

**Post-translational modification of proteins  
by SUMO in *Arabidopsis thaliana***

Inaugural - Dissertation  
zur  
Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Universität zu Köln

vorgelegt von

Ruchika Budhiraja  
aus Chandigarh, Indien

Köln, September 2005

Berichterstatter: Prof. Dr. George Coupland  
Prof. Dr. Jürgen Dohmen

Tag der letzten mündlichen Prüfung: 06.12.2005

## **Acknowledgements**

My sincere thanks to my supervisor Dr. Andreas Bachmair for giving me an opportunity to work in his laboratory. I am grateful for his guidance and stimulating discussions. I am particularly thankful for his constant support and vision throughout my tenure as a doctoral student in his lab. I will always be indebted to him for introducing me to molecular biology and lab techniques during my first year at the Max Planck Institute.

I would like to thank Prof. George Coupland for his role as my PhD advisor and for taking special interest in my research work. He has provided me much scope for thought during our meetings and advisory sessions. I am also grateful to Prof. Jürgen Dohmen from University of Cologne for being my secondary advisor and giving me worthy comments and suggestions on my project's regular progress reports.

My gratitude to Dr. Ralf Petri, Scientific Coordinator of the International Max Planck Research School (IMPRS) at our Institute, for being my secondary supervisor and a wonderful coordinator. His encouragement and assistance made my stay in Cologne a bit easier.

I am grateful to the Max Planck society for financial support during my PhD studies.

I am appreciative of Kerstin Luxa and Michaela Lehnen for their technical assistance and help in the laboratory. Special thanks to Jürgen Schmidt, Anne Bräutigam, Thomas Colby (MPIZ) and Stephan Müller (University of Cologne) for their help in mass spectrometry analysis. I would also like to acknowledge our collaborator Dr. Christian Hardtke (McGill University /Univ. of Lausanne) for providing some research material.

I thank some wonderful friends and colleagues, especially Nicki and Mrs Helga Stein for being great friends and for their unwavering encouragement that has been the most important factor in the completion of this task.

Finally, I would like to thank my family back home, for their undaunting perseverance and support all this while that I have been away from them in the pursuit of my goal.

---

## Table of contents

	Page
<b>Zusammenfassung</b>	
<b>1. INTRODUCTION</b>	
1.1. Polypeptides as protein modifiers	1
1.2. Discovery of SUMO modification	1
1.3. SUMO structure and isoforms	2
1.4. The <i>Arabidopsis</i> SUMO conjugation pathway	7
1.4.1. SUMO-activating enzyme (SAE)	9
1.4.2. SUMO-conjugating enzyme (SCE)	10
1.4.3. SUMO ligases	11
1.4.4. SUMO proteases - desumoylation	13
1.5. Substrate specificity in sumoylation	14
1.6. Substrates and functions of SUMO protein modification	15
1.7. Strategies to analyze SUMO conjugation in <i>Arabidopsis thaliana</i>	17
<b>2. MATERIAL AND METHODS</b>	
2.1. <b>Material</b>	19
2.1.1. Plant plasmid constructs	19
2.1.2. Bacterial plasmid constructs	22
2.1.3. Plasmid vectors used	23
2.1.4. Oligonucleotides used	27
2.1.5. Bacterial strains	29
2.1.6. Plant and bacterial growth media	29
2.1.7. Chemicals and Enzymes	30
2.1.8. Antibiotics	30
2.1.9. Herbicide	31
2.2. <b>Methods</b>	31
2.2.1.1. Isolation of plasmid DNA from <i>E. coli</i>	31
2.2.1.2. Treatment of DNA with restriction endonucleases	32
2.2.1.3. Running of agarose gels and purification of DNA fragments	32
2.2.1.4. DNA ligation	32
2.2.1.5. Transformation of <i>E. coli</i>	33
2.2.1.6. Growth and transformation of <i>Agrobacterium tumefaciens</i>	33
2.2.1.7. Analysis of transformed <i>A. tumefaciens</i>	33
2.2.1.8. <i>A. tumefaciens</i> mediated transformation of <i>Arabidopsis thaliana</i>	34
2.2.1.9. Seed sterilization and selection of transformants	34
2.2.1.10. Crosses	35

	<b>Page</b>
2.2.1.11. Application of estradiol to transgenic <i>Arabidopsis</i> plants	35
2.2.2. Plant growth and analysis of flowering time	35
2.2.3. Biochemical enrichment of SUMO1 conjugates from <i>Arabidopsis</i>	36
2.2.4. Techniques for protein analysis	38
2.2.4.1. Isolation of proteins from <i>A. thaliana</i>	38
2.2.4.2. Measurement of protein concentration with Bradford reagent	39
2.2.4.3. Precipitation of proteins using organic solvents	39
2.2.4.4. Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE)	41
2.2.4.5. Coomassie staining of proteins resolved by SDS-PAGE	41
2.2.4.6. Sypro-Ruby staining of proteins	42
2.2.4.7. Excision of protein bands, in-gel trypsin digest, and MALDI-TOF analysis	42
2.2.4.8. Isoelectric focusing (IEF)	44
2.2.4.9. Two-dimensional gel electrophoresis	45
2.2.4.10. Protein spotting, trypsin digests with a robot and MALDI-TOF analysis	46
2.2.4.11. Western blotting	46
2.2.4.12. Small scale testing of overexpression of proteins in <i>E. coli</i>	47
2.2.4.13. Purification of His-tagged proteins expressed in <i>E. coli</i> using Ni-NTA matrix	47
2.2.4.14. Expression and purification of SUMO-conjugating enzyme (SCE)	49
2.2.5. Production of SCE antibody in rabbit	50
2.2.5.1. Purification of rabbit polyclonal immunoglobulins (IgGs) using affinity chromatography	50
2.2.6. <i>In vitro</i> sumoylation assays	51
2.2.7. Brief protocol for PCR-grade DNA isolation from <i>A. thaliana</i>	52
2.2.8. Polymerase chain reaction (PCR)	52
2.2.9. Nucleic acid hybridization techniques	52
2.2.9.1. Isolation of total RNA from <i>A. thaliana</i>	52
2.2.9.2. Electrophoresis of RNA on denaturing gels	53
2.2.9.3. Northern blotting	54
2.2.9.4. Radioactive labeling of probes	54
2.2.9.5. Hybridization of radio labeled probe to RNA immobilized on nylon membrane	55

### 3. RESULTS

3.1	<b>Characterization of proteins of the sumoylation pathway</b>	56
3.1.1.	Cloning and expression of tagged SUMO isoforms in pET-9d	56
3.1.2.	Purification of SUMO isoforms from <i>E. coli</i>	56
3.1.3.	Cloning of SUMO-activating enzyme	58
3.1.4.	Purification of <i>Arabidopsis</i> recombinant SAE from <i>E. coli</i>	58
3.1.5.	Cloning and expression of <i>Arabidopsis</i> SUMO-conjugating enzyme in pET-9d	59
3.1.6.	Purification of the <i>Arabidopsis</i> SCE from <i>E. coli</i>	59
3.2.	<b><i>In vitro</i> analysis of sumoylation</b>	60
3.2.1.	Activation of <i>Arabidopsis</i> SUMOs with human sumoylation enzymes	60

	<b>Page</b>
3.2.2. Sumoylation of nucleosome assembly factor (NAF)	61
3.3. <b>SUMO-conjugating enzyme mutant</b>	63
3.3.1. Phenotypic characterization of plants overexpressing SCE (C94S)	63
3.3.2. Characterization of the flowering time of plants overexpressing SCE (C94S)	66
3.3.3. Analysis of SCE (C94S) lines by immunoblot assays	66
3.3.4. Purification of rabbit polyclonal immunoglobulins (IgGs)	67
3.3.5. Further characterization of SCE (C94S) mutant	68
3.3.6. mRNA expression pattern of selected genes in SCE (C94S) mutant plants	69
3.4. <b><i>In vivo</i> analysis of sumoylation</b>	71
3.4.1. Expression of SUMO isoforms in <i>Arabidopsis</i>	71
3.4.2. Sumoylation pattern	73
3.4.3. Splice junctions of poorly expressed SUMO isoforms	74
3.4.4. Analysis of C-terminal SUMO variants in <i>Arabidopsis</i>	75
3.4.5. Analysis of SUMO variants potentially inhibiting deconjugation	76
3.4.6. Biochemical enrichment of SUMO1 substrates	78
3.4.6.1. Separation and analysis of enriched extracts	78
3.4.6.2. Identification of sumoylated proteins in enriched fractions by mass spectrometry	80
3.4.6.3. Identification of sumoylated proteins in enriched fractions using antibodies	88
<b>4. DISCUSSION</b>	
4.1. <i>Arabidopsis</i> SUMO proteins	90
4.2. C-terminal SUMO variants expressed in <i>Arabidopsis thaliana</i>	91
4.3. SUMO variants potentially inhibiting deconjugation	91
4.4. Biochemical enrichment of sumoylated proteins of <i>Arabidopsis</i>	92
4.5. <i>Arabidopsis thaliana</i> SCE (C94S) mutant	94
4.6. <i>In vitro</i> sumoylation assays	95
4.7. Sumoylated substrates identified using antibody targeted approach	96
4.8. Future perspective	96
<b>5. SUMMARY</b>	97
<b>6. REFERENCES</b>	99
<b>7. APPENDIX</b>	110
Abbreviations	110
List of Figures and Tables	113
Sequence of <i>Arabidopsis</i> SUMO transgenes	116
Declaration (Erklärung)	119
Curriculum vitae (Lebenslauf)	120

## Zusammenfassung

Das kleine ubiquitin-ähnliche Protein SUMO kann kovalent mit Substratproteinen verknüpft werden. Neben der Ubiquitylierung ist dies der am besten untersuchte Fall einer posttranslationalen Veränderung von Proteinen durch Anheften sogenannter Modifikatorproteine. SUMO besteht aus ca. 100 Aminosäuren und wird in Eukaryonten, z. B. in Hefe, Menschen und Pflanzen, an Substrate angeheftet. Dazu dienen spezifische Enzyme, nämlich SUMO aktivierendes Enzym SAE, SUMO konjugierendes Enzym SCE und SUMO Ligasen. Diese Enzyme sind homolog zur Enzymkaskade E1, E2 und E3, welche die Konjugation von Ubiquitin katalysiert. SUMO stellt sich als vielseitig verwendetes Modifikatorprotein heraus, das eine große Zahl an Substraten in verschiedenen Stoffwechsel und Signaltransduktionswegen aufweist, und die Konsequenzen der Modifikation scheinen so vielfältig zu sein wie die Substratproteine selbst. *Arabidopsis thaliana* hat acht Gene mit Ähnlichkeit zu SUMO aus Pilzen oder Tieren. Phylogenetisch können sie in fünf Gruppen eingeteilt werden: SUMO1/2, SUMO3, SUMO5, SUMO4/6, und SUMO7/8. Vier der *Arabidopsis* SUMO Gene, SUMO1, 2, 3 und 5, werden stark exprimiert und *in vivo* zur Bildung von Konjugaten herangezogen. Um SUMO Substrate in Pflanzen zu identifizieren und zu charakterisieren, wurden SUMO Gene aller Gruppen mit Epitop-Tag versehen und in der Pflanze *Arabidopsis thaliana* exprimiert. Nach Expression von SUMO1 mit Epitop-Tag und biochemischer Anreicherung wurde eine Reihe hochmolekularer Konjugate im Coomassie-gefärbten Gel erhalten. MALDI-TOF Analyse erlaubte die Identifizierung von 25 neuen potentiellen Substraten der Sumoylierung in *Arabidopsis*. Einige dieser Kandidatenproteine enthalten ein sogenanntes Konsensus-Sumoylierungs-Motif. Es konnte gezeigt werden, dass eines dieser Proteine, NAF, auch *in vitro* ein Substrat der Sumoylierung ist.

Alle SUMO Proteine werden zunächst als inaktive Vorstufen hergestellt. Die Reifung erfolgt durch proteolytische Spaltung am Carboxyl-Terminus, an dem ein Di-Glycin Motif freigelegt wird. Um zu testen, ob diese beiden terminalen Aminosäuren für die Konjugation essentiell sind, wurden SUMO Varianten hergestellt und getestet, die

Alanin-Glycin, Glycin-Alanin oder Alanin-Alanin an deren Stelle aufweisen. Anders als erwartet konnten in *Arabidopsis* Konjugate mit allen Varianten nachgewiesen werden, was die hohe Flexibilität des pflanzlichen SUMO Konjugationssystems belegt. Es konnte auch die Anhäufung von Konjugaten nachgewiesen werden, welche nach Expression von SUMO 1 mit einem Glutamin zu Alanin Austausch in Position 93 in *Arabidopsis* entstanden. Weiters wurden Pflanzen untersucht, deren Fähigkeit zur Bildung von SUMO Konjugaten durch Expression eines veränderten SUMO konjugierenden Enzyms beeinträchtigt war. Im veränderten Enzym ist das Cystein des aktiven Zentrums durch Serin ersetzt. Phänotypische Charakterisierung der exprimierenden Pflanzen zeigte Kleinwüchsigkeit und frühzeitige Blühinduktion im Lang - sowie im Kurztag. Analyse mittels Immunoblot ergab, dass die veränderten Pflanzen eine geringere Konzentration an freiem SUMO aufweisen. Insgesamt zeigen die Ergebnisse, dass ähnlich wie in anderen Eukaryonten viele Pflanzenproteine durch SUMO modifiziert werden können und dass dieser Prozess wichtig ist für Entwicklung und Zelluläre Funktionen von *Arabidopsis thaliana*.

## **1. INTRODUCTION**

### **1.1 Polypeptides as protein modifiers**

All organisms use a variety of chemical modifiers for the post-translational control of proteins that effect development, growth and homoeostasis. In these cases, specific amino acid residues of target proteins are modified by molecules such as methanol, acetate, fatty acids, sugars, nucleosides and phosphate. In addition to these molecules, eukaryotes also employ polypeptides as protein modifiers. Ubiquitin (Ub) was one such polypeptide that was first described to covalently attach to other proteins after their synthesis has been completed. This concept of covalent attachment of one polypeptide to another was first realized ~ 25 years ago with the noble prize winning work of Hershko, Ciechanover, Rose and coworkers (Hershko and Ciechanover, 1998). Conjugation of ubiquitin (termed ubiquitination) has a well established role in earmarking proteins for degradation by the 26S proteasome. During the past few years, a panoply of proteins have been discovered in eukaryotes that have sequence similarity to ubiquitin. The ubiquitin like proteins fall into two separate classes (Jentsch and Pyrowolakis, 2000). Proteins of the first class termed as ‘ubiquitin-like modifiers’ (UBLs) function in a manner analogous to that of ubiquitin. Examples include SUMO (small ubiquitin-like modifier), Rub1 (also named Nedd8), Apg8 and Apg12. Proteins of the second class include RAD23, DSK2 and parkin. They bear domains that are related in sequence to ubiquitin (otherwise unrelated to each other) and are designated ‘ubiquitin domain proteins’ (UDPs). In contrast to UBLs, these proteins are not conjugated to other proteins (Hochstrasser *et al.* 2000). In the UBLs class of proteins, some polypeptides are highly similar to each other in amino acid sequence (Ub and Rub1), whereas others share little homology (Ub and Apg 8). Despite their sequence diversity, structure determinations indicate that most of these polypeptides adopt a similar three-dimensional shape (Fig. 1 e.g Ub and SUMO).

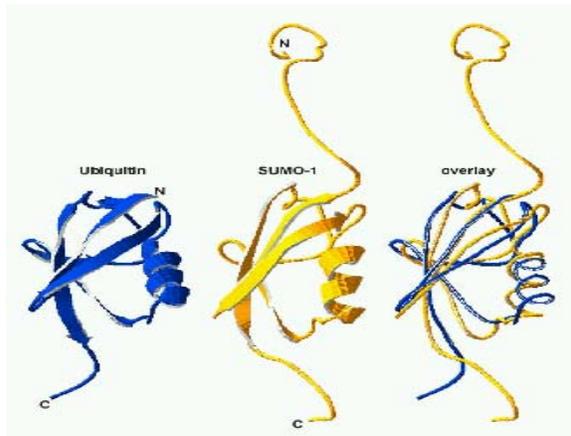
### **1.2 Discovery of SUMO modification**

SUMO was first discovered in animals during studies on nuclear import in mammalian

cells as a covalent modifier of RanGAP1, a protein associated with the nuclear import complex. On amino acid sequencing, RanGAP1 was discovered to have two N-termini, but only one C-terminus, indicating that the protein contained a second polypeptide attached through an isopeptide bond (Matunis *et al.* 1996). The tagged polypeptide was termed SUMO (small ubiquitin-like modifier) based on its weak similarity to ubiquitin (Mahajan *et al.* 1997). This discovery was also facilitated by the fact that RanGAP1 is quantitatively and constitutively modified with SUMO. This modification targets the cytosolic RanGAP1 to the nuclear pore complex where it plays a role in nuclear import by activating the GTPase activity of cytosol/nuclear shuttling factor Ran (Mahajan *et al.* 1998). SUMO was also identified independently in a number of other studies subsequently coining alternate names for the protein as Sentrin, Smt 3, UBL1, GMP or PIC1 (Okura *et al.* 1996; Mannen *et al.* 1996; Shen *et al.* 1996; Boddy *et al.* 1996).

### **1.3 SUMO structure and isoforms**

SUMOs share only ~18 % sequence identity with ubiquitin, although structure analysis by nuclear magnetic resonance (NMR; Bayer *et al.* 1998) revealed that both share a common three dimensional structure that is characterized by a tightly packed globular fold with the  $\beta$ -sheets wrapped around one  $\alpha$ -helix. The three-dimensional folds of SUMO and ubiquitin can be superimposed (Fig. 1). The Gly Gly motif at the C terminus of both the proteins, which is the site of attachment to target proteins, is also positioned alike. However, the surface charge topology of SUMO is quite different from ubiquitin, with distinct positive and negative regions (Müller *et al.* 2001). SUMO proteins contain an unstructured short amino-terminal extension of up to 22 residues not present in ubiquitin, which provides an additional interface for protein-protein interactions (Seeler and Dejean, 2003). This extension varies among different SUMO proteins from 11 to 35 amino acids and is well conserved within, but not between different SUMO families. Interestingly, SUMO proteins do not have the lysine residue corresponding to Lys-48 in the ubiquitin molecule that is required for the formation of poly ubiquitin chains, implying that SUMO does not make the same type of multi chain as ubiquitin (Bayer *et al.* 1998).

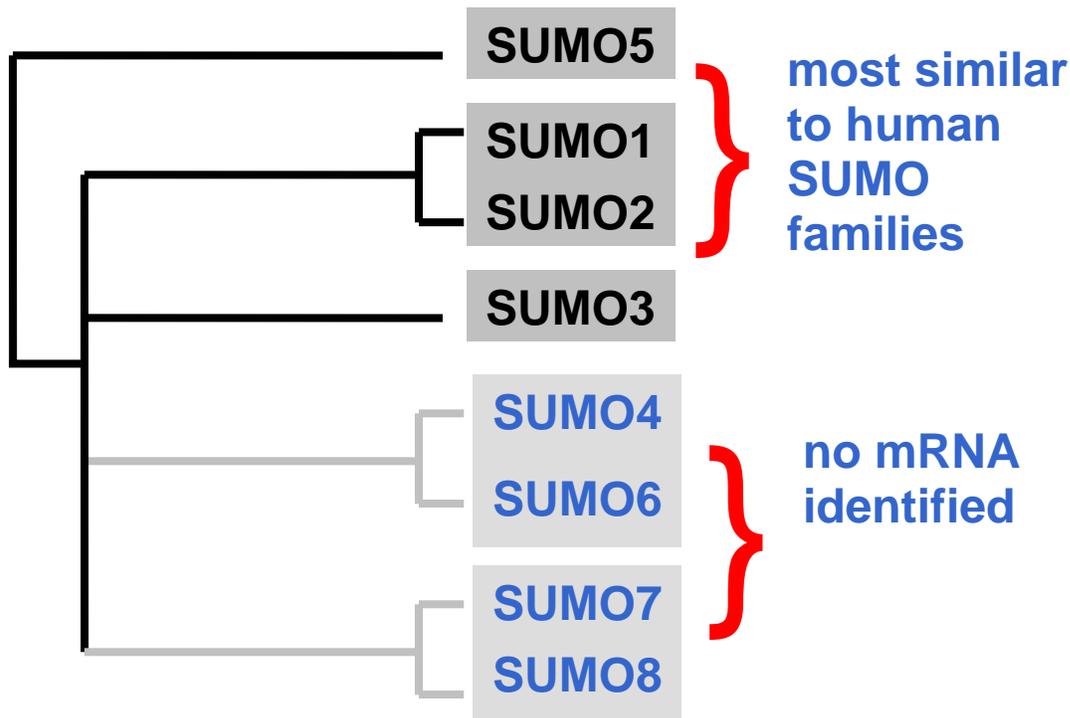


**Fig. 1 Ribbon diagrams highlight the similarity of the three-dimensional structures of ubiquitin and human SUMO1** (taken from Dohmen 2004). Both proteins have a tightly packed  $\beta\beta\alpha\beta\beta\alpha\beta$  fold. Notably SUMO has a N-terminal extension not found in ubiquitin. X-ray crystallography was used for determining the structure of ubiquitin (Vijay-Kumar *et al.* 1987); the structure of SUMO was determined in solution by NMR technique (Bayer *et al.* 1998).

SUMOs are ca. 15 kDa proteins present in all eukaryotic kingdoms and highly conserved from yeast to humans. The yeast and invertebrates studied up to date contain only a single SUMO gene termed as SMT3, originally discovered as a suppressor of mutants in the centromeric protein MIF2 (Meluh and Koshland, 1995). While the SUMO encoding SMT3 is essential for the viability in budding yeast *Saccharomyces cerevisiae*, (Johnson and Blobel, 1997) fission yeast *Schizosaccharomyces pombe* lacking the SUMO gene PMT3 are barely viable and have severe defects in the genome maintenance (Tanaka *et al.* 1999). The mammalian SUMO family members consist of SUMO1, SUMO2 and SUMO3 (Kamitani *et al.* 1998). SUMO1 shares 48% identity with SUMO2 and 46% identity with SUMO3. SUMO2 and SUMO3 share 95% identity, and can be grouped into a subfamily distinct from SUMO1. All three members have distinct N-terminal amino acid sequences and C-terminal extensions. The recently identified fourth isoform, SUMO4, is encoded by a sequence that lies within an intron of the human TAB2 gene (Bohren *et al.* 2004). The divergence of the functions of these various isoforms is just beginning to emerge. Like SUMO1, other members of this family can also be conjugated to target proteins. Recent studies of Saitoh and Hinchey (2000) have shown the functional heterogeneity of SUMO family members. Using an antibody which interacts with SUMO2 and SUMO3, but not with SUMO1, they demonstrated that SUMO2 and SUMO3

are conjugated poorly to RanGAP1, a major SUMO1 target protein. However, conjugation of SUMO 2\3 can be strongly induced in response to various stress conditions such as high temperature. (Saitoh *et al.* 2000). Another difference between SUMO1 and SUMO2\3\4 is that while SUMO2, SUMO3 and SUMO4 contain a SUMO attachment consensus sequence ( $\psi$ KXE\D, where  $\psi$  is a hydrophobic aliphatic residue, X is any residue and K, E and D correspond to the standard one letter symbols for amino acids; K is the lysine where SUMO attaches) in their N-terminal extension, such a site is absent from SUMO1. Consistent with this observation, in contrast to SUMO1, SUMO2\3 as well as SUMO4 have been shown to form SUMO chains *in vivo* and *in vitro* (Tatham *et al.* 2001).

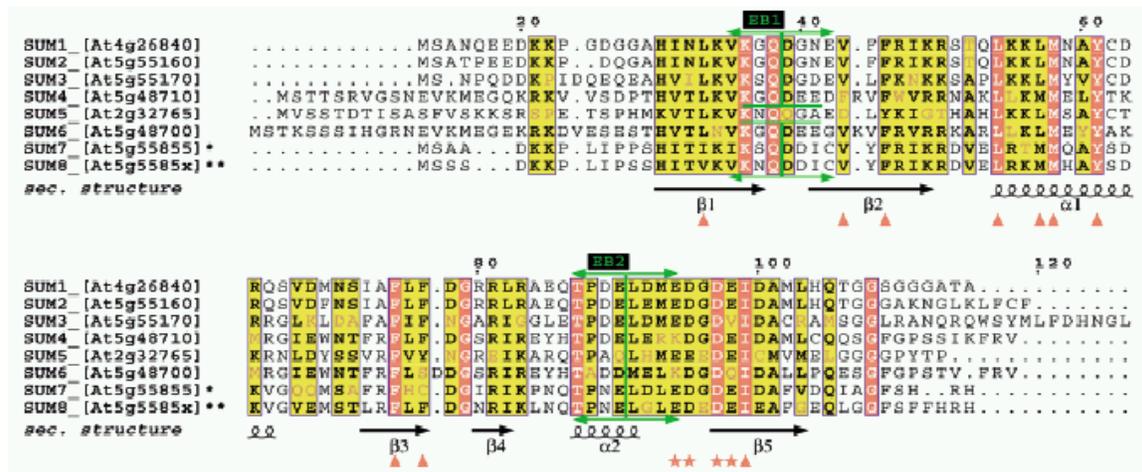
Plants contain even more SUMO genes. The genome of the model plant *Arabidopsis thaliana* has nine genes that show similarity to animal and fungal SUMO proteins. (Table 1; Kurepa *et al.* 2003) One of them, namely SUMO9, is a pseudogene and does not encode a complete SUMO protein. Phylogenetic analysis clustered the eight full-length *Arabidopsis* SUMO proteins into five subfamilies: SUMO1\2, SUMO3, SUMO5, SUMO4\6 and SUMO7\8 (Fig. 2). The SUMO gene family is potentially derived from genome rearrangements. For instance, SUMO2 and SUMO3, as well as SUMO4 and SUMO6 are closely linked and are listed as examples of tandem duplications. The same probably holds true for SUMO7 and SUMO8 (Novatchkova *et al.* 2004). Sequence comparison shows that SUMO1\SUMO2, SUMO4\SUMO6 and SUMO7\SUMO8, respectively, are very similar to each other (Fig. 3). SUMO5 is sequentially most distinct for all other *Arabidopsis* SUMOs. ESTs (expression sequence tags) exist for SUMO1, SUMO2, SUMO3 and SUMO5, providing evidence for *in vivo* expression. The expression levels of SUMO4, SUMO6, SUMO7 and SUMO8, if they do not represent pseudogenes, is very low. Forced expression of an intron-containing SUMO7 construct allowed detection of mRNA (Budhiraja, R. and Bachmair A., unpublished). cDNA isolation indicated the formation of two splice variants, SUMO7 and SUMO7v. The latter has a three amino acid insertion (Glu-Leu-Gln) at the position of the second intron (see Fig. 3). Forced expression of SUMO6 confirms the intron-exon structure as predicted by computer algorithms (Novatchkova *et al.* 2004).



**Fig. 2 Phylogenetic tree of the *Arabidopsis* SUMO family.** SUMO1\2, SUMO3 and SUMO5 are highly expressed. SUMO4, SUMO6, SUMO7, SUMO8 have low expression levels.

As with yeast and animal SUMOs, *Arabidopsis* SUMO1-8 bear additional C-terminal residues beyond the glycine necessary for conjugation. Presumably, these residues are removed post-translationally by SUMO proteases to generate the mature active tags of 96-100 amino acids. Whereas most mature *Arabidopsis* SUMOs, like ubiquitin, are predicted to terminate in a Gly Gly motif, three (SUMO4, 6 and 7) are predicted to end in a single glycine. However, it is not yet known whether this distinction affects the processing and/or activity of these SUMOs. All highly expressed SUMO isoforms in *Arabidopsis* are engaged in conjugation reactions. Antibodies directed against SUMO1\SUMO2 (Lois *et al.* 2003, Murtas *et al.* 2003), and those directed against SUMO3 (Kurepa *et al.* 2003) indicate that these proteins form conjugates *in vivo*. Likewise,

expression of epitope-tagged SUMO5 allows detection of conjugates with this protein (Budhiraja, R. and Bachmair A., unpublished).

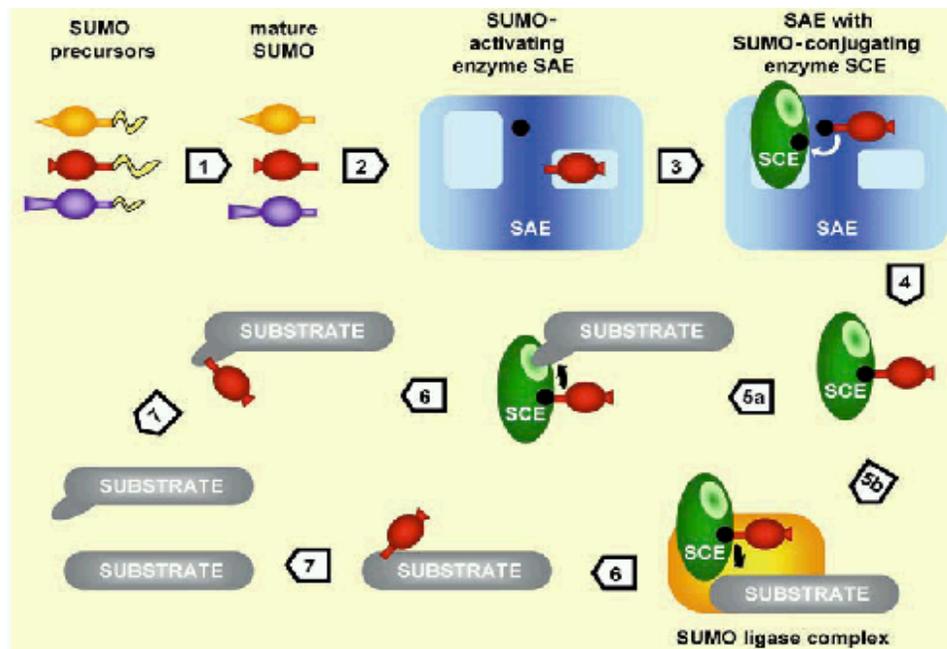


**Fig. 3 Alignment of SUMO protein sequences of *Arabidopsis thaliana*** (taken from Navatchkova M., Budhiraja R., Coupland G., Eisenhaber F., Bachmair A., 2004). Conserved residues in the SUMO core have a yellow background; highly conserved residues are highlighted in red. EB1 and EB2 indicate positions of intron 1 and intron 2. *Sec structure* indicates the predicted secondary structure. Red triangles below the alignment indicate hydrophobic residues important for the stability of the compact ubiquitin-like core of SUMO. Asterisks below the alignment indicate amino acid residues that form a acidic patch over the SUMO surface, a feature that distinguishes the SUMO from other protein modifiers. Dots indicate spaces introduced to optimize alignment. As a maturation step, all SUMO proteins are predicted to be cleaved after the last conserved glycine residue at position 108. A cDNA splicing variant of SUM7 (not shown) contains the three amino-acid insertion Glu-Leu-Gln at position EB2 (see text).

#### 1.4 The *Arabidopsis* SUMO conjugation pathway

The pathway of conjugation of SUMO (termed sumoylation) is mechanistically analogous to ubiquitin conjugation (termed ubiquitylation or ubiquitination). The enzymes of the SUMO pathway, although similar to those of the ubiquitin pathway, are specific for SUMO and have no role in conjugating ubiquitin or any other ubiquitin-like proteins. Like ubiquitin, all SUMO forms are made as inactive precursor proteins. They mature by a carboxyl terminus cleavage event, which generates the mature form of the protein with exposed carboxyl terminus glycine residues. This cleavage reaction is catalyzed by a group of specific cysteine proteases, termed ULPs (ubiquitin-like protein processing enzymes) or SUMO-specific proteases (Li and Hochstrasser, 2000; Schwienhorst *et al.* 2000). Cleavage occurs after a conserved glycine residue (position 108 in Fig. 3). These double glycine residues are required for the formation of the SUMO-substrate linkage. Whereas most plant SUMO proteins have the same glycine-glycine motif at the cleavage site as present in animal and fungal SUMOs, the carboxyl termini of SUMO4, SUMO6 and SUMO7 deviate at the penultimate position. SUMO7 has an alanine-glycine, while SUMO4 and SUMO6 have a serine-glycine instead. The linkage between SUMO and its substrates is an isopeptide bond between the carboxyl terminus of SUMO and an  $\epsilon$ -amino group of a lysine residue in the target protein.

Sumoylation requires a specific set of enzymes which were first characterized in the yeast *Saccharomyces cerevisiae* (Fig. 4). The SUMO pathway begins with SUMO-activating enzyme (SAE), which carries out an ATP-dependent activation of the SUMO carboxyl terminus and then transfers the activated SUMO to a SUMO-conjugating enzyme (SCE) occasionally called Ubc9. SUMO is then transferred from SCE to a lysine residue in the substrate with the assistance of one of the several SUMO-protein ligases (E3). Many of the lysine residues where SUMO becomes attached are in the short consensus sequence  $\psi$ KXE\textbackslash D where  $\psi$  is a large hydrophobic amino acid, generally isoleucine, leucine or valine; K is the lysine residue that is modified; X is any residue; E is glutamic acid and D is aspartic acid (Rodriguez *et al.* 2001). Consistent with structural studies, showing direct



**Fig. 4 The sumoylation cycle** (taken from Navatchkova M., Budhiraja R., Coupland G., Eisenhaber F., Bachmair A., 2004). Multiple isoforms of SUMO exist in plants. All SUMOs are synthesized as inactive precursors that undergo proteolytic cleavage at their carboxyl terminus by SUMO-specific proteases (*Step 1*). Mature SUMO is activated by SUMO-activating enzyme (SAE), a heterodimer that has two large cavities (*light blue boxes*). One of the cavities can bind SUMO for activation (*Step 2*). The carboxyl terminus glycine of mature SUMO is activated by linkage to ATP, forming an AMP-SUMO intermediate. The SUMO carboxyl terminus is subsequently coupled to a cysteine residue of SAE (symbolized by a black dot) in a thioester linkage (*Step 3*). The second cavity of SAE can hold SUMO-conjugating enzyme (SCE). SUMO is transferred to the active-site cysteine residue of SCE, which dissociates from the complex (*Step 4*). SCE can directly bind to the substrates that contain a sumoylation consensus sequence ( $\psi$ KXE/D) in an accessible position (*Step 5a*). So far, this sequence of events is mainly supported by *in vitro* data. Alternatively, SUMO protein ligases form a ternary complex with SCE and substrate, to catalyze sumoylation of substrate proteins at an  $\epsilon$ -amino group of internal lysine residues (*Step 5b*). The sumoylation substrates are released (*Step 6*). SUMO-specific proteases cleave off SUMO for re-use and restore the substrate to its previous form (*Step 7*).

recognition of this consensus sequence motif by SCE, recombinant SAE, SCE and SUMO are sufficient for ATP-dependent SUMO modification of many substrates *in vitro* (Bernier-Villamor *et al.* 2002). SUMO ligases probably enhance specificity by interacting with other features of the substrate or by activating the SUMO-SCE complex (Reverter and Lima, 2005).

