A Novel Proteomics Approach to Identify SUMOylated Proteins and Their Modification Sites in Human Cells*

Frederic Galisson, a,b Louiza Mahrouche, b,c,d,e Mathieu Courcelles, c,d Eric Bonneil, c Sylvain Meloche, c,f Mounira K. Chelbi-Alix, a,g,h and Pierre Thibaultc,d,g,i,j

The small ubiquitin-related modifier (SUMO) is a small group of proteins that are reversibly attached to protein substrates to modify their functions. The large scale identification of protein SUMOylation and their modification sites in mammalian cells represents a significant challenge because of the relatively small number of in vivo substrates and the dynamic nature of this modification. We report here a novel proteomics approach to selectively enrich and identify SUMO conjugates from human cells. We stably expressed different SUMO paralogs in HEK293 cells, each containing a His6 tag and a strategically located trypic cleavage site at the C terminus to facilitate the recovery and identification of SUMOylated peptides by affinity enrichment and mass spectrometry. Tryptic peptides with short SUMO remnants offer significant advantages in large scale SUMOlyome experiments including the generation of paralog-specific fragment ions following CID and ETD activation, and the identification of modified peptides using conventional database search engines such as Mascot. We identified 205 unique protein substrates together with 17 precise SUMOylation sites present in 12 SUMO protein conjugates including three new sites (Lys-380, Lys-400, and Lys-497) on the protein promyelocytic leukemia. Label-free quantitative proteomics analyses on purified nuclear extracts from untreated and arsenic trioxide-treated cells revealed that all identified SUMOylated sites of promyelocytic leukemia were differentially SUMOylated upon stimulation.


The small ubiquitin-like modifier (SUMO)1 proteins are structurally similar to ubiquitin, although they share less than 20% sequence identity (1). Like ubiquitylation, protein SUMOylation is regulated by a cascade of reactions involving SUMO-activating enzymes (SAE1/SAE2), -conjugating enzymes (Ubc9), and one of several SUMO E3 ligases (e.g. PIAS1, PIAS3, PIASXα, PIASβ, PIASγ, RanBP2, and Pci) that covalently attach SUMO to specific protein substrates (2, 3). SUMO proteins are expressed as an immature proform that comprises an invariant Gly-Gly motif followed by a C-terminal stretch of variable length (2–11 amino acids). Removal of this C-terminal extension by sentrin-specific proteases (SENPs) to expose the diglycine motif is necessary for the conjugation of SUMO to protein targets. These SUMO proteases are able to cleave both a peptide bond during the formation of mature SUMO and an isopeptide bond to deconjugate modified protein substrates (4). This covalent modification arises from the formation of an isopeptide bond between the ε-amino group of a lysine within the protein substrate and the C terminus carboxyl group of the SUMO glycine residue. SUMO conjugation frequently occurs at the lysine residue within the consensus motif ψKXE (where ψ is an aliphatic residue and X is any amino acid) that is recognized by Ubc9 (5, 6). Recent studies have also identified a phosphorylation-dependent motif (ΨKXEKKxPSp where pS is phosphoserine) (7) and a negatively charged amino acid-dependent motif (8) that harbor negative charges next to the basic SUMO consensus site to enhance protein SUMOylation. However, several other SUMOylated proteins including proliferating cell nuclear antigen, E2-25K, Daxx (death domain-associated protein), and USP25 are modified at non-consensus sites (9–11). Whether these types of sites are rare nuclear ribonucleoprotein; NB, nuclear body; NTA, nickel nitriolate acid; PARP1, poly(ADP-ribose) polymerase 1; PIAS, protein inhibitor of activated STAT; PML, promyelocytic leukemia; RanBP2, Ran-binding protein 2; RanGAP1, Ran GTPase-activating protein 1; RING, really interesting new gene; RNF4, RING finger 4 protein; RSF1, remodeling and spacing factor 1; SAE, SUMO-activating enzyme; SAFB2, scaffold assembly factor B2; SENP, sentrin-specific protease; TIF-1β, transcription intermediary factor 1-β; Ubc9, ubiquitin-like protein SUMO1-conjugating enzyme; mut, mutant; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; SCX, strong cation exchange; IPI, International Protein Index; HSE, heat shock element; CTBP, C-terminal binding protein.

From the aCNRS FRE3238, Institut André Lwoff, Villejuif 94801, France and bInstitute for Research in Immunology and Cancer and Departments of aBiochemistry, aPharmacology, and aChemistry, Université de Montréal, P. O. Box 6128, Station Centre-ville, Montréal, Québec H3C 3J7, Canada

Received, September 13, 2010, and in revised form, November 8, 2010
Published, MCP Papers in Press, November 22, 2010, DOI 10.1074/mcp.M110.004796

1 The abbreviations used are: SUMO, small ubiquitin-related modifier; As2O3, arsenic trioxide; E2-25K, E2 ubiquitin ligase, 25 kilodaltons; ETD, electron transfer dissociation; hnRNP, heterogeneous nuclear ribonucleoprotein; NB, nuclear body; NTA, nickel nitriolate acid; PARP1, poly(ADP-ribose) polymerase 1; PIAS, protein inhibitor of activated STAT; PML, promyelocytic leukemia; RanBP2, Ran-binding protein 2; RanGAP1, Ran GTPase-activating protein 1; RING, really interesting new gene; RNF4, RING finger 4 protein; RSF1, remodeling and spacing factor 1; SAE, SUMO-activating enzyme; SAFB2, scaffold assembly factor B2; SENP, sentrin-specific protease; TIF-1β, transcription intermediary factor 1-β; Ubc9, ubiquitin-like protein SUMO1-conjugating enzyme; mut, mutant; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; SCX, strong cation exchange; IPI, International Protein Index; HSE, heat shock element; CTBP, C-terminal binding protein.
exceptions or reflect the presence of other E2-conjugating enzymes is presently unknown.

In lower eukaryotes, a single SUMO gene is expressed (Smt3 in Saccharomyces cerevisiae), whereas in vertebrates, three paralogs designated as SUMO1, SUMO2, and SUMO3 are ubiquitously expressed in all tissues. The human genome also encodes a fourth gene for SUMO4 that seems to be uniquely expressed in the spleen, lymph nodes, and kidney (12). However, its role remains enigmatic as its in vivo maturation into a conjugation-competent form still remains unclear (13). Interestingly, SUMO2 and SUMO3 share 97% sequence identity and are expressed at much higher levels than SUMO1 with which they only share about 50% identity (1). Although SUMO paralogs use the same conjugation machinery and have partially overlapping subsets of target proteins, they respond differently to stress (14) and can be distinguished by their ability to form self-modified polymers in vivo and in vitro (15, 16). SUMO1 lacks a consensus modification site and does not form polySUMO1 chains in vivo, although RanBP2 was reported to be hypermodified by SUMO1 chains in vitro (17). In contrast, SUMO2 and SUMO3 can form polymeric chains in vivo and in vitro through their consensus motif (15), whereas SUMO1 forms terminating chain on polySUMO2 or polySUMO3 conjugates (16).

