Targeting of SUMO substrates to a Cdc48-Ufd1-Npl4 segregase and STUbL pathway in fission yeast

Kõhler, Julie Bonne; Tammsalu, Triin; Jørgensen, Maria Louise Mønster; Steen, Nana; Hay, Ronald Thomas; Thon, Genevieve

Published in:
Nature Communications

DOI:
10.1038/ncomms9827

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Targeting of SUMO substrates to a Cdc48–Ufd1–Npl4 segregase and STUbL pathway in fission yeast

Julie Bonne Køhler1,*, Triin Tammsalu2,*, Maria Mønster Jørgensen1, Nana Steen1, Ronald Thomas Hay2 & Geneviève Thon1

In eukaryotes, the conjugation of proteins to the small ubiquitin-like modifier (SUMO) regulates numerous cellular functions. A proportion of SUMO conjugates are targeted for degradation by SUMO-targeted ubiquitin ligases (STUbLs) and it has been proposed that the ubiquitin-selective chaperone Cdc48/p97-Ufd1-Npl4 facilitates this process. However, the extent to which the two pathways overlap, and how substrates are selected, remains unknown. Here we address these questions in fission yeast through proteome-wide analyses of SUMO modification sites. We identify over a thousand sumoylated lysines in a total of 468 proteins and quantify changes occurring in the SUMO modification status when the STUbL or Ufd1 pathways are compromised by mutations. The data suggest the coordinated processing of several classes of SUMO conjugates, many dynamically associated with centromeres or telomeres. They provide new insights into subnuclear organization and chromosome biology, and, altogether, constitute an extensive resource for the molecular characterization of SUMO function and dynamics.
The small ubiquitin-like modifier (SUMO) can be conjugated to hundreds of protein species, following an evolutionarily conserved enzymatic cascade akin to ubiquitylation. SUMO is thereby believed to participate in the regulation of multiple biological processes by changing the interaction properties of the proteins to which it is conjugated. Prior to conjugation, SUMO is processed to expose a conserved diglycine motif (diGly) at its carboxy terminus. Conjugation is then catalysed by the sequential action of an E1-activating, an E2-conjugating and an E3-ligating enzyme, which ultimately leads to the formation of an isopeptide bond between the carboxy terminus of SUMO and the ε amino-group of a lysine in the substrate. Self-modification of SUMO at internal lysine residues can furthermore result in the formation of polymeric SUMO chains. The fission yeast sumoylating enzymes comprise the E1 heterodimer Rad31/Fub2 (Aos1/Uba2), the E2 enzyme Ubc9 (also known as Hus5) and two identified E3s of the SP-RING family; that is, the PIAS homolog PII1 and the Mms21 homolog Nse2 (ref. 3). A fundamental aspect of sumoylation is its reversible nature. Desumoylation is carried out by SUMO-specific proteases (Ulp/SNPs) that by hydrolyzing SUMO isopeptide bonds allow proteins to return to their non-sumoylated form.

An alternative to desumoylation is the elimination of sumoylated proteins through the STUbL pathway. STUbLs, as represented by the Rfp1/Slx8 and Rfp2/Slx8 dimers in Schizosaccharomyces pombe, the Slx5/Slx8 dimer in Saccharomyces cerevisiae, and homo-dimeric RNF4 in mammals, are conserved ubiquitin E3 enzymes that target sumoylated proteins for ubiquitylation and subsequent degradation or recycling. The activity of STUbLs is determined by a RING domain, commonly found in ubiquitin E3 ligases, and by several SUMO-interaction motifs (SIMs) that bind poly-sumoylated species. High-molecular weight SUMO conjugates accumulate in STUbL-deficient cells and this phenotype is associated with genomic instability and general cellular dysfunction. The best described STUbL substrate to date is the mammalian PML protein. Other reported STUbL substrates include the transcriptional regulator Mot1 (ref. 12), the inner kinetochore protein CENP1 (ref. 13), the DNA damage protein MDC1 (refs 10,14) and the cohesion α-kleisin subunit Mcd1 (ref. 15) and the SUMO ligase Siz1 (ref. 16). Given the diversity of SUMO conjugates in the cell, other proteins are likely to be targeted by STUbLs under normal or perturbed conditions. Indeed, STUbLs have been detected at centromeres, nuclear pores, replication forks and at sites of DNA damage.

A recently identified factor that also participates in the processing of sumoylated proteins is the conserved Cdc48–Ufd1–Npl4 complex. Cdc48–Ufd1–Npl4 has well-documented roles in several ubiquitin-related processes. By coupling the ubiquitin-binding properties of Ufd1–Npl4 with the ATPase activity of Cdc48 (p97 in mammals), Cdc48–Ufd1–Npl4 is believed to aid in the extraction of ubiquitylated proteins from higher order complexes. Identification of a SIM motif in both fission and budding yeast Ufd1 revealed that Cdc48–Ufd1–Npl4 is in addition capable of recognizing sumoylated proteins, suggesting the complex might bind protein species co-modified by SUMO and ubiquitin, such as produced by STUbLs. Like STUbL mutants, Ufd1 mutants also accumulate sumoylated proteins that coalesce into subnuclear foci. Physical interactions between Cdc48, Ufd1 and STUbLs further support the idea that these proteins act in concert. Cdc48–Ufd1–Npl4 may help mobilize STUbL substrates to promote their processing by the proteasome or by demodifying enzymes. Despite evidence suggesting a cooperative function, details of the substrates affected by a common STUbL/Cdc48–Ufd1–Npl4 pathway and the extent to which the two activities are coupled remain to be determined.

Mass spectrometry (MS)-based proteomics has become a widely used tool to study protein sumoylation. Hundreds of SUMO conjugates have been identified, and, recently, techniques have been developed to map individual modification sites at a large scale. This level of information is crucial to address the biological outcomes of individual sumoylation events. The identification of sumoylation sites has traditionally been hampered by an inability to enrich for SUMO modifications at the peptide level. Instead, protein-level purifications produce complex peptide mixtures from which site-specific information is difficult to extract. Moreover, the long SUMO branch left on modified peptides after trypsin cleavage gives rise to tandem MS (MS/MS) spectra that are difficult to interpret. Here we use a newly designed two-step enrichment strategy that overcomes these obstacles and permits site-specific identification of SUMO targets at a proteome-wide scale. In combination with Stable Isotope Labelling by Amino acids in Cell culture (SILAC)-based quantitative proteomics, this strategy enabled us to identify >1,000 sumoylated lysines and to compare sumoylation levels for a subset of these sites between fission yeast wild-type and mutant strains deficient in STUbL (slx8-1) or Ufd1 (ufd1ΔC213–342) function. These data suggest that certain substrates associated with centromeres and telomeres are coordinately processed by the combined action of the STUbL and Ufd1 pathways.

