Uncovering global SUMOylation signaling networks in a site-specific manner

Ivo A Hendriks¹, Rochelle C J D’Souza², Bing Yang¹, Matty Verlaan-de Vries¹, Matthias Mann² & Alfred C O Vertegaal¹

SUMOylation is a reversible post-translational modification essential for genome stability. Using high-resolution MS, we have studied global SUMOylation in human cells in a site-specific manner, identifying a total of >4,300 SUMOylation sites in >1,600 proteins. To our knowledge, this is the first time that >1,000 SUMOylation sites have been identified under standard growth conditions. We quantitatively studied SUMOylation dynamics in response to SUMO protease inhibition, proteasome inhibition and heat shock. Many SUMOylated lysines have previously been reported to be ubiquitinated, acetylated or methylated, thus indicating cross-talk between SUMO and other post-translational modifications. We identified 70 phosphorylation and four acetylation events in proximity to SUMOylation sites, and we provide evidence for acetylation-dependent SUMOylation of endogenous histone H3. SUMOylation regulates target proteins involved in all nuclear processes including transcription, DNA repair, chromatin remodeling, precursor-mRNA splicing and ribosome assembly.

Reversible post-translational modification (PTM) of lysine residues in proteins by small ubiquitin-like modifiers (SUMOs) has a key role in genome stability and transcription¹⁻³. Important SUMO-target proteins in the DNA-damage response include PCNA⁴,⁵, BRCA1 (ref. 6) and 53BP1 (ref. 7). SUMOs are conjugated to target proteins via an enzymatic cascade involving a dimeric E1 enzyme (SAE1–SAE2), a single E2 enzyme (Ubc9) and a limited number of E3 enzymes⁸. Mice deficient in Ubc9 die at the early postimplantation stage as a result of chromosome-condensation and chromosome-segregation defects, thus underlining the essential role of SUMOylation in the maintenance of genome stability⁹.

Frequently, SUMOylation regulates the function of target proteins by enabling or stabilizing noncovalent protein-protein interactions via SUMO-interaction motifs (SIMs)⁸. Classical examples of this type of interaction include the binding of SUMOylated RanGAP1 to the nucleoporin RanBP2 (ref. 10) and the binding of the SRS2 helicase enzymes. We enriched SUMOylated peptides from a HeLa cell line stably expressing decahistidine (His10)-tagged SUMO-2. This tag is small and compatible with denaturing buffer conditions. We established these stable cells by using a bicistronic lentivirus encoding His10-SUMO-2 and GFP separated by an internal ribosome entry site (IRES). After infection, we used flow cytometry to obtain a population expressing this construct at low levels. We confirmed low expression levels by immunoblotting (Fig. 1a). As expected, the protein located predominantly in the nucleus (Fig. 1b)³⁰,³¹. The His10 tag, in contrast to the hexahistidine tag commonly used in the field, enabled single-round purification with a high yield and purity (Fig. 1c).

In order to enrich for SUMOylated peptides, we used a SUMO-2 form that is resistant to cleavage by endopeptidase Lys-C. Lysine-deficient (K0) SUMO-2 behaves very similarly to wild-type SUMO-2, except for SUMO polymerization²⁹. Our purification strategy consisted of a first round of His10-SUMO-2 purification, a filter step to concentrate SUMO-2 conjugates while simultaneously separating SUMO-2 conjugates from free SUMO-2, a digestion by Lys-C and a second round of purification followed by a trypsin digest (Fig. 1d,e). The second round of purification enabled enrichment at the site level, greatly reducing the complexity of the final sample.

¹Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, the Netherlands. ²Department of Proteomics and Signal Transduction, Max Planck Institute for Biochemistry, Martinsried, Germany. Correspondence should be addressed to A.C.O.V. (vertegaal@lumc.nl).

Received 3 March; accepted 19 August; published online 14 September 2014; doi:10.1038/nsmb.2890
Figure 1 A strategy for mapping SUMO-2–acceptor lysines in endogenous proteins. (a) Immunoblot (IB) confirming the low-level expression of His$_{10}$–SUMO-2 K0 in HeLa cells. Ponceau-S staining is shown as a loading control. His$_{10}$ pulldown (PD) was performed to enrich SUMOylated proteins, and Ponceau-S staining is shown to indicate high-specificity enrichment. The experiment shown was replicated twice. MW, molecular weight. (b) Confocal fluorescence microscopy image confirming the predominantly nuclear localization of His$_{10}$–SUMO-2 K0. DIC, differential interference contrast. Scale bars, 25 μm. The experiment shown was replicated twice. Asterisk indicates the front of the gel and multiple size markers smaller than 51 kDa. (c) Coomassie stain displaying the efficacy of a single-step His$_{10}$ pulldown performed on approximately 50 million HeLa and His$_{10}$–SUMO-2 K0 cells. The experiment shown was replicated twice. (d) Schematic overview of the His$_{10}$–SUMO-2 K0 SUMOylation-site purification strategy. A direct purification step is followed by concentration of SUMOylated proteins, which are subsequently digested with endopeptidase Lys-C. The His$_{10}$–SUMO-2 K0 bearing the SUMOylated peptide is repurified, concentrated, digested with trypsin and finally analyzed by high-resolution nanoscale LC-MS/MS. (e) Immunoblotting analysis to confirm the efficiency of the purification steps described in (d). The experiment shown was replicated twice. (f) Immunoblotting analysis of total lysates from cells stably expressing His$_{10}$–SUMO-2 K0 which were mock treated or treated with MG-132, PR-619 or heat shock. Ponceau-S staining is shown to indicate differential expression of cells (~20 million). Moreover, unlike previously reported methodologies, it allows cells to be grown under standard conditions. Our data set includes an extensive number of well-known SUMO-targetologies, it allows cells to be grown under standard conditions. Our data set includes an extensive number of well-known SUMO-target.

The C-terminal tryptic fragment of wild-type human SUMO-2 is 32 amino acids (aa) and, owing to its size, is not compatible with efficient mapping of SUMO-2–acceptor lysines. In contrast, yeast SUMO, Smt3, contains a conveniently located arginine that results in a 5-aa C-terminal tryptic fragment. We have generated a SUMO-2 Q87R mutant mimicking yeast SUMO to enable the identification of SUMO-2–acceptor lysines by MS. SUMOylation is a dynamic process that is regulated via cross-talk with the ubiquitin-proteasome system and is sensitive to heat shock and to the broad-range inhibitor of SUMO and ubiquitin proteasomes PR-619. Treatments with the proteasome inhibitor MG-132, heat shock and PR-619 resulted in considerable accumulation of SUMO-2 conjugates (Fig. 1f).

