Global Analysis of Protein Sumoylation in Saccharomyces cerevisiae

James A. Wohlschlegel§§, Erica S. Johnson¶, Steven I. Reed§, and John R. Yates III¶¶
From the §Department of Cell Biology, The Scripps Research Institute, La Jolla, California 92037, the ¶Department of Biochemistry and Molecular Pharmacology, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, and the §§Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037

Although the modification of cellular factors by SUMO is an essential process in Saccharomyces cerevisiae, the identities of the substrates remain largely unknown. Using a mass spectrometry-based approach, we have identified 271 new SUMO targets. These substrates play roles in a diverse set of biological processes and greatly expand the scope of SUMO regulation in eukaryotic cells. Transcription appears to be the most prevalent process associated with sumoylation with novel SUMO substrates found in basal transcription machinery for RNA polymerases I, II, and III, pol II transcriptional elongation complexes, and a variety of chromatin remodeling, chromatin modifying, and chromatin silencing complexes. Additionally, our global analysis has revealed a number of interesting biological patterns in the list of SUMO targets including a clustering of sumylation targets within macromolecular complexes.

Posttranslational modification of proteins is a central mechanism by which biological processes are regulated. One such modification involves the covalent attachment of the small ubiquitin-related polypeptide SUMO (1) to different cellular substrates (1, 2). SUMO conjugation is carried out by a multistep enzymatic pathway consisting of the heterodimeric SUMO-activating enzyme Aos1p/Uba2p, the SUMO-conjugating enzyme Ubc9p, and several different SUMO ligases including Siz1p and Siz2p in yeast (3–7). The end result of this enzymatic cascade is a covalent isopeptide bond linking the C-terminal glycine of SUMO to the ε-amino group of specific lysines in the target protein. Sumoylation of a target protein in this manner can regulate protein function by a number of different mechanisms including altering its subcellular localization, modulating its interaction with other proteins, or by antagonizing its attachment to ubiquitin or other lysine-targeting modifications. Desumoylation of substrates by SUMO isopeptidases, such as Ulp1p and Ulp2p in yeast, is also a central feature in regulation of protein function by SUMO (8–10). Genetic data in both budding and fission yeast, as well as the variety of known SUMO targets in mammalian cells, point to a broad role for sumoylation and desumoylation in the regulation of many biological processes including transcription, cell cycle progression, DNA damage response, and signal transduction (11, 12).

It is known that a large number of proteins are sumoylated in Saccharomyces cerevisiae, but very few of these substrates have been identified to date. This is thought to be due to the low abundance of the sumoylated targets and the fact that only a small fraction of a substrate is sumoylated under a given set of conditions. A critical step in the effort to gain a more complete picture of the role of SUMO in eukaryotic biology will be the identification and characterization of a wider range of SUMO targets. To this end, we have performed a global proteomics analysis of sumoylation in budding yeast.

EXPERIMENTAL PROCEDURES

Yeast Strains—EJ337 contains at the chromosomal SMT3 locus a version of SMT3 encoding eight histidines between the third and fourth residues, making the N-terminal sequence MSDHHHHHHHSEV... This allele was constructed with a HIS3 marker at the 3' end of SMT3 using overlap PCR as described previously (13). The tagged locus was sequenced and contains no additional mutations. Strains expressing Cdc10Δ, Rad6Δ, RpΔ4, Prp45Δ, Rec2p, and Vps72p C-terminally tagged with the HA and His₆-containing sequence GYPYDVPDYAFL- HHHHHHHHH were constructed by transforming strain JD52 with PCR fragments that introduced this sequence followed by a HIS3 marker at the chromosomal loci.

