Evolution of a signalling system that incorporates both redundancy and diversity: Arabidopsis SUMOylation

Renee CHOSED*, Sohini MUKHERJEE*, Luisa Maria LOIS† and Kim ORTH*†

*Department of Molecular Biology, University of Texas Southwestern Medical Center (UTSWMC), 5323 Harry Hines Blvd, Dallas, TX 75390, U.S.A., and †Departamento de Bioquímica i Biologia Molecular, Universitat de Barcelona, Av. Diagonal 645, Barcelona 08028, Spain

The reversible post-translational modifier, SUMO (small ubiquitin-related modifier), modulates the activity of a diverse set of target proteins, resulting in important consequences to the cellular machinery. Conjugation machinery charges the processed SUMO so that it can be linked via an isopeptide bond to a target protein. The removal of SUMO moieties from conjugated proteins by isopeptidases regenerates pools of processed SUMOs and unmodified target proteins. The evolutionarily conserved SUMO-conjugating proteins, E1 and E2, recognize a diverse set of Arabidopsis SUMO proteins using them to modify protein substrates. In contrast, the deSUMOylating enzymes differentially recognize the Arabidopsis SUMO proteins, resulting in specificity of the deconjugating machinery. The specificity of the Arabidopsis deSUMOylating enzymes is further diversified by the addition of regulatory domains. Therefore the SUMO proteins, in this signalling system, have evolved to contain information that allows not only redundancy with the conjugation system but also diversity with the deconjugating enzymes.

Key words: Arabidopsis thaliana, reversible post-translational modification, signalling system, small ubiquitin-related modifier (SUMO), SUMOylation, ubiquitin-like protein protease-1 (ULP1).

INTRODUCTION

SUMO (small ubiquitin-related modifier), like ubiquitin, is an evolutionarily conserved, reversible post-translational modifier that is covalently attached to cellular proteins via an isopeptide bond. One of the first reports on SUMO demonstrates that RanGAP [Ran GAP (GTPase-activating protein)] is modified by M-SUMO-1 (mammalian SUMO-1) [1]. Around the same time, another report implicated SUMO (referred to as sentrin) as a molecule that modified responses triggered by cell-death receptors [2]. Later, the modification by M-SUMO-1 was shown to be critical for the localization of RanGAP-SUMO-1 to the nuclear pore complex [3]. Since its original discovery, SUMO has been cloned from an array of eukaryotic organisms, including animals, fungi and plants. In humans, three additional SUMOs (M-SUMO-2, -3 and -4) were identified that are implicated in a nuclear pore complex [3]. Since its original discovery, SUMO has been cloned from an array of eukaryotic organisms, including animals, fungi and plants.

SUMOylation machinery makes use of one universal E2-conjugating enzyme, Ubc9 (ubiquitin conjugating enzyme 9), which is able to transfer SUMO directly to target proteins. In many cases, an E3 ligase is utilized to facilitate the transfer of the modifier protein to the target protein. Secondly, for SUMO, a weak consensus modification motif in target proteins has been identified as ΨKXE, where Ψ is a large hydrophobic residue, K is the lysine that SUMO is attached to, X is any residue and E is a glutamic acid residue [10]. Ubiquitin can modify a protein with either a poly-ubiquitin chain or with mono-ubiquitin [11]. Modifications by SUMO are most commonly associated with mono-SUMOylation, although previous studies have implicated poly-SUMOylation [12]. While both ubiquitin and SUMO are processed to a mature form by exposing two glycine residues at their C-terminus, a distinct family of cysteine proteases, referred to as ULP1s (ubiquitin-like protein protease-1), process the C-terminus of SUMO [4,13].

ULP1s, like SUMOs, belong to a growing family of proteins that is evolutionarily conserved from yeast to human. Seven ULP1 proteases have been found in humans and are referred to as SENP1 (sentrin-specific protease 1), SENP2, SENP3, SENP5, SENP6, SENP7 and SENP8 [13]. Ulp1 from yeast plays a dual role, acting as a peptidase by cleaving the FL (full-length) SUMO and as an isopeptidase by cleaving SUMO-conjugated target proteins [7]. Kurepa and co-workers [8] have identified a family of four ULP1 genes (AtULP1A–AtULP1D) encoded in the Arabidopsis genome and Murtas and co-workers [14] have identified a fifth ULP1 gene [AtESD4 (where ESD4 is early in short days 4)].

Another ULP1, XopD (Xanthomonas outer protein D), from Xanthomonas campestris pv. vesicatoria, is expressed as a virulence factor by this plant bacterial pathogen [15]. As bacteria do not encode either the ubiquitin or the SUMO signalling machineries, Xanthomonas appears to have usurped the activity of eukaryotic ULP1s and uses this activity to aid in pathogenesis by disrupting...

