Mapping of SUMO sites and analysis of SUMOylation changes induced by external stimuli

Francis Impens[a,b,c], Lilliana Radoshevich[a,b,c], Pascale Cossart[a,b,c,1], and David Ribet[a,b,c,1]

[a]Unité des Interactions Bactéries-Cellules, Institut Pasteur, F-75015 Paris, France; [b]Institut National de la Santé et de la Recherche Médicale, Unité 604, F-75015 Paris, France; and [c]Institut National de la Recherche Agronomique, Unité sous-contrat 2020, F-75015 Paris, France

Contributed by Pascale Cossart, July 22, 2014 (sent for review May 28, 2014)

SUMOylation is an essential ubiquitin-like modification involved in important biological processes in eukaryotic cells. Identification of small ubiquitin-related modifier (SUMO)-conjugated residues in proteins is critical for understanding the role of SUMOylation but remains experimentally challenging. We have set up a powerful and high-throughput method combining quantitative proteomics and peptide immunocapture to map SUMOylation sites and have analyzed changes in SUMOylation in response to stimuli. With this technique we identified 295 SUMO1 and 167 SUMO2 sites on endogenous substrates of human cells. We further used this strategy to characterize changes in SUMOylation induced by listeriolysin O, a bacterial toxin that impairs the host cell SUMOylation machinery, and identified several classes of host proteins specifically deSUMOylated in response to this toxin. Our approach constitutes an unprecedented tool, broadly applicable to various SUMO-regulated cellular processes in health and disease.

posttranslational modification | *Listeria* | cortactin | actin | anillin

Posttranslational modifications (PTMs) are key mechanisms used by both prokaryotes and eukaryotes to regulate protein activity specifically, locally, and temporally. Ubiquitin and ubiquitin-like proteins (UBLs) constitute a specific class of small protein modifiers that can be covalently attached to a target protein via the formation of an isopeptide bond in a reversible manner. Small ubiquitin-related modifier (SUMO), one of these UBLs, is an essential PTM in eukaryotic cells that is involved in various cellular functions including gene expression regulation, DNA repair, intracellular transport, and response to viral and bacterial infections (1–5). The human genome encodes three different functional SUMO isoforms (SUMO1, SUMO2, and SUMO3) that are conjugated to distinct but overlapping sets of target proteins (1, 2, 6). Conjugation of SUMO to its targets in humans requires an E1-activating enzyme (the SAE1/SAE2 heterodimer), an E2-conjugating enzyme (Ubc9), and several E3 SUMO enzymes. Once conjugated to its target, SUMO can be deconjugated by several different SUMO isopeptidases that tightly regulate the SUMOylation levels of proteins (7).

Since the discovery of SUMO two decades ago, much effort has been dedicated to the identification of SUMO-conjugated proteins in different organisms including yeast, plants, and mammals (8). However, isolation of SUMOylated proteins has proven to be challenging. Indeed, for most SUMO substrates, only a small proportion of the total amount of protein is SUMO-modified. In addition, the high activity of SUMO isopeptidases in cell lysates results in the rapid loss of SUMO conjugation in the absence of appropriate inhibitors. Thus, the most common approach used to isolate SUMOylated proteins is based on the expression of His-tagged versions of SUMO allowing the purification of SUMO-conjugated proteins by nickel chromatography under denaturing conditions (8, 9). Denaturing conditions inactivate SUMO isopeptidases and also prevent contamination by proteins interacting noncovalently with SUMO via specific domains such as SUMO-interacting motifs (SIMs) (2). Once SUMOylated proteins have been isolated, their analysis by mass spectrometry (MS) has been widely used to identify SUMO-modified proteins and, albeit less successfully, SUMO-conjugation sites.

Mapping the exact lysine residue to which SUMO is attached in modified proteins is a critical step to get further insight into the function of SUMOylation. Indeed, the identification of SUMO sites allows the generation of non-SUMOylatable mutants and the study of associated phenotypes. Identification of SUMO sites by MS is not straightforward (8). Unlike ubiquitin, which leaves a small diglycine (GG) signature tag on the modified lysine residue after trypsin digestion, SUMO leaves a larger signature that severely hampers the identification of modified peptides.

In addition to the identification of the SUMO site per se, a comparison of the SUMOylation status of sites in different cell-growth conditions is critical for better characterizing the biological implications of SUMOylation. For example, analysis of SUMOylation changes induced after heat shock, arsenic treatment, inhibition of the proteasome, or during the cell cycle has led to numerous insights into the role of SUMOylation in cell physiology (refs. 10–14 and reviewed in ref. 2). Here, we devised a performant approach which combines the use of SUMO variants, peptide immunocapture, and quantitative proteomics for high-throughput identification of SUMO sites. We then show that our approach is able to characterize global changes in the cell SUMOylome in response to a given stimulus, such as exposure to a bacterial toxin, listeriolysin O (LLO).

Results

A Proteomics-Based Strategy to Map SUMO-Modified Lysines. In contrast to ubiquitylated proteins, SUMOylated proteins upon trypsin digestion lead to large signature tags (19 or 32 amino acids for human SUMO1 and SUMO2/3, respectively) on

Significance

Small ubiquitin-related modifier (SUMO) is a posttranslational modification essential for many functions in eukaryotic cells. A better understanding of the role of this ubiquitin-like modification, identification of proteins modified by SUMO, and knowledge of the exact sites of SUMO conjugation are critical but remain experimentally challenging. We have developed an innovative proteomic strategy allowing proteome-wide identification of SUMOylation sites and quantification of cell SUMOylation changes in response to diverse stimuli. Identification of yet unknown SUMO targets and characterization of SUMOylome alterations in response to environmental stresses, drugs, toxins, or bacterial and viral infections will help decipher previously unidentified roles of SUMOylation in cell physiology and disease.

