscher, O., Kroetz, M. B., McConchie, H. F., Sung, P., and Hochstrasser, M. (2007) The yeast Hex3-Slx8 heterodimer is a ubiquitin ligase stimulated by substrate sumoylation. *J. Biol. Chem.* **282**, 34176–34184
- Tatham, M. H., Geoffroy, M. C., Shen, L., Plechanovova, A., Hattersley, N., Jaffray, E. G., Palvimo, J. J., and Hay, R. T. (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nat. Cell Biol.* **10**, 538–546
- Lallemand-Breitenbach, V., Jeanne, M., Benhenda, S., Nasr, R., Lei, M., Peres, L., Zhou, J., Zhu, J., Raught, B., and de Thé, H. (2008) Arsenic degrades PML or PML-RAR $\alpha$  through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat. Cell Biol.* **10**, 547–555

## Cross-talk between SUMO-2/3 and the Proteasome

38. Burgess, R. C., Rahman, S., Lisby, M., Rothstein, R., and Zhao, X. (2007) The Slx5-Slx8 complex affects sumoylation of DNA repair proteins and negatively regulates recombination. *Mol. Cell. Biol.* **27**, 6153–6162
39. Azam, M., Lee, J. Y., Abraham, V., Chanoux, R., Schoenly, K. A., and Johnson, F. B. (2006) Evidence that the *S. cerevisiae* Sgs1 protein facilitates recombinational repair of telomeres during senescence. *Nucleic Acids Res.* **34**, 506–516
40. 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
41. Mullen, J. R., Nallaseeth, F. S., Lan, Y. Q., Slagle, C. E., and Brill, S. J. (2005) Yeast Rmi1/Nce4 controls genome stability as a subunit of the Sgs1-Top3 complex. *Mol. Cell. Biol.* **25**, 4476–4487
42. 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
43. Mukhopadhyay, D., and Dasso, M. (2007) Modification in reverse: the SUMO proteases. *Trends Biochem. Sci.* **32**, 286–295
44. 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
45. 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
46. Hakli, M., Karvonen, U., Janne, O. A., and Palvimo, J. J. (2005) SUMO-1 promotes association of SNURF (RNF4) with PML nuclear bodies. *Exp. Cell Res.* **304**, 224–233
47. 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
48. Wojcik, C., and DeMartino, G. N. (2003) Intracellular localization of proteasomes. *Int. J. Biochem. Cell Biol.* **35**, 579–589
49. Bailey, D., and O'Hare, P. (2005) Comparison of the SUMO1 and ubiquitin conjugation pathways during the inhibition of proteasome activity with evidence of SUMO1 recycling. *Biochem. J.* **392**, 271–281

## **6 Profiling of the SUMO substrates proteome after heat shock**

**MANUSCRIPT: System-wide changes in SUMO modification in response to heat shock**

The following pages contain the submitted version of the manuscript.

It presents the most in-depth characterization of the SUMO substrates proteome and its large scale systems-wide profiling after heat shock and heat shock recovery.

# **System-wide changes in SUMO modification in response to heat shock**

Filip Golebiowski\*<sup>1</sup>, Ivan Matic\*<sup>2</sup>, Michael H. Tatham\*<sup>1</sup>, Christian Cole<sup>3</sup>, Jurgen Cox<sup>2</sup>,  
Akihiro Nakamura<sup>4</sup>, Geoffrey J. Barton<sup>3</sup>, Matthias Mann<sup>2</sup>, Ronald T. Hay<sup>1</sup>

1 – Wellcome Trust Centre for Gene Regulation and Expression, College of Life  
Sciences, University of Dundee, DD1 5EH. Scotland. UK

2 - Proteomics and Signal transduction, Max-Planck-Institute for Biochemistry, Am  
Klopferspitz, D-82152 Martinsried, Germany

3 – School of Life Sciences Research, University of Dundee, Dow Street, Dundee, DD1  
5EH, UK.

4- Department of Molecular and Cellular Pharmacology, Faculty of Medicine, Gunma  
University, Graduate School of Medicine, 3-39-22 Showa-Machi, Maebashi, Gunma 371-  
8511, Japan

\*Authors contributed equally.

Word count; Introductory paragraph - 227

Main body text - 1697

All forms of cellular life need to sense and respond to extreme environmental and pathological conditions. In the absence of an appropriate reaction such stresses can lead to cell damage or death. The heat shock (HS) response is one of the most evolutionarily conserved of such defensive mechanisms, and is characterized by massive induction of HS genes, producing HS proteins (HSPs) which act to protect the cell from the cytotoxic stress(Lindquist, 1986). However, while the role of HSPs has been well documented, early signaling mechanisms have yet to be fully investigated. Post-translational modifications (PTMs) are commonly involved in rapid signal transduction, and conjugation of one such PTM, the small ubiquitin-like modifier 2 (SUMO-2), to an unknown set of cellular substrates is induced by HS(Saitoh & Hinchev, 2000). Here we employ stable isotope labeling of amino acids in cell culture (SILAC)(Mann, 2006), a stringent SUMO purification protocol and advanced mass spectrometric technology to identify over 750 SUMO-2 substrates and quantify changes in their sumoylation after HS. As a response to HS SUMO is polymerized into polySUMO chains and redistributed between a wide variety of proteins involved in cell cycle regulation, apoptosis, protein trafficking, folding and degradation, mRNA transcription and translation, and DNA replication, recombination and repair. This is the most comprehensive substrate proteomic analysis of a ubiquitin-like modifier (Ubl) and identifies a pervasive role for SUMO in the biologic response to hyperthermic stress.

Owing to their dynamic nature, reversible PTMs involving the addition and removal of molecular adducts are capable of rapid transduction of cellular signals. Mass spectrometry is well-suited to the analysis of PTMs in general(Witze et al, 2007) and sumoylation in particular(Matic et al, 2008b; Tatham et al, 2008; Tatham et al, 2001). To better understand the role of SUMO-2 in the HS response we undertook a systems-wide and quantitative proteomic analysis of SUMO-2 targets both during and after HS.

We generated HeLa cells stably expressing SUMO-2 fused to a tandem affinity protein (TAP) tag(Rigaut et al, 1999) (Fig. 1a), that is expressed at similar levels to endogenous SUMO-2, and responds to HS in a similar manner to the endogenous protein (Fig. 1b). Comparisons between TAP only and TAP-SUMO-2 purifications under initially denaturing conditions confirm the high stringency of this method (Fig. 1c). For proteome quantitation, we employed SILAC labeling(Mann, 2006) using arginine and lysine to compare purifications from HeLa cells expressing TAP alone with two pools of TAP-SUMO-2 HeLa cells; either unstressed or subjected to HS (Fig. 1d). This allowed 1159 proteins to be quantified. Among them, the heat-shock transcription factor 1 (HSF1) displays dramatically increased SUMO modification after HS (Fig. 1d), which is consistent with previous studies(Hietakangas et al, 2003). The changes to all proteins were visualized on a triple-SILAC map (tsMap) (Fig. 1e), in which each protein is located by coordinates derived from the TAP-SUMO-2/TAP and TAP-SUMO-2 HS/TAP-SUMO-2 ratios. Although contaminants were almost invisible by silver-staining (Fig. 1c), they were still detected by this approach and separated in the tsMap from SUMO substrates (Supp. Fig. 1a). This analysis allowed the unambiguous identification

of 660 SUMO-2 targets (Supp. File 1). Although the *in vivo* identification of SUMO branched peptides is technically challenging, the known (RanGAP1, SUMO-2 and SUMO-3 Supp. Fig. 2) and novel (Antigen KI-67, SAF-B1 – Supp. Fig. 3) modification sites were identified, by manual screening and interpretation of spectra (Matic et al, 2008b).

During HS, SUMO-2 modification increases and free SUMO-2 is depleted (Saitoh & Hinchey, 2000) (Fig. 1b). Although analysis of the tsMap indicates that HS induces the sumoylation of a significant number of proteins, unexpectedly this occurs concomitantly to the desumoylation of a different subset, which is detectably modified under basal conditions. This is apparent as a close correlation between TAP-SUMO-2/TAP and TAP-SUMO-2-HS/TAP-SUMO-2 ratios (Suppl. Fig. 1b), and suggests not simply an increase in global sumoylation, but a dynamic exchange of SUMO between substrates upon HS.

To better understand the temporal dynamics of the sumoylation response to HS we applied this robust and sensitive protocol to monitor the system-wide changes in sumoylation of proteins both upon HS and after a two-hour recovery (HSR) period (Supp. Fig. 4, Fig. 2a). This experiment quantified 1255 proteins (Fig. 2c), which, after tsMap filtering, yielded 628 SUMO-2 target proteins (Supp. Fig. 5a, Supp. File 1). There was excellent agreement between the two experiments for the HS response (Supp. Fig. 5c&d), and together a total of 765 SUMO-2 substrates were identified. Currently single protein studies have documented 265 SUMO targets, of which this study identified a greater proportion than all eleven previous SUMO substrate proteomics studies taken together (Fig 2 and Supp. Fig. 6). Furthermore, our SUMO-2 substrate proteome contains a

comparable number of  $\psi$ KxE SUMO conjugation consensus motifs(Rodriguez et al, 2001) per protein to the 265 independently verified substrates (Table 1). Significantly, the TAP-SUMO-2 internal contaminants list contains almost 3 fold fewer and is comparable to a human proteome ‘background’ set (Table 1). This strongly suggests that the approach used here was successful in identifying SUMO substrates.