Sumoylation is a reversible and dynamic process. The cleavage of SUMO from its target proteins (termed desumoylation or deconjugation) is catalysed by ULPs that specifically cleave at the carboxyl terminus of SUMO. *In vitro*, ULPs can catalyze the processing of SUMO, yielding its mature form as well as remove SUMO from the isopeptide-linked conjugate.

#### **1.4.1 SUMO-activating enzyme (SAE)**

Like the E1 of the ubiquitin conjugation pathway, the SUMO-activating enzyme activates SUMO at the carboxyl terminus glycine residue (Fig. 4) SUMO-activating enzyme (SAE) is a heterodimer of 40 kDa and 70 kDa (Dohmen *et al.* 1995; Desterro *et al.* 1999). One protein shows similarity to the amino-terminal half, and the other to the carboxyl terminus half of the ubiquitin activating enzyme. *Arabidopsis thaliana* contains two genes for the smaller SAE subunit, SAE1a (At4g24940) and SAE1b (At5g50580). SAE1a and SAE1b are contained in segments that are duplicated between the chromosomes 4 and 5. The larger subunit of SAE, SAE2 (At2g21470) is represented by a single copy gene in the *Arabidopsis* genome (Table 1; Kurepa *et al.* 2003). The available structural data for the activation enzyme of RUB1, another protein modifier, suggests a mechanism for the activation that probably holds true for all the protein modifiers including SUMO (Walden *et al.* 2003). The SUMO-activating enzyme (SAE) catalyzes a three part reaction. First, the C-terminal carboxyl group of SUMO attacks ATP forming a SUMO C-terminal adenylate and releasing pyrophosphate. Next, the thiol group of the active site cysteine in the E1 attacks the SUMO adenylate, releasing AMP and forming a high energy thiolester bond between the E1 and the C terminus of SUMO. In the final reaction, the activated SUMO is transferred to a cysteine residue in the SCE. Most organisms contain a single SUMO-activating enzyme, which is required for the

conjugation of all SUMO isoforms.

#### **1.4.2 SUMO-conjugating enzyme (SCE)**

The second step after activation of SUMO is its transfer to the active site cysteine of SUMO-conjugating enzyme (SCE), to form a SUMO-E2 thiolester intermediate (Johnson and Blobel, 1997; Desterro *et al.* 1997). The SCE serves as a final donor in the final reaction in which SUMO is transferred to the substrate. The substrates are linked to SUMO via an isopeptide bond between the  $\epsilon$ -amino group of an internal lysine residue and the activated SUMO carboxyl terminus.

SCE is the only SUMO-conjugating enzyme in yeast and invertebrates, and most likely in vertebrates as well. *Arabidopsis thaliana* has one pseudogene and one active gene for the SUMO-conjugating enzyme (SCE1a). The enzyme is called Ubc9 in baker's yeast *Saccharomyces cerevisiae* because of its similarity to ubiquitin-conjugating enzymes, and Hus5 in the fission yeast *Schizosaccharomyces pombe*. As the SUMO-activating enzyme, SCE is predominantly a nuclear protein (Seufert *et al.* 1995). The presence of only one gene in *Arabidopsis* is interesting in the light of the fact that the plant has eight distinct SUMO proteins. The presence of only one SCE contrasts with the ubiquitin pathway where the multiple ubiquitin-conjugating enzymes participate in ubiquitylation of distinct sets of substrates (Bachmair *et al.* 2001). The gene encoding SCE is essential in all organisms studied except in *S. pombe*, in which the SUMO conjugation is not required for viability (Seufert *et al.* 1995; Jones *et al.* 2002; Ho and Watts, 2003). Current data suggest that in animals and probably also in plants, SUMO conjugation reactions *in vivo* as well as *in vitro* proceed without the assistance of protein ligases (discussed later in the text). SCE has a strong overall positive charge. A patch surrounding the active site cysteine of SCE binds directly to the  $\psi$ KXE\textbackslash D consensus sequence in the substrate ( $\psi$  is a large hydrophobic amino acid, generally isoleucine, leucine or valine; K is the lysine residue that is modified; X is any residue; E is glutamic acid and D is aspartic acid). In addition to the consensus, other properties of the substrate protein sequence appear necessary. For instance, X-ray structure data of an SCE-substrate complex indicate that in

order to specifically attract SUMO, this consensus sequence has to be positioned in a large and accessible loop (Bernier-Villamor *et al.* 2002). Apart from sumoylation at the consensus sites, an increasing number of examples are found where the sumoylated lysine residues are not positioned in a canonical consensus sequence. These sumoylation events are prime candidates for *in vivo* dependence on SUMO ligases.

### 1.4.3 SUMO ligases

SUMO ligases are enzymes that bind, directly or indirectly, specific protein substrates and promote the transfer of SUMO from a thiolester intermediate to amide linkages with proteins. While SUMO-activating enzyme (SAE) and SUMO-conjugating enzyme (SCE) were shown to be sufficient for sumoylation of various substrates *in vitro*, recent studies have demonstrated that *in vivo* most of the SUMO conjugating reactions require SUMO ligase activity. So far, three types of SUMO ligases have been identified in animals and fungi, namely SIZ\PIAS, RanBP2 and Pc2, all of which interact with SCE and enhance sumoylation both *in vivo* and *in vitro*.

The SIZ group (prototype members are SIZ1 and SIZ2\NF11 of *Saccharomyces cerevisiae*, and the PIAS family of animals) is similar to the major class of ubiquitin ligases in that it uses a RING-like domain for binding of the SCE-SUMO complex (Johnson and Gupta 2001; Kahyo *et al.* 2001). *Arabidopsis* homologs of this group of ligases are listed in Table 1. The second type RanBP2 (Ran-binding protein 2; Pichler *et al.* 2002), does not display any sequence relation to SIZ-type SUMO ligases or ubiquitin ligases. RanBP2 is located at the cytoplasmic filaments of the nuclear pore complex, where it interacts with sumoylated RanGAP and the GTPase. RanGAP was itself one of the SUMO targets to be identified (Saitoh *et al.* 1998). This type of SUMO ligase is probably restricted to animals, because its prominent substrate RanGAP is apparently not sumoylated in fungi, and a similar situation may hold in plants. In particular, the SUMO acceptor domain is lacking in plant RanGAP (Rose and Meier, 2001).

The polycomb group (PcG) protein Pc2 was reported to be the third type of SUMO ligase which is structurally unrelated to SIZ\PIAS or RanBP2 and to ubiquitin ligases (Kagey *et al.* 2003). PcG proteins form large multimeric complexes, which are detectable microscopically

as discrete foci, called PcG bodies within the cell. It is difficult to identify candidate ligases of this type in *Arabidopsis* because a precise definition of the subdomain(s) involved in

Yeast Homologs	Name	BAC locus	Chr. locus	GB accession (protein)	Chromosome Map-view	Domain Architecture	
<b>SUMO</b>							
SMT3p / Q12306	SUM1	F10M23	At4g26840	NP_194414			
	SUM2	MCO15	At5g55160	NP_200327			
	SUM3	MCO15	At5g55170	NP_200328			
	SUM4	K24G6	At5g48710	NP_199682			
	SUM5	F24L7	At2g32765	NP_565752			
	SUM6	K24G6	At5g48700	NP_199681			
	SUM7	MWJ3	At5g55855*	NA			
	SUM8	MWJ3	At5g5585x*	NA			
	SUM9	F5110	At4g0027x*	NA		NA	
<b>SUMO-like domain protein</b>							
Esc2p / NP_010650		T22E19	At1g68185	NP_564924			
<b>SUMO activating enzyme</b>							
Aos1p / NP_015506	SAE1a	F13M23	At4g24940	NP_567712			
	SAE1b	MBA10/ MFB16	At5g50580/ At5g50680	NP_568741			
Uba2p / NP_010678	SAE2	f3k23	At2g21470	NP_179742			
<b>SUMO conjugating enzyme</b>							
Ubc9p / NP_010219	SCE1a	T10K17	At3g57870	NP_191346			
	SCE1b	T1E22	At5g02240*	NA		NA	
<b>SUMO ligase candidates</b>							
Nfi1p / Q12216	SIZ1	MUF9	At5g60410	NP_200849			
	PIASlike1	F7G19	At1g08910	NP_172366			
	PIASlike2	MBK23	At5g41580	NP_198973			
	Ubiquitin-like (PF00240)			Ubiquitin-like (subsignif)			ThiF (PF00899)
	MoeB-similar			UBACT (PF02134)			UQ_con (PF00179)
	SAP (PF02037)			PHD (PF00628)			zf-MIZ (PF02891)

**Table 1. Listing and sequence characterization of *Arabidopsis thaliana* SUMO-related proteins and predicted proteins involved in SUMO conjugation** (taken from Navatchkova M., Budhiraja R., Coupland G., Eisenhaber F., Bachmair A., 2004). Proteins of a potential orthology relationship in *Saccharomyces cerevisiae* and *Arabidopsis thaliana* have been identified in an approach as reciprocal best hits in these two proteomes (Tatusov *et al.* 2003). Further *Arabidopsis thaliana* genome searches have been used to confirm the completeness of the test. The genomic map view has been derived from the NCBI Mapviewer (Wheeler *et al.* 2004). Domain architectures have been determined using Conserved Domain Database (CDD) queries (Wheeler *et al.* 2004). Abbreviations: Chr.- Chromosome, NA - not available.

sumoylation is not available. Furthermore, similarity of *Arabidopsis* proteins to domains common to all polycomb members may be insufficient to define the functional homologs of PcG because most polycomb proteins have no known SUMO ligase activity.

#### **1.4.4 SUMO proteases - desumoylation**

The cleavage of SUMO from its target proteins is catalyzed by SUMO proteases. SUMO cleaving enzymes have at least two functions in sumoylation. They remove SUMO from proteins, making the modification reversible, and also provide a source of free SUMO to be used for conjugation to other proteins. All known SUMO-cleaving enzymes contain a ~ 200 amino acid domain (Ulp domain), which has the SUMO cleaving activity (Mossesso and Lima, 2000). The Ulp domain does not share sequence similarity with the enzymes that cleave ubiquitin. Instead, it is distantly related to number of viral proteases (Li and Hochstrasser, 1999).

Two desumoylating enzymes Ulp1 and Ulp2 (ubiquitin-like modifier proteases) have been identified in baker's yeast. *In vitro*, both Ulp1 and Ulp2 can catalyze the carboxyl terminus processing of SUMO, and both enzymes can remove SUMO from isopeptide-linked conjugates (Li and Hochstrasser, 2000). Seven genes in mammalian genomes encode proteins with Ulp domains. These proteins are called SENPs (sentrin proteases). The homologs of SUMO proteases in *Arabidopsis* are called AtULPs (Kurepa *et al.* 2003) Not all members of the SENP group are specific for SUMO. For instance, SENP8 was found to cleave at the carboxyl terminus of the small ubiquitin-like modifier NEDD8 (Mendoza *et al.* 2003). NEDD8 is called RUB 1 in most organisms including *Arabidopsis* (Rao-Naik *et al.* 1998). In plants, the enzyme specificity is even more difficult to evaluate since data base searches have identified at least 67 genes in *Arabidopsis* with similarity to the SUMO protease domain. Thus, there has been a huge expansion in this class of protease, and it is unlikely that all of them are specific for SUMO. However, one of these protease genes, ESD4 (early in short days 4), has been functionally characterized to encode a SUMO protease (Murtas *et al.* 2003). Subcellular localization has been identified as a critical aspect of SUMO protease function (Huang and Dasso, 2000; Li and Hochstrasser, 2003). A recent report that the plant SUMO protease ESD4 specifically localizes

to the nuclear periphery suggests a similar situation in plants (Murtas *et al* 2003). Table 2 lists seven likely candidates, and groups five genes with yeast proteases Ulp1 and Ulp2. AtUlp1c and AtUlp1d (Kurepa *et al.* 2003) are close homologs, located in segmentally duplicated regions of chromosome 1.

Yeast Homologs	Name	Chr. locus	Domain Architecture
<b>SUMO cleaving protease candidates</b>			
ULP1p / Q02724 	<b>ESD4</b>	<b>At4g15880</b>	
	<b>ULP1a/EL1</b>	<b>At3g06910</b>	
	ULP1b	At4g00690	
ULP2p / P40537 	ULP2like1	At4g33620	
	ULP2like2	At1g09730	
	ULP1c	At1g10570	
	ULP1d	At1g60220	
	SENPl like1	At5g60190	
	 Peptidase_C48 (PF02902)		

**Table 2. SUMO-specific proteases in *Arabidopsis* and related proteins in yeast.**

The prototype enzymes are ULP1 and ULP2 from baker's yeast. Animal enzymes are called SENPs (Sentrin proteases). *Arabidopsis thaliana* has at least 67 genes with similarity to the SUMO-specific protease domain.

### 1.5 Substrate specificity in sumoylation

SUMO is conjugated to most target proteins at a lysine residue in a consensus sequence  $\psi$ KXE\textbackslash D but there are other determinants involved in target selection as well. The glutamic acid is the most highly conserved position other than lysine. In some cases, a glutamine to aspartic acid mutation in the consensus sequence significantly reduces sumoylation (Sapetschnig *et al.* 2002), although a few sequences are sumoylated (Johnson and Blobel, 1999). The  $\psi$ KXE\textbackslash D motif can directly bind the SUMO-conjugating enzyme (Sampson *et al.* 2001). This direct interaction explains why so many

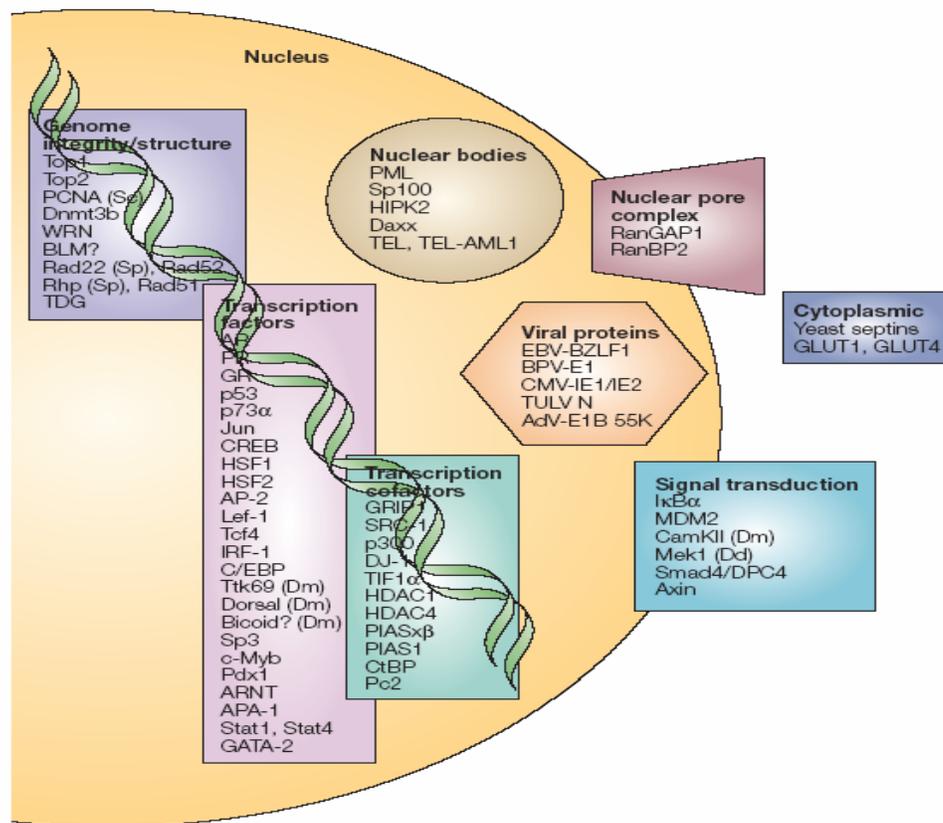
sumoylation substrates have been identified via their interaction with SCE in the yeast two hybrid screen. Furthermore, it accounts for SAE and SCE together being sufficient to sumoylate many proteins at the consensus motif *in vitro* in the absence of SUMO ligases. Remarkably, a  $\psi$ KXEVD sequence and nuclear localization sequence (NLS) are sufficient to target an artificial substrate for sumoylation, indicating that the requirements for SUMO conjugation can be very simple (Rodriguez *et al.* 2001). The  $\psi$ KXEVD motif is very short and found in many proteins, most of which are not targeted by SUMO. For example, of the 5884 open reading frames (ORFs) in *Saccharomyces cerevisiae*, there are 2799 sequences of the form  $\psi$ KXE distributed in 1931 different ORFs. Thus, in addition to the consensus motif, other interactions between the substrate and SCE may be relevant for substrate selection. For instance, the crystal structure of the RanGAP1-SCE complex (substrate-SCE complex) indicates that, in order to specifically attract SUMO, this consensus sequence has to be positioned in a large and accessible loop (Bernier-Villamor *et al.* 2002). Several proteins are also modified at sites other than  $\psi$ KXE. The replication processivity factor PCNA has two sumoylation sites, one conforming to the consensus sequence and other at a TKET sequence (Hoegge *et al.* 2002). Other reports confirm sumoylated proteins TEL, PML, Smad4 and the Epstein barr virus BZLF1 protein to have sumoylation sites at sequences TKED, AKCP, VKYC and VKFT, respectively (Adamson *et al.* 2001; Kamitani *et al.* 1998; Lin *et al.* 2003; Rui *et al.* 2002; Chakrabarti *et al.* 2000). Moreover, some sumoylated proteins such as Mdm2, Daxx, CREB, and CTBP2 do not contain a  $\psi$ KXE sequence, others are still sumoylated when all consensus sites are mutated (Kagey *et al.* 2003; Miyauchi *et al.* 2002; Jang *et al.* 2002; Rangasamy *et al.* 2000; Xirodimas *et al.* 2002, Comerford *et al.* 2003). It is not clear how these nonconsensus sites are recognized.

## **1.6 Substrates and functions of SUMO protein modification**

Since the identification of the first SUMO modified protein, RanGAP in 1996 (Matunis *et al.* 1996), a large number of proteins have been shown to be post-translationally modified by SUMO, and new substrates of SUMO modification continue to be identified at a rapid pace. Some of the proteins known to be modified by SUMO are listed in Fig 5.

Many of the known targets of SUMO are mammalian proteins involved in signal transduction and transcriptional regulation. Others are involved in DNA damage repair (Hoegel *et al.* 2002), chromosome segregation (Khodairy *et al.* 1995), blocking ubiquitin-mediated events (Johnson 2004), etc. In this last capacity for example, I $\kappa$ B, the inhibitor of the nuclear factor  $\kappa$ B transcriptional activator, can be modified by both SUMO and ubiquitin to regulate its half life (Desterro *et al.* 1998). To date, in plants there are no specific proteins reported to be modified by SUMO, but there are possibly many plant SUMO targets that await identification. However, there have been studies and evidence that support a role of SUMO during response of plants to various environmental stresses. When *Arabidopsis* is exposed to heat shock, hydrogen peroxide, ethanol or the amino acid analogue canavanine, the levels of SUMO1\2 conjugates are significantly and reversibly increased (Kurepa *et al.* 2003). *Arabidopsis* plants overexpressing SUMO1 were subsequently shown to be less sensitive to the hormone abscisic acid, further supporting a role for sumoylation in stress response (Lois *et al.* 2003). Several species of pathogenic bacteria for both plants and animals have been recently reported to secrete proteases related to ULPs into the host upon infection, with this secretion attenuating the accumulation of SUMO conjugates *in planta* (Orth *et al.* 2000; Hotson *et al.* 2003). The recent discovery that the *Arabidopsis* ESD4 locus, which promotes early flowering in short days when mutated, encodes a nuclear localized protein related to Ulp family of SUMO specific protease suggests that desumoylation of conjugates may also be important to floral induction (Murtas *et al.* 2003).

What appears to emerge from the analysis of SUMO targets is that their modification alters their activities, their ability to interact with other proteins, or their subcellular localization. An *in vitro* sumoylation system based upon purified proteins of *E. coli* extracts expressing a complete set of enzymes may help to understand how the biochemical properties of a given substrate are changed by sumoylation (Uchimura *et al.* 2004).



**Fig. 5 SUMO modified proteins** (taken from Seeler *et al.* 2003). Some of the many proteins that have been found to be post-translationally modified by SUMO are depicted in the diagram by function/localization. Many SUMO modified proteins function in regulation of transcription, chromatin structure, maintenance of genome and signal transduction. All proteins are of mammalian origin unless specifically indicated by *Saccharomyces cerevisiae* (Sc), *Shizosaccharomyces pombe* (Sp), *Drosophila* (Dm), and *Dictyostelium* (Dd).

### 1.7 Strategies to analyze SUMO conjugation in *Arabidopsis thaliana*

The use of biochemical protocols and molecular genetics in the model plant *Arabidopsis thaliana* has proved a powerful approach to study the SUMO conjugation pathway. *Arabidopsis thaliana*, also known as Mouse-ear or Thale cress, is a member of the

*Brassicaceae* family, and was chosen as a model organism for many reasons. The major advantages for experimental purpose are its small size, short generation time, diploid nature, a small genome (125 Mb) and ability to produce a large number of seeds (Schmidt 1995). Furthermore, after years of research by scientific groups worldwide, thousands of mutants and research material are available. In addition, the entire genome of this non-agronomically relevant plant has been sequenced and information of different metabolic, regulatory pathways and other specific processes is accessible.

This study employs two different strategies to analyze the SUMO conjugation system of *Arabidopsis*. The first approach is aimed at purification and identification of proteins that are sumoylated in this plant. The second strategy to gain more insights into the process of sumoylation in *Arabidopsis*, is by interfering with the genetic components of the sumoylation machinery. It is aimed at manipulation of one or more genes involved in the sumoylation pathway. This may hint at the relevance of sumoylation in growth and development of *Arabidopsis thaliana*.

## 2. MATERIAL AND METHODS

### 2.1 Material

#### 2.1.1 Plant plasmid constructs

The *Arabidopsis thaliana* plants and their ecotypes used in this study are listed in Table 3.

CONSTRUCT	ECOTYPE	DESCRIPTION	OBTAINED FROM
Wild-type	Col	Wild type plants without transgene	Andreas Bachmair
Wild-type	Ws	Wild type plants without transgene	Christian Hardtke
pHi SUMO1GG	Col	SUMO1 transgene with double glycine amino acid residues at the carboxyl terminus; 3x HA and His <sub>6</sub> affinity tags at the amino terminus; constitutive expression	Andreas Bachmair
pHi SUMO1AA	Col	SUMO1 transgene with double alanine amino acid residues at the carboxyl terminus; 3x HA and His <sub>6</sub> affinity tags at the amino terminus; constitutive expression	Andreas Bachmair
pHi SUMO1AG	Col	SUMO1 transgene with alanine-glycine amino acid residues at the carboxyl terminus; 3x HA and His <sub>6</sub> affinity tags at the amino terminus; constitutive expression	Andreas Bachmair
pHi SUMO1GA	Col	SUMO1 transgene with glycine-alanine amino acid residues at the carboxyl terminus; 3x HA and His <sub>6</sub> affinity tags at the amino terminus; constitutive expression	Andreas Bachmair

<b>CONSTRUCT</b>	<b>ECOTYPE</b>	<b>DESCRIPTION</b>	<b>OBTAINED FROM</b>
pHi SUMO1Stop	Col	SUMO1 transgene with deletion of the last two amino acid residues at the carboxyl terminus; 3x HA and His <sub>6</sub> affinity tags at the amino terminus; constitutive expression	Andreas Bachmair
pER8 tag3 SUMO1	Col	SUMO1 transgene with Strep, 3xHA, His <sub>8</sub> affinity tags at the amino terminus; inducible expression (estradiol)	This work
pER8 tag3 SUMO3	Col	SUMO3 transgene with Strep, 3xHA, His <sub>8</sub> affinity tags at the amino terminus; inducible expression (estradiol)	This work
pER8 tag3 SUMO5	Col	SUMO5 transgene with Strep, 3xHA, His <sub>8</sub> affinity tags at the amino terminus; inducible expression (estradiol)	This work
pER8 tag3 SUMO1 GG(Q93A)	Col	SUMO1 GG transgene with substitution of glutamine at position 93 with alanine; Strep, 3xHA, His <sub>8</sub> affinity tags at the amino terminus; inducible expression (estradiol)	This work
pHi SUMO6	Col	SUMO6 transgene with introns and Strep, His <sub>6</sub> , 3xHA affinity tags at the amino terminus; constitutive expression	Andreas Bachmair

<b>CONSTRUCT</b>	<b>ECOTYPE</b>	<b>DESCRIPTION</b>	<b>OBTAINED FROM</b>
pHi SUMO7	Col	SUMO7 transgene with introns and Strep, His <sub>6</sub> , 3x HA affinity tags at the amino terminus; constitutive expression	Andreas Bachmair
pHi SUMO7v	Col	SUMO7 transgene with introns and three amino acid insertion ELQ at the second intron position; Strep, His <sub>6</sub> and 3x HA affinity tags at the amino terminus; constitutive expression	Andreas Bachmair
pHi SUMO1 (Q93A)	Col	SUMO1 transgene with substitution of glutamine at position 93 with alanine; 3xHA, His <sub>6</sub> affinity tags at the amino-terminus; constitutive expression	This work
pHi SUMO1 (Q93D)	Col	SUMO1 transgene with substitution of glutamine at position 93 with aspartic acid; 3xHA, His <sub>6</sub> affinity tags at the amino terminus; constitutive expression	This work
pHi SUMO1 (Q93L)	Col	SUMO1 transgene with substitution of glutamine at position 93 with leucine; 3xHA, His <sub>6</sub> affinity tags at the amino terminus; constitutive expression	This work
pHi SUMO1 (Q93R)	Col	SUMO1 transgene with substitution of glutamine at position 93 with arginine; 3xHA, His <sub>6</sub> affinity tags at the amino-terminus; constitutive expression	This work
35S: SCE	Ws	Constitutively overexpressed SCE	Christian Hardtke

<b>CONSTRUCT</b>	<b>ECOTYPE</b>	<b>DESCRIPTION</b>	<b>OBTAINED FROM</b>
35S: SCE (C94S)	Ws	Constitutively over expressed SCE with a mutated active site cysteine at position 94 in the gene to serine,	Christian Hardtke

**Table 3. Listing of the *Arabidopsis thaliana* plants used in this study**

### 2.1.2 Bacterial plasmid constructs

The bacterial plasmid constructs employed in this study are listed in Table 4.

<b>CONSTRUCT</b>	<b>DESCRIPTION</b>	<b>OBTAINED FROM</b>
pET-9d tag 3 SUMO1	SUMO1 transgene with Strep, 3xHA, His <sub>8</sub> affinity tags at the amino terminus	This work
pET-9d tag 3 SUMO3	SUMO3 transgene with Strep, 3xHA, His <sub>8</sub> affinity tags at the amino terminus	This work
pET-9d tag 3 SUMO5	SUMO5 transgene with Strep, 3xHA, His <sub>8</sub> affinity tags at the amino terminus	This work
pET-9d tag 3 SUMO6	SUMO6 transgene with introns and containing Strep, 3xHA, His <sub>8</sub> affinity tags at the amino terminus	This work
pET-9d tag 3 SUMO7	SUMO7 transgene with introns and containing Strep, 3xHA, His <sub>8</sub> affinity tags at the amino terminus	This work

CONSTRUCT	DESCRIPTION	OBTAINED FROM
pET-9d tag 3 SUMO7v	SUMO7v transgene with introns and the three amino acid insertion in the second intron; Strep, 3xHA, His <sub>8</sub> affinity tags at the amino-terminus	This work
pACYC177-SAE1a	Smaller SAE subunit with His <sub>6</sub> affinity tag	This work
pDEST-17 SAE2	Larger SAE subunit with His <sub>6</sub> affinity tag	Yong-Fu Fu
pSK SCE	SCE	Andreas Bachmair
pQE-30 SCE	SCE with His <sub>6</sub> affinity tag	This work
pET-9d SCE	SCE in <i>E. coli</i> expression vector pET-9d	This work

**Table 4. Listing of bacterial plasmid constructs used in this study**

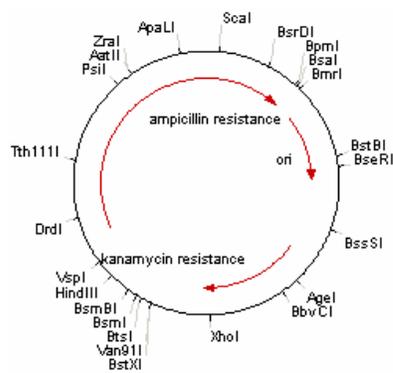
### 2.1.3 Plasmid vectors used

#### pACYC-177 (New England Biolabs)

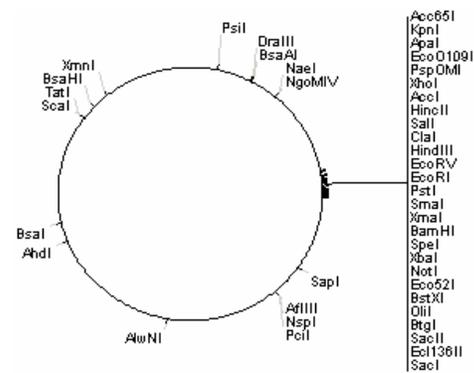
pACYC-177 is an *E. coli* plasmid cloning vector containing the p15A origin of replication. This allows pACYC-177 to coexist in cells with plasmids of the ColE1 compatibility group. It is a low copy number vector, at about 15 copies per cell. Selection markers are ampicillin and kanamycin resistance. The vector is about 3941 bp (Fig. 6).

pBluescript II (Stratagene)

This vector was used for routine cloning procedures. In addition to its large polylinker, it allows  $\alpha$ -complementation of an N-terminal deletion of lac-Z (beta galactosidase gene). Different orientations of the polylinker are available; the selectable marker is ampicillin resistance. The vector is about 2961 bp in size (Fig. 7).



**Fig. 6 Schematic drawing of pACYC-177**  
Representative restriction sites are shown.  
Selective marker is amp<sup>r</sup> and kan<sup>r</sup>  
Vector size is 3941 bp



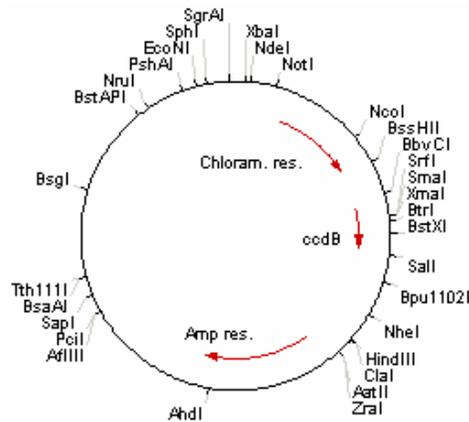
**Fig. 7 Schematic drawing of pBIKS**  
Representative sites are shown  
Selective marker is amp<sup>r</sup>  
Vector size is 2961 bp

pDEST-17 (Invitrogen)

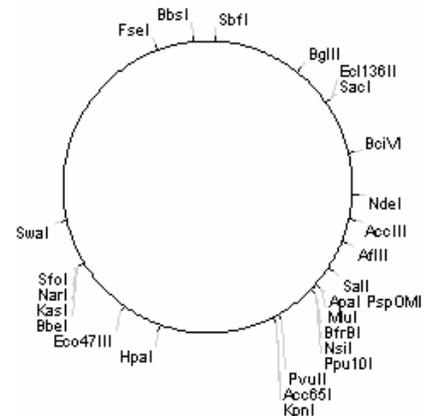
Cloning and overexpression of *Arabidopsis* SAE1 in *E. coli* was done in pDEST-17 expression vector. It allows easy purification of the fusion protein due to the incorporated His<sub>6</sub> sequence. The promoter used contains the *E. coli* phage T7 promoter. It originally contains two selection markers, namely chloramphenicol and ampicillin resistance. However, chloramphenicol resistance is lost when a sequence is inserted by recombination. This vector has a size of 6354 bp (Fig. 8).

pER-8 (Zuo et al. 2000)

pER-8 is a binary T-DNA cloning vector for expression under the control of an estradiol inducible promoter. It is 11784 bp pairs in size. Selectable markers for the vector are spectinomycin (bacteria) and hygromycin (plants) resistance. Both *Stu*I sites (in MCS and ER region) are blocked by *dcm* methylation (Fig. 9).