Author contributions: F.I., P.C., and D.R. designed research; F.I., L.R., and D.R. performed research; F.I., L.R., P.C., and D.R. analyzed data; and F.I., P.C., and D.R. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The mass spectrometry proteomics data reported in this article have been deposited in the ProteomeXchange Consortium, http://proteomecentral.proteomexchange.org [PRoteomics IDEntifications (PRIDE) partner repository dataset identifier PXD000459].

[1]To whom correspondence may be addressed. Email: pcossart@pasteur.fr or david.ribet@pasteur.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1413825111/-/DCSupplemental.

peptides containing the modified lysine residue. These tags generate complex ion patterns during tandem mass spectrometry (MS/MS) fragmentation of the peptides, thus preventing their straightforward identification by common search algorithms. To circumvent this problem, we generated variants of mature human SUMO1 and SUMO2 with one arginine introduced immediately before the C-terminal GG motif (SUMO1 T95R and SUMO2 T91R, respectively) (Fig. 1*A*), thereby mimicking the sequence of human ubiquitin. Trypsin digestion of proteins modified by these SUMO variants generates SUMO-modified peptides with a GG tag easily identifiable by classical LC-MS/MS, as previously described for other similar SUMO variants (Fig. 1*B*) (11, 14–17). We also tagged these SUMO variants with 6xHis stretches, thereby allowing affinity purification of SUMOylated proteins from cell lysates under denaturing conditions (9).

To verify that the SUMO1 T95R and SUMO2 T91R variants behave similarly to their wild-type counterparts, we transfected HeLa cells with wild-type or variant His_6-SUMO1 and His_6-SUMO2 and pulled down SUMOylated proteins from cell lysates using nickel chromatography. Immunoblot analysis of SUMOylation patterns revealed a slight decrease in the global intensity of proteins SUMOylated by SUMO variants as compared with wild type (Fig. 1*C*); this result may indicate that these SUMO variants are conjugated less efficiently by the SUMOylation machinery. Nevertheless, the relative distribution of SUMOylated proteins in these patterns (i.e., the number and size of observed bands) does not differ greatly between wild type and SUMO variants, strongly suggesting that the sets of proteins conjugated by these different SUMO forms are similar. We further verified that the SUMOylation of a known SUMO target, RanGAP1 [Ras-related nuclear protein (Ran) GTPase-activating protein 1], was similar in wild-type and variant SUMOs (Fig. 1*D*). Strikingly, the isoform preference observed with wild-type SUMO (i.e., preferential modification of RanGAP1 by SUMO1 compared with SUMO2) (7) also was observed with the SUMO1 T95R and SUMO2 T91R variants (Fig. 1*D*). Finally, we tested if the interaction of SUMO1 T95R and SUMO2 T91R with SIMs was similar to that of their wild-type counterparts (2). We used an established assay based on split-luciferase complementation to detect noncovalent interactions between SUMO and SIMs from two different proteins, Daxx (Death domain-associated protein) and PIAS2 (Protein inhibitor of activated STAT 2) (18). We observed that wild-type and variant SUMO interact similarly with the different SIMs tested in this assay, strongly suggesting that the introduced mutations in SUMO1 and SUMO2 do not alter their ability to be recognized by SIMs (Fig. S1).

Trypsin digestion of proteins after His purification of SUMOylated proteins generates a mixture of SUMO-modified and nonmodified peptides. Under these conditions, identification of SUMO-modified peptides is challenging, because they represent a very small fraction of the total amount of peptides. Indeed, after direct MS analysis of His–pulled-down samples, we were able to identify only very few GG-modified peptides (<0.1% of total identified peptides). We thus added an additional enrichment step for GG-modified peptides to our protocol by taking advantage of antibodies directed against GG-modified lysines (anti–K-ε-GG antibodies). Although these antibodies initially were used to capture and enrich GG-modified peptides from trypsinized ubiquitylated proteins (19–21), we decided to broaden their application to the study of SUMOylation sites by combining them with the aforementioned SUMO variants.

To rule out the possibility that the GG-modified peptides identified in our study came from endogenous ubiquitin-, Nedd8 (Neural precursor cell expressed developmentally down-regulated protein 8)-, or ISG15 (Interferon-stimulated gene 15)-modified proteins (because these three modifiers also leave GG-tags on modified proteins), we used stable isotope labeling by amino acids in cell culture (SILAC) to compare peptides derived from cells expressing wild-type or variant SUMO. SILAC (22) allows differential isotope labeling of proteins during cell culture by metabolic incorporation of essential amino acids (predominantly lysine and arginine) that carry light or heavy isotopes. After mixing light- and heavy-labeled cell lysates, isolated SUMO-targets are subjected to trypsin digestion, and the resulting peptide mixture is analyzed by MS/MS. Then proteins are identified by searching the recorded MS/MS spectra against protein databases, and quantification is obtained by comparing the light and heavy intensity for each peptide. In our experimental set up, all GG-modified peptides identified from cells expressing wild-type SUMO are expected to correspond to non-SUMO sites, because trypsin digestion of SUMOylated proteins in these conditions does not produce GG tags. Only GG-modified lysines identified specifically with SUMO variants were considered as bona fide SUMOylation sites (Fig. 2). A third labeling condition can be added to this experiment to study SUMOylome changes in response to external stimuli. As described below, we treated cells expressing SUMO variants with the bacterial toxin LLO (Fig. 2).

