Substrates Related to Chromatin and to RNA-Dependent Processes Are Modified by Arabidopsis SUMO Isoforms That Differ in a Conserved Residue with Influence on Desumoylation[1][W][OA]

Ruchika Budhiraja[2], Rebecca Hermkes[2], Stefan Müller, Jürgen Schmidt, Thomas Colby, Kishore Panigrahi, George Coupland, and Andreas Bachmair*

Department of Plant Developmental Biology (R.B., R.H., K.P., G.C., A.B.) and Mass Spectrometry Group (J.S., T.C.), Max Planck Institute for Plant Breeding Research, D–50829 Cologne, Germany; Institute of Biochemistry, University of Cologne, D–50931 Cologne, Germany (S.M.); and Department of Biochemistry, Max F. Perutz Laboratories, University of Vienna, A–1030 Vienna, Austria (A.B.)

The higher plant Arabidopsis (*Arabidopsis thaliana*) has eight genes potentially coding for small ubiquitin-related modifier (SUMO) proteins. However, two well-expressed isoforms differ from fungal and animal consensus in a conserved glutamine (Gln) residue situated four residues from the carboxyl terminus. We tested deviations in this position in the background of SUMO1, the isoform with the highest expression level, and found that changes do not prevent conjugation to substrate proteins in vivo. Replacement of this conserved Gln by alanine resulted in a protein that was less readily removed from a substrate by SUMO protease EARLY IN SHORT DAYS4 in an in vitro reaction and apparently led to higher levels of SUMO conjugates when expressed in vivo. We used the SUMO1 variant with the Gln-to-alanine substitution, as well as SUMO3 and SUMO5 (which carry methionine and leucine, respectively, at this position), to enrich in vivo substrates. Identification of the most abundant proteins contained in these fractions indicated that they are involved in DNA-related, or in RNA-dependent, processes, such as regulation of chromatin structure, splicing, or translation. The majority of the identified bona fide substrates contain predicted sumoylation sites. A subset of the proteins was expressed in *Escherichia coli* and could be sumoylated in vitro.

Protein modification allows for adjustment of protein properties following their synthesis and plays multiple roles in regulation. Among the large and growing number of modification types, attachment of SUMO (for small ubiquitin-related modifier) was found to regulate nuclear and cytoplasmic processes. SUMO is covalently linked to substrate proteins by an enzyme cascade of SUMO-activating enzyme (SAE) and SUMO-conjugating enzyme (SCE) and was shown in several cases to depend on SUMO ligases for substrate selection. SUMO attachment can be reversed by specific Cys proteases, which recognize the SUMO C terminus as exclusive substrate and hydrolyze its linkage to the target protein (Dohmen, 2004; Novatchkova et al., 2004; Kerscher et al., 2006; Dye and Schulman, 2007; Geiss-Friedlander and Melchior, 2007; Miura et al., 2007a). While usually only a minor fraction of a substrate protein is linked to SUMO at a given time, the modification can nonetheless be essential if it constitutes part of an activity cycle (Johnson, 2004). For instance, SUMO has been implicated in the assembly and disassembly processes of protein complexes. In some cases, sumoylation antagonizes or promotes other modifications such as ubiquitylation (Kerscher et al., 2006; Tatham et al., 2008).

In plants, SUMO was recently discovered to influence a variety of responses to the environment. SUMO is involved in tolerance to cold, heat, drought, and salt stress (Kurepa et al., 2003; Yoo et al., 2006; Catala et al., 2007; Miura et al., 2007b; Conti et al., 2008), modulates abscisic acid responses (Lois et al., 2003), has an important role in phosphate homeostasis (Miura et al., 2005), and controls the time of flower initiation (Murtas et al., 2003; Jin et al., 2008). The findings that pathogens manipulate SUMO conjugation (Hotson et al., 2003; Roden et al., 2004) and that sumoylation influences innate immunity (Lee et al., 2007) imply an important role in plant-pathogen relations.

Arabidopsis (*Arabidopsis thaliana*) contains eight SUMO genes (Kurepa et al., 2003; Novatchkova

[1] This work was supported by the Max Planck Society, by the German Research Foundation (grant no. SFB 635 to G.C. and grant nos. BA1158/3–1 and SPP1365 to A.B.), by the Austrian Science Foundation (grant no. P 21215–B12 to A.B.), and by predoctoral fellowships from the International Max Planck Research School to R.B. and R.H.

[2] These authors contributed equally to the article.

* Corresponding author; e-mail andreas.bachmair@univie.ac.at.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Andreas Bachmair (andreas.bachmair@univie.ac.at).

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.108.135053

et al., 2004), at least half of which are expressed to a significant extent. The significance of different isoforms is still unclear, but differences in the activity of SUMO proteases toward different isoforms have been reported (Chosed et al., 2006, 2007; Colby et al., 2006; Hay, 2007; Mukhopadhyay and Dasso, 2007). Here, we show that variations in one conserved residue do not abolish functionality but nonetheless influence the properties of a SUMO protein that contains the variant residue. Furthermore, we identify putative in vivo sumoylation substrates of Arabidopsis by enrichment of proteins conjugated to tagged SUMO isoforms, followed by mass spectrometric identification. Most of the identified proteins are functionally linked to transcription, chromatin modification, RNA metabolism, or translation. SUMO is a predominantly nuclear protein, but sumoylation also occurs in the cytoplasm, and the set of candidate substrates identified in this work encompasses both nuclear and cytoplasmic proteins.

RESULTS

SUMO Isoforms of Arabidopsis Are Differentially Expressed

The Arabidopsis genome encodes eight potential SUMO proteins (Kurepa et al., 2003; Novatchkova et al., 2004). SUMO1, -2, -3, and -5 are highly expressed and therefore probably functional. The relative expression levels are SUMO1 ~ SUMO2 > SUMO3 ~ SUMO5, with SUMO3 and SUMO5 mRNAs being approximately 7-fold less abundant than those of SUMO1 or SUMO2. Putative proteins SUMO4, -6, -7, and -8 are probably rare, because no cDNAs are found in EST libraries. Microarray data suggest low expression of SUMO4 (about 1% of the SUMO1 level) but are not informative regarding SUMO6, -7, or -8 (http://www.arabidopsis.org/), so it is unclear under which conditions these open reading frames are expressed.

SUMO Proteins with N-Terminal Extension Are Conjugated in Vivo

SUMO isoform 1 (At4g26840; Kurepa et al., 2003; Novatchkova et al., 2004) is highly expressed and well represented in EST databases. To analyze in vivo sumoylation, a *SUMO1* gene with an N-terminal extension was constructed. One extension (tag1) contains a hexa-His tag followed by a triple hemagglutinin (HA) tag (see "Materials and Methods"). A second extension (tag3), used for most of the work presented, contains a Strep tag, a triple HA tag, and octa-His (extension sequence shown in Fig. 1A). Mature SUMO proteins containing either of these extensions were expressed in Arabidopsis plants. Figure 2 (lane 3, extract from plants expressing tag1-SUMO1, versus control lanes 1 and 2, and lane 4, extract from plants expressing tag3-SUMO1) demonstrates that SUMO1 carrying either N-terminal extension is incorporated into higher molecular mass material.

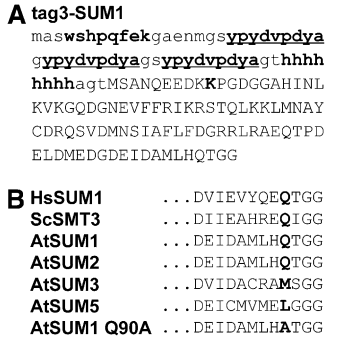

Figure 1. A, N-terminal extension of AtSUM1 (uppercase letters) by tag3 to allow detection and enrichment of sumoylated protein substrates with minimal disruption of functionality. tag3 (single-letter code in lowercase letters) consists of a Strep tag (boldface), three HA tags (boldface, underlined), and an octa-His sequence (boldface). The K residue in position 10 of the SUM1 sequence (boldface) was previously identified as a site of SUMO attachment in SUMO chains. B, Alignment of C termini of mature SUMO proteins (single-letter code). HsSUM1, Human SUMO1; ScSMT3, SUMO of *S. cerevisiae*; AtSUM1, -2, -3, and -5, the four highly expressed SUMO isoforms of Arabidopsis. The Gln at position 4 from the C terminus (boldface; position 90 in AtSUM1) is conserved in animal and fungal SUMO proteins but replaced by a hydrophobic amino acid in AtSUM3 and AtSUM5. In AtSUM1 Q90A, this residue was changed to Ala in the sequence of Arabidopsis SUMO1.

Because SUMO1 and -2 are highly similar, we did not include SUMO2 in these investigations. Functional equivalence of SUMO1 and -2 was indeed recently demonstrated (Saracco et al., 2007). We did, however, use the distinct isoforms SUMO3 and -5 for in vivo and in vitro experiments. We thus constructed transgenes for in planta expression of tag3-SUMO3 (At5g55170) and tag3-SUMO5 (At2g32765). Figure 2 shows that tag3-SUMO3 and tag3-SUMO5 are also incorporated into higher molecular mass material, suggestive of conjugation. There are subtle differences in the pattern of conjugates between SUMO1, -3, and -5. The high number of bands, however, does not permit conclusions regarding whether the differences pertain to relative abundance of conjugates present for all three isoforms or whether any of the substrates of SUMO3 and SUMO5 conjugation are distinct from those of SUMO1. For further work, tag3-SUMO1, -3, and -5 were expressed under the control of a β-estradiol-inducible promoter (vector pER8; Zuo et al., 2000). Figure 2 (lanes 8 and 9) shows the inducible tag3-SUMO5 construct as an example.