**Results**

**Engineering of a His$_{6}$SUMO$^{L109K}$ variant for proteomic studies.** We replaced the endogenous fission yeast SUMO gene (pmt3) with a His$_{6}$-tagged version in which the leucine at position 109 changed to a lysine (His$_{6}$SUMO$^{L109K}$ variant; Fig. 1b). A lysine at position 109 will, on digestion with the endoproteinase Lys-C, leave a diGly remnant on SUMO acceptor sites suitable for the purification of modified peptides with a diGly–Lys-specific antibody (anti-K-$\varepsilon$-GG). This antibody was originally prepared against ubiquitylated peptides for which a diGly remnant is left on trypsin digestion. As ubiquitin contains an arginine at the position upstream its diGly sequence, Lys-C will not cleave at this site. Similarly, Lys-C digestion of proteins conjugated to other ubiquitin-like proteins generates long peptide-branches not recognized by the diGly–Lys antibody (Fig. 1b). Thus, Lys-C digestion of Ni$_{2}^{2+}$-purified His$_{6}$SUMO$^{L109K}$ conjugates allows specific enrichment of SUMO diGly-modified peptides, facilitating MS analysis of the diGly-modified lysines based on their unique mass-to-charge ratio.

Functionality of the His$_{6}$SUMO$^{L109K}$ variant was verified by growth assays, microscopic examination, western blotting and MS. Cells expressing His$_{6}$SUMO$^{L109K}$ as their sole copy of SUMO, from the endogenous pmt3 locus, displayed a wild-type morphology and growth rate (Fig. 1c,d). This is in sharp contrast to the growth defects of pmt3Δ cells. The proteome of strains expressing His$_{6}$SUMO$^{L109K}$ appeared unaltered compared with cells expressing wild-type SUMO (Fig. 1e and Supplementary Data 1). Also, His$_{6}$SUMO$^{L109K}$ cells responded to heat shock as wild-type, by a massive increase in high-molecular weight conjugates (Supplementary Fig. 1a). We combined His$_{6}$SUMO$^{L109K}$ with the slx8-1 STUbL mutation and with the ufd1ΔC213–342 mutation truncating the C-terminal half of Ufd1 (ref. 19). There again, growth and morphological analyses found that slx8-1 and ufd1ΔC213–342 cells expressing His$_{6}$SUMO$^{L109K}$ were phenotypically similar to their wild-type SUMO counterparts at 30°C (Fig. 1c,d). By western blotting, overall conjugation appeared slightly reduced for His$_{6}$SUMO$^{L109K}$ as previously observed for an analogous S. cerevisiae SUMO$^{L196K}$ mutant (Fig. 1f). Reduced sumoylation has been shown to alleviate the growth defect of the slx8-1 mutant at 36°C (refs 5,7).
Figure 1 | Strains engineered to test the effects of STUbLs and Cdc48–Ufd1–Npl4 on the sumoylated proteome of fission yeast. (a) Model for the processing of sumoylated species by STUbLs and Cdc48–Ufd1–Npl4. Sumoylated proteins can be targeted for degradation by STUbL-catalysed ubiquitylation. Cdc48–Ufd1–Npl4 may act directly on sumoylated proteins, or on proteins co-modified by SUMO and ubiquitin, to facilitate their recycling or proteasomal degradation. The strains shown in b–f were designed to identify substrates in these pathways. (b–f) Design and phenotypes of the His6SUMOL109K variant. (b) Sequences of S. pombe ubiquitin-like proteins upstream of their carboxy-terminal diGly motifs. Residues highlighted in blue indicate the branch left on conjugates after endoproteinase Lys-C digestion. (c) Growth analysis of wild-type, ufd1ΔCt213–342 and slx8-1 strains expressing SUMO or His6SUMOL109K. (c, left) Tenfold dilution series of the indicated strains were spotted onto rich medium (YES) and incubated at 33 °C for 3 days. From top to bottom: JK424, JK408, JK425, JK414, JK426 and JK418. (c, right) Growth curves. Cells were propagated at 30 °C in EMM2 minimal medium with required supplements and culture absorbance was recorded at 600 nm at the indicated time intervals. S.e. were computed from three independent cultures for each strain. (d) Micrographs of strains expressing SUMO or His6SUMOL109K. (e) 2D scatter plot showing the log2 relative abundance of 2,266 proteins between the samples His6SUMO/SUMO (x-axis) or His6SUMOL109K/SUMO (y-axis). Each point corresponds to a single protein and is coloured from green to red according to the density estimation value, reflecting the density of data points from less to more compressed. Proteins that are regulated by more than twofold are coloured in purple, however, none of these changes are statistically significant according to the multiple-sample analysis of variance test (Supplementary Data 1). (f) Western blot of wild-type SUMO and His6SUMOL109K conjugates. Whole-cell extracts of the indicated strains were probed with a SUMO antibody. Tubulin was used as a loading control. Cultures were propagated at 30 °C unless indicated otherwise. Asterisks indicate cross-reactivity of the SUMO anti-serum.
and indeed this was also the case with \textit{slx8-1} His6SUMO\textsuperscript{L109K} cells (Supplementary Fig. 1b). However, the sumoylation patterns detected in strains expressing His6SUMO\textsuperscript{L109K} or SUMO were globally very similar (Fig. 1f), in particular, the changes in sumoylation occurring in the \textit{slx8-1} and \textit{ufd1\textit{AC}^{213–342}} mutants compared with wild type were conserved, pointing to the suitability of His6SUMO\textsuperscript{L109K} to globally map sumoylation sites of fission yeast wild-type and mutant cells.

\textbf{Proteome-wide identification of sumoylation sites.} To identify the sumoylated species that accumulate in Ufd1 mutants\textsuperscript{19,20}, we isolated His6SUMO\textsuperscript{L109K}-modified proteins from \textit{ufd1\textit{AC}^{213–342}} cells. Cells were lysed under denaturing conditions and His6SUMO\textsuperscript{L109K} conjugates were isolated by Ni\textsuperscript{2+} - chromatography. The purified material was digested with Lys-C, or Lys-C and Glu-C to generate shorter peptides, and diGly-Lys peptides were immunoprecipitated with anti-K-e-GG antibody prior to MS analysis (Fig. 2a). This strategy identified a total of 841 unique diGly-modified lysines mapped with a localization probability of > 0.75. Together these sumoylated sites defined a total of 414 proteins (Supplementary Data 2).