Proteome-Wide Identification of SUMO1 and SUMO2 Sites. HeLa cells cultivated in different SILAC media (~2 × 10^7 cells per SILAC condition) were transfected with wild-type (light

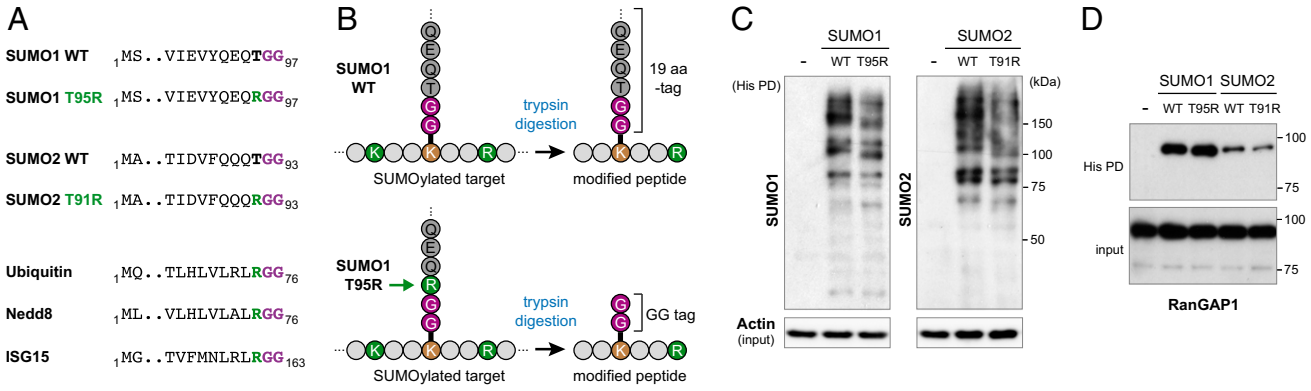

Fig. 1. Generation of SUMO variants compatible with quantitative GG-capture proteomics. (*A*) Comparison of C-terminal sequences in mature SUMO1, SUMO2, ubiquitin, Nedd8, and ISG15. C-terminal GG motifs are highlighted in purple. The positions of mutations introduced in SUMO1 and SUMO2 are indicated in green. (*B*) Schematic representation of the signature tags left after trypsin digestion on peptides modified by either wild-type or T95R SUMO1. (*C*) Comparison of the patterns of SUMO-conjugated protein in cells expressing wild-type or variant SUMOs. SUMOylated proteins from HeLa cells transfected with wild-type or variant His_6-SUMOs were pulled down and analyzed by immunoblot using anti-SUMO1 or anti-SUMO2/3 antibodies. (*D*) Detection of RanGAP1 SUMO-conjugated forms by immunoblot analysis of HeLa cells expressing wild-type or variant His_6-SUMOs. Input fractions are shown as control.

labeling) or variant His_6-SUMO (medium and heavy labeling) (Fig. 2*A*). Two independent experiments were performed to identify SUMO1 and SUMO2 sites, respectively. Two days after transfection, cells were lysed in denaturing buffer, and SUMOylated proteins were affinity purified by nickel chromatography and digested by trypsin. GG-modified peptides then were enriched by peptide immunoprecipitation before MS analysis. With this last step GG-modified peptides were enriched by more than 300 fold, as previously observed in studies focusing on ubiquitin (21).

SUMO sites were determined by comparing the data obtained from cells transfected with wild-type and variant SUMO. To this end, GG-modified peptides were quantified by comparing the intensities of the different SILAC labels in the MS spectra. More than 70% of these peptides showed no detectable signal in the light channel and therefore were considered to be markers of bona fide SUMO sites (Fig. 2). Together, these analyses led to the identification of 295 SUMO1 sites from 227 endogenous targets and 167 SUMO2 sites from 135 endogenous targets (see Datasets S1 and S2 for lists of identified SUMO sites and GG-modified peptides), resulting in a very comprehensive list of human SUMO sites (Dataset S1). Among the 332 different SUMO sites identified in our screen, 130 sites (39%) were found both with SUMO1 and SUMO2, 165 sites (50%) were found only with SUMO1, and 37 sites (11%) were found only with SUMO2 (Fig. 2*D*). These percentages are consistent with previously observed isoform preferences (6).

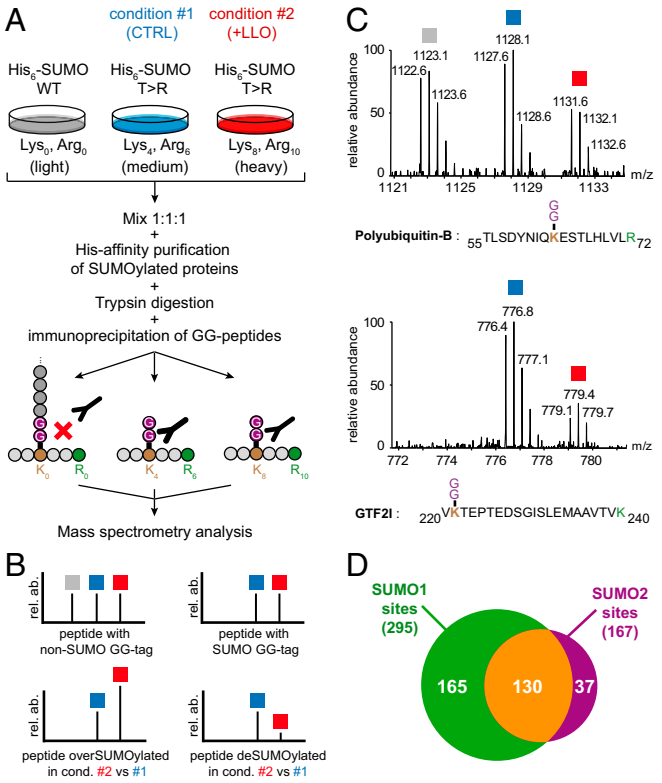

Fig. 2. Method used for quantitative mapping of SUMO sites. (*A*) Schematic representation of the approach. (*B*) Theoretical MS spectra expected for peptides conjugated to SUMO or non-SUMO modifiers and to peptides over-SUMOylated or deSUMOylated in condition #2 versus condition #1. rel. ab., relative abundance, (*C*) MS spectra obtained for two different GG-modified peptides. The peptide from polyubiquitin-B is present in its light-labeled form and therefore was considered a contaminant, most likely derived from K63-linked polyubiquitin chains. The peptide from GTF2I (General transcription factor II-I) displays a SUMO1 site on K221. The absence of the peptide in its light-labeled form excludes GG modification by a non-SUMO modifier (i.e., ubiquitin, Nedd8, or ISG15). The decrease in intensity of the heavy-labeled form of this peptide indicates a decrease in the SUMOylation level of this site after LLO treatment. (*D*) Overlap between SUMO1 and SUMO2 sites identified in this study.

