Small ubiquitin-like modifier (SUMO) modification regulates numerous cellular processes. Unlike ubiquitin, detection of endogenous SUMOylated proteins is limited by the lack of naturally occurring protease sites in the C-terminal tail of SUMO proteins. Proteome-wide detection of SUMOylation sites on target proteins typically requires ectopic expression of mutant SUMOs with introduced tryptic sites. Here, we report a method for proteome-wide, site-level detection of endogenous SUMOylation that uses α-lytic protease, WaLP. WaLP digestion of SUMOylated proteins generates peptides containing SUMO-remnant diglycyl-lysine (KGG) at the site of SUMO modification. Using previously developed immuno-affinity isolation of KGG-containing peptides followed by mass spectrometry, we identified 1209 unique endogenous SUMO modification sites. We also demonstrate the impact of proteasome inhibition on ubiquitin and SUMO-modified proteomes using parallel quantitation of ubiquitylated and SUMOylated peptides. This methodological advancement enables determination of endogenous SUMOylated proteins under completely native conditions.
The family of small ubiquitin-like modifier (SUMO) proteins in humans includes four distinct genes with three types of members: SUMO1, SUMO2/3 (which differ by only three residues), and SUMO4. SUMO proteins regulate the function of various proteins by reversible covalent isopeptide bond attachment between the C terminus of SUMO and a free ε-amine group typically on lysine residues within target proteins, similar to ubiquitin (Ub). Ub conjugation mainly targets proteins for degradation by the proteasome, but has also been implicated in DNA repair, receptor signaling, and cell communication. The function of SUMO conjugation on target proteins is similarly diverse with SUMOylation catalyzing alteration of protein activity for targets involved in gene expression, DNA repair, nuclear import, heat shock, cell motility, and lipid metabolism. SUMO targets are generally low-abundance proteins, and the amount of the modification at steady state is also low. Given its importance in numerous cellular functions, several groups have developed proteomic methods for analysis of SUMOylated proteins.

Three general approaches have been previously utilized to isolate and identify the SUMOylated proteome. Ectopic expression of epitope-tagged SUMO followed by standard isolation techniques and mass spectrometry has been widely used in a variety of organisms. These approaches have yielded various maps of SUMO-interacting proteins but few sites of SUMOylation are identified by this approach and the extent to which exogenous expression of modified SUMOs alters substrate targeting is unknown. Immuno-affinity approaches utilizing antibodies that recognize endogenous SUMO2/3 or SUMO1 have been used to identify SUMO-interacting proteins under endogenous conditions. However, as with the epitope-tagging approach, few actual sites of SUMO modification were identified using this method. Therefore, methods that allow proteome-level identification of endogenous SUMOylation sites are needed.

A robust proteomic method has been developed to measure thousands of endogenous ubiquitylation sites. The method takes advantage of the C-terminal sequence of Ub (RGG) (Fig. 1a). When cleaved with trypsin, ubiquitylated substrate proteins will generate peptides containing a Ub-remnant diglycyl-lysine (KGG) that can be enriched using specific antibodies and identified by tandem mass spectrometry. Instead of the trypsin-friendly arginine residue preceding the C-terminal diglycine sequence observed in the processed Ub sequence, mature human SUMO paralogs have a threonine preceding the C-terminal diglycine sequence and no other trypptic cleavage sites near the C terminus (Fig. 1a). To overcome this problem, various schemes that introduce mutations within the C terminus of SUMO to render it more amenable to trypsin-based cleavage and identification by mass spectrometry have been developed for global profiling of SUMO attachment sites. Several groups have reported global profiling approaches in which mutant SUMOs with various affinity tags and protease recognition sites were introduced into cells. For example, Hendriks et al. introduced a lysine-deficient SUMO-3 with a C-terminal trypsin cleavage site, His19–SUMO-3 K0 Q87R. The SUMOylated proteins were enriched by immobilized metal affinity chromatography (IMAC) and digested with Lys-C. Peptides modified with SUMO were then purified again with IMAC and finally digested with trypsin to generate a five amino acid C-terminal SUMO-remnant modification. This group has compiled all available data resulting in 7327 SUMOylation sites in 3617 proteins. Other groups have used similar engineering approaches to express mutant SUMOs and identify up to 1000 unique SUMOylation sites upon induction of cell stress. While these methods have proved effective in mapping SUMOylation sites, they all require exogenous expression of mutant version of SUMO, which preclude analysis of SUMO-modification sites in native settings or from human tissues. Currently, no method exists for identifying endogenous SUMO sites on a global proteome scale without introduction of mutant SUMO.