**Fig. 8 Schematic drawing of pDEST-17**  
Relevant restriction sites are shown.  
Selective markers are *chl*<sup>r</sup> and *amp*<sup>r</sup>  
Vector size is 6354 bp



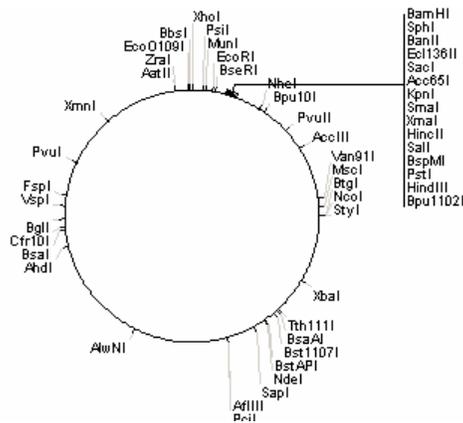
**Fig. 9 Schematic drawing of pER-8**  
Relevant restriction sites are depicted  
Vector size is 11784 bp

pET-9a (Novagen)

The pET-9a expression vector carries a N-terminal T7 tag sequence and a *Bam*HI cloning site. Unique sites have been shown on the circle map (Fig. 10) with pBR322 origin of replication. The promoter used contains the *E. coli* phage T7 promoter. Selective marker is kanamycin resistance and the size of the vector is 4341 bp.



This allows induction of expression using IPTG. At the 3' end of these regulatory sequences, a multiple cloning site can be found. The cassette is terminated by stop codons in all the 3 ORFs and by  $t_0$  transcriptional terminator from phage lambda. The selective marker used is ampicillin resistance (Fig. 12).



**Fig. 12 Schematic drawing of pQE-30**  
 Representative restriction sites are shown.  
 Selective marker is  $amp^r$   
 Vector size is 3462 bp

## 2.1.4 Oligonucleotides used

### SUMO isoforms

SUMO1:

dn: CTAGCCATGGCTCATCATCATCACCATCATATGTCTGCAAACCAGGAGGA  
 up: GATCGGTACCGAGTAGTAGTAGTGGTAGTATACAGACGTTTGGTCCTCCT

SUMO midup: GACTTTGAGATTGATGTGAGCTCCT

SUMO mid dn: GTTCATCCTGCCCCGTCTCCGGCT

SUMO3:

dn: GCCGGTACCATGTCTAACCCTCAAGATGACAAGCCCATC

up: TCGTCTAGATTCAACCACCACTCATCGCCCGGCACGCATCTATCACATC

SUMO5:

dn: GCCGGTACCATGGTCAGTTCCACAGACACAATCTCTGCT

up: TGCTCTAGATTCAGCCACCACCAAGTTCCATGACCATACAGATCTCATC

SUMO6:

dn: GCCGGTACCATGTCAACGAAGAGCAGTAGTATTCATGGAA

up: TGCTCTAGATTCAACCACTTTCTTGAGGCAACAATGCATCGATTTGATC

SUMO7:

dn: GCCGGTACCATGTCCGGCAGCTGACAAAAAACCGTTGATT

up: TGCTCTAGATTCACCTGCTATTTGGTCAACAAATGCATCGATTCATC

C-terminal SUMO variants

GlyGly: CTAGATCAGCCACCAGTCTG

AlaGly: CTAGATCAGCCAGCAGTCTG

GlyAla: CTAGATCAGGCACCAGTCTG

AlaAla: CTAGATCAGGCAGCAGTCTG

Stop: CTAGATCATTA ACTCTG

SUMO variants potentially inhibiting deconjugation

SUMO (Q93A) top: CGATGCGATGCTCCATGCTACTGGTGGCTGATT

SUMO (Q93A) bottom: CTAGAATCAGCCACCAGTAGCATGGAGCATCGCAT

SUMO (Q93D) top: CGATGCGATGCTCCATGATACTGGTGGCTGATT

SUMO (Q93D) bottom: CTAGAATCAGCCACCAGTATCATGGAGCATCGCAT

SUMO (Q93L) top: CGATGCGATGCTCCATTTGACTGGTGGCTGATT

SUMO (Q93L) bottom: CTAGAATCAGCCACCAGTCAAATGGAGCATCGCAT

SUMO (Q93R) top: CGATGCGATGCTCCATAGAACTGGTGGCTGATT

SUMO (Q93R) bottom: CTAGAATCAGCCACCAGTTCTATGGAGCATCGCAT

SUMO activating enzyme

SAE2B dn:

CTGTACACCATGGAGATCTGAGCCTGCTTCTAAGAAGAGAAGACT

SAE link up:

GAGTCTATCTCCGTCCATGGCACCATGGTGTATGATGGTGTATGGGTCATTATTC  
AACTCTTATCTTCTT

SAE2B dn2:

CCATGGTGTACAGGCCAGATCTGAGCCTGCTTCTAAGAAGAGAAGACT

SAE link up2:

GAGCTCATCTCCGTCCATGGCACCATGGTGTATGATGGTGTATGGGTCATTATTC  
AACTCTTATCTTCTT

SUMO conjugating enzyme

SCE 1dn: CGTACCATGGCTAGTGGAATCGCTCGTGGTC

SCE 1up: TCCCCGGGTTAGACAAGAGCAGGATACTGCTTGGACT

Affinity tags

Strep SUMO top: CATGGCTTGGTCTCATCCACAATTCGAAAAGGG

Strep SUMO bottom: CATGCCCTTTTCGAATTGTGGATGAGACCAAGC

### **2.1.5 Bacterial strains**

*Escherichia coli*

XL1-blue (Stratagene) was used for all routine *E. coli* transformations.

Genotype: F' :: Tn10 *proA*<sup>+</sup>*B*<sup>+</sup> *lacI*q Δ(*lacZ*) M15/*recA1endA1 gyrA96* (Nal<sup>r</sup>) *thi hsdR17*  
(r<sub>K</sub><sup>-</sup>m<sub>K</sub><sup>+</sup>) *glnV44 relA1 lac*

BL21 (DE3) or BL21 (DE3) pLysS (Novagen) was used for transformation with expression vectors carrying the insert.

*Agrobacterium tumefaciens*

C58C1-pCV 2260 strain (selection with 50μg/ml Kanamycin and 100μg/ml Rifampicin) was used for transformation.

### **2.1.6 Plant and bacterial growth media**

#### **Plant growth medium**

Germination medium: 4.33g/l Murashige and Skoog salt mixture, 1% (w/v) sucrose,

100 $\mu$ g/ml inositol, 1 $\mu$ g/ml thiamine, 0.5 $\mu$ g/ml pyridoxine, 0.5 $\mu$ g/ml nicotinic acid, 0.5mg/ml 2-[N-Morpholino]-ethanesulphonic acid (MES), 0.9%(w/v) agar, pH 5.7

### **Bacterial growth medium**

#### *Escherichia coli*

2xTY (liquid medium): 1.6% (w/v) Bacto tryptone (Difco), 1% (w/v) Bacto yeast extract (Difco), 0.5 % (w/v) NaCl, pH 7.5

2xTY (solid medium): As above with 1.5% Bacto agar (Difco)

#### *Agrobacterium tumefaciens*

YEB (liquid medium: 0.5%(w/v) Bacto beef extract (Difco), 0.1% (w/v) Bacto yeast extract (Difco), 0.5% (w/v) Bacto Peptone (Difco), 0.5% (w/v) Sucrose, pH 7.2

Supplemented with 10mM MgSO<sub>4</sub> after autoclaving.

YEB (solid medium): As above with 1.5% Bacto agar (Difco)

### **2.1.7 Chemicals and Enzymes**

Unless otherwise stated, all chemicals were supplied by Sigma, Merck, Roth, Invitrogen or Duchefa. Enzymes were supplied by New England Biolabs, Fermentas or Roche.

### **2.1.8 Antibiotics**

<u>Bacterial selection</u>	<u>Working concentration</u>	<u>Solvent</u>
Ampicillin (Amp)	50 $\mu$ g/ml	dissolved as powder in autoclaved medium
Kanamycin (Kan)	50 $\mu$ g/ml	water
Rifampicin (Rif)	50 $\mu$ g/ml	water
Chloramphenicol (Chl)	50 $\mu$ g/ml	ethanol
<u>Plant selection</u>		
Kanamycin	25 $\mu$ g/ml	water
Hygromycin	25 $\mu$ g/ml	supplied as liquid
Gentamycin	50 $\mu$ g/ml	water
Cefotaxime Na-salt	200 $\mu$ g/ml	water

All antibiotics were supplied by Duchefa or Sigma and antibiotics were filtered sterilized before use; except hygromycin (Roche, supplied sterile).

### **2.1.9 Herbicide**

Bialaphos (Basta): active ingredient: glucofosinate-ammonium  
Working concentration: 100mg/l

## **2.2 METHODS**

### **2.2.1.1 Isolation of Plasmid DNA from *E. coli* (Alkaline lysis method)**

A single bacterial colony was inoculated in 3ml of 2xTY (liquid medium) containing appropriate antibiotics, and grown overnight at 37°C. 1.5ml of the culture was transferred to an Eppendorf tube and the bacterial cells were pelleted by centrifugation at 14,000 rpm for two minutes in a microfuge (Eppendorf). The pellet was resuspended in 100µl of GTE buffer (50mM glucose, 25mM Tris.Cl pH 8.0, 10mM EDTA) and mixed properly using a vortex shaker. The resuspended cells were incubated for 5-10 minutes at room temperature. 200µl of freshly prepared alkaline-SDS solution (0.2N NaOH, 1% SDS) was added, mixed slowly by inverting the tube five to six times, and incubated on ice for 5 minutes. Following incubation on ice, 150µl of acetate solution (5M acetic acid is added to 5M potassium acetate solution until pH is ~ 4.8) was added to the tubes, which were incubated for 15 minutes on ice. The samples were then centrifuged at 14,000 rpm for 5 minutes at 4°C to remove the cell debris. The supernatant was transferred to an Eppendorf tube and the DNA was ethanol precipitated by addition of 1ml ethanol (100% ethanol; RT). The DNA pellet was dissolved in 20-30µl dH<sub>2</sub>O. One volume of 5M NH<sub>4</sub>OAc was added to the above dissolved DNA pellet, incubated for five minutes at room temperature and centrifuged at 14,000 rpm for 5 minutes (RT). The supernatant was transferred to a new Eppendorf tube and 0.6 volumes of isopropanol were added and the contents of the tube were mixed gently. Following 10 minute incubation at RT, the isopropanol-precipitated DNA was centrifuged for 10 minutes at 4°C. The DNA pellet was washed with 70% ethanol (cold), re-pelleted and dried in vacuum. The dried DNA pellet was finally dissolved in 20-30µl of dH<sub>2</sub>O or TE buffer (10mM Tris Cl, 1mM EDTA).

### **2.2.1.2 Treatment of DNA with restriction endonucleases**

Cleavage of DNA was performed using restriction enzymes. Digestions were performed in a 1.5ml Eppendorf tube in a total volume of 20 $\mu$ l. Typically, 0.5- 1.0 $\mu$ g of DNA was digested in a buffer at a temperature according to manufactures instructions. The length of the digestion varied from 4 hrs to overnight.

### **2.2.1.3 Running of agarose gels and purification of DNA fragments**

Separation of digested DNA products was performed by electrophoresis using agarose gels. The density of the gels used depended on the product length. DNA samples were prepared for loading by the addition of 10x loading dye (50% glycerol, 0.2M EDTA, pH 8.3, 0.05% w/v Orange G and dH<sub>2</sub>O). To prepare agarose gel for electrophoresis, agarose and 1x TAE (0.04M Tris acetate, 0.001M EDTA) were mixed to a relevant concentration (typically 0.8 % to 1.5% w/v). The solution was heated in a microwave oven until boiling. It was allowed to cool to approximately 50°C and ethidium bromide was added. The solution was poured in a gel mold and allowed to set. At this stage it was placed in a gel tank (BioRad) containing 1x TAE buffer. Following loading, the samples were electrophoresed at 2-20V per cm. Subsequently, the DNA products were visualized in a UV trans-illuminator and the DNA fragments of interest were excised from the gel under UV light with a scalpel. The DNA product was then purified from the gel by use of the Nucleospin Extract II gel extraction kit (Macherey-Nagel). Concentrations of the purified products were determined by running alongside a 100bp ladder (NEB) of known concentration on an agarose gel.

### **2.2.1.4 DNA ligation**

The ligation of DNA inserts into the plasmid vectors was carried out overnight at 16°C using 1 unit of T4 DNA ligase (Roche) and 1x ligase buffer.

### **2.2.1.5 Transformation of *E. coli* (Heat shock method)**

50µl of XL-1 blue competent cells were thawed on ice in an Eppendorf tube. 1-10ng of plasmid DNA was added to the tube and mixed gently by flicking the tube. The sample was then incubated on ice for 30 minutes. Cells were subjected to heat shock at 37°C for 2 minutes and immediately chilled on ice for 30 seconds. Following chilling, 950µl of 2xTY (liquid medium) was added to the cells and incubated for one hour at 37°C. Cells were then plated onto 2xTY-agar containing appropriate antibiotics.

### **2.2.1.6 Growth and transformation of *Agrobacterium tumefaciens***

The transformation technique in *A. tumefaciens* is related to *E. coli* in that the exponentially growing cells are stressed which leads to uptake of plasmid DNA added to that culture. 2ml O/N culture of *Agrobacterium tumefaciens* strain C58C1-pCV 2260 were used to inoculate 100ml of *Agrobacterium* liquid medium (containing rifampicin and MgSO<sub>4</sub>) and incubated at 28°C for 4-5 hours (until O.D reaches 0.27). After pelleting the culture at 4°C, cells were resuspended in 1ml ice cold *Agrobacterium* medium. To the chilled 200µl aliquot, 2-5µg of T-DNA vector was added and the culture was frozen in liquid nitrogen. Thereafter, cells were thawed (at RT), and incubated at 37°C (without shaking) for 5 minutes. Following incubation, 1ml of *Agrobacterium* medium was added and the cells were incubated at 28°C (slight shaking) for two hours. Aliquots of these cells were then plated on selective medium and the plates were incubated at 28°C. Transformed colonies were visible on the plates after 2-3 days of incubation.

### **2.2.1.7 Analysis of transformed *A. tumefaciens***

In order to find out whether the binary T- DNA vector is present in *Agrobacterium*, a plasmid DNA preparation was performed. 1.5ml of *Agrobacterium* culture grown under selective conditions were pelleted, resuspended in 150µl TE containing 0.5% (v/v) sarcosyl and proteinase K (100µg/ml), and incubated at 37°C for one hour. Afterwards, 200µl alkaline SDS solution was added and the Eppendorf tube kept on ice for 5 minutes. After neutralization with 150µl acetate solution, the tube was kept on ice for additional 15 minutes and thereafter centrifuged at 4°C for 15 minutes. The supernatant was transferred

to a fresh tube and phenol: chloroform: isopropanol (PCI) extraction was performed. The DNA was then precipitated with 96% ethanol. The resulting pellet after centrifugation was washed with 70% ethanol, dried in a spin vacuum and resuspended in 30µl of H<sub>2</sub>O. One quarter of such DNA preparations was used for restriction digests. 10µg/ml of RNase A may be added to such digests. After one hour of incubation with restriction enzymes, the digested samples are ready to be analyzed on an agarose gel.

#### **2.2.1.8 *A. tumefaciens* mediated transformation of *Arabidopsis thaliana***

Plants were transformed with *Agrobacterium* using a protocol based on floral dip method (Clough *et al.* 1998). 2ml of *Agrobacterium* culture (from 20 ml starter culture grown at 28°C for 2 days) was used to inoculate 100ml of YEB medium containing appropriate antibiotics. Following an overnight incubation at 28°C with constant shaking, the cells were collected by centrifugation at 4000 rpm for 15 minutes at 4°C. The cells were resuspended in 200ml of infiltration medium (5% sucrose and 0.05% silwet L-77). Plants were submerged in this solution for 2-3 minutes. Excess liquid was removed from the plants by allowing them to dry on their sides. Plants were then transferred to the greenhouse. A plastic sheet was used to cover the plants for 1-2 days to aid infiltration.

#### **2.2.1.9 Seed sterilization and selection of transformants**

##### Bleach sterilization

Seeds (30-40) were taken in an Eppendorf tube and immersed in a solution containing 5% calcium hypochloride and 0.02% triton for 15 minutes with gentle shaking. The seeds were then rinsed 3 times with sterile water, dried in the fume hood and sprinkled onto the germination medium (GM) plates containing the appropriate selective agents. Following stratification (3-5days), the plates were placed into a long day percival cabinet or growth room. Plants that were not sensitive to the respective selective agent were removed from the plates after 15 days of growth, and transferred to soil. Following selfing and siliques formation, the plants were bagged into seed collection bags, and allowed to dry out for 2-

3 weeks. Seed sterilization was repeated and segregation was analyzed by germinating seeds on GM medium with the appropriate selective agents.

#### **2.2.1.10 Crosses**

Following bolting and flower setting, some flowers from the female plant were emasculated by removal of petals, sepals and immature anthers with fine forceps. Pollen from a mature male plant was used to dust the exposed stigma. Following the cross, any new flowers and secondary stems growing in the proximity of the developing carpel were removed. Pollinated stigmas were wrapped in plastic foil for protection and to avoid additional pollination.

#### **2.2.1.11 Application of estradiol to transgenic *Arabidopsis* plants**

A 0.5 $\mu$ M solution of estradiol was sprayed on 45 day old plants (pER8 tag 3 SUMO1; estradiol inducible promoter) for induction of the SUMO1 protein. The plants were induced overnight with estradiol and harvested in the morning for biochemical enrichment of SUMO1 conjugates.

### **2.2.2 Plant growth and analysis of flowering time**

#### Stratification and sowing

Seeds were stratified on moist filter paper in the dark at 4°C for 3-5 days to break seed dormancy. Seeds were then sown on compost composed of, 7 parts peat: 2 parts vermiculite: 1 part sand, fertilized with 1kg\m<sup>3</sup> tribon and 1kg\m<sup>3</sup> osmocote; supplemented with 12g\m<sup>3</sup> intercept (materials supplied by Blumenerdewerk Stender GmbH). Following sowing, the pots were covered with a plastic lid to maintain humidity. The covers were removed once the seeds germinated and had produced leaves.

#### Light conditions

Plants were grown either in the greenhouses or in controlled environmental chambers.

Greenhouses: Plants were grown in natural long days and short days. This was supplemented with artificial light in the winter to ensure a minimum day length of 15 hours. Light could be blocked with rolling shutters during long summer days if short-day light conditions were required.

Environmental chambers: Plants were grown in Percival controlled environment cabinets (Percival Scientific Inc., USA). For short day light conditions, cabinets were provided with 8 hrs of light. Long day light conditions comprised of 16 hrs of light.

#### Measurement of flowering time

Eight to ten plants per plant line were used to measure flowering time. The number of rosette and cauline leaves were counted before the first flower opened. Mean values from at least 8 plants were calculated (Koornneef *et al.* 1991).

### **2.2.3 Biochemical enrichment of SUMO1 conjugates from *Arabidopsis***

SUMO1 conjugates were enriched from *Arabidopsis* plants overexpressing SUMO1 by the following two methods (Fig. 13)

#### **Method 1**

This method employed the use of Ni based immobilized metal affinity chromatography followed by further enrichment using anti-HA antibody coupled to a matrix in a batch procedure.

Extraction buffer A: 6M Guanidinium chloride, 0.1M Na phosphate buffer, 0.1M Tris.Cl, pH 8

Wash buffer B: 8M urea, 0.1M Na phosphate buffer, 0.1M Tris.Cl, pH 8

Elution buffer C: 8M urea, 0.2M acetic acid

Dilution buffer D: 1% Triton, 0.1% SDS, 1mM  $\beta$ -Mercaptoethanol, Protease inhibitors 0.5-10 $\mu$ g/ml E64, 0.1-1mg/ml Pefabloc, 0.7 $\mu$ g/ml Pepstatin

200g of fresh transgenic plant material was ground in 5 volumes of buffer A. At the time of grinding, 20mM  $\beta$ -mercaptoethanol, 10mM sodium metabisulphite, 3-4% PVPP (poly

vinylpolypyrrolidone), 5% sucrose, and 5mM imidazole were added to the plant material. Grinding was done with a blender until the material blended well with the buffer into green homogenous slurry. Following grinding, the green slurry was stirred with a magnetic stirrer for 30 minutes and then centrifuged at 5000 rpm to collect the supernatant. The supernatant was collected by filtering through mira cloth and again centrifuged at 18,000 rpm for one hour. High speed centrifugation was done to pellet particulate matter which may interfere with affinity binding to the matrix. The supernatant was decanted carefully so as not to disturb the green material which settles as a loose pellet upon centrifugation. The resulting supernatant (free of particulate matter) was incubated with 8ml of Ni-NTA resin (50% slurry, Qiagen) on a rotating shaker overnight at RT (batch procedure). The following day, an empty chromatography column was packed with the overnight incubated Ni-NTA plant slurry. Subsequently, the column was washed with 5-10 column volumes of buffer B. The protein was eluted with 1-2 column volumes of elution buffer C. The eluted protein mixture was immediately neutralized with 1M Tris, pH 8 before proceeding to the next affinity purification step. In the cold room (8°C), the neutralized protein mixture was diluted 20 times with dilution buffer D, while stirring with a bar magnet and incubated overnight with 150-200µl of anti-HA affinity matrix (Roche) on a rotating shaker. An empty column was packed with the above protein mixture-anti HA matrix (cold room) and washed with dilution buffer D. Elution was done at 37°C for 15 minutes using 1-2ml of elution buffer C. The protein mixture was immediately neutralized upon elution with 1M Tris, pH 8.0. The eluted protein was concentrated with centrifugal device (Amicon) of 10,000 kDa molecular weigh cut off.

## Method 2

This method involved the use of an immobilized metal affinity chromatography in two step procedure. Crude extracts were first enriched by Ni-NTA matrix followed by the use of Dyna beads to further fish out recombinant histidine tagged proteins.

Binding and wash buffer: 8M urea, 0.1M Na phosphate buffer, 0.1M Tris.Cl, 0.01% Tween<sup>®</sup>- 20 pH 8

Elution buffer C: 8M urea, 0.2M acetic acid, 0.01% Tween<sup>®</sup> - 20

The transgenic plant material was treated in a similar manner as described in method 1, until the protein mixture was eluted from the Ni-NTA column and neutralized with 1M Tris, pH 8.0. The next step in this method employed the use of Dyna beads<sup>®</sup> TALON<sup>™</sup> (DYNAL Biotech), which are magnetizable beads developed for the isolation of recombinant histidine tagged proteins. Dyna beads were thoroughly resuspended prior to use. 50µl (2mg) of Dyna beads TALON solution was transferred to an Eppendorf tube. The tube was placed on a magnetic particle concentrator (DYNAL Biotec) until the beads had migrated to the side of the tube and the liquid was clear. The supernatant was discarded and beads were equilibrated with 700µl of binding and wash buffer. The sample was then added to the equilibrated beads and incubated on rotating shaker for 1 hour at RT (18-20°C). Thereafter, the supernatant was discarded upon separating the beads. The Dyna beads were washed four times with 700µl of binding and wash buffer, with through separation and resuspension between each washing step. The enriched protein mixture was eluted with 250µl of elution buffer upon incubating the suspension on a rotating wheel for 15 minutes at RT. The eluate was collected by separating the beads with a magnet. The protein mixture was immediately neutralized with 1M Tris, pH 8.0 upon elution and concentrated with Centricon centrifugal device prior to analysis.

#### **2.2.4 Techniques for protein analysis**

##### **2.2.4.1 Isolation of proteins from *A. thaliana***

A mini preparation technique was used to isolate proteins from *A thaliana* for Western blot analysis. A fresh leaf (~15mg) was frozen (with a pinch of sand) in liquid nitrogen. 200µl of prewarmed (5 minute at 37°C) Fergusons solution (50mM Tris.Cl pH 6.8, 4% SDS, 10 % β-Mercaptoethanol) was added and the sample was ground with help of an electric homogenizer. Following grinding, the samples were briefly centrifuged (1minute, 10,000 rpm) and the supernatant was transferred to a fresh Eppendorf tube. The supernatant was then heated at 95°C for 10 minutes and centrifuged at 14,000 rpm for 10 minutes at RT. The supernatant after centrifugation contains the total protein extract from

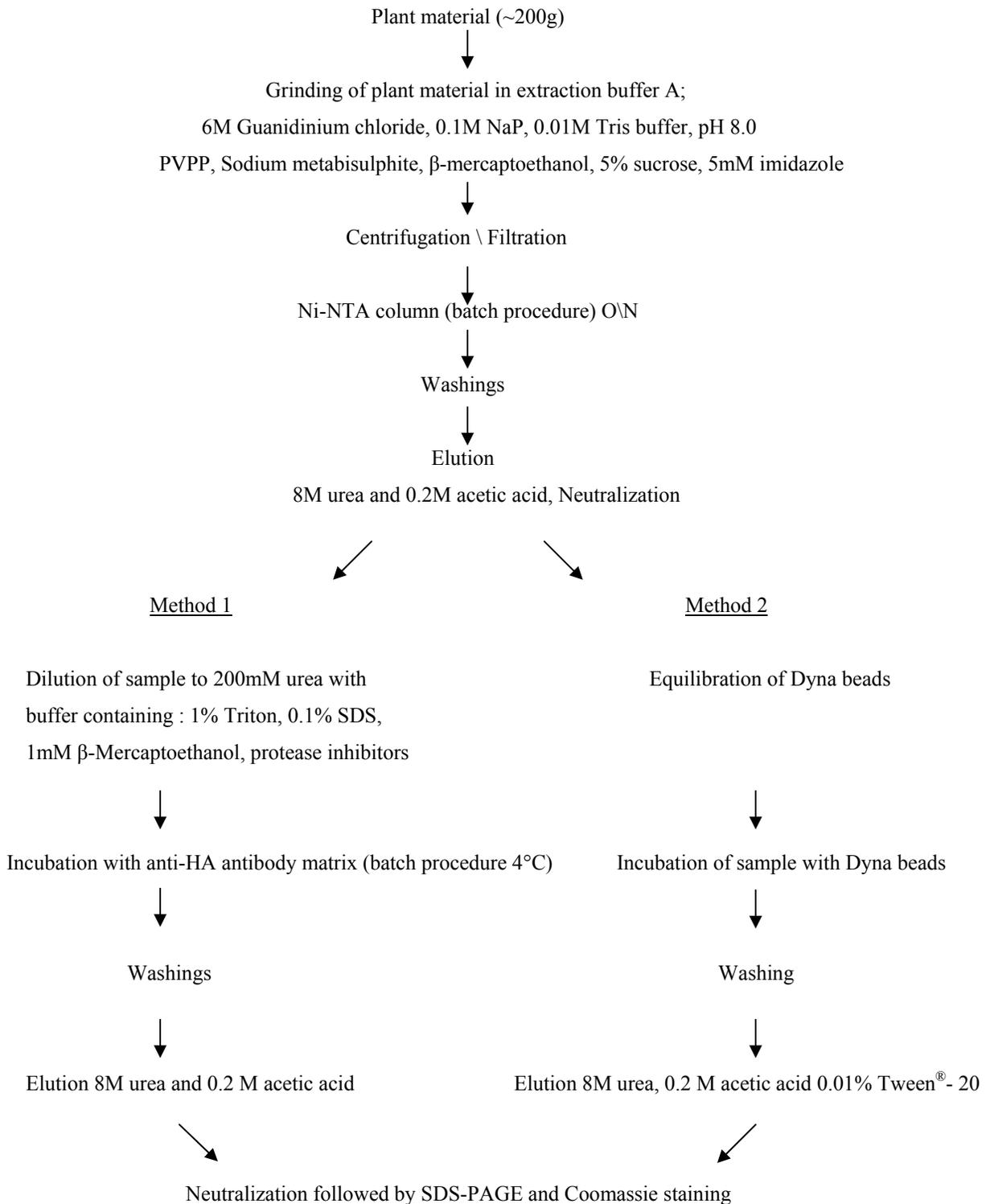
the plant material and it can be used directly for SDS-PAGE after mixing with 2x LSB or stored at -80°C for later use.

#### **2.2.4.2 Measurement of protein concentration with Bradford reagent**

Protein concentration was measured using the 'Bradford Protein Assay'. Bradford solution diluted 1:5 with water was filtered with a 0.45µm filter. 1ml of the filtered solution is added to varying amounts of aqueous protein solutions (50-100µl) in a test tube and incubated at RT. The colour of the solution changed from red to deep blue. OD<sub>595</sub> was measured photometrically and the protein amount was estimated via a calibration curve of a standard protein (BSA).

#### **2.2.4.3 Precipitation of proteins using organic solvents**

Organic solvents such as acetone and trichloroacetic acid (TCA) are useful for precipitating proteins and removing salts and detergents prior to protein analysis. Two volumes of 10% TCA in acetone was added to 1 volume of protein sample in an Eppendorf tube. The mixture was allowed to stand ON at -20°C, followed by centrifugation of the mixture at maximum speed for 20 minutes. The supernatant was immediately decanted and the residual liquid was removed with a pipette. Residual traces of TCA were removed by washing the pellet with cold acetone. The protein pellet was dried on ice at RT. The precipitated protein was dissolved in 1x LSB and analyzed by SDS-PAGE.



**Fig. 13 Schematic representation of the biochemical enrichment of SUMO1 conjugates**

#### **2.2.4.4 Sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE)**

##### Mini gel and large gel

Stacking gel (5%, 2ml): 1.4 ml dH<sub>2</sub>O, 0.33ml acrylamide-bis mix (29:1) 30% w/v (Serva), 0.25ml of 1.5M Tris pH 6.8, 0.02ml 10% SDS, 0.02ml of 10% ammonium persulphate, 0.002ml TEMED

Resolving gel (15%, 5ml) : 1.1 ml dH<sub>2</sub>O, 2.5ml acrylamide-bis mix (29:1) 30% w/v (Serva), 1.3ml of 1.5M Tris pH 8.8, 0.05ml 10%SDS, 0.05ml of 10% ammonium persulphate, 0.002ml TEMED

2x (LSB): 50% glycerol, 20mM DTT (Dithiothreitol), 2% SDS, 125mM Tris. Cl pH 6.8, 0.03% bromophenol blue

2x (LSB) – non-reducing gel. All of above except DTT

5x electrophoresis buffer: 25mM Tris, 192mM Glycine, 0.5% SDS v/v

Proteins were separated on 10-15% SDS-polyacrylamide gels. Samples were incubated in the same volume of 2x Laemmli-sample buffer (LSB) at 65° C then loaded into the wells. For non-reducing gels, 2x Laemmli-sample buffer without the reducing agent (DTT) was used. Wells without the sample were filled with 1x LSB. This reduces ‘gel smiling’ and provides a better running gel front. The stacking gel concentration used was 5%. Electrophoresis was performed at 70 V in the stacking region and at 100V (constant voltage) in the resolving gel with 1x electrophoresis buffer. Mini gel apparatus used was from BioRad and for large gels an apparatus was procured from Whatman Biometra.

##### Invitrogen mini gel apparatus

For some experiments, NuPAGE precast 4-12% bis tris gels were used. The samples were incubated with the loading buffer provided with the system according to the suppliers instructions. Running buffer for the electrophoresis was also prepared according to the suppliers instructions. Electrophoresis was done at 200V (constant voltage).

#### **2.2.4.5 Coomassie staining of proteins resolved by SDS-PAGE**

Gels that were not used for Western blotting were stained with Coomassie stain. Roti-Blue (Roth), a commercial name of a colloidal Coomassie stain was used for staining the protein gels. A solution (60% water, 20% Roti-Blue, 20% methanol) was prepared for

staining the gels for minimum of 7-8 hours. The gels stained with this specific Coomassie stain (Roti-Blue) do not require destaining. The protein bands are stained blue in a transparent background.

#### **2.2.4.6 Sypro-Ruby staining of proteins**

After electrophoresis, the gels were incubated in the fixative solution (10% methanol, 7% acetic acid) at RT for 15 minutes. The fixation step was repeated. Following fixation, 50ml of Sypro-Ruby stain (Molecular probes) was used for staining the gels overnight. Thereafter, to reduce background fluorescence and increase sensitivity, the gels were transferred to a clean staining dish and washed two times in 10% methanol, 7% acetic acid for 30 minutes. The gel was then monitored using UV-illumination and treated for mass spectrometry analysis.

#### **2.2.4.7 Excision of protein bands, in-gel trypsin digest, and MALDI-TOF analysis**

##### Excision of protein bands from polyacrylamide gel and washing

The stained gel was washed for 10 minutes with water (gentle shaking). Stained protein bands (Coomassie stain) of interest were excised, cut to 1 mm-cubes with a sharp knife and transferred to a 0.5µl Eppendorf tube. The gel pieces were then washed with water, followed by a wash with 50mM  $\text{NH}_4\text{HCO}_3$  \acetonitrile 1:1 (v/v) for 15 minutes. All the liquid was removed and the gel pieces were covered with acetonitrile until the gel shrunk and shriveled together. Acetonitrile was removed and the gel was rehydrated in 50mM  $\text{NH}_4\text{HCO}_3$ . After 5 minutes of incubation, an equal volume of acetonitrile was added. All the liquid was removed after 15 minutes. Again acetonitrile was added and removed from the Eppendorf tube after the gel pieces had contracted. The gel pieces were then dried in a vacuum centrifuge.

##### Reduction and alkylation

To the dried gel pieces, 10mM DTT \25mM  $\text{NH}_4\text{HCO}_3$  (freshly prepared) was added and the sample was incubated for 45 minutes at 65°C. After incubation, the Eppendorf tubes

were chilled to RT, the liquid was removed and replaced quickly by the same volume (enough to cover the gel pieces) of freshly prepared 55mM iodoacetamide (light sensitive) in 25mM  $\text{NH}_4\text{HCO}_3$ . The tubes were incubated for 30 minutes at RT in dark. The iodoacetamide solution was removed and the gel pieces were washed with in 50mM  $\text{NH}_4\text{HCO}_3$  \ acetronitrile 1:1 (v\v) for 15minutes. The washing step was repeated twice. The gel pieces were again covered with acetonitrile and after the gel shrunk, liquid was removed and the gel pieces were dried in a vacuum centrifuge.