Abbreviations used: Aos1, activation of Smt3p; At, Arabidopsis thaliana; ESD4, early in short days 4; FL, full-length; GST, glutathione S-transferase; HA, haemagglutinin; GAP, GTPase-activating protein; RanGAP; Ran GAP; RRL, rabbit reticulocyte lysate; SENP, sentrin-specific protease; SMT, suppressor of mif two; SUMO, small ubiquitin-related modifier; M-SUMO, mammalian SUMO; STOP, Stop codon; T-SUMO, tomato SUMO; TEV, tobacco etch virus; UbA2, ubiquitin activating protein 2; UbC9, ubiquitin conjugating enzyme 9; ULP1, ubiquitin-like protein protease-1; XopD, Xanthomonas outer protein D.

1 To whom correspondence should be addressed (email Kim.Orth@utsouthwestern.edu).
host defence signalling in the infected plant cell [15]. These and the aforementioned studies support the importance of SUMO modification in plant signalling [8,9,14,15].

The importance of SUMO modification of target proteins has been established in many cellular processes [13,16]. However, the mechanisms that dictate the specificity, frequency and half-life of this modification are poorly understood. The basic machinery involved in the conjugation of SUMO (E1 and E2) is promiscuous, in that a diverse pool of SUMO molecules are used as substrates. The localization and diversity of the SUMO E3s also play a major role in regulating the conjugation. Three different types of SUMO E3 ligases have been identified that include the PIAS [protein inhibitor of activated STAT (signal transducer and activator of transcription)] family, RanBP2 (Ran binding protein 2) and Pcm2 (Polycomb 2) [13]. As a consequence of subcellular localization, spatial restrictions on the conjugation and the deconjugation machineries are thought to contribute to the specificity of both the addition and removal of particular SUMO moieties. In vitro studies support the notion that the deSUMOylating enzymes are promiscuous and that their localization dictates their specificity [4,13]. In one case, localization of human SENP2 appears to be regulated by differential splicing of its mRNA [17].

To further our understanding of mechanisms utilized for the reversible post-translational modification by SUMO, we have investigated whether eukaryotes (plants and animals) encode such a large number of ULP1s and SUMOs in order to diversify this signalling system or to simply add redundancy to it. Our studies reveal that both of these mechanisms have been incorporated into the Arabidopsis SUMOylation signalling machinery. We have discovered that a great range of specificity for the different SUMO substrates is inherent in the catalytic core of the Arabidopsis ULP1s and that this system is further diversified by the addition of regulatory domains associated with the ULP1s. In contrast, we have observed that the evolutionarily conserved E1 and E2 in the conjugation machinery do not discriminate among the various Arabidopsis SUMO proteins and are able to use all Arabidopsis SUMOs in conjugation reactions. Therefore inherent in the Arabidopsis SUMO proteins is the information that dictates the specificity of their hydrolysis by isopeptidases so that they can be uniformly recognized by the conjugation machinery.

**EXPERIMENTAL**

**Cloning and construction of plasmids**

NCBI Protein Database accession numbers for sequences used in the present study are: AtSUMO-1, At4g26840; AtSUMO-2, At5g55160; AtSUMO-3, At5g55170; AtSUMO-5, At2g32765; AtULP1A, At3g06910; AtULP1C, At1g10570; AtULP1D, At1g60220; and AtESD4, At4g15880.

AtSUMOs (AtSUMO-1, -2, -3 and -5), AtULP1A, AtULP1C, AtULP1D and AtESD4 were cloned from cDNA. Yeast GST–ΔUlp1 (where GST is glutathione S-transferase), HA (haemagglutinin)–RanGAP and T-SUMO (tomato SUMO) constructs were made as described by Hotson et al. [15]. SMT3 (suppressor of mif two 3) was cloned from Saccharomyces cerevisiae cDNA. M-SUMO-1 and -2 were cloned from mammalian cDNA. M-SUMO-4 was a gift from Cong-Yi Wang (Center for Biotechnology and Genomic Medicine, Medical College of Georgia, Augusta, GA, U.S.A.) [6].

For constructing GST–SUMO-Gly-Gly–STOP (where STOP is Stop codon) constructs to be used with in vitro SUMOylation assays, AtSUMO (AtSUMO-1, -2, -3 and -5), T-SUMO, SMT3 and M-SUMO (M-SUMO-1, -2 and -4) were cloned into pGEXr-TEV (where TEV is tobacco etch virus) [18]. These SUMOs were constructed with an N-terminal GST tag and their C-terminus contained a STOP after the Gly-Gly motif. For constructing HisN–SUMO-Gly-Gly–HA constructs to be used with in vitro peptidase assays, AtSUMO (AtSUMO-1, -2 and -3), T-SUMO, SMT3 and M-SUMO (M-SUMO-1, -2 and -4) were cloned into pET15b (Novagen) and pTEF15b (Novagen) and AtSUMO-5 was cloned into pT7-LOH (a gift from J. Clemens, Purdue University, West Lafayette, IN, U.S.A.). These SUMOs were constructed with an N-terminal His tag and their C-terminus contained an HA tag directly following the Gly-Gly motif. HisN–AtSUMO-3-Gly-Gly-X18 and HisN–AtSUMO-5-Gly-Gly-X18, for use in in vitro peptidase assays, were constructed with an N-terminal His tag and their C-terminus contained their native C-terminal extension following the Gly-Gly motif. For constructing GST–ULP1 constructs to be used with in vitro peptidase and isopeptidase assays, AtULP1A, AtULP1C and AtULP1D and ESD4 were cloned into pGEXr-TEV. These ULP1s were constructed with an N-terminal GST tag used for protein purification purposes. The sequences of all DNA constructs were verified by cycle sequencing.