Protein SUMOylation is an essential cellular process conserved from yeast to mammals and plays an important role in the regulation of intracellular trafficking, cell cycle, DNA repair and replication, cell signaling, and stress responses (2, 18, 19). Protein SUMOylation imparts significant structural and conformational changes on the substrate proteins by masking and/or by conferring additional scaffolding surfaces for protein interactions. At present, a few hundred protein substrates are known to be SUMOylated in vivo. These protein targets include regulators of gene expression (e.g., transcription factors, co-activators, and repressors) as well as oncogenes and tumor suppressor genes such as promyelocytic leukemia (PML), murine double minute-2 (Mdm2), c-Myb, c-Jun, and p53 whose misregulation leads to tumorigenesis and metastasis (20). There is growing evidences of cross-talk between protein SUMOylation and ubiquitylation processes (21, 22). Earlier reports indicated that SUMOylation can antagonize the ubiquitylation of nuclear factor κB (23), whereas recent data also suggest that SUMOylation can be a prerequisite for ubiquitylation and subsequent proteasome-dependent degradation. A case in point is the identification of RNF4, an E3 ubiquitin ligase that specifically recognizes and ubiquitylates polySUMO chains of PML (24, 25). Interestingly, PML SUMOylation can also be enhanced using arsenic trioxide (As$_2$O$_3$), a therapeutic agent used for the treatment of acute promyelocytic leukemia (25–28).

The relatively low abundance of protein SUMOylation is a significant analytical challenge for the identification and quantification of this modification in vivo. Recent reports have described the successful identification of SUMO protein candidates by enriching the small subset of SUMOylated proteins using cysteine-targeted purification (29), a tandem affinity tag with His$_6$-SUMO proteins resistant to Lys-C proteolysis (30), and cells stably expressing His$_6$-SUMO constructs (31) and quantifying their proportions using mass spectrometry (MS) and metabolic labeling in cell cultures (32). Despite these significant advances, the identification of SUMOylation sites by MS remains challenging because of their variable stoichiometry and the presence of long SUMO C terminus polypeptides that complicates the interpretation of the corresponding product ion spectra. Upon tryptic digestion, SUMOylated peptides contain a relatively long SUMO remnant chain (up to 32 amino acids for human SUMO2,3) appended on the modified Lys side chain. In contrast to other protein modifications, the SUMO remnant chain gives rise to multiple fragment ions that overlap with those of the target peptide. Accordingly, standard database search engines tailored to identify linear peptides are generally not capable of assigning the correct sequence of the branched peptides because of the complex distribution of overlapping fragment ions. This limitation was also described in an earlier report by Wohlschlegel et al. (33) who indicated that yeast Smt3 mutant with an arginine at the third residue from the C terminus yielded tryptic peptides identical to those of a ubiquitin remnant and facilitated their identification using the database engine SEQUEST. To overcome some of these limitations, different database searching strategies including an automated recognition pattern tool (SUMmOn) (34) and combined search engines (ChopNSpice) (35) were developed to identify potential acceptor sites from MS/MS spectra of these large precursor ions.

Here, we present a new approach to the large scale identification of modified peptides and their conjugation sites using three separate HEK293 cell lines, each stably expressing a SUMO (1, 2, or 3) mutant protein that contains a His$_6$ tag and a strategically located tryptic cleavage site enabling convenient affinity enrichment and MS analyses of distinct SUMO paralogs. We introduced an Arg residue at the sixth position from the C terminus to minimize structural changes with respect to the endogenous proteins. We confirmed the activity and functional properties of the His$_6$-SUMO mutants using in vitro SUMOylation assays and immunofluorescence. We profiled the subcellular distribution of SUMOylated proteins in HEK293 cells and monitored the change in SUMOylation upon As$_2$O$_3$ treatment. MS analyses of NTA-enriched nuclear cell extracts from mock and HEK293 cells stably expressing His$_6$-SUMO3 mutant enabled the identification of unique SUMOylated protein substrates and their modification sites including three novel sites on the protein PML. The separation of His$_6$-containing proteins selectively enriched the mutated SUMOylated proteins from the endogenous counterparts, although their presence in the original cell extracts did not interfere with these analyses. Moreover, we measured the effect of As$_2$O$_3$ on target modification using label-free quantitative proteomics and observed an increase in PML.
SUMOylation in nuclear extracts consistent with previous reports (25, 27). The insertion of an Arg residue at the C terminus of each SUMO paralog not only facilitated the identification of SUMOylated peptides via the observation of paralog-specific fragment ions but also reduced the abundance and distribution of overlapping fragment ions observed in MS/MS spectra of peptides with a long modified Lys side chain. In addition, the five-amo acid ACMO remnant formed during the tryptic digestion of SUMOylated proteins can also be used for immunoaffinity enrichment as recently reported for ubiquitylated tryptic peptides (36).

**EXPERIMENTAL PROCEDURES**

Plasmid Construction and Generation of Stable HEK293 Cells Expressing SUMO Constructs—cDNAs of His6-SUMO wild type (WT) and His6-SUMO mutant were generated by PCR with the forward primer (5′-gaccacagtgtacatggctcatac-3′) containing the His6 tag and KpnI and Ncol restriction sites and the reverse primers containing the stop codon and XhoI restriction site. The WT and SUMO mutant primer sequences were as follows: SUMO1 WT: forward, 5′-gacccaagcttggtaccatggctcatc-3′; reverse, 5′-ctaccgcgtcgaattacccctcgggt-3′; SUMO1 mut: forward, 5′-gacccaagcttggtaccatggctcatc-3′; reverse, 5′-ctaccgcgtcgaattacccctcgggt-3′; SUMO2 WT: forward, 5′-gaccccaagcttggtaccatggctcatc-3′; reverse, 5′-ctaccgcgtcgaattacccctcgggt-3′; SUMO2 mut: forward, 5′-gaccccaagcttggtaccatggctcatc-3′; reverse, 5′-ctaccgcgtcgaattacccctcgggt-3′; SUMO3 WT: forward, 5′-gacccaagcttggtaccatggctcatc-3′; reverse, 5′-ctaccgcgtcgaattacccctcgggt-3′; and SUMO3 mut: forward, 5′-gaccccaagcttggtaccatggctcatc-3′; reverse, 5′-ctaccgcgtcgaattacccctcgggt-3′.

