Covalent conjugation of the small ubiquitin-like modifier (SUMO) proteins to target proteins regulates many important eukaryotic cellular mechanisms. Although the molecular consequences of the conjugation of SUMO proteins are relatively well understood, little is known about the cellular signals that regulate the modification of their substrates. Here, we show that SUMO-2 and SUMO-3 are required for cells to survive heat shock. Through quantitative labeling techniques, stringent purification of SUMOylated proteins, advanced mass spectrometric technology, and novel techniques of data analysis, we quantified heat shock–induced changes in the SUMOylation state of 766 putative substrates. In response to heat shock, SUMO was polymerized into polySUMO chains and redistributed among a wide range of proteins involved in cell cycle regulation; apoptosis; the trafficking, folding, and degradation of proteins; transcription; translation; and DNA replication, recombination, and repair. This comprehensive proteomic analysis of the substrates of a ubiquitin-like modifier (Ubl) identifies a pervasive role for SUMO proteins in the biologic response to hyperthermic stress.

INTRODUCTION

All forms of cellular life need to be able to sense and respond to extreme environmental and pathological conditions. In the absence of an appropriate response, such stresses can lead to cell damage or death. The heat shock response is one of the most evolutionarily conserved of such defense mechanisms, and is characterized by massive induction of the expression of genes that encode heat shock proteins (HSPs), which act to protect the cell from the cytotoxic stress. Although the roles of HSPs are well documented, early signaling mechanisms in the heat shock response have yet to be fully investigated. Because of their dynamic nature, reversible posttranslational modifications that involve the addition and removal of molecular modifiers are capable of rapid transduction of cellular signals; specifically, the reversible attachment to target proteins of the protein ubiquitin is critical to important cellular signaling pathways (7).

Ubiquitin is the founding member of the ubiquitin-like protein modifier (Ubl) family of posttranslational modifying proteins that share sequence and structural similarity with ubiquitin, although they often confer distinct biological consequences (2). A subfamily within this group, the small ubiquitin-like modifiers (SUMOs), are covalently conjugated to various target proteins and have a broad range of functional consequences (2). Three SUMO proteins are found in mammalian cells, with SUMO-2 and SUMO-3 sharing 97% sequence identity with one another and only about 50% identity with SUMO-1 (3). Although paralog-specific features and functions have been described, SUMO-2 and SUMO-3 are thought to have substantially overlapping substrate specificities with SUMO-1 (4), and they display some functional redundancy (5). SUMO-2 and SUMO-3 share with ubiquitin the ability to form self-modified polymers in vivo and in vitro (3), whereas the role of SUMO-1 in such assemblies seems to be limited only to the termination of conjugates of poly-SUMO-2 or poly-SUMO-3 (6).

Another important distinction between the SUMO proteins is that the conjugation of SUMO-2 and SUMO-3, but not that of SUMO-1, is stimulated by cellular stresses such as exposure to heat shock (7). Mass spectrometry is well-suited to the analysis of posttranslational modifications in general (8) and SUMOylation in particular (3, 6, 9). Thus, to further understand the role of SUMO in the heat shock response, we undertook a system-wide, quantitative, proteomic analysis of targets of SUMO-2 during the heat shock response. We developed a highly stringent strategy for the purification of SUMOylated proteins, which, coupled to novel data processing techniques, allowed the identification of 766 putative substrates of SUMO-2. We measured the changes in the SUMOylation states of these targets during and after heat shock and demonstrated that SUMO-2 underwent a rapid and dramatic redistribution between substrates, most of which have functional relevance for the heat shock response. These findings not only identify 574 previously unknown substrates of SUMO proteins, but also implicate SUMO-2 as an important, early signaling molecule in the cellular defense against hyperthermic cytotoxicity.

RESULTS

SUMO-2 and SUMO-3 promote cell survival after heat shock

Although SUMO-2 and SUMO-3 respond to heat shock by forming high molecular weight conjugates (7), the biological relevance of this response has not been established. To investigate this, clonogenic survival assays were performed in U2OS cells after SUMO-2 and SUMO-3 were knocked down with short interfering RNAs (siRNAs). The substantial depletion of SUMO-2 and SUMO-3 (Fig. 1A) caused a sevenfold reduction in cell survival after exposure to heat shock (30 min at 45°C) compared to that in control cells (Fig. 1, B and C), indicating that SUMO-2 and SUMO-3 play roles in the cellular response to heat shock that cannot be completely compensated for by SUMO-1 (Fig. 1A).
Identification of targets of SUMO-2 and measurement of changes in their conjugation state after heat shock

To identify the cellular targets of stress-induced SUMOylation, a quantitative proteomic approach was taken. We generated a HeLa cell line that contained SUMO-2 fused to a tandem affinity protein (TAP) tag (10) (Fig. 2A), which was similar in its abundance and its response to heat shock to that of endogenous SUMO-2 (Fig. 2B). Comparison of the products purified under initially denaturing conditions from HeLa cells containing only the TAP tag with those from cells containing TAP–SUMO-2 confirmed the high stringency of this method (Fig. 2C) and, hence, its suitability for proteomic analysis. For quantitation of the proteome, we used the stable isotope labeling of amino acids in cell culture (SILAC) technique, in which isotopically unique forms of amino acids are used to measure the relative abundance of proteins in cells grown under different conditions (11).