#### In-gel trypsin digest

To the dried gel pieces, freshly prepared enzyme solution (25mM  $\text{NH}_4\text{HCO}_3$  with 5ng\μl of trypsin) was added to cover the gel, which was incubated at 37°C for 30 minutes. The gel pieces soak up the enzyme solution. Enough 25mM  $\text{NH}_4\text{HCO}_3$  was added to keep the gel wet before O\N incubation at 37°C.

#### Extraction of peptides

The gel pieces were sonicated for 10 minutes and the supernatant was recovered. A solution of 50% acetonitrile and 1% trifluoroacetic acid was added to the gel pieces, which were sonicated again for 10 minutes. The supernatant was recovered after sonication and the extraction step with acetonitrile\trifluoroacetic acid was repeated. The supernatant was pooled and concentrated to about 10μl in a vacuum centrifuge.

#### Zip-Tip purification

Zip-Tip (C18, tip size P10, Millipore) was used for purification of the sample before analysis by MALDI-TOF. The Zip-Tip was washed twice with 10μl of 50% acetonitrile and 0.1% trifluoroacetic acid. Thereafter, the tip was equilibrated twice with 10μl of 0.1% trifluoroacetic acid. Peptide binding was done by pipetting the solution 3-10 times through the filter matrix. The tip was then washed four times with 10μl of 0.1% trifluoroacetic acid. Finally, the peptides were eluted with 8μl of 50% acetonitrile and 0.1% trifluoroacetic acid.

## MALDI-TOF

The peptides were analyzed by MALDI-TOF mass spectrometry (Bruker Daltonics, Reflex IV). Protein identification and result assessment was done with data base system (NCBI) using MASCOT (Matrix science) and Proteome Discoverer (Genomic solutions) search engines.

### **2.2.4.8 Isoelectric focusing (IEF)**

Isoelectric focusing was done with ZOOM IPGRunner System (Invitrogen life technologies)

#### Rehydration of ZOOM strips:

Rehydration buffer: 7M Urea, 2M Thiourea, 25mM CHAPS, 0.5 %(v/v) Ampholyte pH 3-7, 0.002% Bromophenol blue, 20µl DDT

Precipitated protein samples (Protein samples were precipitated as described in section 2.2.4.3) were resuspended in 150µl of rehydration buffer. The rehydration buffer containing the protein sample was loaded into the sample loading well in the ZOOM IPGRunner cassette. The protective covering from the IPG strip was peeled off. Each IPG strip has a gel side and a side with printed marking on it. The strip with the gel side upwards and held with forceps on the basic (-) end was gently slid through the sample loading well of the ZOOM IPGRunner cassette. The IPG strip was inserted until the acidic (+) end of the strip touched the end of the channel slot. Care was taken to avoid introducing large air bubbles while sliding the IPG strip into the sample well. All sample loading wells in the IPGRunner cassette were sealed with a sealing tape provided with the kit. This was done to create a sealed environment for rehydration. The ZOOM IPGRunner cassette with the ZOOM strips was incubated for 8-16 hours at RT to hydrate the strips.

#### Focusing

The sealing tape was removed from the ZOOM IPGRunner cassette. Electrode wicks (provided with the kit) were moistened with deionized water and placed over the exposed

adhesives at each end of the ZOOM IPGRunner cassette. The IPGRunner apparatus was then assembled by sliding the ZOOM IPGRunner cassette towards the electrodes of the ZOOM IPGRunner core until the electrode wicks of the ZOOM IPGRunner cassette are in contact with the electrodes of the ZOOM IPGRunner core. The gel tension lever which holds the ZOOM IPGRunner cassette and ZOOM IGRunner core firmly, was inserted into the focusing chamber. The outer chamber of the focusing apparatus was filled with 600ml of deionized water and isoelectric focusing was performed (Voltage ramp: 175V for 15 minutes, 175-2000V ramp for 45 minutes, and 2000V for 20-30 minutes). Following isoelectric focusing, the entire cassette with the strips can be stored in a sealed container at -80°C or the IPG strips can be removed from the cassette and stored (in a sealed container) at -80°C for later follow up experiment.

#### Equilibration of IPG strips

Equilibrating the ZOOM IPG strips in the equilibration buffer prepares the strips for 2D SDS-PAGE.

Equilibration buffer A: 4.5ml 1x NuPAGE LDS sample buffer, 0.5ml NuPAGE sample reducing agent

Equilibration buffer B: 4.5ml 1x NuPAGE LDS sample buffer, 116mg iodoacetamide

Each IPG strip was incubated in a conical tube containing equilibration buffer A for 15 minutes at RT. After 15 minutes of incubation equilibration buffer A was decanted, equilibration buffer B was added to the tube and the strip incubated further for 15 minutes. The strips were ready to proceed for SDS-PAGE.

#### **2.2.4.9 Two-dimensional gel electrophoresis**

A 0.5% agarose solution was prepared and kept warm (55-65°C) until use. The ZOOM IPG gel cassette was disassembled and the IPG strips were gently removed. The plastic ends of the strip were cut off, taking care not to cut any gel pieces. The IPG strip was slid into the channel well of a precast 4-12% bis-tris gel using a thin spatula or a thin plastic ruler. 400µl of 0.5% agarose solution was added to the well containing the IPG strip,

taking care that the agarose did not flow into the well containing the molecular weight marker. The agarose was allowed to solidify. A typical SDS-PAGE was performed using the Invitrogen mini electrophoresis apparatus according to the supplier's instructions.

#### **2.2.4.10 Protein spotting, trypsin digests with a robot and MALDI-TOF analysis**

Samples subjected to two-dimensional gel electrophoresis were subsequently stained using Sypro-Ruby according to the producer's manual (Molecular probes). Fluorescence images were acquired using FLA 3000 (Fuji film). Spots of various intensities were automatically picked (PROTEINEER sp, Bruker) and tryptically digested with a robot (PROTEINEER dp, Bruker). Aliquots of the digest were prepared for MALDI-TOF\MS. Aliquots of more interesting spots were subjected to LC-MS\MS analysis. Protein identification and result assessment was done on a protein scape 1.3 database system that triggered Mascot searches (MATRIX sciences) and simplified the evaluation of MS and LC\MS searches in a gel related context using Bio-Tools 3.0 (Bruker).

#### **2.2.4.11 Western blotting**

Transfer buffer: 190mM glycine, 20mM Tris, 20% methanol, 0.05% SDS

1x ANT: 150mM NaCl, 50mM Tris.Cl pH 8, 0.02% sodium azide (NaN<sub>3</sub>)

NBT( Nitoblu Tetrazolium chloride) solution. 110mM in 70% dimethylformamide(DMF) solution

X-phosphate solution: BCIP (5- Bromo-4 chloro 3-indolephosphate) 90mM in DMF

After electrophoresis, gels were transferred to a trough containing transfer buffer and equilibrated for 30 minutes with gentle shaking. Meantime, a piece of PVDF-membrane (Millipore) was briefly dipped in methanol and also equilibrated in transfer buffer. The transfer was performed in the BioRad transfer apparatus for one hour at 50V (constant voltage) in the cold room. Subsequent to the transfer, the membrane was blocked for 1.5 hours with 20% NCS at room temperature. Following blocking, the blot was incubated with the primary antibody diluted to a final concentration of 1:1000 in the 20% NCS\1xANT. Incubation with the primary antibody was done O\N in the cold room.

After washing the membrane with 1xANT containing 0.05% Tween 20, (4 washes, 15 minutes each) a secondary antibody conjugated to alkaline phosphatase (1:1000) in 20% NCS\1xANT was added and the blot was incubated at RT for 2 hours. After several washing steps (3 times, 10 minutes each) with 1xANT\0.05% Tween 20, the membrane was developed in dark with TE containing 45µl\10ml of NBT solution and 35µl\10ml of X-phosphate solution (BCIP).

#### **2.2.4.12 Small scale testing of overexpression of proteins in *E. coli***

IPTG: (Isopropylthiogalactoside) stock solution 100mM, filter sterilized

To find out whether recombinant expression vectors express the cloned gene properly, a small scale protein preparation was carried out. 4ml of 2xTY medium containing the appropriate antibiotic was inoculated with 400µl of an O\N culture and grown at 37°C to an OD<sub>600</sub> of 0.8. 1mM IPTG was added and the culture was further incubated for 3-4 hours. 100µl of the culture was pelleted by centrifugation and suspended in 10µl of dH<sub>2</sub>O and 10µl of 2x LSB. The resuspended cultures were boiled at 95°C for 10 minutes before use for SDS-PAGE.

#### **2.2.4.13 Purification of His-tagged proteins expressed in *E. coli* using Ni-NTA matrix**

##### Native conditions

Lysis buffer (1 litre): 50mM NaH<sub>2</sub>PO<sub>4</sub>, 300mM NaCl, 10mM imidazole, pH 8

Wash buffer (1 litre): 50mM NaH<sub>2</sub>PO<sub>4</sub>, 300mM NaCl, 20mM imidazole, pH 8

Elution buffer (1 litre): 50mM NaH<sub>2</sub>PO<sub>4</sub>, 300mM NaCl, 250mM imidazole, pH 8

Overexpression of proteins fused to a His tag allows purification using a Ni-NTA matrix. *E. coli* cells containing the fusion vector were grown at 37°C O\N in 10ml of culture medium containing selective antibiotics. 100ml of prewarmed medium (with antibiotics) was inoculated with 5ml of the overnight culture and grown at 37°C with vigorous shaking until an OD<sub>600</sub> of 0.7 was attained. 1mM of IPTG was added to induce the expression of the gene. After 4-5 hours of growth, cells were harvested by centrifugation

at 5000 rpm for 20 minutes. The pellet was stored at -20°C until use. The frozen cell pellet was thawed on ice and the cells were resuspended in lysis buffer at 2-5ml per gram wet weight. 1mg/ml lysozyme was added and the resuspended cells were incubated on ice for 30 minutes. The partially disrupted cells were then sonicated on ice (six 10 second bursts at 200-300W with 10 seconds cooling period after each burst) using a sonicator equipped with a microtip. The lysate was centrifuged at 10,000 x g for 20-30 minutes at 4°C to pellet the cellular debris. Thereafter, the supernatant (cleared lysate) was removed. 1ml of the 50% Ni-NTA slurry was added to 4ml cleared lysate and mixed gently by shaking at 4°C for 1 hour on a rotating shaker. The lysate Ni-NTA mixture was loaded into a column and the column flow through was collected for analysis. The column was washed twice with 4ml wash buffer, and fractions were collected. Thereafter, the protein was eluted 5-6 times with 0.5ml of elution buffer. Fractions of the eluate were collected in different tubes and analyzed by SDS-PAGE. Subsequent to SDS-PAGE analysis, the fractions of interest were further purified by sizing chromatography. FPLC was done using a superdex 200 column and the whole procedure was performed using an automated FPLC system (Amersham). The samples were centrifuged at maximum speed, or filtered through a 0.2µm filter before injecting into the FPLC system.

#### Denaturing conditions

Lysis buffer (1 litre): 100mM NaH<sub>2</sub>PO<sub>4</sub>, 10mM Tris.Cl, 6M Guanidinium chloride, pH 8

Wash buffer (1 litre): 100mM NaH<sub>2</sub>PO<sub>4</sub>, 10mM Tris.Cl, 8M Urea, pH 6.3

Elution buffer (1 litre): 100mM NaH<sub>2</sub>PO<sub>4</sub>, 10mM Tris.Cl, 8M Urea, pH 4.5

*E. coli* cells containing the fusion vector were grown in 10ml of culture medium containing the selective antibiotics at 37°C O\N. 100ml of prewarmed medium (with antibiotics) was inoculated with 5ml of the overnight cultures and grown at 37°C with vigorous shaking until an OD<sub>600</sub> of 0.7 was attained. 1mM of IPTG was added to induce the expression of the gene. After 4-5 hours of growth, cells were harvested by centrifugation at 5000 rpm for 20 minutes. The pellets were stored at -20°C until use. The frozen cell pellet was thawed on ice and the cells were resuspended in lysis buffer at 5ml per gram

wet weight. Cells were stirred at room temperature in the lysis buffer or alternatively lysed by gentle shaking with a vortex mixture, taking care to avoid foaming. The lysate was centrifuged at 10,000xg for 20-30 minutes at 4°C to pellet the cellular debris. Thereafter, the supernatant was decanted. 1ml of the 50% Ni-NTA slurry was added to 4ml cleared lysate and mixed gently for 1 hour (RT) on a rotating shaker. The lysate - Ni-NTA mixture was loaded into an empty column and the column flow through was collected. The column was washed twice with 4ml wash buffer, and fractions were collected. Thereafter, the protein was eluted 5-6 times with 0.5ml of elution buffer. Fractions of the eluate were collected in different tubes and analyzed by SDS-PAGE.

#### **2.2.4.14 Expression and purification of SUMO-conjugating enzyme (SCE)**

This experimental procedure was courtesy of Dr. Frauke Melchior (MPI, Martinsried\Univ. of Göttingen). The SUMO-conjugating enzyme gene cloned in an expression vector (pET-9d) was transformed in BL21(DE3) competent cells. From a 50ml O<sub>2</sub>N culture, 20ml of culture was withdrawn. The cells were harvested by centrifugation, resuspended in 20 ml fresh medium and the culture was used for inoculating 2 liters of 2xTY selective (antibiotic) medium. The cells grown at 37°C were induced with 1mM IPTG after an OD<sub>600</sub> of 0.7 was attained. The culture was further incubated for 3-4 hours at 37°C. Bacteria were harvested by centrifugation and resuspended in 60ml buffer containing 50mM Na-phosphate buffer pH 6.5 and 50mM NaCl. The resuspended bacteria were stored at -80°C overnight (freezing is essential). Cells were thawed on ice and subsequently 1µg/ml protease inhibitors (PMSF, aprotinin, leupeptin, pepstatin) and 1mM DTT were added to the thawed cells. Thereafter, ultracentrifugation was done at 100,000xg for 1 hour at 4°C. 10ml SP-Sepharose beads (SIGMA) were equilibrated by resuspension and centrifugation as follows: 1x with 0.5M Na-Phosphate (pH 6.5); 2x with 0.5M Na-Phosphate, 50mM NaCl (pH 6.5) and 1x with 0.5M Na-Phosphate, 50mM NaCl (pH 6.5), protease inhibitors and 1mM DTT. The bead suspension was loaded onto a 50ml empty column (cold room). The supernatant from the ultracentrifugation was passed through the packed column (discarded flow through). Subsequently, the column was washed with 2-3 bed volumes of 0.5M Na-Phosphate, 50mM NaCl (pH 6.5),

protease inhibitors and 1mM DTT. Protein was eluted with 2-3 column volumes of buffer consisting of 0.5M Na-Phosphate, 300mM NaCl (pH 6.5), protease inhibitors and 1mM DTT. Eluate was collected in 2ml fractions and kept on ice. The fractions were analyzed by SDS-PAGE and stained with Coomassie colloidal stain. The peak fractions were pooled and concentrated to a volume of 1-2ml with a Centricon centrifugal device. The sample was filtered through 0.2 $\mu$ m filter or centrifuged at maximum speed before loading on to a preparative Superdex 200 column (FPLC). The column was equilibrated with buffer containing 20mM Hepes\KOH (pH 7.3), 110mM potassium acetate, 2mM magnesium acetate, 0.5mM EDTA, 1mM DDT, 1 $\mu$ g/ml aprotinin, 1 $\mu$ g/ml each of leupeptin and pepstatin prior to sample loading. The eluate was collected in 5 ml fractions. The fractions were analyzed by SDS-PAGE and the gel was Coomassie stained. The peak fractions were pooled, and small aliquots (15-30 $\mu$ l) were flash frozen in liquid nitrogen.

### **2.2.5 Production of SCE antibody in rabbit**

The SUMO-conjugating enzyme used for antibody production was purified using Ni-NTA resin under denaturing conditions. The concentration of the protein was determined using Bradford assay. A SDS-PAGE was performed and the gel was stained with 4M sodium acetate solution for 2-3 hours. The protein band was excised from the gel and send to Eurogentec (Belgium) for SCE antibody production. 400mg of the protein per rabbit was send to the company. Two rabbits were injected for antibody production.

#### **2.2.5.1 Purification of rabbit polyclonal immunoglobulins (IgGs) using affinity chromatography**

Two thirds of the final bleed obtained from the company Eurogentec was used for purification of the SCE antibody from crude serum using affinity chromatography.

#### **Ammonium sulphate precipitation**

The crude serum was mixed with 1 volume of phosphate buffered saline (PBS), pH 7.4.

The mixture was gently stirred on ice and simultaneously solid ammonium sulphate was added (29.1g\100ml). This preparation was kept at 4°C O\N. The following day, the mixture was centrifuged at 8000rpm at 4°C for 15 minutes and the supernatant was discarded. The pellet was dissolved in 20ml of PBS and dialysed in a dialysis tubing (molecular weight cut off 12000-14000 kDa, Serva) in the cold room. PBS, pH 7.4 was used as dialysis buffer.

#### Ligand coupling and binding of precipitated serum proteins

A 2ml uniform suspension of Affi gel 10 was equilibrated by suspension and centrifugation in 5mM HEPES, 20mM NaCl pH 7.5. 25mg\ml of ligand (purified SCE) was added, and the matrix was gently agitated on a rotating shaker at 4°C O\N. Thereafter, 0.1ml of 1M ethanolamine HCl was added per ml of the gel matrix. The ligand coupled matrix was then loaded onto a 20ml empty chromatography column and equilibrated with PBS. Precipitated serum proteins (dissolved in 20ml PBS) were applied to the column and the flow through was collected. The flow through was passed through the matrix in the column and subsequently the column was washed with 3-4 bed volumes of PBS buffer. Immunoglobulins were eluted with 20ml 0.1M glycine.HCl (pH 3.0) in 2ml fractions, which were immediately neutralized with 1M Tris, pH 8 and analyzed by SDS-PAGE.

#### **2.2.6 *In vitro* sumoylation assays**

The *in vitro* SUMO conjugation system employed in this investigation contains human SAE (0.5µg of SAE1/SAE2, Biomol or Boston Biochem), human SCE (1µg Ubc9), *Arabidopsis* SUMO proteins (1µg) and substrate protein (2-5µg) in sumoylation buffer (20mM Tris PH 7.5, 5mM MgCl<sub>2</sub>, 5mM ATP). The assays were performed in 50µl volume. The sumoylation reactions were incubated at 30°C for 4 hours. Reactions were terminated by boiling the mixture in SDS containing loading buffer. Reaction products were separated on a 12% SDS-PAGE and Western blotting was done using an antibody specific to the epitope tag on the substrate.

### **2.2.7 Brief protocol for PCR-grade DNA isolation from *A. thaliana*.**

30-50 mg of plant material (leaves) was homogenized in 400µl buffer (200mM Tris pH 7.5, 250mM NaCl, 0.5% SDS, 25mM EDTA) with an electric homogenizer (IKA Ultra Turrax, Germany). Cell debris was removed by centrifugation (5 minutes, 14,000 rpm, RT). The supernatant was transferred to a fresh Eppendorf tube, mixed with 1 volume isopropanol and incubated at RT for 5-10 minutes. Precipitated nucleic acids were collected by centrifugation (5 minutes, 14,000 rpm, RT), washed with 70% ethanol and dried in a spin vacuum. The pellet was resuspended in 50µl TE buffer by incubating at 65°C for 10 minutes. 1-3µl of the crude DNA preparation was used in a 50µl PCR reaction.

### **2.2.8 Polymerase chain reaction (PCR)**

PCR reactions were typically carried out in 50 µl reactions. For standard PCR, DNA was amplified by the use of *Taq* DNA polymerase. Typically, a PCR mixture employed 1.25 units Polymerase, 0.4µM forward (Fwd) and reverse (Rev) primers, 200µM dNTPs (dATP, dCTP, dGTP, dTTP) and 1x PCR buffer (10x PCR buffer: 100mM Tris.Cl pH 8.3, 500mM KCl, 15mM MgCl<sub>2</sub>, 0.1% gelatin, 0.5% Tween 20, 250µg/ml BSA). The temperature regime consisted of an initial 2 minute denaturing step at 96°C. This was followed by a cycle of denaturing, annealing and extension conditions. The number of cycles and the temperatures of each step depended on template and primers. A typical PCR had 30 cycles of (i) 30 seconds denaturing at 94°C (ii) 30 seconds annealing at 62°C (iii) 2 minute extension at 72°C. This cycling was followed by a 10 minute final extension. The major variations in PCR conditions involved differences in the number of cycles performed and the use of different annealing temperatures.

### **2.2.9 Nucleic acid hybridization techniques**

#### **2.2.9.1 Isolation of total RNA from *A. thaliana***

RNA extraction was performed using the Qiagen RNeasy plant mini kit extraction. 100 mg of the plant tissue was ground with a pestle and mortar to homogenize the tissue, and

450µl of lysis buffer RLT was added. The sample was mixed on a vortex shaker and incubated at RT for two minutes. It was then added to a QIAshredder spin column and centrifuged for 2 minutes at maximum speed. The supernatant was removed and added to 225µl of 100% ethanol. The sample was added to an RNeasy minicolumn and centrifuged for 15 seconds at 1000 rpm (Eppendorf microcentrifuge). 700µl of wash buffer RW1 was added to the column, and the column was centrifuged for 15 seconds at 1000 rpm. The column was transferred to a new collection tube, and 500µl of wash buffer RPE were added. The sample was centrifuged for 15 seconds at 10,000 rpm. The RPE wash step was repeated. The column was transferred to a new 1.5ml collection tube and RNA was eluted by the addition of 50µl of RNase free water to the column, followed by centrifuging for one minute at 10,000 rpm. Concentration of RNA was analysed by running an agarose gel and comparing the intensity of the RNA bands to a standard RNA marker.

### **2.2.9.2 Electrophoresis of RNA on denaturing gels**

10x MOPS (1L): 41.8 g MOPS, 16.6ml of 3M NaOAc, 20ml of 0.5M EDTA, H<sub>2</sub>O upto 1000 ml, store the solution in dark.

6x RNA loading buffer: 150mg Ficoll, 200µl of 0.25 EDTA, 0.15% w/v orange G

RNA sample buffer (24µl each sample): 2.4µl 10x MOPS, 4.5µl formaldehyde, 12µl deionized formamide, 4.5µl DEPC-H<sub>2</sub>O, 0.75µl of 6x RNA loading buffer, 0.75µl ethidium bromide (7mg/ml)

RNA was fractionated by running through a 1.2% agarose, 3% formaldehyde gel (115 ml gel = 1.40g agarose, 3.5ml formaldehyde, 11.16ml 10x MOPS, 100ml DEOPC treated water; MOPS and formaldehyde were added after the dissolved agarose had cooled to 65°C). 5-10µg of RNA was precipitated with 3M DEPC-NaOAc and 96% ethanol. The RNA pellet was then suspended in 24µl RNA sample buffer. Prior to loading, the samples were heated for 5 minutes to 95°C and then briefly centrifuged at 4°C. Electrophoresis was performed at 5V/cm gel in a fume hood. Electrophoresis buffer was 1x MOPS.

### 2.2.9.3 Northern blotting

20x SSC: 3M NaCl, 0.3M Sodium citrate, pH 7

Following electrophoresis, the formaldehyde\agarose gel was placed in a tray containing water and gently shaken. This was followed by two 15 minute rinses in 10x SSC. The gel was placed on two pieces of Whatman 3MM paper, supported over a reservoir of 20xSSC. The ends of the Whatman 3MM paper were submerged in the 20xSSC solution to form a wick. A piece of nylon membrane (Hybond-N, Amersham) was cut to the same size as the gel, and dipped in 50% DEPC-methanol, rinsed in DEPC-water and subsequently in 10xSSC. The membrane was placed on top of the gel, and any air bubbles present were removed by rolling with a glass pipette. Three more pieces of Whatman 3MM paper were cut to the size of the gel, presoaked in 10xSSC and placed on top of the nylon membrane. A stack of paper towels was then placed on top and a ~500g weight was placed on the towels. This encouraged the transfer of the 20xSSC from the reservoir to the paper towels, through a capillary action. The setup was left overnight, to allow the RNA to thoroughly transfer to the membrane. After the transfer, RNA was bound to the membrane by UV cross linking (12000 $\mu$ J UV light, Stratagene Stratalinker) followed by baking for 2 hours at 80°C.

### 2.2.9.4 Radioactive labeling of probes

60-90 ng of probe DNA was denatured by boiling in an appropriate amount of distilled water (for the total reaction volume of 30 $\mu$ l). The denatured probe was immediately transferred to ice in order to retain its denatured state. 3 $\mu$ l of hexadeoxyribonucleotides (10x, Roche) and 3 $\mu$ l of 5mM dNTP (without dCTP) were added to the denatured probe on ice. Further, 4 $\mu$ l (~40 $\mu$  Ci) of radioactivity [ $\alpha$ -<sup>32</sup>P] was then added to the probe DNA followed by addition of 1 $\mu$ l Klenow fragment (2 units/ $\mu$ l). The reaction was incubated for 1 hour at 37°C. To remove the unincorporated nucleotides from the probe, the probe was purified with a Nucleospin-Extract II gel extraction kit (Machinery-Nagel) according to the supplier's instructions. The purified probe was eluted with 50 $\mu$ l of elution buffer supplied in the kit. A further 50 $\mu$ l of TE was added to the eluted probe. The probe was

then denatured by boiling in an Eppendorf tube for 5 minutes and transferred immediately to ice.

#### **2.2.9.5 Hybridization of radio labeled probe to RNA immobilized on nylon membrane.**

Hybridization buffer: 1% BSA, 1mM EDTA, 0.5M sodium phosphate buffer pH 7.2, 7% SDS

Wash buffer A: 2x SSC, 0.1% SDS

Wash buffer B: 1x SSC, 0.1% SDS

Wash buffer C: 0.5x SSC, 0.1% SDS

Wash buffer D: 0.1x SSC, 0.1% SDS

The nylon membrane was pre-hybridized at 65°C for one hour in Pyrex glass tubes containing 20-25ml of the hybridization buffer. This was performed in a hybridization oven containing a rotating spindle. Radiolabelled denatured probe was then added to the Pyrex tube, and returned to the oven. Hybridization was performed overnight. Thereafter, the membrane was rinsed with low stringency wash buffer A, and then washed for 40 minutes. This wash was followed by another wash of 25 minutes with high stringency buffer B (25ml), then 15 minutes with buffer C in the oven. The final wash was with a high stringency buffer D for 3-5 minutes. Following washing, membranes were sealed in a cling film. The strength of the hybridization signal was assessed with a Giga counter. Thereafter, the membrane was exposed to the screen of a phosphoimager cassette. For re-use, the membrane was stripped to remove the probe by subjecting it to two washes of 2-3 hours in boiling 0.1% SDS solution.

### 3. RESULTS

#### 3.1 Characterization of proteins of the sumoylation pathway

##### 3.1.1 Cloning and expression of tagged SUMO isoforms in pET-9d

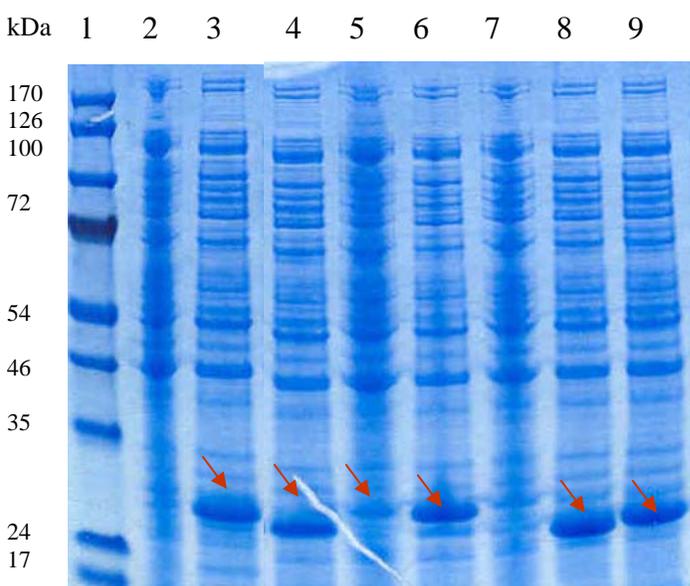
SUMO isoforms, SUMO 1, 3, 5, 6, 7, 7v (see table 4; sequences in appendix) were first tagged with three affinity purification tags, namely Strep tag, triple hemagglutinin affinity tag and octahistidine Ni affinity sequence at the amino terminus of each gene. The pBluescript vector carrying the sequences for the expression of these tags was cleaved with Asp718 and XbaI. Inserts of SUMO1, 3, 5, 7 and 7v (Asp718 and XbaI) were then ligated into the linearized pBluescript plasmid containing the tags, which generated the SUMOs with three affinity tags. The resulting plasmid was called pSK tag3 SUMO insert. To obtain the inserts of SUMO isoforms with the affinity tags, SUMO1, 3, 7 and 7v were cleaved with NotI, the recessed ends were filled in with Klenow enzyme and finally digested with NcoI. For SUMO5 and SUMO6, Ecl136II and NcoI were used for cleavage. The pET-9d expression vector was cleaved with BamHI, followed by Klenow fill in and cleavage with NcoI. Tagged SUMO inserts (1, 3, 5, 7, 7v) were ligated into the linearized pET-9d vector and transformed into *E. coli* (XL1- Blue). The clones obtained were screened for the respective SUMO inserts. The correct clones were used for recombinant DNA isolation and the isolated DNA construct was transformed into *E. coli* (BL 21) cells for expression studies.

Small scale cultures of SUMO1, 3, 5, 6, 7 and 7v were used for the preparation of crude protein extract and subjected to SDS-PAGE. Most of the overexpressed SUMOs except SUMO6 were visible as a prominent band corresponding to their expected molecular weight (Fig. 14) in the Coomassie stained gel.

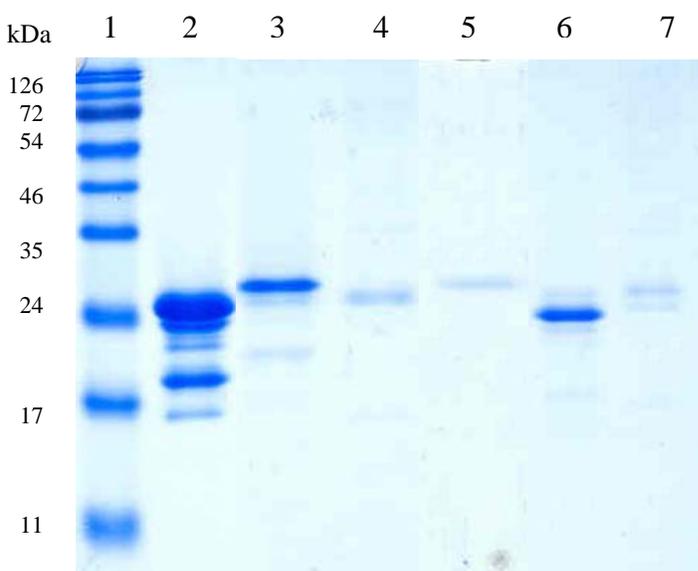
##### 3.1.2 Purification of SUMO isoforms from *E. coli*

SUMO isoforms were purified under native conditions using 100ml of pET-9d tag 3 SUMO (1, 3, 5, 6, 7 and 7v) cultures grown and induced as described in section 2.2.4.13. Purification of the overexpressed protein was performed by metal affinity chromatography (Ni-NTA) according to the protocol described in The QIAexpressionist,

5<sup>th</sup> edition (for detailed description see material and methods). The Ni-NTA purified fraction of the proteins was subjected to size exclusion chromatography (FPLC) using superdex 200 column. Fig.15 shows the purified fractions of the various SUMOs on a Coomassie stained gel. The results indicate that all *Arabidopsis* SUMO proteins were stably expressed and purified from *E. coli*.



**Fig. 14 SDS-PAGE analysis of cell lysates from SUMO expressing *E. coli* clones.** Lane 1 represents: Protein marker, lane 2: non induced SUMO1 expression, lane 3: IPTG induced expression of SUMO1, lane 4: IPTG induced expression of SUMO3, lane 5: non induced SUMO5 expression, lane 6: IPTG induced expression of SUMO5, lane 7: IPTG induced expression of SUMO6, lane 8: IPTG induced expression of SUMO7, lane 9: IPTG induced expression of SUMO7v. The gel was stained with Coomassie blue for visualization of protein bands. Red arrows indicate the various SUMO protein bands.