The different His6-SUMO constructs were generated by inserting the cDNA of SUMO paralogs in pcDNA3 or pET28b in KpnI or Ncol/Xhol, respectively. pcDNA3-PMLIII and pcDNA3-PMLIII-YFP were used as described in Percherancier et al. (37). HEK293 stably expressing SUMO was obtained by transfection with pcDNA3 SUMO constructs and subsequent neomycin selection (0.5 mg/ml). PMLIII WT and a PMLIII 3K construct with mutations at Lys-65, Lys-160, and Lys-490 were obtained as described previously (37).

As2O3 Treatment and Antibodies—As2O3 (Sigma) was prepared in 1 M NaOH and then further diluted to 1 μM in the growth medium, and cells were typically exposed for 4 h unless otherwise specified. Rabbit polyclonal anti-SUMO1 and polyclonal anti-PML antibodies were from Santa Cruz Biotechnology, Inc. Rabbit anti-SUMO2.3 and rabbit anti-NeuN (Chemicon) antibodies were used in this study. Anti-NeuN was purified against mouse PML developed in rabbit and then purified by protein-A affinity chromatography. Rabbit anti-Lamin A/C polyclonal antibody (Santa Cruz Biotechnology, Inc.), polyclonal anti-SUMO2,3 and polyclonal anti-Lamin A/C antibodies used in this study were histone H3 polyclonal antibody (Roche), anti-poly(ADP-ribose) polymerase 1 (PARP1) antibody (Cell Signaling Technology, Inc.), and anti-poly(ADP-ribose) polymerase 1 (PARP1) antibody (Cell Signaling Technology, Inc.).

Enrichment of SUMOylated Proteins—Cells stably expressing His6-SUMO (109 cells) were lysed in denaturing buffer A (6 M guanidinium HCl, 0.1 M NaH2PO4, 0.01 M Tris-HCl, pH 8, 0.01% β-mercaptoethanol), sonicated, centrifuged at 16,000 × g, and incubated with 50 μl of NTA-agarose beads (Invitrogen) for 3 h (31). After washing in buffer A and washing five times in buffer B (8 M urea, 0.1 M NaH2PO4, 0.01 M Tris-HCl, pH 6.3, 10 mM β-mercaptoethanol), the beads were eluted with 300 mM imidazole in 0.15 M Tris-HCl, pH 6.7, 30% glycerol, 0.72 M β-mercaptoethanol. The eluates were subjected to a 4–12% NuPAGE Bis-Tris gel (Invitrogen).

Confocal Microscopy—HEK293 cells, transfected with pcDNA3-His6-SUMO and pcDNA3-PMLIII-YFP, were fixed in 4% paraformaldehyde for 10 min at 4 °C and revealed by anti-His antibody followed by Alexa Fluor 594-conjugated antibody. Confocal images were obtained on a Leica TCS-NT/SP inverted confocal laser scanning microscope using an Apochromat ×63/1.32 oil immersion objective. Co-localization experiments were performed by overlaying images using the Leica Confocal Software, LCS (Heidelberg, Germany). Excitation and emission filters for the different labeled dyes were as follows: YFP (green): λex= 488 nm; λem: 540/25 nm; Texas Red (red): λex= 568 nm; λem: 610/30 nm; and DAPI: λex= 405 nm; 10% power.

Recombinant Hisg-SUMO Protein Production and Purification—Escherichia coli BL21 cells transformed with different pET28-Hisg-SUMO-expressing vectors were induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside for 5 h. Cells were lysed in 20 mM phosphate buffer, pH 7.6, 600 mM NaCl, 30 mM imidazole by successive liquid nitrogen and 37 °C baths followed by sonication. After centrifugation, the supernatant was loaded on a 5-ml NTA HiTrap Chelating HP column (GE Healthcare). The column was washed according to the manufacturer’s instructions, and the sample was eluted using an imidazole gradient of 50–500 mM. The fractions containing most of the Hisg-SUMO recombinant protein were concentrated using Ultra Centricon with a cutoff of 30 kDa (Millipore).

In Vitro SUMOylation Assay—To the reaction buffer (20 mM NH4HCO3, pH 9, 20 mM NaCl, 0.5 mM DTT), 1 μg of recombinant SUMO proteins; 0.5 μg of substrate E2-25K (Boston Biochem), GST-RanGAP fragment 418–587 (Boston Biochem), or GST-PML fragment 485–495 (Biomol International); 0.1 μg of SASE1/SAE2 heterodimer (Boston Biochem); and 0.5 μg of conjugating enzyme hUbC9 (Boston Biochem) were added with or without 5 mM Mg-ATP (Boston Biochem). After incubation at 37 °C for 1 h, the reaction was stopped with 10 mM EDTA. The samples were analyzed by immunoblot, Coomassie staining, or silver staining and MS.

Cell Fractionation and Large Scale Purification of SUMOylated Proteins—HEK293 cells (1010 cells/replicate) stably expressing His6-SUMO mutant were lysed in hypotonic buffer (10 mM Tris-HCl, pH 7.65, 1.5 mM MgCl2, 1 mM DTT, 20 mM N-ethylmalimide, proteases inhibitors) and centrifuged at 3000 × g. The supernatant constituted the cytoplasmic fraction. The pellet was resuspended in buffer A, sonicated, centrifuged at 16,000 × g, and added to 500 μl of NTA-agarose beads for 3 h. After washing and elution steps as described above, an aliquot of the eluate was used for immunoblotting, and the rest (50–60 μg/rePLICATE) was used for MS analyses. Proteins were reduced in 0.5 mM tris(2-carboxyethyl)phosphine (Pierce) for 20 min at 37 °C and then alkylated in 50 mM chloroacetamide (Sigma-Aldrich) for 20 min at 37 °C. A solution of 50 mM dithiothreitol was added to the protein solution to react with the excess chloroacetamide. The total protein amount was quantitated by Bradford protein assay. Proteins were digested in 50 mM ammonium bicarbonate with modified trypsin (Pierce) overnight (1:300 enzyme:substrate ratio) at 37 °C under high agitation speed. The digest mixture was acidified with 0.1% TFA, incubated on an Optiguard SCX trap column (50 μm, 300 Å, 0.5 mm inner diameter, Phenomenex), and then dried down by SpeedVac prior to MS analyses.