Identification of 4,361 SUMO-2 sites in 1,606 proteins

We analyzed tryptic digests of in solution–digested repurified SUMOylated peptides by nanoscale LC-MS/MS, without further fractionation, with 2-h LC gradients. We identified 5,339 SUMOylated peptides corresponding to 4,361 unique SUMOylation sites (Supplementary Table 1) in 1,606 proteins (Supplementary Table 2), at a false discovery rate (FDR) below 1%. Mass accuracy was within 3 p.p.m. for 98.0% of all identified sites and within 6 p.p.m. for all sites, with an average absolute mass error of 0.77 p.p.m. The majority of identified SUMOylation sites had an Andromeda peptide score in the range of 60–100 (Fig. 2a). We pinpointed the precise SUMO-2–acceptor lysines in over 98.8% of the SUMOylated peptides (Supplementary Table 1). The overall purity achieved by our method was demonstrated by the average presence of a SUMO-2–acceptor lysine in 25.0% of the peptides identified in the final purified fractions. For negative controls, we attempted identification of SUMOylated lysines after performing site enrichment on the parental HeLa cell line. In addition, we attempted to identify SUMO sites from HeLa total lysates. In both cases, we did not find a single site.

This methodology represents a step forward for the field because it enables efficient analysis of SUMOylation at a proteome-wide level in a site-specific manner and requires only a relatively small number of cells (~20 million). Moreover, unlike previously reported methodologies, it allows cells to be grown under standard conditions. Our data set includes an extensive number of well-known SUMO-target proteins—for example, RanGAP1, PML, topoisomerases 1, 2α and 2β, PCNA, BLM, BRCA1, RanBP2, RNF168 and SAFB2 (Supplementary Table 2)—thus further confirming the validity of the approach.

In total, we identified 1,069 sites from cells grown under regular cell-culture conditions (Fig. 2b,c). The dynamic nature of SUMOylation was underlined by the identification of 3,292 additional SUMOylation
Further, we validated a set of newly identified SUMO-target proteins—RNF216, SNW1, TCF12 and ZND280D—through pulldown and immunoblot analysis with a cell line stably expressing His10-tagged SUMO-2 wild type (Supplementary Fig. 2b). Additionally, because we used the PR-619 inhibitor of SUMO protease, which had not been previously reported to our knowledge, we were able to validate the quantified change in SUMOylation of various target proteins by using a wild-type SUMO cell line. This set includes four known SUMO targets (FOXM1, HRNPM, RAD18 and SART1) as well as two new SUMO-target proteins (WDR70 and MCM10) (Supplementary Fig. 2c).

SUMOylation sites per protein ranged from a single site in nearly half of all SUMO-target proteins to ten or more sites in 96 proteins and 20 or more sites in 13 proteins: PARP1 (20 sites), ZMYM2 (20 sites), ZNF281 (21 sites), MECOM (22 sites), NSUN2 (22 sites), MKI67 (23 sites), MIS18BP1 (25 sites), TFX2 (26 sites), BLM (27 sites), NKTR (27 sites), FBN1 (31 sites), GTF2I (34 sites) and ZNF451 (40 sites) (Fig. 2e). We found that most proteins (64.2%) were conjugated to only one or two SUMO moieties.

We compared all SUMOylated proteins and sites identified in this study to those from previous studies on SUMOylation. On average, we identified 51% of previously reported MS/MS-identified SUMO-target proteins (Supplementary Fig. 3a and Supplementary Table 4) and expanded the known number of SUMOylated proteins by nearly 1,000. We found 52% of all previously reported MS/MS-identified SUMOylation sites (Supplementary Fig. 3b), as reported in the PhosphoSitePlus database (PSP; http://www.phosphosite.org) and more recently by Schimmel et al. and Tammsalu et al. Our study expands the number of known MS/MS-identified SUMOylation sites by over 3,000, and we identified over 1,000 SUMOylation sites under standard growth conditions.

SUMOylation is a key post-translational modification in all eukaryotes but is absent in prokaryotes. We studied phylogenetic conservation of SUMOylation with respect to conservation of entire proteomes (Supplementary Fig. 4a). SUMOylated proteins are substantially more conserved than total proteomes. Within orthologs, SUMOylation is the most conserved post-translational modification, together with
SUMO is extensively involved in PTM cross-talk. (a) Schematic representation of the overlap of the identified SUMOylated lysines (K) and lysines modified by ubiquitination, acetylation or methylation. Enrichment ratios between observed overlaps and expected overlaps are indicated and were significant according to Fisher’s exact test, with \( P < 1 \times 10^{-10} \). (b) Combined overlap between SUMOylation, ubiquitination, acetylation and lysine methylation. (c) Overview of phosphorylation sites adjacent to SUMOylated lysines, both identified by MS in our screen, and their amino acid spacing in relation to the SUMOylated lysine. Peptides SUMOylated exclusively in conjunction with phosphorylation are in blue. SUMOylated peptides with nonunique phosphorylation sites are in red. (d) Schematic representation of the identified SUMOylation sites on ubiquitin. (e) Schematic representation of the histone H3 peptide as identified by MS/MS, simultaneously modified by SUMOylation on Lys19 and acetylation on Lys24. Identified fragment ions are indicated. A fully annotated high-resolution MS/MS spectrum is available as part of Supplementary Data Set 1. (f) Immunoblot analysis of total lysates and His-tagged samples from HeLa cells stably expressing His10–SUMO-2 wild type or K0 mutant, which were either mock treated, treated with the histone-deacetylase inhibitor trichostatin A (TSA) at the indicated dose in nM or treated with the histone-deacetylase inhibitor curcumin (Cur) at the indicated dose in \( \mu \)M. Ponceau-S staining is shown as a loading control. H3ac, acetylated H3. The experiment shown was replicated twice.

acetylation, which is equally conserved. In proteins with no orthologs in lower eukaryotes, SUMOylation is more conserved than phosphorylation but is less conserved than ubiquitination, acetylation and methylation (Supplementary Fig. 4b). This difference indicates an increased frequency of SUMOylation occurring on proteins that are absent in lower eukaryotes.

