Reversible covalent modification of proteins with a small ubiquitin-related modifier (SUMO) is emerging as an important system contributing to dynamic regulation of protein function. To enhance our understanding of the cell regulatory systems impacted by sumoylation, we used affinity chromatography-coupled high pressure liquid chromatography/tandem mass spectrometry for unbiased identification of candidate cellular SUMO substrate proteins. Here we describe the identification of 21 candidate sumoylated proteins from whole-cell lysates of HEK-293 cells. The nature of the proteins identified is consistent with a role for sumoylation in diverse cell regulatory systems but highlights regulation of chromatin organization and gene expression as major systems targeted by the sumoylation machinery.

In addition to the well recognized role of protein phosphorylation, dynamic post-translational regulation of protein function can be achieved through a combination of a variety of other reversible covalent modifications including methylation, acetylation, ubiquitination, and sumoylation. Ubiquitination and sumoylation involve the ligation of ubiquitin family small polypeptides to one or more lysines of the target protein. Ubiquitination has long been recognized as an address code that sends proteins to the proteasome for degradation. However, apart from its role in protein degradation, ubiquitination also participates in the regulation of gene expression, signal transduction, and intracellular transport. The consequences of sumoylation on target proteins are relatively obscure; although based on the growing list of identified sumoylated proteins, sumoylation has been implicated in a diverse array of cell regulatory functions including the regulation of chromatin structure, subcellular compartmentalization, transcription factor activity, DNA binding, and protein complex assembly (1–3).

There are three known small ubiquitin-like modifiers (SUMO) family members in mammalian cells: SUMO1, -2, and -3. The biochemical pathway mediating protein sumoylation has been characterized and is analogous to that defined for ubiquitination. Namely, the combined action of three proteins (an E1 activating enzyme, an E2 conjugating enzyme, and an E3 ligase) mediates isopeptide bond formation between SUMO and the target lysine on the substrate protein. Importantly, sumoylation is reversible, and a number of SUMO proteases that can cleave SUMO from specific target proteins have been identified (1–3).

The bulk of present knowledge of the role of sumoylation in the regulation of cell behavior has been inferred from the nature of identified SUMO substrate proteins (3). The list of known SUMO substrates has been compiled by a combination of serendipitous observations and direct candidate analysis. It is not at all clear how well this list reflects the contribution of dynamic sumoylation to cell regulatory events. Therefore, we wished to assess the feasibility of a proteomics-based approach for an unbiased sampling of cellular SUMO target proteins. Here, we report the use of affinity chromatography-coupled HPLC/MS/MS analysis for identification of sumoylated proteins. Our observations suggest a major role for SUMO in the modulation of proteins controlling chromatin structure and gene expression.

MATERIALS AND METHODS

Purification of Cellular SUMO1-conjugated Proteins—HA-tagged SUMO1 and SUMO1-QT were transiently expressed in human embryonic kidney-293 (HEK-293) cells. HEK-293 cells were cultured in 10% fetal bovine serum/Dulbecco’s modified Eagle’s medium. Five μg of pCDNA3-HA-SUMO1 or pCDNA3-HA-SUMO1-QT were transfected per 10-cm plate of cells using LipofectAMINE PLUS reagent (Invitrogen). 24 h post-transfection, HEK-293 cells were serum-deprived for 24 h. Fifteen 10-cm plates of the transfected HEK-293 cells were used for the subsequent immunopurification of expressed HA-tagged proteins. Cells were lysed in boiling phosphate-buffered saline, 4% SDS. Lysates were sonicated and cleared by centrifugation at 17,000 × g for 30 min. SDS was diluted to a final concentration of 0.5% SDS in 20 mM Tris, pH 7.4, 137 mM NaCl, 10% glycerol, 1% Triton X-100, 0.5% sodium deoxycholate, 2 mM EDTA, with a protease inhibitor mixture. Lysates again were cleared by centrifugation at 17,000 × g for 30 min prior to immunoprecipitation of HA-tagged proteins with Sepharose beads conjugated to monoclonal anti-HA antibody (Santa Cruz Biotechnology). Rabbit polyclonal anti-HA antibody (BAbCO) was used for immunoblot analysis of sumoylated proteins. Goat polyclonal anti-TFII-I (Santa Cruz Biotechnology) was used to detect TFII-I in anti-HA immunoprecipitates.