We recently described the application of wild-type α-lytic protease (WaLP) to proteome digestion for shotgun proteomics. Although relatively relaxed specificity was observed, WaLP prefers to cleave after threonine residues and rarely cleaved after arginine. In addition, WaLP generates peptides of the same average length as trypsin despite its more relaxed substrate specificity. We show here that WaLP cleaves at the C-terminal TGG sequence (all SUMO paralogs) leaving a SUMO-remnant KGG at the position of SUMO attachment in target proteins. The resulting KGG-containing peptides can then be identified using methods already developed for profiling the Ub-modified proteome as described above (Fig. 1c). The method allows identification of SUMO attachment sites under completely native conditions using the Ub-profiling workflow by simply substituting WaLP for trypsin. The same sample can be subjected to analysis of both the Ub- and SUMO-modified proteomes simply by digesting the sample with either trypsin or WaLP, respectively. We demonstrate the effectiveness of this parallel identification of ubiquitylation and SUMOylation sites in cells treated with...
proteasome inhibitors. We provide the description a unique method of identifying proteins containing individual lysine residues that are modified by both SUMO and Ub from the same sample. This method can be simply applied to any sample, including human tissues samples, to identify endogenous SUMOylation sites.

Fig. 2 Summary of SUMOylation site identifications from this study and comparison with previous studies. a Venn diagram showing overlap between the 1209 SUMO sites identified in this study from both Hela and HCT116 cell lines (blue) and all SUMO modifications from either Hendriks et al. (red) or Phosphosite Plus (yellow). b Identified spectrum of a novel SUMO modification site identified in DNA-PK that would have resulted in a 57-amino-acid tryptic peptide. The sequence above the spectrum shows the tryptic cleavage sites in blue, the modification site is in red, and the matched sequence from the WaLP digest is underlined. c Identified spectrum of a previously uncharacterized SUMO modification site in Sp100 that would result in only a six-residue tryptic peptide.
Results

SUMOylation site profiling by WaLP digestion. Previous studies demonstrated the utility of WaLP digestion in shotgun proteomics platforms21. Our observation that threonine was among the preferred amino acids in the P1 position for WaLP digestion21 led us to hypothesize that digestion of SUMOylated proteins with WaLP, which should cleave after the threonine in the SUMO C-terminal sequence, TGG, would generate a SUMO-remnant diglycyl-lysine (KGG) at SUMOylation sites (Fig. 1a). The same workflow used for Ub-remnant profiling could then be used to globally profile SUMO attachment sites (Fig. 1b, c). Although WaLP can simply be substituted for trypsin during sample preparation, identification of non-tryptic peptides produced from WaLP digestion is challenging because search engines score on the basis of b and y ion series that are expected from tryptic peptides with a C-terminal positive charge. We previously demonstrated that identification of peptides arising from WaLP digestion benefits from electron transfer dissociation (ETD)22, whereas higher-energy collisional dissociation generates internal ions that can complicate peptide spectral matching21. WaLP cleaves after at least four different amino acids requiring the use of “no enzyme” specificity in database searches, which is challenging for many publicly available search algorithms. To facilitate mapping of SUMOylation sites using WaLP digestion, we used the MS-GF+ search engine whose scoring function can be trained using identifications from an initial search23,24.

Identification of novel SUMOylation sites using WaLP. As an initial test of the method, we digested lysates generated from human cell lines with WaLP and enriched for KGG-containing peptides using established protocols for Ub-modified peptide enrichment14. Subsequent analysis by mass spectrometry and database searching using MS-GF+ resulted in the identification of 2051 unique KGG-containing peptides, which were confidently localized (PTMProphet score >0.9) to 1209 unique sites (Supplementary Data 1). Comparison of the SUMOylation sites identified using our WaLP digestion approach with the SUMOylation sites amalgamated by Hendriks et al.17 revealed an overlap of only 30% (Fig. 2a). When compared to previously reported SUMOylation sites from Hendriks et al. (7710 sites)17, Phosphosite Plus (780 sites)25, and Uniprot (1863 sites)26, 826 were novel. Peptides containing SUMO remnants from WaLP digestion may correspond to sequences that are not covered by tryptic digestion due to the abundance or lack of nearby tryptic cleavage sites21. We found several examples of such cases. For example, we identified a novel SUMOylation site in DNA-PK (PRKDC, P78527), which would have been in a tryptic peptide of length 57 (Fig. 2b), and another in Sp100 (P23497), which would have been a tryptic peptide of only six amino acids (Fig. 2c).