For cloning details, see Supplementary Table S1 at http://www.BiochemJ.org/bj/398/bj3980521add.htm.

**Protein expression and purification**

GST–ΔUlp1 was purified as described by Hotson et al. [15]. GST–AtSUMO-1, GST–AtSUMO-2, GST–AtSUMO-3, GST–AtSUMO-5, GST–T-SUMO, GST–Smt3, GST–M-SUMO-1, GST–M-SUMO-2, GST–M-SUMO-4 and all GST-tagged AtULP1 family members were expressed in Escherichia coli BL21/DE3 cells and then purified by standard GST affinity chromatography [19]. Briefly, cells were grown to exponential phase in 2YT medium [1.6 % (w/v) tryptone, 1 % (w/v) yeast extract and 0.5 % (w/v) NaCl] and then induced with 0.5 µM isopropyl β-D-thiogalactoside (Roche) for 4 h at 30°C. The cells were lysed in PBS (pH 8), 1 % Triton X-100 (Fisher), 0.1 % 2-mercaptoethanol (Bio-Rad) and 1 mM PMSF (Sigma) using a cell disruptor (Emulsiflex C5; Avestin). The protein was bound to glutathione–agarose beads and then eluted with 10 mM GSH as previously described [19]. Overnight cleavage at room temperature (25°C) with HisN–TEV protease was used to remove the GST tag from GST–AtSUMO-3 and GST–AtSUMO-5 in SUMOylation assays where noted. Purified proteins were analysed by SDS/PAGE and quantified using a modification of the Lowry procedure [20].

**In vitro peptidase assays**

In vitro SUMOylation assays

In vitro SUMOylation of 35S-labelled mammalian HA–RanGAP were in vitro translated in the TNT (transcription and translation) coupled RRL (rabbit reticulocyte lysate) system (Promega) with L-[35S]methionine (Amersham). For each in vitro peptidase assay, 2 µl of the 35S-labelled translation reaction mixture was added to 18 µl of either glutathione elution buffer (10 mM GSH, 50 mMTris, pH 8.0, 150 mMNaCl and 0.1 % 2-mercaptoethanol) without glutathione or 0.5 mg/ml of purified enzyme for 1 h at 30°C. The samples were then resolved by SDS/PAGE and then the gels were incubated for 10 min at 25°C with Amplify fluorographic reagent (Amersham) and analysed by autoradiography.
Figure 1  SUMO family members can be utilized by SUMO conjugation machinery in vitro

(A) Sequence alignment of amino acids of yeast Smt3, T-SUMO, AtSUMO-1, AtSUMO-2, AtSUMO-3, and AtSUMO-5. Numbering is shown with respect to yeast Smt3. Black circles above the alignment denote residues in yeast Smt3 that are in direct contact with yeast ΔUlp1 as described by Mossessova and Lima [23]. The arrow indicates where SUMO substrates are cleaved by ULP1s. (B) [35S]Mammalian RanGAP was in vitro translated in an RRL. In the RRL, some of the RanGAP is SUMOylated by endogenous SUMOylation machinery (lane 1). The in vitro translated product was then used in an in vitro SUMOylation assay using purified recombinant GST–AtSUMO-1, GST–AtSUMO-2, GST–AtSUMO-3, GST–AtSUMO-5, GST–T-SUMO, GST–Smt3, GST–M-SUMO-1, GST–M-SUMO-2 and GST–M-SUMO-4 to produce GST–SUMO-modified RanGAP. A volume of 5 µl of the SUMO-modified RanGAP from each reaction was then added to 5× SDS sample buffer [250 mM Tris, pH 6.8, 50% glycerol, 5% (w/v) SDS, 5% 2-mercaptoethanol and Bromophenol Blue dye] and the samples were resolved on SDS/8% polyacrylamide gels and visualized by autoradiography.