Protein Digestion—The affinity-purified proteins present in the antibody-conjugated beads were eluted using 1× SDS sample buffer; the resulting proteins were resolved in 4–12% SDS-polyacrylamide gel. Protein bands specific to natural SUMO1 were excised from the gel slab. Each protein band was cut into 1-mm³ pieces and destained with ubiquitin-protein isopeptide ligase; HPLC, high pressure liquid chromatography; MS, mass spectrometry; HA, hemagglutinin; HEK, human embryonic kidney; TF, transcription factor; ms/ms, tandem mass spectrometry; Pol, polymerase.
25 mM NH₄HCO₃ in 50% methanol, 50% water (v/v) three times for 10 min. Then they were washed with 10% acetic acid, 50% methanol, 40% water (v/v/v) three times for 1 h and swollen in water twice for 20 min. After that, the gel pieces were dehydrated with acetonitrile and dried in a SpeedVac (Thermo Savant, Holbrook, NY). The gel pieces were rehydrated with modified porcine trypsin (Promega, Madison, WI) at the concentration of 10 ng/ml in 50 mM NH₄HCO₃ and then were subjected to trypsin proteolytic digestion at 37 °C overnight. Tryptic peptides were sequentially extracted with 50% acetonitrile, 45% water, 5% trifluoroacetic acid (v/v/v) and 75% acetonitrile, 24.9% water, 0.1% trifluoroacetic acid (v/v/v) solutions. The peptide extracts were combined and dried in SpeedVac. The peptide samples were cleaned with a ZipTip C18 (Millipore, Bedford, MA) prior to HPLC/MS/MS analysis.

Nano-HPLC/MS Mass Spectrometry for Protein Identification—HPLC/MS/MS analysis was performed in a LCQ Deca XP ion trap mass spectrometer (ThermoFinnigan, San Jose, CA) equipped with a nano electrospray ionization source. The electrospray source was coupled online with an Agilent 1100 series nano flow HPLC system (Agilent, Palo Alto, CA). Two μl of the peptide solution in buffer A (2% acetonitrile, 97.9% water, 0.1% acetic acid (v/v/v)) was manually loaded into a capillary HPLC column (50-mm length × 75 μm (inner diameter), 5-μm
tolerance in MS/MS data. Protein hits with more than a three-peptide cleavage, a maximum 4-Da error tolerance in MS, and an 0.5-Da error used in all MASCOT searches: a maximum of one missed trypsin algorithm for protein identification. The following search parameters were