Differences between SUMO Isoforms in a Usually Conserved Residue Do Not Prevent Conjugation But May Influence Isoform Characteristics

Comparison of the SUMO gene family of Arabidopsis indicated that, interestingly, some of the SUMO isoforms of Arabidopsis do not contain residues con-

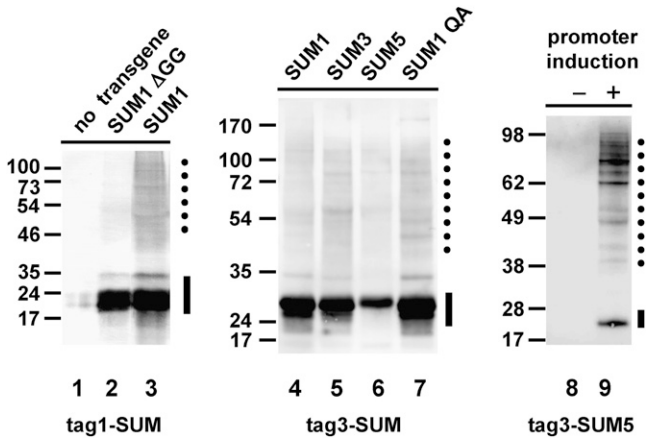

Figure 2. SUMO isoforms with N-terminal extension are conjugated to protein substrates in vivo. Extracts from plants expressing different SUMO constructs were used for protein blotting and detection with antibody directed against the HA tag of the extension. SUMO1 (extended by tag1 at the N terminus as indicated in the text; lane 3) was compared with extract from plants expressing SUMO1 without the two C-terminal residues (SUMO1 ΔGG; lane 2), indicating dependence of the conjugation reaction on an intact C terminus. Lane 1 shows an extract from nontransgenic plants. SUMO isoforms SUMO3 and -5 (lanes 5 and 6, respectively), as well as a variant of SUMO1 with change Q90A (lane 7), N-terminally extended by tag3 (compare with Fig. 1), can also form conjugates in vivo. The conjugate patterns of SUMO1 and SUMO1 Q90A are similar. Lanes 8 and 9 show examples of SUMO expression by an inducible promoter. Extracts from cells transformed with an inducible tag3-SUMO5 construct contain conjugates to this protein only after induction (lane 9 versus lane 8). Bars at right indicate the position of unconjugated tag-SUMO, and the dotted line indicates the positions of SUMO conjugates. Molecular mass markers in kD are indicated at left.

served in animal or fungal SUMO proteins (Kurepa et al., 2003; Novatchkova et al., 2004). In particular, the usually conserved Gln residue four amino acids from the C terminus is replaced by Met and Leu, respectively, in Arabidopsis isoforms SUMO3 and SUMO5 (Fig. 1B). With one exception (vertebrate SUMO4), deviation from the Gln consensus is apparently restricted to flowering plants (Supplemental Fig. S1). Based on structural data, this Gln was predicted to be important for interaction with SUMO-specific proteases (Mossessova and Lima, 2000) and may thus influence the efficiency of deconjugation. Detection of in vivo conjugates (Fig. 2) indicates that changes at this position, as present in SUMO3 (Met instead of Gln) and SUMO5 (Leu instead of Gln), do not prevent conjugation, suggesting that the conserved Gln residue is not essential for the conjugation reaction. We engineered different residues into this position of SUMO1. Mutation of Gln to a hydrophobic (Leu), a basic (Arg), an acidic (Asp), or a small (Ala) residue was tested (Budhiraja, 2005). Interestingly, high level overexpression of (N-terminally extended) SUMO1 with the Gln-to-Leu or the Gln-to-Ala change had a deleterious effect on plant growth (shown for the Ala variant in Supplemental Fig. S2A; Budhiraja, 2005), whereas similar overexpression of wild-type SUMO1 (with N-terminal extension) resulted in only a slight retardation of growth (Supplemental Fig. S2B). We concluded that change of Gln to a hydrophobic residue, or to Ala, leads to changes in biological properties. In order to avoid possible negative effects of high overexpression, we used lower expression levels and an inducible promoter system for the enrichment of in vivo substrates described below.

As a first biochemical characterization of SUMO1 Q90A, we capitalize in the following on the consequences of the SUMO1 Gln-to-Ala (Q90A) mutation on conjugation and deconjugation reactions. Figure 2 (lane 7) shows that SUMO1 Q90A expression leads to a similar conjugate pattern as SUMO1 expression. However, we consistently observed a slightly more intense antibody staining (Fig. 2; data not shown; Coomassie Brilliant Blue staining of control gels served to adjust to equal protein loading), suggestive of an increased presence of the variant in conjugates.

The SUMO1 Q90A Mutation Decreases Desumoylation by SUMO Protease EARLY IN SHORT DAYS4

Previous data had indicated that the activity of SUMO proteases toward different SUMO isoforms can differ significantly (Chosed et al., 2006, 2007; Colby et al., 2006; Hay, 2007; Mukhopadhyay and Dasso, 2007). The availability of sumoylated, recombinant plant proteins as described below allowed an in vitro investigation of the differences between SUMO1 and SUMO1 Q90A regarding desumoylation. The activity of desumoylating enzyme EARLY IN SHORT DAYS4 (ESD4; Murtas et al., 2003) on SUMO1 versus SUMO1 Q90A was assessed in a qualitative way. An in vitro sumoylation reaction was established using nucleosome assembly factor (NAF; At2g19480; see below and Table I) as a substrate and either SUMO1 or SUMO1 Q90A. Sumoylated and nonsumoylated substrate proteins were removed from sumoylation enzymes using affinity beads specific for the Flag tag on NAF. The isolated, bead-associated proteins were incubated with ESD4 protease fragment and analyzed by western blotting. Figure 3A shows that, whereas SUMO1 Q90A-NAF was gradually desumoylated over a 60-min incubation period, all SUMO1-NAF had been deconjugated by 15 min under the same conditions, suggesting that SUMO1 deconjugation was at least four times faster. In another experiment with less protease, we observed deconjugation of 70% of the SUMO1-NAF conjugate after 30 min of incubation. However, no deconjugation of SUMO1 Q90A could be observed under these conditions (Fig. 3B). In the latter experiment, secondary antibodies carried an infrared dye, which allowed quantification by measurement of emitted light after excitation. These experiments were repeated and were consistent with the notion that SUMO1 Q90A can be deconjugated by SUMO proteases but that its deconjugation is significantly delayed compared with the unmutated SUMO1. The exact

Table I. *Putative in vivo sumoylation substrates of Arabidopsis*

Criteria for inclusion of a protein were robust detection in preparations, absence from control preparations, and apparent molecular mass, as detailed in the text. n.t., Not tested.

Identifier	Characteristics	ΨKxD/E[a]	In Vivo Conjugate to:	In Vitro Sumoylation[b]
RNA-dependent processes				
At1g29400	AML5; has one RRM2 domain, two RRM1 domains	1 (1)	SUM3	Yes
At3g56860	Contains two RRM1 domains	1 (0)	SUM1	Yes
At2g43970	Contains an LA domain	0 (1)	SUM1	No
At2g47020	Translation release factor	1 (2)	SUM3	n.t.
At3g60240	eIF4G	4 (2)	SUM3	n.t.
Chromatin-related processes				
At1g55300	TAF7 (TFIID subunit)	0 (0)	SUM3, -5	Yes
At2g19480	Nucleosome assembly factor	2 (0)	SUM1	Yes
At5g08450	Contains an Rxt3 domain	3 (10)	SUM3	Yes
At5g39660	zfDOF transcription factor	1 (1)	SUM1	n.t.
At5g43130	TFIID subunit	2 (4)	SUM1	n.t.
Others				
At1g27430	Contains a GYF domain	1 (4)	SUM1, -3	n.t.
At1g64330	Contains a KIP domain, coiled-coil region	2 (1)	SUM1, -3	n.t.
At5g13480	FY; WD40 repeat protein	0 (0)	SUM1, -3	n.t.
At5g52300	RD29B desiccation induced	3 (2)	SUM1	n.t.

[a]In addition to canonical SUMO attachment sites (Ψ is an aliphatic hydrophobic residue, x indicates any residue, D/E is either Asp or Glu), many proteins contain nearly canonical sequences as further potential SUMO attachment sites (number of such sites is given in parentheses; www.abgent.com/doc/sumoplot). [b]The in vitro sumoylation reactions referred to here used SUMO1 and its variant SUMO1 Q90A with identical results.

magnitude of the difference remains to be determined and may differ for different SUMO proteases.

Enrichment and Mass Spectrometric Identification of SUMO Conjugates

Based on tag-SUMO-expressing plants, an enrichment procedure was established to facilitate mass spectrometric identification of in vivo sumoylated proteins. As it turned out that deconjugation was a major problem during enrichment, denaturing conditions were used whenever possible. Figure 4 documents steps of the enrichment procedure. Both anti-SUMO1 and anti-HA antibodies allowed detection of SUMO conjugates in enriched fractions. However, because the polyclonal anti-SUMO1 serum showed additional bands (Fig. 4A), the monoclonal HA antibody was used preferentially to follow the enrichment. Protein extracts from mature whole plants (see "Materials and Methods") were incubated with nickel affinity resin to bind proteins with hexa-His tag, which were subsequently eluted. Comparison of the amount of free tag-SUMO contained in crude extracts versus proteins eluted from nickel affinity resin suggested that there was an at least 50-fold enrichment (western blot using anti-HA antibody; Fig. 4B). Figure 4C shows a comparison of a Coomassie Brilliant Blue-stained gel of nickel resin eluate (lane 5) with a parallel western blot using anti-HA antibody (lane 6). While the western blot documents the enrichment of sumoylated proteins, these proteins do not in general correspond to visible bands in the Coomassie Brilliant Blue-stained preparation. However, application of proteins eluted from the nickel column onto anti-HA resin allows a further enrichment. Figure 4D shows a comparison of Coomassie Brilliant Blue staining (lane 7) and western blotting (lane 8) with such a fraction. In this case, many major bands of the western blot coincide with visible bands on the Coomassie Brilliant Blue-stained gel. Therefore, we concluded that the protein fraction obtained from the nickel affinity step contains a significant amount of sumoylated proteins but that these proteins are usually overlaid by protein contaminants. Because the anti-HA affinity step requires seminative conditions (the sample was diluted to 0.3 M urea in order to allow binding), the procedure results in considerable loss of signal (as judged by western blotting) and variable yields (see "Discussion"). This step, while useful to assess the abundance of sumoylated proteins in enriched fractions, was therefore not used to generate material for mass spectrometry. Figure 4E shows a typical preparative gel, which was cut into pieces according to molecular mass and used for mass spectrometry.