Purification of SUMO Conjugates for Mass Spectrometry—The His₆-Smt3p and control purifications were performed essentially as described (13). EJ337 and the JD53 control strains were each grown in 6 liters of YPD (1% yeast extract, 2% peptone, 2% dextrose) to an A₅₆₄₀ of 1.8. Cells were harvested by centrifugation and lysed in 1.85N NaOH and 7.5% β-mercaptoethanol, and the proteins in the resulting lysate were precipitated by the addition of trichloroacetic acid to a final concentration of 20%. The precipitate was collected by centrifugation, washed in acetone, and resuspended in Buffer A (6 M guanidine HCl, 100 mM Tris-HCl pH 8.0, 20 mM imidazole). Lysates were clarified by centrifugation with pH adjusted to 8.0 using NaOH and incubated with Ni-nitroacetic acid NTA-agarose (Qiagen) for 2 h at room temperature. Ni-NTA-agarose was added to the His₆-Smt3p or control lysates and rotated overnight at 4°C. After washing with low salt buffer (30 mM Tris-HCl pH 8.0, 200 mM NaCl, 10% glycerol, 10 mM imidazole), proteins were eluted using Buffer B (8 M urea, 100 mM Tris-HCl, pH 8.0, 200 mM imidazole). Eluates were concentrated using an ultracentrifugal concentrator to a final concentration of 1 mg/ml, and stored at –80°C until analysis.

Identification of Proteins by Mass Spectrometry—To determine the identities of the substrates, eluates from the purification were precipitated with trichloroacetic acid and digested using Lys-C and trypsin proteases as described (13). The resulting tryptic peptides were then analyzed using multidimensional chromatography and tandem mass spectrometry. Mass spectra obtained for the peptides were analyzed using the Mascot search engine (Matrix Science) with a database of S. cerevisiae entries (50,138) to identify proteins that were sumoylated.

This work was supported by an American Cancer Society Postdoctoral Fellowship (to J. A. W.) and National Institutes of Health Grants RR1823-08 (to J. R. Y.), ES012021-02 (to J. R. Y.), CA5487 (to S. I. R.), and GM62685-04 (to E. S. J.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The abbreviations used are: SUMO, small ubiquitin-like modifier; HA, hemagglutinin; Ni-NTA, nickel-nitrilotriacetic acid; MS, mass spectrometry; TAP, tandem affinity purification.

Posttranslational modification of proteins is a central mechanism by which biological processes are regulated. One such modification involves the covalent attachment of the small ubiquitin-related polypeptide SUMO to different cellular substrates (1, 2). SUMO conjugation is carried out by a multistep enzymatic pathway consisting of the heterodimeric SUMO-activating enzyme Aos1p/Uba2p, the SUMO-conjugating enzyme Ubc9p, and several different SUMO ligases including Siz1p and Siz2p in yeast (3–7). The end result of this enzymatic cascade is a covalent isopeptide bond linking the C-terminal glycine of SUMO to the ε-amino group of specific lysines in the target protein. Sumoylation of a target protein in this manner can regulate protein function by a number of different mechanisms including altering its subcellular localization, modulating its interaction with other proteins, or by antagonizing its attachment to ubiquitin or other lysine-targeting modifications. Desumoylation of substrates by SUMO isopeptidases, such as Ulp1p and Ulp2p in yeast, is also a central feature in regulation of protein function by SUMO (8–10). Genetic data in both budding and fission yeast, as well as the variety of known SUMO targets in mammalian cells, point to a broad role for sumoylation and desumoylation in the regulation of many biological processes including transcription, cell cycle progression, DNA damage response, and signal transduction (11, 12).

It is known that a large number of proteins are sumoylated in Saccharomyces cerevisiae, but very few of these substrates have been identified to date. This is thought to be due to the low abundance of the sumoylated targets and the fact that only a small fraction of a substrate is sumoylated under a given set of conditions. A critical step in the effort to gain a more complete picture of the role of SUMO in eukaryotic biology will be the identification and characterization of a wider range of SUMO targets. To this end, we have performed a global proteomics analysis of sumoylation in budding yeast.