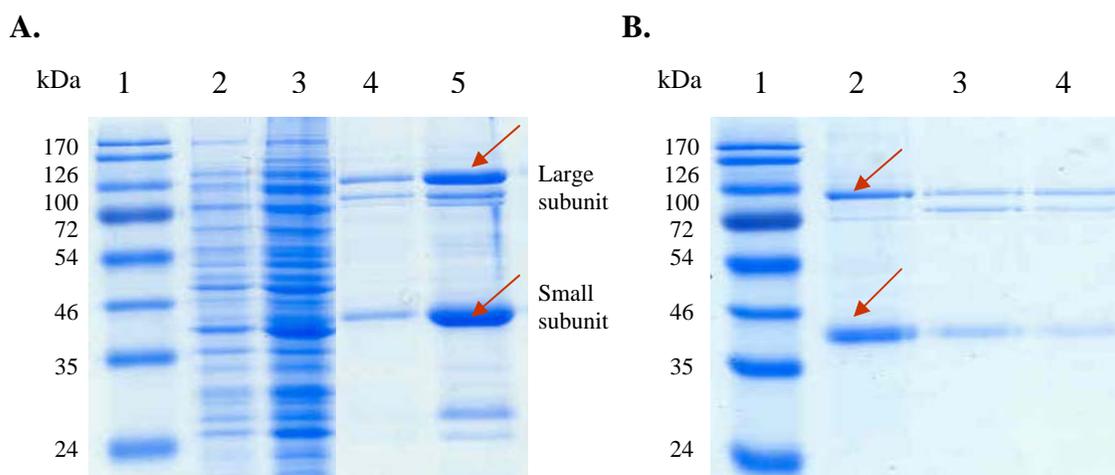
**Fig. 15 Purification of recombinant SUMO proteins.** Recombinant proteins were purified from *E.coli* as described in section 2.2.4.13, subjected to SDS-PAGE and stained with Coomassie blue. Purifications were done using Ni-NTA affinity purification followed by size exclusion chromatography. Lane 1: protein marker, lane 2: SUMO1, lane 3: SUMO3, lane 4: SUMO 5, lane 5: SUMO6, lane 6: SUMO7, lane 7: SUMO7v

### 3.1.3 Cloning of SUMO-activating enzyme

*Arabidopsis* SAE is a heterodimer consisting of one smaller SAE subunit, SAE1a or SAE1b, and the larger subunit of SAE, SAE2. The smaller SAE subunit genes were fragments obtained from gateway clone pDEST-SAE1a\b (obtained from Dr. Yong-Fu Fu) digested with NheI and SgrAI and inserted into low copy vector pACYC177. For insertion into pACYC177, the vector was first cleaved with ScaI and BanI and re-ligated to destroy the Amp resistance marker before cleaving with NheI and SgrAI. The larger subunit SAE2 fragment in pDEST-17 vector was also obtained from Dr. Yong-Fu Fu. Both vectors for SAE1a and SAE2 expression were co-transformed into *E. coli* for the expression and subsequent purification of SAE.

### 3.1.4 Purification of *Arabidopsis* recombinant SAE from *E. coli*

*Arabidopsis* SUMO-activating enzyme was purified from *E. coli* using the Ni-NTA matrix and the eluted fractions were further purified by size exclusion chromatography. Fig. 16 shows the cell lysates from IPTG-induced and non-induced cultures. The red arrows mark the large and the small subunit of the protein in the eluted fraction.



**Fig. 16 Fractionation of the purified *Arabidopsis* SUMO-activating enzyme by SDS-PAGE (A)** The protein was purified using metal (Ni-NTA) affinity chromatography Lane 1: stained marker, lane 2: cell lysate of non induced SAE expressing culture, lane 3: cell lysate of IPTG-induced culture, lanes 4 and 5: eluted fractions from Ni-NTA column. The red arrows indicate the large and small subunit of SAE after purification in the Coomassie stained gel. **(B)** Purified fractions after size exclusion chromatography of the His<sub>6</sub> affinity fraction of *Arabidopsis* SUMO activating enzyme. Lanes 1, 2 and 3 show the various fractions of the eluted protein. The large and small subunit of the protein have been marked with red arrows.

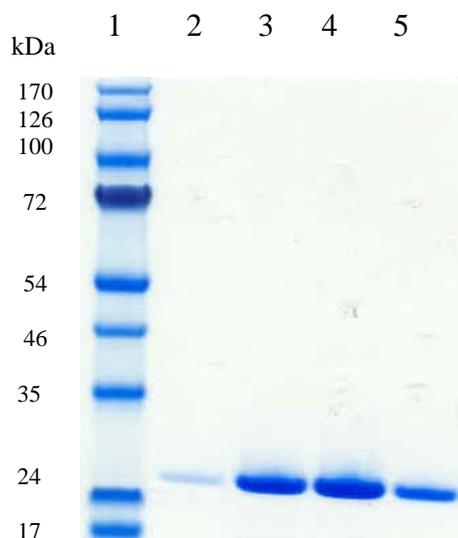
### 3.1.5 Cloning and expression of *Arabidopsis* SUMO-conjugating enzyme in pET-9d

The SCE insert was generated from the pSK-SCE (pBluescript-SCE) by cleavage with SmaI at RT and followed by NcoI at 37°C. The fragment was excised and purified from the gel. pET-9d plasmid vector was cleaved with BamHI, followed by Klenow fill in and a final restriction digest with NcoI. The SCE insert was ligated into the linearized expression vector and the recombinant DNA (pET-9d SCE) was transformed in *E. coli* (XL1- Blue). DNA was isolated from the transformed cells by the alkaline lysis method and screened for the SCE insert by restriction digests. DNA from positive clones was transformed in BL 21 cells for further expression studies.

Some pET-9d SCE expression clones in BL21 cells were tested for the expression of the protein by induction with IPTG. The expected size of the protein was about 19 kDa. Using precast 4-12% bis-tris gradient gels, the protein migrated slowly and the band appeared at around 26 kDa. However, on a self-made Tris buffer 12 % polyacrylamide gel, the band migrated to a position corresponding to 19 kDa.

### 3.1.6 Purification of the *Arabidopsis* SCE from *E. coli*

The *Arabidopsis* SUMO conjugating enzyme was purified from 2L of pET-9d SCE culture as described in the Material and method section 2.2.4.15. The protein leaks out into the supernatant upon freezing and thawing of the cells. The lysate was subjected to ion exchange chromatography (SP-Sepharose beads) and further purified by FPLC using a Superdex 200 column. The protein eluted in several fractions which were tested by SDS-PAGE and the gel stained with Coomassie stain to visualize the bands. Fig. 17 shows the SDS-PAGE analysis of the purified *Arabidopsis* SUMO-conjugating enzyme.



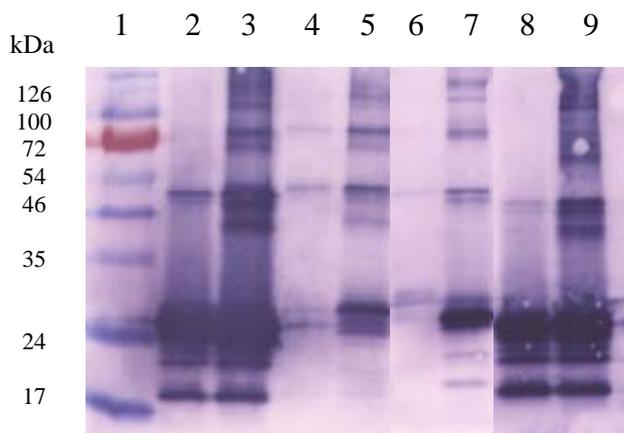
**Fig. 17 Purification of *Arabidopsis* SCE.** The protein was purified using ion exchange chromatography followed by further purification with a sizing chromatography column. The picture shows a Coomassie stained gel of the purified *Arabidopsis* SCE fractions after size exclusion chromatography. Lane 1: Protein marker, lanes 2, 3, 4 and 5 are fractions of the eluted protein from Superdex 200 (size exclusion) column.

## 3.2 *In vitro* analysis of sumoylation

### 3.2.1 Activation of *Arabidopsis* SUMOs with human sumoylation enzymes

The ability of various *Arabidopsis* SUMOs to form thioester with human SCE was characterized *in vitro* in presence of MgATP and recombinant human SAE. The products of the *in vitro* reaction were analyzed by SDS-PAGE under non-reducing conditions. A range of higher molecular weight bands were visualized on a Western blot with anti-HA antibody (SUMOs carry an HA tag). The higher molecular weight bands possibly represent thioester to SCE, SAE and isopeptide bonds of the *Arabidopsis* SUMOs. Interestingly, SUMO7 which is not a highly expressed protein in plants was also found to be activated by the human sumoylation enzymes *in vitro*. In addition, thioester bond was

also formed between SUMO1 (Q93A) mutant and human SCE (Fig. 18), which validate our results *in vivo* (discussed in section 3.4.5) where we were able to detect conjugates in plants expressing the SUMO (Q93A) mutant protein.

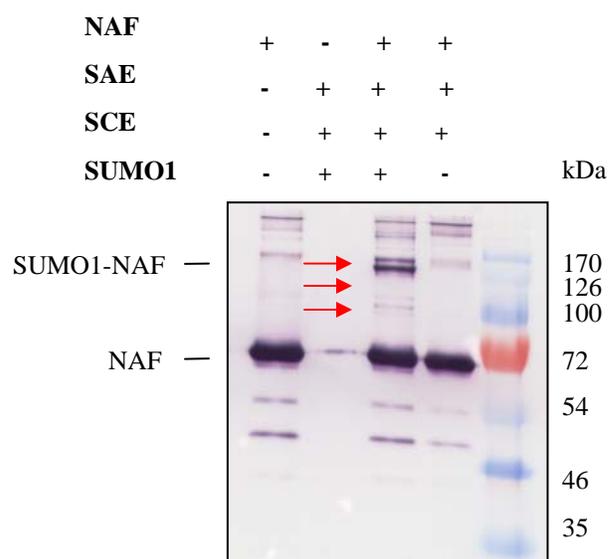


**Fig. 18 *Arabidopsis* SUMO isoforms and SUMO1 (Q94A) mutant are activated by human sumoylation enzymes.** Assay for thioester formation between SUMO and human SCE. Lane 1: marker, lane 2: SUMO1, lane3: SUMO1, human SAE, human SCE, lane 4: SUMO3, lane 5: SUMO3, human SAE, human SCE, lane 6: SUMO7, lane7: SUMO7, human SAE, human SCE, lane 8: SUMO1 (Q93A), Lane 9: SUMO1 (Q93A), human SAE, human SCE. The thioester adducts and potential isopeptide bound forms with SAE and SCE proteins and SUMO itself can be visualized as higher molecular weight bands with anti-HA antibody on a Western blot.

### 3.2.2 Sumoylation of nucleosome assembly factor (NAF)

Our attempts to enriched and identify SUMO targets *in planta* (discussed in section 3.4.6) revealed nucleosome assembly factor (NAF), a factor, involved in the biogenesis of nucleosomes, as a potential SUMO substrate *in vivo*. Inspection of the sequence of *Arabidopsis* NAF revealed the presence of a consensus sumoylation site. To test whether sumoylation of NAF occurs *in vitro*, we used an *in vitro* assay in which purified recombinant NAF (FLAG-NAF) was incubated in the presence of recombinant human SUMO-activating enzyme (SAE), SUMO-conjugating enzyme (SCE) and bacterially expressed and purified *Arabidopsis* SUMO1. SDS-PAGE analysis and subsequent immunoblotting with anti-FLAG antibody of the *in vitro*

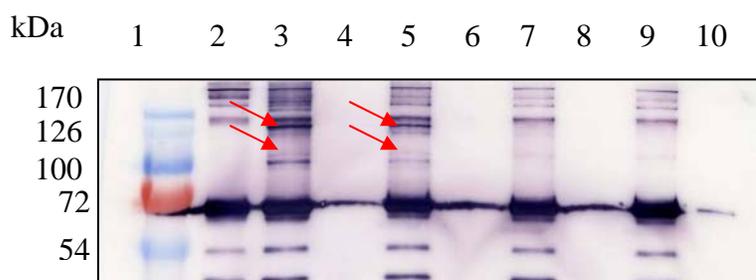
sumoylated products, revealed three slower migrating bands of sumoylated NAF (Fig. 19, marked with red arrow). These higher molecular weight bands were absent when the enzymes of the sumoylation pathway or the substrate itself were not included in the reaction mixture. Several bands detected at a lower molecular weight than the NAF which probably represent the various degraded forms of the protein.



**Fig. 19 Sumoylation of NAF *in vitro*.** Purified NAF was incubated in a sumoylation buffer in presence of human E1 and E2 and recombinant *Arabidopsis* SUMO1. Reactions were separated by SDS-PAGE, and Western blotting was performed using anti-FLAG antibody. Reactions where NAF or SUMO1 were not included in the sumoylation buffer, or where all the components of the SUMO conjugation pathway were absent, served as negative controls. The red arrows mark the position of sumoylated NAF.

Since, SUMO1 was found to target NAF *in vitro*, we also tested the ability of the SUMO1 (Q93A) mutant to conjugate NAF under the same *in vitro* reaction conditions. Western blot analysis with anti-FLAG antibody of the reaction products again showed the presence of slower migrating bands, indicating that the mutant SUMO (Q93A) was equally active in conjugating NAF as the wild type SUMO. In contrast, SUMO3 and SUMO7, which were also shown to form thioester bonds and

possibly SUMO-SUMO isopeptide bonds with human SCE, were not able to sumoylate NAF *in vitro* (Fig. 20). Anti-FLAG antibody could not detect any higher molecular weight band of sumoylated NAF with SUMO3\7, indicating that these proteins may have a different conjugation dynamics or may not modify NAF. Whether NAF is a SUMO1 specific target remains to be determined in a homologous *in vitro* conjugation system, and by *in vivo* studies.



**Fig. 20 Sumoylation of NAF *in vitro* by SUMO isoforms.** Lane 1: prestained protein marker, lane 2: NAF, lane 3: NAF, SUMO1 (WT), human SAE, human SCE, lane 4: SUMO1 (WT), human SAE, human SCE, lane 5: NAF, SUMO1 (Q93A), human SAE, human SCE, lane 6: SUMO1 (Q93A), human SAE, human SCE, lane 7: NAF, SUMO3, human SAE, human SCE, lane 8: SUMO3, human SAE, human SCE, lane 9: NAF, SUMO7, human SAE, human SCE, lane 10: SUMO7, human SAE, human SCE. Western blot was probed with anti-FLAG antibody.

### 3.3 SUMO-conjugating enzyme SCE (C94S) mutant

#### 3.3.1 Phenotypic characterization of plants overexpressing SCE (C94S)

In collaboration with a research group from McGill University, Montreal, we received some *Arabidopsis* transgenic seeds. The transgenic material consisted of lines overexpressing the normal wild type SUMO-conjugating enzyme (SCE) and the lines in which the SCE has been inactivated (SCEIA) due to a mutational change in the active site cysteine at position 94 to serine:

35S: SCEWT > wild type SCE

35S: SCEIA > inactive SCE (C94S)

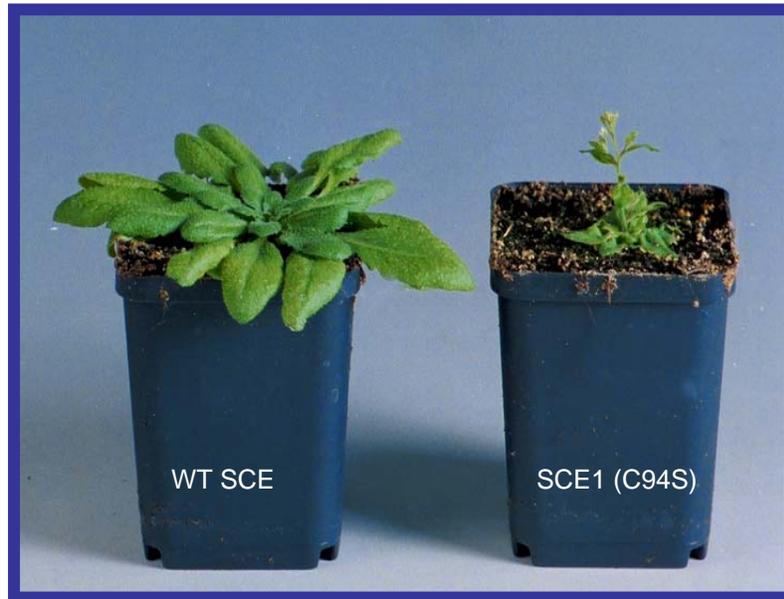
Phenotypic characterization of SCE mutant was done under long and short day light conditions. The growth characteristics of SCE mutant was compared to the plants overexpressing the wild type SCE. Two homozygous lines from each of the wild type and mutant SCE overexpressing lines were used for this study. Seeds from the selected homozygous lines, both of the overexpressing SCE mutant and the wild type SCE were stratified on moist filter paper in the dark at 4°C for 3-5 days to break seed dormancy and then transferred to soil.

- i) Phenotypic characterization of these mutant lines on soil, showed stunted morphology as compared to the wild type (Fig. 21)
- ii) Growth under short day and long day conditions revealed that these plants flower early both under long day and short day light conditions in contrast to the wild type counterparts (Fig. 22). However, the flowering phenotype of the SCE (C94S) mutant was more pronounced under short day growth conditions.



**Fig. 21 Photograph illustrating the phenotype of the SCE (C94S) mutant.** Plants overexpressing SCE1 (C94S) do not grow as well as wild type plants. Picture on the left is the plant overexpressing the wild type SCE. The plant overexpressing the mutant SCE (C94S) on the right shows poor growth.

A.



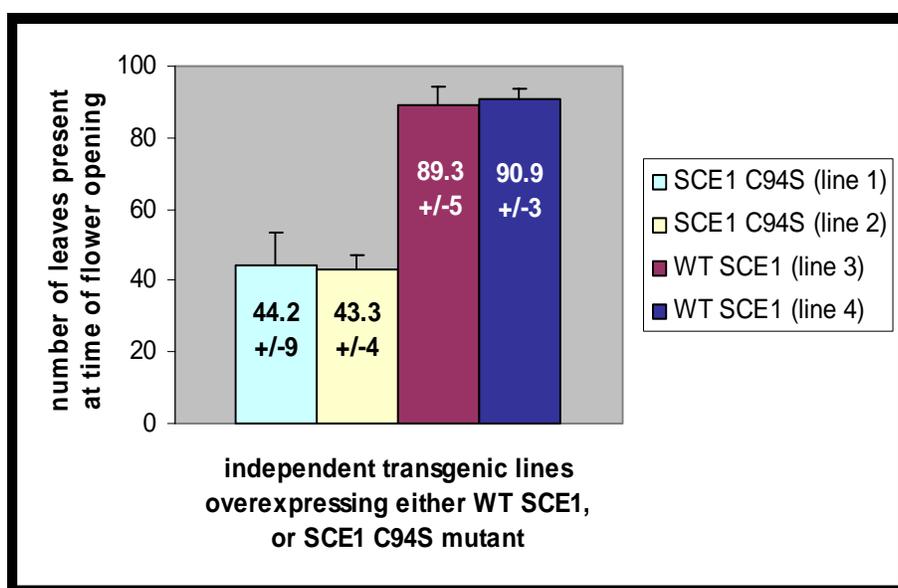
B.



**Fig. 22 Phenotypic characterization of SCE (C94S) mutant under different light conditions.** Plants overexpressing the mutant SCE flower earlier than the plants expressing the wild type SCE both under (A) short day light conditions (8hrs) (B) long day light conditions (16 hrs).

### 3.3.2 Characterization of the flowering time of plants overexpressing SCE (C94S)

At least 8 plants from each of the wild type and mutant SCE overexpressing lines were used to measure flowering time. The number of rosette and cauline leaves were counted before the first flower opened and expressed as a mean  $\pm$  s.d (standard deviation). Fig. 23 shows the statistical data in a graphic representation of the early flowering phenotype of the SCE (C94S) mutant.

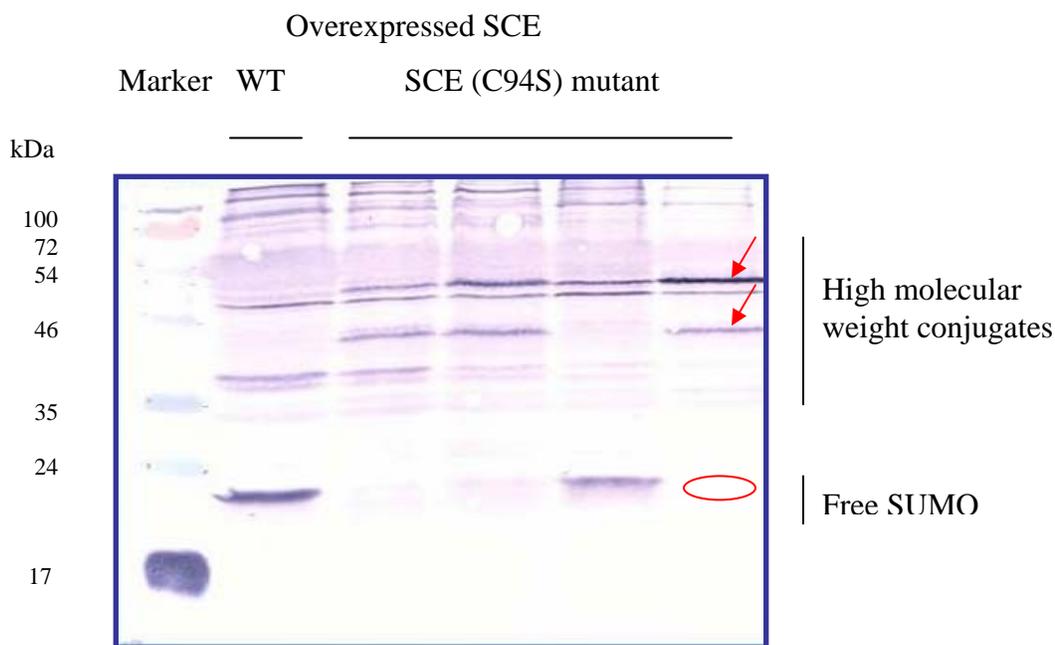


**Fig. 23 Histogram comparing the number of rosette and cauline leaves in the wild type SCE and SCE (C94S) mutant lines.** Plants were grown under short day light conditions.

### 3.3.3 Analysis of the SCE (C94S) lines by immunoblot assays

Total protein was isolated from the leaves of wild type and mutant SCE lines. The protein mixture was separated by SDS-PAGE (12% gel). Western blotting was performed using anti-SUMO1 antibodies to probe crude leaf extract from these lines. A difference in the assortment of SUMO conjugates was evident. Some SUMO conjugates (Fig. 24, bands

marked with red arrow) present in the mutant lines were absent in the wild type. Moreover, depletion in the level of free SUMO was observed in the mutant lines (encircled red in Fig. 24). These observations lead to an hypothesis that the expression of inactive SCE (mutant) seems to trap SUMO in one or two predominant dead end products, e.g. the conjugate of SUMO to SCE (C94S). Thus, we developed antibodies against SUMO-conjugating enzyme (SCE) in order to test the validity of this hypothesis.

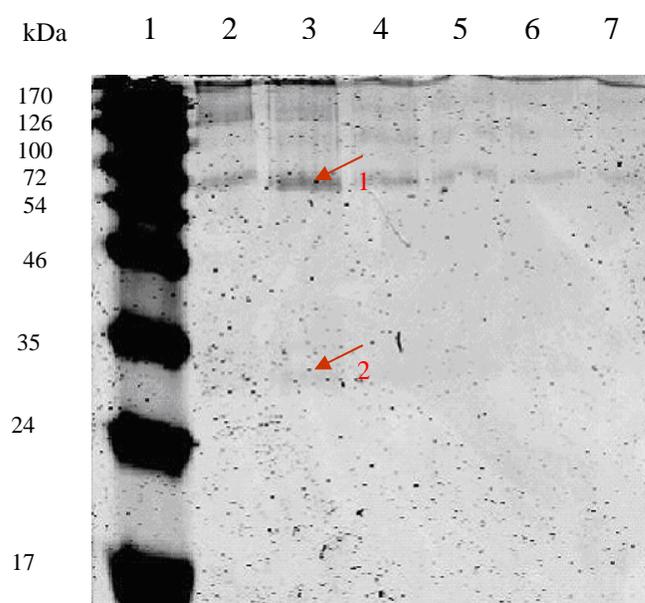


**Fig. 24 Immunoblot assay with protein extracts from SCE overexpressing lines using anti-SUMO1 antibody.** Some SUMO conjugates (bands marked with red arrow) present in the mutant lines were absent from the wild type. Moreover, depletion in the level of free SUMO and lower levels of high molecular weight SUMO conjugates was observed in the mutant lines

### 3.3.4 Purification of rabbit polyclonal immunoglobulins (IgGs)

SUMO-conjugating enzyme antibody developed in rabbit (see section 2.2.5) was purified from a mixture of crude serum proteins using affinity chromatography technique. The serum proteins were first precipitated by ammonium sulphate. The resuspended precipitated serum proteins were applied to an affinity column containing a matrix-bound

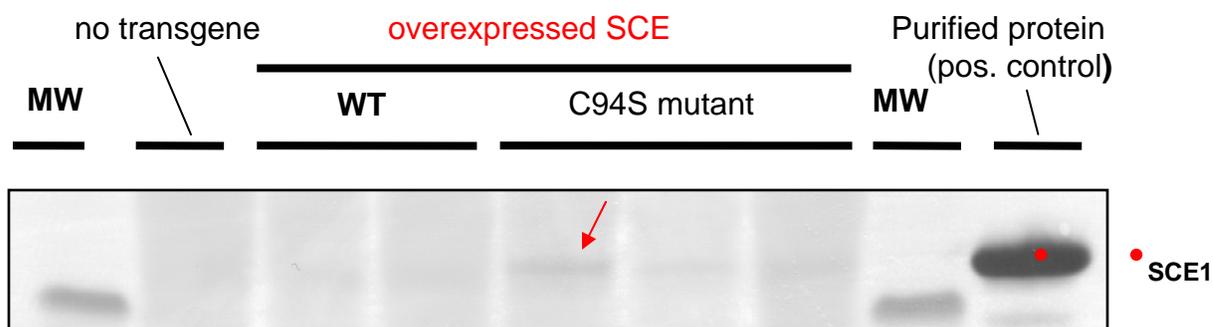
SUMO-conjugating enzyme. The flow through was passed over the column several times for proper binding of the IgGs to the affinity matrix. Bound immunoglobulins were eluted from the column with 0.2M glycine HCl (pH 3.0). The eluted fractions were brought to higher pH and analyzed by SDS-PAGE. Subsequently, the gel was subjected to Coomassie staining. The heavy chain of the IgG was visible in the Coomassie stained gel at a size of ~65 kDa and the light chain migrated at ~ 30 kDa (Fig. 25).



**Fig. 25 Purification of SUMO-conjugating enzyme antibody.** Ammonium sulphate precipitated serum proteins were loaded onto a column containing the ligand (SCE) bound to a matrix. The bound immunoglobulins were eluted in fractions (lanes 2-7) with a low pH buffer. Aliquotes were resolved by SDS-PAGE and visualized by Coomassie staining. Arrow 1 in the gel shows the heavy chain of the purified IgG and arrow 2 points to the light chain.

### 3.3.5 Further characterization of SCE (C94S) mutant

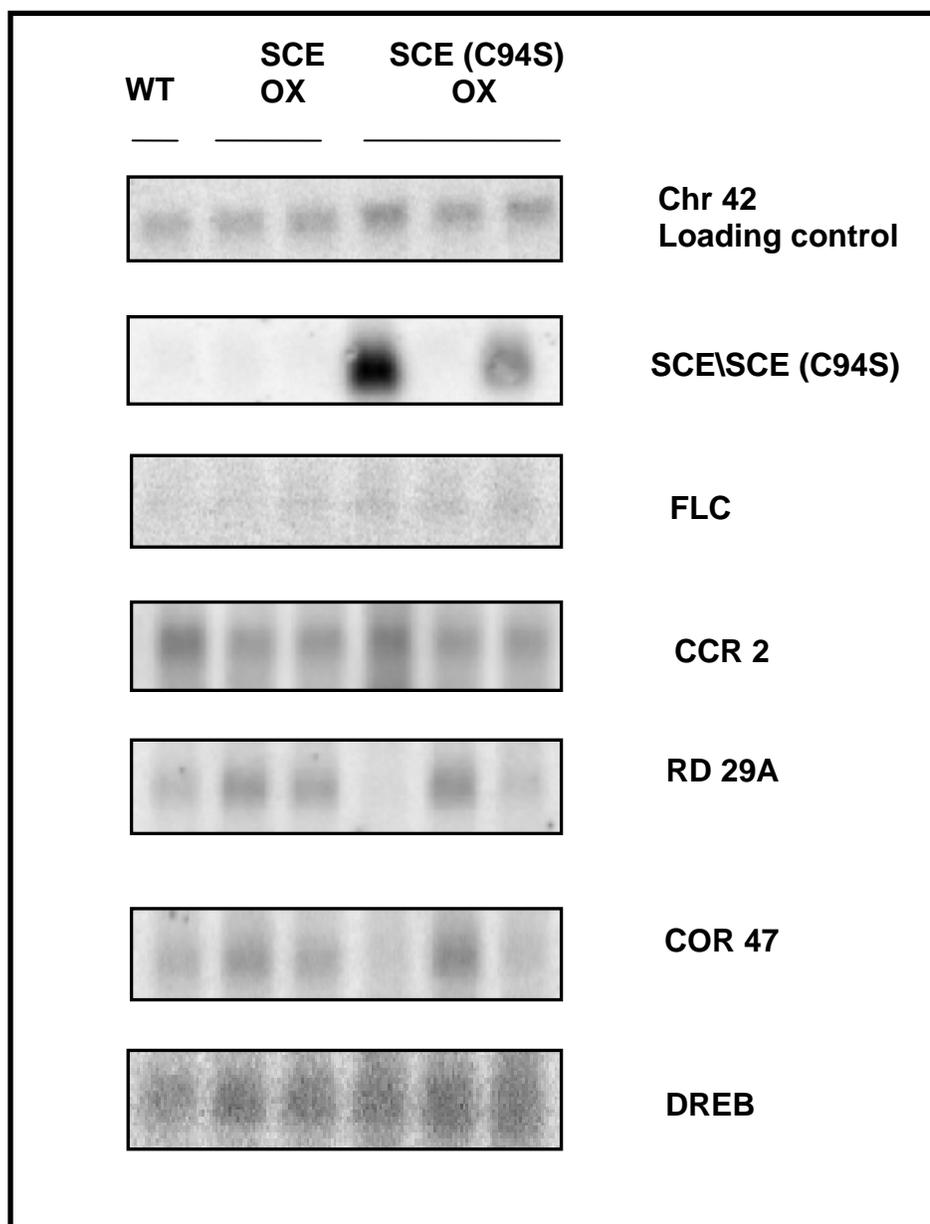
The antibody developed against SCE was used to further probe the SCE transgenic lines. When Immunoblot assays were performed with the same transgenic lines probed with anti-SCE antibody, a prominent (marked with red arrow in Fig. 26) band, likely the free SUMO-conjugating enzyme corresponding to the same molecular weight as the positive control (purified SCE), appeared in the mutant SCE (C94S) line (Fig. 26). Accumulation of free SCE in the mutant overexpressing lines ruled out our hypothesis that SUMO was being trapped as SUMO-SCE conjugate. Therefore, the SCE (C94S) inhibits conjugation probably by non-covalent interactions.



**Fig. 26 Immunoblot with protein extracts from SCE overexpressing lines using anti-SCE1 antibody.** Free SUMO-conjugating enzyme (marked with red arrow) corresponding to the same molecular weight as the positive control (purified SCE) accumulated in the mutant SCE (C94S).

### 3.3.6 mRNA expression pattern of selected genes in SCE (C94S) mutant plants

The SCE (C94S) mutant has a pleiotropic phenotype, which strongly indicated an interference with other gene regulation pathways. We tested the mRNA level of selected marker genes namely FLC (a flowering suppressor), CCR2 (a circadian regulated gene and clock regulator), RD29A, COR47, DREB (cold stress responsive genes) in the wild type SCE overexpressing lines as well as in the lines expressing the mutant SCE (C94S). The mRNA of gene Ch42 was used as a loading control. Total RNA was extracted from the SCE lines, and immobilized on a nylon membrane. DNA of the marker genes was used as a radiolabelled probe. The mRNA level of SCE was variable in each of the homozygous lines used for the study, which explained the varying pattern of expression of the genes under investigation. We did not observe a significant difference in FLC mRNA levels compared with control plants. Likewise, the circadian rhythm gene, CCR2 is expressed similarly in wild type and mutant plants. However, analysis of mRNA levels of two stress responsive genes (RD29A, COR47) revealed that their expression levels were down-regulated in SCE overexpressing plants compared to the plant where no transgene was expressed (Fig. 27).



**Fig. 27 Northern analysis of SCE overexpressing wild type and mutant lines using FLC, CCR2, RD29A, COR47, DREB probes.** RNA was extracted using 20 day old plants from two SCE wild type and three SCE mutant expressing lines. RNA blots were hybridized with probes derived from DNA of FLC, CCR2, RD29, COR 47, DREB genes, respectively. Endogenous Chlorata 42 mRNA which is ubiquitously expressed was used as a loading control and for quantification of the result. OX: overexpression

Transgenic lines	Ratio of signals of probe genes\chlorata 42 control					
	SCE\ SCE(C94S)	FLC	CCR2	RD29A	COR4	DREB
WT	1.74	0.21	61.01	7.35	3.21	0.35
SCE OX T2 line 4	1.87	0.21	37.61	12.6	4.70	0.46
SCE OX T2 line2	2.32	0.21	35.25	9.05	3.25	0.39
SCE(C94S) OX T2 line 5	7.26	0.19	42.40	<u>2.11</u>	<u>1.44</u>	0.33
SCE(C94S) OX T2 line 6	2.0	0.21	33.11	11.96	5.05	0.43
SCE(C94S) OX T2 line 3	2.02	0.21	28.95	<u>5.69</u>	<u>2.05</u>	0.36

**Table 5. Quantification of the genes regulated in plants overexpressing wild type and mutant SCE.** Data indicates that cold stress responsive genes RD29A and COR47, are down-regulated in SCE over expressing plants. OX: overexpression

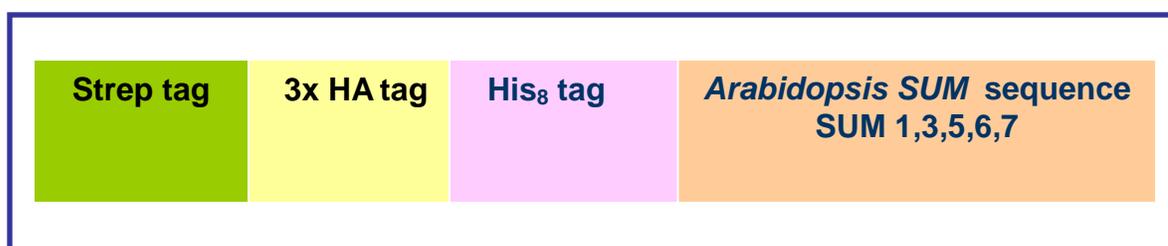
### 3.4 *In vivo* analysis of sumoylation

#### 3.4.1 Expression of SUMO isoforms in *Arabidopsis*

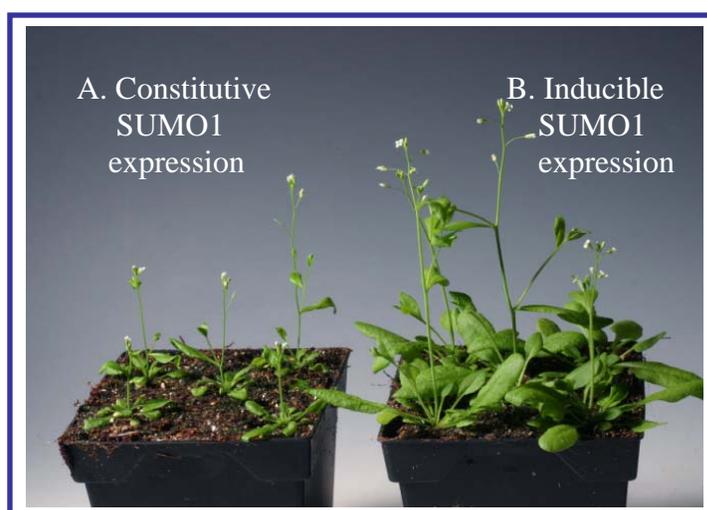
mRNA for four of the eight *Arabidopsis* SUMO genes, SUMO1, SUMO2, SUMO3 and SUMO5 (Fig. 3; Table 1) were identified by RT-PCR of the total RNA from *Arabidopsis* plants. We could not detect mRNAs for SUMO4, SUMO6, SUMO7 and SUMO8 in total RNA, suggesting that they were either not expressed, expressed at low levels, or expression was confined to specific developmental stages of the plant not tested here.