The His6SUMO\textsuperscript{L109K} variant was then employed in conjunction with SILAC-based proteomics\textsuperscript{33} to quantitatively compare the sumoylated proteomes of wild type and mutants defective in Ufd1 (\textit{ufd1\textit{AC}^{213–342}}) or STUBL (\textit{slx8-1}). Two SILAC experiments were conducted in parallel, one experiment comparing a heavy lysine (Lys8) \textit{slx8-1} culture with a light lysine (Lys0) wild-type culture, and the other experiment comparing the Lys8 \textit{slx8-1} culture with a Lys0 \textit{ufd1\textit{AC}^{213–342}} culture (Fig. 2b).

The experiments were conducted in triplicate starting from isolated colonies, propagating cells at 30°C and shifting them to 33°C for the last 12 h to induce the \textit{slx8-1} temperature-sensitive phenotype at the semi-permissive temperature of 30°C (ref. 5). The \textit{slx8-1} cultures were split, mixed separately in 1:1 ratios with wild-type and \textit{ufd1\textit{AC}^{213–342}} cultures, and lysed. Small samples of cell lysate were subjected to in-solution digestion with Lys-C for total proteome analysis, while His6SUMO\textsuperscript{L109K}-conjugated proteins and diGly-modified peptides were isolated from the remaining material. In total, the SILAC experiments detected 618 distinct sumoylation sites in 301 proteins. About 400 sites identified in the \textit{ufd1\textit{AC}^{213–342}} strain had already been identified in the pilot experiment performed under non-SILAC conditions, indicating a high degree of experimental reproducibility (Supplementary Fig. 2). The higher complexity of SILAC peptide samples, where very intense SILAC pairs can hinder detection of less abundant species, might account for the fact that fewer sites were identified in the SILAC experiments than in the non-SILAC experiment even though each experiment used approximately the same amount of starting material. Together, the SILAC and non-SILAC experiments produced a comprehensive list of 1,028 sumoylation sites distributed between 468 proteins. This list is shown in Supplementary Data 2, the SILAC results are shown in Supplementary Data 3, and the total proteome analyses are shown in Supplementary Data 4.

Comparing the effects of Ufd1 and STUBLs with SILAC. Among the diGly-modified peptides identified in the SILAC experiments, only peptides identified in both heavy and light forms in at least two biological replicates of either \textit{slx8-1}/wild type or \textit{slx8-1}/\textit{ufd1\textit{AC}^{213–342}} were used for quantification. This resulted in the calculation of heavy-to-light ratios for 327 distinct peptides (Supplementary Data 3). For the total proteome samples, peptides quantified in at least one biological replicate lead to the quantification of 3,115 crude proteins (Supplementary Data 4). SILAC ratios for crude proteins were normalized using the MaxQuant software\textsuperscript{36}, while the mean of their non-normalized ratios was used to normalize the ratios from the purified material\textsuperscript{37}. Normalized SILAC ratios were used to derive median values for peptides measured in 3 biological replicates, or mean values for peptides quantified in only 2 biological replicates, resulting in the quantification of 243 and 261 diGly-modified peptides in the \textit{slx8-1}/wild type and \textit{slx8-1}/\textit{ufd1\textit{AC}^{213–342}} experiments, respectively (Supplementary Data 3). As 171 of these peptides were quantified for both \textit{slx8-1}/wild type and \textit{slx8-1}/\textit{ufd1\textit{AC}^{213–342}} (Fig. 2c) their relative abundance in \textit{ufd1\textit{AC}^{213–342}} and wild type could also be calculated. Figure 2d compares the \textit{slx8-1}/wild type and \textit{ufd1\textit{AC}^{213–342}}/wild-type ratios. Sumoylation was increased ≥2-fold for 51 sites (in 33 proteins) in the \textit{slx8-1} mutant and for 40 sites (in 31 proteins) in the \textit{ufd1\textit{AC}^{213–342}} mutant. Conversely, sumoylation was reduced ≥2-fold for 14 and 21 sites in the 2 mutants, respectively. Regression analysis showed the two data sets are correlated with a Pearson coefficient of \(r = 0.49\). Even with a great degree of variation, this supports the existence of common tendencies in sumoylation dynamics in the two strains.

\textbf{Comparison of sumoylation and total protein levels.} Should proteins with increased sumoylation in \textit{slx8-1} and/or \textit{ufd1\textit{AC}^{213–342}} be normally targeted for degradation by STUBL-mediated ubiquitylation, these proteins may be stabilized in the mutant backgrounds. We could quantify protein levels for a subset of proteins with quantified sumoylation sites (112 and 79 proteins were quantified in both crude and diGly-Lys immunoprecipitated samples for \textit{slx8-1}/wild type and \textit{ufd1\textit{AC}^{213–342}}/wild type, respectively). For these proteins, changes in sumoylation were not generally correlated with a change in protein abundance (Fig. 2e,f).

Importantly, the abundance of SUMO itself was unchanged in the mutants excluding the possibility that protein sumoylation is increased in the \textit{slx8-1} or \textit{ufd1\textit{AC}^{213–342}} mutant due to elevated SUMO concentration. Two exceptions were Stl1, an uncharacterized orphan protein, which displayed a nearly twofold increase in \textit{ufd1\textit{AC}^{213–342}} cells compared with a threefold increase in sumoylation, and Smc5, which increased at the protein level by nearly threefold in both mutants.

\textbf{General features of the SUMO proteome at the residue level.} As our study is the first global analysis of sumoylation in \textit{S. pombe}, and one of the very first to identify sumoylation sites globally in any organism, it also provided important insights into sumoylation \textit{per se}. Nearly half the proteins identified in our combined experiments displayed multiple sumoylation sites (Fig. 3a). Some proteins contained many sites with, for example, 14 sites identified in Uba2, 17 sites in the HMG-box protein Hmo1 or 22 sites in the topoisomerase Top2. We identified seven SUMO–SUMO branching points (K14, K30, K39, K51, K60, K63 and K71, with co-sumoylation of K14 and K30), and multiple sumoylation sites in ubiquitin (K6, K11, K27, K48 and K63). Doubly or triply modified peptides were detected for 22 proteins even though our approach detects co-modification only for sequential lysines, pointing to widespread multi-sumoylation. Altogether these observations are consistent with features of sumoylation recently proposed for mammalian cells\textsuperscript{29–31}. Half of the sumoylation sites identified conformed to the forward (\textit{PKXED})\textsuperscript{32} or reversed (D/ExK)\textsuperscript{27} consensus motif.
proposed in other organisms, while the remaining sites did not conform to any identifiable motif (Fig. 3b).