The (moderately) increased level of conjugates with SUMO1 Q90A compared with SUMO1 (Fig. 2; data not shown) motivated the use of this SUMO1 variant to enrich SUMO1 conjugates for identification. SUMO3 and SUMO5 were used in similar experiments. Because we found that high-level expression of SUMO1 Q90A, SUMO3, or SUMO5 with a constitutive promoter delayed growth and could even cause senes-

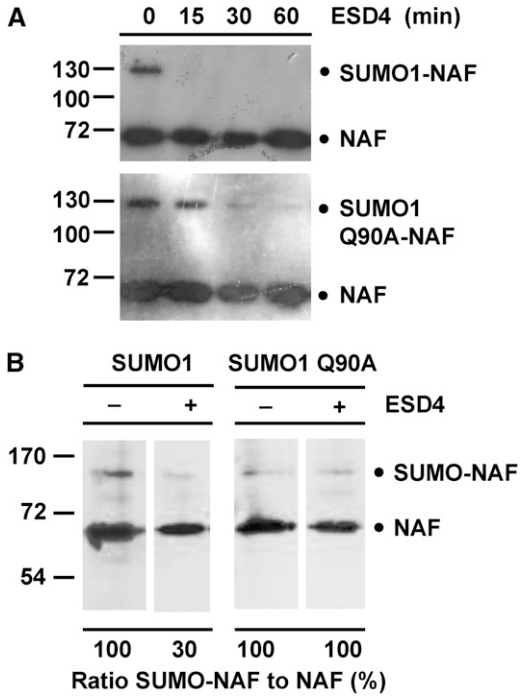

Figure 3. Qualitative assessment of in vitro desumoylation of a substrate conjugated to SUMO1 versus SUMO1 Q90A. NAF protein with Flag tag was purified from *E. coli* and subjected to in vitro sumoylation, using either SUMO1 or SUMO1 Q90A. After reisolation with anti-Flag resin, the material was incubated with SUMO protease ESD4 and harvested at the indicated times. Reaction products were detected on western blots using anti-Flag antibody. SUMO1 Q90A-NAF (A, bottom panel) was gradually desumoylated, while SUMO1-NAF (A, top panel) was deconjugated quantitatively within 15 min. At lower ESD4 concentration (B), 70% of SUMO1-NAF was desumoylated in a 30-min incubation, while SUMO1 Q90A-NAF was almost unchanged. Band intensities in B were quantified with secondary antibody coupled to infrared dye and detection of light emission. Unmodified NAF protein was used as an internal standard.

cence-like symptoms on leaf margins (Supplemental Fig. S2; Budhiraja, 2005), we used estradiol-inducible constructs for these experiments (Fig. 2, lanes 8 and 9). Control experiments indicated that the expression levels achieved with the inducible system do not compromise plant growth (Supplemental Fig. 2; data not shown). Greenhouse-grown mature transgenic plants were induced by spraying with 5 μM β-estradiol. At 16 h after spraying, the material was harvested and processed as described above and in "Materials and Methods." When the procedure was applied to nontransgenic Arabidopsis plants, the amount of protein recovered was slightly smaller than the material recovered from plants induced to express tag3-SUMO1 Q90A (Fig. 4E, lane 10 versus lane 9). Gel lanes were cut into pieces, to represent defined protein size fractions, and processed for protein identification by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Similar protein extracts were prepared from the plants expressing tag3-SUMO3 and tag3-SUMO5.

A number of proteins were identified in extracts from plants that did not contain SUMO transgenes with hexa-His tag. Control experiments suggested that the major contaminant, Rubisco (visible as the strongest band at the 54-kD marker in Fig. 4, C–E, but present in each size fraction of the gel), is significantly retained on the Sephadex-based matrix irrespective of the nickel-nitrilotriacetic acid agarose (Ni-NTA) affinity ligand and constitutes the majority of the nonspecific background. Other protein contaminants, such as catalases, may be natural chelators of the nickel affinity ligand. For some proteins, such as At1g80480 (annotated as PRL1-interacting protein L), the latter property was also supported by poly-His tract(s) in the protein sequence. Phospholipase D α1 (At3g15730) was another prominent contaminant.

Next to proteins enriched and identified from wild-type plant extracts, which were obvious contaminants in preparations from transgenic plants, there were many proteins specific for preparations from tag3-SUMO-expressing plants. The candidate list contains only proteins for which at least three polypeptides were sequenced from enriched fractions but no peptide was present in control preparations. The apparent molecular mass of identified proteins, as determined by the position on the gel, was a further criterion to decide whether the protein was a likely SUMO substrate. The conjugation of transgenic SUMO to a protein is supposed to increase its apparent molecular mass. Thus, only proteins that migrated 20 to 40 kD above their expected molecular mass were retained on candidate lists. Table I shows putative sumoylation targets with the highest scores. The identification method has an obvious bias for more abundant proteins. For instance, peptides from transcription factors were frequently identified (e.g. from IAA18/At1g51950, from Perianthia/At1g685409, or from FWA/At4g25530), but more than two peptides were rarely identified on one gel piece, so that scores for these candidate substrates stayed relatively low. Likewise, peptide WSVIAR, which occurs in several Myb domain transcription factors, was consistently obtained.

Canonical or high-probability sumoylation sites (sumoplot www.abgent.com) are statistically enriched among candidate proteins (see "Discussion"). The most noteworthy candidate in this respect is At5g08450, which contains more than 10 high-probability sumoylation sites. The presence of an Rxt3 domain in this protein suggests participation in histone deacetylation. An exception to the trend is At1g55300 (TAF7), a subunit of general transcription factor TFIID (Thomas and Chiang, 2006), which has no predicted high-probability sumoylation site.

In Vitro Sumoylation of Candidate Proteins

The mass spectrometric information regarding sumoylation substrates, as listed in Table I, was complemented by in vitro sumoylation assays with a subset of

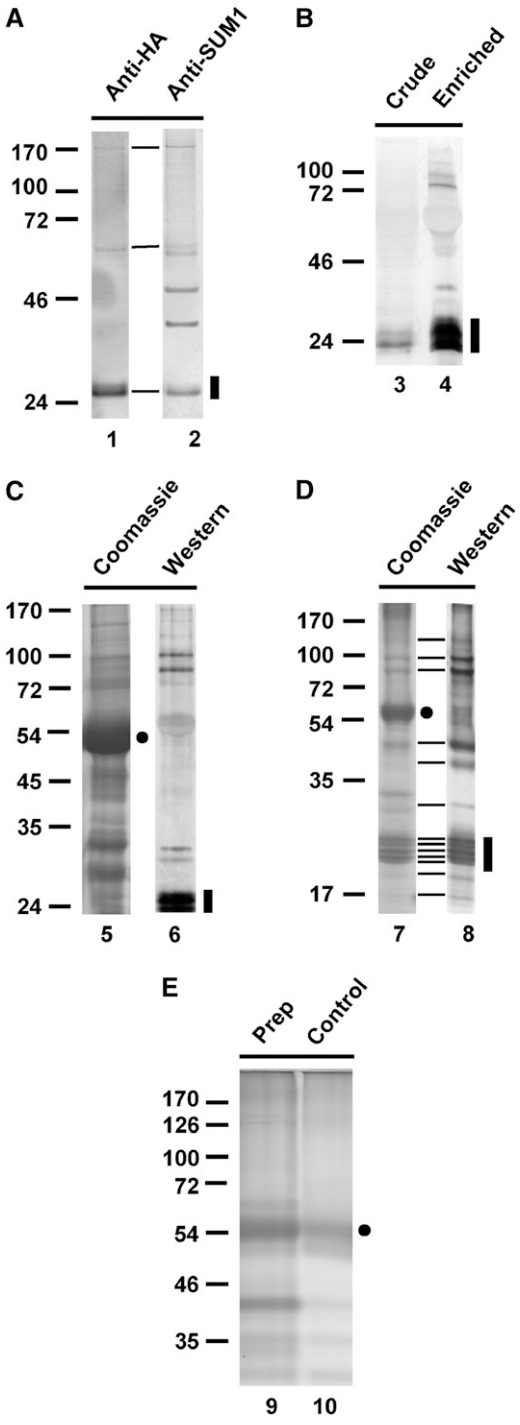

Figure 4. Enrichment of sumoylated proteins. A, An enriched protein fraction from tag-SUMO-expressing plants was used for antibody detection using either rat monoclonal anti-HA antibody (lane 1) or polyclonal anti-SUMO antiserum (lane 2). The polyclonal serum detects additional bands. B, Crude plant extract (lane 3) was compared with an enriched fraction (lane 4) using gel blot and anti-HA detection. The intensity difference of the tag-SUMO band (bar at right) suggests an approximately 50-fold enrichment of SUMO and sumoylated proteins. C, Side-by-side comparison of a Coomassie Brilliant Blue-stained preparative gel of an enriched protein fraction (lane 5) with a western blot of the same material (lane 6). The blot demonstrates enrichment of sumoylated proteins, but the bands detected by the anti-HA antibody do not comigrate with prominent bands of the Coomassie Brilliant Blue-stained gel. D, Material as shown in C was further purified using anti-HA affinity chromatography. The Coomassie Brilliant Blue-stained gel (lane 7) has prominent bands at the position of the most intense bands of the western blot (lane 8), indicating that preparations as shown in C do contain significant amounts of SUMO conjugates. E, Coomassie Brilliant Blue-stained gel loaded with protein extract enriched by one step (Ni^{2+} affinity; compare with C), as used for mass spectrometric protein identification. Lane 9, extract from tag-SUMO transgenic plants (Prep); lane 10, material from nontransgenic plants (background control). In A and D, prominent bands visualized by two distinct detection methods are linked by horizontal lines. Dots show the major contaminant Rubisco large subunit; bars at right show the position of tag-SUMO. Molecular mass markers in kD are indicated at left.