We investigated the potential overlap between the identified SUMO-acceptor lysines and other post-translational lysine modifications. For this purpose, we extracted all known human MS/MS-identified ubiquitination, acetylation and lysine-methylation sites from PSP and cross-compared modification sites (Fig. 3a,b and Supplementary Table 5). SUMOylation is known to compete with ubiquitin for acceptor lysines in target proteins\(^ {27} \). However, the extent of this cross-talk is currently unclear. We compared the identified SUMO-acceptor lysines to acceptor lysines for ubiquitin in the PSP database and found that nearly one-quarter (22.4%) of SUMOylation sites are also known to be ubiquitinated, thus indicating extensive cross-talk between SUMOylation and ubiquitination (Fig. 3a,b). Overlap between SUMOylation and acetylation and between SUMOylation and lysine methylation occurs less frequently, although this could be related to the smaller number of acetylation and methylation sites currently identified. From the perspective of all known ubiquitination, acetylation and methylation sites, SUMOylation occurs at roughly 4% of all these sites and does not favor one PTM over the other. Considering the observed amount of overlap between SUMO and other PTMs on the same lysines, as compared to the total lysines in the human proteome, the observed overlap is substantial.

Interestingly, cross-talk between SUMOylation and other post-translational modifications includes regulation of enzymatic components including 46 kinases, 33 proteins with intrinsic phosphatase activity, 29 ubiquitin protein ligase family members, 7 ubiquitin proteases, 7 acetyltransferases, 10 deacetylases, 22 methyltransferases and 13 demethylases (Supplementary Table 6).

We found 23 peptides SUMOylated exclusively in conjunction with phosphorylation and an additional 47 peptides SUMOylated together with phosphorylation in a ‘nonunique’ fashion (Supplementary Table 7). Phosphorylation occurred relatively close to the lysine and was present both upstream and downstream of the SUMOylation lysine. We found that phosphorylation occurred predominantly at or near to the +5 position (Fig. 3c), a result in agreement with the earlier described PDSM motif\(^ {29,38} \). We observed five phosphorylation sites at position +2 relative to the SUMOylated lysine; these sites, instead of glutamate or aspartate, could serve as the negative charge required for efficient SUMOylation.

Furthermore, we found direct modification of endogenous ubiquitin by SUMO-2 on Lys11, Lys48 and Lys63 under control conditions (Fig. 3d). After cellular treatments, we additionally observed SUMO-2 modification of ubiquitin on Lys6 and Lys27 (Fig. 3d). Thus, mixed-chain formation between ubiquitin and ubiquitin-like family members is more extensive than was previously thought. Furthermore, we detected mixed-chain formation between all SUMO family members (Supplementary Table 1).

We found three peptides SUMOylated exclusively in conjunction with acetylation, with two of these events occurring on histones...
H3 (Fig. 3e) and H4 (Supplementary Table 7). We detected an additional SUMOylated peptide from PML together with acetylation in a nonunique fashion. These sites could indicate acetylation-dependent SUMOylation, thus suggesting a new type of cross-talk between these two major modifications. In order to further investigate such a dependency, we treated HeLa cells expressing either wild-type or

![Image](https://example.com/image1)

**Figure 4** New insight into the SUMOylation consensus motif. (a) IceLogo representation of all SUMOylation sites identified under control conditions. The height of the amino acid letters corresponds to fold change. All amino acid changes were significant with \( P < 0.05 \) by two-tailed Student’s \( t \) test (\( n = 1,069 \) sites). (b) SubLogo representation of various consensus motifs. The height of amino acid letters represents the percentage change of enrichment or depletion between the motif set and the reference set. All amino acid changes were significant with \( P < 0.05 \) by two-tailed Student’s \( t \) test. KxE, \( n = 1,300 \) sites compared to 3,061 sites; (ED)xK, \( n = 900 \) sites compared to 3,461 sites; KxD, \( n = 245 \) sites compared to 4,116 sites; (VI)K, \( n = 959 \) sites compared to 3,402 sites; (VI)xxK, \( n = 460 \) sites compared to 3,901 sites; KQ, \( n = 264 \) sites compared to 4,097 sites. Asterisks mark the position of the SUMOylated lysines. (c) As in a, but in heat-map format. Blue color indicates a statistical enrichment as compared to randomly expected, and red indicates a depletion. All amino acid changes were significant with \( P < 0.05 \) by two-tailed Student’s \( t \) test (\( n = 1,069 \) sites). (d) Overview of the number of SUMOylation sites matching the short consensus motif KxE in different subsets of sites corresponding to different cellular treatments. Additionally, per subset, the top 25% of most intense sites, the top 50% of most intense sites and all sites are shown. Matching of ubiquitin sites to the motif and the randomly expected frequency is also shown. (e) As in d, but for the short consensus motif (IVML)K. (f) Cysteine frequency close to all identified control SUMOylation sites as well as other PTM sites, ranging from −10 to +10 amino acids around the modified lysine. For all PTMs, a second-order polynomial trend line was calculated. The background cysteine frequency is indicated.
lysine-deficient SUMO with the histone deacetylase inhibitor trichostatin A (TSA), known for increasing acetylation of histones, or the histone acetyltransferase inhibitor curcumin, known for decreasing acetylation of histones. More specifically, we performed SUMO enrichment and investigated the SUMOylation state of histone H3 (Fig. 3f). To our knowledge, this is the first visualization of the modification of endogenous histone H3 by SUMO. Furthermore, we found that the SUMOylation of histone H3 increased upon TSA treatment, in correlation with an increase in histone H3 acetylation. Conversely, after treatment with curcumin SUMOylation of histone decreased, in correlation with a decrease in histone H3 acetylation. The total pool of conjugated SUMO was mildly increased regardless of the inhibitors used, thus ruling out a nonspecific change in the SUMOylation state of histone H3 (Fig. 3f).