In addition to multiply sumoylated peptides, sumoylated peptides bearing other post-translational modifications were also detected (Supplementary Data 6), including branched SUMO peptides phosphorylated on T20 and T24. Co-modifications suggestive of cross-talks were observed for the histones H3, H2A and H2B. In particular, concomitant phosphorylation of S121 and sumoylation of K119 occurred on H2A-alpha and -beta. H2AS121 is phosphorylated by Bub1, mediator of the spindle assembly checkpoint. H2AS121 phosphorylation might facilitate sumoylation at K119 by introducing a negative charge.

Figure 2 | Quantitative identification of sumoylated residues in wild-type and slx8-1 and ufd1D C t213–342 mutants. (a) Two-step enrichment for sumoylated peptides using His6SUMO109K strains. (b) Experimental design for SILAC experiments (see text for details). (c) Overview of peptides quantified in the triplicate SILAC experiments. Only diGly–Lys peptides for which SILAC ratios were obtained in at least two biological replicates were used in the analyses. In total, 171 distinct diGly–Lys peptides were quantified in both the slx8-1/ wild-type and slx8-1/ufd1D C t213–342 IPs, which enabled us to calculate ufd1D C t213–342/wild-type ratios. See also Supplementary Fig. 2. (d) The normalized log2 ratios of quantified diGly–Lys-modified peptides for ufd1D C t213–342/wild type and for slx8-1/wild type were plotted on a tsMAP, giving an overview of the relative change in abundance of individual peptides in each mutant. The two data sets correlate with a Pearson coefficient of r = 0.49. Coloured data points represent peptides originating from the same proteins or from functionally related proteins, identified in Fig. 5. (e, f) Changes in protein sumoylation are not generally correlated with changes in protein concentration in either the slx8-1 or ufd1D C t213–342 mutant. Among the proteins with quantified sumoylation sites, 112 and 79 were also quantified in the total proteome analyses of crude lysates for slx8-1/wild type and ufd1D C t213–342/wild type, respectively. Log2 SILAC ratios for sumoylated peptides were plotted against quantified ratios obtained for these proteins, yielding the maps in e and f.
**Figure 3 | General features of sumoylation at the residue level.**

**a)** Distribution of sites between sumoylated proteins. **(b)** Amino-acid context of sumoylated lysines. Four pLogo graphs display the amino-acid sequence context of all sumoylated lysines; of KxE/KxD consensus-site lysines; of lysines within the ExK/DxK inverted consensus site and of sites not conforming to either consensus (residual sites). The numbers in parentheses above each graph indicate the total number of sites in the category (first number) and, among them, the number of sites that covered the full 13-amino acids window required for the pLogo analysis (second number). The log-odds of the binomial probability (\( p \)) are represented on the y-axis. Significance threshold values of 3.68 (\( P < 0.05 \)) are shown in red. **(c)** Sumoylated residues identified in this study are shown in yellow in a cartoon structure of *S. cerevisiae* topoisomerase 2 (PDB#4GFH; 94% identical to *S. pombe*). The portion of the protein whose structure was solved is coloured in the *S. cerevisiae* Top2 representation on the right. Enlarged cartoon views show the tendency of sumoylation sites to fall into coils or end of helices, the proximity of sumoylation sites in three-dimension and the proximity of sites to DNA. **(d)** Distribution of 82 sumoylated lysines identified in this study and mapped onto existing crystal structures using PyMOL. The structures are shown in Supplementary Figs 4 and 5 and the sites are listed in Supplementary Data 7. ‘End-of-helix’ and ‘end-of-sheet’ refer to lysines at three or less amino acids from the end of a helix or sheet. **(e)** Distribution of lysines in secondary structures predicted with PSIPRED\(^70\) for respective sumoylated lysines, and for all the lysines in proteins with a sumoylated lysine. *P* < 0.001. P values were calculated with a hypergeometric probability test. See also Supplementary Data 8.

---

**ARTICLE NATURE COMMUNICATIONS | DOI: 10.1038/ncomms9827**

© 2015 Macmillan Publishers Limited. All rights reserved.
NATURE COMMUNICATIONS | DOI: 10.1038/ncomms9827

Functional clustering of sumoylated proteins. Gene ontology analyses of the ensemble of identified SUMO targets (Supplementary Data 2) found an over-representation of proteins participating in chromosome organization and segregation, transcription, as well as DNA recombination and repair (Fig. 4a). Also enriched were proteins associated with ribosome biogenesis and nucleo-cytoplasmic transport. These enrichments are not participating in chromosome organization and segregation, transcription, as well as DNA recombination and repair (Fig. 4a).

Corroborating another hallmark of sumoylation, many of the S. pombe proteins targeted for sumoylation were part of heavily sumoylated macromolecular complexes (Supplementary Data 2 and Fig. 4b). For example, the SWI/SNF chromatin-remodelling complex, the histone acetylase SAGA complex and the TFIIID transcription factor and RNA polymerase core complexes were sumoylated on multiple subunits. The nucleolar U3 sno-RNA-containing complex was sumoylated at 15 residues, on 9 subunits. Other complexes highly targeted by sumoylation were the core cohesion, SMC5/6, and condensin complexes, with the latter modified on all 5 components with a total of 17 sumoylation sites.

Functional classes of conjugates targeted by STUbL or Ufd1. The SILAC data sets (Supplementary Data 3) were analysed to identify cellular functions relying on the processing of SUMO conjugates by STUbLs or Cdc48–Npl4–Ufd1. The peptides whose sumoylation was most affected by the slx8-1 or ufd1AC213–342 mutants are displayed in bar graphs in Fig. 5. Peptides were loosely grouped in the graphs according to whether their sumoylation was increased in both mutants (Fig. 5a; the largest and most strongly affected class), increased specifically in one mutant (Fig. 5b) or decreased in either mutant (Fig. 5c). Peptides originating from the same proteins were placed side-by-side, revealing strongly correlated patterns of sumoylation within proteins. Some proteins contained sites whose sumoylation was below the cutoff used in Fig. 5; in general sumoylation at these sites changed with the same trends as for the peptides on display, just with a reduced amplitude (Supplementary Data 3). Peptides whose abundance was measured only in the slx8-1/wild-type SILAC experiment are shown in Fig. 5d. Several of these peptides originated from proteins represented in the upper graphs, in all cases corroborating the observed changes.