Of our 332 identified sites, 86 (26%) were previously reported in proteomic screens for SUMO sites (11, 14, 16, 17, 23, 24), and 82 (25%) are reported in the PhosphoSitePlus database (a resource for human PTMs, including SUMOylation) (25), thereby validating our approach (Dataset S1). Thus, to our knowledge, 227 SUMO sites (203 for SUMO1 and 82 for SUMO2) were identified here for the first time (Dataset S1). To validate the reproducibility of our approach, we repeated our analysis for SUMO1 in a smaller-scale experiment ($\sim 1 \times 10^7$ cells per SILAC condition). We identified 132 SUMO1 sites, of which 115 (87%) were in common with the 295 SUMO1 sites found in the first analysis, indicating a high degree of reproducibility (Fig. S2).

To validate further the SUMOylated proteins identified in our screens and their associated SUMO sites, we selected several candidates: the transcriptional repressors ZBTB20 (Zinc finger and BTB domain containing 20), HMBOX1 (Homeobox containing protein 1), NACC1 (Nucleus accumbens-associated protein 1), the transcription factor TFAP2A (Transcription factor AP-2 alpha), the microtubule-binding protein MAP7 (Microtubule-associated protein 7)/ensconsin, and the lamin-B1 protein LMNB1. ZBTB20, HMBOX1, NACC1, MAP7, and LMNB1 were previously reported to be SUMOylated, but their SUMOylation sites had not been characterized (6, 12, 13, 26). We generated expression vectors for the HA-tagged version of each of these six proteins and mutated the SUMO-modified lysines identified in our screens into arginines to obtain non-SUMOylatable mutants. We cotransfected HeLa cells with plasmids encoding the different HA-tagged candidates and the corresponding mutants and His_6-tagged SUMO1 or SUMO2. After cell lysis, SUMOylated proteins were nickel purified, and the presence of the SUMOylated forms of the different candidates was assayed by immunoblotting experiments using anti-HA antibodies. For each tested candidate, slower-migrating bands corresponding to SUMO-modified forms of these proteins were detected in the His–pulled-down fraction from cells expressing His_6-SUMO1 or His_6-SUMO2, thus demonstrating that all of these proteins are indeed modified by SUMO (Fig. 3). As already observed for many other SUMO targets (2), the percentage of SUMOylated versus non-SUMOylated proteins is rather low (~5% for ZBTB20, ~2% for TFAP2A, and below 1% for HMBOX1, NACC1, MAP7, and LMNB1). For ZBTB20, HMBOX1, and TFAP2A, we did not observe SUMO-modified forms in cells expressing the different non-SUMOylatable mutants, thus confirming that the sites identified in our proteomic screen are bona fide SUMO sites (Fig. 3). For NACC1, MAP7, and LMNB1, expression of non-SUMOylatable mutants led to a decrease in SUMOylation of these proteins, even though some SUMOylated forms still could be detected (Fig. 3). These data argue that the mutated residues correspond to real SUMO sites but also suggest either that additional SUMO sites are present in these targets or that compensatory events may lead to SUMOylation of different lysines in these mutants. Together, these experiments confirm the previously described SUMO site of TFAP2A (27) and provide the first identification (to our knowledge) of SUMO sites for ZBTB20, HMBOX1, NACC1, MAP7, and LMNB1.

We then classified the identified SUMOylated proteins from our screens by Gene Ontology (GO) analysis. Proteins annotated as "nuclear," "nuclear lumen," or "nucleolus" were significantly enriched in the list of SUMO-identified proteins relative to the whole human proteome (Fig. 4*A*). This result confirmed the well-established finding that a high proportion of SUMOylated proteins are nuclear factors. In particular, we identified several SUMO-conjugated DNA-binding proteins and transcription factors that are consistent with the essential role of SUMO in the regulation of gene expression (1, 28). Apart from these classes of proteins, we identified a significant enrichment for cytoskeletal proteins in our list of identified SUMOylated proteins (modified Fisher exact P value = 4.1×10^{-4} for SUMO1 and 1.6×10^{-2} for

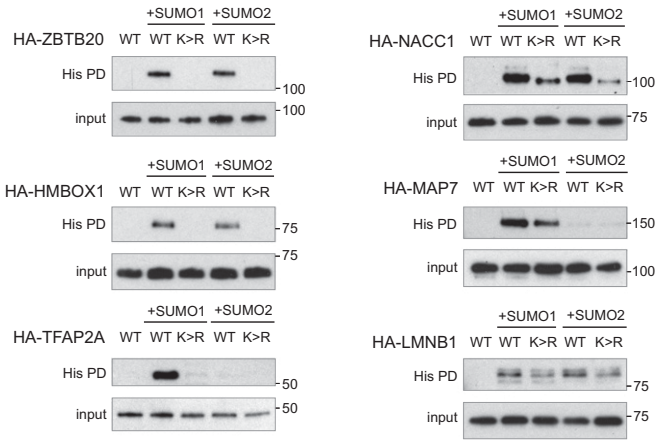

Fig. 3. Validation of identified SUMOylated proteins and SUMO sites. HeLa cells were cotransfected with His_6-SUMO1 or -SUMO2 and with HA-tagged wild-type or K-to-R mutant expression vectors for selected candidates. Cell lysates were subjected to His pulldown (His PD), and the presence of SUMOylated forms of the different candidates was assayed using anti-HA antibodies. Input fractions are shown as control.