Mass Spectrometry—All LC-MS/MS analyses were performed using an LTQ-Orbitrap Velos hybrid mass spectrometer with a nano-electrospray ion source (ThermoFisher, San Jose, CA) coupled with an Eksigent nano-LC 2D pump (Dublin, CA). Peptides were separated on an Optiguard SCX trap column (5 μm, 300 Å, 0.5 mm inner diameter × 23 mm; Optimize Technologies, Oregon City, OR) and eluted on line to a 360-μm-inner diameter × 4-mm C18 trap column prior to separation on a 150-μm-inner diameter × 10-cm nano-LC column (Jupiter C18, 3 μm, 300 Å, Phenomenex). Tryptic digests were

**Proteomics Identification of Human SUMOylated Proteins**

**Molecular & Cellular Proteomics 10.2**

10.1074/mcp.M110.004796–3
loaded on the SCX trap and sequentially eluted using salt plugs of 0, 75, 250, 500, 1, and 2 M ammonium acetate, pH 3.5. After adsorption on the C 18 precolumn, peptides were separated on the analytical column using a linear gradient of 5–40% acetonitrile (0.2% formic acid) in 53 min and a flow rate of 600 nl/min. The conventional MS spectra (survey scan) were acquired in profile mode at a resolution of 60,000 at m/z 400. MS/MS spectra were acquired in the ion trap using collision-induced dissociation (CID) only or by combining CID and electron transfer dissociation (ETD) with supplemental activation mode in a decision tree data-dependent fashion (38) for multiply charged ions exceeding a threshold of 10,000 counts. Mass calibration used an internal lock mass (protonated (Si(CH 3)2O))6; m/z 445.12057) and typically provided mass accuracy within 5 ppm for precursor ion mass measurements. ETD activation was triggered for precursor ions with m/z < 650 (3 + ), m/z < 900 (4 + ), and m/z 950 (5 + ). For ETD, the precursor cation automatic gain control target was set at 50,000, whereas a value of 100,000 was used for the fluoranthene anion population. Ion/ion reaction duration was fixed at 200 ms. The dynamic exclusion of previously acquired precursor ions was enabled (repeat count, 1; repeat duration, 30 s; exclusion duration, 120 s).

Orbitrap raw LC-MS data files were transformed into peptide maps using in-house peptide detection and clustering software (39). Peptide maps belonging to one experiment were clustered and aligned using clustering parameters of enabled (repeat count, 1; repeat duration, 30 s; exclusion duration, 5 min). Peptide clusters were aligned with Mascot identification files to assign sequence identity.

Protein Identification and Bioinformatics Analyses—MS data were acquired using the Xcalibur software (version 2.0 SR1). Peak lists were then generated using the Mascot distiller software (version 2.1.1, Matrix Science, London, UK), and MS processing was achieved using the LCC_plus zoom script. Database searches were performed against a nonredundant IPI human database containing 150,858 sequences (version 3.54, released January 2009) using Mascot (version 2.1, Matrix Science). A Mascot search against a concatenated target/decoy database consisting of combined forward and reverse versions of the IPI human database was performed to establish a cutoff score threshold of typically 25 for CID or ETD with a false-positive rate of less than 2% (p < 0.02). The error windows for precursor and fragment ion mass values were set to 0.02 and 0.5 Da, respectively. The number of allowed missed cleavage sites for trypsin was set to 1 and the number of allowed missed cleavage sites for trypsin was set to 1 and a flow rate of 600 nl/min. The conventional MS spectra (survey scan) were acquired in profile mode at a resolution of 60,000 at m/z 400. MS/MS spectra were acquired in the ion trap using collision-induced dissociation (CID) only or by combining CID and electron transfer dissociation (ETD) with supplemental activation mode in a decision tree data-dependent fashion (38) for multiply charged ions exceeding a threshold of 10,000 counts. Mass calibration used an internal lock mass (protonated (Si(CH 3)2O))6; m/z 445.12057) and typically provided mass accuracy within 5 ppm for precursor ion mass measurements. ETD activation was triggered for precursor ions with m/z < 650 (3 + ), m/z < 900 (4 + ), and m/z 950 (5 + ). For ETD, the precursor cation automatic gain control target was set at 50,000, whereas a value of 100,000 was used for the fluoranthene anion population. Ion/ion reaction duration was fixed at 200 ms. The dynamic exclusion of previously acquired precursor ions was enabled (repeat count, 1; repeat duration, 30 s; exclusion duration, 120 s).

To facilitate in vivo identification of SUMOylated proteins, we developed pcDNA3-His-SUMO expression vectors comprising strategically located mutations at the C terminus of each SUMO paralog (Fig. 1a). These mutations confer important properties to the stably expressed protein products. First, His 6-SUMO mutants with an Arg substitution introduce a convenient tryptic cleavage site on the side chain of modified Lys residues whereby individual paralogs can be identified by mass-specific signature fragment ions. Second, the short five-amino acid segment appended to the modified Lys residues (e.g. EQTGG for SUMO1 mut; Fig. 1a) result in fewer fragment ions from the Lys side chain, a property that favors the identification of SUMOylated peptides using conventional database search engines. An additional advantage of the short SUMO remnant is the availability of an epitope to which antibodies can be raised and used in large scale immunofinity experiments. In the context of this study, we devised an affinity enrichment strategy whereby SUMOylated proteins are first isolated under denaturing conditions using NTA columns prior to their tryptic digestion and subsequent identification of SUMO-modified peptides using MS (Fig. 1b).

His 6-SUMO Mutants Are Functional and Can Be Used to Monitor Protein SUMOylation in Vitro and ex Vivo—To determine that site-directed mutagenesis did not impair the transfer of His 6-SUMO mutants by the SAE1/2-Ubc9 conjugation machinery, we conducted in vitro assays using well-established protein SUMOylation substrates. We compared the SUMOylation profiles of His 6-SUMO1,2,3 WT and the corresponding mutants using RanGAP1 and the ubiquitin-conjugating enzyme E2 (E2-25K), two proteins that are SUMOylated ex vivo and in vitro (11, 40). An intact E2-25K and a GST-tagged C-terminal RanGAP1 protein fragment (amino acids 418–587) were SUMOylated in vitro as described previously (40, 41) (Fig. 2a). The silver-stained gels of the corresponding reactions indicated that all His 6-SUMO mutants showed conjugation efficiencies comparable with those of WT SUMO proteins. In the presence of ATP, almost all protein substrates were converted to the SUMOylated RanGAP1 and E2-25K products. Interestingly, we observed polySUMOylated chains for all His 6-SUMO mutants and WT proteins conjugated to RanGAP1 including SUMO1, which does not contain a ϕ5KXE consensus motif. MS analyses of in vitro digestion products confirmed the SUMOylation of RanGAP1 for all SUMO paralogs on the residue corresponding to Lys-524 in the full protein construct (supplemental Fig. S1). PolySUMOylation was identified on Lys-11 for His 6-SUMO2,3 mutants and their WT proteins. We also observed SUMOylation of His 6-SUMO1 on sites Lys-23, Lys-37, Lys-39 and Lys-48, two of which (Lys-37 and Lys-39) were previously reported by Cooper et al. (42) during in vitro experiments (supplemental Fig. S2) In vitro SUMOylation analyses confirmed the covalent attachment of His 6-SUMO1–3 mutants to Lys-14 of E2-25K (Fig. 2b). Each His 6-SUMO mutant was uniquely identified by specific fragment ions (b 2′, b 3′, b 2′ − H 2 O, b 3′ − H 2 O, etc.) arising from the cleavage of the SUMO side chain for precursor ions of m/z < 750 fragmented in the ion trap (1/6 rule for fragment ion transmission). The correlation of these unique fragment ions using extracted ion chromato-