The function of the proteins whose sumoylation status strongly depends on Slx8 or Ufd1 was determined using the Pombase43 and Pubmed databases, allowing us to annotate the graphs in Fig. 5 accordingly. Protein localization data43 were also used, although not systematically displayed. Functions could be assigned for most proteins. Colour was added in both Fig. 5 and Fig. 2d to label peptides according to their functional groups.

Remarkable functional clustering was observed. More than a third of the proteins whose sumoylation coordinate was increased in both mutants (7/20; Fig. 5a) are proteins associated with centromeres or telomeres. Figure 6 displays all the SUMO conjugates identified in our study localizing to these chromosomal regions and labels conjugates quantified in the SILAC experiments. In total, the sumoylation of 17 centromere- or telomere-associated proteins was quantified in both mutants. The sumoylation status of 13 of them was significantly increased in Ufd1 or Slx8 (Fig. 6a).

It was also notable that condensins and cohesins constituted more than half of the proteins whose sumoylation was specifically increased in the slx8-1 mutant (Fig. 5b), representing nearly all condensins and cohesins measured in the SILAC experiments. The sumoylation state of numerous transcription factors and chromatin remodelers also changed in the mutants, albeit differently. SUMO branches increased greatly in both mutants (Fig. 5a). Sumoylation of ubiquitin increased significantly inside the ufd1AC213–342 mutant (Fig. 5b).

Comparison with sumoylation in S. cerevisiae. We compared our S. pombe data (Supplementary Data 2) with five S. cerevisiae large-scale proteomic studies of sumoylation42,45–46 in YeastMine47, filtering for proteins that have orthologues in both species (Supplementary Fig. 6 and Supplementary Data 9). As noted previously, the five studies are in poor concordance with each other. Comparing them individually with our more comprehensive data sets indicates the poor concordance is partly due to limited coverage. Merging four of the S. cerevisiae lists finds that sumoylation is conserved between the two species for at least 42% of the 412 S. pombe proteins listed in Supplementary Data 2 for which orthologues were found in S. cerevisiae.

The study by Albuquerque et al. is of particular interest since it identifies potential STUbL substrates in S. cerevisiae, using the slx5A mutant. Comparing with the potential substrates identified here for Slx8 reveals strong evolutionary conservation for some classes of substrates in particular cohesins, condensins and for the Smc5/6 complex (Supplementary Fig. 7). The kinetochoore and DASH complex components identified here in the case of Slx8 were not identified for Slx5, but another kinetochoore protein, CBP2, and regulators of kinetochoore-microtubule attachment (SL15 and BIR1) were identified, suggesting STUbLs and Cdc48–Ufd1–Npl4 might have conserved roles in chromosome segregation. Divergence was observed for septins, a major class of proposed substrates for S. cerevisiae STUbLs, for which sumoylation appears much less prominent in S. pombe, at the +2 position and effectively creating a consensus modification site.

To investigate the structural features of SUMO modification sites in proteins, we mapped sumoylation sites onto known protein structures. Of the 22 sites in Top2 (Fig. 3c), we mapped 10 to the DNA gate, with 6 at the interface of the ATnPase domain and DNA-binding and cleavage core (Fig. 3c). One sumoylated lysine (K388 in S. pombe or K333 in S. cerevisiae) is in the K loop, whose interaction with DNA has been proposed to trigger ATP hydrolysis and release of DNA from Top2 (ref. 39). In addition, 12 sumoylation sites are in the last 302 amino acids of S. pombe Top2, in a part of the protein not present in the crystal structure but whose sumoylation is also conserved29,30,40 and likely to interfere with the functioning of the C gate. Supplementary Figure 4 displays sites in Cdc48 and in the nucleosome, and Supplementary Fig. 5 displays a total of 73 additional sites in 9 S. pombe proteins and 34 highly conserved proteins from other organisms (82 sites in total; Supplementary Data 7). Several conclusions could be reached from this analysis. First, nearly all sites were accessible at the surface of proteins, indicating sumoylation occurs on folded proteins. Second, sites located in different domains of a protein could be very close in three-dimensional structures as for K388 and K978 in Top2, or in the case of the nucleosome where modification sites on distinct histones (H3K79 and H4K79) are close together in the nucleoprotein complex. Finally, sites were predominantly in loops or coils or at the very end of alpha helices or beta sheets (Fig. 3d). Consistently, a PSI-PRED secondary structure prediction for the proteins in our list (Supplementary Data 2) found sumoylated lysines preferentially in coils, with an under-representation in helices and coils or at the very end of alpha helices or beta sheets (Fig. 3d). Consistently, a PSI-PRED secondary structure prediction for the proteins in our list (Supplementary Data 2) found sumoylated lysines preferentially in coils, with an under-representation in helices (Fig. 3e and Supplementary Data 8).
possibly reflecting diverged modes of septation between fission and budding yeast.

**Discussion**

Using a recently developed purification strategy, we have analysed the SUMO proteome of fission yeast wild-type and mutant cells, identifying over a thousand unique sumoylation sites in 468 different proteins in conditions where SUMO was expressed at endogenous levels. Our data sets highlight the participation of SUMO in higher order protein complexes and the widespread links between SUMO and chromosome transactions. They provide a quantitative overview of the effects of STUbLs and Ufd1 on the sumoylated proteome, showing that both STUbLs and Cdc48–Ufd1–Npl4 target discrete classes of sumoylated proteins and identifying potential substrates sequentially processed by STUbLs and Cdc48–Ufd1–Npl4. The modification status of SUMO itself, and of ubiquitin, suggest a sequential action of Slx8 and Ufd1 that would result in the accumulation of SUMO–SUMO linkages in both mutants, and of mixed ubiquitin–SUMO chains specifically in the Ufd1 mutant. Some of the protein groups affected individually or in a correlated manner by STUbLs and Ufd1 are discussed below.

Many potential targets of a STUbL/Ufd1 pathway were at centromeres/kinetochores, telomeres and at the nuclear envelope (NE) (Figs 5a and 6a). They include Mis17, Ask1 and Dam1. Mis17 associates with the inner kinetochore plate together with the Mis6–Mal2–Sim4 complex (CENP-I/-O/-K homologues) throughout the cell cycle48, while Ask1 and Dam1 are part of the DASH complex that binds the outer kinetochore only during mitosis, to facilitate the attachment of kinetochores to the mitotic spindle and subsequent chromosome segregation49,50. In human cells, RNF4-mediated turnover of the inner kinetochore homologue CENP-I has been proposed to help form the inner kinetochore plate13. Our finding of elevated Mis17 sumoylation in STUbL and Ufd1 mutant cells...
indicates a conserved role for STUbLs and Ufd1 in regulating inner kinetochore structures. STUbL/Ufd1-mediated turnover of outer kinetochore components may similarly contribute to normal spindle and chromosomes dynamics, possibly accounting for some of the mitotic defects observed in STUbL mutants. 