Multiple SUMOylation sites were found in over 60% of the proteins identified, with 5% having more than six SUMOylation sites. The SUMOylation sites identified in this study correspond to the expected motifs as described by others. We found that 31% of the SUMO attachments occurred at the “forward” sequence motif (ΨKX[ED]I) (Supplementary Fig. 1a), 9% corresponded to the “inverted” motif ([EID]XK) (Supplementary Fig. 1b), and 60% did not correspond to either consensus but these sites were somewhat enriched in acidic residues (Supplementary Fig. 1c). Gene ontology term enrichment was performed to functionally annotate the SUMO-modified proteins. As reported previously, SUMOylation sites were primarily found in nuclear proteins and SUMOylated proteins were enriched for proteins involved in chromatin biology, RNA metabolism, and transcription (Supplementary Fig. 2).

It is possible that the same lysine within an individual protein can be either modified by Ub or SUMO and that SUMOylation may antagonize ubiquitylation27–29. In fact, 30% of the sites identified in this study were either previously reported to be ubiquitylated in the Phosphosite Plus database or were found in our experiments to be ubiquitylated. This overlap is similar to that previously reported by Hendriks et al.18.

Validation of SUMOylation sites using in vitro deSUMOylation. Our previous studies on WaLP digestion specificity indicated that WaLP rarely cleave after arginine residues (Fig. 1b) leads to the possibility of using WaLP to generate diGlycine-remnant peptides from SUMO-modified proteins and not Ub-modified proteins. However, when using WaLP to isolate and identify SUMOylated peptides in this study, we occasionally identified peptides whose MS/MS spectra matched best to peptides that resulted from cleavage after arginine either on the N or C terminus. These peptides were rare, with percentages varying from 3–8%, which we attribute at least partly to our inability to completely remove the trypsin used to detach cells during scale-up. Trypsin was avoided in cell harvesting, scraping was used instead. Even still, the fact that the P1 Arg was usually present in a sequence that had multiple positively charged residues in a row led us to question whether WaLP digestion could occasionally result in the generation of a KGG-peptide from a ubiquitylated protein. To evaluate this possibility and to validate our approach, native cell lysates were either untreated or treated with recombinant deSUMOylating enzymes, SENP1 and SENP2, prior to digestion with either trypsin or WaLP. Western blot analysis revealed that in vitro treatment with SENP1/2 resulted in a robust reduction of SUMO1- and SUMO2/3-modified proteins while leaving ubiquitylated proteins unaltered (Fig. 3a). Digestion of SENP1/2-treated lysates with trypsin or WaLP and subsequent analysis by mass spectrometry revealed a dramatic reduction in the abundance of KGG-modified peptides from WaLP-digested samples compared to untreated samples (Fig. 3b). Importantly, the abundance of KGG-peptides resulting from trypsin cleavage, which would arise from ubiquitylated substrates, was unaltered by SENP1/2 treatment (Fig. 3b). The observation that 88% of sites identified after WaLP digestion were decreased at least twofold upon SENP1/2 treatment validated our approach. Conversely, treatment of lysates with a promiscuous deubiquitylating enzyme, Usp2cc, would result in a specific decrease in the amount of a KGG-peptide from a ubiquitylated site without altering SUMOylated peptides. Consistent with our prediction, less than 2% of KGG-peptides generated from WaLP-digested cells decreased upon Usp2cc treatment (a similar percentage to the FDR), whereas 97% of the KGG-containing peptides generated from trypsin digestion decreased after Usp2cc treatment (Fig. 3c). Taken together, we conclude that WaLP digestion and subsequent immuno-affinity enrichment of KGG-modified peptides specifically identifies endogenous SUMOylated peptides.

Parallel Ub and SUMO site identification. A unique feature of our method is the ability to identify ubiquitylation and SUMOylation sites in parallel from the same cell or tissue lysate. The lysate can be simply split in half, with one half digested with WaLP and the other half digested with trypsin (Fig. 1c, d). The subsequent immuno-affinity enrichment steps are identical, although mass spectrometry and data processing were optimized for tryptic peptides for Ub and non-tryptic peptides for SUMO. As is already well established, SUMOylation is a lower abundance modification than ubiquitylation, and indeed, we identified 6472 Ub sites in the same samples in which 1209 SUMO sites were identified. Qualitative comparisons of the data revealed all
possible types of dual modification scenarios. We found 2713 proteins that were ubiquitylated, 768 proteins that were SUMOylated, and 407 proteins that had both SUMO and Ub. Some had a large number of ubiquitylated lysines and few SUMOylated lysines while others had larger numbers of SUMO-modified lysines and few ubiquitylated sites. Those for which we observed both modifications mostly carried those modifications on different residues. We observed 243 lysines that were found to be Ub modified in the trypsin-digested samples and SUMO modified in the WaLP-digested samples. These results indicate that our method can identify proteins that are modified by both SUMO and Ub on the same lysine residues in the same sample.