Epitope-tagged *Arabidopsis* SUMO genes namely; SUMO1, SUMO3, SUMO5, SUMO6, SUMO7 (for SUMO6 and SUMO7, intron-containing genomic constructs were used) under the control of a constitutively expressing vector (pHi) were transformed into *Arabidopsis thaliana* by *Agrobacterium* mediated floral dip transformation. The tags included Strep-tag, triple hemagglutinin tag and octa-histidine tag at the amino terminus of the transgene (Fig. 28).

The transformants were selected on *Arabidopsis* germination medium containing hygromycin and 15 day old transgenic seedlings were then transferred onto soil for seed production. Plants expressing intron containing SUMO6 and SUMO7 constructs had normal growth characteristics. Overexpression of SUMO1 also did not result in any dramatically altered phenotype of the plant. However, overexpression of SUMO3 and SUMO5 was toxic to the plant. Thus, inducible transgenic lines both for SUMO3 and 5 (estradiol inducible expression) were established. For comparative purpose, SUMO1 inducible transgenic plants were also generated (Fig. 29).



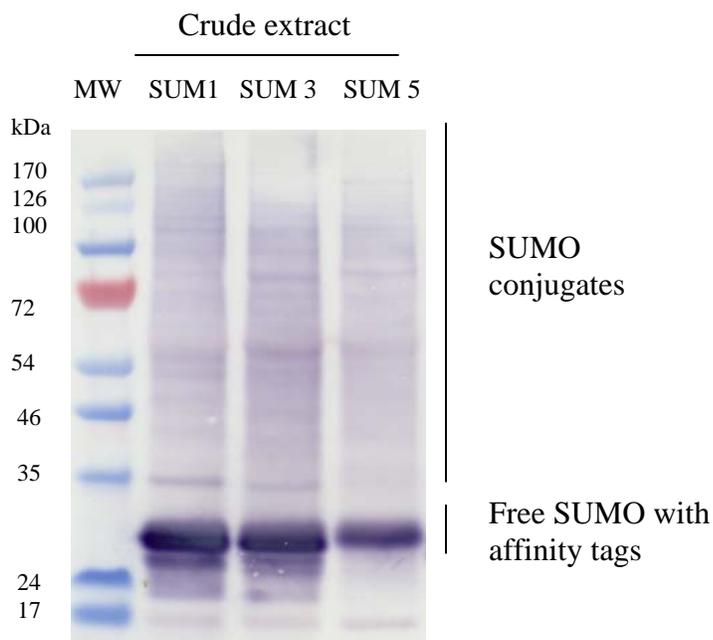
**Fig. 28 Schematic representation of various SUMO transgenes with three affinity tags at the amino terminus.** The SUMO isoforms were individually transformed into *Arabidopsis thaliana*.



**Fig. 29 Photograph illustrating plants overexpressing tagged SUMO1 transgene under the control of (A) constitutive promoter and (B) inducible promoter (estradiol induction; un-induced state).** Plants constitutively expressing SUMO1 have a smaller size than plants in which the expression of SUMO1 can be induced.

### 3.4.2 Sumoylation pattern

Each of the highly expressing SUMO isoforms i.e SUMO1\2, SUMO3, SUMO5 were analyzed for their conjugates using Western blotting technique. Since expression of SUMO3 and 5 resulted in plant death, we established transgenic lines that were under the control of estradiol as an inducer. For comparative purpose, estradiol inducible lines for SUMO1 were also made. Affinity-tagged SUMO1, 3 and 5 constructs with estradiol inducible promoter (pER8) were transformed into *Arabidopsis* by floral dip technique. Transformants were selected on *Arabidopsis* growth medium containing hygromycin and 15 day old transgenic seedlings were then transferred onto soil. Seeds were collected from the F1 generation and sown on soil after confirming their resistance to the selective agent (hygromycin) on germination medium. These transgenic lines were further characterized for the level of expression of the respective transgene after induction with estradiol. 10 days after transfer from germination medium to soil, the transgenic plants were induced ON by spraying with 5 $\mu$ M estradiol (see section 2.2.1.11). Total protein was extracted from each line, and the proteins were resolved by SDS-PAGE. The resolved proteins were transferred to a nylon membrane and probed with anti-HA antibody. The expression of epitope-tagged SUMO1, 3 and 5 allowed detection of conjugates with all these SUMO isoforms (Fig. 30). We did not observe a striking difference in the pattern of conjugates among the three SUMO isoforms. At this point, it is an open question whether or not the various isoforms have a different spectrum of substrates. Although it is difficult to predict the differences in the conjugates of these SUMO isoforms from an immunoblot, we hypothesize that variability in the level of deconjugation among the various SUMOs may be a distinct possibility.



**Fig. 30 Conjugation pattern of SUMO1, 3 and 5 to substrates *in vivo*.** Total protein was extracted from 10 day old transgenic plants expressing tagged versions of SUMO1, 3, and 5 upon O\N induction with 5 $\mu$ M estradiol. SDS-PAGE resolved proteins were transferred to nylon membrane and probed with anti-HA antibody.

### 3.4.3 Spliced junctions of poorly expressed SUMO isoforms

In a recent publication (Kurepa *et al.* 2003), the sequence of presumed intron - exon structures for all the *Arabidopsis* SUMO proteins were depicted based on search for animal SUMO homologs in the *Arabidopsis* protein and DNA databases. Since we had cloned and expressed both SUMO6 and SUMO7 in *Arabidopsis* using genes with putative introns, we experimentally analyzed the intron and exon pattern in the respective genes and found some misassignments in the published data. Forced expression of the intron-containing SUMO7 construct allowed detection of mRNA (Budhiraja R. and Bachmair A., unpublished). The cDNA isolated indicated the formation of two spliced variants, SUM7 and SUM7v. SUM7v has a three amino acid insertion (Glu - Leu - Gln) at the position of the second intron (Fig. 31). Forced expression of SUMO6 confirmed the intron - exon structure predicted by computer algorithms.

```

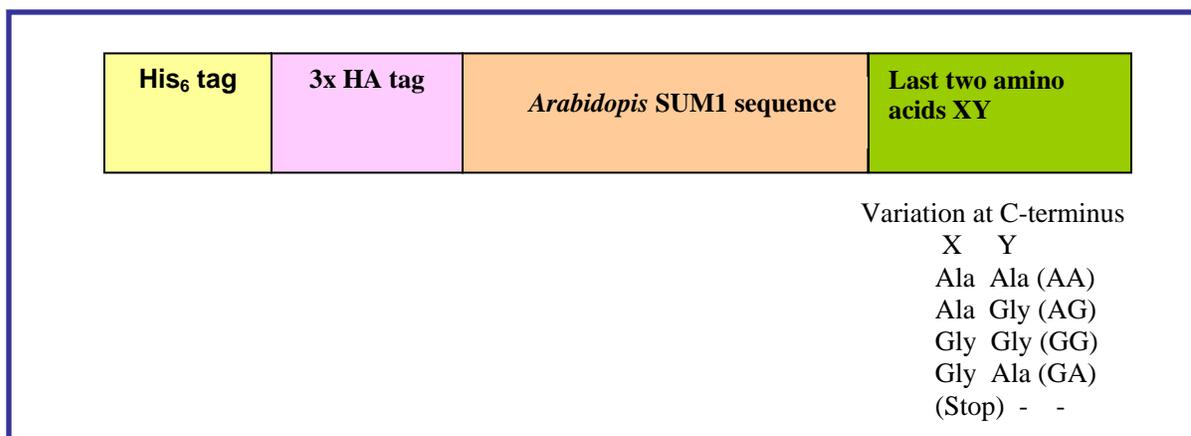
SUM7      MSAADRRPLPPSHITIRISQDDICV.YFRIRRDVELRTMMQAY.SDRVGQQMSAFRFHCDGIRIRPNQTPNE...LDLEDGDEIDA FVDQITG
SUM7v    MSAADRRPLPPSHITIRISQDDICV.YFRIRRDVELRTMMQAY.SDRVGQQMSAFRFHCDGIRIRPNQTPNELQLDLEDGDEIDA FVDQITG

```

**Fig. 31 Spliced variant of SUMO7 differs by a three amino acid insertion.** Amino acid alignment of SUMO7 and SUMO7v of *Arabidopsis thaliana* shows conserved residues with yellow background. A cDNA splicing variant of SUMO7, SUMO7v, contains the three-amino-acid insertion ELQ at the second intron position.

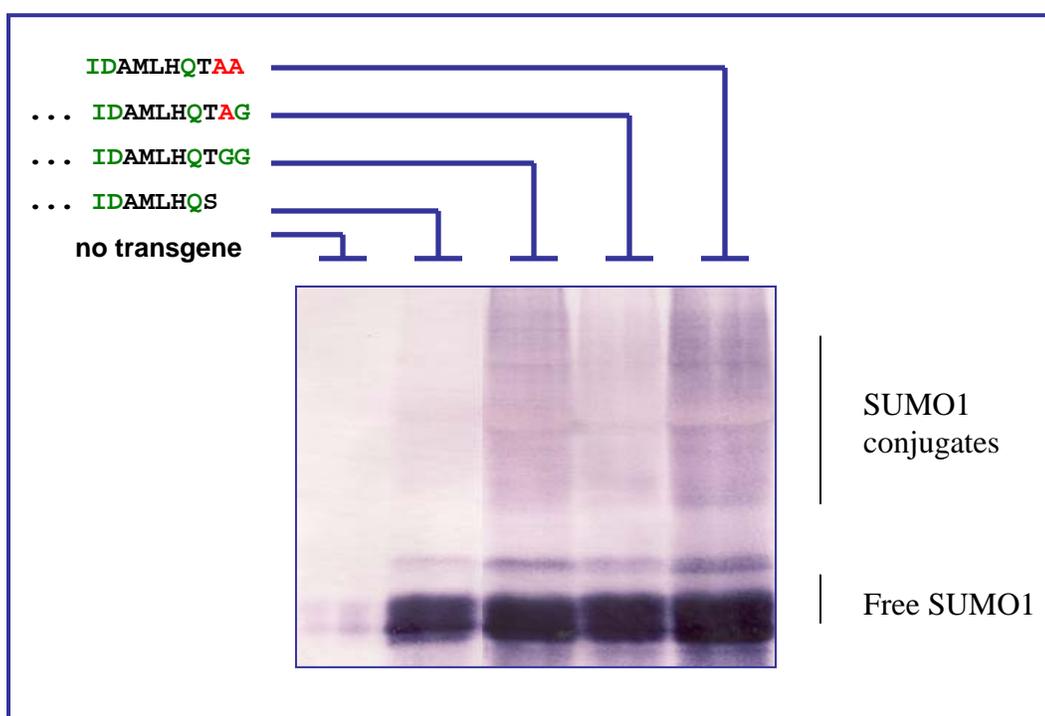
#### 3.4.4 Analysis of C-terminal SUMO variants in *Arabidopsis*

The glycine - glycine (GG) residues that are known to be required at the carboxyl terminus of mature SUMO for its conjugation to the substrates were substituted by the amino acid residues alanine - alanine (AA), alanine - glycine (AG), glycine - alanine (GA), or by the deletion of the last two amino acids in the sequence context of the SUMO1 transgene, respectively (Fig. 32). These individually modified SUMO1 variants were expressed in *Arabidopsis*.



**Fig. 32 Diagrammatic representation of SUMO1 transgene with affinity tags at the amino terminus and variation in the amino acids at the carboxyl terminus.** Individual SUMO variant constructs were expressed in *Arabidopsis thaliana*.

Western blot analysis with crude extracts of the *Arabidopsis* lines with various amino acid residues at the carboxyl terminus showed an array of SUMO1 conjugates with substitutions other than glycine-glycine (GG) (Fig. 33). This finding suggests that the terminal di- Gly motif is not essential for SUMO1 conjugation in *Arabidopsis* which indicates an unexpected flexibility of the plant SUMO conjugation system. This finding perhaps holds true for all other highly expressing SUMO isoforms as well.



**Fig. 33 Immunoblot with anti-HA antibody demonstrates conjugation of SUMO variants to substrate proteins.** Individual constructs of tagged SUMO1 transgene with amino acid variations at the carboxyl terminus were expressed in *Arabidopsis thaliana*.

### 3.4.5 Analysis of SUMO variants potentially inhibiting deconjugation

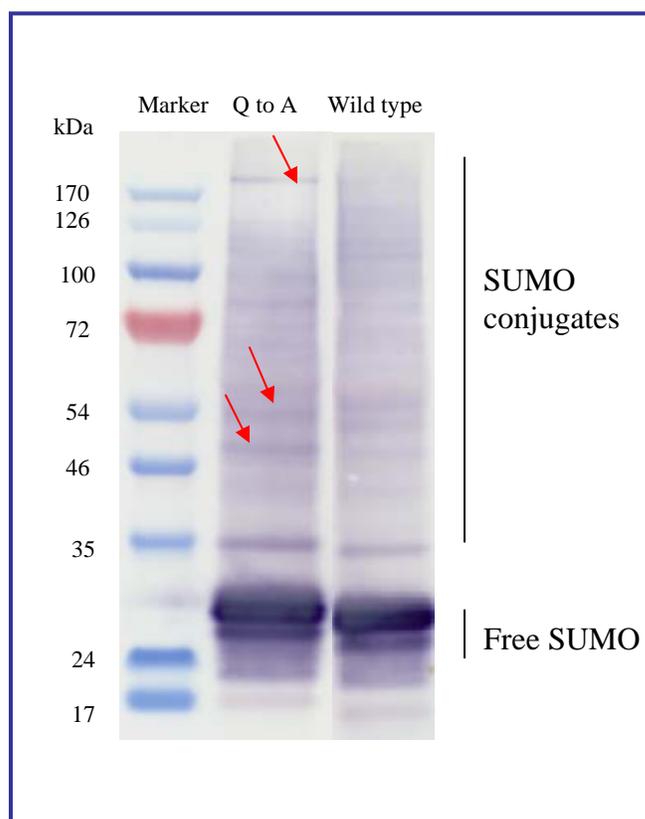
Keeping in mind the information that is available about SUMO and the SUMO protease, especially the aspect that the SUMO specific proteases have close contact to the conserved

glutamine (Q) residue at position 93 of the SUMO protein, this residue was substituted with alanine (A), aspartic acid (D), leucine (L) or arginine (R), respectively, and expressed in *Arabidopsis thaliana*. The expression of some of these constructs (via constitutive promoter) seemed to be fairly harmless to the plant. However, in some cases e.g. the expression of (Q93A) was rather toxic and resulted in poor plant development (Fig. 34). Inducible transgenic lines were made with these SUMO variants (estradiol inducible).



**Fig. 34 Photograph illustrating the phenotype of SUMO (Q93A) mutant.** Expression of mutant SUMO (Q93A) results in poor plant development and early senescence ultimately leading to plant death.

Analysis of these SUMO variant transgenic lines by Western blot analysis detected increased accumulation of certain SUMO conjugates in the (Q93A) variant as compared to the wild type (Fig. 35, red arrows) and the other variants namely (Q93D), (Q93R) and (Q93L), respectively (data not shown).

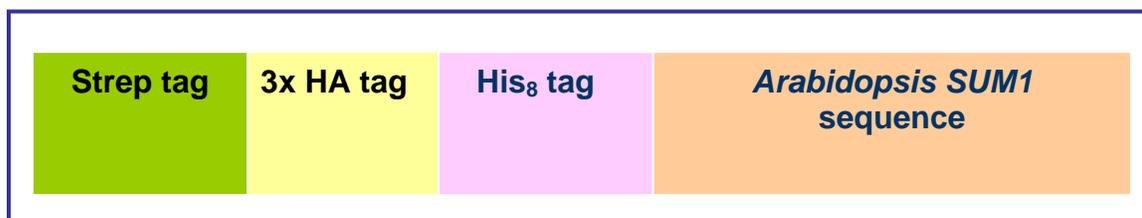


**Fig. 35 Immunoblot with anti-HA antibody shows accumulation of SUMO conjugates in the SUMO (Q93A) mutant.** Constructs with altered residue at a site potentially critical for the action of desumoylation enzymes were expressed in *Arabidopsis thaliana*.

### 3.4.6 Biochemical enrichment of SUMO1 substrates

#### 3.4.6.1 Separation and analyses of enriched extracts

The enrichment of SUMO1 substrates was done using transgenic plants overexpressing the SUMO1 gene. Although, we had established transgenic lines with all isoforms of SUMO, our efforts were mainly concentrated on enriching proteins targeted by SUMO1. Enrichment procedures were based on employing the affinity tags present at the amino terminus of the SUMO1 gene (Fig. 36)



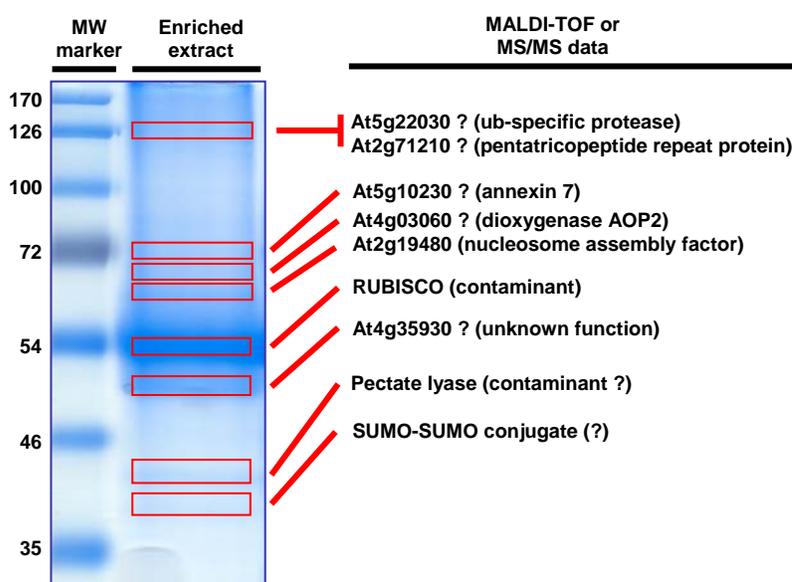
**Fig. 36 Schematic representation of SUMO1 transgene with three affinity tags at the amino terminus.** Expression of amino-terminally tagged SUMO1 facilitated enrichment of conjugates.

The transgene depicted in Fig. 36 was expressed both under a constitutive (pHi) and inducible promoter (pER8) in *Arabidopsis*. A constitutively high expressing line (SUMO1) of the F1 generation was used for biochemical enrichment of the sumoylated proteins. ~200g of SUMO1 overexpressing plants were ground in presence of 6M guanidinium chloride (denaturing conditions). The crude plant extract was cleared of cell debris by centrifugation and filtration. Subsequently, a batch procedure was followed for affinity purification, using first the Ni-NTA resin followed by Strep affinity matrix. However, the second purification step with Strep affinity was unsuccessful and we could not enrich the potentially sumoylated proteins using the Strep Tactin resin. We therefore changed our enrichment strategy and employed HA tag as our second affinity purification step after enriching with the Ni-NTA matrix (Method 1 in material and methods section). Over the time, the constitutive expression of SUMO1 gene was lowered, perhaps due to the silencing of the gene. Hence, we used inducible *Arabidopsis* plants (estradiol induction) overexpressing SUMO1 for our further efforts to enrich SUMO conjugates. The enriched fraction of the tagged SUMO conjugates after Ni-NTA and anti-HA matrix affinity purification was concentrated with a Centricon centrifugal device (Amicon) of 10,000 kDa molecular cut off to a volume of ~50 $\mu$ l. The samples were resolved by one-dimensional SDS-PAGE followed by Coomassie staining of the gel (Fig. 37).

### 3.4.6.2 Identification of sumoylated proteins in enriched fractions by mass spectrometry

#### One-dimensional gel electrophoresis

An array of proteins was observed in the Coomassie stained gel of the enriched SUMO1 conjugates. Distinct bands were cut out from the Coomassie stained gel and the remaining smeared portion was excised into equal size gel pieces. The samples were then digested with trypsin and prepared for MALDI-TOF/MS or LC/MS/MS analysis (section 2.2.4.7). Fig. 37 shows the proteins identified in the mass spectrometric analysis. This analysis identified some candidate SUMO1 targets marked in Fig. 37. We also recovered proteins which were likely contaminants, for example Rubisco peptides were identified in high amounts in most of the enrichment experiments.

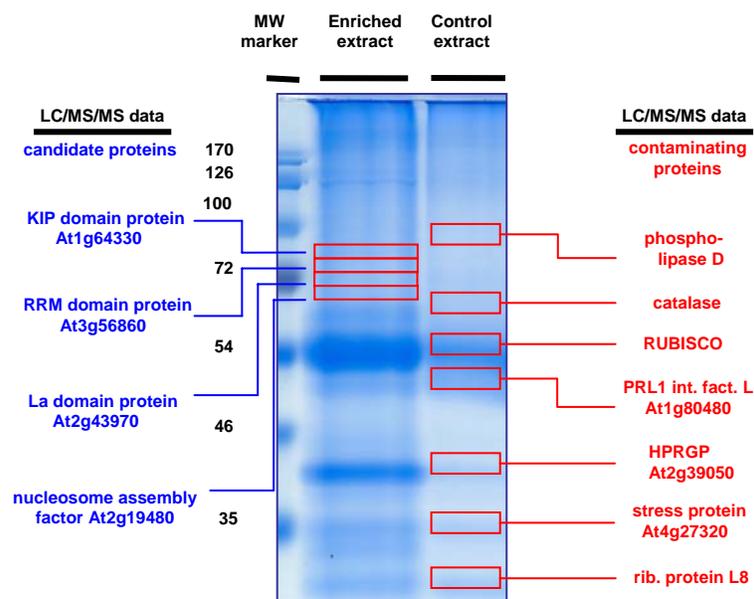


**Fig. 37** Coomassie staining of the sumoylated proteins enriched from *Arabidopsis* plants overexpressing SUMO1. The sumoylated proteins were enriched as described in the text. Red lines point to the identified proteins listed on the right.

---

To rule out the possibility of recovering and enriching contaminants, we performed the same enrichment procedure with wild type *Arabidopsis* plants expressing no transgene as control. The enriched proteins from both the SUMO1 overexpressing and control plants were resolved by a one-dimensional SDS-PAGE (Fig. 38). Corresponding distinct Coomassie stained bands from each of the sample lanes (wild type and SUMO1 enrichment) were excised from the gel and the remaining gel was cut into equal sized 1mm cubes. The proteins were subjected to in-gel digestion and analyzed by LC\MS\MS. Whereas, matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry technique measures mass-to-charge ratio yielding the molecular weight of the peptides, liquid chromatography-mass spectrometry (LC\MS\MS) technique yields the masses and fragmentation pattern of peptides derived from proteins.

LC\MS\MS analysis identified proteins in the control sample that also appeared at the same molecular weight in the SUMO1 enriched fractions. This approach clearly highlighted the contaminants and helped to focus our search and identification on the truly SUMO1 modified proteins. Potential SUMO1 targets were identified from the enriched fractions from plants overexpressing the SUMO1 protein (Table 6 and Fig. 38). Because the covalent attachment of SUMO (15-20 kDa) to a substrate protein is expected to add to its apparent molecular mass by at least 15 kDa, comparison of the observed molecular weight of a candidate SUMO substrate to its theoretical molecular weight provides an immediate indication of whether it is a likely sumoylated substrate. In ideal cases, a SUMO peptide mass should be identifiable among the masses of the identified peptides. Based on this criterion, all these proteins were candidate sumoylation targets. In addition, these specific proteins were not detected in the wild type control enrichment.



**Fig. 38** Coomassie staining of enriched proteins from SUMO1 overexpressing plants and non transgenic wild type as control. Proteins listed on the left are sumoylated candidates. The listing on the right in red represents contaminating proteins.

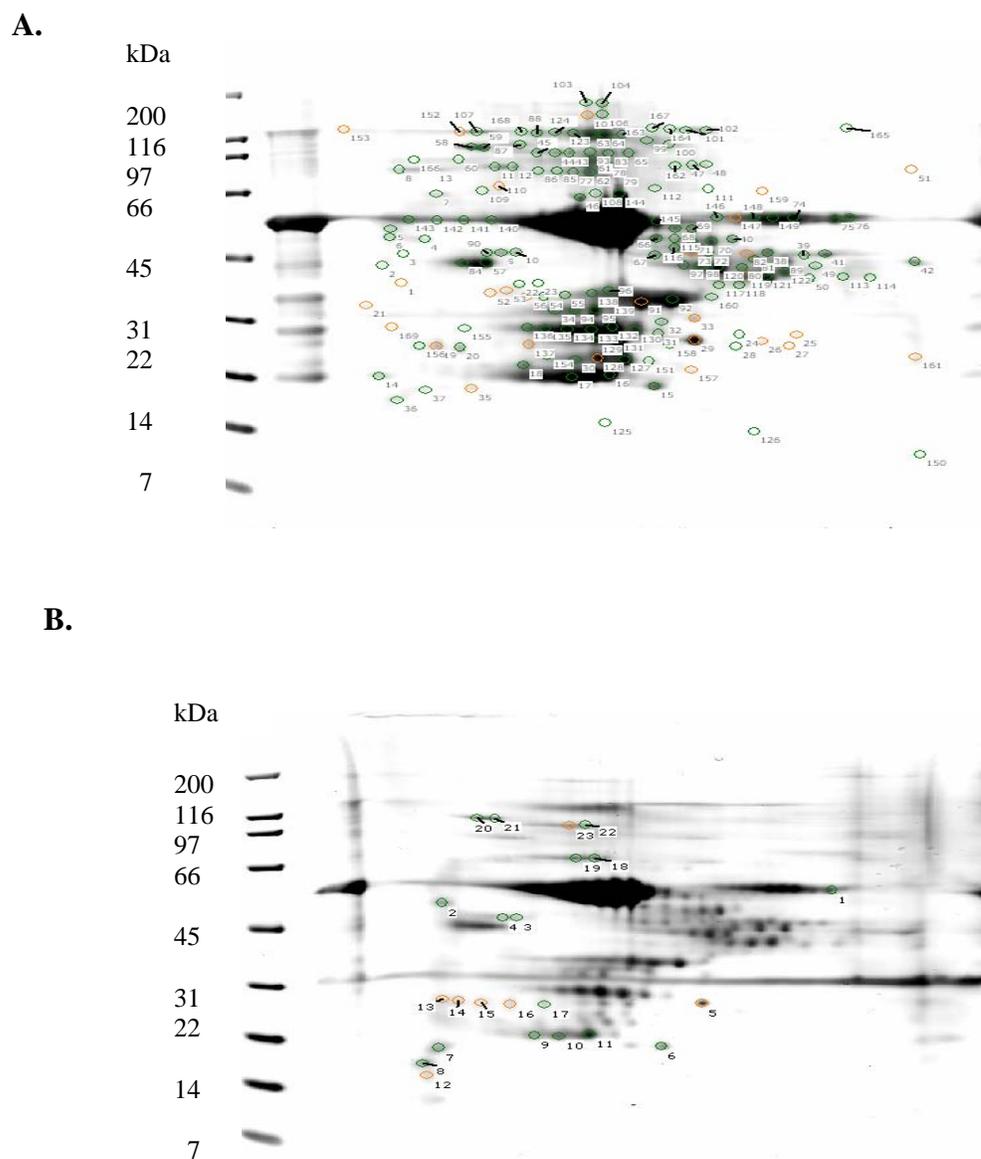
### Two-dimensional gel electrophoresis

To further analyze and identify enriched SUMO1 targets, we performed two-dimensional gel electrophoresis with precipitated samples of the enriched fractions. The precipitated protein mixture pellet was resuspended in sample rehydration buffer and isoelectric focusing was performed. Proteins were then resolved by two-dimensional electrophoresis and stained with Sypro-Ruby stain (Fig. 39, sections 2.2.4.9 and 2.2.4.6). Spots marked 1-168 in panel (A) of Fig. 39 and spots 1-22 in panel (B) were automatically picked, in-gel trypsin digests were done with a robot and analyzed by MALDI-TOF. Almost all spots in the wild type (Fig. 39B) were identified as Rubisco enzyme. Likewise, in the SUMO1 enrichment (Fig. 39A), most of the gel was masked by Rubisco as a contaminating protein. The spots which were identified other than Rubisco were not suitably fitting the criteria chosen for proteins to be labeled as SUMO1 target.

Candidate gene	putative function	MW	MW in gel	score	analysis method	ΨKxE/D
F5O11.15	kinesin-like protein	98.7	ca. 130	29	MALDI-TOF	no
At5g22030	ubiquitin-specific protease (UBP8)	102.6	ca. 130	27	MALDI-TOF	yes
At2g20010	unknown function	107.2	ca. 130	26	MALDI-TOF	no
At1g71210	pentatricopeptide repeat protein	ca. 100	ca. 130	39/25	MALDI-TOF	no
At5g23080	has RNA binding domains	101,6	ca. 130	29	MALDI-TOF	yes
At5g10230	annexin 7	ca. 36	ca. 72	30	MALDI-TOF	yes
At1g33900	avirulence-induced (AIG1)	36.8	ca. 72	30	MALDI-TOF	yes
At3g56720	unknown function	44.3	ca. 72	29	MALDI-TOF	no
At5g63640	vesicular traffic ?	49.5	ca. 72	24	MALDI-TOF	yes
At4g03060	2-oxoglutarate dependent dioxygenase (AOP2)	ca. 48	ca. 70	43	MALDI-TOF	no
At2g19480	nucleosome assembly factor	43.7	ca. 62	183	MS/MS	yes
F28O16.5	Oxophytadienoate reductase (OPR1)	41.2	ca. 62	33	MALDI-TOF	no
At1g11580	pectin methyl esterase	39.4	ca. 62	25	MALDI-TOF	no
At1g27000	bZIP transcription factor	33	ca. 62	24	MALDI-TOF	yes
At4g35930	unknown function	33.7	ca. 50	37	MALDITOF	Yes

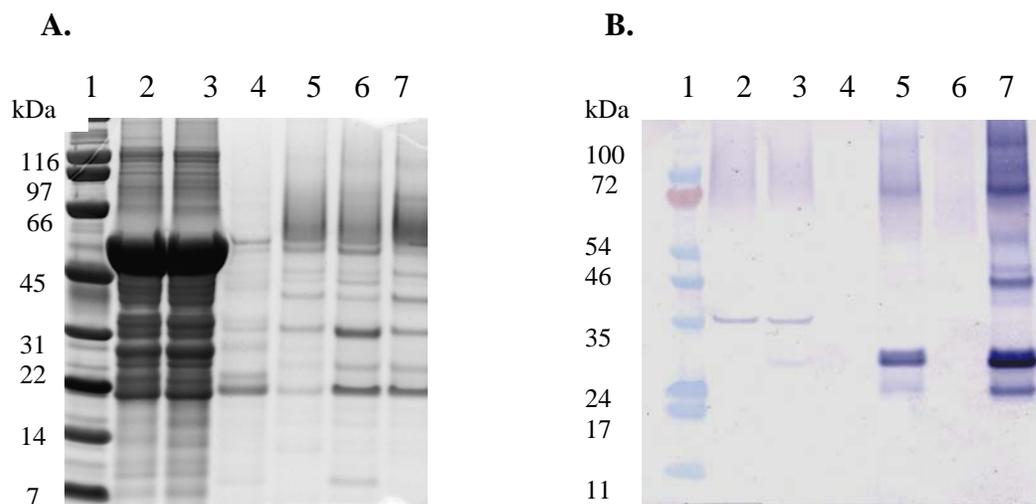
Candidate gene	putative function	MW	MW in gel	score	analysis method	ΨKxE/D
At1g64330	Kinase interaction domain 1 (KIP 1 domain) protein	64,6	ca. 75	108	LC/MS/MS	yes
At2g43970	La domain protein (RNA binding ?)	60	ca. 72	361	LC/MS/MS	no
At3g56860	contains 2 RRM domains (RNA binding)	ca. 50	ca. 70	67	LC/MS/MS	no
At2g41060	contains 2 RRM domains (sister gene to At3g56860)	ca. 50	ca. 70	67	LC/MS/MS	no
At1g27430	GYF domain protein (F17L21.22)	ca. 170	ca. 170	225	LC/MS/MS	NA
At5g43130	TATA binding protein associated factor 4b	ca. 90	ca. 110	129	LC/MS/MS	NA
At5g06790	ABC transporter homolog PrATH-like protein	67	73	93	LC/MS/MS	NA
At3g55760	(F1116.170) No assigned function or domain	56	65	83	LC/MS/MS	NA
At5g08440	(F8L15.170) No assigned function or domain	60-80	ca. 110	62	LC/MS/MS	NA
At1g09730	SUMO protease ULP2like2	113	ca. 180	13	LC/MS/MS	NA

**Table 6. List of potential *Arabidopsis* SUMO1 substrates**



**Fig. 39** Sypro-Ruby stained two-dimensional gels of enriched sumoylated proteins using Ni-NTA and anti-HA antibody matrix. Sypro-Ruby staining of the purified proteins from (A) plants overexpressing the SUMO1 protein and (B) wild type non transgenic plants. 168 spots marked in panel A and 22 spots in Panel B were automatically picked, in gel digested with a robot and analyzed by mass spectrometry. The spots in green colour were identified as protein entities by MALDI-TOF. Spots marked in red colour were not identifiable by MALDI-TOF analysis.