**Figure 5** | Sumoylated peptides with a $\geq 2$-fold change in abundance in the slx8-1 or ufd1/DCT213-342 mutant. Colours highlight peptides originating from the same proteins or from proteins with related function. The same peptides are coloured in the tsMAP in Fig. 2d. (a) Peptides of increased abundance in both mutants. (b) Peptides of increased abundance in single mutants. (c) Peptides of decreased abundance in either mutant. (d) Peptides quantified for only slx8-1/wild type.
At telomeres, sumoylation of the shelterin proteins Rap1 and Ccq1 was increased in the three to sixfold range in both mutants (Fig. 5a,d). The shelterin complex protects the ends of chromosomes and regulates their replication by telomerase. In *S. cerevisiae*, non-functional poly-sumoylated versions of Rap1 are cleared by Uls1, a dual STUbL and DNA-dependent ATPase/translocase, to sustain Rap1-dependent inhibition of telomeric end-to-end fusions53. A similar function could be carried out by the combination of Ssk8/Rtp STUbL and Cdc48–Ufd1–Npl4-dependent ATPase activity.

Not only were kinetochore and shelterin components heavily sumoylated in the *slx8-1* and *ufd1ΔCt213–342* mutants, but their anchors at the nuclear periphery were also increasingly sumoylated in the Ud1 mutant. In *S. pombe*, the association of telomerases with the NE relies on interactions between telomeric shelterin proteins (Taz1 and Rap1) and the inner nuclear membrane proteins Bqt3 and Bqt4 (refs 52,53). During mitosis, these associations are released to allow faithful chromosome segregation. Other factors implicated in telomere–NE anchoring include the integral nuclear membrane protein of the animal LEM (Lap2/Emerin/Man1) subfamily of lamin-associated proteins, Lem2 and Man1 (ref. 54). We detected more than fourfold increase in the sumoylation of Lem2 at K121, a lysine situated in the nucleoplasm, in the *ufd1ΔCt213–342* background. Additional five sumoylation sites, clustered near the amino-terminal chromatin-binding Helix–Extension–Helix domain of Lem2, were detected exclusively in the *ufd1ΔCt213–342* strain (Fig. 6b and Supplementary Data 2). In addition, a total of nine sumoylation sites were identified in Bqt4; another, solely in *ufd1ΔCt213–342*. Telomere anchoring to the nuclear periphery is known to depend on sumoylation in *S. cerevisiae* and Caenorhabditis elegans, but the targets of sumoylation are not precisely defined55,56. Our findings suggest undiscovered modes of telomere–NE regulation through sumoylation of inner membrane proteins. Such regulation is likely to also apply to other heterochromatic regions associating with the NE, such as centromeres. For instance, Lem2 is highly enriched at the centromere/spindle pole body region54 and we detected several sumoylation sites on the two-centromere NE-anchoring proteins Sad1 (SUN domain protein) and Csi1 (Fig. 6b). Again, Sad1 and Csi1 sumoylated peptides were almost exclusively detected in the *ufd1ΔCt213–342* background. Since the Cdc48–Ufd1–Npl4 ATPase-driven translocation of membrane-bound proteins have been well-established at the endoplasmic reticulum, one may speculate on similar roles for Cdc48–Ufd1–Npl4 in the turnover of NE-associating proteins, possibly via Ufd1 SUMO–STi2L interactions.

In addition to proteins associated with centromeres or telomeres, multiple chromatin modifiers and transcription factors were targeted by both Ssk8 and Ufd1 including Mot1, a Swi2/Snf2 family ATPase that displaces the TATA-box-binding protein Tbp1 from promoters. Mot1 is an established STUbL substrate in *S. cerevisiae*12. We observed increased sumoylation of *S. pombe* Mot1 at K101 and K105 in *slx8-1* and *ufd1ΔCt213–342* mutants (Fig. 5a), in a domain of the protein corresponding to the region sumoylated in *S. cerevisiae*. This implicates Cdc48–Ufd1–Npl4 in the processing of sumoylated Mot1. We also detected sumoylation of Tbp1 at K209, within the DNA-binding domain (Supplementary Fig. 5). Other global controllers of transcription and chromatin structure co-regulated by Ssk8 and Ufd1 included the SWI/SNF complex subunit Snf59—the histone H3K9 methyltransferase Clr4, the co-repressor Tup12 and the iron-responsive transcription factor Fep1 known to associate with Tup12.

The accumulation of other classes of SUMO conjugates was specific to the *slx8-1* mutant. This was the case for the DNA-directed RNA polymerase I and III complex subunits, for the Cnd2, Cut3 and Cut14 condensin proteins, for the Psm1 and Psm3 core cohesion proteins and for the nuclear protein Dnt1. Strikingly, Dnt1 showed 1.5- to 7-fold increased sumoylation on 7 different lysines in *slx8-1*. The *S. cerevisiae* Dnt1 homologue, Tof2, is an established SUMO target. Tof2 is part of a network of proteins that tether the ribosomal DNA (rDNA) to the inner nuclear membrane and sumoylation may modulate these interactions57. Proper localization of the rDNA at the nuclear periphery is required for retaining rDNA structure and silencing. Consequently, loss of sumoylation leads to decreased rDNA silencing58, and an inability to form SUMO chains has been linked to gross defects in rDNA compaction/organization58. Moreover, rDNA silencing is also alleviated in STUbL-deficient budding yeast cells39, suggesting that a pathway involving both polysumoylation and STUbL activity is required for normal rDNA function. The identification of Dnt1 as a highly upregulated SUMO target in *slx8-1* cells suggests that Dnt1 may constitute a substrate for such a pathway.

Increased expression of the Psm1 and Psm3 core cohesion proteins in *slx8-1* is particularly interesting in light of a recent study conducted in *S. cerevisiae*. Here the other core cohesion subunit Mcd1 (Rad21 homologue) was shown to be targeted for STUbL-mediated degradation in the absence of the cohesin-associated factor Pds5, leading to premature sister chromatid separation15. By promoting cohesin proteolysis, STUbLs may allow efficient chromosome segregation during anaphase in wild-type cells15 and/or they might regulate cohesion functions outside mitosis such as during DNA repair where the sumoylation of cohesins plays important, yet incompletely defined, roles40.