**Parallel quantification of Ub and SUMO sites.** Previous studies indicated that SUMO modification could be stimulated by cell stresses such as heat shock and proteasome inhibition\(^ {30,31} \). Using SILAC-based quantitative proteomics, and our method allows for the parallel capture, identification, and quantification of SUMO- and Ub-modified proteins in response to various cell perturbations. As it is well established that proteasome inhibition results in global alteration of the Ub-modified proteome, we evaluated the dynamic response of both the Ub- and SUMO-modified proteome to proteasome inhibition. Metabolically labeled cells were treated with MG132 and mixed with unlabeled cells prior to cell lysis. Again, the cell lysates were split and half was digested with trypsin for Ub analysis and the other half was digested with WaLP for SUMO analysis. We identified and quantified 330 unique SUMOylated sites and 2621 unique ubiquitylation sites from biological replicate samples. The results for both experiments were normally distributed (Fig. 4) and the results from the Ub analysis were comparable to previously published data in
which proteasome inhibitors were utilized\cite{14,32}. For the SUMO analysis, 30 sites increased in abundance by at least twofold and 72 sites decreased in abundance by at least twofold upon MG132 treatment\cite{18}. Changes in modification abundance are not likely due to changes in protein levels as protein levels were not found to change significantly during a 4-h MG132 treatment\cite{14}. We observed 103 proteins for which both SUMOylation and ubiquitylation could be quantified. These proteins had a total of 591 modified lysines. Of the total 591 modified lysines, 53 were found to be both SUMOylated and ubiquitylated.

The proteins with the most abundant sites that were both SUMOylated and ubiquitylated were SUMO and Ub themselves, so we analyzed these to see if any sites were reciprocally regulated. It is important to note that we observed multiple modified peptides with more than one SUMO modification on SUMO proteins only in the WaLP digest. This indicates the presence of SUMO molecules that are simultaneously SUMOylated at different lysine residues and may have a branched-chain architecture. Interestingly, SUMO-modified ubiquitin was only observed on single lysine residues, suggesting that multiply SUMO-modified ubiquitin is a rare or nonexistent event. Conversely, peptides with more than one Ub modification on ubiquitin were observed only in the trypsin digest. These results are consistent with the presence of poly-SUMO (detected by

---

**Fig. 4** Parallel identification and quantitation of the SUMO and Ub-modified proteome upon MG132 treatment of HCT116 cells. Plot of normalized log(2) SILAC ratios (heavy/light) for all quantified unique Ub modification sites **a** or SUMO modification sites **b** upon MG132 treatment. The left graphs depict the ordered distribution of the SILAC ratios and right graphs depict the histogram of SILAC ratios. **c** Comparison of the SILAC ratios from biological replicates for ubiquitylation sites (black symbols) or SUMOylation sites (green symbols).
WaLP) and poly-Ub (detected by trypsin) but only mono-modifications of the opposite protein. We quantified ubiquitylation and SUMOylation of three lysine residues in SUMO2/3 upon MG132 treatment (Fig. 5a). Neither ubiquitylation nor SUMOylation of lysines 33 or 45 in SUMO2/3 were dramatically altered upon proteasome inhibition. However, ubiquitylation of lysine 11 on SUMO2/3 increased, whereas SUMOylation on this site decreased upon MG132 treatment (Fig. 5a). Cross-modification of SUMO2/3 by ubiquitin has been previously observed, and consistent with these results, ubiquitylation of SUMO2/3 on lysine 11 increases upon proteasome inhibition.