**TABLE I**

| Identified proteins           | Molecular mass (kDa) | No. of peptides | NCBI No. |
|------------------------------|----------------------|-----------------|----------|
| Sumoylation machinery        |                      |                 |          |
| RanBP2*                      | 357.9                | 6               | NP_006258|
| SUMO1 E1 C subunit           | 71.2                 | 5               | AAD12784 |
| SUMO1 E1 N subunit           | 38.4                 | 8               | NP_005491|
| Sumo1 E2I                    | 20.4                 | 5               | AAH00744 |
| Sumo1 E2II                   | 14.7                 | 10              | XP_212611|
| Sumo1                        | 11.5                 | 16              | NP_003343|
| Chromatin-associated factors |                      |                 |          |
| PML*                         | 90.6                 | 2               | NP_150242|
| TIF1-β                       | 88.5                 | 2               | NP_005753|
| TFIH-I                       | 110.4                | 4               | NP_127494|
| NXP-2                        | 108.2                | 1               | Q14149   |
| Zinc finger protein 262      | 136.2                | 2               | NP_005086|
| DNA topoisomerase 3           | 66.6                 | 2               | I1PQ_A   |
| BTB-like domain proteins     |                      |                 |          |
| Similar to MGC25497          | 83.4                 | 1               | NP_945342|
| MGC2628                      | 26.5                 | 3               | NP_076981|
| RNA polymerases              |                      |                 |          |
| RNA Pol III subunit 155 kDa  | 155.6                | 3               | NP_008996|
| RNA Pol III subunit RPC2     | 127.6                | 3               | NP_060552|
| RNA Pol III subunit RPC5     | 77.4                 | 2               | BAA95976 |
| RNA Pol III subunit 62 kDa   | 60.6                 | 2               | NP_006459|
| RNA Pol III RPC8             | 19.6                 | 1               | AAH17248 |
| RNA Pol I subunit            | 38.6                 | 1               | NP_004866|
| RNA Pol II subunit RB5       | 24.5                 | 1               | AAC0239 |
| Other                        |                      |                 |          |
| RanGAP1*                     | 67.3                 | 8               | BAB47464 |
| HSP70.1                      | 70                   | 4               | P08107   |
| Erp72                        | 72.9                 | 4               | NP_004902|
| ErbB3                        | 147.8                | 1               | NP_001973|
| Polymyositis/scleroderma     | 100.7                | 1               | XP_001527|
| autoantigen 2                |                      |                 |          |
| NOP5                         | 52.4                 | 2               | AAP29084 |

* Previously unidentified SUMO target.

**TABLE II**

| Identified proteins           | AVLLFWM/KXX |
|------------------------------|-------------|
| TIF1-β                       | VKE, IKEE   |
| TFIH-I                       | VKTE, VKEE, VKEE, IKQE, VKEE, VKQE, VKQE, IKEE, IKTE |
| NXP-2                        | IKQA, IKQE, VKKE, VKEE, VKKE, VKEE, IKEE, IKTE |
| Zinc finger protein 262      | AKEE, ARSE |
| Similar to MGC25497          | VKSE, IKQE |
| MGC2628                      | None        |
| RNA Pol III subunit 155kD    | AKA, AKYE   |
| RNA Pol III subunit RPC2     | None        |
| RNA Pol III subunit RPC5     | AKE, IKEE   |
| RNA Pol III subunit 62kD     | None        |
| RNA Pol III RPC8             | None        |
| RNA Pol I subunit            | None        |
| RNA Pol II subunit RB5       | None        |
| HSP70.1                      | None        |
| Erp72                        | VKEE, FKGE, VKEE |
| ErbB3                        | IKRE        |
| Polymyositis/scleroderma     | AKE, IKSE, IKQE |
| autoantigen 2                |             |
| NOP5                         | AKE, AKTE, VKE, IKEE |

**RESULTS AND DISCUSSION**

To assist in the affinity purification of SUMO-modified cellular proteins, we expressed HA-tagged SUMO or a non-ligatable SUMO variant, HA-SUMO-QT (4), in HEK-293 cells. Sumoylation is a dynamic modification controlled by the balance between the activities of SUMO ligase and SUMO protease machinery (2). Therefore, we reasoned that although ectopic expression of SUMO may enrich the population of sumoylated proteins in the cell, it likely would have little effect on the specificity of the reaction. As shown in Fig. 1, HA-SUMO was incorporated into multiple high molecular weight proteins and, as expected, HA-SUMO-QT was not. Affinity purification of HA-tagged proteins from denatured whole-cell lysates recovered a complex mixture from HA-SUMO-expressing cells, whereas only the monomer was isolated from HA-SUMO-QT-expressing cells (Fig. 2).