Although the SUMO-2 substrate proteome contains proteins from a broad range of molecular weights, HS induces the accumulation of very high molecular-weight SUMO conjugates (Figs. 1b&2a). Analysis of substrate distribution throughout the gel in the second SILAC experiment shows that HS-induced changes are most prominent in the upper regions (Fig. 3a left), with slices over 100kDa being predominantly populated with proteins of lower predicted molecular weight (Fig. 3a right). Also, individual proteins whose conjugation increases with HS, are generally present in broader molecular weight ranges than would be expected for conjugation to one or two SUMO adducts (Fig. 3b). As SUMO-2 and SUMO-3 share with ubiquitin the ability to form polymers(Matic et al, 2008b; Tatham et al, 2001), a potential explanation for this unexpectedly exaggerated protein migration is the formation of polySUMO-2/-3 adducts on target proteins. This is confirmed by quantitation of SUMO-2/SUMO-2 branched peptides which indicate a strong induction of SUMO polymerization via lysine 11 upon HS, that returns to basal levels after recovery (Fig. 3c). Whether in some cases these signal for degradation via the ubiquitin-proteasome pathway(Tatham et al, 2008), remains to be investigated.

Many proteins known to be directly involved in the cellular HS response such as heat shock transcription factors 1 and 2(Lindquist, 1986), members of the HSP 40kDa,

60kDa and 70kDa families(Lindquist, 1986), the co-chaperonins Sti1, RUVBL1 and RUVBL2(Pearl et al, 2008), the HS signal transduction kinase p38-MAPK14(Dorion & Landry, 2002) and the translocated promoter region (TPR) protein(Skaggs et al, 2007), are SUMO-2 targets (Supp. File 1). More generally, the HS response triggers changes in transcription, translation, apoptosis, cell cycle control, protein folding and protein degradation, that coordinately determine whether a cell exposed to hyperthermia will die or survive and become stress tolerant(Beere, 2005; Bond, 2006; Lindquist, 1986). Proteins involved in all of these processes are significantly over-represented in the SUMO-2 modified proteome, in addition to proteins involved in DNA replication, recombination and repair (Fig. 4a, Supp. Figs. 7a and 8). This latter observation is likely to be a consequence of HS-induced DNA double-strand breaks(Kaneko et al, 2005; Takahashi et al, 2004), and there is significant overlap with the phosphoproteome of the DNA damage responsive kinases ATM/ATR(Matsuoka et al, 2007) (Supp. Figs. 9 b&c) with notable exception of proteins involved in ubiquitin-dependent protein turnover (Supp. Fig. 10).

While the global increase in SUMO conjugation observed by Western-blotting in response to HS is rapid, with all free SUMO being conjugated into high molecular weight adducts after 5 minutes, the deconjugation of SUMO during the recovery period is much slower (Supp. Fig. 4). Systems-wide analysis of individual proteins by tsMap reveals a general trend (Supp. Fig. 5b) whereby substrates sumoylated upon HS become deconjugated during recovery (e.g. glucocorticoid receptor – Fig. 2b), but those demodified during HS do not tend to recover their conjugation state after two hours (eg.

SAF-B1 – Fig. 2b). To extract biologically relevant information from these data we developed an analysis which maps protein members from significantly clustering GO groups onto tsMaps, allowing changes in their sumoylation with respect to the entire dataset to be easily visualized (Fig. 4b-g). Inspection of these GOtsMaps shows that like most large GO terms the transcription factors (Fig. 4b) do not exhibit significant group regulation (Supp. Table 1), however, sub-categories do show coordinated desumoylation (Fig. 4c). This group includes the dimerising transcription factors Jun, FOS and the V-maf musculoaponeurotic fibrosarcoma oncogene homolog (MAF) proteins (Fig. 4h), which act upstream of the oxidative stress-response transcription factor NRF2 (Supp Fig. 11). As Jun is activated in response to HS(Dai et al, 1995) and repressed by sumoylation(Muller et al, 2000), it seems likely that rapid desumoylation of Jun plays a role in the early response to HS.

Transcriptional control during hyperthermic stress also acts more globally to facilitate the expression of HSPs while repressing non-HS genes (Lindquist, 1986). Consistent with this many SUMO-2 substrates are involved in chromatin remodelling. For example the SWI/SNF related matrix associated actin dependent regulators of chromatin (SMARCs) (Fig. 4i), which are known to be involved in HS-induced chromatin remodeling(Shivaswamy & Iyer, 2008) are all deconjugated upon HS. Recently, clear roles for SUMO modification in the recruitment of repressive complexes have been established for KRAB-ZFP(Ivanov et al, 2007) and Sp3(Stielow et al, 2008). SUMO modification of Sp3 or the KRAB-ZPF co-repressor KAP1 is thought to mediate successive recruitment of the NuRD complex to deacetylate histones and the SETDB1

complex to methylate histones, which in turn bind HP1 proteins to maintain the silent chromatin state. The remarkable finding from our analysis is that almost all of the components of the above mentioned protein complexes are SUMO modified: transcription factors Sp3 and KRAB-ZFP; corepressor KAP1; NuRD complex (HDAC1, Mi2 $\alpha$ , Mi2 $\beta$ , MTA2 and RbAP46); SETDB1 complex (SETDB1, MBD1 and ATF7IP) and HP1 $\alpha$  and  $\gamma$ . An alternative route to histone deacetylation is recruitment of the SIN3 complex and again virtually all of its components are SUMO modified (Sin3b, SDS3, SAP18, SAP130 and SAP180). Consistent with previous work (Nathan et al, 2006; Shiio & Eisenman, 2003) H2A, H2B and H4 histone subfamilies were identified as SUMO substrates (Supp. File 1), and are coordinately deconjugated from SUMO upon HS (Supp. Fig. 7b). While histone modification represents one path to epigenetic modification, DNA methylation is also critical (Ooi & Bestor, 2008). In this respect the DNA methyltransferase DNMT1 is subject to SUMO modification and its associated cofactor UHRF1 displays a 45 fold increase in SUMO modification after heat shock. Together these observations suggest a considerably more extensive role for SUMO modification in the regulation of transcriptional repression than was previously suspected.

RNA binding proteins cluster to two distinct regions of the tsMap (Fig. 4d) indicating significant conjugation or deconjugation in response to HS. Almost all the small nuclear ribonucleoproteins (snRNPs) are desumoylated (Fig. 4e & 1), while the heterogeneous ribonucleoproteins (HNRNPs) show both increased and decreased sumoylation with HS (Fig. 4l). Many importins and nuclear pore complex proteins also

seem to be SUMO-2 targets (Fig. 4n), indicating that both protein and RNA nucleocytoplasmic transport features are regulated by SUMO.

Proteins involved in DNA repair (Fig. 4f&j, Supp. Fig. 12b), DNA recombination (Fig. 4g) and DNA damage response (Fig. 4m, Supp. Fig. 12a) almost exclusively increase in sumoylation in response to HS. A striking functional group within these are those involved in the initiation of DNA replication (Fig. 4k). This group contains three components of the origin recognition complex (Orcs 2L, 3L and 6L) and four subunits of the MCM replicative helicase (MCMs 2, 3, 4 and 7). The targeting of these proteins for sumoylation after HS suggests a hitherto unrecognized role for SUMO in the regulation of DNA replication, and is consistent with the observation that delaying S-phase progression rescues cells from heat-induced S-phase hypertoxicity (VanderWaal et al, 2001).

In summary, a combination of high stringency purification, high-resolution mass spectrometry, quantitative proteomic methods and novel data analysis allowed the identification of over 550 novel SUMO targets and shows that regulated sumoylation of important cellular mechanisms occurs during hyperthermic stress. We predict that this study will provide both a platform for further investigations into the role of SUMO in a wide variety of cellular functions, and a template for the systems-wide study of ubiquitin and other Ubls.