For ubiquitin, we were able to quantify Ub and SUMO modification of lysines 27, 29, 33, 48, and 63 (Fig. 5b). Both SUMOylation and ubiquitylation of ubiquitin responded similarly to MG132 treatment (Fig. 5b). Both SUMO and Ub modification of lysines 48, 29, and 33 increased in abundance upon MG132 treatment. Conversely, both SUMO and Ub modifications on lysine 63 of Ub decreased in abundance in the MG132-treated cells. Similarly, we observed cross-modification of another ubiquitin-like (Ubl) protein, Nedd8. Lysines 11 and 48 on Nedd8 were observed to be modified by both Ub and SUMO and the extent of modification was largely unperturbed by proteasome inhibition (Fig. 5c). Trypsin digestion followed by KGG immuno-affinity enrichment cannot distinguish between ubiquitylation or neddylation due to the presence of an arginine preceding the C-terminal diGlycine in both Ub and Nedd8. As such, it is possible...
that Nedd8 may be poly-neddylated at these sites rather than ubiquitylated. Consistent with this observation, DCN1 was observed to be diGlycine-modified at a lysine residue previously characterized to be a neddylation site and this same lysine residue was SUMOylated (Fig. 5c). Interestingly, the abundance of both SUMOylated and neddylated DCN1 decreased upon proteasome inhibition. These observations suggest that cross-modification of Ub proteins is prevalent and may possibly be antagonistic modifications.

Separate from the highly abundant Ub proteins, we observed a few examples of proteins that were modified by both SUMO and Ub on the same lysine residue with differential responses to proteasome inhibition. For example, we quantified two SUMOylation and nine ubiquitylation sites on annexin A1 (Fig. 5d). We observed SUMOylation at Lys 147, which has not been previously reported, and at Lys 312, a known SUMOylation and ubiquitylation site. This C-terminal lysine appears to experience reciprocal regulation by MG132 as the abundance of the SUMO-modified form was significantly decreased, whereas the ubiquitylation was significantly increased (Fig. 5d). Further, lysine 26 in caveolin-1 and lysine 140 in the large ribosomal protein 15 were modified by both Ub and SUMO and these modifications displayed divergent abundance alterations upon proteasome inhibition (Fig. 5e, f). These results indicate that reciprocal modification by SUMO and Ub on the same lysine within a target protein may be a relatively rare event, but that it clearly does occur at some sites in the proteome and that these modification events may impart differential functional outputs for the substrate proteins. These results also point to a unique strength of our method in that it allows for parallel mapping of regulated ubiquitylation and SUMOylation sites from a single sample.

Identification of SUMOylation in tissue samples. One advantage of our method is that it allows for the identification of endogenous SUMOylation sites from native tissues without exogenous SUMO protein expression. To establish this application of our method, we processed murine tissue from brain, heart, muscle, and liver to establish this application of our method, we processed murine SUMOylated proteins in vivo. We identified 144 unique SUMOylation sites across the four tissues (Supplementary Data 2). Overall, muscle and liver had the most similar SUMOylated proteins, sharing ~70% of the total sites observed in those tissues, whereas brain had the most unique SUMOylated proteins that were not found in any other tissue type. We identified the well-characterized SUMOylated protein RANGAP1 in all tissues validating our approach. Interestingly, we again identified ubiquitin as a SUMO-modified protein with lysines 48 and 63 serving as the SUMOylation sites in all tissues. This result suggests that ubiquitin is SUMOylated at critical lysine residues in vivo and validates that our approach can be successfully applied toward the identification of endogenous SUMOylation sites in tissues.

Discussion

We set out to develop a method for global profiling of native SUMOylation events by taking advantage of the propensity of WaLP for cleavage after threonine. By simply substituting WaLP for trypsin, it was possible to immunopurify and identify a large number of KGG-containing peptides corresponding to SUMO remnants. A large number of the identified sites corresponded to as-yet-unreported SUMOylation events. Several reasons could explain the large number of new sites identified. First, as we reported previously, the orthogonal specificity of WaLP allows cleavage of proteins at sites that may not be accessible to trypsin. Second, although previous studies attempted to achieve minimal expression of their mutant SUMO construct, it is possible that slight overexpression of SUMO or the presence of mutant sequences could cause unnatural SUMO attachment. Third, our method does not differentiate between SUMO1–4, whereas Hendriks et al. examined only SUMOylation sites utilizing SUMO-3 attachment. One caveat is that WaLP also cleaves after Leu and to some extent Ile, so the method does not distinguish between SUMOylated, FatIylated, or FubIylated proteins. However, the observation that 88% of WaLP diGlycine proteins were reduced upon SENP1/2 treatment argues that the vast majority of observed KGG sites arise from SUMOylation. Additionally, analysis of the observed sites recapitulates previous reports of the expected motifs of SUMOylation and GO term enrichment.