EXPERIMENTAL PROCEDURES

Yeast Strains—EJ337 contains at the chromosomal SMT3 locus a version of SMT3 encoding eight histidines between the third and fourth residues, making the N-terminal sequence MSDHHHHHHHSEV... This allele was constructed with a HIS3 marker at the 3' end of SMT3 using overlap PCR as described previously (13). The tagged locus was sequenced and contains no additional mutations. Strains expressing Cdc10Δ, Rad6Δ, RpΔ4, Prp45Δ, Rec2p, and Vps72p C-terminally tagged with the HA and His₆-containing sequence GYPYDVPDYAFL-HHHHHHHHH were constructed by transforming strain JD52 with PCR fragments that introduced this sequence followed by a HIS3 marker at the chromosomal loci.

Purification of SUMO Conjugates for Mass Spectrometry—The His₆-Smt3p and control purifications were performed essentially as described (13). EJ337 and the JD53 control strains were each grown in 6 liters of YPD (1% yeast extract, 2% peptone, 2% dextrose) to an A₅₆₄₀ of 1.8. Cells were harvested by centrifugation and lysed in 1.85N NaOH and 7.5% β-mercaptoethanol, and the proteins in the resulting lysate were precipitated by the addition of trichloroacetic acid to a final concentration of 20%. The precipitate was collected by centrifugation, washed in acetone, and resuspended in Buffer A (6 M guanidine HCl, 100 mM Tris-HCl pH 8.0, 20 mM imidazole). Lysates were clarified by centrifugation with pH adjusted to 8.0 using NaOH and incubated with Ni-nitroacetic acid NTA-agarose (Qiagen) for 2 h at room temperature. Ni-NTA-agarose was then washed extensively in Buffer A followed by Buffer B (8 M urea, 100 mM Tris-HCl, pH 8.0, 200 mM imidazole). Proteins bound to the Ni-NTA were eluted by the addition of Buffer C (Buffer B + 500 mM imidazole), trichloroacetic acid-precipitated, and analyzed by mass spectrometry. Small aliquots of these purifications were also analyzed by immunoblotting with α-Smt3p antisera and silver staining (Fig. 1B).

Identification of Proteins by Mass Spectrometry—Trichloroacetic acid precipitates from the His₆-Smt3p and control purifications were resuspended in digestion buffer (8 M urea, 100 mM Tris-HCl, pH 8.5) and digested by the sequential addition of Lys-C and trypsin proteases as previously described (14). The digests were then fractionated using microscale two-dimensional liquid chromatography and analyzed by tandem mass spectrometry using a Finnigan LTQ ion trap mass spectrometer. Multidimensional chromatography was performed on-line ac-
according to MacCoss et al. (15). Tandem mass spectra were collected in data-dependent fashion by collecting one full MS scan ($m/z$ range = 600–1600) followed by MS/MS spectra of the eight most abundant peptides with $X_{c}$corr values greater than 2.0 for $+1$ spectra, 2.2 for $+2$ spectra, and 3.75 for $+3$ spectra. The filtering for the control purification used the same $X_{c}$corr thresholds but required only one unique fully tryptic peptide for identification.

Confirmation of SUMO Substrates by Immunoblotting—Strains expressing TAP-tagged forms of putative SUMO targets were purchased from Open Biosystems. These strains were transformed with a LEU2-marked centromeric plasmid expressing His$_{8}$-FLAG-tagged Smt3p from a galactose-inducible promoter (pRS315-P$_{gal}$-His$_{8}$-FLAG-Smt3p) (4). 50-ml cultures of TAP strains $\pm$ pRS315-P$_{gal}$-His$_{8}$-FLAG-Smt3p were grown on selective medium containing 2% galactose to an $A_{600}$ ~ 1.2. Cells were collected by centrifugation, and lysates were prepared as described above. Purifications were performed using Ni-NTA spin columns (Qiagen) according to the manufacturer’s instructions. Ni-NTA eluates were trichloroacetic acid-precipitated, resuspended in Laemmli loading buffer, and analyzed by SDS-PAGE. Immunoblotting for the presence of the TAP-tagged protein was performed using $\alpha$-calmodulin-binding peptide antiserum kindly provided by Jonathan Weissman. HA-His$_{8}$-tagged version of Cdc11p, Rad51p, Rpb4p, Prp45p, Rsc2p, and Vps72p were affinity-purified from the appropriate strains using essentially the same approach as described above. Purifications were analyzed by SDS-PAGE and immunoblotted with either $\alpha$-HA or $\alpha$-Smt3p.