It may be expected that sumoylated proteins normally targeted for degradation by STUbL-catalysed ubiquitylation would accumulate in the *slx8-1* and *ufd1ΔCt213–342* backgrounds. However, we found changes in protein sumoylation generally not to be correlated with a change in protein abundance. This lack of correlation, as determined for at least a subset of *slx8-1* and/or *ufd1ΔCt213–342*-regulated SUMO substrates, is in line with the notion that only a minor fraction of a protein is typically modified by SUMO at any given time28,61. Consequently, the total pool of a protein likely to be subjected to STUbL-Ufd1-mediated regulation is proportionally low and any change in total concentration of the target protein may thus go undetected. It is also possible that ubiquitin is added by STUbLs not to degrade sumoylated proteins (that is, K63 rather than K48-linked chains), but rather to allow their extraction from large complexes.

In summary, our SILAC experiments identified a number of proteins whose sumoylation is increased in the *slx8-1* and/or *ufd1ΔCt213–342* backgrounds, of which only some were discussed above. We cannot exclude that some conjugates might accumulate through indirect effects of the *slx8-1* or *ufd1ΔCt213–342* mutations, for instance, effects on the cell cycle might result in the enrichment of certain cell-cycle-regulated SUMO conjugates. Conversely, our data set might not include particular SUMO-modified peptides, whose chemical composition precludes identification using standard MS-based bottom-up proteomics approaches. This might be the case for a predicted sumoylated peptide62,63 that was not identified in this study. We trust that the proteins identified include bona fide STUbL substrates, based on the finding of already known STUbL targets in the list, and of proteins localizing to subcellular structures or protein complexes previously shown to be affected by STUbL activity. Comparatively few sites displayed reduced sumoylation in either or both mutants. Reductions were limited in both amplitude and number of sites affected, with the remarkable exception of Hmo1 and a predicted ZF-GATA-type transcription factor (SPCC1393.08), which together account for nearly all sites of
reduced sumoylation in both the slx8-1 and ufd1ΔCt213–342 mutants (Fig. 5). Common tendencies in sumoylation dynamics between the slx8-1 and ufd1ΔCt213–342 mutant strains lend support to recent genetics observations of a functional relationship between STUbLs and the Cdc48–Ufd1–Npl4 segregase, and should motivate the further characterization of the specific substrates through which STUbL and Cdc48–Ufd1–Npl4 act either in concert or individually.

Methods

Strain constructions. The strains used in this study are listed in Supplementary Table 1. The endogenous S. pombe SUMO gene, pmt3, was engineered to encode a SUMO variant with an N-terminal His6-tag and a single amino-acid substitution changing the leucine at position 109, immediately upstream the C-terminal diGly sequence, into a lysine (His6SUMOL109K allele). This was performed using a 'loop-in-loop-out' strategy, where a plasmid containing His6SUMOL109K, pmt3+, 5' and 3' flanking regions and a ura4+ gene (pJBK142) was integrated into the pmt3-flanking region by homologous recombination in a ura-D18 strain. 'Loop-outs' where His6SUMOL109K remained as the single pmt3 genomic copy were subsequently selected by 5'-fluoroorotic acid resistance and they were verified by PCR amplification of the locus and sequencing. To produce the pJBK142 'loop-in' plasmid, the S. pombe SUMO (pmt3) open reading frame and its flanking regions were amplified in two separate PCR reactions: The open reading frame and its 5' region were amplified with the forward primer GTO-492 (5'-GCTCCAAATTCCTCACAAGGCACCC-3') and reverse primer GTO-548 (5'-GCATGGATCCCTAAAGGCATAGATGGGTGCAACCACC-TTTCTGTTCTAAG-3'), changing the codon for leucin 109 (TTA) to a lysine codon (TTT), as underlined. The pmt3 downstream region was amplified with GTO-549 (5'-GCATGGATCCAGTACCACTGGTACGACGGAGTAC-3').

Figure 6 | Sumoylated proteins targeted by Slx8 or Ufd1 at centromeres, telomeres and nuclear envelope. Only proteins for which sumoylation sites were identified in this study are represented. (a) Examples of interacting protein complexes with multiple sumoylated subunits. The number of sumoylation sites are indicated in brackets. S. pombe centromeres cluster at the spindle pole body (SPB) in interphase cells through inner and outer nuclear membrane (INM and ONM) proteins. Proteins associating with the centromeres and kinetochores of mitotic chromosomes are displayed on the right. The sumoylation of circled proteins was quantified in the SILAC experiments. Proteins with increased sumoylation specifically in slx8-1, specifically in ufd1ΔCt213–342 or in both mutant backgrounds are highlighted in orange, purple or pink, respectively. (b) Schematic representation of sumoylation sites identified in nuclear-envelope-associated proteins involved in centromere and/or telomere anchoring. Yellow circles specify sumoylated lysines. Transmembrane domains are indicated with blue squares. Sumoylation sites were exclusively detected in the nucleoplasmic parts of the proteins, and almost exclusively in the ufd1ΔCt213–342 mutant. HeH, helix-extension-helix domain; SUN, Sad1p UNC-84 domain.)
CAAGTATTTAAGCTGTTTC-3' and GTO-550 (5'-GATCCATGCGGCC CAGCTCAAGGATCCTGATCGACG-3'). Prior to chromosomal integration the plasmids were digested with SspI. After SUMO109K allele replacements had been identified, a ura4+ gene was inserted at a tightly linked location 5' of pmr1 to follow the allele through subsequent crosses. HisSUMO109K strain PBK111 (ura4) was crossed into slx8-1 and ufl1AC13-342 mutant backgrounds and combined with the lys3-27 allele using standard genetic crosses.

**Anti-SUMO western blots.** The indicated strains were grown in yeast-extract medium at 30 °C and harvested at an A600nm < 0.5. Cells were lysed with the glass bead method in a denaturing buffer containing 500 mM NaCl, 100 mM Tris-HCl pH 7.5 and 50 mM -NTA agarose beads (Qiagen) pre-equilibrated in binding buffer, and the slurry was incubated at 4 °C overnight. Beads were collected and washed three times with 200 μl of wash buffer pH 8.0 (maximum injection time 1,000 ms) with a resolution 35,000 at 1,800 Th with a resolution of 70,000 at 66 Th. Enriched protein samples were concentrated onto three filters. Each filter was washed twice with 200 μl UA buffer (8 M urea; 100 mM Tris-HCl pH 7.5) and treated with 50 mM chloroacetamide in UA buffer for 20 min in the dark. After two additional washes with UA buffer, the samples were equilibrated three times with 200 μl of IP buffer (50 mM MOPS-NaOH pH 7.2; 10 mM Na2HPO4; 50 mM NaCl). Proteins were digested with Lys-C (1:10 enzyme-to-protein ratio) to recover high-molecular weight peptides retained on filters, another digestion step was performed with Glu-C (1:10 enzyme-to-protein ratio) in 50 μl IP buffer for 16 h at 37 °C. Peptides collected after Lys-C or Lys-C and Glu-C digests were kept separately.