the identified proteins. Arabidopsis has two forms of SAE that share the same large subunit but differ in the smaller subunit (Kurepa et al., 2003; Novatchkova et al., 2004). Previous knowledge suggested that if produced in *Escherichia coli*, both subunits should be coexpressed for optimal enzyme activity (F. Melchior, personal communication). Therefore, we constructed a dicistronic mRNA that expressed the larger SAE subunit, SAE2 (At2g21470; Kurepa et al., 2003; Novatchkova et al., 2004), as the first reading frame, followed by either SAE1a or SAE1b with an N-terminal hexa-His tag. The two open reading frames (ORFs) were linked as found in a polycistronic mRNA of *E. coli* that encodes ribosomal proteins (Yates and Nomura, 1980). We expected this arrangement to result in minimal ribosome disassembly at the end of the first ORF, allowing close to equimolar production of both proteins. SCE (At3g57870), a single-copy gene in Arabidopsis, was expressed and purified without extension. SUMO1 and SUMO1 Q90A were expressed with the N-terminal extension peptide as used for expression in planta (Fig. 1A). In most experiments, a fragment of SUMO ligase SIZ1 (At5g60410; Miura et al., 2005) was included in the reaction. This SIZ1 clone contained the 478 C-terminal residues that include the SP-RING. It lacks a putative chromatin-binding domain to increase its solubility in *E. coli* but also lacks the PHD finger that was recently shown to contribute to certain sumoylation reactions (Garcia-Dominguez et al., 2008).

SUMO chain formation was used as a test of sumoylation components (Fig. 5). These experiments allowed the following conclusions. First, SAE preparations containing SAE1b fared as well as, and occasionally even better than, those containing SAE1a (SAE1b lanes versus SAE1a lanes of reaction with SUM1 in Fig. 5). The former isoform was used in substrate-based assays shown in the following. Second, consistent with previous reports, SUMO1 formed chains in vitro (Colby et al., 2006; putative sumoylation site indicated in Fig. 1A). It is possible, therefore, that SUMO1 chain formation is functionally important in plants, similar to the recently reported role in yeast and animals (Uzunova et al., 2007; Tatham et al., 2008).

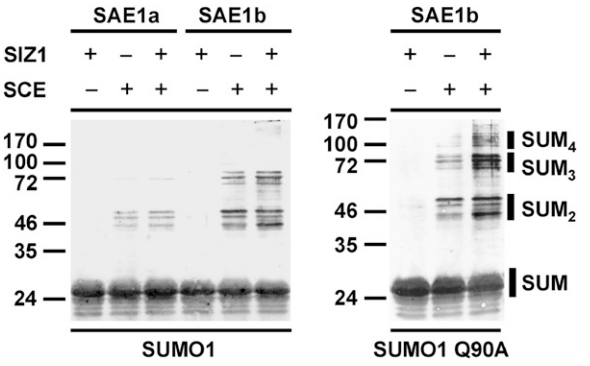

Figure 5. Test of sumoylation enzymes by formation of SUMO chains. SAEs SAE1a/SAE2 or SAE1b/SAE2 (labeling at top) were incubated with SCE (+ label) or without SCE (− label) together with tag3-SUMO1 (left panel) or tag3-SUMO1 Q90A (right panel) in the presence (+) or absence (−) of SUMO ligase SIZ1. Note that SUMO1 can form SUMO-SUMO linkages in a pattern that is identical to that of SUMO1 Q90A. Bars at right indicate the positions of di-, tri-, and tetra-SUMO (several bands, presumably due to conformation differences or to different SUMO attachment sites). Bands were detected with antibody against the HA epitope present on SUMO.

In contrast, experiments with SUMO3 and SUMO5 suggested that these latter two isoforms have a much lower tendency for chain formation in the same assay (data not shown; see also substrate-based assays in Fig. 7 below). Third, SUMO ligase SIZ1 enhanced in vitro conjugation (+SIZ1 lanes versus –SIZ lanes in Fig. 5). This effect was also found in most substrate-based assays. Fourth, SUMO1 variant SUMO1 Q90A behaved similar to SUMO1 in these conjugation reactions (left panel versus right panel in Fig. 5). This finding could be confirmed in all substrate-based assays (see below).

Protein tags used for immunological detection of potential substrates were tested in the in vitro sumoylation system to find out whether they contained sumoylation sites (Fig. 6). Whereas the Flag peptide, used as a C-terminal extension, seems to be neutral regarding SUMO1 attachment, the glutathione *S*-transferase (GST) tag as encoded in vector pET42c may be sumoylated in certain contexts in the S peptide, which is part of the sequence connecting GST core and polylinker. Therefore, we used only clones without the S peptide sequence in the following.

Candidate proteins obtained from mass spectrometric analysis that could be expressed in *E. coli* were used for in vitro sumoylation. Results are shown in Figure 7 and listed in Table I. Sumoylation patterns were identical between SUMO1 and SUMO1 Q90A. In contrast, SUMO3 gave a different pattern. Consistent with the lower tendency to form SUMO3-SUMO3 chains, only one higher molecular mass band was obtained with NAF (At2g19480) or with TAF7 (At1g55300). The protein encoded by At5g08450 (labeled RXT3 in Fig. 7) contains 13 predicted high-probability sumoylation sites. Its in vitro sumoylation results in multiple bands, consistent with SUMO attachment at several positions. Isolation and mass spectrometric analysis of SUMO1-NAF and SUMO1-RRM confirmed that sumoylation did not occur in the Flag tag of these proteins (T. Colby, unpublished data).

DISCUSSION

In this work, we describe biochemical properties of the Arabidopsis sumoylation system based on in vivo and in vitro experiments and compile a short list of potential in vivo sumoylation substrates of Arabidopsis.

Arabidopsis encodes eight SUMO proteins (Kurepa et al., 2003; Novatchkova et al., 2004). Four of them are expressed to a considerable extent. Because they differ in an amino acid residue that is conserved in fungal and most animal SUMO proteins, we tested amino acid changes in the sequence context of SUMO1, the most abundant isoform, which carries the standard residue in this position. SUMO3 and -5 have large

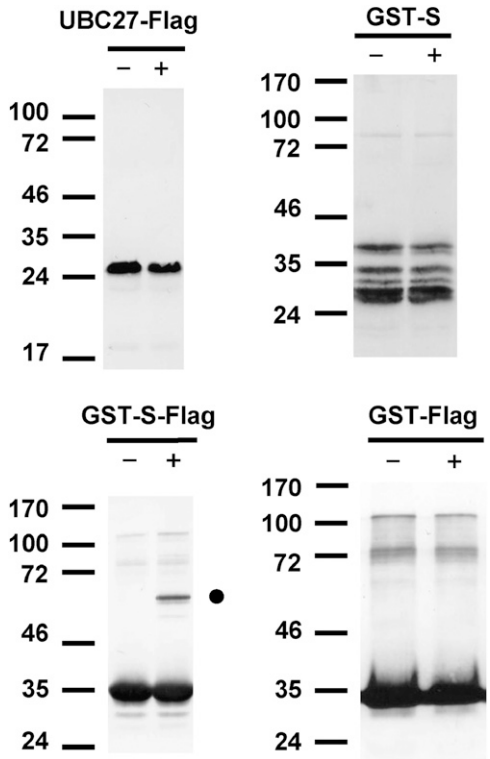

Figure 6. In vitro sumoylation of protein tags. UBC27-Flag, a ubiquitin-conjugating enzyme with a Flag tag extension, or GST encoded by vector pET42c (GST-S) can serve as a negative control. However, extension of GST-S by a Flag tag converts GST into a substrate (GST-S-Flag; dot at right indicates the position of the sumoylated form). Sumoylation is abolished by deletion of the S peptide from the linker region between the GST core and Flag peptide (GST-Flag). Control lanes (−) contain an inactive (Cys-94-to-Ser change) version of SCE; + indicates the presence of wild-type SCE. Proteins were detected by western blotting, using antibodies against Flag, of GST tags.