RESULTS

Purification of Smt3p-conjugated Proteins from His$_{8}$-Smt3p-expressing Yeast—To affinity-purify Smt3p-conjugated proteins from S. cerevisiae, we first constructed a yeast strain in which His$_{8}$-tagged Smt3p was expressed from its original chromosomal location under the control of its endogenous promoter. Importantly, sumoylation patterns in this strain as assessed by immunoblotting of whole cell extracts with anti-Smt3p antisera showed a similarity to sumoylation patterns in the wild-type parental strain in both asynchronous cells and in cells treated with nocodazole (Fig. 1A). The small mobility shift in the sumoylation patterns between the two strains is due to the presence of the His$_{8}$ affinity tag on Smt3p. We have previously found that sumoylation patterns in yeast overexpressing Smt3p can differ significantly from sumoylation patterns in wild-type yeast, which can lead to the spurious identification of SUMO substrates (data not shown). Furthermore, we have found that the use of even a slightly larger epitope tag such as His$_{8}$-FLAG or HA appeared to reduce SUMO attachment to some proteins, even when the tagged SUMO was overexpressed (data not shown). SUMO-conjugated proteins were then affinity-purified from the His$_{8}$-Smt3p-expressing yeast strain under denaturing conditions using Ni-NTA-agarose. A mock purification from an untagged Smt3p strain was also performed as a control. Immunoblotting with anti-Smt3p antisera showed a high degree of enrichment of Smt3p conjugates after the affinity purification compared with the control purification (Fig. 1B). Both control and His$_{8}$-Smt3p affinity purifications were also analyzed by silver staining (Fig. 1B). Interestingly, very few differences in staining patterns could be found between the two samples. Thus, even though Smt3p was highly enriched after affinity purification, the abundance of the SUMO conjugates was still quite low compared with the non-sumoylated yeast proteins that bound nonspecifically to the affinity matrix.

Identification of Smt3p Substrates Using Multidimensional Protein Identification Technology—The extremely low abundance of the enriched SUMO conjugates compared with the large amount of contaminants makes the identification of sumoylated targets technically challenging. By relying on multidimensional protein identification technology (MudPIT) previously developed in our laboratory coupled with newer more sensitive instrumentation, we were able to analyze both the control and His$_{8}$-Smt3p affinity-purified proteins (18, 19). Each protein mixture was digested with Lys-C and trypsin proteases, fractionated using on-line multidimensional microscale chromatography, and analyzed by tandem mass spectrometry. Peptide sequences and their corresponding proteins were identified by correlating the fragmentation spectra from the predicted fragmentation spectra from a yeast protein data base using the SEQUEST algorithm (16). In total, 5047 pep-
tides from 704 proteins were identified from the His$_6$-Smt3p sample using minimum criteria of two fully tryptic peptides (peptides whose sequence at both termini is consistent with the known cleavage specificity of trypsin) per locus and Xcorr cutoffs of 2.0 for +1 spectra, 2.2 for +2 spectra, and 3.75 for +3 spectra. 2171 peptides from 574 proteins were identified in the control sample using identical Xcorr cutoffs but only one fully tryptic peptide per locus. Upon comparison of the two samples, 271 proteins were found to be unique to the His$_6$-Smt3p sample and represent putative sumoylation targets. A complete list of all of the proteins found in these samples can be found in supplementary material.