10.1074/mcp.M110.004796–4
grams facilitated the identification of potential SUMOylated peptides in complex trypic digests. It is noteworthy that all SUMOylated proteins examined including the fusion protein GST-PML(485–495) were efficiently deSUMOylated by SENP1, thus confirming that the mutation site did not impair the enzymatic activity of the SUMO isopeptidases (supplemental Fig. S3).

We compared the ex vivo SUMOylation efficiencies of the His$_6$-SUMO mutants with their WT counterparts in HEK293 cells. We first examined the changes in SUMOylation of PML upon treatment with 1 μM As$_2$O$_3$ for 4 h. As$_2$O$_3$ is known to enhance the SUMOylation and the subsequent degradation of PML and the PML-retinoic acid receptor α fusion proteins in acute promyelocytic leukemia cells (25–28). Three
SUMOylation sites (Lys-65, Lys-160, and Lys-490) within PML have been reported previously (43), although only Lys-160 is required for As$_2$O$_3$-triggered degradation (28, 44). Immunoblots showed increased polySUMOylation of PML for the His$_6$-SUMO WT and mutants upon As$_2$O$_3$ treatment (Fig. 3a, upper panel). The increase in PML SUMOylation was also

Fig. 2. a, in vitro SUMOylation assays of RanGAP (Lys-524) and E2 ligase (Lys-14). b, MS/MS confirmation of modified Lys-14 in E2-25K for individual SUMO paralogs.
accompanied by the depletion of the unmodified PML in both WT and mutant SUMO3. The SUMOylation of PML was clearly evidenced in immunoblots from protein extracts purified using NTA columns (Fig. 3a, bottom panel). It is noteworthy that PML showed an increase SUMOylation by SUMO3 and SUMO2 compared with SUMO1, a situation

FIG. 3. **Comparison of His-SUMO WT and mutants to SUMOylate PMLIII and to co-localize with PML on NBs.** a, immunoblots of input (upper panel) and His pulldown (lower panel) of extracts from HEK293 cells co-transfected with PMLIII and WT or mutant SUMO and treated or not with As2O3. b, immunofluorescence of HEK293 cells co-transfected with YFP-PMLIII and His-SUMO WT or mutant revealed using anti-His antibody. Scale bars, 10 μm.
that results in ubiquitin-dependent proteolytic degradation of PML (45).

PML is the organizer of subnuclear structures of 0.2–1.0 μm named PML nuclear bodies (NBs). PML NBs are present in most mammalian cell nuclei and require SUMO for their formation. These NBs are involved in the regulation of different cellular processes including the induction of apoptosis and cellular senescence, inhibition of proliferation, maintenance of genomic stability, and antiviral responses (46, 47). We examined the recruitment by PML of each His6-SUMO mutant and the corresponding WT on NBs in HEK293 cells co-transfected by His6-SUMO and YFP-PMLIII constructs (Fig. 3b). Immuno-fluorescence staining of His6-SUMO mutants confirmed their co-localization with YFP-PML in multiple and dense nuclear foci characteristic of PML NBs similar to that observed for His6-SUMO WT. In addition, As2O3 induced aggregation of SUMOylated PML in NBs for both His6-SUMO1 WT and mutant compared with untreated cells (supplemental Fig. S4). Altogether, these experiments established that His6-SUMO mutants have functional characteristics similar to those of their WT counterparts.

**Subcellular Distribution and Induction of Protein SUMOylation by As2O3**—To determine the global distribution of SUMOylated proteins, we performed subcellular fractionation to isolate cytosol and nuclear extracts from HEK293 cells stably expressing mutant His6-SUMO1,3. Immunoblot analyses of these extracts using anti-His antibody revealed that a higher proportion of SUMOylated proteins was found in nuclear fractions of cells expressing His6-SUMO1 and His6-SUMO3 mutants (Fig. 4a, lanes 4 and 6). Although poly-SUMOylation chains were observed for high molecular weight bands of these two paralogs, higher polymerization levels were noted for proteins modified with His6-SUMO3 consistent with previous reports (48). Interestingly, free His6-SUMO1 and His6-SUMO3 were more abundant in the cytosol compared with nuclear extracts (Fig. 4a, lanes 3 and 5) as noted previously by Seeler and Dejean (49). It is noteworthy that anti-His immunoblots also revealed the presence of nonspecific proteins in nuclear extracts of mock HEK293 cells (Fig. 4a, lanes 1 and 2). MS analyses of these NTA-purified nuclear extracts (see below) identified several nonspecific proteins including Forkhead box and homeobox proteins, Pit-Oct-Unc (POU) domain transcription factors, histidine triad nucleotide-binding proteins that contain multi-His sequences, and zinc metal-binding proteins known to bind to Ni2+ ions (supplemental Table S1).

Overall changes in protein SUMOylation were also evaluated in NTA-purified nuclear extracts from cells treated or not with As2O3 (Fig. 4b). Enhanced protein SUMOylation was noted for both His6-SUMO1 and His6-SUMO3 mutants resulting in multiple band patterns of high molecular weight proteins for the corresponding SUMO1 and SUMO2,3 immunoblots. It is noteworthy that multimerization can be obtained with mixed SUMO chains from endogenous and mutant proteins, the distribution of which depends on their relative proportion ex vivo. Interestingly, PML immunoblots of the same NTA affinity-purified extracts clearly showed an increase in the SUMOylation of endogenous PML upon As2O3 treatment (Fig. 4b, upper panel). The corresponding banding pattern is almost superimposable to that of both SUMO1 and SUMO2,3 immunoblots, suggesting that PML represents a primary
SUMOylation substrate upon As$_2$O$_3$ treatment (Fig. 4b, bottom panel). Note that in the absence of As$_2$O$_3$, endogenous PML is barely detectable even when using longer exposure periods (data not shown). Although we identified TIF-1β, lamin A/C, PARP1, and histone H3 as proteins modified by SUMO (see below), we could not detect the SUMOylation of these endogenous proteins by immunoblots (data not shown) presumably because of the low abundance of their SUMOylated counterparts in nuclear extracts and/or the inaccessibility of epitope for antibody binding. Among all the proteins tested, only the endogenous PML showed enhanced SUMOylation in response to As$_2$O$_3$ (Fig. 4b, bottom panel).