SUMO2) (Fig. 4*A* and Dataset S1). In particular, we identified several SUMO targets implicated in the architecture of the actin cytoskeleton, such as actin itself, anillin, cortactin, or RhoGDI. Of note, actin already was reported to be SUMOylated (albeit on different residues) (29), but cortactin and anillin were not known to be SUMOylated. We also identified several intermediate filament proteins as SUMO targets, including keratins, lamins, nestin, and vimentin, for which we provide data on yet uncharacterized SUMO sites. Finally, in addition to nuclear and cytoskeletal proteins, we identified examples of other classes of proteins, such as plasma membrane proteins or proteins from intracellular organelles that also can be targeted by SUMOylation (Dataset S1).

Consensus and Nonconsensus SUMO Sites. Analysis of the amino acids surrounding SUMOylated lysines showed that 65% of SUMO sites lie in a Kx[DE] environment and that 48% conform to the previously established SUMO consensus motif [FILMV]Kx[DE] (Fig. 4*B*) (30, 31). It has been established that residues constituting this SUMO consensus motif interact directly with the E2 SUMO-conjugating enzyme (reviewed in ref. 2). For SUMO sites located in such a motif, we observed a marked preference for valine or isoleucine for the hydrophobic residue preceding the modified lysine (Fig. 4*C*).

In addition to this general SUMO consensus motif, extended motifs have been characterized in several SUMO targets. Among them, phosphorylation-dependent SUMO motifs (PDSMs) are characterized by the presence of a phosphorylated residue downstream of a classical consensus motif; the presence of this residue was shown to increase SUMOylation efficiency by mediating interactions between the target and a basic patch on Ubc9 (32, 33). To study the link between SUMOylation and phosphorylation further, we also analyzed our MS data, taking into account phosphorylation as a possible peptide modification. We identified 13 SUMO-conjugated peptides that may be phosphorylated. Interestingly, six of these peptides were found only in their phosphorylated state, probably reflecting the importance of phosphorylation in the SUMOylation efficiency of the corresponding targets (Dataset S3). For each of these SUMOylated/phosphorylated peptides, the phosphorylated residue is a serine located downstream of a lysine lying in a SUMO consensus motif. As previously reported for some SUMOylated/phosphorylated targets, the sequence motif surrounding the phosphorylated residue diverges slightly from the initially described PDSM (ΨKxExxSP, in which K is the SUMO-conjugated lysine and S the phosphorylated serine) (17, 32, 34). Indeed, we observed that the distance between the phosphorylated and SUMOylated residues can vary between 4 and 14 residues, and, even though we noticed a marked preference for proline after the phosphorylated residue, glutamate or aspartate residues were present also (Dataset S3).

In parallel to PDSM, negatively charged residues also can be found downstream of SUMO-conjugated lysines, which may replace the phosphorylated serine side chain of PDSMs and maintain a constitutively active motif for SUMO conjugation. Among our list of identified SUMO sites, we frequently observed aspartate and glutamate residues in positions +4 to +8 downstream of the modified lysine, and, accordingly, 58 sites (37% of sites in SUMO consensus motif) are located in negatively charged amino acid-dependent SUMO motifs (Fig. 4*C* and Dataset S1) (35).

Finally, 29 of our identified sites (18% of the sites in the SUMO consensus motif) match the hydrophobic cluster SUMOylation motif, which is characterized by the presence of at least three residues with hydrophobic properties upstream of the modified lysine instead of the single hydrophobic residue usually present (Dataset S1) (17).

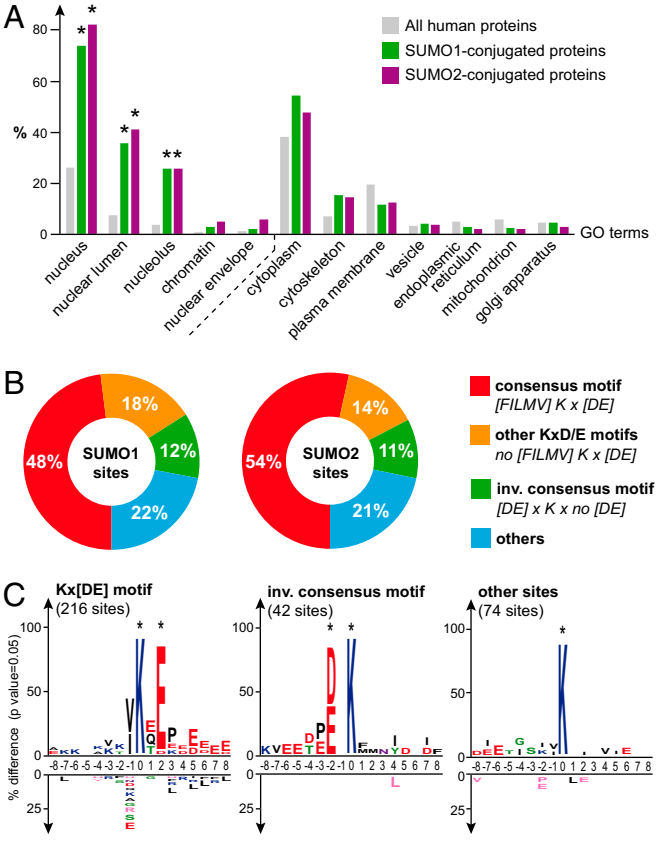

Fig. 4. Analysis of identified SUMO1- and SUMO2-conjugation sites. (*A*) GO terms enrichment analysis of proteins conjugated to SUMO1 or SUMO2 compared with all human proteins. Bars correspond to the percentage of proteins annotated with each GO term. Asterisks indicate groups with significant enrichment (modified Fisher exact *P* value < 0.0001). (*B*) Distribution of SUMO1 and SUMO2 sites over different types of SUMOylation motifs. (*C*) IceLogo (41) representations showing the amino acids surrounding the SUMO-conjugated lysine residue in different motifs. Imposed amino acids for each type of motif are marked by an asterisk. The frequency of nonimposed amino acids at every subsite is compared with sampled frequencies in the human proteins stored in the UniProt/Swiss-Prot database (negative control). Only residues that are statistically overrepresented (upper part of the iceLogo) or underrepresented (lower part of the iceLogo) at a 95% confidence level are depicted. Residues that never were observed at specific positions are shown in pink.