As a whole, the list of putative SUMO substrates appears to be of high quality. Most previously identified SUMO targets were identified in this approach including Cdc3p, Cdc11p, Shs1p, Pol30p, and Top2p (13, 20–22). Pds5p and Ycs4p were the only published Smt3p conjugates not identified by our approach (23, 24). The identification of previously identified SUMO targets strongly supports the validity of our approach and suggests that most of the 271 proteins identified in our study are likely to be bona fide SUMO substrates.

Although the 271 proteins uniquely identified in the Smt3p purification provide a plethora of new substrates to be characterized, this list is not comprehensive and it is likely that there are some proteins erroneously classified as negatives. There are three potential sources of false negatives in this study. The first potential is that some sumoylated proteins may be of extremely low abundance and below the level of detection in our analysis. This is the probable reason that we were unable to identify Pds5p and Ycs4p in our investigation. A second potential cause of these misclassified negatives is that, if a protein had bound to Ni-NTA-agarose in the strain where Smt3p was not tagged, it was excluded from our list of putative SUMO targets. However, it is likely that some of these proteins, which include almost 10% of yeast proteins, are still authentic sumoylated substrates. We have found that, whereas a protein may bind to Ni-NTA-agarose in both strains at low levels, its relative abundance as estimated by spectral count may be significantly higher in Smt3p-tagged strains versus the control strain. For example, although Cdc48p was present in both Smt3p and control purifications, 15 unique peptides and 35 different spectra were identified for Cdc48p in the Smt3p purification compared with only 4 unique peptides and 15 spectra in the control purification. As recent work has shown that differences in spectra count can reflect differences in protein abundance, it is likely that Cdc48p was enriched to a greater extent in the SUMO purification and is potentially a SUMO substrate (25, 26). Indeed, the immunoblotting experiments presented later in Fig. 2A identify Cdc48p as a legitimate sumoylation substrate. Nonetheless, the mere presence of these potential SUMO conjugates in the control purification prevents global conclusions from being made regarding their sumoylation status and makes it necessary to analyze them on an individual basis. A third possible cause for false negatives is the filtering criteria used in this analysis. The criteria for filtering the Smt3p purification data was very stringent, requiring a correct protein identification to have at least two fully tryptic peptide identifications above the Xcorr cutoffs described earlier. In this case, many proteins identified by a single unique but high quality peptide identification were excluded from our analysis, although many of these are still potentially sumoylation targets but probably present at very low levels. A list of these one-peptide identifications can be found in supplementary material. To ensure reasonable quality of these one peptide identifications, we included only those peptides that were independently identified by at least two different spectra and by two separate search algorithms (SEQUEST and Pep_Probe) (16, 27).

As in any large scale analysis, it is also important to note that some of the putative SUMO substrates identified in our analysis will be incorrect (false positives). We have taken several steps in an attempt to minimize the false positive rate in our study. We have used very stringent filtering criteria as described earlier. The false positive rate for peptide identification using these criteria was estimated using a decoy data base approach to be <1% (28). A false positive could also arise when a protein that binds nonspecifically to the affinity matrix is not identified in the control purification. To partially address this concern, we lowered the criteria for protein identifications from the control purification as compared with the Smt3p purification (one peptide minimum versus two peptide minimum) to generate a more comprehensive but less reliable list of background proteins.