## **Methods Summary**

### *Cell culture and SILAC labeling*

Quantitative proteomic experiments were performed using the SILAC technique as described (Ong & Mann, 2007): “SILAC experiment 1” compared TAP only cells, with unstimulated TAP-SUMO-2 cells and TAP-SUMO-2 cells stimulated by heat-shock (43°C for 30 minutes). Cells were grown in modified Dulbecco’s modified Eagle’s medium lacking amino-acids with 10% FCS dialysed against PBS with 3500 molecular weight cut-off membrane. Medium was enriched with all amino-acids except L-lysine and L-arginine, which were replaced with stable isotope (SILAC) forms (Cambridge Isotope Laboratories) depending on the treatment: Untreated TAP-only cells were grown in the presence of isotopically normal lysine (Lys0) and arginine (Arg0). TAP-SUMO-2 cells cultured at 37°C only were grown in the presence of 4,4,5,5-D<sub>4</sub> lysine (Lys4) and <sup>13</sup>C<sub>6</sub> arginine (Arg6), and heat shocked cells were grown in the presence of <sup>13</sup>C<sub>6</sub> <sup>15</sup>N<sub>2</sub> lysine (Lys8) and <sup>13</sup>C<sub>6</sub> <sup>15</sup>N<sub>4</sub> arginine (Arg10).

“SILAC experiment 2” used only TAP-SUMO-2 HeLa cells at 37°C (Lys0, Arg0), 30 minutes HS at 43°C (Lys4, Arg6) and HS followed by 2 hours of recovery in 37°C (Lys8, Arg10). Cells were harvested at 90-100% confluency.

#### *Tandem affinity purification (TAP) of SUMO-2 conjugates*

Cells were harvested under denaturing conditions (50mM Tris/HCl pH 8.0, 2% SDS, 10mM iodoacetamide, 1mM EDTA with complete protease inhibitor cocktail (Roche)) and mixed in equal protein ratios. Cleared lysates were refolded by dilution into 25 volumes of cold renaturation buffer (RB) (50 mM Tris-HCl pH 8.0, 0.75 M NaCl, 1% NP-40, 2mM iodoacetamide, 0.5 mM EDTA), then purified in two stages. Firstly by IgG-sepharose affinity chromatography, followed by elution with TEV protease. In the second

step, TEV eluates were purified by calmodulin sepharose affinity chromatography. Eluted proteins were TCA precipitated and resuspended to 30 $\mu$ L volume before analysis by Coomassie-stained SDS-PAGE (Invitrogen, NuPAGE 10% Bis-Tris). The lane containing SUMO-2 purified proteins was then sliced into 6 sections and followed by in-gel tryptic digestion. See 'Supplementary methods' for a detailed description.

#### *Quantitative mass spectrometry*

Peptide mixtures were analyzed by LC-MS/MS with a LTQ-Orbitrap mass spectrometer equipped with a nanoelectrospray ion source (Proxeon Biosystems) and coupled to an Agilent 1100 nanoflow system (Agilent Technologies) fitted with an in-house made 75  $\mu$ m reverse phase C18 column. The instrument was operated with the "lock mass" option and in the data-dependent mode to automatically switch between MS and MS/MS. Raw files were processed with the in house developed quantitative processing software MaxQuant (version 1.0.7.1) (Cox & Mann, 2008b) and using the Mascot search engine (Matrix Science). The data were searched against a target/decoy human IPI database (version 3.24) (Kersey et al, 2004). 1% false discovery rate was required at both the protein and peptide level. The SUMO-2 branched peptides were manually found in the unmatched list, containing quantified but non-identified peptides (Matic et al, 2008b).

#### **Acknowledgements**

F.G. is funded through a Marie Curie research fellowship and M.H.T is funded by CRUK. I.M. is supported by the EU FP6 Rubicon consortium and an EMBO short-term fellowship. C.C is funded by the Scottish Funding Council via the Scottish

Bioinformatics Research Network (SBRN)). We would like to acknowledge help from Catherine Botting, Robin Antrobus, David Martin and Michelle Scott.

## Tables

<b>Dataset<sup>+</sup></b>	<b>Number of proteins</b>	<b>Number of predicted consensus sites</b>	<b>Number of proteins without consensus</b>	<b>Average consensus sites per protein</b>
Published SUMO substrates	264	517	49	2.0
TAP-SUMO-2 proteome	759	1681	195	2.2
Human proteome <sup>§</sup>	43,964	28,078	28,571	0.6
TAP-SUMO-2 internal contaminants	594	458	312	0.8

### **Table 1. Predicted SUMO consensus site frequency of proteins identified by**

**different studies.** SUMO consensus sites of the form  $\psi$ KxE as predicted by

SUMOsp2.0(Xue et al, 2006) using ‘High’ threshold. §Using a non-redundant human proteome data set from UniRef90. Datasets were filtered for redundancy. The shorter of two sequences with >90% pairwise sequence identity over >90% of the sequence length were removed.

## References

1. Lindquist, S. The heat-shock response. *Annu Rev Biochem* **55**, 1151-1191 (1986).
2. Saitoh, H. & Hinchev, J. Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *The Journal of biological chemistry* **275**, 6252-6258 (2000).
3. Mann, M. Functional and quantitative proteomics using SILAC. *Nature reviews* **7**, 952-958 (2006).
4. Witze, E.S., Old, W.M., Resing, K.A. & Ahn, N.G. Mapping protein post-translational modifications with mass spectrometry. *Nature methods* **4**, 798-806 (2007).
5. Matic, I. *et al.* *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 (2008).
6. Tatham, M.H. *et al.* RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nature cell biology* **10**, 538-546 (2008).
7. Tatham, M.H. *et al.* Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *The Journal of biological chemistry* **276**, 35368-35374 (2001).
8. Rigaut, G. *et al.* A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**, 1030-1032 (1999).

9. Hietakangas, V. *et al.* Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. *Molecular and cellular biology* **23**, 2953-2968 (2003).
10. Rodriguez, M.S., Dargemont, C. & Hay, R.T. SUMO-1 conjugation *in vivo* requires both a consensus modification motif and nuclear targeting. *The Journal of biological chemistry* **276**, 12654-12659 (2001).
11. Pearl, L.H., Prodromou, C. & Workman, P. The Hsp90 molecular chaperone: an open and shut case for treatment. *The Biochemical journal* **410**, 439-453 (2008).
12. Dorion, S. & Landry, J. Activation of the mitogen-activated protein kinase pathways by heat shock. *Cell Stress Chaperones* **7**, 200-206 (2002).
13. Skaggs, H.S. *et al.* HSF1-TPR interaction facilitates export of stress-induced HSP70 mRNA. *The Journal of biological chemistry* **282**, 33902-33907 (2007).
14. Beere, H.M. Death versus survival: functional interaction between the apoptotic and stress-inducible heat shock protein pathways. *J Clin Invest* **115**, 2633-2639 (2005).
15. Bond, U. Stressed out! Effects of environmental stress on mRNA metabolism. *FEMS yeast research* **6**, 160-170 (2006).
16. Kaneko, H., Igarashi, K., Kataoka, K. & Miura, M. Heat shock induces phosphorylation of histone H2AX in mammalian cells. *Biochemical and biophysical research communications* **328**, 1101-1106 (2005).
17. Takahashi, A. *et al.* Evidence for the involvement of double-strand breaks in heat-induced cell killing. *Cancer research* **64**, 8839-8845 (2004).

18. Matsuoka, S. *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160-1166 (2007).
19. Dai, T. *et al.* Stress-activated protein kinases bind directly to the delta domain of c-Jun in resting cells: implications for repression of c-Jun function. *Oncogene* **10**, 849-855 (1995).
20. Muller, S. *et al.* c-Jun and p53 Activity Is Modulated by SUMO-1 Modification. *The Journal of biological chemistry* **275**, 13321-13329 (2000).
21. Shivaswamy, S. & Iyer, V.R. Stress-dependent dynamics of global chromatin remodeling in yeast: dual role for SWI/SNF in the heat shock stress response. *Molecular and cellular biology* **28**, 2221-2234 (2008).
22. Ivanov, A.V. *et al.* PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Molecular cell* **28**, 823-837 (2007).
23. Stielow, B., Sapetschnig, A., Wink, C., Kruger, I. & Suske, G. SUMO-modified Sp3 represses transcription by provoking local heterochromatic gene silencing. *EMBO reports* (2008).
24. Nathan, D. *et al.* Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications. *Genes & development* **20**, 966-976 (2006).
25. Shio, Y. & Eisenman, R.N. Histone sumoylation is associated with transcriptional repression. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 13225-13230 (2003).

26. Ooi, S.K. & Bestor, T.H. The colorful history of active DNA demethylation. *Cell* **133**, 1145-1148 (2008).
27. VanderWaal, R.P., Griffith, C.L., Wright, W.D., Borrelli, M.J. & Roti, J.L. Delaying S-phase progression rescues cells from heat-induced S-phase hypertoxicity. *J Cell Physiol* **187**, 236-243 (2001).
28. Ong, S.E. & Mann, M. Stable isotope labeling by amino acids in cell culture for quantitative proteomics. *Methods in molecular biology (Clifton, N.J)* **359**, 37-52 (2007).
29. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* (2008).
30. Kersey, P.J. *et al.* The International Protein Index: an integrated database for proteomics experiments. *Proteomics* **4**, 1985-1988 (2004).
31. Xue, Y., Zhou, F., Fu, C., Xu, Y. & Yao, X. SUMOsp: a web server for sumoylation site prediction. *Nucleic acids research* **34**, W254-257 (2006).
32. Thomas, P.D. *et al.* PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic acids research* **31**, 334-341 (2003).