A powerful advantage of our method is that it allows for simultaneous determination of ubiquitylation and SUMOylation in the same sample. The same population of cells or tissue can be subjected to analysis of both Ub attachment and SUMO attachment simply by splitting the sample in two and digesting half with trypsin and the other half with WaLP. The samples can then be processed in parallel to immunopurify peptides for the presence of the KGG modification, and sequence by mass spectrometry. For the mass spectrometry, it is best to use optimized ionization approaches and data analysis tailored to the non-tryptic WaLP peptides. Although the commercial antibody used in this study was developed for enrichment of KGG-peptides from Ub modification, it appears to also efficiently immunopurify KGG-peptides from SUMO modification. Finally, the method allows for identification of SUMOylated proteins under native conditions including from tissue samples.

Methods

WaLP. WaLP was expressed from Lysobacter enzymogenes type 495 (ATCC) using Bachovchin’s media supplemented with MEM vitamins and 60 g l−1 sucrose. L. enz. was grown at 30 °C with shaking at 100 rpm for 3 days. WaLP was purified from the culture supernatant as described previously. Brieﬂy, the protease was captured from the supernatant by batch binding on SP-sephrose, which is washed extensively and then eluted with high salt glycine buffer. After buffer exchange to pH 7.2, the enzyme was loaded onto an FPLC monoS 10/10 column eluted with a gradient of 10 mM NaHPO4, pH 7.2 to the same buffer containing 250 mM sodium acetate over 1 h.

SENP1 and SENP2 desSUMOylation. To verify the in vitro desSUMOylation and deubiquitylation assay, untreated HCT116 cells were harvested and lysed with denaturing lysis buffer (8 M urea, 150 mM NaCl, 50 mM Tris pH 7.8, 1 mM sodium orthovanadate, 1 mM NaF, 1 mM sodium 2-glycerophosphate, protease inhibitor tablet (Roche), 5 mM N-Ethylmaleimide (NEM, made fresh in methanol). Lysates were sonicated and centrifuged at 20,000×g for 1 h and the supernatant was then subjected to a deSUMOylation assay. The antibodies used for immunoblotting were against ubiquitin (MAB1510, EMD Millipore, 1:1000 dilution), SUMO1 (4930, Cell Signaling Technology, 1:1000 dilution), and sequence by mass spectrometry.

Hela cells were cultured in MEM media with 10% fetal bovine serum and 1% penicillin/streptomycin at 37 °C with 5% CO2. Heat shock was performed at 43°C for 1 h, then Hela cells were washed with cold PBS, and harvested with 8 M urea lysis buffer (50 mM Tris pH 8.0, 8 M urea, 1 mM vanadate, 2.5 mM sodium pyrophosphate, 1 mM beta-glycerophosphate). Extracts were sonicated and centrifuged at 20,000×g for 15 min to remove insoluble material. Protein concentrations were measured by Bradford Assay.
PA) at 10 units per milligram of lysate protein and incubating with lysate for overnight at room temperature. In vitro de-ubiquitination reaction was performed by adding USP2c (Enzo Life Sciences) at an enzyme-to-substrate ratio of 1:500 and incubating at RT for 4 h. Untreated control lysate samples were also prepared. Efficiency of SUMoylation cleavage by SUMO proteases and ubiquitin cleavage by USP2c was monitored by using specific antibodies for SUMO1 (Upstate, 9390), SUMO2/3 (9471), and ubiquitin (9393) from Cell Signalling Technology. The lysates were digested by WAP or trypsin, an enzyme-to-substrate ratio of 1:100 at 37 °C overnight with slow rotation. Digestion was stopped by adding 20% TFA solution to final TFA concentration of 1%. Peptides were then subject to C18 cleaning by Sep-Pak cartridges (Waters Corp.) and lyophilized.

**Enrichment of KGG-containing peptides.** For each 5–7 mg peptide sample, 10 μl of Ub-concanavalin A (20 μl of slurry, Cell Signaling Technology) was used to immunoisolate the KGG-containing peptides according to the CST protocols. First, the samples were diluted fourfold using the Ubiscan beads at 2 h with rocking at 4 °C. After isolating the beads by centrifugation at 1000×g for 1 min, the beads were washed two times with 1× IAP buffer, then four times with HPLC-grade water. The peptides were eluted in two steps. The beads were incubated at room temperature with 55 μl of 0.15% TFA for 10 min, centrifuged at 3500×g for 1 min, and the supernatant was collected. Then the beads were incubated for an additional 10 min with follow by 45 μl of 0.15% TFA for 10 min, centrifuged at 3500×g for 1 min, and the supernatant was combined with the first elution. The samples were analyzed by nLC-MS/MS on an orbitrap Fusion or LUMOS for WAP-digested samples, and on a Q-Exactive for Ub-digested samples.