Large Scale Identification of Protein SUMOylation—To identify SUMOylated proteins present in nuclear extracts, we performed large scale NTA affinity purification experiments from HEK293 cells expressing His$_6$-SUMO3 mutant exposed or not to As$_2$O$_3$. Similar experiments were also performed on mock HEK293 cells to identify proteins binding nonspecifically to the NTA affinity column. We typically obtained 40–60 µg of NTA-purified proteins from 10$^8$ HEK293 cells in any of the conditions and biological replicates examined. Protein extracts following NTA purification (2 µg/injection) were subjected to MS analyses using a two-dimensional nano-LC system (SCX/C$_18$) coupled to an LTQ-Orbitrap Velos instrument. Tandem mass spectra were acquired using CID and ETD in a decision tree manner to enhance the overall number of identifications (38). In total, we acquired more than 15,000 MS/MS spectra corresponding to 6282 unique peptides identified using the Mascot database search engine. To reduce the number of ambiguous identifications, we compared proteins that were identified by at least two peptides in each condition with a false discovery rate of less than 2%. By using these conservative selection criteria, we identified a total of 639 unique proteins of which 232 proteins (36%) were common to all three different cell extracts (Fig. 5a). Common proteins were assigned to nonspecific binders co-purified from NTA columns and included proteins containing multi-His sequences and zinc metal-binding proteins. It is noteworthy that some of these proteins such as E3 SUMO-protein ligase RanBP2, zinc finger protein OZF, and DNA topoisomerase 1 (Top1) were previously reported to be SUMOylated, and their fortuitous isolation from mock HEK293 NTA extract could not definitively rule out their endogenous SUMOylation. More importantly, we identified 205 proteins specific to extracts from HEK293 SUMO3 mutant (with and without As$_2$O$_3$) that comprised known SUMOylated substrates such as PML, TIF-1β, PARP1, heterogeneous nuclear ribonucleoprotein (hnRNP) isoform F, and hnRNP C1/C2. A list of proteins found in protein extracts from each condition is provided in supplemental Table S1. It should be noted that unambiguous assignment of SUMOylated proteins relies on the identification of tryptic peptides comprising the five-amino acid SUMO remnant on the modified Lys residues, a situation that only applies to a smaller subset of the potential SUMOylated proteins (see below). We also performed gene ontology analysis for terms associated with biological processes enriched in the HEK293 SUMO3 mutant extracts using the software Protein Analysis through Evolutionary Relationships (PANTHER) (http://www.pantherdb.org/). These analyses revealed that potential SUMOylated targets were significantly enriched in proteins involved in chromatin remodeling, organelle organization, and nuclear transport (supplemental Fig. S5).

The comparison of NTA-enriched protein extracts from cells expressing His$_6$-tag SUMO3 mutant enabled the identification of at least 205 SUMOylated protein candidates that were not detected in mock HEK293 cells. To identify the location of SUMOylation sites on protein substrates, we used Mascot, SUMmOn, and ChopNSpice search engines. We also developed a script to make use of the specific SUMO fragment ions to retrieve all MS/MS spectra of potential SUMOylated peptide candidates (“Experimental Procedures”). Altogether, we identified 17 unique SUMOylation sites on 12 different protein substrates from these large scale proteomics experiments (Table I). All MS/MS spectra were validated manually (supplemental Fig. S8) and comprised fragments characteristic of the SUMO3 side chain (GGTQN). A distribution of the number of identified residues according to the three different database search engines is provided in supplemental Table S2. We confirmed previously known SUMOylation sites on proteins such as PML (Lys-490) and TIF-1β (Lys-750 and Lys-779) and cross-link sites on SUMO2,3 (Lys-11 and Lys-41 for SUMO3 and Lys-42 for SUMO2). These analyses also revealed new SUMOylation sites on previously unreported nuclear substrates such as histone H3.1 (Lys-24), lamin (Lys-420), SAFB2 (Lys-524), RSF1 (Lys-287), and WIZ1 (Lys-1523) and cross-link sites with SUMO4 (Lys-11) and ubiquitin (Lys-11).

To profile proteins that showed differential regulation upon cell treatment with As$_2$O$_3$, we compared the abundance of peptide ions identified in digests of extracts from HEK293 cells expressing His$_6$-SUMO3 mutant from control and stimulated cells. The distribution of abundance for 6790 peptide ions is shown in the scatter plot of Fig. 5b and indicated that more than 92% of all ions showed less than a 3-fold change in abundance upon cell stimulation. The most significant change was observed for PML, a protein that showed more than a 15-fold increase upon As$_2$O$_3$ treatment. We obtained a sequence coverage of 43% for this protein, and several PML tryptic peptides were found to be modified with SUMO3 mutant. For example, the product ion of the quadruply protonated peptide ion ($m/z$ 523.5) acquired using ETD fragmentation is shown in Fig. 5c and confirmed the SUMOylation of the Lys-490 residue. The MS/MS spectrum is dominated by c- and z-type fragment ions from the peptide backbone and by specific fragment ions arising from the cleavage of the SUMO3 mutant side chain (e.g. $c_2^*$, $c_3^*$, and $c_4^*$). As indicated in Fig. 5b, the abundance of this peptide was increased in samples from
cells treated with \( \text{As}_2\text{O}_3 \). Residue Lys-490 is one of three known PML SUMOylation sites; the other two are Lys-65 and Lys-160 (43). However, we could not identify tryptic peptides harboring these two modified residues in any of the cell extracts examined presumably because of the relatively large molecular weight and hydrophobicity of the corresponding peptides that precluded their successful separation by C18 chromatography.