## Figure Legends

**Figure 1.** *Identification of 660 putative SUMO-2 target proteins from HeLa cells.* (a) Schematic representation of the TAP-SUMO-2 construct stably transfected into HeLa cells for this study. The Tandem Affinity Protein (TAP) tag consists of a Protein A domain separated from a calmodulin binding protein (CBP) domain by the tobacco etch virus (TEV) protease site. This is N-terminally tagged to SUMO-2 (residues 1-92) (NCBI Entrez protein CAG46970), which can be conjugated directly to target proteins via a covalent bond. (b) Anti-SUMO-2 western blot of crude cell lysates from TAP-SUMO-2 HeLa cells under normal conditions (37°C) and after heat-shock for 30 minutes (43°C). (c) Silver-stained SDS PAGE gel showing a TAP purification products from HeLa cells expressing TAP alone and TAP-SUMO-2. (d) Overview of the quantitative proteomic experiment comparing TAP alone HeLas, with TAP-SUMO-2 HeLas either untreated (37°C) or heat-shocked (43°C). TAP only-expressing cells were grown in isotopically normal ('light') medium, TAP-SUMO-2 cells grown under normal conditions were cultured in SILAC medium containing Lys4 and Arg6 ('medium') isotopic forms, and heat-shocked TAP-SUMO-2 cells were grown in Lys8 and Arg10 ('heavy') SILAC medium. After denaturing lysis the pooled lysates were subjected to TAP purification, separation by SDS-PAGE, in-gel digestion with trypsin, analysis by high resolution LC-MS/MS and data processing using MaxQuant software(Cox & Mann, 2008b) (see Materials and Methods for further details). Examples are shown of raw MS data showing spectra from representative peptides for the internal contaminant glyceraldehyde phosphate dehydrogenase (GAPDH), and the genuine target, heat shock transcription factor 1 (HSF1). (e) Triple-label SILAC quantitation map with marginal histograms of

1164 quantified proteins from this experiment. This was filtered to remove putative contaminants to yield 660 putative SUMO-2 substrates (See Supplementary Methods and Supp. Fig. 1).

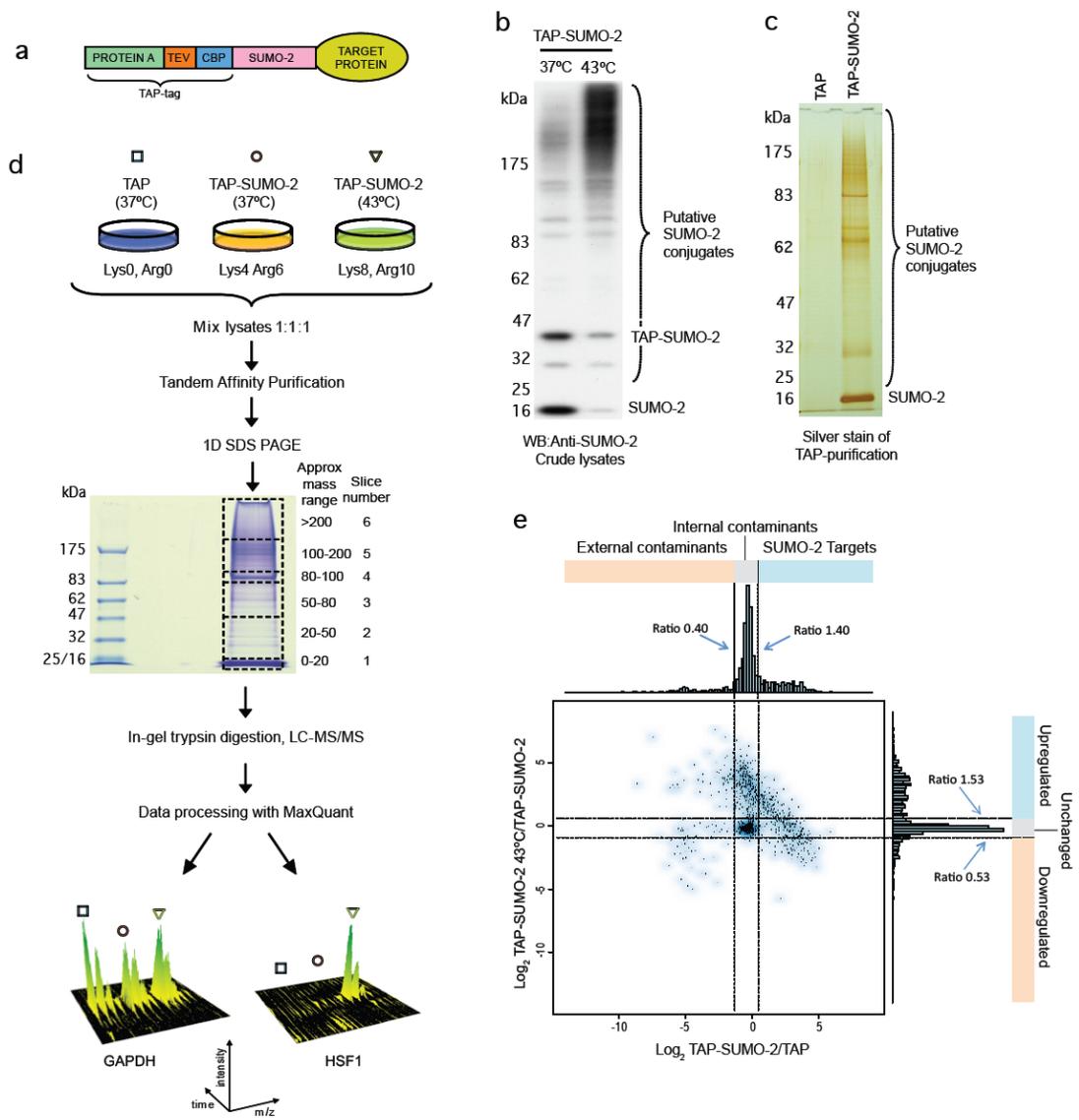
**Figure 2.** *The effect of recovery from heat-stress on cellular SUMO-2 conjugates.* (a) Anti-SUMO-2 Western-blot of individual lysates of TAP-SUMO-2 HeLas which were untreated (37°C), Heat-shocked at 43°C for 30 minutes (HS) or allowed to recover for 2 hours at 37°C after heat-shock, (HSR) using SILAC labels as indicated. (b) Raw MS data showing spectra from representative peptides for scaffold attachment factor B1 (SAF-B1) and the glucocorticoid receptor. (c) Smoothed colour density scatter plot with marginal histograms of the quantitation data of 1255 identified proteins. (d) Comparison of the overlap among the proteins identified in this study (TAP-SUMO-2 proteome), previously published SUMO targets ('Confirmed' SUMO substrates) and proteins identified by previous proteomics studies (Published SUMO proteomes).

**Figure 3.** *Stimulation of SUMO-2 polymerisation by heat stress* (a) Left charts; relative comparison of the frequency of  $\log_2(\text{TAP-SUMO-2 Recovery}/\text{TAP-SUMO-2})$  (green lines) against  $\log_2(\text{TAP-SUMO-2 heatshock}/\text{TAP-SUMO-2})$  (red lines) for each of the 6 gel slices from SILAC experiment 2. Note the progressive tendency to higher values for the red traces compared with the green as the abundances of proteins of high molecular weight accumulate with HS. Right histograms; Frequencies of protein molecular weights found in each of the six slices. Predicted molecular weight range of each slice is shown by red boundaries (b) Heatmap analysis of 40 SUMO-2 targets comparing for each protein the contribution to the overall ratio data of an individual slice. Black indicates