For our first proteomic experiment, three separate pools of HeLa cells were grown in SILAC culture medium such that control cells that contained only the TAP tag were grown under isotopically normal (light, L) conditions, cells containing TAP–SUMO-2 proteins grown at 37°C were cultured with Lys⁴ and Arg⁶ (medium, M) isotopic forms, and cells containing TAP–SUMO-2 proteins that were treated for 30 min at 43°C were grown with Lys⁸ and Arg¹⁰ (heavy, H) amino acids (Fig. 3A). Cells were lysed under denaturing conditions and TAP-tagged proteins were purified from pooled lysates, yielding a single protein sample that was fractionated by SDS-polyacrylamide gel electrophoresis (PAGE) (Fig. 3A). The SDS-PAGE gel was excised into six slices, each of which was subjected to in-gel tryptic digestion, and the resulting peptides were analyzed by liquid chromatography–tandem mass spectrometry (LC-MS/MS). To provide quantitative information on the relative abundance of individual peptides from the light, medium, and heavy culture conditions, MaxQuant data analysis was used (12), which gave quantitative data for 1159 proteins (Supplementary File 1). This approach was validated as a method for monitoring changes in SUMOylation after heat shock by the finding that, consistent with previous studies (13), the heat-shock transcription factor 1 (HSF1) displayed markedly increased SUMOylation after heat shock (Fig. 3A).

Due to the size of the data set, a clear graphical representation method was developed to allow the evaluation of the quantitative data for all proteins in a single chart, a triple-SILAC map (tsMap), which displays quantitative data for all of the 1159 proteins identified in this analysis (Fig. 3B). Each protein is represented by a single point on an x/y scatter plot with coordinates derived from the M:L (TAP–SUMO-2:TAP) and H:M (TAP–SUMO-2:HS:TAP–SUMO-2) ratios. Proteins that originated from outside the experiment (“external rejects”) would only be found in the unlabeled form, which results in their having low M:L and H:M ratios, and so they would be located in the bottom left portion of the tsMAP (fig. S1A). Those proteins that are likely to be nonspecific contaminants from the purification process would be equally abundant in all three isotopic forms, giving them M:L and H:M ratios of about 1, and so they would be found close to the origin of the tsMAP (fig. S1A). Putative substrates of SUMO-2 are therefore those proteins that remain after the rejection of both former groups of proteins (fig. S1). This analysis enabled the unambiguous identification of 662 putative targets of SUMO-2 (Supplementary File 1), of which 90 were identified by a single unique peptide. By performing experiments under three SILAC conditions rather than two, we were able to filter the data twice, thereby increasing the sensitivity of the experiment. Moreover, hnRNPM (14), the glucocorticoid receptor (15), and the Werner syndrome helicase (16), “known” substrates of SUMO, would have been omitted had this approach not been taken.

Western blotting analysis showed that during heat shock, the number of proteins modified by SUMO-2 increased and the amount of free SUMO-2 or control siRNA grown under normal conditions (37°C) or heat stressed at 45°C for 30 min (see Materials and Methods). Surviving cells were determined by counting colony numbers as judged by Giemsa staining. The experiment was performed twice in triplicate; one replicate data set is shown. (C) Summary of triplicate data from (B).

![Image](http://www.sciencemag.org/content/full/26/72/ra24/F1.large.jpg)
Changes to the SUMOylation state of proteins as cells recover from heat shock

To better understand the temporal dynamics of the SUMOylation response to heat shock, we applied the same quantitative approach to monitor the system-wide changes in SUMOylation of proteins both in response to heat shock and after a 2-hour heat shock recovery period (fig. S2 and Fig. 4A).

Fig. 2. TAP purification of SUMO-2 from cultured human cells. (A) Schematic representation of the TAP–SUMO-2 construct with which HeLa cells were stably transfected. The TAP tag consists of a protein A domain separated from a calmodulin-binding protein (CBP) domain by the tobacco etch virus (TEV) protease site. This was N-terminally tagged to SUMO-2 (amino acid residues 1 to 92) (NCBI Entrez protein CAG46970), which can be conjugated directly to target proteins through a covalent bond. (B) An antibody against SUMO-2 was used for the analysis of Western blots of crude cell lysates from TAP–SUMO-2-containing HeLa cells under normal conditions (37°C) and after heat shock for 30 min (43°C). This Western blot is representative of five separate experiments. (C) Silver-stained SDS-PAGE gel showing TAP purification products from HeLa cells containing TAP alone or TAP–SUMO-2. This stained gel is representative of six separate experiments.

This second SILAC experiment quantified 1255 proteins (Fig. 4, B and C, and Supplementary File 1), which, after tsMap filtering (fig. S3A) yielded 672 proteins targeted by SUMO-2 (Supplementary File 1), of which 130 were identified by a single unique peptide. There was excellent agreement between this experiment and the first SILAC experiment with regard to analysis of the heat shock response (fig. S3, C and D), and taking SILAC experiments 1 and 2 together, a total of 766 putative substrates of SUMO-2 were identified, of which 568 were common to both experiments. Crucially, quantitative analysis of unpurified cell lysates (fig. S4, A to C) indicated that differences between the light, medium, and heavy SILAC conditions in the TAP-purification experiment reflected changes in the SUMOylation state of targets rather than simply changes in the overall quantities of these proteins. Analysis of the tsMAP of these crude data (fig. S4D) showed how changes in the global abundance of proteins were relatively small compared to changes in the number of proteins conjugated to SUMO-2 (compare with Fig. 4C).