RESULTS AND DISCUSSION

Arabidopsis SUMOs are conjugated to RanGAP by evolutionarily conserved E1 and E2

The SUMO substrates used in the present study are aligned, including the four AtSUMOs (AtSUMO-1, AtSUMO-2, AtSUMO-3 and AtSUMO-5), yeast Smt3, T-SUMO, M-SUMO-1, M-SUMO-2 and M-SUMO-4, in Figure 1A. Although mining of the Arabidopsis genome revealed eight SUMOs, only four are encoded by mRNA [8]. AtSUMO-1 and AtSUMO-2 share 89% sequence identity with each other and 83% with T-SUMO. For all of these proteins, the two conserved glycine residues in the C-terminus mark the end of the processed form of the SUMOs.

To assess whether the various SUMOs can be used as substrates by the conjugation machinery, we constructed and purified recombinant N-terminally tagged GST–fusion proteins with the processed form of the Arabidopsis SUMOs, M-SUMOs, T-SUMO and yeast Smt3 (GST–SUMO-Gly-Gly–STOP). We utilized RanGAP that was in vitro transcribed and translated in an RRL as a target protein for the SUMOylation. As observed in Figure 1B (lane 1), a fraction of the translated [35S]-labelled RanGAP from each reaction was then added to 5× SDS sample buffer [250 mM Tris, pH 6.8, 50% glycerol, 5% (w/v) SDS, 5% 2-mercaptoethanol and Bromophenol Blue dye] and the samples were resolved on SDS/8% polyacrylamide gels and visualized by autoradiography.
Figure 2 Sequence alignments of the ULP1 family members

(A) Sequence alignment of amino acids in the catalytic core of yeast Ulp1, AtULP1A, AtULP1C, AtULP1D and AtESD4. Numbering is shown with respect to yeast Ulp1. Asterisks above the alignment denote the catalytic histidine, aspartic acid and cysteine residues. Black circles above the alignment denote residues in yeast Ulp1 that are in direct contact with yeast Smt3 as described by Mossessova and Lima [23]. The catalytic domains of all Arabidopsis ULP1s are extended to their native C-terminus. (B) Phylogenetic tree representation of amino acid sequence distance among AtULP1s used in the present study. The tree was calculated using ClustalW (v1.4) algorithm in the MacVector program.

In crystallography studies, Reverir and Lima [22] identified residues in Smt3 that interact with residues in its E2, Ubc9 (Figure 1A). Smt3 uses Glu93 (P5) to make a salt bridge via its Arg63 (P35) with UBC9 Glu122 [22]. These residues are conserved in ArSUMO-5, M-SUMO-1, M-SUMO-2 and M-SUMO-4. However, in T-SUMO, ArSUMO-1 and ArSUMO-2, the residue at P35 has been changed from a glutamic acid (or glutamine) residue to a histidine residue and the corresponding residue at P35 has been changed from an arginine to an alanine. We predict, with this configuration, that a histidine residue at P5 would directly interact with Ubc9 Glu122 to form a salt bridge. In the case of ArSUMO-3, both P5 and P35 are replaced with an alanine and, therefore, these changes do not hinder the conjugation process, as observed in Figure 1(B). We propose that the residues that are required for recognition by Ubc9 alter in a manner that maintains the recognition of the SUMO molecules for Ubc9.

The family of Arabidopsis ULP1s share limited sequence identity with yeast Ulp1

Initially, we attempted to see whether SUMO specificity of ArULp1-like enzymes could be predicted by alignment with various ULP1 family members and, as a guide, we used the co-crystal structure of the N-terminally deleted yeast Ulp1 (ΔUlp1) with yeast Smt3 [23]. In Figure 2(A), we aligned the catalytic core (Δ) of the ArULP1 proteases with yeast Ulp1 and designated the residues that are important for the interaction between yeast ΔUlp1 and its substrate, yeast SUMO, Smt3. Alignment of all of the enzymes was used to identify the N-terminal boundary of the catalytic domain for each ArULP1 (Figure 2A). The four FL ArULP1s share limited sequence identity with yeast Ulp1 (16–29% sequence identity), thereby making any predictions on substrate specificity difficult. In contrast, when the catalytic cores of Arabidopsis ULP1 family members are compared with one another, greater similarities are found to exist and the ArULP1s can be grouped into two pairs: ΔArULP1A and ΔArULP1D with 65% sequence identity and ΔArULP1C and ΔArULP1D with 72% sequence identity (Figure 2B). On the basis of only contact residues identified in the yeast ΔUlp1:Smt3 structure, yeast ΔUlp1 is found to be most similar to ΔArULP1A (Figure 2A). Overall, it is difficult to determine which ArULP1-like enzymes either share a limited range of SUMO-specificity as is observed with Xanthomonas XopD or exhibit more promiscuous activity as is seen with yeast Ulp1 [15].
To date, all enzymatic and structural studies have used only the C-terminal catalytic domain of ULP1s. However, recent studies have implicated the N-terminal regulatory domain in playing a role in substrate specificity [24]. In the present study, we have expressed and analysed both the FL and the catalytic core (Δ) of Arabidopsis ULP1 family members to assess the impact of the regulatory domain on the specificity and activity of the catalytic domain.