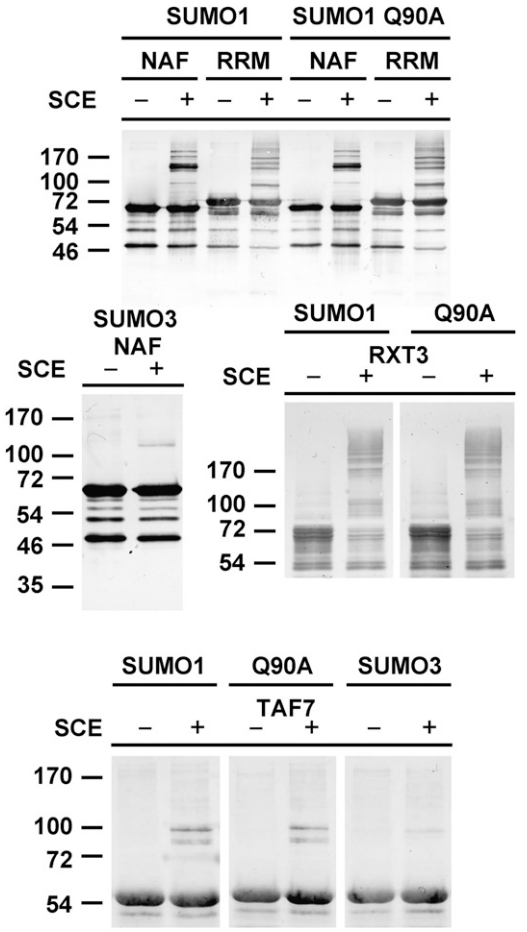

Figure 7. In vitro sumoylation of potential in vivo substrate proteins. Proteins identified as potential in vivo substrates were expressed in *E. coli*, purified, and incubated with sumoylation enzymes. NAF, At2g19480; RRM, At3g56860; RXT3, At5g08450; TAF7, At1g55300. Control lanes labeled SCE (−) contain an inactive (Cys-94-to-Ser change) version of SCE. SUMO1 and SUMO1 Q90A give identical patterns, while SUMO3 results in a single, considerably weaker sumoylation band of NAF or TAF7. Proteins NAF and RRM have a C-terminal Flag peptide that was used for antibody-based detection, whereas RXT3 and TAF7 have an N-terminal GST tag (without the S peptide) and were detected by anti-GST antibody.

hydrophobic residues instead of a conserved Gln four residues from the C terminus (Novatchkova et al., 2004). Mutations at this position in SUMO1 were tested (Budhiraja, 2005; this work). One of them, SUMO1 Q90A, was used for further characterization. This variant had an in vivo conjugation pattern similar to SUMO1 (Fig. 2), and in vitro sumoylation reactions resulted in an identical pattern of conjugate bands for all proteins tested (Fig. 7). However, the change resulted in decreased desumoylation of the variant (Fig. 3). This finding is consistent with structural data (Mossessova and Lima, 2000), and in fact, testing variants at this position was in part motivated by the prospects to obtain a "noncleavable" SUMO1 variant. Deconjugation was assessed by qualitative assays using an in vitro sumoylated substrate and a fragment of the Arabidopsis SUMO protease ESD4 (Murtas et al., 2003). The data shown in Figure 3, and additional experiments, indicate that SUMO1 Q90A is cleaved off the model substrate NAF (At2g19480) but at least four times slower than SUMO1. This result has obvious implications for SUMO3 and -5, which have Met and Leu, respectively, in this position. We hypothesize that the previously found differences in in vitro deconjugation between different Arabidopsis SUMO isoforms (Chosed et al., 2006; Colby et al., 2006) are to a significant extent due to the nature of this critical residue. Distinct deconjugation rates of different isoforms should translate into different lifetimes for conjugated substrate proteins in vivo, thus implying different roles for different isoforms, even if the spectrum of substrates as such does not differ.

It is likely that changes close to the SUMO C terminus, as tested in the context of substrate conjugation and deconjugation, also have an influence on the maturation of SUMO precursors, which is necessary to expose the mature SUMO C terminus. Assuming that the half-life of SUMO isoforms SUMO3 and -5 largely exceeds the average half-life of their conjugates, a slow maturation rate would still allow the majority of SUMO protein to be in the mature (i.e. active) form and thus would not have a significant influence on protein properties. However, it is also possible that maturation of SUMO3 and -5 is a regulated step, so that SUMO3/5 precursors can accumulate under certain circumstances. In this latter case, the amount of SUMO3/5 available for conjugation would differ from expectations based on mRNA abundance, but the properties of mature SUMO isoforms remain as suggested by this work. This proposition differs from findings concerning a human SUMO isoform, hSUMO4, which has Pro at position −4. In contrast to Arabidopsis SUMO1 Q90A, which can be slowly desumoylated in vitro (Fig. 3), the proteolytic processing of human SUMO1 with Pro at position −4, or of hSUMO4, could not be demonstrated in vitro. Therefore, it was proposed that hSUMO4 is not conjugated in vivo (Owerbach et al., 2005).

Arabidopsis SUMO1 variant Q90A, as well as SUMO3 and SUMO5, were expressed in vivo with N-terminal tags that served for enrichment of potential SUMO conjugates. Experiments to obtain fractions enriched for SUMO conjugates have been described in other organisms (Li et al., 2004; Panse et al., 2004; Wohlschlegel et al., 2004; Zhao et al., 2004; Zhou et al., 2004; Hannich et al., 2005). Some of these publications, dealing with *Saccharomyces cerevisiae*, describe a one-step purification scheme employing a hexa-His tag. However, a similarly straightforward approach in Arabidopsis appeared more challenging due to the more complex proteome and concomitant decrease in signal-to-noise ratios. Whereas use of nickel affinity in combination with denaturing protein extract preparation resulted in a consistent enrichment of sumoylated proteins as documented in Figure 4, subsequent use of

the HA tag in a two-step procedure resulted in variable yield and significant losses. A possible reason was that HA affinity purification with an antibody column does not tolerate completely denaturing conditions, and coenrichment of proteases (e.g. At1g09730, a predicted SUMO-specific protease, could be identified as a low-scoring component of enriched fractions) might pose a constant danger. For this reason, single-step enrichment via hexa-His tag under completely denaturing conditions was used for mass spectrometric protein identification. The major contaminant, Rubisco large subunit, contributes as much as half of the protein mass of these preparations. Peptides with the expected mass of the tryptic Rubisco digest, therefore, were automatically discarded and not used for fragmentation. As a conservative estimate, we assume that the enriched fraction contains between 2% and 5% sumoylated proteins. This estimate is based on protein gels and gel blots as shown in Figure 4. The estimate is also consistent with the data collected by mass spectrometry, in which more than one in 10 peptides were unique to the sample fraction and not present in fractions prepared from nontransgenic plants.

Potential sumoylation substrates were identified based on the following criteria. First, the proteins copurified with in vivo expressed tagged versions of SUMO proteins but were absent from control preparations from nontransgenic plants. Second, at least three peptides derived from a putative substrate could be sequenced by LC-MS/MS, and none of these peptides was present in control preparations. Third, the putative substrate proteins enriched in the preparations migrated by SDS-PAGE at a molecular mass that was 20 to 40 kD higher than their predicted position, suggestive of covalent modification by SUMO.

Many high-scoring candidates contain an RNA-binding domain or are involved in DNA-dependent or chromatin-related processes. A common theme is that candidate substrates participate in assembly and disassembly processes. This is true for NAF, for components of the transcription machinery, for RNA-binding proteins with a putative role in splicing, and for components of the translation machinery. Incidentally, the latter proteins are the first candidates, to our knowledge, for cytoplasmic substrates of sumoylation from the model plant.

A subset of the proteins listed in Table I could be expressed in *E. coli* and was subjected to in vitro sumoylation, further supporting their status as candidate substrates. Only one of the *E. coli* expressed proteins from Table I, encoded by At2g43970, tested negative in the in vitro reaction. Many of the putative substrates contained consensus sumoylation sites, which were statistically enriched in the ensemble of candidate substrates. Whereas, for instance, 40% of the human proteome was reported to contain a sumoylation consensus motif (Zhou et al., 2005), 70% of the proteins listed in Table I contain such a motif, and more than 40% have more than one canonical sumoylation sequence.

A particularly interesting case was At5g08450 (labeled RXT3 in Fig. 7), a protein with similarity to yeast histone deacetylase component Rxt3. This protein contains more than 10 high-probability sumoylation sites and can be efficiently sumoylated in vitro. In contrast, At1g55300, the TAF7 subunit of general transcription factor TFIID (Thomas and Chiang, 2006), does not contain a single site, was identified in enriched preparations, and could nonetheless be sumoylated in vitro (TAF7 in Fig. 7). Interestingly, TAF7 contains several sequences that resemble SUMO interaction motifs (Hecker et al., 2006), so that association between the SUMO-SCE complex and TAF7 might rely on interaction with the SUMO moiety, not with SCE (which has affinity to exposed sumoylation consensus motifs).

While experiments regarding the Gln residue at position −4 imply a difference between SUMO isoforms 1, 3, and 5 in deconjugation, the question of whether different isoforms differ in substrate specificity (i.e. have distinct conjugation rates for certain substrates) remains open. Generally, in vitro sumoylation by SUMO1 was possible with most substrates tested (Table I), and in vitro conjugation by SUMO3 worked less efficiently even with substrates identified as in vivo SUMO3 substrates. Therefore, the in vitro system did not display strong differences in substrate selection by different isoforms. Similarly, our in vivo data had no bearing on this question, because substrate identification from enriched fractions was far below saturation. The absence of a particular substrate from preparations made with one isoform versus another, therefore, by no means implied that this substrate does not form in vivo conjugates with the other isoform.

In summary, this work presented an investigation of Arabidopsis SUMO isoforms and introduced a set of plant substrates of the SUMO conjugation system. Whereas many of these substrates operate in the nucleus, some are cytoplasmic proteins or occur in the nucleus and the cytoplasm (Table I). They can be subjects of further studies and may serve as standards in future in vitro and in vivo work. The in vitro sumoylation system used in this work is based entirely on Arabidopsis components and can be used to characterize additional components, such as new substrates, SUMO ligases, or SUMO proteases. Finally, the SUMO conjugate enrichment procedure described can be easily adapted to other ubiquitin-like protein modifiers.