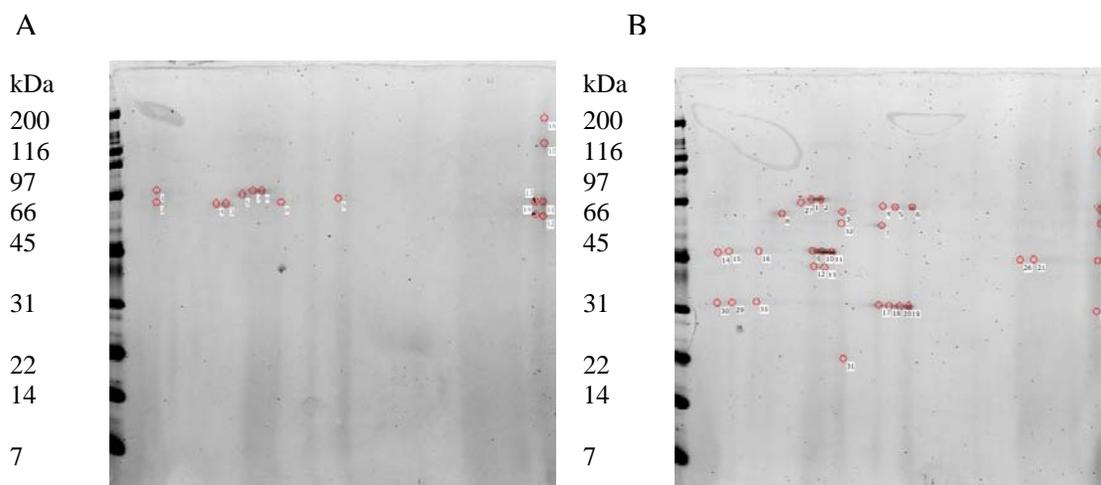
Another method for enriching SUMO1 conjugates (Method 2 in material and methods section 2.2.3) employed using Dyna beads<sup>®</sup> TALON<sup>™</sup> (DYNAL Biotech) which are magnetizable beads developed for the isolation of recombinant histidine tagged proteins (see Fig.13). The biochemical enrichment using this approach was carried out with plants harvested under two different external light parameters for analyzing differences in accumulation or the type of SUMO1 targets, under different external environmental conditions (absence or presence of light). SUMO1 overexpressing plants as well as the wild type control plants were harvested in dark after 8hrs of dark period and also at noon, when the metabolic rate was high. For analytical purpose, enrichment procedure using Ni-NTA and Dyna beads was first carried out with ~ 50 g of plant material harvested in dark (after 8 hrs of dark period) and at noon. Aliquots of enriched fractions were subjected to one-dimensional gel electrophoresis followed by Sypro-Ruby staining of the gel (Fig. 40 A).



**Fig. 40 (A) Sypro-Ruby stained one-dimensional gel showing SUMO1 conjugates after enrichment with Ni-NTA matrix and Dyna beads (B) Western blot of the same with anti-HA antibody.** Plants expressing tagged SUMO 1 transgene and wild type plants were harvested in dark, after 8 hrs of dark period and during noon. Lane 1: molecular weight marker, lane 2: enriched fraction of wild type plants with Ni-NTA matrix (noon), lane 3: enriched fraction of SUMO1 expressing plants with Ni-NTA matrix (noon), lane 4: enriched fraction of wild type plants with Ni-NTA matrix and Dyna beads (dark), lane 5: enriched fraction of SUMO1 expressing plants with Ni-NTA matrix and Dyna beads (dark), lane 6: enriched fraction of wild type plants with Ni-NTA matrix and Dyna beads (noon), lane 7: enriched fraction of SUMO1 expressing plants with Ni-NTA matrix and Dyna beads (noon).

Western blotting was done with the same samples using anti-HA antibody (SUMO contains a 3x HA tag, Fig. 40 B). Clearly, SUMO1 conjugates were visualized on the Sypro-Ruby stained gel using this method and the level of accumulation of SUMO1 conjugates was higher in enriched fractions from SUMO1 overexpressing plants that were harvested at noon.

Since plants harvested under light conditions (at noon) showed an increase in the level of SUMO1 conjugates, the enrichment using this method (see section 2.2.3, method 2) was done with ~250 g of plant material harvested at noon (SUMO1 and wild type). The enriched fraction was concentrated with a Centricon centrifugal device (Amicon) of 10,000 kDa molecular cut off to a volume of ~50 $\mu$ l. The samples were resolved by two-dimensional SDS-PAGE followed by Sypro-Ruby staining of the gel (Fig. 41).



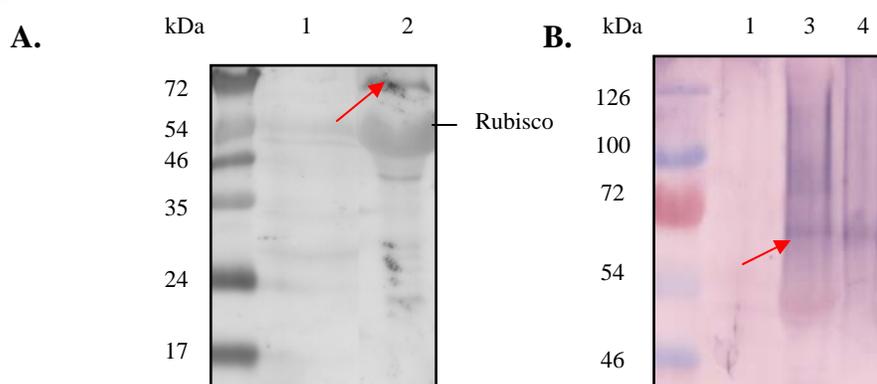
**Fig. 41 Sypro-Ruby stained two-dimensional gels of enriched sumoylated proteins using Ni-NTA matrix and Dyna beads.** Sypro-Ruby staining of proteins from (A) plants overexpressing the SUMO1 protein. (B) Wild type-non transgenic plants. 15 spots marked in panel A and 33 spots in Panel B were automatically picked, in-gel digested with a robot and analyzed by mass spectrometry.

Sypro-Ruby stained spots from the gel were automatically picked; in-gel trypsin digests were performed with a robot and subsequently analyzed by MALDI-TOF for identification of the peptides. This approach, using a double step purification of enriching

histidine tagged proteins exclusively, was successful in the sense that we were able to get rid of Rubisco as a contaminating protein. However, results from the mascot searches did not identify any protein to be suitable characterized as a potential SUMO substrate.

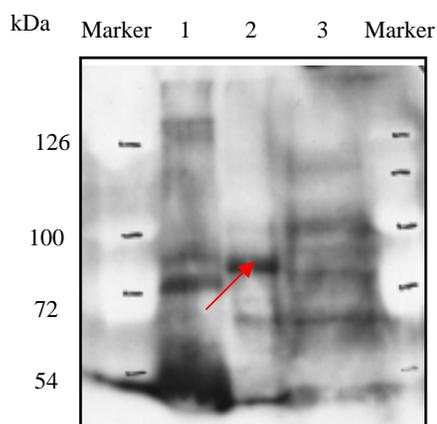
### 3.4.6.3 Identification of sumoylated proteins in enriched fractions using antibodies

We asked whether some SUMO substrates could be identified by methods other than mass spectrometry. PROPORZ1 (PRZ1), is a putative *Arabidopsis* transcriptional adaptor protein essential for the developmental switch from cell proliferation to differentiation in response to intrinsic factors such as phytohormones (Sieberer *et al.* 2003). To verify PRZ1 as sumoylated substrate, we crossed plants overexpressing myc tagged PRZ to plants expressing SUMO1 protein. The progeny containing the selected alleles was used for the enrichment of PRZ1 by affinity purification. The first purification step involved the use of Ni-NTA matrix followed by further enriching the proteins from the Ni-NTA pull down with anti-myc antibody bound matrix. Affinity-enriched proteins were separated by SDS-PAGE. When developed after Western blotting using anti-PRZ1 antibody (Fig. 42), a band of 15-20 kDa higher than the actual molecular weight of the PRZ1 protein was visualized. A higher size band was observed not only in enriched fractions from plants overexpressing the myc tagged PRZ1 and SUMO1, but also from plants expressing tagged SUMO1 alone.



**Fig. 42 Sumoylation of PRZ1 *in vivo*.** Affinity-purified fractions from myc tagged PRZ1 and SUMO1 overexpressing plants were probed with anti-PRZ1 antibody. Red arrows show the position of the sumoylated PRZ1 protein. Lane 1: crude extract from PRZ1 overexpressing plants, lane 2: fraction enriched by Ni-NTA affinity from SUMO1 expressing plants (A), lane 3: Ni-NTA enriched from plants overexpressing PRZ1 and SUMO1, lane 4: double enrichment, Ni-NTA followed by anti-myc antibody bound matrix from plants overexpressing myc-PRZ1 and SUMO (B).

Another potential SUMO target tested for enrichment in the purified fractions was CONTITUTIVE PHOTOMORPHOGENIC 1 (COP1). COP1 is a regulatory protein that suppresses photomorphogenesis in darkness (Deng *et al.* 1991). PVDF membranes with the blotted enriched fractions were sent to Dr. X.W Deng's laboratory at Yale University for probing with COP1 antibody. Western blots with anti-COP1 antibody revealed a band, higher by ~10 kDa than COP1 itself in enriched fractions from plants expressing SUMO1 (Fig. 43). This band was absent in enrichment samples from wild type (non transgenic) plants serving as control.



**Fig. 43 Sumoylation of COP1 *in vivo*.** Enriched fractions of SUMO 1 expressing plants and wild type as control were tested with COP1 antibody. The band marked with red arrow represents the sumoylated COP1. End lanes in the picture are the molecular weight markers. Lane 1: enriched fraction from wild type non transgenic plants and lane 2: represents the enriched fraction from SUMO1 expressing plants, lane 3: Crude extract from wild type non transgenic plants.

## 4. DISCUSSION

### 4.1 *Arabidopsis* SUMO proteins

The family of SUMO proteins recently discovered in yeast and animals represents a class of polypeptide tags that post-translationally modify and thus regulate numerous intracellular proteins in plants. Through searches of the *Arabidopsis* genome data-bases, 25 *Arabidopsis* loci were identified that appear functional as follows: 8 genes encoding SUMO, three SAE, one for SCE, at least one potential E3, and a twelve gene family encoding putative SUMO proteases (Novatchkova *et al.* 2004). The similarity of these protein sequences to those found in yeast and animals indicates that the SUMO pathway has been strongly conserved during eukaryotic evolution. Of the nine *Arabidopsis* SUMO loci, eight (SUMO1-8) are predicted to encode the expected full length proteins of 95-103 amino acids. By comparison, yeast encodes a single SUMO (Smt3), whereas humans encode only four, SUMO1, 2, 3 and 4. By contrast, SUMO9 appears to be a pseudogene. Such pseudogenes have been described in other organisms, for example, there are three predicted SUMO1 pseudogenes in humans and two in mice (Howe *et al.* 1998). Like their yeast and animal counterparts, the *Arabidopsis* proteins bear a long N-terminal region extending beyond the ubiquitin fold and terminate in a 5-10 amino acid sequence that caps the C-terminal glycine essential for attachment. Two proteins (SUMO4, 6) have a consensus sumoylation site ( $\psi$ KXEVD) near their N termini, suggesting that they may be involved in forming polymeric SUMO chains (Tatham *et al.* 2001). Expression studies indicate that at least four of the *Arabidopsis* SUMO genes, SUMO1, 2, 3 and 5, are transcribed and highly expressed, implying that a complex assortment of SUMO isoforms may exist in each cell type. The expression levels of SUMO4, 6 and 8 are presumably much lower. Among the transgenic plants generated, overexpression of SUMO3 and 5 was toxic and resulted in plant death. Plants with increased levels of SUMO1 conjugates showed no gross impairment of general growth (Fig. 29). It is possible that these plants display physiological adaptation to higher levels of sumoylation or that the modified SUMO1 targets are not involved in any basic growth process. Immunoblot assays with antibodies directed against SUMO1\2 (Kurepa *et al.* 2003, Murtas *et al.* 2003) and also those

directed against SUMO3 indicate that these proteins form conjugates *in vivo*. Similarly, expression of our epitope tagged SUMO3 and SUMO5 allows detection of conjugates with these proteins. Thus, all highly expressed SUMO forms in *Arabidopsis* are involved in conjugation reactions. However, an unresolved question is whether various SUMO isoforms have different spectra of substrates. Our experiments with NAF as *in vitro* target of *Arabidopsis* SUMOs showed that NAF is possibly a SUMO1 target and cannot be conjugated to *Arabidopsis* SUMO3 and SUMO7 *in vitro* (Fig. 20). Further investigation and empirical data would be necessary to assess whether or not the various isoforms have diverse spectra of substrates.

#### **4.2 C-terminal SUMO variants expressed in *Arabidopsis thaliana***

Like ubiquitin, SUMO proteins are expressed as precursors that need to be proteolytically processed by proteases to make the C-terminal Gly Gly motif available for conjugation. This cleavage generates the mature form of the protein and occurs after a conserved Gly residue. Most plant SUMO proteins terminate in the same Gly Gly motif at the cleavage site as present in animal and fungal SUMOs, whereas the carboxyl termini of three SUMO proteins (SUMO4, 6 and 7) deviate at the penultimate position. SUMO7 has Ala-Gly, SUMO4 and SUMO6 have Ser-Gly instead. Interestingly, SUMO1 fusion proteins with Ala-Gly or Gly-Ala at this position cannot be processed by SUMO-specific protease ESD4 (Murtas *et al.* 2003). However, when expressed in *Arabidopsis*, mature SUMO1 carrying a Ala Gly at this position is still conjugated to substrates (Figs. 32, 33). This indicates that the changes present in SUMO4, 6 and 7 do not necessarily compromise functionality, although critical parameters of sumoylation and desumoylation may differ from the Gly Gly terminal SUMO isoforms.

#### **4.3 SUMO variants potentially inhibiting deconjugation**

Sumoylation is a dynamic modification controlled by the balance between the activities of SUMO conjugation and SUMO deconjugation machinery. SUMO cleaving enzymes that bring about the desumoylation of existing enzymes are a source of free SUMO and critical for maintaining normal levels of SUMO conjugations because cellular pools of conjugated

SUMOs are presumably very low. From what is known about SUMO and SUMO specific proteases, especially that the SUMO deconjugation enzymes have a close proximity to the glutamine residue at position 93 in the SUMO gene, we examined the role of SUMO proteases in plants by generating mutants with possible defect in deconjugation. Substitution of glutamine at position 93 in the SUMO1 gene with aspartic acid, arginine and leucine (SUMO1-Q93D, Q93R, Q93L) and subsequent expression in *Arabidopsis* was harmless to the plants. However, SUMO1 (Q93A) expression was toxic and resulted in poor growth of the plant with necrotic leaves (Fig. 34). Total protein extracts from inducible plant lines expressing SUMO1 (Q93A) immunoblotted with anti-HA antibody, showed an increase in the population of sumoylated proteins as compared to the unmodified SUMO1 (Fig. 35). Therefore, we reasoned that some SUMO proteases may be less efficient in binding the substituted residue at position 93 in the SUMO protein, which would lead to the inhibition of deconjugation and hence accumulation of the SUMO conjugates.

#### **4.4 Biochemical enrichment of sumoylated proteins of *Arabidopsis***

Numerous SUMO substrates have been identified either individually or through proteomic efforts in mammals and yeast. Two approaches have largely been used for identification of target proteins for SUMO modification. One approach is the detection of slow migrating proteins in sodium dodecylsulphate polyacrylamide gels, since sumoylation should increase the size of target proteins (Matunis *et al.* 1996, Mahajan *et al.* 1997, Buschmann *et al.* 2000). The other approach is the use of the yeast two-hybrid screening method for interaction between SUMO1\Ubc9 and target proteins (Boddy *et al.* 1996, Shen *et al.* 1996, Desterro *et al.* 1999, Müller *et al.* 2000). A variety of sumoylated proteins have been identified with these approaches. The identities of these substrates implicate sumoylation in diverse cellular processes. Intriguingly, many transcription factors\cofactors and components of the chromatin remodeling complexes have been shown to be sumoylated. In plants however, there is very little information on substrates targeted by SUMO. We used an *in vivo* oriented approach, based on detection of slow

migrating proteins in SDS polyacryamide gels to identify plant substrates of the sumoylation pathway. To assist in affinity purification of SUMO modified cellular proteins, we expressed the Strep-HA-His tagged SUMO1 in *Arabidopsis thaliana*. A biochemical approach was undertaken to enrich SUMO1 targets under denaturing conditions using 6M guanidinium chloride and 20mM  $\beta$ -mercaptoethanol in the extraction buffer. The strategy of the enrichment procedure has been illustrated in Fig.13. Briefly, two methods were applied for the enrichment of SUMO1 targets. With both methods, the crude extract of the plants expressing SUMO1 was subjected to batch purification using Ni-NTA resin for enriching proteins carrying a His tag. The partially purified fraction was either further enriched by anti-HA antibody matrix purification (first method) or using Dyna beads (second method). The same enrichment protocol was also followed with the wild type plants that expressed no transgene. This approach gave us a better insight into the contaminating proteins that were pulled down from the purification columns along with the SUMO1 substrate(s). One-dimensional SDS-PAGE and subsequent mass spectrometric analysis identified 25 novel potentially sumoylated proteins *in vivo*. Consistent with several recent studies in other eukaryotic systems, the majority of SUMO1 substrates identified in our screen are involved in transcription, RNA processing and maintenance of genome integrity. However, there are no previous reports of any of the identified proteins as being SUMO1 targets in mammalian, yeast or even plant systems. Analysis of the sequences of the MS identified peptides showed that one half of the identified proteins contain a consensus sumoylation site, although presence or absence of the consensus site does not make the protein a likely SUMO target. Except for At2g19840 (nucleosome assembly factor), every individual attempt at enriching SUMO1 substrates and subsequent analysis by MALDI-TOF and LS\MS\MS, identified a new set of potential SUMO targets, raising the question of reproducibility of our experiments. However, it can be argued that the whole process of SUMO conjugation could be acting through a dynamic cycle of sumoylation and desumoylation, rather than by persistent attachment of SUMO to the substrate. Likewise, the physiological conditions of the plants or exogenous environmental parameters at the time harvest may have a role in differential conjugation of SUMO to various substrates. Another reason for identification

of different substrates in different experiments can be a limiting amount of material, which allows unambiguous identification only for a small percentage of the enriched proteins.

Attempts to resolve the enriched SUMO1 substrate proteins (using the second method of biochemical enrichment, section 2.3.3) by two-dimensional gel electrophoresis succeeded in clearing Rubisco as a major contaminating protein from our enrichments. However, Mascot searches with the mass spectrometry data from this methodology approach did not yet give convincing results to label the identified proteins as SUMO targets. In view of the fact, that the use of Dyna beads as our second methodology approach for enriching SUMO1 substrates, accomplished clearing the enriched fractions from major contaminating proteins such as Rubisco, this step can possibly be adopted as a very initial step for purification from crude extracts. Thereafter, the enriched fractions can be further purified employing the Ni-NTA resin or anti-HA antibody matrix.

#### **4.5 *Arabidopsis thaliana* SCE (C94S) mutant**

Analysis of the transgenic plants with defects in SUMO conjugation elucidated important links of the sumoylation pathway in growth and development of *Arabidopsis*. These plants showed obvious growth impairment. Interestingly, the SCE (C94S) overexpressing plants initiated early flowering as compared to wild type overexpressors. Histone deacetylase complexes (HDAC) which are involved in chromatin modification are known to act as regulators of flowering in *Arabidopsis* (Amasino 2004), and HDAC in animals is subject to post-translational modification by SUMO (David *et al.* 2002). One likely explanation for the flowering phenotype of SCE (C94S) expressing plants is that the role of the SUMO proteins in chromatin structure regulation is necessary for proper flowering time.

Free SUMO pool was depleted in the SCE (C94S) mutants, and less high molecular weight conjugates could be observed in Western blots probed with anti-SUMO antibody (Fig. 24). One possibility was that the mutant version of AtSCE was trapping the free SUMO in a dead end product (SUMO-SCE complexes) and thereby depleting the pools of endogenous free SUMO. However, immunoblots with anti-AtSCE antibody apparently

ruled out this explanation, as bands corresponding to the molecular weight of purified SCE were visualized, indicating the presence of free SCE in the mutant lines. Thus, *in vivo*, conjugation of SUMO was probably inhibited by non covalent interactions.

RNA gel blot analysis showed that mRNA levels of stress responsive genes RD29A and COR47 were down-regulated in the plants overexpressing the mutant SCE (C94S) compared with wild type plants, suggesting a specialized role of SUMO during stress conditions. Differences in the expression of other genes (FLC and CCR2) were less dramatic, probably because the SUMO targets involved in the regulation of these genes are not present or not perturbed sufficiently in plants expressing the mutant SCE (C94S).

#### 4.6 *In vitro* sumoylation assays

*In vitro* data revealed that *Arabidopsis* SUMO proteins 1, 3 and 7 can be activated by the enzymes (human) of the sumoylation pathway resulting in thioester bond formation between SUMO and SAE or SCE, as well as isopeptide bonds between SUMO itself (Fig. 18). Intriguingly, SUMO7 was also found to be active *in vitro* in presence of the human sumoylation machinery although the protein is expressed at a relatively low level *in planta*. Furthermore, using recombinant human SAE and SCE, we report that NAF, a factor involved in assembly of nucleosomes in all eukaryotes, is a potential *in vitro* sumoylation target of SUMO1 in the absence of any SUMO ligase. NAF was identified by mass spectrometric analysis as a potential sumoylated target *in vivo* carrying the consensus  $\psi$ KXE sumoylation site and its being an *in vitro* sumoylation target confirmed the validity of our screen. Relying on SDS-PAGE to resolve potentially sumoylated products, we could visualize 3-4 bands for the sumoylated NAF in immunoblots probed with anti-FLAG antibody, suggesting that NAF is multi-sumoylated. NAF also undergoes modification by sumoylation with SUMO1 (Q93A) deconjugation defective mutant, supporting our results *in vivo*, where SUMO1 (Q93A) mutant was shown to be involved in conjugation reactions. Our mass spectrometric methods however could not define the sumoylation site(s) in NAF, although  $\psi$ KXE is apparently not sufficient to ensure sumoylation. *In vitro* assays of sumoylation of NAF by SUMO3 and SUMO7 did not reveal any slow migrating bands of sumoylated NAF. Thus, it is likely that NAF is

specifically a SUMO1 target *in vitro*, although its conjugation to other SUMO isoforms may depend on conjugation dynamics of the specific isoform.

#### **4.7 Sumoylated substrates identified using antibody targeted approach**

Identification of potential sumoylation substrates using antibodies against the putative target proteins in enriched fractions is an easy and simple method for screening sumoylation proteins. Our studies using this approach with antibodies showed that PRZ1 and COP1 proteins are potential targets for sumoylation in plants. The increased molecular weight of the proteins (ca. by 15-20 kDa) as observed on Western blots probed with the antibody specific against the protein under investigation gave a reasonable indication for being a SUMO target. Although, such analysis offers a faster alternative for tracking sumoylation targets, the availability of specific antibodies against the proteins may impose restrictions and hamper such investigations.

#### **4.8 Future perspective**

The approach presented in this work is the basis on which further strategy toward investigating SUMO conjugation in plants can be designed. Although we have identified some novel potential SUMO substrates *in planta*, investigations of how the sumoylation of these proteins affects the biological processes in plants needs to be deciphered. Several features of the SUMO system, including the low level of modification, the presence of protease activity in case of native lysates, and a number of complex interactions among different enzymes and substrates, combine to make functional analysis challenging. In fact, for some proteins that have been reported to be sumoylated, it is not clear that there is a function, or even that the protein is really sumoylated under endogenous expression levels of SUMO pathway enzymes.

The most important experiment in studying the function of SUMO conjugation to a particular protein is mutational elimination of the SUMO attachment site(s). Overexpression, dominant negative or knockdown experiments involving SUMO pathway enzymes can complement these results, but it is imperative that such experiments are performed with both wild type substrate and the substrate that cannot be sumoylated to confirm correctness of the results.

## 5. SUMMARY

Among the systems that modify protein structure, the covalent attachment of small ubiquitin like modifier - SUMO protein to its substrates (sumoylation) represents after ubiquitylation, the best studied example of post-translational modification by a protein modifier. SUMO is a peptide of approximately 100 amino acids that modifies proteins of many organisms including yeast, humans and plants. SUMO is covalently linked to other proteins via a set of specific enzymes, namely SUMO-activating enzyme (SAE), SUMO-conjugating enzyme (SCE) and SUMO ligases. These enzymes are homologous to E1 - E2 - E3 cascade that operates in ubiquitylation. SUMO is emerging as a versatile modifier for a large number of proteins in many different pathways and the consequences of this modification seem to be as diverse as its targets. *Arabidopsis thaliana* has eight full-length genes with significant similarity to animal and fungal SUMO proteins. Phylogenetic analysis clustered the *Arabidopsis* SUMO proteins into five subfamilies: SUMO1\2, SUMO3, SUMO5, SUMO 4\6 and SUMO 7\8. To identify and characterize SUMO substrates in plants, we developed transgenic *Arabidopsis thaliana* lines expressing tagged versions of all the SUMO isoforms. Of all these genes expressed in *Arabidopsis*, four (SUMO1\2, SUMO3 and SUMO5) are highly expressed and form conjugates with substrate proteins *in vivo*. Following expression of the affinity tagged SUMO1 and subsequent biochemical enrichment, an array of high molecular weight SUMO1 substrates was revealed on a Coomassie stained gel. MALDI-TOF analysis identified 25 novel potential sumoylation targets in *Arabidopsis*, some of which carried the consensus sumoylation motif. We were able to demonstrate that one of these identified sumoylation substrates (NAF) is an *in vitro* target for sumoylation.

All SUMO isoforms are made as inactive precursors. They mature by a carboxyl-terminal proteolytic cleavage which yields the mature modifier with exposed carboxyl terminus di-glycine motif. In order to test the requirement for these residues in conjugation reactions, we substituted the double glycine residues with alanine-glycine, glycine-alanine and alanine-alanine in a SUMO1 transgene. These individually modified SUMO1 transgenes were expressed in *Arabidopsis*. Contrary to the expectation, substrates

were visualized on immunoblotting with all the expressed SUMO variants, which demonstrates the extreme flexibility of the plant SUMO conjugation system. Moreover, an accumulation of SUMO substrates was evident from immunoblot experiments with transgenic *Arabidopsis* plants expressing a SUMO1 (Q93A) mutant. Furthermore, we investigated *Arabidopsis* plants with decreased capacity to conjugate SUMO to its target substrates. These plants expressed a mutated version of SUMO conjugating enzyme (AtSCE) in which the active site cysteine residue was changed to serine (C94S). Phenotypic characterization of these plants deficient in sumoylation showed stunted morphology and early flowering characteristics both under short and long day light conditions as compared to the wild type counterparts. Immunoblot analysis revealed that these transgenic lines had lower levels of free endogenous SUMO. Generally, the results suggest that similar to other eukaryotic organisms, many proteins in plants also undergo post-translational modification via sumoylation and this process has functional significance for development and cell biology of *Arabidopsis thaliana*.

## 6. References

Adamson A.L. and Kenney S. (2001) The Epstein - Barr virus immediate-early protein BZLF1 is SUMO-1 modified and disrupts promyelocytic leukemia (PML) bodies. *J. Virol.* 74, 1224-1233

Al-Khodairy F., Enoch T., Hagan I.M., Carr A.M. (1995) The *Schizosaccharomyces pombe* *hus5* gene encodes a ubiquitin conjugating enzyme required for normal mitosis. *J. Cell Sci.* 108, 475-486

Amasino R. (2004) Take a cold flower. *Nat. Genet.* 26, 111-112

Bachmair A., Novatchkova M., Potuschak T., Eisenhaber F. (2001) Ubiquitylation in plants: a post genomic look at a post-translational modification. *Trends Plant Sci.* 6, 463-470

Bayer P., Arndt A., Metzger S., Mahajan R., Melchoir F., Jaenicke R., Becker J. (1998) Structure determination of the small ubiquitin-related modifier SUMO-1. *J. Mol. Biol.* 280(2), 275-286

Bernier-Villamor V., Sampson D.A., Matunis M.J., Lima C.D. (2002) Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* 108; 345-356

Boddy M.N., Howe K., Etkin L.D., Solomon E., Freemont P.S. (1996) PIC1, a novel ubiquitin-like protein which interacts with the PML component of the multiprotein complex that is disrupted in acute promyelocytic leukemia. *Oncogene* 13(5), 971-982

Bohren, K.M., Nadkarni V., Song J.H., Gabbay K.H., Owerbach D. (2004) A M55V polymorphism in a novel SUMO gene (SUMO-4) differently activates heat shock transcription factors and is associated with susceptibility to type I diabetes mellitus. *J. Biol. Chem.* 279 (26) 27233-27238

Buschmann T., Lerner D., Lee C.G., Ronai Z. (2001) The Mdm-2 amino terminus is required for Mdm2 binding and SUMO-1 conjugation by the E2 SUMO-1 conjugating enzyme Ubc9. *J. Biol. Chem.* 276(44), 40389-40395

Chakrabarti S.R., Sood R., Nandi S., Nucifora G. (2000) Post translational modification of TEL and Tel\AML1 by SUMO-1 and the cell cycle-dependent assembly into nuclear bodies. *Proc. Natl. Acad. Sci. USA* 97, 13281-13285

Clough S.J. and Bent A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735-743

Comerford K.M., Leonard M.O., Karhausen J., Carey R., Colgan S.P., Taylor C.T. (2003) Small ubiquitin-related modifier-1 modification mediates resolution of CREB-dependent responses to hypoxia. *Proc. Natl. Acad. Sci. USA* 100, 986-991

David G., Neptune M.A., DePinho R.A. (2002) SUMO-1 modification of histone deacetylase 1 (HDAC1) modulates its biological activity. *J. Biol. Chem.* 277, 23658-23663

Deng X.-W., Casper T., Quail P.H. (1991) Cop1: a regulatory locus involved in light controlled development and gene expression in *Arabidopsis*. *Genes Dev.* 5, 1172-1182

Desterro J.M., Rodriguez M.S., Hay R.T. (1998) SUMO-1 modification of I $\kappa$ B $\alpha$  inhibits NF- $\kappa$ B activation. *Mol. Cell.* 2, 233-239

Desterro J.M., Rodriguez M.S., Kemp G.D., Hay R.T. (1999) Identification of the enzyme required for the activation of the small ubiquitin like protein SUMO-1. *J. Biol. Chem.* 274, 10618-10624

- 
- Desterro J.M., Thomson J., Hay R.T. (1997) Ubch9 conjugates SUMO but not ubiquitin. *FEBS Lett.* 417, 297-300
- Dohmen R.J. Stappen R., McGrath J.P., Forrova H., Kolarov J. *et al.* (1995) An essential yeast gene encoding a homolog of ubiquitin-activating enzyme. *J. Biol. Chem.* 270, 18099-18109
- Dohmen R.J. (2004) SUMO protein modification. *Biochimica et Biophysica Acta* 1695, 113-131
- Hershko A. and Ciechanover A. (1998) The ubiquitin system. *Ann. Rev. Biochem.* 67, 425-479
- Ho J.C. and Watts F.Z. (2003) Characterization of the SUMO-conjugating enzyme mutants in *Schizosaccharomyces pombe* identifies a dominant negative allele that severely reduces SUMO conjugation. *Biochem. J.* 372, 97-104
- Hochstrasser M. (2000) Evolution and function of ubiquitin-like protein-conjugation systems. *Nature Cell Biol.* 2, E153-E157
- Hoegel C., Pfander B., Moldovan G.L., Pyrowolakis G., Jentsch S. (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* 419, 135-141
- Hotson A., Chosed R., Shu H.J., Orth K., Mudgett M.B. (2003) Xanthomonas type III effector XopD targets SUMO-conjugated proteins in planta. *Mol. Microbiol.* 50, 377-389
- Howe K. (1998) The ubiquitin-homology gene PIC1: characterization of mouse (Pic1) and human (UBC1) genes and pseudogenes. *Genomics* 47, 92-100
- Howe K., Williamson J., Boddy N., Sheer D., Freemont P., Solomon E. (1998) The ubiquitin-homology gene PIC1: characterization of mouse (PIC1) and human (UBL1) genes and pseudogenes. *Genomics* 47, 92-100

Huang J. and Dasso M. (2000) Association of the human SUMO-1 protease SENP2 with the nuclear pore. *J. Biol. Chem.* 277, 19961-19966

Jang M.S., Ryu S.W., Kim E. (2002) Modification of Daxx by small ubiquitin-related modifier-1. *Biochem. Biophys. Res. Commun.* 295, 495-500

Jentsch S. and Pyrowolakis G. (2000) Ubiquitin and its kin: how close are the family ties? *Trends Cell Biol.* 10, 335-342

Johnson E.S. (2004) Protein modification by SUMO. *Annu. Rev. Biochem.* 73, 355-382

Johnson E.S. and Blobel G. (1997) Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. *J. Biol. Chem.* 272, 26799-26802

Johnson E.S. and Blobel G. (1999) Cell cycle-regulated attachment of the ubiquitin-related protein SUMO to the yeast septins. *J. Cell. Biol.* 147, 981-994

Johnson E.S. and Gupta A.A. (2001) An E-3 like factor that promotes SUMO conjugation to the yeast septins. *Cell* 106, 735-744

Johnson E.S., Schweinhorst I., Dohmen. R.J., Blobel G. (1997) The ubiquitin-like protein Smt3 is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *EMBO J.* 16, 5509-5519

Jones D., Crowe E., Stevens T.A., Candido E.P. (2002) Functional and phylogenetic analysis of the ubiquitylation enzymes in *Caenorhabditis elegans*: ubiquitin-conjugating enzymes, ubiquitin activating enzymes, and the ubiquitin-like proteins. *Genome Biol.* 3, research 0002.1-0002.15

- 
- Kagey H.M., Melhuish T.A., Wotton D. (2003) The polycomb protein Pc2 is a SUMO E3. *Cell* 113, 127-137
- Kahyo T., Nishida T., Yasuda H. (2001) Involvement of PIAS1 in the sumoylation of tumor suppressor p53. *Mol. Cell* 8, 713-718
- Kamitani T., Kito K., Nguyen H.P., Fukuda-Kamitani T., Yeh T.H. (1998) Characterization of the second member of the sentrin family of ubiquitin-like proteins. *J. Biol. Chem.* 273(18), 11349-11353
- Kamitani T., Kito K., Nguyen H.P., Wada H., Fukuda-Kamitani T., Yeh E.T. (1998) Identification of three major sentrinization sites in PML. *J. Biol. Chem.* 273, 26675-26682
- Koornneef M., Hanhart C.J., van der Veen J.H. (1991) A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol. Gen. Genet.* 229, 57-66
- Kurepa J., Walker J.M., Smalle J., Gosink, M.M., Davis S.J., Durham T.L., Sung, D-Y., Vierstra R.D. (2003) The small ubiquitin-like modifier (SUMO) protein modification system in *Arabidopsis*. *J. Biol. Chem.* 278, 6862-6872
- Li S.J. and Hochstrasser M. (1999) A new protease required for cell cycle progression in yeast. *Nature* 398, 246-251
- Li S.J. and Hochstrasser M. (2000) The yeast *ULP2 (SMT4)* gene encodes a novel protease specific for the ubiquitin-like Smt3 protein. *Mol. Cell. Biol.* 20, 2367-2377
- Li S.J. and Hochstrasser M (2003) The Ulp1 SUMO isopeptidase: distinct domains required for viability, nuclear envelope localization, and substrate specificity. *J. Cell. Biol.* 160, 1069-1081.