The peptidase activity of ULP1 family members with Arabidopsis SUMOs reveals differences in substrate specificity

To analyse the peptidase activity of the ULP1s, we designed an assay whereby the SUMO substrates (SUMO-Gly-Gly–HA) are cleaved after the C-terminal Gly-Gly residues, resulting in a product that migrates faster than the substrate on SDS/polyacrylamide gels (Figures 3 and 4). A recent report by Reverter et al. [25] suggested that residues C-terminal of the conserved Gly-Gly motif of SUMOs might play a role in the specificity of the ULP1 family members. We have observed only slight variability in the efficiency of the processing activity of yeast ΔULP1 due to the residues C-terminal to the Gly-Gly motif (R. Chosed and K. Orth, unpublished work). To eliminate variability in the efficiency of cleavage due to the differences at the C-terminus of SUMO substrates, thereby allowing the focus to be on the differences in the mature SUMO proteins, we have utilized SUMO constructs that encode an HA tag directly following the Gly-Gly motif.

GST fusion proteins of yeast ΔULP1 and the FL and the catalytic core of ArULP1A, ArULP1C, ArULP1D and ArESD4 were expressed in E. coli and purified using standard GST–glutathione affinity chromatography. All enzymes were expressed as soluble proteins and the yield of the affinity purified proteins varied between 2 and 10 mg/l. Molar excess of each ULP1 is incubated with each 35S-labelled SUMO substrate for 1 h at 30°C, to ensure cleavage of any potential substrate (Figures 3 and 4). Partial cleavage of specific SUMO substrates is denoted (Figures 3 and 4).

Within the first set of in vitro peptidase assays, we observed an unpredicted requirement for the regulatory domain of the Arabidopsis enzymes. Two of the four Arabidopsis ULP1s (ArULP1A and ArESD4) required their N-terminal regulatory domain for peptidase activity in vitro (Figure 3B, assays I–IV, lanes 2, 3, 8 and 9). Yet, for the remaining two Arabidopsis enzymes, ArULP1C and ArULP1D, the activity of the catalytic core was unaffected by the presence of the regulatory domain (Figure 3B, assays I–IV, lanes 4–7). Yeast ΔULP1 was able to cleave both ArSUMO-1 and ArSUMO-2 regardless of whether or not its N-terminal domain is present (Figure 3B, assays I and II, lane 10) (R. Chosed and K. Orth, unpublished work). Only ArULP1A was able to cleave ArSUMO-3, albeit weakly, and none of the ULP1-like enzymes cleaved ArSUMO-5 (Figure 3B, assays III and IV, lane 2). Owing to the weak and limited processing of ArSUMO-3 by the ULP1s, as well as the inability to detect any processing of ArSUMO-5, we wanted to eliminate the possibility that the use of the HA tag at the C-terminus of these SUMO proteins was interfering with the processing. Thus constructs of ArSUMO-3 and ArSUMO-5 were made with the native C-terminus of each protein followed by the Gly-Gly motif in place of the HA tag.
of the HA tag as used before (see Figure 1A for sequence of each C-terminal extension). As shown in Figure 3(B) (assays V and VI), these two constructs (ArSUMO-3-Gly-Gly-X\textsubscript{8} and ArSUMO-5-Gly-Gly-X\textsubscript{3}) were not processed by any of the ULP1s tested, suggesting that the HA tag was not impeding the processing of these SUMOs.

From these results, we observe a number of striking features on the specificity of the enzymes for various substrates. We observed similar activity profiles for the two ArULP1s that encode catalytic cores that are most similar to one another (ArULP1C and ArULP1D; 72% sequence identity) (Figure 1B). All of the active ArULP1s, and as predicted yeast \( \Delta \text{ULp1} \), were able to process ArSUMO-1 and ArSUMO-2, which is not surprising, based on the substrate’s 89% sequence identity. We observed a requirement for the N-terminal regulatory domain for the activity of ArULP1A and ArESD4, but not for ArULP1C and ArULP1D. These observations support the proposal that the N-terminal regulatory domains are able to affect the catalytic activity of the enzymes.