Validation of the SUMO-2 substrate proteome

To date, targets of SUMO proteins have been identified through either proteomic approaches or more empirical techniques such as the purification of SUMOylated proteins from cells followed by Western blotting analysis for a specific protein. Currently, the empirical approaches have documented more than 260 protein targets of all three SUMO proteins, of which our study has identified a greater proportion (34%) than did all 11 previous SUMO substrate proteomics studies taken together (28%) (fig. S5). Moreover, those proteins identified by MaxQuant but excluded from our analysis as internal rejects contained only 6% of the “confirmed” targets (fig. S5). Sequence analysis revealed that our SUMO-2 substrate proteome contains a comparable number of ψKXE SUMO conjugation consensus motifs (17) per protein to that contained in the 265 independently verified targets (Table 1). The TAP–SUMO-2 internal rejects list contains almost threefold fewer SUMO consensus sites per protein than did the TAP–SUMO-2 proteome and is comparable to a human proteome “background” set (Table 1).

The TAP–SUMO-2 putative substrate proteome contains many well-known targets of SUMO proteins. Despite being almost exclusively modified by SUMO-1 (7), the nuclear pore complex protein RanGAP1 was identified with 74% sequence coverage in our SUMO-2 purification system. The conjugation state of RanGAP1 remained largely unchanged throughout heat shock and the recovery period, which is consistent with the idea that SUMOylation of RanGAP1 is relatively stable. Notably, we identified a number of SUMO E3 ligases that not only accelerate the SUMOylation of target proteins but can also be SUMOylated themselves. These included PIAS [protein inhibitor of activated signal transducer and activator of transcription (STAT)] 1, 2, 3, and 4, Ran-binding protein 2 (RanBP2) (18), and topoisomerase 1–binding, arginine- and lysine-rich protein (TOPORS) (19), all of which were increased in conjugation state after heat shock, with the abundance of SUMOylated PIAS4 increasing 20-fold after heat shock compared to that in control cells.

Other interesting groups of proteins that are known substrates of SUMO proteins include three components of promyelocytic leukemia (PML) bodies, PML, DAXX, and SP100, although DAXX was the only protein whose conjugation status was significantly altered (increased 24-fold compared to that in control cells). DNA topoisomerases 1, 2A, and 2B were also identified as targets of SUMO proteins, although their conjugation state was largely unchanged after heat shock. SUMO-1 and SUMO-3 were purified along with TAP–SUMO-2, which was likely a consequence of the purification of hetero-SUMO-polymers (3, 6), SUMO-2 conjugates that were also identified by other SUMOs at multiple lysines, or both. To confirm some of these identifications and to further validate this method for purifying SUMO-2–conjugated proteins, the SUMO-substrate branched peptides for four of these known targets (RanGAP1, SUMO-2, SUMO-3, and SAFB) and a
previously unidentified target (Ki67 antigen), were identified by manual screening and interpretation of spectra (6) (figs. S6 and S7).

Although the SUMO-2 substrate proteome contains proteins with a broad range of molecular weights, exposure to heat shock induced the accumulation of very high molecular weight SUMO conjugates (Figs. 2B and 4A). Analysis of target distribution throughout the gel in the second SILAC experiment showed that heat shock–induced changes were most prominent in the upper regions (Fig. 5A, left), with slices more than 100 kD being predominantly populated with proteins of lower predicted molecular weight (Fig. 5A, right). In addition, individual proteins whose conjugation state increased with heat shock were generally present in broader molecular weight ranges than would be expected if they had been conjugated to only one or two SUMO adducts (Fig. 5B). Consistent with the purification of all three SUMOs in the experiment, a potential explanation for this aberrant protein migration is the formation of polySUMO-2 or polySUMO-3 adducts on target proteins. This seems likely because quantitation of SUMO-2–SUMO-2 branched peptides indicated a strong induction of SUMO polymerization through Lys11 upon exposure to heat shock, which returned to a basal abundance after recovery (Fig. 5C). Whether in some cases these are signals for degradation through the ubiquitin-proteasome pathway (9) remains to be investigated. Altogether, these analyses and observations strongly suggest that the approach used here was successful in identifying substrates of SUMO proteins.

Functional characteristics of substrates of SUMO-2

Many proteins that are directly involved in the cellular response to heat shock—such as HSF-1 and HSF-2 (20); members of the HSP 40-, 60-,
and 70-kD families (20); the co-chaperonins Sti1, RUVBL1, and RUVBL2 (21); the heat shock signal transduction kinase p38-MAPK14 (22); and the translocated promoter region (TPR) protein, which facilitates specific transport of HSP70 messenger RNA (mRNA) (23)—are all targets of SUMO-2 (Supplementary File 1). The relationship between the biological consequences of heat shock and the SUMO-2 substrate proteome also continues at a general functional level; heat shock causes damage to both proteins (20) and DNA (24, 25), which, depending on the severity of the exposure, results in either programmed cell death or survival and desensitization to future stress (20). Accordingly, the heat shock response involves regulation of transcription, translation, apoptosis, cell cycle control, protein folding, and protein degradation (20, 26, 27). Proteins involved in all of these processes and those associated with DNA replication, recombination, and repair were significantly overrepresented in the SUMO-2 modified proteome (Fig. 6A and figs. S8A and S9). Notably, there was significant overlap with the phosphoproteome of the DNA damage responsive kinases ATM-ATR (28) (fig. S10, A to C), suggesting many common biological goals. The major difference between these is the involvement of SUMO-2 in ubiquitin-dependent protein turnover (figs. S10 and S11).