**M5 analysis of in vitro deSUMOylation experiment.** Immunoprecipitated peptides were reduced in 0.125% formic acid and analyzed by an Orbitrap Fusion Lumos to determine the mass spectrometer (Thermo) coupled to an EASY-nLC 1200 (Thermo). Each sample was split and analytical replicate injections were run to increase the number of identifications and provide metrics for analytical reproducibility of the method. Standard peptide mix (MassPREP® Protein Digestion Standard Mix 1, Waters) was spiked in each sample vial in a total quantity of 100 fmol per injection prior to LC-MS/MS analysis. The sample was loaded onto an EASY-Spray® analytical column (PepMap® C18, 75 μm × 50 cm, C18, 2 μm, 100 Å, Thermo), which was connected to an EASY-Spray® ionization source (Thermo). The column was heated to 45 °C for all runs. Mobile phase solvent A was composed of 0.1% formic acid and water. Mobile phase solvent B was composed of 0.1% formic acid, 93.5% acetonitrile, and water. Peptides were separated using a gradient from 5% B to 32% B over 90 min and continued to 53% B over 5 min at a constant flow rate of 300 nl min⁻¹. Full MS scans were obtained with a range of m/z 300–1500 at a mass resolution of 120,000 (m/z 200), with an AGC target value of 4.0E5 and maximum injection time of 50 ms. MS/MS data were acquired for MS/MS analysis, and ETD in the ion trap with the rapid scan rate, resulting in two separate spectra for precursors with intensity greater than 5000 were fragmented sequentially with CID and fragmentation of the CID fragment ions was achieved with ETD in the ion trap with the rapid scan rate, resulting in two separate spectra for each selected precursor ion.

**Preparation of tissue samples.** Murine tissue samples such as brain, heart, and muscle, and liver were obtained from mature BALB/c mice (Cell Signaling Technology). Tissue was homogenized in 8 M urea lysis buffer (50 mM Tris pH 8.0, 8 M urea, 1 mM β-mercaptoethanol, 10 mM β-mercaptoethanol, 10 mM EDTA, plus protease inhibitors). Lysate was reduced by 4.5 mM DTT for 30 min at 55 °C. Reduced lysate was alkylated with 10 mM iodoacetamide for 15 min at 25 °C in the dark. Sample was diluted fourfold with 50 mM Tris, pH 8, and digested overnight with WAP at a ratio of 1:100 at 37 °C overnight with slow rotation. Digested peptide lysate was precipitated by centrifugation at 20,000×g at 4 °C for 1 h. The pellets were washed two times with 1× IAP buffer, then four times with HPLC-grade water. The peptides were eluted in two steps. The beads were incubated at room temperature with 55 μl of 0.15% TFA for 10 min, centrifuged at 3500×g for 1 min, and the supernatant was collected. Then the beads were incubated for an additional 10 min with follow by 45 μl of 0.15% TFA for 10 min, centrifuged at 3500×g for 1 min, and the supernatant was combined with the first elution. The samples were analyzed by nLC-MS/MS on an orbitrap Fusion or LUMOS for WAP-digested samples, and on a Q-Exactive for Ub-digested samples.
when available from the Fusion data. The .mzid output from the MS-GF+ output, before the 1% FDR background.

17. Hendriks, I. A. & Vertegaal, A. C. A comprehensive compilation of SUMO protein sites with localization scores below 0.9 were removed. The log(2) of each ratio was computed. The log(2) of localization sites with localization scores in response to changing.

18. Hendriks, I. A. et al. Uncovering global SUMOylation signaling networks in a site-specific manner. Nat. Struct. Mol. Biol. 21, 927–936 (2014).

19. Lamelatte, F. et al. Large-scale analysis of lysine SUMOylation by SUMO remnant immunofluorescence profiling. Nat. Commun. 5, 5409 (2014).

20. Tammalsu, T. et al. Proteome-wide identification of SUMO2 modification sites. Sci. Signal. 7, rs2.1–10 (2014).

21. Meyer, J. G. et al. Expanding proteome coverage with orthogonal-specificity α-lytic proteases. Mol. Cell. Proteomics 13, 823–835 (2014).

22. Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J. & Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc. Natl. Acad. Sci. USA 101, 9528–9533 (2004).

23. Gauthals, A., Boucher, C. & Bandeira, N. The generating function approach for peptide identification in spectral networks. J. Comput. Mol. Biol. 22, 353–366 (2014).