Insight into the SUMOylation consensus motif
SUMOylation is known to occur on the classical consensus motif ΨKx(K/E) [Ψ, X, (K/E)]50,51, where Ψ is a large hydrophobic amino acid. Previously, we found that other residues are also used at the Ψ position59. Our data set provides an opportunity to obtain further insight into the SUMOylation consensus motif. More than half of the identified sites from untreated cells matched the consensus motif KxE, whereas KxD-type sites were not enriched over background frequency (Fig. 4a). We studied the KxE-type SUMOylation motif in RanGAP1, a highly SUMOylated protein10,42. E526D replacement resulted in a notable
A drop in SUMOylation (Supplementary Fig. 5a,b). For a negative control, we included our previously described ∆GL RanGAP1 mutant29. The acidic residue is not necessarily located two positions downstream of the SUMOylated lysine; it can also be found two positions upstream, with aspartate occurring at a higher frequency at this position than at the regular SUMOylation motif (Fig. 4b), in agreement with the inverted consensus motif EDxK that we proposed on the basis of a very small number of identified SUMOylation sites29. We confirmed the relevance of the Glu300 residue in the inverted motif covering Tel SUMOylated at Lys302 (Supplementary Fig. 5c,d). Replacing this residue with alanine eliminated SUMOylation, whereas replacing it with aspartate did not completely abolish SUMOylation, a result in agreement with the inverted consensus motif (ED)xK.

We superimposed the 1,069 SUMOylation sites identified under control conditions, including a sequence window ranging from −15 to +15 aa, displaying the amino acid frequencies normalized to the randomly expected frequencies across the human proteome (Fig. 4a,c). We observed the highest degree of enrichment for valine and isoleucine at −1, and glutamate at +2, with over half of all control SUMOylation sites adhering to this consensus. Among the top 50% and top 25% of the most abundant SUMOylation sites for glutamate at +2, this frequency increased to 60% and 67%, respectively, indicating that the lysines situated in SUMOylation consensus motifs are efficiently SUMOylated (Fig. 4d). We observed a similar trend for the hydrophobic amino acids at −1 (Fig. 4e). A further expanded consensus motif, taking statistical local enrichments into account, was (IVML)-K-(EQMTP)-E-P. Interestingly, the adherence to the consensus motif dropped moderately for sites mapped exclusively after heat-shock treatment and decreased drastically after MG-132 and PR-619 treatment, after which sites matching KxE reached as low as 13%, barely higher than the ubiquitin or random-lysine frequencies (Fig. 4d).

Figure 6 SUMO modifies highly interconnected functional networks of proteins. (a) STRING-network analysis of all identified SUMOylated proteins, with a STRING interaction confidence of 0.7 or greater. MCODE was used to extract the most highly interconnected functional clusters from the network, which are indicated in different colors. (b) Overview of relative STRING-network score corresponding to Table 1. This score was computed through multiplication of the interaction enrichment ratio, protein network connectivity and average STRING confidence of all interactions. (c) Schematic overview of the three highest-scoring MCODE subclusters from a. The size and color of the individual proteins correspond to the number of SUMOylation sites identified in the protein. The six additional MCODE clusters are available in Supplementary Figure 7. 

NATURE STRUCTURAL & MOLECULAR BIOLOGY  VOLUME 21  NUMBER 10  OCTOBER 2014 933
Table 1: Subsets of SUMO-target proteins are differentially interconnected

| Selection     | Proteins | Interactions | Enrichment | Connected |
|---------------|----------|--------------|------------|-----------|
| All           | 1,584    | 6,801        | 9.4        | 60.0%     |
| All (2+ sites)| 860      | 2,765        | 11.6       | 57.6%     |
| Control       | 790      | 2,035        | 12.1       | 58.7%     |
| Control (2+ sites) | 539   | 1,202        | 13.4       | 56.4%     |
| Heat          | 1,222    | 5,240        | 10.2       | 65.1%     |
| Heat (LFQ 1+) | 424      | 987          | 11.9       | 56.8%     |
| Heat (LFQ 3+) | 188      | 242          | 14.6       | 52.1%     |
| PR-619        | 1,197    | 4,349        | 10.5       | 59.0%     |
| PR-619 (LFQ 1+) | 423   | 793          | 11.7       | 49.4%     |
| PR-619 (LFQ 3+) | 151   | 133          | 12.3       | 41.7%     |
| MG-132        | 1,208    | 4,263        | 9.7        | 56.3%     |
| MG-132 (LFQ 1+) | 431   | 651          | 11.2       | 45.0%     |
| MG-132 (LFQ 3+) | 202   | 145          | 11.6       | 32.2%     |
| All conditions| 596      | 1,347        | 12.2       | 57.9%     |

Overview of STRING analyses of different subsets of SUMOylated proteins. Enrichment is a ratio derived from the observed number of interactions divided by the expected number of interactions. Connected refers to the percentage of input proteins connected to the core cluster. P values for all individual analyses are <1 × 10−15 by Fisher’s exact test. The network displayed in Figure 6a corresponds to ‘all’.

The 30-aa region flanking SUMOylated lysines was enriched with lysine and glutamate (Fig. 4a,c). Thus, SUMOylated lysines are frequently located in regions that are enriched with charged residues and are therefore probably solvent exposed. Interestingly, SUMOylated regions are deficient in phenylalanine, tryptophan, tyrosine, leucine and, most notably, cysteine (Fig. 4a,c). Because SUMOs are transferred along an enzymatic cascade via thioester formation, the reduced frequency of cysteines near SUMOylated lysines under standard conditions could help to prevent the formation of thioesters between SUMOs and target proteins. Interestingly, a reduced frequency of cysteine could also be observed for ubiquitinated regions, probably for the same reason as proposed for SUMOylation (Fig. 4f). Reduced frequencies of cysteine were less pronounced for regions flanking methylated or acetylated lysines (Fig. 4f).

Insight into SUMOylated protein groups

Protein domains that are frequently SUMOylated include the Krüppel-associated box (KRAB) domain, which is a repressor domain found in many zinc-finger protein–based transcription factors (Fig. 5a). Other domains included zinc fingers, PHD fingers and RRM1, which all have important roles in binding of DNA, RNA or other proteins and are often found in nuclear or chromatin-associated proteins.

We investigated the subcellular localization of SUMOylated proteins by Gene Ontology (GO) cellular compartments classes and plotted all identified proteins and sites (Fig. 5b). We found SUMOylation to be an almost exclusively nuclear modification, with cytoplasmic modification occurring primarily on proteins that are also annotated as nuclear. Enrichment analysis showed the highest ratio for chromatin-associated proteins, and this was closely followed by the ratio for nuclear proteins (Fig. 5b). In contrast, cytoplasmic and membrane proteins were depleted. Correspondingly, the number of SUMO sites per protein was also higher in chromatin-associated and nuclear proteins.