CELL BIOLOGY

By analyzing sites that lack a Kx[DE] motif, we confirmed the existence of an "inverted SUMOylation consensus motif" (defined as [DE]xKx[no DE]) (17) for 42 sites (13% of total SUMO sites) (Fig. 4 *B* and *C*). Interestingly, this inversion concerns not only the D/E residue but also the negatively charged residues, which in this case frequently are found upstream of the SUMO-conjugated lysine sites (Fig. 4*C*). A significant number (73; 22%) of sites did not lie in either a Kx[DE] motif or an inverted SUMOylation consensus motif. Analysis of these sites did not reveal any marked preference for surrounding amino acids, as is the case for ubiquitin sites (Fig. 4*C*). Further investigation will be required to determine whether the specificity of the modified lysine is directed by the flanking residues of these nonconsensus SUMO sites interacting with Ubc9 or by other factors such as SUMO E3 enzymes. The absence of motifs for these nonconsensus sites renders their prediction by bioinformatic analysis almost impossible, thus strengthening the utility of an untargeted approach, such as the one described here, for their identification.

Finally, we observed no significant differences in the frequencies of the various motifs identified for SUMO1 and SUMO2 sites (Fig. 4*B*). This result indicates that the selection of the SUMO-conjugated lysine is not influenced by the sequence of the SUMO isoform itself but probably is imposed by the interactions between the target, the E2 and the E3 SUMO enzymes.

We cross-referenced our data with previously established databases of other PTMs (25). Interestingly, we observed that important fractions of our identified SUMO sites also are reported to be acetylated (20%) or ubiquitylated (37%) (Dataset S1). This finding highlights the importance of the cross-talk between SUMOylation and other posttranslational modifications that either can compete for the same lysine or occur sequentially at a given site.

Analysis of SUMOylome Changes Induced by an External Stimulus. One powerful aspect of our strategy is that it enables the comparison of the SUMO landscape of two cell populations in different biological conditions. We thus decided to apply our technique to elucidate changes in SUMOylation induced by a bacterial toxin, LLO. This toxin is secreted by the Gram-positive bacterial pathogen *Listeria monocytogenes*, the etiological agent of human listeriosis (36), and has a potent signaling activity through pore formation in the host plasma membrane during infection (37). Among the different cellular responses triggered by LLO, we previously have shown that this toxin induces the degradation of Ubc9, the unique E2 enzyme of the host SUMOylation machinery (38). This degradation leads to a blockade in SUMOylation and a global decrease in the level of host SUMOylated proteins, an effect that is beneficial for efficient infection by *Listeria* (38). To obtain further insight into the extent of deSUMOylation events in response to LLO and to identify deSUMOylated proteins, we added to the protocol described above a third SILAC condition corresponding to cells transfected with variant His_6-SUMO that had been treated with a sublytic concentration of LLO for a short time (i.e., 3 nM for 20 min) (Fig. 2). Immunoblot analysis using antibodies specific for SUMO1 and SUMO2/3 confirmed that these conditions led to a global decrease in the level of SUMO1- and SUMO2-conjugated proteins (Fig. 5*A*). Quantitative analysis of our proteomic data confirmed this global decrease in SUMO-conjugated proteins (Dataset S4). More specifically, we identified 35 SUMO1 sites and 90 SUMO2 sites with a medium/heavy (M/H) ratio >2, indicative of a decrease in SUMOylation in response to LLO (Dataset S4). In contrast, very few proteins are associated with M/H ratios <1, indicating that de novo SUMOylation of host proteins in response to LLO is very limited.

Analysis of SUMO motif frequencies for sites highly deSUMOylated in response to LLO shows an overrepresentation of lysines lying in SUMO consensus motifs compared with the total list of identified SUMOylated sites (Fig. S3). Functional enrichment analysis of highly deSUMOylated proteins [31 proteins for SUMO1 (M/H ratio > 2) and 35 proteins for SUMO2 (M/H ratio > 4)] showed that several classes of targets annotated "nucleus," "DNA-binding," "zinc-finger," or "transcription regulation" are significantly overrepresented in the list of highly deSUMOylated proteins relative to the total list of identified SUMOylated proteins (Fig. 5*B*). This result suggests that SUMOylated host factors are not similarly sensitive to LLO-induced loss of SUMOylation. In particular, several nuclear factors are strongly deSUMOylated in response to LLO, whereas other classes of proteins (e.g., cytoskeletal proteins) are less affected. SUMOylation of transcription factors is known to regulate their transcriptional activity either positively or negatively (1, 28). Thus LLO-induced modification of the SUMOylation state of these proteins may alter the transcriptome of the cell in response to toxin exposure. Alteration of the host transcriptome in response to LLO has been described previously (39), but the exact role of deSUMOylation in this response remains to be determined. The identification of proteins deSUMOylated upon LLO treatment thus provides candidate host factors for which SUMO-regulation may play an important role during the establishment and persistence of *Listeria* infection.

Discussion

In the last decade several strategies have been developed to identify SUMOylation sites. Site-directed mutagenesis of lysine residues in the SUMOylated target constitutes one of the most common strategies for identifying SUMO-conjugated lysines. However, this technique is time-consuming and often is limited to SUMO sites predicted by the analysis of SUMO consensus motifs. In addition, this technique does not formally differentiate between bona fide SUMO sites and residues involved in SUMOylation of distal lysines (e.g., residues mediating interactions with the SUMOylation machinery that themselves are not modified). MS constitutes an untargeted and high-throughput approach to identify SUMO sites. Different strategies have been used to circumvent the difficulties in identifying SUMO-conjugated peptide arising from the complexity of the associated MS/MS spectra. These approaches rely essentially on elucidation of the complex MS/MS spectra or on the use of modified SUMO versions that leave simpler tags on SUMO-modified peptides and thus are more easily identified by classical MS (refs. 11, 14–17, 23, 24 and reviewed in ref. 8). Despite these efforts, only a limited number of SUMO sites have been identified by these different approaches thus far.