- Lin X., Liang M., Liang Y.Y., Brunicardi F.C., Melchior F., Feng X.H. (2003) Activation of transforming growth factor-beta signaling by SUMO-1 modification of tumor suppressor smad4/DPC4. *J. Biol. Chem.* 278, 18714-18719
- Lois L.M., Lima C.D., Chua N-H (2003) Small ubiquitin-like modifier modulates abscisic acid signaling in *Arabidopsis*. *Plant Cell* 15, 1-13
- Mahajan R., Delphin C., Guan T., Gerace L., Melchior F. (1997) A small ubiquitin related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* 88, 97-107
- Mahajan R., Gerace L., Melchior F. (1998) Molecular characterization of the SUMO-1 modification of RanGAP1 and its role in nuclear envelope association. *J. Cell Biol.* 140, 259-270
- Mannen H., Tseng H.M., Cho C.L., Li S.S. (1996) Cloning and expression of human homolog HSMT3 to yeast SMT3 suppressor of MIF2 mutations in a centromere protein gene. *Biochem. Biophys. Res. Commun.* 222(1), 178-180
- Matunis M.J., Coutavas E., Blobel G. (1996) A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP 1 between the cytosol and the nuclear pore complex. *J. Cell Biol.* 135, 1457-1470
- Meluh P.B. and Koshland D. (1995) Evidence that the MIF2 gene of *Saccharomyces cerevisiae* encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Mol. Biol. Cell.* 6(7), 793-807
- Mendoza H.M., Shen L-n, Blotting C., Lewis A., Chen J., Ink B., Hay R.T. (2003) NEDP1, a highly conserved cysteine protease that deneddylates cullins. *J. Biol. Chem.* 278; 25637-25643

- Miyauchi Y., Yogosawa S., Honda R., Nishida T., Yasuda H., (2002) Sumoylation of Mdm2 by PIAS and RanBP2 enzymes. *J. Biol. Chem.* 18, 50131-50136
- Mossesso E. and Lima C.D. (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
- Müller S., Berger M., Lehembre F., Seeler J.S., Haupt Y., Dejean A. (2000) c-Jun and p53 activity is modulated by SUMO-1 modification. *J. Biol. Chem.* 275(18); 13321-13329
- Müller S., Hoege C., Pyrowolakis G., Jentsch S. (2001) SUMO, Ubiquitin's mysterious cousin. *Nature Rev. Cell Biol.* 2, 202-210
- Murtas G., Reeves P.H., Fu Y-F, 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
- Navatchkova M., Budhiraja R., Coupland G., Eisenhaber F., Bachmair A. (2004) SUMO conjugation in plants. *Planta* 220, 1-8
- Okura T., Gong L., Kamitani T., Wada T., Okura I., Wie C.F., Chang H.M., Yeh E.T. (1996) protection against Fas\APO-1- and tumor necrosis factor-mediated cell death by a novel protein, sentrin. *J. Immunol.* 157 (10), 4277-4281
- Orth K., Xu Z., Mudgett M.B., Bao Z.Q., Palmer L.E., Bliske J.B., Mangel W.F., Staskawicz B., Dixon J.E. (2000) Disruption of signaling by *Yersinia* effector YopJ, a ubiquitin-like protein protease. *Science* 290, 1594-1597

- Pichler A., Gast A., Seeler J.S., Dejean A., Melchior F. (2002). The nucleoporin RanBP2 has a SUMO1 E3 ligase activity. *Cell* 108, 109-120
- Rangasamy D., Woytek K., Khan S.A., Wilson V. G. (2000) SUMO-1 modification of bovine papillomavirus E1 protein is required for intranuclear accumulation. *J. Biol. Chem.* 275, 37999-38004
- Rao-Naik C., de la Cruz W., Laplaza J.M., Tan S., Callis J., Fisher A.J. (1998) The Rub family of ubiquitin-like proteins- crystal structure of *Arabidopsis* Rub1 and expression of multiple Rubs in *Arabidopsis*. *J. Biol. Chem.* 273, 34976-34982
- Reverter D. and Lima C.D. (2005) Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature* 435, 687-692
- Rodriguez M.S., Dargemont C., Hay R.T. (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J. Biol. Chem.* 276, 12654-12659
- Rose A. and Meier I. (2001) A domain unique to plant RanGAP is responsible for its targeting to the plant nuclear rim. *Proc. Natl. Acad. Sci. USA* 98, 15377-15382
- Rui H.L., Fan E., Zhou H. M., Xu Z., Zhang Y., Lin S. C. (2002) SUMO-1 modification of the C-terminal KVEKVD of axin is required for JNK activation but has no effect on Wnt signaling. *J. Biol. Chem.* 277, 42981-42986.
- Saitoh H. and Hinchev J. (2000) Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J. Biol. Chem* 275, 6252-6258
- Saitoh H., Sparrow D.B., Shiomi T., Pu R.T., Nishimoto T., Mohun T.J., Dasso M. (1998) Ubc9p and the conjugation of SUMO-1 to RanGAP1 and RanBP2. *Curr. Biol.* 8, 121-124

Sampson D.A., Wang M., Matunis M.J. (2001) The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. *J. Biol. Chem.* 276, 21664-21669

Sapetschnig A., Rischitor G., Braun H., Doll A., Schergaut M., Melchior F., Suske G. (2002) Transcription factor Sp3 is silenced through SUMO modification by PIAS1. *EMBO J.*, 21, 5206-5215

Schlögelhofer P. and 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

Schmidt R., West J., Love K., Lenehan Z., Lister C., Thompson H., Bouchez D., Dean C. (1995) Physical map and organization of *Arabidopsis thaliana* chromosome 4 *Science* 270, 480-483

Schwienhorst I., Johnson E.S., Dohmen R.J. (2000) SUMO conjugation and deconjugation. *Mol. Gen. Genet.* 263, 771-786

Seeler, J.S. and Dejean, A. (2003) Nuclear and Unclear functions of SUMO. *Nature Rev. Cell Biol.* 4, 690-699

Seufert W., Futcher B., Jentsch S. (1995) Role of ubiquitin-conjugating enzyme in degradation of S- and M-phase cyclins. *Nature* 373, 78-81

Sieberer T., Hauser M. T., Seifert G.J., Luschnig C. (2003) PROPORZI, a putative *Arabidopsis* transcriptional adaptor protein, mediates auxin and cytokinin signals in the control of cell proliferation. *Curr. Biol.* 13, 837-842

- Shen Z., Pardington-Purtymun P.E., Comeaux J.C., Moyzis R.K., Chen D.J. (1996) UBL1, a human ubiquitin-like protein associating with human RAD51\RAD52 proteins. *Genomics* 36(2), 271-279
- Tanaka. K., Nishide J., Okazaki K., Kato H., Niwa O., Nakagawa T., Matsuda H., Kawamukai M., Murakami Y. (1999) Characterization of a fission yeast SUMO-1 homologue, pmt3p, required for multiple nuclear events, including the control of telomere length and chromosome segregation. *Mol. Cell Biol.* 19, 8660-8672
- Tatham M.H., Jaffray E., Vaughan O.A., Desterro J.M., Botting C.H., Naismith J.H., Hay R.T. (2001) Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1\SAE2 and Ubc9. *J. Biol. Chem.* 276, 35368-35374
- Tatusov R.L., Fedorova N.D., Jackson J.D., Jacobs A.R. Kiryutin B., Koonin E.V., Krylov D.M., Mazumdar R., Mekhedov S.L., Nikolskaya A.N., Rao B.S., Smirnov A.V., Vasudevan S., Wolf Y.I., Yin J.J., Natale D.A. (2003) The COG database: an updated version includes eukaryotes *BMC Bioinformatics* 4, 41
- Uchimura Y., Nakao M., Saitoh H. (2004) Generation of SUMO-1 modified proteins in E.coil: toward understanding the biochemistry\structural biology of the SUMO-1 pathway. *FEBS Lett.* 564, 85-90
- Vijay-Kumar S., Bugg C.E., Wilkinson, K.D., Vierstra R.D., Hatfield P.M, Cook W.J. (1987) Comparison of the three-dimensional structure of human, yeast, and oat ubiquitin. *J. Biol. Chem.* 262, 6396-6399.
- Walden H., Podgorski M.S., Schulman B.A. (2003) Insights into the ubiquitin transfer cascade from the structure of the activating enzyme for Nedd8. *Nature* 422, 330-334

Wheeler D.L., Chruch D.M. Edgar R., Federhen S., Helmberg W., Madden T.L., Pontius J.U., Schuler G.D., Schriml L.M., Sequeira F., Suzek T.O., Tatusova T.A., Wagner L. (2004) Database resources of the National Center for Biotechnology Information: update. *Nucleic Acids Res.* 32, D35-D40

Xirodimas D.P., Chisholm J., Desterro J. M., Lane D.P., Hay R.T. (2002) P14ARF promotes accumulation of SUMO-1 conjugated (H)Mdm2. *FEBS Lett.* 528, 207-211

Zuo J., Niu Q.W., Chua N.H. (2000) An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J.* 24, 265-273

## 7. APPENDIX

### Abbreviations

#### General abbreviations

APS	ammonium peroxodisulfate
Amp	ampicillin
AMP	adenosine monophosphate
At	<i>Arabidopsis</i>
ATP	adenosine triphosphate
BAC	bacterial artificial chromosome
BCIP	5-bromo-4-chloro-3-indolyl-phosphate
bp	base pair
BSA	bovine serum albumin
CDD	conserved domain database
cDNA	complementary DNA from transcribed RNA
Ci	Curie
Chr.	chromosome
Col	Columbia
Da	daltons
DEPC	diethylpyrocarbonate
DMF	di-methylformamide
dn	down
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
Fig	figure
FPLC	fast protein liquid chromatography
GM	germination medium
GTE	glucose, tris, EDTA
HA	haemagglutinin (of influenza virus)
HEPES	(N-[2-hydroxyethyl]piperazine-N'-[2' ethanesulfonic acid])
His	histidine
IPTG	isopropylthiogalactoside
Kan	kanamycin
kb	kilobase
kDa	kilodaltons
KOAc	potassium acetate
LC-MS	liquid chromatography-mass spectrometry
LD	long days
LSB	Laemmli sample buffer

---

mcs	multiple cloning site
mRNA	messenger RNA
MALDI -TOF	matrix assisted laser desorption ionisation time of flight
Mb	megabase
MOPS	3-(N-mopholino) propanesulfonic acid
MPIZ	Max-Planck-Institut für Züchtungsforschung
NA	not available
NaOAc	sodium acetate
NBT	4-nitro blue tetrazolium chloride
NCBI	National Center for Biotechnology Information
NCS	newborn calf serum
Ni-NTA	nitriolo triacetic acid matrix charged with Ni <sup>++</sup>
NLS	nuclear localization sequence
NMR	nuclear magnetic resonance
OD	optical density
O\N	overnight
ORF	open reading frame
PCI	phenol: chloroform: isopropanol
PcG	polycomb group
Rif	rifampicin
RNA	ribonucleic acid
Rpm	revolutions per minute
RT	room temperature
SAE	SUMO activating enzyme
SCE	SUMO conjugating enzyme
SD	short days
SDS	sodium dodecylsulphate
SENPs	Sentrin proteases
Strep	streptavidin
SUMO	Small ubiquitin-like modifier
TE	tris, EDTA
Tris	tris(hydroxymethyl)aminomethane
Ub	Ubiquitin
UBL	Ubiquitin-like
UDP	Ubiquitin domain proteins
ULPs	Ubiquitin-like protein processing enzymes
v\v	volume per volume
WT	wild type
Ws	Wassilewskija

**Gene abbreviations**

APG 8	Autophagy protein 8
APG12	Autophagy protein 12
CCR2	Cold circadian rhythm
COR47	Cold regulated 47
CREB	CRE binding protein
CTBP	C-terminal binding protein
Daxx	Fas death domain associated proteins
DREB	DRE (dehydration responsive elements) binding protein
DSK	Drosulfakinin
EB1	End binding protein 1
EB2	End binding protein 2
ESD4	Early in short days 4
FLC	Flowering locus C
GMP	GAP modifying protein 1
HDAC	Histone deacetylase complex
MDM2	Mouse double minute 2
NEDD8	Neural precursor cell expressed developmentally downregulated 8
PC2	Polycomb2
PIAS	Protein inhibitor of activated STAT
PIC	PML interacting protein
PML	Pro myelocytic leukemia
RAD23	Radiation sensitive 23
RanBP2	Ran binding protein
RanGAP1	Ran GTPase-activating protein 1
RD29A	Response to dehydration 29A
RUB1	Related to ubiquitin 1
SMAD4	Mothers against decapentaplegic (Drosophila) homolog 4
SMT3	Suppressor of maintenance 3
SIZ	SAP and Miz
TEL	Translocation E26 transforming-specific leukaemia protein
ULP1	Ubiquitin-like protease 1
ULP2	Ubiquitin-like protease 2

---

## List of Figures and Tables

### List of Figures

- Figure 1** Ribbon diagrams highlight the similarity of the three-dimensional structures of ubiquitin and human SUMO1.
- Figure 2** Phylogenetic tree of the *Arabidopsis* SUMO family.
- Figure 3** Alignment of SUMO protein sequences of *Arabidopsis thaliana*.
- Figure 4** The sumoylation cycle.
- Figure 5** SUMO modified proteins.
- Figure 6** Schematic drawing of pACYC-177.
- Figure 7** Schematic drawing of pBIKS.
- Figure 8** Schematic drawing of pDEST-17.
- Figure 9** Schematic drawing of pER-8.
- Figure 10** Schematic drawing of pET-9a.
- Figure 11** Schematic drawing of pHi.
- Figure 12** Schematic drawing of pQE-30.
- Figure 13** Schematic representation of the biochemical enrichment of SUMO1 conjugates.
- Figure 14** SDS-PAGE analysis of cell lysates from SUMO expressing *E. coli* clones
- Figure 15** Purification of recombinant SUMO proteins.
- Figure 16** Fractionation of the purified *Arabidopsis* SUMO-activating enzyme by SDS-PAGE.
- Figure 17** Purification of *Arabidopsis* SCE.
- Figure 18** *Arabidopsis* SUMO isoforms and SUMO1 (Q94A) mutant are activated by human sumoylation enzymes.
- Figure 19** Sumoylation of NAF *in vitro*.
- Figure 20** Sumoylation of NAF *in vitro* by SUMO isoforms.
- Figure 21** Photograph illustrating the phenotype of the SCE (C94S) mutant.

- 
- Figure 22** Phenotypic characterization of SCE (C94S) mutant under different light conditions.
- Figure 23** Histogram comparing the number of rosette and cauline leaves in the wild type SCE and SCE (C94S) mutant lines.
- Figure 24** Immunoblot assay with protein extracts from SCE overexpressing lines using anti-SUMO antibody.
- Figure 25** Purification of SUMO conjugating enzyme antibody.
- Figure 26** Immunoblot with protein extracts from SCE overexpressing lines using anti-SCE1 antibody.
- Figure 27** Northern analysis of SCE overexpressing wild type and mutant lines using FLC, CCR2, RD29A, COR47, DREB probes.
- Figure 28** Schematic representation of various SUMO transgenes with three affinity tags at the amino terminus.
- Figure 29** Photograph illustrating plants overexpressing tagged SUMO1 transgene under the control of (A) constitutive promoter and (B) inducible promoter (estradiol induction; un-induced state).
- Figure 30** Conjugation pattern of SUMO1, 3 and 5 to substrates *in vivo*.
- Figure 31** Spliced variant of SUMO7 differs by a three amino acid insertion.
- Figure 32** Diagrammatic representation of SUMO1 transgene with affinity tags at the amino terminus and variation in the amino acids at the carboxyl terminus.
- Figure 33** Immunoblot with anti-HA antibody demonstrates conjugation of SUMO variants to substrate proteins.
- Figure 34** Photograph illustrating the phenotype of SUMO (Q93A) mutant.
- Figure 35** Immunoblot with anti-HA antibody shows accumulation of SUMO conjugates in the SUMO (Q93A) mutant.
- Figure 36** Schematic representation of SUMO1 transgene with three affinity tags at the amino terminal.
- Figure 37** Coomassie staining of sumoylated proteins enriched from *Arabidopsis* plants overexpressing SUMO1.

- Figure 38** Coomassie staining of enriched proteins from SUMO1 overexpressing plants and non transgenic wild type as control.
- Figure 39** Sypro-Ruby stained two dimensional gels of enriched sumoylated proteins using Ni-NTA and anti-HA antibody matrix.
- Figure 40** A) Sypro-Ruby stained one-dimensional gel showing SUMO1 conjugates after enrichment with Ni-NTA matrix and Dyna beads (B) Western blot of the same with anti -HA antibody.
- Figure 41** Sypro-Ruby stained two dimensional gels of enriched sumoylated proteins using Ni-NTA matrix and Dyna beads.
- Figure 42** Sumoylation of PRZ1 *in vivo*.
- Figure 43** Sumoylation of COP1 *in vivo*.

### **List of Tables**

- Table 1** Listing and sequence characterization of *Arabidopsis thaliana* SUMO-related proteins and predicted proteins involved in SUMO conjugation.
- Table 2** SUMO specific proteases in *Arabidopsis* and related proteins in yeast.
- Table 3** Listing of the *Arabidopsis thaliana* plants used in this study.
- Table 4** Listing of bacterial plasmid constructs used in this study.
- Table 5** Quantification of the genes regulated in plants overexpressing wild type and mutant SCE.
- Table 6** List of potential *Arabidopsis* SUMO1 substrates.

## Sequence of *Arabidopsis* SUMO transgenes

### pSK tag 3 SUMO1 insert (524bps)

```

1  CTCGAGAATTACTATTTACAATTACCATGGCTTCATGGTCTCATCCACAATTCGAAAAGG
61  GTGCTGAAAATATGGGATCCTACCCATACGATGTTCCCTGACTATGCGGGCTATCCCTATG
121  ACGTCCCGGACTATGCAGGATCCTATCCATACGACGTTCCAGATTACGCTGGTACTCATC
181  ACCATCATCATCACCATCATGCTGGTACCCATATGTCTGCAAACCAGGAGGAAGACAAGA
241  AGCCAGGAGACGGAGGAGCTCACATCAATCTCAAAGTCAAGGGACAGGATGGAAACGAGG
301  TTTTCTTTAGGATCAAGAGAAGCACTCAGCTCAAGAAGCTGATGAATGCTTACTGTGACC
361  GGCAATCTGTGGACATGAACCCATTGCTTTCTGTTTGTATGGGCGTCGCTTTCGTGCTG
421  AGCAAACCTCCCGATGAGCTTGACATGGAGGATGGTGTATGAGATCGATGCGATGCTCCATC
481  AGACTGGTGGCTGATCTAGAGCGGCCGCCACCCGGTGGAGCTC

```

### pSK tag3 SUMO3 insert (522 bps)

```

1  CTCGAGAATTACTATTTACAATTACCATGGCTTCATGGTCTCATCCACAATTCGAAAAGG
61  GTGCTGAAAATATGGGATCCTACCCATACGATGTTCCCTGACTATGCGGGCTATCCCTATG
121  ACGTCCCGGACTATGCAGGATCCTATCCATACGACGTTCCAGATTACGCTGGTACTCATC
181  ACCATCATCATCACCATCATGCTGGTACCATGTCTAACCCCTCAAGATGACAAGCCCATCG
241  ACCAAGAACAAGAAGCTCATGTCAATTCCAAGGTCAAGAGCCAGGATGGAGACGAAGTCT
301  TATTTAAGAACAACAAAAAGCGCTCCACTTAAAAAGCTCATGTATGTTTACTGCGACCGCC
361  GAGGTTTGAAAATTAGACGCATTCGCTTTCATTTTTAATGGAGCTCGTATAGGTGGCCTGG
421  AGACTCCAGATGAGCTTGATATGGAAGATGGAGATGTGATAGATGCGTGCCGGGCGATGA
481  GTGGTGGTTGATTCTAGAGCGGCCGCCACCCGGTGGAGCTC

```

### pSK tag3 SUMO5 insert (552bps)

```

1  CTCGAGAATTACTATTTACAATTACCATGGCTTCATGGTCTCATCCACAATTCGAAAAGG
61  GTGCTGAAAATATGGGATCCTACCCATACGATGTTCCCTGACTATGCGGGCTATCCCTATG
121  ACGTCCCGGACTATGCAGGATCCTATCCATACGACGTTCCAGATTACGCTGGTACTCATC
181  ACCATCATCATCACCATCATGCTGGTACCATGGTGGAGTTCACAGACACAATCTCTGCTT
241  CATTTGTATCAAAGAAGTCTCGATCTCCTGAAACATCACCCATATGAAAGTCACTCTCA
301  AGGTCAAGAACCAACAGGGAGCAGAGGATTTGTATAAAAATGGAACCTCATGCACATCTAA
361  AGAAACTAATGAGTGCTTACTGTACGAAGAGAACTTAGATTACAGTTCTGTCCGATTCCG
421  TCTACAATGGTAGAGAAATCAAAGCTCGACAGACTCCTGCTCAGCTGCACATGGAAGAAG
481  AAGATGAGATCTGTATGGTTCATGGAACCTGGTGGTGGCTGATCTAGAGCGGCCGCCACC
541  GCGGTGGAGCTC

```

**pSK tag3 SUMO6 insert (567bps)**

1 CTCGAGAATTACTATTTACAATTACCATGGCTTCATGGTCTCATCCACAATTTCGAAAAGG  
 61 GTGCTGAAAATATGGGATCCTACCCATACGATGTTCCCTGACTATGCGGGCTATCCCTATG  
 121 ACGTCCCGGACTATGCAGGATCCTATCCATACGACGTTCCAGATTACGCTGGTACTCATC  
 181 ACCATCATCATCACCATCATGCTGGTACCATGTCAACGAAGAGCAGTAGTATTCATGGAA  
 241 GGAATGAAGTGAAGATGGAAGGGGAAAAGCGTAAAGACGTTGAGTCGGAATCAACTCATG  
 301 TTACTCTGAATGTGAAGGGTCAAGATGAGGAAGGGGTTAAAGTCTTCCGGGTGAGAAGGA  
 361 AGGCTAGGCTTCTTAAATGATGGAATATTACGCCAAAATGAGAGGCATAGAGTGGAACA  
 421 CATTTGCTTTCTATCTGATGATGGCTCAAGAATTCGAGAGTATCATAACAGCTGATGACA  
 481 TGGAGCTGAAAAGATGGAGATCAAATCGATGCATTGTTGCCTCAAGAAAGTGGTTGATTCT  
 541 AGAGCGGCCGCCACCCGGTGGAGCTC

**pSK tag3 SUMO7 insert (513bps)**

1 CTCGAGAATTACTATTTACAATTACCATGGCTTCATGGTCTCATCCACAATTTCGAAAAGG  
 61 GTGCTGAAAATATGGGATCCTACCCATACGATGTTCCCTGACTATGCGGGCTATCCCTATG  
 121 ACGTCCCGGACTATGCAGGATCCTATCCATACGACGTTCCAGATTACGCTGGTACTCATC  
 181 ACCATCATCATCACCATCATGCTGGTACCATGTCCGACGCTGACAAAAAACCGTTGATTC  
 241 CGCCGTCACATATCACCATCAAAATCAAAAAGTCAGGATGACATATGTGTGACTTCCGGA  
 301 TTAAGAGGGACGTTGAGCTCCGTACGATGATGCAAGCATATTCCGACAAAGTTGGACAAC  
 361 AAATGTCAGCTTTTAGGTTTCACTGCGATGGAATTAGGATCAAACCCAATCAAACCTCCA  
 421 ATGAGCTTGATCTTGAAGACGGAGATGAAATCGATGCATTGTTGACCAAATAGCAGGGT  
 481 GATTCTAGAGCGGCCGCCACCCGGTGGAGCTC

**SUMO6+introns+flanking sequences (660bps)**

1 ACGTAATCATGCTTTGAGAAGTGAATTATCTGATCCAACCAGTGGAGATAATGTCAACGA  
 61 AGAGCAGTAGTATTCATGGAAGGAATGAAGTGAAGATGGAAGGGGAAAAGCGTAAAGACG  
 121 TTGAGTCGGAATCAACTCATGTTACTCTGAATGTGAAGGGTCAAGTAAGATTTGATTGTT  
 181 GAAACTATACAATTTCTATATTAGCATGCATGACTTAATATAACATGTGTGGATTTATGT  
 241 AGGATGAGGAAGGGGTTAAAGTCTTCCGGGTGAGAAGGAAGGCTAGGCTTCTTAAATTGA  
 301 TGGAAATATTACGCCAAAATGAGAGGCATAGAGTGGAACACATTTTCGCTTTCTATCTGATG  
 361 ATGGCTCAAGAATTCGAGAGTATCATAACAGCTGATGACGTATGTATTATCATTCAAGCTT  
 421 AAAGATTCCTTATAATATATAAAGATAAATCATTCATAAATGATCTTTTTTCACTGTGTTT  
 481 TCTCAGATGGAGCTGAAAGATGGAGATCAAATCGATGCATTGTTGCCTCAAGAAAGTGGT  
 541 TTTGGTCTTCTACAGTATTTAGAGTTTGTGTTTCTGGTGTGAATGTAATGTTGGAAT  
 601 TGGTCAAAGTTGAAGTCGTTTTCTGAGTATTTTCGTATCACGTTTAGCTGAATAAGTTAC

**SUMO7+introns+flanking sequences (720bps)**

1 TGTGAAGATGTCGGCAGCTGACAAAAACCGTTGATTCCGCCGTCACATATCACCATCAA  
 61 AATCAAAAGTCAGGTACTTGTAAATAATCTTAATGTTATTATAAGTTAAATTTTTTCTAA  
 121 AAAGACCACCTTCTTGTATGCTATTGTGAAAACAATAAAATTTCCATACACAAGTATTT  
 181 GGAGATTATGCAAAACAGCAAAGAATTCATTGATTCTAAGTGCTAGGTTTCTAGGTTTTT  
 241 TTCTGTTGGTTAAATCGTTGACAGAAAATAACATATTTTCTATAAGTGAATACTTTCTAC  
 301 AACCTCAAAGATGATAATATTTTAAATTTTAAACCAATCTAAAATTAAGCTGAGTTAACAT  
 361 GCGTATAGATTAACATGTTAAATTTTTGTAAATATAATAAAATAGGATGACATATGTGTG  
 421 TACTTCCGGATTAAGAGGGACGTTGAGCTCCGTACGATGATGCAAGCATATTCGGACAAA  
 481 GTTGGACAACAAATGTCAGCTTTTAGGTTTCACTGCGATGGAATTAGGATCAAACCAAT  
 541 CAAACTCCCAATGAGGTATTTATTTATTTTGTGCATAGAAATGTTACATATGAATCATAT  
 601 TGAGAAACAAAAGGAATTGCAGCTTGATCTTGAAGACGGAGATGAAATCGATGCATTTGT  
 661 TGACCAATAGCAGGGTTCAGCCATCGCCATTAAGAGTATCTTAACGATATAAAAAGTCA

**Mature SUMO7 ORF from cDNA (270bps)**

1 ATGTCGGCAGCTGACAAAAACCGTTGATTCCGCCGTCACATATCACCATCAAAATCAA  
 61 AGTCAGGATGACATATGTGTGACTTCCGGATTAAGAGGGACGTTGAGCTCCGTACGATG  
 121 ATGCAAGCATATTCGGACAAAGTTGGACAACAAATGTCAGCTTTTAGGTTTCACTGCGAT  
 181 GGAATTAGGATCAAACCAATCAAACCTCCCAATGAGGATGATCTTGAAGACGGAGATGAA  
 241 ATCGATGCATTTGTTGACCAAAATAGCAGGG

**Mature SUMO7v ORF from cDNA (279bps)**

1 ATGTCGGCAGCTGACAAAAACCGTTGATTCCGCCGTCACATATCACCATCAAAATCAA  
 61 AGTCAGGATGACATATGTGTGACTTCCGGATTAAGAGGGACGTTGAGCTCCGTACGATG  
 121 ATGCAAGCATATTCGGACAAAGTTGGACAACAAATGTCAGCTTTTAGGTTTCACTGCGAT  
 181 GGAATTAGGATCAAACCAATCAAACCTCCCAATGAGGAATTGCAGCTTGATCTTGAAGAC  
 241 GGAGATGAAATCGATGCATTTGTTGACCAAAATAGCAGGG

**Mature SUMO6 ORF from cDNA (324bps)**

1 ATGTCAACGAAGAGCAGTAGTATTCATGGAAGGAATGAAGTGAAGATGGAAGGGGAAAAG  
 61 CGTAAAGACGTTGAGTCGGAATCAACTCATGTTACTCTGAATGTGAAGGGTCAAGATGAG  
 121 GAAGGGGTTAAAGTCTTCCGGGTGAGAAGGAAGGCTAGGCTTCTTAAATTGATGGAATAT  
 181 TACGCCAAAATGAGAGGCATAGAGTGAACACATTTTCGCTTCTATCTGATGATGGCTCA  
 241 AGAATTCGAGAGTATCATACAGCTGATGACATGGAGCTGAAAGATGGAGATCAAATCGAT  
 301 GCATTGTTGCCTCAAGAAAGTGGT

## **ERKLÄRUNG**

“Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. George Coupland betreut worden.”

Köln, den 29.09.2005

Ruchika Budhiraja

## **Teilpublikationen**

M. Navatchkova, **R. Budhiraja**, G. Coupland, F. Eisenhaber and A. Bachmair (2004)  
*Planta* 220 (1) 1-8. SUMO conjugation in plants.

## **LEBENS LAUF**

### **Name**

Ruchika Budhiraja

### **Adresse**

Max-Planck-Institut für Züchtungsforschung  
Carl-von-Linné-Weg 10  
D-50829 Köln, Deutschland  
Tel: +49 (0) 221 5062 266  
Fax: +49 (0) 221 5062 207  
E-Mail: ruchika@mpiz-koeln.mpg.de

### **Geburtsdatum und ort**

19.10.1977, Chandigarh  
Indien

### **Nationalität**

Indisch

### **Ausbildung**

1984-1993 Grundschule in U.S.A und Indien  
1993-1995 Gymnasium in Indien  
1995-1998 Biologiestudium an der Universität zu Punjab (Chandigarh, Indien)  
1998-2000 M. Sc an der Universität zu Kurukshetra, (Kurukshetra, Indien)  
Spezialisierung: Microbiologie ,  
01.09.2002 Beginn der Doktorarbeit am Max-Planck-Institut für Züchtungsforschung  
in Köln, Abteilung Prof. Dr.George Coupland, Arbeitsgruppe Dr. Andreas  
Bachmair

### **Publikationen**

M. Navatchkova, **R. Budhiraja**, G. Coupland, F. Eisenhaber and A. Bachmair (2004)  
*Planta* 220 (1) 1-8. SUMO conjugation in plants.

(Ruchika Budhiraja)