**Confirmation of Smt3p Targets by Immunoblotting**—To further establish the legitimacy of our proteomics results, we confirmed that a subset of the putative sumoylation substrates could be verified by immunoblotting. Strains in which putative targets were chromosomally TAP-tagged were transformed with a plasmid expressing His$_6$- and FLAG-tagged Smt3p under the control of a galactose-inducible promoter. TAP-Cdc11, a septin known to be sumoylated, was used as a positive control (13). These strains were grown in medium containing galactose and lysed under denaturing conditions. Smt3p conjugates were then purified using Ni-NTA-agarose and analyzed by immunoblotting with antibodies against the calmodulin-binding peptide portion of the TAP tag. TAP-tagged strains not expressing tagged Smt3p were used as negative controls. Fig. 2A shows the presence of TAP-Cdc11p, TAP-Pre2p, TAP-Nop2p, TAP-Rsc5p, TAP-Cdc14p, TAP-Cdc48p, TAP-Swe1p, and TAP-Pop3p in the Ni-NTA purification from His$_6$-FLAG-Smt3p-expressing strains but not in the corresponding control strains, indicating that these candidate SUMO targets are indeed authentic in vivo SUMO substrates. To ensure that all of the TAP-tagged proteins are not sumoylated in this assay, we also tested several TAP-tagged proteins not found to be sumoylated in our proteomic analysis (TAP-Cor1p, TAP-Nde1p, TAP-Arf1p, TAP-Pho88p) and were unable to detect their sumoylation (data not shown).

As a second approach to confirming our proteomics analysis, we also constructed yeast strains in which a His$_6$-HA tag was fused to the C terminus of Rpb4p, Prp45p, Rac2p, and Vps72p. Rad51p and Cdc10p, which were not identified in our proteomic analysis, were also tagged as negative controls. These fusion proteins were purified using Ni-NTA-agarose, analyzed by SDS-PAGE, and immunoblotted with either α-Smt3 or α-HA antibodies. Rpb4p, Prp45p, Rac2p, and Vps72p were all found to be sumoylated under these conditions, whereas Cdc10p sumoylation was extremely weak and no sumoylation of Rad51p was detected (Fig. 2B). Together, these results further authenticate our mass spectrometry data and provide added confidence in the quality of our list of putative SUMO substrates.

**Categorization of SUMO Targets by Gene Ontology (GO) Annotations and Subcellular Localization**—A detailed analysis of the sumoylation substrates identified in our study reveals a number of interesting trends and patterns that could not be ascertained by the investigation of individual proteins. Fig. 3A shows a histogram of the sumoylation targets categorized by GO-Slim biological process annotations, which is composed of very broad functional classifications. This histogram clearly demonstrates the diversity of processes potentially associated with sumoylation. Beyond the expected role of
SUMO in transcription, DNA metabolism, and transport, these results suggest other roles for sumoylation in a wide range of processes including RNA metabolism, protein catabolism, and protein biosynthesis. Moreover, a large number of sumoylation targets have not yet been ascribed a role in any known biological process, making them interesting candidates for further investigation.

Categorizing proteins by GO-Slim annotations separates proteins into very broad categories. To search for the clustering of SUMO substrates into more specific processes, we used the Gene Ontology Term finder tool on the Saccharomyces Genome Data bank to look for the categorical overrepresentation of GO annotations for a specific process in the SUMO candidate list compared with the entire yeast genome (www.yeastgenome.org) (29). A binomial distribution is used to calculate the probability that the fraction of sumoylation candidates mapped to a given biological function (GO process) could occur by chance given the fraction of all of the yeast proteins that map to that function. A representative subset of the biological processes that are associated with the sumoylation target list with a p value of <0.05 are shown in Table I. A complete list of the GO processes associated with sumoylation as well as the sumoylation substrates that map to those annotations can be found in supplementary material.

Table I displays several interesting features with respect to the regulation of biological processes by sumoylation. Although Fig. 3A demonstrates the connection between sumoylation and transcription, Table I clearly shows the prevalence of sumoylation in many different facets of transcription including transcription by all three RNA polymerase complexes, transcriptional elongation by RNA polymerase II, and modulation of chromatin structure through both chromatin modification and remodeling. This is significant in that that prior work linking transcription to sumoylation has focused primarily on the sumoylation of transcription factors and some chromatin-modifying enzymes such as histone deacetylases (30, 31). Our data provide the first report of SUMO directly modifying basal transcription machinery, transcriptional elongation complexes, and histone acetylases (SAGA, Spt-Ada-Gcn5-acetyltransferase complex) as well as multiple chromatin-remodeling complexes including Swr1p, Ino80p, and SWI/SNF. Clearly, the regulation of transcription by sumoylation is far more widespread than previously realized.