MATERIALS AND METHODS

Plant Growth and Transformation

Arabidopsis (*Arabidopsis thaliana*) plants were grown in the greenhouse under long-day conditions (16 h of light, 8 h of darkness). Transgenes were introduced by floral dip transformation. Lines with high expression level were selected after western blotting of protein extracts. Vector pER8-based SUMO constructs were induced by spraying mature plants (grown on soil for 45 d) with water containing 5 μM β-estradiol. After overnight exposure, plants were harvested on the following day.

Plasmid Construction

Plasmid pSK-TagSUMO-GG contains the sequence of mature SUMO1 with N-terminal extension tag1 (5′-MA**HHHHHH**MGS**YPYDVPDYA**G**YPYDVPDYA**GS**YPYDVPDYA**SI-3′; HA tags are in boldface and underlined, and the hexa-His tag is in boldface). The extension sequence is preceded by an *Xho*I site followed by part of the Ω translational enhancer of *Tobacco mosaic virus*. The stop codon of the SUMO ORF is followed by a unique *Xba*I site. The construct was obtained by first amplifying a fragment of SUMO1 from mRNA using oligonucleotides and standard reverse transcription and PCR methods. The SUMO1 fragment ended with the *Bst*XI site close to the C terminus followed by a newly introduced *Xba*I site. Thereafter, sequences encoding His_6 and triple HA tags were inserted, and the missing portion of the carboxyl end was inserted as an oligonucleotide. For plant expression shown in Figure 2, an *Xho*I-*Xba* fragment was transferred into *Xho*I-*Xba*I-digested vector pHi (Schlögelhofer and Bachmair, 2002). Plasmid pSK-Tag3 contains the extension of all other SUMO constructs used in this work (see Fig. 1A for sequence). The ORF of the extension is preceded by an *Xho*I site, followed by part of the Ω translational enhancer of *Tobacco mosaic virus*. The extension ends with a *Kpn*I site followed by an *Xba*I site. SUMO3 and SUMO5 cDNA sequences were inserted as *Kpn*I-*Xba*I fragments. To obtain SUMO1 variants with changes at position 90, oligonucleotides spanning the *Cla*I-*Xba*I sites of a pSK-based SUMO1 construct were used for cassette exchange, and mutated fragments were transferred thereafter into pSK-Tag3. *Xho*I-*Xba*I fragments of Tag3-containing clones were inserted into *Xho*I-*Xba*-digested pHi or into an *Xho*I-*Spe*I-digested variant of pER8 that had a reduced number of enhancer sequences due to a promoter deletion spanning nucleotides 3,970 to 4,047 (pER8 accession no. is AF309825). For expression in *Escherichia coli*, *Not*I Klenow/*Nco*I fragments containing the coding regions from pSK-Tag3-SUMO1, pSK-Tag3-SUMO1 Q90A, pSK-Tag3-SUMO3, and pSK-Tag3-SUMO5 were inserted into *Bam*HI Klenow/*Nco*I-digested vector pET9d (Novagen).

For expression of SAE in *E. coli*, cDNAs encoding SAE2, SAE1a, and SAE1b were obtained from the RIKEN depository (pda10760, pda07771, and pda08247, respectively; http://www.brc.riken.jp/lab/epd/Eng/; Seki et al., 2002). Using oligonucleotides 5′-CCATGGTGTACAGGCCAGATCTGAGCCTGCTTCTAAGAAGAGAAGACT-3′ and 5′-GAGCTCATCTCCGTCCATGGCACCATGGTGATGATGGTGATGGGTCATTATTCAACTCTTATCTTCTT-3′, a C-terminal fragment of SAE2 was amplified that connected to SAE1a with sequence TAATG (TAA as stop codon of SAE2, ATG as start codon of SAE1a) and simultaneously extended the SAE1a reading frame by His_6 codons. The PCR fragment was inserted into *Ecl*136II-digested pSK. A *Sca*I-*Sac*I fragment from cDNA pda08247 was inserted into the *Spe*I (recessive ends filled with Klenow DNA polymerase)-*Sac*I-digested pSK-based vector. The latter vector was digested with *Bsr*GI and *Bgl*II, and a *Bsr*GI-*Bam*HI fragment from SAE2 cDNA was inserted. Vector pET9d was digested with *Nco*I and *Bam*HI, and an *Nco*I-*Bam*HI fragment from SAE2 cDNA was inserted. Thereafter, the pSK-based plasmid was digested with *Bam*HI and *Bsr*GI, and the insert was combined with the *Bam*HI-*Bsr*GI-digested pET-based vector. Finally, a spurious frameshift mutation in the SAE2 part was removed by inserting an *Afl*II-*Bsr*GI fragment amplified from cellular mRNA by reverse transcription and PCR. The clone was called pETSAE1b2. To obtain the analogous clone with isoform SAE1a (called pETSAE1a2), a *Pvu*II-*Sal*I fragment from pETSAE1b2 was inserted into *Ecl*136II-*Sal*I-digested vector pSK. The resulting clone was digested with *Nco*I and *Sal*I, and oligonucleotides 5′-CATGCGAGCATGGACGGAGAAGAGCCCGGGATCC-3′ and 5′-TCGAGGATCCCGGGCTCTTCTCCGTCCATGCTCG-3′ were annealed and inserted. The resulting plasmid was digested with *Sap*I and *Sma*I, and a *Sap*I-*Pme*I fragment from the SAE1a cDNA was inserted. Finally, a *Bsr*GI-*Bam*HI fragment from this construct was used to replace the *Bsr*GI-*Bam*HI fragment of pETSAE1b2, resulting in pETSAE1a2. For expression of SCE1 in *E. coli*, a cDNA fragment was amplified by reverse transcription and PCR from cellular mRNA. The fragment was flanked by *Nco*I (at the N terminus) and *Sma*I (after the stop codon) and cloned between *Nco*I and filled-in *Bam*HI sites of vector pET9d. A similar construct with a Cys-to-Ser (TGT-to-AGT) change (mutated clone kindly provided by Dr. C. Hardtke, University of Lausanne) was also prepared and called pETSCE C94S. A fragment of SUMO ligase SIZ1, spanning amino acid residues 370 to 873 of the protein, was inserted into vector pDEST17 for expression in *E. coli*.

Protein substrates for in vitro sumoylation assays were obtained in the following way. GST-S was expressed by induction of pET42c (Novagen). For UBC27Flag expression, UBC27 was amplified from cDNA pda06125 (RIKEN) using oligonucleotides 5′-CGGCCCGTCATATGATAGATTTCAGTCGAATC-3′ and 5′-GGCCTCGAGCTTGTCATCGTCGTCCTTGTAGTCGGTACCAGCAGAACAGAGCTTTTCC-3′. The resulting fragment was digested with *Nde*I and *Xho*I and inserted into *Nde*I-*Xho*I-digested vector pET42c (Novagen) to give pETUBC27FlagHis. pETGST-S-Flag, the vector for expression of GST-S-Flag, was constructed by inserting an *Nde*I-*Kpn*I fragment from pET42c into *Nde*I-*Kpn*I-digested vector pETUBC27FlagHis and replacing a *Swa*I-*Spe*I fragment of the resulting construct by a *Swa*I-*Spe*I fragment that was generated using oligonucleotides 5′-CCAGAACCACTAGTTGAACCATCCGAGCGTGGAGGAT-3′ and 5′-GCTGAAAATGTTCGAAGATCGTTT-3′ and pET42c template. For expression of GST-Flag, vector pETGST-S-Flag was digested with *Mfe*I and *Bgl*II, treated with Klenow fragment of *E. coli* DNA polymerase, and religated to give pET42c-FlagmutΔS. For expression of NAF (ORF of At2g19480 extended by a Flag tag), oligonucleotides 5′-CGGCGTACGCATATGAGCAACGACAAGGACAGCAT-3′ and 5′-CGGCCCCGGGAATTTTGCTTCCATCTCATC-3′ were used to amplify a fragment from cDNA pda08885 (RIKEN). After digestion with *Nde*I, the fragment was inserted into *Nde*I-*Eco*RV-digested vector pETUBC27FlagHis. Another fragment was amplified with oligonucleotides 5′-CCCGCCCGGGAGGACGACGATGATGAGATT-3′ and 5′-CGCCGGTACCCTGCTGCTTACATTCCGGT-3′ from cDNA pda08885, digested with *Sma*I and *Kpn*I, and inserted into the latter construct after digestion with *Sma*I-*Kpn*I. Finally, this vector was digested with *Afl*II and *Bsp*MI, and an *Afl*II-*Bsp*MI fragment from cDNA pda08885 was inserted to give vector pETNAF. The protein called RRM in Figure 7 (ORF of At3g56860 extended by a Flag tag) was expressed from vector pETRRM1, which was constructed similar to pETNAF, using cDNA pda08474 (RIKEN) as a starting clone, oligonucleotides 5′-CGGCGTACGCATATGACAAAGAAGAGAAAGCT-3′ and 5′-CGGCCCCGGGTCTTCTTGATTGCCAGATC-3′ to amplify a 5′ portion of the ORF, oligonucleotides 5′-CCCGCCCGGGCT GGTTACGGTACTCAAGCT-3′ and 5′-CGCCGGTACCGTGACCCATGTAAGGAGTAC-3′ to amplify a 3′ portion, and restriction enzymes *Aat*II and *Sex*AI to insert the central fragment from pda08474. To express the ORF of At5g08450 (RXT3) with the N-terminal GST tag (without the S peptide), cDNA pda08750 was used as template together with oligonucleotides 5′-CCGCGCCAATTGGTCATATGAGTGGTGTTCCAAAGAGAT-3′ and 5′-CCGCGCGGTACCTTAACCACCGTTCTAGATT-3′, and the resulting PCR fragment was digested with *Mfe*I and *Kpn*I and inserted into *Mfe*I-*Kpn*I-digested vector pET42c. After digestion of the latter vector with *Bam*HI and partial digestion with *Xba*I, a *Bam*HI-*Xba*I fragment from cDNA pda08750 was inserted to give vector pETRXT3. For expression of TAF7 as a GST fusion, an *Nco*I-*Not*I fragment from pUNI clone U63389 (Yamada et al., 2003) was inserted between *Nco*I and *Not*I sites of vector pET42c. Thereafter, the resulting vector was digested with *Bgl*II and *Mfe*I, treated with Klenow fragment, and religated to obtain pETTAFΔS.