Whereas the global increase in SUMO conjugation in response to heat shock observed by Western blotting experiments was rapid, with almost all free SUMO proteins being conjugated into high molecular weight adducts after 5 min, deconjugation during the recovery period was much slower (fig. S2). System-wide analysis of the SUMO-2 substrate proteome by tsMap (fig. S3B) revealed a general trend whereby proteins SUMOylated upon heat shock became deconjugated during the recovery period [for

Fig. 4. The effect of recovery from heat shock on the conjugation states of substrates of SUMO-2. (A) An antibody against SUMO-2 was used in the analysis of Western blots of individual lysates of TAP–SUMO-2–containing HeLa cells that were either untreated (37°C), exposed to heat shock at 43°C for 30 min (HS), or allowed to recover for 2 hours at 37°C after heat-shock (HSR) cultured in the indicated SILAC labels. (B) Raw MS data showing spectra from representative peptides for scaffold attachment factor B1 (SAF-B1) and the glucocorticoid receptor. (C) tsMap of 1255 identified proteins. Ratio cutoffs were determined graphically by isolating proteins that are unchanged in both conditions.
example, the glucocorticoid receptor (GR) (Fig. 4B)), but those proteins that were deconjugated during the heat shock response mostly did not recover their conjugation state after 2 hours [for example, SAF-B1 (Fig. 4B)]. To extract biologically relevant information from these data, we developed an analysis that mapped protein members from significantly clustering gene ontology (GO) groups (29) onto tsMaps, which allowed changes in their SUMOylation status with respect to the entire data set to be easily visualized (Fig. 6, B to G). Inspection of these GoTsMaps showed that, like most large groups, transcription factors (Fig. 6B) did not exhibit significant regulation as a group (table S1); however, subcategories of proteins did show coordinated deSUMOylation (Fig. 6C). These included the dimerizing transcription factors Jun, FOS, and the V-MAF musculoaponeurotic fibrosarcoma (MAF) oncogene homolog proteins (Fig. 7B), which act upstream of the oxidative stress-response transcription factor NRF2 (fig. S12). Because Jun is activated in response to heat shock (30) and is repressed by SUMOylation (31), it seems likely that rapid deSUMOylation depresses Jun as an early response to heat shock.

Transcriptional control during hyperthermic stress also acts more globally to facilitate the production of HSPs while repressing the expression of genes that do not encode HSPs (20). Consistent with this, many proteins identified in this study as putative substrates of SUMO-2 are involved in chromatin remodeling. For example, the switch-sucrose nonfermentable (SWI/SNF)-related matrix–associated actin-dependent regulators of chromatin (SMARCs) (fig. S14A), which are involved in heat shock–induced chromatin remodeling (32), were largely conjugated to SUMO upon exposure to heat shock. Recently, clear roles for SUMOylation in the recruitment of repressive complexes have been established for Kruppel-associated box zinc finger protein (KRAB-ZFP) (33) and Sp3 (34). SUMOylation of the transcription factor Sp3 or the KRAB-ZFP corepressor KAP1 is thought to mediate successive recruitment of the nucleosome remodeling and deacetylating (NuRD) complex to deacetylate histones and the Suvar 3-9 and Enhancer of zeste domain bifurcated 1 (SETDB1) complex to methylate histones, which in turn bind to HP1 proteins to maintain the chromatin silencing. The finding from our analysis is that almost all of the components of the above-mentioned protein complexes are SUMOylated: the transcription factors Sp3 and KRAB-ZFP; the corepressor KAP1; the NuRD complex (HDAC1, Mi2, Mi2β, Mi2γ, MTA2, and RAP46); the SETDB1 complex (SETDB1, MBD1, and ATFIIP); and HP1α and HP1γ.

An alternative route to histone deacetylation is the recruitment of the SIN3 complex; again, we found that almost all of its components were SUMOylated after heat shock, including Sin3B, Sd3, Sin3β-associated protein 18 (SAP18), SAP130, and SAP180. Furthermore, consistent with previous work (35, 36), the H2A, H2B, and H4 histone subfamilies were identified as targets of SUMO in our analysis (Supplementary File 1) and, together with histone deacetylase (HDAC) and histone acetyltransferase (HAT) proteins, were coordinately deconjugated from SUMO proteins upon heat shock (fig. S8B). Although modification of histone proteins represents one path to epigenetic regulation, methylation of DNA is also critical (37). In this respect, the DNA methyltransferase DNMT1 was subject to SUMO modification, and its associated cofactor UHRF1 displayed a 45-fold increase in SUMO modification after heat shock. Together, these observations suggest a considerably more extensive role for SUMO modification in the regulation of transcriptional repression than was previously suspected.