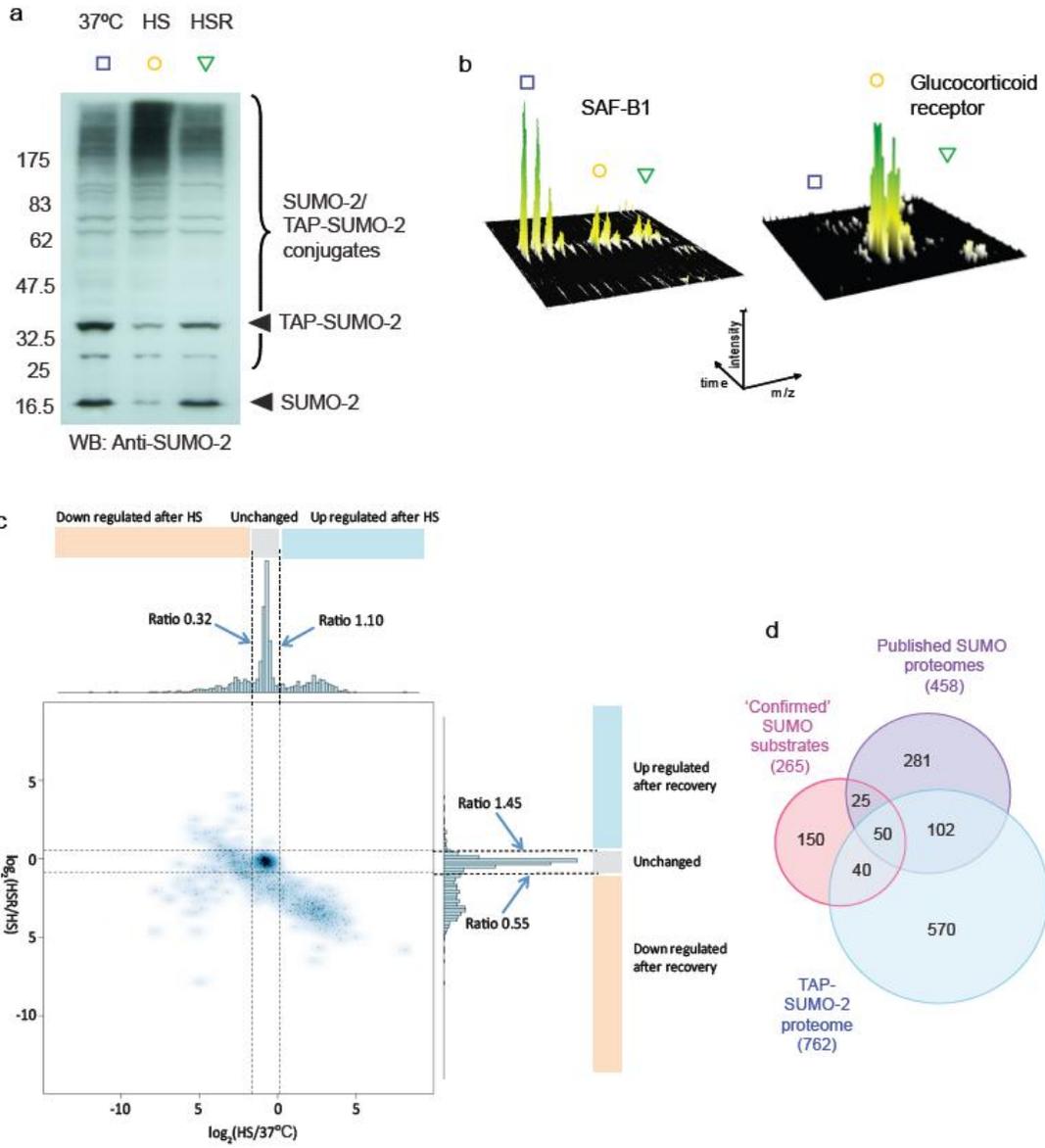
that the protein ratio is equal to the combined ratio, red that it is higher and green, lower. Data produced by Kernel density estimation. (c) 3D plot of the MS peak corresponding to the peptide fragment of the SUMO-2-SUMO-2 linkage via lysine 11 from SILAC experiment 2.

**Figure 4.** *Co-regulation of protein functions by SUMO-2 after heat-stress.* (a) PANTHER(Thomas et al, 2003) ‘Biological process’ ontology comparison between the relative abundance of TAP-SUMO-2 targets compared to the human genome. Categories shown where p values of the significance of the difference compared to the human genome (shown in brackets) was less than 0.05. Groups are arranged from most significantly over-represented (left) to most significantly under-represented (right). (b-g) GOsMaps showing the relationship between the TAP-SUMO-2 HS/TAP-SUMO-2 and TAP-SUMO-2 HSR/TAP-SUMO-2 HS ratios for the indicated functions (red), compared with the entire list of identified SUMO-2 targets (grey), from SILAC experiment 2 (Figure 2). Red numbers indicate the number of members of each category and black numbers show the significance of clustering in the x-axis and y-axis. See Supp. Table 2 for lists of proteins under each GO term. (h-n) Examples of protein modules identified in this study as being regulated by SUMO-2 and heat-stress. Red and green nodes are either more or less conjugated to SUMO-2 after HS. Labels are gene names, node shapes indicate protein function; trapezium – transporter, diamond – enzyme, point down triangle – kinase, point up triangle – phosphatase, ellipse – transcriptional regulator, circle – other function. Previously identified SUMO substrates are highlighted in bold.

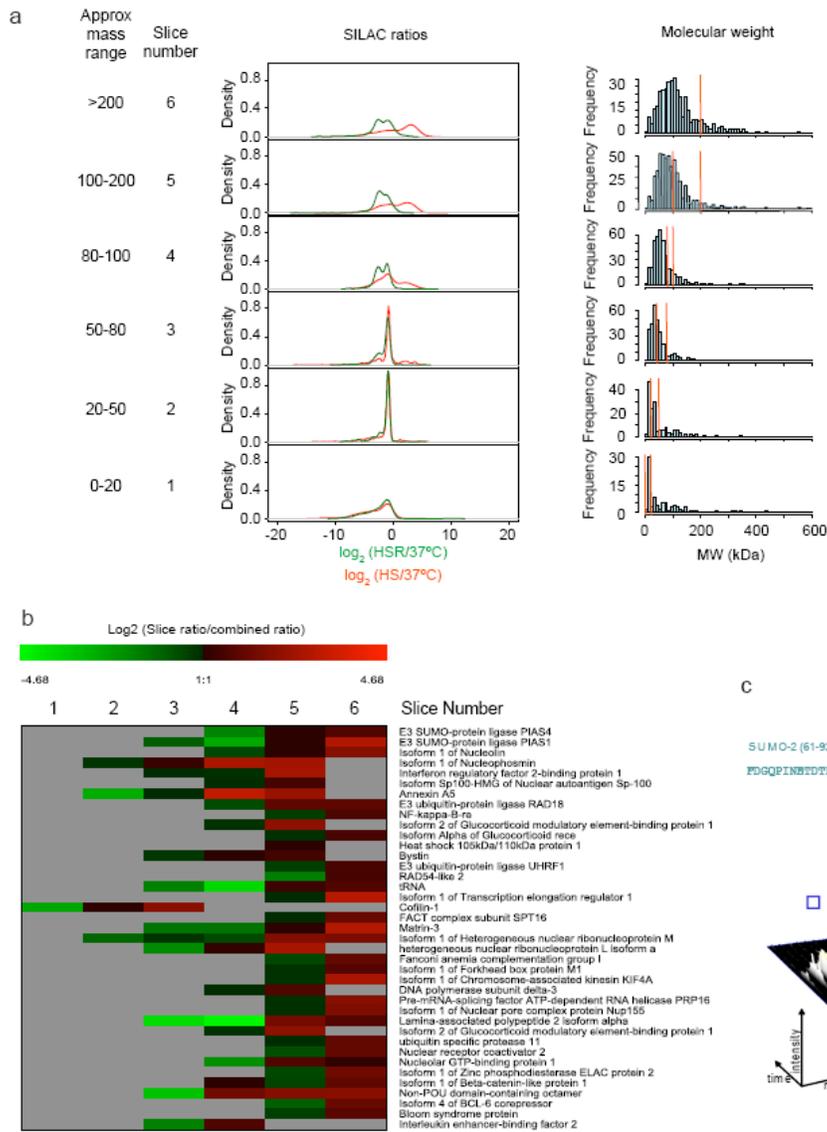
Solid lines indicate direct interactions and broken lines, indirect. Arrow from A to B shows that A acts on B.



Hay Figure 1



Hay Figure 2



Hay Figure 3



## 7 Conclusion and perspectives

Virtually all the processes that control the fate of cells and organisms are regulated by sumoylation and many SUMO modified proteins are involved in disease states. SUMO functions include modulation of gene expression, DNA replication and repair, cancer development and cell cycle regulation. SUMO conjugation of specific proteins in signaling pathways has been investigated mainly by classical biological techniques and approaches. Despite their obvious success in elucidating mechanisms of sumoylation and showing the fundamental role of SUMO in regulating functionally important proteins, there are still many areas, such as novel substrates and PTMs of SUMO itself, remain to be explored. Mass spectrometry is ideally suited to the hypothesis-free analysis of PTMs and to reveal non-anticipated aspects of SUMO dynamics.

The projects of this thesis apply the most recent advances in mass spectrometry-based proteomics to qualitative and quantitative investigation of the SUMO system. The *in vivo* characterization of the primary structure of SUMO revealed its phosphorylation, resulting in the first report of such modification on a protein modifier (Matic et al, 2008a), its ubiquitination (Schimmel et al, 2008) and the existence of mixed SUMO polymers (Matic et al, 2008b). Future SILAC-based quantitative proteomics studies will give additional insights into the role of these PTMs by profiling their relative abundances in different cellular conditions.