Insight was gained not only on the substrate specificity of the enzymes, but also on the ability of a substrate to be recognized as a substrate by a particular enzyme. Because none of the ULPI family members tested can process the most distantly related ArSUMO-5, it is possible to predict those residues that may impede the processing of this SUMO. By scanning residues at the P\textsubscript{1}–P\textsubscript{7} positions, we observe that the conserved glutamine residue at position P\textsubscript{4} is a leucine in ArSUMO-5 and that at position P\textsubscript{3} there is an uncharged valine residue (Figure 1A). These residues may account for the lack of enzyme recognition of this SUMO substrate, since residues in these positions were observed to make hydrogen bonds between yeast \( \Delta \text{ULp1} \) and Smt3. In addition, many of the residues in Smt3 (R72E and D83H) that conjugated to proteins in cultured cells [5,6]. The most obvious difference distinguishing this SUMO from others is the proline residue found at the P\textsubscript{4} position, as opposed to the conserved glutamine residue (Figure 1A). When M-SUMO-1 is mutated from a glutamine to a proline residue at the P\textsubscript{4} position, yeast \( \Delta \text{ULp1} \) is unable to cleave the mutated M-SUMO-1, indicating that the proline residue at P\textsubscript{4} has caused a change in the structure of the M-SUMO-1 that is not compatible with substrate recognition (R. Chosed and K. Orth, unpublished work). However, when M-SUMO-4 is changed from a proline at P\textsubscript{4} to a glutamine, yeast \( \Delta \text{ULp1} \) is still unable to cleave M-SUMO-4 (R. Chosed and K. Orth, unpublished work). Thus this suggests the presence of other residues in M-SUMO-4 that are necessary for substrate recognition by yeast \( \Delta \text{ULp1} \).

**Isopeptidase activity of ArULP1 family members**

All assays presented so far have examined the SUMO specificity of ULPIs using peptidase assays. To further characterize the activity and specificity of this panel of proteases, we analysed their specificity using isopeptidase assays. We utilized, as a substrate, RanGAP modified by various recombinant GST–SUMO proteins (Figure 1B). While RanGAP may not be the native substrate that is modified by these SUMO proteins, RanGAP is a commonly used substrate for in vitro SUMOylation assays. As observed above (Figure 1B), a fraction of the translated \(^{35}\text{S}\)-labelled RanGAP is modified by the RRL endogenous SUMOylation machinery (\(^{35}\text{S}\)-RanGAP-RRL-SUMO). Further modification by purified recombinant GST–M-SUMO-1, GST–T-SUMO, GST–ArSUMO-1, GST–ArSUMO-2, GST–ArSUMO-3 and GST–ArSUMO-5 produced a population of GST–SUMOylated \(^{35}\text{S}\)-labelled RanGAP (Figures 1B and 5B). Therefore each isopeptidase reaction includes radiolabelled \(^{35}\text{S}\)-RanGAP, \(^{35}\text{S}\)-RanGAP-RRL-SUMO, \(^{35}\text{S}\)-RanGAP-GST–SUMO and a test protease. Each protease is tested for activity with the same target protein, RanGAP, which is modified by either RRL-SUMO (as an internal control for each reaction) or recombinant GST–M-SUMO (Figure 5A). We selected GST–M-SUMO-1, GST–T-SUMO, GST–ArSUMO-1, GST–ArSUMO-2, GST–ArSUMO-3 and GST–ArSUMO-5 for in vitro SUMOylation assays. In these assays, GST–M-SUMO-1 serves as a GST control for the endogenous SUMO added to RanGAP during the in vitro translation reaction. The use of these six SUMOs provides a range of SUMO substrates in an internally controlled isopeptidase assay.

In the first isopeptidase assay, \(^{35}\text{S}\)-RanGAP-RRL-SUMO is used as substrate for the panel of ULPIs (Figure 5B, assay I). Only FL ArULP1A and ArESD4, \( \Delta \text{ArULP1C} \) and yeast \( \Delta \text{ULp1} \) process the endogenous SUMO from RanGAP (Figure 5B, assay I, lanes 2, 5, 8 and 10). This profile is similar to what was observed in the peptidase assays with M-SUMO-1.

In the second isopeptidase assay, both \(^{35}\text{S}\)-RanGAP-RRL-SUMO and \(^{35}\text{S}\)-RanGAP-GST–M-SUMO-1 are used as substrates. The results are identical with the first isopeptidase assay with the exception of one enzyme, \( \Delta \text{ArULP1C} \) (Figure 5B, assay II, lanes 2, 5, 8 and 10). \( \Delta \text{ArULP1C} \) is able to cleave the RRL-SUMO but not the GST–M-SUMO-1 (Figure 5B, assay II, lane 5). Therefore \( \Delta \text{ArULP1C} \) is able to recognize M-SUMO-1 as a substrate only in the peptide assay. The differences observed for \( \Delta \text{ArULP1C} \) could be due to the fact that rabbit SUMO (RRL-SUMO) varies enough from the human SUMO (GST–M-SUMO-1) such that it cannot be cleaved by \( \Delta \text{ArULP1C} \). In contrast, FL ArULP1D was able to recognize and partially cleave \(^{35}\text{S}\)-RanGTP-GST–M-SUMO-1 in the isopeptidase assay (Figure 5B, assay II, lane 6), but neither FL nor \( \Delta \text{ArULP1D} \) was able to recognize...
Arabidopsis SUMOylation