Here, by combining SILAC-based quantitative proteomics and immunocapture of SUMO-modified peptides, we developed a powerful method for identifying SUMO sites. Using this

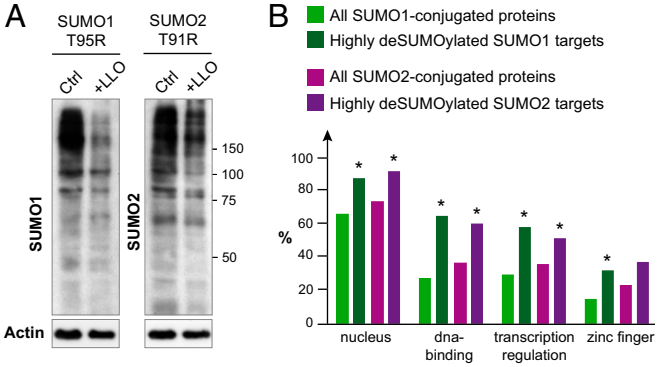

Fig. 5. Analysis of targets deSUMOylated in response to the LLO toxin. (*A*) Immunoblot analysis of SUMO-conjugated patterns from cells transfected with His_6-SUMOs variants and exposed to LLO. (*B*) Functional enrichment analysis of highly deSUMOylated proteins after LLO treatment compared with all SUMO-conjugated proteins. Bars correspond to the percentage of proteins annotated with each Splice Pattern-Protein Information Resources (SP-PIR) keyword. Asterisks indicate groups with significant enrichment (modified Fisher exact *P* value < 0.05).

approach, we identified 295 SUMO1 and 167 SUMO2 sites in human endogenous proteins. Of note, we identified 227 SUMO sites that were not previously described and that will provide a useful database for the SUMO community. Furthermore, by taking advantage of quantitative proteomics, our method allows SUMOylome comparison between two different cell populations and thus may open new avenues for studying the role of SUMOylation in response to variations in environmental conditions, exposure to drugs or toxins, or infection by various pathogens.

While this manuscript was in review, a similar strategy using cells stably expressing a His_6-SUMO2 T91K variant and allowing mapping of SUMO2 sites was published (40). In this study, His_6-SUMO2 T91K–conjugated proteins were isolated by nickel affinity chromatography and then digested by endoproteinase Lys-C, leaving a specific GG tag on SUMO2 T91K-modified peptides (whereas proteins modified by endogenous ubiquitin, Nedd8, or ISG15 give rise to peptides modified by a larger tag, because endoproteinase Lys-C cuts only after lysine and not after arginine residues) (Fig. 1*A*) (40). GG-modified peptides then were enriched using anti–K-ε-GG immunoprecipitation before MS analysis. This approach was used successfully to identify SUMO2 sites in proteins from HEK293 cells after heat shock, a stress condition known to enhance SUMO2 conjugation strongly (10, 13, 40). The independent success of both our approach and that of Tammsalu et al. (40) demonstrates that the use of SUMO variants in combination with enrichment of resulting SUMO-modified peptides constitutes a powerful and broadly applicable strategy to study SUMOylation. The additional use of SILAC-based quantitative proteomics in our approach allows further analysis of SUMOylome changes in response to external stimuli.

Finally, these methods, although developed for the study of SUMO, can be broadened further to investigate conjugation sites of other UBL proteins by using similar combinations of variants leading to the presence or absence of GG-modified peptides after digestion by trypsin or by endoproteinase Lys-C. These approaches therefore constitute a generalizable tool to study ubiquitin-like modifications and to provide further insight in the role of these PTMs in cell physiology.

Methods

Immunocapture of diglycine (GG)-modified peptides was performed using the PTMScan Ubiquitin Remnant Motif (K-e-GG) Kit (Cell Signaling Technology) and peptides were analyzed on an LTQ Orbitrap Velos mass spectrometer (Thermo). Descriptions of plasmids, bacterial strains, and antibodies, as well as detailed experimental procedures used in this study are provided in *SI Methods*. Bacterial strains are listed in Table S1.

ACKNOWLEDGMENTS. We thank Edith Gouin and Marie-Anne Nahori for help with production of antibodies and plasmids and Anne Dejean for reagents and the Pasteur Proteomics platform. This work received financial support from Institut Pasteur, Institut National de la Santé et de la Recherche Médicale (INSERM), Institut National de la Recherche Agronomique (INRA), Agence Nationale de Recherches (ERANET Pathogenomics LISTRESS), Grant ANR-10-LABX-62-IBEID from the French Government's Investissement d'Avenir program, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases," Advanced Grant 233348 MODELIST from the European Research Council (ERC), the Fondation le Roch les Mousquetaires, and the Fondation Louis-Jeantet. F.I. received financial support from a Pasteur-Roux Fellowship and L.R. received financial support from the Human Frontiers Science Program. P.C. is a Senior International Research Scholar of the Howard Hughes Medical Institute and D.R. is an INSERM Research Associate.