![Figure 5](image-url)
Table I

<table>
<thead>
<tr>
<th>Proteins</th>
<th>Site**</th>
<th>Statusa</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histone 3.1</td>
<td>Lys-24</td>
<td>Unknown</td>
<td>Nucleosome assembly</td>
</tr>
<tr>
<td>Histone 4</td>
<td>Lys-14</td>
<td>Unknown</td>
<td>Nucleosome assembly</td>
</tr>
<tr>
<td>HSF4B</td>
<td>Lys-288</td>
<td>Unknown</td>
<td>DNA-binding protein; binds HSE&lt;sup&gt;+&lt;/sup&gt;; regulates transcription</td>
</tr>
<tr>
<td>Lamin A/C</td>
<td>Lys-420</td>
<td>Unknown</td>
<td>Nuclear lamina</td>
</tr>
<tr>
<td>PML</td>
<td>Lys-380, Lys-400, Lys-490, Lys-497</td>
<td>Unknown/known (43)</td>
<td>Organizer of PML nuclear bodies</td>
</tr>
<tr>
<td>RSF1</td>
<td>Lys-287</td>
<td>Unknown</td>
<td>Assembly of regular nucleosome by the RSF chromatin-remodeling complex</td>
</tr>
<tr>
<td>SAFB2</td>
<td>Lys-524</td>
<td>Unknown</td>
<td>Binds to scaffold/matrix attachment region DNA; can also inhibit cell proliferation</td>
</tr>
<tr>
<td>TIF-1β</td>
<td>Lys-750, Lys-779</td>
<td>Known (56)</td>
<td>Forms a complex with a KRAB domain TF and increases KRAB-mediated repression</td>
</tr>
<tr>
<td>WIZ1</td>
<td>Lys-1523</td>
<td>Unknown</td>
<td>May link EHMT1 and EHMT2 to the CTBP co-repressor machinery</td>
</tr>
<tr>
<td>Ubiquitin</td>
<td>Lys-11</td>
<td>Unknown</td>
<td>Cross-link</td>
</tr>
<tr>
<td>SUMO2</td>
<td>Lys-11, Lys-42</td>
<td>Known (15, 42)</td>
<td>Cross-link</td>
</tr>
<tr>
<td>SUMO3</td>
<td>Lys-11, Lys-41</td>
<td>Known (15, 42)</td>
<td>Cross-link</td>
</tr>
<tr>
<td>SUMO4</td>
<td>Lys-11</td>
<td>Expected</td>
<td>Cross-link</td>
</tr>
</tbody>
</table>

<sup>a</sup> Sites refer to protein sequence including initiating Met residue. Bold type indicates a known site.

<sup>b</sup> Status indicates whether or not sites were previously reported along with the corresponding reference.

Our MS analyses also revealed the presence of three new PML SUMOylation sites at Lys-380, Lys-400, and Lys-497. Residues Lys-380 and Lys-400 were identified previously as sites of polyubiquitylation in response to As<sub>2</sub>O<sub>3</sub> (25). Site-directed mutagenesis indicated that mutation of Lys-400 delayed but did not prevent PML ubiquitylation and its subsequent proteasome-mediated degradation (25). Residues Lys-380 and Lys-400 are located between the B box domains and the nuclear localization sequence, whereas Lys-497 is next to the nuclear localization signal of PML (Fig. 6a). To confirm the identification of these new SUMOylation sites, we first examined possible site specificities of different SUMO paralogs with transfected PML and SUMOylation site mutants thereof. We compared the ex vivo SUMOylation efficiency by each SUMO WT paralog in PML III WT and a PML III 3K mutant (K65R, K160R, K490R) in extracts from HEK293 cells co-transfected with the different PML and SUMO constructs (Fig. 6b). Anti-PML immunoblots showed an increase in PML III WT SUMOylation by SUMO2 and SUMO3 when cells were exposed to As<sub>2</sub>O<sub>3</sub> consistent with results presented in Fig. 3a. This was clearly evidenced for the same protein extracts purified using NTA columns (Fig. 6b, see His pulldown). It is noteworthy that similar experiments performed with more sensitive ECL immunoblots revealed the SUMOylation of PML by SUMO1 but to a lower level than that observed for SUMO2 and SUMO3 (supplemental Fig. S6). In all cases, PML showed significantly higher SUMOylation levels by SUMO3 compared with other SUMO paralogs. In contrast, PML III 3K displayed one band in immunoblot analysis of extracts from cells co-transfected with PML III 3K and His<sub>6</sub>-SUMO WT treated or not with As<sub>2</sub>O<sub>3</sub> (see Fig. 6b, upper right panel, see input). Interestingly, NTA protein extracts of these samples indicated that residual SUMOylation of PML III 3K was observed with His-SUMO3 and that this modification was enhanced upon As<sub>2</sub>O<sub>3</sub> treatment (see Fig. 6b, middle right panel, see His pulldown). These experiments indicated that in the absence of the three known SUMOylation sites (Lys-65, Lys-160, and Lys-490) PML is still SUMOylated, although the extent of this modification is significantly lower than that observed for PML III WT. The new sites identified could thus account for the remaining SUMOylation observed in PML.

LC-MS/MS analyses of tryptic digests from NTA-purified protein extracts of HEK293 His<sub>6</sub>-SUMO3 mutant cells identified three additional PML SUMOylation sites modulated by As<sub>2</sub>O<sub>3</sub>. The extracted ion chromatograms of multiply charged ions corresponding to modified tryptic peptides at residues Lys-490, Lys-400, and Lys-380 are shown as supplemental data and yielded fragmentation patterns superimposable to those of the native peptides, thereby confirming two previously unknown SUMOylation sites on PML (supplemental Fig. S7).
The identification of SUMOylation sites on protein substrates represents an important analytical challenge because of the relatively low stoichiometry and the dynamic nature of this modification. In mammalian cells, this difficulty is further exacerbated by the large peptide remnant of SUMO paralogs left on the modified lysine residue upon tryptic digestion. In this context, the construction of SUMO paralogs that comprise a His tag and a strategically located Arg residue on the C terminus provides a convenient approach to enrich and