For protein identification, we transfected ~3 × 10^7 cells with HA-SUMO or HA-SUMO-QT and affinity-purified HA-conjugated proteins from denatured whole-cell lysates. Our method of lysates and precipitation made it highly unlikely that SUMO-specific proteins would be recovered as a consequence of non-covalent association with SUMO or other SUMO-modified proteins. The immunoprecipitates were resolved by SDS-PAGE (Fig. 2), excised, in-gel digested, and analyzed by HPLC/MS/MS (Fig. 3). Protein identifications that were specific to the HA-SUMO immunoprecipitate are shown in Table I.

This small scale analysis identified four previously characterized SUMO target proteins and 18 novel candidate SUMO targets. In addition, SUMO E1, E2, and E3 enzymes were identified. The latter identification likely is a consequence of the purification of proteins conjugated to HA-SUMO through thioester linkages as no reducing agents were included in the lysates or immunoprecipitation buffer. All gel slices also produced peptides corresponding to SUMO1 and SUMO2. SUMO2 is from the native pool of SUMO-like proteins and suggests that individual proteins can be modified by a mixture of SUMO family members. For novel candidate sumoylated proteins, we searched the amino acid sequences for matches to a SUMO target consensus sequence (hydrophobic KXE, where X is any amino acid) present in many but not all previously identified...
sumoylated proteins (Table II and Ref. 1).

Six proteins were identified that fall into a general class of chromatin-associated factors. In yeast, deletion of the SUMO ortholog results in a number of chromatin-related defects including aberrant mitosis, telomere elongation, and chromosome missegregation (1). In mammalian cells, a number of transcription factors and histone modification enzymes recently have been identified as SUMO targets (3). In addition to the classical SUMO target PML and the previously identified DNA topoisomerase I (5), we identified three predicted chromatin-associated proteins and the HDAC-3-associated transcription factor TFII-I (6). NXP-2 is a member of the MORC family of nuclear matrix-associated proteins (7). In mice, MORC is required for progression through meiotic prophase (8). Like NXP-2, zinc finger protein 262 contains a predicted chromatin-association domain (9). A highly related family member, znf261, is involved in chromosome translocation and may be the XQ13.1 defect responsible for X-linked retardation (10). TIF1-α is a homolog of the known SUMO target TFII-1. TIF1-α and -β have been implicated as negative regulators of nuclear hormone receptors (11). TIF1-β forms a complex with a histone H3-specific methyltransferase and a KRAB domain transcription factor, increasing the efficiency of transcriptional repression (12). Interestingly, TFII-I has been shown to interact with PIASx-β, a SUMO E3 ligase, and HDAC-3, one of four class I histone deacetylases involved in transcriptional repression (13). To validate modification of TFII-I, we immunoblotted the anti-HA immunoprecipitates with anti-TFII-1 antibody. Unmodified TFII-I migrated at 135 kDa. As shown in Fig. 4, TFII-I in the HA-SUMO immunoprecipitate migrated much slower with a predicted 40-kDa increase in molecular weight suggesting that TFII-I is a bona fide SUMO target protein.

In an intriguing parallel to the identification of TFII-I as a SUMO target, two proteins that share the presence of a predicted BTB/POZ domain were identified. This domain is present in a variety of proteins that interact with histone deacetylase complexes and mediate transcriptional repression (14). A founding member of the protein family defining the BTB/POZ domains, the Drosophila protein Tramtrack 69 is a bona fide SUMO substrate (15).

We were surprised to identify a large class of RNA polymerase II subunits as well as RNA Pol I and RNA Pol II subunits. It has been suggested that one function of SUMO may be localization of proteins at PML domains. Recently it has been reported that PML domains dynamically localize at sights of active transcription in a cell cycle-regulated manner (16). It will be interesting to see whether sumoylation regulates the activity and/or subnuclear compartmentalization of RNA polymerases.

Five additional novel candidate SUMO substrates were identified. Two chaperones, HSP70.1 and the endoplasmic reticulum resident protein disulfide isomerase Erp72 (17), were isolated. NOP5 is a small ribonucleotide protein involved in ribosome biosynthesis (18). Polymyositis/scleroderma autoantigen 2 is a component of the exosome, an mRNA quality control complex (19). Finally, the epidermal growth factor-like receptor ErbB3 was identified. Interestingly, this family of mitogenic receptors is known to be down-regulated by ubiquitin-mediated degradation (20). We are currently examining the impact of SUMO modification on that regulatory step.