The first identified SUMOylation site, Lys524 in the nuclear-pore component RanGAP1, is located in an unstructured region of the protein[3]. SUMOylation is thought to occur predominantly in unstructured regions[4]. To investigate the localized structural properties of proteins around sites of SUMOylation, we folded all 4,361 sites in silico, including the 30-aa sequence window, as well as over 5,000 lysines randomly chosen from SUMOylated proteins as a reference set. We performed secondary-structure prediction of the modified lysine and classified the structures as α-helix, β-sheet or otherwise coiled (Fig. 5c). Our results indicate a modest reduction in SUMOylation of α-helices and a significant increase in SUMOylation of β-sheets compared to background frequencies. This trend was most striking for KxEx-type SUMOylation sites, where we observed a significant decrease in unstructured regions, as a trade-off for an increase in β-sheets. We additionally observed an increased tendency for SUMOylated regions to be solvent exposed (Fig. 5d).

SUMO modifies highly interconnected networks of proteins

Genome stability, transcription and translation are three important biological processes as evidenced by term enrichment analysis for GO biological processes involving the identified SUMO targets (Fig. 5e). Furthermore, nucleic acid metabolism, chromosome organization, DNA repair, cell-cycle regulation, RNA splicing, histone modification and nuclear-body organization are among the most enriched processes. For GO molecular functions, in absolute numbers, we identified 673 DNA-binding proteins and 484 zinc-binding proteins as the largest functional groups of SUMO-target proteins (Fig. 5f). SUMO also substantially modified subunits from known complexes in the Comprehensive Resource of Mammalian Protein Complexes (CORUM) database, including Nop56p precursor rRNA, SIN2–SAP25, BRF53–BRCA2, LARG, BHC, MeCP1 and HDAC1 and HDAC2 protein complexes (Fig. 5g). Additional analyses by keywords and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms also highlighted SUMO’s regulation of many important cellular processes (Fig. 5h), including an enrichment of proteins known to be involved in cancer pathways (Supplementary Table 8), such as TP53, MTF, VHL, BRCA2, STAT1, FOS, JUN and SMAD4. The complete term enrichment analysis (Supplementary Table 8) and a fully annotated list of all SUMOylated proteins (Supplementary Table 9) are available online.

SUMOylated proteins form a very complex, highly organized network of interacting proteins, as we visualized with search tool for the retrieval of interacting genes and/or proteins (STRING) network analysis (Fig. 6a). 60% of all identified proteins are part of one main functional cluster, at high STRING confidence. We performed STRING analyses on a per-treatment basis and at high STRING confidence to assess protein–protein–interaction enrichment ratios and network participation (Table 1). Overall, we observed ten times more interactions than expected, with the SUMO-target proteins from untreated cells and with multiple sites showing the highest degree of enrichment. The strength of the STRING networks, derived from interaction enrichment as well as participation of all proteins, was greatest for proteins identified under control conditions (Fig. 6b). Strikingly, we observed proteins identified after heat shock and found that many proteins were significantly upregulated in SUMOylation after heat shock to form a highly coherent STRING network. This stands in contrast to proteins upregulated in SUMOylation after MG-132 or PR-619 treatments, which we found to be less related in their interactions.

Subsequently, we performed MCODE analysis, which revealed highly interconnected subclusters within the core network, including nine subclusters with interconnectivity scores ranging from 9 to 39 (Fig. 6a). Three dominant clusters involve many functionally related proteins from the spliceosome and the ribosome as well as cell cycle–related factors (Fig. 6c). Other clusters contain chromatin-remodeling proteins, histone deacetylases, histone methyltransferases, regulators of mitotic prometaphase, regulators of ubiquitin protein...
ligases and proteins involved in ribonucleoprotein–complex formation (Supplementary Fig. 7).

We found sites of SUMOylation, ubiquitination and acetylation to overlap substantially (Fig. 3a,b). To further investigate this, we performed STRING analysis on the subset of proteins containing these lysines. We found that 70% of these proteins were situated in a single functional network (Supplementary Fig. 6a–c). We found that, beyond the observed overlap between modification sites, these clusters of proteins were highly modified by SUMO, averaging over five SUMOylation sites per protein (Supplementary Fig. 6d).

Additionally, in these clusters as compared to the full SUMO network, we observed a higher degree of enrichment for protein–protein interactions than expected (Supplementary Fig. 6e) and a much higher degree of network participation of all proteins. This resulted in some of the highest-scoring networks for proteins containing lysines modified by SUMO and either ubiquitin or acetyl groups, and by far the strongest network for proteins that are modified on the same lysines by all three of these major PTMs (Supplementary Fig. 6f). Thus, SUMOylation appears to function in concert with other major PTMs, and it co-regulates a tight functional cluster of heavily modified and dynamic proteins.

**DISCUSSION**

We have developed a new methodology for identification of global SUMOylation sites, which enabled us to map over 4,300 SUMO-acceptor lysines in over 1,600 proteins and has provided detailed insight into the function of SUMOylation. All nuclear processes are orchestrated by SUMOylation, including transcription, DNA repair, chromatin remodeling, precursor-mRNA splicing and ribosome assembly. Our data set represents an improvement over the relatively small number of SUMOylation sites that were found previously in cells cultured in standard growth conditions and provides a basis for the scientific community to perform follow-up studies. Our data set expands the number of known SUMOylation sites by over 3,000, reconfirms many of the 1,000 sites that were recently mapped in response to heat shock and provides over 1,000 new sites identified under standard growth conditions. This is particularly relevant for the half of the identified sites that comprised sites not located in a SUMOylation consensus motif and that therefore have eluded in silico prediction. Furthermore, we have obtained new insight into the consensus motif for SUMOylation, finding that SUMOylation sites are frequently located in domains enriched in charged residues. Moreover, protein regions with SUMOylated lysines are deficient in cysteines, possibly to limit thioester formation between SUMOs and target proteins.