Plant Extraction

SUMO substrates for mass spectrometric analysis were obtained from induced plants by immersion of typically 200 g of plant material in 5 volumes of buffer A (6 M guanidine hydrochloride, 0.1 M Tris-Cl, and 0.1 M Na phosphate buffer, pH 8, supplemented with 20 mM 2-mercaptoethanol, 10 mM Na metabisulfite, 3% polyvinylpyrrolidone, 5% [w/v] Suc, and 5 mM imidazole). The material was homogenized with a blender (Polytron PT 2100; Kinematica). The slurry was stirred for 30 min and then centrifuged at 3,000*g* for 30 min. The supernatant was filtered through Miracloth, centrifuged again (30,000*g*, 60 min), and incubated overnight at room temperature (constant agitation) with 4 mL of Ni-NTA resin (Qiagen). The resin was filled into a disposable column (Poly-Prep; Bio-Rad) and washed with 5 to 10 volumes of buffer B (8 M urea, 0.1 M Na phosphate, and 0.1 M Tris-Cl, pH 8). Protein was eluted with 1 to 2 column volumes of buffer C (8 M urea and 0.2 M acetic acid). The eluate was neutralized with 1 M Tris-Cl, pH 8, and concentrated by ultrafiltration (Centriprep and Centricon tubes, 10-kD cutoff; Millipore). To remove background, the eluate of Ni-NTA resin was in some cases reapplied onto a fresh batch of Ni-NTA resin, followed by elution as described above. For the experiment shown in Figure 4D, the Ni-NTA eluate was diluted to 0.3 M urea. Triton X-100 (1% final concentration), SDS (0.1% final concentration), protease inhibitor cocktail (Roche), and 2-mercaptoethanol (1 mM) were added, and the material was applied to anti-HA resin (Roche). Bound proteins were eluted using 8 M urea in Gly buffer (100 mM, pH 2.2), followed by neutralization of the eluate.

Mass Spectrometric Analysis

In-Gel Digestion

Proteins were separated on standard SDS-polyacrylamide gels. Coomassie Brilliant Blue-stained bands were excised from the gel and treated as described (Shevchenko et al., 1996), except that no $CaCl_2$ was added during digestion with trypsin (sequencing grade; Promega) and final extraction of peptides was carried out with 100 μL of 1% trifluoroacetic acid (TFA) for 30 min at 37°C. The extraction was repeated once with 100 μL of 0.1% TFA: acetonitrile (1:2). The volume of the combined supernatants was reduced to 5 μL in a vacuum centrifuge, and 20 μL of 0.1% TFA was added to each sample.

LC-MS/MS of In-Gel-Digested Proteins

LC-MS data were acquired on a quadrupole-time of flight mass spectrometer (Q-Tof II; Micromass) equipped with a Z spray source. Samples were introduced by the Ultimate nano-LC system (LC Packings) equipped with a Famos autosampler and a Switchos column switching module. The column setup comprises a 0.3-mm × 1-mm trapping column and a 0.075-mm × 150-mm analytical column, both packed with 3 μm of Atlantis dC18 (Waters). A total of 10 μL was injected onto the trap column and desalted for 1 min with 0.1% TFA at a flow rate of 10 μL min^{-1}. Peptides were eluted onto the analytical column by a gradient of 2% acetonitrile in 0.1% formic acid to 40% acetonitrile in 0.1% formic acid over 55 min at a column flow rate of approximately 200 nL min^{-1}, resulting from a 1:1,000 split of the 200 μL min^{-1} flow delivered by the pump. The electrospray ionization interface comprised an uncoated 10-μm i.d. PicoTip spray emitter (New Objective) linked to the HPLC flow path using a 7-μL dead volume stainless steel fitting mounted onto the PicoTip holder assembly (New Objective). Stable nanospray was established by the application of 1.7 to 2.4 kV to the stainless steel union. The data-dependent acquisition of MS and MS/MS spectra was controlled by the Masslynx 4.0. Survey scans of 1 s covered the range from mass-to-charge ratio 360 to 1,200. Doubly and triply charged ions rising above a given threshold were selected for MS/MS experiments. In MS/MS mode, the mass range from mass-to-charge ratio 50 to 1,200 was scanned in 1 s, and three scans were added up for each experiment. Micromass-formatted peak lists were generated from the raw data using the Proteinlynx software module. Proteins were identified by searching the NCBI nr public database (National Center for Biotechnology Information) using a local installation of MASCOT 1.9 (Matrix Science). A mass deviation of 0.5 D was allowed for peptide and fragment ions.

Protein Expression

Proteins were expressed in *E. coli* strain Rosetta (DE3) pLysS (Novagen) or in strain BL21 (AtSIZ1 fragment). Purification via Ni-NTA resin (Qiagen) or gluthatione-agarose (Novagen) followed the resin manufacturer's recommendations. If necessary, buffer changes and concentration of protein samples were carried out by centrifugation (Vivaspin 500; Sartorius). SCE1 was purified from pelleted, induced *E. coli* cells. After addition of one-thirtieth volume of lysis buffer (50 mM Na phosphate, pH 6.5, 50 mM NaCl, 1 mM dithiothreitol, and Complete protease inhibitor mix [Roche]), the cells were subjected to freeze-thaw cycles and centrifuged (1 h at 100,000*g*, 4°C). The supernatant was applied to SP Sepharose (GE Healthcare) preequilibrated with lysis buffer. After washing (3 bed volumes of lysis buffer), batch elution was carried out with lysis buffer plus 300 mM NaCl. Active fractions were subjected to buffer change by ultrafiltration (20 mM HEPES/KOH, pH 7.3, 10 mM potassium acetate, 2 mM Mg acetate, 0.5 mM EGTA, 1 mM dithiothreitol, and protease inhibitors aprotinin, leupeptin and pepstatin [1 μg mL^{-1} each]).

In Vitro Sumoylation

Typical sumoylation reactions were carried out for 4 h or overnight at 30°C in 50 μL of reaction buffer containing 5 mM ATP, 5 mM Mg^{2+}, 20 mM Tris, pH 7.5, 100 μg of tag-SUMO, 4 μg of SAE, 0.6 μg of SCE, and 0.15 μg of SIZ1 protein fragment. Between 1 and 10 μg of substrate was added. Aliquots of 10 μL were separated by SDS-PAGE and processed by western blotting to detect substrate-specific bands using either Flag antibody-alkaline phosphate conjugate (Sigma) or anti-GST antibody followed by alkaline phosphatase-coupled secondary antibody.

In Vitro Desumoylation

An in vitro sumoylation reaction with substrate NAF was incubated with anti-Flag resin (Sigma) in buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 1 mM EDTA, and 1% Triton X-100) and washed twice with Tris-buffered saline and finally with sumoylation assay buffer. *E. coli*-expressed ESD4 fragment was purified using Ni^{2+} affinity, and the eluate in buffer containing 50 mM NaH_2PO_4, 300 mM NaCl, and 250 mM imidazole (pH 8.0) was added to the resin-bound substrate. Aliquots were withdrawn, and the reaction was terminated by addition of *N*-ethylmaleimide to a final concentration of 2 mM at the indicated times. Sample buffer was added, and aliquots were applied for gel electrophoresis and western blotting.

Protein Detection

Proteins were separated on laboratory-made minigels, except for those in Figure 2, lanes 8 and 9, for which a precast gradient gel (3%–12%; Invitrogen) was used, and detected by western blotting essentially as described (Stary et al., 2003) using Immobilon (Millipore) membrane for protein support, followed by detection with alkaline phosphatase-coupled secondary antibody (Sigma). Primary antibodies were directed against protein tags HA (Roche), Flag (Sigma), and GST (Sigma) or against SUMO (Murtas et al., 2003; Abcam). For experiments shown in Figure 3, proteins were transferred to nitrocellulose membrane and detected using anti-Flag primary antibody, followed by either horseradish peroxidase-coupled secondary antibody (Amersham ECL, GE Healthcare; Fig. 3A) or IR Dye 800-coupled secondary antibody (Rockland; detection by the Odyssey Infrared Imager [Li-Cor]; Fig. 3B) as described (Garzón et al., 2007). Protein concentration in plant extracts was assessed by gel electrophoresis of aliquots, followed by Coomassie Brilliant Blue staining. Adjusted volumes with equal protein content were applied to the blotting gels.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Extended alignment of C termini of mature SUMO proteins.

Supplemental Figure S2. Phenotypes of plants expressing SUMO constructs.

ACKNOWLEDGMENTS

We thank K. Luxa, M. Lehnen, A. Krull, and A. Harzen for expert technical help, Dr. Y.-F. Fu for SAE2 cDNA and the SIZ1 expression clone, Dr. C. Hardtke for the SCE1 C94S clone, Dr. F. Melchior for advice on SCE and SAE purification, and Dr. N. Elrouby for reading of the manuscript.

Received December 29, 2008; accepted January 10, 2009; published January 16, 2009.