In conclusion, our observations suggest that a proteomics-based approach is a viable mechanism to derive unbiased profiles of cellular sumoylated proteins. We have examined a single growth condition for this study. However, sumoylation is a dynamic process, and substrate profiles may change dramatically depending upon the cell regulatory context. The majority of candidate SUMO targets identified here appear to be involved with the regulation of chromatin and gene expression. Future studies coupling this approach with multiple cell regulatory phases likely will help reveal the major nodes of regulation impacted by dynamic SUMO modification.

Acknowledgment—We thank Kim Orth for generously providing SUMO expression constructs.

REFERENCES

1. Melchior, F. (2000) Annu. Rev. Cell Dev. Biol. 16, 591–626
2. Melchior, F., Schergau, M., and Pichler, A. (2003) Trends Biochem. Sci. 28, 612–618
3. Verger, A., Perdomo, J., and Crossley, M. (2003) EMBO Rep. 4, 137–142
4. Orih, K., Xu, Z., Madgett, M. J., Ilac, Z. Q., Palmer, L. H., Bielka, J. B., Mangel, W. F., Staskawicz, B., and Dixon, J. E. (2000) Science 290, 1594–1597
5. Mao, Y., Sun, M., Desai, S. D., and Liu, L. F. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 4046–4051
6. Wen, Y. D., Cress, W. D., Roy, A. L., and Seto, E. (2003) J. Biol. Chem. 278, 1841–1847
7. Kimura, Y., Sakai, F., Nakano, O., Kisaki, O., Sugimoto, H., Sawamura, T., Sadano, H., and Osumi, T. (2002) J. Biol. Chem. 277, 20611–20617
8. Inoue, N., Hess, K. D., Moreadith, R. W., Richardson, L. L., Handel, M. A., Watson, M. L., and Zinn, A. R. (1999) Hum. Mol. Genet. 8, 1201–1207
9. Smalley, D., Hamoudi, R., Lu, Y. J., Cooper, C., and Shipley, J. (1999) Genomics 60, 244–247
10. van der Maarel, S. M., Scholten, I. H., Huber, I., Philippe, C., Suijkerbuijk, R. F., Gilgenkrantz, S., Kere, J., Cremers, F. P., and Ippers, H. H. (1996) Hum. Mol. Genet. 5, 887–897
11. Le Douarin, B., Nielsen, A. L., Garnier, J. M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. (1996) EMBO J. 15, 6701–6715
12. Hoos, C., Poletti, P., Feizi, T., Srinivasan, M., Haugejorden, S. M., and Green, M. (1998) Proc. Natl. Acad. Sci. U. S. A. 99, 12807–12812
13. de Vries, J., de Vries, J., and van der Maarel, S. M. (1999) J. Biol. Chem. 273, 27695–27702
14. Lehembre, F., Badenhorst, P., Muller, S., Travers, A., Schweiguth, F., and Dejean, A. (2000) Mol. Cell. Biol. 20, 1072–1082
15. Kieslich, A., von Mikecz, A., and Hemmerich, P. (2002) J. Struct. Biol. 140, 167–179
16. Mazzarella, R. A., Srinivasan, M., Haugejorden, S. M., and Green, M. (1999) J. Biol. Chem. 265, 1094–1101
17. Lyman, S. K., Gerace, L., and Baserga, S. J. (1999) RNA (N. Y.) 5, 1597–1604
18. Almang, C., Pefalski, E., Podtelejnikov, A., Mann, M., Tollervey, D., and Mitchell, P. (1999) Genes Dev. 13, 2148–2158
19. Qiu, X. B., and Goldberg, A. L. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 14843–14848