**DISCUSSION**

The identification of SUMOylation sites on protein substrates represents an important analytical challenge because of the relatively low stoichiometry and the dynamic nature of this modification. In mammalian cells, this difficulty is further
identify short peptides bearing the SUMO-specific remnant moiety. The judicious location of a C terminus Arg residue was important to maintain the function of SUMO paralogs while minimizing structural changes compared with endogenous proteins. Accordingly, plasmids were constructed with single base substitutions at the C-terminal end of SUMO paralogs to generate mature proteins with an Arg residue at the sixth position from the C terminus similar to that found in the yeast Smt3 protein. In vitro SUMOylation assays with well known protein substrates such as RanGAP1 and E2-25K confirmed the functionality of the His6-SUMO mutants and their efficient transfer by the SAE1/2-Ubc9 conjugation machinery. MS analyses of the enzymatic products not only determined the expected modification sites on these substrates but also identified sites of polySUMOylation for each paralog. In addition to the site Lys-11 reported previously on SUMO2 and SUMO3 proteins, we also identified polySUMOylation sites on SUMO1 at Lys-23, Lys-37, Lys-39, and Lys-48, although the physiological relevance of SUMO1 polymerization chains is presently unknown. Indeed, SUMO1, which lacks the consensus motif RXK(E/D), has not been reported to form polymeric chains, although recent studies indicated that it can cap polySUMO2,3 chains (50). Our analyses also indicated that MS/MS spectra of modified peptides afforded backbone sequence ions together with short SUMO-specific fragment ions from CID and ETD activation that can be advantageously exploited to confirm SUMOylated peptides. The presence of paralog-specific fragment ions such as m/z 240.1 (b*3 – H2O), 341.2 (b*3 – H2O) for SUMO1 mutant, and m/z 243.1 (b*3) and 343.2 (b*3) for SUMO3 mutant can be used for confirmation purposes as shown here or for targeted identification of these modified peptides via data-dependent acquisition. The relatively limited number of fragment ions originating from the modified Lys side chain also facilitated protein identification using common database search engines such as Mascot (supplemental Table S2).

The subcellular localization of SUMOylated proteins using immunoblotting and immunofluorescence experiments revealed that a large proportion of substrates are nuclear, an observation that also accounts for the significant role of this modification in transcription, DNA repair, nuclear bodies, and nucleocytoplasmic transport. This distribution is partly attributed to the enrichment of SUMO-modifying enzymes in this compartment, although a sizable number of substrates are also present in the cytoplasm, plasma membrane, mitochondria, and endoplasmic reticulum (51). We performed large scale proteomics analyses of nuclear protein extracts from mock and HEK293 cells stably expressing His6-SUMO3 mutant to identify the nature of SUMOylated substrates including those that could be regulated by As2O3. By using strict comparison criteria, we found more than 205 proteins unique to the His6-SUMO3 mutant such as proteins involved in chromatin remodeling, organelle organization, and nuclear transport (supplemental Fig. S5). Interestingly, we found several proteins involved in the regulation of ribosome biogenesis including hnRNPs, RNA helicases, and ribosomal subunits, suggesting that SUMO3 modification may regulate the assembly of these macromolecular complexes. Recent reports indicated that several of these substrates were identified in nucleolus extracts and appeared to be regulated through the ubiquitin-proteasome pathway, suggesting that SUMOylation may target unassembled ribosomal proteins for degradation (52, 53).

Our proteomics analyses also enabled the identification of several new SUMOylation sites in proteins such as H3.1 (Lys-24), lamin (Lys-420), SAFB2 (Lys-524), RSF1 (Lys-287), and WIZ1 (Lys-1523) and cross-link sites with SUMO4 (Lys-11) and ubiquitin (Lys-11). Several of the modified proteins are involved in transcription such as TIF-1β, HSF4B, and PML. Of particular interest is PML, a protein that localize to NBs where it also acts as a tumor suppressor through the regulation of p53 response to oncogenic signals (54). Quantitative proteomics revealed that PML showed more than a 15-fold increase in abundance upon cell stimulation with As2O3. In response to As2O3, PML is phosphorylated through the mitogen-activated protein kinase (MAPK) pathway leading to its transfer from the nucleoplasm to the nuclear matrix and to an increase in PML SUMOylation and NB size (44, 55). SUMOylated PML recruits the RING domain-containing ubiquitin E3 ligase RNF4, resulting in its degradation through the ubiquitin-proteasome pathway (24, 25).

Interestingly, a PMLIII 3K construct in which all three known sites of SUMOylation were mutated to Arg is still transferred to the nuclear matrix but is resistant to As2O3-induced PML degradation (44). The exact mechanism by which PML is transferred to the nuclear matrix in a SUMO-independent manner upon As2O3 treatment is still unclear but could involve its prior phosphorylation. It is noteworthy that the SUMOylated forms of PML were barely detectable when total extracts from control and As2O3-treated cells expressing PMLIII 3K and SUMO paralogs were analyzed by immunoblot with anti-PML antibody (Ref. 44 and Fig. 6b). Furthermore, we observed that NTA enrichment of protein extracts revealed residual SUMOylation of PMLIII 3K by SUMO3 and that SUMOylation of PMLIII 3K by SUMO2 and SUMO3 increased in response to As2O3. Our data demonstrated that As2O3-mediated SUMOylation of PMLIII could still occur at sites other than the three known residues Lys-65, Lys-160, and Lys-490. Detailed proteomics analyses enabled the unambiguous identification of Lys-380, Lys-400, and Lys-497 as additional SUMOylation sites regulated by As2O3 treatment (24). Further investigations are required to determine the significance of these new sites on PML functions.

The availability of functional SUMO mutants that can be stably expressed in human cells opens up new avenues for
large scale SUMOylome analyses. The enrichment of SUMOylated peptides can be achieved using a single affinity purification with an NTA column as demonstrated here or by combining an additional immunoaffinity enrichment step using specific antibodies that recognize the SUMO remnant of modified tryptic peptides. We anticipate that the combination of a dual affinity enrichment approach will yield a larger proportion of SUMOylated tryptic peptides than that achievable with NTA alone. The presence of a short paralog-specific peptide segment on the side chain of modified Lys residues provides characteristic fragment ions that facilitate the identification and confirmation of SUMOylated peptides from complex cell digests. The present approach provides a convenient tool to identify substrates modified by each SUMO paralog and to determine the nature of polySUMOylation chains on a global scale or for specific proteins separated by SDS-PAGE. This method is complementary to a recent approach described by Matic et al. (30) that uses a Lys-C resistant SUMO2 mutant that also harbors an Arg residue on the C terminus of this protein. These fragment ions can be used advantageously in the design of MS data-dependent experiments to target their identification more efficiently. Although MS/MS spectra of these modified tryptic peptides contain spectral features specific to the SUMO remnants, a database search of the corresponding peptides can be efficiently achieved using conventional search engines like Mascot. The analytical advantages of the present approach and the possibility of conducting quantitative proteomics analyses will greatly facilitate large scale experiments to unveil the complex regulation of protein SUMOylation.

Acknowledgments—We thank Dr. Brian Raught and Tharan Srikumar (University Health Network, Toronto, Canada) for valuable assistance in the identification of SUMOylated substrates using the software SUMMOn. We are most grateful to Dr. Muriel Aubry (Université de Montréal, Montreal, Canada) for valuable comments and discussions. The Institute for Research in Immunology and Cancer is supported in part by the Canadian Centre of Excellence in Commercialization and Research, the Canadian Foundation for Innovation, and the Fonds de la Recherche en Santé du Québec.

REFERENCES