RNA-binding proteins clustered in two distinct regions of the GoTsMap (Fig. 6D). Among these proteins, almost all of the small nuclear ribonucleoproteins (snRNPs) were deSUMOylated (Fig. 6E and fig. S14B), whereas the heterogeneous nuclear ribonucleoproteins (hnRNPs) showed both increased and decreased SUMOylation state with heat shock (fig. S14B). In addition to these proteins, many importins and nuclear pore complex proteins also seemed to be targets of SUMO-2 (fig. S14C), indicating that SUMO proteins seem to take a multifaceted approach to control mRNA translation, consistent with observed effects of heat shock on mRNA metabolism (38).

Proteins involved in DNA repair (Fig. 6F, Fig. 7C, and fig. S13B), DNA recombination (Fig. 6G), and the DNA-damage response (Fig. 7D and fig. S13A) were almost exclusively SUMOylated in response to heat shock. A striking functional group within these proteins was that involved in the initiation of DNA replication (Fig. 7A). This group contains three components of the origin recognition complex (Orcs 2L, 3L, and 6L) and four subunits of the mini chromosome maintenance (MCM) replicative helicase (MCMs 2, 3, 4, and 7). The targeting of these proteins for SUMOylation after exposure to heat shock suggests a hitherto unrecognized role for SUMO proteins in the regulation of DNA replication and is consistent with the observation that delaying the progression of S-phase rescues cells from heat-induced S-phase hypertoxicity (39).

### Table 1. Predicted frequency of SUMO consensus sites of proteins identified by different studies. SUMO consensus sites of the form yKxE as predicted by SUMOSp2.0 (52) with a “High” threshold.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Number of proteins</th>
<th>Number of predicted consensus sites</th>
<th>Number of proteins without consensus</th>
<th>Average consensus sites per protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Published SUMO substrates</td>
<td>264</td>
<td>517</td>
<td>49</td>
<td>2.0</td>
</tr>
<tr>
<td>TAP–SUMO-2 proteome</td>
<td>759</td>
<td>1,681</td>
<td>195</td>
<td>2.2</td>
</tr>
<tr>
<td>Human proteome†</td>
<td>43,964</td>
<td>28,078</td>
<td>28,571</td>
<td>0.6</td>
</tr>
<tr>
<td>TAP–SUMO-2 internal rejects</td>
<td>594</td>
<td>458</td>
<td>312</td>
<td>0.8</td>
</tr>
</tbody>
</table>

†Number of proteins included in this analysis.  †From a nonredundant human proteome data set from UniRef90. Data sets were filtered for redundancy. The shorter of two sequences with >90% pairwise sequence identity over >90% of the sequence length were removed.
Although advances in mass spectrometry have led to substantial increases in instrument sensitivity, it is apparent that the combined use of the TAP purification protocol and the double quantitative filtering with tSMaps provided the greatest advantage of our approach over previous methods. Specifically, nickel affinity purifications of proteins from cells containing 6His-SUMO-2 that were analyzed with the same equipment and quantitation software that we used, but that used only single quantitative filtering, identified 113 putative SUMO-2 conjugates. The purification and analytical techniques used and developed here not only provide a platform for future system-wide analysis of SUMO conjugates, but should also...

Fig. 5. Heat shock stimulates the polymerization of SUMO proteins. (A) Left: relative comparison of the frequency of log\(_2\) (TAP–SUMO-2 HSR:TAP–SUMO-2) (green lines) against log\(_2\) (TAP–SUMO-2 HS:TAP–SUMO-2) (red lines) for each of the six gel slices from SILAC experiment 2. Note the progressive tendency to higher values for the red traces compared with the green traces as the abundances of proteins of high molecular weight accumulate with heat shock. Right: frequencies of protein molecular weights found in each of the six slices. Predicted molecular weight range of each slice is shown by red boundaries. (B) Heat-map analysis of 40 targets of SUMO-2 comparing for each protein the contribution to the overall ratio data of an individual slice. Black indicates that the protein ratio is equal to the combined ratio; red indicates that the protein ratio is higher than the combined ratio, whereas green indicates it is lower. (C) Three-dimensional plot of the MS peak corresponding to the peptide fragment of the SUMO-2–SUMO-2 linkage through Lys11 from SILAC experiment 2.
be broadly applicable to the study of other Ubls. Furthermore, we expect that the use of tsMaps and GOtsMaps for filtering and presenting three-condition SILAC data is likely to have wider applications than in the Ubl field alone.

Our TAP–SUMO-2 substrate proteome identified 34% of the proteins already defined in the literature as substrates of SUMO (fig. S5). Although many substrates of SUMO-2 present in our cell lysates may have not been detected in our analysis, this apparently low figure is also likely to be a consequence of the fact that the reference set contains proteins that are not modified by SUMO-2. There are two reasons for this. First, because of the very low fraction of the cellular pool of any substrate that is modified...
by SUMO at any one time (41), few proteins have been shown to be SUMOylated without resorting to overexpression of SUMO, the protein substrate, the conjugation enzymes, or some combination of all three. Such a strategy has the potential to identify the wrong SUMO paralog as the physiological modifier, or it may even drive the SUMOylation of a protein that is not actually a target. Second, it is not clear from the literature precisely how much redundancy there is between SUMO paralogs. A relatively small-scale comparative proteomics study (4) suggests that about three-quarters of the targets of SUMO-1 are not shared with SUMO-2 and SUMO-3; however, genetic studies show that SUMO-2 and SUMO-3 appear to be able to compensate sufficiently for SUMO-1 in SUMO-1–deficient mice (5, 42). As such, with our current knowledge, we cannot definitively categorize many proteins as having a SUMO paralog–specific modification. These are important points, and further comparative studies of SUMO-1 and SUMO-2 and large-scale detection of SUMO-substrate branched peptides will be key to resolving them.