LITERATURE CITED

Budhiraja R (2005) Post-translational modification of proteins by SUMO in *Arabidopsis thaliana*. PhD thesis. University of Cologne, Cologne, Germany

Catala R, Ouyang J, Abreu IA, Hu Y, Seo H, Zhang X, Chua NH (2007) The *Arabidopsis* E3 SUMO ligase SIZ1 regulates plant growth and drought responses. Plant Cell **19:** 2952–2966

Chosed R, Mukherjee S, Lois LM, Orth K (2006) Evolution of a signalling system that incorporates both redundancy and diversity: Arabidopsis sumoylation. Biochem J **398:** 521–529

Chosed R, Tomchick DR, Brautigam CA, Mukherjee S, Negi VS, Machius M, Orth K (2007) Structural analysis of *Xanthomonas* XopD provides insights into substrate specificity of ubiquitin-like protein proteases. J Biol Chem **282:** 6773–6782

Colby T, Matthäi A, Boeckelmann A, Stuible HP (2006) SUMO-conjugating and SUMO-deconjugating enzymes from Arabidopsis. Plant Physiol **142:** 318–332

Conti L, Price G, O'Donnell E, Schwessinger B, Dominy P, Sadanandom A (2008) Small ubiquitin-like modifier proteases OVERLY TOLERANT TO SALT1 and -2 regulate salt stress responses in *Arabidopsis*. Plant Cell **20:** 2894–2908

Dohmen RJ (2004) SUMO protein modification. Biochim Biophys Acta **1695:** 113–131

Dye BT, Schulman B (2007) Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins. Annu Rev Biophys Biomol Struct **36:** 131–150

Garcia-Dominguez M, March-Diaz R, Reyes JC (2008) The PHD domain of plant PIAS proteins mediates sumoylation of bromodomain GTE proteins. J Biol Chem **283:** 21469–21477

Garzón M, Eifler K, Faust A, Scheel H, Hofmann K, Koncz C, Yephremov A, Bachmair A (2007) PRT6/At5g02310 encodes an Arabidopsis ubiquitin ligase of the N-end rule pathway with arginine specificity and is not the CER3 locus. FEBS Lett **581:** 3189–3196

Geiss-Friedlander R, Melchior F (2007) Concepts in sumoylation: a decade on. Nat Rev Mol Cell Biol **8:** 947–956

Hannich JT, Lewis A, Kroetz MB, Li SJ, Heide H, Emili A, Hochstrasser M (2005) Defining the SUMO-modified proteome by multiple approaches in *Saccharomyces cerevisiae*. J Biol Chem **280:** 4102–4110

Hay RT (2007) SUMO-specific proteases: a twist in the tail. Trends Cell Biol **17:** 370–376

Hecker CM, Rabiller M, Haglund K, Bayer P, Dikic I (2006) Specification of SUMO1- and SUMO2-interacting motifs. J Biol Chem **281:** 16117–16127

Hotson A, Chosed R, Shu H, Orth K, Mudgett MB (2003) Xanthomonas type III effector XopD targets SUMO-conjugated proteins in planta. Mol Microbiol **50:** 377–389

Jin BJ, Jin YH, Lee J, Miura K, Yoo CY, Kim WY, Van Oosten M, Hyun Y, Somers DE, Lee I, et al (2008) The SUMO E3 ligase, AtSIZ1, regulates flowering by controlling a salicylic acid-mediated floral promotion pathway and through effects on FLC chromatin structure. Plant J **53:** 530–540

Johnson ES (2004) Protein modification by SUMO. Annu Rev Biochem **73:** 355–382

Kerscher O, Felberbaum R, Hochstrasser M (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins. Annu Rev Cell Dev Biol **22:** 159–180

Kurepa J, Walker JM, Smalle J, Gosink MM, Davis SJ, Durham TL, Sung DY, Vierstra RD (2003) The small ubiquitin-like modifier (SUMO) protein modification system in Arabidopsis. J Biol Chem **278:** 6862–6872

Lee J, Nam J, Park HC, Na G, Miura K, Jin JB, Yoo CY, Baek D, Kim DH, Jeong JC, et al (2007) Salicylic acid-mediated innate immunity in Arabidopsis is regulated by SIZ1 SUMO E3 ligase. Plant J **49:** 79–90

Li T, Evdokimov E, Shen RF, Chao CC, Tekle E, Wang T, Stadtman ER, Yang DCH, Chock PB (2004) Sumoylation of heterogeneous nuclear ribonucleoproteins, zinc finger proteins, and nuclear pore complex proteins: a proteomic analysis. Proc Natl Acad Sci USA **101:** 8551–8556

Lois LM, Lima CD, Chua NH (2003) Small ubiquitin-like modifier modulates abscisic acid signaling in *Arabidopsis*. Plant Cell **15:** 1–13

Miura K, Jin JB, Hasegawa P (2007a) Sumoylation, a post-translational regulatory process in plants. Curr Opin Plant Biol **10:** 1–8

Miura K, Jin JB, Lee J, Yoo CY, Stirm V, Miura T, Ashworth EN, Bressan RA, Yun DJ, Hasegawa PM (2007b) SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A expression and freezing tolerance in *Arabidopsis*. Plant Cell **19:** 1403–1414

Miura K, Rus A, Sharkhuu A, Yokoi S, Karthikeyan AS, Raghothama KG, Baek D, Koo YD, Jin JB, Bressan RA, et al (2005) The *Arabidopsis* SUMO E3 ligase *SIZ1* controls phosphate deficiency responses. Proc Natl Acad Sci USA **102:** 7760–7765

Mossessova E, Lima CD (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. Mol Cell **5:** 865–876

Mukhopadhyay D, Dasso M (2007) Modification in reverse: the SUMO proteases. Trends Biochem Sci **32:** 286–295

Murtas G, Reeves PH, Fu YF, Bancroft I, Dean C, Coupland G (2003) A nuclear protease required for flowering-time regulation in *Arabidopsis* reduces the abundance of small ubiquitin-related modifier conjugates. Plant Cell **15:** 2308–2319

Novatchkova M, Budhiraja R, Coupland G, Eisenhaber F, Bachmair A (2004) SUMO conjugation in plants. Planta **220:** 1–8

Owerbach D, McKay EM, Yeh ETH, Gabbay KH, Bohren KM (2005) A proline-90 residue unique to SUMO-4 prevents maturation and sumoylation. Biochem Biophys Res Commun **337:** 517–520

Panse VG, Hardeland U, Werner T, Kuster B, Hurt E (2004) A proteome-wide approach identifies sumoylated substrate proteins in yeast. J Biol Chem **279:** 41346–41351

Roden J, Eardley L, Hotson A, Cao Y, Mudgett MB (2004) Characterization of the *Xanthomonas* AvrXv4 effector, a SUMO protease translocated into plant cells. Mol Plant Microbe Interact **17:** 633–643

Saracco SA, Miller MJ, Kurepa J, Vierstra RD (2007) Genetic analysis of sumoylation in Arabidopsis: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. Plant Physiol **145:** 119–134

Schlögelhofer P, Bachmair A (2002) A test of fusion protein stability in the plant, Arabidopsis thaliana, reveals degradation signals from ACC synthase and from the plant N-end rule pathway. Plant Cell Rep **21:** 174–179

Seki M, Narusaka M, Kamiya A, Ishida J, Satou M, Sakurai T, Nakajima M, Enju A, Akiyama K, Oono Y, et al (2002) Functional annotation of a full-length *Arabidopsis* cDNA collection. Science **296:** 141–145

Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal Chem **68:** 850–858

Stary S, Yin XJ, Potuschak T, Schlögelhofer P, Nizhynska V, Bachmair A (2003) PRT1 of Arabidopsis is a ubiquitin protein ligase of the plant N-end rule pathway with specificity for aromatic amino-terminal residues. Plant Physiol **133:** 1360–1366

Tatham MH, Geoffroy MC, Shen L, Plechanovova A, Hattersley N, Jaffray EG, Palvimo JJ, Hay RT (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. Nat Cell Biol **10:** 538–546

Thomas MC, Chiang CM (2006) The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol **41:** 105–178

Uzunova K, Göttsche K, Miteva M, Weisshaar SR, Glanemann C, Schnellhardt M, Niessen M, Scheel H, Hofmann K, Johnson ES, et al (2007) Ubiquitin-dependent proteolytic control of SUMO conjugates. J Biol Chem **282:** 34167–34175

Wohlschlegel JA, Johnson E, Reed SI, Yates III Jr (2004) Global analysis of protein sumoylation in *Saccharomyces cerevisiae*. J Biol Chem **279:** 45662–45668

Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, et al (2003) Empirical analysis of transcriptional activity in the Arabidopsis genome. Science **302:** 842–846

Yates JL, Nomura M (1980) E. coli ribosomal protein L4 is a feedback regulatory protein. Cell **21:** 517–522

Yoo CY, Miura K, Jin JB, Lee J, Park HC, Salt DE, Yun DJ, Bressan RA, Hasegawa PM (2006) SIZ1 small ubiquitin-like modifier E3 ligase facilitates basal thermotolerance in Arabidopsis independent of salicylic acid. Plant Physiol **142:** 1548–1558

Zhao Y, Kwon SW, Anselmo A, Kaur K, White MA (2004) Broad spectrum identification of cellular small ubiquitin-related modifier (SUMO) substrate proteins. J Biol Chem **279:** 20999–21002

Zhou F, Xue Y, Lu H, Chen G, Yao X (2005) A genome-wide analysis of sumoylation-related biological processes and functions in human nucleus. FEBS Lett **579:** 3369–3375

Zhou W, Ryan JJ, Zhou H (2004) Global analyses of sumoylated proteins in *Saccharomyces cerevisiae*. J Biol Chem **279:** 32262–32268

Zuo J, Niu QW, Chua NH (2000) An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. Plant J **24:** 265–273