**Filter-aided sample preparation and protein digestion.** HisSUMO109K-conjugated proteins (~300 μg as determined by ultraviolet–visible spectrophotometry) were isolated on 30-kDa agrose beads (Qiagen) pre-equilibrated in binding buffer, the slurry was incubated at 4 °C overnight. Beads were collected and washed on gravity columns with 50 ml of each buffer in the following order: binding buffer, wash buffer pH 8.0 (8 M urea; 100 mM NaHPO4, pH 8.0, 10 mM Tris-HCl, pH 8.0; 20 mM imidazole; 5 mM β-mercaptoethanol), wash buffer pH 6.3 (8 M urea; 100 mM NaHPO4, pH 6.3, 10 mM Tris-HCl, pH 6.3; 20 mM imidazole; 5 mM β-mercaptoethanol) and once more with wash buffer pH 8.0. Proteins were eluted in sequential steps with 5 ml of wash buffer pH 8.0 containing 200 mM imidazole and the three elution fractions of each biological sample were combined for further processing.

**MS analysis.** All peptide samples were desalted and concentrated with homemade reverse-phase C18 (Empore) Stop and Go extraction Tips®38, resuspended in 0.1% TFA and analysed by LC/MS/MS. An EASY-nLC 1000 liquid chromatography system (Thermo Scientific) coupled to a Q Exactive mass spectrometer (Thermo Scientific) through an EASY-Spray ion source (Thermo Scientific) was used and purified peptides were separated on an EASY-Spray column (75 μm × 500 mm). Samples containing SUMO remnant-modified peptides were separated on a 0.1% formic acid, respectively, at a flow rate of 250 ml per min, and the majority of peptides eluted during a 72-, 220-, 130-min acetonitrile window from 5 to 40% or from 5 to 50%. MS and MS/MS spectra of the eluting peptides were acquired on the Q Exactive mass spectrometer operated in a data-dependent mode. Precursor ion full-scan MS spectra were acquired at a target value of 1,000,000 ions (maximum injection time 20 ms) over a scan range of 300 to 1,600 Th or 300 to 1,800 Th with a resolution of 70,000 at m/z 400. For samples following high LC–MS/MS spectra for complete proteome and LFQ analyses at a target value of 500,000 ions (maximum injection time 60 ms) with a resolution 17,500 at m/z 400. Normalized collision energy was set to 30%, singly or highly (> 8) charged ions or ions with unassigned charge states were rejected, mz values of acquired ions were added to the exclusion list for 40 s and peptide match option was set to preferred.
Data analysis. Raw mass spectrometric data files were processed with MaxQuant software (version 1.3.0.5)\(^{36}\) and searched against UniprotKB S. pombe reference protein sequences (downloaded on 14 December 2013). UniprotKB gene names were used in Tables except for a few names that were updated with Pombase gene names. Protease specificity was set for Lys-C or Lys-C and Glu-C permitting up to three of five missed cleavages, respectively. Carbamidomethylation of Cys residues was added as a fixed modification and oxidation of Met, acetylation of protein N- and diGly adduction to internal Lys in the peptide sequence were set as variable modifications. In addition, phosphorylation of Ser, Thr and Tyr or acetylation of Lys were added as variable modifications and the minimum peptide length was set to seven amino acids. For SILAC samples, modification was set to 2, and Lys2 was selected as a heavy labelled counterpart. LFQ files were analysed using MaxQuant software version 1.5.2.8. Multiplicity was set to 1, trypsin was selected as an enzyme and up to two missed cleavages were permitted. LFQ analysis and the calculation of protein LFQ ratios were performed in the MaxQuant software. The raw LFQ ratios (Supplementary Dataset 2) were used as input data for a false discovery rate (FDR) and a localization probability calculation. Potential contaminants and proteins identified based on reverse decoy database search. LFQ values of His6SUMO/SUMO and His6SUMOL109K/SUMO were expressed as a log2.

Complete proteome results (Supplementary Data 4) report only proteins that were quantified in at least one biological replicate. Potential contaminants and proteins identified based on reverse decoy database search were excluded. For each quantified protein, the normalized protein ratio of individual biological replicates was averaged and expressed as a log2. The number of ratio occurrences and the variance of these ratios between the individual biological replicates of each experimental condition are reported in Supplementary Data 4. Mean normalized ratio of JK414/JK408 was derived from the mean ratios of JK418/JK408 and JK418/ JK414, and was expressed as a log2. The list of proteins quantified in the LFQ experiment (Supplementary Data 1) includes proteins that were quantified in at least two biological replicates of one experimental condition, omitting potential contaminants and proteins identified based on reverse decoy database search. LFQ values of His6SUMO/SUMO and His6SUMOL109K/SUMO were averaged and the ratio of the median values of His6SUMO/SUMO and His6SUMOL109K/SUMO were expressed as a log2. Each protein in the 2D scatter plot was coloured using a gradient from green to red according to the density estimation values calculated using Perseus software package (300 points per dimension; Pe(x,y) density function; version 1.5.2.6; downloaded from http://141.61.102.17/perseus_doku/doku.php?id=start). Perseus software was also used to calculate multiple-sample analysis of variance test — sum of the log P values (false discovery rate = 0.01; se = 0.14). The list of modification sites (Supplementary Data 2) contains peptides encompassing a GlyGly-modified lysine identified from any of the immunoaffinity purification experiments. Potential contaminants and proteins identified based on reverse decoy database search. For each quantified site, the localization probability of the additional modification is

Analysis of His\(_6\)SUMO\(^{109K}\)-modified sites. Sequence motifs were visualized using pLogo\(^{69}\). Sequence windows of 13 amino acids centred on all SUMO-modified lysines identified were used as input. Protein N- or C-terminal sequences containing 145,132 lysines was used as a background data set and residues were scaled relative to their Bonferroni-corrected statistical significance.

Mapping of sumoylation sites onto existing protein structures. The ensemble of S. pombe proteins for which sumoylation sites were identified in this study (Supplementary Data 2) were used in a Blast search at NCBI against the Protein Data Bank. PDB files were retrieved for proteins with high identity with their S. pombe homologues. Lysines that are sumylated in S. pombe were mapped onto structures with the PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC. The data are displayed in Fig. S3c, Supplementary Figs 4 and 5 and in Supplementary Data 7.

**References**