Figure 5  AtULP1 family members exhibit isopeptidase activity in vitro

(A) Schematic diagram of in vitro SUMO isopeptidase assay. (B) The ULP1 family of proteases shows specificity for the SUMO moiety of SUMOylated substrates. [35S]Mammalian RanGAP was in vitro translated in an RRL. In the RRL, some of the RanGAP is SUMOylated by endogenous SUMOylation machinery. The in vitro translated product was then used in an in vitro SUMOylation assay using purified recombinant GST–M-SUMO-1, GST–T-SUMO, GST–AtSUMO-1, GST–AtSUMO-2, GST–AtSUMO-3 and GST–AtSUMO-5 to produce GST–SUMO-modified RanGAP. The SUMO-modified RanGAP reaction was then incubated with a buffer (50 mM Tris, pH 8.0, 150 mM NaCl and 0.1 % 2-mercaptoethanol) or 0.5 mg/ml of FL AtULP1A, AtULP1C, AtULP1D and AtESD4, and the catalytic core (∆) of AtULP1A, AtULP1C, AtULP1D, AtESD4, or yeast Ulp1 for 1 h at 30°C. The samples were then resolved on SDS/8 % polyacrylamide gels and visualized by autoradiography. Roman numerals indicate each isopeptidase assay number as referred to in the text. Asterisks (*) indicate partial cleavage of GST–T-SUMO-RanGAP by FL AtULP1A and GST–M-SUMO-1-RanGAP by FL AtULP1D. (C) The GST moiety attached to AtSUMO-3 and AtSUMO-5 is not responsible for the lack of ULP1 activity towards these two SUMO substrates. As in (B), [35S]mammalian RanGAP was in vitro translated in an RRL where some of the RanGAP is SUMOylated by endogenous SUMOylation machinery. The in vitro translated product was then used in an in vitro SUMOylation assay using purified recombinant AtSUMO-3 and AtSUMO-5 (the GST tag was removed from these proteins using TEV protease) to produce AtSUMO-modified RanGAP. This AtSUMO-modified RanGAP migrates just below the RRL-SUMO-modified RanGAP. The SUMO-modified RanGAP reaction was then incubated with a buffer (50 mM Tris, pH 8.0, 150 mM NaCl and 0.1 % 2-mercaptoethanol) or 0.5 mg/ml of FL AtULP1A, AtULP1C, AtULP1D and AtESD4, and the catalytic core (∆) of AtULP1A, AtULP1C, AtULP1D and AtESD4 for 1 h at 30°C. The samples were then resolved on SDS/8 % polyacrylamide gels and visualized by autoradiography.

M-SUMO-1 as a substrate in peptidase assays (Figure 4, lanes 6 and 7).

In the third isopeptidase assay, the panel of ULP1s is tested using 35S-RanGAP-RRL-SUMO and 35S-RanGAP-GST–T-SUMO as substrates (Figure 5B, assay III). All enzymes, with the exception of ∆AtULP1C that recognized T-SUMO as a peptidase substrate, were able to cleave 35S-RanGAP-GST–T-SUMO. Similarly, in the fourth and fifth isopeptidase assays, the
panel of ULP1s is tested using 35S-RanGAP-RRL-SUMO and 35S-RanGAP-GST–AtSUMO-1 or 35S-RanGAP-GST–AtSUMO-2 as substrates (Figure 5B, assays IV and V). Again all enzymes that recognized AtSUMO-1 or AtSUMO-2 as peptidase substrates were able to cleave 35S-RanGAP-GST–AtSUMO-1 or 35S-RanGAP-GST–AtSUMO-2 with the exception of ΔAtULP1C.

In the sixth and seventh isopeptidase assays, none of ULP1 family members tested were able to remove the GST–AtSUMO-3 or GST–AtSUMO-5 modification from RanGAP (Figure 5B, assays VI and VII). In all isopeptidase assays, the profile of cleavage of the internal control, 35S-RanGAP-RRL-SUMO, was consistent. None of the ULP1s were able to remove the GST–AtSUMO-3 or GST–AtSUMO-5 from RanGAP, which could be due to the presence of the GST tag on the SUMOs. To address this concern, the GST tag of GST–AtSUMO-3 or GST–AtSUMO-5 was removed using TEV protease. The untagged purified recombinant AtSUMO-3 and AtSUMO-5 were then used in the in vitro SUMOylation assays and in the isopeptidase assays as before. As shown in Figure 5(C), the untagged AtSUMO-3 and AtSUMO-5 were used by the SUMOylation system to modify RanGAP, but none of the ULP1s were able to remove the untagged AtSUMO-3 or AtSUMO-5 from RanGAP. This assay demonstrates that the GST tag was not interfering with the ability of the ULP1s to remove the SUMOs from RanGAP.

Some obvious differences and similarities are seen when comparing the in vitro isopeptidase and peptidase activities of these enzymes. FL AtULP1A is extremely efficient at cleaving peptide bonds, but less efficient with isopeptide bonds. In contrast, AtESD4 cleaves SUMO-conjugated proteins better than SUMO peptides. Both the peptidase and isopeptidase activities of the ULP1s are affected by the presence or absence of their N-terminal regulatory domain. The modulation of an enzymatic domain by a regulatory domain is a recurring theme in many signalling systems, including phosphatases and kinases [26,27].