Approaches applied to our in-depth analysis of changes of sumoylated proteins will be a template for future time-resolved quantitative studies of SUMO conjugates upon various stimuli, which will provide a wealth of hypotheses to follow and a starting point for

systems biology modeling. This study represents the largest and high-quality SUMO substrates proteome and reveals a larger role for SUMO in signaling pathway than previously considered. It will constitute a valuable resource for everyone interested in the emerging SUMO field.

A challenge in the analysis of substoichiometric post-translational modifications, such as SUMO, is their enrichment from a complex biological sample. Whereas the enormous success of phosphoproteomics is due to the relative ease of purifying phosphorylated peptides, similarly powerful enrichment strategies are not currently available for peptides modified by ubiquitin, SUMO or other UbIs. In case of sumoylated peptides, the long SUMO part of the branched peptides could be used as an epitope in immunoprecipitation with antibodies recognizing the C-terminal peptide of SUMO. Even partial separation of sumoylated peptides from other tryptic peptides would allow the sequencing of a large number of endogenous SUMO branched peptides, previously lost in the ‘sea’ of background peptides.

I have successfully detected mixed SUMO polymerization sites *in vivo* by manual screening of spectra (Matic et al, 2008b) and applied the methods developed in that study to quantify SUMO polymerization after treatment with proteasome inhibitor (Schimmel et al, 2008). However this laborious approach is not applicable to the computational analysis of the large amounts of data generated by modern MS instrumentation. A direct implementation of a SUMO fragment ion recognition tool into MaxQuant, an increasingly popular quantitative proteomics platform, in combination with the above described method to enrich SUMO modified peptides will greatly expand the applicability of MS-based proteomics to the SUMO field. It may also make SUMO one of the central foci of MS-based proteomics, similarly to phosphoproteomics today.

## 8 Bibliography

Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. *Nature* **422**: 198-207

Anckar J, Sistonen L (2007) SUMO: getting it on. *Biochemical Society transactions* **35**: 1409-1413

Andersson L, Porath J (1986) Isolation of phosphoproteins by immobilized metal (Fe<sup>3+</sup>) affinity chromatography. *Analytical biochemistry* **154**: 250-254

Baek SH (2006) A novel link between SUMO modification and cancer metastasis. *Cell cycle (Georgetown, Tex)* **5**: 1492-1495

Bayer P, Arndt A, Metzger S, Mahajan R, Melchior F, Jaenicke R, Becker J (1998) Structure determination of the small ubiquitin-related modifier SUMO-1. *J Mol Biol* **280**: 275-286

Beere HM (2005) Death versus survival: functional interaction between the apoptotic and stress-inducible heat shock protein pathways. *J Clin Invest* **115**: 2633-2639

Bernier-Villamor V, Sampson DA, Matunis MJ, Lima CD (2002) Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* **108**: 345-356

Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, Kannouche P, Wider G, Peter M, Lehmann AR, Hofmann K, Dikic I (2005) Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science (New York, NY)* **310**: 1821-1824

Bond U (2006) Stressed out! Effects of environmental stress on mRNA metabolism. *FEMS yeast research* **6**: 160-170

- Catic A, Ploegh HL (2005) Ubiquitin--conserved protein or selfish gene? *Trends in biochemical sciences* **30**: 600-604
- Colinge J, Bennett KL (2007) Introduction to computational proteomics. *PLoS computational biology* **3**: e114
- Coon JJ, Ueberheide B, Syka JE, Dryhurst DD, Ausio J, Shabanowitz J, Hunt DF (2005) Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 9463-9468
- Cooper HJ, Tatham MH, Jaffray E, Heath JK, Lam TT, Marshall AG, Hay RT (2005) Fourier transform ion cyclotron resonance mass spectrometry for the analysis of small ubiquitin-like modifier (SUMO) modification: identification of lysines in RanBP2 and SUMO targeted for modification during the E3 autoSUMOylation reaction. *Analytical chemistry* **77**: 6310-6319
- Cox J, Mann M (2007) Is proteomics the new genomics? *Cell* **130**: 395-398
- Cox J, Mann M (2008a) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* **26**: 1367-1372
- Cox J, Mann M (2008b) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*
- Dai T, Rubie E, Franklin CC, Kraft A, Gillespie DA, Avruch J, Kyriakis JM, Woodgett JR (1995) Stress-activated protein kinases bind directly to the delta domain of c-Jun in resting cells: implications for repression of c-Jun function. *Oncogene* **10**: 849-855
- Dasso M (2008) Emerging roles of the SUMO pathway in mitosis. *Cell Division* **3**: 5
- Desterro JM, Rodriguez MS, Hay RT (1998) SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. *Mol Cell* **2**: 233-239
- Dorion S, Landry J (2002) Activation of the mitogen-activated protein kinase pathways by heat shock. *Cell Stress Chaperones* **7**: 200-206
- Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature methods* **4**: 207-214
- Farrell PJ, Broeze RJ, Lengyel P (1979) Accumulation of an mRNA and protein in interferon-treated Ehrlich ascites tumour cells. *Nature* **279**: 523-525

Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science (New York, NY)* **246**: 64-71

Geiss-Friedlander R, Melchior F (2007) Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol*

Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 6940-6945

Gill G (2005) Something about SUMO inhibits transcription. *Current Opinion in Genetics & Development* **15**: 536-541

Girdwood DW, Tatham MH, Hay RT (2004) SUMO and transcriptional regulation. *Seminars in cell & developmental biology* **15**: 201-210

Glickman MH, Ciechanover A (2002) The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiological reviews* **82**: 373-428

Haglund K, Dikic I (2005) Ubiquitylation and cell signaling. *The EMBO journal* **24**: 3353-3359

Hay RT (2005) SUMO: a history of modification. *Mol Cell* **18**: 1-12

Hay RT (2007) SUMO-specific proteases: a twist in the tail. *Trends in cell biology* **17**: 370-376

Hayakawa F, Privalsky ML (2004) Phosphorylation of PML by mitogen-activated protein kinases plays a key role in arsenic trioxide-mediated apoptosis. *Cancer cell* **5**: 389-401

Hecker CM, Rabiller M, Haglund K, Bayer P, Dikic I (2006) Specification of SUMO1- and SUMO2-interacting motifs. *J Biol Chem* **281**: 16117-16127

Hershko A, Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* **67**: 425-479

Hietakangas V, Ahlskog JK, Jakobsson AM, Hellesuo M, Sahlberg NM, Holmberg CI, Mikhailov A, Palvimo JJ, Pirkkala L, Sistonen L (2003) Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. *Molecular and cellular biology* **23**: 2953-2968

Hietakangas V, Anckar J, Blomster HA, Fujimoto M, Palvimo JJ, Nakai A, Sistonen L (2006) PDSM, a motif for phosphorylation-dependent SUMO modification. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 45-50

- Hochstrasser M (2000) Evolution and function of ubiquitin-like protein-conjugation systems. *Nat Cell Biol* **2**: E153-E157
- Hochstrasser M (2001) SP-RING for SUMO: new functions bloom for a ubiquitin-like protein. *Cell* **107**: 5-8
- Hoegge C, Pfander B, Moldovan GL, Pyrowolakis G, Jentsch S (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **419**: 135-141
- Hunter T (2007) The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol Cell* **28**: 730-738
- Ikeda F, Dikic I (2008) Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series. *EMBO reports* **9**: 536-542
- Ishov AM, Sotnikov AG, Negorev D, Vladimirova OV, Neff N, Kamitani T, Yeh ET, Strauss JF, 3rd, Maul GG (1999) PML is critical for ND10 formation and recruits the PML-interacting protein daxx to this nuclear structure when modified by SUMO-1. *The Journal of cell biology* **147**: 221-234
- Ivanov AV, Peng H, Yurchenko V, Yap KL, Negorev DG, Schultz DC, Psulkowski E, Fredericks WJ, White DE, Maul GG, Sadofsky MJ, Zhou MM, Rauscher FJ, 3rd (2007) PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Molecular cell* **28**: 823-837
- Johnson ES (2004) Protein modification by SUMO. *Annu Rev Biochem* **73**: 355-382
- Johnson ES, Gupta AA (2001) An E3-like factor that promotes SUMO conjugation to the yeast septins. *Cell* **106**: 735-744
- Johnson LN, Barford D (1993) The effects of phosphorylation on the structure and function of proteins. *Annual review of biophysics and biomolecular structure* **22**: 199-232
- Kahyo T, Nishida T, Yasuda H (2001) Involvement of PIAS1 in the sumoylation of tumor suppressor p53. *Mol Cell* **8**: 713-718
- Kaneko H, Igarashi K, Kataoka K, Miura M (2005) Heat shock induces phosphorylation of histone H2AX in mammalian cells. *Biochemical and biophysical research communications* **328**: 1101-1106

- Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry* **60**: 2299-2301
- Karin M, Ben-Neriah Y (2000) Phosphorylation meets ubiquitination: the control of NF- $\kappa$ B activity. *Annual review of immunology* **18**: 621-663
- Kerscher O (2007) SUMO junction-what's your function? New insights through SUMO-interacting motifs. *EMBO reports* **8**: 550-555
- Kerscher O, Felberbaum R, Hochstrasser M (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annual review of cell and developmental biology* **22**: 159-180
- Kersey PJ, Duarte J, Williams A, Karavidopoulou Y, Birney E, Apweiler R (2004) The International Protein Index: an integrated database for proteomics experiments. *Proteomics* **4**: 1985-1988
- Kirkin V, Dikic I (2007) Role of ubiquitin- and Ubl-binding proteins in cell signaling. *Current opinion in cell biology* **19**: 199-205
- Kirkpatrick DS, Denison C, Gygi SP (2005) Weighing in on ubiquitin: the expanding role of mass-spectrometry-based proteomics. *Nature cell biology* **7**: 750-757
- Krappmann D, Scheidereit C (2005) A pervasive role of ubiquitin conjugation in activation and termination of IkappaB kinase pathways. *EMBO reports* **6**: 321-326
- Kurepa J, Walker JM, Smalle J, Gosink MM, Davis SJ, Durham TL, Sung DY, Vierstra RD (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
- Lallemand-Breitenbach V, Jeanne M, Benhenda S, Nasr R, Lei M, Peres L, Zhou J, Zhu J, Raught B, de The H (2008) Arsenic degrades PML or PML-RARalpha through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat Cell Biol* **10**: 547-555
- Lallemand-Breitenbach V, Zhu J, Puvion F, Koken M, Honore N, Doubeikovsky A, Duprez E, Pandolfi PP, Puvion E, Freemont P, de The H (2001) Role of promyelocytic leukemia (PML) sumolation in nuclear body formation, 11S proteasome recruitment, and As2O3-induced PML or PML/retinoic acid receptor alpha degradation. *The Journal of experimental medicine* **193**: 1361-1371
- Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jorgensen TJ (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* **4**: 873-886

- Lieberman AP (2004) SUMO, a ubiquitin-like modifier implicated in neurodegeneration. *Experimental neurology* **185**: 204-207
- Lin D, Tatham MH, Yu B, Kim S, Hay RT, Chen Y (2002) Identification of a substrate recognition site on Ubc9. *J Biol Chem* **277**: 21740-21748
- Lindquist S (1986) The heat-shock response. *Annu Rev Biochem* **55**: 1151-1191
- Mahajan R, Delphin C, Guan T, Gerace L, Melchior F (1997) A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **88**: 97-107
- Makarov A (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Analytical chemistry* **72**: 1156-1162
- Mann M (2006) Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol* **7**: 952-958
- Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. *Science (New York, NY)* **298**: 1912-1934
- Matic I, Macek B, Hilger M, Walther TC, Mann M (2008a) Phosphorylation of SUMO-1 occurs in vivo and is conserved through evolution. *Journal of proteome research* **7**: 4050-4057
- Matic I, van Hagen M, Schimmel J, Macek B, Ogg SC, Tatham MH, Hay RT, Lamond AI, Mann M, Vertegaal AC (2008b) 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
- Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER, 3rd, Hurov KE, Luo J, Bakalarski CE, Zhao Z, Solimini N, Lerenthal Y, Shiloh Y, Gygi SP, Elledge SJ (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**: 1160-1166
- Matunis MJ, Coutavas E, 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. *The Journal of cell biology* **135**: 1457-1470
- Meulmeester E, Melchior F (2008) Cell biology: SUMO. *Nature* **452**: 709-711
- Moldovan GL, Pfander B, Jentsch S (2007) PCNA, the maestro of the replication fork. *Cell* **129**: 665-679

- Muller S, Berger M, Lehembre F, Seeler JS, Haupt Y, Dejean A (2000) c-Jun and p53 Activity Is Modulated by SUMO-1 Modification. *The Journal of biological chemistry* **275**: 13321-13329
- Nathan D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanovic M, Dorsey JA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, Johnson ES, Berger SL (2006) Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications. *Genes & development* **20**: 966-976
- Nesvizhskii AI, Vitek O, Aebersold R (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nature methods* **4**: 787-797
- Nielsen ML, Vermeulen M, Bonaldi T, Cox J, Moroder L, Mann M (2008) Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nature methods* **5**: 459-460
- Olsen JV, Blagoev B, Gnadt F, Macek B, Kumar C, Mortensen P, Mann M (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**: 635-648
- Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, Pesch R, Makarov A, Lange O, Horning S, Mann M (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Molecular & Cellular Proteomics* **4**: 2010-2021
- Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M (2007) Higher-energy C-trap dissociation for peptide modification analysis. *Nature methods* **4**: 709-712
- Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (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
- Ong SE, Mann M (2005) Mass spectrometry-based proteomics turns quantitative. *Nature chemical biology* **1**: 252-262
- Ong SE, Mann M (2007) Stable isotope labeling by amino acids in cell culture for quantitative proteomics. *Methods in molecular biology (Clifton, NJ)* **359**: 37-52
- Ooi SK, Bestor TH (2008) The colorful history of active DNA demethylation. *Cell* **133**: 1145-1148
- Palvimo JJ (2007) PIAS proteins as regulators of small ubiquitin-related modifier (SUMO) modifications and transcription. *Biochemical Society transactions* **35**: 1405-1408

- Pearce MJ, Mintseris J, Ferreyra J, Gygi SP, Darwin KH (2008) Ubiquitin-like protein involved in the proteasome pathway of Mycobacterium tuberculosis. *Science (New York, NY)* **322**: 1104-1107
- Pearl LH, Prodromou C, Workman P (2008) The Hsp90 molecular chaperone: an open and shut case for treatment. *The Biochemical journal* **410**: 439-453
- Pedrioli PG, Raught B, Zhang XD, Rogers R, Aitchison J, Matunis M, Aebersold R (2006) Automated identification of SUMOylation sites using mass spectrometry and SUMMOn pattern recognition software. *Nature methods* **3**: 533-539
- Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP (2003) A proteomics approach to understanding protein ubiquitination. *Nature biotechnology* **21**: 921-926
- Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**: 3551-3567
- Perkins ND (2007) Integrating cell-signalling pathways with NF-kappaB and IKK function. *Nat Rev Mol Cell Biol* **8**: 49-62
- Pfander B, Moldovan GL, Sacher M, Hoegge C, Jentsch S (2005) SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. *Nature* **436**: 428-433
- Pichler A, Knipscheer P, Saitoh H, Sixma TK, Melchior F (2004) The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type. *Nature structural & molecular biology* **11**: 984-991
- Pinkse MW, Uitto PM, Hilhorst MJ, Ooms B, Heck AJ (2004) Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Analytical chemistry* **76**: 3935-3943
- Reverter D, Lima CD (2005) Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature* **435**: 687-692
- Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**: 1030-1032
- Rodriguez MS, Dargemont C, Hay RT (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J Biol Chem* **276**: 12654-12659

- Ross S, Best JL, Zon LI, Gill G (2002) SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization. *Mol Cell* **10**: 831-842
- Sachdev S, Bruhn L, Sieber H, Pichler A, Melchior F, Grosschedl R (2001) PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. *Genes & development* **15**: 3088-3103
- Saitoh H, Hinchey J (2000) Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *The Journal of biological chemistry* **275**: 6252-6258
- Salomoni P, Pandolfi PP (2002) The role of PML in tumor suppression. *Cell* **108**: 165-170
- Schimmel J, Larsen KM, Matic I, van Hagen M, Cox J, Mann M, Andersen JS, Vertegaal AC (2008) The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle. *Mol Cell Proteomics* **7**: 2107-2122
- Scigelova M, Makarov A (2006) Orbitrap mass analyzer--overview and applications in proteomics. *Proteomics* **6 Suppl 2**: 16-21
- Shalizi A, Gaudilliere B, Yuan Z, Stegmuller J, Shirogane T, Ge Q, Tan Y, Schulman B, Harper JW, Bonni A (2006) A calcium-regulated MEF2 sumoylation switch controls postsynaptic differentiation. *Science (New York, NY)* **311**: 1012-1017
- Shiio Y, Eisenman RN (2003) Histone sumoylation is associated with transcriptional repression. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 13225-13230
- Shivaswamy S, Iyer VR (2008) Stress-dependent dynamics of global chromatin remodeling in yeast: dual role for SWI/SNF in the heat shock stress response. *Molecular and cellular biology* **28**: 2221-2234
- Skaggs HS, Xing H, Wilkerson DC, Murphy LA, Hong Y, Mayhew CN, Sarge KD (2007) HSF1-TPR interaction facilitates export of stress-induced HSP70 mRNA. *The Journal of biological chemistry* **282**: 33902-33907
- Song J, Durrin LK, Wilkinson TA, Krontiris TG, Chen Y (2004) Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 14373-14378
- Steen H, Mann M (2004) The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* **5**: 699-711

Stielow B, Sapetschnig A, Wink C, Kruger I, Suske G (2008) SUMO-modified Sp3 represses transcription by provoking local heterochromatic gene silencing. *EMBO reports*