In contrast to our relatively thorough understanding of the biochemistry of the SUMOylation of proteins, the biological stimuli that regulate SUMOylation are poorly understood. Here, we showed that the regulated SUMOylation of many proteins that participate in important cellular processes occurred during hyperthermic stress. There was a broad continuum in the extent of conjugation or deconjugation after heat shock, ranging from more than 100-fold increased SUMOylation to 20-fold reduction in SUMO conjugation. Moreover, the reversal of these trends during the recovery period after heat shock was not as strictly temporally coupled. In general, those substrates that were conjugated upon exposure to heat shock were deconjugated 2 hours later, whereas deconjugated proteins tended not to recover their modification status over the same period (Fig. 4C and fig. S3B). Together, these features give the SUMO system the ability not only to rapidly respond to hyperthermic stress, but also to have a broad range of signal durations. For example, many transcription factors and chromatin-remodeling proteins were deconjugated from SUMO-2 upon heat shock, and this state of deconjugation persisted even 2 hours after removal of the stress. Conversely, many proteins involved in DNA damage response and repair were rapidly SUMOylated after heat shock and then deconjugated within 2 hours of recovery. This appears to be consistent with a role for SUMO in both short-term damage repair and longer-term protection from future insult, which is characteristic of the heat shock response (20).

These findings raise the important question of how the stress signal is transduced to lead to altered SUMOylation. Because modification by SUMO proteins was already almost maximal 5 min after exposure to heat shock (fig. S2), this suggests that transmission of the message from the sensor of increased temperature to the enzymes that influence SUMO modification must be fairly direct. This might involve either increasing the activity of

---

**Fig. 7.** Functional network analysis of SUMO-2 substrates. (A to D) Examples of protein modules identified in this study as being regulated by SUMO-2 and heat shock. Red and green nodes represent proteins that were more or less conjugated, respectively, to SUMO-2 after heat shock. Labels are gene names. The shapes of the nodes indicate protein function: rhombus, enzyme; ellipse, transcriptional regulator; circle, other function. Previously identified SUMO substrates are highlighted in bold. Solid lines indicate direct interactions, whereas broken lines indicate indirect interactions. An arrow from node X to Y indicates that X acts on Y.
the conjugation machinery or decreasing the activity of the SUMO-specific proteases (SENP) that are responsible for deconjugation of target proteins, although the hypermodification of SUMO E3 ligases after heat shock hints at the former explanation.

MATERIALS AND METHODS

Clonogenic cell survival assays
U2OS cells grown to 30% confluency in 38-mm-diameter wells of six-well plates were transfected with 10 nM siRNAs specific for SUMO-2 and SUMO-3 (Dharmacon) or a nontarget control with Lipofectamine RNAiMax (Invitrogen) as described by the manufacturer. Seventy-two hours later, cells were heat shocked at 45°C for 30 min and then incubated at 37°C for 10 to 14 days to allow colonies to form. Colonies were stained with Giemsa and counted.

DNA cloning and generation of stable cell lines
Human SUMO-2 [National Center for Biotechnology Information (NCBI) Entrez protein CAG46970] complementary DNA (cDNA) (encoding amino acid residues 1 to 93) was amplified by polymerase chain reaction (PCR) from pCDNA3.1 and subcloned into a bacterial expression vector pCMV5/TAP (provided by D. Alessi, University of Dundee, UK) at the BamHI and Hind III sites to generate pCMV5/TAP–SUMO-2. The TAP–SUMO-2 fragment was then subcloned into the eukaryotic expression vector pEFIRE-P (43) at the XhoI and XbaI sites. pEFIRE-P–TAP–SUMO-2 was transfected into HeLa cells and clones expressing TAP–SUMO-2 recombinant protein at an abundance comparable to that of endogenous SUMO-2 were selected. This procedure was repeated to produce negative control cells that contained only the TAP tag. All constructs were verified by automated DNA sequencing.

Cell culture and SILAC labeling
Quantitative proteomic experiments were performed with the SILAC technique as described (44). Briefly, cells were grown in Dulbecco’s modified Eagle’s medium lacking all amino acids except l-lysine and l-arginine, which were replaced with stable isotope (SILAC) forms (Cambridge Isotope Laboratories) depending on the treatment (see below). Medium was supplemented with 10% dialyzed fetal calf serum (FCS). SILAC experiment 1 compared TAP-containing cells (Lys6 and Arg6) with unstimulated TAP–SUMO-2–containing cells (4,4,5,5-D3-Lysine, Lys6, and Arg6) and TAP–SUMO-2–containing cells stimulated by heat shock at 43°C for 30 min (13C6-Lysine, Lys6, and 15N2-arginine, Arg6). SILAC experiment 2 included only TAP–SUMO-2–containing HeLa cells at 37°C (Lys6, Arg6), 30 min of heat shock at 43°C (Lys6, Arg6), and heat shock followed by 2 hours of recovery at 37°C (Lys6, Arg6). Cells were grown in 150-mm-diameter dishes and 50 dishes used per SILAC condition.