We also identified an unprecedented number of phosphorylation events occurring in proximity to SUMOylation, with 70 of these events almost matching the total number of SUMOylation sites mapped in our previously reported screen. Moreover, the identification of four SUMO and acetyl comodified peptides, and of the acetylation-dependent modification of endogenous histone H3, suggests an interesting and new prospect for cross-talk between these two major PTMs in their regulation of nucleosomes. Previously, it was demonstrated that SUMOylation of histones is linked to transcriptional repression and that, beyond the observed overlap between modification sites, these clusters of proteins were highly modified by SUMO, averaging over five SUMOylation sites per protein (Supplementary Fig. 6d).

In contrast, acetylation is associated with transcriptional activation. Acetylation-dependent SUMOylation of histone H3 is thus a surprising type of cross-talk with unclear significance. Hypothetically, these differences could be reconciled if histone H3 were to be first acetylated to activate transcription and later SUMOylated to generate a transient activation pattern. It would be interesting to study whether acetylated and SUMOylated histone H3 is located on the genome.
ONLINE METHODS

Plasmids. The His10–SUMO–2 K0 Q87R that we described and used in this manuscript has the following amino acid sequence: MAHHHHHHHHHHHGSMSERPRREGVRQKEHHNLVRAGDGSVGQFQIRRHTTPLSRLMARCYCERQGLSMRQRFRFDFQQPINETDTPALEMEDEDTDIDVFQFQTFGG. The His10–SUMO–2 wild type that we described and used in this manuscript has the following amino acid sequence: MAHHHHHHHHHHHGSMSERPKPEGVKTECTHHNLVRAGDGSVGQFQIRRHTTPLSRLMARCYCERQGLSMRQRFRFDFQQPINETDTPALEMEDEDTDIDVFQFQTFGG. The corresponding nucleotide sequences were cloned in between the PstI and XhoI sites of the plasmid pLV-CMV-RES–GFP46.

Cell culture and cell-line generation. HeLa and U2-OS cells were cultured in Dulbecco’s Modified Eagle’s Medium (DMEM) supplemented with 10% FBS and 100 U per mL penicillin and streptomycin (Invitrogen). HeLa cells stably expressing His10–SUMO–2 or His10–SUMO–2 K0 Q87R were generated through lentiviral infection with a virus encoding His10–SUMO–2–RES–GFP or His10–SUMO–2 K0 Q87R–RES–GFP. Two weeks after infection, cells were fluorescence-scored for a low expression level of GFP with a FACS Aria II (BD Biosciences). Cells were passed through a 100-µm capillary at a pressure of 138 kPa and selected for 7.5 × 10² to 3 × 10³ GFP BF 530/30-A intensity, which was 2.5–10 times higher than the background cellular autofluorescence of 3 × 10¹.

Treatments, transfection and lentiviral infection. In order to accumulate SUMOylated proteins, cells were treated with 10 µM MG-132 (Sigma) dissolved in DMSO for 7 h, treated with 20 µM PR-619 (Millipore) dissolved in DMSO for 7 h, or incubated at 43 °C for 1 h (heat shock). For increasing acetylation of histones, trichostatin A (TSA, Sigma) was used at a concentration of 150 nM or 600 nM for 18 h49. In order to decrease acetylation of histones, curcumin (Sigma) was used at a concentration of 25 µM or 50 µM for 18 h49. For transfection, cells were cultured in DMEM lacking penicillin and streptomycin. Transfections were performed with 2.5 µg of polyethylenimine (PEI) per 1 µg of plasmid DNA, with 1 µg of DNA per 1 million cells. Transfection reagents were mixed in 150 mM NaCl and incubated for 15 min before direct addition to the cells. Cells were split after 24 h and investigated after 48 h. Lentiviruses were generated essentially as described previously49. Infections were performed with a multiplicity of infection of two and with a concentration of 8 µg per mL polybrene in the medium. 24 h after infection, the medium was replaced.

Purification of His10–SUMO–2 and His10–SUMO–2 K0 Q87R, stage 1. Per single MS/MS run to identify SUMO–2 sites, one single fully confluent 15-cm dish of His10–SUMO–2 K0 Q87R (~20 million cells) was prepared. Cells were washed three times on the plate with ice-cold PBS before being scraped and collected in a 15-mL tube. Cells were centrifuged at 250 r.c.f. and resuspended in ice-cold PBS. Subsequently, the cell pellets were lysed in ten pellet volumes of guanidine lysis buffer (6 M guanidine–HCl, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 8.0). Lysates were subjected to sonication with a microtip sonicator at a power of 30 W. Sonication bursts of 5 s per 5 mL lystate were used, to a total sonication time of 15 s. Subsequently, lysates were supplemented by addition of imidazole to 50 mM and β-mercaptoethanol to 5 mM. 20 µL (dry volume) Ni–NTA agarose beads (Qiagen) were prepared per 1 mL sample. The equilibrated beads were added to the lysates and allowed to tumble at 4 °C for 5 h. After incubation, beads were pelleted by centrifugation at 500 r.c.f. and washed for 15 min with at least five bead volumes of the following wash buffers in order: wash buffer 1 (6 M guanidine–HCl, 0.1% Triton X-100, 10 mM imidazole, 5 mM β-mercaptoethanol, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 8.0), wash buffer 2 (8 M urea, 0.1% Triton X-100, 10 mM imidazole, 5 mM β-mercaptoethanol, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 8.0), wash buffer 3 (8 M urea, 10 mM imidazole, 5 mM β-mercaptoethanol, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 6.3), wash buffer 4 (8 M urea, 5 mM β-mercaptoethanol, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 6.3), and wash buffer 5 (same as wash buffer 4). Subsequently, all wash buffer was removed from the beads, and proteins were eluted for 20 min with one bead volume of elution buffer (7 M urea, 300 mM imidazole, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 7.0). The elution procedure was repeated twice, and all eluates were pooled and passed through 0.45-µM filters (Ultrafree, Millipore). Next, samples were concentrated on 10 kDa–cutoff spin filters (Vivacon 500, Sartorius Stedim) at 20 °C and at 14,000 r.c.f. Concentration was performed until <10–50 µL of sample remained. After concentration, the proteins remaining on the filters were washed once with 250 µL of elution buffer minus imidazole and recombinant. Final concentrated SUMOylated peptides were removed from the filters and were snap frozen and stored at –80 °C. The single stage–purified SUMOylated proteins are compatible with both in-gel and in-solution digestion protocols and subsequent MS analysis and may also be target-specifically investigated by SDS-PAGE and immunoblotting analysis. Alternatively, repurification of SUMOylated peptides is performed according to stage 2 of the protocol.