Takahashi A, Matsumoto H, Nagayama K, Kitano M, Hirose S, Tanaka H, Mori E, Yamakawa N, Yasumoto J, Yuki K, Ohnishi K, Ohnishi T (2004) Evidence for the involvement of double-strand breaks in heat-induced cell killing. *Cancer research* **64**: 8839-8845

Takahashi Y, Kahyo T, Toh EA, Yasuda H, Kikuchi Y (2001) Yeast Ull1/Siz1 is a novel SUMO1/Smt3 ligase for septin components and functions as an adaptor between conjugating enzyme and substrates. *J Biol Chem* **276**: 48973-48977

Tatham MH, Geoffroy MC, Shen L, Plechanovova A, Hattersley N, Jaffray EG, Palvimo JJ, Hay RT (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nature cell biology* **10**: 538-546

Tatham MH, Jaffray E, Vaughan OA, Desterro JM, Botting CH, Naismith JH, Hay RT (2001) Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *The Journal of biological chemistry* **276**: 35368-35374

Tempe D, Piechaczyk M, Bossis G (2008) SUMO under stress. *Biochemical Society transactions* **36**: 874-878

Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, Vandergriff JA, Doremieux O (2003) PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic acids research* **31**: 334-341

Ulrich HD (2005) Mutual interactions between the SUMO and ubiquitin systems: a plea of no contest. *Trends in cell biology* **15**: 525-532

VanderWaal RP, Griffith CL, Wright WD, Borrelli MJ, Roti JL (2001) Delaying S-phase progression rescues cells from heat-induced S-phase hypertoxicity. *J Cell Physiol* **187**: 236-243

Varadan R, Walker O, Pickart C, Fushman D (2002) Structural properties of polyubiquitin chains in solution. *J Mol Biol* **324**: 637-647

Voges D, Zwickl P, Baumeister W (1999) The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu Rev Biochem* **68**: 1015-1068

Welchman RL, Gordon C, Mayer RJ (2005) Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat Rev Mol Cell Biol* **6**: 599-609

Witze ES, Old WM, Resing KA, Ahn NG (2007) Mapping protein post-translational modifications with mass spectrometry. *Nature methods* **4**: 798-806

Wotton D, Merrill JC (2007) Pc2 and SUMOylation. *Biochemical Society transactions* **35**: 1401-1404

Xirodimas DP, Saville MK, Bourdon JC, Hay RT, Lane DP (2004) Mdm2-mediated NEDD8 conjugation of p53 inhibits its transcriptional activity. *Cell* **118**: 83-97

Xue Y, Zhou F, Fu C, Xu Y, Yao X (2006) SUMOsp: a web server for sumoylation site prediction. *Nucleic acids research* **34**: W254-257

Yang SH, Galanis A, Witty J, Sharrocks AD (2006) An extended consensus motif enhances the specificity of substrate modification by SUMO. *The EMBO journal* **25**: 5083-5093

Zhang XD, Goeres J, Zhang H, Yen TJ, Porter AC, Matunis MJ (2008) SUMO-2/3 modification and binding regulate the association of CENP-E with kinetochores and progression through mitosis. *Mol Cell* **29**: 729-741

Zhong S, Muller S, Ronchetti S, Freemont PS, Dejean A, Pandolfi PP (2000) Role of SUMO-1-modified PML in nuclear body formation. *Blood* **95**: 2748-2752

Zubarev R, Mann M (2007) On the proper use of mass accuracy in proteomics. *Mol Cell Proteomics* **6**: 377-381

## 9 Acknowledgments

I deeply thank my PhD supervisor Matthias Mann for giving me the opportunity to pursue my scientific interests in his department. I am particularly grateful for the huge freedom and flexibility to conduct cutting-edge research independently in an area that I am passionate about, and for support I received from the beginning of my PhD study. In addition to the opportunity to learn the most advanced format of mass spectrometry-based proteomics, he encouraged me to participate in many international conferences and meetings and conduct research at the Wellcome Trust Centre for Gene Regulation and Expression in Dundee and at the University of California, San Francisco.

Many thanks to Boris Macek for mentoring me during my starting phase in the group and introducing me to mass spectrometry.

I am also grateful to Alfred Vertegaal and his group, Boris Macek, Juergen Cox and Maximliane Hilger for the enjoyable collaboration.

I wish to thank all members of our department for support, fruitful discussion and great time together.

I am very thankful to Ron Hay and his group for a wonderful and fruitful collaboration and for inviting me to work in his group at the Wellcome Trust Centre for Gene Regulation and Expression, University of Dundee. In particular, I thank Mike Tatham and Dimitris Xirodimas for stimulating discussion, their friendship and nice time during my stay in Dundee. The research in the group of Ron Hay was supported by the EMBO short term fellowship.

Furthermore, I thank Peter Walter for giving me the opportunity to work in his group at the University of California, San Francisco. Special thanks to Paige Nittler for her invaluable help during the experimental work.

Many thanks to Tobias Walther and his group for insightful discussion, collaboration and nice time.

The research projects described in this thesis were supported by the RUBICON network of excellence (EU FP6).

At the end I wish to thank immensely my family and friends for always encouraging me.

# 10 Curriculum vitae

## PERSONAL INFORMATION

Name	Ivan Matic
Date of birth	October 12, 1980
Place of birth	Šibenik, Croatia
Marital status	unmarried
Gender	male
Nationality	Croatian



## EDUCATION

2006 – 2009	PhD study, Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried, Germany
2000 – 2005	Laurea magistrale (master degree: 5-years program) in Pharmaceutical Biotechnology (110/110 cum laude), University of Bologna, Faculty of Pharmacy, Bologna, Italy.

## TRAINING/INTERNSHIPS

07/2003 – 07/2005	Internship at the Department of Anatomical Science, University of Bologna and Research Center “Codivilla Putti” IOR Bologna, Italy. Thesis supervisor: Prof. Lucio Cocco.
09/2004	Practical training at the Biology laboratory (director Dr. Roberto Antolovic), PLIVA Research Institute Ltd, Zagreb, Croatia.
09/2005 – 10/2005	Short-term project at the Molecular Signaling group of Prof. Ivan Dikic, Goethe University, Faculty of Medicine, Frankfurt am Main, Germany.
01/2006 – 03/2006	Short-term project at the Department of Proteomics and Signal Transduction (director Prof. Matthias Mann), Max Planck Institute for Biochemistry, Martinsried, Germany.
07/2008	Short-term project at the group of Prof. Peter Walter, University of California, San Francisco, USA.
10/2008 – 12/2008	Short-term project at the group of Ronald T. Hay, Wellcome Trust Centre for Gene Regulation and Expression, College of Life Sciences, University of Dundee, UK.

## LIST OF PUBLICATIONS

- Filip Golebiowski, **Ivan Matic\***, Michael H. Tatham, Christian Cole, Akihiro Nakamura, Geoffrey J. Barton, Matthias Mann & Ronald T. Hay. System-wide changes in the SUMO-2 proteome in response to heat stress. (*under review*).
- Jürgen Cox, **Ivan Matic**, Maximiliane Hilger, Nagarjuna Nagaraj, Matthias Selbach, Jesper V. Olsen and Matthias Mann. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics (*accepted in Nature Protocols*).
- Schimmel J, Larsen KM, **Matic I\***, van Hagen M, Cox J, Mann M, Andersen JS, Vertegaal AC (2008) The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle. *Mol Cell Proteomics* **7**: 2107-2122
- **Matic I\***, Macek B, Hilger M, Walther TC & Mann M (2008) Phosphorylation of SUMO-1 occurs *in vivo* and is conserved through evolution. *J Proteome Res* **7**, 4050-4057.
- **Matic I\***, van Hagen M, Schimmel J, Macek B, Ogg SC, Tatham MH, Hay RT, Lamond AI, Mann M & Vertegaal AC (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.
- Bavelloni A, Faenza I, Cioffi G, Piazzini M, Parisi D, **Matic I**, Maraldi NM & Cocco L (2006) Proteomic-based analysis of nuclear signaling: PLCbeta1 affects the expression of the splicing factor SRp20 in Friend erythroleukemia cells. *Proteomics* **6**, 5725-5734.

\* first author or first shared author

## BOOK CHAPTER

Andersen JS, **Matic I**, Vertegaal AC (2009) Identification of SUMO target proteins by quantitative proteomics. *Methods in molecular biology (Clifton, NJ)* **497**: 19-31

## **AWARDS/FELLOWSHIPS**

- Three awards (23/04/1998, 18/03/1999, 28/03/2000) granted by “Opera Moreschini” Foundation for excellent marks (high school).
- Three scholarships granted by ARSTUD for academic years 2001/2002, 2003/2004, 2004/2005.
- Fellowship granted by Biotechnology school, University of Bologna, for training in the groups of Prof. Ivan Dikic (September - October 2005) and Prof. Matthias Mann (January - March 2006).
- EMBO short term fellowship for a collaborative project in the group of Prof. Ron Hay (October-December 2008)