Tandem affinity purification (TAP) of SUMO-2 conjugates
When at around 90% confluence, cells were washed twice with phosphate-buffered saline (PBS) before being harvested in lysis buffer [50 mM tris-HCl (pH 8.0), 2% SDS, 10 mM iodoacetamide, 1 mM EDTA, complete protease inhibitor cocktail (Roche)] and stored at -80°C. Once thawed, lysates were sonicated, protein concentrations were determined by BCA (bicinchoninic acid) assay (Pierce), and equivalent amounts of protein from each lysate were mixed to give a volume of about 150 ml, which contained about 450 mg of total protein. The pooled lysate was centrifuged at 180,000 g at 20°C for 1 hour. The cloudy upper phase was removed, clarified by filtration, and added to the remaining supernatant. Proteins in this solution were refolded by 25-times dilution into renaturation buffer (RB) [50 mM tris-HCl (pH 8.0), 0.75 M NaCl, 1% NP-40, 2 mM iodoacetamide, and 0.5 mM EDTA]. TAP–SUMO-2 proteins were purified in two stages, first by protein A–Sepharose affinity chromatography. The refolded protein solution was passed over a 3.5-ml immunoglobulin G–Sepharose 6 Fast Flow column (GE) at 4°C, then washed with RB followed by tobacco etch virus (TEV) protease buffer [50 mM tris-HCl (pH 8.0), 0.75 M NaCl, 1% NP-40, 1 mM dithiothreitol (DTT), 0.5 mM EDTA]. Proteins were eluted by incubation of the beads in a 1:1 slurry of TEV protease buffer with 600 µg of TEV protease for 24 hours at 4°C. The supernatant from these beads was then used in the second step of purification, calmodulin affinity chromatography. For this stage, six volumes of calmodulin binding buffer (CBB) [50 mM tris-HCl (pH 8.0), 0.75 M NaCl, 1 mM DTT, 1 mM magnesium acetate, 1 mM imidazole, 1% NP-40, and 2 mM CaCl2], was added to the TEV flow-through and the final concentration of CaCl2 was adjusted to 2 mM. This mixture was then combined with 1.2 ml of calmodulin Sepharose 4B beads (GE) and mixed for 3 hours at 4°C. The calmodulin beads were then washed with CBB, and calmodulin-bound proteins were eluted with 3 ml and another 2 ml of calmodulin elution buffer (CEB) [50 mM tris-HCl (pH 8.0), 0.75 M NaCl, 1 mM DTT, 1 mM magnesium acetate, 1 mM imidazole, 1% NP-40, and 10 mM EGTA]. Eluted proteins were precipitated with trichloroacetic acid and resuspended to 30 µl before analysis by Coomassie-stained SDS-PAGE (Invitrogen, NuPAGE 10% Bis-tris). The lane containing SUMO-2–purified proteins was then sliced into six sections before undergoing in-gel tryptic digestion.

Quantitative mass spectrometry
Protein bands were subjected to in-gel digestion with trypsin (Promega), essentially as described previously (45). The resulting peptide mixtures were desalted and concentrated with self-made reversed-phase C18 stop and go extraction (STAGE) tips (46). Mass spectrometric analysis was performed by LC-MS/MS with a linear ion trap–orbitrap hybrid mass spectrometer (LTQ-Orbitrap, Thermo Fisher Scientific) equipped with a nano-electrospray ion source (Proxeon Biosystems) and coupled to an Agilent 1100 nano-HPLC system (Agilent Technologies). Peptides were injected into an in-house made 75-µm reversed-phase C18 column with a flow rate of 500 nl/min and eluted with a 120-min linear gradient of 98% solvent A (0.5% acetic acid in water) to 50% solvent B (80% acetonitrile, 0.5% acetic acid in water) and a flow rate of 250 nl/min. The instrument was operated with the “lock mass” option to improve the mass accuracy of precursor ions and data were acquired in the data-dependent mode to automatically switch between MS and MS/MS acquisition. Full-scan spectra [mass/charge ratio (m/z) 300 to 1600] were acquired in the orbitrap with resolution $R = 60,000$ at $m/z$ 400 (after accumulation to a target value of 1,000,000). The five most intense ions were fragmented by collision-induced dissociation and recorded in the linear ion trap (target value of 5000) based on a low-resolution ($R = 15,000$) preview of the survey scan.