Purification of His10–SUMO–2 K0 Q87R, stage 2. Sequencing-grade endoprotease Lys-C (Wako) was added to the samples in a 1:25 enzyme/protein ratio and incubated for 4 h at room temperature, still and in the dark. Subsequently, another 10 mM of fresh β-mercaptoethanol was added to the first sample, and this was followed by an additional amount of Lys-C equal to the first amount. The second incubation was performed overnight, at room temperature, still and in the dark. Next, digests were transferred to 15-mL tubes and diluted with an amount of guanidine lysis buffer equal to half the amount used to lyse the initial cell pellet. The samples were then supplemented with addition of imidazole to 50 mM and β-mercaptoethanol to 5 mM. Next, 40 µL (dry volume) Ni–NTA agarose beads (Qiagen) were added to the lysates and allowed to tumble at 4 °C for 5 h. After incubation, beads were pelleted by centrifugation at 500 r.c.f. and washed for 15 min with at least five bead volumes of the following wash buffers in order: wash buffer 1 (6 M guanidine–HCl, 0.1% Triton X-100, 10 mM imidazole, 5 mM β-mercaptoethanol, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 8.0), wash buffer 2 (8 M urea, 0.1% Triton X-100, 10 mM imidazole, 5 mM β-mercaptoethanol, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 8.0), wash buffer 3 (8 M urea, 10 mM imidazole, 5 mM β-mercaptoethanol, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 6.3), wash buffer 4 (8 M urea, 5 mM β-mercaptoethanol, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 6.3), and wash buffer 5 (same as wash buffer 4). Subsequently, all wash buffer was removed from the beads, and proteins were eluted for 20 min with one bead volume of elution buffer (7 M urea, 300 mM imidazole, 100 mM sodium phosphate, and 10 mM Tris, buffered at pH 7.0). The elution procedure was repeated twice, and all elutions were pooled and passed through 0.45-µM filters (Ultrafree, Millipore). Next, samples were concentrated on 10 kDa–cutoff spin filters (Vivacon 500, Sartorius Stedim) at 20 °C and at 14,000 r.c.f. Concentration was performed until <10–25 µL of sample remained. After concentration, the proteins remaining on the filters were washed twice with 250 µL of elution buffer minus imidazole and recombinant. Final concentrated SUMOylated peptides were removed from the filters and were snap frozen and stored at –80 °C. The double-purified SUMOylated peptides are compatible with in-solution digestion protocols and subsequent MS analysis aimed at determination of specific sites of protein SUMOylation.

In-solution digestion. SUMOylated peptides were supplemented with ammonium bicarbonate (ABC) to 50 mM. Subsequently, dithiothreitol (DTT) was added to a concentration of 1 mM, and samples were left to incubate at room temperature for 30 min. Next, chloroacetic acid was added to a concentration of 5 mM, and samples were incubated at room temperature for 30 min. After alkylation, another 5 mM of DTT was added, and samples were left to incubate at room temperature for 30 min. At this point, samples were gently diluted four-fold with 50 mM ABC. Subsequently, an amount of sequencing-grade modified trypsin (Promega) was added equal to 25% of the Lys-C initially used in a single digestion step. Digestion with trypsin was performed overnight, at room temperature, still and in the dark.

LC-MS/MS analysis. In solution–digested peptides were cleaned, desalted and concentrated on triple-disc C18 reverse-phase StageTips50, before being eluted twice with 25 µL 80% acetonitrile in 0.1% formic acid. Desalted peptides were vacuum centrifuged at room temperature until 10% of the original volume remained, before online nanoflow liquid chromatography–tandem MS. The analysis of in solution–digested peptides was performed with an EASY-nLC system (Proxeon) connected to a Q-Exactive (Thermo) with higher collisional.
dissociation (HCD) fragmentation. Separation of peptides was performed with 20-cm-long analytical columns (ID 75 μm, Polymeric Avantes) packed in house with 1.8 μm C18 beads (Reprospher 100), with a 120-min gradient from 5% to 75% acetonitrile in 0.1% formic acid and a flow rate of 250 nL/min. The mass spectrometer was operated in data-dependent acquisition mode with a top-ten method. Full-scan MS spectra were acquired with a target value of 3E5 and a resolution of 70,000, with a scan range from 300 to 1,750 m/z. HCD tandem MS/MS spectra were acquired with a target value of 1E5, a resolution of 17,500, and a normalized collision energy of 25%. All charges lower than two and higher than six were rejected, and all unknown charges were rejected. The underfill ratio was set to 0.1%, and a dynamic exclusion of 20 s was used. Alternatively, the underfill ratio was set to 1.0% with the dynamic exclusion time set to 10 s.