In addition to the diversity of biological processes associated...
Fig. 3. Categorization of SUMO substrates by GO-Slim annotation and subcellular localization. A, histogram showing the identification of SUMO substrates in a diverse range of biological processes. B, histogram showing the identification of SUMO substrates in a wide range of subcellular locations. ER, endoplasmic reticulum.
with SUMO, we find the presence of SUMO conjugates in a diverse range of subcellular locations. Recent work by Huh et al. (32) used genome-wide green fluorescent protein tagging of yeast open reading frames followed by fluorescence microscopy to determine the subcellular localization of >4100 yeast proteins. Taking advantage of this localization data, Fig. 3 shows a histogram depicting the number of SUMO substrates found in different subcellular compartments. Whereas the nucleus is by far the most abundant location for SUMO substrates, a large number of candidate targets can also be found in the cytoplasm and other subcellular structures. This is in good agreement with previous work suggesting that sumoylation is a predominantly nuclear process with a few notable cytoplasmic exceptions (12). An important caveat is that sumoylation is known to regulate the subcellular localization of many of its substrates, so that the subcellular localization of the sumoylated target may differ significantly in some cases from the localization of its unsumoylated counterpart.

**Clustering of Sumoylation in Macromolecular Complexes**—An intriguing feature discovered in our sumoylation analysis is the clustering of sumoylation targets in macromolecular complexes. Although this clustering has been previously noted for the septins in budding yeast in which Cdc3p, Cdc11p, and Shs1p are all present in the same complex and independently sumoylated, it was not clear whether this was a recurring theme (13). In our analysis, we find that septins are not alone in this regard and that there are multiple examples in which well defined protein complexes contain multiple sumoylated components (Table II). For example, eight proteins in the Ino80p chromatin-remodeling complex, seven different components of the Swr1p chromatin-remodeling complex, and six different factors of the SAGA complex are found to be sumoylated in our analysis. Importantly, the Smt3p purification was done without technical difficulties because sumoylation is a highly dynamic process and only a very small fraction of SUMO substrates is modified at a given time. By coupling traditional affinity purification methods with high throughput ultrasensitive protein identification technologies, we have been able to overcome these difficulties and report in this paper the identification of 271 proteins that are modified by SUMO in *S. cerevisiae*. The vast majority of these proteins has not previously been associated with sumoylation and implicates SUMO for the first time in the regulation of these biochemical pathways. The identification of these substrates in such a diverse array of complexes and processes raises new questions about the biological role of SUMO and sets the stage for the in-depth analysis of the regulation of these new substrates by this important but poorly understood modification.