Quantitative data analysis
For the purposes of analysis, the data from the raw mass spectrometric data files for SILAC experiments 1 and 2 were collated into a single data set and were processed with the quantitative processing software MaxQuant (version 1.0.9.7) (12) and with the Mascot search engine (Matrix Science, version 2.1.04). Further details of the algorithms and principles applied in the MaxQuant package have been described previously (12). Enzyme specificity was set to that of trypsin, allowing for cleavage N-terminal to proline residues and between aspartic acid and proline residues. This considers not only tryptic products but also accounts for the fact that the amide bond between aspartic acid and proline is the weakest peptide bond and results in a breakdown product in acidic conditions (47). Setting the enzyme specificity by allowing for cleavage between aspartic acid and proline enabled the identification of peptides produced by this nonenzymatic cleavage. Cysteine carbamidomethylation was selected as a fixed modification, and methionine oxidation,
protein N-acetylation, and deamidation of asparagine and glutamine residues were searched as variable modifications. The data were searched against a target-decoy human International Protein Index (IPI) database (version 3.24) (49). The initial maximum allowed mass deviation was set to 7 parts per million (ppm) for peptide masses and 0.5 D for MS/MS peaks. The maximum peptide length was set to six amino acid residues and a maximum of three missed cleavages and three labeled amino acids were allowed. MS/MS spectra determined to be SILAC labeled in MaxQuant presearch were searched with the fixed modifications Arg\(^{10}\) and Lys\(^{10}\) or Arg\(^{15}\) and Lys\(^{15}\); for spectra with a SILAC state that was not determinable before the database search, Arg\(^{2}\), Arg\(^{3}\), Lys\(^{10}\), and Lys\(^{15}\) were taken as variable modifications. A false discovery rate (FDR) of 1% was required at both the protein and peptide levels. In addition to the FDR threshold, proteins were considered to be identified if they had at least one unique peptide, and they were considered quantified if they had at least two quantified SILAC pairs. The SUMO-2 branched peptides were manually found in the unmatched list, containing quantified but nonidentified peptides, essentially as described previously (6), and 1924 protein hits were identified (see Supplementary Materials).

A single parent filtered list was generated for both sets (SILAC experiments 1 and 2) that eliminated hits on the basis of a number of criteria, which included non-real proteins from the “Reverse” database; proteins with insufficient data (that is, zero value for any ratio for both experiments, zero unique peptides for both experiments); proteins defined as external rejects in the first SILAC experiment (that is, from the “contaminant” library, or with both M:L and H:L ratios <0.15); proteins defined as internal rejects in the first SILAC experiment (both 0.4 < M:L < 1.4 and 0.53 < H:M < 1.53); and proteins suspected to be contaminants but not already excluded (including keratins, immunoglobulin proteins, and calmodulin). This parent list of proteins was then used to create a list of targets of SUMO-2 for SILAC experiment 1 by removing all proteins lacking any numerical value for M:L, H:L, or H:M ratio for this experiment. This yielded 662 putative substrates of SUMO-2. For SILAC experiment 2, the parent list was edited to remove proteins that fulfilled specific criteria in terms of data for SILAC experiment 2. These included: proteins lacking any numerical value for M:L, H:L, or H:M; external rejects (both M:L<0.15 and H:L<0.15); internal rejects (both 0.4 < M:L < 0.84 and 0.35 < H:L < 0.75), unless they were defined as substrates in the first experiment; and proteins rejected from the first SILAC experiment. This gave 672 putative substrates of SUMO-2. Supplementary data set 1 contains the data for all 766 proteins proposed to be substrates of SUMO-2 by this analysis, as well as those rejected. The effect of filtering on the data from SILAC experiments 1 and 2 are also shown (Figs. S1 and S3A).

For GOrTsMAP clustering analysis, the IPI protein codes for proteins identified by MaxQuant were matched to their appropriate Gene Ontology (GO) annotations with mappings provided by the GO annotation (GOA) service at the European Bioinformatics Institute (EBI) (49). The GO terms were matched to their parent terms with code used by the Gotcha (50) server (provided by D. Martin). Proteins with shared GO terms were clustered together and their log\(_2\) ratios from the M:L and H:M SILAC experiments were compared to those of all the proteins used as a background. Significant differences between the GO term–clustered proteins and the background were determined with the Wilcoxon test (51), and all GO terms with a \(P\) value < 0.05 were reported. The software used to generate GOrTsMaps will be described in more detail in a future publication. Protein network analysis was performed by using Ingenuity pathways analysis (Ingenuity Systems, www.ingenuity.com).

**SUPPLEMENTARY MATERIALS**

www.sciencesignal.org/cgi/content/full/2/72/ra24/DC1

Fig. S1. Double quantitative filtering with tsMaps.

Fig. S2. Effect of recovery from heat shock on SUMO-2 conjugates in HeLa cells.

**REFERENCES AND NOTES**


RESEARCH ARTICLE

31. S. Muller, M. Berger, F. Lehembre, J. S. Seeler, Y. Haupt, A. Dejean, c-Jun and p53 are stochastically larger than the other.
46. F. G. was funded through an EU Marie Curie research fellowship. Work in the R.T.H. laboratory was funded by Cancer Research UK, Wellcome Trust, and the Biotechnology and Biological Sciences Research Council. I.M. was supported by the EU FP6 Rubicon consortium and a European Molecular Biology Organization short-term fellowship. C.C. was funded by the Scottish Funding Council through the Scottish Bioinformatics Research Network. We would like to acknowledge help from C. Botting, R. Antrobus, D. Martin, and M. Scott.

Submitted 19 February 2009
Accepted 7 May 2009
Final Publication 26 May 2009
10.1126/science.1200028