Data processing. MaxQuant version 1.4.1.2 was used to analyze all RAW data. The control-condition experiment was performed in sextuplicate and measured as 11 technical replicates. The heat shock-experiment was performed in triplicate and measured as 6 technical replicates. The MG-132 experiment was performed in triplicate and measured as 6 technical replicates. The PR-619 experiment was performed in quintuplicate and measured as 9 technical replicates. The heat shock-experiment was performed as a lysine-specific modification, with a monoisotopic mass of 471.20776, and and pyro-QQTGG were set as variable peptide modifications. QQTGG was set as a lysine acetylation, peptide N-terminal carbamylation, methionine oxidation, QQTGG modifications, lysine acetylation and serine-threonine-tyrosine phosphorylation were individually added as further variable modifications in separate searches. For pQQTGG, b5-pQQTGG, b4-pQQTG, b3-pQQT and b2-pQQ were accepted as diagnostic peaks. For pQQTGG, b5-pQQTGG, b4-pQQTG, b3-pQQT and b2-pQQ were accepted as diagnostic peaks. In addition to the above variable modifications, l-cysteine carboxamidomethylation was set as a fixed peptide modification. Protein N-terminal acetylation, peptide N-terminal carbamylation, methionine oxidation, QQTGG and pyro-QQTGG were set as variable peptide modifications. QQTGG was set as a lysine-specific modification, with a monoisotopic mass of 471.20776, and was not allowed to occur at the C-terminal ends of peptides. Pyro-QQTGG may spontaneously form out of the tryptic QQTGG remnant as a result of cyclization of the N-terminal glutamine. Pyro-QQTGG (pQQTGG) was set as a lysine-specific modification, with a monoisotopic mass of 454.18212, and was not allowed to occur at the C-terminal ends of peptides. In order to increase identification confidence, diagnostic peaks were searched within MS/MS spectra corresponding to SUMOylated peptides. To this end, candidate MS/MS spectra were searched for peaks corresponding in m/z to fragmentation of the tryptic QQTGG or pQQTGG remnant present on all SUMOylated peptides. For QQTGG, b5-QQTGG, b4-QQT, b3-QQT and b2-QQ were accepted as diagnostic peaks. For pQQTGG, b5-pQQTGG, b4-pQQTG, b3-pQQT and b2-pQQ were accepted as diagnostic peaks. In addition to the above variable modifications, l-cysteine acetylation and serine-threonine-tyrosine phosphorylation were individually added as further variable modifications in separate searches. For protein identification, peptides with all above variable modifications were accepted, and protein identification by at least one single SUMO-modified unique peptide was performed. Peptides were accepted with a minimum length of 6 aa, a maximum size of 5 kDa, and a maximum charge of six. The processed data were filtered by posterior error probability (PEP) to achieve a protein false discovery rate (FDR) of below 1% and a peptide-spectrum match FDR of below 1%, and in filtering by posterior error probability (PEP) to achieve a protein false discovery rate (FDR) of below 1% and a peptide-spectrum match FDR of below 1%, and in filtering by posterior error probability (PEP) to achieve a protein false discovery rate (FDR) of below 1% and a peptide-spectrum match FDR of below 1%, and in filtering by posterior error probability (PEP) to achieve a protein false discovery rate (FDR) of below 1% and a peptide-spectrum match FDR of below 1%.

Statistics. Statistical methods used in this manuscript are detailed in Supplementary Note.

Primary antibodies. Primary antibodies used in this study were mouse anti-SUMO-2 (ab81371, Abcam, 1:2,000), rabbit anti-SUMO-2 (raised against the C-terminal part of SUMO-2, 1:5,000)35, mouse anti-His(HIS-1, H-1029, Sigma, 1:2,500), mouse anti-HA(HA-11, MMS-101R, Sanbio,1:1,000), rabbit anti–SART-1 (raised against SART-1 peptides, 1:1,000)33, rabbit anti–histone H3 (4499S, Cell Signaling Technology, 1:500), rabbit anti-H3ac (06–599, Upstate, 1:2,500), rabbit anti-RNF216 (A304–111A, Bethyl, 1:2,500), rabbit anti–TCF12 (11822S, Cell Signaling Technology, 1:1,000), rabbit anti–WDR70 (A301–871A, Bethyl, 1:2,500), rabbit anti–FOXIM1 (C–20, sc–502, Santa Cruz, 1:1,000), mouse anti–HNRNPM (HL374, R3902, Sigma, 1:5,000), rabbit anti–RAD18 (A301–340A, Bethyl, 1:2,500), and rabbit anti–MCM10 (A300–131A, Bethyl, 1:2,500).

Validation of antibodies is provided on the manufacturers’ websites, in the cited references and in Antibodypedia.

Electrophoresis and immunoblotting. Protein samples were size-fractionated on Novex 4–12% Bis–Tris gradient gels with MOPS running buffer (Invitrogen). Size-separated proteins were transferred to Hybond-C membranes (Amersham Biosciences) with a subclass memory system (Invitrogen). Gels were Coomassie stained according to the manufacturer’s instructions (Invitrogen). Membranes were stained for total protein loading with 0.1% Protean-S in 5% acetic acid (Sigma). Membranes were blocked with PBS containing 0.1% Tween-20 (PBS-T) and 5% milk powder for 1 h. Subsequently, membranes were incubated with primary antibodies as indicated, in blocking solution. Incubation with primary antibody was performed overnight at 4 °C. Afterwards, membranes were washed three times with PBS-T and briefly blocked again with blocking solution. Next, membranes were incubated with secondary antibodies (donkey anti-rabbit HRP, 31458, Pierce, 1:2,500 and goat anti-mouse HRP, 31432, Pierce, 1:2,500) for 1 h, before being washed three times with PBS-T and twice with PBS. Membranes were then treated with ECL2 (Pierce) per the manufacturer’s instructions, and chemiluminescence was captured with BioMax XAR film (Kodak). A compilation of all uncropped images corresponding to all scans of gels, membranes and films displayed throughout this manuscript is available as Supplementary Data Set 1.

Microscopy. Cells were seeded on glass coverslips and fixed 24 h later for 10 min in 3.7% paraformaldehyde in PHEM buffer (60 mM PIPES, 25 mM HEPES, 10 mM EGTA, and 2 mM MgCl2, pH 6.9) at 37 °C. After being washed with PBS, cells were permeabilized with 0.1% Triton X-100 for 10 min, washed with PBS-T, and blocked with TBNB (100 mM Tris, pH 7.5, 150 mM NaCl, and 0.5% Blocking Reagent (Roche)) for 30 min. Cells were incubated with primary antibody as indicated, in TBNB for 1 h. Subsequently cells were washed five times with PBS-T and incubated with secondary antibodies (goat anti-mouse Alexa 594 (A-11005, Invitrogen, 1:500)) in TBNB for 1 h. Next, cells were washed five times with PBS-T and dehydrated with alcohol, embedded in Citifluor (Agar Scientific) containing 400 ng per microliter DAPI (Sigma) and sealed on the slides with nail varnish. Images were recorded on a Leica SP5 confocal microscope system with 488-nm and 561-nm lasers for excitation, and a 63× lens for magnification, and were analyzed with Leica confocal software.

48. Vellinga, J. et al. A system for efficient generation of adenovirus protein X-producing helper cell lines. J. Gene Med. 8, 147–154 (2006).
49. Tiscornia, G., Singer, O. & Verma, I.M. Production and purification of lentiviral vectors. Nat. Protoc. 1, 241–245 (2006).
50. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. 2, 1896–1906 (2007).
51. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. 10, 1794–1805 (2011).
52. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Methods 8, 1367–1372 (2008).
53. Vertegaal, A.C. et al. Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. Mol. Cell. Proteomics 5, 2298–2310 (2006).