**Both redundancy and diversity are exhibited by the Arabidopsis SUMOylation system**

The aforementioned observations demonstrate that the deSUMOylation of protein SUMO conjugates is a highly regulated and complex system. While cellular localization of proteases may account for part of the regulation of SUMOylated proteins, we observe that, intrinsic to these proteases, is encoded another layer of specificity that dictates substrate recognition and cleavage. From our characterization of ULP1 family members expressed in Arabidopsis, we conclude that this family of proteases exhibits substrate specificity, both for the processing of SUMO and for the cleavage of SUMO conjugates. The substrate specificity of the enzymes used in the present study is summarized in Figure 6. Assays with both the catalytic core and the FL ULP1 proteins reveal variability in the requirement for the N-terminal regulatory domain.

The most striking examples of enzyme regulation are observed by the inhibitory and the activating effects of the regulatory domain on the catalytic domain. In one case, the activity of AtULP1C was inhibited by the presence of its regulatory domain when processing certain substrates (Figure 4). These observations are reminiscent of the classic regulation of Src kinase activity by its N-terminal domain [27]. In another case, the regulatory domain is required for both peptidase and isopeptidase activities for two of the AtULP1s (AtULP1A and AtESD4); however, for another, it is necessary only for its peptidase activity (AtULP1D). Thus the role for the regulatory domain in these proteases varies from one protein to another, further demonstrating the diversity in the ULP1 family of enzymes.

In our studies, none of the AtULP1s tested were able to recognize AtSUMO-5, and only one enzyme was able to recognize AtSUMO-3. The AtULP1s used in our studies had been classified based on sequence identity with known ULP1s. Recently, a de-ubiquitinating enzyme (DUB) from herpes simplex virus 1 was discovered, and interestingly, this enzyme showed no sequence homology with any known de-ubiquitinating enzyme [28]. Based on this finding, it is possible that other proteins in Arabidopsis have ULP1 activity, but do not share sequence homology with known ULP1s. Thus a ULP1-like protein may exist in Arabidopsis that is able to recognize and hydrolyse AtSUMO-5.

Our results support the hypothesis that not only is there sequence diversity but also mechanistic diversity in this family of proteins. Our findings support the hypothesis that although genome duplication has occurred, these protein families have diversified and are not simply functionally redundant copies of one another. The four Arabidopsis ULP1s studied here have unique catalytic profiles with respect to their substrate specificity.
Expression of the catalytic core is not always sufficient for observing the activity of a ULP1. In addition, in some cases, the regulatory domains can influence the specificity of these enzymes.

While we have observed specificity among the deSUMOylating enzymes for their SUMO substrates, we have also observed that the evolutionarily conserved SUMOylation machinery is promiscuous in its choice of SUMO substrate. We observed that all SUMOs can be used as substrates for conjugation, which supports the proposal that all SUMOs used in the present study have conserved residues utilized for their interaction with Ubc9. Thus differences in the amino acid sequence of each SUMO determine which Ulp1 will process which SUMO, yet similarity is maintained to allow for their recognition by the SUMOylation machinery. Furthermore, we predict that the residues important for dictating specificity by the ULP1s may also be important for the SUMO E3 ligases in specifying which SUMO is added to a target substrate.

Undoubtedly, the growing family of E3 SUMO ligases contribute to the specificity of the SUMO conjugation machinery for diverse substrates. During the evolution of the SUMO signalling system, surface residues in SUMOs have been conserved so that the substrates are recognized by the conjugation machinery, while other surface residues have been diversified, thereby adding complexity to this type of post-translational modification.

Our studies contribute to the basic understanding of the complexity of this transient and reversible system of regulation and are reminiscent of many other systems involved in reversible post-translational modifications.

We appreciate the generosity and support of our colleagues at University of Texas Southwestern, especially Alisha Tizenor, George DeMartino, Hong Tao Yu, Margaret Phillips, Diana Tischick, Yuh Min Chook and all the members of the Orth laboratory. We also thank Chris Lima (Structural Biology Program, Sloan-Kettering Institute, New York, NY, U.S.A.) for his generosity and support, and Cong-Yi Wang for his generosity in supplying reagents. This work was funded in part by NIH (National Institutes of Health, York, NY, U.S.A.) for his generosity and support, and Cong-Yi Wang for his generosity in supplying reagents. This work was funded in part by NIH (National Institutes of Health, Bethesda, MD, U.S.A.) AID grant AI056404 (to K. O.) and Welch Research Foundation grant AID grant AI056404 (to K. O.) and Welch Research Foundation grant.

REFERENCES


Received 20 March 2006; revised 2 June 2006
Published as BJ Immediate Publication 2 June 2006, doi:10.1042/BJ20060426

© 2006 Biochemical Society