24. Klar, S. I. & Pevzner, P. A. Next generation sequencing makes progress towards a universal data search tool for proteomics. Nat. Commun. 3, 1–10 (2012).

25. Hornbeck, P. V. et al. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic Acids Res. 40, D261–D270 (2012). (Database issue).

26. Consortium, T. U. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 45, D158–D169 (2017).

27. Meulmeester, E., Kunze, M., Hsiao, H. H., Urlaub, H. & Melchior, F. Mechanism and consequences for paralog-specified sumoylation of ubiquitin-specific protease 23. Mol. Cell 30, 610–619 (2008).

28. Deusch, A., Bochem-Chosa, R., Marfany, G. & Gonzalez-Duarte, R. & KappalAlpha inhibits NF-kappaB activation. Mol. Cell. 2, 233–239 (1998).

29. Uzunova, K. et al. Ubiquitin-dependent proteolytic control of SUMO conjugates. J. Biol. Chem. 282, 34167–34175 (2007).

30. Schimmel, J. et al. The ubiquitin-proteasome system is a key component of the SUMO-2/3 cycle. Mol. Cell Proteomics 7, 2107–2122 (2008).

31. Gendron, J. M. et al. Using the ubiquitin-modified proteome to monitor distinct and spatially restricted protein homeostasis dysfunction. Mol. Cell Proteomics 15, 2576–2593 (2016).

32. Ma, J. C. & Agard, D. A. Kinetic and structural characterization of mutations of glycine 216 in alpha-lytic protease: a new target for engineering substrate specificity. J. Mol. Biol. 254, 720–736 (1995).

33. Olsen, J. V. et al. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. Mol. Cell Proteomics 4, 2010–2021 (2005).

34. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to tandem mass spectral data of peptides with amino acid sequences in a protein database. J. Am. Soc. Mass Spectrom. 5, 976–989 (1994).

35. Huttlin, E. L. et al. A tissue-specific atlas of mouse protein phosphorylation and expression. Cell 143, 1174–1189 (2010).

36. Villen, J., Beausoleil, S. A., Gerber, S. A. & Gygi, S. P. Large-scale phosphorylation analysis of mouse liver. Proc. Natl Acad. Sci. USA 104, 1488–1493 (2007).

37. Villen, J. & Gygi, S. P. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat. Protoc. 3, 1630–1638 (2008).

38. Kessler, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 24, 2534–2536 (2008).

39. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria, 2011) ISBN 3-900051-07-0. http://www.R-project.org/.

40. Shetyenyberg, D. et al. PTMProphet: TPP software for validation of modified site locations on post-translationally modified peptides. 60th American Society for Mass Spectrometry (ASMS) Annual Conference (ASMS, Vancouver, Canada, 2012).

41. Anwyl, A., Nathaniel, J. A., Fobs, J. G. & Shteynberg, D. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. 75, 7538–7542 (2002).

42. Shteynberg, D. et al. PhoIPP: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. Mol. Cell Proteomics 10, M111.019760 (2011).

43. Shetyenyberg, D. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. 74, 5383–5392 (2002).

44. Shteynberg, D. et al. PhoIPP: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. Mol. Cell Proteomics 10, M111.019760 (2011).

45. Klar, S. I. & Pevzner, P. A. Next generation sequencing makes progress towards a universal data search tool for proteomics. Mol. Cell Proteomics 12, 2383–2393 (2013).

Acknowledgements

This work was supported by NSF MCB 1244506 (I.A.K.), NIH T32 EB009380 (J.G.M.), NIH T32 GM 008326 (R.J.L.), New Scholar awards from the Sidney Kimmel Foundation
for Cancer Research and the Ellison Medical Foundation, a Hellman Fellowship, and a NIH New Innovator Award (DP2 GM119132) (E.J.B.).

Author contributions
E.A.K. conceived the study. R.J.L., A.S.A., and E.J.B. performed all mass spectrometry-based methods and analysis. H.G., Y.Z., and M.L. performed the in vitro SENP1/2 treatment and subsequent KGG enrichment and LC-MS/MS analysis. R.J.L., E.A.K., and E.J.B. performed data analysis. J.G.M. and R.J.L. developed custom scripts to facilitate computational analysis. K.R.C. assisted in spectra validation and computational support. The manuscript was written by E.J.B. and E.A.K. with assistance from R.J.L.

Additional information
Supplementary Information accompanies this paper at doi:10.1038/s41467-017-01271-3.

Competing interests: H.G. and Y.Z. are employees of Cell Signaling Technology. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017