**DISCUSSION**

The modification of cellular factors by SUMO is known to play a key role in many cellular processes. However, it has been technically challenging to identify many of these targets because sumoylation is a highly dynamic process and only a very small fraction of SUMO substrates is modified at a given time. By coupling traditional affinity purification methods with high throughput ultrasensitive protein identification technologies, we have been able to overcome these difficulties and report in this paper the identification of 271 proteins that are modified by SUMO in *S. cerevisiae*. The vast majority of these proteins has not previously been associated with sumoylation and implicates SUMO for the first time in the regulation of these biochemical pathways. The identification of these substrates in such a diverse array of complexes and processes raises new questions about the biological role of SUMO and sets the stage for the in-depth analysis of the regulation of these new substrates by this important but poorly understood modification.
An important aspect of our findings is the correlation between the phenotypes of yeast mutants defective in SUMO conjugation or deconjugation and the SUMO substrates identified in our study. For example, mutations in Ubc9 or Ulp1 show defects in cell cycle progression, arresting in G2/M phase and, in our proteomic analysis, we find the sumoylation of eleven proteins involved in different aspects of mitotic progression (4, 8, 33). These proteins can now be tested for possible roles in SUMO-dependent cell cycle progression. A ulp2Δ mutant shows defects in chromosome cohesion and condensation, highlighting the potential significance of the sumoylation of the cohesin subunits Mcd1p and Smc1p and of the condensin components Brn1p and Smc4p identified in our analysis (20, 34). We have also identified a number of proteins such as Rad16p, Sgs1p, Smc5p, Pol30p, Top1p, Top2p, and multiple components of the Ino80p chromatin-remodeling complex, that may play a role in the sensitivity to DNA-damaging agents exhibited by many yeast strains deficient in sumoylation or desumoylation (2, 35). In total, our results provide a large number of candidate substrates that may eventually explain the pleiotropic phenotypes seen in yeast mutants defective in various aspects of SUMO conjugation and deconjugation.

In our proteomic analysis, we also find noteworthy examples of the evolutionary conservation of sumoylation between budding yeast and higher eukaryotes. This conservation occurs at the level of both biological processes and individual SUMO substrates. For example, although the role of sumoylation in transcription has been well documented in higher eukaryotes, no such evidence suggested that it played a similar role in S. cerevisiae. However, our findings suggest a strong link between sumoylation and transcriptional regulation in budding yeast with the largest number of new substrates identified playing key roles in transcriptional control. This evolutionary conservation can also be seen at the level of individual substrates. Previous results have identified topoisomerase II as the only SUMO substrate shown to be modified in both yeast and mammalian cells. We find that the sumoylation of histone H4 and Werner helicase in mammalian cells as described in the literature is conserved in budding yeast where we find the sumoylation of their homologues Hhf1p/Hhf2p and Sgs1p, respectively (36, 37).

As demonstrated in Table II, our work provides multiple examples of an intriguing clustering of sumoylation in well defined macromolecular complexes. The high degree of incidence of this phenomenon in our data strongly suggests that substrate clustering constitutes an important facet of the regulation of these complexes by sumoylation. These results may also hint at a role for sumoylation in regulating processes by targeting macromolecular complexes rather than individual proteins as typical of other posttranslational modifications. It may be that the sumoylation of entire complexes is used as a co-regulatory mechanism to coordinate the activity of different components in a complex. Such coordinated sumoylation could function to relocalize an entire complex to a specific subcellular structure or to stabilize the whole complex by antagonizing the ubiquitination of its components. Clearly, additional work aimed at elucidating the functional strategy underlying this mechanism will provide important insight into how sumoylation regulates many cellular processes.

In addition to our work, two recent studies have also attempted to identify new SUMO substrates in S. cerevisiae (38, 39). Although both were successful in identifying new targets using their different approaches, the overall number of SUMO targets identified in these studies were considerably smaller than the number of substrates identified in our study. The larger scale of our sumoylation analysis provides the important benefit of being able to identify biological patterns and trends that may go unrealized when fewer proteins are analyzed. Together, these three studies complement each other and offer a plethora of new substrates to be characterized and future studies directed toward understanding the functional consequences of the sumoylation of these factors will undoubtedly provide important insight into the biological role of sumoylation.

Acknowledgment—We thank Alison Reindle for technical assistance.

REFERENCES

29. Deleted in proof
Global Analysis of Protein Sumoylation in *Saccharomyces cerevisiae*

James A. Wohlschlegel, Erica S. Johnson, Steven I. Reed and John R. Yates III


doi: 10.1074/jbc.M409203200 originally published online August 23, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M409203200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2004/09/01/M409203200.DC1

This article cites 38 references, 14 of which can be accessed free at http://www.jbc.org/content/279/44/45662.full.html#ref-list-1