1. Geiss-Friedlander R, Melchior F (2007) Concepts in sumoylation: A decade on. *Nat Rev Mol Cell Biol* 8(12):947–956.
2. Gareau JR, Lima CD (2010) The SUMO pathway: Emerging mechanisms that shape specificity, conjugation and recognition. *Nat Rev Mol Cell Biol* 11(12):861–871.
3. Ribet D, Cossart P (2010) Pathogen-mediated posttranslational modifications: A re-emerging field. *Cell* 143(5):694–702.
4. Ribet D, Cossart P (2010) SUMOylation and bacterial pathogens. *Virulence* 1(6):532–534.
5. Wimmer P, Schreiner S, Dobner T (2012) Human pathogens and the host cell SUMOylation system. *J Virol* 86(2):642–654.
6. Becker J, et al. (2013) Detecting endogenous SUMO targets in mammalian cells and tissues. *Nat Struct Mol Biol* 20(4):525–531.
7. Hickey CM, Wilson NR, Hochstrasser M (2012) Function and regulation of SUMO proteases. *Nat Rev Mol Cell Biol* 13(12):755–766.
8. Da Silva-Ferrada E, Lopitz-Otsoa F, Lang V, Rodríguez MS, Matthiesen R (2012) Strategies to Identify Recognition Signals and Targets of SUMOylation. *Biochem Res Int* 2012:875148.
9. Tatham MH, Rodriguez MS, Xirodimas DP, Hay RT (2009) Detection of protein SUMOylation in vivo. *Nat Protoc* 4(9):1363–1371.
10. Saitoh H, Hinchey J (2000) Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J Biol Chem* 275(9):6252–6258.
11. Galisson F, et al. (2011) A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells. *Mol Cell Proteomics* 10(2):M110 004796.
12. Tatham MH, Matic I, Mann M, Hay RT (2011) Comparative proteomic analysis identifies a role for SUMO in protein quality control. *Sci Signal* 4(178):rs4.
13. Golebiowski F, et al. (2009) System-wide changes to SUMO modifications in response to heat shock. *Sci Signal* 2(72):ra24.
14. Schimmel J, et al. (2014) Uncovering SUMOylation dynamics during cell-cycle progression reveals FoxM1 as a key mitotic SUMO target protein. *Mol Cell* 53(6):1053–1066.
15. Knuesel M, Cheung HT, Hamady M, Barthel KK, Liu X (2005) A method of mapping protein sumoylation sites by mass spectrometry using a modified small ubiquitin-like modifier 1 (SUMO-1) and a computational program. *Mol Cell Proteomics* 4(10):1626–1636.
16. Blomster HA, et al. (2010) In vivo identification of sumoylation sites by a signature tag and cysteine-targeted affinity purification. *J Biol Chem* 285(25):19324–19329.
17. Matic I, et al. (2010) Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. *Mol Cell* 39(4):641–652.
18. Hirohama M, et al. (2014) Assay methods for small ubiquitin-like modifier (SUMO)-SUMO-interacting motif (SIM) interactions in vivo and in vitro using a split-luciferase complementation system. *Anal Biochem* 448:92–94.
19. Xu G, Paige JS, Jaffrey SR (2010) Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat Biotechnol* 28(8):868–873.
20. Kim W, et al. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol Cell* 44(2):325–340.
21. Wagner SA, et al. (2011) A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Mol Cell Proteomics* 10(10):M111 013284.
22. Ong SE, et al. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1(5):376–386.
23. Hsiao HH, Meulmeester E, Frank BT, Melchior F, Urlaub H (2009) "ChopNSpice," a mass spectrometric approach that allows identification of endogenous small ubiquitin-like modifier-conjugated peptides. *Mol Cell Proteomics* 8(12):2664–2675.
24. Lamoliatte F, et al. (2013) Targeted identification of SUMOylation sites in human proteins using affinity enrichment and paralog-specific reporter ions. *Mol Cell Proteomics* 12(9):2536–2550.
25. Hornbeck PV, et al. (2012) PhosphoSitePlus: A comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* 40(Database issue):D261–D270.
26. Tirard M, et al. (2012) In vivo localization and identification of SUMOylated proteins in the brain of His6-HA-SUMO1 knock-in mice. *Proc Natl Acad Sci USA* 109(51):21122–21127.
27. Eloranta JJ, Hurst HC (2002) Transcription factor AP-2 interacts with the SUMO-conjugating enzyme UBC9 and is sumolated in vivo. *J Biol Chem* 277(34):30798–30804.
28. Cubeñas-Potts C, Matunis MJ (2013) SUMO: A multifaceted modifier of chromatin structure and function. *Dev Cell* 24(1):1–12.
29. Hofmann WA, et al. (2009) SUMOylation of nuclear actin. *J Cell Biol* 186(2):193–200.
30. Rodriguez MS, Dargemont C, Hay RT (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J Biol Chem* 276(16):12654–12659.
31. Sampson DA, Wang M, Matunis MJ (2001) The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. *J Biol Chem* 276(24):21664–21669.
32. Yang XJ, Grégoire S (2006) A recurrent phospho-sumoyl switch in transcriptional repression and beyond. *Mol Cell* 23(6):779–786.
33. Mohideen F, et al. (2009) A molecular basis for phosphorylation-dependent SUMO conjugation by the E2 UBC9. *Nat Struct Mol Biol* 16(9):945–952.
34. Hietakangas V, et al. (2006) PDSM, a motif for phosphorylation-dependent SUMO modification. *Proc Natl Acad Sci USA* 103(1):45–50.
35. Yang SH, Galanis A, Witty J, Sharrocks AD (2006) An extended consensus motif enhances the specificity of substrate modification by SUMO. *EMBO J* 25(21):5083–5093.
36. Cossart P (2011) Illuminating the landscape of host-pathogen interactions with the bacterium Listeria monocytogenes. *Proc Natl Acad Sci USA* 108(49):19484–19491.
37. Hamon MA, Ribet D, Stavru F, Cossart P (2012) Listeriolysin O: The Swiss army knife of Listeria. *Trends Microbiol* 20(8):360–368.
38. Ribet D, et al. (2010) Listeria monocytogenes impairs SUMOylation for efficient infection. *Nature* 464(7292):1192–1195.
39. Hamon MA, et al. (2007) Histone modifications induced by a family of bacterial toxins. *Proc Natl Acad Sci USA* 104(33):13467–13472.
40. Tammsalu T, et al. (2014) Proteome-wide identification of SUMO2 modification sites. *Sci Signal* 7(323):rs2.
41. Colaert N, Helsens K, Martens L, Vandekerckhove J, Gevaert K (2009) Improved visualization of protein consensus sequences by iceLogo. *Nat Methods* 6(11):786–787.